University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Michał Brzozowski

Sharp weighted inequalities for martingales

Supervisor
dr hab. Adam Osękowski
Institute of Mathematics
University of Warsaw

May 2020

Author's declaration:
I hereby declare that this dissertation is my own work.

....
Date Signature

Supervisor's declaration:
The dissertation is ready to be reviewed

....
Date Signature

Summary

The thesis is devoted to the study of weak-type, strong-type and maximal estimates for martingale transforms in the weighted context. We put a particular emphasis to the following two aspects of the subject:

I. To identify the optimal dependence of the constants on Muckenhoupt's characteristics of the weights involved;
II. To establish the results with the use of the Bellman function method.

Our interest in I. stems from the applications in the PDEs. For example, the aforementioned dependence on the characteristic is related to the regularity of solutions to certain elliptic differential equations; via extrapolation, this topic is also linked to sharpness of some other estimates. The motivation for II. comes from the unweighted setting, in which the Bellman function method has been developed very intensively during the last forty years. We expect that the special functions and the approach developed in the thesis can be applied in the study of related bounds in probability theory and harmonic analysis.

It should be emphasized that the content of the thesis is quite technical. Even in the simpler unweighted setting the analysis of the special functions of two or three variables can be quite elaborate, as evidenced in many papers in the literature. The passage to the truly weighted context adds two extra variables to the picture and makes the calculations really involved at many places. In some cases these computations seem unavoidable, but for some estimates we manage to develop methods which enable significant simplification.

The material is organized as follows. The next chapter has a preliminary character and contains some motivation, the description of the necessary background and notation. The main contribution has been placed in Chapters 2, 3, 4 and 5.

Chapter 2 is concerned with the analysis of tight weighted strong-type (p, p) estimates $(1 < p < \infty)$ for martingale transforms. The contents of this part is taken from the joint paper with R. Bañuelos and A. Osękowski [5].

Chapter 3 studies the tight weighted weak-type (p, p) estimates $(1 < p < \infty)$ for martingale transforms. It is the most technical part of the thesis, the analysis of the Bellman function constructed there is quite intricate and rests on the careful investigation of appropriate 4×4 matrices. The material is taken from a joint work with A. Osękowski (in preparation).

Chapter 4 is the continuation of Chapter 3 and contains the investigation of weighted weak-type (p, p) estimates in the endpoint case $p = \infty$. We have decided to insert the analysis in a separate chapter, since the arguments are different from those in Chapter 3 and exploit certain novel composition/extrapolation of simpler Bellman functions. The material is taken from the paper [12] written jointly with A. Osękowski.

In the final part of the thesis, Chapter 5, we investigate weighted maximal L^1 estimates for martingale transforms, which can be regarded as endpoint complements of the result from Chapter 2. The material is taken from the work [13].

We have tried to keep this thesis as self-contained as possible, providing unproved statements with convenient references.

2010 Mathematics Subject Classification. Primary: 60G44. Secondary: 60G42.
Keywords and phrases: martingale, weight, differential subordination, Bellman function.

Streszczenie

Niniejsza rozprawa poświęcona jest zbadaniu nierówności słabego typu, silnego typu oraz nierówności maksymalnych dla transformat martyngałowych w kontekście ważonym. W szczególności interesują nas następujące dwa problemy:

I. Zidentyfikowanie optymalnej zależności stałych od charakterystyki wagi;
II. Uzyskanie wyników z wykorzystaniem metody Bellmana.

Motywacją dla zbadania problemu I. są zastosowania w teorii równań różniczkowych cząstkowych. Wspomniana zależność od charakterystyki wagi jest powiązana z regularnością rozwiązań pewnych eliptycznych równań różniczkowych; jest ona również powiązana z optymalnością pewnych innych oszacowań. Powodem, dla którego istotna jest kwestia II. są nierówności bezwagowe, w których z powodzeniem korzystano z metody Bellmana w przeciągu ostatnich czterdziestu lat. Oczekujemy, że funkcje specjalne i argumentacja przedstawiona w rozprawie mogą zostać zastosowane także w badaniu pokrewnych problemów w rachunku prawdopodobieństwa i analizie harmonicznej.

Należy pokreślić, że rozprawa jest miejscami dosyć techniczna. Nawet w prostszym przypadku bezwagowym, analiza pojawiających się funkcji specjalnych dwóch lub trzech zmiennych bywa złożona, co odzwierciedla wiele prac z tej dziedziny. Przejście do kontekstu ważonego dodaje dwie kolejne zmienne i sprawia, że dowody stają się bardzo skomplikowane. Czasami nie udało się nam uniknąć złożonych, technicznych rozważań, jednak w niektórych przypadkach dowód został znacząco uproszczony dzięki wykorzystaniu nowych konstrukcji i argumentów.

Rozprawa jest zorganizowana w następujący sposób. Pierwszy rozdział ma wstępny charakter i zawiera motywację oraz niezbędne definicje. Główne wyniki są zamieszczone w Rozdziałach 2, 3, 4 oraz 5.

Rozdział 2 poświęcony jest analizie ważonych nierówności mocnego typu (p, p) dla transformat martyngałowych, gdzie $1 < p < \infty$. Rozdział jest oparty na wynikach ze wspólnego artykułu z R. Bañuelosem i A. Osękowskim [5].

W Rozdziale 3 badamy ważone oszacowania słabego typu (p, p) dla transformat martyngałowych, gdzie $1 < p < \infty$. Jest to najbardziej techniczna część rozprawy, analiza skonstruowanej funkcji Bellmana jest dość złożona i opiera się na precyzyjnej analizie rozmaitych macierzy 4×4. Wynik pochodzi ze wspólnej pracy z A. Osękowskim (w przygotowaniu).

Rozdział 4 stanowi kontynuację Rozdziału 3 i zawiera analizę nierówności słabego typu (p, p) w przypadku krańcowym $p = \infty$. Zdecydowaliśmy się umieścić ten wynik w osobnym rozdziale, ponieważ zastosowana metoda różni się od tej w Rozdziale 3 i wykorzystuje pewne nowe złożenia/ekstrapolację za pomocą prostszych funkcji Bellmana. Rozdział opiera się na publikacji [12] napisanej wspólnie z A. Osękowskim.

W ostatniej części pracy, Rozdziale 5, badamy ważoną nierówność maksymalną L^1 dla transformat martyngałowych, którą można interpretować jako krańcowy przypadek wyniku z Rozdziału 2. Materiał zaczerpnięto z pracy [13].

2010 Mathematics Subject Classification. Primary: 60G44. Secondary: 60G42.
Słowa kluczowe: martyngał, waga, silna dominacja, funkcja Bellmana.

Acknowledgments

I would like to express my gratitude to my supervisor Adam Osękowski for his patience, guidance and valuable insights. I am deeply grateful for introducing me to the Bellman function method and for all interesting discussions about probability and harmonic analysis.

I also want to thank my family for support and encouragement.

The research presented in this thesis was partially supported by the National Science Centre, Poland, via the Sonata grant no. 2014/14/E/ST1/00532.

Contents

Chapter 1

Preliminaries

1.1. Basic definitions and motivation

Inequalities for the Haar system. The thesis studies several weighted extensions of the celebrated inequalities for martingale transforms established by Burkholder more than fifty years ago. The topic has its roots in the following properties of the classical Haar system $(h_n)_{n\geq 0}$. Recall that this collection consists of functions on $[0,1)$ given by

$$
\begin{aligned}
h_0 &= [0,1), & h_1 &= [0,1/2) - [1/2,1),\\
h_2 &= [0,1/4) - [1/4,1/2), & h_3 &= [1/2,3/4) - [3/4,1),\\
h_4 &= [0,1/8) - [1/8,1/4), & h_5 &= [1/4,3/8) - [3/8,1/2),\\
h_6 &= [1/2,5/8) - [5/8,3/4), & h_7 &= [3/4,7/8) - [7/8,1)
\end{aligned}
$$

and so on (here we have identified a set with its indicator function). Schauder proved in [74] that this family is a basis in the space $L^p(0,1)$ (with the underlying Lebesgue's measure) for any $1 \leq p < \infty$. A classical inequality due to Marcinkiewicz [50], based on the earlier work of Paley [67], can be stated as follows. If $1 < p < \infty$, then there is a finite constant c_p depending only on p such that for all $n \geq 0$ and any numbers $a_0, a_1, \ldots, a_n \in \mathbb{R}$, $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n \in \{-1,1\}$ we have

$$\left\| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right\|_{L^p(0,1)} \leq c_p \left\| \sum_{k=0}^{n} a_k h_k \right\|_{L^p(0,1)}. \tag{1.1.1}$$

This remarkable property is called unconditionality and plays a significant role in harmonic analysis, approximation theory and geometry of Banach spaces: see e.g. the monographs [38, 75, 82] and consult the references therein.

Inequalities for transforms of discrete-time martingales. In 1966 Burkholder showed that the inequality (1.1.1) can be reformulated and generalized to the probabilistic setting. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space endowed with a discrete-time filtration, that is, a non-decreasing sequence $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots$ of sub-σ-algebras of $\mathcal{F}$. Let $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ be adapted real-valued martingales, i.e., sequences of integrable random variables satisfying $\mathbb{E}(f_{n+1}|\mathcal{F}_n) = f_n$ and $\mathbb{E}(g_{n+1}|\mathcal{F}_n) = g_n$ for each $n \geq 0$. Define the associated difference sequences $df = (df_n)_{n\geq 0}$, $dg = (dg_n)_{n\geq 0}$ by $df_0 = f_0$ and $df_n = f_n - f_{n-1}$ for $n = 1, 2, \ldots$, and analogously for dg. Equivalently, we have the identities

$$f_n = \sum_{k=0}^{n} df_k, \qquad g_n = \sum_{k=0}^{n} dg_k, \qquad n = 0,\ 1,\ 2,\ \ldots.$$

We say that g is a martingale transform of f, if there exists a predictable sequence $v = (v_n)_{n\geq 0}$ such that $dg_n = v_n df_n$ for each $n \geq 0$. Here by predictability of v we mean that for each $n \geq 0$, the random variable v_n is measurable with respect to $\mathcal{F}_{(n-1)\vee 0}$. Moreover, if the values of v are in the interval $[-1, 1]$, then we will say that g is a ± 1-transform of f. There is a very natural general problem: to compare magnitudes of a martingale f and its ± 1-transform g, where the size is measured in terms of norms in various function spaces (e.g., L^p-norms, weak-L^p norms, etc.). The aforementioned result of Burkholder [15] is the following.

Theorem 1.1. *For any $1 < p < \infty$, there is a constant C_p depending only on p such that if f is a martingale and g is its ± 1-transform, then*

$$||g_n||_{L^p} \leq C_p ||f_n||_{L^p}, \qquad n = 0, 1, 2, \ldots. \tag{1.1.2}$$

This does extend (1.1.1). Indeed, if the probability space is the interval $[0, 1)$ with its Borel subsets and Lebesgue's measure, then the Haar system $(h_n)_{n\geq 0}$, and hence also any collection of the form $(a_n h_n)_{n\geq 0}$, are martingale differences sequences (relative to the natural filtration). It turns out that the above theorem has many profound applications in harmonic analysis, for example it yields inequalities for the Hilbert transform, Riesz transforms and a wider class of singular operators and Fourier multipliers (see e.g. [4, 51]). Because of these applications, it is of interest and importance to identify the optimal (i.e., the least possible) value of the constant C_p. This problem was successfully handled by Burkholder in [19]. He developed a unified approach to estimates of such type and applied it to obtain the following statement.

Theorem 1.2. *Let $1 < p < \infty$ and $p^* = \max\left\{p, \frac{p}{p-1}\right\}$. For every martingale f and its ± 1-transform g, we have*

$$||g_n||_{L^p} \leq (p^* - 1)||f_n||_{L^p}, \qquad n = 0,\ 1,\ 2,\ \ldots. \tag{1.1.3}$$

The constant $p^* - 1$ is optimal, even in the special case (1.1.1) *(in other words, the unconditional L^p-constant of the Haar system is equal to $p^* - 1$).*

The above moment inequality fails to hold for $p \in \{1, \infty\}$, which gives rise to the question about some substitutes in these endpoint cases. It is natural to take a look at the corresponding weak-type estimates. We state them below in the full range $1 \leq p \leq \infty$:

$$||g_n||_{L^{p,\infty}} \leq C'_p ||f_n||_{L^p}, \qquad n = 0,\ 1,\ 2,\ \ldots. \tag{1.1.4}$$

Here $||\xi||_{L^{p,\infty}} = \sup(\lambda^p \mathbb{P}(|\xi| \geq \lambda))^{\frac{1}{p}}$ stands the weak p-norm of ξ, $1 \leq p < \infty$; the definition in the case $p = \infty$ requires some additional notation and will be presented in the next paragraph (in the literature, sometimes one puts $L^{\infty,\infty} = L^\infty$; however, we will work with a different space weak-L^∞). The case $p = 1$ of (1.1.4) was handled by Burkholder [18] in the late seventies: the optimal value of the constant C'_1 is equal to 2. The paper [19] contains the extension of this result to the range $1 \leq p \leq 2$, and the version for $p > 2$ (which turned out to be much more complicated) was successfully tackled by Suh in [81].

Theorem 1.3. *The optimal value of the constant C'_p equals $\left(\frac{2}{\Gamma(p+1)}\right)^{1/p}$ for $1 \leq p \leq 2$ and $\left(\frac{p^{p-1}}{2}\right)^{1/p}$ for $2 \leq p < \infty$. This is already optimal in the context of the Haar system.*

We turn our attention to the version of (1.1.4) for $p = \infty$. Let ξ be a random variable on a non-atomic probability space. We define ξ^*, the decreasing rearrangement of ξ, by

$$\xi^*(t) = \inf\{\lambda \geq 0 : \mathbb{P}(|\xi| > \lambda) \leq t\}.$$

Then $\xi^{**} : (0,1] \to [0,\infty)$, the maximal function of ξ, is given by the formula

$$\xi^{**}(t) = \frac{1}{t}\int_0^t \xi^*(s)\mathrm{d}s, \qquad t \in (0,1],$$

or equivalently,

$$\xi^{**}(t) = \frac{1}{t}\sup\left\{\int_E |\xi|\mathrm{d}\mathbb{P} : E \in \mathcal{F}, \mathbb{P}(E) = t\right\}.$$

These objects enable the introduction of the weak-L^∞ space. Following Bennett, DeVore and Sharpley [10], we set

$$||\xi||_{weak(\mathbb{P})} = \sup_{t\in(0,1]} (\xi^{**}(t) - \xi^*(t))$$

and define $weak(\mathbb{P}) = \{\xi : ||\xi||_{weak(\mathbb{P})} < \infty\}$. The main motivation behind the introduction of $weak(\mathbb{P})$ comes from interpolation theory, we will discuss this later in Chapter 4. Coming back to martingale context, we have the following statement proved by Osękowski in [62].

Theorem 1.4. *The least value of the constant C'_∞ in* (1.1.4) *equals* 2*. This constant is already the best possible in the context of the Haar system.*

We would like to mention here a result which can be regarded as yet another substitute for Burkholder's inequality (1.1.3) in the endpoint case $p = 1$. The idea is to enlarge the right-hand side of this estimate, by replacing the first norm of f_n by the first norm of the maximal function $|f|_n^* = \sup_{0\leq k\leq n} |f_k|$, $n = 0, 1, 2, \ldots$. Then we have the following fact, established by Burkholder in [21].

Theorem 1.5. *If f is a martingale and g is its ± 1-transform, then*

$$||g_n||_{L^1} \leq \eta|| \, |f|_n^*||_{L^1}, \qquad n = 0, 1, 2, \ldots, \tag{1.1.5}$$

where $\eta = 2.536\ldots$ is the unique solution of the equation $\eta - 3 = -\exp\left(\frac{1-\eta}{2}\right)$. The constant is the best possible.

Differential subordination. All the results formulated above can be studied in the less restrictive setting in which the assumption of g being a martingale transform of f is relaxed. Suppose that f and g are adapted martingales. Following Burkholder [19], we say that g is differentially subordinate to f, if for any $n \geq 0$ we have $|dg_n| \leq |df_n|$ almost surely. It is immediate that if g is a martingale transform of f with respect to some predictable sequence with values in $[-1,1]$, then g is differentially subordinate to f. On the other hand, one easily constructs a pair (f,g) satisfying the differential subordination such that g is not a transform of f. It turns out that Theorems 1.2, 1.3, 1.4 and 1.5 remain valid in this more general context, with unchanged constants (which, of course, remain best possible).

Inequalities for continuous-time martingales. The notions and results discussed above can be also considered in the context of continuous-time processes. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$

is a complete probability space equipped with a right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing family of sub-σ-algebras of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all the events of probability 0. Let $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ be adapted martingales: for each $s \geq 0$ the variables X_s, Y_s are integrable and we have $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$, $\mathbb{E}(Y_t|\mathcal{F}_s) = Y_s$ whenever $t > s$. As usual, we assume that both processes are càdlàg: their trajectories are right-continuous and have limits from the left.

Suppose that $Y = (Y_t)_{t\geq 0}$ is of special form: it is the stochastic integral, with respect to $X = (X_t)_{t\geq 0}$, of some predictable process $H = (H_t)_{t\geq 0}$ which takes values in $[-1, 1]$:

$$Y_t = H_0 X_0 + \int_{0+}^{t} H_s \mathrm{d}X_s, \qquad t \geq 0.$$

We will write $Y = H \cdot X$ in such a case. This is the obvious analogue of a martingale transform from the discrete-time setting. One can also extend the notion of the differential subordination to the case of continuous processes. To this end, we need to introduce the square bracket $[X] = ([X]_t)_{t\geq 0}$ associated with X: for $t \geq 0$, $[X]_t$ is the limit in probability of the sums $Q_{n,t}$ as $n \to \infty$, where

$$Q_{n,t} = |X_0|^2 + \sum_{k=0}^{2^n-1} |X_{t(k+1)2^{-n}} - X_{tk2^{-n}}|^2.$$

See Dellacherie and Meyer [27] for the much more on the subject. Now, rewrite the definition of the discrete-time differential subordination in the form

$$(S_n^2(f) - S_n^2(g))_{n\geq 0} \text{ is nondecreasing and nonnegative as a function of } n,$$

where $S_n(f) = \left(\sum_{k=0}^{n} |df_k|^2\right)^{1/2}$ denotes the nth term of the square function of f. The above definition of $[X]_t$ shows that the square bracket is the continuous-time analogue of the square function. This observation suggests the following extension of the differential subordination: given two arbitrary martingales X and Y, we say that Y is differentially subordinate to X, if the process $([X]_t - [Y]_t)_{t\geq 0}$ is nondecreasing and nonnegative as a function of t. If we treat two discrete-time martingales $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ as continuous-time processes with the use of identities $X_t = f_{\lfloor t\rfloor}$ and $Y_t = g_{\lfloor t\rfloor}$, $t \geq 0$, then the above domination of X over Y reduces to the classical differential subordination of g to f; in other words, the new definition is consistent with the previous one.

It can be shown that Theorems 1.2, 1.3, 1.4 and 1.5 remain valid in the context of continuous-time differentially subordinate martingales. Quite interestingly, if we additionally assume that X and Y have continuous paths, then the constant $\eta = 2.536\ldots$ in Theorem 1.5 can be improved to $\sqrt{2}$ (cf. [59]); the constants in the remaining three theorems are still optimal.

Bellman function method. All the sharp estimates discussed above can be obtained by using a very efficient and flexible method, which reduces the whole problem to that of finding a special function satisfying appropriate majorization and concavity-type requirements. The technique has its origins in the theory of stochastic optimal control developed by Bellman in the fifties, but it was Burkholder who first noticed the possibility of modifying and applying the approach in the more general context of martingales and the Haar system. In the last twenty years, after the appearance of several works of Nazarov, Treil and Volberg in the nineties, the method has become a powerful tool widely used in probability theory and harmonic analysis. It would be impossible to describe here the technique in full generality, so in the thesis we will focus on its specialized form which will be sufficient for our purposes. For the sake of clarity and the reader's convenience, in the next section we describe the approach in a very basic setup. This will give us some insight and intuition, and we will be ready for various modifications and extensions of the technique developed in the next chapters.

Contribution of this thesis. Our principal purpose is to establish weighted extensions of the inequalities (1.1.2), (1.1.4), (1.1.5) formulated above. More specifically, suppose that w is a weight, that is, a nonnegative integrable random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let f, g be discrete-time martingales adapted to some filtration $(\mathcal{F}_n)_{n\geq 0}$, such that g is a ± 1-transform of f. We will study the above estimates in which the corresponding strong/weak L^p-norms are computed with respect to the new measure $w\mathrm{d}\mathbb{P}$. There are two natural problems:

1. Characterize those w, for which an inequality under investigation holds with some finite constant C_w.
2. If w is as in 1., estimate C_w (efficiently, or even better, optimally) from above, in terms of an appropriate size characteristic of w.

One can ask analogous questions in the continuous-time setting. It turns out that the first question is easy and gives rise to the probabilistic version of the so-called Muckenhoupt's A_p condition (see Subsection 1.3 below for further motivation and details). The main difficulty lies in the second question. It will be handled with the use of Bellman function method in Chapters 2-5.

1.2. Bellman function method

In this section we will describe a version of the Bellman function method which enables an efficient study of inequalities for martingale transforms. The main idea is to construct a special function satisfying the appropriate majorization and concavity-type conditions. The existence of such an object turns out to be sufficient for (actually: even equivalent to) the validity of an inequality in question. For convenience and to present the main idea behind the approach, we will also include some simple proofs. The reader interested in the more detailed presentation is referred to the articles [19, 61] and the monograph [60].

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space with a filtration $(\mathcal{F}_n)_{n\geq 0}$, and let $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ be adapted martingales such that g is a ± 1-transform of f. For technical reasons, in what follows, we will also assume that f and g are simple martingales: for every n, the random variables f_n, g_n take only a finite number of values. Let $G : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be a given function and suppose that we are interested in the inequality

$$\mathbb{E}G(f_n, g_n) \leq 0, \tag{1.2.1}$$

for all n and all pairs (f, g) as above. We do not assume that G is Borel or even measurable; there is no problem with integrability of $G(f_n, g_n)$, since both f and g are simple. The idea is to find a supermartingale $(B_n)_{n\geq 0}$ majorizing $(G(f_n, g_n))_{n\geq 0}$ and satisfying the condition $B_0 \leq 0$. The existence of such an object immediately implies (1.2.1): indeed, then we have

$$\mathbb{E}G(f_n, g_n) \leq \mathbb{E}B_n \leq \mathbb{E}B_0 \leq 0.$$

Actually, one searches for $(B_n)_{n\geq 0}$ in the class of special supermartingales of the form $(B(f_n, g_n))_{n\geq 0}$ for some function B, the Bellman function associated with (1.2.1). As we will see in the moment, the aforementioned properties of $(B_n)_{n\geq 0}$ translate into the language of B as follows:

1° (Initial condition) We have $B(x,y)\leq 0$ if $|y|\leq |x|$.

2° (Majorization condition) We have $B(x,y)\geq G(x,y)$.

3° (Concavity) The function B is concave along any line of slope in $[-1,1]$.

Theorem 1.6. *If there is a function B satisfying 1°, 2° and 3°, then (1.2.1) holds.*

Proof. Let (f,g) be as above. By 3°, the process $(B(f_n,g_n))_{n\geq 0}$ is a supermartingale:

$$\begin{aligned}\mathbb{E}(B(f_{n+1},g_{n+1})|\mathcal{F}_n) = \mathbb{E}(B(f_n+df_n,g_n+dg_n)|\mathcal{F}_n) &= \mathbb{E}(B(f_n+df_n,g_n+v_ndf_n)|\mathcal{F}_n)\\ &\leq B(f_n,g_n),\end{aligned}$$

where the last passage is due to conditional Jensen's inequality. So, by 2° and then 1°,

$$\mathbb{E}G(f_n,g_n)\leq \mathbb{E}B(f_n,g_n)\leq \mathbb{E}B(f_0,g_0)\leq 0. \quad \square$$

One of the beautiful and remarkable aspects of the Bellman function method is that the statement of the above theorem can be reversed: the validity of (1.2.1) implies the existence of a function satisfying the aforementioned three requirements. To show this, we need some additional notation. In the definition of a martingale transform we have assumed that the condition $dg_n = v_ndf_n$ is satisfied for each $n = 0, 1, 2, \ldots$. In what follows, we will need a wider class of sequences. For every $(x,y)\in\mathbb{R}^2$, let $M(x,y)$ denote the family of all pairs (f,g) of simple martingales satisfying $(f_0,g_0)=(x,y)$ and $dg_n = v_ndf_n$ for $n = 1, 2, \ldots$, where $(v_n)_{n\geq 0}$ is a predictable sequence bounded in absolute value by 1. (In other words, we allow the martingale transform condition to be violated at the initial time.) In the definition of $M(x,y)$, the underlying filtration is allowed to vary, as well as the probability space.

For any given and fixed $G:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, introduce an abstract function $\mathcal{B}^0:\mathbb{R}\times\mathbb{R}\to\mathbb{R}\cup\{\infty\}$ by the formula

$$\mathcal{B}^0(x,y)=\sup\big\{\mathbb{E}G(f_n,g_n)\,:\,(f,g)\in M(x,y),\ n=0,\,1,\,2,\,\ldots\big\}.$$

Theorem 1.7. *If $\mathcal{B}^0$ is finite, then it is the least function satisfying 2° and 3°.*

Proof. The condition 2° is easy to check: it is enough to consider the constant martingales $f\equiv x$ and $g\equiv y$ in the definition of $\mathcal{B}^0$. The concavity 3° is proved by the so-called "splicing argument" of Burkholder. We must show that for all $x,y\in\mathbb{R}$, $\varepsilon\in[-1,1]$ and any $\alpha\in(0,1)$, t_1, $t_2\in\mathbb{R}$ such that $\alpha t_1+(1-\alpha)t_2=0$, we have

$$\alpha\mathcal{B}^0(x+t_1,y+\varepsilon t_1)+(1-\alpha)\mathcal{B}^0(x+t_2,y+\varepsilon t_2)\leq\mathcal{B}^0(x,y).$$

To this end, pick arbitrary pairs (f^j,g^j) from $M(x+t_j,y+\varepsilon t_j)$, $j=1, 2$. Suppose that these pairs are given on two disjoint probability spaces $(\Omega^1,\mathcal{F}^1,\mathbb{P}^1)$ and $(\Omega^2,\mathcal{F}^2,\mathbb{P}^2)$. Let us "glue" these pairs into one pair of martingales using the parameter α. Namely, set $\Omega=\Omega^1\cup\Omega^2$, $\mathcal{F}=\sigma(\mathcal{F}^1,\mathcal{F}^2)$ and define the probability measure $\mathbb{P}$ on $\mathcal{F}$ by requiring that $\mathbb{P}(A_1\cup A_2)=\alpha\mathbb{P}(A_1)+(1-\alpha)\mathbb{P}(A_2)$ for any $A_1\in\mathcal{F}^1$ and $A_2\in\mathcal{F}^2$. In addition, introduce the pair $(f,g)=((f_n,g_n))_{n\geq 0}$ by $(f_0,g_0)=(x,y)$ and

$$(f_n(\omega),g_n(\omega))=\begin{cases}(f^1_{n-1}(\omega),g^1_{n-1}(\omega)) & \text{if } \omega\in\Omega^1,\\ (f^2_{n-1}(\omega),g^2_{n-1}(\omega)) & \text{if } \omega\in\Omega^2,\end{cases}$$

for $n = 1, 2, \ldots$. We have the identity

$$\mathbb{E}((f_1, g_1)|\mathcal{F}_0) = \mathbb{E}(f_1, g_1) = \alpha(x + t_1, y + \varepsilon t_1) + (1 - \alpha)(x + t_2, y + \varepsilon t_2) = (x, y).$$

From this and the fact that (f^1, g^1) and (f^2, g^2) are (two-dimensional) martingales, we infer that the pair (f, g) is a martingale as well. Actually, it is not difficult to see that (f, g) belongs to $M(x, y)$. Consequently, by the very definition $\mathcal{B}^0$,

$$\mathcal{B}^0(x, y) \geq \mathbb{E}G(f_n, g_n) = \alpha\mathbb{E}^1 G(f^1_{n-1}, g^1_{n-1}) + (1 - \alpha)\mathbb{E}^2 G(f^2_{n-1}, g^2_{n-1}),$$

where $\mathbb{E}^i$ is the expectation with respect to the probability measure $\mathbb{P}^i$. Therefore, taking the supremum over all n and all pairs (f^1, g^1), (f^2, g^2) as above, we obtain

$$\mathcal{B}^0(x, y) \geq \alpha\mathcal{B}^0(x + t_1, y + \varepsilon t_1) + (1 - \alpha)\mathcal{B}^0(x + t_2, y + \varepsilon t_2).$$

This is the desired condition 3°. It remains to show that $\mathcal{B}^0$ is the smallest function satisfying 2° and 3°. Let B be an arbitrary function satisfying these two conditions, pick $(x, y) \in \mathbb{R}^2$ and fix a pair (f, g) from $M(x, y)$. We already know from the proof of Lemma 1.6 that $(B(f_n, g_n))_{n\geq 0}$ is a supermartingale, and hence we get

$$\mathbb{E}G(f_n, g_n) \leq \mathbb{E}B(f_n, g_n) \leq \mathbb{E}B(f_0, g_0) = B(x, y).$$

It remains to take the supremum over all $(f, g) \in M(x, y)$ and all $n = 0, 1, 2, \ldots$ to obtain the estimate $\mathcal{B}^0(x, y) \leq B(x, y)$. This completes the proof of the theorem. □

Combining the above two facts, we deduce the following statement.

Theorem 1.8. *The inequality* (1.2.1) *is valid if and only if there exists a function* $B : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *satisfying* 1°, 2° *and* 3°.

Proof. By Theorem 1.6, if there is a function B satisfying 1°, 2° and 3°, then the inequality (1.2.1) holds. It remains to establish the reverse implication. We start with the observation that the validity of (1.2.1) immediately yields 1°. Therefore, by the previous theorem, it is enough to show that $\mathcal{B}^0$ is finite. The initial condition 1° (which we have just checked) guarantees the inequality $\mathcal{B}^0(x, y) \leq 0 < \infty$ for every (x, y) satisfying $|y| \leq |x|$. Suppose that $|x| \neq |y|$ and take an arbitrary pair (f, g) from $M(x, y)$. Next, consider another martingale (f', g'), which starts from the point $((x + y)/2, (x + y)/2)$ and, in the first step, moves to (x, y) or (y, x) with probability 1/2. If it goes to the point (y, x), it stops ultimately; otherwise, we determine (f', g') by the assumption that the conditional distribution of $(f'_n, g'_n)_{n\geq 1}$ with respect to $\mathcal{F}_1$ coincides with the unconditional distribution of $(f_n, g_n)_{n\geq 0}$. Then it is easy to check that g' is a ± 1-transform of f', and hence, for any $n \geq 1$,

$$0 \geq \mathbb{E}G(f'_n, g'_n) = \frac{1}{2}G(y, x) + \frac{1}{2}\mathbb{E}G(f_{n-1}, g_{n-1}).$$

It remains to take the supremum over f, g and n to obtain that $\mathcal{B}^0(x, y) \leq -G(y, x)$. □

Theorem 1.8 contains a very basic version of the Bellman function method, which is adapted for the study of discrete-time martingales and their transforms. Despite the simplicity, we have encountered here all the relevant aspects of the technique, which are worth highlighting here once again: the validity of a given estimate is equivalent to the existence of a certain special function which is neither too big, nor too small, and satisfies an appropriate concavity requirement. In the later chapters, we will see how to adjust the technique so that it is applicable to:

· weighted inequalities,
· inequalities for continuous-time martingales and their stochastic integrals,
· inequalities for differentially subordinate martingales,
· maximal inequalities for martingale transforms.

1.3. Weighted inequalities in analysis

Before we proceed with the weighted estimates in the martingale context, we must present some classical results from the analytic setting. They will provide us with some additional insight and motivation. For simplicity, we will restrict ourselves to the Euclidean setting, but most of the machinery described below works in the more general context of spaces of homogeneous and non-homogeneous type (see e.g. [26, 56]).

A_p classes. Let w be a weight in $\mathbb{R}^d$, i.e., a positive, locally integrable function w on $\mathbb{R}^d$. This weight induces a measure, again denoted by w (with no risk of confusion), which is defined by $w(E) = \int_E w(x)\mathrm{d}x$ for any measurable subset E of $\mathbb{R}^d$. For every $1 \leq p \leq \infty$, let $L^p(\mathbb{R}^d, \mathrm{d}x)$ and $L^{p,\infty}(\mathbb{R}^d, \mathrm{d}x)$ be the standard L^p and weak-L^p spaces associated with the Lebesgue measure. The weighted L^p space associated with w is denoted by $L^p(w)$, and consists of all equivalence classes of functions f for which

$$\|f\|_{L^p(w)} = \left(\int_{\mathbb{R}^d} |f(x)|^p w(x)\mathrm{d}x\right)^{1/p} \tag{1.3.1}$$

is finite (for $p = \infty$, one takes the usual modification involving the essential supremum). For every $1 \leq p < \infty$ and a weight w, we define the weak space $L^{p,\infty}(w)$ to be the function space with the quasi-norm

$$\|f\|_{L^{p,\infty}(w)} = \sup_{\lambda \geq 0} \lambda w(\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\})^{1/p}, \qquad 1 \leq p < \infty. \tag{1.3.2}$$

One can also introduce a version of this space for $p = \infty$ (as in Subsection 1.1 above), but we will not need it here.

Roughly speaking, the weighted theory concerns the boundedness of classical analytic operators in various weighted function spaces. Throughout the thesis, we will focus on the one-weight case: that is, the estimates we will study will involve the same weight appearing on both sides. For the sake of completeness, we should mention here that the two-weight theory (devoted to the study of inequalities whose left- and the right-hand sides depend on different weights) is a separate challenging area, with an independent set of tools and applications.

To proceed, suppose that T is some classical operator: for instance, one can take the Hardy-Littlewood maximal function, a singular integral, an area integral, etc. Furthermore, fix an exponent p. Typically, the first problem is to characterize the class of those weights w, for which the strong- or weak-type estimates

$$\int_{\mathbb{R}^d} |Tf(x)|^p w(x)\mathrm{d}x \leq C_{p,d,w} \int_{\mathbb{R}^d} |f(x)|^p w(x)\mathrm{d}x, \tag{1.3.3}$$

$$\lambda^p w(\{x \in \mathbb{R}^d : |Tf(x)| \geq \lambda\}) \leq C'_{p,d,w} \int_{\mathbb{R}^d} |f(x)|^p w(x)\mathrm{d}x \tag{1.3.4}$$

hold for all f with some finite constants $C_{p,d,w}$, $C'_{p,d,w}$. For example, assume that M is the Hardy-Littlewood maximal function on $\mathbb{R}^d$: this operator acts on locally integrable functions f by

$$Mf(x) = \sup_Q \frac{1}{|Q|} \int_Q |f(y)| \mathrm{d}y, \qquad x \in \mathbb{R}^d,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^d$, containing x, which have sides parallel to the axes. A celebrated result of Muckenhoupt [52] gives the full characterization of weights for which (1.3.3) and (1.3.4) hold. Given $1 < p < \infty$, we say that a weight satisfies Muckenhoupt's condition A_p / belongs to the class A_p (and denote this by $w \in A_p$), if there exists a constant K such that for every cube $Q \subset \mathbb{R}^d$ with sides parallel to the axes,

$$\left(\frac{1}{|Q|} \int_Q w(x) \mathrm{d}x\right) \left(\frac{1}{|Q|} \int_Q (w(x))^{1/(1-p)} \mathrm{d}x\right)^{p-1} \leq K < \infty. \tag{1.3.5}$$

The smallest possible value of K is called the A_p characteristic of w and denoted by $[w]_{A_p}$. There are analogues of Muckenhoupt's conditions in the boundary case $p \in \{1, \infty\}$, which can be obtained with a limiting argument. Namely, the weight w satisfies condition A_1, if there exists a constant K such that for almost every $x \in \mathbb{R}^d$, $Mw(x) \leq Kw(x)$ (here M is still the Hardy-Littlewood maximal operator). As previously, the smallest possible value of K is called the A_1 characteristic of w and denoted by $[w]_{A_1}$. Finally, a weight w satisfies A_∞ condition (belongs to the class A_∞) if there exists a constant K such that for ever cube $Q \subset \mathbb{R}^d$,

$$\left(\frac{1}{|Q|} \int_Q w(x) \mathrm{d}x\right) \exp\left(-\frac{1}{|Q|} \int_Q \log(w(x)) \mathrm{d}x\right) \leq K < \infty.$$

The smallest possible value K is called A_∞ characteristic of w and is denoted by $[w]_{A_\infty}$.

We are ready to formulate Muckenhoupt's result.

Theorem 1.9. *(i) Let $1 < p < \infty$ be fixed. The Hardy-Littlewood maximal function is bounded as an operator on $L^p(w)$ if and only if w belongs to the A_p class.*

(ii) Let $1 \leq p < \infty$ be fixed. The Hardy-Littlewood maximal function is bounded as an operator from $L^p(w)$ to $L^{p,\infty}(w)$ if and only if w belongs to the A_p class.

This result became a landmark in the theory of weighted inequalities when it turned out that the condition A_p characterizes the boundedness of other important classical operators in harmonic analysis. Let us present several selected results in this direction, the full list is much larger and impossible to survey here. The sufficiency of Muckenhoupt's condition for the boundedness of the Hilbert transform was proved by Hunt, Muckenhoupt and Wheeden [36], while the setting of Riesz transforms (actually, even a more wider class of singular integrals) was handled by Coifman and Fefferman [24]. (Proofs of the necessity can be found in [33, 36, 80].) The weighted estimates for fractional and Poisson integrals were investigated by Sawyer [72, 73], the analysis of square function operators can be found in the works of Buckley [14], Chanillo and Wheeden [23] and Lerner [48].

On dependence of the norms on the characteristic. Another important aspect of the theory of weighted inequalities was initiated with the thesis of Buckley [14], published in 1993. It can be regarded as a stronger quantitative version of Muckenhoupt's result and will play a crucial role in our further considerations.

Theorem 1.10. *Let $1 < p < \infty$ and $w \in A_p$. Then the Hardy-Littlewood operator M satisfies*

$$\|M\|_{L^p(w)\to L^p(w)} \leq C_{p,d}[w]_{A_p}^{1/(p-1)},$$

where the constant $C_{p,d}$ depends only on p and the dimension d. Moreover, the exponent $1/(p-1)$ is optimal in the sense that for any $\kappa < 1/(p-1)$ and any $K > 0$, there is an A_p weight w on $\mathbb{R}$ such that

$$||M||_{L^p(w)\to L^p(w)} > K[w]_{A_p}^{\kappa}.$$

There are many different proofs of this beautiful result, e.g. a very elementary argument can be found in Lerner's paper [47]. Similar questions about the optimal dependence on Muckenhoupt characteristic has been asked for other classical operators. A paper which gave the impetus to the development of this direction of research was the work of Petermichl and Volberg from 2002. In [69], they established the sharp weighted inequality for the Beurling-Ahlfors operator (a complex-valued analog of the Hilbert transform) and applied this result to solve a famous borderline regularity problem for the Beltrami equation. Since then, the problem of finding optimal exponents in the weighted inequalities has become one of the main themes in the theory of weights (see e.g. [35, 37, 44–46, 48, 70, 71]).

Selected properties of A_p weights. Suppose that $1 < p < \infty$. By Jensen's inequality, if w is a weight on $\mathbb{R}^d$, then we have

$$\left(\frac{1}{|Q|}\int_Q w(x)\mathrm{d}x\right)\left(\frac{1}{|Q|}\int_Q (w(x))^{1/(1-p)}\mathrm{d}x\right)^{p-1} \geq 1$$

for every cube $Q \subset \mathbb{R}^d$. Hence $[w]_{A_p} \geq 1$ and Muckenhoupt's condition amounts to saying that above inequality can be reversed, up to a constant. Similar phenomenon occurs in the boundary cases $p = 1$ and $p = \infty$ as well. Thus, it is not surprising that the weights satisfying the condition A_p have a regular and "balanced" structure. Let us take a look at some fundamental properties of Muckenhoupt's weights which have been extensively applied in proving weighted inequalities (and which will be important to us later).

Proposition 1.11. *(i) If $1 \leq p < q \leq \infty$, then $A_p \subset A_q$ and $[w]_{A_q} \leq [w]_{A_p}$.*

(ii) If $1 < p < \infty$, then $w \in A_p$ if and only if $w^{1-p'} \in A_{p'}$. Furthermore, we have the equality $[w]_{A_p} = [w^{1-p'}]_{A_{p'}}^{p-1}$. (Here $p' = p/(p-1)$ is the conjugate exponent to p.)

(iii) If $w_0, w_1 \in A_1$, then $w_0 w_1^{1-p} \in A_p$ and $[w_0 w_1^{1-p}]_{A_p} \leq [w_0]_{A_1}[w_1]_{A_1}^{p-1}$.

(iv) A weight w belongs to A_∞ if and only if w satisfies reverse Hölder inequality: there exists $\varepsilon > 0$ and a constant K such that for every cube $Q \subset \mathbb{R}^d$,

$$\left(\frac{1}{|Q|}\int_Q w(x)^{1+\varepsilon}dx\right)^{1/(1+\varepsilon)} \leq K\frac{1}{|Q|}\int_Q w(x)dx.$$

(v) If $w \in A_p$, $p > 1$, then there exists ε, $0 < \varepsilon < p-1$, such that $w \in A_{p-\varepsilon}$.

(vi) If $w \in A_p$, $p > 1$, then there exist $w_1, w_2 \in A_1$ such that $w = w_1 w_2^{1-p}$.

The first three conditions have elementary proofs which exploit nothing but Hölder's inequality. The remaining three properties require more work. The reverse Hölder inequality (iv) was proved by Coifman and Fefferman [24]. The property (v) is an immediate consequence of reverse Hölder inequality for the weight $w^{1-p'}$. The property (vi) is the deep result named factorization theorem, it was conjectured by Muckenhoupt and proved by Jones in [42]. The original argument was quite complicated and was subsequently simplified by Coifman, Jones and Rubio de Francia [25].

Factorization of weights is strictly connected to one of the most surprising and profound techniques in the weighted theory: the extrapolation. Let us start with the simple observation which follows from factorization and Riesz-Thorin interpolation theorem.

Theorem 1.12. *Let $1 \leq p_0 < p < p_1 < \infty$. Assume that T is a linear operator which is bounded on $L^{p_0}(w)$ for all $w \in A_{p_0}$ and which is bounded as an operator on $L^{p_1}(w)$ for all $w \in A_{p_1}$. Then T is bounded on $L^p(w)$ for all $w \in A_p$.*

Proof. Let w be an arbitrary A_p weight. By factorization, we have $w = w_0 w_1^{1-p}$ for some $w_0, w_1 \in A_1$. Observe that from the assumptions, the operator mapping f to $w_1^{-1}T(w_1 f)$ is bounded on $L^{p_0}(w_0 w_1)$ and on $L^{p_1}(w_0 w_1)$. Indeed, we have

$$\|w_1^{-1}T(w_1 f)\|_{L^{p_0}(w_0 w_1)} = \|T(w_1 f)\|_{L^{p_0}\left(w_0 w_1^{1-p_0}\right)} \leq \|T\|_{L^{p_0}\left(w_0 w_1^{1-p_0}\right)} \|f\|_{L^{p_0}(w_0 w_1)}.$$

Similarly we show boundedness on $L^{p_1}(w_0 w_1)$. Hence, from Riesz-Thorin interpolation theorem, the operator is also bounded on $L^p(w_0 w_1)$, so

$$\|T(w_1 f)\|_{L^p(w)} = \|w_1^{-1}T(w_1 f)\|_{L^p(w_0 w_1)} \leq C\|f\|_{L^p(w_0 w_1)} = C\|w_1 f\|_{L^p(w)},$$

for some constant C. This completes the proof. □

Rubio de Francia observed that just one of the above boundedness assumptions is sufficient for the result. That is, it is possible to "extrapolate" instead of merely "interpolate".

Theorem 1.13 (Rubio de Francia extrapolation theorem)**.** *Given an operator T, suppose that for some p_0, $1 \leq p_0 < \infty$, and for every $w \in A_{p_0}$, there exists a constant C depending only on $[w]_{A_{p_0}}$ such that*

$$\int_{\mathbb{R}^d} |Tf(x)|^{p_0} w(x)dx \leq C \int_{\mathbb{R}^d} |f(x)|^{p_0} w(x)dx.$$

Then for every $1 < p < \infty$, and every $w \in A_p$ there exists a constant C depending only on $[w]_{A_p}$ such that

$$\int_{\mathbb{R}^d} |Tf(x)|^{p} w(x)dx \leq C \int_{\mathbb{R}^d} |f(x)|^{p} w(x)dx.$$

The central idea in the proof of the extrapolation theorem (and the simplest proofs of Jones factorization theorem as well) is to use the iteration method called Rubio de Francia algorithm. The detailed discussion would take us too far, so we will not present it here and refer the reader to the monograph [33].

Instead, let us formulate an extension of the above statement which is due to Duoandikoetxea [31]. It turns out that the extrapolation holds for pairs of functions and the appearance of the linear operator T is not required.

Theorem 1.14. *Assume that for some pair (f, g) of nonnegative functions, for some $p_0 \in [1, \infty)$ and all $w \in A_{p_0}$ we have the estimate*

$$||g||_{L^{p_0}(w)} \leq CN\left([w]_{A_{p_0}}\right)||f||_{L^{p_0}(w)},$$

where N is an increasing function and C does not depend on w. Then for all $1 < p < \infty$ and all weights $w \in A_p$ we have

$$||g||_{L^p(w)} \leq CK(w)||f||_{L^p(w)},$$

where

$$K(w) = \begin{cases} N\left([w]_{A_p}(2||M||_{L^p(w)\to L^p(w)})^{p_0-p}\right) & \text{if } p < p_0, \\ N\left([w]_{A_p}^{\frac{p_0-1}{p-1}}\left(2||M||_{L^{p'}(w^{1/(1-p)})\to L^{p'}(w^{1/(1-p)})}\right)^{\frac{p-p_0}{p-1}}\right) & \text{if } p > p_0. \end{cases}$$

In particular, $K(w) \leq C_1 N\left(C_2[w]_{A_p}^{\max\{\frac{p_0-1}{p-1},1\}}\right)$ *for* $w \in A_p$.

Coming back to the context of linear operators, we have the following corollary concerning the extrapolation of weak type estimates. We provide the simple proof.

Corollary 1.15. *Suppose that* T *is an operator acting on measurable functions on* $\mathbb{R}^d$ *and let* $p_0 \in [1,\infty)$ *be a fixed parameter. Let* C *be a finite constant and let* N *be an increasing function on* $[1,\infty)$. *If we have the estimate* $||T||_{L^{p_0}(w)\to L^{p_0,\infty}(w)} \leq CN([w]_{A_p})$ *for all weights* $w \in A_{p_0}$, *then for all* $1 < p < \infty$ *we have* $||T||_{L^p(w)\to L^{p,\infty}(w)} \leq CK(w)$, *where* K *is as in the previous statement.*

Proof. Fix $\lambda > 0$. We have that

$$\|\lambda \mathbb{1}_{\{|Tf|\geq\lambda\}}\|_{L^{p_0}(w)} = \lambda(w(|Tf| \geq \lambda))^{1/p_0} \leq \|Tf\|_{L^{p_0,\infty}(w)} \leq CN([w]_{A_{p_0}})\|f\|_{L^{p_0}(w)}$$

for any A_p weight w. Therefore, by the previous theorem we obtain

$$\|\lambda \mathbb{1}_{\{|Tf|\geq\lambda\}}\|_{L^p(w)} \leq CK(w)\|f\|_{L^p(w)}.$$

It remains to take the supremum over all λ to get the assertion. □

We conclude by observing that the extrapolation theorem can also be used in the study of unweighted inequalities. For example, if we can prove the weighted L^2 estimate for A_2 weights, then the Rubio de Francia extrapolation theorem immediately yields the weighted L^p estimate for any $w \in A_p$; hence, by restricting to constant weights, one gets the unweighted L^p estimate as well. The described phenomenon also gives some heuristic insight into why proving weighted $L^2(w)$ inequalities (for a *single* exponent 2) is significantly more difficult than establishing the unweighted L^p versions for *all* p.

1.4. Martingale weights

Equipped with the definitions and motivation from analysis, we are ready to return to the stochastic context.

Continuous-time setting. In contrast to our previous considerations in the unweighted realm, here it is more convenient to start with the continuous-time case. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space filtered by $(\mathcal{F}_t)_{t\geq 0}$, a continuous-time right-continuous filtration such that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$ and their complements only. For any adapted martingale $X = (X_t)_{t\geq 0}$, its maximal function is given by $|X|^* = \sup_{t\geq 0}|X_t|$ and its one-sided version is $X^* = \sup_{t\geq 0} X_t$. Any positive and integrable random variable W is called a weight, and such an object gives rise to the measure on $(\Omega, \mathcal{F})$ (also denoted by W), which is given by $W(A) = \mathbb{E}W\mathbb{1}_A$. The associated L^p and weak L^p spaces are defined as in (1.3.1) and (1.3.2), just replacing $\mathbb{R}^d$ with Ω and the Lebesgue measure with $\mathbb{P}$. The weight W induces the martingale $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$, which will be denoted by W as well. This slight abuse of notation (W denotes a random variable, a measure and a martingale...) is widely accepted in the literature. We also assume (for technical reasons: see discussion below) that the martingale $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$ generated by the weight W has continuous trajectories while the auxiliary process V will be allowed to have jumps.

We are ready to define a probabilistic version of the Muckenhoupt condition.

Definition 1.1. (A_1) A weight W is called a Muckenhoupt A_1 weight (or W belongs to the A_1 class), if there exists a finite deterministic constant C such that

$$W_t^* \leq CW_t$$

almost surely for all $t \geq 0$. The smallest possible C above is called the A_1 characteristic of W and denoted by $[W]_{A_1}$.

(A_p) Let $1 < p < \infty$. A weight W is called a Muckenhoupt A_p weight (or W belongs to the A_p class), if the random variable $V = W^{1/(1-p)}$ is integrable and there exists a finite deterministic constant C such that

$$W_t V_t^{p-1} \leq C$$

almost surely for all $t \geq 0$. Here $(V_t)_{t\geq 0}$ is the martingale given by $V_t = \mathbb{E}(W^{1/(1-p)}|\mathcal{F}_t)$. The smallest possible C above is called the A_p characteristic of W and denoted by $[W]_{A_p}$.

(A_∞) A weight W is called a Muckenhoupt A_∞ weight (or W belongs to the A_∞ class), if the random variable $V = \log(W)$ is integrable and there exists a finite deterministic constant C such that

$$W_t \exp(-V_t) \leq C$$

almost surely for all $t \geq 0$. Here $(V_t)_{t\geq 0}$ is a martingale given by $V_t = \mathbb{E}(\log(W)|\mathcal{F}_t)$. The smallest possible C above is called the A_∞ characteristic of W and denoted by $[W]_{A_\infty}$.

Note that by Jensen's inequality we have $[W]_{A_p} \geq 1$ for all p. The basic properties of analytic weights listed in Proposition 1.11 are also satisfied for probabilistic weights [43].

Let us present a certain geometric interpretation of A_p weights, which follows directly from the above definition. Namely, any A_1 weight W gives rise to a two-dimensional process (W, W^*), which takes values in the angle $\{(x, y) \in (0, \infty)^2 : x \leq y \leq [W]_{A_1} x\}$. Observe that this pair has the following dynamics: when away from the diagonal $x = y$, the process moves horizontally "in a martingale manner" and if $W = W^*$, then the second coordinate might increase infinitesimally. For $1 < p < \infty$, any A_p weight W can be identified with with the martingale pair (W, V), which takes values in the hyperbolic strip $\{(x, y) \in (0, \infty)^2 : 1 \leq xy^{p-1} \leq [W]_{A_p}\}$ and terminates at the lower boundary. The case $p = \infty$ is similar: we identify any A_∞ weight with the martingale pair (W, V) taking values in the logarithmic (or rather exponential) strip $\{(x, y) \in (0, \infty) \times \mathbb{R} : 1 \leq xe^{-y} \leq [W]_{A_\infty}\}$ and terminating at the upper boundary.

In what follows, we will need a version of the extrapolation Theorem 1.14 for probabilistic weights. The proof is similar to the one in the analytic setting and can be found in [6].

Theorem 1.16 (Extrapolation theorem for probabilistic weights)**.** *Let $p_0 \in (1, \infty)$ and $C, \kappa > 0$ be fixed parameters. Suppose that f, g are nonnegative random variables such that for any weight $W \in A_{p_0}$ we have*

$$||g||_{L^{p_0}(W)} \leq C[W]_{A_{p_0}}^{\kappa} ||f||_{L^{p_0}(W)}.$$

Then for all $1 < p < \infty$ and all $W \in A_p$ we have

$$||g||_{L^p(W)} \leq K(W)||f||_{L^p(W)},$$

where

$$K(W) = \begin{cases} 2^{1-p/p_0} C \left(2||M||_{L^p(W)}\right)^{\kappa(p_0-p)} [W]_{A_p}^{\kappa} & \text{if } p < p_0, \\ 2^{p'(1-p_0/p)/p_0} C \left(2||M||_{L^{p'}(W^{1/(1-p)})}\right)^{\frac{\kappa(p-p_0)}{p-1}} [w]_{A_p}^{\frac{\kappa(p_0-1)}{p-1}} & \text{if } p > p_0. \end{cases}$$

In particular, $K(W) \leq C_1 [W]_{A_p}^{\kappa \max\{\frac{p_0-1}{p-1},1\}}$ for some C_1 not depending on W.

Finally, let us address the issue of the regularity of trajectories. We have considered above only those weights such that the generated martingale $(W_t)_{t\geq 0}$ has continuous paths. This assumption, which might seem restrictive, appeared in first works on this topic by Izumisawa-Kazamaki [41] and is still frequently used in weighted inequalities (see for example [6,7,64]). There are several reasons for considering only path-continuous martingales. As we have observed above, Muckenhoupt's weights can be identified with two-dimensional martingales with values in certain hyperbolic/logarithmic domains. Because these domains are not convex, some difficulties might arise when applying Bellman function method for martingales with jumps: roughly speaking, the main problem is that for non-convex sets, the local concavity does not imply the usual concavity. Moreover, in the general setting of jump processes, some basic properties of weights do not hold true, for example the classical self-improvement property fails [11]. Finally, it is known that some fundamental martingale inequalities are valid only under the additional restrictions on the regularity of jumps. For instance, consider the classical Burkholder-Davis-Gundy inequality comparing (quasi-)norms of maximal and square functions of martingales:

$$c_p \mathbb{E}([X, X]^{p/2}) \leq \mathbb{E}(|X|^*)^p \leq C_p \mathbb{E}([X, X]^{p/2}).$$

It is known that this inequality holds for general càdlàg martingales for $1 \leq p < \infty$ [17]. However for continuous-path martingales or martingales with jumps satisfying additional regularity conditions, the above inequality holds true for all $0 < p < \infty$ [16]. In the last chapter of this thesis we will consider another example of the inequality which is true for continuous processes, but does not hold for general martingales with jumps.

Discrete-time setting. Now let us describe weighted inequalities for discrete-time martingales. Again, we start with basic definitions and notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with the discrete-time filtration $(\mathcal{F}_n)_{n\geq 0}$. For any adapted martingale $f = (f_n)_{n\geq 0}$, its maximal function is given by $|f|^* = \sup_{n\geq 0} |f_n|$ and its one-sided counterpart is $f^* = \sup_{t\geq 0} f_t$. Any positive and integrable random variable w is called a weight; such a variable generates the measure on $(\Omega, \mathcal{F})$ (which, following the standard convention, is also denoted by w), which is given by $w(A) = \mathbb{E} w \mathbb{1}_A$. The weighted L^p and weak L^p spaces are defined analogously to (1.3.1) and (1.3.2). The weight w also gives rise to the uniformly integrable martingale $(w_n)_{n\geq 0} = (\mathbb{E}(w|\mathcal{F}_n))_{n\geq 0}$, which is again denoted by w.

Now, the definitions of Muckenhoupt's classes and A_p characteristic are the same as in the continuous-time case and, as before, we have $[w]_{A_p} \geq 1$ for all $1 \leq p \leq \infty$. The properties (i), (ii) and (iii) from Proposition 1.11 are still satisfied for probabilistic weights. Quite remarkably, if we do not assume any additional regularity of the underlying filtration, some fundamental properties may not hold anymore. See the above discussion on the possibility of jumps of continuous-time weights; a similar phenomenon occurs in the discrete-time as well. Because of that, in what follows, we will sometimes impose the following regularity assumption on the filtration.

Definition 1.2. Let $\theta \in (0, 1/2]$ be a fixed parameter. A filtration $(\mathcal{F}_n)_{n\geq 0}$ is said to be θ-regular, if $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and for any $n \geq 0$, every atom A of $\mathcal{F}_n$ splits into a finite number A_1, A_2, ..., A_k of atoms of $\mathcal{F}_{n+1}$ satisfying $\mathbb{P}(A_j) \geq \theta\mathbb{P}(A)$, $j = 1, 2, \ldots, k$.

Regular filtrations are natural extensions of dyadic filtrations used widely in harmonic analysis. For a fixed dimension d, the dyadic filtration of the space $\big([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|\big)$ is 2^{-d}-regular in the above sense.

Chapter 2

Weighted strong-type inequalities for stochastic integrals

2.1. Statement of the result, some historical comments

Let f be a discrete-time martingale and let g be its ± 1-transform. A celebrated result of Burkholder [19] asserts that for $1 < p < \infty$ we have the sharp bound

$$\|g\|_{L^p} \leq (p^* - 1)\|f\|_{L^p}, \tag{2.1.1}$$

where $p^* = \max\{p, p/(p-1)\}$. Here we have used the notation $\|f\|_{L^p} = \sup_{n\geq 0} \|f_n\|_{L^p}$ (and similarly for $\|g\|_{L^p}$). Alternatively, if the martingale $f = (f_n)_{n\geq 0}$ is L^p-bounded, then it converges almost surely to a limit which is denoted by f or by f_∞, and then $\|f\|_{L^p}$ is the usual L^p-norm of this limit random variable.

The inequality (2.1.1) was proved using the Bellman function technique: as we have described in Chapter 1, it suffices to find $B : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that

1° $B(x, y) \leq 0$ if $|y| \leq |x|$;
2° $B(x, y) \geq |y|^p - (p^* - 1)|x|^p$
3° for any x, y and d, e with $|e| \leq |d|$, the function $t \mapsto B(x + td, y + te)$ is concave on $\mathbb{R}$.

Burkholder proved that the function

$$B(x, y) = \alpha_p(|y| - (p^* - 1)|x|)(|x| + |y|)^{p-1}$$

(where α_p is a certain constant depending only on p) enjoys all the properties. It turns out that this function can be applied in seemingly unrelated areas of mathematics. Namely, there is a deep and unexpected connection of B with the geometric function theory, particularly with the theory of quasiconformal mappings, rank-one convex functionals and the properties of Beurling-Ahlfors operator: see [1–3, 39, 40] and consult the references therein. In other words, although the function B originates in the probabilistic estimate (2.1.1), its explicit formula is of independent interest and importance in contexts far and beyond martingale theory.

The above observation was one of the motivations for our research. There is a natural and interesting question about the following weighted version of (2.1.1):

$$\|g\|_{L^p(w)} \leq C_{p,w}\|f\|_{L^p(w)}, \qquad w \in A_p,$$

to be valid for all uniformly integrable martingales f, g such that g is a ± 1-transform of f. Here $\|f\|_{L^p(w)}$, $\|g\|_{L^p(w)}$ are the weighted L^p-norms of the limit random variables $\lim_{n\to\infty} f_n$, $\lim_{n\to\infty} g_n$, and the multiplicative constant $C_{p,w}$ depends on p and the weight w. Actually, motivated by the discussion in the analytic setting presented in the previous chapter, one can ask about the extraction of the sharp dependence of $C_{p,w}$ on the A_p characteristic of w. Namely, the problem is to identify the least exponent κ_p for which

$$\|g\|_{L^p(w)} \leq C_p[w]_{A_p}^{\kappa_p}\|f\|_{L^p(w)},$$

where C_p depends on p only. For martingales on one-dimensional probability space (the interval $[0,1]$ with Lebesgue's measure and the dyadic filtration), this problem was answered by Wittwer [87]: the optimal κ_p is equal to $\max\{1, 1/(p-1)\}$. The proof was based on several reductions and depended heavily on the dyadic structure. The general case for arbitrary filtrations was established independently by Thiele, Treil and Volberg [83] and Lacey [44]. The extension of the inequality to continuous-time, differentially subordinate and uniformly integrable martingales was obtained in the recent work of Domelevo and Petermichl [29]:

$$\|Y\|_{L^p(W)} \leq c_p[W]_{A_p}^{\max\{1,1/(p-1)\}}\|X\|_{L^p(W)},$$

for every $1 < p < \infty$. Here, as before, $\|X\|_{L^p(W)}$ and $\|Y\|_{L^p(W)}$ are the usual weighted L^p-norms of the limiting random variables $\lim_{t\to\infty} X_t$ and $\lim_{t\to\infty} Y_t$. The first step is to use extrapolation: Theorem 1.16 shows that it is enough to study the above inequality for the case $p = 2$ only:

$$||Y||_{L^2(W)} \leq c_2[W]_{A_2}||X||_{L^2(W)}. \tag{2.1.2}$$

The proof of this special bound, presented in [29], rests on duality and a number of complicated Bellman functions (involving six variables). There is a natural question whether the estimate (2.1.2) can be established directly, in the spirit of Burkholder's approach described above. The principal purpose of this chapter is to answer this question in the affirmative. The appearance of A_2 weights forces the introduction of two additional arguments into the function and hence the problem reduces to the construction of an explicit function of four variables, enjoying the appropriate concavity and size conditions similar to 1°-3° above. Here is the result for continuous-time, continuous-path martingales, obtained by Bañuelos, Osękowski and the author [5]. As discussed in the introduction we will assume that the martingale $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$ generated by the weight W, has continuous paths. Interestingly, as a by-product, our approach will allow us to obtain the following stronger maximal estimate.

Theorem 2.1. *Suppose that W is an A_p weight and X, Y are continuous-path martingales such that Y is stochastic integral, with respect to X, of some predictable process H taking values in [-1,1]. Then for any $1 < p < \infty$ there is finite constant C_p depending only on p such that*

$$|||Y|^*||_{L^p(W)} \leq C_p[W]_{A_p}^{\max\{1,1/(p-1)}||X||_{L^p(W)}. \tag{2.1.3}$$

The exponent $\max\{1, 1/(p-1)\}$ *is the best possible.*

In our considerations below, we will focus on the proof of (2.1.3) in the case $p = 2$; as we have already noted above, this special case yields the L^p bound for all p via extrapolation.

The optimality of the exponent in (2.1.3) is well-known and can be extracted from [29, 44, 83]. We will not address this issue here.

2.2. Bellman function method for weighted inequalities

Let $1 < p < \infty$. As we have described in the previous chapter, any A_p weight W satisfying $[W]_{A_p} \leq c$ can be identified with a two-dimensional martingale (W, V) taking values in the domain

$$\mathcal{D}_c = \big\{(w, v) \in (0,\infty) \times (0,\infty) : 1 \leq wv^{p-1} \leq c\big\}$$

and terminating at the lower boundary of this set: $W_\infty V_\infty^{p-1} = 1$ almost surely. Now, let $c \geq 1$ be a fixed parameter. Suppose that $G : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ is a given Borel function and assume that we are interested in showing that

$$\mathbb{E}G(X_t, Y_t, W_t, V_t) \leq 0, \qquad t \geq 0. \tag{2.2.1}$$

Here (X, Y) is an arbitrary pair of bounded martingales such that Y is the stochastic integral, with respect to X, of some predictable process with values in $[-1, 1]$, and (W, V) is a pair associated with some A_p weight of characteristic not bigger than c. A key to handle this problem is to consider a C^2 function $B : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ which satisfies the following conditions:

1° (Initial condition) We have $B(x, y, w, v) \leq 0$ if $|y| \leq |x|$ and $1 \leq wv^{p-1} \leq c$.
2° (Majorization property) We have $B \geq G$ on $\mathbb{R}^2 \times \mathcal{D}_c$.
3° (Concavity-type property) For any $(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c$ and $d, e, r, s \in \mathbb{R}$ satisfying $|e| \leq |d|$, the function

$$\xi_B(t) := B(x + td, y + te, w + tr, v + ts),$$

given for those t, for which $1 \leq (w + tr)(v + ts)^{p-1} \leq c$, is locally concave.

The connection between the existence of such a function and the validity of (2.2.1) is described in the lemma below.

Lemma 2.2. *Let $1 < p < \infty$ and $c \geq 1$ be fixed. If there is a function B satisfying conditions 1°, 2° and 3°, then the inequality (2.2.1) holds true for all X, Y, W and V as above.*

Proof. The argument rests on Itô's formula. Consider a process $Z = (X, Y, W, V)$. Since B is of class C^2, we may write

$$B(Z_t) = I_0 + I_1 + I_2/2 + I_3,$$

where

$$\begin{aligned}
I_0 &= B(Z_0),\\
I_1 &= \int_{0+}^t B_x(Z_u)\mathrm{d}X_u + \int_{0+}^t B_y(Z_u)\mathrm{d}Y_u + \int_{0+}^t B_w(Z_u)\mathrm{d}W_u + \int_{0+}^t B_v(Z_u)\mathrm{d}V_u,\\
I_2 &= \int_{0+}^t D^2B(Z_u)\mathrm{d}[Z^c]_u,\\
I_3 &= \sum_{0<s\leq t} (B(Z_u) - B(Z_{u-}) - B_v(Z_{u-})\Delta V_u).
\end{aligned}$$

Here ξ^c denotes continuous part of a semimartingale, D^2B is the Hessian matrix of B and in definition of I_2 we have used the shortened notation for the sum of all second-order terms. The appearance of the term I_3 is due to the fact that X, Y, W have continuous paths (as we have assumed), but V does not have to possess this property. Let us study the properties of the terms I_0, I_1, I_2 and I_3. The first of them is nonpositive because of the condition 1°. The expectation of I_1 is zero, by the properties of stochastic integrals. To handle the next term, note that by a simple differentiation, 3° implies that

$$\langle D^2B(x, y, w, v)(d, e, r, s), (d, e, r, s)\rangle \leq 0$$

for any $(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c$ and any $(d, e, r, s) \in \mathbb{R}^4$ satisfying $|e| \leq |d|$. This implies $I_2 \leq 0$ by a straightforward approximation of the integral by Riemann sums. Indeed, pick an arbitrary partition $0 < s_0 < s_1 < \ldots < s_N < t$ and an integer $i \in \{0, \ldots, N-1\}$. Next, set $\Delta X = X_{s_{i+1}} - X_{s_i}$ and define $\Delta W, \Delta V^c$ similarly. By the concavity condition

$$\left\langle D^2 B(Z_{s_i})(\Delta X, H_{s_i}\Delta X, \Delta W, \Delta V^c), (\Delta X, H_{s_i}\Delta X, \Delta W, \Delta V^c)\right\rangle \leq 0.$$

Summing over all i and letting the diameter of the partition (s_i) go to zero, one obtains $I_2 \leq 0$. Finally, I_3 is also nonpositive, which is again due to 3°: setting $e = d = r = 0$, we see that for any x, y, z, w and v, the function $s \mapsto B(x, y, w, v + s)$ is concave (on the set of s satisfying $1 \leq w(v + s)^{p-1} \leq c$).

Putting all the above observations together, we get $\mathbb{E}B(Z_t) \leq 0$, which combined with the majorization condition 2° gives the assertion. □

We conclude this section with three observations.

Remark 2.1. The above statement is true without the assumption that B is of class C^2: it is enough to ensure that B is continuous. Indeed, the condition 3° guarantees that any possible "cusp" of B is of concave type and hence the argument works. More precisely, this can be proved by standard mollification argument (consult e.g. Domelevo and Petermichl [29] or Wang [86]). There are also several other methods of handling the irregularity of B: one can use the appropriate extension of Itô's formula developed in [68]; alternatively, one can first establish the corresponding estimate for (discrete-time) martingales and use approximation: see [19] for details.

Remark 2.2. The above approach works also in the unweighted case, which corresponds to the choice $c = 1$. Then the processes W and V are constant, and hence the special function B depends only on the variables x, y. This brings us back to the original setting considered by Burkholder.

Remark 2.3. The above approach is very flexible and can be easily modified to other contexts. For example, suppose that we are interested in the maximal bound of the form

$$\mathbb{E}G(X_t, Y_t, Y_t^*, W_t, V_t) \leq 0, \qquad t \geq 0,$$

for all X, Y, W and V as in (2.2.1) (recall that $Y_t^* = \max_{0 \leq s \leq t} Y_s$ is the truncated one-sided maximal function of Y). Then it is enough to construct $B : \{(x, y, z, w, v) \in \mathbb{R}^3 \times \mathcal{D}_c : y \leq z\} \to \mathbb{R}$ satisfying

1° (Initial condition) We have $B(x, y, y, w, v) \leq 0$ if $|y| \leq |x|$ and $1 \leq wv^{p-1} \leq c$.
2° (Majorization property) We have $B \geq G$.
3° (Concavity-type property) For any $(x, y, z, w, v) \in \mathbb{R}^3 \times \mathcal{D}_c$ and $d, e, r, s \in \mathbb{R}$ satisfying $y < z$ and $|e| \leq |d|$, the function

$$\xi_B(t) := B(x + td, y + te, z, w + tr, v + ts),$$

given for those t, for which $1 \leq (w + tr)(v + ts)^{p-1} \leq c$, is locally concave.
4° We have $B_z(x, y, y, w, v) \leq 0$ for all (x, y, y, w, v) from the domain of B.

Again, the proof rests on Itô's formula. The extra requirement 4° enables us to handle the additional stochastic integral $\int_{0+}^t B_z(X_s, Y_s, Y_s^*, W_s, V_s)\mathrm{d}Y_s^*$ (and guarantees that this integral is nonpositive).

2.3. Construction of a special function: a toy example and some rough arguments

From now on, we restrict ourselves to the special case $p = 2$. The aim of this section is to describe some informal steps which led us to the Bellman function corresponding to (2.1.2). The "value" function G is given by

$$G(x, y, w, v) = y^2 w - Cc^2 x^2 w.$$

A toy example. To gain some intuition about the problem, let us study the weaker case in which the weight W is replaced by W^α for some fixed $\alpha \in (0, 1)$:

$$\|Y\|_{L^2(W^\alpha)} \leq C\|X\|_{L^2(W^\alpha)} \tag{2.3.1}$$

(where C depends on $[W]_{A_2}$). At the first glance, it is absolutely not clear how to construct, or guess, the candidate for the Bellman function and a natural idea is to consider first the simpler unweighted setting. The unweighted Burkholder L^2 inequality

$$\|Y\|_{L^2} \leq \|X\|_{L^2}$$

is trivial (we have $\|Y\|_{L^2} = \|[Y]\|_{L^1}^{1/2} \leq \|[X]\|_{L^1}^{1/2} = \|X\|_{L^2}$), it can also be proved with the Bellman function of two variables, equal to

$$b(x, y) = y^2 - x^2.$$

We can expect that the passage to the weighted setting requires adding the variable w to this function. How to do this? The function G corresponding to (2.3.1) is given by

$$G(x, y, w, v) = y^2 w^\alpha - Cx^2 w^\alpha. \tag{2.3.2}$$

Since W is an A_2 weight, we have $W_\infty V_\infty = 1$ and hence replacing G with

$$(x, y, w, v) \mapsto y^2 v^{-\alpha} - Cx^2 v^{-\alpha}$$

would also lead to the desired estimate. Actually, one could replace G with any function of the form

$$(x, y, w, v) \mapsto y^2 w^\alpha \phi(wv) - Cx^2 w^\alpha \psi(wv), \tag{2.3.3}$$

for some functions $\phi : [1, c] \to \mathbb{R}$ and $\psi : [1, c] \to \mathbb{R}$ satisfying $\phi(1) = \psi(1) = 1$ (the appearance of the intervals $[1, c]$ is enforced by the domain of the Bellman function or, which is the same, by the condition $1 \leq W_t V_t \leq c$). While we will stick to the definition (2.3.2), the formula (2.3.3) is a nice guess for the Bellman function B.

How to find ϕ and ψ? We want B to satisfy 3°, so, roughly speaking, we want the term $y^2 w^\alpha \phi(wv)$ to be (appropriately) concave and $x^2 w^\alpha \psi(wv)$ to be (appropriately) convex. It is easy to *guess* ψ: it is well known that for $\alpha \in (0, 1]$, the function $(x, v) \mapsto x^2 v^{-\alpha}$ is (fully) convex on $\mathbb{R} \times (0, \infty)$, so we take $\psi(s) = s^{-\alpha}$. For ϕ the situation is a little more difficult. The appearance of the factor y^2 shows that it is *impossible* to make $y^2 w^\alpha \phi(wv^{p-1})$ become a concave function. Anyhow, we can hope that the "partial convexity" induced by this term will be overpowered by the concavity generated by the second term. We make the simple choice $\phi(s) = 1$, which transforms the first term of the Bellman function into $y^2 w^\alpha$. Hence we have obtained the candidate

$$B(x, y, w, v) = y^2 w^\alpha - Cx^2 v^{-\alpha}.$$

Let us verify rigorously the condition 3°. We compute that the Hessian $D^2B(x,y,w,v)$ is

$$\begin{bmatrix} -2Cv^{-\alpha} & 0 & 0 & 2\alpha Cxv^{-\alpha-1} \\ 0 & 2w^{\alpha} & 2\alpha yw^{\alpha-1} & 0 \\ 0 & 2\alpha yw^{\alpha-1} & \alpha(\alpha-1)y^2w^{\alpha-2} & 0 \\ 2\alpha Cxv^{-\alpha-1} & 0 & 0 & -\alpha(\alpha+1)Cx^2v^{-\alpha-2} \end{bmatrix}.$$

We must prove that $\langle D^2B(x,y,w,v)(d,e,r,s),(d,e,r,s)\rangle \le 0$ provided $|e|\le|d|$. We easily check that the 2×2 matrix created by removing the second and third rows/columns:

$$\begin{bmatrix} -2Cv^{-\alpha} & 2\alpha Cxv^{-\alpha-1} \\ 2\alpha Cxv^{-\alpha-1} & -\alpha(\alpha+1)Cx^2v^{-\alpha-2} \end{bmatrix}$$

is negative definite. Actually, it is "strongly negative definite" in the sense that we can increase the term in the upper-left corner by $\frac{2(1-\alpha)}{1+\alpha}Cv^{-\alpha}$, and this sign property will remain unchanged. Now, this surplus can be used to enforce the matrix

$$\begin{bmatrix} 2w^{\alpha} & 2\alpha yw^{\alpha-1} \\ 2\alpha yw^{\alpha-1} & \alpha(\alpha-1)y^2w^{\alpha-2} \end{bmatrix} \tag{2.3.4}$$

(obtained from the Hessian $D^2B(x,y,w,v)$ by removing the first and last rows/columns) to be seminegative definite. Here we use the fact that $|e|\le|d|$: the surplus $\frac{2(1-\alpha)}{1+\alpha}Cv^{-\alpha}$, with the appropriate choice of C, turns the term $2w^{\alpha}$ into an appropriately small negative number. One can check that C must be at least $C'c^2$, for some universal C'; this is how the desired linear dependence on the characteristic $[W]_{A_2}$ comes into play.

The return to (2.1.2). The above reasoning does not work for $\alpha=1$: there is no chance of making the matrix (2.3.4) seminegative definite. However, we will use some observations from the previous setting. First, we try to take

$$B(x,y,w,v) = y^2w\phi(wv) - x^2w\psi(wv). \tag{2.3.5}$$

As above, one could rewrite the concavity condition 3° as $D^2B\le 0$ (on appropriate quadruples (d,e,r,s)) and take a look at the resulting differential inequalities for ϕ and ψ. After a lot of further experimentation, it turns out that the choices (up to some multiplicative constants depending only on c)

$$\phi(t) = 2 - \frac{1}{t} - \frac{\ln(t)}{2c}, \qquad \psi(t) = (t\phi(t))^{-1}$$

lead to the Bellman function satisfying 1°, 2° and 3°, but the majorization holds with

$$\tilde{G}(x,y,w,v) = \kappa(y^2w - Cc^3x^2w).$$

Hence, the above construction implies the inequality of the form

$$\|Y\|_{L^2(W)} \le C[W]_{A_2}^{3/2}\|X\|_{L^2(W)},$$

with suboptimal exponent 3/2. Unfortunately, to get the sharper inequality, we need some further modifications. It turns out that the function constructed above would lead to the optimal inequality if we could restrict the domain to the "lower angular" subset of

$\mathbb{R}^2 \times \mathcal{D}_c$ given by $\{(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c : |y| \leq C_1|x|\}$. On the other hand, in the "upper angular" subset $\{(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c : |y| \geq C_2c|x|\}$ we can use the function of the form

$$(x, y, w, v) \mapsto y^2 w\phi(wv) - \frac{y^2}{2v} - Kc^2x^2v^{-1}.$$

Then it remains to find an appropriate function in the "middle angular" subset $\{(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c : C_1|x| \leq |y| \leq C_2c|x|\}$. Unfortunately, it seems that functions of the form (2.3.5) do not work there and we are forced to use a mixed summand depending on both x and y, of the form $|x||y|wf(wv)$ for some function f.

In the following section we will give a precise definition of the function B and check that it satisfies all required conditions.

2.4. A special function

Throughout this section, $c > 1$ is a fixed parameter (which corresponds to the "truly" weighted context). The main result of this section is the following.

Theorem 2.3. *Let $p = 2$. There is a continuous function $B : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ satisfying $1°$-$3°$ with $G(x, y, w, v) = \kappa(y^2w - C^2c^2x^2v^{-1})$ for some positive constants κ and C.*

Note that the assertion is slightly stronger than we need: the function G appearing above majorizes the function $(x, y, w, v) \mapsto \kappa(y^2w - C^2c^2x^2w)$ which would be sufficient for (2.1.2). However, we will need this stronger form in our considerations below, e.g. when passing from the continuous- to the discrete-time case.

The above statement combined with Lemma 2.2 yields the validity of (2.1.2), by passing $t \to \infty$ and using standard limiting arguments. A slightly stronger, maximal estimate announced in Introduction will be proved at the end of this section. Assume the D^1, D^2, D^3 are the 'angular' subsets of $\mathbb{R}^2 \times \mathcal{D}_c$, given by

$$\begin{aligned}
D^1 &= \{(x, y, w, v) : |y| \geq 20c|x|(c/t)^{1-\beta}\},\\
D^2 &= \{(x, y, w, v) : 10|x| \leq |y| \leq 20c|x|(c/t)^{1-\beta}\},\\
D^3 &= \{(x, y, w, v) : |y| \leq 10|x|\},
\end{aligned}$$

where $\beta = 3/4$. Here and in what follows, we denote $t = wv$. Define the functions $b_i : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ by

$$\begin{aligned}
b_1(x, y, w, v) &= y^2w\phi(wv),\\
b_2(x, y, w, v) &= y^2(2v)^{-1},\\
b_3(x, y, w, v) &= c^2x^2v^{-1},\\
b_4(x, y, w, v) &= c^\beta|x||y|w^{1-\beta}v^{-\beta},\\
b_5(x, y, w, v) &= c^\beta y^2w^{1-\beta}v^{-\beta},\\
b_6(x, y, w, v) &= c^2x^2w\psi(wv),
\end{aligned}$$

where ϕ, ψ are functions from $[1, c]$ to $\mathbb{R}$ given by

$$\phi(t) = 2 - \frac{1}{t} - \frac{\ln(t)}{2c}, \qquad \psi(t) = (t\phi(t))^{-1}.$$

Furthermore, set $U(x,y,w,v) = b_1 - b_2 - 320000b_3 - 294400b_6$. Now we are finally ready to introduce the explicit formula for the desired Bellman function B:

$$B(x,y,w,v) = \begin{cases} B_1(x,y,w,v) & \text{on } D^1, \\ B_2(x,y,w,v) & \text{on } D^2, \\ B_3(x,y,w,v) & \text{on } D^2, \end{cases}$$

where $B_1, B_2, B_3 : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ are given by

$$\begin{aligned} B_1(x,y,w,v) &= U(x,y,w,v) + 6400c^2x^2v^{-1}, \\ B_2(x,y,w,v) &= U(x,y,w,v) + 320b_4, \\ B_3(x,y,w,v) &= U(x,y,w,v) + 32b_5. \end{aligned}$$

As we have already announced earlier, this function has quite a complicated form, it is actually defined with three different formulas on three separate domains. Let us verify that it satisfies conditions 1°-3° listed in the formulation of Lemma 2.2. The first two properties are relatively easy to prove; the main difficulty lies in establishing the concavity condition. We start with the easy part.

Lemma 2.4. *The function B satisfies the properties 1° and 2°. Furthermore, we have $B_1 \leq B_2$ on D^1, $B_2 \leq \min(B_1, B_3)$ on D^2 and $B_3 \leq B_2$ on D^3.*

Proof. To check the initial condition, note that for $|y| \leq |x|$ we have $(x,y,w,v) \in D^3$. Furthermore, from $\phi(t) \leq 2$, we obtain

$$B_3(x,y,w,v) \leq b_1 + 32b_5 - 320000b_3 \leq x^2w(2 + 32(c/t)^\beta - 320000(c/t)c) \leq 0.$$

Let us now study the majorization. Observe that

$$b_1 - b_2 = y^2w\left[2 - \frac{1}{t} - \frac{\ln t}{2c} - \frac{1}{2t}\right] \geq \frac{1}{2}y^2w,$$

because the function in the square bracket is increasing and have its minimum at the point $t = 1$. Now, from the estimate $\phi(t) \geq 1$, we have that $\psi(t) \leq 1/t$ and, as a consequence,

$$320000b_3 + 294400b_6 \leq c^2x^2w\left(320000\frac{1}{t} + 294400\frac{1}{t}\right) \leq 614400c^2x^2v^{-1}.$$

Finally, we have

$$\begin{aligned} B \geq b_1 - b_2 - 320000b_3 - 294400b_6 &\geq \frac{1}{2}y^2w - 614400c^2x^2v^{-1} \\ &= \frac{1}{2}(y^2w - 1228800c^2x^2v^{-1}), \end{aligned}$$

so the condition 2° in Theorem 2.3 is satisfied with $\kappa = 1/2$ and $C = (1228800)^{1/2} < 1109$.

It remains to verify the relations between B_1, B_2 and B_3. If $(x,y,w,v) \in D^1$, then

$$320b_4 = 320c^\beta|x||y|w^{1-\beta}v^{-\beta} \geq 6400c^2x^2v^{-1},$$

so $B_2 \geq B_1$. If $(x,y,w,v) \in D^2$, we have reverse inequality $B_2 \leq B_1$. Furthermore, on D^2 we have

$$320b_4 = 320c^\beta|x||y|w^{1-\beta}v^{-\beta} \leq 32c^\beta|y|^2w^{1-\beta}v^{-\beta} = 32b_5,$$

which is exactly $B_2 \leq B_3$. To finish the proof, observe that the above estimate is reversed on D_3. □

We turn our attention to the crucial condition 3°. From symmetry (and the equality $B_x(0, y, w, v) = 0$ for all y, w, v), without loss of generality, we may only consider points $(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c$ such that x and y are nonnegative. Furthermore, by the second part of the above lemma, it is enough to verify the version of concavity "localized" to each D^i. More precisely, it suffices to show that for each i, each $(x, y, w, v) \in D^i$ and any $d, e, r, s \in \mathbb{R}$ satisfying $|e| \le |d|$, the function

$$\xi_{B_i}(t) := B_i(x + td, y + te, w + tr, v + ts),$$

given for those t, for which $(x + td, y + te, w + tr, v + ts) \in D^i$, satisfies $\xi''_{B_i}(0) \le 0$. This localized concavity will be accomplished by a careful analysis of the derivatives $\xi''_{b_j}(0)$ of the building blocks $b_j, j = 1, 2, \dots, 6$. In the next lemma we gather estimates for the parts b_1 and b_6.

Lemma 2.5. *We have following estimates on the quadratic forms associated with the functions b_1 and b_6.*

(a) $\xi''_{b_1}(0) \le 80cwe^2$,
(b) $\xi''_{b_1}(0) \le 4we^2 + 8y|e||r|$,
(c) $\xi''_{b_6}(0) \ge (1/16)cv^{-3}x^2s^2$.

Proof. (a) It is equivalent to showing nonpositive-definiteness of the matrix

$$\mathbb{A}(y, w, v) = \begin{pmatrix} 2w\phi(t) - 80cw & 2y\phi(t) + 2yt\phi'(t) & 2yw^2\phi'(t) \\ 2y\phi(t) + 2yt\phi'(t) & 2y^2v\phi'(t) + y^2tv\phi''(t) & 2y^2w\phi'(t) + y^2tw\phi''(t) \\ 2yw^2\phi'(t) & 2y^2w\phi'(t) + y^2tw\phi''(t) & y^2w^3\phi''(t) \end{pmatrix}.$$

From Sylvester's criterion it is enough to prove that

$$y^2w^3\phi''(t) \le 0, \tag{2.4.1}$$

$$\det \begin{pmatrix} 2y^2v\phi'(t) + y^2tv\phi''(t) & 2y^2w\phi'(t) + y^2tw\phi''(t) \\ 2y^2w\phi'(t) + y^2tw\phi''(t) & y^2w^3\phi''(t) \end{pmatrix} \ge 0 \tag{2.4.2}$$

and

$$\det \mathbb{A}(y, w, v) \le 0. \tag{2.4.3}$$

The inequality (2.4.1) follows immediately from $t \in [1, c]$ and the estimate

$$y^2w^3\phi''(t) = -\frac{y^2w^3}{2ct^3}(4c - t) \le 0.$$

The inequality (2.4.2) is equivalent to $\phi'(t)(2\phi'(t) + t\phi''(t)) \le 0$ and follows from

$$\phi'(t) = \frac{1}{2ct^2}(2c - t) \ge 0 \qquad \text{and} \qquad 2\phi'(t) + t\phi''(t) = -\frac{1}{2ct} \le 0.$$

In order to show (2.4.3) we simplify the matrix $\mathbb{A}$ by carrying out some elementary operations. The determinant of $\mathbb{A}$ has the same sign as

$$\det \begin{pmatrix} -80cw & 2\phi(t) + 2t\phi'(t) & 0 \\ 2\phi(t) & 0 & 2\phi'(t) \\ 2w\phi'(t) & 2\phi'(t) + t\phi''(t) & w\phi''(t) \end{pmatrix}$$
$$= 4w\left[\left(2(\phi'(t))^2 - \phi(t)\phi''(t)\right)(\phi(t) + t\phi'(t)) + 40c\phi'(t)(2\phi'(t) + t\phi''(t))\right].$$

We compute that

$$\phi(t) + t\phi'(t) = 2 - \frac{\ln(t)}{2c} - \frac{1}{2c} \leq 2, \tag{2.4.4}$$

$$2(\phi'(t))^2 = \phi'(t)\frac{2c-t}{ct^2} \leq \frac{2\phi'(t)}{t}$$

and, since $\phi(t) \leq 2$,

$$-\frac{\phi(t)\phi''(t)}{\phi'(t)} \leq \frac{2\left(\frac{2}{t^3} - \frac{1}{2ct^2}\right)}{\frac{1}{t^2} - \frac{1}{2ct}} \leq \frac{8}{t}. \tag{2.4.5}$$

Combining these facts we obtain

$$(2(\phi'(t))^2 - \phi(t)\phi''(t))(\phi(t) + t\phi'(t)) \leq \frac{20\phi'(t)}{t},$$

and since

$$40c\phi'(t)(2\phi'(t) + t\phi''(t)) = -\frac{20\phi'(t)}{t},$$

the inequality (2.4.3) is satisfied. This completes the proof of the part (a).

(b) Firstly, observe that it is sufficient to prove nonpositive-definiteness of the matrix

$$\mathbb{B}(y,w,v) = \begin{pmatrix} 2w\phi(t) - 4w & 0 & 2yw^2\phi'(t) \\ 0 & 2y^2v\phi'(t) + y^2tv\phi''(t) & 2y^2w\phi'(t) + y^2tw\phi''(t) \\ 2yw^2\phi'(t) & 2y^2w\phi'(t) + y^2tw\phi''(t) & y^2w^3\phi''(t) \end{pmatrix}.$$

Indeed, we have the estimate

$$\begin{aligned} \xi''_{b_1}(0) &= \langle \mathbb{B}(y,w,v)(e,r,s),(e,r,s)\rangle + 4we^2 + 2(2y\phi(t) + 2yt\phi'(t))er \\ &\leq 4we^2 + 4(y\phi(t) + yt\phi'(t))|e||r| \leq 4we^2 + 8y|e||r|, \end{aligned}$$

where the last inequality follows from (2.4.4).

From Sylvester's criterion nonpositive-definiteness of the matrix $\mathbb{B}$ is equivalent to inequalities (2.4.1) and (2.4.2) (which we have already shown in the proof of the (a) part of the lemma) and the estimate

$$\det \mathbb{B}(y,w,v) \leq 0. \tag{2.4.6}$$

By carrying out some elementary operations we show that the determinant of $\mathbb{B}$ has the same sign as

$$\begin{aligned} &\det \begin{pmatrix} 2w\phi(t) - 4w + 2wt\phi'(t) & 0 & 2w\phi'(t) \\ 0 & 0 & 2\phi'(t) \\ 2w\phi'(t) + wt\phi''(t) & 2\phi'(t) + t\phi''(t) & w\phi''(t) \end{pmatrix} \\ &= -(2\phi'(t) + t\phi''(t))2\phi'(t)(2w\phi(t) - 4w + 2wt\phi'(t)). \end{aligned}$$

Now, we compute that

$$2\phi'(t) + t\phi''(t) = -\frac{1}{2tc} \leq 0.$$

Hence, since $\phi(t)+t\phi'(t)\leq 2$ and $\phi'(t)\geq 0$, the inequality (2.4.6) is satisfied.

(c) In analogy to the above considerations, we must show that the matrix

$$\mathcal{C}(x,w,v) = \begin{pmatrix} 2c^2w\psi(t) & 2xc^2(\psi(t)+t\psi'(t)) & 2xc^2w^2\psi'(t) \\ 2xc^2(\psi+t\psi'(t)) & x^2c^2(2v\psi'(t)+wv^2\psi''(t)) & x^2c^2(2w\psi'(t)+w^2v\psi''(t)) \\ 2xc^2w^2\psi'(t) & x^2c^2(2w\psi'(t)+w^2v\psi''(t)) & x^2c^2w^3\psi''(t)-\frac{1}{16}cv^{-3}x^2 \end{pmatrix}$$

is nonnegative-definite. For notational convenience, let us define the function $\widehat{\psi}:[1,c]\longrightarrow\mathbb{R}$ as $\widehat{\psi}(t)=t\psi(t)=(\phi(t))^{-1}$ and set

$$d(x,t)=x^2c^2\left(2v^{-3}\widehat{\psi}-2v^{-2}w\widehat{\psi}'+v^{-1}w^2\widehat{\psi}''\right)-\frac{1}{16}v^{-3}cx^2.$$

Then we can rewrite the matrix $\mathcal{C}$ as

$$\begin{pmatrix} 2c^2\widehat{\psi}(t)v^{-1} & 2c^2x\widehat{\psi}'(t) & 2xc^2\left(\widehat{\psi}'(t)wv^{-1}-\widehat{\psi}(t)v^{-2}\right) \\ 2c^2x\widehat{\psi}'(t) & c^2vx^2\widehat{\psi}''(t) & c^2wx^2\widehat{\psi}''(t) \\ 2xc^2\left(\widehat{\psi}'(t)wv^{-1}-\widehat{\psi}(t)v^{-2}\right) & c^2wx^2\widehat{\psi}''(t) & d(x,t) \end{pmatrix},$$

Again, from Sylvester's criterion, we reduce the problem to checking the signs of appropriate minors. More precisely, we will show that

$$2c^2\widehat{\psi}v^{-1}\geq 0, \tag{2.4.7}$$

$$\det\begin{pmatrix} 2c^2\widehat{\psi}(t)v^{-1} & 2c^2x\widehat{\psi}'(t) \\ 2c^2x\widehat{\psi}'(t) & c^2vx^2\widehat{\psi}''(t)\end{pmatrix}\geq 0, \tag{2.4.8}$$

$$\det\mathcal{C}(x,w,v)\geq 0. \tag{2.4.9}$$

The inequality (2.4.7) is obvious. Condition (2.4.8) is equivalent to $\widehat{\psi}(t)\widehat{\psi}''(t)-2(\widehat{\psi}'(t))^2\geq 0$, which is a consequence of the definition $\widehat{\psi}(t)=(\phi(t))^{-1}$ and the inequality $\phi''(t)\leq 0$. To show (2.4.9), we perform certain elementary operations on the columns and rows of the matrix to prove that the determinant of $\mathcal{C}$ has the same sign as

$$\det\begin{pmatrix} 2\widehat{\psi}(t)v^{-1} & 2\widehat{\psi}'(t) & 0 \\ 2\widehat{\psi}'(t) & v\widehat{\psi}''(t) & 2v^{-1}\widehat{\psi}'(t) \\ -2\widehat{\psi}(t)v^{-2} & 0 & -2wv^{-2}\widehat{\psi}'(t)-\frac{1}{16}v^{-3}c^{-1}\end{pmatrix}$$
$$=2v^{-3}\left(\left(-2\widehat{\psi}'(t)t-\frac{1}{16}c^{-1}\right)(\widehat{\psi}(t)\widehat{\psi}''(t)-2(\widehat{\psi}'(t))^2)-4\widehat{\psi}(t)(\widehat{\psi}'(t))^2\right).$$

We compute that

$$\widehat{\psi}'(t)=-\frac{\phi'(t)}{\phi^2(t)},$$

$$\widehat{\psi}''(t)=-\frac{\phi(t)\phi''(t)-2(\phi'(t))^2}{\phi^3(t)}$$

and

$$\widehat{\psi}(t)\widehat{\psi}''(t) - 2(\widehat{\psi}'(t))^2 = -\frac{\phi''(t)}{\phi^3(t)}.$$

Hence, we need to show the estimate

$$-2\phi'(t)\phi''(t)t + \frac{1}{16}c^{-1}\phi''(t)(\phi(t))^2 - 4(\phi'(t))^2 \geq 0.$$

Now observe that

$$-2\phi'(t)\phi''(t)t - 4(\phi'(t))^2 = -2\phi'(t)(2\phi'(t) + \phi''(t)t) = \phi'(t)c^{-1}t^{-1}$$

and from (2.4.5) and $\phi(t) \leq 2$

$$\frac{1}{16}c^{-1}\phi''(t)(\phi(t))^2 \geq \frac{-\phi(t)\phi'(t)}{2ct} \geq -\frac{\phi'(t)}{ct},$$

which completes the proof of the lemma. □

In the series of three lemmas below we will show that the function B satisfies the required concavity condition. Let us start with the domain D^1.

Lemma 2.6. *We have $\xi''_{b_1-b_2-160b_3}(0) \leq 0$ for any $(x, y, w, v) \in D^1$ and any (d, e, r, s) such that $|e| \leq |d|$.*

Remark 2.4. This lemma handles the property 3° on the domain D^1. Indeed, the additional summands $-313440b_3$ and $-294400b_6$ are concave functions (concavity of $-b_3$ is easy to check, concavity of $-b_6$ follows from part (c) of Lemma 2.5).

Proof of Lemma 2.6. We have that

$$\xi''_{b_2}(0) = \frac{1}{v}\left(e - \frac{ys}{v}\right)^2, \qquad \xi''_{b_3}(0) = \frac{2c^2}{v}\left(d - \frac{xs}{v}\right)^2.$$

Now consider two cases. If $|d - \frac{xs}{v}| \geq d/2$, then from above formulas and part (a) of Lemma 2.5 we obtain

$$\begin{aligned}\xi''_{b_1-b_2-160b_3}(0) &\leq 80cwe^2 - \frac{1}{v}\left(e - \frac{ys}{v}\right)^2 - \frac{320c^2}{v}\left(d - \frac{xs}{v}\right)^2 \\ &\leq 80cwe^2 - 80c^2v^{-1}d^2 \\ &\leq 80cwd^2 - 80cwd^2 \\ &= 0\end{aligned}$$

If $|d - \frac{xs}{v}| < d/2$, then

$$\begin{aligned}\frac{ys}{dv} - \frac{e}{d} = \frac{ys}{dv} - \frac{y}{x} + \frac{y}{x} - \frac{e}{d} &= y\left(\frac{s}{dv} - \frac{1}{x}\right) + \frac{y}{x} - \frac{e}{d} = \frac{y}{x}\left(\frac{sx}{vd} - 1 + 1 - \frac{e}{d}\frac{x}{y}\right) \\ &\geq \frac{y}{x}\left(-\frac{1}{2} + 1 - \frac{1}{20c}\right) \geq 20c\left(\frac{1}{2} - \frac{1}{20c}\right) = 10c - 1 \geq 9c.\end{aligned}$$

Hence

$$\xi''_{b_1-b_2-160b_3}(0) \leq \xi''_{b_1-b_2}(0) \leq 80cwe^2 - \frac{1}{v}d^2 9^2 c^2 \leq 80cwd^2 - 81cwd^2 \leq 0.$$

The proof is complete. □

The next lemma discusses the concavity condition in the middle domain D^2.

Lemma 2.7. *We have* $\xi''_{b_1+320b_4-320000b_3}(0) \leq 0$ *for any* $(x, y, w, v) \in D^2$ *and any* (d, e, r, s) *such that* $|e| \leq |d|$.

Remark 2.5. This lemma handles the property 3° on the domain D^2. Indeed, functions $-b_2$ and $-294400b_6$ are concave, so they do not affect the condition 3°.

Proof of Lemma 2.7. Let $D = \frac{d}{x}$, $E = \frac{e}{y}$, $R = \frac{r}{w}$ and $S = \frac{s}{v}$. We compute that

$$\begin{aligned}\xi_{b_4-1000b_3}(0) = c^\beta xyw^{1-\beta}v^{-\beta}\left\langle A_1(D,E,R,S),(D,E,R,S)\right\rangle \\ -1000c^2x^2v^{-1}\left\langle A_2(D,S),(D,S)\right\rangle,\end{aligned}$$

where matrices A_1 and A_2 are defined as

$$A_1 = \begin{pmatrix} 0 & 1 & 1-\beta & -\beta \\ 1 & 0 & 1-\beta & -\beta \\ 1-\beta & 1-\beta & \beta(\beta-1) & \beta(\beta-1) \\ -\beta & -\beta & \beta(\beta-1) & \beta(\beta+1) \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}.$$

From the assumption $y \geq 10x$ and differential subordination ($|e| \leq |d|$) we obtain

$$|E| = \left|\frac{e}{y}\right| \leq \frac{1}{10}\left|\frac{d}{x}\right| = \frac{1}{10}|D|,$$

so $E = \lambda D$, where λ is a constant with absolute value bounded by $\frac{1}{10}$. We can reduce Hessians to three variables (D,R and S): we have

$$\begin{aligned}\xi''_{b_4-1000b_3}(0) = c^\beta xyw^{1-\beta}v^{-\beta}\left\langle A_3(D,R,S),(D,R,S)\right\rangle \\ -1000c^2x^2v^{-1}\left\langle A_4(D,R,S),(D,R,S)\right\rangle,\end{aligned} \tag{2.4.10}$$

where matrices A_3 and A_4 are defined as

$$A_3 = \begin{pmatrix} 2\lambda & (1-\beta)(1+\lambda) & -\beta(1+\lambda) \\ (1-\beta)(1+\lambda) & \beta(\beta-1) & \beta(\beta-1) \\ -\beta(1+\lambda) & \beta(\beta-1) & \beta(\beta+1) \end{pmatrix}$$

and

$$A_4 = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 2 \end{pmatrix}.$$

Now from $y \leq 20cx(c/t)^{1-\beta}$ we have the estimate

$$1000c^2x^2v^{-1} \geq 50xyv^{-1}c(c/t)^{\beta-1} = 50c^\beta xyw^{1-\beta}v^{-\beta}.$$

Obviously, A_2 is nonnegative-definite. Hence, from the above inequality and (2.4.10), we obtain

$$\xi''_{b_4-1000b_3}(0) \leq c^\beta xyw^{1-\beta}v^{-\beta}\left\langle (A_3-50A_4)(D,R,S),(D,R,S)\right\rangle.$$

It is enough to show that

$$\langle (A_3 - 50A_4)(D,R,S),(D,R,S)\rangle \le -\frac{1}{20}D^2 - \frac{1}{40}|D||R|. \tag{2.4.11}$$

Indeed, from the above inequalities and part (b) of Lemma 2.5 we have that

$$\begin{aligned}\xi''_{b_1+320(b_4-1000b_3)}(0) &\le 4we^2 + 8y|e||r| - 16c^\beta xyw^{1-\beta}v^{-\beta}D^2 - 8c^\beta xyw^{1-\beta}v^{-\beta}|D||R|\\ &\le 4wd^2 + 8y|d||r| - 16wd^2(c/t)^\beta(y/x) - 8y|d||r|(c/t)^\beta\\ &\le 0.\end{aligned}$$

To finish the proof of the lemma, observe that the estimate (2.4.11) is equivalent to nonpositive-definiteness of the matrix

$$\begin{pmatrix} 2\lambda - 100 + \frac{1}{20} & \frac14(1+\lambda) \pm \frac{1}{80} & -\frac34(1+\lambda)+100\\ \frac14(1+\lambda)\pm\frac{1}{80} & -\frac{3}{16} & -\frac{3}{16}\\ -\frac34(1+\lambda)+100 & -\frac{3}{16} & \frac{21}{16}-100\end{pmatrix},$$

for every $|\lambda| \le 1/10$, which we check by straightforward calculation (determinant of this matrix is convex as a function of λ, so it is sufficient to check only two endpoint cases $\lambda = 1/10$ and $\lambda = -1/10$). □

Finally, we prove the concavity condition in the domain D^3 in the last lemma.

Lemma 2.8. *We have* $\xi''_{b_1+32b_5-4600b_3-294400b_6}(0) \le 0$ *for any* $(x,y,w,v) \in D^3$ *and any* (d,e,r,s) *such that* $|e| \le |d|$.

Remark 2.6. This lemma handles the property 3° on the domain D^3. Indeed, the additional summands $-b_2$ and $-315400b_3$ are concave, so they do not affect the condition 3°.

Proof of Lemma 2.8. We use the same notation for relative changes D, E, R and S as in the proof of the previous lemma. We start with the analysis of the part b_3. We have that

$$\xi''_{b_3}(0) = \frac{2c^2}{v}\left(d - \frac{xs}{v}\right)^2 = \frac{2c^2x^2}{v}(D-S)^2 \ge \frac{2cx^2}{v}(D-S)^2 = \frac{2cx^2w}{t}(D-S)^2.$$

From the part (c) of Lemma 2.5 and the above estimate we obtain

$$\begin{aligned}\xi''_{b_3+64b_6}(0) &\ge 2\left(\frac{c}{t}x^2w(D^2-2DS+S^2) + 2cv^{-3}x^2s^2\right)\\ &= 2\left(\frac{c}{t}x^2w(D^2-2DS+S^2) + 2\left(\frac{c}{t}\right)x^2wS^2\right)\\ &= 2\left(\frac{c}{t}\right)x^2w(D^2-2DS+3S^2)\\ &\ge \left(\frac{c}{t}\right)x^2w(D^2+2S^2)\\ &\ge \left(\frac{c}{t}\right)y^2w\left(E^2 + \frac{2}{100}S^2\right),\end{aligned}$$

hence

$$\xi''_{16^{-1}\cdot 23\cdot 100(b_3+64b_6)}(0) \ge \left(\frac{c}{t}\right)y^2w16^{-1}[2300E^2 + 46S^2]. \tag{2.4.12}$$

Now we turn our attention to the analysis of the part b_5. We have that

$$\xi''_{b_5}(0) = \left(\frac{c}{t}\right)^{\beta} y^2 w \langle A_5(E,R,S),(E,R,S)\rangle,$$

where

$$A_5 = \begin{pmatrix} 2 & 2(1-\beta) & -2\beta \\ 2(1-\beta) & \beta(\beta-1) & \beta(\beta-1) \\ -2\beta & \beta(\beta-1) & \beta(\beta+1) \end{pmatrix}.$$

We check by straightforward calculation that

$$\begin{pmatrix} 2 & 2(1-\beta) & -2\beta \\ 2(1-\beta) & \beta(\beta-1) & \beta(\beta-1) \\ -2\beta & \beta(\beta-1) & \beta(\beta+1) \end{pmatrix} \leq 16^{-1} \begin{pmatrix} 192 & \pm 2 & 0 \\ \pm 2 & 0 & 0 \\ 0 & 0 & 46 \end{pmatrix}.$$

Hence

$$\xi''_{b_5}(0) \leq \left(\frac{c}{t}\right)^{\beta} y^2 w 16^{-1}(192E^2 - 4|E||R| + 46S^2)$$

and, from (2.4.12),

$$\begin{aligned} \xi''_{b_5 - 16^{-1}\cdot 23\cdot 100(b_3+64b_6)}(0) &\leq \left(\frac{c}{t}\right)^{\beta} y^2 w 16^{-1}(-2108E^2 - 4|E||R|) \\ &\leq y^2 w 16^{-1}(-2108E^2 - 4|E||R|). \end{aligned}$$

Now, from the above estimate an part (b) of Lemma 2.5, we have that

$$\xi''_{b_1+32b_5-4600b_3-294400b_6} \leq -4212we^2 \leq 0,$$

which concludes the proof.

□

We conclude by proving the maximal inequality formulated in the introductory section.

Proof of (2.1.3). By the extrapolation argument, it is enough to show the estimate for $p=2$ only. Fix $c>1$ and consider the function $\mathbb{B} : \{(x,y,z,w,v) \in \mathbb{R}^3 \times \mathcal{D}_c : x \leq z\} \to \mathbb{R}$ given by $\mathbb{B}(x,y,z,w,v) = B(x,y-z,w,v)$ This new object enjoys the properties listed in Remark 2.3 above (in 2°, it majorizes $\mathbb{G}(x,y,z,w,v) = G(x,y-z,w,v) = \kappa((y-z)^2 w - C^2c^2x^2v^{-1})$); in particular, we have $\mathbb{B}_z(x,y,y,w,v) = B_y(x,0,w,v) = 0$, by the symmetry of B. Hence, we obtain

$$\mathbb{E}(Y_t^* - Y_t)^2 W_t \leq C^2c^2 \mathbb{E}X_t^2 V_t^{-1}, \qquad t \geq 0,$$

for any X,Y such that Y is a stochastic integral of X and any pair (W,V) associated with an A_2 weight of characteristic not exceeding c. The function $(x,v) \mapsto x^2v^{-1}$ is convex on $\mathbb{R}\times(0,\infty)$, so the conditional Jensen inequality yields $\mathbb{E}X_t^2V_t^{-1} \leq \mathbb{E}X_\infty^2 V_\infty^{-1} = \mathbb{E}X_\infty^2 W_\infty$. Letting $t\to\infty$ and using Fatou's lemma, we get

$$||Y^* - Y||_{L^2(W)} \leq C[W]_{A_2}||X||_{L^2(W)}.$$

It remains to use the fact that the two-sided maximal function $|Y|^*$ satisfies $|Y|^* \leq |Y^*| + |(-Y)^*|$, which implies

$$|||Y|^*||_{L^2(W)} \leq ||Y^*||_{L^2(W)} + ||(-Y)^*||_{L^2(W)} \leq 2C[W]_{A_2}||X||_{L^2(W)}.$$

The proof is complete. □

2.5. Inequality for discrete-time martingales and regular filtrations

Now let us discuss the weighted L^2 inequality in the discrete-time setting. Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, filtered by $(\mathcal{F}_n)_{n\geq 0}$, a non-decreasing sequence of sub-σ-algebras of $\mathcal{F}$. We are interested in the estimate

$$||g||_{L^2(w)} \leq C[w]_{A_2}||f||_{L^2(w)}, \tag{2.5.1}$$

where f, g are martingales such that g is *differentially subordinate* to f. Can we prove this estimate with the use of the Bellman function B constructed above? It seems that the answer is negative. In the course of proof we would need to show that the process $(B(f_n, g_n, w_n, v_n))_{n\geq 0}$ is a supermartingale, basing on the concavity of B. However, the condition 3° is only local and is not sufficient to establish the supermartingale property for processes with jumps: this is due to non-convexity of the hyperbolic strip $\{(w, v) \in \mathbb{R}^2 : 1 \leq wv \leq c\}$. However, using geometric properties of this set, we can obtain (2.5.1) for the special case, where the underlying filtration is θ-regular for some $\theta \in (0, 1/2]$.

Concerning discrete-time concavity of B, we will prove the following lemma.

Lemma 2.9. *Consider an arbitrary line segment with endpoints $P = (x, y, w, v), Q = (x + d, y + e, w + r, v + s), |d| \leq |e|$, which is entirely contained in $\mathcal{D}_c$. For such a segment we have*

$$\begin{aligned} B(Q) \leq B(P) &+ B_x(P)(Q_x - P_x) + B_y(P)(Q_y - P_y) \\ &+ B_w(P)(Q_w - P_w) + B_v(P)(Q_v - P_v). \end{aligned} \tag{2.5.2}$$

Here if $P \in D^1 \cap D^2$, the partial derivatives are understood to appropriate derivatives of the function B_1; if $P \in D^2 \cap D^3$, the derivatives come from B_2.

Proof. The concavity condition 3° implies that the function $G(t) = B(x + td, y + te, w + tr, v + ts)$ is concave on $[0, 1]$. Consequently, we have $G(1) \leq G(0) + G'(0+)$, and this immediately gives (2.5.2). □

The following geometric lemma shows that regularity of filtration implies some control over jumps of martingale (w_n, v_n).

Lemma 2.10. *Assume that $c > 1$, $\alpha \in [\theta, 1 - \theta]$ and suppose that points P, Q and $R = \alpha P + (1 - \alpha)Q$ lie in $\mathcal{D}_c$. Then the whole line segment PQ is contained within $\mathcal{D}_{\theta^{-1}c}$.*

Proof. Using a simple geometrical argument, it is enough to consider the case when the points P and R lie on the curve $wv = c$ (the upper boundary of $\mathcal{D}_c$) and Q lies on the curve $wv = 1$ (lower boundary of $\mathcal{D}_c$). Then the line segment RQ is contained within $\mathcal{D}_c$, and hence also within $\mathcal{D}_{\theta^{-1}c}$, so it is enough to ensure that the segment RQ is contained in $\mathcal{D}_{\theta^{-1}c}$. Let $P = (P_w, P_v)$, (Q_w, Q_v) and $R = (R_w, R_v)$. We consider two cases. If $P_w < R_w$, then

$$P_v = \alpha^{-1}R_v - (\alpha^{-1} - 1)Q_v < \alpha^{-1}R_v \leq \theta^{-1}R_v,$$

so the segment PR is contained in the quadrant $\{(w, v) : w \leq R_w, v \leq \theta^{-1}R_v\}$. Consequently, PR lies below the hyperbola $wv = \theta^{-1}c$ passing through $(R_w, \theta^{-1}R_v)$. This proves the assertion in the case $P_w < R_w$. In the case $P_w \geq R_w$ the reasoning is similar. Indeed, we check easily that the line segment PR lies below the hyperbola $wv = \theta^{-1}c$ passing through $(\theta^{-1}R_w, R_v)$. □

Now we will exploit the properties of B to obtain the proof of weighted L^2 inequality for regular filtrations.

Theorem 2.11. *If the filtration is θ-regular, then we have the weighted L^2 bound*

$$||g||_{L^2(w)} \leq C\theta^{-1}[w]_{A_2}||f||_{L^2(w)}, \tag{2.5.3}$$

under the assumption that g is differentially subordinate to f. Here C is a universal constant.

Proof. Let w be an A_2 weight and set $c = \theta^{-1}[w]_{A_2}$. Furthermore, let f, g be two martingales such that g is differentially subordinate to f; we may assume that $||f||_{L^2(w)} < \infty$, since otherwise there is nothing to prove. Let $B = B_c$ be the function constructed in the previous section. For a fixed $n \geq 0$, we apply (2.5.2) with $x = f_n, y = g_n, w = w_n, v = v_n$ and $d = df_{n+1}, e = dg_{n+1}, r = dw_{n+1}$ and $s = dv_{n+1}$. The application is allowed: indeed, we have $|dg_{n+1}| \leq |df_{n+1}|$ (by differential subordination), furthermore, the points (x, y, w, v) and $(x + d, y + e, w + r, v + s)$ belong to $\mathcal{D}_{\theta c}$ and hence the interval which joins them is entirely contained in $\mathcal{D}_c$ by Lemma 2.10. To explain the latter statemant, fix an atom A of $\mathcal{F}_n$: the random variable (w_n, v_n) is constant on A, denote its value by P. Let $A_1, A_2, \ldots, A_k$ be the collection of all atoms of $\mathcal{F}_{n+1}$ into which the event A is split. For any fixed $j \in \{1, 2, \ldots, k\}$, the random variable (w_{n+1}, v_{n+1}) restricted to A_i belongs to the convex set $\{(w, v) \in (0, \infty)^2 : wv \geq 1\}$, and hence so does the average

$$R = (R_w, R_v) := \frac{1}{\mathbb{P}(\Omega \setminus A_j)} \int_{\Omega \setminus A_j} (w_{n+1}, v_{n+1}) \mathrm{d}\mathbb{P}.$$

We know that P and Q belong to $\mathcal{D}_{\theta c}$ and we need to show that PQ is entirely contained in $\mathcal{D}_c$. If $R \in \mathcal{D}_{\theta c}$, then this follows from Lemma 2.10; the appropriate bound on the number α appearing in the assumption is due to the regularity of the filtration. However, if $R \notin \mathcal{D}_{\theta c}$, then necessarily $R_w R_v > \theta c$ (we know that $R_w R_v < 1$ cannot hold) and hence the whole line segment PQ must be contained $\mathcal{D}_{\theta c}$: otherwise, we would have $P \notin \mathcal{D}_{\theta c}$, by convexity of the set $\{(w, v) \in (0, \infty)^2 : wv > \theta c\}$.

Thus, we have shown that the use of (2.5.2) is permitted, and as the result we get

$$B(f_{n+1}, g_{n+1}, w_{n+1}, v_{n+1}) \leq B(f_n, g_n, w_n, v_n) + I,$$

where $\mathbb{E}(I|\mathcal{F}_n) = 0$. Indeed, the term I is a linear combination of $df_{n+1}, dg_{n+1}, dw_{n+1}$ and dv_{n+1} with $\mathcal{F}_n$ coefficients. Hence, we obtain that the sequence $(\mathbb{E}B(f_n, g_n, w_n, v_n))_{n\geq 0}$ is nondecreasing, so the majorization and the initial condition give

$$\frac{1}{2}(\mathbb{E}g_n^2 w_n - 1109^2 c^2 \mathbb{E}f_n^2 v_n^{-1}) \leq \mathbb{E}B(f_n, g_n, w_n, v_n) \leq \mathbb{E}B(f_0, g_0, w_0, v_0) \leq 0.$$

The function $(x, v) \mapsto x^2 v^{-1}$ is convex, so $\mathbb{E}f_n^2 v_n^{-1} \leq \mathbb{E}f^2 v^{-1} = \mathbb{E}f^2 w$. Furthermore, we have $w_n = \mathbb{E}(w|\mathcal{F}_n)$, so the above inequality implies

$$\mathbb{E}g_n^2 w \leq 1109^2 c^2 \mathbb{E}f^2 w.$$

It remains to let $n \to \infty$ and apply Fatou's lemma to get the claim. □

Now we turn our attention to discrete version of maximal inequality (2.1.3). We have the following lemma.

Lemma 2.12. *Suppose that the filtration is θ-regular for some $\theta \in (0,1/2]$. Then for any martingales f, g such that g is differentially subordinate to f, we have the weighted L^2-bound*

$$||g^* - g||_{L^2(w)} \leq C\theta^{-1}[w]_{A_2}||f||_{L^2(w)}. \tag{2.5.4}$$

Here C is a universal constant.

Proof. The argumentation splits naturally into two parts

Step 1. An auxilliary function and its properties. Introduce the function $\mathbb{B} : \mathbb{R}^3 \times \mathcal{D}_c \to \mathbb{R}$ of five variables given by

$$\mathbb{B}(x,y,z,w,v) = B(x, y - (y \vee z), w, v).$$

Consider an arbitrary line segment with endpoints $(x,y,w,v), (x+d, y+e, w+r, v+s)$, which is entirely contained in $\mathbb{R}^2 \times \mathcal{D}_c$. We will prove that for such a segment, if $y \leq z$ and $P = (x,y,z,w,v)$ and $Q = (x+d, y+e, w+r, v+s)$, then

$$\begin{aligned} \mathbb{B}(Q) \leq \mathbb{B}(P) &+ \mathbb{B}_x(P)(Q_x - P_x) + \mathbb{B}_y(P)(Q_y - P_y) + \mathbb{B}_w(P)(Q_w - P_w) \\ &+ \mathbb{B}_v(P)(Q_v - P_v). \end{aligned} \tag{2.5.5}$$

(The problematic parts are handled with as in (2.5.2): if P is such that $(x, y-z, w, v) \in D^1 \cap D^2$, we compute the appropriate derivatives with the use of the formula for B on the subdomain D^1; etc.). To show the above inequality, consider the function $G(t) = \mathbb{B}(x+td, y+te, z, w+tr, v+ts)$; it is enough to prove that it is concave on the interval $[0,1]$, since this implies $G(1) \leq G(0) + G'(0+)$, which is precisely (2.5.5). If $e = 0$, concavity of G follows at once from 3°, so we may assume that $e \neq 0$. By the definition of U,

$$G(t) = \begin{cases} B(x+td, y+te-z, w+tr, v+ts) & \text{if } y+tk \leq z, \\ B(x+td, 0, w+tr, v+ts) & \text{if } y+tk \geq z \end{cases}$$

and directly from 3°, the function G is concave on each of the intervals $\{t \in [0,1] : y+te \leq z\}$ and $\{t \in [0,1] : y+te \geq z\}$. Furthermore, if $y + t_0 k = z$, then

$$G(t_0-) - G(t_0+) = B_y(x+t_0 d, 0, w+t_0 r, v+t_0 s)k = 0.$$

This establishes (2.5.5).

Step 2. Proof of (2.5.4). Suppose that w is an A_2 weight and let $c = \theta^{-1}[w]_{A_2}$. Let f, g be martingales such that g is differentially subordinate to f; again, we may assume that $||f||_{L^2(w)}$ is finite. Fix $n \geq 0$ and let us apply (2.5.5) with $x = f_n, y = g_n, z = g_n^*$ (the one sided maximal function), $w = w_n$, $v = v_n$ and $d = df_{n+1}$, $e = dg_{n+1}$, $r = dw_{n+1}$ and $s = dv_{n+1}$. Arguing as previously, we check that the application is allowed. Since $\mathbb{B}(x,y,z,w,v) = \mathbb{B}(x,y,y\vee z,w,v)$, we get

$$\begin{aligned} \mathbb{B}(f_{n+1}, g_{n+1}, g_{n+1}^*, w_{n+1}, v_{n+1}) &= \mathbb{B}(f_{n+1}, g_{n+1}, g_n^*, w_{n+1}, v_{n+1}) \\ &\leq \mathbb{B}(f_n, g_n, g_n^*, w_n, v_n) + I, \end{aligned}$$

where $\mathbb{E}(I|\mathcal{F}_n) = 0$. This yields that the sequence $(\mathbb{E}\mathbb{B}(f_n, g_n, g_n^*, w_n, v_n))_{n\geq 0}$ is nonincreasing, and hence by 2° and then 1°,

$$\begin{aligned} \frac{1}{2}\left(\mathbb{E}(g_n - g_n^*)^2 w_n - 1109^2c^2\right)\mathbb{E}f_n^2 v_n^{-1}) &\leq \mathbb{E}B(f_n, g_n - g_n^*, w_n, v_n) \\ &= \mathbb{E}\mathbb{B}(f_n, g_n, g_n^*, w_n, v_n) \\ &\leq \mathbb{E}\mathbb{B}(f_0, g_0, g_0^*, w_0, v_0) = \mathbb{E}B(f_0, 0, w_0, v_0) \leq 0. \end{aligned}$$

Repeating the arguments from the previous lemma, this implies the estimate $\mathbb{E}(g-g^*)^2 \leq 1109^2 c^2 \mathbb{E} f^2 w$, which is the claim. □

We are ready for the proof of the maximal inequality for discrete-time martingales.

Theorem 2.13. *Fix $\theta \in (0, 1/2]$. Let f, g be martingales adapted to a θ-regular filtration such that g is differentially subordinate to f. Then for any A_2 weight w we have*

$$|||g|^*||_{L^2(w)} \leq C\theta^{-1}[w]_{A_2}||f||_{L^2(w)} \tag{2.5.6}$$

for some universal constant C.

Proof. Fix a weight w and a pair (f, g) of martingales such that g is differentially subordinate to f. We have $|g|^* \leq |g - g_0|^* + |g_0| \leq (g - g_0)^* + (g_0 - g)^* + |g_0|$ and hence

$$|||g|^*||_{L^2(w)} \leq ||(g - g_0)^*||_{L^2(w)} + ||(g_0 - g)^*||_{L^2(w)} + ||g_0||_{L^2(w)}.$$

But the martingales $g^1 = g - g_0$ and $g^2 = g_0 - g$ are also differentially subordinate to f, so by Theorem 2.11 and (2.5.6),

$$||g^{1*}||_{L^2(w)} \leq ||g^1 - g^{1*}||_{L^2(w)} + ||g^1||_{L^2(w)} \leq 2218 \cdot \theta^{-1}[w]_{A_2}||f||_{L^2(w)}$$

and similarly for g^2. Furthermore, we have

$$||g_0||^2_{L^2(w)} \leq ||f_0||^2_{L^2(w)} = \mathbb{E}f_0^2 w_0 \leq [w]_{A_2}\mathbb{E}f_0^2 v_0^{-1} \leq \theta^{-1}[w]_{A_2}||f||^2_{L^2(w)}$$

(in the third passage we have used the estimate $w_0 v_0 \leq [w]_{A_2}$, in the fourth we exploited convexity of the function $(x, v) \to x^2 v^{-1}$ and the last is due to $v^{-1} = w$). Putting all the above facts together, we get $|||g|^*||_{L^2(w)} \leq 4437 \cdot \theta^{-1}[w]_{A_2}||f||_{L^2(w)}$, which is the desired assertion. □

2.6. Further discussion

In this section we will make some additional comments about the weighted L^p inequality. By the direct application of the Bellman function method we proved that

$$||Y||_{L^2} \leq C[W]_{A_2}||X||_{L^2(W)} \tag{2.6.1}$$

holds for continuous-path martingales X, Y such that $Y = H \cdot X$ for some predictable process H with values in $[-1, 1]$. Concerning discrete-time martingales, we showed that

$$||g||_{L^2} \leq C\theta^{-1}[W]_{A_2}||f||_{L^2(W)} \tag{2.6.2}$$

is satisfied for differentially subordinate martingales adapted to a θ-regular filtration. Both these results can be generalized. The estimate (2.6.1) holds for differentially subordinate martingales and (2.6.2) is valid for arbitrary filtrations (see [29]). On the other hand, we emphasize that the result in [29] does not contain our result (2.1.3) involving the maximal function. Combining the inequality (2.6.1) with the sharp weighted L^2 bound for the maximal function (Theorem 1.10) yields the maximal estimate (2.1.3), but with the worse quadratic dependence on the A_2 characteristic. Thus, the maximal estimate (2.1.3) is essentially stronger than (2.6.1). We also note that this sharp maximal result was obtained recently using a different technique called sparse domination (see [44] for the discrete-time setting and [30] for the continuous-time setting).

There is a natural question about an explicit Bellman function associated with (2.6.1) and (2.6.2) for these more general settings mentioned above. In other words, can we use the function B to prove the inequality for differentially subordinate continuous-time martingales? And similarly, can we relax the regularity assumption for the discrete-time setting? Naturally, these problems translate into the question whether B satisfies stronger concavity than that appearing in 3°. Recall that the function B satisfies the condition

$$\langle D^2B(x,y,w,v)(d,e,r,s),(d,e,r,s)\rangle \leq 0, \tag{2.6.3}$$

for any $(x,y,w,v)\in\mathbb{R}^2\times\mathcal{D}_c$ and any $(d,e,r,s)\in\mathbb{R}^4$ satisfying $|e|\leq|d|$. While this is sufficient to establish the inequality for stochastic integrals, in order to prove the estimate for differentially subordinate martingales, one would require the following stronger concavity: there is a nonnegative function A on $\mathbb{R}^2\times\mathcal{D}_c$ such that for any $(x,y,w,v)\in\mathbb{R}^2\times\mathcal{D}_c$ and any $(d,e,r,s)\in\mathbb{R}^4$ we have

$$\langle D^2B(x,y,w,v)(d,e,r,s),(d,e,r,s)\rangle \leq A(x,y,w,v)(e^2-d^2). \tag{2.6.4}$$

To see this, one repeats the argument based on Itô's formula presented above. The only problematic term is the integral $I_2=\int_{0+}^t D^2B(Z_u)\mathrm{d}[Z^c]_u$. To show that it is nonpositive, we again use an approximation by Riemann sums. Pick an arbitrary partition $0<s_0<s_1<\ldots<s_N=t$ and an integer $i\in\{0,\ldots,N-1\}$. Next, set $\Delta X=X_{s_{i+1}}-X_{s_i}$ and define ΔY, ΔW, ΔV^c similarly. By (2.6.4),

$$\langle D^2B(Z_{s_i})(\Delta X,\Delta Y,\Delta W,\Delta V^c),(\Delta X,\Delta Y,\Delta W,\Delta V^c)\rangle \leq A(Z_{s_i})(|\Delta Y|^2-|\Delta X|^2).$$

Summing over i and letting the diameter of the partition (s_i) go to zero, one obtains

$$I_2\leq\int_0^t A(Z_s)\mathrm{d}([Y]_s-[X]_s)\leq 0,$$

by the differential subordination and the fact that A is nonnegative. The remaining part of the proof goes along the same lines.

Of course, concavity (2.6.4) contains (2.6.3), but it is stronger, since it imposes an upper bound for the second derivative of B also when $|e|>|d|$. We do not know if the function B constructed in this chapter satisfies this stronger concavity which would yield the estimate (2.6.1) for differentially subordinate martingales.

Now we turn our attention to the estimate (2.6.2) for discrete-time martingales. As we have already mentioned, it is known from other proofs that the regularity assumption is not necessary here: the estimate is valid for arbitrary filtrations. The Bellman function method rests on the construction of a function satisfying appropriate majorizations and such that the process $(B(f_n,g_n,w_n,v_n))_{n\geq0}$ is a supermartingale. This last condition is equivalent to the following "full" concavity:

3°' For any $(x,y,u,v)\in\mathbb{R}^2\times\mathcal{D}_c$, any positive integer n and any sequences $(\alpha_j)_{j=1}^n$, $(e_j)_{j=1}^n$, $(d_j)_{j=1}^n$, $(r_j)_{j=1}^n$, $(s_j)_{j=1}^n$ satisfying

$$|e_j|\leq|d_j|,$$

$$\alpha_j\in(0,1)\qquad\text{and}\qquad\sum_{j=1}^n\alpha_j=1,$$

$$\sum_{j=1}^n \alpha_j d_j = \sum_{j=1}^n \alpha_j e_j = \sum_{j=1}^n \alpha_j r_j = \sum_{j=1}^n \alpha_j s_j = 0$$

and

$$(x + d_j, y + e_j, u + r_j, v + s_j) \in \mathcal{D}_c,$$

we have

$$B(x, y, u, v) \geq \sum_{j=1}^n \alpha_j B(x + d_j, y + e_j, u + r_j, v + s_j).$$

Indeed, from this we would immediately obtain the supermartingale property:

$$\begin{aligned}\mathbb{E}[B(f_n, g_n, w_n, v_n)|\mathcal{F}_{n-1}] &= \mathbb{E}[B(f_{n-1} + df_n, g_{n-1} + dg_n, w_{n-1} + dw_n, v_{n-1} + dv_n)|\mathcal{F}_{n-1}] \\ &\leq B(f_{n-1}, g_{n-1}, w_{n-1}, v_{n-1}).\end{aligned}$$

Unfortunately, the function B we have constructed above cannot satisfy this stronger concavity. In the proof it was crucial that the building block function $(x, w, v) \mapsto -x^2 w\psi(wv)$ is concave as a function of three variables. While it is true for local concavity, as we have checked in Lemma 2.5 (c), this is no longer the case for "full" concavity (see Remark 5.3 in Chapter 5).

Chapter 3

Weighted weak-type inequalities for stochastic integrals

3.1. Statement of the result, some historical comments

This chapter is the most technical part of the thesis. It is devoted to the weighted weak-type (p, p) inequalities for stochastic integrals in the range $1 < p < \infty$. Our starting point is the unweighted version of these estimates, due to Burkholder [19], Suh [81] and Wang [86].

Theorem 3.1. *Suppose that X and Y are martingales such that Y is a stochastic integral, with respect to X, of some predictable process H with values in $[-1, 1]$. Then for any $1 < p < \infty$ we have*

$$||Y||_{L^{p,\infty}} \leq C_p||X||_{L^p}, \tag{3.1.1}$$

where the optimal constant satisfies $C_p^p = 2/\Gamma(p+1)$ for $1 < p \leq 2$ and $C_p^p = p^{p-1}/2$ for remaining p.

We will study the following weighted extension of (3.1.1). We work in the same continuous-time setting as in the previous chapter, in particular, we assume that processes X, Y and $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$ have continuous paths while $(V_t)_{t\geq 0}$ is allowed to have jumps.

Theorem 3.2. *Fix $1 < p < \infty$. Let X, Y be martingales such that Y is a stochastic integral, with respect to X, of some predictable process H with values in $[-1, 1]$. Then for any A_p weight W we have*

$$\|Y\|_{L^{p,\infty}(W)} \leq C_p[W]_{A_p}\|X\|_{L^p(W)}, \tag{3.1.2}$$

for some constant C_p depending only on p. The linear dependence on the A_p characteristic of W is optimal: for any $\kappa < 1$ and any $K > 0$, there is a weight W, a real-valued martingale X and a predictable process H with values in $\{-1, 1\}$ such that the stochastic integral $Y = H \cdot X$ satisfies $\|Y\|_{L^{p,\infty}(W)} > K[W]_{A_p}^{\kappa}\|X\|_{L^p(W)}$.

A related result for bounded martingales was obtained by Osękowski in [64]: it was shown there that

$$\|Y\|_{L^{p,\infty}(W)} \leq C_p[W]_{A_p}\|X\|_{L^\infty(W)}\|W\|_{L^1} \tag{3.1.3}$$

and the linear dependence on the characteristic is optimal (in the above sense). Obviously, we have $\|X\|_{L^p(W)} \leq \|X\|_{L^\infty(W)}\|W\|_{L^1}$, so (3.1.2) is an extension of (3.1.3). Hence in particular, the optimality of the linear dependence in (3.1.2) follows at once from the analogous sharpness in (3.1.3).

So, all we need is the proof of (3.1.2). Moreover, it is enough to show this only in the range $1 < p < 2$, because the case $2 \leq p < \infty$ follows at once from strong-type bound

$$\|Y\|_{L^p(W)} \leq C_p[W]_{A_p}\|X\|_{L^p(W)}$$

established in the previous chapter. Two further comments are in order.

Remark 3.1. There is also another method of proving (3.1.2) in the range $p > 2$. Note that if (3.1.2) is true for some exponent p_0, with the linear dependence on the weight $W \in A_{p_0}$, then this estimate holds, with the linear dependence, for every $p > p_0$. Indeed, this is a consequence of the extrapolation Theorem 1.16 (see Corollary 1.15 about extrapolating weak-type estimates). Thus, it is enough to study the weak-type bound only for $p \in (1, 2)$.

Remark 3.2. The unweighted weak-type inequality (3.1.1) is valid also in the endpoint $p = 1$, and there is an interesting question about the sharp dependence on the characteristic $[W]_{A_1}$. In fact, in the light of the previous remark, if this bound held with the linear dependence, we would immediately get the estimate in the full range $1 \leq p < \infty$ via extrapolation. Quite unexpectedly, the weak-type (1,1) estimate is valid with an LlogL dependence on the characteristic: motivated by analogous statements [49, 54] in the analytic setting, Osękowski [65] proved that

$$||Y||_{L^{1,\infty}(W)} \leq K[W]_{A_1} \log(1 + [W]_{A_1})||X||_{L^1(W)}$$

and the logarithmic factor cannot be omitted. Therefore, the weak-type bound for the critical case $p = 1$ is essentially different than for the case $1 < p < \infty$. This boundary case will not be treated in this thesis.

3.2. Bellman function method for weighted weak-type inequality

The Bellman function method described in the previous chapter works perfectly in the context of weak-type estimates. Fix $1 < p < 2$ and $c \geq 1$. As in the previous chapter, we distinguish the hyperbolic strip

$$\mathcal{D}_c = \left\{(w, v) \in (0, \infty) \times (0, \infty) : 1 \leq wv^{p-1} \leq c\right\}.$$

We want to prove the weak-type inequality

$$\sup_{\lambda > 0} \lambda^p W(|Y| \geq \lambda) \leq C_p^p c^p \mathbb{E}|X|^p W \tag{3.2.1}$$

for any weight W such that $[W]_{A_p} \leq c$ and any martingales $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ such that $Y = H \cdot X$ is a stochastic integral, with respect to X, of some predictable process H with values in $[-1, 1]$. By homogeneity, we may and do assume that $\lambda = 1$; then (3.2.1) can be rewritten as

$$\mathbb{E}\left[W 1_{\{|Y|\geq 1\}} - C_p^p c^p |X|^p W\right] \leq 0. \tag{3.2.2}$$

By a simple limiting argument (which will be described later), it is enough to show that

$$\mathbb{E}\left[W_t 1_{\{|Y_t|\geq 1\}} - C_p^p c^p |X_t|^p V_t^{1-p}\right] \leq 0 \tag{3.2.3}$$

for all $t \geq 0$, where (W, V) is the martingale pair associated with W. Clearly, (3.2.3) is of the form (2.2.1), with

$$G(x, y, w, v) = w 1_{\{|y|\geq 1\}} - C_p^p c^p |x|^p v^{1-p}, \tag{3.2.4}$$

and hence the Bellman function approach can be exploited. Suppose that $\mathfrak{B} : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ is a C^2 function satisfying

1° (Initial condition) We have $\mathfrak{B}(x,y,w,v) \le 0$ if $|y| \le |x|$ and $1 \le wv^{p-1} \le c$.

2° (Majorization property) There is a positive constant C_p such that

$$\mathfrak{B}(x,y,w,v) \ge w\mathbb{1}_{\{|y|\ge 1\}} - C_p^p c^p |x|^p v^{1-p} \qquad \text{for } (x,y,w,v) \in \mathbb{R}^2 \times \mathcal{D}_c. \tag{3.2.5}$$

3° (Concavity-type property) For any $(x,y,w,v) \in \mathbb{R}^2 \times \mathcal{D}_c$ and $d,e,r,s \in \mathbb{R}$ satisfying $|e| \le |d|$, the function

$$\xi_{\mathfrak{B}}(t) := \mathfrak{B}(x+td, y+te, w+tr, v+ts),$$

given for those t, for which $1 \le (w+tr)(v+ts)^{p-1} \le c$, is locally concave.

The existence of such a special function yields the desired estimate (3.2.3). As previously, by Remark 2.1, the C^2 requirement can be relaxed: it is enough to assume that B is continuous.

3.3. Construction of a special function: some rough arguments

In this section we describe some of the informal steps which lead to the Bellman function associated with (3.2.3). As before, to gain some intuition, it is instructive to consider first the simpler unweighted setting. The corresponding endpoint weak-type estimate reads

$$||Y||_{L^{1,\infty}} \le 2||X||_{L^1},$$

for which the gain function is equal to $G(x,y) = \mathbb{1}_{\{|y|\ge 1\}} - 2|x|$. This sharp inequality was proved by Burkholder and the associated Bellman function is given by the formula (see [61])

$$b(x,y) = \begin{cases} y^2 - x^2 & \text{for } |x|+|y| \le 1, \\ 1-2|x| & \text{for } |x|+|y| > 1. \end{cases}$$

For some more or less formal arguments which lead to the discovery of this function, see e.g. [60, 61]. We see that b is built of two components: when the argument is close to the origin, then b coincides with the special function $y^2 - x^2$ corresponding to the L^2 estimate. For remaining arguments, the Bellman function is an affine expression (in $|x|$), which is almost equal to the gain function G. One easily checks 1° and 2°; to verify 3°, one rewrites the above formula as

$$b(x,y) = \begin{cases} \min\left\{y^2 - x^2, 1-2|x|\right\} & \text{if } |x| \le 1, \\ 1-2|x| & \text{if } |x| > 1, \end{cases} \tag{3.3.1}$$

from which it is clear that the concavity holds: both $(x,y) \mapsto y^2 - x^2$ and $(x,y) \mapsto 1-2|x|$ are concave along the lines of slope ± 1, and hence so is b, being essentially the minimum of the two.

We will try to construct the Bellman function $\mathfrak{B}$ associated with (3.2.3) using the observations above. Namely, $\mathfrak{B}$ will be defined (essentially) by two different formulas: the formula B corresponding to (x,y) close to zero and another, $\bar{B}$, for remaining (x,y). In the light of the above discussion, guessing the correct $\bar{B}$ is not difficult: a look at the formula (3.2.4) suggests the choice $\bar{B}(x,y,w,v) = w - Kc^p|x|^p v^{1-p}$, for some constant K.

The main technical difficulty lies in the discovery of the formula B. A natural idea is to take the function developed in the previous chapter, or rather its certain modification. In a sense, this idea turns out to work: we have managed to find the appropriate formula after many lengthy calculations and elaborate experimentation. We will present these computations in the next section.

3.4. A special function

Throughout this section, $c > 1$ and $p \in (1, 2)$ are fixed parameters. The main result of this section is the following.

Theorem 3.3. *There is a continuous function* $\mathfrak{B} : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ *satisfying* $1° - 3°$.

Similarly to the previous chapter, we start by defining auxiliary parameters and building block functions. Let us denote

$$\alpha_p = 36\sqrt{\frac{2}{(p-1)(3-p)}}, \qquad \beta = \frac{p+1}{2p}.$$

Define the functions $b_i : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ by

$$\begin{aligned}
b_1(x, y, w, v) &= 3|y|^3 w^{1-1/6c} v^{-(p-1)/6c} - |y|^3 v^{1-p},\\
b_2(x, y, w, v) &= |y|^3 v^{1-p},\\
b_3(x, y, w, v) &= \alpha_p^p c^p |x|^p v^{1-p},\\
u(x, y, w, v) &= \min\left\{c^\beta |x| y^2 w^{1-\beta} v^{-\beta(p-1)} - \alpha_p^{p-1} c^p |x|^p v^{1-p}, 0\right\},\\
b_4(x, y, w, v) &= u(x, y, w, v) - (\alpha_p^p - \alpha_p^{p-1}) c^p |x|^p v^{1-p},\\
b_5(x, y, w, v) &= c|x|^3 v^{1-p},\\
b_6(x, y, w, v) &= cx^2 v^{1-p}.
\end{aligned}$$

Let $B : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ be given by

$$B = b_1 - b_2 - b_3 + \frac{216(p+1)}{p-1}(b_4 - 1882 b_5 - 9 b_6).$$

We will show later that the above function has appropriate concavity and satisfies the initial condition 1°. This will be one component of the Bellman function $\mathfrak{B}$, corresponding to the points near the origin. (It cannot be used on the whole domain, since the majorization 2° fails when (x, y) are far from the origin; see the previous section for a similar phenomenon in the unweighted case). As we have noted above, the second component $\bar{B} : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ is given by

$$\bar{B}(x, y, w, v) = w - Kc^p |x|^p v^{1-p},$$

where the multiplicative constant equals

$$K = 2 \cdot 10^6 \cdot \frac{\alpha_p^p}{p-1}.$$

It is easy to check that the function $\bar{B}$ enjoys the required concavity and satisfies the majorization condition with $C_p = K^{1/p}$. (However, now the initial condition breaks near the origin, so this function does not work on the whole domain either). Now we will splice the two components so that the resulting Bellman function $\mathfrak{B}$ is continuous. First, let us describe the boundary between these two components. Given $(w, v) \in \mathcal{D}_c$, let

$$y_0 = y_0(w, v) = \inf\{y > 0 : B(0, y, w, v) \geq \bar{B}(0, y, w, v)\} = \left(\frac{t}{3t^{1-1/6c} - 2}\right)^{1/3},$$

where $t = wv^{p-1}$. Observe that $y_0 \leq 1$, which follows from $1 \leq t \leq c$ and a direct differentiation. Next, define the function $\xi : [0, y_0] \to \mathbb{R}$ by the following formula

$$\xi(y) = \inf\{x > 0 : B(x, y, w, v) \geq \bar{B}(x, y, w, v)\}.$$

Let us show that ξ is well defined.

Lemma 3.4. *For every $y \in [0, y_0]$ we have $0 \leq \xi(y) \leq 1$.*

Proof. It is sufficient to show that for every $y \in [0, y_0]$ we have

$$B(1, y, w, v) \geq \bar{B}(1, y, w, v).$$

This follows from the estimate

$$\begin{aligned} B(1, y, w, v) &\geq -2y^3v^{1-p} - \alpha_p^p c^p v^{1-p} - \frac{216(p+1)}{p-1}\alpha_p^p c^p v^{1-p} \\ &\quad - \frac{216 \cdot 1882(p+1)}{p-1} cv^{1-p} - \frac{216 \cdot 9(p+1)}{p-1} cv^{1-p} \\ &\geq w - Kc^p v^{1-p}. \end{aligned}$$

Here in the last inequality is important that we have chosen K large enough. More precisely,

$$K \geq 1 + 2 + \alpha_p^p + \frac{216(p+1)}{p-1}\alpha_p^p + \frac{216 \cdot 1882(p+1)}{p-1} + \frac{216 \cdot 9(p+1)}{p-1}. \quad \square$$

Next, define the auxiliary sets

$$D = \left\{(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c : 0 \leq |y| \leq y_0,\ 0 \leq |x| \leq \xi(|y|)\right\}, \qquad \bar{D} = \mathbb{R}^2 \times \mathcal{D}_c \setminus D.$$

The Bellman function $\mathfrak{B} : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ is given by

$$\mathfrak{B}(x, y, w, v) = \begin{cases} B(x, y, w, v) & \text{for } (x, y, w, v) \in D, \\ \bar{B}(x, y, w, v) & \text{for } (x, y, w, v) \in \bar{D}. \end{cases}$$

It is immediate from this construction that $\mathfrak{B}$ is continuous. Now we will check that the function $\mathfrak{B}$ satisfies required conditions. Before we proceed, let us make an important observation. From Lemma 3.4 we see that D is contained in the square region $[-1, 1]^2 \times \mathcal{D}_c$. The initial and majorization properties are relatively straightforward to prove, the main difficulty lies in establishing the concavity condition. We start with the easy part.

Theorem 3.5. *The function $\mathfrak{B}$ satisfies the properties 1° and 2°.*

Proof. Because of the symmetry, we may assume that x and y are nonnegative. Let us consider the initial condition 1°. Assume that $0 \leq y \leq x$. When $(x, y, w, v) \in D$, we use the estimate $b_4 \leq 0$ which gives

$$\mathfrak{B}(x, y, w, v) \leq 3x^3 w(wv^{p-1})^{-1/6c} - \frac{216 \cdot 1882(p+1)}{p-1} cx^3 w(wv^{p-1})^{-1} \leq 0.$$

On the other hand, when $(x, y, w, v) \in \bar{D}$, we obtain

$$\mathfrak{B}(x, y, w, v) = \bar{B}(x, x, w, v) = w - Kc^p x^p v^{1-p} \leq w - Kc^p x_0^p v^{1-p}, \tag{3.4.1}$$

where x_0 is defined as follows

$$x_0 = \inf\{x > 0 : B(x, x, w, v) \geq \bar{B}(x, x, w, v)\}.$$

We have that

$$w - Kc^p x_0^p v^{1-p} = \bar{B}(x_0, x_0, w, v) = B(x_0, x_0, w, v) \leq 0,$$

because $(x_0, x_0, w, v) \in D$ and we have already checked the initial condition for the function B. This concludes the proof of the initial condition 1°. Now we consider the majorization condition 2°. We need to show that

$$\mathfrak{B}(x, y, w, v) \geq w\mathbb{1}_{\{|y| \geq 1\}} - C_p^p c^p |x|^p v^{1-p} \tag{3.4.2}$$

for $C_p = K^{1/p}$. If $(x, y, w, v) \in D$, then we have

$$\begin{aligned}
&\mathfrak{B}(x, y, w, v) \\
&\geq B(x, 0, w, v) \\
&= \left(-\alpha_p^p - \frac{216(p+1)}{p-1}\alpha_p^p\right) c^p x^p v^{1-p} - \frac{216 \cdot 1882(p+1)}{p-1} cx^3 v^{1-p} - \frac{216 \cdot 9(p+1)}{p-1} cx^2 v^{1-p} \\
&\geq -C_p^p c^p |x|^p v^{1-p},
\end{aligned}$$

where the last inequality follows from $x \leq 1$ (this a consequence of Lemma 3.4) and the estimate

$$C_p^p = K \geq \alpha_p^p + \frac{216(p+1)}{p-1}\alpha_p^p + \frac{216 \cdot 1882(p+1)}{p-1} + \frac{216 \cdot 9(p+1)}{p-1}.$$

On the other hand, if $(x, y, w, v) \in \bar{D}$, we may write

$$\mathfrak{B}(x, y, w, v) = \bar{B}(x, y, w, v) = w - Kc^p x^p v^{1-p} = w - C_p^p c^p x^p v^{1-p},$$

so (3.4.2) is evident. This completes the proof of the theorem. □

We turn our attention to the crucial condition 3°. By symmetry (and the equality $\mathfrak{B}_x(0, y, w, v) = \mathfrak{B}_y(x, 0, w, v) = 0$ for all x, y, w, v), we may consider only those points $(x, y, w, v) \in \mathbb{R}^2 \times \mathcal{D}_c$, for which x and y are nonnegative. Concavity of the function $\bar{B}$ is immediate: it is easy to check that the function $(x, v) \mapsto -|x|^p v^{1-p}$ is concave (actually, we have used this several times in the previous chapter). The verification of the concavity of B is much more challenging. We must show that for each $(x, y, w, v) \in D$ and any $d, e, r, s \in \mathbb{R}$ satisfying $|e| \leq |d|$, the function

$$\xi_B(t) := B(x + td, y + te, w + tr, v + ts),$$

given for those t, for which $(x+td, y+te, w+tr, v+ts) \in D$, satisfies $\xi_B''(0) \le 0$. This localized concavity will be accomplished by a careful analysis of the derivatives $\xi_{b_j}''(0)$ of the building blocks $b_j, j = 1, 2, \ldots, 6$. Again, we use the notation $D = d/x, E = e/y, R = r/w, S = s/v$ for the relative changes (we hope that the slight abuse of notation - the letter D has two meanings - will not lead to a confusion). For convenience, we will use the following shortened notation for Hessian matrices and quadratic forms associated with them. Namely, for any function $b : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$, the symbol $\xi_b(D, E, R, S)$ denotes corresponding Hessian matrix of b written in terms of D, E, R, S, that is

$$\xi_b(D,E,R,S) = \begin{pmatrix} \frac{\partial^2 b}{\partial x^2}x^2 & \frac{\partial^2 b}{\partial x \partial y}xy & \cdots \\ \frac{\partial^2 b}{\partial x \partial y}xy & \ddots & \\ \vdots & & \end{pmatrix} = \begin{pmatrix} x \\ y \\ w \\ v \end{pmatrix} \cdot D^2 b(x,y,w,v) \cdot \begin{pmatrix} x \\ y \\ w \\ v \end{pmatrix}^T .$$

Then we have the equality of quadratic forms

$$\big\langle \xi_b(D,E,R,S) \cdot (D,E,R,S), (D,E,R,S) \big\rangle = \big\langle D^2 b(x,y,w,v) \cdot (d,e,r,s), (d,e,r,s) \big\rangle .$$

Moreover, we will often identify the matrix $\xi_b(D, E, R, S)$ with the associated quadratic form $\langle \xi_b(D, E, R, S) \cdot (D, E, R, S), (D, E, R, S)\rangle$. Sometimes we will also use the above notation with a reduced set of variables. For example, we will write

$$\xi_b(D,R) = \begin{pmatrix} x \\ w \end{pmatrix} \cdot D^2_{x,w} b(x,y,w,v) \cdot \begin{pmatrix} x \\ w \end{pmatrix}^T .$$

where $D^2_{x,w} b$ is the Hessian of b considered as a function of variables x and w only. The above notation might seem a little unfriendly, but it enables to write all the computations in a relatively concise form.

We start with the following technical lemma.

Lemma 3.6. *The matrix*

$$A = \begin{bmatrix} -\alpha_p p(p-1) & 1-\beta & (p-1)(p\alpha_p - \beta) \\ 1-\beta & \beta(\beta-1) & (p-1)\beta(\beta-1) \\ (p-1)(p\alpha_p-\beta) & (p-1)\beta(\beta-1) & (p-1)(\beta(\beta(p-1)+1) - p\alpha_p) \end{bmatrix}$$

is less or equal to any of the matrices

$$-\frac{p-1}{4(p+1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad -\frac{p-1}{4(p+1)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad -\frac{p-1}{4(p+1)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In particular, A is less or equal to any convex combination of these matrices.

Proof. The inequality

$$A \le -\frac{p-1}{4(p+1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is equivalent to saying that the matrix

$$\begin{bmatrix} \frac{p-1}{4(p+1)} - \alpha_p p(p-1) & 1-\beta & (p-1)(p\alpha_p - \beta) \\ 1-\beta & \beta(\beta-1) & (p-1)\beta(\beta-1) \\ (p-1)(p\alpha_p-\beta) & (p-1)\beta(\beta-1) & (p-1)(\beta(\beta(p-1)+1) - p\alpha_p) \end{bmatrix}$$

is nonpositive definite. Since $\alpha_p \geq 36$, the entry in the upper-left corner is negative. Furthermore, since $p-1 = 2p(1-\beta)$, we compute that

$$\det \begin{bmatrix} \frac{p-1}{4(p+1)} - \alpha_p p(p-1) & 1-\beta \\ 1-\beta & \beta(\beta-1) \end{bmatrix} = (1-\beta)^2 \left[p(p+1)\alpha_p - 5/4\right] \geq 0$$

and hence it is enough to prove that the determinant of the full matrix is nonpositive. Let us perform some elementary operations on the rows and columns, to make the determinant easier to compute. First add, to the third column, the second column multiplied by $1-p$; then add, to the third row, the second row multiplied by $1-p$. The obtained matrix

$$\begin{bmatrix} \frac{p-1}{4(p+1)} - \alpha_p p(p-1) & 1-\beta & (p-1)(p\alpha_p - 1) \\ 1-\beta & \beta(\beta-1) & 0 \\ (p-1)(p\alpha_p-1) & 0 & p(p-1)(\beta-\alpha_p) \end{bmatrix}$$

has the same determinant, equal to

$$(p-1)^2(1-\beta)\left[\frac{\alpha_p}{4}\left(p^2 - 2p - \frac{1}{2}\right) + \frac{3}{16}\frac{p+1}{p}\right] < 0,$$

since $p^2 - 2p - 1/2 \leq -1/2$ and $\alpha_p \geq 36$. This shows the first inequality of the lemma. The remaining two estimates are proved in the same manner; actually, by the above analysis we have $A \leq 0$, and hence it is enough to show that the determinants of the (full) matrices

$$\begin{bmatrix} -\alpha_p p(p-1) & 1-\beta & (p-1)(p\alpha_p - \beta) \\ 1-\beta & \frac{p-1}{2(p+1)} + \beta(\beta-1) & (p-1)\beta(\beta-1) \\ (p-1)(p\alpha_p - \beta) & (p-1)\beta(\beta-1) & (p-1)(\beta(\beta(p-1)+1) - p\alpha_p) \end{bmatrix}$$

and

$$\begin{bmatrix} -\alpha_p p(p-1) & 1-\beta & (p-1)(p\alpha_p - \beta) \\ 1-\beta & \beta(\beta-1) & (p-1)\beta(\beta-1) \\ (p-1)(p\alpha_p - \beta) & (p-1)\beta(\beta-1) & \frac{p-1}{4(p+1)} + (p-1)(\beta(\beta(p-1)+1) - p\alpha_p) \end{bmatrix}$$

are nonpositive. The first matrix has determinant zero: the sum of the first and the third row is proportional to the second row. Concerning the second matrix, we compute that its determinant is equal to

$$-\beta(1-\beta)(p-1)^2\left[\frac{p\alpha_p}{2(p+1)}\left(2p + \frac{1}{2} - p^2\right) + \frac{1}{4(p+1)^2} - \frac{1}{2}\right] < 0,$$

since

$$\frac{p\alpha_p}{2(p+1)}\left(2p + \frac{1}{2} - p^2\right) \geq \frac{\alpha_p}{6} \cdot \frac{1}{2} > 3.$$

The proof is complete. □

The next lemma concerns the function $b_1(x, y, w, v) = 3|y|^3 w^{1-1/6c} v^{-(p-1)/6c} - |y|^3 v^{1-p}$.

Lemma 3.7. *We have*

$$\xi_{b_1}(E, R, S) \leq 3y^3 w^{1-1/6c} v^{-(p-1)/6c} \begin{bmatrix} 6 & 3(1-\frac{1}{6c}) & 0 \\ 3(1-\frac{1}{6c}) & -\frac{1}{12c}(1-\frac{1}{6c}) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.4.3}$$

Proof. We compute that

$$\xi_{b_1}(E,R,S) = 3y^3w^{1-1/6c}v^{-(p-1)/6c}\begin{bmatrix} 6 & 3(1-\frac{1}{6c}) & \frac{1-p}{2c} \\ 3(1-\frac{1}{6c}) & -\frac{1}{6c}(1-\frac{1}{6c}) & \frac{1-p}{6c}(1-\frac{1}{6c}) \\ \frac{1-p}{2c} & \frac{1-p}{6c}(1-\frac{1}{6c}) & \frac{p-1}{6c}(1+\frac{p-1}{6c}) \end{bmatrix}$$
$$-y^3v^{1-p}\begin{bmatrix} 6 & 0 & 3(1-p) \\ 0 & 0 & 0 \\ 3(1-p) & 0 & p(p-1) \end{bmatrix}.$$

Since $v^{1-p} \geq w^{1-1/6c}v^{-(p-1)/6c}/c$ and the last matrix above is nonnegative-definite, we see that $\xi_{b_1}(E,R,S)$ is not bigger than

$$3y^3w^{1-1/6c}v^{-(p-1)/6c}\begin{bmatrix} 6-\frac{2}{c} & 3(1-\frac{1}{6c}) & \frac{p-1}{2c} \\ 3(1-\frac{1}{6c}) & -\frac{1}{6c}(1-\frac{1}{6c}) & \frac{1-p}{6c}(1-\frac{1}{6c}) \\ \frac{p-1}{2c} & \frac{1-p}{6c}(1-\frac{1}{6c}) & \frac{p-1}{6c}(1+\frac{p-1}{6c}) - \frac{p(p-1)}{3c} \end{bmatrix}. \tag{3.4.4}$$

It remains to bound appropriately the latter matrix. Observe that

$$\frac{p-1}{6c}\left(1+\frac{p-1}{6c}\right) - \frac{p(p-1)}{3c} \leq -\frac{17}{6}\cdot\frac{(p-1)^2}{6c} \leq -\frac{(p-1)^2}{8c} - \frac{(p-1)^2}{3c}.$$

Indeed, the second inequality reduces to the obvious bound $\frac{17}{36} \geq \frac{1}{8}+\frac{1}{3}$, while the first is equivalent to

$$1+\frac{p-1}{6c} \leq 2p - \frac{17}{6}(p-1),$$

which holds true because of $c \geq 1$ and $p \leq 2$. Consequently, the matrix in (3.4.4) is not bigger than $I_1+I_2+I_3$, where

$$I_1 = \begin{bmatrix} -\frac{2}{c} & 0 & \frac{p-1}{2c} \\ 0 & 0 & 0 \\ \frac{p-1}{2c} & 0 & -\frac{(p-1)^2}{8c} \end{bmatrix}, \quad I_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{12c}(1-\frac{1}{6c}) & \frac{1-p}{6c}(1-\frac{1}{6c}) \\ 0 & \frac{1-p}{6c}(1-\frac{1}{6c}) & -\frac{(p-1)^2}{3c} \end{bmatrix}$$

and

$$I_3 = \begin{bmatrix} 6 & 3(1-\frac{1}{6c}) & 0 \\ 3(1-\frac{1}{6c}) & -\frac{1}{12c}(1-\frac{1}{6c}) & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

It remains to note that I_1, $I_2 \leq 0$. The claim follows. □

Remark 3.3. Observe that the matrix

$$\begin{bmatrix} 6 & 3(1-\frac{1}{6c}) & 0 \\ 3(1-\frac{1}{6c}) & -\frac{1}{12c}(1-\frac{1}{6c}) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

appearing in (3.4.3) is less than any of the matrices

$$J_1 = \begin{bmatrix} 108c & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad J_2 = \begin{bmatrix} 12 & 3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Roughly speaking, our further construction will ensure that the quadratic form ξ of the remaining part of the Bellman function will be bigger than J_1 on some part of the domain and bigger than J_2 on the remaining part of the domain.

Now we turn our attention to the functions $b_2(x,y,w,v)=|y|^3v^{1-p}$ and $b_3(x,y,w,v)=\alpha_p^pc^p|x|^pv^{1-p}$.

Lemma 3.8. *The functions b_2 and b_3 are convex. Furthermore, if $|y|\geq\alpha_pc|x|$, then*

$$\xi_{b_2+b_3}(E,S)\geq 3y^3w^{1-1/6c}v^{-(p-1)/6c}\begin{bmatrix}108c & 0\\ 0 & 0\end{bmatrix}. \tag{3.4.5}$$

Proof. Convexity of b_2 and b_3 is evident: the corresponding Hessian matrices are nonnegative-definite. To show (3.4.5), note that

$$\begin{aligned}\xi_{b_2}(E,S)=|y|^3v^{1-p}\begin{bmatrix}6 & 3(1-p)\\ 3(1-p) & p(p-1)\end{bmatrix}&\geq|y|^3v^{1-p}\begin{bmatrix}0 & 0\\ 0 & \frac{(p-1)(3-p)}{2}\end{bmatrix}\\ &=|y|^3v^{1-p}\cdot\frac{(p-1)(3-p)}{2}|S|^2\end{aligned}$$

and

$$\xi_{b_3}(D,S)=p(p-1)\alpha_p^pc^p|x|^pv^{1-p}(D-S)^2.$$

Now, if $|D-S|\leq|D|/2$, then

$$|S|\geq|D|/2=\frac{|d|}{2|x|}\geq|E|\cdot\left|\frac{y}{x}\right|\cdot\frac{1}{2}\geq|E|\cdot\frac{\alpha_pc}{2}$$

and hence

$$\begin{aligned}\xi_{b_2+b_3}(D,E,S)\geq\xi_{b_2}(E,S)&\geq|y|^3v^{1-p}\cdot\frac{(p-1)(3-p)}{2}|S|^2\\ &\geq 3|y|^3v^{1-p}\cdot 108c^2E^2\\ &\geq 3|y|^3w^{1-1/6c}v^{-(p-1)/6c}\cdot 108cE^2.\end{aligned}$$

Here in the last passage we have used the estimate $cv^{1-p}\geq w\geq w^{1-1/6c}v^{-(p-1)/6c}$. This establishes (3.4.5). On the other hand, if $|D-S|\geq|D|/2$, then

$$\begin{aligned}\xi_{b_2+b_3}(D,E,S)\geq\xi_{b_3}(D,S)&\geq p(p-1)\alpha_p^pc^p|x|^pv^{1-p}D^2/4\\ &=p(p-1)\frac{\alpha_p^pc^2v^{1-p}}{4}\cdot(c|x|)^{p-2}d^2.\end{aligned}$$

But $\alpha_pc|x|\leq|y|$ and $|e|\leq|d|$, so the latter expression is not smaller than

$$p(p-1)\frac{\alpha_p^2c^2v^{1-p}}{4}\cdot y^{p-2}e^2=p(p-1)\frac{\alpha_p^2c^2v^{1-p}}{4}\cdot y^pE^2\geq p(p-1)\frac{\alpha_p^2c^2v^{1-p}}{4}y^3E^2.$$

But $p(p-1)\alpha_p^2/4\geq 36^2/4=3\cdot 108$ and, as previously, we have used the estimate $cv^{1-p}\geq w\geq w^{1-1/6c}v^{-(p-1)/6c}$, so we finally obtain

$$\xi_{b_2+b_3}(D,E,S)\geq 3|y|^3w^{1-1/6c}v^{-(p-1)/6c}\cdot 108cE^2.$$

□

The analysis of the next function $b_4(x,y,w,v)=u(x,y,w,v)-(\alpha_p^p-\alpha_p^{p-1})c^p|x|^pv^{1-p}$ will be a little more complicated. Consider the region

$$D_1=\left\{(x,y,w,v)\,:\,|y|\leq 1\ \text{ and }\ y^2\leq\left(\frac{c}{wv^{p-1}}\right)^{1-\beta}(\alpha_pc|x|)^{p-1}\right\}.$$

Since $c \geq wv^{p-1}$, the "remaining problematic region"

$$\{(x,y,w,v) : |y| \leq 1 \text{ and } |y| \leq \alpha_p c|x|\}$$

is contained in D_1. Note that the function

$$u(x,y,w,v) = \min\left\{c^\beta |x| y^2 w^{1-\beta} v^{-\beta(p-1)} - \alpha_p^{p-1} c^p |x|^p v^{1-p}, 0\right\}$$

vanishes outside D_1 and is negative in the interior of D_1.

Lemma 3.9. *Assume that $(x,y,w,v) \in D_1$. If $|y| \geq \frac{60(p+1)}{p-1}|x|$, then*

$$\xi_{b_4}(D,E,R,S) \leq -\frac{p-1}{4(p+1)}|y|^3 w^{1-1/6c} v^{-(p-1)/6c}\left[3E^2 + \frac{1}{3}|ER|\right]. \tag{3.4.6}$$

If $|y| < \frac{60(p+1)}{p-1}|x|$, then

$$\begin{aligned}&\xi_{b_4}(D,E,R,S)\\ &\leq c^\beta w^{1-\beta} v^{-\beta(p-1)}\left\{\left(82|x|^3 + 4x^2|y|\right)D^2 - \frac{p-1}{12(p+1)}|y|^3|ER|\right\}.\end{aligned} \tag{3.4.7}$$

Proof. We have

$$\xi_{b_4}(D,E,R,S) = \xi_{b_4}(D,R,S) + \left(\xi_{b_4}(D,E,R,S) - \xi_{b_4}(D,R,S)\right),$$

and $\xi_{b_4}(D,R,S)$ is given by

$$\begin{aligned}&c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)}\begin{bmatrix} 0 & 1-\beta & \beta(1-p) \\ 1-\beta & \beta(\beta-1) & (p-1)\beta(\beta-1) \\ \beta(1-p) & (p-1)\beta(\beta-1) & \beta(p-1)(\beta(p-1)+1)\end{bmatrix}\\ &-\alpha_p^p c^p x^p v^{1-p}\begin{bmatrix} p(p-1) & 0 & p(1-p) \\ 0 & 0 & 0 \\ p(1-p) & 0 & p(p-1)\end{bmatrix}.\end{aligned}$$

On D_1 we have $\alpha_p^p c^p x^p v^{1-p} \geq \alpha_p \cdot c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)}$. Therefore, since the latter matrix is nonnegative-definite, we see that $\xi_{b_4}(D,R,S)$ is not bigger than $c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)} A$, where A is the matrix from Lemma 3.6. By this lemma, we have

$$\xi_{b_4}(D,R,S) \leq -\frac{p-1}{4(p+1)} c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)}\left[\frac{8}{10}D^2 + \frac{1}{10}\cdot 2R^2 + \frac{1}{10}S^2\right] \tag{3.4.8}$$

(the right-hand side is the convex combination, with weights 8/10, 1/10 and 1/10, of the three majorizing matrices from Lemma 3.6). Similarly,

$$\xi_{b_4}(D,R,S) \leq -\frac{p-1}{4(p+1)} c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)}\left[\frac{1}{6}D^2 + \frac{1}{3}\cdot 2R^2 + \frac{1}{2}S^2\right]. \tag{3.4.9}$$

Let us now analyze the remaining part

$$\begin{aligned}&\xi_{b_4}(D,E,R,S) - \xi_{b_4}(D,R,S)\\ &\quad = c^\beta x y^2 w^{1-\beta} v^{-\beta(p-1)}\left[2E^2 + 4(1-\beta)ER - 4\beta(p-1)ES + 4ED\right].\end{aligned} \tag{3.4.10}$$

We consider two cases, corresponding to (3.4.6) and (3.4.7). For convenience, let us denote $\lambda = \frac{60(p+1)}{p-1}$. If $|y| \geq \lambda|x|$, then $|E| = |e/y| \leq \lambda^{-1}|d/x| = \lambda^{-1}|D|$ and hence the expression above is not bigger than

$$\lambda^{-1}c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\big[6D^2 + 4(1-\beta)|DR| + 4\beta(p-1)|DS|\big].$$

Since $2|DR| \leq D^2 + R^2$, $2|DS| \leq D^2 + S^2$, $2(1-\beta) \leq 1/2$ and $2\beta(p-1) \leq 3/2$, we get

$$\xi_{b_4}(D,E,R,S) - \xi_{b_4}(D,R,S) \leq \frac{p-1}{60(p+1)}c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\big[8D^2 + R^2/2 + 3S^2/2\big].$$

Combining this with (3.4.8), we see that

$$\begin{aligned}
\xi_{b_4}(D,E,R,S) &\leq -\frac{p-1}{4(p+1)}c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\left[\frac{8}{30}D^2 + \frac{1}{6}R^2\right] \\
&= -\frac{p-1}{4(p+1)}c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\left[\frac{1}{10}D^2 + \frac{1}{6}(D^2+R^2)\right] \\
&\leq -\frac{p-1}{4(p+1)}c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\left[\frac{1}{10}D^2 + \frac{1}{3}|DR|\right] \\
&\leq -\frac{p-1}{4(p+1)}\left(\frac{c}{wv^{p-1}}\right)^\beta y^3w\left[3E^2 + \frac{1}{3}|ER|\right],
\end{aligned}$$

where in the last line we have used the inequality $|x|D^2/|y| \geq |DE| \geq 60E^2$. Since $1 \leq wv^{p-1} \leq c$, we get

$$\xi_{b_4}(D,E,R,S) \leq -\frac{p-1}{4(p+1)}y^3w^{1-1/6c}v^{-(p-1)/6c}\left[3E^2 + \frac{1}{3}|ER|\right],$$

which is (3.4.6). Now suppose that $|y| \leq \frac{60(p+1)}{p-1}|x|$. Observe that

$$\frac{2}{p}|ER| \leq \frac{R^2}{8(p+1)} + \frac{8(p+1)}{p^2}E^2 \leq \frac{R^2}{8(p+1)} + 16E^2$$

and

$$\frac{2(p+1)}{p}|ES| \leq \frac{S^2}{8(p+1)} + \frac{8(p+1)^3}{p^2}E^2 \leq \frac{S^2}{8(p+1)^2} + 64E^2,$$

which is equivalent to saying that

$$4(1-\beta)|ER| \leq \frac{p-1}{8(p+1)}R^2 + 16(p-1)E^2 \leq \frac{p-1}{8(p+1)}R^2 + 16E^2$$

and

$$4\beta(p-1)|ES| \leq \frac{p-1}{4(p+1)}\cdot\frac{S^2}{2} + 64(p-1)E^2 \leq \frac{p-1}{4(p+1)}\cdot\frac{S^2}{2} + 64E^2.$$

Combining these estimates with (3.4.9) and (3.4.10), we get that

$$\begin{aligned}
\xi_{b_4}(D,E,R,S) &= \xi_{b_4}(D,R,S) + \big(\xi_{b_4}(D,E,R,S) - \xi_{b_4}(D,R,S)\big) \\
&\leq c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\left\{82E^2 + 4|ED| - \frac{p-1}{24(p+1)}(D^2+R^2)\right\} \\
&\leq c^\beta xy^2w^{1-\beta}v^{-\beta(p-1)}\left\{82E^2 + 4|ED| - \frac{p-1}{12(p+1)}|DR|\right\} \\
&\leq c^\beta w^{1-\beta}v^{-\beta(p-1)}\left\{82x^3D^2 + 4x^2|y|D^2 - \frac{p-1}{12(p+1)}y^3|ER|\right\}.
\end{aligned}$$

This proves the claim. □

The final lemma is much simpler. We will analyze functions $b_5(x,y,w,v) = c|x|^3 v^{1-p}$ and $b_6(x,y,w,v) = cx^2 v^{1-p}$.

Lemma 3.10. *We have*

$$\xi_{b_5}(D,S) \geq c|x|^3 v^{1-p} D^2, \qquad \xi_{b_6}(D,S) \geq \frac{2(2-p)}{p} c x^2 v^{1-p} D^2. \tag{3.4.11}$$

Furthermore, if $|y| \leq \frac{60(p+1)}{p-1}|x|$, *then*

$$\xi_{1882b_5+9b_6}(D,S) \geq c^\beta w^{1-\beta} v^{-\beta(p-1)} \left\{82|x|^3 D^2 + 6x^2 y D^2\right\}. \tag{3.4.12}$$

Proof. We have

$$\xi_{b_5}(D,S) = cx^3 v^{1-p} \begin{bmatrix} 6 & 3(1-p) \\ 3(1-p) & p(p-1) \end{bmatrix} \geq cx^3 v^{1-p} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

which gives the first inequality in (3.4.11). The second inequality is equivalent to the estimate

$$\begin{bmatrix} 2 & 2(1-p) \\ 2(1-p) & p(p-1) \end{bmatrix} \geq \begin{bmatrix} 2(2-p)/p & 0 \\ 0 & 0 \end{bmatrix},$$

which is easy to check. Therefore, we have the estimate

$$82\xi_{b_5}(D,S) \geq 82cx^3 v^{1-p} D^2 \geq 82c^\beta x^3 w^{1-\beta} v^{-\beta(p-1)} D^2.$$

Furthermore, if $p \leq 3/2$, then $2(2-p)/p \geq 2/3$, so

$$6c^\beta x^2 |y| w^{1-\beta} v^{-(p-1)\beta} D^2 \leq 6c^\beta x^2 w^{1-\beta} v^{-(p-1)\beta} D^2 \leq 9\xi_{b_6}(D,S).$$

If $p > 3/2$, then $60(p+1)/(p-1) \leq 300$ and

$$6c^\beta x^2 |y| w^{1-\beta} v^{-(p-1)\beta} D^2 \leq 1800 c^\beta x^3 w^{1-\beta} v^{-(p-1)\beta} D^2 \leq 1800\xi_{b_5}(D,S).$$

The proof is complete. □

Remark 3.4. The combination of the two lemmas above shows that in the set D_1:

$$\xi_{b_4-1882b_5-9b_6}(D,E,R,S) \leq -\frac{p-1}{4(p+1)}|y|^3 w^{1-1/6c} v^{-(p-1)/6c} \left[3E^2 + \frac{1}{3}|ER|\right].$$

Proof. When $|y| \geq \frac{60(p+1)}{p-1}x$, this follows immediately from (3.4.6). For $|y| < \frac{60(p+1)}{p-1}x$, we apply (3.4.7) and (3.4.12) to obtain the estimate

$$\begin{aligned}\xi_{b_4-1882b_5-9b_6} &\leq c^\beta w^{1-\beta} v^{-\beta(p-1)} \left[4x^2|y|D^2 - 6x^2|y|D^2 - \frac{p-1}{12(p+1)}|y|^3|ER|\right] \\ &\leq -\frac{p-1}{4(p+1)}|y|^3 c^\beta w^{1-\beta} v^{-\beta(p-1)} \left[\frac{8(p+1)}{p-1}E^2 + \frac{1}{3}|ER|\right] \\ &\leq -\frac{p-1}{4(p+1)}|y|^3 w^{1-1/6c} v^{-(p-1)/6c} \left[3E^2 + \frac{1}{3}|ER|\right].\end{aligned}$$

In the last line we again used the inequality $1 \leq wv^{p-1} \leq c$. □

Recall that $B : \mathbb{R}^2 \times \mathcal{D}_c \to \mathbb{R}$ is defined as

$$B = b_1 - b_2 - b_3 + \frac{216(p+1)}{p-1}(b_4 - 1882b_5 - 9b_6).$$

Theorem 3.11. *The function B satisfies the concavity property 3° on D.*

Proof. This is just a combination of (3.4.3), (3.4.5), (3.4.6), (3.4.7) and (3.4.11). We may assume that $(x, y, w, v) \notin \partial D_1$. Suppose first that $|y| \geq \alpha_p c|x|$. From (3.4.3), (3.4.5) and Remark 3.3 we infer that $\xi_{b_1-b_2-b_3}(D, E, R, S) \leq 0$. Moreover, from Remark 3.4, the convexity of b_5 and b_6, and the fact that b_4 is concave outside D_1 (because u vanishes there), we obtain that

$$\xi_{b_4-1882b_5-9b_6}(D, E, R, S) \leq 0.$$

These two inequalities yield the desired concavity property in the angular region $|y| \geq \alpha_p c|x|$. Now suppose $|y| < \alpha_p c|x|$: then $(x, y, w, v) \in D_1$. We use Remark 3.4, apply Remark 3.3 to estimate the matrix by J_2 and exploit (3.4.6) to show that

$$\xi_{b_1+\frac{216(p+1)}{p-1}(b_4-1882b_5-9b_6)}(D, E, R, S) \leq 0.$$

Indeed, we have that

$$\begin{aligned}\xi_{\frac{216(p+1)}{p-1}(b_4-1882b_5-9b_6)} &\leq -54|y|^3 w^{1-1/6c} v^{-(p-1)/6c}\left(3E^2 + \frac{1}{3}|ER|\right)\\ &\leq -3|y|^3 w^{1-1/6c} v^{-(p-1)/6c}\left(12E^2 + 6|ER|\right)\\ &\leq \xi_{-b_1}\end{aligned}$$

Because b_2 and b_3 are convex, the function B does satisfy the concavity condition. □

Theorem 3.12. *The function $\mathfrak{B}$ satisfies the concavity-type condition 3°.*

Proof. Both components have the appropriate concavity, so it remains to check the behavior of $\mathfrak{B}$ on the boundary ∂D. However, it follows from the construction that any point $(x, y, w, v) \in \partial D$ has a neighbourhood in which the function $\mathfrak{B}$ is equal to $\min(B, \bar{B})$. Thus, the condition 3° is satisfied: the minimum of concave functions is concave. □

We verified that the function $\mathfrak{B}$ satisfies all the required conditions. Hence, we get

$$\mathbb{E}\left[W_t \mathbb{1}_{\{|Y_t|\geq 1\}} - C_p^p [W]_{A_p}^p |X_t|^p V_t^{1-p}\right] \leq 0, \qquad t \geq 0. \tag{3.4.13}$$

It remains to use simple limiting arguments to obtain (3.2.2). The function $(x, v) \mapsto |x|^p v^{1-p}$ is convex, so $\mathbb{E}|X_t|^p V_t^{1-p} \leq \mathbb{E}|X|^p V^{1-p} = \mathbb{E}|X|^p W$. Furthermore, we have $W_t = \mathbb{E}(W|\mathcal{F}_t)$, so the above inequality implies $\mathbb{E}W\mathbb{1}_{\{|Y_t|\geq 1\}} \leq C_p^p [W]_{A_p}^p \mathbb{E}|X|^p W$. The claim follows by letting $t \to \infty$ and applying Fatou's lemma.

3.5. Further discussion

In the last section of this chapter we will discuss some extensions and corollaries of the weak-type bound. We have proved that

$$\|Y\|_{L^{p,\infty}(W)} \leq C_p [W]_{A_p} \|X\|_{L^p(W)}$$

for every $1 < p < \infty$. Recall that for the strong-type inequality, we actually obtained a stronger maximal estimate with $|Y|^*$ on the left-hand side. It was handled by constructing a new Bellman function with an additional variable. For the weak-type inequality such a maximal version also holds, which follows directly from a well-known stopping time argument. Namely, let us define a stopping time $\tau = \inf\{t : |Y_t| \geq 1\}$. The stopped process Y^τ is a stochastic integral of X^τ, so

$$\mathbb{E}W_{\tau\wedge t}\mathbb{1}_{\{|Y_{\tau\wedge t}|\geq 1\}} \leq C_p^p[W]_{A_p}^p\mathbb{E}|X_{\tau\wedge t}|^p V_{\tau\wedge t}^{1-p}, \qquad t \geq 0.$$

Hence, Fatou's lemma and the convexity of the function $(x,v) \mapsto |x|^p v^{1-p}$ yield, for any $a > 1$,

$$\begin{aligned}\mathbb{E}W\mathbb{1}_{\{|Y|^*\geq a\}} \leq \lim_{t\to\infty}\mathbb{E}W_{\tau\wedge t}\mathbb{1}_{\{|Y|_{\tau\wedge t}\geq 1\}} &\leq C_p^p[W]_{A_p}^p \lim_{t\to\infty}\mathbb{E}|X_{\tau\wedge t}|^p V_{\tau\wedge t}^{1-p}\\ &\leq C_p^p[W]_{A_p}^p\mathbb{E}|X|^p W.\end{aligned}$$

Letting $a \to 1$, we see that we have proved the following theorem.

Theorem 3.13. *Fix $1 < p < \infty$. Let X, Y be path-continuous martingales such that Y is is a stochastic integral of some predictable process H with values in $[-1,1]$ with respect to X. Then for any A_p weight W we have*

$$\||Y|^*\|_{L^{p,\infty}(W)} \leq C_p[W]_{A_p}\|X\|_{L^p(W)},$$

where C_p is some constant depending only on p. The linear dependence on $[W]_{A_p}$ is optimal.

There is a natural question about the weak-type bound for discrete-time martingales. Using the same method as in the previous chapter, we can prove this estimate for θ-regular filtrations.

Theorem 3.14. *Fix $1 < p < \infty$ and $\theta \in (0, 1/2]$. Let f, g be martingales adapted to a θ-regular filtration such that $|dg_n| \leq |df_n|$ for each $n = 0, 1, \ldots$. Then for any A_p weight w we have*

$$\||g|^*\|_{L^{p,\infty}(w)} \leq C_p\theta^{-1}[w]_{A_p}\|f\|_{L^p(w)},$$

where C_p is some constant depending only on p. The linear dependence on $[w]_{A_p}$ is optimal.

Combining the weak-type bound with the interpolation theorem gives the following interesting corollary.

Theorem 3.15. *Fix $1 < q < p < \infty$. Let X, Y be path-continuous martingales such that $Y = H \cdot X$ for some predictable process H with values in $[-1,1]$. Then for any A_q weight W we have*

$$\|Y\|_{L^p(W)} \leq C_{p,q}[W]_{A_q}\|X\|_{L^p(W)}, \tag{3.5.1}$$

where $C_{p,q}$ is a constant depending only on the parameters indicated.

Proof. Let T be a linear operator given by a stochastic integral $TX = H \cdot X$ for a fixed H as above. By Theorem 3.2, we have that

$$||TX||_{L^{q,\infty}(W)} \leq C_q[W]_{A_q}||X||_{L^q(W)},$$

for any $X \in L^q(W)$. Since $[W]_{A_{p+1}} \le [W]_{A_q}$, we also obtain

$$||TX||_{L^{p+1,\infty}(W)} \le C_{p+1}[W]_{A_q}||X||_{L^{p+1}(W)},$$

for any $X \in L^{p+1}(W)$. It remains to apply Marcinkiewicz interpolation theorem for $p \in (q, p+1)$. □

For $p \in [2, \infty)$ we can obtain the above result from the strong-type inequality. Indeed, using (2.1.3) and monotonicity $[W]_{A_q} \le [W]_{A_p}$ we obtain

$$||Y||_{L^{p,\infty}(W)} \le C_p[W]_{A_q}^{\max\{1,1/(p-1)\}}||X||_{L^p(W)}.$$

However, for $p \in (1, 2)$ this estimate is weaker than the linear dependence in (3.5.1).

Chapter 4

Weighted weak type (∞, ∞) inequality for differentially subordinate martingales

4.1. Statement of the result, some historical comments

In this chapter we continue the study of weighted weak-type inequalities in the endpoint case $p = \infty$. Given a random variable ξ, recall the definitions of nondecreasing rearrangement ξ^* and its maximal function ξ^{**} presented in Chapter 1. Following Bennett, DeVore and Sharpley [10], we set

$$||\xi||_{weak(\mathbb{P})} = \sup_{t \in (0,1]} (\xi^{**}(t) - \xi^*(t))$$

and define $weak(\mathbb{P}) = \{\xi : ||\xi||_{weak(\mathbb{P})} < \infty\}$. The main motivation behind the introduction of $weak(\mathbb{P})$ comes from the interpolation theory (the reader interested in the more detailed presentation is referred to [10]). First, note that this space contains $L^\infty(\Omega)$. Next, one can show that if a linear operator T is bounded from L^1 to $L^{1,\infty}$ and from L^∞ to $weak(\mathbb{P})$, then it can be extended to a bounded operator on all L^p spaces, $1 < p < \infty$. In other words, in the context of the above weak-L^∞ spaces, we have a substitute of Marcinkiewicz interpolation theorem for operators which are unbounded on L^∞. There is a further phenomenon which explains why the space $weak(\mathbb{P})$ can be regarded as a substitute for the weak L^∞. Namely, the Peetre K-functional for the pair (L^1, L^∞) (consult (Butzer and Berens [22], p. 184)) can be expressed explicitly in the form

$$K(\xi, t; L^1, L^\infty) = \int_0^t \xi^*(s)\mathrm{d}s = t\xi^{**}(t), \qquad t \in (0, 1],$$

and the weak-L^1 norm equals $||\xi||_{L^{1,\infty}} = \sup_{t \in (0,1]} t\xi^*(t) = \sup_{t \in (0,1]} t\frac{d}{dt}K(\xi, t; L^1, L^\infty)$. It seems plausible to define the weak-L^∞ norm by interchanging the roles of L^1 and L^∞ in the latter formula. Since $K(\xi, t; L^\infty, L^1) = tK(\xi, t^{-1}; L^1, L^\infty)$, we compute that

$$\sup_{t \in (0,1]} t\frac{d}{dt}K(\xi, t; L^\infty, L^1) = \sup_{t \in (0,1]} [\xi^{**}(t) - \xi^*(t)] = ||\xi||_{weak(\mathbb{P})},$$

as desired.

Coming back to the martingale context, let us recall the main result of Osękowski [62].

Theorem 4.1. *Let X, Y be martingales such that Y is differentially subordinate to X. Then we have the inequality*

$$||Y||_{weak(\mathbb{P})} \le 2||X||_{L^\infty} \tag{4.1.1}$$

and the constant 2 is the best possible.

The main goal of this chapter is to study the following weighted extension of (4.1.1):

$$||Y||_{weak(W)} \le C_W ||X||_{L^\infty(W)}. \tag{4.1.2}$$

under the assumption that W belongs to the class A_∞. We will also study the optimal dependence of C_W on the characteristic of the weight. We work in the same continuous-time setting as in the previous two chapters, in particular, we assume that processes X, Y and $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$ have continuous paths while $(V_t)_{t\geq 0}$ is allowed to have jumps.

Here is our main result, obtained jointly with Osękowski [12].

Theorem 4.2. *Let X, Y be continuous-path martingales such that X is bounded and Y is differentially subordinate to X. Then for any A_∞ weight W with $\|W\|_1 = 1$, we have*

$$||Y||_{weak(W)} \leq 97[W]_{A_\infty}||X||_{L^\infty}. \tag{4.1.3}$$

The linear dependence on the A_∞ characteristics is optimal.

A weaker result for A_1 weights was obtained in [63]: it was shown there that we have

$$||Y||_{L^1(W)} \leq C[W]_{A_1}||X||_{L^\infty}, \tag{4.1.4}$$

and the linear dependence on the characteristic is optimal (in the above sense). Since $[W]_{A_\infty} \leq [W]_{A_1}$ and $||Y||_{L^1(W)} \leq ||Y||_{weak(W)}$, (4.1.3) is indeed a generalization. Furthermore, the optimality of the linear dependence in the above theorem follows at once from the analogous sharpness in (4.1.4). Thus in what follows, we will focus on the proof of (4.1.3).

4.2. Bellman function method for weak type (∞,∞) inequality

Let us briefly discuss the version of Bellman function method which can be used to establish estimates involving the A_∞ class. The argumentation requires only some small modifications. As we already mentioned in Chapter 1, an A_∞ weight W of characteristic less or equal to c gives rise to a two-dimensional uniformly integrable martingale (W,V) taking values in the logarithmic domain

$$\mathcal{D}_c = \left\{(w,v) \in (0,\infty)\times\mathbb{R} : 1 \leq we^{-v} \leq c\right\}.$$

Let $G : [-1,1]\times\mathbb{R}\times\mathcal{D}_c \to \mathbb{R}$ is a given Borel function and assume that we want to show that

$$\mathbb{E}G(X_t,Y_t,W_t,V_t) \leq 0, \qquad t \geq 0, \tag{4.2.1}$$

for all (X,Y,W,V), where X,Y are continuous-path martingales such that $|X| \leq 1$, Y is differentially subordinate to X, W is a continuous-path martingale and V is a càdlàg martingale for which $1 \leq W_te^{-V_t} \leq c$ almost surely (for all $t \geq 0$). To study this problem, consider a C^2 function $B : [-1,1]\times\mathbb{R}\times\mathcal{D}_c \to \mathbb{R}$ with the following properties:

1° (Initial condition) We have $B(x,y,w,v) \leq 0$ if $|y|\leq|x|$ and $1 \leq we^{-v} \leq c$.

2° (Majorization property) There is a positive constant $\kappa > 0$ such that

$$B(x,y,w,v) \geq \kappa G(x,y,w,v) \qquad \text{for } (x,y,w,v) \in [-1,1]\times\mathbb{R}\times\mathcal{D}_c.$$

3° (Concavity-type property) There is a nonnegative function A on $\mathcal{D}_c$ such that for any $(x,y,w,v) \in [-1,1]\times\mathbb{R}\times\mathcal{D}_c$ and any $d,e,r,s \in \mathbb{R}$ we have

$$\left\langle D^2_{xywv}B(x,y,w,v)(d,e,r,s),(d,e,r,s)\right\rangle \leq A(x,y,w,v)(e^2-d^2),$$

where $D^2_{xywv}B$ is the Hessian matrix of the function B.

Lemma 4.3. *Let $c \geq 1$ be fixed. If there is a function B satisfying conditions 1°, 2° and 3°, then the inequality* (4.2.1) *holds true for all X, Y, W and V as above.*

Proof. The argument rests on Itô's formula applied to the function B and the process $Z = (X, Y, W, V)$. The process Z takes values in $\mathcal{D}_c$ and the function B is of class C^2, so

$$B(Z_t) = I_0 + I_1 + I_2/2 + I_3, \tag{4.2.2}$$

where

$$\begin{aligned} I_0 &= B(Z_0),\\ I_1 &= \int_{0+}^t B_x(Z_{s-})\mathrm{d}X_s + \int_{0+}^t B_y(Z_{s-})\mathrm{d}Y_s + \int_{0+}^t B_w(Z_{s-})\mathrm{d}W_s + \int_{0+}^t B_v(Z_{s-})\mathrm{d}V_s,\\ I_2 &= \int_{0+}^t D_{xywv}B(Z_{s-})\mathrm{d}[Z^c]_s,\\ I_3 &= \sum_{0<s\leq t} \left(B(Z_s) - B(Z_{s-}) - B_v(Z_{s-})\Delta V_s\right). \end{aligned}$$

Here the appearance of the term I_3 is due to the fact that X, Y, W have continuous paths (as we have assumed), but V does not have to possess this property. We have $I_0 \leq 0$, by the initial condition 1°. The stochastic integrals in I_1 have expectation 0. The expression I_2 is nonpositive, which follows directly from 3° by approximation. Finally, I_3 is also nonpositive, which is again due to 3°: setting $e = d = r = 0$, we see that for any x, y, z, w and v, the function $s \mapsto B(x, y, z, w, v+s)$ is concave (on the set of s satisfying $1 \leq w\exp(-v-s) \leq c$). Putting all the above facts together and taking the expectation in (4.2.2), we obtain $\mathbb{E}B(X_t, Y_t, W_t, V_t) \leq 0$. It remains to apply the majorization 2° to get the desired estimate (4.2.1). □

4.3. Construction of a special function: rough arguments

In this section we will informally describe the reasoning which led us to the correct Bellman function for a weak-type (∞,∞) inequality. As usual, to gain some intuition, we consider first the simpler unweighted setting. The proof of the estimate (4.1.1) rests on establishing the following intermediate result (see [62]):

$$\mathbb{E}(|Y_t| - \lambda)_+ \leq 2\mathbb{P}(|Y_t| > \lambda), \qquad \lambda > 0. \tag{4.3.1}$$

A similar pattern is true in the weighted case. We will prove the following fact.

Theorem 4.4. *Suppose that X and Y are martingales with continuous paths such that $||X||_{L^\infty} \leq 1$ and Y is differentially subordinate to X. Let W be an A_∞ weight with characteristic $[W]_{A_\infty}$. Then for any $\lambda \geq 0$ and $t \geq 0$ we have*

$$\mathbb{E}W(|Y_t| - \lambda)_+ \leq 97[W]_{A_\infty} W(|Y_t| > \lambda). \tag{4.3.2}$$

The inequality (4.3.2) is of the form (4.2.1), hence it remains to construct an appropriate special function. Once again, it is natural to get some insight by returning to the unweighted setting. For a fixed $\lambda > 0$, the two-dimensional Bellman function corresponding to the estimate (4.3.1) is equal to:

$$b(x,y) = \begin{cases} \frac{1}{2}(|y| - \lambda - 1)^2 - \frac{1}{2}|x|^2 & \text{for } |x| + |y| \geq \lambda + 1,\\ 0 & \text{for } |x| + |y| < \lambda + 1. \end{cases}$$

Our weighted function $\mathfrak{B}$ will be similar. Roughly speaking, we want $\mathfrak{B}(x,y,w,v)$ to be equal to 0 for (x,y) close to the origin, and to coincide with some concave function $B(x,y,w,v)$ elsewhere. Of course, the crucial problem now is to construct an appropriate function B satisfying the concavity. Fortunately, this will be much easier than in the previous two chapters and will not require any tedious decompositions of the domain. The key idea is to use the fact that composing a concave and increasing function with a concave function is also concave. More precisely, we first find a 3-dimensional concave function $F : [0,\infty)\times\mathcal{D}_c \to \mathbb{R}$ which is nondecreasing with respect to the first variable. Let $U : [-1,1]\times\mathbb{R} \to [0,\infty)$ be a two-dimensional function such that $(U(X_t,Y_t))_{t\geq 0}$ is a supermartingale for any pair of (X,Y) satisfying the differential subordination. The desired component B of the Bellman function is, essentially, the composition $F(U(x,y),w,v)$. Finding U is relatively easy - we can use numerous functions corresponding to the (unweighted) martingale inequalities which are known to possess the required property. The main reason why this composition argument works here - contrary to the previous two chapters - is the condition $|x|\leq 1$, which is a consequence of the boundedness assumption $||X||_{L^\infty}\leq 1$. We will formally describe the above construction in details in the next section.

4.4. A special function

Let $c=[W]_{A_\infty}$. To show (4.3.2), we need to find a C^2 function $B : [-1,1]\times\mathbb{R}\times\mathcal{D}_c \to \mathbb{R}$ which satisfies the conditions $1^\circ - 3^\circ$ with $G(x,y,z,w,v) = (|y|-\lambda-97c)w\mathbb{1}_{\{|y|>\lambda\}}$. To keep the notation short, define the constants

$$\alpha = 1-(2c)^{-1}, \qquad \beta=(6c)^{-1}, \qquad a=3/4, \qquad p=6c$$

and the following auxiliary functions: for $(u,w,v)\in(0,\infty)^2\times\mathbb{R}$ with $1\leq we^{-v}\leq c$, set

$$F(u,w,v) = u^\beta(we^{-v}-a)^\alpha e^v.$$

Note that $p\geq 6>2$. For any x, $y\in\mathbb{R}$, introduce

$$\mathcal{U}(x,y) = p(1-1/p)^{p-1}\big(|y|-(p-1)|x|\big)\big(|x|+|y|\big)^{p-1}.$$

This is the celebrated special function invented by Burkholder [20] to establish sharp L^p bounds for differentially subordinate martingales. Actually, this object is also meaningful if x, y are vectors from $\mathbb{R}^2$ (or even a separable Hilbert space), if we interpret $|\cdot|$ as the Euclidean norm. Burkholder proved that $\mathcal{U}$ enjoys the following properties.

Lemma 4.5. *If (x,y), $(d,e)\in(\mathbb{R}^n)^2$, then we have*

$$\langle D^2\mathcal{U}(x,y)(d,e),(d,e)\rangle \leq p(p-1)\cdot p(1-1/p)^{p-1}(|x|+|y|)^{p-2}(|e|^2-|d|^2)$$

and

$$\mathcal{U}(x,y)\geq |y|^p-(p-1)^p|x|^p.$$

The function $\mathcal{U}$ (for $n=1$) is not sufficient for our purposes, since it is not of class C^2 (the points at the x-axis are problematic). To overcome this difficulty, we take $\mathcal{U}$ in the dimension $n=2$ and define the C^2 function $U:\mathbb{R}^2\to\mathbb{R}$ by

$$U(x,y) = \mathcal{U}((x,0),(y,1)) = p(1-1/p)^{p-1}\big(\tilde{y}-(p-1)|x|\big)\big(|x|+\tilde{y}\big)^{p-1}.$$

Here and below, we will use the notation $\tilde{y}=(y^2+1)^{1/2}$. In the series of lemmas below we will analyze size and concavity properties of the auxiliary function U and F.

Directly from Lemma 4.5, we get

Lemma 4.6. *For $(x,y)\in\mathbb{R}^2$ and any $d,e\in\mathbb{R}$ we have*

$$\langle D^2U(x,y)(d,e),(d,e)\rangle \leq p(p-1)\cdot p(1-1/p)^{p-1}(|x|+\tilde{y})^{p-2}(e^2-d^2)$$

and

$$U(x,y)\geq |y|^p-(p-1)^p|x|^p.$$

Concerning F, we start with the following fact.

Lemma 4.7. *For any $(u,w,v)\in(0,\infty)\times\mathcal{D}_c$ with $1\leq we^{-v}\leq c$, we have*

$$\frac{1}{4}wu^\beta \leq F(u,w,v)\leq wu^\beta. \tag{4.4.1}$$

Proof. Recall that $\alpha=1-(2c)^{-1}$ and $a=3/4$. We must show that

$$\frac{1}{4}\leq \frac{(we^{-v}-a)^\alpha}{we^{-v}}\leq 1.$$

Firstly observe that the function $t\longmapsto (t-a)^\alpha/t$ is increasing. Indeed, we have

$$\left(\frac{(t-a)^\alpha}{t}\right)'=\frac{(t-a)^{\alpha-1}((\alpha-1)t+a)}{t^2}\geq 0.$$

Thus, the assertion follows from the trivial estimates $1/4\leq(1-a)^\alpha$ and $(c-a)^\alpha/c\leq 1$. □

Lemma 4.8. *The Hessian matrix of $-F$ is nonnegative-definite. (That is, the function $-F$ is a locally convex function).*

Proof. For brevity, set $\phi(t)=(t-a)^\alpha$ for $t\geq a$. We will also write $t=we^{-v}$. The proof rests on Sylvester's criterion. First, observe that $F_{ww}(u,w,v)=u^\beta e^{-v}\phi''(t)$ is negative, because the function ϕ is concave. Next, since $F_{wv}(u,w,v)=-u^\beta we^{-v}\phi''(t)$ and $F_{vv}(u,w,v)=u^\beta e^v\phi(t)+u^\beta w^2e^{-v}\phi''(t)-u^\beta w\phi'(t)$, we derive that

$$\det\begin{pmatrix}F_{ww} & F_{wv}\\ F_{vw} & F_{vv}\end{pmatrix}=u^{2\beta}\phi(t)\phi''(t)-u^{2\beta}t\phi'(t)\phi''(t)=u^{2\beta}\phi''(t)(\phi(t)-t\phi'(t)).$$

This is positive, because the expression $\phi(t)-t\phi'(t)=(t-a)^{\alpha-1}(t(1-\alpha)-a)$ is negative when $t\leq c$. It remains to show that the determinant of the full Hessian is nonpositive:

$$\det\begin{pmatrix}F_{ww} & F_{wv} & F_{wu}\\ F_{vw} & F_{vv} & F_{vu}\\ F_{uw} & F_{uv} & F_{uu}\end{pmatrix}\leq 0.$$

Add to the second column the first column multiplied by w; then add to the second row the first row multiplied by w. Then the above inequality amounts to saying that

$$\det\begin{pmatrix}u^\beta e^{-v}\phi''(t) & 0 & \beta u^{\beta-1}\phi'(t)\\ 0 & e^vu^\beta(\phi(t)-t\phi'(t)) & \beta u^{\beta-1}\phi(t)e^v\\ \beta u^{\beta-1}\phi'(t) & \beta u^{\beta-1}\phi(t)e^v & \beta(\beta-1)u^{\beta-2}e^v\phi(t)\end{pmatrix}\leq 0.$$

Observe that powers of u and the factors e^v, e^{-v} do not change the sign of the determinant. The above inequality is equivalent to

$$\det\begin{pmatrix}\phi''(t) & 0 & \beta\phi'(t)\\ 0 & \phi(t)-t\phi'(t) & \beta\phi(t)\\ \beta\phi'(t) & \beta\phi(t) & \beta(\beta-1)\phi(t)\end{pmatrix}$$
$$=\beta(\beta-1)\phi(t)\phi''(t)(\phi(t)-t\phi'(t))-\beta^2(\phi'(t))^2(\phi(t)-t\phi'(t))-\beta^2(\phi(t))^2\phi''(t)\leq 0,$$

or, after some manipulations, $(\alpha-1)^2t+(\alpha+\beta\alpha-1)a\le 0$. This inequality is the strongest when $t=c$ and then reads

$$\beta\le\frac{(1-\alpha)((\alpha-1)c+a)}{a\alpha}.$$

This is easily shown to be true: simply plug the formulas for α and β. □

We are ready for the main step: we will construct the special function $\mathfrak{B}$ and check that it satisfies the conditions 1°, 2° and 3°. The function is given by the formula

$$\mathfrak{B}(x,y,w,v)=\begin{cases}F\left(U\left(x,|y|-\mu\right)+(p-1)^p,w,v\right)-12cw, & \text{for } (x,y,w,v)\in D_{c,1},\\ 0, & \text{for } (x,y,w,v)\in D_{c,2},\end{cases}$$

where $\mu=\lambda+49c$. Note that $U(x,y-\mu)+(p-1)^p>0$ (see the second estimate in Lemma 4.5), so the function is well-defined. The sets $D_{c,1}$ and $D_{c,2}$ are constructed so that B is continuous. More precisely, let $\xi:[0,1]\to[0,\mu]$ be given by

$$\xi(x)=\sup\{y\in[0,\mu]:F\left(U\left(x,|y|-\mu\right)+(p-1)^p,w,v\right)-12cw>0\}.$$

Sets $D_{c,1}$ and $D_{c,2}$ are defined as follows:

$$D_{c,2}=\{(x,y,w,v)\in[-1,1]\times\mathbb{R}\times\mathcal{D}_c:|y|\le\xi(|x|)\},$$

$$D_{c,1}=[-1,1]\times\mathbb{R}\times\mathcal{D}_c\backslash D_{c,2}.$$

From the definition it is not obvious that $\xi(x)$ is well defined. It will be proved in the next lemma.

Lemma 4.9. *For every $x\in[0,1]$ we have that* $\max(1,\lambda)<\xi(x)<\mu$.

Proof. First observe that the function $y\mapsto F\left(U\left(x,|y|-\mu\right)+(p-1)^p,w,v\right)-12cw$ considered on the domain $[0,\mu]$ is continuous. Hence, it is sufficient to show that for every $x\in[0,1]$:

$$F\left(U\left(x,\lambda-\mu\right)+(p-1)^p,w,v\right)-12cw>0, \tag{4.4.2}$$

$$F\left(U\left(x,1-\mu\right)+(p-1)^p,w,v\right)-12cw>0 \tag{4.4.3}$$

and

$$F\left(U\left(x,\mu-\mu\right)+(p-1)^p,w,v\right)-12cw<0. \tag{4.4.4}$$

It is easy to check that the function $U(x,y)$ is nonincreasing as a function of $x>0$. Hence, it is sufficient to prove (4.4.2) at the point $x=1$, which is equivalent to

$$\left(U\left(1,\lambda-\mu\right)+(p-1)^p\right)^\beta\frac{\left(we^{-v}-a\right)^\alpha}{we^{-v}}>12c.$$

Now we use Lemma 4.6 and Lemma 4.7, obtaining

$$\left(U\left(1,\lambda-\mu\right)+(p-1)^p\right)^\beta\frac{\left(we^{-v}-a\right)^\alpha}{we^{-v}}\ge|\lambda-\mu|^{p\beta}\frac{1}{4}>12c.$$

Similarly we show (4.4.3). Again by the monotonicity with respect to x, it is sufficient to show (4.4.4) at the point $x=0$; this is equivalent to

$$\left(U\left(0,0\right)+(p-1)^p\right)^\beta\frac{\left(we^{-v}-a\right)^\alpha}{we^{-v}}<12c.$$

Now we use Lemma 4.7 to obtain that

$$(\mathcal{U}(1,0)+(p-1)^p)^\beta \frac{(we^{-v}-a)^\alpha}{we^{-v}} < 2(p)^{p\beta} = 12c.$$

This completes the proof of the lemma. □

In the next series of lemmas we will show that the function $\mathfrak{B}$ satisfies the conditions 1°, 2° and 3°.

Lemma 4.10. *The function $\mathfrak{B}$ satisfies the initial condition 1°.*

Proof. From the previous lemma we know that $\xi(x) \geq 1$. Hence, for $|y| \leq |x|$ we have $(x,y,w,v) \in D_{c,2}$ and $\mathfrak{B}(x,y,w,v) = 0$. □

Lemma 4.11. *The function $\mathfrak{B}$ satisfies the majorization property*

$$\mathfrak{B}(x,y,w,v) \geq \frac{1}{4}(|y|-\lambda-97c)w\mathbb{1}_{|y|\geq\lambda}.$$

Proof. When $(x,y,w,v) \in D_{c,2}$, Lemma 4.9 gives

$$|y| \leq \mu < \lambda + 97c,$$

so the condition is satisfied. On the other hand, for $(x,y,w,v) \in D_{c,1}$ we have that $|y| \geq \lambda$. The second inequality of Lemma 4.6 implies that $(U(x,y)+(p-1)^p)^\beta \geq |y|$ and, as we have proved in Lemma 4.7, we also have

$$\frac{(we^{-v}-a)^\alpha}{we^{-v}} \geq \frac{1}{4}.$$

Consequently, we obtain that $\mathfrak{B}(x,y,w,v) \geq (||y|-\mu|-48c)w/4$ and it is sufficient to show that $||y|-\mu|-48c \geq |y|-\lambda-97c$. We consider two cases. When $|y| \leq \mu$ we have

$$||y|-\mu|-48c \geq -48c = \mu-\lambda-97c \geq |y|-\lambda-97c,$$

while for $|y| > \mu$ we see that $||y|-\mu|-48c = |y|-\mu-48c = |y|-\lambda-97c$. □

Lemma 4.12. *The function $\mathfrak{B}$ satisfies the condition 3°.*

Proof. Fix $(x,y,w,v) \in D_{c,1}\backslash\partial D_{c,2}$ and $d,e,r,s \in \mathbb{R}$. To shorten the notation, define

$$\begin{aligned} P &= (U(x,|y|-\mu)+(p-1)^p, w, v), \\ P_\varepsilon &= (U(x+\varepsilon d, |y+\varepsilon e|-\mu)+(p-1)^p, w+\varepsilon r, v+\varepsilon s). \end{aligned}$$

By concavity of F, for sufficiently small $\varepsilon > 0$,

$$\mathfrak{B}(x+\varepsilon d, y+\varepsilon e, w+\varepsilon r, v+\varepsilon s)+12c(w+\varepsilon r) = F(P_\varepsilon) \leq F(P)+\langle\nabla F(P), P_\varepsilon - P\rangle,$$

$$\mathfrak{B}(x-\varepsilon d, y-\varepsilon e, w-\varepsilon r, v-\varepsilon s)+12c(w-\varepsilon r) = F(P_{-\varepsilon}) \leq F(P)+\langle\nabla F(P), P_{-\varepsilon} - P\rangle.$$

Let us add these inequalities, substract $2F(P)$ from both sides, divide by ε^2 and let $\varepsilon \to 0$. As the result, we get

$$\begin{aligned} &\left\langle D^2_{xywv}\mathfrak{B}(x,y,w,v)(d,e,r,s),(d,e,r,s)\right\rangle \\ &\quad \leq F_u(P)\left\langle D^2_{xy}U(x,y)(d,e),(d,e)\right\rangle \\ &\quad \leq F_u(P)p(p-1)\cdot p(1-1/p)^{p-1}(|x|+(\tilde{y}))^{p-2}(e^2-d^2), \end{aligned}$$

where the last line follows from Lemma 4.6. Since $F_u \geq 0$, we are done.

For $(x,y,w,v) \in D_{c,2}\backslash\partial\mathcal{D}_{c,1}$ the function $\mathfrak{B}$ is constant and the condition 3° is satisfied. □

Proof of (4.3.2). We start with an important observation concerning the regularity of $\mathfrak{B}$. Obviously, this function is not of class C^2 (which is required when applying the Bellman function method: we use Itô's formula there). However, this can handled as before. Directly from the construction we see that every point (x,y,w,v) in the domain of $\mathfrak{B}$ has the neighbourhood such that the function $\mathfrak{B}$ restricted to this domain is the minimum of the finite number of C^2 functions. This and the concavity-type property is enough to show that

$$\mathbb{E}\mathfrak{B}(X_t,Y_t,W_t,V_t)\leq \mathbb{E}\mathfrak{B}(X_0,Y_0,W_0,V_0),$$

which by 1° and 2° yields

$$\mathbb{E}W_t(|Y_t|-\lambda)_+\leq 97[W]_{A_\infty}\mathbb{E}W_t1\!\!1_{(|Y_t|>\lambda)}.$$

Since $W_t=\mathbb{E}(W|\mathcal{F}_t)$, we have that that $\mathbb{E}W(|Y_t|-\lambda)_+\leq 97[W]_{A_\infty}W(|Y_t|>\lambda)$. Now we use Fatou's lemma and Lebesgue's dominated convergence theorem to obtain

$$\begin{aligned}\mathbb{E}W(|Y_\infty|-\lambda)_+ &\leq \lim_{t\to\infty}\mathbb{E}W(|Y_t|-\lambda)_+\leq \lim_{t\to\infty}97[W]_{A_\infty}W(|Y_t|>\lambda)\\ &=97[W]_{A_\infty}W(|Y_\infty|>\lambda).\end{aligned}$$

This proves (4.3.2). □

Proof of Theorem 4.2. Without loss of generality, we may and do assume that $||X||_{L^\infty(W)}\leq 1$. Pick arbitrary $s\in[0,\infty], t\in(0,1]$ and recall the alternative definition of Y_s^{**}:

$$Y_t^{**}=\sup\left\{\frac{1}{W(E)}\mathbb{E}W|Y_s|1\!\!1_E : E\in\mathcal{F}, W(E)=t\right\}.$$

It follows that

$$Y_s^{**}(t)-Y_s^*(t)=\sup\left\{\frac{1}{W(E)}\mathbb{E}W(|Y_s|-Y_s^*(t))1\!\!1_E : E\in\mathcal{F}, W(E)=t\right\}.$$

However, by the definition of $Y_s^*(t)$, we have $W(|Y_s|>Y_s^*(t))\leq t$. Hence, the above formula implies

$$Y_s^{**}(t)-Y_s^*(t)\leq\frac{1}{W(|Y_s|>Y_s^*(t))}\mathbb{E}W(|Y_s|-Y_s^*(t))_+\leq 97[W]_{A_\infty},$$

where the latter bound follows from (4.3.2). This gives the desired bound. □

Chapter 5

Weighted maximal inequality for martingale transforms

5.1. Statement of the result, some historical comments

As we have already seen above, the Bellman function method can be modified to prove maximal martingale inequalities. The modification was introduced by Burkholder in [21] who applied it to establish the following result.

Theorem 5.1. *If f, g are martingales satisfying $dg_n = \varepsilon_n df_n$, $n = 0, 1, 2, \dots$ for some predictable sequence $\varepsilon = (\varepsilon_n)_{n\geq 0}$ with values in $[-1, 1]$, then*

$$||g||_{L^1} \leq \eta|| \, |f|^*||_{L^1}, \tag{5.1.1}$$

where $\eta = 2.536\dots$ is the unique solution of the equation $\eta - 3 = -\exp\left(\frac{1-\eta}{2}\right)$. The constant is the best possible.

See also [57], [58] and [59] for related results and generalizations. The main theme of this chapter is to study the following weighted extension of (5.1.1):

$$|| \, |g|^*||_{L^1(w)} \leq C_{\theta,w}|| \, |f|^*||_{L^1(w)}. \tag{5.1.2}$$

Note that the maximal functions appear on both sides of the estimate. We will show that if w belongs to the class A_∞, then (5.1.2) holds for all martingales f and their ± 1-transforms. In addition, we will study the sharp dependence of the constant C on the characteristics $[w]_{A_\infty}$.

The main result of this chapter was obtained by Osękowski and the author in [13]. Here is the precise statement.

Theorem 5.2. *Fix $\theta \in (0, 1/2]$. Let f, g be martingales adapted to a θ-regular filtration such that g is a transform of f by a predictable sequence with values in $[-1, 1]$. Then for any A_∞ weight w we have*

$$|| \, |g|^*||_{L^1(w)} \leq 769\, \theta^{-2}[w]_{A_\infty}|| \, |f|^*||_{L^1(w)}. \tag{5.1.3}$$

The dependence on the A_∞ characteristics of the weight is optimal.

A weaker result for Haar multipliers and A_p weights was obtained in [66]. It was shown there that we have

$$\left|\left| \max_{0\leq n\leq N} \left|\sum_{k=0}^{n} \varepsilon_k a_k h_k\right|\right|\right|_{L^1(w)} \leq C_p[w]_{A_p} \left|\left| \max_{0\leq n\leq N} \left|\sum_{k=0}^{n} a_k h_k\right|\right|\right|_{L^1(w)}, \tag{5.1.4}$$

where $1 < p < \infty$, w is a dyadic A_p weight, N is a nonnegative integer, $a_0, a_1, \dots, a_N$ are real numbers, $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_N$ is a sequence of signs and $(h_k)_{k=0}^\infty$ is the Haar system on $[0, 1)$. Moreover, it was proved in [66] that the linear dependence on the characteristic is optimal. Observe that (5.1.3) generalizes this result in two directions. Firstly, we consider the more general case of θ-regular filtrations. Secondly, since $[w]_{A_\infty} \leq [w]_{A_p}$, the estimate (5.1.3) is stronger; hence, the optimality of the linear dependence in (5.1.3) follows at once from the analogous sharpness in (5.1.4) and all we need is the proof of (5.1.3).

Let us now make an important comment on the θ-regularity of the underlying filtration. The multiplicative constant in the inequality (5.1.3) depends on θ and goes to infinity when θ tends to 0. We will prove that this dependence is necessary, even if we consider the weaker estimate for A_p weights for any given $p > 1$. Here is the precise formulation.

Theorem 5.3. *Let $p > 1$ and let K be an arbitrary positive constant. Then there is a positive integer d, a martingale f on d-dimensional dyadic probability space, an A_p weight w satisfying $[w]_{A_p} \leq 2$ and a predictable sequence v with values in $\{-1, 1\}$ such that the associated martingale transform g satisfies*

$$||g||_{L^1(w)} > K|||f|^*||_{L^1(w)}.$$

We conclude the discussion by passing to the continuous-time context. We will establish the following statement. As in the previous chapters, we assume that the process $(W_t)_{t\geq 0} = (\mathbb{E}(W|\mathcal{F}_t))_{t\geq 0}$ has continuous paths while $(V_t)_{t\geq 0}$ is allowed to have jumps.

Theorem 5.4. *Let X, Y be continuous-path martingales such that $Y = H \cdot X$ for some predictable process H with values in $[-1, 1]$. Then for any A_∞ weight W giving rise to a continuous martingale $(W_t)_{t\geq 0}$, we have*

$$||\,|Y|^*||_{L^1(W)} \leq 3076\,[W]_{A_\infty}||\,|X|^*||_{L^1(W)}.$$

The linear dependence on the A_∞ characteristics is optimal.

5.2. Bellman function method for weighted maximal inequalities

In this section we will present a version of the Bellman function method modified to establish a maximal inequality for discrete-time martingales adapted to regular filtrations. Suppose that $G : \mathbb{R}^2 \times [0, \infty) \times (0, \infty) \to \mathbb{R}$ is a given continuous function and assume that we want to show that

$$\mathbb{E}G(f_n, g_n, |f|_n^*, w_n) \leq 0, \qquad n \geq 0, \tag{5.2.1}$$

where f, g are martingales such that g is the ± 1-transform of f and w is an A_∞ weight satisfying $[w]_{A_\infty} \leq c$. We additionally assume that all these processes are adapted to a θ-regular filtration on some probability space; this in particular implies that for each n, the variables f_n, g_n and w_n take only a finite number of values. The key to handle the estimate (5.2.1) is to construct a function B defined on the five-dimensional domain

$$\mathcal{D} = \{(x, y, z, u, v) \in \mathbb{R}^2 \times (0, \infty)^2 \times \mathbb{R} : |x| \leq z, 1 \leq ue^{-v} \leq c\}$$

and enjoying the following three properties:

1° (Initial condition) We have $B(x, y, |x|, u, v) \leq 0$ if $|y| \leq |x|$, $|x| > 0$ and $1 \leq ue^{-v} \leq c$.

2° (Majorization property) We have

$$B(x, y, z, u, v) \geq G(x, y, z, u) \qquad \text{for } (x, y, z, u, v) \in \mathcal{D}.$$

3° (Concavity-type property) For any $(x, y, z, u, v) \in \mathcal{D}$, any $\varepsilon \in [-1, 1]$, any positive integer $k \leq 1/\theta$ and any sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying

$$\alpha_j \in [\theta, 1) \qquad \text{and} \qquad \sum_{j=1}^k \alpha_j = 1, \tag{5.2.2}$$

$$\sum_{j=1}^k \alpha_j h_j = \sum_{j=1}^k \alpha_j r_j = \sum_{j=1}^k \alpha_j s_j = 0$$

and

$$(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j) \in \mathcal{D},$$

we have

$$B(x, y, z, u, v) \geq \sum_{j=1}^k \alpha_j B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j).$$

The relation between functions satisfying the above special properties and the validity of (5.2.1) is described in the statement below.

Theorem 5.5. *If there is a function satisfying 1°, 2° and 3°, then the inequality (5.2.1) holds true.*

Proof. We repeat the reasoning from Chapter 1. By a standard limiting argument (using continuity of G and the fact that the variables f_n, g_n, $\ldots$ take only a finite number of values), we may and do assume that $|f_0| > 0$ almost surely; then the process $z_n = (f_n, g_n, |f|_n^*, w_n, \sigma_n)$ takes values in $\mathcal{D}$. The key fact is that the process $(B(z_n))_{n \geq 0}$ is a supermartingale, which is an immediate consequence of the concavity-type condition 3°:

$$\begin{aligned}
&\mathbb{E}[B(f_n, g_n, |f|_n^*, w_n, \sigma_n)|\mathcal{F}_{n-1}] \\
&= \mathbb{E}[B(f_{n-1} + df_n, g_{n-1} + dg_n, |f|_{n-1}^* \vee |f_n + df_n|, w_{n-1} + dw_n, \sigma_{n-1} + d\sigma_n)|\mathcal{F}_{n-1}] \\
&\leq B(f_{n-1}, g_{n-1}, |f|_{n-1}^*, w_{n-1}, \sigma_{n-1}).
\end{aligned}$$

Therefore, if we apply the majorization 2° and then the initial condition 1°, we get

$$\mathbb{E}G(f_n, g_n, |f|_n^*, w_n) \leq \mathbb{E}B(f_n, g_n, |f|_n^*, w_n, \sigma_n) \leq \mathbb{E}B(f_0, g_0, |f|_0^*, w_0, \sigma_0) \leq 0,$$

which is the desired inequality (5.2.1). □

As in the unweighted case, the implication of the above theorem can be reversed.

Theorem 5.6. *If the inequality* (5.2.1) *holds true (for all f, g and all weights w with $[w]_{A_\infty} \leq c$), then there is a special function satisfying 1°, 2° and 3°.*

Proof. Define $B : \mathcal{D} \to \mathbb{R}$ by the abstract formula

$$B(x, y, z, u, v) = \sup \mathbb{E}G(f_n, g_n, |f|_n^* \vee z, w_n).$$

Here the supremum is taken over all n, all A_∞ weights w satisfying $[w]_\infty \leq c$, $w_0 = u$, $\mathbb{E} \log w = v$ and all martingale pairs (f, g) satisfying $(f_0, g_0) = (x, y)$ and $dg_k = \varepsilon_k df_k$, $k \geq 1$, for some predictable sequence $\varepsilon = (\varepsilon_k)_{k \geq 1}$ with values in $[-1, 1]$. Here the probability space as well as the θ-regular filtration are also assumed to vary. We will show that the function B satisfies conditions 1°-3°. The initial condition 1° follows immediately from the assumed inequality (5.2.1). The majorization condition is also easy: it suffices to compute the expression in the definition of B for $n = 0$. The most difficult issue is the concavity-type condition 2°. We will use the so-called "splicing" argument. Fix the parameters x, y, z, u, v, k, $\ldots$ as in the formulation of 3° and, for each $j = 1, 2, \ldots, k$, pick arbitrary martingales (f^j, g^j, w^j) as in the definition of $B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j)$. We may assume that these martingales are given on k pairwise disjoint probability spaces $(\Omega^j, \mathcal{F}^j, \mathbb{P}^j)$. Now we "glue" these spaces and the martingale triples into one space and one triple using the parameters $(\alpha_j)_{j=1}^k$. Namely, let $\Omega = \Omega^1 \cup \Omega^2 \cup \ldots \cup \Omega^k$, $\mathcal{F} = \sigma(\mathcal{F}^1, \mathcal{F}^2, \ldots, \mathcal{F}^k)$ and define the probability measure $\mathbb{P}$ on $\mathcal{F}$ by requiring that $\mathbb{P}(\bigcup_{j=1}^k A_j) = \sum_{j=1}^k \alpha_j \mathbb{P}(A_j)$ for any $A_j \in \mathcal{F}^j$, $j = 1, 2, \ldots, k$. Next, define (f, g, w) by $(f_0, g_0, w_0) = (x, y, w)$ and

$$(f_n(\omega), g_n(\omega), w_n(\omega)) = (f_{n-1}^j(\omega), g_{n-1}^j(\omega), w_{n-1}^j(\omega)),$$

if $\omega \in \Omega^j$. Finally, let $(\mathcal{F}_n)_{n\geq 0}$ be the natural filtration of (f, g, w).

Let us now study the properties of the object we have just constructed. Directly from the above definition, we see that

$$\mathbb{E}((f_1, g_1, w_1)|\mathcal{F}_0) = \mathbb{E}(f_1, g_1, w_1) = \sum_{j=1}^k \alpha_j (x + h_j, y + \varepsilon h_j, u + r_j) = (x, y, u).$$

Furthermore, since (f^j, g^j, w^j) are martingales, the triple (f, g, w) has this property as well. In addition,

$$\mathbb{E} \log(w) = \sum_{j=1}^k \alpha_j \mathbb{E}^j \log(w^j) = \sum_{j=1}^k \alpha_j (v_j + s_j) = v,$$

where $\mathbb{E}^j$ is the expectation with respect to probability measure $\mathbb{P}^j$. Our next observation is that w is an A_∞ weight with $[w]_{A_\infty} \leq c$. Indeed, we have $w_0 e^{-\sigma_0} = ue^{-v} \leq c$, and for $n \geq 1$ the pointwise estimate $w_n e^{-\sigma_n} \leq c$ follows from condition $[w^j]_{A_\infty} \leq c$. Consequently, by the very definition of B,

$$B(x, y, z, u, v) \geq \mathbb{E} G(f_n, g_n, |f_n| \vee z, w_n) = \sum_{j=1}^k \alpha_j \mathbb{E}^j G(f_{n-1}^j, g_{n-1}^j, |f_{n-1}^j| \vee z, w_{n-1}^j),$$

so taking the supremum over all n and all triples (f^j, g^j, w^j) as above, we obtain

$$B(x, y, z, u, v) \geq \sum_{j=1}^k \alpha_j B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j).$$

This is precisely the desired condition 3°. □

Two important comments are in order.

Remark 5.1. The above method works for A_p weights as well: the only change concerns the definition of the domain $\mathcal{D}$, in which the double estimate $1 \leq ue^{-v} \leq c$ should be changed to $1 \leq uv^{p-1} \leq c$.

Remark 5.2. Suppose that we are interested in the estimate (5.2.1) in the d-dimensional dyadic context. Then the above approach can be modified easily: we consider the function B given by the abstract formula as above,

$$B(x, y, z, u, v) = \sup \mathbb{E} G(f_n, g_n, |f|_n^* \vee z, w_n).$$

Here the supremum taken over all martingales as in the above proof, the essential difference is that the probability space is fixed to be $([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|)$ and the filtration is assumed to be dyadic. Thanks to the fractal, self-similar structure of the dyadic filtration, the above splicing argument is valid, and the function B satisfies 1°, 2° and a weaker version of 3°, with all α_j's equal to 2^{-d}. A similar modification can be applied in the context of A_p weights (see the previous remark). This observation will be crucial in the last section where we show that (5.1.3) cannot hold universally, i.e., with a constant independent of θ.

5.3. Construction of a special function

This section contains some informal reasoning which leads to the special function corresponding to (5.1.3). As we will see later, the main difficulty lies in proving the estimate

$$||g_n||_{L^1(w)} \leq C[w]_{A_\infty} |||f|_n^*||_{L^1(w)}, \qquad n \geq 0,$$

which is slightly weaker than (5.1.3), since it does not involve the maximal function of g on the left. This inequality is of the form (5.2.1) with $G(x, y, z, u, v) = |y|u - Cczu$, where $c = [w]_{A_\infty}$, and hence all we need is an appropriate special function B. As usual, at the first glance it is not clear how to search for this object. Quite unexpectedly, there will be plenty of similarities between this Bellman function and the function constructed in the preceding chapter.

Anyhow, as previously, we hope to gain some intuition from the analysis of the unweighted case. We start with the non-maximal $L^\infty \to L^2$ inequality (as we will see in a moment, it will be of key importance): if f, g are martingales such that $\|f\|_\infty \leq 1$ and g is a ± 1-transform of f, then we have $\|g\|_2 \leq \|f\|_2 \leq \|f\|_\infty \leq 1$. This trivial result can be proved by Bellman function method: the corresponding $u : [-1, 1] \times \mathbb{R} \to \mathbb{R}$ is

$$u(x, y) = y^2 - x^2.$$

Next, we turn towards maximal estimates (still in the unweighted setting). As shown in [59], the special function $\mathcal{U} : \{(x, y, z) : |x| \leq z\} \to \mathbb{R}$ corresponding to a continuous analogue of (5.1.1) is given by

$$\mathcal{U}(x, y, z) = \frac{y^2 - x^2 - z}{z} = z\left(u\left(\frac{x}{z}, \frac{y}{z}\right) - 1\right). \tag{5.3.1}$$

This special function uses two components: the multiplicative constant z which controls the maximal function of f, and the special function u on the strip which handles the $L^\infty \to L^2$ estimate.

A natural idea is to try to follow this path in the weighted setting. Suppose that w is an A_∞ weight. The main problem is to find an appropriate weighted analogue of the function u above; indeed, having found such an object (let us denote it by $\bar{u}$, it is a function of variables x, y, u and v), it seems plausible to put

$$B(x, y, z, u, v) = z\left(\bar{u}\left(\frac{x}{z}, \frac{y}{z}, u, v\right) - Lu\right),$$

for some constant L to be found. The function $\bar{u}$ should encode the $L^\infty(W) \to L^2(W)$ inequality, or rather $L^\infty(W) \to L^q(W)$ estimate for some q, for martingale transforms. Fortunately, some indications towards its discovery can be extracted from [64] (we also encountered similar objects in the previous chapter). Roughly speaking, to obtain $L^\infty(W) \to L^q(W)$ estimates in this context, the procedure is as follows. Take a special function U_r associated with non-maximal and unweighted $L^r \to L^r$ bound (this problem is well-understood, see Burkholder [19]) and then put

$$\bar{u}(x, y, u, v) = (U_r(x, y) + \kappa)^\beta (uv^{p-1} - a)^\alpha v^{1-p}$$

for some parameters α, β, κ and a. In the present chapter we want to take $p = \infty$, so some change is needed. It turns out that the right choice for $\bar{u}$ is

$$\bar{u}(x, y, u, v) = (U_r(x, y) + \kappa)^\beta (ue^{-v} - a)^\alpha e^v$$

(compare this to the Bellman functions used in the previous chapter).

5.4. A special function

In order to prove the inequality (5.1.3), we will first prove the weaker estimate

$$||g_n||_{L^1(w)} \leq C[w]_{A_\infty} || \, |f|_n^* ||_{L^1(w)}, \qquad n = 0, 1, \ldots. \tag{5.4.1}$$

From the previous section it is sufficient to find a function $B : \mathcal{D} \to \mathbb{R}$ which satisfies conditions 1°-3° with $G(x, y, z, u, v) = |y|u - Cczu$. Define the constants

$$\beta = \theta(8c(1-\theta))^{-1}, \qquad \alpha = 1 - (2c)^{-1}, \qquad a = 3/4, \qquad p = 1/\beta, \qquad A = 4/\theta - 1$$

and the auxiliary functions motivated by the discussion in the previous section. Observe that $p = 8c(1/\theta - 1) \geq 8$. Recall the domain $\mathcal{D}_c = \{(u, v) \in (0, \infty) \times \mathbb{R} : 1 \leq ue^{-v} \leq c\}$ and, for $(r, u, v) \in (0, \infty) \times \mathcal{D}_c$, set

$$F(r, u, v) = r^\beta (ue^{-v} - a)^\alpha e^v.$$

Furthermore, for any x, $y \in \mathbb{R}$ we define

$$U(x, y) = \begin{cases} p(1 - 1/p)^{p-1}\big(|y| - (p-1)|x|\big)\big(|x| + |y|\big)^{p-1} & \text{if } |y| \geq (p-1)|x|, \\ |y|^p - (p-1)^p |x|^p & \text{if } |y| < (p-1)|x|. \end{cases}$$

This is the celebrated special function invented by Burkholder [20] to establish sharp L^p bounds (1.1.3). An important comment is in order. In general, Bellman functions (even those leading to sharp constants) are not unique. We have already used before a simpler Bellman function associated with the L^p bound: it was given by $(x, y) \mapsto p(1-1/p)^{p-1}\big(|y| - (p-1)|x|\big)\big(|x| + |y|\big)^{p-1}$ (see the previous chapter). The reason why we need here a slightly more complicated function U (which is actually the smallest Bellman associated with the L^p bound) is the following additional property of U:

$$U(x, y) \geq U(x, 0) \qquad \text{for every } (x, y) \in \mathbb{R}^2.$$

This monotonicity will be crucial later to establish the stronger bound (5.1.3).

Burkholder proved that U enjoys the following properties.

Lemma 5.7. *Function U has the following properties*

(i) (Initial condition) We have $U(x,y)\le 0$ if $|y|\le|x|$.

(ii) (Majorization property) We have $U(x,y)\ge |y|^p-(p-1)^p|x|^p$.

(iii) (Concavity-type property) For any $(x,y)\in\mathbb{R}^2$, any $\varepsilon\in[-1,1]$, any positive integer k and any sequences $(\alpha_j)_{j=1}^k,(h_j)_{j=1}^k$ satisfying

$$\alpha_j\in[0,1),\qquad \sum_{j=1}^k\alpha_j=1\qquad \text{and}\qquad \sum_{j=1}^k\alpha_jh_j=0,$$

we have

$$U(x,y)\ge\sum_{j=1}^k\alpha_jU(x+h_j,y+\varepsilon h_j).$$

We are ready to construct a Bellman function described in the previous section. Let $B:\mathcal{D}\to\mathbb{R}$ be given by the formula

$$\begin{aligned}B(x,y,z,u,v)&=\left[F\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2(p-1)^pA^p,u,v\right)-3Apu\right]z\\&=F\left(U\left(x,y\right)+2(p-1)^pA^pz^p,u,v\right)-3Apuz.\end{aligned}$$

Here the second equality follows from homogeneity of the function U (we have $U(\lambda x,\lambda y)=|\lambda|^pU(x,y)$) and the relation $\beta=1/p$. Now we will prove certain properties of the functions U and F which will be exploited later.

Lemma 5.8. *For any $\varepsilon\in[-1,1], t\ge 0$ and $\eta\in\mathbb{R}$ we have*

$$\left(U(1,\eta)+2(p-1)^pA^p\right)^{\beta-1}\left(U(1,\eta)+2(p-1)^pA^p+\beta U_y(1,\eta)(\varepsilon-\eta)\right)\le 3Ap.$$

Proof. Recall that $\beta=1/p$. If $|\eta|<p-1$, we use the second formula in the definition of U and calculate that $U_y(1,\eta)=p\mathrm{sgn}(\eta)|\eta|^{p-1}$. Hence

$$\begin{aligned}U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta)=\varepsilon|\eta|^{p-1}\mathrm{sgn}(\eta)-(p-1)^p&\le(p-1)^{p-1}-(p-1)^p\\&\le 0\le p(1+|\eta|)^{p-1}.\end{aligned}\tag{5.4.2}$$

Next, consider the case $|\eta|\ge p-1$. We use the first formula in the definition of U and calculate that $U_y(1,\eta)=p(1-1/p)^{p-1}\mathrm{sgn}(\eta)(1+|\eta|)^{p-2}(p|\eta|+p(2-p)|x|)$. Consequently,

$$\begin{aligned}U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta)=p(1-1/p)^{p-1}(1+|\eta|)^{p-2}\Big[&(|\eta|-(p-1))(1+|\eta|)\\&+\mathrm{sgn}(\eta)\beta p(|\eta|+2-p)(\varepsilon-\eta)\Big].\end{aligned}$$

The expression in a square bracket is equal to $\varepsilon\eta-1+(\varepsilon\mathrm{sgn}(\eta)+1)(2-p)$. So if $|\eta|\ge p-1$, then

$$\begin{aligned}U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta)&=p(1-1/p)^{p-1}(1+|\eta|)^{p-2}\left(\varepsilon\eta-1+(\varepsilon\mathrm{sgn}(\eta)+1)(2-p)\right)\\&\le p(1-1/p)^{p-1}(1+|\eta|)^{p-1}\le p(1+|\eta|)^{p-1},\end{aligned}\tag{5.4.3}$$

where in the first inequality we have used the estimates $\varepsilon\eta-1\le 1+|\eta|$ and $(\varepsilon\mathrm{sgn}(\eta)+1)(2-p)\le 0$. By (5.4.2) and (5.4.3), every $\eta\in\mathbb{R}$ satisfies

$$U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta)\le p(1+|\eta|)^{p-1}.\tag{5.4.4}$$

Hence, from the above estimate and the second part of Lemma 5.7 (recall that the exponent $\beta - 1 = 1/p - 1$ is negative), it is sufficient to show that

$$\left(\eta^p - (p-1)^p + 2(p-1)^p A^p\right)^{\beta-1}\left(2(p-1)^p A^p + p(1+\eta)^{p-1}\right) \leq 3Ap \tag{5.4.5}$$

for every $\eta \geq 0$. The derivative of the expression on the left (with respect to η) equals

$$\begin{aligned} &p(\eta^p + (2A^p-1)(p-1)^p)^{\beta-2}\Big[\eta^{p-1}(\beta-1)2(p-1)^p A^p \\ &+ \eta^{p-1}p(\beta-1)(1+\eta)^{p-1} + (p-1)(1+\eta)^{p-2}(\eta^p - (p-1)^p + 2(p-1)^p A^p)\Big]. \end{aligned}$$

From the identity $p(\beta-1) = 1-p$, we obtain that the expression in the square bracket is equal to

$$2(\beta-1)(p-1)^p A^p \eta^{p-1} + (1+\eta)^{p-2}((1-p)\eta^{p-1} + (2A^p-1)(p-1)^p(p-1)).$$

Since the summand $(1+\eta)^{p-2}((1-p)\eta^{p-1} - (p-1)^{p+1})$ is negative, the expression above does not exceed $2A^p(p-1)^p((\beta-1)\eta^{p-1} + (1+\eta)^{p-2}(p-1))$. This is nonpositive for $\eta \geq 4(p-2)$. Indeed, we have

$$\begin{aligned} (1+\eta)^{p-2}(p-1) = \eta^{p-2}p(1-\beta)(1+1/\eta)^{p-2} &\leq \eta^{p-2}p(1-\beta)e^{1/4} \\ &\leq \eta^{p-2}(1-\beta)p\frac{4(p-2)}{p} \leq (1-\beta)\eta^{p-1}. \end{aligned}$$

Here in the first and third inequality we have exploited the assumption $\eta \geq 4(p-2)$, and the second inequality follows from the bound $p \geq 8$. We proved that the expression on the left-hand side of (5.4.5) is decreasing for $\eta \geq 4(p-2)$. Hence to establish (5.4.5), it is enough to prove this for $\eta \in [0, 4(p-2))$. We estimate the left-hand side of this inequality from above by

$$\begin{aligned} &((2A^p-1)(p-1)^p)^{\beta-1}(2(p-1)^p A^p + p(4p-7)^{p-1}) \\ &= A(2-A^{-p})^{\beta-1}\left(2(p-1) + p\left(\frac{4p-7}{Ap-A}\right)^{p-1} A^{-1}\right) \\ &\leq A(2-4^{-p})^{\beta-1}(2(p-1) + pA^{-1}) \leq 3Ap. \end{aligned}$$

Here in the equality we used the identity $p(\beta-1) = 1-p$, the first inequality is due to the bound $A \geq 4$, while the last step follows from $2 - 4^{-p} \geq 1$ and the estimation of the expression in a second bracket (from above) by $3p$. This completes the proof. □

Concerning F, we start with the following fact, which we have already proved in the previous chapter (see Lemma 4.7).

Lemma 5.9. *For any $(r,u,v) \in (0,\infty) \times \mathcal{D}_c$ with $1 \leq ue^{-v} \leq c$, we have*

$$\frac{1}{4}ur^\beta \leq F(r,u,v) \leq ur^\beta. \tag{5.4.6}$$

We will also need a certain concavity-type property of F. In Lemma 4.8 we have established the *local* concavity of F, however it is not sufficient here. Because of the discrete-time, θ-regular setting, we need the following stronger condition.

Lemma 5.10. *Function F is θ-concave. That is, for any x, x_1, x_2, ..., $x_n \in (0,\infty) \times \mathcal{D}_c$ and any sequence α_1, α_2, ..., α_n of numbers in $[\theta, 1)$ satisfying $\sum_{j=1}^n \alpha_j = 1$ and $\sum_{j=1}^n \alpha_j x_j = x$, we have*

$$F(x) \geq \sum_{j=1}^n \alpha_j F(x_j).$$

Proof. Observe that by the homogeneity of the function F we may and do assume that $v = 0$. In other words, it suffices to prove the inequality

$$r^\beta (u-a)^\alpha \geq \sum_{j=1}^n \alpha_j r_j^\beta \left(u_j e^{-v_j} - a\right)^\alpha e^{v_j}, \tag{5.4.7}$$

where $\sum_{j=1}^n \alpha_j (r_j, u_j, v_j) = (r, u, 0)$ and $(r,u,0), (r_1, u_1, v_1), \ldots (r_n, u_n, v_n) \in (0,\infty) \times \mathcal{D}_c$. Because $\alpha + \beta + \beta < 1$, the function $(0,\infty) \times (1,\infty) \times (0,\infty) \ni (k,s,t) \mapsto k^\beta (s-a)^\alpha t^\beta$ is concave. Hence, we obtain that

$$\sum_{j=1}^n \left(r_j e^{-v_j}\right)^\beta \left(u_j e^{-v_j} - a\right)^\alpha \left(e^{v_j}\right)^\beta \frac{e^{v_j}\alpha_j}{P} \leq \left(\frac{r}{P}\right)^\beta \left(\frac{u}{P} - a\right)^\alpha \left(\frac{Q}{P}\right)^\beta,$$

where $P = \sum_{j=1}^n \alpha_j e^{v_j}$ and $Q = \sum_{j=1}^n \alpha_j e^{2v_j}$. Thus, to prove (5.4.7), it is sufficient to establish the inequality

$$P^{1-\beta} \left(\frac{u}{P} - a\right)^\alpha \left(\frac{Q}{P}\right)^\beta \leq (u-a)^\alpha. \tag{5.4.8}$$

We will need the following estimate involving expressions P and Q:

$$Q \leq \frac{1}{\theta} P^2 - \frac{1-\theta}{\theta}.$$

This follows from the assumption $\alpha_j \in [\theta, 1)$, by applying twice the convexity of the function e^x. Indeed, we have that

$$P^2 - \theta Q = \sum_{j=1}^n \alpha_j e^{v_j} \left((\alpha_j - \theta) e^{v_j} + \sum_{k \neq j} \alpha_k e^{v_k} \right) \geq \sum_{j=1}^n \alpha_j e^{v_j} (1-\theta) e^{-v_j \theta/(1-\theta)} \geq 1 - \theta.$$

Hence, to prove (5.4.8) it is enough to show that

$$P \left(\frac{u}{P} - a\right)^\alpha \left(\frac{1}{\theta} - \frac{1-\theta}{\theta P^2}\right)^\beta \leq (u-a)^\alpha.$$

Moreover, we know that $u \in [1, c]$ and $P \in [1, u]$ (here the lower bound is just the convexity of e^x, while the upper bound follows from the conditions $u_j e^{-v_j} \geq 1$ for each j). Let us denote $s = 1/P$. It is enough to show that the inequality

$$s^{-1} (us - a)^\alpha \left(1 - (1-\theta)s^2\right)^\beta \leq \theta^\beta (u-a)^\alpha$$

holds for any $u \in [1, c]$ and $s \in [1/u, 1]$. Observe that for $s = 1$ both sides are equal. Hence, it is sufficient to show that the function $s \mapsto s^{-1}(us-a)^\alpha \left(1 - (1-\theta)s^2\right)^\beta$ is nondecreasing. Differentiating, we see that we must prove that

$$((\alpha - 1)us + a)\left(1 - (1-\theta)s^2\right) - 2\beta(1-\theta)s^2(us - a) \geq 0.$$

But $s \leq 1$, so the expression on the left is greater than

$$\begin{aligned} &((\alpha-1)us+a)\theta - 2\beta(1-\theta)(us-a) \\ &= \left(\alpha - 1 - 2\beta\frac{1-\theta}{\theta}\right)\theta us + \left(1+2\beta\frac{1-\theta}{\theta}\right)\theta a \geq \left(\alpha-1-2\beta\frac{1-\theta}{\theta}\right)\theta c + \theta a = 0. \end{aligned}$$

Here in the last step we just plugged values of the parameters: $a = 3/4$, $\alpha = 1 - 1/2c$ and $\beta = \theta(8c(1-\theta)^{-1}$. This completes the proof. □

Two comments are in order.

Remark 5.3. The regularity assumption $\alpha_j \geq \theta$ is necessary here. In other words, the function F does not satisfy the concavity condition if we do not assume any lower bound on α_j's. Indeed, fix $\theta \in (0, 1/2]$ and consider the points $P = (P_u, P_u, P_v)$, $Q = (c, c, 0)$ and $R = (R_u, R_u, R_v)$ in $(0, \infty) \times \mathcal{D}_c$ satisfying the equations

$$P_u e^{-P_v} = 1, \qquad R_u e^{-R_v} = c \quad \text{and} \quad Q = (1-\theta)P + \theta R.$$

Now assume that the function F is θ-concave for every $\theta \in (0, 1/2]$. We have that $F(Q) \geq (1-\theta)F(P) + \theta F(R)$, or equivalently

$$c^{1+\beta}\frac{(c-a)^\alpha}{c} \geq (1-\theta)P_u^{1+\beta}\frac{(1-a)^\alpha}{1} + \theta R_u^{1+\beta}\frac{(c-a)^\alpha}{c}. \tag{5.4.9}$$

Now observe that this inequality cannot hold for small θ. We have that $\lim_{\theta\to 0} P_u = 1$, $\lim_{\theta\to 0} R_u = +\infty$ and $\lim_{\theta \to 0} \theta R_u = c - 1$. Hence, the left-hand side of (5.4.9) is constant but the right-hand side goes to ∞ when θ goes to 0. This shows that the function F cannot be θ-concave for arbitrary small θ. This is one of the reasons why the regularity condition of the underlying filtration is crucial (and cannot be removed).

Remark 5.4. In order to establish θ-concavity of F we have used the direct approach. The alternative method would be to show first local concavity of F (this is easy - see Lemma 4.8) and then obtain θ-concavity via the geometric argument similar to that in Lemma 2.10. However this idea cannot work, because there is no substitute for the latter lemma for A_∞ weights.

We are ready for the main step: we will check that the function B satisfies conditions 1°, 2° and 3°.

Lemma 5.11. *The function B satisfies the initial condition 1°.*

Proof. Recall the initial condition 1°: for every $(x, y, |x|, u, v) \in \mathcal{D}$ such that $|y| \leq |x|$ and $1 \leq ue^{-v} \leq c$ we have $B(x, y, |x|, u, v) \leq 0$. By the definition of B, this is equivalent to showing the estimate

$$\left[F\left(U\left(\frac{x}{|x|}, \frac{y}{|x|}\right) + 2(p-1)^p A^p, u, v\right) - 3Apu\right]|x| \leq 0. \tag{5.4.10}$$

Using Lemma 4.7, we get that

$$\frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leq 1.$$

Recall the key identity $p\beta = 1$. From the condition (i) in Lemma 5.7, if $|y| \leq |x|$, then $U(x/|x|, y/|x|) \leq 0$ and hence

$$\left(U\left(\frac{x}{|x|}, \frac{y}{|x|}\right) + 2(p-1)^p A^p\right)^\beta \frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leq 2A(p-1) \leq 3Ap,$$

which is precisely the required estimate (5.4.10). □

Lemma 5.12. *The function B satisfies the majorization condition*

$$B(x,y,z,u,v) \geq \frac{1}{4}(|y|u - 12Apzu).$$

Proof. From the second part of Lemma 5.7, the estimate $|x|/z \leq 1 \leq A$ and the identity $p\beta = 1$, we obtain

$$\left(U\left(\frac{x}{z},\frac{y}{z}\right) + 2(p-1)^p A^p\right)^\beta \geq \left(\left(\frac{|y|}{z}\right)^p - (p-1)^p\left(\frac{|x|}{z}\right)^p + 2(p-1)^p A^p\right)^\beta \geq \frac{|y|}{z}.$$

In addition, we have proved in Lemma 4.7 that

$$\frac{(ue^{-v} - a)^\alpha}{ue^{-v}} \geq \frac{1}{4}. \tag{5.4.11}$$

Putting these facts together, we see that

$$B(x,y,z,u,v) \geq \frac{1}{4}|y|u - 3Apuz = \frac{1}{4}(|y|u - 12Apzu). \qquad \square$$

It remains to check the most difficult condition 3°. For appropriate numbers x, y, $z, \ldots$, we have

$$B(x,y,z,u,v) \geq \sum_{j=1}^{k} \alpha_j B(x+h_j, y+\varepsilon h_j, |x+h_j| \vee z, u+r_j, v+s_j). \tag{5.4.12}$$

We have already shown that auxiliary functions U and F have appropriate concavity properties formulated in Lemmas 5.7 and 5.10. These properties imply that B satisfies (5.4.12) for points satisfying the additional condition $|x+h_j| \leq z$. The main difficulty is to remove this assumption. To handle this problem, we consider the extension of B on the domain

$$\bar{\mathcal{D}} = \{(x,y,z,u,v) \in \mathbb{R}^2 \times (0,\infty) \times (0,\infty) \times \mathbb{R} : |x| \leq Az, 1 \leq ue^{-v} \leq c\},$$

which is given by the same formula as for the initial B:

$$\bar{B}(x,y,z,u,v) = \left[F\left(U\left(\frac{x}{z},\frac{y}{z}\right) + 2A^p(p-1)^p, u, v\right) - 3Apu\right] z.$$

In the next theorem we will prove the concavity and monotonicity properties of $\bar{B}$.

Theorem 5.13. *The function $\bar{B}$ has the following properties:*

1) (Concavity-type property) For any $(x,y,z,u,v) \in \bar{\mathcal{D}}$, any $\varepsilon \in [-1,1]$, any positive integer $k \leq 1/\theta$ and sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying

$$\alpha_j \in [\theta, 1) \qquad \textit{and} \qquad \sum_{j=1}^{k} \alpha_j = 1,$$

$$\sum_{j=1}^{k} \alpha_j h_j = \sum_{j=1}^{k} \alpha_j r_j = \sum_{j=1}^{k} \alpha_j s_j = 0$$

and

$$(x+h_j, y+\varepsilon h_j, z, u+r_j, v+s_j) \in \bar{\mathcal{D}},$$

we have

$$\bar{B}(x,y,z,u,v) \geq \sum_{j=1}^{k} \alpha_j \bar{B}(x+h_j, y+\varepsilon h_j, z, u+r_j, v+s_j). \tag{5.4.13}$$

2) (Vertical monotonicity) We have $\bar{B}_z(x,y,z,u,v) \leq 0$ for every $(x,y,z,u,v) \in \bar{\mathcal{D}}$.

3) (Diagonal monotonicity) Let $(x_1,y_1,|x_1|,u,v),(x_2,y_2,|x_2|,u,v) \in \mathcal{D}$. If $|x_2| < |x_1|$ and $|y_2-y_1| \leq |x_2-x_1|$, then $B(x_1,y_1,|x_1|,u,v) \leq B(x_2,y_2,|x_2|,u,v)$.

Proof. The first part of the theorem follows immediately from Lemma 5.7, Lemma 5.10 and the monotonicity $F_r \geq 0$. Indeed, we have:

$$\begin{aligned} &F\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2A^p(p-1)^p,u,v\right)\\ &\geq F\left(\sum_{j=1}^{k}\alpha_j U\left(\frac{x+h_j}{z},\frac{y+\varepsilon h_j}{z}\right)+2A^p(p-1)^p, u+\sum_{j=1}^{k}\alpha_j r_j, v+\sum_{j=1}^{k}\alpha_j s_j\right)\\ &\geq \sum_{j=1}^{k}\alpha_j F\left(U\left(\frac{x+h_j}{z},\frac{y+\varepsilon h_j}{z}\right)+2A^p(p-1)^p, u+r_j, v+s_j\right), \end{aligned}$$

which is equivalent to the desired inequality. It remains to show the monotonicity properties. We start with 2). By symmetry, we may assume that $x \geq 0$. Because $\beta = 1/p$, the condition $\bar{B}_z(x,y,z,u,v) \leq 0$ is equivalent to the inequality

$$\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2A^p(p-1)^p\right)^{\beta-1} 2A^p(p-1)^p \frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leq 3Ap.$$

From the second part of Lemma 5.7 and Lemma 4.7, the left hand side is smaller than

$$\left(\left(\frac{|y|}{|z|}\right)^p - (p-1)^p\left(\frac{|x|}{|z|}\right)^p + 2A^p(p-1)^p\right)^{\beta-1} 2A^p(p-1)^p \leq 2Ap.$$

This gives the assertion. To handle 3), we first apply the symmetry and homogeneity to assume that $x_1 > 0$ and $x_2 = 1$. Consider the function $\phi : [0,\infty) \mapsto \mathbb{R}$ given by the formula

$$\phi(t) = B(1+t, y+\varepsilon t, 1+t, u, v),$$

where $\varepsilon \in [-1,1]$ and $(u,v) \in \mathcal{D}_c$ are fixed. It is sufficient to show that $\phi'(t) \leq 0$. This is equivalent to proving that the expression

$$\begin{aligned} (U(1,\eta)+2(p-1)^pA^p)^{\beta-1}\left(U(1,\eta)+2(p-1)^pA^p+\beta U_y(1,\eta)(\varepsilon-\eta)\right)\frac{(ue^{-v}-a)^\alpha}{ue^{-v}}u\\ -3Apu, \end{aligned}$$

where $\eta = (y+\varepsilon t)/(1+t)$, is nonpositive. This follows from Lemma 5.8 and Lemma 4.7. This completes the proof of the theorem. □

We are ready to prove that the function B satisfies the concavity-type condition.

Proof of (5.4.12). The function B satisfies the homogeneity condition $B(\lambda x, \lambda y, \lambda z, u, v) = \lambda B(x,y,z,u,v)$ for every $\lambda > 0$. Hence, we can divide both sides of (5.4.12) by z to obtain the equivalent form

$$B(x/z,y/z,1,u,v) \geq \sum_{j=1}^{k}\alpha_j B(x/z+h_j/z, y/z+\varepsilon h_j/z, |x/z+h_j/z| \vee 1, u+r_j, v+s_j).$$

Thus, to prove the general case, it is sufficient to prove the statement for $z = 1$.

For convenience let us denote $x_j = x + h_j$, $y_j = y + \varepsilon h_j$, $u_j = u + r_j$ and $v_j = v + s_j$ and sort the points in the increasing order, that is: $x_1 \le x_2 \le \ldots \le x_k$. We will divide the proof into two steps.

Step 1. Let us consider two special cases: when $x_1 \ge -1$ or when $x_k \le 1$. From symmetry it is enough to solve only the first one, the second is analogous. From $x_1 \ge -1$ and the lower bound on probabilities, we can deduce that x_k cannot be large. More precisely:

$$x_k = \left(x - \sum_{j=1}^{k-1} \alpha_j x_j \right) / \alpha_k \le (x + 1 - \alpha_k)/\alpha_k = (x+1)/\alpha_k - 1 \le 2/\theta - 1 \le A.$$

Hence $(x_j, y_j, 1, u_j, v_j) \in \bar{\mathcal{D}}$ and from the second part of Theorem 5.13 we obtain that

$$\bar{B}(x_j, y_j, |x_j| \vee 1, u_j, v_j) \le \bar{B}(x_j, y_j, 1, u_j, v_j).$$

Combining this with the first part of Theorem 5.13, we get

$$\begin{aligned} \sum_{j=1}^{k} \alpha_j B(x_j, y_j, |x_j| \vee 1, u_j, v_j) \le \sum_{j=1}^{k} \alpha_j \bar{B}(x_j, y_j, 1, u_j, v_j) &\le \bar{B}(x, y, 1, u, v) \\ &= B(x, y, 1, u, v). \end{aligned}$$

Step 2. In this step we will reduce the general case to the one considered before. Assume that $x_1 < -1$ and $x_k > 1$. The idea is to replace x_1, y_1, x_k, y_k by $\hat{x}_1, \hat{y}_1, \hat{x}_k, \hat{y}_k$ in such a way that:

1° We "pull" the points closer to the center: $\hat{x}_1 \in (x_1, -1]$ and $\hat{x}_k \in [1, x_k)$.
2° We have that $\hat{y}_1 - y_1 = \varepsilon(\hat{x}_1 - x_1)$ and $\hat{y}_k - y_k = \varepsilon(\hat{x}_k - x_k)$.
3° The average is preserved: $\alpha_1 x_1 + \alpha_k x_k = \alpha_1 \hat{x}_1 + \alpha_k \hat{x}_k$.

Then in the light of the third part of Theorem 5.13, we have $B(x_1, y_1, |x_1|, u_1, v_1) \le B(\hat{x}_1, \hat{y}_1, |\hat{x}_1|, u_1, v_1)$ and $B(x_k, y_k, |x_k|, u_k, v_k) \le B(\hat{x}_k, \hat{y}_k, |\hat{x}_k|, u_k, v_k)$. Hence the replacement does not change the left hand side of (5.4.12) and does not decrease the right hand side making the inequality stronger. Moreover we will also ensure that

4° We "pull" the points as close as possible: $\hat{x}_1 = -1$ or $\hat{x}_k = 1$.

Now we repeat the replacement procedure until all the first coordinates $x_1, \ldots, x_k$ are contained either in the set $[-1, \infty)$ or in the set $(-\infty, 1]$, which is the case solved in Step 1. The condition 4° ensures that this algorithm will stop after at most $n - 1$ replacements. It remains to find the points $\hat{x}_1, \hat{y}_1, \hat{x}_k, \hat{y}_k$ satisfying conditions 1°-4°. This will be done explicitly. Let us consider two cases. If $\alpha_1 x_1 + \alpha_k x_k \ge \alpha_k - \alpha_1$, then we put $\hat{x}_1 = -1$, $\hat{x}_k = (\alpha_1 x_1 + \alpha_k x_k + \alpha_1)/\alpha_k$, $\hat{y}_1 = \varepsilon(\hat{x}_1 - x_1) + y_1$ and $\hat{y}_k = \varepsilon(\hat{x}_k - x_k) + y_k$. Conditions 1°-4° are easy to check. The case $\alpha_1 x_1 + \alpha_k x_k < \alpha_k - \alpha_1$, is analogous. We put $\hat{x}_k = 1$, $\hat{x}_1 = (\alpha_1 x_1 + \alpha_k x_k - \alpha_k)/\alpha_1$, $\hat{y}_1 = \varepsilon(\hat{x}_1 - x_1) + y_1$ and $\hat{y}_k = \varepsilon(\hat{x}_k - x_k) + y_k$. Again it is easy to check the required conditions. This completes the proof. □

We have shown that B satisfies the conditions 1°-3°. By the method of Section 2, this yields the estimate (5.4.1) with $C = 12Apc^{-1} \le 384\theta^{-2}$.

5.5. Proof of the main inequality

To prove the main inequality (5.1.3) we will construct the function of *six* variables.

$$\mathfrak{D} = \{(x, y, z, r, u, v) \in \mathbb{R}^2 \times (0, \infty) \times \mathbb{R} \times (0, \infty) \times \mathbb{R} : |x| \le z, y \le r, 1 \le ue^{-v} \le c\}.$$

The additional variable r is associated with the one-sided maximal function defined as $g_n^* = \sup_{n\ge 0} g_n$. We define Burkholder's function $\mathfrak{B} : \mathfrak{D} \to \mathbb{R}$ by

$$\begin{aligned}\mathfrak{B}(x, y, z, r, u, v) &= \left[F\left(U\left(\frac{x}{z}, \frac{r-y}{z}\right) + 2A^p(p-1)^p, u, v\right) - 12cu\right] z\\ &= B(x, r-y, z, u, v).\end{aligned}$$

This new function satisfies the following properties:

1° (Initial condition) We have that $\mathfrak{B}(x, y, |x|, y, u, v) \le 0$ if $1 \le ue^{-v} \le c$.

2° (Majorization property) For any $(x, y, z, r, u, v) \in \mathfrak{D}$ we have

$$\mathfrak{B}(x, y, z, r, u, v) \ge \frac{1}{4}((r-y)u - 12Apzu).$$

3° (Concavity-type property) For any $(x, y, z, r, u, v) \in \mathfrak{D}$, any $\varepsilon \in [-1, 1]$, any positive integer $k \le 1/\theta$ and sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying

$$\alpha_j \in [\theta, 1) \qquad \text{and} \qquad \sum_{j=1}^k \alpha_j = 1,$$

$$\sum_{j=1}^k \alpha_j h_j = \sum_{j=1}^k \alpha_j r_j = \sum_{j=1}^k \alpha_j s_j = 0$$

and

$$(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, (y + \varepsilon h_j) \vee r, u + r_j, v + s_j) \in \mathfrak{D},$$

we have

$$\mathfrak{B}(x, y, z, r, u, v) \ge \sum_{j=1}^k \alpha_j \mathfrak{B}(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, (y + \varepsilon h_j) \vee r, u + r_j, v + s_j).$$

Conditions 1° and 2° are immediate consequences of analogous properties of B. Now consider the concavity-type condition. It is easy to check that Burkholder's function U has the following property: $U(x, y) \ge U(x, 0)$ for every $(x, y) \in \mathbb{R}^2$. Hence

$$B(x, y, z, u, v) \ge B(x, 0, z, u, v),$$

for every $(x, y, z, u, v) \in \mathcal{D}$. From the above estimate and the inequality (5.4.12) we have

$$\begin{aligned}&\sum_{j=1}^k \alpha_j \mathfrak{B}(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, (y + \varepsilon h_j) \vee r, u + r_j, v + s_j)\\ &= \sum_{j=1}^k \alpha_j B(x + h_j, (r - y - \varepsilon h_j) \vee 0, |x + h_j| \vee z, u + r_j, v + s_j)\\ &\le \sum_{j=1}^k \alpha_j B(x + h_j, r - y - \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j)\\ &\le B(x, r - y, z, u, v) = \mathfrak{B}(x, y, z, r, u, v).\end{aligned}$$

Now we repeat, word-by-word, the reasoning of Section 2: the only change is that the process $(z_n)_{n\geq 0}$ is six-dimensional and involves the one-sided maximal function of g: $z_n = (f_n, g_n, |f|_n^*, g_n^*, w_n, \sigma_n)$. Hence, we obtain

$$\mathbb{E}\left(g_n^* - g_n\right) w_n \leq 12Ap[w]_{A_\infty}\mathbb{E}|f|_n^* w_n \leq 384\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n$$

and, by symmetry, $\mathbb{E}\left((-g)_n^* + g_n\right) w_n \leq 384\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n$. Add these two bounds to get

$$\mathbb{E}(g_n^* + (-g)_n^*)w_n \leq 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n. \tag{5.5.1}$$

Now observe that if g started from 0, we would have the pointwise inequality $|g|_n^* \leq g_n^* + (-g)_n^*$ and (5.5.1) would give

$$\mathbb{E}|g|_n^* w_n \leq 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n,$$

as desired (see the limiting argument below). To prove (5.1.3) in full generality, note that if $(g_n)_{n\geq 0}$ is a ± 1-transform of f, then the martingale $\tilde{g} = (g_n - g_0)_{n\geq 0}$ also has this property and additionally starts from 0. Hence, by the above estimate,

$$\mathbb{E}|g|_n^* w_n \leq \mathbb{E}(|\tilde{g}|_n^* + |g_0|)w_n \leq 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n + \mathbb{E}|f_0|w_n \leq 769\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n.$$

Since $w_n = \mathbb{E}(w|\mathcal{F}_n)$, this gives $\mathbb{E}|g|_n^* w \leq 769\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w$ and the claim follows by letting $n \to \infty$ and applying Lebesgue's monotone convergence theorem.

The function $\mathfrak{B}$ can be used to establish the analogous inequality for continuous-path martingales. We conclude this section with the proof of Theorem 5.4.

Proof of Theorem 5.4. Let $\mathfrak{B} : \mathfrak{D} \to \mathbb{R}$ be the Bellman function constructed above for $\theta = 1/2$. We showed that $\mathfrak{B}$ satisfies conditions 1°, 2° and the following midpoint concavity: for appropriate numbers $x, y, z, \ldots$, we have

$$\begin{aligned}\mathfrak{B}(x, y, z, r, u, v) \geq\ & \mathfrak{B}(x + h, y + \varepsilon h, |x + h| \vee z, (y + \varepsilon h) \vee r, u + r_1, v + s)/2 \\ & + \mathfrak{B}(x - h, y - \varepsilon h, |x - h| \vee z, (y - \varepsilon h) \vee r, u - r_1, v - s)/2.\end{aligned} \tag{5.5.2}$$

Once again, the argument rests on Itô's formula applied to $\mathfrak{B}(X_t, Y_t, |X|_t^*, Y_t^*, W_t, V_t)$. Clearly, the above midpoint concavity condition is stronger than local concavity. To handle the additional integral $\int_{0+}^t \mathfrak{B}_z(X_s, Y_s, |X|_s^*, Y_s^*, W_s, V_s)\mathrm{d}|X|_s^*$ we need to show that $\mathfrak{B}_z(z, y, z, r, u, v) \leq 0$ for $z > 0$ and $1 \leq ue^{-v} \leq c$. Observe that this condition is also contained in (5.5.2). Indeed, fix (z, y, z, r, u, v), $r_1 = s = \varepsilon = 0$ and some small $h > 0$. We have:

$$\mathfrak{B}(z, y, z, r, u, v) - \mathfrak{B}(z + h, y, z + h, r, u, v)/2 - \mathfrak{B}(z - h, y, z, r, u, v)/2 \leq 0.$$

It remains to divide both sides by h and pass $h \to 0$ to obtain precisely the required estimate $\mathfrak{B}_z(z, y, z, r, u, v) \leq 0$. Similarly we show that the additional stochastic integral $\int_{0+}^t \mathfrak{B}_r(X_s, Y_s, |X|_s^*, Y_s^*, W_s, V_s)\mathrm{d}Y_s^*$ is nonpositive. The problem with lack of C^2 condition of the function $\mathfrak{B}$ is handled as before - see Remark 2.1. Hence we obtain the Theorem 5.4 with a constant $769\theta^{-2} = 3076$. □

5.6. Necessity of the θ-regularity condition

The purpose of this section is to establish Theorem 5.3, and from now on we work with dyadic filtrations only. We could prove the theorem by constructing appropriate examples, but these seem to have quite involved, fractal-type structure and their analysis is a little complicated. Our approach will rest on Remark 5.2, which enables us to avoid most of these technical issues. Roughly speaking, the argument is as follows. First we assume, on contrary, that the inequality does hold universally, i.e., with the constant independent of the dimension. Then the Bellman method yields the existence of an abstract function satisfying the appropriate size and concavity requirements. Finally, we exploit these properties in the right order to obtain a contradiction (with the assumption that the constant involved is dimension-free). So, suppose that there is $1 < p < \infty$ and a constant K depending only on p, such that for any dimension d, any martingales f and g adapted to the d-dimensional dyadic filtration on $[0,1)^d$ such that $dg_n = v_n dg_n$ for predictable sequence of signs v_n, and any A_p weight w on $[0,1)^d$ with $[w]_{A_p} \leq 2$, we have

$$||g||_{L^1(w)} \leq K|||f|^*||_{L^1(w)}. \tag{5.6.1}$$

Fix d and let B be the associated Bellman function, given by

$$B(x,y,z,u,v) = \sup \mathbb{E}\Big\{|g_n|w - K(|f_n|^* \vee z)w\Big\}.$$

Here the probability space is equal to $([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|)$, the filtration is dyadic and the above supremum is taken over:

· all adapted martingale pairs (f,g) satisfying $(f_0, g_0) = (x,y)$ and $dg_k = v_k df_k$ for all $k \geq 1$, for some deterministic sequence v_1, v_2, ... of signs.

· all dyadic A_p weights w satisfying $[w]_{A_p} \leq 2$, $\mathbb{E}w = u$ and $\mathbb{E}w^{1/(1-p)} = v$.

This Bellman function enjoys the appropriate initial, majorization and concavity conditions, proved in Section 2. We will also need the following additional properties which follow from the special form of the function G.

Theorem 5.14. *(i) We have*

$$B(x,y,z,u,v) = B(|x|,|y|,|x| \vee z,u,v). \tag{5.6.2}$$

(ii) For any $\lambda \neq 0$ and any $\mu > 0$ we have

$$B(\lambda x, \lambda y, |\lambda| z, \mu u, \mu^{-1/(p-1)} v) = |\lambda| \mu B(x,y,z,u,v). \tag{5.6.3}$$

(iii) We have

$$B(x,y,z,u,v) \geq B(x,0,z,u,v). \tag{5.6.4}$$

Proof. The symmetry $B(x,y,z,u,v) = B(|x|,|y|,z,u,v)$ follows directly from the definition. Indeed, if f, g, w are arbitrary martingales as in the definition of $B(x,y,z,u,v)$, then $-f, g, w$ satisfies all the requirements needed in the definition of $B(-x,y,z,u,v)$, so

$$B(-x,y,z,u,v) \geq \mathbb{E}\Big\{|-g_n|w - K(|f_n|^* \vee z)w\Big\} = \mathbb{E}\Big\{|-g_n|w - K(|f_n|^* \vee z)w\Big\}.$$

Taking the supremum over all f, g and w we get $B(-x,y,z,u,v) \geq B(x,y,z,u,v)$, and the passage from x to $-x$ shows that we actually have equality here. The identities $B(x,y,z,u,v) = B(x,-y,z,u,v)$ is shown in the same manner, and the equality $B(|x|,|y|,z,u,v) = B(|x|,|y|,|x| \vee z,u,v)$ follows from the fact that

$$\mathbb{E}\Big\{|g_n|w - K(|f_n|^* \vee z)w\Big\} = \mathbb{E}\Big\{|g_n|w - K(|f_n|^* \vee |f_0| \vee z)w\Big\}.$$

The proof of the homogeneity property (ii) is analogous: pick arbitrary martingales f, g, w as in the definition of $B(x,y,z,u,v)$. Then λf has the average λx, λg has the average λy, while μw is an A_p weight with the characteristics bounded by 2 satisfying $\mathbb{E}\mu w = \mu u$ and $\mathbb{E}(\mu w)^{-1/(p-1)} = \mu^{-1/(p-1)}v$. Consequently,

$$\begin{aligned} B(\lambda x, \lambda y, |\lambda| z, \mu u, \mu^{-1/(p-1)}v) &\geq \mathbb{E}\Big\{|\lambda g_n|(\mu w) - K(|\lambda f_n|^* \vee |\lambda z|)(\mu w)\Big\} \\ &= |\lambda|\mu\mathbb{E}\Big\{|g_n|w - K(|f_n|^* \vee z)w\Big\}. \end{aligned}$$

Hence, taking the supremum over all f, g and w as above, we get

$$B(\lambda x, \lambda y, |\lambda| z, \mu u, \mu^{-1/(p-1)}v) \geq |\lambda|\mu B(x,y,z,u,v). \tag{5.6.5}$$

To get the reverse bound, apply the above estimate to the point $(\lambda x, \lambda y, |\lambda|z, \mu u, \mu^{-1/(p-1)}v)$ in the place of (x,y,z,u,v) and the numbers λ^{-1}, μ^{-1} in the place of λ and μ.

Finally, to check (iii), we will prove that the function $y \mapsto B(x,y,z,u,v)$ is convex; together with its symmetry (which is guaranteed by (i)), we will get the claim. Pick $\alpha \in (0,1)$, two real numbers y_1, y_2 and set $y = \alpha_1 y_1 + (1-\alpha_1)y_2$. If f, g, w are martingales as in the definition of $B(x,y,z,u,v)$, then the convexity of the function $t \mapsto |t|$ yields

$$\begin{aligned} &\mathbb{E}\Big\{|g_n|w - K(|f_n|^* \vee z)w\Big\} \\ &\leq \alpha_1\mathbb{E}\Big\{|y_1 - y + g_n|w - K(|f_n|^* \vee z)w\Big\} + \alpha_2\mathbb{E}\Big\{|y_2 - y + g_n|w - K(|f_n|^* \vee z)w\Big\} \\ &\leq \alpha_1 B(x,y_1,z,u,v) + \alpha_2 B(x,y_2,z,u,v). \end{aligned}$$

Therefore, taking the supremum over all f, g, w and n gives the desired convexity. □

We will exploit the concavity of B in appropriate directions; to this end, we need the following auxiliary geometrical fact, taken from [66]. We provide an easy proof for the sake of completeness.

Lemma 5.15. *Suppose that N is a huge positive integer, $u = 1$ and $v = 2^{1/(p-1)}$. Then there are two points $R, T \in \mathbb{R}^2$ such that $R = (R_x, R_y)$ lies on the curve $xy^{p-1} = 2$, $T = (T_x, T_y)$ lies on the curve $xy^{p-1} = 1$, $R_x \leq T_x$ and*

$$(1 - (1-2^{-d})^N)R + (1-2^{-d})^N T = (u,v). \tag{5.6.6}$$

Furthermore,

$$(1 - (1-2^{-d})^N)2^d R_x < 1/2 \tag{5.6.7}$$

provided d is sufficiently large.

Proof. The existence of the points R, T follows from a very simple continuity argument. Pick any point $R = (R_x, R_y)$ on the curve $xy^{p-1} = 2$, such that $R_x \leq u$ and let T be defined by the condition (5.6.6) (then of course $R_x \leq u \leq T_x$). Note that T is a continuous function of R. Furthermore, if R_y is huge, then T_y is negative, so T lies below the curve $xy^{p-1} = 1$. On the other hand, when $R_y = v$, then $R = T = (u,v)$, so T lies above the curve $xy^{p-1} = 1$. Thus, by Darboux property, there must be a point R for which the desired configuration is satisfied.

To show (5.6.7), we exploit (5.6.6). Recall that $u = 1$. We have

$$1 = (1 - (1-2^{-d})^N)R_x + (1-2^{-d})^N T_x$$

and since $R_x < 1 < T_x$,

$$2^{1/(p-1)} = (1-(1-2^{-d})^N)\left(\frac{2}{R_x}\right)^{1/(p-1)} + (1-2^{-d})^N T_x^{-1/(p-1)}$$
$$< (1-(1-2^{-d})^N)\left(\frac{2}{R_x}\right)^{1/(p-1)} + (1-2^{-d})^N,$$

which implies

$$R_x < \left(\frac{1-(1-2^{-d})^N}{1-(1-2^{-d})^N/2^{1/(p-1)}}\right)^{p-1}.$$

Thus, if $d \to \infty$, then $R_x \to 0$; on the other hand we have $(1-(1-2^{-d})^N)2^d \leq N$ for each d. This proves the assertion. □

Let u, v, R and T be as in (ii) above. In what follows, we will also exploit the points $T_0, T_1, \ldots, T_N$ given by $T_0 = (u, v)$ and the recursive equation

$$T_k = 2^{-d}R + (1-2^{-d})T_{k+1}. \tag{5.6.8}$$

By straightforward induction, we see that $(u,v) = (1-2^{-d})^k T_k + (1-(1-2^{-d})^k)R$ for each k and hence in particular $T_N = T$.

Proof of Theorem 5.3. We will sometimes use the following notation: if $x \in \mathbb{R}$, $y \in \mathbb{R}$, $z \geq 0$ and $P = (u, v) \in \mathcal{D}$, we will write $B(x, y, z; P) = B(x, y, z, u, v)$. Let $\bar{x} = 1/(2^{d+1}-1)$. As we have shown in Section 2 (see Remark 5.2), the function B satisfies the initial condition 1°: for every $|y| \leq |x|$ we have $B(x, y, |x|, u, v) \leq 0$. Now observe that this condition combined with Theorem 5.14 (iii) gives

$$0 \geq B\left(1,1,1,1,2^{1/(p-1)}\right) \geq B\left(1,0,1,1,2^{1/(p-1)}\right). \tag{5.6.9}$$

Next, the concavity property combined with (5.6.8) yields, for each k,

$$B(\bar{x}, 2k\bar{x}, \bar{x}; T_k) \geq 2^{-d}B(1, (2k+1)\bar{x}-1, \bar{x}; R) + (1-2^{-d})B(-\bar{x}, 2(k+1)\bar{x}, \bar{x}; T_{k+1}).$$

By part (i) of Theorem 5.14, this expression is equal to

$$2^{-d}B(1,(2k+1)\bar{x}-1,1;R) + (1-2^{-d})B(\bar{x}, 2\bar{x}(k+1), \bar{x}; T_{k+1}),$$

which, by parts (ii) and (iii) is not smaller than

$$2^{-d}R_x B\left(1,0,1,1,2^{1/(p-1)}\right) + (1-2^{-d})B\left(\bar{x}, 2\bar{x}(k+1), \bar{x}; T_{k+1}\right).$$

Hence, by induction, we obtain

$$\begin{aligned}\bar{x}B\left(1,0,1,1,2^{1/(p-1)}\right) &= B\left(\bar{x},0,\bar{x};T_0\right)\\ &\geq (1-2^{-d})^N B(\bar{x}, 2\bar{x}N, \bar{x}; T_N)\\ &\quad + \sum_{k=0}^{N-1}(1-2^{-d})^k 2^{-d} R_x B\left(1,0,1,1,2^{1/(p-1)}\right)\\ &= (1-2^{-d})^N T_x \bar{x} B(1, 2N, 1, 1, 1)\\ &\quad + (1-(1-2^{-d})^N)R_x B\left(1,0,1,1,2^{1/(p-1)}\right).\end{aligned}$$

Now we assume that d is large; if we apply (5.6.9) and (5.6.7), we obtain

$$B(1,2N,1,1,1) \leq \frac{\bar{x} - \left(1 - \left(1-2^{-d}\right)^N\right) R_x}{\left(1-2^{-d}\right)^N T_x \bar{x}} B\left(1,0,1,1,2^{1/(p-1)}\right) \leq 0. \tag{5.6.10}$$

As we have shown in Section 2 (see Remark 5.2), the function B satisfies the majorization condition 2°: $B(x,y,z,u,v) \geq |y|u - Kzu$, where K is a finite constant in our key assumption (5.6.1). Hence, the left-hand side of (5.6.10) is greater than $2N - K$. This implies $2N - K \leq 0$, a contradiction, since N was arbitrary. The claim is proved. □

Bibliography

[1] K. Astala, T. Iwaniec, I. Prause and E. Saksman, *Burkholder integrals, Morrey's problem an quasi-conformal mappings*, J. Amer. Math. Soc. **25** (2012), 507-531.

[2] K. Astala, T. Iwaniec, I. Prause and E. Saksman, *A hunt for sharp Lp-estimates and rank-one convex variational integrals*, Filomat **29** (2015), 245-261.

[3] A. Baernstein II and S. J. Montgomery-Smith, *Some conjectures about integral means of* ∂f *and* $\bar{\partial} f$, Complex Analysis and Differential Equations, Acta Univ. Upsaliensis Skr. Uppsala Univ. C. Organ. Hist. **64**, Uppsala Univ., Uppsala, (1999), 92-109.

[4] R. Bañuelos, *Martingale transforms and related singular integrals*, Trans. Amer. Math. Soc., **293** (1986), 547–563.

[5] R. Bañuelos, M. Brzozowski and A. Osękowski, *Burkholder's function and a weighted* L^2 *bound for stochastic integrals*, to appear in Proceedings of the American Mathematical Society.

[6] R. Bañuelos and A. Osękowski, *Weighted square function estimates*, arXiv:1711.08754v1 [math.PR] 23 Nov 2017.

[7] R. Bañuelos and A. Osękowski, *Sharp Weighted* L^2 *inequalities for square functions*, Transactions of the American Mathematical Society **370** (2018), 2391-2422.

[8] R. Bañuelos and G. Wang, *Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms*, Duke Math. J. **80** (1995), no. 3, 575-600.

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

[10] C. Bennett, R. A. DeVore and R. Sharpley, *Weak-*L^∞ *and BMO*, Ann. of Math. **113** (1981), 601-611.

[11] A. Bonami and D. Lepingle, *Fonction maximale et variation quadratique des martingales en presence d'un poids*, Séminaire de Probabilités XIII, Univ. Strasbourg 1977/78, Lect. Notes Math. **721**, (1979), 294-306.

[12] M. Brzozowski and A. Osękowski, *Sharp weighted weak type* (∞, ∞) *inequality for differentially subordinate martingales.* Statistics and Probability Letters **155** (2019), 108561.

[13] M. Brzozowski and A. Osękowski, *Weighted maximal inequalities for martingale transforms*, to appear in Probability Math. Statist.

[14] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253-272.

[15] D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. **37** (1966), 1494-1504.

[16] D.L. Burkholder and R.F. Gundy, *Extrapolation and interpolation for convex functions of operators on martingales* Acta Math. , **124** (1970), 249-304

[17] D.L. Burkholder, B. Davis and R.F. Gundy, *Integral inequalities for convex functions of operators on martingales* , Proc. 6th Berkeley Symp. Math. Statistics and Probability, **2** (1972), 223-240

[18] D.L. Burkholder, *A sharp inequality for martingale transforms*, Ann. Probab. **7** (1979), 858-863.

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

[20] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals*, Colloque Paul Lévy (Palaiseau, 1987), Astérisque **157-158** (1988), 75-94.

[21] 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.

[22] P. L. Butzer and H. Berens. *Semi-Groups of Operators and Approximation.* Springer Verlag, New York (1967).

[23] S. Chanillo and R. L. Wheeden, Some weighted norm inequalities for the area integral, Indiana Univ. Math. J.**36** (1987), 277–294.

[24] R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math., **51**:241-250, (1974).

[25] R. R. Coifman, P. W. Jones, and J. L. Rubio de Francia, *Constructive decomposition of BMO functions and factorization of Ap weights*. Proc. Amer. Math. Soc., **87(4)**:675-676, (1983).

[26] R. R. Coifman and G. Weiss, Analyse Harmonique Non-Commutative sur Certains Espaces Homogénes, Lecture Notes in Math. **242**, Springer–Verlag, Berlin and New York 1971.

[27] C. Dellacherie and P. A. Meyer, *Probabilities and potential* B, North-Holland, Amsterdam, (1982).

[28] C. Doléans, *Variation quadratique des martingales continues 'a droite*, Ann. Math. Statist. **40** (1969), 284-289.

[29] K. Domelevo and S. Petermichl, *Differential subordination under change of law*, Ann. Probab. **47** (2019), 896-925.

[30] K. Domelevo and S. Petermichl, *Continuous-time sparse domination*, arXiv:1607.06319v2 [math.PR] 20 Apr 2019.

[31] J. Duoandikoetxea, *Extrapolation of weights revisited: new proofs and sharp bounds*, J. Funct. Anal. **260**(2011), 1886-1901.

[32] J. Duoandikoetxea, *Forty years of Muckenhoupt weights. Function spaces and inequalities* Spring Lectures in Analysis, Paseky nad Jizerou, 2013. Charles University and Academy of Sciences, Prague 2013.

[33] J. Garcia-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, volume 116 of North-Holland Mathematics Studies. North-Holland Publishing Co., Amsterdam 1985.

[34] J. Garcia-Cuerva, *José Luis Rubio de Francia (1949-1988).* Collect. Math., **38(1)**:3-15, (1987).

[35] S. Hukovič, S. Treil, A. Volberg, *The Bellman functions and sharp weighted inequalities for square functions*, Complex Analysis, Operators, and Related Topics, Oper. Theory Adv. Appl., vol. **113**, Birkhäuser, Basel, 2000, 97–113.

[36] R. Hunt, B. Muckenhoupt, and R. L. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform* . Trans. Amer. Math. Soc., **176**:227-251, (1973).

[37] T. P. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators, Ann. Math. **175** (2012), 1473–1506.

[38] T. Hytönen, J.M.A.M. van Neerven, M.C. Veraar, and L. Weis, *Analysis in Banach Spaces. Volume I: Martingales and Littlewood-Paley Theory*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, (2016).

[39] T. Iwaniec, *Extremal inequalities in Sobolev spaces and quasiconformal mappings*, Z. Anal. Anwendungen **1** (1982), 1-16.

[40] T. Iwaniec, *Nonlinear Cauchy-Riemann operators in* $\mathbb{R}^n$, Trans. Amer. Math. Soc.**354**(2002), 1961-1995.

[41] M. Izumisawa and N. Kazamaki, *Weighted norm inequalities for martingales*, Tohoku Math. J. **29** (1977), 115-124.

[42] P. W. Jones, *Factorization of Ap weights*. Ann. of Math. (2), **111(3)**:511-530, 1980.

[43] N. Kazamaki, *Continuous exponential martingales and BMO*, Lecture Notes in Math. **1579**, Springer-Verlag, Berlin, Heidelberg, 1994.

[44] M. T. Lacey, *An elementary proof of the* A_2 *bound*, Israel Journal of Mathematics, **217**: 181-195, (2017).

[45] M. T. Lacey, K. Moen, C. Pérez and R. H. Torres, *Sharp weighted bounds for fractional integral operators*, J. Funct. Anal. **259** (2010), 1073–1097.

[46] M. T. Lacey, S. Petermichl and M.C. Reguera, Sharp A_2 inequality for Haar Shift Operators, Math. Ann. **348** (2010), 127–141.

[47] A. K. Lerner, *An elementary approach to several results on the Hardy-Littlewood maximal operator*. Proc. Amer. Math. Soc., **136(8)**:2829-2833, (2008).

[48] A. K. Lerner, On some weighted norm inequalities for Littlewood-Paley operators, Illinois J. Math. **52** (2009), 653–666.

[49] A. K. Lerner, S. Ombrosi and C. Pérez, *Sharp* A_1 *bounds for Calderón-Zygmund operators and the relationship with a problem of Muckenhoupt and Wheeden*, Int. Math. Res. Not. IMRN **6** (2008), Art. ID rnm161, 11 p.

[50] J. Marcinkiewicz, *Quelques théorèmes sur les séries orthogonales*, Ann. Soc. Polon. Math. **16** (1937), 84-96.

[51] T. McConnell. On Fourier multiplier transformations of Banach-valued functions. Trans. Amer. Math. Soc. **285** (1984), 739–757.

[52] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc **165** (1972), 207-226.

[53] B. Muckenhoupt, *The equivalence of two conditions for weight functions*, Studia Math. **49** (1973/74), 101-106.

[54] F. L. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg, *Weak norm estimates of weighted singular operators and Bellman functions.* Manuscript (2010).

[55] 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.

[56] F. L. Nazarov, S. Treil and A. Volberg, *The Tb-theorem on nonhomogeneous spaces*, Acta Math. **190** (2003), 151–239.

[57] A. Osękowski, *Sharp maximal inequality for stochastic integrals*, Proc. Amer. Math. Soc. **136** (2008), 2951-2958.

[58] A. Osękowski, *Sharp maximal inequality for martingales and stochastic integrals*, Electron. Commun. Probab. **14** (2009), 17-30.

[59] A. Osękowski, *Maximal inequalities for continuous martingales and their differential subordinates*, Proc. Amer. Math. Soc. **139** (2011), 721-734.

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

[61] A. Osękowski, *Survey article: Bellman function method and sharp inequalities for martingales*, Rocky Mountain Journal of Mathematics **43** (2013), 1759-1823.

[62] A. Osękowski, *Sharp Weak-Type Inequality for the Haar System, Harmonic Functions and Martingales*, Bulletin of the Polish Academy of Sciences Mathematics **62** (2014), pp. 187-196.

[63] A. Osękowski, *Weighted maximal inequality for differentially subordinate martingales*, Electron. Commun. Probab. 21 (2016), paper no. **23**, 1-10.

[64] A. Osękowski, *Weighted weak-type inequality for martingales*, Bulletin of the Polish Academy of Sciences Mathematics **65** (2017), 165-175.

[65] A. Osękowski, *Weighted inequalities for martingale transforms and stochastic integrals*, Mathematika **63** (2017), 433-450.

[66] A. Osękowski, *Weighted maximal inequalities for the Haar system*, Monatshefte für Mathematik **186** (2018), 321-336.

[67] R. E. A. C. Paley, *A remarkable series of orthogonal functions*, Proc. London Math. Soc. **34** (1932), 241-264.

[68] G. Peskir, *A Change-of-Variable Formula with Local Time on Surfaces.* In: Donati-Martin C., Émery M., Rouault A., Stricker C. (eds) Séminaire de Probabilités XL. Lecture Notes in Mathematics, vol **1899**. Springer, Berlin, Heidelberg, 2007.

[69] S. Petermichl and A. Volberg, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular.* Duke Math. J., **112(2)**:281-305, (2002).

[70] S. Petermichl, The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p- characteristic, Amer. J. Math., **129** (2007), no. 5, 1355–1375.

[71] S. Petermichl, The sharp weighted bound for the Riesz transforms, Proc. Amer. Math. Soc., **136** (2008), no. 4, 1237–1249.

[72] E. T. Sawyer, *A two weight weak type inequality for fractional integrals*, Trans. Amer. Math. Soc. **281** (1984), 339–345.

[73] E. T. Sawyer. *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), 533–545.

[74] J. Schauder, *Eine Eigenschaft des Haarschen Orthogonalsystems*, Math. Z. **28** (1928), pp. 317-320.

[75] I. Singer, Bases in Banach Spaces, Grundlehren der mathematischen Wissenschaften **154**, Springer Verlag, 2010.

[76] L. Slavin and V. Vasyunin, *Sharp results in the integral-form John-Nirenberg inequality*, Trans. Amer. Math. Soc. **363** (2011), 4135-4169.

[77] F. Soria and G. Weiss, *A remark on singular integrals and power weights* Indiana Univ. Math. J., **43(1)**: 187-204, (1994).

[78] E. M. Stein, *On certain operators on Lp spaces* PhD thesis, University of Chicago, Chicago IL, (1955).

[79] E. M. Stein, *Note on singular integrals.* Proc. Amer. Math. Soc., **8**:250-254, (1957).

[80] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume **43** of Princeton Mathematical Series. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[81] 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.
[82] V. Temlyakov, *Sparse Approximation with Bases*, Springer Basel, 2015.
[83] C. Thiele, S. Treil, and A. Volberg, *Weighted martingale multipliers in the non-homogeneous setting and outer measure spaces*, Advances in Mathematics, **285**:1155-1188, (2015).
[84] 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), no. 1, 49-79.
[85] V. Vasyunin and A. Volberg, *Monge-Ampére equation and Bellman optimization of Carleson Embedding Theorems*, Amer. Math. Soc. Transl. (2), vol. **226**, Linear and Complex Analysis, 2009, pp. 195-238.
[86] G.Wang, *Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities*, Ann. Probab. **23** (1995), 522-551.
[87] J. Wittwer, *A sharp estimate on the norm of the martingale transform*, Math. Res. Lett. **7(1)**, 1-12 (2000).