Mathematical
Inequalities
& Applications
Volume 3, Number 4 (2000), 519–537

ON MIXED HÖLDER–MINKOWSKI INEQUALITIES AND TOTAL CONVEXITY OF CERTAIN FUNCTIONS IN $\mathscr{L}^p(\Omega)$

C. A. ISNARD AND A. N. IUSEM *

(*communicated by J. Borwein*)

Abstract. We prove the following mixed Hölder-Minkowski-type inequalities for all $x,z \in \mathscr{L}^p(\Omega)$:

$$0 \leqslant \frac{||z||_p^{p-s}}{s}\Big[\Big(||x||_p+||z||_p\Big)^s - ||x+z||_p^s\Big] \leqslant ||x||_p||y||_q - \mathrm{Re}(\langle x,y\rangle) \quad \text{if} \quad 1 \leqslant s \leqslant p \leqslant 2,$$

$$\frac{||z||_p^{p-s}}{s}\Big[\Big(||x||_p+||z||_p\Big)^s - ||x+z||_p^s\Big] \geqslant ||x||_p||y||_q - \mathrm{Re}(\langle x,y\rangle) \geqslant 0 \quad \text{if} \quad 2 \leqslant p \leqslant s,$$

where $y \in \mathscr{L}^q(\Omega)$ is defined as $y(\xi) = |z(\xi)|^{p-2}z(\xi)$, if $z(\xi) \neq 0$, $y(\xi) = 0$ otherwise, and $1/p + 1/q = 1$. Next we consider the Bregman distance $D_f : \mathscr{L}^p(\Omega) \times \mathscr{L}^p(\Omega) \to \mathbf{R}$ defined as $D_f(x,y) = f(x) - f(y) - \langle f'(y), x-y\rangle$, with $f(x) = ||x||_p^s$ ($s,p > 1$), and prove that $\inf\{D_f(u,z) : ||u-z||_p = t\} > 0$, $\sup\{D_f(u,z) : ||u-z||_p = t\} < \infty$, for all $p,s > 1$, all $z \in \mathscr{L}^p(\Omega)$ and all $t > 0$, so that the Bregman distance induced by $f(x) = ||x||_p^s$ and the metric distance $d(x,y) = ||x-y||_p$ are topologically equivalent. As a consequence, this f can be used in projection algorithms for the convex feasibility problem and generalized proximal point methods for convex optimization in $\mathscr{L}^p(\Omega)$.

1. Introduction

The purpose of this paper is twofold. In section 2. we consider the space $\mathscr{L}^p(\Omega)$ of measurable complex valued functions defined on a measure space $(\Omega, \mathscr{A}, \mu)$ with the p-norm given by $||x||_p = \big[\int_\Omega |x(\xi)|^p d\mu(\xi)\big]^{1/p} < \infty$ and prove that the following inequalities hold for all $x,z \in \mathscr{L}^p(\Omega)$, with $p \geqslant 1$:

$$0 \leqslant \frac{||z||_p^{p-s}}{s}\big[\big(||x||_p+||z||_p\big)^s - ||x+z||_p^s\big] \leqslant ||x||_p||y||_q - \mathrm{Re}(\langle x,y\rangle) \quad \text{if} \quad 1 \leqslant s \leqslant p \leqslant 2,$$

$$\frac{||z||_p^{p-s}}{s}\big[\big(||x||_p+||z||_p\big)^s - ||x+z||_p^s\big] \geqslant ||x||_p||y||_q - \mathrm{Re}(\langle x,y\rangle) \geqslant 0 \quad \text{if} \quad 2 \leqslant p \leqslant s,$$

where $y \in \mathscr{L}^q(\Omega)$ is defined as $y(\xi) = |z(\xi)|^{p-2}z(\xi)$ if $z(\xi) \neq 0$, $y(\xi) = 0$ otherwise, $1/p + 1/q = 1$, and $\langle x,y\rangle = \int_\Omega x(\xi)\overline{y(\xi)}d\mu(\xi)$. These inequalities relate

Mathematics subject classification (1991): 46B10, 46B25, 46E30.

Key words and phrases: Banach spaces, Hölder's inequality, Minkowski's inequality.

* Research of this author is partially supported by CNPq grant 301280/86.

Paper MIA-03-50

Hölder's and Minkowski's ones and are in fact stronger than these. The particular case of $s = p$ has been previously considered in [16].

In section 3. we use these inequalities to prove that the function $f(x) = ||x||_p^s$ in $\mathscr{L}^p(\Omega)$ with $p, s > 1$ enjoys the property of *total convexity*, which we describe next (a complete development appears in section 3). If f is a Fréchet differentiable and convex real function defined in a Banach space B, then $D_f : B \times B \to \mathbf{R}$, defined as $D_f(x,y) = f(x) - f(y) - \langle f'(y), x-y \rangle$, where $f'(y) \in B^*$ is the Fréchet derivative of f at y, introduces a sort of "distance" in B, called *Bregman distance*. This notion has its origin in [4] and has been extensively used in its finite dimensional version for convex optimization algorithms (e.g. [9], [13], [18]), and in particular, and more recently, for extensions of the proximal point method (e.g. [10], [11], [14], [15], [17]). Bregman distances have been considered in Hilbert spaces in [5], and in Banach spaces in [6], [7] and [8].

It is natural to compare the Bregman distance with the distance induced by the norm of the Banach space, which leads to the notion of *local modulus of convexity* of f, introduced in [6] and defined as $\nu_f(x,t) = \inf\{D_f(u,x) : ||u-x|| = t\}$. A function f is said to be *totally convex* if and only if $\nu_f(x,t) > 0$ for all $x \in B$ and all $t > 0$. Total convexity is a necessary assumption for convergence of projection algorithms for the convex feasibility problem ([6] and [8]) and of the proximal point method in Banach spaces ([2] and [7]), which use, as an auxiliary device, the Bregman distance associated with f. Therefore, it is important to identify totally convex functions which can be used in these algorithms, for relevant Banach spaces like $\mathscr{L}^p(\Omega)$.

Total convexity is a more demanding property that strict convexity (in fact it is equivalent to strict convexity in finite dimension) but it is weaker that either uniform or strong convexity (see definitions in section 3). Since there are no strongly convex and smooth functions in Banach spaces which are not Hilbert ones (e.g. $\mathscr{L}^p(\Omega)$ with $p \neq 2$, see section 3), and even uniformly convex functions are scarce in these spaces (e.g. $f(x) = ||x||_p^p$ is uniformly convex in $\mathscr{L}^p(\Omega)$ for $p \geqslant 2$, but not for $p < 2$), totally convex functions seem to be the right cathegory which extends the notion of strict convexity to infinite dimensional spaces.

Total convexity of $f(x) = ||x||_p^p$ in $\mathscr{L}^p(\Omega)$ with $p > 1$ has been proved, with arguments different from those used in this paper, in [6]. Here we establish total convexity of $f(x) = ||x||_p^s$ for all $p, s > 1$, providing closed formulae or explicit lower bounds for $\nu_f(x,t)$ in terms of p, s, $||x||_p$ and t. As a consequence, this familiy of functions can be used in the algorithms studied in [6], [7] and [8].

Throughout the paper $\mathbf{R}^n_+ = \{x \in \mathbf{R}^n : x_j \geqslant 0 \ (1 \leqslant j \leqslant n)\}$ and $\mathbf{R}^n_{++} = \{x \in \mathbf{R}^n : x_j > 0 \ (1 \leqslant j \leqslant n)\}$.

2. Some Hölder-Minkowski-type inequalities in $\mathscr{L}^p(\Omega)$

We will prove the following Hölder-Minkowski-type inequalities. Take $p \geqslant 1$, $x, z \in \mathscr{L}^p(\Omega)$, and define $y \in \mathscr{L}^q(\Omega)$ as $y(\xi) = |z(\xi)|^{p-2} z(\xi)$ if $z(\xi) \neq 0$, $y(\xi) = 0$

otherwise, with $1/p+1/q=1$. Then

$$0 \leqslant \frac{||z||_p^{p-s}}{s}\left[\left(||x||_p+||z||_p\right)^s-||x+z||_p^s\right] \leqslant ||x||_p||y||_q-\mathrm{Re}(\langle x,y\rangle) \quad \text{if } 1 \leqslant s \leqslant p \leqslant 2, \tag{2.1}$$

$$\frac{||z||_p^{p-s}}{s}\left[\left(||x||_p+||z||_p\right)^s-||x+z||_p^s\right] \geqslant ||x||_p||y||_q-\mathrm{Re}(\langle x,y\rangle) \geqslant 0 \quad \text{if } 2 \leqslant p \leqslant s, \tag{2.2}$$

(2.1)–(2.2) generalize the inequalities established in [16], which correspond to the case of $s=p$. We prove first a result in finite dimension, and then get (2.1)–(2.2) through a standard limiting argument. We recall that for $x \in \mathbf{C}^n$ the p-norm is defined as $||x||_p = \left[\sum_{j=1}^n |x_j|^p\right]^{1/p}$, and $\langle x,y\rangle = \sum_{j=1}^n x_j\overline{y_j}$. We exclude for the time being the case of $s=1$ for which a simple proof will be given in Theorem 2.1 (in fact the argument in the proof of the following proposition does not apply to this case).

PROPOSITION 2.1. *Take $x,z \in \mathbf{C}^n$, $p,q>1$ such that $1/p+1/q=1$ and define $y \in \mathbf{C}^n$ as $y_j=|z_j|^{p-2}z_j$ if $z_j \neq 0$, $y_j=0$ otherwise. Then*

$$0 \leqslant \sup_{t>0}\left\{\frac{||z||_p^{p-s}}{st}\left[\left(t||x||_p+||z||_p\right)^s-||tx+z||_p^s\right]\right\} = ||x||_p||y||_q - Re(\langle x,y\rangle) \tag{2.3}$$

if $1<s\leqslant p\leqslant 2$, and

$$\inf_{t>0}\left\{\frac{||z||_p^{p-s}}{st}\left[\left(t||x||_p+||z||_p\right)^s-||tx+z||_p^s\right]\right\} = ||x||_p||y||_q - Re(\langle x,y\rangle) \geqslant 0 \tag{2.4}$$

if $2\leqslant p\leqslant s$.

Proof. The result holds trivially if either x or z vanish. Thus, we assume that $x\neq 0\neq z$.

Define $\psi:\mathbf{R}_+\to\mathbf{R}$ as

$$\psi(t) = \left(t||x||_p+||z||_p\right)^s-||tx+z||_p^s. \tag{2.5}$$

We will prove that ψ is concave when $1<s\leqslant p\leqslant 2$ and convex when $2\leqslant p\leqslant s$. Let $I=\{j : x_j=z_j=0\}$ and $J=\{1,\dots,n\}\setminus I$. Then (2.5) can be rewritten as

$$\psi(t) = \left(t||x||_p+||z||_p\right)^s-\Big(\sum_{j\in J}|tx_j+z_j|^p\Big)^{s/p}. \tag{2.6}$$

It is easy to check that ψ is differentiable in $\mathbf{R}_+$ and that

$$\psi'(t) = s\left(t||x||_p+||z||_p\right)^{s-1}||x||_p - s||tx+z||_p^{s-p}\sum_{j\in J}|tx_j+z_j|^{p-2}\mathrm{Re}[(tx_j+z_j)\overline{x_j}], \tag{2.7}$$

with the convention that if $tx_j+z_j=0$ then the j-th term in the summation vanishes, and if $tx+z=0$ then

$$\psi'(t) = s\left(t||x||_p+||z||_p\right)^{s-1}||x||_p.$$

We claim now that ψ' is continuous in $\mathbf{R}_+$. Note that

$$|tx_j+z_j|^{p-2}\mathrm{Re}[(tx_j+z_j)\overline{x_j}] \leqslant |tx_j+z_j|^{p-1}|x_j|, \tag{2.8}$$

so that when t approaches a value such that $tx_j+z_j=0$ (if any) then the j-th term in the summation of (2.7) approaches 0, because $p>1$. In view of (2.8), the second term in the right hand side of (2.7) is bounded above by $s||tx+z||_p^{s-1}||x||_1$, and as a consequence, since $s>1$, this second term approaches 0 as t approaches a value such that $tx+z=0$ (if any). This settles the issue for these exceptional values of t, and the right hand side of (2.7) is clearly continuous for values of t other than these.

On the other hand, the second derivative of ψ is not defined at such exceptional values. Let $T=\{t>0 : tx_j+z_j=0 \ \text{ for some } j\}$. It is clear that T is either empty of finite. For $t\notin T$, ψ'' can be computed as

$$\begin{aligned}\psi''(t) &= s(s-1)(t||x||_p+||z||_p)^{s-2}||x||_p^2 \\ &\quad -s(s-p)||tx+z||_p^{s-2p}\Big\{\sum_{j\in J}|tx_j+z_j|^{p-2}\mathrm{Re}[(tx_j+z_j)\overline{x_j}]\Big\}^2 \\ &\quad -s||tx+z||_p^{s-p}\sum_{j\in J}\Big\{(p-2)|tx_j+z_j|^{p-4}\big[\mathrm{Re}((tx_j+z_j)\overline{x_j})\big]^2+|tx_j+z_j|^{p-2}|x_j|^2\Big\}.\end{aligned} \tag{2.9}$$

We consider first the case of $2\leqslant p\leqslant s$. Since $s-p\geqslant 0$, $p-2\geqslant 0$ and $\mathrm{Re}\,(a)\leqslant |a|$ for all $a\in\mathbf{C}$, we get from (2.9),

$$\begin{aligned}\psi''(t) &\geqslant s(s-1)\big(t||x||_p+||z||_p\big)^{s-2}||x||_p^2-s(s-p)||tx+z||_p^{s-2p}\Big[\sum_{j\in J}|tx_j+z_j|^{p-1}|x_j|\Big]^2 \\ &\quad -s(p-1)||tx+z||_p^{s-p}\sum_{j\in J}|tx_j+z_j|^{p-2}|x_j|^2.\end{aligned} \tag{2.10}$$

Take vectors v, w, $\hat{v}$, $\hat{w}\in\mathbf{R}_+^{|J|}$ defined as $v_j=|tx_j+z_j|^{p-2}$, $w_j=|x_j|^2$, $\hat{v}_j=|tx_j+z_j|^{p-1}$, $\hat{w}_j=|x_j|$.

By Hölder's inequality,

$$\sum_{j\in J}|tx_j+z_j|^{p-1}|x_j| = \langle\,\hat{v},\hat{w}\rangle \leqslant ||\hat{v}||_q||\hat{w}||_p = ||tx+z||_p^{p-1}||x||_p. \tag{2.11}$$

and, if $p>2$,

$$\sum_{j\in J}|tx_j+z_j|^{p-2}|x_j|^2 = \langle\, v,w\rangle \leqslant ||v||_{p/(p-2)}||w||_{p/2} = ||tx+z||_p^{p-2}||x||_p^2, \tag{2.12}$$

If $p=2$, then

$$\sum_{j\in J}|tx_j+z_j|^{p-2}|x_j|^2 = ||x||_2^2 = ||tx+z||_p^{p-2}||x||_p^2. \tag{2.13}$$

Since $s-p \geqslant 0$ and $p-1 \geqslant 0$, we get from (2.10), using (2.11) and (2.12) or (2.13),

$$\begin{aligned}\psi''(t) \geqslant & s(s-1)\big(t||x||_p+||z||_p\big)^{s-2}||x||_p^2 \\ & -s(s-p)||tx+z||_p^{s-2}||x||_p^2 - s(p-1)||tx+z||_p^{s-2}||x||_p^2 \\ = & s(s-1)||x||_p^2\big[\big(t||x||_p+||z||_p\big)^{s-2} - ||tx+z||_p^{s-2}\big] \geqslant 0, \end{aligned} \tag{2.14}$$

uisng Minkowski's inequality, since $s-2 \geqslant 0$. So $\psi''(t) \geqslant 0$ for $t \in \mathbf{R}_+ \setminus T$. Since ψ' is continuous in $\mathbf{R}_+$ and its derivative is nonnegative excepting at most at a finite set of points, we conclude that ψ' is nondecreasing in $\mathbf{R}_+$. Noting that $\psi(0) = 0$, we have, by the Mean Value Theorem,

$$\frac{\psi(t)}{t} = \frac{\psi(t)-\psi(0)}{t} = \psi'(\widehat{t}),$$

for some $\widehat{t} \in (0,t)$. Since ψ' is nondecreasing in $\mathbf{R}_+$, we get

$$\frac{\psi(t)}{t} \geqslant \psi'(0) \tag{2.15}$$

for all $t \geqslant 0$. Observe that $\lim_{t\to 0^+} \psi(t)/t = \lim_{t\to 0^+}(\psi(t)-\psi(0))/t = \psi'(0)$. In view of (2.15) we conclude that

$$\begin{aligned}\inf_{t>0}\left\{\frac{\psi(t)}{t}\right\} &= \psi'(0) = s||z||_p^{s-1}||x||_p - s||z||_p^{s-p}\sum_{j\in J}|z_j|^{p-2}\mathrm{Re}(z_j\overline{x_j}) \\ = s||z||_p^{s-1}||x||_p - s||z||_p^{s-p}\mathrm{Re}\left[\sum_{j=1}^{n}|z_j|^{p-2}z_j\overline{x_j}\right] &= s||z||_p^{s-p}\left[||z||_p^{p-1}||x||_p - \mathrm{Re}(\langle y,x\rangle)\right] \\ &= s||z||_p^{s-p}\left[||y||_q||x||_p - \mathrm{Re}(\langle x,y\rangle)\right], \end{aligned} \tag{2.16}$$

using (2.7) in the second equality. Now the equality in (2.4) follows immediately from (2.16). The inequality in (2.4) is Hölder's inequality, and so (2.4) holds.

Next we consider the case of $1 < s \leqslant p \leqslant 2$. In this case $s-p \leqslant 0$, $p-2 \leqslant 0$, and so the inequality in (2.10) is reversed. (2.11) and (2.13) still hold but the inequality in (2.12) is reversed as a consequence of Hölder's reverse inequality (e.g. [19], p. 99), which can be applied because $p/(2-p) < 0$ and v, w are nonnegative vectors. Since now we have $p-1 \geqslant 0$, $s-p \leqslant 0$ and $s-2 \leqslant 0$, the inequalities in (2.14) are reversed, and we conclude that ψ' is nonincreasing in $\mathbf{R}_+$. It follows that the inequality in (2.15) is reversed, and, using again the fact that $\lim_{t\to 0^+}\psi(t)/t = \psi'(0)$, we get

$$\sup_{t>0}\left\{\frac{\psi(t)}{t}\right\} = s||z||_p^{s-p}\left[||x||_p||y||_q - \mathrm{Re}(\langle x,y\rangle)\right]. \tag{2.17}$$

The equality in (2.3) follows immediately from (2.17) and the inequality is a consequence of Minkowski's inequality. □

COROLLARY 2.1. *(2.1) and (2.2) hold for* $x, z \in \mathbf{C}^n$ *with* $y \in \mathbf{C}^n$ *defined as* $y_j = |z_j|^{p-2} z_j$ *if* $z_j \neq 0$, $y_j = 0$ *otherwise.*

Proof. Take $t = 1$ in (2.3)–(2.4). □

We proceed now to extend (2.1)–(2.2) from $\mathbf{C}^n$ to $\mathscr{L}^p(\Omega)$, $\mathscr{L}^q(\Omega)$. We remark that there are several technical difficulties to go through the proof of Proposition 2.1 directly in these spaces, in the case $1 < s \leqslant p \leqslant 2$. Basically, the set T of singularities of ψ'' may be infinite: ψ'' diverges for all t such that $tx(\xi) + z(\xi) = 0$ for ξ in some subset of Ω of positive measure (if any), and, even if $tx(\xi) + z(\xi) \neq 0$ for all $\xi \in \Omega$, the scalar product $\langle v, w \rangle$ may diverge for certain values of t. Next we shall prove the infinite dimensional case by using the Dominated Convergence Theorem. We need first a preliminary lemma. A function $h : \Omega \to \mathbf{C}$ is said to be *simple* if it is measurable and its image is finite.

LEMMA 2.1. *Given any complex valued and measurable function* $h : \Omega \to \mathbf{C}$, *there exists a sequence of simple functions* $h_n : \Omega \to \mathbf{C}$ *such that for each* $\xi \in \Omega$ *it holds that* $\lim_{n\to\infty} h_n(\xi) = h(\xi)$ *and additionally* $|Re\ h_n(\xi)| \leqslant |Re\ h(\xi)|$, $|Im\ h_n(\xi)| \leqslant |Im\ h(\xi)|$ *(and therefore* $|h_n(\xi)| \leqslant |h(\xi)|$*) for all* $\xi \in \Omega$ *and all* n.

Proof. The result is well known at least when h is real valued and nonnegative, in which case the domination condition reduces to $0 \leqslant h_n(\xi) \leqslant h(x)$. The general case for real valued h can be proved by considering the restrictions of h and $-h$ to the sets $\{\xi \in \Omega : h(\xi) > 0\}$ and $\{\xi \in \Omega : h(\xi) < 0\}$ respectively, and the complex case by considering separately the real and imaginary parts, since any linear combination of simple functions is still a simple function. □

THEOREM 2.1. *Take* $p \geqslant 1$ *and* $x, z \in \mathscr{L}^p(\Omega)$. *Define* $y : \Omega \to \mathbf{C}$ *as* $y(\xi) = |z(\xi)|^{p-2} z(\xi)$ *if* $z(\xi) \neq 0$, $y(\xi) = 0$ *otherwise. Then* $y \in \mathscr{L}^q(\Omega)$ *with* $1/p + 1/q = 1$, *and* (2.1)–(2.2) *hold.*

Proof. The fact that y belongs to $\mathscr{L}^q(\Omega)$ is immediate from the definition of y. We proceed to prove the inequalities. For (2.1) we consider first the case of $1 < s \leqslant p \leqslant 2$, i.e. we leave aside the case of $s = 1$, which will be considered later on.

We start with the case when x and z are simple functions. In this situation there exists a partition $\Omega = \cup_{i=0}^m \Omega_i$ in pairwise disjoint measurable sets such that both x and z are constant on each Ω_i. By convention, $\Omega_0 = \{\xi \in \Omega : x(\xi) = 0 = z(\xi)\}$ (Ω_0 is possibly empty). Let α_i, β_i be the values of x and z respectively in the set Ω_i, and $\gamma_i = \beta_i |\beta_i|^{p-2}$ if $\beta_i \neq 0$, $\gamma_i = 0$ otherwise. Then $||x||_p^p = \sum_{i=1}^n \alpha_i \mu(\Omega_i)$, $||z||_p^p = \sum_{i=1}^n \beta_i \mu(\Omega_i)$ so that $\mu(\Omega_i) < \infty$ for $i \geqslant 1$, because $\alpha_i \neq 0$ or $\beta_i \neq 0$. By definition of y, we have $y(\xi) = \gamma_i$ for all $\xi \in \Omega_i$. Now define $\hat{x}$, $\hat{y}$, $\hat{z} \in \mathbf{C}^m$ as $\hat{x}_i = \alpha_i \mu(\Omega_i)^{1/p}$, $\hat{z}_i = \beta_i \mu(\Omega_i)^{1/p}$ and $\hat{y}_i = \gamma_i \mu(\Omega_i)^{1/q}$. $\hat{y}$ and $\hat{z}$ are related as requested in the assumptions of Corollary 2.1 and so (2.1) and (2.2) hold with $\hat{x}$, $\hat{y}$ and $\hat{z}$ substituting for x, y and z. We observe that $||x||_p = ||\hat{x}||_p$, $||y||_q = ||\hat{y}||_q$, $||z||_p = ||\hat{z}||_p$, $\langle x, y \rangle = \langle \hat{x}, \hat{y} \rangle$ and conclude that (2.1) and (2.2) hold for x, y and z. This settles the issue for simple functions. Now we consider arbitrary $x, z \in \mathscr{L}^p(\Omega)$.

By Lemma 2.1 there exist sequences of simple functions $\{x_n\}$ and $\{z_n\}$ defined on Ω such that for all $\xi \in \Omega$ it holds that $\lim_{n\to\infty} x_n(\xi) = x(\xi)$ and $\lim_{n\to\infty} z_n(\xi) = z(\xi)$, with $|x_n(\xi)| \leqslant |x(\xi)|$, $|z_n(\xi)| \leqslant |z(\xi)|$ for all $\xi \in \Omega$ and all n. By Lebesgue's Dominated Convergence Theorem we have

$$\lim_{n\to\infty} ||x_n||_p = ||x||_p, \quad \lim_{n\to\infty} ||z_n||_p = ||z||_p, \quad \lim_{n\to\infty} ||x_n + z_n||_p = ||x+z||_p, \quad (2.18)$$

because $|x_n(\xi) + z_n(\xi)| \leqslant |x(\xi)| + |z(\xi)|$. Let now $y_n(\xi) = |z_n(\xi)|^{p-2} z_n(\xi)$ if $z_n(\xi) \neq 0$, $y_n(\xi) = 0$ otherwise. Then

$$|y_n(\xi)| = |z_n(\xi)|^{p-1} \leqslant |z(\xi)|^{p-1} = y(\xi).$$

Hence, by the Dominated Convergence Theorem again, we have

$$\lim_{n\to\infty} ||y_n||_q = ||y||_q, \quad \lim_{n\to\infty} \langle x_n, y_n \rangle = \langle x, y \rangle, \quad (2.19)$$

because $|x_n(\xi)\overline{y_n(\xi)}| \leqslant |x(\xi)\overline{y(\xi)}|$. Since the result holds for simple functions, we know that (2.1) and (2.2) hold with x_n, y_n and z_n substituting for x, y and z respectively. The required inequalities (excepting for the case of $s = 1$) follow by taking limits as n goes to ∞, in view of (2.18) and (2.19).

Finally, we consider the case of $s = 1$. First, we remark that for $p = 1$ we get $y \in \mathscr{L}^\infty(\Omega)$. Note that $||y||_q = ||z||_p^{p-1}$ and $|z(\xi)|^p = \overline{y(\xi)} z(\xi)$ for all $\xi \in \Omega$, so that

$$\langle z, y \rangle = ||z||_p^p = ||z||_p ||y||_q. \quad (2.20)$$

Since $s = 1$, (2.1) is equivalent to

$$\mathrm{Re}(\langle x, y \rangle) + ||z||_p ||y||_q \leqslant ||x+z||_p ||y||_q. \quad (2.21)$$

In view of (2.20), we have

$$\mathrm{Re}(\langle x, y \rangle) + ||y||_q ||z||_p = \mathrm{Re}(\langle x+z, y \rangle) \leqslant ||x+z||_p ||y||_q,$$

by Hölder's inequality, so that (2.21) (and consequently (2.1) for the case $s = 1$) holds. □

3. Total convexity of $f(x) = ||x||_p^s$ in $\mathscr{L}^p(\Omega)$

Let B be a Banach space, B^* its dual, $f : B \to \mathbf{R}$ a convex function and $\partial f(x) \subset B^*$ the subdifferential of f at x, i.e.

$$\partial f(x) = \{\eta \in B^* : \langle \eta, x - y \rangle \leqslant f(y) - f(x) \text{ for all } y \in B\},$$

where $\langle \cdot, \cdot \rangle : B^* \times B \to \mathbf{R}$ denotes the usual duality pairing.

By convexity of f, $\partial f(x)$ is nonempty and bounded for all $x \in B$ (see [12]). We define $D_f : B \times B \to \mathbf{R}$ as

$$D_f(x, y) = f(x) - f(y) - \inf\{\langle \eta, x - y \rangle : \eta \in \partial f(y)\}.$$

Boundedness of $\partial f(y)$ guarantees that $D_f(x,y)$ is well defined for all $(x,y) \in B \times B$. If f is Fréchet differentiable, then

$$D_f(x,y) = f(x) - f(y) - \langle f'(y), x-y \rangle, \tag{3.1}$$

where $f'(y)$ is the Fréchet derivative of f at y. By definition of ∂f, $D_f(x,y) \geqslant 0$ for all $x,y \in B$ and $D_f(x,x) = 0$ for all $x \in B$. In general, it is not true that $D_f(x,y) = D_f(y,x)$ and the triangular inequality does not hold. D_f is called the *Bregman distance* associated with f.

The *local modulus of convexity* of f, denoted as $\nu_f : B \times \mathbf{R}_+ \to \mathbf{R}$ has been defined in [6] as

$$\nu_f(z,t) = \inf\{D_f(u,z) : ||u-z|| = t\}, \tag{3.2}$$

and$f$ is said to be *totally convex* if and only if $\nu_f(x,t) > 0$ for all $z \in B$ and all $t > 0$.

We comment next on the relation between total convexity and other variations of convexity.

f is said to be *strictly convex* if and only if $f(x) - f(y) - \langle \eta, x-y \rangle > 0$ for all $x,y \in B$ such that $x \neq y$ and all $\eta \in \partial f(y)$.

f is said to be *uniformly convex* if and only if $\langle \xi - \eta, x-y \rangle > \gamma ||x-y||^p$ for all $x,y \in B$, all $\xi \in \partial f(x)$, all $\eta \in \partial f(y)$ and some $\gamma > 0$, some $p > 1$ (see [1] for a slightly less restrictive definition of uniform convexity, and [20] for a related but different definition).

f is said to be *strongly convex* if and only if $\langle \xi - \eta, x-y \rangle > \gamma ||x-y||^2$ for all $x,y \in B$, all $\xi \in \partial f(x)$, all $\eta \in \partial f(y)$ and some $\gamma > 0$.

Clearly, strong convexity implies uniform convexity. It has been proved in [6] that uniform convexity implies total convexity, and the fact that total convexity implies strict convexity is immediate. In finite dimension, a simple compactness argument shows that strict convexity is equivalent to total convexity, but in infinite dimensional spaces there are strictly convex functions which are not totally convex (see [6]). In Banach spaces which are not Hilbert ones, there are no strongly convex functions which are sufficiently smooth: it has been proved in [3] that if $f : B \to \mathbf{R}$ is a strongly convex function which is twice differentiable at least at some $z \in B$, then B is isomorphic to a Hilbert space, because $\left[D^2 f(z)(x,x)\right]^{1/2}$ (where $D^2 f(z)$ denotes the bilinear form of the second derivative of f at z) defines a Hilbertian norm which is equivalent to the given norm in B. In view of this result, it is relevant to study the existence of totally convex functions in Banach spaces which are not Hilbert, e.g. in $\mathscr{L}^p(\Omega)$. Also, total convexity if necessary for convergence of several algorithms in Banach spaces. As an example, we mention the proximal point method for minimizing a convex function $g : B \to \mathbf{R}$. This method starts with some $x^0 \in B$ and generates a sequence $\{x^k\} \subset B$ through the iteration

$$x^{k+1} = \operatorname{argmin}\left\{g(x) + D_f(x, x^k)\right\},$$

where f is an auxiliary convex function, and D_f is as in (3.1). Convergence of $\{x^k\}$ to a minimizer of g is established in [7] under hypotheses on f which include total convexity.

In this section we will prove that $f : \mathscr{L}^p(\Omega) \to \mathbf{R}$ defined as $f(x) = ||x||_p^s$ is totally convex for all $s,p > 1$. We consider the case of the space $\mathscr{L}^p(\Omega)$ of complex valued functions from Ω to $\mathbf{R}$. We present closed formulae or explicit lower bounds for $\nu_f(x,t)$ in terms of p, s, $||x||_p$ and t. Obviously, these bounds are also valid for the real case.

We will consider several cases, depending on the values of s and p. The first one corresponds to $1 < s \leqslant p \leqslant 2$. In this case we give a closed formula for $\nu_f(x,t)$ and the point u^* which realizes the infimum in the definition of ν_f. This result is a consequence of Theorem 2.1. The formula of ν_f extends the one established in [6] for the particular case of $s = p$.

From now on we will use the following notation: $\nu^{ps}(z,t) = \nu_f(z,t)$ and $D^{ps}(x,y) = D_f(x,y)$ for $f(x) = ||x||_p^s$ in $\mathscr{L}^p(\Omega)$, and $U_t = \{x \in \mathscr{L}^p(\Omega) : ||x||_p = t\}$ for $t \in \mathbf{R}_+$.

PROPOSITION 3.1. *Take s, p such that $1 < s \leqslant p \leqslant 2$. Then*

i) $\nu^{ps}(z,t) = \big(t + ||z||_p\big)^s - ||z||_p^s - s||z||_p^{s-1}t = D_\varphi\big(t + ||z||_p, ||z||_p\big) > 0$ *for all* $t > 0$ *with* $\varphi : \mathbf{R}_+ \to \mathbf{R}$ *defined as* $\varphi(t) = t^s$,

ii) if $z \neq 0$ *then* $\nu^{ps}(z,t) = D^{ps}(u^*,z)$ *with* $u^* = \big(1 + t/||z||_p\big)z$.

Proof. A simple computation, using (3.1), shows that

$$D^{ps}(u^*,z) = \big(t + ||z||_p\big)^s - ||z||_p^s - s||z||_p^{s-1}t = D_\varphi\big(t + ||z||_p, ||z||_p\big),$$

if $z \neq 0$. The rightmost inequality in (i) follows from strict convexity of φ in $\mathbf{R}_{++}$. It suffices therefore to prove that $D^{ps}(u,z) \geqslant D^{ps}(u^*,z)$ for all $u \in B$ such that $||u - z||_p = t$, or equivalently

$$D^{ps}(x + z, z) \geqslant \big(t + ||z||_p\big)^s - ||z||_p^s - s||z||_p^{s-1}t \tag{3.3}$$

for all $x \in U_t$. It is easy to get from (3.1) the explicit expression for the left hand side of (3.3), namely

$$D^{ps}(x + z, z) = ||x + z||_p^s - ||z||_p^s - s||z||_p^{s-p}\mathrm{Re}(\langle y, x\rangle), \tag{3.4}$$

with $y(\xi) = |z(\xi)|^{p-2}z(\xi)$ if $z(\xi) \neq 0$, $y(\xi) = 0$ otherwise. In view of (3.4), it turns out that, since $||x||_p = t$ and $||z||_p^{p-1} = ||y||_q$, (3.3) is a direct consequence of (2.1), which holds by virtue of Theorem 2.1. □

The computation of ν^{ps} for values of s, p other than those considered in Proposition 3.1 is much harder, and we will not obtain a closed formula for $\nu^{ps}(z,t)$, but rather a positive lower bound. We start with a lower bound for ν^{pp} with $p \geqslant 2$.

As discussed in [6], $f(x) = ||x||_p^p$ is totally convex for $p \geqslant 2$, because it is uniformly convex, as proved in [20]. Our next proposition gives an explicit positive lower bound for $\nu^{pp}(z,t)$ with $p \geqslant 2$. We need first two elementary results, the first of which can be seen as a chain rule for D_f.

LEMMA 3.1. *Take $\varphi : \mathbf{R} \to \mathbf{R}$ convex, differentiable and nondecreasing, and $g : B \to \mathbf{R}$ convex. Define $f : B \to \mathbf{R}$ as $f(x) = \varphi(g(x))$. Then*

$$D_f(x,y) = D_\varphi(g(x), g(y)) + \varphi'(g(y))D_g(x,y). \tag{3.5}$$

Proof. See [16], Proposition 3. □

LEMMA 3.2. *Take $p \geqslant 2$ and define $\phi_p : \mathbf{C} \to \mathbf{R}$ as $\phi_p(t) = |1+t|^p - |t|^p - p|t|^{p-2}Re(t)$. Let $c_p = \inf_{t\in\mathbf{C}} \phi_p(t)$. Then, $c_p = \min_{||t||\leqslant 1/2} \phi_p(t) > 0$ and $p2^{1-p} \geqslant c_p$, so that $c_p < 1$ if $p > 2$.*

Proof. For $p = 2$ we compute $c_p = \phi_2(t) = 1$ via $|1+t|^2 = 1+2\mathrm{Re}(t)+|t|^2$. For $p > 2$ we define $g_p : \mathbf{C} \to \mathbf{R}$ as $g_p(t) = |t|^p$. It follows that $\phi_p(t) = D_{g_p}(1+t,t)$ and $g_p(t) = \big(g_2(t)\big)^{p/2}$. Since $p/2 > 1$, the function $\psi : \mathbf{R} \to \mathbf{R}$ defined by $\psi(r) = r^{p/2}$ is convex. Therefore, we get from Lemma 3.1

$$\phi_p(t) = D_{g_p}(1+t,t) \geqslant \psi'\big(g_2(t)\big)D_{g_2}(1+t,t) = \frac{p}{2}\big(|t|^2\big)^{p/2-1}\phi_2(t) = \frac{p}{2}|t|^{p-2}. \quad (3.6)$$

By (3.6),

$$\inf_{|t|\geqslant 1/2} \phi_p(t) \geqslant \frac{p}{2}2^{2-p} = p2^{1-p} \geqslant \min_{|t|\leqslant 1/2} \phi_p(t) = c_p. \quad (3.7)$$

It follows from (3.7) that $c_p > 0$, because $\{t \in \mathbf{C} : |t| \leqslant 1/2\}$ is compact, $\phi_p(t)$ is continuous, $\phi_p(t) \geqslant (p/2)|t|^{p-2} > 0$ if $t \neq 0$ and $\phi_p(0) = 1$.

If $p > 2$ we have $2^{p-1} > p$, because the function $\beta : \mathbf{R} \to \mathbf{R}$ defined as $\beta(r) = r^{p-1}$ is strictly convex in $(0,+\infty)$, and so we have $\beta(2) > \beta(1) + \beta'(1) = p$. It follows that $c_p < 1$. □

REMARK 3.1. In the Appendix we establish the following estimate:

$$p2^{1-p} \geqslant c_p > \left[1 + (2p-1)^{1/(1-p)}\right]^{1-p} \geqslant 2^{1-p}(2-1/p),$$

which gives also lower bounds for c_p.

PROPOSITION 3.2. *If $p \geqslant 2$ then $\nu^{pp}(z,t) = c_p t^p > 0$, with $c_p > 0$ as given by Lemma 3.2.*

Proof. By (3.1), for all $x, z \in \mathscr{L}^p(\Omega)$

$$D^{pp}(x+z,z) = \int_\Omega \big[|x(\xi)+z(\xi)|^p - |z(\xi)|^p - p|z(\xi)|^{p-2}\mathrm{Re}\big(z(\xi)\overline{x(\xi)}\big)\big]d\mu(\xi). \quad (3.8)$$

Let $\Omega(x) = \{\xi \in \Omega : x(\xi) \neq 0\}$ and $u(\xi) = z(\xi)/x(\xi)$ for $\xi \in \Omega(x)$. Since $x \in \mathscr{L}^p(\Omega)$, $\Omega(x)$ is measurable, and, since the integrand in (3.8) vanishes for $\xi \notin \Omega(x)$, we get

$$\begin{aligned} D^{pp}(x+z,z) &= \int_{\Omega(x)} |x(\xi)|^p\big[|1+u(\xi)|^p - |u(\xi)|^p - p|u(\xi)|^{p-2}\mathrm{Re}\big(u(\xi)\big)\big]d\mu(\xi) \\ &= \int_{\Omega(x)} |x(\xi)|^p \phi_p(u(\xi))d\mu(\xi), \end{aligned} \quad (3.9)$$

with ϕ_p as in Lemma 3.2. By Lemma 3.2, for all $x \in U_t$,

$$D^{pp}(x+z,z) \geqslant c_p \int_{\Omega(x)} |x(\xi)|^p d\mu(\xi) = c_p t^p, \quad (3.10)$$

with equality occurring if $\phi_p\big(u(\xi)\big) = c_p$ for $\xi \in \Omega(x)$, i.e. if $x(\xi) = \lambda^{-1}u(\xi)$, where, for $p > 2$, $\lambda \neq 0$ is such that $\phi_p(\lambda) = c_p$ and, for $p = 2$, λ is arbitrary, because in such cases we have $\phi_p(t) = 1 > c_p$ by Lemma 3.2 for $p > 2$ and $\phi_2(t) = 1 = c_2$ for all t.

The result follows by taking infimum over $x \in U_t$ in (3.10). □

We mention that, though in this case we have a lower bound for v^{ps} rather than an exact formula, the result is in some sense stronger than for the case of $1 < s \leqslant p \leqslant 2$, because our lower bound for $p \geqslant 2$ is independent of z, so that $v^{pp}(z,t)$ is a global, rather than local, modulus of convexity, confirming the fact that $f(x) = ||x||_p^p$ is uniformly convex for $p \geqslant 2$ (see [20]). It is not possible to get a positive lower bound independent of z for $v^{ps}(z,t)$ with $1 < s \leqslant p \leqslant 2$ because for any fixed $t > 0$ the value of $v^{ps}(z,t)$ given by Proposition 3.1(i) converges to 0 as $||z||_p$ goes to ∞, as can be easily verified.

We will use the values of v^{ps} obtained in Propositions 3.1 and 3.2 in order to compute lower bounds for v^{ps} with values of p and s not covered in these propositions. The first case is $s \geqslant p$, for which we will use Lemma 3.1.

PROPOSITION 3.3. *If $s \geqslant p > 1$ and $t > 0$ then*

$$v^{ps}(0,t) = t^s > 0, \tag{3.11}$$

$$v^{ps}(z,t) \geqslant \frac{s}{p}||z||_p^{s-p}v^{pp}(z,t) > 0 \qquad \textit{if} \ \ z \neq 0. \tag{3.12}$$

Proof. The formula for D^{ps} given in (3.4) holds in fact for all $p, s > 1$. (3.11) is a direct consequence of (3.2) and (3.4). The rightmost inequality in (3.12) follows from Proposition 3.2 for $p > 2$ and from Proposition 3.1 with $s = p$ for $p \leqslant 2$. We proceed to establish the leftmost inequality in (3.12). Take $\varphi(t) = t^{s/p}$, $f(x) = ||x||_p^s$ and $g(x) = ||x||_p^p$. Since $s \geqslant p$, φ is convex and nondecreasing. Thus, we can apply Lemma 3.1, getting

$$D^{ps}(x+z,z) = D_\varphi(||x+z||_p^p, ||z||_p^p) + \frac{s}{p}||z||_p^{s-p}D^{pp}(x+z,z). \tag{3.13}$$

Since $D_\varphi(||x+z||_p^p, ||z||_p^p) \geqslant 0$ by convexity of φ, we get from (3.13)

$$D^{ps}(x+z,z) \geqslant \frac{s}{p}||z||_p^{s-p}D^{pp}(x+z,z), \tag{3.14}$$

and the result follows taking infima over $x \in U_t$ on both sides of (3.14). □

The remaining case, i.e. $p > \max\{s,2\}$, is the hardest, and we will estimate v^{ps} by partitioning U_t into three subsets and getting lower bounds for $D^{ps}(x+z,z)$ with x in each one of them separately. Let $y(\xi) = |z(\xi)|^{p-2}z(\xi)$ if $z(\xi) \neq 0$, $y(\xi) = 0$ otherwise, and define, for $t > 0$:

$$V_t^1 = \{x \in U_t : \mathrm{Re}(\langle x,y\rangle) \geqslant 0\}, \tag{3.15}$$

$$V_t^2 = \{x \in U_t : ||x+z||_p \geqslant ||z||_p, \mathrm{Re}(\langle x, y\rangle) < 0\}, \tag{3.16}$$

$$V_t^3 = \{x \in U_t : ||x+z||_p < ||z||_p\}. \tag{3.17}$$

It is clear that $U_t = \cup_{i=1}^3 V_t^i$. We remark that V_t^3 might be empty, e.g. if $t > 2||z||_p$. We need next a preliminary result.

LEMMA 3.3. *Let $\phi : \mathbf{R}_+ \to \mathbf{R}$ be a convex and differentiable function. Then, for any fixed $t_0 \in \mathbf{R}_+$ the function $t \mapsto D_\phi(t, t_0)$ is nonincreasing in $[0, t_0]$ and nondecreasing in $[t_0, +\infty)$*

Proof. By (3.1)

$$\frac{d}{dt} D_\phi(t, t_0) = \phi'(t) - \phi'(t_0). \tag{3.18}$$

By convexity of ϕ the right hand side of (3.18) is nonpositive if $t \leqslant t_0$ and nonnegative if $t \geqslant t_0$. $\square$

PROPOSITION 3.4. *If $1 < s \leqslant p$ and $p \geqslant 2$ then, for all $x \in V_t^1$,*

$$||x+z||_p \geqslant \left(||z||_p + c_p t^p\right)^{1/p}, \tag{3.19}$$

$$D^{ps}(x+z, z) \geqslant D_\varphi\left(\left[2^{1-p} t^p + ||z||_p^p\right]^{1/p}, ||z||_p\right), \tag{3.20}$$

with $\varphi : \mathbf{R}_+ \to \mathbf{R}$ defined as $\varphi(t) = t^s$ and c_p as in Lemma 3.2.

Proof. Take $x \in V_t^1$. Since $\mathrm{Re}(\langle x, y\rangle) \geqslant 0$ by (3.15), we get from Proposition 3.2 and (3.4) with $s = p$,

$$||x+z||_p^p \geqslant D^{pp}(x+z, z) + ||z||_p^p \geqslant v^{pp}(z, t) + ||z||_p^p \geqslant c_p t^p + ||z||_p^p. \tag{3.21}$$

By (3.21),

$$||x+z||_p \geqslant \left(c_p t^p + ||z||_p^p\right)^{1/p}, \tag{3.22}$$

and thus (3.19) holds. Now we use Lemma 3.1 with $f(x) = ||x||_p^s$, $g(x) = ||x||_p$ and $\varphi(t) = t^s$. Since g is convex, $D_g(x+z, z) \geqslant 0$, and also $\varphi'(||z||_p) \geqslant 0$, because $s > 0$. Therefore

$$D^{ps}(x+z, z) = D_\varphi(||x+z||_p, ||z||_p) + \varphi'(||z||_p) D_g(x+z, z) \geqslant D_\varphi(||x+z||_p, ||z||_p). \tag{3.23}$$

By Lemma 3.3 and (3.22),

$$D_\varphi(||x-z||_p, ||z||_p) \geqslant D_\varphi\left(c_p t^p + ||z||_p^p\right)^{1/p}, ||z||_p\big). \tag{3.24}$$

(3.20) follows from (3.19), (3.24) and (3.23). $\square$

PROPOSITION 3.5. *If* $1 < s \leqslant p$, $p \geqslant 2$, *then, for all* $x \in V_t^2$,

$$D^{ps}(x+z,z) \geqslant \frac{s}{p}\big(t + ||z||p\big)^{s-p} c_p t^p, \tag{3.25}$$

with c_p *as in Lemma 3.2.*

Proof. Take $x \in V_t^2$. Let $a = ||x+z||_p^p$, $b = ||z||_p^p$, $c = p||z||_p^{-p}\mathrm{Re}(\langle x,y\rangle)$ and $r = s/p$, with y as in the definition of the sets V_t^i. Then, by (3.4),

$$D^{ps}(x+z,z) = a^r - b^r - rb^r c, \tag{3.26}$$

$$D^{pp}(x+z,z) = a - b - bc. \tag{3.27}$$

Let $\psi(t) = t^r$. Note that $r \in (0,1]$ because $s \leqslant p$, so that ψ is concave in $\mathbf{R}_+$. It follows that $\psi(b) \leqslant \psi(a) + \psi'(a)(b-a)$, implying

$$a^r - b^r \geqslant ra^{r-1}(a-b). \tag{3.28}$$

By (3.26), (3.27) and (3.28),

$$\begin{aligned} D^{ps}(x+z,z) &\geqslant ra^{r-1}(a-b) - rb^r c = ra^{r-1}(a-b-bc) + ra^{r-1}bc - rb^r c \\ &= ra^{r-1}D^{pp}(x+z,z) + rbc(a^{r-1} - b^{r-1}). \end{aligned} \tag{3.29}$$

Since $x \in V_t^2$, we have that $a \geqslant b$ and $c < 0$, so that $a^{r-1} \leqslant b^{r-1}$, because $r-1 \leqslant 0$, and hence $rbc(a^{r-1} - b^{r-1}) \geqslant 0$. Therefore, using (3.29) and Proposition 3.2,

$$\begin{aligned} D^{ps}(x+z,z) &\geqslant ra^{r-1}D^{pp}(x+z,z) \\ &\geqslant \frac{s}{p}||x+z||_p^{s-p}\nu^{pp}(z,t) = \frac{s}{p}||x+z||_p^{s-p}c_p t^p. \end{aligned} \tag{3.30}$$

Since $s - p \leqslant 0$, Minkowski's inequality implies that

$$||x+z||_p^{s-p} \geqslant \big(||x||_p + ||z||_p\big)^{s-p} = \big(t + ||z||_p\big)^{s-p}. \tag{3.31}$$

The result follows from (3.30) and (3.31). □

Finally, we establish a lower bound for $D^{ps}(x+z,z)$ over $x \in V_t^3$. We need first an elementary result.

LEMMA 3.4. *Define* $\tau : \mathbf{R}_{++} \to \mathbf{R}$ *as* $\tau(t) = \beta + (\alpha^t - 1)/t$, *with* $\alpha \in \mathbf{R}_{++}$, $\beta \in \mathbf{R}$. *Then* τ *is concave if* $\alpha \leqslant 1$ *and convex if* $\alpha \geqslant 1$.

Proof. Consider $\gamma : \mathbf{R}_{++} \to \mathbf{R}$ defined as $\gamma(t) = t(\log t)^2 - 2t\log t + 2t - 2$. Since $\gamma'(t) = (\log t)^2 > 0$, we get

$$\gamma(t) \leqslant \gamma(1) = 0 \qquad \text{if } t \in (0,1], \tag{3.32}$$

$$\gamma(t) \geqslant \gamma(1) = 0 \qquad \text{if } t \geqslant 1. \tag{3.33}$$

On the other hand, differentiating τ,

$$\tau''(t) = t^{-3}\Big[\alpha^t\big(\log\alpha^t\big)^2 - 2\alpha^t\log\alpha^t + 2\alpha^t - 2\Big] = t^{-3}\gamma(\alpha^t). \tag{3.34}$$

Since $t > 0$, we have that $\alpha^t \geqslant 1$ if and only if $\alpha \geqslant 1$. The result follows from (3.32), (3.33) and (3.34). □

PROPOSITION 3.6. *If* $1 < s \leqslant p$, $p \geqslant 2$, *then for all* $z \neq 0$ *and all* $x \in V_t^3$,

$$D^{ps}(x+z,z) \geqslant \frac{s(s-1)}{p(p-1)}||z||_p^{s-p} c_p t^p. \tag{3.35}$$

Proof. Take $x \in V_t^3$. We remark that (3.4) holds also for $s = 1$ when $z \neq 0$, because $||\cdot||_p$ is differentiable everywhere except at 0. We consider τ as in Lemma 3.4 with $\alpha = ||x+z||_p/||z||_p$, $\beta = -||z||_p^{-p}\mathrm{Re}(\langle x, y\rangle)$ and y as in the definition of the sets V_t^i. It follows easily from (3.4) that

$$D^{ps}(x+z,z) = s||z||_p^s \tau(s). \tag{3.36}$$

Observe that $s = \frac{p-s}{p-1} + \left(\frac{s-1}{p-1}\right)p$ with $\frac{p-s}{p-1} \in [0,1)$ and $\frac{p-s}{p-1} + \frac{s-1}{p-1} = 1$. Since $||x+z||_p/||z||_p \in (0,1)$ by definition of V_t^3, τ is concave by Lemma 3.4, and therefore

$$\tau(s) \geqslant \left(\frac{p-s}{p-1}\right)\tau(1) + \left(\frac{s-1}{p-1}\right)\tau(p). \tag{3.37}$$

By (3.36) and (3.37),

$$\begin{aligned} \frac{||z||_p^{-s}}{s} D^{ps}(x+z,z) &\geqslant \left(\frac{p-s}{p-1}\right)||z||_p^{-1} D^{p1}(x+z,z) + \left(\frac{s-1}{p-1}\right)\frac{||z||_p^{-p}}{p} D^{pp}(x+z,z) \\ &\geqslant \frac{(s-1)}{p(p-1)}||z||_p^{-p} D^{pp}(x+z,z), \end{aligned} \tag{3.38}$$

using nonnegativity of D^{p1}, which follows from convexity of $g(x) = ||x||_p$, in the rightmost inequality of (3.38). By (3.38), using Proposition 3.2,

$$\begin{aligned} D^{ps}(x+z,z) &\geqslant \frac{s(s-1)}{p(p-1)}||z||_p^{s-p} D^{pp}(x+z,z) \\ &\geqslant \frac{s(s-1)}{p(p-1)}||z||_p^{s-p} \nu^{pp}(z,t) \geqslant \frac{s(s-1)}{p(p-1)}||z||_p^{s-p} c_p t^p. \end{aligned}$$

□

Now we establish a common lower bound for those found in Propositions 3.4, 3.5 and 3.6.

PROPOSITION 3.7. *If* $1 < s \leqslant p$, $p \geqslant 2$ *and* $t > 0$ *then*

$$\nu^{ps}(0,t) = t^s > 0, \tag{3.39}$$

$$\nu^{ps}(z,t) \geqslant \left(1 + \frac{t}{||z||_p}\right)^{s-p} D_\varphi\left(\left[c_p t^p + ||z||_p^p\right]^{1/p}, ||z||_p\right) > 0 \quad \text{if } z \neq 0, \tag{3.40}$$

with $\varphi(t) = t^s$ *and* c_p *as in Lemma 3.2.*

Proof. (3.39) holds with the same proof as in Proposition 3.3. We proceed to prove (3.40). Let $\rho_1(z,t) = D_\varphi\left(\left[c_p t^p + ||z||_p^p\right]^{1/p}, ||z||_p\right)$, $\rho_2(z,t) = \left(t + ||z||_p\right)^{s-p} c_p t^p$ and

$\rho_3(z,t) = \frac{s(s-1)}{p(p-1)}||z||_p^{s-p} c_p t^p$. By Propositions 3.4, 3.5 and 3.6, since $U_t = \cup_{i=1}^3 V_t^i$, in order to establish the leftmost inequality in (3.40) it suffices to prove that

$$\min_{1 \leqslant i \leqslant 3} \rho_i(z,t) \geqslant \left(1 + \frac{t}{||z||_p}\right)^{s-p} \rho_1(z,t). \tag{3.41}$$

Let $\delta = 1 + t/||z||_p$. Note that

$$\rho_2(z,t) = \delta^{s-p} \frac{p(p-1)}{s(s-1)} \rho_3(z,t) \geqslant \delta^{s-p} \rho_3(z,t), \tag{3.42}$$

because $[p(p-1)]/[s(s-1)] \geqslant 1$. We claim that

$$\rho_3(z,t) \geqslant \rho_1(z,t). \tag{3.43}$$

We proceed to prove the claim. Let $\alpha = \left(1 + c_p t^p/||z||_p^p\right)^{1/p}$. By (3.1),

$$\begin{aligned}
\rho_1(z,t) &= D_\varphi\left(\left[c_p t^p + ||z||_p^p\right]^{1/p}, ||z||_p\right) \\
&= \left(c_p t^p + ||z||_p^p\right)^{s/p} - ||z||_p^s - s||z||_p^{s-1} \left[\left(c_p t^p + ||z||_p^p\right)^{1/p} - ||z||_p\right] \\
&= ||z||_p^s \left[\alpha^s - 1 - s(\alpha - 1)\right].
\end{aligned} \tag{3.44}$$

On the other hand,

$$\rho_3(z,t) = \frac{s(s-1)}{p(p-1)} ||z||_p^s \left(\frac{c_p t^p}{||z||_p^p}\right) = \frac{s(s-1)}{p(p-1)} ||z||_p^s \left(\alpha^p - 1\right). \tag{3.45}$$

Consider τ as in Lemma 3.4 with $\alpha = \left(1 + c_p t^p/||z||_p^p\right)^{1/p}$, $\beta = 0$. Then, by (3.44) and (3.45),

$$\begin{aligned}
\rho_3(z,t) - \rho_1(z,t) &= s||z||_p^s \left[\left(\frac{s-1}{p-1}\right)\tau(p) - \tau(s) + \tau(1)\right] \\
&\geqslant s||z||_p^s \left[\left(\frac{s-1}{p-1}\right)\tau(p) - \tau(s) + \left(\frac{p-s}{p-1}\right)\tau(1)\right],
\end{aligned} \tag{3.46}$$

because $(p-s)/(p-1) \leqslant 1$, since $s > 1$, and $\tau(1) = \alpha - 1 > 0$. τ is convex by Lemma 3.4, because $\alpha \geqslant 1$, and therefore, writing s as $\left(\frac{s-1}{p-1}\right)p + \left(\frac{p-s}{p-1}\right)$,

$$\tau(s) \leqslant \left(\frac{s-1}{p-1}\right)\tau(p) + \left(\frac{p-s}{p-1}\right)\tau(1). \tag{3.47}$$

By (3.46) and (3.47), $\rho_3(z,t) - \rho_1(z,t) \geqslant 0$ and the claim is established.

Since $\delta \geqslant 1$ and $s - p \leqslant 0$, we have that $1 \geqslant \delta^{s-p}$. Therefore, by (3.42) and (3.43),

$$\rho_3(z,t) \geqslant \rho_1(z,t) \geqslant \delta^{s-p}\rho_1(z,t), \tag{3.48}$$

$$\rho_2(z,t) \geqslant \delta^{s-p}\rho_3(z,t) \geqslant \delta^{s-p}\rho_1(z,t). \tag{3.49}$$

In view of the definition of δ , (3.41) follows from (3.48) and (3.49), so that the leftmost inequality in (3.40) holds. The rightmost one follows from positivity of $c_p t^p$ and strict convexity of φ , which implies that $D_\varphi(\widehat{t}, \bar{t}) > 0$ for $\widehat{t} \neq \bar{t}$. $\square$

COROLLARY 3.1. *Take* $p, s > 1$ *and consider* $f : \mathscr{L}^p(\Omega) \to \mathbf{R}$ *defined as* $f(x) = ||x||_p^s$. *Then* f *is totally convex.*

Proof. Follows from Propositions 3.2, 3.3 and 3.7, and the definition of total convexity. □

For the sake of completeness, we analyze now $\mu_f(z,t) := \sup\{D_f(u,z) : ||u - z||_p = t\}$, for $f(x) = ||x||_p^s$. We will prove that $\mu_f(z,t) < \infty$ for all $z \in \mathscr{L}^p(\Omega)$ and all $t \in \mathbf{R}_+$. While positivity of $\nu_f(z,t)$ says that each metric ball (i.e. a ball induced by the p-norm) contains a Bregman ball (i.e. one induced by the Bregman distance associated with $f(x) = ||x||_p^s$), finiteness of $\mu_f(z,t)$ says that each metric ball is contained in a Bregman ball. As a consequence the metric distance and the Bregman distance associated with this f are topologically equivalent, in the sense that for any sequence $\{x^k\} \subset \mathscr{L}^p(\Omega)$ and any $x \in \mathscr{L}^p(\Omega)$ it holds that $\lim_{k\to\infty} D_f(x^k, x) = 0$ if and only if $\lim_{k\to\infty} ||x^k - x||_p = 0$. Reversing the situation of ν_f, we obtain a closed formula for μ_f when $2 \leqslant p \leqslant s$ and a finite upper bound otherwise.

PROPOSITION 3.8. *i) For* $2 \leqslant p \leqslant s$, *the supremum of* $D_f(u,z)$ *over the set* $\{u \in \mathscr{L}^p(\Omega) : ||u - z|| = t\}$ *is attained at* $u^* = (1 + t/||z||_p)z$ *with value* $D_f(u^*, z) = \mu_f(z,t) = \left(t + ||z||_p\right)^s - ||z||_p^s - s||z||^{s-1}t$.

ii) For all $p, s > 1$, $\mu_f(z,t) \leqslant \left(t + ||z||_p\right)^s - ||z||_p^s - s||z||_p^{s-1}t < \infty$.

Proof. i) The argument in the proof of Proposition 3.1, with the inequalities reversed, and using (2.2) instead of (2.1) in the final step, establishes the result.

ii) By (3.4), with $x \in U_t$ and $y(\xi) = |z(\xi)|^{p-2}z(\xi)$ if $z(\xi) \neq 0$, $y(\xi) = 0$ otherwise,

$$\begin{aligned} D^{ps}(x+z,z) &= ||x+z||_p^s - ||z||_p^s - s||z||_p^{s-p}\mathrm{Re}(\langle y,x\rangle) \\ &\leqslant \left(||x||_p + ||z||_p\right)^s - ||z||_p^s + s||z||_p^{s-p}|\langle y,x\rangle| \\ &\leqslant \left(t + ||z||_p\right)^s - ||z||_p^s + s||z||_p^{s-p}||y||_q||x||_p \\ &= \left(t + ||z||_p\right)^s - ||z||_p^s + s||z||_p^{s-1}t < \infty, \end{aligned} \tag{3.50}$$

using Minkowski's inequality in the first inequality and Hölder's one in the second one. The result follows by taking supremum over $x \in U_t$ in (3.50). □

4. Appendix

We present here some estimates of c_p which provide a more explicit lower bound for ν^{ps} in terms of s, p, $||z||_p$ and t, when substituted in the bound given in Proposition 3.7.

LEMMA 4.1. *Take* $p \geqslant 2$ *and define* ϕ_p *as in Lemma 3.2,* $c_p = \inf_{z\in\mathbf{C}} \phi_p(z)$. *Then*

$$\begin{aligned} p2^{1-p} \geqslant c_p &= \inf\left\{\phi_p(x) : x \in \mathbf{R}, -1/2 \leqslant x < 0\right\} \\ &\geqslant \left[1 + (2p-1)^{-1/(p-1)}\right]^{1-p} \geqslant 2^{1-p}(2 - 1/p). \end{aligned} \tag{4.1}$$

Proof. The leftmost inequality in (4.1) has been established in Lemma 3.2. We prove first the equality. Since $\phi_2(z) = 1$ for all $z \in \mathbf{C}$, we may assume that $p > 2$. From Lemma 3.2 we have that $c_p = \min\{\phi_p(z) : z \in \mathbf{C}, |z| \leqslant 1/2\}$. Let $z = x + yi$, with x and y real. Then, from $|z|^2 = x^2 + y^2$ and $|1+z|^2 = (1+x)^2 + y^2$, we get

$$\frac{\partial}{\partial x}\left(|z|^p\right) = p|z|^{p-2}x, \qquad \frac{\partial}{\partial y}\left(|z|^p\right) = p|z|^{p-2}y,$$

$$\frac{\partial}{\partial x}\left(|1+z|^p\right) = p|1+z|^{p-2}(1+x), \qquad \frac{\partial}{\partial y}\left(|1+z|^p\right) = p|1+z|^{p-2}y.$$

As a consequence,

$$\frac{\partial}{\partial x}\phi_p(z) = p\left(|1+z|^{p-2} - |z|^{p-2}\right)(1+x) - p(p-2)|z|^{p-4}x^2, \tag{4.2}$$

$$\frac{\partial}{\partial y}\phi_p(z) = p\left(|1+z|^{p-2} - |z|^{p-2}\right)y - p(p-2)|z|^{p-4}xy. \tag{4.3}$$

At the point z where ϕ_p attains its minimum on $\mathbf{C}$ we have

$$0 = \frac{\partial}{\partial x}\phi_p(z) = \frac{\partial}{\partial y}\phi_p(z), \tag{4.4}$$

so that we get from (4.3) that either $y = 0$ or

$$|1+z|^{p-2} - |z|^{p-2} = (p-2)|z|^{p-4}x. \tag{4.5}$$

If (4.5) holds, then, substituting (4.5) in (4.2) and (4.4), we obtain

$$0 = \frac{\partial}{\partial x}\phi_p(z) = p(p-2)|z|^{p-4}x(1+x) - p(p-2)|z|^{p-4}x^2,$$

so that either $z = 0$ or $x(1+x) - x^2 = 0$. It follows that in any case $x = 0$, which, substituted in (4.5), gives $|1+z|^{p-2} = |z|^{p-2}$, implying that $|1+z|^2 = |z|^2$. Since $x = 0$, we conclude that $1 + y^2 = y^2$, which is a contradiction. Thus (4.5) cannot hold and therefore we have $y = 0$, i.e. $z = x \in \mathbf{R}$.

Suppose now that $z = x \geqslant 0$. From (4.2) and (4.4) we get

$$0 = \left[(1+x)^{p-2} - x^{p-2}\right](1+x) - (p-2)x^{p-2} = (x+1)^{p-1} - x^{p-1} - (p-1)x^{p-2},$$

which is a contradiction, because the function $x \mapsto x^{p-1}$ is strictly convex in $(0, +\infty)$. From this contradiction and Lemma 3.2 we conclude that $z = x \in [-1/2, 0)$, which proves the equality in (4.1).

We proceed to prove the second and third inequalities in (4.1). Let $x = -\xi$ with $\xi \in (0, 1/2]$. Then,

$$\phi_p(-\xi) = (1-\xi)^p - \xi^p + p\xi^{p-1}.$$

Since $p\xi^{p-1} \geqslant 2p\xi^p$, because $1 \geqslant 2\xi$, we get, defining $\alpha(\xi) = (1-\xi)^p + (2p-1)\xi^p$,

$$\phi_p(-\xi) \geqslant \alpha(\xi). \tag{4.6}$$

Observe that

$$\alpha'(\xi) = -p(1-\xi)^{p-1} + (2p-1)p\xi^{p-1}. \tag{4.7}$$

It follows easily from (4.7) that $\alpha'(\xi) \leqslant 0$ if and only if $\xi \leqslant \xi^*$, with $\xi^* = \left[1 + (2p-1)^{1/(p-1)}\right]^{-1}$. So, we get from (4.6)

$$\alpha(\xi^*) = \min\{\alpha(z) : \xi \in (0, 1/2]\} \leqslant c_p. \tag{4.8}$$

The definitions of α and ξ^* give, after some simple algebra,

$$\alpha(\xi^*) = \left[1 + (2p-1)^{-1/(p-1)}\right]^{1-p}. \tag{4.9}$$

Take r such that $(1/r) + [1/(p-1)] = 1$. Since $p - 1 > 1$, we get by Hölder's inequality applied to $\left(1, (2p-1)^{-1/(p-1)}\right) \in \mathbf{R}^2$ and $(1,1) \in \mathbf{R}^2$,

$$1 + (2p-1)^{-1/(p-1)} \leqslant \left[1 + (2p-1)^{-1}\right]^{1/(p-1)} 2^{1/r} = 2\left(\frac{p}{2p-1}\right)^{1/(p-1)}, \tag{4.10}$$

using the definition of r in the equality. By (4.9) and (4.10)

$$\alpha(\xi^*) \geqslant 2^{1-p}\left(\frac{2p-1}{p}\right) = 2^{1-p}(2 - 1/p). \tag{4.11}$$

The second and third inequalities in (4.1) follow from (4.8), (4.9) and (4.11), observing that $\phi_p(-1/2) = p2^{1-p} \geqslant c_p$. □

REFERENCES

[1] ALBER, YA. I., *Metric and generalized projection operators in Banach spaces: properties and applications*, In Theory and Applications of Nonlinear Operators of Monotone and Accretive Type (A. Kartsatos, editor), Marcel Dekker, New York (1996), 15–50.

[2] ALBER, YA. I, BURACHIK, R. S., IUSEM, A. N., *A proximal point method for nonsmooth convex optimization problems in Banach spaces*, Abstract and Applied Analysis **2** (1997), 97–120.

[3] ARAUJO, A., *The non-existence of smooth demands in general Banach spaces*, Journal of Mathematical Economics **17** (1988), 309–319.

[4] BREGMAN, L. M., *The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming*, USSR Computational Mathematics and Mathematical Physics **7** (1967), 200–217.

[5] BURACHIK. R. S., IUSEM, A. N., *A generalized proximal point algorithm for the variational inequality problem in a Hilbert space*, (to be published in SIAM Journal on Optimization).

[6] BUTNARIU, D., IUSEM, A. N., *Local moduli of convexity and their application to finding almost common points of measurable families of operators*, AMS Contemporary Mathematics **204** (1997), 61–91.

[7] BUTNARIU, D., IUSEM, A. N., *On a proximal point method for convex optimization in Banach spaces*, Numerical Functional Analysis and Optimization **18** (1997), 723–744.

[8] BUTNARIU, D., IUSEM A. N., BURACHIK, R. S., *Iterative methods of solving stochastic convex feasibility problems and applications*, Computational Optimization and Applications **15** (2000), 269–307.

[9] CENSOR, Y., DE PIERRO, A. N., ELFVING, T., HERMAN, G. T., IUSEM, A. N., *On iterative methods for linearly constrained entropy maximization*, In *Numerical Analysis and Mathematical Modelling* (A. Waculicz, editor). *Banach Center Publication Series*, Banach Center, Warsaw **24** (1990), 145–163.

[10] CENSOR, Y., ZENIOS, S., *The proximal minimization algorithm with D-functions*, Journal of Optimization Theory and Applications **73** (1992), 451–464.

[11] CHEN, G., TEBOULLE, M., *Convergence analysis of proximal-like optimization algorithm using Bregman functions*, SIAM Journal on Optimization **3** (1993), 538–543.
[12] CLARKE, F. H., *Optimization and Nonsmooth Analysis*. John Wiley, New York (1983).
[13] DE PIERRO, A. R., IUSEM, A. N., *A relaxed version of Bregman's method for convex programming*, Journal of Optimization Theory and Applications **51** (1986), 421–440.
[14] ECKSTEIN, J., *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Mathematics of Operations Research, **18** (1993), 202–226.
[15] IUSEM, A. N., *On some properties of generalized proximal point methods for quadratic and linear programming*, Journal of Optimization Theory and Applications **85** (1995), 593–612.
[16] IUSEM, A. N., ISNARD, C. A., BUTNARIU, D., *A mixed Hölder and Minkowski inequality*, Proceedings of the American Mathematical Society **127** (1999), 2405–2415.
[17] KIWIEL, K. C., *Proximal minimization methods with generalized Bregman functions*, SIAM Journal on Control and Optimization **35** (1997) 1142–1168.
[18] KIWIEL, K. C., *Free-steering relaxations methods for problems with strictly convex costs and linear constraints*, Mathematics of Operations Research **22** (1997) 326–349.
[19] MITRINOVIĆ, D. S., PEČARIĆ, J. E., FINK, A. M., *Classical and New Inequalities in Analysis*, Kluwer, Dordrecht (1993).
[20] VLADIMIROV, A. A., NESTEROV, Y. E., CHEKANOV, Y. N., *Uniformly convex functionals*, Vestnik Moskovskogo Universiteta, Series Mathematika i Kybernetika **3** (1978) 12–23.

(Received May 21, 1999)

C. A. Isnard and A. N. Iusem
Instituto de Matemática Pura e Aplicada
Estrada Dona Castorina 110
Jardim Botânico
Rio de Janeiro, RJ, CEP 22460-320
Brazil
e-mail: usp@impa.br