COLLOQUIUM MATHEMATICUM

Online First version

AN OPTIMAL ESTIMATE FOR MARTINGALE TRANSFORMS

BY

BARBARA BUGAŁA

Abstract. Let $1 < p < \infty$ be a fixed exponent. Suppose that $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are real-valued martingales satisfying $f_0 \equiv x$, $g_0 \equiv y$, $\|(f_n)_{n\geq 0}\|_p = F$ and $g_n - g_{n-1} = v_n(f_n - f_{n-1})$, $n = 1, 2, \ldots$, for some predictable sequence $(v_n)_{n\geq 1}$ taking values in $[-1, 1]$. The purpose of this paper is to determine the optimal (i.e. the smallest) constant $B = B_{1,p}(x, y, F)$ for which $\|(g_n)_{n\geq 0}\|_1 \leq B_{1,p}(x, y, F)$.

1. Introduction. The aim of this paper is to investigate a certain martingale inequality, whose roots go back to the works of Ville [18], Burkholder (cf. [4, 5]) and Choi [7]. Consider a player A engaging in an infinite fair game, where fairness is understood to mean that the sequence of winnings per round, $(df_n)_{n\geq 0}$, forms a martingale difference sequence. In other words, the sequence $(f_n)_{n\geq 0}$, representing the player's capital after round n, constitutes a martingale. We assume that the player starts with an initial capital $x \in \mathbb{R}$ (that is, $f_0 = df_0 = x$ almost surely, allowing for the possibility of negative capital), and that for some fixed $F \geq |x|$ and $1 < p < \infty$, the condition

$$\|(f_n)_{n\geq 0}\|_p = F$$

is satisfied. Here and below, we use the notation $\|(f_n)_{n\geq 0}\|_p = \sup_{n\geq 0} \|f_n\|_p$. Thus, we have the complete knowledge about the initial value and the size of the martingale $(f_n)_{n\geq 0}$, measured in terms of the pth norm.

Next, we introduce an additional player B, who starts with capital y and interacts with player A: in each round n (for $n = 1, 2, \ldots$), player B adapts and modifies the strategy of player A. Specifically, player B picks a predictable sequence $(v_n)_{n\geq 1}$ of random variables taking values in $[-1, 1]$, and her/his winnings in the nth round are given by $v_n df_n$. For instance, setting $v_n = 1$ corresponds to the case in which player B replicates the strategy of player A in the nth game, while $v_n = -1$ means that player B follows the opposite strategy. Another natural possibility is to set $v_n = 0$,

---

2020 *Mathematics Subject Classification*: Primary 60G42; Secondary 60G46.

Key words and phrases: martingale, martingale transform, Burkholder's method, optimal constant.

Received 9 September 2025; revised 10 March 2026.

DOI: 10.4064/cm9708-4-2026

in which case player B abstains in the nth round. In the above setup, we see that $(g_n)_{n\geq 0}$, the capital of player B, also forms a martingale, with the sequence of differences $(dg_n)_{n\geq 0}$ given by

$$dg_n = v_n df_n, \qquad n = 1, 2, \ldots.$$

A natural question arises: given $f_0 = x$ and $\|(f_n)_{n\geq 0}\|_p = F$, how large can the martingale $(g_n)_{n\geq 0}$ be in the L^p norm? That is, how much advantage can player B gain over player A by appropriately selecting the sequence $(v_n)_{n\geq 0}$? This problem was resolved by Burkholder [4], who provided an explicit (quite complicated) formula for the function

$$\tilde{B}_{p,p}(x, y, F) = \sup_{v} \|(g_n)_{n\geq 0}\|_p.$$

In this expression, the supremum is taken over all pairs $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ starting from the points x and y, respectively, satisfying the conditions $dg_n = v_n df_n$, $n = 1, 2, \ldots$, where $(v_n)_{n\geq 0}$ is a predictable sequence taking values in $[-1, 1]$, and $\|(f_n)_{n\geq 0}\|_p = F$. See also [14, 16]. The above problem can be generalized in various directions. For instance, one may ask about the version of the above result in the limiting case $p = 1$. It turns out that the function $\tilde{B}_{1,1}$ takes the form

$$\tilde{B}_{1,1}(x, y, F) = \begin{cases} \infty & \text{if } F > |x|, \\ |y| & \text{if } F = |x|, \end{cases}$$

so the advantage of player B is infinite in this case. One can also consider different functionals quantifying the size of the martingale $(g_n)_{n\geq 0}$. Choi [7] identified an explicit formula for the function

$$\tilde{B}_{1,1}(x, y, F) = \sup\left\{\mathbb{P}(g_n \geq 1) : f_0 = x,\ g_0 = y,\ \|(f_n)_{n\geq 0}\|_1 = F\right\}.$$

See [15] for a two-sided extension, and [11] for an analogous result for nonnegative submartingales.

The aim of the present paper is to study a related problem, in which we seek an optimal bound on the size of the martingale $(g_n)_{n\geq 0}$ in the L^1 norm. More precisely, for a fixed $1 < p < \infty$, we investigate the function

$$\tilde{B}_{1,p}(x, y, F) = \sup\left\{\|(g_n)_{n\geq 0}\|_1 : f_0 = x,\ g_0 = y,\ \|(f_n)_{n\geq 0}\|_p = F\right\}.$$

Recall that we must necessarily have $F \geq |x|$, since otherwise, by the Hölder inequality, there are no martingales $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ satisfying the above requirements and the function $\tilde{B}_{1,p}$ is not well-defined. The analysis of the above problem will involve the construction of a special function satisfying appropriate majorization and concavity conditions. Actually, as we will see below, we will not address $\tilde{B}_{1,p}$ directly, but rather concentrate on a family of related, simpler special functions of two variables.

Our main result, which identifies an explicit form of this optimal bound, is the following theorem:

THEOREM 1.1. *Let* $1 < p < \infty$ *and suppose* $(f_n)_{n\geq 0}$ *and* $(g_n)_{n\geq 0}$ *are real-valued martingales such that* $f_0 \equiv x$, $g_0 \equiv y$, *and* $\|(f_n)_{n\geq 0}\|_p = F$. *Assume that* $(g_n)_{n\geq 0}$ *is a martingale transform of* $(f_n)_{n\geq 0}$ *with a predictable sequence* $(v_n)_{n\geq 0}$ *taking values in* $[-1,1]$. *Then*

$$\|(g_n)_{n\geq 0}\|_1 \leq B_{1,p}(x,y,F),$$

where

$$B_{1,p}(x,y,F) = \inf_{\lambda>0}\{\lambda^{-1}\mathcal{B}^0(\lambda x, \lambda y) + \lambda^{p-1}F^p\}.$$

Here $\mathcal{B}^0$ *denotes the Bellman function corresponding to* $V(x,y) = |y| - |x|^p$.

The explicit form of $\mathcal{B}^0$ depends on the value of p and will be derived in Sections 3 and 4.

A brief outline of the structure of this paper is as follows. In the next section we describe our approach to the study of the above problem. Sections 3 and 4 are devoted to the analysis of the function $\tilde{B}_{1,p}$ in the cases $1 < p \leq 2$ and $2 < p < \infty$, respectively.

2. Basic concepts

2.1. Bellman function method. The main technique used in this paper is the so-called Burkholder method, or Bellman function method. It is a powerful tool for studying various extremal problems and inequalities in probability and analysis. In particular, when applied carefully, it can lead to the determination of optimal constants involved. The technique originated in the fundamental work of Richard Bellman on the theory of optimal stochastic control [3], and the applicability of this approach to problems in martingale theory and harmonic analysis was first observed by Donald Burkholder in the 1980s [4]. This direction has further been explored by many authors (see [1, 13] for more on the subject). An important extension of the method, which allowed the study of various analytic contexts, was proposed by Nazarov, Treil, and Volberg [8, 9] in the 1990s. Since then, the method has been applied to the study of numerous problems arising in probability and analysis. See [13, 17] for a more systematic presentation.

Roughly speaking, the Bellman function method connects the validity of a given estimate with the existence of a certain special function satisfying some majorization and concavity conditions. The form of these conditions depends on the structure of the inequality under consideration and the dynamics of the processes involved. Below, we will work with a version corresponding to martingale transforms, developed by Burkholder.

We start by introducing some terminology and notation. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a fixed probability space equipped with a filtration $(\mathcal{F}_n)_{n\geq 0}$. Assume further that $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are two adapted, real-valued martingales, with difference sequences $(df_n)_{n\geq 0}$ and $(dg_n)_{n\geq 0}$. We say that $(g_n)_{n\geq 0}$ is a *transform* of $(f_n)_{n\geq 0}$ if there is a predictable sequence $v = (v_n)_{n\geq 0}$ such that $dg_n = v_n df_n$ for $n = 0, 1, 2, \ldots$. In what follows, the symbol f_∞ will denote the almost sure limit (if it exists) of $(f_n)_{n\geq 0}$ as $n \to \infty$. A martingale $(f_n)_{n\geq 0}$ is called *simple* if for each n the variable f_n takes only a finite number of values and there is a deterministic index N such that $f_N = f_{N+1} = \cdots$ with probability 1.

Now, assume that $V : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a given function and suppose that we are interested in the value

$$\mathcal{B}^0 = \sup \mathbb{E}V(f_\infty, g_\infty), \tag{2.1}$$

where the supremum is taken over all pairs $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ of simple martingales such that $(g_n)_{n\geq 0}$ is a transform of $(f_n)_{n\geq 0}$ by a predictable sequence $(v_n)_{n\geq 0}$ taking values in $[-1, 1]$. Here the underlying probability space and filtration may vary, but we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is nonatomic. Observe that, since the martingales $(f_n)_{n\geq 0}$ and $(g_n)_{n\geq 0}$ are simple, we do not need to concern ourselves with the measurability of $V(f_\infty, g_\infty)$ nor with the existence of the expected values appearing in the definition of $\mathcal{B}^0$.

To study (2.1) efficiently, we specialize the problem to martingales which start from arbitrary deterministic positions $(x, y) \in \mathbb{R}^2$, that is, we assume that $f_0 \equiv x$ and $g_0 \equiv y$. Note that this actually forces us to go beyond the context of martingale transforms. Indeed, if $|y| > |x|$, then $(g_n)_{n\geq 0}$ *cannot* be a transform of $(f_n)_{n\geq 0}$, since $df_0/dg_0 \notin [-1, 1]$. This motivates the following definition: we let $\mathcal{B}^0 : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$\mathcal{B}^0(x, y) = \sup \mathbb{E}V(f_\infty, g_\infty),$$

where this time the supremum is taken over all pairs $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ of simple martingales such that

- $f_0 \equiv x$ and $g_0 \equiv y$ almost surely;
- $dg_n = v_n df_n$ for $n = 1, 2, \ldots$ for some predictable sequence $(v_n)_{n\geq 0}$ with values in $[-1, 1]$.

Note that in the latter condition we have omitted the equality $dg_0 = v_0 df_0$.

Having extended (2.1) to arbitrary starting positions, we distinguish a certain class of special functions. It consists of all $B : \mathbb{R}^2 \to \mathbb{R}$ satisfying the following two conditions:

1° (Majorization condition) For all $x, y \in \mathbb{R}$, we have

$$B(x, y) \geq V(x, y).$$

2° (Concavity condition) For all $x, y \in \mathbb{R}$, $\varepsilon \in [-1, 1]$, and any $\alpha \in (0, 1)$, $t_1, t_2 \in \mathbb{R}$ satisfying $\alpha t_1 + (1 - \alpha)t_2 = 0$, the following estimate holds:

$$\alpha B(x + t_1, y + \varepsilon t_1) + (1 - \alpha)B(x + t_2, y + \varepsilon t_2) \leq B(x, y).$$

Here is the link between functions B as above and the problem (2.1).

THEOREM 2.1. *Suppose that B satisfies conditions 1° and 2°. Let $(v_n)_{n\geq 1}$ be a predictable sequence with values in $[-1, 1]$ and let $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ be simple martingales satisfying $dg_n = v_n df_n$ for $n = 1, 2, \dots$. Then*

$$\mathbb{E}V(f_\infty, g_\infty) \leq \mathbb{E}B(f_0, g_0).$$

In particular, $\mathcal{B}^0 \leq B$ on $\mathbb{R}^2$.

Proof. From the concavity condition it follows that $(B(f_n, g_n))_{n\geq 0}$ forms a supermartingale: for every $n \geq 0$ we have

$$\begin{aligned}\mathbb{E}[B(f_{n+1}, g_{n+1}) \mid \mathcal{F}_n] &= \mathbb{E}[B(f_n + df_{n+1}, g_n + dg_{n+1}) \mid \mathcal{F}_n] \\ &= \mathbb{E}[B(f_n + df_{n+1}, g_n + v_{n+1}df_{n+1}) \mid \mathcal{F}_n].\end{aligned}$$

Combining this with the majorization condition, we get

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}B(f_n, g_n) \leq \mathbb{E}B(f_0, g_0).$$

Letting $n \to \infty$, we obtain the result. ■

Our next observation will allow us to reduce the analysis of special functions to the first quadrant $\{(x, y) : x, y \geq 0\}$.

LEMMA 2.1. *If V satisfies the symmetry condition*

$$V(x, y) = V(|x|, |y|) \quad \text{for all } (x, y) \in \mathbb{R}^2,$$

then $\mathcal{B}^0$ enjoys this property as well.

Proof. Fix $(x, y) \in \mathbb{R}^2$; let $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ be an arbitrary pair of simple martingales as in the definition of $\mathcal{B}^0(x, y)$. Then the pair $(-f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ is among those used in the supremum defining $\mathcal{B}^0(-x, y)$, and hence

$$\mathcal{B}^0(-x, y) \geq \mathbb{E}V(-f_\infty, g_\infty) = \mathbb{E}V(f_\infty, g_\infty).$$

Taking the supremum over all $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ as above, we get

$$\mathcal{B}^0(-x, y) \geq \mathcal{B}^0(x, y) \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

Passing from x to $-x$, we get the reverse bound and hence $\mathcal{B}^0(-x, y) = \mathcal{B}^0(x, y)$ for all x, y. The symmetry with respect to y is proved analogously, by switching the sign of $(g_n)_{n\geq 0}$. ■

2.2. Passing from $\mathcal{B}^0$ to $\tilde{B}_{1,p}$. Recall that we are interested in the function

$$\tilde{B}_{1,p}(x, y, F) = \sup \|(g_n)_{n\geq 0}\|_1,$$

where the supremum is taken over all pairs $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ starting from the points x and y, respectively, satisfying $dg_n = v_n df_n$, $n = 1, 2, \dots$ (where

$(v_n)_{n\geq 0}$ is a predictable sequence taking values in $[-1,1]$, as usual), and such that $\|(f_n)_{n\geq 0}\|_p = F$. It is not difficult to show that in the above supremum we may restrict ourselves to simple martingales. Furthermore, we may and will assume from now on that $F > |x|$. Indeed, if $F = |x|$, then the martingale $(f_n)_{n\geq 0}$ must be constant almost surely and thus $\tilde{B}_{1,p}(x,y,F) = |y|$.

Clearly, the above problem does not fall within the scope of the Bellman function discussed above, because of the appearance of the pth moment of $(f_n)_{n\geq 0}$. To handle this difficulty, we consider a slightly different problem: simply maximize the quantity $\mathbb{E}\big(|g_\infty| - |f_\infty|^p\big)$, without imposing any assumption on the pth norm of $(f_n)_{n\geq 0}$. This can be studied with the Bellman function method. Let

$$\mathcal{B}^0(x,y) = \sup \mathbb{E}V(f_\infty, g_\infty),$$

where $V(x,y) = |y| - |x|^p$, be the associated special function. Here is the relation between $\mathcal{B}^0$ and the desired function $\tilde{B}_{1,p}$.

Theorem 2.2. *We have*

$$\tilde{B}_{1,p}(x,y,F) = \inf_{\lambda>0}\{\lambda^{-1}\mathcal{B}^0(\lambda x,\lambda y) + \lambda^{p-1}F^p\}.$$

Proof. To prove the "$\leq$" part, assume that $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are martingales as in the definition of $\tilde{B}_{1,p}(x,y,F)$. Then, for any $\lambda > 0$, the pair $(\lambda f_n)_{n\geq 0}$, $(\lambda g_n)_{n\geq 0}$ is a pair of martingales as in the definition of $\mathcal{B}^0(\lambda x,\lambda y)$, so we obtain

$$\begin{aligned}\|(g_n)_{n\geq 0}\|_1 = \mathbb{E}|g_\infty| &= \lambda^{-1}\mathbb{E}[|\lambda g_\infty| - |\lambda f_\infty|^p + (\lambda F)^p]\\ &\leq \lambda^{-1}\mathcal{B}^0(\lambda x,\lambda y) + \lambda^{p-1}F^p.\end{aligned}$$

Taking the infimum over $\lambda > 0$, and then the supremum over $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$, we obtain the "$\leq$" part of the theorem.

To prove the reverse inequality, fix an arbitrary point (x,y) and $F \geq |x|$. It will be shown below that for any $\varepsilon > 0$ there exists $\lambda > 0$ and a pair $(\lambda f_n)_{n\geq 0}$, $(\lambda g_n)_{n\geq 0}$ as in the definition of $\mathcal{B}^0(\lambda x,\lambda y)$ such that $\|(\lambda f_n)_{n\geq 0}\|_p^p \geq (\lambda F)^p - \lambda\varepsilon$ and

$$\mathcal{B}^0(\lambda x,\lambda y) \leq \mathbb{E}\big(|\lambda g_\infty| - |\lambda f_\infty|^p\big) + \lambda\varepsilon.$$

Then we will have

$$\mathcal{B}^0(\lambda x,\lambda y) \leq \lambda\|(g_n)_{n\geq 0}\|_1 - (\lambda F)^p + \lambda\varepsilon$$

and hence

$$\|(g_n)_{n\geq 0}\|_1 \geq \lambda^{-1}\mathcal{B}^0(\lambda x,\lambda y) + \lambda^{p-1}F^p - \varepsilon.$$

Thus the claim will follow directly from the definition of $\tilde{B}_{1,p}$ and the fact that ε was chosen arbitrarily. ■

Thus, we have successfully reduced the dimension of the problem: all we need is to find an explicit formula for the function $\mathcal{B}^0$.

2.3. Search. Let $\mathcal{B}^0$ be the Bellman function associated with the function $V(x,y) = |y| - |x|^p$. Now we will describe some informal steps which can be used in the search for $\mathcal{B}^0$. The proof of Theorem 2.1 provides certain hints in this direction. Suppose that for each (x,y) the supremum

$$\mathcal{B}^0(x,y) = \sup \mathbb{E} V(f_\infty, g_\infty)$$

is attained by some pair $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$. It seems plausible that the predictable sequence for this extremal pair should take extreme values, i.e., from the set $\{-1,1\}$. Furthermore, we have two possibilities. First, it may happen that the extremal pair consists of constant martingales: in this case, we have $\mathcal{B}^0(x,y) = V(x,y)$. However, for some (x,y), achieving the supremum requires the use of nontrivial martingales, in which case $\mathcal{B}^0(x,y) > V(x,y)$. Moreover, since for such a nontrivial pair we have the equality

$$\mathcal{B}^0(x,y) = \mathbb{E} V(f_\infty, g_\infty),$$

all transitions in the proof of Theorem 2.1 must also be equalities. In other words, the random variable (f_∞, g_∞) must satisfy $V(f_\infty, g_\infty) = \mathcal{B}^0(f_\infty, g_\infty)$ almost surely, and at every step where we use the conditional Jensen inequality (which leads to the supermartingale property), we must have equality. The latter condition means that, locally, the function $\mathcal{B}^0$ should be locally linear along segments of slope -1 or $+1$ passing through (x,y).

Summarizing, we expect that the domain $\mathbb{R}^2$ of $\mathcal{B}^0$ is the union of three sets:

$$\begin{aligned} C &= \{(x,y) : \mathcal{B}^0(x,y) = V(x,y)\}, \\ D_- &= \{(x,y) : \mathcal{B}^0 \text{ is locally linear along segments of slope } -1\}, \\ D_+ &= \{(x,y) : \mathcal{B}^0 \text{ is locally linear along segments of slope } +1\}. \end{aligned}$$

Let us note that these sets may overlap: for instance, if $\mathcal{B}^0$ has the form $\mathcal{B}^0(x,y) = y^2 - x^2$ on some subdomain, then this subdomain is contained in both D_- and D_+.

The problem now reduces to determining the sets C, D_-, and D_+ and finding an expression for the function satisfying conditions 1° and 2°. In what follows, we will make use of additional assumptions and conjectures, which will lead us to a *candidate* for the special function, denoted by B. Once we obtain this object, we will verify rigorously that it coincides with the desired function $\mathcal{B}^0$. More specifically, the estimate $B \geq \mathcal{B}^0$ will be proved by applying Theorem 2.1, and the reverse bound will be established by constructing appropriate examples.

It turns out that the special functions corresponding to the cases $1 < p \leq 2$ and $2 < p < \infty$ have different structures. Therefore we analyse them in two separate sections.

3. The case $1 < p \le 2$. We start with the search for a candidate for the Bellman function. Having constructed the candidate B, we will verify the inequalities $B \ge \mathcal{B}^0$ and $B \le \mathcal{B}^0$ in the subsequent two subsections. We emphasize that during the search we are allowed to impose additional assumptions and make guesses: the only purpose is to obtain a candidate, whose key properties will be rigorously checked later.

3.1. Search for the candidate. It is convenient to split the analysis into three steps.

STEP 1: *The shape of the sets C and $D_\pm$.* Let us rewrite the definition of the Bellman function:

$$\mathcal{B}^0(x,y) = \sup \mathbb{E}(|g_\infty| - |f_\infty|^p).$$

When p is close to 1, the expression inside the expectation is close to the difference between the pth powers of $|f_\infty|$ and $|g_\infty|$. A related problem was studied by Burkholder [6], who showed that the function

$$\tilde{\mathcal{B}}^0(x,y) = \sup \mathbb{E}\big\{|g_\infty|^p - (p-1)^{-p}|f_\infty|^p\big\}$$

is given by the formula

$$\tilde{\mathcal{B}}^0(x,y) = \begin{cases} |y|^p - (p-1)^{-p}|x|^p & \text{if } |y| \ge (p-1)^{-1}|x|, \\ \alpha_p(|y| - (p-1)^{-1}|x|)(|x|+|y|)^{p-1} & \text{if } |y| < (p-1)^{-1}|x|, \end{cases}$$

for some constant α_p depending only on p. In this case, we have

$$\begin{aligned} C &= \{(x,y) : |y| \ge (p-1)^{-1}|x|\}, \\ D_- &= \{(x,y) : |y| < (p-1)^{-1}|x| \text{ and } xy > 0\}, \\ D_+ &= \{(x,y) : |y| < (p-1)^{-1}|x| \text{ and } xy < 0\}. \end{aligned}$$

See Figure 1. It seems plausible to expect that in our setting the geometry of these sets should be similar. Thus, we postulate the existence of a certain smooth curve $\gamma : [0,\infty) \to [0,\infty)$ for which

$$\begin{aligned} C &= \{(x,y) : |y| \ge \gamma(|x|)\}, \\ D_- &= \{(x,y) : |y| < \gamma(|x|) \text{ and } xy > 0\}, \\ D_+ &= \{(x,y) : |y| < \gamma(|x|) \text{ and } xy < 0\}. \end{aligned}$$

See Figure 2. Furthermore, as in Burkholder's function above, we will assume that our candidate is of class C^1 on $\mathbb{R}^2$ and satisfies the symmetry condition $B(x,y) = B(|x|,|y|)$ for $(x,y) \in \mathbb{R}^2$. Thanks to these assumptions and Lemma 2.1, we may restrict our analysis to the first quadrant.

STEP 2: *Formula for γ.* Suppose that $(x,y) \in [0,\infty)^2$. From the proposed form of C, we infer that if $y \ge \gamma(x)$, then

$$B(x,y) = y - x^p, \tag{3.1}$$

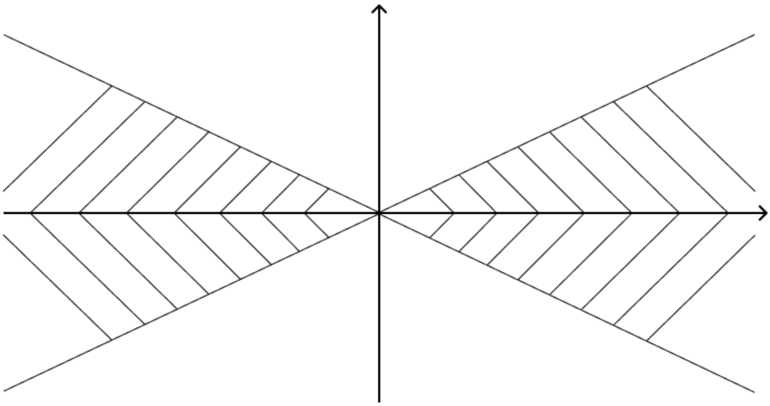

Fig. 1. Domain decomposition of Burkholder's function $\tilde{\mathcal{B}}^0$ corresponding to L^p estimates, $1 < p \leq 2$. The segments indicate the directions along which $\tilde{\mathcal{B}}^0$ is linear. In C, such segments do not exist.

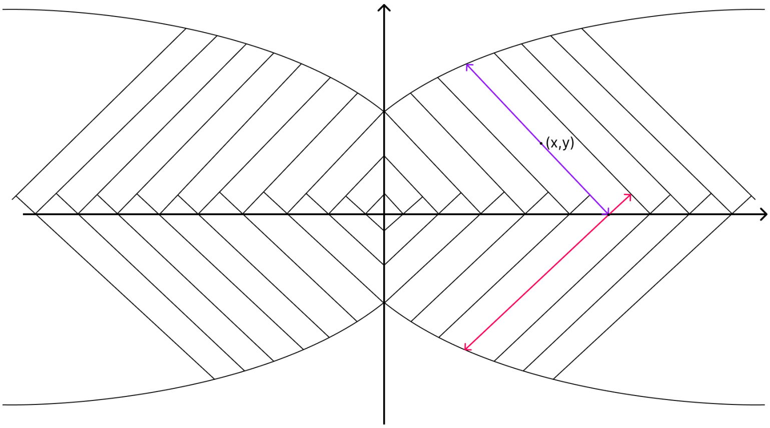

Fig. 2. The segments indicate the directions along which B is locally linear.

and below γ the function B should be linear along segments of slope -1. Since $B \in C^1$, we can write a formula for B in the set $\{(x,y) : \gamma(0)-x \leq y < \gamma(x)\}$: if $t \in [0, \gamma(x)]$, we must have

$$\begin{aligned} B(x+t,\gamma(x)-t) &= B(x,\gamma(x)) + B_x(x,\gamma(x)) \cdot t + B_y(x,\gamma(x)) \cdot (-t) \\ &= \gamma(x) - x^p - px^{p-1}t - t. \end{aligned} \tag{3.2}$$

Furthermore, from the assumed symmetry of the Bellman function, we deduce that $B_y(x,0) = 0$. Differentiating (3.2) with respect to x gives

$$\begin{aligned} \frac{\partial}{\partial x} B(x+t,\gamma(x)-t) &= B_x(x+t,\gamma(x)-t) + B_y(x+t,\gamma(x)-t) \cdot \gamma'(x) \\ &= \gamma'(x) - px^{p-1} - p(p-1)x^{p-2}t, \end{aligned}$$

so substituting $\gamma(x)=t$, we obtain

$$B_x(x+\gamma(x),0)=\gamma'(x)-px^{p-1}-p(p-1)x^{p-2}\gamma(x). \tag{3.3}$$

Similarly, differentiating (3.2) with respect to t gives

$$\frac{\partial}{\partial t}B(x+t,\gamma(x)-t)=B_x(x+t,\gamma(x)-t)-B_y(x+t,\gamma(x)-t)=-1-px^{p-1},$$

and setting $\gamma(x)=t$ yields $B_x(x+\gamma(x),0)=-1-px^{p-1}$. Combining this with (3.3) gives the following differential equation for γ:

$$\gamma'(x)=p(p-1)x^{p-2}\gamma(x)-1. \tag{3.4}$$

The general solution is given by

$$\gamma(x)=e^{px^{p-1}}\int_x^\infty e^{-pt^{p-1}}dt+ce^{px^{p-1}},$$

where $c\in\mathbb{R}$ is some parameter. To identify its value, let us study the asymptotic behavior of γ' at infinity. We check that

$$\gamma'(x)=p(p-1)x^{p-2}ce^{px^{p-1}}+p(p-1)x^{p-2}e^{px^{p-1}}\int_x^\infty e^{-pt^{p-1}}dt-1.$$

Now, the first term of this sum diverges to plus or minus infinity unless $c=0$. Furthermore, some straightforward calculations show that the remaining part of $\gamma'(x)$ converges to 0. Consequently, the only nontrivial (i.e., finite) limit is obtained for $c=0$. We will henceforth *assume* that this holds, leading to the final formula

$$\gamma(x)=e^{px^{p-1}}\int_x^\infty e^{-pt^{p-1}}dt.$$

Before we proceed to an explicit formula for B, let us establish the monotonicity of γ, which will play a crucial role later. We have

$$\gamma'(x)=p(p-1)x^{p-2}e^{px^{p-1}}\int_x^\infty e^{-pt^{p-1}}dt-1$$

and the substitution $u=pt^{p-1}-px^{p-1}$ gives

$$\gamma'(x)=\int_0^\infty\left(1+\frac{u}{px^{p-1}}\right)^{\frac{2-p}{p-1}}e^{-u}\,du-1>\int_0^\infty e^{-u}\,du-1=0.$$

Here we have used the estimates $\frac{2-p}{p-1}\geq 0$ and $1+\frac{u}{px^{p-1}}>1$ for $x>0$ and $u>0$. Thus, γ is strictly increasing.

Step 3: *Explicit candidate.* Equation (3.2) gives a formula for the Bellman function on the set $\{(x,y)\in\mathbb{R}^2:|x|+|y|\geq\gamma(0)\}$. It remains to handle the square $\{(x,y):|x|+|y|<\gamma(0)\}$. We assume that in this square, the function B is linear both along segments of slope -1 and along segments of slope +1. This choice is motivated by the form of Burkholder's function for the so-called weak-type inequality [4] (see also [12]). This assumption implies that B has the form $B(x,y)=\alpha(y^2-x^2)+\beta$ for some $\alpha,\beta\in\mathbb{R}$.

To determine α and β, recall that the desired Bellman function should be concave along segments with slopes in $[-1,1]$; in particular, for each y, the function $y\mapsto B(x,y)$ must be concave. Hence, the condition $\alpha\geq 0$ must hold. To determine the values of α and β, we use the continuity of B: setting $y=-x+\gamma(0)$ and referring to (3.2), we obtain

$$\gamma(0)-x=\alpha((-x+\gamma(0))^2-x^2)+\beta=\alpha(-2x\gamma(0)+\gamma^2(0))+\beta$$

for all $x\in[0,\gamma(0)]$. Since $\gamma(0)=\int_0^\infty e^{-pt^{p-1}}dt=p^{-1/(p-1)}\mathit{\Gamma}(p/(p-1))$, this forces us to take

$$\alpha=\left(2p^{-1/(p-1)}\mathit{\Gamma}\left(\frac{p}{p-1}\right)\right)^{-1},\qquad \beta=\frac{p^{-1/(p-1)}}{2}\mathit{\Gamma}\left(\frac{p}{p-1}\right).$$

Summarizing, we have obtained the following formula for the candidate B:

$$B(x,y)=\begin{cases}|y|-|x|^p & \text{if } |y|\geq\gamma(x),\\ \alpha(y^2-x^2)+\beta & \text{if } |x|+|y|\leq\gamma(0),\\ |y|+(p-1)f(|x|+|y|)^p-|x|pf(|x|+|y|)^{p-1} & \text{otherwise,}\end{cases}$$

where $f(x)=(x+\gamma(x))^{-1}$ is the inverse function of $x\mapsto x+\gamma(x)$. Here the last case is a simple reformulation of (3.2).

3.2. The inequality $\mathcal{B}^0\leq B$. In the light of Theorem 2.1, it is enough to check that the candidate B meets conditions 1° and 2°.

Proof of 2°. We must show that for any $(a,b)\in\mathbb{R}^2$ and $\varepsilon\in[-1,1]$, the function

$$G_{a,b,\varepsilon}(s)=B(a+s,b+\varepsilon s)$$

is concave on $\mathbb{R}$. Since $G_{a,b,\varepsilon}$ is C^1, it suffices to check that $G''_{a,b,\varepsilon}(s)\leq 0$ for all s where the second derivative exists. Moreover, due to the shift property $G_{a,b,\varepsilon}(s+t)=G_{a+t,b+\varepsilon t,\varepsilon}(s)$, it is enough to check that $G''_{a,b,\varepsilon}(0)\leq 0$ provided the second derivative exists. Taking into account the smoothness and symmetry, we may restrict ourselves to the case $a,b>0$.

We will first verify concavity for the boundary values $\varepsilon\in\{-1,1\}$. If $b\geq\gamma(a)$, there is nothing to prove, as $B(a,b)=b-a^p$. If $b<\gamma(a)$, then concavity along segments of slope -1 (in fact, even local linearity) follows directly from the construction. It remains to check concavity along segments of slope 1; we may restrict ourselves to the case $a+b>\gamma(0)$ (when $a+b\leq\gamma(0)$, local linearity holds in both directions ±1). We write G instead of $G_{a,b,\varepsilon}$; we have the explicit formula

$$G(s)=B(a+s,b+s)=b+s+(p-1)f(a+b+2s)^p-p(a+s)f(a+b+2s)^{p-1}.$$

After some lengthy, but rather straightforward calculations we obtain

$$G''(0) = p(p-1)x^{p-3}\bigg[(p-1)x\bigg(\frac{2}{1+\gamma'(x)}\bigg)^2 - \frac{4x^2\gamma''(x)}{(1+\gamma'(x))^3} - \frac{4x}{1+\gamma'(x)} \\ - (x+t)(p-2)\bigg(\frac{2}{1+\gamma'(x)}\bigg)^2 - (x+t)x\cdot\frac{-4\gamma''(x)}{(1+\gamma'(x))^3}\bigg],$$

where $x = f(a+b)$ and $t = a-x$. We will check that the expression in square brackets is nonpositive. For convenience, let us multiply this expression by $(1+\gamma'(x))^3 > 0$ and omit the remaining strictly positive factors, which does not affect the sign: after some tedious calculations, we obtain

$$-x\gamma'(x)(1+\gamma'(x)) + t(2-p)(1+\gamma'(x)) + tx\gamma''(x).$$

Plugging in

$$\gamma'(x) = p(p-1)x^{p-2}\gamma(x) - 1,$$
$$\gamma''(x) = p(p-1)(p-2)x^{p-3}\gamma(x) + p(p-1)x^{p-2}\gamma'(x)$$

and manipulating a little, we get

$$-p(p-1)x^{p-1}\gamma(x)\gamma'(x) + tp(p-1)x^{p-1}\gamma'(x) \\ = p(p-1)x^{p-1}\gamma'(x)(t-\gamma(x)).$$

Since $t \le \gamma(x)$, we conclude that $G''(0) \le 0$, which gives the concavity of B along lines of slope ±1.

We now consider the general form of concavity: we assume that the slope ε lies in the interval $[-1,1]$ (rather than in the set $\{-1,1\}$, as before). Assume that B is of class C^2 in a neighborhood of the point (a,b) (this holds for a dense subset of the first quadrant). We have

$$\frac{\partial^2}{\partial t^2}B(a+t, b+\varepsilon t)\bigg|_{t=0} = B_{xx}(a,b) + 2\varepsilon B_{xy}(a,b) + \varepsilon^2 B_{yy}(a,b).$$

This is a quadratic function of ε, and it is nonpositive for $\varepsilon \in \{-1,1\}$, as shown above. So, the general concavity will hold if we manage to prove that $B_{yy}(a,b) \ge 0$. Again, it suffices to restrict our analysis to the set $\{(a,b) : a > 0,\ b > 0,\ a+b > \gamma(0),\ b < \gamma(a)\}$, on which $B_{yy}(a,b)$, up to a positive factor, is equal to

$$\frac{p(p-1)x^{p-1}}{(1+\gamma'(x))^3}[\gamma(x) + t\gamma'(x)],$$

where, as before, $x = f(a+b)$ and $t = a-x$. Clearly, this expression is positive, which gives $B_{yy}(a,b) > 0$ and the concavity condition 2° follows. ■

Proof of 1°. By the symmetry of the functions B and V, we may restrict ourselves to the first quadrant. It suffices to observe that for each $x > 0$, the function

$$y \mapsto B(x,y) - V(x,y), \qquad y > 0,$$

is of class C^1, convex, and equals zero for sufficiently large arguments. If the function were to take negative values, its convexity would be violated because it would have to "bend" upward with a strictly increasing slope to eventually reach and maintain the zero value for large arguments. ■

3.3. The inequality $\mathcal{B}^0 \geq B$. Clearly, we may restrict ourselves to the set $\{(x,y) : |y| \leq \gamma(|x|)\}$, since on its complement we have $\mathcal{B}^0 \geq V = B$. Furthermore, by symmetry, we may assume that $x, y \geq 0$.

To establish the majorization $\mathcal{B}^0(x,y) \geq B(x,y)$, it suffices to construct an example of a pair of simple martingales $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ for which the value of the expression $\mathbb{E}V(f_\infty, g_\infty)$ is as close to $B(x,y)$ as desired. In the considerations below, we will be rather brief; for a similar detailed argument, we refer the interested reader to [2, 10]. The general idea is to move the pair $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$ over line segments along which B is linear.

To be more precise, suppose first that $x + y \geq \gamma(0)$ (and $y < \gamma(x)$, as we have already assumed above). Fix a small parameter $\delta > 0$ and consider a Markov, two-dimensional martingale $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$, whose distribution is uniquely determined by the following requirements:

(i) $(f_0, g_0) = (x, y)$ almost surely.

(ii) In the first step, (f_1, g_1) jumps along a segment of slope -1 so that it reaches the point on the curve $y = \gamma(x)$ or the point $(x+y, 0)$ on the x axis. That is, the pair makes a jump along a maximal line segment along which B is linear. Now, if the pair reaches the curve γ, it stops (we have entered the "stopping set" C). If it jumps to the x axis, the evolution continues.

(iii) If the pair reaches $(x+y, 0)$, it can no longer move along a segment of linearity of B: $(x+y, 0)$ is not in the interior of any such segment. Thus we are forced to accept a small deficit (it will be clear later what we mean by that). Namely, the next step of $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$ is taken along the slope 1, either to the curve $y = -\gamma(x)$, or to the point $(x + y + \delta, \delta)$.

(iv) If the process reaches the curve $y = -\gamma(x)$, it terminates; otherwise, it continues, in a similar fashion to (ii). Specifically, this time the point $(x+y+\delta, \delta)$ "splits" along a segment of slope -1, with endpoints on the curve $y = \gamma(x)$ and on the x axis; in this transition, there is no deficit. The pattern is then repeated.

(v) To ensure that the pair obtained forms a sequence of simple martingales, we assume that the process terminates after a certain (large) number N of steps.

Let us study the behavior of the sequence $(\mathbb{E}B(f_n,g_n))_{n\geq 0}$. By (i), this sequence starts from $B(x,y)$. Then, since in (ii) the pair moves along a segment over which B is linear, we deduce that $\mathbb{E}B(f_1,g_1)=B(x,y)$. Now, the deficit mentioned in (iii) means that $\mathbb{E}B(f_2,g_2)=\mathbb{E}B(f_1,g_1)+o(\delta)=B(x,y)+o(\delta)$. Here we use the fact that B is of class C^1. Repeating this argument, after N steps we obtain

$$\mathbb{E}B(f_N,g_N)=B(x,y)+(N-1)\cdot o(\delta).$$

Consequently, if N is odd, we may explicitly compute that

$$\begin{aligned}\mathbb{E}B(f_N,g_N)&=\mathbb{E}B(f_N,g_N)1_{\{|g_N|=\gamma(f_N)\}}+\mathbb{E}B(f_N,g_N)1_{\{|g_N|=0\}}\\&=\mathbb{E}V(f_N,g_N)1_{\{|g_N|=\gamma(f_N)\}}+B(x+y+(N-1)\delta,0)\mathbb{P}(|g_N|=0)\\&=\mathbb{E}V(f_N,g_N)+\big(B(x+y+(N-1)\delta,0)\\&\qquad\qquad-V(x+y+(N-1)\delta,0)\big)\mathbb{P}(|g_N|=0).\end{aligned}$$

To study this last quantity, note that since $0\geq B(a,0)\geq -|a|^p$ and $V(a,0)=-|a|^p$, we have

$$\big|B(x+y+(N-1)\delta,0)-V(x+y+(N-1)\delta,0)\big|\leq (x+y+(N-1)\delta)^p.$$

Next, the probability $\mathbb{P}(g_N=0)$ can be computed as follows. Note that on the set $\{g_N=0\}$ the martingale $f_0,f_1,\ldots,f_N$ is increasing; furthermore, we have

$$\begin{aligned}p_1&:=\mathbb{P}(f_1=x+y)=\frac{y}{x+y-f(x+y)},\\p_n&:=\mathbb{P}(f_n=x+y+(n-1)\delta\,|\,f_{n-1}=x+y+(n-2)\delta)\\&=\begin{cases}1-\frac{\delta}{x+y+(n-1)\delta-f(x+y+(n-2)\delta)} & \text{if } n\geq 2 \text{ is even,}\\ 1-\frac{\delta}{x+y+(n-1)\delta-f(x+y+(n-1)\delta)} & \text{if } n\geq 2 \text{ is odd.}\end{cases}\end{aligned}$$

Consequently, using $p_1\leq 1$, $1-x=e^{-x}+o(x)$ for x close to 0, and the approximation of integrals by Riemann sums, we get

$$\mathbb{P}(g_N=0)=p_1p_2\ldots p_N\leq \exp\bigg(-\int_{x+y}^{x+y+(N-1)\delta}\frac{1}{s-f(s)}\,ds\bigg)+N\cdot o(\delta).$$

The integral equals

$$\begin{aligned}\int_{x+y}^{x+y+N\delta}\frac{1}{s-f(s)}\,ds&=\int_{f(x+y)}^{f(x+y+(N-1)\delta)}p(p-1)t^{p-2}\,dt\\&=p\big(f(x+y+(N-1)\delta)^{p-1}-(x+y)^{p-1}\big).\end{aligned}$$

Putting all the above observations together, we handle the "error term" as follows:

$$\begin{aligned}|B(x+y+(N-1)\delta,0)-V(x+y+(N-1)\delta,0)|\mathbb{P}(|g_N|=0)\\ \le |x+y+(N-1)\delta|^p\\ \times\big[\exp\big(-p\big(f(x+y+(N-1)\delta)^{p-1}-(x+y)^{p-1}\big)\big)+N\cdot o(\delta)\big].\end{aligned}$$

However, recall that f is the inverse to the function $s\mapsto s+\gamma(s)$, and γ is a concave function with $\gamma'(s)\to 0$ for $s\to\infty$. This implies $f(s)>s/2$ for large s, and hence the error term can be made arbitrarily small if $N\delta$ is sufficiently large.

Thus, we proceed as follows. We fix an arbitrary $\varepsilon>0$ and pick a huge positive number K such that for $K'>K$ we have

$$|x+y+K'|^p\exp\big(-p\big(f(x+y+K')^{p-1}-(x+y)^{p-1}\big)\big)<\varepsilon.$$

Next, we pick a small δ and let N be the smallest integer for which $N\delta>K$. Then the error term above can be bounded by 2ε if δ is sufficiently small. Thus, returning to our previous calculations, we obtain

$$|\mathbb{E}V(f_N,g_N)-B(x,y)|\le 3\varepsilon,$$

on decreasing δ if necessary. This proves the desired sharpness, since ε was arbitrary.

If the initial point (x,y) lies in the square $|x|+|y|<\gamma(0)$, then the corresponding pair $((f_n)_{n\ge0},(g_n)_{n\ge0})$ moves according to the same rules as above, the only difference being that the first step is carried out over a line segment of slope 1, passing through (x,y) and joining two points on the boundary of the square. The structure of the processes is best illustrated in Figure 3. The comparison of $\mathbb{E}V(f_N,g_N)$ and $B(x,y)$ uses the same arguments as before.

There is an important observation, which was needed in the proof of Theorem 2.2 above.

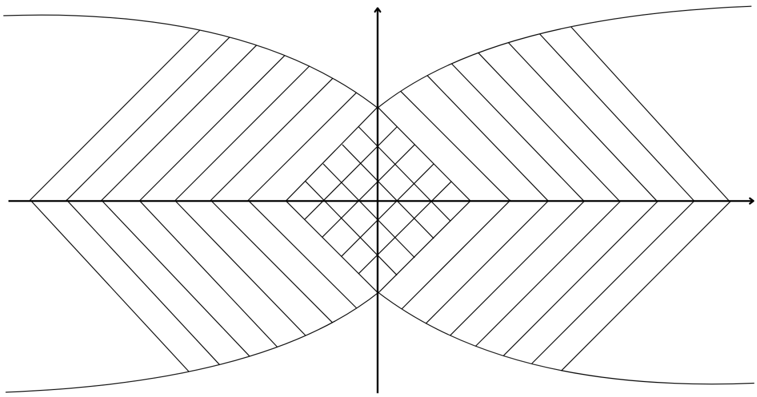

Fig. 3. The segments indicate directions along which B is locally linear.

THEOREM 3.1. *Suppose that* $(x, y) \in \mathbb{R}^2$, $\varepsilon > 0$ *and* $F \geq |x|$. *Then there exists* $\lambda > 0$ *and an* (*almost*) *extremal pair* $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$ *as above with the starting position* $(\lambda x, \lambda y)$ *and satisfying* $\|(f_n)_{n\geq 0}\|_p^p \geq (\lambda F)^p - \lambda\varepsilon$.

Proof. Fix the initial position $(x, y) \in \mathbb{R}^2$ of the (almost extremal) martingale pair $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$. It is clear from the above construction that if δ (the size of the jump) approaches zero and K tends to ∞, then the L^p norm of $(f_n)_{n\geq 0}$ converges to some constant $C_p(x, y)$; furthermore, the function C_p is continuous. It is not difficult to find a formula for C_p, but this is not necessary. All we need is that $C_p(x, y) = |x|^p$ if $|y| \geq \gamma(|x|)$ and $C_p(0, 0) > 0$. Therefore, for a given (x, y), the function

$$\xi(\lambda) = \frac{C_p(\lambda x, \lambda y)}{\lambda^p}, \qquad \lambda > 0,$$

satisfies $\lim_{\lambda\to 0+} \xi(\lambda) = \infty$ and $\lim_{\lambda\to\infty} \xi(\lambda) = |x|^p$. Consequently, for any $F \geq |x|$ we may find λ such that $\xi(\lambda) = F^p$, and then consider the pair $((f_n)_{n\geq 0}, (g_n)_{n\geq 0})$ corresponding to $(\lambda x, \lambda y)$. ∎

4. The case $p > 2$.

The analysis of this case will be carried out similarly to the case $1 < p \leq 2$. We begin by attempting to determine a candidate B for the Bellman function. We assume that B is of class C^1 on $\mathbb{R}^2$ and satisfies the symmetry condition $B(x, y) = B(|x|, |y|)$ for all x, y. This allows us to restrict our analysis to the first quadrant.

STEP 1: *The shape of the sets* C *and* $D_\pm$. As before, we assume the existence of a function $\gamma : [0, \infty) \to [0, \infty)$ such that $B(x, y) = y - x^p$ for $y \geq \gamma(x)$ and $B(x, y) > y - x^p$ for $y < \gamma(x)$. To gain some additional intuition, it is instructive to look at the behavior of γ in the case $1 < p < 2$ (which has already been analyzed above) for $p \uparrow 2$. It is easy to see that as p grows to 2, the curve γ increases slower and slower, and in the boundary case $p = 2$ it is constant. Thus, it is not surprising that for $p > 2$, the function γ will actually be decreasing. This "phase transition" for $p = 2$ affects also the behavior of the sets $D_\pm$. Some experimentation shows that now in the first quadrant the linearity (if it holds) should be experienced along line segments of slope 1. See Figure 1 for illustration.

STEP 2: *Formula for* γ. It will be convenient to divide our analysis into two parts, corresponding to the subregions $\{(x, y) : x \geq y\}$ and $\{(x, y) : x < y\}$.

The case $x \geq y$. As stated above, we postulate that for $y \geq \gamma(x)$, we have

$$B(x, y) = y - x^p. \tag{4.1}$$

Next, we require that below the curve γ the function B be linear along certain segments. A little experimentation shows that in the case $p>2$, we should consider segments of slope 1. This leads to the following formula for B in the region $\{(x,y): y\leq x,\, y<\gamma(x)\}$:

$$B(x-t,\gamma(x)-t)=B(x,\gamma(x))+B_x(x,\gamma(x))\cdot(-t)+B_y(x,\gamma(x))\cdot(-t) \\ =\gamma(x)-x^p+px^{p-1}t-t \tag{4.2}$$

for $t\in[0,\gamma(x))$. Differentiating the above equation with respect to x, we obtain

$$\frac{\partial}{\partial x}B(x-t,\gamma(x)-t)=B_x(x-t,\gamma(x)-t)+B_y(x-t,\gamma(x)-t)\cdot\gamma'(x) \\ =\gamma'(x)-px^{p-1}+p(p-1)x^{p-2}t. \tag{4.3}$$

Substituting $t=\gamma(x)$ and noting that $B_y(x-\gamma(x),0)=0$ (which comes from the symmetry of the Bellman function) we obtain

$$B_x(x-\gamma(x),0)=\gamma'(x)-px^{p-1}+p(p-1)x^{p-2}\gamma(x). \tag{4.4}$$

Let us emphasize that here we crucially use the assumption that $x\geq y$: thanks to this condition, the point $(x-\gamma(x),0)$ remains in the first quadrant.

Similarly, if we differentiate the formula for B with respect to t and set $t=\gamma(x)$, we find that $B_x(x-\gamma(x),0)=1-px^{p-1}$. Combining this with (4.4), we obtain the differential equation for $\gamma(x)$:

$$\gamma'(x)=-p(p-1)x^{p-2}\gamma(x)+1. \tag{4.5}$$

Its general solution is

$$\gamma(x)=e^{-px^{p-1}}\int_0^x e^{pt^{p-1}}dt+ce^{-px^{p-1}},$$

where $c\in\mathbb{R}$ is a parameter. To find it, it will be helpful to study its interplay with the objects coming from the case $y>x$.

The case $y>x$. We proceed as before: we know that for $y\geq\gamma(x)$, equation (3.1) holds, while for $y<\gamma(x)$, formula (4.2) is true. Thus, in particular, for $y<\gamma(x)$ the identity (4.3) is also satisfied. Setting $t=x$ and using the equation $B_x(0,\gamma(x)-x)=0$ (which follows from the symmetry of B), we obtain

$$B_y(0,\gamma(x)-x)\cdot\gamma'(x)=\gamma'(x)-px^{p-1}+p(p-1)x^{p-1}. \tag{4.6}$$

Similarly, differentiating (4.2) with respect to t and setting $t=x$ yields

$$B_y(0,\gamma(x)-x)=1-px^{p-1}. \tag{4.7}$$

Combining (4.6) with (4.7), we obtain the differential equation

$$\gamma'(x)(1-px^{p-1})=\gamma'(x)-px^{p-1}+p(p-1)x^{p-1}, \tag{4.8}$$

which simplifies to $\gamma'(x) = 2 - p$. Therefore, $\gamma(x) = (2-p)x + a$, where $a \in \mathbb{R}_+$ is another parameter.

The values of a and c. From the analysis in the previous section, we have determined the formula for γ up to the parameters a and c. Namely, if $\gamma(x) > x$, we have

$$\gamma(x) = \gamma_1(x) = (2-p)x + a, \tag{4.9}$$

while for $\gamma(x) \leq x$, we have obtained

$$\gamma(x) = \gamma_2(x) = e^{-px^{p-1}} \int_0^x e^{pt^{p-1}}\,dt + ce^{-px^{p-1}}. \tag{4.10}$$

To determine a and c, we will smoothly “glue” the functions γ_1 and γ_2 along the line $y = x$ (i.e., ensuring that their derivatives also match). To this end, note that the equality $\gamma(x) = x$ is equivalent to $x = a/(p-1)$; that is, the curves γ_1 and γ_2 meet at the point $(a/(p-1), a/(p-1))$. It remains to exploit the equalities $\gamma_1'\big(\frac{a}{p-1}\big) = \gamma_2'\big(\frac{a}{p-1}\big)$ and $\gamma_2\big(\frac{a}{p-1}\big) = \frac{a}{p-1}$. Since $\gamma_1'(x) = 2 - p$ and

$$\gamma_2'(x) = -p(p-1)x^{p-2}\gamma_2(x) + 1,$$

these two equalities yield

$$a = (p-1)p^{\frac{1}{1-p}} \tag{4.11}$$

and

$$p^{\frac{1}{1-p}} = \gamma_2(p^{\frac{1}{1-p}}) = e^{-pp^{-1}} \int_0^{p^{\frac{1}{1-p}}} e^{pt^{p-1}}\,dt + ce^{-pp^{-1}},$$

or after some straightforward manipulations,

$$c = ep^{\frac{1}{1-p}} - p^{\frac{1}{1-p}} \int_0^1 e^{s^{p-1}}\,ds. \tag{4.12}$$

We conclude this part by studying some important properties of γ. First, by a simple application of the L'Hospital rule, we check that both

$$\gamma(x) = e^{-px^{p-1}} \int_0^x e^{pt^{p-1}}\,dt + ce^{-px^{p-1}}$$

and

$$\gamma'(x) = -p(p-1)x^{p-2}ce^{-px^{p-1}} - p(p-1)x^{p-2}e^{-px^{p-1}} \int_0^x e^{pt^{p-1}}\,dt + 1$$

converge to zero as $x \to \infty$. Secondly, it is not difficult to verify that γ is a nonincreasing function. Indeed, for $x < p^{1/(1-p)}$ we have $\gamma'(x) = 2 - p < 0$, while for $x > p^{1/(1-p)}$, we check that the inequality $\gamma'(x) \le 0$ is equivalent to

$$h(x) = \int_0^x e^{pt^{p-1}}dt + c - \frac{x^{2-p}e^{px^{p-1}}}{p(p-1)} \ge 0.$$

It suffices to verify that $h(p^{\frac{1}{1-p}}) \ge 0$ and $h'(x) \ge 0$ for $x > p^{1/(1-p)}$. We compute

$$h(p^{\frac{1}{1-p}}) = \int_0^{p^{\frac{1}{1-p}}} e^{pt^{p-1}}dt + c - \frac{p^{\frac{2-p}{1-p}}e^{-1}}{p(p-1)}$$

and plug in the value of c (see (4.12)) to obtain

$$h(p^{\frac{1}{1-p}}) = \frac{p^{\frac{1}{1-p}}}{e(p-1)}\left(e^2(p-1)-1\right) \ge \frac{p^{\frac{1}{1-p}}}{e(p-1)}(e^2-1) \ge 0.$$

Furthermore,

$$h'(x) = e^{px^{p-1}} - e^{px^{p-1}} - \frac{(2-p)x^{1-p}e^{px^{p-1}}}{p(p-1)} = \frac{(p-2)x^{1-p}e^{px^{p-1}}}{p(p-1)} \ge 0.$$

This establishes the monotonicity of γ.

Step 3: *The explicit candidate.* Summarizing our analysis so far, we have obtained a formula for the candidate B. If $|y| \ge \gamma(|x|)$, then

$$B(x,y) = |y| - |x|^p.$$

If $|x| \ge |y|$ and $|y| < \gamma(|x|)$, then

$$B(x,y) = y + (p-1)f(x-y)^p + xpf(x-y)^{p-1} - (p-1)p^{\frac{p}{1-p}}.$$

Finally, for $|x| \le |y| < \gamma(x)$, we have

$$\begin{aligned} B(x,y) = y + (p-1)&\left[\frac{y-x-(p-1)p^{\frac{1}{1-p}}}{1-p}\right]^p \\ &- xp\left[\frac{y-x-(p-1)p^{\frac{1}{1-p}}}{1-p}\right]^{p-1}, \end{aligned}$$

where the function $f(x) = (x-\gamma(x))^{-1}$ is the inverse of $\xi : [a/(p-1),\infty) \to [0,\infty)$, given by $\xi(x) = x - \gamma(x)$.

In the remaining part of this section, we will verify that B coincides with the Bellman function $\mathcal{B}^0$ by proving the inequalities $\mathcal{B}^0 \le B$ and $\mathcal{B}^0 \ge B$.

4.1. The inequality $\mathcal{B}^0 \le B$. We will show that the candidate satisfies conditions 1° and 2°.

Proof of 2° *for* $\varepsilon = \pm 1$. As in the case $1 < p \le 2$, first we show that for any $(a,b) \in \mathbb{R}^2$ and any $\varepsilon \in \{-1,1\}$, the function

$$G(s) = B(a+s, b+\varepsilon s)$$

is concave on $\mathbb{R}$. Arguing as before, it is enough to prove that $G''(0) \leq 0$ provided the derivative exists. By symmetry, we may assume that $a, b > 0$. If $b \geq \gamma(a)$, the result is immediate since $B(a, b) = b - a^p$. If $b < \gamma(a)$, then concavity along lines of slope $+1$ follows directly from the construction. Thus it remains to verify concavity along lines of slope -1.

Now, if $b \leq a$, we compute

$$\begin{aligned} G''(0) = p(p-1)x^{p-3}\Bigg[&(p-1)x(\frac{2}{1-\gamma'(x)})^2 + x^2\frac{4\gamma''(x)}{(1-\gamma'(x))^3} - 2x\frac{2}{1-\gamma'(x)} \\ &- (x+t)(p-2)\left(\frac{2}{1-\gamma'(x)}\right)^2 - (x-t)x\frac{4\gamma''(x)}{(1-\gamma'(x))^3}\Bigg], \end{aligned}$$

where $x = f(a-b)$ and $t = x - a$. We must show that the expression in square brackets is nonpositive. Multiplying by $(1-\gamma'(x))^3$ and using the identities

$$\begin{aligned} \gamma'(x) &= -p(p-1)x^{p-2}\gamma(x) + 1, \\ \gamma''(x) &= -p(p-1)(p-2)x^{p-3}\gamma(x) - p(p-1)x^{p-2}\gamma'(x), \end{aligned}$$

we obtain

$$p(p-1)x^{p-1}\gamma(x)\gamma'(x) - tp(p-1)x^{p-1}\gamma'(x) = p(p-1)x^{p-1}\gamma'(x)(\gamma(x) - t).$$

It remains to note that $\gamma(x) - t > 0$ (this is equivalent to $b > 0$) and $\gamma'(x) \leq 0$: this gives $G''(0) \leq 0$.

On the other hand, if $b \geq a$, after some straightforward calculations we obtain

$$G''(s) = -\frac{4(a+s)p(p-2)}{p-1}\left[\frac{b-a-2s-(p-1)p^{\frac{1}{1-p}}}{1-p}\right]^{p-3}.$$

Since $s \geq 0$ and $b - a = (1-p)x + (p-1)p^{\frac{1}{1-p}}$, we have

$$\frac{b-a-2s-(p-1)p^{\frac{1}{1-p}}}{1-p} = \frac{(1-p)x-2s}{1-p} = x + \frac{2s}{p-1} \geq 0,$$

so the term in square brackets is nonnegative. Finally, since $a, s \geq 0$, we see that $G''(s) \leq 0$, and concavity follows. ∎

Proof of 2°, *the general case.* Arguing as in the case $1 < p \leq 2$, it suffices to verify that $B_{yy} \geq 0$ in the first quadrant below the curve γ. We consider two cases. If $b \leq a$, then

$$B_y(a, b) = 1 - p(p-1)f(a-b)^{p-1}f'(a-b) + ap(p-1)f(a-b)^{p-2}f'(a-b)$$

and

$$\begin{aligned} B_{yy}(a, b) = p(p-1)f(a-b)^{p-3}[&(p-1)f(a-b)f'(a-b)^2 \\ &+ f(a-b)^2f''(a-b) - a(p-2)f'(a-b)^2 - af(a-b)f''(a-b)]. \end{aligned}$$

Since

$$f'(x) = -\frac{1}{1-\gamma'(f(x))} \quad \text{and} \quad f''(x) = \left(\frac{1}{1-\gamma'(f(x))}\right)^3 \cdot \gamma''(f(x)),$$

we check that $B_{yy}(a,b)/(p(p-1)x^{p-3})$ simplifies to

$$\frac{x}{(1-\gamma'(x))^2} + \frac{t(p-2)}{(1-\gamma'(x))^2} + \frac{tx\gamma''(x)}{(1-\gamma'(x))^3}.$$

Now we plug in

$$\begin{aligned}\gamma'(x) &= -p(p-1)x^{p-2}\gamma(x) + 1,\\ \gamma''(x) &= -p(p-1)(p-2)x^{p-3}\gamma(x) - p(p-1)x^{p-2}\gamma'(x)\end{aligned}$$

to deduce that the above expression equals

$$\frac{p(p-1)x^{p-1}}{(1-\gamma'(x))^3}[\gamma(x) - t\gamma'(x)],$$

which is clearly positive.

It remains to study the case $b > a$. A direct differentiation shows that

$$B_{yy}(a,b) = \frac{p}{p-1}\left[\frac{b-a-(p-1)p^{\frac{1}{1-p}}}{1-p}\right]^{p-3}(-b+a(3-p)+(p-1)p^{\frac{1}{1-p}}).$$

Note that the expression in square brackets is positive: rewriting it in terms of x and t gives

$$\frac{b-a-(p-1)p^{\frac{1}{1-p}}}{1-p} = \frac{(2-p)x+p^{\frac{1}{1-p}}-t-x-t-p^{\frac{1}{1-p}}}{1-p} = x + \frac{2t}{p-1} \geq 0$$

for $x, t \geq 0$. As for the sign of the expression $-b+a(3-p)+(p-1)p^{\frac{1}{1-p}}$, we pass to the variables x and t again, obtaining

$$\begin{aligned}&-b+a(3-p)+(p-1)p^{\frac{1}{1-p}}\\ &\qquad = -(2-p)x-(p-1)p^{\frac{1}{1-p}}+t+(x-t)(3-p)+(p-1)p^{\frac{1}{1-p}}\\ &\qquad = x+t(p-2) \geq x-t \geq 0.\end{aligned}$$

This proves the inequality $B_{yy}(a,b) \geq 0$ and completes the proof of 2°. ■

Proof of 1°. The argument is the same as in the case of $1 < p \leq 2$. ■

4.2. The inequality $\mathscr{B}^0 \geq B$. In order to establish the inequality $\mathscr{B}^0(x,y) \geq B(x,y)$ (for a fixed $(x,y) \in \mathbb{R}^2$), it suffices to construct an example of a pair of simple martingales $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ for which $\mathbb{E}V(f_\infty, g_\infty)$ is as close to $B(x,y)$ as desired. Once again, we omit tedious computational details and instead provide an outline of the structure of such extremal martingales. As before, the idea is to enforce the movement of the martingales along the directions where the function B is linear, until they reach the set $C = \{(x,y) : B(x,y) = V(x,y)\}$. See Figure 4.

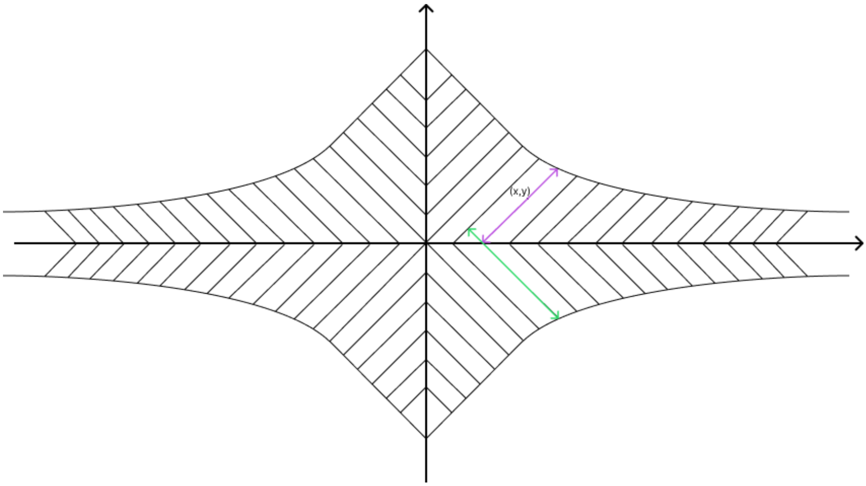

Fig. 4. The segments indicate the directions along which B is locally linear.

Again, we will restrict ourselves to the case $x, y \geq 0$. Let $\delta > 0$ be a small parameter. In the present case, the requirements are:

(i) The pair starts from the point (x, y).

(ii) The first jump is performed along a segment with slope -1 (along which B is linear) until its endpoints, i.e., either to the curve $y = \gamma(x)$ or to one of the axes. If the pair reaches $\gamma(x)$, it stops (since it has reached the set C).

(iii) We consider two cases. If, as a result of (ii), the pair arrives at the point $(x-y, 0)$ on the x-axis, the next jump is performed along a segment with slope -1, either to the curve $y = -\gamma(x)$ or to the point $(x - y - \delta, \delta)$. The second case corresponds to the situation in which the pair comes to $(0, y - x)$, as a result of (ii). Then it moves along a line segment of slope -1 until it gets to the curve $y = \gamma(-x)$ or to $(\delta, y - x - \delta)$.

(iv) If at any point the pair reaches $(0, 0)$, then in the next step it jumps along the line $y = x$ until it reaches one of the points $(x, \gamma(x))$ or $(-x, -\gamma(x))$ (for the appropriate x). The pair then terminates its evolution.

(v) The pattern is then repeated: the points on the curve $|y| = \gamma(|x|)$ are absorbing, and for other points we apply the above rules (ii) and (iii).

(vi) To ensure that the constructed pair is a simple martingale pair, we assume that the evolution process terminates after a certain number N of steps.

As in the case $1 < p \leq 2$, it can be verified that by choosing sufficiently large N and sufficiently small δ, the difference between $\mathbb{E}V(f_\infty, g_\infty)$ and $B(x, y)$ can be made arbitrarily small. This establishes the desired inequality $\mathcal{B}^0 \geq B$. Furthermore, it is not difficult to see that Theorem 3.1 remains valid.

Acknowledgements. I would like to sincerely thank Professor Adam Osękowski for all his support and guidance throughout the writing of this paper. His incredible patience, clear explanations, and constant encouragement made a huge difference, helping me understand even the most challenging parts. I am truly grateful for the time and effort he devoted to helping me, and for making this a rewarding and enjoyable learning experience.

REFERENCES

[1] R. Bañuelos, *The foundational inequalities of D. L. Burkholder and some of their ramifications*, Illinois J. Math. 54 (2010), 789–868.

[2] R. Bañuelos and A. Osękowski, *Stability in Burkholder's differentially subordinate martingales inequalities and applications to Fourier multipliers*, J. Math. Pures Appl. 119 (2018), 1–44.

[3] R. Bellman, *Dynamic Programming*, Princeton Landmarks in Math., Princeton Univ. Press, Princeton, NJ, 2010 (reprint of the 1957 edition).

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

[5] D. L. Burkholder, *An extension of a classical martingale inequality*, in: Probability Theory and Harmonic Analysis, Dekker, New York, 1986, 21–30.

[6] D. L. Burkholder, *A sharp and strict L^p-inequality for stochastic integrals*, Ann. Probab. 15 (1987), 268–273.

[7] K. P. Choi, *Some sharp inequalities for martingale transforms*, Trans. Amer. Math. Soc. 307 (1988), 279–300.

[8] 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.

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

[10] A. Osękowski, *Sharp LlogL inequalities for differentially subordinated martingales and harmonic functions*, Illinois J. Math. 52 (2008), 745–756.

[11] A. Osękowski, *Sharp tail inequalities for nonnegative submartingales and their strong differential subordinates*, Electron. Comm. Probab. 15 (2010), 508–521.

[12] A. Osękowski, *Sharp moment inequalities for differentially subordinated martingales*, Studia Math. 201 (2010), 103–131.

[13] A. Osękowski, *Sharp Martingale and Semimartingale Inequalities*, Monograf. Mat. (N.S.) 72, Birkhäuser, Basel, 2012.

[14] A. Osękowski, *Survey article: Bellman function method and sharp inequalities for martingales*, Rocky Mountain J. Math. 43 (2013), 1759–1823.

[15] A. Osękowski, *Some sharp estimates for the Haar system and other bases in $L^1(0,1)$*, Math. Scand. 115 (2014), 123–142.

[16] V. Vasyunin and A. Volberg, *Burkholder's function via Monge–Ampère equation*, Illinois J. Math. 54 (2010), 1393–1428.
[17] V. Vasyunin and A. Volberg, *The Bellman Function Technique in Harmonic Analysis*, Cambridge Stud. Adv. Math. 186, Cambridge Univ. Press, Cambridge, 2020.
[18] J. Ville, *Étude critique de la notion de collectif*, Gauthier-Villars, Paris, 1939.

Barbara Bugała
University of Warsaw
02-097 Warszawa, Poland
E-mail: basiab.bugala@gmail.com