ON THE BEST CONSTANT IN THE L^p ESTIMATE FOR THE SHARP MAXIMAL FUNCTION

ŁUKASZ KAMIŃSKI AND ADAM OSĘKOWSKI

Abstract. Let I be a fixed subinterval of $\mathbb{R}$ and let $1 \leq p < \infty$ be a fixed exponent. Suppose that $f \in L^p(I)$ is an arbitrary function of integral zero and let $f^{\#}$ be the BLO-based sharp maximal function of f. The paper contains the identification of the best constant C_p in the estimate

$$\|f\|_{L^p} \leq C_p \|f^{\#}\|_{L^p}.$$

The proof rests on the explicit evaluation of the Bellman function associated with the above estimate.

1. Introduction

The paper is devoted to the study of the best L^p-constant for the sharp maximal function, an important object arising in analysis and interpolation theory. We start the discussion on the subject presenting the necessary background and notation. A locally integrable function $f : \mathbb{R} \to \mathbb{R}$ belongs to the space BMO, the class of functions of bounded mean oscillation, if the quantity

$$\|f\|_{BMO} := \sup_I \left\langle |f - \langle f \rangle_I| \right\rangle_I, \tag{1.1}$$

where the supremum is taken over all intervals $I \subset \mathbb{R}$, is finite. Here $\langle f \rangle_I = \frac{1}{|I|}\int_I f$ denotes the average of f over I, calculated with respect to the Lebesgue measure. The space BMO was introduced in the paper [10] by John and Nirenberg, and it has proved to be a fundamental object in analysis and probability theory. Let us mention its two basic properties, for a more systematic overview we refer the reader to the monographs [8, 9, 11]. First, in some contexts, the class BMO can be regarded as a convenient substitute for the space L^∞. Namely, many important operators (e.g. Calderón-Zygmund singular integrals) are bounded on L^p if $1 < p < \infty$; for $p = \infty$ this boundedness fails, but the operators map L^∞ into BMO. The second important property is that BMO provides insight into the structure of Hardy spaces: as Fefferman proved in [7], we have the duality $BMO = (H^1)^*$.

There is a related, smaller class of functions which will be crucial for our further considerations. A function f is said to have bounded lower oscillation, if

$$\|f\|_{BLO} := \sup_I \left[\langle f \rangle_I - \operatorname*{essinf}_I f\right] < \infty, \tag{1.2}$$

where the supremum is as before. In comparison to (1.1), we see that the term $\langle f \rangle_I$ has been replaced with $\operatorname{essinf}_I f$, the essential infimum of f over I. This class first appeared in the paper of Coifman and Rochberg [6], who used it to prove the following decomposition property of BMO: any function of bounded

2010 *Mathematics Subject Classification.* Primary: 42A05. Secondary: 49K20.

Key words and phrases. sharp function; bounded lower oscillation; best constant.

mean oscillation can be written as a difference of two BLO functions. In addition, the BLO class arises naturally while studying the action of the Hardy-Littlewood maximal operator $\mathcal{M}$ on BMO spaces; for instance, Bennett [2] proved that $f \in BLO$ if and only if it is of the form $\mathcal{M}F + h$, where F is a function of bounded mean oscillation satisfying $\mathcal{M}F < \infty$ almost everywhere, and h is bounded. See also Korenovskii [12] for a variety of related results in this direction.

We proceed to the introduction of another important notion. Given a locally integrable function f on $\mathbb{R}$, its sharp maximal function $f^{\#} : \mathbb{R} \to \mathbb{R}_+$ is defined by

$$f^{\#}(x) = \sup \big\langle |f - \langle f\rangle_I| \big\rangle_I,$$

where the supremum is taken over all intervals $I \subset \mathbb{R}$ containing x. There is an evident relation of this object to the BMO seminorm (1.1): indeed, we have $f \in BMO$ if and only if $f^{\#} \in L^\infty$. The sharp maximal function arises in the interpolation theory and one of its key properties is the validity of the estimate

$$\|f\|_{L^p(\mathbb{R})} \leq K_p \|f^{\#}\|_{L^p(\mathbb{R})}, \qquad 1 < p < \infty. \tag{1.3}$$

One can ask about the optimal value of the constant K_p above, and this is the main theme of this paper. We will study the version of this problem in which the sharp maximal function is defined a little differently. Namely, motivated by the notion (1.2) of the space BLO, we let

$$\widetilde{f^{\#}}(x) = \sup_I \Big(\langle f\rangle_I - \operatorname*{essinf}_I f \Big),$$

where, as previously, the supremum is taken over all intervals $I \subset \mathbb{R}$ which contain x. To the best of our knowledge, this object has not been studied in the literature before. The sharp maximal function $f^{\#}$ is dominated pointwise by $\widetilde{f^{\#}}$, up to a multiplicative constant 2: indeed, for any I we have

$$\big\langle |f - \langle f\rangle_I| \big\rangle_I = \frac{2}{|I|}\int_I \big(f - \langle f\rangle_I\big)_+ \leq \frac{2}{|I|}\int_I \big(f - \operatorname*{essinf}_I f\big)_+ = 2\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big).$$

Therefore, the estimate (1.3) yields the related bound

$$\|f\|_{L^p(\mathbb{R})} \leq 2K_p \big\|\widetilde{f^{\#}}\big\|_{L^p(\mathbb{R})}, \qquad 1 < p < \infty.$$

All the objects discussed above - BMO/BLO spaces and the sharp maximal functions - can be considered in the localized setting, in which the functions are defined on a half-line or an arbitrary interval. This requires only a minor modification in the definitions: namely, one needs to take the corresponding suprema over all intervals I contained in the new base space. Most of our considerations below will concern the case in which the functions are given on a fixed interval: the restriction to such localized setting will enable the efficient study of best-constant inequalities.

We are ready for the formulation of our main result. For $1 \leq p < \infty$, introduce the constant

$$C_p = \begin{cases} \left(\int_0^\infty e^{-u} |u-1|^p \mathrm{d}u \right)^{1/p} & \text{if } 1 \leq p < 2, \\ p - 1 & \text{if } p \geq 2. \end{cases}$$

Theorem 1.1. *If $1 \leq p < \infty$, I is an arbitrary interval and $f \in L^p(I)$ is a function of integral zero, then we have*

$$\|f\|_{L^p(I)} \leq C_p \left\| \widetilde{f^\#} \right\|_{L^p(I)}. \tag{1.4}$$

The constant C_p is the best possible.

Interestingly, the reverse bound does not hold with any finite constant, which can be seen by the following simple example. Take an $\varepsilon \in (0, 1/2)$ and consider the function $f = \chi_{[0,\varepsilon]} - \chi_{(\varepsilon,2\varepsilon]}$. Then f has integral zero, satisfies $\|f\|_{L^p([0,1])} = (2\varepsilon)^{1/p}$ and for any $x \in [0, 1]$ we have $\widetilde{f^\#}(x) \geq \langle f \rangle_{[0,1]} - \operatorname{essinf}_{[0,1]} f = 1$. This gives $\|\widetilde{f^\#}\|_{L^p([0,1])}/\|f\|_{L^p([0,1])} \geq (2\varepsilon)^{-1/p} \to \infty$ as $\varepsilon \to 0$. This should be compared to the inequality (1.3), the reverse to which does hold true (the maximal function $f^\#$ is controlled by the Hardy-Littlewood maximal function: see [9]).

By a standard restriction/stretching argument (see [20], for example), we see that (1.4) is true if we replace the base interval I by an arbitrary half-line contained in $\mathbb{R}$, or the real line $\mathbb{R}$ itself. However, then it is not clear to us whether the constant C_p is still optimal: the enlargement of the underlying base space might increase $\widetilde{f^\#}$ significantly. There seem to be no efficient transference arguments which would cover the above setting. To the best of our knowledge, the only related results have been obtained very recently in [18] in the context of BMO estimates, and the methods developed there do not apply in our setting. This is an interesting topic for the further research.

As an application of the above statement, one obtains the estimate

$$\|f\|_{L^p(I)} \leq C_p \|f\|_{BLO(I)}, \qquad 1 \leq p < \infty. \tag{1.5}$$

This inequality is sharp in the case $1 \leq p \leq 2$. For $p > 2$, one can show that the best constant is given by $\left(\int_0^\infty e^{-u}|u-1|^p \mathrm{d}u\right)^{1/p}$ (which is the same expression as in the previous case), the details will appear elsewhere. See also the first half of Section 3. As another application, one can deduce a version of (1.3) for nonincreasing functions on $\mathbb{R}_+$. It is not difficult to check that for such an f we have $\widetilde{f^\#}(x) \leq 2f^\#(2x)$ (see [12], p. 58), which gives

$$\|f\|_{L^p(\mathbb{R}_+)} \leq 2^{1-1/p} C_p \|f^\#\|_{L^p(\mathbb{R}_+)}.$$

This is probably not sharp, but we believe that the constant is not far from optimal.

A few words about our approach are in order. We will exploit the so-called Bellman function method, a well-known tool used widely in analysis and probability theory for the investigation of various extremal problems. Roughly speaking, the technique allows to deduce a given estimate from the existence of a certain special function, enjoying appropriate size and concavity requirements. A nice feature of the approach is that the implication can be reversed in many settings, i.e., if we know a priori that the estimate holds, then it can be proved with the use of Bellman functions. Thus, in particular, the technique allows the search of optimal constants involved. Furthermore, the explicit formula for the Bellman function carries a lot of additional information, e.g. it provides the insight into the structure of extremizers, that is, functions for which the equality is attained, or almost (asymptotically) attained. From the historical point of view, the roots of the approach go back to the classical works of Bellman [1] in the 60s, concerning the theory of optimal control. The first applications of the method in the harmonic analysis and probability theory are due to Burkholder, who proved sharp estimates for the Haar system and martingale transforms (cf. [3]). Burkholder's line of research was continued in many papers on semimartingale theory, see e.g. [4, 5, 19, 21], consult also the monograph [16] and the references therein. A significant extension of the Bellman function method was obtained in the series of subsequent works of Nazarov, Treil and Volberg in the mid-90's (see [14, 15]), who proved that the approach can be applied in a much wider context, allowing a successful study of various problems in harmonic analysis. Since then, the technique has been implemented in the exploration of weighted theory, the properties of BMO spaces, and many more.

In the next section we describe the Bellman setup which links the estimate (1.4) with an appropriate abstract special function. This object - the associated Bellman function - is explicitly evaluated in Sections 3 and 4. The final part of the paper contains some informal steps which has led us to the discovery of the explicit formula.

2. The associated Bellman function

Suppose that $1 \leq p < \infty$ is a fixed parameter. The estimate (1.4) can be rewritten in the slightly more complicated form

$$\sup\left\{\langle |f|^p\rangle_I \,:\, \left\langle \left(\widetilde{f^\#}\right)^p\right\rangle_I \leq w\right\} \leq C_p^p w.$$

In a sense, the Bellman function is defined as the expression on the left. However, to enable a certain iterative argumentation, one needs to introduce a few additional parameters which control the self-similarity of the problem. To this end, we begin by the identification of the "processes" which appear, explicitly or implicitly, in the above estimate. Obviously, there are three of them: the average, the essential infimum and the sharp maximal function of f. We assign the variables x, y and z to these objects, viewing them as the starting positions of the processes. Precisely, consider the function $\mathfrak{B}^{(p)}$ given by the formula

$$\mathfrak{B}^{(p)}(x,y,z,w) = \sup\left\{\langle |f|^p\rangle_I \,:\, \langle f\rangle_I = x,\, \operatorname*{essinf}_I f = y,\, \left\langle \left(\widetilde{f^\#}\vee z\right)^p\right\rangle_I \leq w\right\}$$

(here and below, we use the notation $a \vee b = \max\{a,b\}$). As we will see later, the interval I will be subject to consecutive divisions into finer and finer families of subintervals, which will translate into appropriate evolution of the parameters x, y, z and w, which, in turn, will be controlled by appropriate concavity of $\mathfrak{B}^{(p)}$.

Formally, $\mathfrak{B}^{(p)}$ is defined for the set of all quadruples $(x,y,z,w) \in \mathbb{R}^4$ satisfying $x \geq y$, $z \geq 0$ and $w \geq z^p$. The relation of this object to the desired inequality (1.4) is evident: having identified the explicit formula for the function $\mathfrak{B}^{(p)}$, we see that the optimal constant C_p is equal to $\sup\left\{\mathfrak{B}^{(p)}(0,y,0,w)/w \,:\, y \leq 0,\, w > 0\right\}$. Obviously, the discovery of the explicit formula for $\mathfrak{B}^{(p)}$ brings much more information about the estimate (1.4), as it encodes the optimal interplay between the L^p norms of f and $\widetilde{f^\#}$ under additional restrictions on these functions. Note that the definition of $\mathfrak{B}^{(p)}$ does not depend on I, which can be easily seen by performing a standard affine transformation argument.

In what follows, we will implement a certain reduction trick, which goes back to [17]. The Bellman function above involves four variables, which makes it quite challenging to handle. To simplify the analysis, we will consider a slightly different definition, which allows pulling the L^p norm of the sharp maximal function under the maximized integral and removes the variable w. Specifically, fix some constant C, distinguish the domain $\Omega = \{(x,y,z)\in\mathbb{R}^3 \,:\, x\geq y\}$ and consider the function $\mathbb{B}^{(p,C)}:\Omega\to\mathbb{R}$ given by

$$\mathbb{B}^{(p,C)}(x,y,z) = \sup\left\{\langle |f|^p\rangle_I - C^p\left\langle\left(\widetilde{f^\#}\vee z\right)^p\right\rangle_I \,:\, \langle f\rangle_I = x,\ \operatorname*{essinf}_I f = y\right\}. \tag{2.1}$$

As previously, this object is intimately related to (1.4): we need to find the smallest C for which the associated function $\mathbb{B}^{(p,C)}$ satisfies $\mathbb{B}^{(p,C)}(0,y,0)\leq 0$ for all $y\leq 0$. Clearly, the new function is a simpler object, as the number of variables is reduced to three. There are two simple properties of the above Bellman function which follow directly from the definition. First, as before, the function $\mathbb{B}^{(p,C)}$ does not depend on the choice of the underlying interval I. Second, since $\widetilde{f^\#}\geq \langle f\rangle_I - \operatorname{essinf}_I f = x-y$, we obtain

$$\mathbb{B}^{(p,C)}(x,y,z) = \mathbb{B}^{(p,C)}\big(x,y,(x-y)\vee z\big),$$

so all the essential information about the function $\mathbb{B}^{(p,C)}$ is encoded in the domain

$$\Omega_+ = \{(x,y,z)\in\mathbb{R}^3 \,:\, z\geq x-y\geq 0\}, \tag{2.2}$$

strictly contained in Ω.

The main result of this paper is the identification of the optimal constant C_p in (1.4) and the evaluation of the Bellman function $\mathbb{B}^{(p,C_p)}$. Introduce a family of auxiliary functions $b^{(p)}:\mathbb{R}\to\mathbb{R}$ as follows: for $1\leq p\leq 2$, put

$$b^{(p)}(s) = e^s\int_s^\infty e^{-u}|u-1|^p\mathrm{d}u,$$

while for $p>2$, we let

$$b^{(p)}(s) = \begin{cases} e^s\displaystyle\int_s^p e^{-u}|u-1|^p\mathrm{d}u + (p^2+p-1)(p-1)^{p-1}e^{s-p} & \text{if } s<p,\\ (s-1)^p + p(p-1)(s-1)^{p-1} + p(p-1)^{p-1} & \text{if } s\geq p.\end{cases}$$

Next, consider $B^{(p)}:\Omega_+\to\mathbb{R}$ given by the formula

$$B^{(p)}(x,y,z) = z^p\left[\frac{x-y}{z}b^{(p)}\left(\frac{y}{z}+1\right) + \left(1-\frac{x-y}{z}\right)\left|\frac{y}{z}\right|^p - C_p^p\right].$$

We will establish the following statement.

Theorem 2.1. *For any $1\leq p<\infty$ and $(x,y,z)\in\Omega_+$, we have*

$$\mathbb{B}^{(p,C_p)}(x,y,z) = B^{(p)}(x,y,z).$$

Let us briefly discuss the relation between this statement and the validity of Theorem 1.1. It will be proven in Lemma 4.1 below that for fixed x and z, the function $B^{(p)}$ nonincreasing as a function of $y\in[x-z,x]$. Consequently, if we establish Theorem 2.1, we will obtain $\mathbb{B}^{(p,C_p)}(0,y,z)\leq \mathbb{B}^{(p,C_p)}(0,-z,z)=0$ for all $y<0$ and $z\geq 0$. Thus by the very definition of $\mathbb{B}^{(p,C_p)}$, for an arbitrary function $f:I\to\mathbb{R}$ of integral zero, we will get

$$\int_I |f|^p - C_p^p\int_I (\widetilde{f^\#})^p\leq 0,$$

which is (1.4). To see that the constant C_p is the best possible, we will show in the next section that the supremum defining $\mathbb{B}^{(p,C_p)}(0,-1,1)$ is actually attained at a nontrivial function. More precisely, there is a function $f:[0,1]\to\mathbb{R}$ satisfying

$\langle f\rangle_{[0,1]}=0$, $\operatorname{essinf}_{[0,1]} f=-1$ and $\|f^\#\|_{L^p(0,1)}\neq 0$, for which $\int_0^1|f|^p-C_p^p\int_0^1(\widetilde{f^\#})^p=0$. This will obviously show that the constant C_p cannot be decreased.

To simplify the notation, from now on we assume that p is a fixed parameter and we skip the upper indices, writing $\mathbb{B}$, B and b instead of $\mathbb{B}^{(p,C_p)}$, $B^{(p)}$ and $b^{(p)}$.

3. Proof of the lower bound $\mathbb{B}\geq B$

In this section we establish a simpler half of Theorem 2.1: we will exhibit extremal examples which yield the pointwise inequality $\mathbb{B}\geq B$. As we mentioned in the previous section, the definition of the Bellman function $\mathbb{B}$ does not depend on the interval I, so it is enough to construct appropriate examples on $(0,1]$. At the first glance, the formulas below might look a bit mysterious; in Section 5 we will present a detailed and elementary explanation showing the origins of these objects.

Fix x, y, $z\in\mathbb{R}$ satisfying $z\geq x-y\geq 0$. If $x=y$, then the constant function $f\equiv x$ gives the desired bound: we have $\langle f\rangle_{(0,1]}=x$ and $\operatorname{essinf}_{(0,1]} f=y$, so

$$\mathbb{B}(x,y,z)\geq\int_0^1\big(|f(s)|^p-C_p^p(\widetilde{f^\#}(s)\vee z)^p\big)\mathrm{d}s=|x|^p-C_p^pz^p=B(x,y,z).$$

Hence, from now on, we may assume that $x>y$ and $z>0$. Let us first analyze the case $1<p\leq 2$, in which the calculations are a bit simpler. Consider the function $f:(0,1]\to\mathbb{R}$ given by

$$f(s)=\max\left\{y-z\ln\frac{sz}{x-y},y\right\}=\begin{cases}y-z\ln\dfrac{sz}{x-y} & \text{if } 0<s<\dfrac{x-y}{z},\\ y & \text{if } \dfrac{x-y}{z}\leq s\leq 1.\end{cases}\tag{3.1}$$

This function is nonincreasing, which greatly simplifies the analysis of $\widetilde{f^\#}$. Indeed, for any $s\in(0,1]$ we have

$$\begin{aligned}\widetilde{f^\#}(s)&=\sup\left\{\frac{1}{b-a}\int_a^b f(u)\mathrm{d}u-f(b)\,:\,s\in(a,b)\subset(0,1]\right\}\\&=\sup\left\{\frac{1}{b}\int_0^b f(u)\mathrm{d}u-f(b)\,:\,b\in(0,1]\right\}.\end{aligned}$$

However, we easily compute directly that

$$\frac{1}{b}\int_0^b f(u)\mathrm{d}u-f(b)=\begin{cases}z & \text{if } 0<b<\dfrac{x-y}{z},\\ \dfrac{x-y}{b} & \text{if } \dfrac{x-y}{z}\leq b<1\end{cases}$$

does not exceed z for any b. This shows that $\widetilde{f^\#}(s)\vee z=z$ on $(0,1]$; furthermore, $\langle f\rangle_{(0,1]}=\int_0^1 f-f(1)+f(1)=x$, by the above calculation. Since $\operatorname{essinf}_{(0,1]} f=y$,

$$\begin{aligned}\mathbb{B}(x,y,z)&\geq\int_0^1\big(|f(s)|^p-C_p^p(\widetilde{f^\#}(s)\vee z)^p\big)\mathrm{d}s\\&=\int_0^{(x-y)/z}\left|y-z\ln\frac{sz}{x-y}\right|^p\mathrm{d}s+\int_{(x-y)/z}^1|y|^p\mathrm{d}s-C_p^pz^p=B(x,y,z)\end{aligned}$$

(to see the latter identity, make the substitution $u=\ln(sz/(x-y))$ in the second integral above). This completes the analysis of the case $1<p\leq 2$.

Before we proceed, observe that as a by-product, the above calculations imply that the best constant in (1.5) is at least $\left(\int_0 e^{-u}|u-1|^p \mathrm{d}u\right)^{1/p}$, in the full range of p. Indeed, take a function f as above, corresponding to $x=0$, $y=-1$ and $z=1$. Then we have $\|f\|_{BLO(0,1)} = \|\widetilde{f^\#}\|_{L^\infty(0,1)} \leq 1$ and $\|f\|_{L^p(0,1)} = \left(\int_0 e^{-u}|u-1|^p\mathrm{d}u\right)^{1/p}$, so the lower bound for the constant in (1.5) is established.

We return to the estimate $\mathbb{B} \geq B$ and assume now that $p>2$. Suppose first that $y \geq (p-1)z$; then $py = y + (p-1)y \geq (p-1)(z+y) \geq (p-1)x$. Let $f:(0,1] \to \mathbb{R}$ be defined by

$$f(s) = \begin{cases} y\left(\dfrac{py-(p-1)x}{sy}\right)^{1/p} & \text{if } 0 < s \leq p - \dfrac{(p-1)x}{y}, \\ y & \text{if } p - \dfrac{(p-1)x}{y} < s \leq 1. \end{cases}$$

Again, this function is nonincreasing and hence

$$\widetilde{f^\#}(s) = \sup\left\{ \frac{1}{b}\int_0^b f(u)\mathrm{d}u - f(b) \,:\, b \in (0,1]\right\}.$$

But for any $b\in(0,1]$ we see that the expression

$$\frac{1}{b}\int_0^b f(u)\mathrm{d}u - f(b) = \begin{cases} \dfrac{y}{p-1}\left(\dfrac{py-(p-1)x}{by}\right)^{1/p} & \text{if } 0 < b \leq p - \dfrac{(p-1)x}{y}, \\ \dfrac{py-(p-1)x}{b(p-1)} & \text{if } p - \dfrac{(p-1)x}{y} < b \leq 1 \end{cases}$$

is decreasing with b. Consequently, we have the identity

$$\widetilde{f^\#}(s) \vee z = \begin{cases} \dfrac{y}{p-1}\left(\dfrac{py-(p-1)x}{sy}\right)^{1/p} & \text{if } 0 < s \leq p - \dfrac{(p-1)x}{y}, \\ \dfrac{py-(p-1)x}{s(p-1)} & \text{if } p - \dfrac{(p-1)x}{y} < s \leq \dfrac{py-(p-1)x}{(p-1)z}, \\ z & \text{if } \dfrac{py-(p-1)x}{(p-1)z} < s \leq 1. \end{cases}$$

Note that in particular, $f(s) = (p-1)\big(\widetilde{f^\#}(s)\vee z\big)$ for $0 < s \leq p-(p-1)x/y$. Therefore, we may write

$$\begin{aligned} \mathbb{B}(x,y,z) &\geq \int_0^1 \Big(|f(s)|^p - (p-1)^p(\widetilde{f^\#}(s)\vee z)^p\Big)\mathrm{d}s \\ &= \int_{p-(p-1)x/y}^1 \Big(|f(s)|^p - (p-1)^p(\widetilde{f^\#}(s)\vee z)^p\Big)\mathrm{d}s = B(x,y,z), \end{aligned}$$

after some straightforward calculations. Finally, suppose that $y<(p-1)z$ and consider the function $f:(0,1]\to\mathbb{R}$ defined by

$$f(s) = \begin{cases} (p-1)z\left(\dfrac{s_0}{s}\right)^{1/p} & \text{if } 0 < s \leq s_0, \\ y - z\ln\dfrac{sz}{x-y} & \text{if } s_0 < s \leq \dfrac{x-y}{z}, \\ y & \text{if } \dfrac{x-y}{z} < s \leq 1. \end{cases}$$

Here $s_0 = \dfrac{x-y}{z}\exp\left(\dfrac{y}{z}-p+1\right)$. Since f is nonincreasing, we may analyze $\widetilde{f\#}$ as before. Namely, we have the equality $f(s) = (p-1)\widetilde{f\#}(s)$ for $s \in (0, s_0)$ (the calculations are the same as previously) and $f(s) > (p-1)z$ for such s, so we get $f = (p-1)\widetilde{f\#} \vee z$ on $(0, s_0]$. Next, observe that

$$\int_0^{s_0} f(s)\mathrm{d}s = \int_0^{s_0}\left(y - z\ln\frac{sz}{x-y}\right)\mathrm{d}s,$$

which implies that $\widetilde{f\#}$ and the sharp maximal function considered in the case $1 < p \leq 2$, coincide on $(s_0, 1]$. Indeed: the expressions $\frac{1}{b}\int_0^b f(u)\mathrm{d}u - f(b)$ are the same for $b \in (s_0, 1]$, and hence give rise to the same suprema. In particular, this yields $\widetilde{f\#} \vee z = z$ on $(s_0, 1]$ and we obtain

$$\begin{aligned}\mathbb{B}(x,y,z) &\geq \int_0^1 \left(|f(s)|^p - (p-1)^p(\widetilde{f\#}(s)\vee z)^p\right)\mathrm{d}s\\ &= \int_{s_0}^1 \left(|f(s)|^p - (p-1)^p(\widetilde{f\#}(s)\vee z)^p\right)\mathrm{d}s = B(x,y,z).\end{aligned}$$

This establishes the desired lower bound.

4. Proof of the upper bound $\mathbb{B} \leq B$

It is convenient to split the reasoning into two separate sections.

4.1. **Crucial properties of B.**

In the two lemmas below, we establish some important estimates for the function B.

Lemma 4.1. *We have $B_y(x,y,z) \leq 0$ in the interior of Ω_+.*

Proof. By homogeneity, we may and do assume that $z = 1$. A direct differentiation reveals that

$$\begin{aligned}&B_y(x,y,1)\\ &= -b(y+1) - yb'(y+1) + |y|^p + p|y|^{p-2}(1+y) + x\Big[b'(y+1) - p|y|^{p-2}y\Big].\end{aligned} \tag{4.1}$$

We will show that the expression in the square brackets is nonnegative. We consider the cases $1 < p \leq 2$ and $p > 2$ separately. In the first case, this is equivalent to

$$\int_{y+1}^\infty e^{-u}|u-1|^p\mathrm{d}u - e^{-y-1}(|y|^p + p|y|^{p-2}y) \geq 0$$

and is due to the following observation: the left-hand side converges to zero as $y \to \infty$, and its derivative (with respect to y) is equal to $-p(p-1)e^{-y-1}|y|^{p-2} < 0$. If $p > 2$ and $y \leq p-1$, the nonnegativity of the square bracket in (4.1) amounts to

$$\int_{y+1}^p e^{-u}|u-1|^p + (p^2+p-1)(p-1)^{p-1}e^{-p} - e^{-y-1}(|y|^p + p|y|^{p-2}y) \geq 0.$$

The same calculation as above shows that the left-hand side is a decreasing function of y; furthermore, its value at $y = p-1$ equals $p(p-1)^p e^{-p} \geq 0$. For $y > p-1$, the nonnegativity is equivalent to $py^{p-1} + p(p-1)^2y^{p-2} \geq py^{p-1}$, which is trivial. Therefore, coming back to (4.1), we see that it suffices to show $B_y(x,y,z) \leq 0$ for the largest x, i.e., for $x = y+1$. However, we have

$$B_y(y+1, y, 1) = b'(y+1) - b(y+1) + |y|^p,$$

which is equal to zero for $1 < p \leq 2$, or for $p > 2$ and $y \leq p-1$. It remains to compute that for $p > 2$ and $y > p-1$, we have

$$b'(y+1) - b(y+1) + |y|^p = p(p-1)\left\{(p-1)y^{p-2} - \frac{p-2}{p-1}y^{p-2} - (p-1)^{p-2}\right\}.$$

The expression in the parentheses is nonpositive, by a direct application of Young's inequality. □

Lemma 4.2. *We have $B_z(x,y,z) \leq 0$ in the interior of Ω_+.*

Proof. Note that for fixed y and z, the function $x \mapsto B_z(x,y,z)$ is linear; furthermore, its limit as $x \to y$ is equal to $-pC_p^p z^{p-1}$. Thus, it is enough to prove the assertion for $x = y+z$. Furthermore, by homogeneity, we may and do assume that $z = 1$. Then the desired bound reads

$$(p-1)b(y+1) - yb'(y+1) + |y|^p - pC_p^p \leq 0. \tag{4.2}$$

The case $1 < p \leq 2$. Since $b'(y+1) + |y|^p = b(y+1)$, we may rewrite the claim in the form $\xi(s) := sb'(s) - p(b(s) - b(0)) \geq 0$, where $s = y+1$. Clearly, we have $\xi(0) = 0$, so we will be done if we show that ξ is nondecreasing on $[0,\infty)$. A direct differentiation shows that

$$\xi'(s) = p(p-1)e^s\left[s\int_s^\infty e^{-u}|u-1|^{p-2}\mathrm{d}u - \int_s^\infty e^{-u}|u-1|^{p-2}(u-1)\mathrm{d}u\right].$$

Denote the expression in the square brackets by $\zeta(s)$. Clearly, $\zeta(s) \to 0$ as $s \to \infty$. Furthermore, we have

$$\zeta(0) = -\int_0^\infty e^{-u}|u-1|^{p-2}(u-1)\mathrm{d}u = e\left(\int_0^1 e^u u^{p-1}\mathrm{d}u - \int_0^\infty e^{-u}u^{p-1}\mathrm{d}u\right),$$

by splitting the first integral into two, over $[0,1]$ and $[1,\infty)$, and making some simple substitutions. Therefore, $\zeta(0) \geq 0$, because

$$\int_0^1 e^u u^{p-1}\mathrm{d}u \geq \int_0^1 e^u u\mathrm{d}u = 1 \geq \int_0^\infty e^{-u}u^{p-1}\mathrm{d}u,$$

where the latter bound follows from the convexity of the function $p \mapsto u^{p-1}$ and the identities $\int_0^\infty e^{-u}\mathrm{d}u = \int_0^\infty e^{-u}u\mathrm{d}u = 1$. Finally, we compute that for $s \neq 1$,

$$\zeta''(s) = (2-p)e^{-s}|s-1|^{p-4}(s-1), \tag{4.3}$$

so ζ is concave on $(0,1)$ and convex on $(1,\infty)$. Putting all the above facts together, we get that $\zeta \geq 0$ on $[0,\infty)$, which implies the desired monotonicity of ξ and yields the claim.

The case $p > 2$. First, note that if $y \geq p-1$, then both sides of (4.2) are equal. If $y < p-1$, we use the identity $b'(y+1) = b(y+1) - |y|^p$ again and rewrite the claim in the form $\xi(s) := sb'(s) - p(b(s) - (p-1)^p) \geq 0$, where, as before, $s = y+1$. As we have just mentioned above, we have $\xi(p) = 0$, so we will be done if we prove the bound $\xi'(s) \leq 0$ for $s \in (0,p)$. After a little computation, one obtains that this is equivalent to

$$s\int_s^p e^{-u}|u-1|^{p-2}\mathrm{d}u - \int_s^p e^{-u}|u-1|^{p-2}(u-1)\mathrm{d}u + (p-1)^{p-1}e^{-p}(s-p) \leq 0.$$

Denoting the left-hand side by ζ, we check that $\zeta(p) = 0$ and $\zeta'(p) = (p-2)(p-1)^{p-2}e^{-p} \geq 0$. Furthermore, (4.3) holds for $s \neq 1$, so ζ is convex on $(0,1)$ and concave on $(1,p)$. Thus, the proof will be complete if we show that $\zeta(0)\leq 0$. Integrating by parts, this is equivalent to saying that

$$(p-1)\int_0^p e^{-u}|u-1|^{p-2}\mathrm{d}u+(p-1)^pe^{-p}\geq 1. \tag{4.4}$$

Let us start with the observation that

$$\int_0^2 e^{-u}|u-1|^{p-2}\mathrm{d}u=e^{-1}\int_{-1}^1 e^s|s|^{p-2}\mathrm{d}s=\frac{2}{e}\sum_{k=0}^\infty\frac{1}{(2k)!(2k+p-1)}.$$

In the last passage we expanded e^s into its Taylor series and noted that the contribution of summands corresponding to odd powers of s is zero, by symmetry. In particular, the identity is true for $p=2$, so

$$\begin{aligned}\int_0^2 e^{-u}(|u-1|^{p-2}-1)\mathrm{d}u&=\frac{2}{e}\sum_{k=0}^\infty\left(\frac{1}{(2k)!(2k+p-1)}-\frac{1}{(2k)!(2k+1)}\right)\\&=-\frac{2(p-2)}{e}\sum_{k=0}^\infty\frac{1}{(2k+1)!(2k+p-1)}\\&\geq-\frac{2(p-2)}{e(p-1)}\sum_{k=0}^\infty\frac{1}{(2k+1)!}\\&=-\frac{2(p-2)}{e(p-1)}\cdot\frac{e-e^{-1}}{2}\geq-\frac{p-2}{p-1}.\end{aligned}$$

Consequently, we may proceed with (4.4) as follows:

$$\begin{aligned}(p-1)\int_0^p e^{-u}|u-1|^{p-2}\mathrm{d}u&\geq(p-1)\int_0^2 e^{-u}|u-1|^{p-2}\mathrm{d}u\\&=(p-1)\int_0^2 e^{-u}(|u-1|^{p-2}-1)\mathrm{d}u+(p-1)(1-e^{-2})\\&\geq-(p-2)+(p-1)(1-e^{-2})=1-(p-1)e^{-2}.\end{aligned}$$

Thus, we will be done if we show that $(p-1)^pe^{-p}\geq(p-1)e^{-2}$, or $((p-1)/e)^{p-1}\geq e^{-1}$. However, both sides of the latter estimate are equal for $p=2$, and the derivative of the left-hand side (with respect to p) is equal to $((p-1)/e)^{p-1}\ln(p-1)\geq 0$. This completes the proof. □

The above two properties immediately give the following concavity-type condition for the function B.

Corollary 4.3. *Suppose that (x,y,z), $(x_\pm,y_\pm,z_\pm)\in\Omega_+$ are three points satisfying $y_\pm\geq y$, $z_\pm\geq z$ and $x_\pm\leq y+z$. Assume further that there are $\alpha_\pm\in[0,1]$ summing up to 1 such that $x=\alpha_-x_-+\alpha_+x_+$. Then*

$$B(x,y,z)\geq\alpha_-B(x_-,y_-,z_-)+\alpha_+B(x_+,y_+,z_+). \tag{4.5}$$

Proof. For fixed y and z, the function $x\mapsto B(x,y,z)$ is linear, so we have

$$B(x,y,z)=\alpha_-B(x_-,y,z)+\alpha_+B(x_+,y,z)$$

(note that $(x_\pm,y,z)\in\Omega$, by the assumptions of the corollary). It remains to apply the two lemmas above to obtain

$$B(x_-,y,z)\geq B(x_-,y_-,z_-)\quad\text{and}\quad B(x_+,y,z)\geq B(x_+,y_+,z_+).$$

The proof is complete. □

4.2. **Proof of $\mathbb{B} \leq B$.** Now we will establish a splitting lemma, which will be fundamental for the Bellman induction argument.

Lemma 4.4. *Let $\eta \in (1,2)$ be a fixed parameter. Assume further that I is an arbitrary interval contained in $\mathbb{R}$ and let $f : I \to \mathbb{R}$ be a given integrable function. Then there is a splitting $I = I_- \cup I_+$ with the property that*

$$\begin{aligned} \langle f\rangle_{I_-} - \operatorname*{essinf}_I f &\leq \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big), \\ \langle f\rangle_{I_+} - \operatorname*{essinf}_I f &\leq \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big) \end{aligned} \tag{4.6}$$

and $|I_\pm| \leq \eta^{-1}|I|$.

Remark 4.5. In the assertion, all the essential infima are taken over I. Obviously, if the splitting as above exists, then we also have the weaker property

$$\langle f\rangle_{I_-} - \operatorname*{essinf}_{I_-} f \leq \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big), \qquad \langle f\rangle_{I_+} - \operatorname*{essinf}_{I_+} f \leq \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big).$$

However, the stronger version (4.6) will be crucial for our argumentation below.

Proof of Lemma 4.4. With no loss of generality we may assume that $I = [0,1]$; thus we search for the splitting of the form $I_- = [0,a]$ and $I_+ = (a,1]$ for some $a \in [1-\eta^{-1}, \eta^{-1}]$. We start by taking the half-splitting $a = 1/2$: if both above estimates hold, then we are done. Suppose conversely that (at least) one of them fails: let us assume that

$$\langle f\rangle_{I_-} - \operatorname*{essinf}_I f > \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big)$$

(if the second inequality in (4.6) is not true, the reasoning is similar). This in particular implies that $\langle f\rangle_I \neq \operatorname{essinf}_I f$. Consider the continuous function $a \mapsto \langle f\rangle_{[0,a]} - \operatorname{essinf}_I f$. Its value for $a = 1$ is equal to $\langle f\rangle_I - \operatorname{essinf}_I f \leq \eta(\langle f\rangle_I - \operatorname{essinf}_I f)$ and hence, by Darboux property, there is $a \in (0,1)$ such that

$$\langle f\rangle_{[0,a]} - \operatorname*{essinf}_I f = \eta(\langle f\rangle_I - \operatorname*{essinf}_I f).$$

We will show that this choice for a has all the required properties. Clearly, the first estimate in (4.6) is satisfied (both sides are equal). Furthermore,

$$\begin{aligned} \langle f\rangle_{(a,1]} - \operatorname*{essinf}_I f &= (1-a)^{-1}(\langle f\rangle_I - a\langle f\rangle_{[0,a]}) - \operatorname*{essinf}_I f \\ &= \frac{1-a\eta}{1-a}\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big) \leq \eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big), \end{aligned}$$

so the second inequality in (4.6) holds as well. It remains to prove that $1-\eta^{-1} \leq a \leq \eta^{-1}$. The left estimate is obvious since $a \geq 1/2$. To prove the right bound, note that

$$\begin{aligned} \langle f\rangle_I - \operatorname*{essinf}_I f &= a\big(\langle f\rangle_{[0,a]} - \operatorname*{essinf}_I f\big) + (1-a)\big(\langle f\rangle_{(a,1]} - \operatorname*{essinf}_I f\big) \\ &\geq a\eta\big(\langle f\rangle_I - \operatorname*{essinf}_I f\big) \end{aligned}$$

and hence $a\eta \leq 1$, as desired. □

Remark 4.6. By a straightforward induction argument, the lemma above leads to the following construction. Let $\eta > 1$ be a fixed parameter and let $f : I \to \mathbb{R}$ be a given integrable function. Then there is an increasing sequence $(\mathcal{I}_n)_{n\geq 0}$ of partitions of I, satisfying the following properties:

(i) $\mathcal{I}_0 = \{I\}$.

(ii) For $n \geq 0$, each $J \in \mathcal{I}_n$ is split into two intervals J_-, $J_+ \in \mathcal{I}_{n+1}$ such that

$$\langle f\rangle_{J_-} - \operatorname*{essinf}_J f \leq \eta\big(\langle f\rangle_J - \operatorname*{essinf}_J f\big), \qquad \langle f\rangle_{J_+} - \operatorname*{essinf}_J f \leq \eta\big(\langle f\rangle_J - \operatorname*{essinf}_J f\big)$$

and $|J_\pm| \leq \eta^{-1}|J|$.

Proof of the estimate $\mathbb{B} \leq B$. Fix an arbitrary interval I, an integrable function $f : I \to \mathbb{R}$ satisfying $\langle f\rangle_I = x$, $\operatorname{essinf}_I f = y$ and a parameter $\eta > 1$. Let $(\mathcal{I}_n)_{n\geq 0}$ be the sequence of partitions described in Remark 4.6. Consider the functional sequences $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ and $(h_n)_{n\geq 0}$ given as follows. For any $n \geq 0$ and $\omega \in I$, let

$$f_n(\omega) = \frac{1}{|I_n(\omega)|}\int_{I_n(\omega)} f, \qquad g_n(\omega) = \operatorname*{essinf}_{I_n(\omega)} f, \qquad h_n = \max_{0\leq k\leq n}\,(f_k(\omega) - g_k(\omega)) \vee z.$$

Here $I_n(\omega)$ is the unique element of $\mathcal{I}_n$ which contains ω. We will show that the sequence

$$\left(\int_I B(f_n, g_n, \eta h_n)\right)_{n\geq 0} \tag{4.7}$$

is nonincreasing. To this end, fix $n \geq 0$, let us pick an element $J \in \mathcal{I}_n$ and its children $J_\pm \in \mathcal{I}_{n+1}$. By the very definition, f_n, g_n and ηh_n are constant on J, while f_{n+1}, g_{n+1}, ηh_{n+1} are constant on J_- and J_+. We have $f_{n+1} - g_n \geq f_{n+1} - g_{n+1}$ and, by Remark 4.6, $f_{n+1} - g_n \leq \eta h_n$ on J. Therefore, we may apply the estimate (4.5) with $(x, y, z) = (f_n, g_n, \eta h_n)|_J$ and $(x_\pm, y_\pm, z_\pm) = (f_{n+1}, g_{n+1}, \eta h_{n+1})|_{J_\pm}$. As the result, we obtain the inequality

$$\int_J B(f_n, g_n, \eta h_n) \geq \int_{J_-} B(f_{n+1}, g_{n+1}, \eta h_{n+1}) + \int_{J_+} B(f_{n+1}, g_{n+1}, \eta h_{n+1}).$$

Consequently, summing over all J, we obtain the monotonicity of the sequence (4.7). In particular, this implies

$$\int_I B(f_n, g_n, \eta h_n) \leq \int_I B(f_0, g_0, \eta h_0).$$

But $B(f_0, g_0, \eta h_0) = B(x, y, z)$ (since $h_0 = (x - y) \vee z = z$) and by Lemma 4.1,

$$B(f_n, g_n, \eta h_n) \geq B(f_n, f_n, \eta h_n) = |f_n|^p - C_p^p\eta^p h_n^p.$$

Furthermore, we have $h_n \leq \widetilde{f^\#} \vee z$, by the very definition of the sharp function. Putting all the above observations together, we obtain the estimate

$$\frac{1}{|I|}\int_I (|f_n|^p - C_p^p\eta^p(\widetilde{f^\#} \vee z)^p) \leq B(x, y, z).$$

If we let $n \to \infty$, then f_n converges to f almost everywhere, which follows directly from Lebesgue's differentiation theorem and the fact that the splitting ratios in Lemma 4.4 are bounded away from zero. Consequently, exploiting Fatou's lemma, we see that the above estimate yields

$$\frac{1}{|I|}\int_I (|f|^p - C_p^p\eta^p(\widetilde{f^\#} \vee z)^p)\mathrm{d}\omega \leq B(x, y, z).$$

Since $\eta > 1$ was arbitrary, this gives the desired assertion. □

5. On the search for the Bellman function

As we have just seen, the discovery of the function $\mathbb{B}$ was based on the following two-step procedure. First we *guessed* the extremal functions f which have provided us with the candidate for the Bellman function; then, using the concavity properties of this object and an appropriate inductive argument, we have proved that this candidate actually coincides with $\mathbb{B}$. Thus, the main difficulty of our problem was hidden in the appropriate choice of extremizers, and the purpose of this section is to describe the reasoning which leads to these objects (and, in turn, produces the function $B^{(p)}$).

Actually, as we will see, it is more convenient to start with the direct search for the Bellman function, and then extract the extremizers from its structure. Fix a parameter $1 \leq p < \infty$. Suppose that we are interested in the discovery of the (a priori unknown) best constant C in the estimate

$$\|f\|_{L^p(I)} \leq C \left\| \widetilde{f^\#} \right\|_{L^p(I)} \tag{5.1}$$

and, more generally, the identification of the associated Bellman function defined in (2.1). How can we address this problem? Suppose that $\mathcal{B} : \Omega \to \mathbb{R}$ is a function enjoying the concavity property of Corollary 4.3 and the majorization

$$\mathcal{B}(y, y, z) \geq |y|^p - C^p z^p, \qquad \text{for all } y \in \mathbb{R} \text{ and } z \geq 0. \tag{5.2}$$

Then the approach presented in the previous section yields the pointwise estimate $B^{(p,C)} \leq \mathcal{B}$. Furthermore, one might hope for equality if $\mathcal{B}$ is chosen to be the least function with the above properties. Now we present an *informal* reasoning which will lead to this object, exploiting a number of additional assumptions and guesses. For the sake of clarity, we split the analysis into several intermediate steps.

Step 1. Initial reductions. We may and do assume that $\mathcal{B}$ is homogeneous of order p, replacing it with the function $(x, y, z) \mapsto \inf_{\lambda > 0} \lambda^{-p} \mathcal{B}(\lambda x, \lambda y, \lambda z)$ if necessary. Second, we may assume that (5.2) holds with equality: if this is not the case, we replace $\mathcal{B}$ with

$$\tilde{\mathcal{B}}(x, y, z) = \begin{cases} \mathcal{B}(x, y, z) & \text{if } x > y, \\ |y|^p - C^p z^p & \text{if } x = y, \end{cases}$$

which inherits the concavity of Corollary 4.3.

Step 2. Key differential equations for $\mathcal{B}$. We begin with testing the estimate (4.5) with some special choices of the variables (x, y, z), $(x_\pm, y_\pm, z_\pm)$ and $\alpha_\pm$. First, plugging $y = y_- = y_+$ and $z = z_- = z_+$ transforms the inequality into

$$\mathcal{B}(\alpha_- x_- + \alpha_+ x_+, y, z) \geq \alpha_- \mathcal{B}(x_-, y, z) + \alpha_+ \mathcal{B}(x_+, y, z).$$

This implies that for given y and z, the function $x \mapsto \mathcal{B}(\cdot, y, z)$ is concave on $[y, y + z]$. Second, let us take $x = x_- = x_+$, $y = y_- < y_+$, $z = z_- = z_+$ and let $\alpha_\pm$ be arbitrary. Then (4.5) gives

$$\mathcal{B}(x, y, z) \geq \mathcal{B}(x, y_+, z),$$

that is, $\mathcal{B}$ is nonincreasing with respect to the variable y. A similar argument shows the analogous monotonicity with respect to the variable z: we have

$$\mathcal{B}(x, y, z) \geq \mathcal{B}(x, y, z_+),$$

provided $(x, y, z) \in \Omega_+$ and $z_+ > z$. Assuming that $\mathcal{B}$ is of class C^2, the above properties imply that

$$\mathcal{B}_{xx}(x, y, z) \leq 0, \tag{5.3}$$

$$\mathcal{B}_y(x, y, z) \leq 0 \tag{5.4}$$

and

$$\mathcal{B}_z(x, y, z) \leq 0. \tag{5.5}$$

Since we want to find the extremal $\mathcal{B}$, it seems plausible to assume that the estimates (5.3), (5.4), (5.5) become equalities, at least for a large set of (x, y, z). Motivated by related results from the literature, we assume that the second-order condition degenerates everywhere, i.e, we have the equality $\mathcal{B}_{xx} = 0$ on the whole Ω_+. Then $\mathcal{B}$ admits the formula

$$\begin{aligned} \mathcal{B}(x, y, z) &= z^p \mathcal{B}(x/z, y/z, 1) \\ &= z^p \left[\frac{x-y}{z} \cdot \mathcal{B}\left(\frac{y}{z} + 1, \frac{y}{z}, 1\right) + \left(1 - \frac{x-y}{z}\right) \cdot \mathcal{B}\left(\frac{y}{z}, \frac{y}{z}, 1\right) \right] \\ &= z^p \left[\frac{x-y}{z} \cdot b\left(\frac{y}{z} + 1\right) + \left(1 - \frac{x-y}{z}\right) \cdot \left|\frac{y}{z}\right|^p - C^p \right], \end{aligned} \tag{5.6}$$

where $b(s) = \mathcal{B}(s+1, s, 1) + C^p$. This should be compared to (2.1).

Step 3. On the derivatives $\mathcal{B}_y$ and $\mathcal{B}_z$. We see that for any y, z the function $x \mapsto \mathcal{B}(x, y, z)$ is linear on $[y, y+z]$ and hence $\mathcal{B}_y$ and $\mathcal{B}_z$ also have this property. Consequently, (5.4) and (5.5) hold true if and only if

$$\mathcal{B}_y(y, y, z) \leq 0, \qquad \mathcal{B}_y(y+z, y, z) \leq 0 \tag{5.7}$$

and

$$\mathcal{B}_z(y, y, z) \leq 0, \qquad \mathcal{B}_z(y+z, y, z) \leq 0. \tag{5.8}$$

The further analysis will be based on the assumption that one of the right inequalities in (5.7) and (5.8) becomes an equality.

Step 4. The case $\mathcal{B}_y(y+z, y, z) = 0$. Combining this equation with (5.6) leads to the ordinary differential equation

$$b'(s) = b(s) - |s-1|^p, \qquad s \in \mathbb{R}, \tag{5.9}$$

which is easily solved: for some real parameter κ,

$$b(s) = e^s \left(\int_s^\infty e^{-u} |u-1|^p \mathrm{d}u + \kappa \right).$$

However, the first inequality in (5.7) implies $-b(s) + |s-1|^p + p|s-1|^{p-2}(s-1) \leq 0$ for all s, which gives $\kappa \geq 0$. Assuming that κ is actually equal to 0, we essentially obtain the Bellman function B, up to the choice of the constant C. The value of this constant is determined by the requirement $\mathcal{B}(0, -z, z) \leq 0$, which must hold in the light of the postulated validity of (5.1). It remains to note that the estimate $\mathcal{B}(0, -z, z) \leq 0$ is equivalent to $C \geq \left(\int_0^\infty e^{-u}|u-1|^p \mathrm{d}u\right)^{1/p}$. Assuming equality, we obtain the candidate for the desired Bellman function. Now, a direct verification reveals that the second inequality (5.8) holds if and only if $1 \leq p \leq 2$. So, when $p > 2$, it seems plausible to switch the partial derivative: this is done in the next step.

Step 5. The case $\mathcal{B}_z(y+z,y,z)=0$. Here the analysis will be more complicated, as the candidate we will end up with, will not satisfy the required properties and will need to be further modified. Nevertheless, one can try to proceed as previously: the equality $\mathcal{B}_z(y+z,y,z)=0$, plugged into (5.6) gives

$$(p-1)b(s+1)-sb'(s+1)+|s|^p-pC^p=0.$$

The general solution to this equation is given by

$$b(s+1)=|s|^p+\kappa|s|^{p-1}+\frac{pC^p}{p-1}, \tag{5.10}$$

where κ is an arbitrary parameter. To identify this parameter and guess the value of the optimal constant C, we inspect the second inequality in (5.7). In the light of (5.6), this estimate is equivalent to

$$b'(s+1)-b(s+1)+|s|^p\leq 0. \tag{5.11}$$

If $s>0$, this can be rewritten in the form

$$(\kappa-p)s^{p-1}-\kappa(p-1)s^{p-2}+\frac{pC^p}{p-1}\geq 0.$$

We must have $\kappa>p$, since otherwise the estimate will fail for large s. For such κ, the left-hand side, considered as a function of s, attains its minimum for

$$s_0=\frac{\kappa(p-2)}{\kappa-p}. \tag{5.12}$$

Plugging this value into the previous estimate and manipulating a little bit, we obtain

$$C^p\geq p^{-1}(p-1)(p-2)^{p-2}\cdot\kappa^{p-1}(\kappa-p)^{2-p}.$$

Since we are interested in the least possible value of C, it is natural to minimize the right-hand side over κ and assume equality. This gives $\kappa=p(p-1)$ and $C=p$. For these choices, it is easy to check that (5.11) holds also for negative s and thus one might hope that b leads to an appropriate candidate. Unfortunately, the function $\mathcal{B}$ thus obtained does not satisfy the inequality $\mathcal{B}(0,-z,z)\leq 0$, which is necessary for the validity of (5.1).

Step 6. Correcting $\mathcal{B}$ *for* $p>2$. The assumption that the equation $\mathcal{B}_z(y+z,y,z)=0$ holds for all y, z brought us to a function which is too big (at least on some part of its domain). To overcome this difficulty, it seems natural to impose the condition $\mathcal{B}_z(y+z,y,z)=0$ only for some y, z, and require that for the remaining y, z, the other derivative vanishes: $\mathcal{B}_y(y+z,y,z)=0$. The analysis performed at the previous step indicates the appropriate "transition point". Namely, take the function

$$b(s+1)=|s|^p+p(p-1)|s|^{p-1}+\frac{p^{p+1}}{p-1}$$

obtained above (we have plugged $\kappa=p(p-1)$ and $C=p$ into (5.10)). We have checked that the corresponding $\mathbb{B}$ satisfies $\mathbb{B}_y(y+z,y,z)\leq 0$, with equality for $y=s_0=p-1$ (where s_0 is defined in (5.12)). So, we assume that the above formula for b holds true for $s\geq p-1$; for $s<p-1$, we determine b by the requirement $\mathbb{B}_y(y+z,y,z)=0$. The calculations are similar to those presented in Step 4, and we leave the details to the reader.

Step 7. Identification of the extremizers. We begin with summarizing the idea of the proof of the inequality $\mathbb{B}\leq B$. Given $(x,y,z)\in\Omega_+$ and an arbitrary function

$f \in L^p$ satisfying $\langle f\rangle_I = x$, $\operatorname{essinf}_I f = y$, we have constructed certain functional sequences $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ and $(h_n)_{n\geq 0}$ for which

$$\begin{aligned}\int_I \Big(|f_n|^p - C_p^p(\widetilde{f^\#}\vee z)^p\Big)\mathrm{d}\omega &\leq \int_I B(f_n,g_n,h_n)\mathrm{d}\omega\\ &\leq \int_I B(f_{n-1},g_{n-1},h_{n-1})\mathrm{d}\omega\\ &\leq \int_I B(f_0,g_0,h_0)\mathrm{d}\omega = B(x,y,z)\end{aligned}$$

(we omitted the auxiliary factor η which was standing in front of h_n, since it is irrelevant for the main idea). Recall that the second inequality above is due to the concavity property described in Corollary 4.3. To find the extremal function f, we need to make sure that all the intermediate estimates above actually become equalities (or almost equalities). In particular, we need to guarantee that

$$\int_I B(f_n,g_n,h_n)\mathrm{d}\omega = \int_I B(f_{n-1},g_{n-1},h_{n-1})\mathrm{d}\omega$$

for all n. Motivated by the definitions of f_n, g_n, h_n, we propose the following algorithm. Suppose that J is a given subinterval of I, $(x,y,z)\in\Omega_+$ is fixed and we are interested in the construction of the extremal function $f : J\to\mathbb{R}$ for which the supremum defining $\mathbb{B}(x,y,z)$ is attained. First, note that if $x = y$, then this task is trivial: the only function (up to a set of measure zero) which satisfies the condition $\langle f\rangle_J = \operatorname{essinf}_J f = x$ is the constant function $f\equiv x$. If $x > y$, then we search for the "degenerated direction" of the concavity condition of Corollary 4.3. That is, we identify appropriate nontrivial $(x_\pm, y_\pm, z_\pm)$ and $\lambda_\pm$ for which

$$B(x,y,z) = \lambda_- B(x_-,y_-,z_-) + \lambda_+ B(x_+,y_+,z_+).$$

Then we split J into two subintervals J_-, J_+ such that $|J_\pm|/|J| = \lambda_\pm$. Now we restrict our attention to $J_\pm$ and construct there the restrictions $f|_{J_\pm}$ as the extremizers corresponding to $B(x_\pm,y_\pm,z_\pm)$, iterating the above algorithm. As we shall see in a moment, it is possible to pick the splitting points $(x_\pm,y_\pm,z_\pm)$ so that $x_+ = y_+$. This implies that $f|_{J_+}$ is identically equal to x_+ and the iterative procedure described above converges to a well-defined function.

We will present the detailed analysis in the case $1\leq p\leq 2$ only; the range $p > 2$ can be dealt with similarly. Fix $(x,y,z)\in\Omega_+$ and let $\delta > 0$ be an auxiliary parameter (which will eventually be sent to zero). We consider three operations.

1° If $x = y$, then the extremal function $f : I\to\mathbb{R}$ is $f\equiv x$.

2° If $y < x < y+z$, we write down the identity

$$B(x,y,z) = \frac{x-y}{z}\cdot B(y+z,y,z) + \left(1-\frac{x-y}{z}\right)\cdot B(y,y,z).$$

According to the above algorithm, we split I into two subintervals $I_\pm$ such that $|I_-|/|I| = (x-y)/z$ and $|I_+|/|I| = 1-(x-y)/z$, and define $f = y$ on I_+. To understand the construction of f on I_-, we use the operation 3° below.

3° If $x = y+z$, then we use the "almost" equality

$$B(x,y,z)\approx (1+\delta)^{-1}B(x+\delta,y+\delta,z) + \delta(1+\delta)^{-1}B(y,y,z).$$

(The error corresponding to the above approximation is of order $o(\delta)$ as $\delta\to 0$). We split I into two subintervals $I_\pm$ such that $|I_-|/|I| = (1+\delta)^{-1}$ and $|I_+|/|I| =$

$\delta(1+\delta)^{-1}$. We set $f \equiv y$ on I_+, while for the definition of f on I_- we iterate the operation 3° (with the new point $(x+\delta, y+\delta, z)$ and the underlying interval I_-).

Now, suppose that the underlying interval I is equal to $[0,1]$ and suppose that $x > y$. It is not difficult to see that the above procedure gives the following function:

$$f(s) = \begin{cases} y + n\delta & \text{if } \frac{x-y}{z}(1+\delta)^{-n-1} < s \le \frac{x-y}{z}(1+\delta)^{-n} \text{ for some } n, \\ y & \text{if } s > \frac{x-y}{z}. \end{cases}$$

In the limit $\delta \to 0$, we obtain precisely the function defined in (3.1).

Acknowledgments

The authors would like to thank an anonymous Referee for the careful and insightful reading of the first version of the paper, for many comments and helpful suggestions. The results were obtained during the realization of the program supported by Polish Ministry of Science and Higher Education: MNiSW/2019/394/DIR/KH.

References

[1] R. Bellman, *Dynamic programming*, Reprint of the 1957 edition. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2010.

[2] C. Bennett, *Another characterization of BLO*, Proc. Amer. Math. Soc. **85** (1982), no. 4, 552–556.

[3] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. **12** (1984), 647-702.

[4] D. L. Burkholder, *Explorations in martingale theory and its applications*, École d'Ete de Probabilités de Saint-Flour XIX—1989, 1–66, Lecture Notes in Math., 1464, Springer, Berlin, 1991.

[5] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994), 343–358, Proc. Sympos. Appl. Math., **52**, Amer. Math. Soc., Providence, RI, 1997.

[6] R. R. Coifman and R. Rochberg, Another characterization of BMO, Proc. Amer. Math. Soc. **79** (1980), pp. 249–254.

[7] C. Fefferman, *Characterization of bounded mean oscillation*, Bull. Amer. Math. Soc. **77** (1971), 587–588.

[8] A. Garsia, *Martingale Inequalities*, Seminar Notes on Recent Progress, New York, Benjamin 1973.

[9] L. Grafakos, Modern Fourier analysis. Third edition. Graduate Texts in Mathematics, 250. Springer, New York, 2014.

[10] F. John and L. Nirenberg, On functions of bounded mean oscillation, Comm. Pure and Appl. Math. **14** (1961), 415–426.

[11] N. Kazamaki, *Continuous exponential martingales and BMO*,Lect. Notes in Math. **1579**, Springer, Berlin, 1994.

[12] A. A. Korenovskii, *Mean oscillations and equimeasurable rearrangements of functions*, Lecture Notes of the Unione Matematica Italiana, **4**. Springer, Berlin; UMI, Bologna, 2007.

[13] A. D. Melas, *The Bellman functions of dyadic-like maximal operators and related inequalities*, Adv. Math. **192** (2005), 310–340.

[14] F. L. Nazarov and S. R. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis,* St. Petersburg Math. J. **8** (1997), 721–824.

[15] F. L. Nazarov, S. R. Treil and A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc., **12** (1999), 909-928.

[16] A. Osękowski, Sharp martingale and semimartingale inequalities, Monografie Matematyczne **72** (2012), Birkhäuser, 462 pp.

[17] A. Osękowski, *A splitting procedure for Bellman functions and the action of dyadic maximal operators on L^p*, Mathematika **61** (2015), 199–212.

[18] D. M. Stolyarov and P. B. Zatitskiy, *Sharp transference principle for BMO and A_p*, available at https://arxiv.org/pdf/1908.09497.pdf.

[19] Y. Suh, *A sharp weak type (p,p) inequality ($p > 2$) for martingale transforms and other subordinate martingales*, Trans. Amer. Math. Soc. **357** (2005), 1545–1564.

[20] V. Vasyunin, *The exact constant in the inverse Hölder inequality for Muckenhoupt weights* (Russian), Algebra i Analiz 15 (2003), 73–117; translation in St. Petersburg Math. J. 15 (2004), 49–79.

[21] G. Wang, *Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities*, Ann. Probab. **23** (1995), 522-551.

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

Email address: L.Kaminski5@students.mimuw.edu.pl

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

Email address: A.Osekowski@mimuw.edu.pl