

## A MIXED HÖLDER AND MINKOWSKI INEQUALITY

ALFREDO N. IUSEM, CARLOS A. ISNARD, AND DAN BUTNARIU

(Communicated by Palle E. T. Jorgensen)

ABSTRACT. Hölder's inequality states that  $\|x\|_p \|y\|_q - \langle x, y \rangle \geq 0$  for any  $(x, y) \in \mathcal{L}^p(\Omega) \times \mathcal{L}^q(\Omega)$  with  $1/p + 1/q = 1$ . In the same situation we prove the following stronger chains of inequalities, where  $z = y|y|^{q-2}$ :

$$\|x\|_p \|y\|_q - \langle x, y \rangle \geq (1/p) [(\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p] \geq 0 \quad \text{if } p \in (1, 2],$$

$$0 \leq \|x\|_p \|y\|_q - \langle x, y \rangle \leq (1/p) [(\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p] \quad \text{if } p \geq 2.$$

A similar result holds for complex valued functions with  $\text{Re}(\langle x, y \rangle)$  substituting for  $\langle x, y \rangle$ . We obtain these inequalities from some stronger (though slightly more involved) ones.

### 1. INTRODUCTION

Hölder's inequality states that

$$\|x\|_p \|y\|_q - \langle x, y \rangle \geq 0$$

for all  $p, q \geq 1$  such that  $1/p + 1/q = 1$ , and all  $x \in \mathcal{L}^p(\Omega)$ ,  $y \in \mathcal{L}^q(\Omega)$ , where  $\langle \cdot, \cdot \rangle$  denotes the duality pairing.

We will prove the following stronger inequalities (for the case of  $p, q > 1$  and real valued functions):

$$\|x\|_p \|y\|_q - \langle x, y \rangle \geq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \geq 0 \quad \text{if } p \in (1, 2],$$

$$0 \leq \|x\|_p \|y\|_q - \langle x, y \rangle \leq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \quad \text{if } p \geq 2,$$

for all  $t > 0$ , where  $z = y|y|^{q-2}$ . Equivalently,  $z = \|y\|_q^{q-2} J(y)$ , where  $J : B \rightarrow B^*$  is the normalized duality mapping with  $B = \mathcal{L}^q(\Omega)$  (see, e.g., [3, p. 72]). In the case of complex valued functions, these inequalities also hold with  $\text{Re}(\langle x, y \rangle)$  substituting for  $\langle x, y \rangle$ .

We will present two proofs of these inequalities. The second one rests on the results on local moduli of convexity developed in [1]. This was our initial proof, and has some interest on its own because it relates these Hölder-Minkowski type

---

Received by the editors December 5, 1996 and, in revised form, November 13, 1997.

1991 *Mathematics Subject Classification*. Primary 46B10, 46B25, 46E30.

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

The first author's research for this paper was partially supported by CNPq grant no. 301280/86.

Work by the third author was done during his visit to the Department of Mathematics of the University of Texas at Arlington.

inequalities with the local modulus of convexity of  $f(x) = \|x\|_p^p$ . On the other hand, the first proof is direct and makes the paper selfcontained. Afterwards we give an application of our technique to inequalities between Bregman distances induced by compositions of convex functions. These are of interest because Bregman distances have been used to solve convex feasibility problems, and particularly convex optimization problems, in  $\ell_p$  and  $\mathcal{L}^p$  (see [2]). In the final remarks we present some alternative statements of the inequalities.

## 2. THE NEW INEQUALITIES

For a Banach space  $B$ ,  $B^*$  will denote its dual and  $\langle \cdot, \cdot \rangle : B \times B^* \rightarrow \mathbf{C}$  the duality pairing. Given  $p > 1$  and a measure space  $(\Omega, \mathcal{A}, \mu)$ ,  $\mathcal{L}^p(\Omega)$  is the set of measurable functions  $h : \Omega \rightarrow \mathbf{C}$  such that  $|h(x)|^p$  is integrable, and  $\|h\|_p = (\int_{\Omega} |h(x)|^p d\mu)^{1/p}$ . We will also use this notation, with the same formula, for any  $0 \neq p \in \mathbf{R}$ , though  $\|\cdot\|_p$  is not a norm for  $p < 1$  (rather it is a pseudonorm).  $\mathbf{R}_+^n$  and  $\mathbf{R}_{++}^n$  will denote the nonnegative and positive orthants of  $\mathbf{R}^n$  respectively.

We will prove the inequalities with the help of an estimation which is indeed stronger and quite interesting on its own, and will be established in Lemma 1. We will start with the case of  $\mathbf{C}^n$  provided with the  $p$ -norm, and then extend the result to  $\mathcal{L}^p(\Omega)$  through a standard limiting argument. Later on, we will comment on the difficulties confronted when a direct proof of the result in  $\mathcal{L}^p(\Omega)$  is attempted. We recall that in  $\mathbf{C}^n$ , the  $p$ -norm is defined as  $\|x\|_p = (\sum_{j=1}^n |x_j|^p)^{1/p}$ . In this case  $\langle \cdot, \cdot \rangle$  denotes the Euclidean inner product.

**Lemma 1.** *Take  $p, q > 0$  such that  $1/p + 1/q = 1$ . Take  $x, y \in \mathbf{C}^n$ , and define  $z \in \mathbf{C}^n$  as  $z_j = y_j |y_j|^{q-2}$  (with the convention that  $z_j = 0$  if and only if  $y_j = 0$ )*  
Then

$$(1) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \sup_{t>0} \left\{ \frac{1}{tp} \left[ (t\|x\|_p + \|z\|_p)^p - \|tx + z\|_p^p \right] \right\} \geq 0$$

if  $p \in (1, 2]$ ,

$$(2) \quad 0 \leq \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \inf_{t>0} \left\{ \frac{1}{tp} \left[ (t\|x\|_p + \|z\|_p)^p - \|tx + z\|_p^p \right] \right\}$$

if  $p \geq 2$ .

*Proof.* The inequalities in (1) and (2) are a consequence of Minkowski's and Hölder's inequalities respectively, so that it suffices to prove the equalities. Observe that for  $p = 2$  we get  $y = z$  and the middle expressions in (1) and (2) become independent of  $t$  and trivially equal to  $\|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle)$ , so we assume that  $p \neq 2$ . Define  $\varphi : \mathbf{R}_+ \rightarrow \mathbf{R}$  as

$$(3) \quad \varphi(t) = \frac{1}{p} \left[ (t\|x\|_p + \|z\|_p)^p - \|tx + z\|_p^p \right].$$

It is easy to compute the first derivative of  $\varphi$  by the chain rule, obtaining

$$(4) \quad \varphi'(t) = (t\|x\|_p + \|z\|_p)^{p-1} \|x\|_p - \sum_{j=1}^n \operatorname{Re} \left( |tx_j + z_j|^{p-2} (tx_j + z_j) \bar{x}_j \right).$$

We remark that  $\varphi'(t)$  is well defined for all  $t > 0$  and continuous, because

$$|tx_j + z_j|^{p-2} (tx_j + z_j) = |tx_j + z_j|^{p-1} \arg(tx_j + z_j).$$

The computation of  $\varphi''$  from (4) gives

$$\begin{aligned} \varphi''(t) &= (p-1)(t\|x\|_p + \|z\|_p)^{p-2} \|x\|_p^2 - \\ (5) \quad &\sum_{j=1}^n |tx_j + z_j|^{p-4} \{ (p-2)[\operatorname{Re}[(tx_j + z_j)\bar{x}_j]]^2 + |(tx_j + z_j)x_j|^2 \}, \end{aligned}$$

Observe that (5) holds for all  $t > 0$  excepting for some of those values of  $t$  such that  $tx_j + z_j = 0$  for some  $j$ , in the case of  $p \in (1, 2)$ , for which  $\varphi''(t) = -\infty$ . Observe also that

$$(6) \quad [\operatorname{Re}[(tx_j + z_j)\bar{x}_j]]^2 \leq |(tx_j + z_j)x_j|^2.$$

In view of (5) and (6)

$$(7) \quad \varphi''(t) \geq (p-1) \left[ (t\|x\|_p + \|z\|_p)^{p-2} \|x\|_p^2 - \sum_{j=1}^n |tx_j + z_j|^{p-2} |x_j|^2 \right] \quad \text{if } p > 2,$$

$$(8) \quad \varphi''(t) \leq (p-1) \left[ (t\|x\|_p + \|z\|_p)^{p-2} \|x\|_p^2 - \sum_{j=1}^n |tx_j + z_j|^{p-2} |x_j|^2 \right] \quad \text{if } p \in (1, 2).$$

We claim that  $\varphi$  is convex for  $p > 2$  and concave for  $p \in (1, 2)$ . Define  $v, w \in \mathbf{R}_+^n$  as  $v_j = |tx_j + z_j|^{p-2}$  (excepting for the exceptional values of  $t$  mentioned above, if any),  $w_j = |x_j|^2$ ,  $r = p/(p-2)$ ,  $s = p/2$ . Note that  $1/r + 1/s = 1$ ,  $\|x\|_p^2 = \|w\|_s$  and  $\sum_{j=1}^n |tx_j + z_j|^{p-2} x_j^2 = \langle v, w \rangle$ . Now we consider the case  $p > 2$ . It follows from (7) that

$$(9) \quad \varphi''(t) \geq (p-1) \left[ (t\|x\|_p + \|z\|_p)^{p-2} \|w\|_s - \langle v, w \rangle \right].$$

Note also that  $\|tx + z\|_p^{p-2} = \|v\|_r$ . Since  $p > 2$ ,

$$(t\|x\|_p + \|z\|_p)^{p-2} \geq \|tx + z\|_p^{p-2} = \|v\|_r,$$

using Minkowski's inequality and  $p-2 > 0$ . As a consequence, we get from (9)

$$(10) \quad \varphi''(t) \geq (p-1) [\|v\|_r \|w\|_s - \langle v, w \rangle] \geq 0,$$

using Hölder's inequality, since  $1/r + 1/s = 1$  and  $r, s > 1$ . We conclude that in this case  $\varphi$  is convex.

If  $p \in (1, 2)$ , then we get from (8)

$$(11) \quad \varphi''(t) \leq (p-1) \left[ (t\|x\|_p + \|z\|_p)^{p-2} \|w\|_s - \langle v, w \rangle \right].$$

In this case,

$$(t\|x\|_p + \|z\|_p)^{p-2} \leq \|tx + z\|_p^{p-2} = \|v\|_r,$$

using again Minkowski's inequality, but noting that now  $p-2 < 0$ . Thus, we get from (11)

$$(12) \quad \varphi''(t) \leq (p-1) [\|v\|_r \|w\|_s - \langle v, w \rangle] \leq 0,$$

using now the reverse Hölder’s inequality (see [4, p. 99]), since  $1/r + 1/s = 1$ ,  $s \in (0, 1)$ ,  $r < 0$  and  $v, w$  are nonnegative vectors. (12) holds for all values of  $t > 0$ , excepting the exceptional values where, as noted before,  $\varphi''(t) = -\infty$ . Therefore,  $\varphi''$  is always nonpositive. We conclude that in this case  $\varphi$  is concave, completing the proof of the claim.

Note that  $\varphi(0) = 0$ . By the Mean Value Theorem

$$(13) \quad \varphi(t)/t = [\varphi(t) - \varphi(0)]/t = \varphi'(s)$$

for some  $s \in (0, t)$ . Since  $\varphi'$  is nondecreasing in  $\mathbf{R}_+$  when  $p > 2$  and nonincreasing in  $\mathbf{R}_+$  when  $p \in (1, 2)$ , by (10) and (12), we get from (13)

$$(14) \quad \varphi(t)/t \geq \varphi'(0) \quad \text{if } p > 2,$$

$$(15) \quad \varphi(t)/t \leq \varphi'(0) \quad \text{if } p \in (1, 2),$$

for all  $t > 0$ . Now we observe that

$$(16) \quad \begin{aligned} \varphi'(0) &= \|z\|_p^{p-1} \|x\|_p - \sum_{j=1}^n \operatorname{Re}(|z_j|^{p-2} z_j \bar{x}_j) = \\ & \|x\|_p \|z\|_p^{p-1} - \operatorname{Re} \left[ \sum_{j=1}^n |z_j|^{p-2} z_j \bar{x}_j \right] = \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle). \end{aligned}$$

by (4). Note also that, since  $\varphi(0) = 0$ , it holds that  $\varphi'(0) = \lim_{t \rightarrow 0^+} \varphi(t)/t$ . Therefore, taking infimum over  $\mathbf{R}_{++}$  in (14) and supremum in (15), we get the result immediately from (16).  $\square$

We mention also that if we try to use the argument above directly in  $\mathcal{L}^p(\Omega)$ , we confront the problem that  $\varphi$  is differentiable but not twice differentiable for  $p \in (1, 2)$ , because the same happens with  $f(x) = \|x\|_p^p$ .

Now we present our main result.

**Theorem 1.** *Take  $p, q > 1$  such that  $1/p + 1/q = 1$ ,  $x \in \mathcal{L}^p(\Omega)$ , and  $y \in \mathcal{L}^q(\Omega)$ , and define  $z = y|y|^{q-2}$  (with the convention that  $z(\xi) = 0$  if and only if  $y(\xi) = 0$ ). Then  $z$  belongs to  $\mathcal{L}^p(\Omega)$  and*

$$(17) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \geq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \geq 0 \quad \text{if } p \in (1, 2],$$

$$(18) \quad 0 \leq \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \leq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \quad \text{if } p \geq 2.$$

*Proof.* The fact that  $z$  belongs to  $\mathcal{L}^p(\Omega)$  can be established either by direct computation or by noting that  $z = \|y\|_q^{q-1} J(y)$  where  $J : \mathcal{L}^q(\Omega) \rightarrow \mathcal{L}^p(\Omega)$  is the normalized duality mapping (see, e.g. [3, p. 72]).

Note that (17) and (18) hold for  $x, y, z \in \mathbf{C}^n$  by Lemma 1, taking  $t = 1$  in (1) and (2) respectively.

Now we consider the case of  $\mathcal{L}^p(\Omega)$  and we start with the case when  $x$  and  $y$  are simple functions, i.e. when both  $x$  and  $y$  have finite images. In this situation, there exists a partition  $\Omega = \bigcup_{i=0}^m \Omega_i$  in pairwise disjoint and measurable sets such that both  $x$  and  $y$  are constant in each  $\Omega_i$ . Let  $\alpha_i, \beta_i$  be the values of  $x, y$  respectively in  $\Omega_i$ . By convention  $\Omega_0 = \{\xi \in \Omega : x(\xi) = y(\xi) = 0\}$ , i.e.  $\alpha_0 = \beta_0 = 0$  ( $\Omega_0$  is possibly

empty). Since  $x \in \mathcal{L}^p(\Omega)$ ,  $y \in \mathcal{L}^q(\Omega)$ , it follows easily that  $\mu(\Omega_i) < \infty$  for all  $i \geq 1$ . By definition of  $z$ , we have  $\gamma_i := z(\xi) = \beta_i |\beta_i|^{q-2}$  if  $\beta_i \neq 0$ ,  $\gamma_i := z(\xi) = 0$  otherwise, for all  $\xi \in \Omega_i$ . Now we define  $\hat{x}, \hat{y}, \hat{z} \in \mathbf{C}^m$  as  $\hat{x}_i = \alpha_i \mu(\Omega_i)^{1/p}$ ,  $\hat{y}_i = \beta_i \mu(\Omega_i)^{1/q}$ ,  $\hat{z}_i = \gamma_i \mu(\Omega_i)^{1/p}$ , for  $1 \leq i \leq m$ . Then, it is clear that  $\hat{x}, \hat{y}, \hat{z}$  are exactly as requested in Lemma 1, so that (1) and (2) hold with  $\hat{x}, \hat{y}, \hat{z}$  and  $m$  substituting for  $x, y, z$  and  $n$ . Now we observe that

$$\|x\|_p = \|\hat{x}\|_p, \quad \|y\|_q = \|\hat{y}\|_q, \quad \|z\|_p = \|\hat{z}\|_p, \quad \langle x, y \rangle = \langle \hat{x}, \hat{y} \rangle,$$

and conclude that (17) and (18) hold for  $x, y$  and  $z$ .

We proceed to consider arbitrary  $x \in \mathcal{L}^p(\Omega)$ ,  $y \in \mathcal{L}^q(\Omega)$ . It is well known that for any measurable, real and nonnegative  $h \in \mathcal{L}^p(\Omega)$  there exists a sequence of simple functions  $\{h^n\} \subset \mathcal{L}^p(\Omega)$  such that for all  $\xi \in \Omega$ , it holds that

$$(19) \quad \lim_{n \rightarrow \infty} h^n(\xi) = h(\xi)$$

and

$$(20) \quad 0 \leq h^n(\xi) \leq h(\xi).$$

For a measurable and real  $h \in \mathcal{L}^p(\Omega)$ , we get a similar approximation, considering separately the positive and negative parts of  $h$ , where  $|h^n(\xi)| \leq |h(\xi)|$  substitutes for (20). Finally, for a measurable and complex  $h \in \mathcal{L}^p(\Omega)$ , we consider separately its real and imaginary parts, and get again a similar pointwise approximation, where now

$$(21) \quad |h^n(\xi)| \leq |\operatorname{Re}[h(\xi)]| + |\operatorname{Im}[h(\xi)]|$$

substitutes for (20).

We consider now sequences of simple functions  $\{x^n\} \subset \mathcal{L}^p(\Omega)$ ,  $\{y^n\} \subset \mathcal{L}^q(\Omega)$  such that (19) and (21) hold with  $h = x$ ,  $h^n = x^n$  and with  $h = y$ ,  $h^n = y^n$ . In view of (21) we can apply the Dominated Convergence Theorem, and therefore we have

$$(22) \quad \lim_{n \rightarrow \infty} \|x^n\|_p = \|x\|_p, \quad \lim_{n \rightarrow \infty} \|y^n\|_q = \|y\|_q, \quad \lim_{n \rightarrow \infty} \langle x^n, y^n \rangle = \langle x, y \rangle.$$

Let now  $z^n(\xi) = |y^n(\xi)|^{q-2} y^n(\xi)$  if  $y^n(\xi) \neq 0$  and  $z^n(\xi) = 0$  otherwise. It is easy to check that  $z^n$  belongs to  $\mathcal{L}^p(\Omega)$  for all  $n$ . Also, since

$$|z^n(\xi)| = |y^n(\xi)|^{q-1} \leq (|\operatorname{Re}[y(\xi)]| + |\operatorname{Im}[y(\xi)]|)^{q-1},$$

because  $q-1 > 0$ , we can conclude again from the Dominated Convergence Theorem that

$$(23) \quad \lim_{n \rightarrow \infty} \|z^n\|_p = \|z\|_p, \quad \lim_{n \rightarrow \infty} \|x^n + z^n\|_p = \|x + z\|_p.$$

Note that  $z^n = |y^n|^{q-2} y^n$ , so that we can apply the result already established for simple functions and conclude that (17) and (18) hold with  $x^n, y^n, z^n$  substituting for  $x, y, z$ , respectively. The result of the theorem follows by taking limits as  $n$  goes to  $\infty$  in (17) and (18) after such substitution, in view of (22) and (23).  $\square$

We remark that the result in Theorem 1 concerning the spaces  $\mathcal{L}^p(\Omega)$  is in principle weaker than the results in Lemma 1 for  $\mathbf{C}^n$  (it corresponds to the particular case of  $t = 1$ ). We will now establish (1) and (2) in  $\mathcal{L}^p(\Omega)$ . In this case we will follow an inverse procedure: we start with the already established case of  $t = 1$  (i.e.

(17)–(18)) and proceed from there to the case of the infimum or supremum over  $t$ , using the fact that  $f(x) = \|x\|_p^p$  is once (though not twice) differentiable in  $\mathcal{L}^p(\Omega)$ .

**Lemma 2.** *Take  $p, q > 0$  such that  $1/p + 1/q = 1$ . Take  $x \in \mathcal{L}^p(\Omega)$ ,  $y \in \mathcal{L}^q(\Omega)$  and define  $z = y|y|^{q-2}$  (with the same convention as in Theorem 1). Then*

$$(24) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \sup_{t>0} \left\{ \frac{1}{tp} \left[ (t\|x\|_p + \|z\|_p)^p - \|tx + z\|_p^p \right] \right\} \geq 0$$

if  $p \in (1, 2]$ ,

$$(25) \quad 0 \leq \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \inf_{t>0} \left\{ \frac{1}{tp} \left[ (t\|x\|_p + \|z\|_p)^p - \|tx + z\|_p^p \right] \right\}$$

if  $p \geq 2$ .

*Proof.* Again, it suffices to prove the equalities in (24) and (25). Consider  $\varphi$  defined as in (3), but with  $x$  and  $z$  in  $\mathcal{L}^p(\Omega)$ . Note that the equalities in (24) and (25) can be rewritten as

$$(26) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \sup_{t>0} \left\{ \frac{\varphi(t)}{t} \right\} \quad \text{if } p \in (1, 2],$$

$$(27) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) = \inf_{t>0} \left\{ \frac{\varphi(t)}{t} \right\} \quad \text{if } p \geq 2.$$

Replace now  $x$  by  $tx$ , with  $t > 0$ , in (17)–(18), use homogeneity in  $x$  of  $\|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle)$ , divide both (17) and (18) by  $t$  and conclude that

$$(28) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \geq \sup_{t>0} \left\{ \frac{\varphi(t)}{t} \right\} \quad \text{if } p \in (1, 2],$$

$$(29) \quad \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \leq \inf_{t>0} \left\{ \frac{\varphi(t)}{t} \right\} \quad \text{if } p \geq 2,$$

so that in order to establish (26) and (27) it suffices to prove that the inequalities in (28) and (29) are indeed equalities.

It is well known that  $f : \mathcal{L}^p(\Omega) \rightarrow \mathbf{R}$  defined as  $f(x) = \|x\|_p^p$  is Fréchet differentiable and that its derivative at  $x$  is the  $\mathbf{R}$ -linear (though not  $\mathbf{C}$ -linear) operator  $f'(x) : \mathcal{L}^p(\Omega) \rightarrow \mathbf{R}$  defined as

$$(30) \quad f'(x)(v) = \operatorname{Re} \left( \left\langle p|x|^{p-2}x, v \right\rangle \right)$$

(see e.g. [3, 72]). In view of (30), we can compute the derivative of  $\varphi$  by the chain rule, obtaining

$$(31) \quad \begin{aligned} \varphi'(t) &= (t\|x\|_p + \|z\|_p)^{p-1} \|x\|_p - \int_{\Omega} \operatorname{Re} \left( |tx(\xi) + z(\xi)|^{p-2} (tx(\xi) + z(\xi)) \bar{x}(\xi) \right) d\mu \\ &= (t\|x\|_p + \|z\|_p)^{p-1} \|x\|_p - \operatorname{Re} \left( \left\langle |tx + z|^{p-2} (tx + z), x \right\rangle \right), \end{aligned}$$

with  $|tx + z|^{p-2} (tx + z) \in \mathcal{L}^q(\Omega)$ ,  $x \in \mathcal{L}^p(\Omega)$ . Note that  $\varphi(0) = 0$  by (3), so that, using (31),

$$(32) \quad \lim_{t \rightarrow 0^+} \varphi(t)/t = \varphi'(0) = \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle).$$

In view of (32), the inequalities in (28) and (29) are indeed equalities, so that (26) and (27) hold and the result is established.  $\square$

It is interesting to consider (17) in the case of  $p = 1$  (i.e.  $q = \infty$ ). In this case our definition of  $z$  as  $y|y|^{q-2}$  becomes meaningless, but similar inequalities hold when  $y$  belongs to a subset of  $\mathcal{L}^q(\Omega)$ , as we discuss next. For  $p > 1$  the relation  $z = y|y|^{q-2}$  can be inverted to  $y = z|z|^{p-2}$ . For  $p = 1$  the application from  $\mathcal{L}^1(\Omega)$  to  $\mathcal{L}^\infty(\Omega)$  given by  $z \rightarrow z|z|^{p-2}$  is still well defined, but it is not onto any more: with the convention of Lemma 1, its image consists of those  $y \in \mathcal{L}^\infty(\Omega)$  such that for all  $\xi \in \Omega$  either  $y(\xi) = 0$  or  $|y(\xi)| = 1$ . In other words, defining  $y$  in terms of  $z$  through the relation just mentioned, (17) becomes  $\|x\|_1 \|y\|_\infty - \operatorname{Re}(\langle x, y \rangle) \geq [\|x\|_1 + \|z\|_1 - \|x + z\|_1] \geq 0$  with  $y = z|z|^{-1}$ . We proceed to prove this inequality.

**Proposition 1.** *Take complex valued functions  $x, z \in \mathcal{L}^1(\Omega)$  and define  $y = z|z|^{-1}$ , with the convention of Lemma 1. Then  $y \in \mathcal{L}^\infty(\Omega)$  and*

$$(33) \quad \|x\|_1 \|y\|_\infty - \operatorname{Re}(\langle x, y \rangle) \geq \|x\|_1 + \|z\|_1 - \|x + z\|_1 \geq 0.$$

*Proof.*  $y$  belongs to  $\mathcal{L}^\infty(\Omega)$  because  $|y(\xi)|$  is equal to either 1 or 0 for all  $\xi \in \Omega$ . The rightmost inequality in (33) is trivial. We proceed to prove the leftmost one. We claim that for all  $\xi \in \Omega$  it holds that

$$(34) \quad \operatorname{Re}[x(\xi)\bar{y}(\xi)] + |z(\xi)| \leq |x(\xi) + z(\xi)|.$$

If  $z(\xi) = 0$ , then we have  $y(\xi) = 0$  by definition, and (34) holds. Otherwise, let  $\beta = x(\xi)/z(\xi)$ . Clearly,  $1 + \operatorname{Re}(\beta) = \operatorname{Re}(1 + \beta) \leq |1 + \beta|$ , which, after multiplying by  $|z(\xi)|$  and rearranging terms, becomes

$$\begin{aligned} |x(\xi) + z(\xi)| &\geq |z(\xi)| + \operatorname{Re}[|z(\xi)|\beta] = |z(\xi)| + \operatorname{Re}\left[\frac{|z(\xi)|x(\xi)}{z(\xi)}\right] = \\ &|z(\xi)| + \operatorname{Re}\left[\frac{x(\xi)\bar{z}(\xi)}{|z(\xi)|}\right] = |z(\xi)| + \operatorname{Re}[x(\xi)\bar{y}(\xi)], \end{aligned}$$

establishing the claim. From (34)

$$(35) \quad |x(\xi)| + |z(\xi)| - |x(\xi) + z(\xi)| \leq |x(\xi)| - \operatorname{Re}[x(\xi)\bar{y}(\xi)] = \|y\|_\infty |x(\xi)| - \operatorname{Re}[x(\xi)\bar{y}(\xi)],$$

since  $\|y\|_\infty = 1$ , because  $|y(\xi)|$  is equal to either zero or 1 (if  $y(\xi) = 0$  for all  $\xi$ , in which case  $z(\xi) = 0$  for all  $\xi$ , then (33) is trivial). Then the result follows integrating (35) over  $\Omega$ .  $\square$

We give now an alternative proof of the chain in inequalities in (17) for the real case, based on the results of [1], which are indeed the source of our results in this paper. Given a real Banach space  $B$  and a convex and Gateaux differentiable  $f : B \rightarrow \mathbf{R}$ , we define  $D_f : B \times B \rightarrow \mathbf{R}$  as

$$(36) \quad D_f(x, y) = f(x) - f(y) - \langle f'(y), x - y \rangle,$$

where  $f'(y) \in B^*$  is the Gateaux derivative of  $f$  at  $y$ . If  $f$  is convex but not differentiable we can extend (36) to

$$(37) \quad D_f(x, y) = f(x) - f(y) - \inf\{\langle \eta, x - y \rangle : \eta \in \partial f(y)\},$$

where  $\partial f(y) \subset B^*$  denotes the set of subgradients of  $f$  at  $y$ . In both cases we have  $0 \leq D_f(x, y)$  for all  $x, y \in B$ , and  $D_f(x, x) = 0$  for all  $x \in B$ . Also,  $D_f(x, y)$  is

always finite because the subdifferential of a real convex function at any point is bounded (e.g. [5, p. 7]), so that the infimum in (37) cannot be  $-\infty$ . Finally, we remark that  $D_f(\cdot, y)$  is convex for all  $y \in B$ , because the infimum of linear functions is concave.  $D_f$  is called the *Bregman distance* associated with  $f$ , though in general  $D_f$  is not symmetric and the triangular inequality does not hold.

Next we define the *local modulus of convexity* of  $f$ ,  $\nu_f : B \times \mathbf{R}_+ \rightarrow \mathbf{R}_+$  as

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

This notion was introduced in [1], where the following result has been proved for real functions:

**Proposition 2.** *Take  $p \in (1, 2]$ ,  $B = \mathcal{L}^p(\Omega)$ ,  $f(x) = \|x\|_p^p$ . For all  $t > 0$  and all  $z \in B$ ,  $z \neq 0$ , the infimum of  $D_f(u, z)$  subject to  $\|u - z\|_p = t$  is attained at  $u^* = (1 + t/\|z\|_p)z$  with value  $\nu_f(z, t) = D_f(u^*, z) = (\|z\|_p + t)^p - \|z\|_p^p - p\|z\|_p^{p-1}t$ .*

We present now the alternative proof of the chain of inequalities in (17) based on Proposition 2.

*Alternative proof of (17) for real functions.* The rightmost inequality in (17) follows from Minkowski's inequality, and so it suffices to establish the leftmost one. Note that  $z = 0$  if and only if  $y = 0$  and that the result is trivial if either  $x$  or  $y$  are zero. We assume therefore that  $x \neq 0 \neq z$ . We prove that  $z$  belongs to  $\mathcal{L}^p(\Omega)$  with the same argument as in Theorem 1. It is easy to check that

$$(38) \quad \|z\|_p^{p-1} = \|y\|_q,$$

either by direct computation or using the properties of the normalized duality mapping  $J$  (e.g. [3, p. 27]). Take  $f(x) = \|x\|_p^p$ ,  $t = \|x\|_p > 0$ ,  $u = z + x$ ,  $u^* = (1 + t/\|z\|_p)z$ . Then  $\|u - z\|_p = t$ , and, by Proposition 1, we have  $D_f(u, z) \geq D_f(u^*, z) = (\|x\|_p + \|z\|_p)^p - \|z\|_p^p - p\|z\|_p^{p-1}\|x\|_p$ , which becomes, after expanding  $D_f(u, z)$ ,

$$(39) \quad \|x + z\|_p^p - \|z\|_p^p - p\langle x, y \rangle \geq (\|x\|_p + \|z\|_p)^p - \|z\|_p^p - p\|z\|_p^{p-1}\|x\|_p.$$

By (38) and (39),

$$(40) \quad \begin{aligned} \|x + z\|_p^p - p\langle x, y \rangle &\geq (\|x\|_p + \|z\|_p)^p - p\|x\|_p\|z\|_p^{p-1} \\ &= (\|x\|_p + \|z\|_p)^p - p\|x\|_p\|y\|_q \end{aligned}$$

and (17) follows easily from (40).  $\square$

### 3. AN APPLICATION TO INEQUALITIES BETWEEN BREGMAN DISTANCES

As an application, we discuss now some results on the behavior of  $D_f$  when  $f$  is a composition of two functions. First we present a general result of some interest on its own.

**Proposition 3.** *Let  $B$  be a Banach space,  $g : B \rightarrow \mathbf{R}$  a convex function and  $\phi : \mathbf{R} \rightarrow \mathbf{R}$  a convex, differentiable and nondecreasing function. Let  $f(x) = \phi(g(x))$ . Then  $f$  is convex and*

$$D_f(x, y) \geq D_\phi(g(x), g(y))$$

for all  $x, y \in B$ .

*Proof.* It is easy to get convexity of  $f$  directly from the definition of convexity. Also, it follows from the definition of subgradient that for any  $z \in B$ ,  $\omega \in \partial f(z)$  if and only if  $\omega = \phi'(g(z))\eta$  for some  $\eta \in \partial g(z)$ . Then

$$D_f(x, y) = \phi(g(x)) - \phi(g(y)) - \inf\{\phi'(g(y)) \langle \eta, x - y \rangle : \eta \in \partial g(y)\} =$$

$$(41) \quad \phi(g(x)) - \phi(g(y)) - \phi'(g(y)) \inf\{\langle \eta, x - y \rangle : \eta \in \partial g(y)\},$$

using nonnegativity of  $\phi'$ . From (41),

$$D_f(x, y) = \phi(g(x)) - \phi(g(y)) - \phi'(g(y))[g(x) - g(y)] +$$

$$\phi'(g(y))[g(x) - g(y) - \inf\{\langle \eta, x - y \rangle : \eta \in \partial g(y)\}] =$$

$$D_\phi(g(x), g(y)) + \phi'(g(y))D_g(x, y) \geq D_\phi(g(x), g(y)),$$

using nonnegativity of  $\phi'$  and  $D_g$ . □

We have the following immediate corollary of Proposition 3.

**Corollary 1.** *Under the hypotheses of Proposition 3,*

$$D_f(z, x + z) \geq D_\phi(g(z), g(x + z))$$

for all  $x, z \in B$ .

We look now at the relation between  $D_f(z, x + z)$  and  $D_\phi(g(z), g(x) + g(z))$ , which is more interesting, in a certain sense, than the relation between  $D_f(z, x + z)$  and  $D_\phi(g(z), g(x + z))$  established in Corollary 1, because  $D_\phi(g(z), g(x + z))$  vanishes if  $g(z) = g(x + z)$ , while  $D_\phi(g(z), g(x) + g(z)) > 0$  whenever  $g(x) \neq 0$  and  $\phi$  is strictly convex, as a consequence of (36). We will analyze the case of  $B = \mathbf{R}^n$ ,  $f(x) = \|x\|_p^p$  and  $\phi(t) = t^p$  with the help of the technique developed in Lemma 1.

**Lemma 3.** *Let  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ ,  $\phi : \mathbf{R}_+ \rightarrow \mathbf{R}$  be defined as  $f(x) = \|x\|_p^p$ ,  $\phi(t) = t^p$  with  $p > 1$ . Then, for all  $x, z \in \mathbf{R}^n$ ,*

$$D_f(z, x + z) \leq D_\phi(\|z\|_p, \|x\|_p + \|z\|_p) \quad \text{if } p \in (1, 2),$$

$$D_f(z, x + z) \geq D_\phi(\|z\|_p, \|x\|_p + \|z\|_p) \quad \text{if } p > 2.$$

*Proof.* Let  $g(x) = \|x\|_p$ . Then  $g$  is convex,  $\phi$  is convex, nondecreasing and differentiable, and we have

$$D_f(z, x + z) - D_\phi(g(z), g(x) + g(z))$$

$$= \|z\|_p^p - \|x + z\|_p^p + p \sum_{j=1}^n |x_j + z_j|^{p-2} (x_j + z_j)x_j -$$

$$(42) \quad \|z\|_p^p + (\|x\|_p + \|z\|_p)^p - p(\|x\|_p + \|z\|_p)^{p-1} \|x\|_p$$

$$= p[\varphi(1) - \varphi'(1)] = p[\varphi'(s) - \varphi'(1)]$$

for some  $s \in (0, 1)$ , with the notation of Lemma 1, using (3), (4) and (13). Since it has been proved in Lemma 1 that  $\varphi'$  is nondecreasing for  $p > 2$  and nonincreasing for  $p \in (1, 2)$ , the result follows from (42). □

4. FINAL REMARKS

1. Clearly, all our results hold in  $\mathbf{R}^n$  or  $\mathbf{C}^n$  with the  $p$ -norm, or in the Banach spaces  $\ell_p$  of summable sequences, since these spaces can be seen as particular cases of  $\mathcal{L}^p(\Omega)$  when  $\Omega$  is the set of the first  $n$  natural numbers, or the set of all of them, and all singletons have unit measure. In the case of the  $\mathcal{L}^p(\Omega)$  with real functions, (17) and (18) take the form:

$$\|x\|_p \|y\|_q - \langle x, y \rangle \geq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \geq 0 \quad \text{if } p \in (1, 2],$$

$$0 \leq \|x\|_p \|y\|_q - \langle x, y \rangle \leq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \quad \text{if } p \geq 2.$$

2. (17) and (18) can be condensed in the following single chain of inequalities, assuming, without loss of generality, that  $p \leq q$ :

$$\begin{aligned} 0 &\leq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \leq \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \\ &\leq \frac{1}{q} \left[ (\|y\|_q + \|z'\|_q)^q - \|y + z'\|_q^q \right], \end{aligned}$$

with  $z = y|y|^{q-2}$ ,  $z' = x|x|^{p-2}$ .

3. (17) and (18) can be rewritten exclusively in terms of  $x$  and  $y$ , using the normalized duality mapping  $J$ , as:

$$\|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \geq \frac{1}{p} \left[ (\|x\|_p + \|y\|_q^{q-1})^p - \|x + \|y\|_q^{q-2} J(y)\|_p^p \right] \geq 0$$

if  $p \in (1, 2)$ ,

$$0 \leq \|x\|_p \|y\|_q - \operatorname{Re}(\langle x, y \rangle) \leq \frac{1}{p} \left[ (\|x\|_p + \|y\|_q^{q-1})^p - \|x + \|y\|_q^{q-2} J(y)\|_p^p \right] \geq 0$$

if  $p > 2$ .

Alternatively, we can write (17) and (18) exclusively in terms of  $x$  and  $z$  as:

$$(43) \quad \|x\|_p \|z\|_p^{p-1} - \operatorname{Re}(\langle x, |z|^{p-2} z \rangle) \geq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right] \geq 0$$

if  $p \in (1, 2]$ ,

$$(44) \quad 0 \leq \|x\|_p \|z\|_p^{p-1} - \operatorname{Re}(\langle x, |z|^{p-2} z \rangle) \leq \frac{1}{p} \left[ (\|x\|_p + \|z\|_p)^p - \|x + z\|_p^p \right]$$

if  $p \geq 2$ .

4. For  $p = 2$  we get  $y = z$  so that the leftmost inequality in (17) and the rightmost one in (18) become trivial equalities, and our result gives no additional information in Hilbert spaces. For  $p \neq 2$  the inequalities in (17) and (18) are strict unless  $x = \delta z$  or  $z = \delta x$  with  $\delta \geq 0$ , in which case all expressions in (17)–(18) vanish.

5. For a complex function  $h \in \mathcal{L}^p(\Omega)$  the following inequalities hold:

$$\begin{aligned} \min \{ & \|\operatorname{Re}(h)\|_p \|\operatorname{Im}(h)\|_p^{p-1}, \|\operatorname{Re}(h)\|_p^{p-1} \|\operatorname{Im}(h)\|_p \} \geq \\ & \frac{1}{p} [(\|\operatorname{Re}(h)\|_p + \|\operatorname{Im}(h)\|_p)^p - \|h\|_p^p] \quad \text{if } p \in (1, 2], \\ \max \{ & \|\operatorname{Re}(h)\|_p \|\operatorname{Im}(h)\|_p^{p-1}, \|\operatorname{Re}(h)\|_p^{p-1} \|\operatorname{Im}(h)\|_p \} \leq \\ & \frac{1}{p} [(\|\operatorname{Re}(h)\|_p + \|\operatorname{Im}(h)\|_p)^p - \|h\|_p^p] \quad \text{if } p \geq 2. \end{aligned}$$

These inequalities follow from (43) and (44) respectively, taking  $x = \operatorname{Re}(h)$ ,  $z = i \operatorname{Im}(h)$ , and then  $x = i \operatorname{Im}(h)$ ,  $z = \operatorname{Re}(h)$ , because in such cases  $\operatorname{Re}(\langle x, y \rangle) = 0$ .

6. The proof of Proposition 1 given in [1], after reversing the appropriate inequalities, can be used to prove that for  $p > 2$  and  $f(x) = \|x\|_p^p$ , the supremum of  $D_f(u, z)$  subject to  $\|u - z\|_p = t$  is attained at  $u^* = (1 + t/\|z\|_p)z$  with value

$$D_f(u^*, z) = (\|z\|_p + t)^p - \|z\|_p^p - p \|z\|_p^{p-1} t,$$

enabling us to prove (18) with the same argument as in the alternative proof of (17) given above. The result just mentioned is not explicitly presented in [1], which is devoted to the properties of  $\nu_f$ , defined as the infimum rather than the supremum.

7. The result of Lemma 3 holds indeed for  $f : \mathcal{L}^p(\Omega) \rightarrow \mathbf{R}$  defined as  $f(x) = \|x\|_p^p$ . We have presented it in its finite dimensional version to fit the statement of Lemma 1.

#### REFERENCES

- [1] Butnariu, D., Iusem, A.N. Local moduli of convexity and their application to finding almost common points of measurable families of operators *Recent Developments in Optimization Theory and Nonlinear Analysis*, edited by Y. Censor and S. Reich, AMS Contemporary Mathematics Series, vol. 204, 1997, 61–91. CMP 97:11
- [2] Butnariu, D., Iusem A.N., Burachik, R. Iterative methods of solving stochastic convex feasibility problems and applications (to be published).
- [3] Cioranescu, I. *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*, Kluwer, Dordrecht (1990). MR **91m**:46021
- [4] Mitrinović, D.S., Pečarić, J.E., Fink, A.M. *Classical and New Inequalities in Analysis*, Kluwer, Dordrecht (1993). MR **94c**:00004
- [5] Phelps, R.R. *Convex Functions, Monotone Operators and Differentiability*. Springer Verlag, Berlin (1993). MR **94f**:46055

INSTITUTO DE MATEMÁTICA PURA E APLICADA, ESTRADA DONA CASTORINA 110, JARDIM BOTÂNICO, RIO DE JANEIRO, RJ, CEP 22460-320, BRAZIL

*E-mail address:* `iusp@impa.br`

UNIVERSITY OF HAIFA, DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, MOUNT CARMEL, 31905 HAIFA, ISRAEL

*E-mail address:* `dbutnaru@mathcs2.haifa.ac.il`