



---

# Equivalent definitions of dyadic Muckenhoupt and reverse Hölder classes in terms of Carleson sequences, weak classes, and comparability of dyadic $L \log L$ and $A_\infty$ constants

Oleksandra Beznosova and Alexander Reznikov

---

**Abstract.** In the dyadic case the union of the reverse Hölder classes,  $\cup_{p>1} RH_p^d$ , is strictly larger than the union of the Muckenhoupt classes,  $\cup_{p>1} A_p^d = A_\infty^d$ . We introduce the  $RH_1^d$  condition as a limiting case of the  $RH_p^d$  inequalities as  $p$  tends to 1 and show the sharp bound on the  $RH_1^d$  constant of the weight  $w$  in terms of its  $A_\infty^d$  constant.

We also examine the summation conditions of the Buckley type for the dyadic reverse Hölder and Muckenhoupt weights and deduce them from an intrinsic lemma which gives a summation representation of the bumped average of a weight. We also obtain summation conditions for continuous reverse Hölder and Muckenhoupt classes of weights and both continuous and dyadic weak reverse Hölder classes. In particular, we prove that a weight belongs to the class  $RH_1$  if and only if it satisfies Buckley's inequality. We also show that the constant in each summation inequality of Buckley type is comparable to the corresponding Muckenhoupt or reverse Hölder constant. To prove our main results we use the Bellman function technique.

## 1. Definitions and main results

Recently novel approaches to the dyadic and continuous  $A_\infty$  classes have yielded essential improvements relevant to the famous  $A_2$  conjecture. The improvement, called the  $A_p$ - $A_\infty$  bound for Calderón–Zygmund operators, was obtained by means of the observation that if a weight  $w$  belongs to the Muckenhoupt class  $A_p$ , then it belongs to a larger class  $A_\infty$  and a certain sequence satisfies the Carleson property. We refer the reader to the papers [10] and [11] for the precise proof of the  $A_2$ - $A_\infty$  bound (in [11] it is not formulated, but can be seen from the proof), and to [9] for a full proof of the  $A_p$ - $A_\infty$  bound.

---

*Mathematics Subject Classification* (2010): 42B20, 42B25.

*Keywords:*  $A_\infty$  weights,  $RH_1$  weights, reverse Hölder condition, sharp estimates, elliptic PDE, Bellman function, Carleson measure.

Carleson sequences are related to  $A_p$  weights and have appeared in many papers related to the boundedness of singular operators. Many such results were proved using the Bellman function method. Using this method, the Carleson embedding theorem was proved in [13]. Results related to Carleson measures (partially proved with certain Bellman functions) also appeared in [14], [15], and [26]. Also, the “easy” case, [24], of the two weight inequality is a certain summation condition, and was also obtained by means of Bellman function. Most of our proofs will use very natural (but not totally sharp) Bellman functions.

We explain our results in more detail. In this paper we present equivalent definitions of Muckenhoupt classes  $A_p$  and reverse Hölder classes  $RH_p$ , and prove sharp inequalities, that show that these definitions are indeed equivalent. One of these definitions is given in terms of Carleson sequences. Also, we define limiting cases  $A_\infty$  and  $RH_1$ , which in the continuous case appear to yield the same sets (see [1]), but in the dyadic case the class  $RH_1$  is strictly larger. We give equivalent definitions of these classes in terms of certain Carleson sequences; besides this, we give a sharp estimate for the so-called  $A_\infty$  and  $RH_1$  constants, which appears to be much harder than in the continuous case (and, actually, somehow uses the continuous result).

The paper is organized as follows. We start by following paper [1], providing all the main definitions of dyadic reverse Hölder and Muckenhoupt classes and state several equivalent definitions of the class  $RH_1^d$ . Also in Section 1 we state our first main result, Theorem 1.7, in which we establish the sharp dependence of the  $RH_1^d$  constant of a weight on its  $A_\infty^d$  constant.

In Section 2 we study summation conditions, introduced first in [6] and [2]. Our second and third main results of this paper are, in fact, Lemmas 2.2 and 2.4, two intrinsic lemmas from which we deduce Theorem 2.6 about the comparability of sums in Buckley’s summation condition and certain bumped averages of the weight  $w$ . The reader should note that even though Theorem 2.6 turns out to be extremely strong and is very handy for Hölder and Muckenhoupt classes, our lemmas, especially Lemma 2.2 are much more general and could be potentially applied to large class of bumped averages of any nonnegative function  $w$  and every interval  $J \subset \mathbb{R}$ . We show how Theorem 2.6 follows from our lemmas and how Buckley’s theorem follows from Theorem 2.6. The reader will see that Theorem 2.6 is substantially stronger than Buckley’s theorem, because it is true for *any* weight, not necessarily from  $A_p$  or  $RH_p$ . This is illustrated in Theorem 2.7, where the comparability of constants in summation conditions and corresponding Hölder and Muckenhoupt constants of the weight is established in both continuous and dyadic cases.

In Section 3 we discuss weak reverse Hölder and Muckenhoupt classes. We start by giving definitions of these classes and state another consequence of Theorem 2.6, namely Theorem 3.5, which contains a version of Buckley’s theorem but for the weak reverse Hölder weights. The proof of Theorem 3.5 is essentially the same as the proof of Theorem 2.6, so we omit most of the details.

All the Bellman function proofs can be found in Section 4. We start with the proof of Lemma 2.2, which we think is the simplest of the three Bellman function proofs given in this paper and is a nice introduction to the Bellman function

technique. The Bellman function technique is not new, but as far as we know this is the first place where Bellman function technique is applied in such an “intrinsic” setup. By “intrinsic” here we mean that a lemma has a function  $A(x)$  as one of the parameters, and convexity properties of the function  $A$  are then used to build the Bellman function for the inequality. The proof of Lemma 2.2 is followed by the proof of Lemma 2.4 which we hope will be easy to digest after having seen the proof of Lemma 2.2. The proof of Theorem 1.7 is the hardest and occupies the last half of the section. The proof itself is in fact very similar to the proof of the continuous version of Theorem 1.7, which can be found in [1]. This dyadic proof is longer than the continuous one because in the dyadic case we have to deal with many details that are specific to the dyadic Bellman function proof in a nonconvex domain. We encourage the reader to understand the proof of Theorem 1.1 from [1] before reading our proof of Theorem 1.7.

All results of this paper apply to the one-dimensional case only.

**Acknowledgements.** The authors are grateful to A. Volberg for useful suggestions related to proving Theorem 1.7 and to V. Vasyunin for useful discussions. We would also like to express our gratitude to C. Thiele, I. Uriarte-Tuero, and A. Volberg for organizing the Summer School 2010 in UCLA, where this paper originated, and to C. Pérez and R. Espínola for organizing the Summer School 2011 in Seville and AIM workshop, where we finished this paper. We also thank Paul Hagelstein, Cristina Pereyra, and Vasili Vasyunin for useful corrections, that made the presentation a lot nicer. Finally, we would like to thank the referee for several useful remarks about the organization of the text.

**1.1. First definitions**

Let  $\mathcal{D}$  be the dyadic grid  $\mathcal{D} := \{I \subset \mathbb{R} : I = [k2^{-j}, (k + 1)2^{-j}); k, j \in \mathbb{Z}\}$ .

We say that  $w$  is a *weight* if it is a locally integrable function on the real line and positive almost everywhere (with respect to Lebesgue measure). Let  $\langle w \rangle_J$  be the average of a weight  $w$  over a given interval  $J \subset \mathbb{R}$ ,

$$\langle w \rangle_J := \frac{1}{|J|} \int_J w \, dx$$

and let  $\Delta_J w$  be defined by

$$\Delta_J w := \langle w \rangle_{J^+} - \langle w \rangle_{J^-},$$

where  $J^+$  and  $J^-$  are the left and right dyadic children of the interval  $J$ .

**Definition 1.1.** A weight  $w$  belongs to the dyadic *Muckenhoupt class*  $A_p^d$  whenever its dyadic Muckenhoupt constant  $[w]_{A_p^d}$  is finite:

$$(1.1) \quad [w]_{A_p^d} := \sup_{J \in \mathcal{D}} \langle w \rangle_J \langle w^{-1/(p-1)} \rangle_J^{p-1} < \infty.$$

**Remark 1.2.** The inequality (1.1) can be rewritten in the following way:

$$0 \leq \langle w^{-1/(p-1)} \rangle_J - \langle w \rangle_J^{-1/(p-1)} \leq ([w]_{A_p^d}^{1/(p-1)} - 1) \langle w \rangle_J^{-1/(p-1)}.$$

Later we will use this formulation in writing the definitions Hölder and Muckenhoupt classes in the proof of Theorem 2.7.

Note that by Hölder’s inequality,  $[w]_{A_p^d} \geq 1$  holds for all  $1 < p < \infty$ , as well as the following inclusion:

$$\text{if } 1 < p \leq q < \infty \text{ then } A_p^d \subseteq A_q^d, \quad [w]_{A_q^d} \leq [w]_{A_p^d}.$$

Hence, for  $1 < p < \infty$ , the Muckenhoupt classes  $A_p^d$  form an increasing chain. There are two natural limits to consider: as  $p$  approaches 1 and as  $p$  goes to  $\infty$ . We will be interested in the limiting case as  $p \rightarrow \infty$ ,  $A_\infty^d = \bigcup_{p>1} A_p^d$ . There are several equivalent definitions of  $A_\infty^d$ . We state the one that we use (the natural limit of the  $A_p^d$  conditions, that also defines the  $A_\infty^d$  constant of the weight  $w$ ); for other equivalent definitions see [7], [8], or [20]:

$$(1.2) \quad w \in A_\infty^d \iff [w]_{A_\infty^d} := \sup_{J \in \mathcal{D}} \langle w \rangle_J e^{-(\log w)_J} < \infty,$$

where  $\log$  stands for the regular natural logarithm.

**Remark 1.3.** The inequality (1.2) can be rewritten in the following way:

$$0 \leq \log \langle w \rangle_J - \langle \log w \rangle_J \leq \log [w]_{A_\infty^d}.$$

Note also that if a weight  $w$  belongs to the Muckenhoupt class  $A_p^d$  for some  $p > 1$ , or, equivalently, to the class  $A_\infty^d$ , then  $w$  has to be a dyadically doubling weight, i.e., its dyadic doubling constant

$$\mathcal{D}^d(w) := \sup_{I \in \mathcal{D}} \frac{\langle w \rangle_{F(I)}}{\langle w \rangle_I},$$

where  $F(I)$  stands for the dyadic parent of the interval  $I$ , has to be finite.

**Definition 1.4.** A weight  $w$  belongs to the *dyadic reverse Hölder class*  $RH_p^d$  ( $1 < p < \infty$ ) if

$$(1.3) \quad [w]_{RH_p^d} := \sup_{J \in \mathcal{D}} \frac{\langle w^p \rangle_J^{1/p}}{\langle w \rangle_J} < \infty.$$

**Remark 1.5.** The inequality (1.3) can be rewritten in the following way:

$$0 \leq \langle w^p \rangle_J - \langle w \rangle_J^p \leq ([w]_{RH_p^d}^p - 1) \langle w \rangle_J^p.$$

Note that by Hölder’s inequality the dyadic reverse Hölder classes satisfy:

$$\text{if } 1 < p \leq q < \infty, \text{ then } RH_q^d \subseteq RH_p^d \text{ and } 1 \leq [w]_{RH_p^d} \leq [w]_{RH_q^d},$$

which is similar to the inclusion chain of the  $A_p^d$  classes, except inclusion runs in the opposite direction. Similarly we can consider the two limiting cases  $RH_\infty^d$  (the smallest) and  $RH_1^d$  (the largest). As in the case of Muckenhoupt classes we are more interested in the largest one, call it  $RH_1^d := \bigcup_{p>1} RH_p^d$ .

The natural limit as  $p \rightarrow 1^+$  of the reverse Hölder inequalities is the condition, which we will take as a definition of the class  $RH_1^d$ ,

$$(1.4) \quad w \in RH_1^d \iff [w]_{RH_1^d} := \sup_{J \in \mathcal{D}} \left\langle \frac{w}{\langle w \rangle_J} \log \frac{w}{\langle w \rangle_J} \right\rangle_J < \infty,$$

where  $\log$  is a regular logarithm base  $e$ , which could be negative. Nevertheless, by the Jensen inequality the  $RH_1$  constant defined this way is always nonnegative.

The  $RH_1^d$  constant of the weight  $w$  is the natural limit of the  $RH_p^d$  constants in the sense that for every interval  $I \in \mathcal{D}$

$$(1.5) \quad \left\langle \frac{w}{\langle w \rangle_I} \log \frac{w}{\langle w \rangle_I} \right\rangle_I = \lim_{p \rightarrow 1^+} \frac{p}{p-1} \log \frac{\langle w^p \rangle_I^{1/p}}{\langle w \rangle_I}$$

**Remark 1.6.** The inequality (1.4) can be rewritten as:

$$\langle w \log(w) \rangle_J \leq \langle w \rangle_J \log \langle w \rangle_J + Q \langle w \rangle_J \quad \forall J \in \mathcal{D},$$

where  $Q = [w]_{RH_1}$ . Note that since the function  $x \log x$  is concave, by Jensen’s inequality we also have

$$\langle w \rangle_J \log \langle w \rangle_J \leq \langle w \log(w) \rangle_J.$$

In the continuous case, for  $A_\infty$  and  $RH_1$ , in 1974 Coifman and Fefferman showed that  $A_\infty = \bigcup_{p>1} RH_p = RH_1$ . In the dyadic case this is not true. One can only claim the inclusion  $A_\infty^d \subset RH_1^d$ . As for the other inclusion, it only holds for the dyadically doubling weights since, unlike the  $A_p^d$  weights, dyadic reverse Hölder weights do not have to be doubling. An example of such a weight can be found in Buckley [2].

**Different ways to define the  $RH_1^d$  constant of the weight  $w$**

First, observe that, trivially, the logarithm in the definition of the  $RH_1^d$  constant can be replaced by  $\log^+(x)$ , ( $\log^+(x) = \max(\log x, 0)$ ) or  $\log(e + x)$ , which will, however, increase the  $RH_1^d$  constant slightly.

Secondly, from the Stein lemma (see [19]), we know that

$$3^{-n} \langle M(f \chi_I) \rangle_I \leq \left\langle f \log \left( e + \frac{f}{\langle f \rangle_I} \right) \right\rangle_I \leq 2^n \langle M(f \chi_I) \rangle_I,$$

where  $Mf$  is the maximal function of  $f$ . Thus an equivalent way to define the  $RH_1$  constant is

$$(1.6) \quad [w]_{RH_1^{d'}} := \sup_{I \in \mathcal{D}} \frac{1}{w(I)} \int_I M(w \chi_I) dx,$$

which, indeed, is one of the ways to define the class  $A_\infty$ ; see, for example, [25] or [10]. The constant  $[w]_{RH_1^{d'}}$  is also called the Wilson  $A_\infty^d$  constant of the weight.

One can also define dyadic reverse Hölder and Muckenhoupt constants using Luxemburg norms. The same is true for the  $RH_1^d$  constant. We first define the Luxemburg norm of a function in the following way: for an Orlicz function  $\Phi: [0, \infty) \mapsto [0, \infty)$ , we define  $\|w\|_{\Phi(L), I}$  by

$$\|w\|_{\Phi(L), I} := \inf \left\{ \lambda > 0: \frac{1}{|I|} \int_I \Phi\left(\frac{|w|}{\lambda}\right) \leq 1 \right\}.$$

Iwaniec and Verde in [12] showed that for every  $w$  and  $I \subset \mathbb{R}^n$

$$\|w\|_{L \log L, I} \leq \int_I w \log \left( e + \frac{w}{\langle w \rangle_I} \right) dx \leq 2 \|w\|_{L \log L, I},$$

so another equivalent definition of the  $RH_1^d$  constant of the weight  $w$  is

$$(1.7) \quad [w]_{RH_1^{d''}} := \sup_{I \in \mathcal{D}} \frac{\|w\|_{L \log L, I}}{\|w\|_{L, I}}.$$

### 1.2. First main result of the paper

In this section we state the first result of the paper, and then explain the other questions we study.

We prove the following sharp relationship between the  $RH_1^d$  and  $A_\infty^d$  constants:

**Theorem 1.7 (Main result 1: comparability of the  $RH_1^d$  and  $A_\infty^d$  constants).** *If a weight  $w$  belongs to the Muckenhoupt class  $A_\infty^d$ , then  $w \in RH_1^d$ . Moreover,*

$$(1.8) \quad [w]_{RH_1^d} \leq C [w]_{A_\infty^d},$$

where the constant  $C$  can be taken to be  $\log(16)$  ( $C = \log(16)$ ). Moreover, the constant  $C = \log(16)$  is the best possible.

A Bellman function proof of this theorem can be found in Section 4.5. An independent proof of the analogue of this theorem for the constant  $[w]_{RH_1^{d'}}$  was recently independently obtained in [10].

Note that all of the above is true in the continuous case and can be found in [1] (with the sharp constant  $C = e$ , and with a double exponential lower bound). Note also that the lower bound (Theorem 1.2 in [1]) in the dyadic case cannot possibly hold since the class  $RH_1^d$  is strictly larger than  $A_\infty^d$ .

## 2. Summation conditions on weights

In this section we will introduce and discuss an important set of inequalities that characterize the dyadic reverse Hölder and Muckenhoupt classes. We are mostly interested in the dyadic results here, so we will follow Buckley [4]. Note that the inequalities we discuss in this section have continuous analogues, and many facts and questions here apply to the continuous case as well (see [6]).

As we discussed earlier,  $RH_1^d \neq A_\infty^d$  because all dyadic Muckenhoupt conditions imply that the weight is dyadically doubling, while dyadic reverse Hölder conditions allow nondoubling weights (in the continuous case both reverse Hölder and Muckenhoupt conditions imply the continuous doubling property). For the dyadically doubling weights the  $RH_1^d$  and  $A_\infty^d$  conditions are equivalent.

We now state a theorem that characterizes the dyadic reverse Hölder and Muckenhoupt classes via summation conditions. We attribute this theorem to Buckley, however all parts but the Buckley inequality (part (2)) in the continuous case and part (4) in the dyadic case first appeared in [6] and are due to Fefferman, Kenig, and Pipher.

**Theorem 2.1** (Buckley '93). *Suppose  $1 < p < \infty$  and  $w$  is a doubling weight. Then:*

(1)  $w \in RH_p^d$  if and only if on every dyadic interval  $J$

$$(2.1) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq K \langle w \rangle_J^p.$$

Moreover,  $K \leq C[w]_{RH_p^d}^p$ .

(2) (Buckley's inequality)  $w \in RH_1^d$  if and only if for some  $K > 0$  on every dyadic interval  $J$

$$(2.2) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I |I| \leq K \langle w \rangle_J.$$

(3)  $w \in A_p^d$  if and only if on every dyadic interval  $J$

$$(2.3) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 (\langle w \rangle_I)^{-1/(p-1)} |I| \leq K \langle w \rangle_J^{-1/(p-1)}.$$

(4) (Fefferman–Kenig–Pipher inequality)  $w \in A_\infty^d$  if and only if on every dyadic interval  $J$

$$(2.4) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 |I| \leq C \log[w]_{A_\infty^d}.$$

Buckley's inequality (part (2)) is the one of the most interest here since as we will see later it characterizes the class  $RH_1^d$ ; it is also the only one stated without the sharp constant. In [26], Wittwer showed that, in the case  $w \in A_2^d$ , Buckley's

inequality holds with  $K = C[w]_{A_2^d}$  and this linear dependence on the  $A_2^d$  constant of the weight  $w$  is sharp, which is the best known result for Buckley’s inequality. Also, in the Fefferman–Kenig–Pipher inequality the sharp constant is  $C = 8$ ; this was obtained by Vasyunin in [22] using the Bellman function method.

Using the method of Bellman functions we are going to show that, in Buckley’s inequality,  $K \leq C[w]_{RH_1^d}$ . We also show that the assumption that  $w$  is a doubling weight can be dropped. Finally, we show that the above four sums also satisfy the lower bound estimates in terms of the corresponding constants. We state our second main result now.

We start with the following lemma, from which Theorem 2.1 will follow.

**Lemma 2.2.** *Let  $A(x)$  be a convex twice differentiable function on  $(0, \infty)$  such that for all numbers  $x$  and  $t$ , where  $x, x \pm t$  are in the domain of  $A$ , the inequality*

$$(2.5) \quad A(x) - \frac{A(x-t) + A(x+t)}{2} + \alpha t^2 A''(x) \leq 0$$

holds for some constant  $\alpha > 0$  independent of  $x$  and  $t$ . Then for every weight  $w$  and any interval  $J$  the following inequality holds:

$$(2.6) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I| \leq C(\langle A(w) \rangle_J - A(\langle w \rangle_J)).$$

Moreover, if the second derivative of  $A$  satisfies the inequality

$$(2.7) \quad \int_{-1}^1 (1 - |t|) A''(x + \varepsilon t) dt \geq q A''(x)$$

for every  $x \in (0, \infty)$  and every  $\varepsilon \geq 0$  with some positive constant  $q$  uniformly in  $x$  and  $\varepsilon$ , then the inequality (2.6) holds with the constant  $C = 8(1/q)$ .

The Bellman function proof of Lemma 2.2 can be found in Section 4.3.

**Remark 2.3.** Note that if the second derivative of  $A$  is a monotone function then (2.7) holds trivially with the constant  $q = 1/2$ , which makes Lemma 2.2 applicable to a large class of functions producing a number of new inequalities of Buckley’s type. In particular, the function  $A(x)$  can be taken to be  $A(x) = x^p$ ,  $p > 1$ ;  $A(x) = x \log x$ ;  $A(x) = x^{-1/(p-1)}$  with  $p > 1$ ; or  $A(x) = -\log x$ . In what follows we will see how these choices of the function  $A(x)$  imply Buckley’s theorem.

Now we want to introduce the “reverse” lemma, which is true for particular (those most interesting for us) choices of the function  $A$ .

**Lemma 2.4.** *Suppose  $A(x)$  is a function bounded from below.*

(1) *Let  $A(x)$  be a function defined on  $(0, \infty)$  such that*

$$(2.8) \quad A(x) - \frac{A(x-t) + A(x+t)}{2} + \beta t^2 A''(x) \geq 0,$$

holds with some positive constant  $\beta$  independent of  $x$  and  $t$ .



Then for every weight  $w$  and interval  $J$

$$(2.9) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I| \geq C(\langle A(w) \rangle_J - A(\langle w \rangle_J)).$$

Moreover, condition (2.8) holds for  $A(x) = x^p$ , for all  $p > 1$ , and for  $A(x) = x \log x$ .

(2) Suppose  $A$  satisfies the inequality

$$(2.10) \quad A(x) - \frac{A(x-t) + A(x+t)}{2} + \beta t^2 A''(x) \geq 0$$

whenever  $0 < t < \frac{C-1}{C}x$  (for  $C > 1$ ,  $\beta$  depends on  $C$ ). Then for every doubling weight  $w$  and any interval  $J$  we have

$$(2.11) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I| \geq C(\langle A(w) \rangle_J - A(\langle w \rangle_J)),$$

where the constant  $C$  depends on the doubling constant of  $w$ .

Moreover, condition (2.10) holds for  $A(x) = x^{-1/(p-1)}$  for all  $p > 1$ .

A Bellman function proof of Lemma 2.4 can be found in Section 4.4.

**Remark 2.5.** Note that in Lemma 2.4, similarly to Lemma 2.2, we can also write conditions (2.8) and (2.10) in integral form, but in this case (2.8) and (2.10) are easier to check at least for the functions we are of interest here.

From our lemmas, by taking  $A(x) = x^p$  and  $A(x) = x^{-1/(p-1)}$ ,  $p > 1$ ,  $A(x) = x \log x$ , and  $A(x) = -\log(x)$ , we derive the following theorem.

**Theorem 2.6 (Main result 2: representation of bumped averages).** *Suppose  $1 < p < \infty$  and  $w$  is a weight. Then:*

(1) *Case  $A(x) = x^p$ ,  $p > 1$ . There are positive real constants  $c$  and  $C$  independent of the weight  $w$ , such that for every interval  $J$*

$$(2.12) \quad c(\langle w^p \rangle_J - \langle w \rangle_J^p) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left(\frac{\Delta_I w}{\langle w \rangle_I}\right)^2 \langle w \rangle_I^p |I| \leq C(\langle w^p \rangle_J - \langle w \rangle_J^p).$$

(2) *Case  $A(x) = x \log x$ . There are positive real constants  $c$  and  $C$  independent of the weight  $w$ , such that for every interval  $J$*

$$(2.13) \quad c(\langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left(\frac{\Delta_I w}{\langle w \rangle_I}\right)^2 \langle w \rangle_I |I| \leq C(\langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J).$$

- (3) Case  $A(x) = x^{-1/(p-1)}$ . There is a positive real constant  $C$  independent of  $w$  such that, for every interval  $J$ ,

$$(2.14) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^{-1/(p-1)} |I| \leq C (\langle w^{-1/(p-1)} \rangle_J - \langle w \rangle_J^{-1/(p-1)}).$$

Moreover, if  $w$  is a doubling weight, then there exists a constant  $c$  that may depend on the doubling constant of the weight  $w$ , such that for every interval  $J$ ,

$$(2.15) \quad c (\langle w^{-1/(p-1)} \rangle_J - \langle w \rangle_J^{-1/(p-1)}) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^{-1/(p-1)} |I|.$$

- (4) Case  $A(x) = -\log x$ . There is a positive real constant  $C$  independent of  $w$ , such that for every interval  $J$ ,

$$(2.16) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 |I| \leq C (\log \langle w \rangle_J - \langle \log w \rangle_J).$$

Theorem 2.6 follows immediately from Lemma 2.2, the remark after it and Lemma 2.4. We will leave its proof to the reader. Instead, we show how Theorem 2.6 implies Buckley’s theorem in the dyadic and continuous cases and in the case of weak reverse Hölder classes.

In order to write our results in a more compact way we start by giving another way of defining reverse Hölder and Muckenhoupt constants. We call them Buckley’s constants and denote them by  $[w]_{RH_p^{d,B}}$  and  $[w]_{A_p^{d,B}}$ . Namely, for  $p \geq 1$ , we define

$$[w]_{RH_p^{d,B}} := \inf \left\{ Q > 1 \text{ s.t. } \forall J \in \mathcal{D} \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq Q \langle w \rangle_J^p \right\},$$

and similarly we can define continuous Buckley’s reverse Hölder constants

$$[w]_{RH_p^B} := \inf \left\{ Q > 1 \text{ s.t. } \forall J \subset \mathbb{R} \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq Q \langle w \rangle_J^p \right\}.$$

Similarly, for  $1 < p < \infty$ , we define dyadic and continuous Buckley’s Muckenhoupt constants

$$[w]_{A_p^{d,B}} := \inf \left\{ Q > 0: \forall J \in \mathcal{D}, \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^{-1/(p-1)} |I| \leq Q \langle w \rangle_J^{-1/(p-1)} \right\}$$

and

$$[w]_{A_p^B} := \inf \left\{ Q > 0: \forall J \subset \mathbb{R}, \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^{-1/(p-1)} |I| \leq Q \langle w \rangle_J^{-1/(p-1)} \right\}.$$

In the  $A_\infty$  case we have

$$[w]_{A_\infty^{d,B}} := \inf \left\{ Q > 0: \forall J \in \mathcal{D}, \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 |I| \leq Q \right\}$$

and

$$[w]_{A_\infty^B} := \inf \left\{ Q > 0: \forall J \subset \mathbb{R}, \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 |I| \leq Q \right\}.$$

Note that in the reverse Hölder case of Buckley’s constants we do not need to define the  $RH_1$  constants separately. We are now ready to state the result about the comparability of Buckley’s constants and the regular reverse Hölder and Muckenhoupt constants.

**Theorem 2.7 (Comparability of constants in summation conditions).**

- (1) *Suppose  $1 < p < \infty$ . Then there are positive constants  $C$  and  $c$  such that, for every weight  $w$ ,*

$$c \left( [w]_{RH_p^d}^p - 1 \right) \leq [w]_{RH_p^{d,B}} \leq C \left( [w]_{RH_p^d}^p - 1 \right)$$

and

$$c \left( [w]_{RH_p}^p - 1 \right) \leq [w]_{RH_p^B} \leq C \left( [w]_{RH_p}^p - 1 \right).$$

- (2) *In the case  $p = 1$  there are positive constants  $C$  and  $c$  such that, for every weight  $w$ ,*

$$c[w]_{RH_1^d} \leq [w]_{RH_1^{d,B}} \leq C[w]_{RH_1^d}$$

and

$$c[w]_{RH_1} \leq [w]_{RH_1^B} \leq C[w]_{RH_1}.$$

- (3) *For any  $1 < p < \infty$  there is a positive constant  $C$  such that, for every weight  $w$ ,*

$$[w]_{A_p^{d,B}} \leq C \left( [w]_{A_p^d}^{1/(p-1)} - 1 \right) \quad \text{and} \quad [w]_{A_p^B} \leq C \left( [w]_{A_p}^{1/(p-1)} - 1 \right).$$

- (4) *In the case  $p = \infty$  there is a positive constant  $C$  such that, for every weight  $w$ ,*

$$[w]_{A_\infty^{d,B}} \leq C \log [w]_{A_\infty^d} \quad \text{and} \quad [w]_{A_\infty^B} \leq C \log [w]_{A_\infty}.$$

Moreover, if  $w$  is a doubling weight then:

- (5) *For any  $1 < p < \infty$ ,*

$$c_d \left( [w]_{A_p^d}^{1/(p-1)} - 1 \right) \leq [w]_{A_p^{d,B}} \quad \text{and} \quad c \left( [w]_{A_p}^{1/(p-1)} - 1 \right) \leq [w]_{A_p^B}$$

hold with positive constants  $c_d$  and  $c$  that depend on the (dyadic) doubling constant of the weight  $w$ .

We now show how Theorem 2.7 follows from Theorem 2.6. Note also that the pair of constants  $c_d$  and  $c$  above is different in part (5) because one depends on the dyadic doubling constant of the weight  $w$  and the other depends on the continuous doubling constant of  $w$ .

*Proof.* We prove case (1). The other cases are proved in a similar way with only minor changes, and are left to the reader.

We will show that the first part of Theorem 2.7 follows from the first part of Theorem 2.6, from which we know that there are constants  $c$  and  $C$  such that for any weight  $w$  and interval  $J \subset \mathbb{R}$

$$(2.17) \quad c (\langle w^p \rangle_J - \langle w \rangle_J^p) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq C (\langle w^p \rangle_J - \langle w \rangle_J^p).$$

First, we assume that  $w \in RH_p^{(d)}$  (dyadic or continuous), which means, by Remark 1.5, that for every (dyadic) interval  $J \subset \mathbb{R}$  we have

$$0 \leq \langle w^p \rangle_J - \langle w \rangle_J^p \leq ([w]_{RH_p^{(d)}}^p - 1) \langle w \rangle_J^p.$$

So, by inequality (2.17), we have that

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq C ([w]_{RH_p^{(d)}}^p - 1) \langle w \rangle_J^p.$$

Hence  $[w]_{RH_p^{(d),B}} \leq C ([w]_{RH_p^{(d)}}^p - 1)$ .

Second, assume that  $w \in RH_p^{(d),B}$ , so for each (dyadic) interval  $J \subset \mathbb{R}$  we have

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq [w]_{RH_p^{(d),B}} \langle w \rangle_J^p.$$

Then from (2.17) we deduce that

$$\langle w^p \rangle_J - \langle w \rangle_J^p \leq \frac{1}{c} [w]_{RH_p^{(d),B}} \langle w \rangle_J^p,$$

which means that  $w \in RH_p^{(d)}$  and  $c ([w]_{RH_p^{(d)}}^p - 1) \leq [w]_{RH_p^{(d),B}}$ .

Parts (2), (3), and (4) of Theorem 2.7 are proved in exactly the same way, using Remarks 1.6, 1.2 and 1.3 and the corresponding parts of Theorem 2.6. The doubling assumptions in (3) and (4) also come from Theorem 2.6.  $\square$

Theorem 2.7 obviously implies Buckley’s theorem (Theorem 2.1), but our Theorem 2.6 is even stronger. Since Theorem 2.6 shows comparability of summations for a given weight with its bumped averages, we can also write summation conditions for the weak reverse Hölder classes in a similar way.

### 3. Summation conditions for weak reverse Hölder classes

In this section we discuss the weak reverse Hölder class  $RHW_p$ ,  $p \geq 1$ . We recall the definition of the  $RHW_p$ -constant. For simplicity, we drop the superscript  $d$ , that referred to the dyadic case.

All of the above is true in the continuous case as well, when all suprema are taken over any interval  $J \subset \mathbb{R}$ . We do not repeat all the definitions and refer the reader to [1].

We also give the definition of the so-called “weak” reverse Hölder class  $RHW_p^d$ .

**Definition 3.1.** In the dyadic case let  $J^*$  stand for the dyadic parent of  $J \in \mathcal{D}$ . Then a weight  $w$  belongs to the dyadic weak reverse Hölder class  $RHW_p^d$ ,  $p > 1$ , if and only if its weak reverse Hölder constant is finite:

$$(3.1) \quad w \in RHW_p^d \iff [w]_{RHW_p^d} := \sup_{J \in \mathcal{D}} \frac{\langle w^p \rangle_J^{1/p}}{\langle w \rangle_{J^*}} < \infty.$$

For  $p = 1$  we define the  $RHW_1^d$  class by

$$(3.2) \quad w \in RHW_1^d \iff [w]_{RHW_1^d} := \sup_{J \in \mathcal{D}} \left\langle \frac{w}{\langle w \rangle_{J^*}} \log \frac{w}{\langle w \rangle_J} \right\rangle_J < \infty.$$

In the continuous case, for any interval  $J \subset \mathbb{R}$  let  $2J$  stand for the interval of twice the length of  $J$  having the same center as  $J$ . Then the weak reverse Hölder classes  $RHW_p$ ,  $p > 1$  and  $RHW_1$  are defined by

$$(3.3) \quad w \in RHW_p \iff [w]_{RHW_p} := \sup_{J \subset \mathbb{R}} \frac{\langle w^p \rangle_J^{1/p}}{\langle w \rangle_{2J}} < \infty$$

and

$$(3.4) \quad w \in RHW_1 \iff [w]_{RHW_1} := \sup_{J \subset \mathbb{R}} \left\langle \frac{w}{\langle w \rangle_{2J}} \log \frac{w}{\langle w \rangle_J} \right\rangle_J < \infty.$$

We again note that it is important that in the definitions of  $RHW_1^d$  and  $RHW_1$ , inside the log we divide by the average of  $w$  over the interval  $J$ , not by the average over  $J^*$  or  $2J$ .

**Remark 3.2.** We now explain why the definition (3.2) of the weak reverse Hölder constant makes sense. In fact, in the spirit of the formula above, we can define this constant as

$$[w]_{RHW_1^{d''}} := \sup_{I \in \mathcal{D}} \frac{\|w\|_{L \log L, I}}{\|w\|_{L, I^*}}.$$

**Remark 3.3.** Also note that as in the strong case we can rewrite (3.1), (3.2), (3.3) and (3.4) as:

$$(3.5) \quad w \in RHW_p^d \iff 0 \leq \langle w^p \rangle_J \leq [w]_{RHW_p^d}^p \langle w \rangle_{J^*}^p, \quad \forall J \in \mathcal{D},$$

$$(3.6) \quad w \in RHW_1^d \iff 0 \leq \langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J \leq [w]_{RHW_1^d} \langle w \rangle_{J^*}, \quad \forall J \in \mathcal{D},$$

$$(3.7) \quad w \in RHW_p \iff 0 \leq \langle w^p \rangle_J \leq [w]_{RHW_p}^p \langle w \rangle_{2J}^p, \quad \forall J \subset \mathbb{R},$$

$$(3.8) \quad w \in RHW_1^d \iff 0 \leq \langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J \leq [w]_{RHW_1^d} \langle w \rangle_{2J}, \quad \forall J \subset \mathbb{R}.$$

We are ready to define dyadic and continuous weak Buckley reverse Hölder constants now in the most natural way.

**Definition 3.4.** For any  $p \geq 1$  let

$$[w]_{RHW_p^{d,B}} := \inf \left\{ Q > 0 : \forall J \in \mathcal{D} \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq Q \langle w \rangle_{J^*}^p \right\}$$

and

$$[w]_{RHW_p^B} := \inf \left\{ Q > 0 : \forall J \subset \mathbb{R} \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq Q \langle w \rangle_{J^*}^p \right\}.$$

Now we are ready to state the following theorem, which is also a consequence of Theorem 2.16.

**Theorem 3.5.** *A weight  $w$  belongs to  $RHW_p^d$  if and only if the weak Buckley constant  $[w]_{RHW_p^{d,B}}$  is finite. Moreover, there exists a positive constant  $C_1$  that does not depend on  $w$  and  $p$  and a positive constant  $C_2$  that may depend on  $p$ , such that for any  $p > 1$*

$$[w]_{RHW_p^{d,B}} \leq C_1 [w]_{RHW_p^d}^p \quad \text{and} \quad [w]_{RHW_p^d} \leq C_2 ([w]_{RHW_p^{d,B}} + 1)^{1/p}$$

and, similarly, in the continuous case

$$[w]_{RHW_p^B} \leq C_1 [w]_{RHW_p}^p \quad \text{and} \quad [w]_{RHW_p} \leq C_2 ([w]_{RHW_p^B} + 1)^{1/p}.$$

In the case  $p = 1$  there are positive constants  $C$  and  $c$  such that

$$c [w]_{RH_1^{d,B}} \leq [w]_{RH_1^d} \leq C [w]_{RH_1^{d,B}}$$

and

$$c [w]_{RH_1} \leq [w]_{RH_1^B} \leq C [w]_{RH_1}.$$

*Proof.* The proofs for the continuous and dyadic cases are identical, so we treat both cases simultaneously.

For  $p > 1$ , by part (1) of Theorem 2.6 we know that for any weight  $w$  and any interval  $J$  there holds

$$(3.9) \quad c(\langle w^p \rangle_J - \langle w \rangle_J^p) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq C(\langle w^p \rangle_J - \langle w \rangle_J^p).$$

Note that  $\langle w \rangle_J$  is nonnegative, so, if  $w$  belongs to the (dyadic or continuous) class  $RHW_p^d$ , by (3.5) or (3.7) we have that, for every (dyadic) interval  $J \subset \mathbb{R}$ ,

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq C \langle w^p \rangle_J \leq C [w]_{RHW_p^{(d)}}^p \langle w \rangle_{F(J)}^p,$$

where  $F(J)$  is either the dyadic parent of  $J$  or  $2J$ . So

$$[w]_{RHW_p^{(d),B}} \leq C [w]_{RHW_p^{(d)}}^p.$$

To prove the reverse inequality we assume that  $w$  is in (dyadic or continuous)  $RHW_p^{(d),B}$ . Then

$$c(\langle w^p \rangle_J - \langle w \rangle_J^p) \leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I^p |I| \leq [w]_{RHW_p^{(d),B}} \langle w \rangle_{F(J)}^p,$$

from which we conclude that

$$\langle w^p \rangle_J \leq \frac{1}{c} [w]_{RHW_p^{(d),B}} \langle w \rangle_{F(J)}^p + \langle w \rangle_J^p.$$

Note that  $\langle w \rangle_J \leq 2 \langle w \rangle_{F(J)}$ , so

$$\langle w^p \rangle_J \leq \left( \frac{1}{c} [w]_{RHW_p^{(d),B}} + 2^p \right) \langle w \rangle_{F(J)}^p,$$

which implies that  $[w]_{RHW_p^{(d)}} \leq \left( \frac{1}{c} [w]_{RHW_p^{(d),B}} + 2^p \right)^{1/p}$  and completes the proof of the theorem for  $p > 1$ .

For  $p = 1$  we use the comparability (part (2) of Theorem 2.6)

$$\begin{aligned} c(\langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J) &\leq \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left( \frac{\Delta_I w}{\langle w \rangle_I} \right)^2 \langle w \rangle_I |I| \\ &\leq C(\langle w \log w \rangle_J - \langle w \rangle_J \log \langle w \rangle_J) \end{aligned}$$

together with the definitions (3.6) and (3.8) of the dyadic and continuous  $RHW_1^{(d)}$ . This proof is similar to the continuous case and is left to the reader.  $\square$

**Remark 3.6.** Observe that we have proved the theorem for any pairs  $(J, F(J))$ , that satisfy the following two conditions:

- (1)  $J \subset F(J)$ , and
- (2)  $|J| \geq c|F(J)|$ .

## 4. Bellman function proofs

### 4.1. Some history

We now present the Bellman-type proofs. Before we do this, we give a historical overview.

The Bellman function, related to investigation of weights on their own (i.e., not related to linear operators in weighted spaces), has been exploited in various papers. Topics such as reverse Hölder,  $L^p$  estimates, and distribution functions of  $A_p$  weights were investigated in [5], [17], [21]. In all these articles the Bellman function was found for continuous  $A_p$ . We enthusiastically refer the curious reader to these papers, since the search for a Bellman function and extremal examples are given there in details.

Aside from these three papers, the theory of BMO weights was developed in [18]. In this paper, together with the continuous BMO, the authors considered the dyadic case. The dyadic problem appeared to be much more delicate in some sense, and required a lot of additional calculations. In what follows, we use several parts of the dyadic proof from [18]. It appears that in our case the same steps give the proof. However, some parts of our proof are more delicate.

We now point out two difficulties that we have. First, the functions in [18] were explicit. In our case, as the reader will see, many ingredients are given implicitly, which makes things a little more complicated.

The main difficulty, though, is not the fact that we have implicit functions. In [18] the authors noticed that the domain of their Bellman function has the following property: it can be enlarged, with a good estimate of this “enlargement”, such that if the endpoints and center of some interval are in the smaller domain, then the whole interval is in the enlarged domain. This immediately implies that, if we do not care about sharp constants, we can get some nice estimates in the dyadic case directly from the continuous case.

In Remark 4.11 we prove that the domain of our Bellman function does not have this property. Therefore, without additional work, we can not make any dyadic statements. This means that we are “forced” to care about best constants and develop a variant of the proof from [18].

We also refer the reader to another paper, [23], treating a dyadic problem. There the authors obtained the exact Bellman function as well. However, the domain of their function was convex, and, therefore, obviously had the above property.

In [1], the authors introduced a particular function of two variables, that enables the proof of the continuous case of the inequality.

We sketch the definition and application of this function and discuss the main difficulty in the dyadic problem.

We will start with the Bellman function proofs of the summation conditions (inequalities (2.4)–(2.1)) since they are simpler than the proof of Theorem 1.7, which is much harder.



**4.2. Technical proposition**

**Proposition 4.1.** 1) Suppose  $\varepsilon > 0$ . For any monotone nonnegative function  $f(x)$  the following inequality holds for some absolute constant  $C$ :

$$\int_{-1}^1 (1 - |t|) f(x + \varepsilon t) dt \geq C f(x).$$

2) If  $A(x)$  satisfies

$$\int_{-1}^1 (1 - |t|) A''(x + \varepsilon t) dt \geq q A''(x)$$

then for some  $\alpha > 0$

$$A(x) - \frac{A(x - t) + A(x + t)}{2} + \alpha t^2 A''(x) \leq 0.$$

3) If  $A(x) = x^p$ ,  $p > 1$ , then for some  $\beta > 0$

$$A(x) - \frac{A(x - t) + A(x + t)}{2} + \beta t^2 A''(x) \geq 0$$

4) Let  $C > 1$  and  $A(x) = x^{-1/(p-1)}$ . Then there exists a number  $\beta$ , depending only on  $p$  and  $C$ , such that the following inequality holds for any  $t$ ,  $0 < t < \frac{C-1}{C} x$ :

$$A(x) - \frac{A(x + t) + A(x - t)}{2} + \beta t^2 A''(x) \geq 0.$$

Moreover, one can take

$$\beta = \frac{p-1}{p'} \left( \left( \frac{C}{C-1} \right)^2 \cdot \frac{\left( \frac{2C-1}{C} \right)^{-1/(p-1)} + \left( \frac{1}{C} \right)^{-1/(p-1)}}{2} - \left( \frac{C}{C-1} \right)^2 \right).$$

We prove Proposition 4.1 in Section 5.

**4.3. Bellman function proof of Lemma 2.2**

Recall that  $A(x)$  be a convex twice differentiable function on  $(0, \infty)$  such that for every  $x \in (0, \infty)$  the second derivative of  $A$  satisfies the inequality

$$(4.1) \quad \int_{-1}^1 (1 - |t|) A''(x + \varepsilon t) dt \geq Q A''(x)$$

with some positive constant  $Q$  uniformly on  $x \in (0, \infty)$  and  $\varepsilon$ .

Then for every weight  $w$  and any interval  $J$ ,

$$(4.2) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} (\langle w \rangle_{I^+} - \langle w \rangle_{I^-})^2 A''(\langle w \rangle_I) |I| \leq 8 \frac{1}{Q} (\langle A(w) \rangle_J - A(\langle w \rangle_J)).$$

*Proof.* Take a function of two variables  $B(u, v) = v - A(u)$ . Then, as we have proved,

$$B(u, v) - \frac{B(u - t, v - s) + B(u + t, v + s)}{2} = -\left(A(u) - \frac{A(u - t) + A(u + t)}{2}\right) \geq \alpha t^2 A''(u),$$

whenever  $B$  is defined at the points  $(u, v)$  and  $(u \pm t, v \pm s)$ .

Now we take a weight  $w$ . Then  $\langle w \rangle_{I_+} + \langle w \rangle_{I_-} = 2 \langle w \rangle_I$  and so  $\langle w \rangle_{I_{\pm}} = \langle w \rangle_I \pm t$ . Therefore,

$$B(\langle w \rangle_J, \langle A(w) \rangle_J) \geq \frac{1}{2} (B(\langle w \rangle_{J_+}, \langle A(w) \rangle_{J_+}) + B(\langle w \rangle_{J_-}, \langle A(w) \rangle_{J_-})) + \alpha (\Delta_J w)^2 A''(\langle w \rangle_J).$$

We rewrite this inequality in the following form:

$$|J| B(\langle w \rangle_J, \langle A(w) \rangle_J) \geq |J_+| B(\langle w \rangle_{J_+}, \langle A(w) \rangle_{J_+}) + |J_-| B(\langle w \rangle_{J_-}, \langle A(w) \rangle_{J_-}) + \alpha (\Delta_J |w|)^2 A''(\langle w \rangle_J) |J|.$$

Now we repeat this estimate down to the  $n$ th descendants of  $J$ . We denote this family by  $\mathcal{D}_n(J)$ . We get

$$|J| B(\langle w \rangle_J, \langle A(w) \rangle_J) \geq \sum_{I \in \mathcal{D}_n(J)} |I| B(\langle w \rangle_I, \langle A(w) \rangle_I) + \alpha \sum_{k \leq n} \sum_{I \in \mathcal{D}_k(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I|.$$

Using that  $B \geq 0$  whenever  $v \geq A(u)$ , which in our case is just Jensen’s inequality, we get

$$|J| B(\langle w \rangle_J, \langle A(w) \rangle_J) \geq \alpha \sum_{k \leq n} \sum_{I \in \mathcal{D}_k(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I|.$$

Since the last estimate is true for any  $n$ , we pass to the limit and get

$$|J| B(\langle w \rangle_J, \langle A(w) \rangle_J) \geq \alpha \sum_{I \in \mathcal{D}(J)} (\Delta_I w)^2 A''(\langle w \rangle_I) |I|.$$

This is exactly what we want. Our proof is finished. □

**4.4. Proof of the “inverse” Lemma 2.4**

*Proof.* We follow the scheme of the proof of Lemma 2.2. Take a function

$$B(u, v) = \alpha v + A(u).$$

Then  $B$  satisfies the following inequality:

$$\begin{aligned}
 B(x) - \frac{B(x+t, y+s-t^2 A''(x)) + B(x-s, y-s-t^2 A''(x))}{2} \\
 = A(x) - \frac{A(x-t) + A(x+t)}{2} + \alpha t^2 A''(x) \geq 0.
 \end{aligned}$$

The last inequality is true for  $A(x) = x^p$  or  $A(x) = x \log x$  without additional assumptions, or for  $A(x) = x^{-1/(p-1)}$ , if  $t < \frac{C-1}{C} x$ .

Now we take a weight  $w$ . If  $w$  is doubling (which we need only for the second part), then there exists a constant  $D(w)$ , such that for any dyadic interval  $J$  there holds

$$\langle w \rangle_J \leq D(w) \langle w \rangle_{J_{\pm}}.$$

If now  $\langle w \rangle_{J_{\pm}} = \langle w \rangle_J \pm t = x \pm t$ , then

$$x \leq D(w) (x - t),$$

which implies

$$t \leq \frac{C-1}{C} x$$

for  $C = D(w)$ .

Now we define

$$u_I = \langle w \rangle_I \quad \text{and} \quad v_I = \frac{1}{|I|} \sum_{R \in \mathcal{D}(I)} (\Delta_R(w))^2 A''(\langle w \rangle_R) |R|.$$

We notice that if  $u_{I_{\pm}} = u_I \pm t$  then

$$v_I - \frac{v_{I_+} + v_{I_-}}{2} = (\Delta_I(w))^2 A''(\langle w \rangle_I) = t^2 A''(\langle w \rangle_I) = t^2 A''(u_I).$$

So,  $v_{I_{\pm}} = v_I \pm s - t^2 A''(\langle w \rangle_I)$ . Therefore, by our inequality for  $B$ , we get

$$B(u_J, v_J) - \frac{B(u_{J_+}) + B(u_{J_-})}{2} \geq 0.$$

By the Bellman iteration procedure, we get

$$|J| B(u_J, v_J) \geq \sum_{I \in \mathcal{D}_n(J)} |I| B(u_I, v_I).$$

We now introduce a sequence of step functions. For a fixed  $n$  we take the family  $\{I: I \in \mathcal{D}_n(J)\}$ , and

$$u_n(t) = u_I, \quad t \in I, \quad v_n(t) = v_I, \quad t \in I.$$

Then the last inequality is the same as

$$|J| B(u_J, v_J) \geq \int_J B(u_n(t), v_n(t)) dt.$$

We now notice that  $B(u, v) = \alpha v + A(u) \geq A(u)$ , so

$$|J| B(u_J, v_J) \geq \int_J A(u_n(t)) dt.$$

We want to use the Fatou lemma. It is true for nonnegative functions, but it is clearly true as well for functions bounded from below (if the measure is finite). Indeed, our  $x^p$  and  $x^{-1/(p-1)}$  are nonnegative, and  $x \log x$  is bounded by  $-1/e$ . We also recall that in the “inverse” lemma we assume that  $A$  is bounded from below. Thus by Fatou’s lemma,

$$|J| B(u_J, v_J) \geq \liminf_n \int_J A(u_n(t)) dt \geq \int_J \liminf_n A(u_n(t)) dt = \int_J A(w(t)) dt.$$

The last equality is true since for almost every  $t$ , by the Lebesgue differentiation theorem, we have  $u_n(t) \rightarrow w(t)$ , and because  $A$  is a continuous function.

Dividing by  $|J|$ , we finally get  $\alpha v_J + A(u_J) \geq \langle A(w) \rangle_J$ , which finishes our proof. □

**4.5. Proof of main Theorem 1.7**

**4.5.1. Notation and definition of the function  $B$ .** For a point  $z = (x, y) \in \mathbb{R}^2$  we let  $[z] = xe^{-y}$ . For any number  $Q > 1$ , we define the domain  $\Omega_Q$  by:

$$\Omega_Q = \{z = (x, y) : 1 \leq [z] \leq Q\}.$$

The boundaries of  $\Omega_Q$  are

$$\Gamma = \{z : [z] = 1\} \quad \text{and} \quad \Gamma_Q = \{z : [z] = Q\}.$$

With any point  $z \in \Omega_Q$  we associate the two numbers  $v$  and  $a$ . We take our point  $z$  and consider the line  $\ell(z)$ , tangent to  $\Gamma_Q$ , that “kisses”  $\Gamma_Q$  on the right-hand side from  $z$ , and such that  $z \in \ell(z)$ . The point  $\ell(z) \cap \Gamma_Q$  is denoted by  $(a, \log(a/Q))$ . Now we extend  $\ell(z)$  to the left until it intersects  $\Gamma$ . The point of intersection is denoted by  $(v, \log(v))$ . Notice that  $v \leq x \leq a$ .

More carefully, let  $\gamma = \gamma(Q)$ ,  $\gamma \leq 1$ , be the smaller solution of equation

$$\gamma - \log(\gamma) - 1 = \log(Q).$$

Then the line  $\ell(z)$  is given by a formula

$$y = \frac{\gamma \cdot x}{v} + \log(v) - \gamma.$$

This equation defines a unique  $v$  such that  $v \leq x$ . Moreover,  $a$  is given by  $v = \gamma \cdot a$ . We are ready to introduce the Bellman Function. We give an explicit formula:

$$B_Q(z) = B_Q(x, y) = x \cdot \log(v) + \frac{x - v}{\gamma}.$$

**Remark 4.2.** The equation for  $t$ ,  $t - \log(t) = \log(u)$  is famous and well developed. In the mathematical program Maple this solution can be obtained using the command

$$-\text{LambertW}\left(-\frac{1}{u}\right).$$

Several of the inequalities in subsequent sections can be checked by graphing related functions. The second author wants to emphasize his gratitude to the developers of Maple.

**4.5.2. Main theorems and discussion.** The following theorem was proved in [1].

**Theorem 4.3.** *The function  $B_Q(z)$  has following properties:*

1.  $B_Q(v, \log(v)) = v \log(v)$ .
2.  $B$  is smooth in  $\Omega_Q$ , and locally concave in  $\Omega_Q$ . Namely, if  $z_1, z_2 \in \Omega_Q$ ,  $z = sz_1 + (1-s)z_2$  for some  $s \in [0, 1]$ , and  $\{tz_1 + (1-t)z_2\} \subset \Omega_Q$  then

$$B(z) \geq sB(z_1) + (1-s)B(z_2).$$

3. For every point  $z = (x, y) \in \Omega_Q$  there exists a function  $w$ ,  $[w]_{A_\infty^d} \leq Q$ , such that  $\langle w \rangle = x$ ,  $\langle \log(w) \rangle = y$ , and  $\langle w \log(w) \rangle = B(x, y)$ .

This theorem implies the following (see [1]).

**Theorem 4.4.** *The following equality holds:*

$$(4.3) \quad B_Q(x, y) = \sup \{ \langle w \log w \rangle : \langle w \rangle = x, \langle \log w \rangle = y, [w]_{A_\infty^d} \leq Q \}.$$

We sketch the proof of this theorem. We warn the reader that we skip some details. We emphasize those parts of the proof that fail in the dyadic setting and that need more careful treatment.

*Proof.* The third property of  $B$  implies that  $B_Q(x, y)$  is not strictly bigger than the right-hand side of (4.3). For the other direction, we take a point  $z = (x, y) \in \Omega_Q$  and a function  $w$ , such that  $\langle w \rangle = x$ ,  $\langle \log(w) \rangle = y$ , and  $[w]_{A_\infty^d} \leq Q$ . Now we take two intervals  $I_\pm$ , such that  $I_+ \cup I_- = I$ ,  $I_+ \cap I_- =$  right end of  $I_-$ . We take

$$(4.4) \quad z_\pm = (x_\pm, y_\pm) = \left( \langle w \rangle_{I_\pm}, \langle \log(w) \rangle_{I_\pm} \right) \in \Omega_Q.$$

Assuming that the interval  $[z_-, z_+]$  lies in  $\Omega_Q$ , we write

$$B(z) \geq \frac{|I_-|}{|I|} B(z_-) + \frac{|I_+|}{|I|} B(z_+).$$

Repeating this procedure, we split  $I_\pm$  in two intervals, etc. On the  $n$ th step we obtain  $N = 2^n$  intervals, which we denote by  $I^n$ . Using the concavity of  $B$ , we write

$$B(z) \geq \sum_{n=1}^N \frac{|I^n|}{|I|} B(z_n),$$

where  $z_n = (\langle w \rangle_{I^n}, \langle \log(w) \rangle_{I^n})$ . We now introduce a pair of step functions. Let

$$u_N(t) = \sum_{n=1}^N \langle w \rangle_{I^n} \chi_{I^n}(t) \quad \text{and} \quad v_N(t) = \sum_{n=1}^N \langle \log(w) \rangle_{I^n} \chi_{I^n}(t).$$

Then we have

$$B(z) \geq \int_I B(u_N(t), v_N(t)) dt.$$

If  $w$  is separated from 0 and  $\infty$  then, by the Lebesgue differentiation theorem, we get that

$$u_N(t) \rightarrow w(t) \text{ a.e.} \quad \text{and} \quad v_N(t) \rightarrow \log(w(t)) \text{ a.e.}$$

Therefore,

$$B(z) \geq \int_I B(w(t), \log(w(t))) dt = \int_I w(t) \log(w(t)) dt = \langle w \log(w) \rangle.$$

In the chain above we used that  $B$  is bounded on compact sets, so we can apply the Lebesgue dominated convergence theorem, and the second property of the function  $B$ . The proof is finished. □

**Remark 4.5.** A careful reader can see two gaps in the proof above. First, we never introduced a proper procedure for choosing the intervals  $I_{\pm}$ . Second, we focused on bounded and separated from 0 functions  $w$  without saying anything about the general case. We refer to the paper [1], where all details are given.

**Remark 4.6.** We explain the main difficulty in the dyadic case. In the proof above we had a formula (4.4). We claimed that  $z_{\pm} = (x_{\pm}, y_{\pm}) = (\langle w \rangle_{I_{\pm}}, \langle \log(w) \rangle_{I_{\pm}}) \in \Omega_Q$ . It was true for *any* intervals  $I_{\pm}$ , thus we had a lot of freedom in choosing these intervals  $I_{\pm}$ . This was used in [1]. In the dyadic case we are *forced* to take  $I_{\pm}$  to be *dyadic* intervals! Therefore, we do not have any procedure for choosing  $I_{\pm}$  except for splitting  $I$  in two halves, etc. The main problem now is that we can never be sure that the segment  $[z_-, z_+]$  lies entirely in the domain  $\Omega_Q$ .

Now we state the main theorem, that works in the dyadic setting. Let  $B_Q$  be the function, described above, defined in the domain  $\Omega_Q$ . For any  $Q_0 > Q$  we define  $\Omega_{Q_0} = \Omega_0$ ,  $\gamma_{Q_0} = \gamma_0$ ,  $v_{Q_0} = v_0$ ,  $a_{Q_0} = a_0$ , and  $B_{Q_0}(z) = B_0(z)$ , as we did for  $Q$ .

**Theorem 4.7.** *There exists a constant  $C$ , which does not depend on  $Q$ , and a number  $Q_0$ , such that  $1 < Q < Q_0 < CQ$ , and such that the function  $B_0$  has the following additional property: whenever  $z, z_+, z_- \in \Omega_Q$ , and  $z = (z_+ + z_-)/2$ , the following inequality holds:*

$$2B_0(z) \geq B_0(z_+) + B_0(z_-).$$

If  $r = \sqrt{1 - 1/Q}$  then  $Q_0$  is given by equation

$$(1 - r) \log(\gamma_0) + \frac{1 - r}{\gamma_0} - (1 - r) - (1 - r) \log(1 - r) - (1 + r) \log(1 + r) = 0.$$

**Remark 4.8.** Notice that this equation defines  $\gamma_0$ , which immediately defines  $Q_0$ .

**Remark 4.9.** We claim that we can take a larger domain and a function  $B_{Q_0}$ , which is bigger than  $B_Q$ , and which has the property: if the three points  $z, z_{\pm}$ , described above, lie in the *small* domain  $\Omega_Q$ , then  $2B_0(z) \geq B_0(z_+) + B_0(z_-)$ , even though the interval  $[z_-, z_+]$  does not lie even in  $\Omega_{Q_0}$ .

The fact that the solution  $Q_0$  of the equation above can be bounded by  $CQ$  will be proved later. To emphasize the difficulty of the problem we prove a lemma, that shows the difference of our problem from the problem solved in [18].

**Lemma 4.10.** *For any constant  $C > 0$ , there exists a number  $Q > 1$ , and three points  $z, z_+, z_- \in \Omega_Q$ , such that  $2z = z_+ + z_-$  and such that for some value of  $t \in (0, 1)$  we have*

$$[t z_+ + (1 - t)z_-] \geq C Q.$$

**Remark 4.11.** Lemma 4.10 shows that for any given  $Q_0 = CQ$  there can be three points in  $\Omega_Q$ , such that the interval  $[z_-, z_+]$  does not lie entirely in  $\Omega_{Q_0}$ .

We note that in [18] there existed a constant  $C$ , such that

$$z_{\pm}, z = \frac{z_+ + z_-}{2} \in \Omega_Q \implies [z_-, z_+] \subset \Omega_{CQ}.$$

This did not simplify the search for the best constant, but would immediately give us linear dependence on  $[w]_{A_{\infty}^d}$ . If this were the case, we would simply take  $Q_0 = CQ$ , and since  $z_{\pm}, z \in \Omega_Q$  implies  $[z_-, z_+] \in \Omega_{Q_0}$ , where  $B_{Q_0}$  is locally concave, we would get  $2B_{Q_0}(z) \geq B_{Q_0}(z_+) + B_{Q_0}(z_-)$ .

Since such a  $C$  does not exist, we are forced to continue our investigation.

*Proof.* This is an easy calculation. In fact, these points are  $z_- = (1 - r, \log \frac{1-r}{Q})$ ,  $z = (1, \log \frac{1}{Q})$ , and  $z_+ = (1 + r, \log(1 + r))$ . □

As a consequence of Theorem 4.7 we get the following.

**Theorem 4.12.** *For every point  $z = (x, y) \in \Omega_Q$  the following inequality holds:*

$$B_0(z) \geq \sup \{ \langle w \log(w) \rangle : \langle w \rangle = x, \langle \log(w) \rangle = y, [w]_{A_{\infty}^d} \leq Q \}.$$

*Proof.* We take a point  $z$ ,  $1 \leq [z] \leq Q$ , and a function  $w$ , such that  $\langle w \rangle = x$ ,  $\langle \log(w) \rangle = y$ , and  $[w]_{A_{\infty}^d} \leq Q$ .

*Case 1:  $w$  is bounded away from 0 and  $\infty$ .* We take  $I^0 = I$ , and we take  $I_{1,2}^1$  to be the left and right halves of  $I$ . Then  $I_{1,2,3,4}^2$  are quarters of  $I$ , etc. For every  $k$  and  $n$  we have  $z_n^k (\langle w \rangle_{I_n^k}, \langle \log(w) \rangle_{I_n^k}) \in \Omega_Q$ , and every  $z_n^k$  is the center of an interval that corresponds to a “child” of  $I_n^k$ . Therefore, for a fixed  $k$ ,

$$B_0(z) \geq \sum_n \frac{|I_n^k|}{|I|} B_0(z_n^k) = \int_I B_0(u_k(t), v_k(t)) dt,$$

where

$$u_k(t) = \sum_n \langle w \rangle_{I_n^k} \chi_{I_n^k}(t), \quad \text{and} \quad v_k(t) = \sum_n \langle \log(w) \rangle_{I_n^k} \chi_{I_n^k}(t).$$

Since  $w$  is separated from 0 and  $\infty$ , we get

$$u_k(t) \rightarrow w(t), \text{ a.e.} \quad \text{and} \quad v_k(t) \rightarrow \log(w(t)), \text{ a.e.}$$

Since we have countably many intervals  $\{I_n^k\}_{n,k}$ , the set  $Z = \{z_n^k\}$  is compact, and thus the function  $B_0$  is bounded on  $Z$ . Therefore, by the Lebesgue Dominated Convergence Theorem,

$$B_0(z) \geq \langle w \log(w) \rangle.$$

*Case 2: arbitrary  $w$ .* Here we sketch the proof, as it is the same as in [1]. We take

$$w_n(t) = \begin{cases} n, & w(t) \geq n \\ w(t), & 1 \leq w(t) \leq n \\ 1, & w(t) \leq 1. \end{cases}$$

Then, as follows from [16],  $[w_n]_{A_\infty^d} \leq Q$ . By the case 1 we get

$$B_0(z) \geq \langle w_n \log(w_n) \rangle = \int_{\{t: w(t) \geq 1\}} w_n \log(w_n).$$

On the set  $\{t: w(t) \geq 1\}$  the sequence  $w_n(t) \log(w_n)(t)$  increases to  $w(t) \log(w(t))$ , and passing to the limit, we get

$$B_0(z) \geq \int_{\{t: w(t) \geq 1\}} w(t) \log(w(t)) dt \geq \int_I w(t) \log(w(t)) dt = \langle w \log(w) \rangle.$$

The last inequality holds simply because on the set  $w(t) < 1$  we have

$$w(t) \log(w(t)) \leq 0. \quad \square$$

The rest of this section is devoted to the proof of the Theorem 4.7. The impatient reader can skip this proof since it does not involve any weight theory.

**4.5.3. Proof of the Theorem 4.7: reminder.** First we would like to recall some notation. We fix a number  $Q, Q > 1$ . In what follows the number  $Q_0$  is always bigger than  $Q$ .

For every point  $z = (x, y)$  such that  $xe^{-y} \in [1, Q]$  we write  $[z] = xe^{-y}$ . Moreover, numbers  $\gamma_0, v = v_0$ , and  $a = a_0$  are defined implicitly by

$$\gamma_0 - \log(\gamma_0) = 1 + \log(Q_0), \quad y = \frac{\gamma_0 \cdot x}{v} + \log(v) - \gamma_0, \quad \text{and} \quad a = \frac{v}{\gamma_0}.$$

In what follows points the  $z_\pm$  are such that  $2z = z_+ + z_-$ , and  $v_\pm$  and  $a_\pm$  are defined as above for these points. Our “larger” function is defined by

$$B_0(x, y) = x \cdot \log(v) + \frac{x - v}{\gamma_0}.$$

Furthermore,

$$\Gamma_Q = \{z: [z] = Q\}, \quad \Gamma_{Q_0} = \{z: [z] = Q_0\}, \quad \text{and} \quad \Gamma = \Gamma_1 = \{z: [z] = 1\}.$$

We sometimes refer to  $\Gamma_Q$  as a  $Q$ -boundary and to  $\Gamma_{Q_0}$  as a  $Q_0$ -boundary.



We start with the following easy lemma.

**Lemma 4.13.** *Suppose*

$$F(x, y, x_+, y_+, x_-, y_-) = 2B_0(x, y) - B_0(x_-, y_-) - B_0(x_+, y_+).$$

If  $F(x, y, x_+, y_+, x_-, y_-) \geq 0$ , then for every number  $C > 0$ , there holds

$$F(Cx, y + \log(C), Cx_+, y_+ + \log(C), Cx_-, y_- + \log(C)) \geq 0.$$

*Proof.* This follows immediately from the homogeneity of  $B_0$ , namely,  $B_0(Cx, y + \log(C)) = Cx \log(C) + CB_0(x, y)$ .  $\square$

Lemma 4.13 allows us to choose  $C = 1/x$  and always think that  $x = 1$ .

We first start with positions of  $z$  and  $z_{\pm}$  that are “worst” in some sense. Later we shall see that in fact the next section is not needed at all. However, for the sake of completeness we keep it.

**4.5.3.1. Remark about notation.** Abusing notation, we always denote by  $\Delta$  the expression

$$\Delta = 2B_0(z) - B_0(z_+) - B_0(z_-).$$

However, in different sections the same letter  $\Delta$  will depend on (and be differentiated in) different variables. We will always specify on which variables it depends.

**4.5.4. Proof of Theorem 4.7. First step.** We start our investigation with the case when  $z_{\pm}$  and  $z$  are on the boundary of  $\Omega_Q$ . Since  $z$  and  $z_{\pm}$  are fixed,  $\Delta$  depends only on  $Q_0$ .

Our first case is when two of these points are on  $\Gamma_Q$  and the third is on  $\Gamma$ . Our second case is when two of them are on  $\Gamma$  and the third is on  $\Gamma_Q$ . Moreover  $z \in \Gamma_Q$  always.

We remind the reader that in 4.5.4.1, 4.5.4.2 and 4.5.4.3 we will have **different**  $\Delta(Q)$ . The reason is that we are going to fix different points  $z_{\pm}$  and  $z$ .

**4.5.4.1.  $z_- \in \Gamma_Q$  and  $z_+ \in \Gamma$ .** We have  $z = (1, \log(1/Q))$ .

We define  $z_+ = (1+r, \log(1+r))$  and  $z_- = (1-r, \log((1-r)/Q))$ ,  $r \geq 0$ . Then, since  $2y = \frac{y_+ + y_-}{2}$ , we obtain  $2 \log(1/Q) = \log((1-r^2)/Q)$ , so  $r^2 = 1 - 1/Q$ , and thus  $r = \sqrt{1 - 1/Q}$ . Then we have:

$$z_- = \left(1 - r, \log \frac{1-r}{Q}\right), \quad z = \left(1, \log \frac{1}{Q}\right), \quad z_+ = (1 + r, \log(1 + r)).$$

**Theorem 4.14.** *Take  $Q_0 = Q$ . Then we get  $\gamma_0 = \gamma$ ,  $B_0 = B$ , and  $v$ , associated with  $Q$ . Define*

$$\Delta = \Delta(Q) = 2B(z) - B(z_-) - B(z_+).$$

*Then  $\Delta \geq 0$ .*

We notice that now  $\Delta$  depends on  $Q$ , and the variable  $Q$  is not smaller than 1. Theorem 4.14, together with next lemma, gives us what we want.

**Lemma 4.15.** *For fixed points  $z, z_{\pm} \in \Omega_Q, \Delta(Q_0) = 2B_0(z) - B_0(z_-) - B_0(z_+)$  is an increasing function with respect to  $Q_0$  on the set  $\{Q_0: Q_0 \geq Q\}$ .*

Lemma 4.15 shows that if our initial  $B$  was “concave” enough, then the “enlarged”  $B_0$  is also “concave” enough.

*Proof of Lemma 4.15.* By definition,

$$z_- = \left(1 - r, \log \frac{1-r}{Q}\right), \quad z = \left(1, \log \frac{1}{Q}\right), \quad z_+ = (1 + r, \log(1 + r)).$$

We have points  $v, v_{\pm} \in \Gamma$ , associated with  $z, z_{\pm}$  and calculated in the enlarged domain. Namely,

$$(4.5) \quad \gamma_0 - \log(\gamma_0) = 1 + \log(Q_0)$$

$$(4.6) \quad \log \frac{1}{Q} = \frac{\gamma_0}{v} + \log(v) - \gamma_0$$

$$(4.7) \quad \log \frac{1-r}{Q} = \frac{\gamma_0(1-r)}{v_-} + \log(v_-) - \gamma_0,$$

$$(4.8) \quad v_+ = 1 + r.$$

In particular we see that  $v_- = (1 - r)v$ . Since

$$B_0(z) = x \log(v) + \frac{x - v}{\gamma_0},$$

and since  $2x - x_+ - x_- = 0$ , one gets

$$\begin{aligned} \Delta(Q_0) &= 2 \log v - (1 - r) \log(v_-) - (1 + r) \log(v_+) - \frac{1}{\gamma_0} (2v - v_- - v_+) \\ &= 2 \log v - (1 - r) \log(v) - (1 - r) \log(1 - r) - (1 + r) \log(1 + r) \\ &\quad - \frac{1}{\gamma_0} (2v - (1 - r)v - (1 + r)) \\ &= (1 + r) \left( \log(v) - \frac{v - 1}{\gamma_0} \right) - (1 - r) \log(1 - r) - (1 + r) \log(1 + r). \end{aligned}$$

The last two terms do not depend on  $Q_0$  at all, so we consider only

$$f(Q_0) = \log(v) - \frac{v - 1}{\gamma_0}.$$

We clearly have

$$\gamma'_0 - \frac{\gamma'_0}{\gamma_0} = \frac{1}{Q_0},$$

so

$$\gamma'_0 = \frac{\gamma_0}{(\gamma_0 - 1)Q_0}.$$

Differentiating the equality

$$\log \frac{1}{Q} = \frac{\gamma_0}{v} + \log(v) - \gamma_0$$

with respect to  $Q_0$ , we get

$$0 = \frac{\gamma'_0}{v} - \frac{\gamma_0}{v^2}v' + \frac{v'}{v} - \gamma'_0,$$

so

$$\begin{aligned} 0 &= \frac{v'}{v} \left(1 - \frac{\gamma_0}{v}\right) - \gamma'_0 \left(1 - \frac{1}{v}\right), & \frac{v'}{v} \frac{v - \gamma_0}{v} &= \frac{v - 1}{v} \frac{\gamma_0}{(\gamma_0 - 1)Q_0}, \\ \frac{v'}{v} &= \frac{1 - v}{v - \gamma_0} \frac{1}{1 - \gamma_0} \frac{\gamma_0}{Q_0}. \end{aligned}$$

Now we differentiate  $f(Q_0)$ . We recall that

$$f(Q_0) = \log(v) - \frac{v - 1}{\gamma_0} = \log(v) + \frac{1 - v}{\gamma_0},$$

so

$$\begin{aligned} f'(Q_0) &= \frac{v'}{v} + \frac{-v'\gamma_0 - \gamma'_0(1 - v)}{\gamma_0^2} \\ &= \frac{1 - v}{v - \gamma_0} \frac{1}{1 - \gamma_0} \frac{\gamma_0}{Q_0} - \frac{v'}{\gamma_0} - \frac{\gamma_0(1 - v)}{(\gamma_0 - 1)Q_0} \frac{1}{\gamma_0^2} \\ &= \frac{1 - v}{v - \gamma_0} \frac{1}{1 - \gamma_0} \frac{\gamma_0}{Q_0} - \frac{1 - v}{v - \gamma_0} \frac{1}{1 - \gamma_0} \frac{v}{Q_0} + \frac{1 - v}{1 - \gamma_0} \frac{1}{Q_0\gamma_0} \\ &= \frac{1 - v}{(1 - \gamma_0)Q_0} \left( \frac{\gamma_0}{v - \gamma_0} - \frac{v}{v - \gamma_0} + \frac{1}{\gamma_0} \right) = \frac{1 - v}{(1 - \gamma_0)Q_0} \left( \frac{1}{\gamma_0} - 1 \right) \geq 0, \end{aligned}$$

since  $v < 1$  and  $\gamma_0 < 1$ . This finishes the proof. □

*Proof of Theorem 4.14.* We return to  $Q$ ,  $\gamma$ ,  $B$ , and  $v$ , calculated for  $\gamma$ . We recall that in the statement of the theorem,  $Q_0 = Q$ .

Recall that

$$z_- = \left(1 - r, \log \frac{1 - r}{Q}\right), \quad z = \left(1, \log \frac{1}{Q}\right), \quad z_+ = (1 + r, \log(1 + r)).$$

Our  $v$  and  $v_{\pm}$  can be written explicitly in terms of  $\gamma$ . Indeed,

$$v_- = \gamma(1 - r) \quad v = \gamma \quad \text{and} \quad v_+ = 1 + r.$$

Then

$$\begin{aligned} \Delta &= 2B(z) - B(z_-) - B(z_+) \\ &= 2 \left( \log(\gamma) + \frac{1-\gamma}{\gamma} \right) - \left( (1-r) \log(\gamma(1-r)) + \frac{1-r-(1-r)\gamma}{\gamma} \right) \\ &\quad - (1+r) \log(1+r) \\ &= 2 \log(\gamma) + \frac{2}{\gamma} - 2 - (1-r) \log(\gamma) - (1-r) \log(1-r) - \frac{1-r}{\gamma} + (1-r) \\ &\quad - (1+r) \log(1+r) \\ &= (1+r) \log(\gamma) + \frac{1+r}{\gamma} - (1+r) - (1-r) \log(1-r) - (1+r) \log(1+r). \end{aligned}$$

We notice that  $1-r^2 = 1/Q$ , so  $\log(1+r) = \log \frac{1}{Q} - \log(1-r)$ . Therefore,

$$\begin{aligned} \Delta &= (1+r) \left( \log(\gamma) + \frac{1}{\gamma} - 1 - \log \frac{1}{Q} \right) - (1-r) \log(1-r) + (1+r) \log(1-r) \\ &= (1+r) \left( \log(\gamma) + \log(Q) + \frac{1}{\gamma} - 1 \right) + 2r \log(1-r) \\ &= (1+r) \left( \gamma + \frac{1}{\gamma} - 2 \right) + 2r \log(1-r). \end{aligned}$$

We would like to know that  $\Delta \geq 0$ . Surprisingly, we can show it. Here is the chain of awful estimates. We define

$$f(Q) = \Delta.$$

Notice that

$$\gamma'(Q) = \frac{\gamma}{Q(\gamma-1)} \quad \text{and} \quad r'(Q) = \frac{1}{2rQ^2}.$$

The second equality is true since  $r^2 - 1 = -1/Q$ .

We notice that if  $Q = 1$  then  $r = 0$  and  $\gamma = 1$ , so  $f(1) = 0$ . We claim that  $f'(Q) \geq 0$ , which will give the desired result.

We have

$$f'(Q) = \left( \gamma + \frac{1}{\gamma} - 2 \right) \frac{1}{2rQ^2} + (1+r) \left( 1 - \frac{1}{\gamma^2} \right) \frac{\gamma}{Q(\gamma-1)} + \left( 2 \log(1-r) - \frac{2r}{1-r} \right) \frac{1}{2rQ^2}.$$

We notice that  $\gamma + 1/\gamma - 2 \geq 0$  and we discard it. Therefore,

$$\begin{aligned} f'(Q) &\geq (1+r) \left( 1 - \frac{1}{\gamma^2} \right) \frac{\gamma}{Q(\gamma-1)} + \left( 2 \log(1-r) - \frac{2r}{1-r} \right) \frac{1}{2rQ^2} \\ &= \frac{1+r}{Q} \frac{\gamma+1}{\gamma} + \frac{\log(1-r)}{rQ^2} - \frac{1}{Q^2} \frac{1}{1-r}. \end{aligned}$$

We now use that

$$\frac{1}{1-r} = \frac{1+r}{1-r^2} = Q(1+r),$$

thus

$$\begin{aligned} f'(Q) &\geq \frac{1}{Q} \left[ 1 + r + \frac{1+r}{\gamma} + \frac{\log(1-r)}{rQ} - \frac{Q(1+r)}{Q} \right] \\ &= \frac{1}{Q} \left[ 1 + r + \frac{1}{Q\gamma(1-r)} + \frac{\log(1-r)}{rQ} - (1+r) \right] \\ &= \frac{1}{Q^2 r} \left[ \frac{r}{\gamma(1-r)} + \log(1-r) \right]. \end{aligned}$$

Finally,  $0 \leq 1-r < 1$ , so  $r/(\gamma(1-r)) > r/\gamma$ , and therefore

$$f'(Q) \geq \frac{1}{Qr^2} \left[ \frac{r}{\gamma} + \log(1-r) \right].$$

We now define

$$g(Q) = \frac{r}{\gamma} + \log(1-r).$$

Again,  $g(1) = 0$ . We are going to prove that  $g'(Q) \geq 0$ . Indeed,

$$g'(Q) = \frac{1}{\gamma} \frac{1}{2rQ^2} - \frac{r}{\gamma^2} \frac{\gamma}{Q(\gamma-1)} - \frac{1}{1-r} \frac{1}{2rQ^2} = \frac{1}{2rQ^2} \left[ \frac{1}{\gamma} - \frac{1}{1-r} - \frac{2r^2Q}{\gamma(\gamma-1)} \right].$$

But  $r^2Q = (1 - 1/Q)Q = Q - 1$ , which implies

$$\begin{aligned} g'(Q) &= \frac{1}{2rQ^2} \left[ \frac{1}{\gamma} - \frac{1}{1-r} - (Q-1) \frac{2}{\gamma(\gamma-1)} \right] \\ &= \frac{1}{2rQ^2} \left[ \frac{1}{\gamma} - \frac{1}{1-r} - 2(Q-1) \left( \frac{1}{\gamma-1} - \frac{1}{\gamma} \right) \right] \\ &= \frac{1}{2rQ^2} \left[ \frac{1}{\gamma} - \frac{1}{1-r} - 2 \frac{Q-1}{\gamma-1} + \frac{2Q}{\gamma} - \frac{2}{\gamma} \right] \end{aligned}$$

Again  $1/(1-r) = Q(1+r)$ , so

$$2rQ^2 \cdot g'(Q) = -\frac{1}{\gamma} + 2 \frac{Q-1}{1-\gamma} + \frac{2Q}{\gamma} - Q(1+r) = \frac{Q-1}{\gamma} + 2 \frac{Q-1}{1-\gamma} + Q \left( \frac{1}{\gamma} - r - 1 \right).$$

The first two terms are clearly nonnegative. To check that the last one is nonnegative we do the following. The number  $\gamma$  satisfies the equation  $\varphi(t) - \log(Q) = 1$ , where  $\varphi(t) = t - \log(t)$ . The function  $\varphi$  is a decreasing function if  $t \in (0, 1]$ . So if we prove that  $\varphi(1/(1+r)) - \log(Q) \leq 1$  then we get that  $1/(1+r) \geq \gamma$ , which means that  $1/\gamma \geq 1+r$ . Thus,

$$\varphi\left(\frac{1}{1+r}\right) - \log(Q) = \frac{1}{1+r} + \log(1+r) - \log(Q);$$

the derivative of this expression with respect to  $Q$  is

$$\left( \frac{1}{1+r} - \frac{1}{(1+r)^2} \right) \frac{1}{2rQ^2} - \frac{1}{Q} = \frac{1}{Q} \left[ \frac{r}{(1+r)^2} \frac{1}{2rQ^2} - 1 \right] = \frac{1}{Q} \left[ \frac{1}{2Q^2(1+r)^2} - 1 \right].$$

Since  $Q > 1$  and  $r > 0$ , we have  $2Q^2(1+r)^2 > 2$ , so the derivative is negative, and therefore

$$\varphi\left(\frac{1}{1+r}\right) - \log(Q) \leq \varphi(1) = 1.$$

This completes the proof. □

**4.5.4.2.**  $z_- \in \Gamma, z_+ \in \Gamma_Q$ . In this case we still have  $r = \sqrt{1 - 1/Q}$ , but

$$z_- = (1 - r, \log(1 - r)), \quad z = \left(1, \log \frac{1}{Q}\right), \quad z_+ = \left(1 + r, \log \frac{1+r}{Q}\right),$$

and

$$v_- = 1 - r, \quad v = \gamma, \quad v_+ = \gamma(1 + r).$$

So,

$$\begin{aligned} \Delta(Q) &= 2 \log(\gamma) + \frac{2}{\gamma} - 2 - (1 - r) \log(1 - r) \\ &\quad - \left( (1 + r) \log(\gamma(1 + r)) + \frac{1 + r - (1 + r)\gamma}{\gamma} \right) \\ &= 2 \log(\gamma) + \frac{2}{\gamma} - 2 - (1 - r) \log(1 - r) - (1 + r) \log(\gamma) \\ &\quad - (1 + r) \log(1 + r) - \frac{1 + r}{\gamma} + (1 + r) \\ &= (1 - r) \log(\gamma) + \frac{1 - r}{\gamma} - (1 - r) - (1 - r) \log(1 - r) - (1 + r) \log(1 + r). \end{aligned}$$

Unfortunately, this expression is negative. To prove it one can take  $Q$  very large and calculate the asymptotics of the last expression. We have no intention to do it. However, the interested reader can draw the graph of  $\Delta(Q)$  in, say, Maple, and see that the function is negative. We now fix our choice of  $Q_0$ .

**Definition 4.16.** We define  $Q_0$  as the solution of  $\Delta(Q_0) = 0$  such that  $Q_0 \geq Q$ .

We notice that this choice of  $Q_0$  is as in Theorem 4.7.

This definition leaves two questions: whether such a  $Q_0$  exists and, more complicated, whether there is a uniform estimate  $Q_0 \leq C \cdot Q$ , where  $C$  does not depend on  $Q$ . Fortunately the answers to both questions are affirmative.

**Lemma 4.17.** *If  $z$  and  $z_{\pm}$  are as above, then for every point  $(u, v) \in [z_-, z_+]$  the following holds:  $u e^{-v} \leq CQ$ , where  $C$  is some uniform constant.*

This lemma shows that if we take  $Q_0 = CQ$  then the function  $B_{Q_0}$  will be locally concave in the domain  $\Omega_{Q_0}$ , and the line segment  $[z_-, z_+]$  lies in this domain. Since  $\Delta(Q) \leq 0$  and  $\Delta(CQ) \geq 0$ , we immediately get that between  $Q$  and  $CQ$  there is some  $Q_0$  for which  $\Delta(Q_0) = 0$ . We prove Lemma 4.17 in Section 5.

**4.5.4.3.  $z_{\pm} \in \Gamma$ .** In this case we change our choice of  $r$ . We have  $z_{\pm} = (1 \pm r, \log(1 \pm r))$  and  $z = (1, \log(1/Q))$ . Since  $\log(1 - r^2) = 2 \log(1/Q)$ , we get  $1 - r^2 = 1/Q^2$ , or  $r = \sqrt{1 - 1/Q^2}$ .

As in the first case, we prove two propositions.

**Lemma 4.18.**  $\Delta(Q) \geq 0$  and for every  $Q_0 \geq Q$  we have  $\Delta(Q_0) \geq \Delta(Q)$ .

*Proof.* We start with the second fact. We always have  $v_{\pm} = 1 \pm r$ , and so

$$\Delta(Q_0) = 2 \log v + 2 \frac{1-v}{\gamma_0} - (1-r) \log(1-r) - (1+r) \log(1+r).$$

We have already seen that the sum of first two terms increases when  $Q_0$  increases, and last two terms do not depend on  $Q_0$ .

For the first part, notice that when  $Q_0 = Q$  we have  $v = \gamma$ , and so

$$\Delta(Q) = 2 \log(\gamma) + 2 \frac{1-\gamma}{\gamma} - (1-r) \log(1-r) - (1+r) \log(1+r).$$

We have

$$\gamma' = \frac{\gamma}{Q(\gamma-1)}, \quad r' = \frac{1}{rQ^3}.$$

The second equality is new because  $r$  is different from what it was in cases 1 and 2. So,

$$\Delta' = \frac{2}{Q\gamma} + \frac{1}{rQ^3} \log \frac{1-r}{1+r} = \frac{2}{\gamma Q} \frac{1}{rQ^3} \log \frac{(1-r)^2}{1-r^2} = \frac{2}{Q} \left[ \frac{1}{\gamma} + \frac{\log(Q - \sqrt{Q^2 - 1})}{Q\sqrt{Q^2 - 1}} \right].$$

We leave the proof that this expression is positive as an easy exercise. We note that for large values of  $Q$  the second term in brackets is negative, and so the first term “pulls” the whole expression above zero. Finally, we get  $\Delta(Q) \geq \Delta(1) = 0$ , and we are done.  $\square$

**4.5.5. Proof of Theorem 4.7: change of variables.**

**4.5.5.1. Discussion.** We remind the reader that in the general case we basically have four variables:  $x_{\pm}$  and  $y_{\pm}$ . Then the center point  $z = (1, y)$  is given by  $2 = x_+ + x_-$  and  $2y = y_+ + y_-$ . The first equation lets us get rid of  $x_-$ , and so we have three variables:  $x_+$ ,  $y_-$ , and  $y_+$ . These variables have rather a complicated domain. Here are the inequalities that define this domain:

$$x_+ e^{-y_+} \in [1, Q], \quad (2 - x_+) e^{-y_-} \in [1, Q] \quad \text{and} \quad e^{-(y_+ + y_-)/2} \in [1, Q].$$

This domain is somewhat inconvenient for us. The explanation is the following. We want to minimize some function on this domain. In the interior we will be able to do it, but then we should switch to the boundary, that is “curved”.

It would be more convenient to introduce different variables, for example,  $x_+ e^{-y_+}$  and  $x_- e^{-y_-}$ . Their domain is  $[1, Q] \times [1, Q]$ , which looks better. However, these variables are still not good enough. Now we introduce the “best” variables.

**4.5.5.2. New variables.** We define

$$\alpha = y - \log \frac{1}{Q_0} = y + \log(Q_0), \quad \alpha_+ = y_+ - \log \frac{x_+}{Q_0}, \quad \alpha_- = y_- - \log \frac{2 - x_+}{Q_0}.$$

In fact,  $\alpha$  and  $\alpha_{\pm}$  are the vertical distances from the points  $z$  and  $z_{\pm}$  to  $\Gamma_{Q_0}$ . For a fixed  $\alpha$  we have *three* variables:  $x_+, \alpha_+$  and  $\alpha_-$ . They are related by equation

$$(4.9) \quad 2\alpha = \alpha_+ + \alpha_- + \log(x_+) + \log(2 - x_+).$$

So  $\alpha_{\pm}$  and  $x_+$  are on some manifold, and to minimize a function of these three variables we should use Lagrange multipliers.

**4.5.5.3. New domain.** Fix  $\alpha \in [\log(Q_0/Q), \log(Q_0)]$ . We have following inequalities for  $\alpha_{\pm}$  and  $x_+$ :

$$\alpha_{\pm} \in [\log(Q_0/Q), \log(Q_0)], \quad x_+ \in [1, 2), \quad \text{and} \quad \alpha_+ + \alpha_- \geq 2\alpha.$$

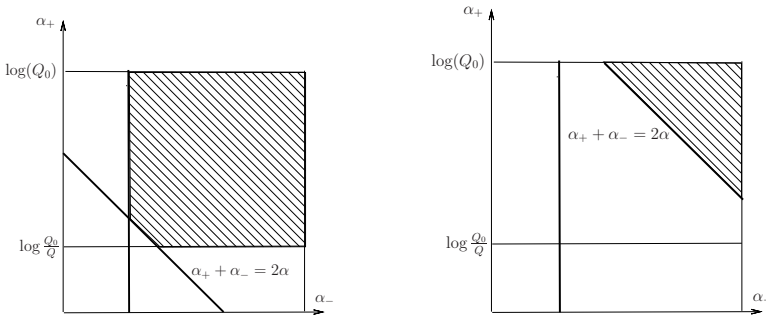
The last inequality follows from the fact that  $\log(x_+) + \log(2 - x_+) \leq 0$ .

We also notice that in fact  $x_+$  can not assume all values in  $[1, 2)$ . We ignore this fact, because from the (4.9),  $x_+$  can be calculated in terms of  $\alpha_{\pm}$ . During the proof the reader will see what we mean.

So for any fixed  $\alpha$  we pay attention only to the domain of  $\alpha_{\pm}$ . We investigate how the domain looks. We notice that, since  $\alpha \geq \log(Q_0/Q)$ , the line  $\alpha_+ + \alpha_- = 2\alpha$  intersects the square  $[\log(Q_0/Q), \log(Q_0)] \times [\log(Q_0/Q), \log(Q_0)]$  (in the  $(\alpha_-, \alpha_+)$  plane).

We notice that domain will look different when  $\alpha \geq \log(Q_0) - \frac{1}{2} \log(Q)$  and when  $\alpha$  is smaller than this number. The reason is that the vertex  $\alpha_- = \log(Q_0)$ ,  $\alpha_+ = \log(Q_0/Q)$  may lie below the line  $\alpha_+ + \alpha_- = 2\alpha$ .

Therefore, the domain of  $\alpha_-$  and  $\alpha_+$  looks as follows.



We are going to study these two cases together. We shall prove that if the global minimum of  $2B_0(z) - B_0(z_+) - B_0(z_-)$  is strictly negative then it is obtained neither in the interior, nor in the interiors of the edges. Then we will investigate the vertices. As the reader can see, edges and vertices where  $\alpha_+ + \alpha_- = 2\alpha$  correspond to vertical segments  $[z_-, z_+]$  and therefore are trivial.

Thus, the second case yields one interesting case  $\alpha_+ = \alpha_- = \log(Q_0)$ , and the first case will give the same vertex and  $\alpha_+ = \log(Q_0/Q)$ ,  $\alpha_- = \log(Q_0)$ .

After this summary, we give all details of searching for possible global minima.



**4.5.5.4. Old variables and new variables.** We now need to recalculate the old variables in terms of the new ones. In particular, we need to relate  $v$  and  $v_{\pm}$  with  $\alpha$  and  $\alpha_{\pm}$  respectively. We will show in a moment that this is possible. The reason is that  $\alpha$  is closely related to the number  $a$ , the first coordinate of a point, where the tangent line to  $\Gamma_{Q_0}$ ,  $\ell(z)$ , “kisses”  $\Gamma_{Q_0}$ .

Take any point  $z = (x, y)$  in  $\Omega_Q$ . We for some time forget that  $x = 1$ , and do calculations for arbitrary  $x$ . We do this because then the same calculations will serve for  $z_{\pm}$ .

We say one more time that now  $v$  and  $a$  correspond to  $Q_0$ , so we should write  $v_0$  and  $a_0$ , but to keep the notation simple we do not do this.

We write the equation of the line  $\ell(z)$ , tangent to  $\Gamma_{Q_0}$  as

$$y = \frac{\gamma_0 x}{v} + \log(v) - \gamma,$$

so

$$\alpha = y - \log \frac{x}{Q_0} = \frac{\gamma_0 x}{v} + \log(v) - \log(x) + \log(Q_0) - \gamma_0.$$

Using the definition of  $\gamma_0$ , we obtain

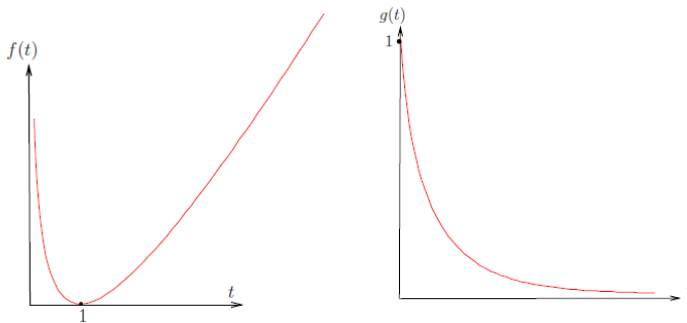
$$\alpha = \frac{\gamma_0 x}{v} + \log(v) - \log(x) - 1 - \log(\gamma_0) = \frac{\gamma_0 x}{v} - \log \frac{\gamma_0 x}{v} - 1.$$

We now introduce the function

$$f(t) = t - \log(t) - 1, \quad t > 0.$$

This function has already appeared in the definition of  $\gamma_0$ . The function  $f$  is decreasing from  $+\infty$  to 0 when  $t \in (0, 1]$  and therefore has an inverse

$$g(t) = f^{-1}(t), \quad g: [0, \infty) \rightarrow (0, 1].$$



We now have an equation

$$\alpha = f\left(\frac{\gamma_0}{x} v\right).$$

We notice that  $x \leq a$ , so  $\gamma_0 x \leq \gamma_0 a = v$ , and so  $g(f(\gamma_0 x/v)) = \gamma_0 x/v$ . Therefore, we write

$$\frac{\gamma_0 x}{v} = g(\alpha),$$

or

$$v = \frac{\gamma_0 x}{g(\alpha)}.$$

In particular we notice that  $g(\alpha) = x/a$ . Basically this is the geometric meaning of  $\alpha$ .

The above equation with the particular points  $z = (1, y)$  and  $z_{\pm}$  gives

$$v = \frac{\gamma_0}{g(\alpha)}, \quad v_+ = \frac{\gamma_0 x_+}{g(\alpha_+)}, \quad \text{and} \quad v_- = \frac{\gamma_0 x_-}{g(\alpha_-)}.$$

We are now ready to introduce the function that we want to minimize.

**4.5.5.5. The function  $\Delta$  in the new variables.** We remind the reader that we fix  $\alpha$  and have three variables  $x_+, \alpha_+,$  and  $\alpha_-$  on the manifold

$$2\alpha = \alpha_+ + \alpha_- + \log(x_+) + \log(2 - x_+).$$

We also remind the reader that  $x_- = 2 - x_+$  and  $x = 1$ . Therefore, our function  $\Delta$  will be

$$\begin{aligned} \Delta(x_+, \alpha_+, \alpha_-) &= 2B_0(z) - B_0(z_+) - B_0(z_-) \\ &= 2\left(\log(v) + \frac{1-v}{\gamma_0}\right) - \left(x_+ \log(v_+) + \frac{x_+ - v_+}{\gamma_0}\right) \\ &\quad - \left((2 - x_+) \log(v_-) + \frac{2 - x_+ - v_-}{\gamma_0}\right). \end{aligned}$$

We now want to rewrite the last expression in terms of  $\alpha_{\pm}$  and  $x_+$ . We get

$$\begin{aligned} \Delta &= (2x \log v - x_+ \log v_+ - x_- \log v_-) - \frac{1}{\gamma_0} (2v - v_+ - v_-) \\ &= 2 \log \frac{1}{g(\alpha)} - x_+ \log \frac{x_+}{g(\alpha_+)} - (2 - x_+) \log \frac{2 - x_+}{g(\alpha_-)} - \frac{2}{g(\alpha)} \\ &\quad + \frac{x_+}{g(\alpha_+)} + \frac{2 - x_+}{g(\alpha_-)}. \end{aligned}$$

Due to the importance of this function, we write the final result separately:

$$\begin{aligned} \Delta(x_+, \alpha_+, \alpha_-) &= 2 \log \frac{1}{g(\alpha)} - x_+ \log \frac{x_+}{g(\alpha_+)} - (2 - x_+) \log \frac{2 - x_+}{g(\alpha_-)} \\ &\quad - \frac{2}{g(\alpha)} + \frac{x_+}{g(\alpha_+)} + \frac{2 - x_+}{g(\alpha_-)}. \end{aligned}$$

Now we will to minimize it.

**Theorem 4.19.** 1) For a fixed  $\alpha \leq \log(Q_0) - \frac{1}{2} \log(Q)$  there holds

$$(4.10) \quad \min \Delta(x_+, \alpha_+, \alpha_-) = \min \left[ 0, \Delta(\hat{x}_+, \log(Q_0), \log(Q_0)), \Delta\left(\tilde{x}_+, \log \frac{Q_0}{Q}, \log(Q_0)\right) \right],$$

where  $\hat{x}_+$  is a solution of  $2\alpha = 2 \log(Q_0) + \log(x_+) + \log(2 - x_+)$ ,  $x_+ \geq 1$ , and  $\tilde{x}_+$  is a solution of  $2\alpha = \log(Q_0) + \log(Q_0/Q) + \log(x_+) + \log(2 - x_+)$ ,  $x_+ \geq 1$ .

2) For a fixed  $\alpha > \log(Q_0) - \frac{1}{2} \log(Q)$  the following holds:

$$\min \Delta(x_+, \alpha_+, \alpha_-) = \min [0, \Delta(\hat{x}_+, \log(Q_0), \log(Q_0))].$$

**Remark 4.20.** We notice that a nonzero minimum may be attained only on the vertices.

**4.5.5.6. Derivatives of  $\Delta$ .** Before we form the Lagrangian, we calculate the derivatives of  $\Delta$  with respect to  $\alpha_+$ ,  $\alpha_-$  and  $x_+$ . First of all,

$$g'(t) = \frac{1}{f'(g(t))} = \frac{g(t)}{g(t) - 1}.$$

So,

$$\frac{\partial \Delta}{\partial \alpha_+} = \frac{x_+}{g(\alpha_+)} \frac{g(\alpha_+)}{g(\alpha_+) - 1} - \frac{x_+}{g(\alpha_+)^2} \frac{g(\alpha_+)}{g(\alpha_+) - 1} = \frac{x_+}{g(\alpha_+)}.$$

Similarly,

$$\frac{\partial \Delta}{\partial \alpha_-} = \frac{2 - x_+}{g(\alpha_-)}.$$

Finally, we take the derivative with respect to  $x_+$ .

$$\begin{aligned} \frac{\partial \Delta}{\partial x_+} &= -\log \frac{x_+}{g(\alpha_+)} - 1 + \log \frac{2 - x_+}{g(\alpha_-)} + 1 + \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)} \\ &= -\log \frac{x_+}{g(\alpha_+)} + \log \frac{2 - x_+}{g(\alpha_-)} + \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)}. \end{aligned}$$

**4.5.5.7. Step 1: interior of the domain.** Suppose we are in the interior of domain of  $\alpha_+$  and  $\alpha_-$ . We define a Lagrangian:

$$L(x_+, \alpha_+, \alpha_-, \lambda) = \Delta(x_+, \alpha_+, \alpha_-) - \lambda \cdot (\alpha_+ + \alpha_- + \log(x_+) + \log(2 - x_+) - 2\alpha).$$

Differentiating  $L$  with respect to  $\alpha_{\pm}$ , we obtain

$$\frac{x_+}{g(\alpha_+)} = \frac{2 - x_+}{g(\alpha_-)} = \lambda.$$

These equalities mean that

$$g(\alpha_+) = \frac{x_+}{\lambda}, \quad g(\alpha_-) = \frac{2 - x_+}{\lambda}.$$

Applying  $f$  to both sides, and recalling that  $f(g(t)) = t$ , we get

$$\begin{aligned} \alpha_+ &= f\left(\frac{x_+}{\lambda}\right) = \frac{x_+}{\lambda} - \log(x_+) + \log(\lambda) - 1, \\ \alpha_- &= f\left(\frac{2 - x_+}{\lambda}\right) = \frac{2 - x_+}{\lambda} - \log(2 - x_+) + \log(\lambda) - 1. \end{aligned}$$

We substitute these equalities into

$$\alpha_+ + \alpha_- - 2\alpha + \log(x_+) + \log(2 - x_+) = 0.$$

By direct calculation,

$$\alpha = \frac{1}{\lambda} + \log(\lambda) - 1 = f\left(\frac{1}{\lambda}\right).$$

We notice that  $1/\lambda = g(\alpha_+)/x_+ \leq g(\alpha_+) \leq 1$ , and so  $g(f(1/\lambda)) = 1/\lambda$ .

Notice that this would not be true if  $\lambda$  was less than 1.

So,  $g(\alpha) = 1/\lambda$  (in fact, from this equation we find  $\lambda$ ). Now we can calculate  $\Delta$  at our point:

$$\Delta = 2 \log(\lambda) - x_+ \log(\lambda) - (2 - x_+) \log(\lambda) - 2\lambda + \lambda + \lambda = 0.$$

**4.5.5.8. Conclusion.** From the calculation above we conclude the following: either the global minimum of  $\Delta$  is zero, or the global minimum is attained on the boundary.

**4.5.5.9. Step 2: reduction to the case  $\alpha_- \geq \alpha_+$ .** We now prove a technical but very useful lemma. It will show that it is sufficient to minimize  $\Delta$  only on half of our domain, when  $\alpha_- \geq \alpha_+$ . This will show that we do not need to consider the edges  $\alpha_+ = \log(Q_0)$  and  $\alpha_- = \log(Q_0/Q)$ , except for vertices.

**Lemma 4.21.** Fix  $x_+$  and let

$$\begin{aligned} \Delta(\alpha_+, \alpha_-) = \Delta(x_+, \alpha_+, \alpha_-) &= 2 \log \frac{1}{g(\alpha)} - x_+ \log \frac{x_+}{g(\alpha_+)} - (2 - x_+) \log \frac{2 - x_+}{g(\alpha_-)} \\ &\quad - \frac{2}{g(\alpha)} + \frac{x_+}{g(\alpha_+)} + \frac{2 - x_+}{g(\alpha_-)}. \end{aligned}$$

If  $u > v$  then  $\Delta(u, v) \geq \Delta(v, u)$ .

**Remark 4.22.** Hence, if  $\alpha_+ > \alpha_-$  then  $\Delta(\alpha_+, \alpha_-) \geq \Delta(\alpha_-, \alpha_+)$ , and so if the global minimum is attained on the boundary, it is necessarily attained on the part where  $\alpha_+ \leq \alpha_-$ .

**Remark 4.23.** Notice that this lemma is natural. As we have seen from the investigation of the cases where  $z$  and  $z_{\pm}$  are on the boundary, the worst case happens when  $z_- \in \Gamma$  and  $z_+ \in \Gamma_Q$ . This corresponds to  $\alpha_- = \log(Q_0)$  and  $\alpha_+ = \log(Q_0/Q)$ , which is smaller than  $\alpha_-$ .

*Proof of Lemma 4.21.* First, since  $u > v$  we have  $g(u) < g(v) \leq 1$ . We define  $t = g(u)$  and  $s = g(v)$ , so  $t < s \leq 1$ . We have

$$\begin{aligned} \Delta(u, v) - \Delta(v, u) &= x_+ \log(t) + (2 - x_+) \log(s) + \frac{x_+}{t} + \frac{2 - x_+}{s} \\ &\quad - \left( x_+ \log(s) + (2 - x_+) \log(t) + \frac{x_+}{s} + \frac{2 - x_+}{t} \right) \\ &= (2x_+ - 2) \log(t) + \frac{2x_+ - 2}{t} + (2 - 2x_+) \log(s) + \frac{2 - 2x_+}{s} \\ &= (2x_+ - 2) \left( \frac{1}{t} + \log(t) - \frac{1}{s} - \log(s) \right). \end{aligned}$$

Define  $\varphi(x) = 1/x + \log(x)$ . Then  $\varphi'(x) = 1/x - 1/x^2 = (x - 1)/x^2 < 0$  when  $x \leq 1$ . Since  $t < s \leq 1$ , we get

$$\Delta(u, v) - \Delta(v, u) \geq 0. \quad \square$$

**4.5.5.10. Step 3:  $\alpha_+ + \alpha_- = 2\alpha$ .** In this case  $x_+ = 1$ , and so  $2 - x_+ = 1$ , and we have a vertical line segment  $[z_-, z_+]$ . It lies entirely in  $\Omega_Q$ , where the function  $B_0$  is locally concave. Therefore,  $\Delta \geq 0$ .

**4.5.5.11. Step 4:  $\alpha_- = \log(Q_0)$ .** In this case our manifold is

$$2\alpha = \alpha_+ + \log(Q_0) + \log(x_+) + \log(2 - x_+).$$

Keeping in mind that  $\alpha_-$  is fixed and we cannot differentiate with respect to it, we write the same Lagrangian as before, and take derivatives with respect to  $\alpha_+$  and  $x_+$ . We have

$$L(x_+, \alpha_+, \alpha_-, \lambda) = \Delta(x_+, \alpha_+, \alpha_-) - \lambda \cdot (\alpha_+ + \alpha_- + \log(x_+) + \log(2 - x_+) - 2\alpha),$$

and so

$$\frac{x_+}{g(\alpha_+)} = \lambda.$$

In particular, we again get that  $\lambda \geq 1$ . We now differentiate with respect to  $x_+$ , getting

$$-\log \frac{x_+}{g(\alpha_+)} + \log \frac{2 - x_+}{g(\alpha_-)} + \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)} - \lambda \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right) = 0.$$

Using the equality  $x_+/g(\alpha_+) = \lambda$ , we obtain

$$-\log(\lambda) + \log \frac{2 - x_+}{g(\alpha_-)} - \frac{1}{g(\alpha_-)} + \frac{\lambda}{2 - x_+} = 0,$$

and thus

$$f\left(\frac{\lambda}{2 - x_+}\right) = f\left(\frac{1}{g(\alpha_-)}\right).$$

We notice that  $2 - x_+ \leq 1$  and  $\lambda \geq 1$ , so  $\lambda/(2 - x_+) \geq 1$ . Since  $f(t)$  increases when  $t \geq 1$ , we obtain

$$\frac{\lambda}{2 - x_+} = \frac{1}{g(\alpha_-)}.$$

The same equation we had when we were investigating the interior. As we have seen, this equation implies  $\Delta = 0$ .

**4.5.5.12. Step 5:  $\alpha_+ = \log(Q_0/Q)$ .** This edge is more delicate. Here we differentiate with respect to  $\alpha_-$  and  $x_+$ :

$$\frac{2 - x_+}{g(\alpha_-)} = \lambda, \quad -\log \frac{x_+}{g(\alpha_+)} + \log \frac{2 - x_+}{g(\alpha_-)} + \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)} - \lambda \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right) = 0.$$

Substituting the first equality into the second, we get

$$-\log \frac{x_+}{g(\alpha_+)} + \log(\lambda) + \frac{1}{g(\alpha_+)} - \frac{\lambda}{x_+} = 0.$$

Similar to the previous step, we get  $f(\lambda/x_+) = f(1/g(\alpha_+))$ . However now we cannot say that  $\lambda/x_+ \geq 1$ , and so we cannot conclude that  $\lambda/x_+ = 1/g(\alpha_+)$ . We show how to finish the proof without this conclusion. We also warn the reader that this proof would not work in the previous step because it is tied to the fact that  $x_+$  is on the  $Q$ -boundary of  $\Omega_Q$ .

We now proceed as follows:

$$\frac{x_+}{g(\alpha_+)} - x_+ \log \frac{x_+}{g(\alpha_+)} = \lambda - x_+ \log(\lambda).$$

Substituting this in  $\Delta$ , we get

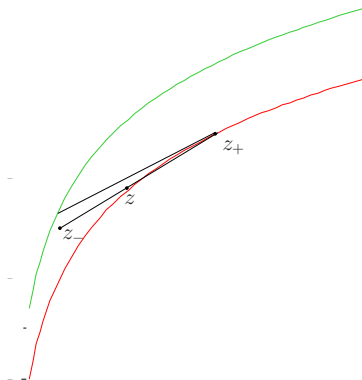
$$\begin{aligned} \Delta &= 2 \log \frac{1}{g(\alpha)} + \lambda - x_+ \log(\lambda) - 2 \log(\lambda) - \frac{2}{g(\alpha)} + \lambda = 2 \left( f(\lambda) - f\left(\frac{1}{g(\alpha)}\right) \right) \\ &= 2 \left( f\left(\frac{2-x_+}{g(\alpha_-)}\right) - f\left(\frac{1}{g(\alpha)}\right) \right). \end{aligned}$$

Notice that if  $(2-x_+)/g(\alpha_-) \geq 1/g(\alpha) \geq 1$  then, due to the monotonicity of  $f(t)$ , we get that  $\Delta \geq 0$ . Therefore, we should prove that  $\Delta$  is nonnegative when  $(2-x_+)/g(\alpha_-) < 1/g(\alpha)$ . The following lemma proves this fact.

**Lemma 4.24.** *If  $(2-x_+)/g(\alpha_-) < 1/g(\alpha)$  then the line segment  $[z_-, z_+]$  lies entirely in  $\Omega_{Q_0}$  and, consequently,  $2B_0(z) - B_0(z_+) - B_0(z_-) \geq 0$ .*

Before proving this lemma we need an observation related to the geometry of  $\Omega_Q$ .

Take the point  $z_+$ , which in our case lies on  $\Gamma_Q$ , and take the tangent to  $\Gamma_Q$  at this point. Since we assume that  $x_+ > x_-$  and  $y_+ > y_-$ , we get the following: if the segment  $[z_-, z_+]$  passes above this tangent line, then it lies entirely in  $\Omega_Q$ , and the fact stated in the lemma is true. Hence the only interesting case is when  $[z_-, z_+]$  passes below this tangent. This means that it leaves  $\Omega_Q$  near  $z_+$ , and then returns before it ‘‘hits’’ the point  $z$ . Therefore, the segment  $[z_-, z]$  lies in  $\Omega_Q$ , so the only problem can occur between  $z$  and  $z_+$ .



Lemma 4.24 is a consequence of the following one.

**Lemma 4.25.** *Suppose  $p \geq a \geq 1$ ,  $\alpha = 1/a - \log(1/a) - 1$ , and  $\alpha_+ = p/a - \log(p/a) - 1$ . If the line segment  $[z_-, z_+]$  does not lie entirely in  $\Omega_{Q_0}$  then  $x_+ \geq p$ .*

*Proof.* Such  $a$  and  $p$  exist, because for every  $u > 0$  the equation  $t - \log(t) - 1 = u$  has two solutions, one of which is less than 1, and the other of which is larger than 1.

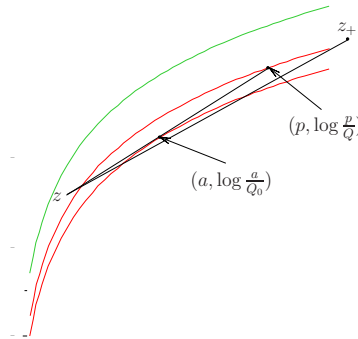
We take our point  $(1, y)$  and draw the tangent to  $\Gamma_{Q_0}$  that goes to the right. Since the only possibility for  $[z_-, z_+]$  to be outside of  $\Gamma_Q$  is that part of  $[z, z_+]$  is outside, we do not care about  $z_-$ . If  $[z, z_+]$  goes above this tangent, then it is in  $\Omega_{Q_0}$ , and so the only “bad” case is when  $[z, z_+]$  goes below the tangent. Suppose that the tangent “kisses”  $\Gamma_{Q_0}$  at point  $(a, \log(a/Q_0))$ . Then the equation (in the  $(x_1, x_2)$  plane) is

$$x_2 - y = \frac{1}{a} (x_1 - 1).$$

Since  $a$  satisfies the equation, and since  $\alpha = y + \log(Q_0)$ , we get

$$\frac{1}{a} - \log \frac{1}{a} - 1 = \alpha.$$

Now take the point  $(p, \log(p/Q))$  – the point where the tangent intersects  $\Gamma_Q$  for the second time. This is the first time when the segment  $[z, z_+]$  can return to  $\Omega_Q$  (if it ever left). Since  $z_+$  is on the right-hand side of the reentry point, we have  $x_+ > p$ . Let us find  $p$ .



We have

$$\log \frac{p}{Q} - y = \frac{p}{a} - \frac{1}{a},$$

and so, since  $\alpha_+ = \log(Q_0/Q)$ , we get

$$\alpha_+ = \frac{p}{a} - \log \frac{p}{a} - 1.$$

Hence both  $a$  and  $p$  are as in the statement of the lemma, which finishes the proof. □

Now we prove the Lemma 4.24.

*Proof.* Suppose that  $[z_-, z_+]$  does not lie in  $\Omega_{Q_0}$ . Then  $x_+ \geq p$ , which implies  $x_+/a \geq p/a > 1$ , so  $f(x_+/a) \geq \alpha_+$ . Next, we have

$$2\alpha = \alpha_+ + \alpha_- + \log(x_+) + \log(2 - x_+) \leq \frac{x_+}{a} - \log \frac{x_+}{a} - 1 + \alpha_- + \log(x_+) + \log(2 - x_+).$$

Recall that  $1/a - \log(1/a) - 1 = \alpha$ , so

$$2\alpha \leq \frac{x_+ - 1}{a} + \frac{1}{a} - \log \frac{1}{a} - 1 + \alpha_- + \log(2 - x_+),$$

thus

$$\alpha \leq \frac{x_+ - 1}{a} + \alpha_- + \log(2 - x_+).$$

Using the equation for  $a$  and  $\alpha$  again, we get

$$\frac{2 - x_+}{a} - \log \frac{2 - x_+}{a} - 1 \leq \alpha_-,$$

so

$$f\left(\frac{2 - x_+}{a}\right) \leq \alpha_-.$$

We apply  $g$  to both sides. We see that  $2 - x_+ = x_-$ , while  $a > 1$ , so  $g(f((2 - x_+)/a)) = (2 - x_+)/a$ , therefore

$$\frac{2 - x_+}{a} \geq g(\alpha_-),$$

and so

$$\frac{2 - x_+}{g(\alpha_-)} \geq a.$$

We know that  $\alpha = f(1/a)$  and  $a > 1$ , so  $g(\alpha) = 1/a$  implies

$$\frac{2 - x_+}{g(\alpha_-)} \geq \frac{1}{g(\alpha)}.$$

This contradicts the assumption of our lemma. □

We claim that Theorem 4.19 is proved. Indeed, the global minimum is either 0 or attained on the boundary. On the boundary it is either again 0 or is attained on vertices. However, the vertices where  $\alpha_+ + \alpha_- = 2\alpha$  give a nonnegative result, so the minimum can be attained only at the vertices indicated in the statement of the theorem.

**4.5.5.13. Step 6: Vertex  $\alpha_+ = \alpha_- = \log(Q_0)$ .** In this case,

$$\alpha = \log(Q_0) + \frac{1}{2} \log(x_+(2 - x_+)).$$



We obtain bounds for  $x_+$ . Clearly,  $x_+ \geq 1$ , and this bound is accessible when  $x_+ = x_- = 1$ . Since  $\alpha \geq \log(Q_0/Q)$ , we get  $x_+(2 - x_+) \geq 1/Q^2$ , which means that  $x_+ \leq 1 + \sqrt{1 - 1/Q^2}$ . As we know from Section 4.5.4.3, this is also accessible when  $x \in \Gamma_Q$ . Hence,  $x_+ \in [1, 1 + r]$ , where  $r = \sqrt{1 - 1/Q^2}$ .

We now treat  $\alpha_+$  as a function of  $x_+$  and, therefore, our  $\Delta$  becomes a function of  $x_+$ . We have

$$\Delta(x_+) = 2 \log \frac{1}{g(\alpha)} - \frac{2}{g(\alpha)} - x_+ \log(x_+) - (2 - x_+) \log(2 - x_+) - 2 \log \frac{1}{g(\alpha_+)} + \frac{2}{g(\alpha_+)}.$$

Again from Section 4.5.4.3 we have that  $\Delta(1 + r) \geq 0$ . We intend to prove that  $\Delta' \leq 0$ . Then we will be done with this case. We first notice that

$$\frac{\partial \alpha}{\partial x_+} = \frac{1}{2} \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right).$$

Therefore,

$$\begin{aligned} \Delta'(x_+) &= \frac{2}{g(\alpha)} \cdot \frac{1}{2} \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right) - \log(x_+) + \log(2 - x_+) \\ &= \frac{1}{g(\alpha)x_+} - \log(x_+) - \frac{1}{g(\alpha)(2 - x_+)} + \log(2 - x_+) \\ &= \frac{1}{g(\alpha)x_+} + \log \frac{1}{g(\alpha)x_+} - \left( \frac{1}{g(\alpha)(2 - x_+)} + \log \frac{1}{g(\alpha)(2 - x_+)} \right). \end{aligned}$$

The last equality is obtained by adding and subtracting  $\log(1/g(\alpha))$ . We notice that the function  $s \mapsto 1/s + \log(1/s)$  is decreasing, and  $g(\alpha)x_+ \geq g(\alpha)(2 - x_+)$ . Therefore,  $\Delta'(x_+) \leq 0$ , which finishes the proof in this case.

**4.5.5.14. The vertex  $\alpha_+ = \log(Q_0/Q)$ ,  $\alpha_- = \log(Q_0)$ .** Now we set  $\alpha_+ = \log(Q_0/Q)$  and  $\alpha_- = \log(Q_0)$ , so

$$\alpha = \log(Q_0) - \frac{1}{2} \log(Q) + \frac{1}{2} \log(x_+(2 - x_+)).$$

The bounds for  $x_+$  in this case are  $1 \leq x_+ \leq 1 + r$ , where  $r = \sqrt{1 - 1/Q}$ . We know from Section 4.5.4.2 that they are accessible, and that  $\Delta(1 + r) = 0$ , as this how  $Q_0$  was chosen, and this is the first and the only time when we use this. Again we would like to prove that  $\Delta$  is decreasing. The difficulty is that now  $\alpha_+ \neq \alpha_-$ , and so  $\Delta$  does not have nice cancelations. We have

$$\begin{aligned} \Delta(x_+) &= 2 \log \frac{1}{g(\alpha)} - \frac{2}{g(\alpha)} - x_+ \log(x_+) - (2 - x_+) \log(2 - x_+) \\ &\quad + x_+ \log(g(\alpha_+)) + (2 - x_+) \log(g(\alpha_-)) + \frac{x_+}{g(\alpha_+)} + \frac{2 - x_+}{g(\alpha_-)}, \end{aligned}$$

and so

$$\begin{aligned} \Delta'(x_+) &= \frac{1}{g(\alpha)} \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right) - \log(x_+) + \log(2 - x_+) \\ &\quad + \frac{1}{g(\alpha_+)} - \log \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)} + \log \frac{1}{g(\alpha_-)}. \end{aligned}$$

From the investigation of the previous vertex we know that

$$\frac{1}{g(\alpha)} \left( \frac{1}{x_+} - \frac{1}{2 - x_+} \right) - \log(x_+) + \log(2 - x_+) \leq 0.$$

This did not depend on the choice of  $\alpha_{\pm}$ . Finally,  $\alpha_+ < \alpha_-$ , so  $g(\alpha_+) > g(\alpha_-)$ , thus  $1/g(\alpha_-) > 1/g(\alpha_+) \geq 1$ , and

$$f\left(\frac{1}{g(\alpha_-)}\right) > f\left(\frac{1}{g(\alpha_+)}\right).$$

This means exactly that

$$\frac{1}{g(\alpha_+)} - \log \frac{1}{g(\alpha_+)} - \frac{1}{g(\alpha_-)} + \log \frac{1}{g(\alpha_-)} < 0.$$

Thus the proof is finished.

### 5. Proofs of the Proposition 4.1 and Lemma 4.17

In this section we give detailed (and rather technical) proofs of Proposition 4.1 and Lemma 4.17.

*Proof of the Proposition 4.1.* (1) Suppose  $f$  is increasing. Then

$$\int_{-1}^1 (1 - |t|) f(x + \varepsilon t) dt \geq \int_0^1 (1 - |t|) f(x + \varepsilon t) dt \geq f(x) \int_0^1 (1 - |t|) dt.$$

If  $f$  is decreasing then we consider the integral over  $(-1, 0)$ , which finishes the proof of the first part.

(2) Let  $x(s) = x + st$ , and  $a(s) = A(x(s))$ . We want to estimate the quantity

$$\begin{aligned} a(0) - \frac{a(1) + a(-1)}{2} &= -\frac{1}{2} \int_{-1}^1 (1 - |s|) a''(s) ds \\ &= -\frac{1}{2} \cdot (t)^2 \int_{-1}^1 (1 - |s|) A''(x(s)) ds = -\frac{1}{2} \cdot t^2 \int_{-1}^1 (1 - |s|) A''(x + st) ds. \end{aligned}$$

Thus,

$$A(x) - \frac{A(x - t) + A(x + t)}{2} \leq -c \cdot t^2 A''(x),$$

which is exactly what we want.

(3) Due to the homogeneity, this inequality is equivalent to

$$f(u) := u^p - \frac{(u + 1)^p + (u - 1)^p}{2} + \beta u^{p-2} \geq 0, \quad u > 1$$

We notice that  $f$  is continuous, and  $\lim_{u \rightarrow \infty} f(u)/u^{p-2}$  is finite. Therefore, such a  $\beta$  exists.

(4) Again using homogeneity, we reduce our problem to the following: the function

$$f_0(u) = u^{-1/(p-1)} - \frac{1}{2}(u-1)^{-1/(p-1)} - (u+1)^{-1/(p-1)} + \gamma u^{-2-1/(p-1)}$$

should be nonnegative when  $u \geq C/(C-1)$ . Here  $\gamma = \alpha p'/(p-1)$ .

We multiply by  $u^{2+1/(p-1)}$  and, letting  $v = u^{-1}$ , we need

$$f_1(v) = \frac{1}{v^2} - \frac{(1-v)^{-1/(p-1)} + (1+v)^{-1/(p-1)}}{2v^2} + \gamma \geq 0,$$

or the function

$$f(v) = \frac{1}{v^2} - \frac{(1-v)^{-1/(p-1)} + (1+v)^{-1/(p-1)}}{2v^2}$$

should be bounded from below whenever  $0 < v < (C-1)/C$ .

We prove the following:

**Lemma 5.1** (Sublemma).  *$f(v)$  is decreasing.*

If we prove the sublemma, we get

$$f(v) \geq f\left(\frac{C-1}{C}\right),$$

and, therefore,

$$\gamma = -f\left(\frac{C-1}{C}\right). \quad \square$$

*Proof of the sublemma.* We prove this proposition by straightforward differentiation. First,

$$v^2 f(v) = 1 - \frac{(1-v)^{-1/(p-1)} + (1+v)^{-1/(p-1)}}{2},$$

and so

$$2v f(v) + v^2 f'(v) = \frac{1}{p-1} \frac{(1+v)^{-1-1/(p-1)} - (1-v)^{-1-1/(p-1)}}{2}.$$

Thus

$$v^2 f'(v) = \frac{1}{p-1} \frac{(1+v)^{-1-1/(p-1)} - (1-v)^{-1-1/(p-1)}}{2} - \frac{2}{v} + \frac{(1-v)^{-1/(p-1)} + (1+v)^{-1/(p-1)}}{v}.$$

We want to prove that  $f'(v) < 0$  or, equivalently, that the right-hand side is negative. We multiply by  $v$  to get (after simple algebra)

$$\begin{aligned} v^3 f'(v) &= (1+v)^{1-p'} \frac{p'+1}{2} + (1-v)^{1-p'} \frac{p'+1}{2} - ((1+v)^{-p'} + (1-v)^{-p'}) \frac{1}{2(p-1)} - 2 \\ &=: \psi(v). \end{aligned}$$

Clearly,  $\psi(0) = 0$ . Next,

$$\begin{aligned} \psi'(v) &= \frac{(1-p')(p'+1)}{2} (1+v)^{-p'} - \frac{(1-p')(p'+1)}{2} (1-v)^{-p'} \\ &\quad + \frac{p'}{2(p-1)} \left( (1+v)^{-1-p'} - (1-v)^{-1-p'} \right), \\ \psi''(v) &= \frac{p'(p'+1)}{2(p-1)} \cdot v \cdot \left( (1+v)^{-2-p'} - (1-v)^{-2-p'} \right). \end{aligned}$$

Thus,  $\psi''(v) \leq 0$ , so  $\psi'(v) \leq \psi'(0) = 0$ , and therefore  $\psi(v) \leq \psi(0) = 0$ , which is what we want.  $\square$

*Proof of Lemma 4.17.* The segment  $[z_-, z_+]$  has the parametrization

$$u(t) = t x_+ (1-t) x_-, \quad v(t) = t y_+ + (1-t) y_-.$$

Then

$$\varphi(t) = u(t) \exp(-v(t)) = (t(x_+ - x_-) + x_-) \exp(-t(y_+ - y_-) - y_-).$$

We would like to prove that there exists a constant  $C$  that does not depend on  $Q$ , and such that  $\varphi(t) \leq CQ$ ,  $t \in [0, 1]$ . First we have  $\varphi(0) = 1$  and  $\varphi(1) = Q$ , so we need to check the local extrema.

$$\varphi'(t) = (x_+ - x_-) \exp(\dots) - (y_+ - y_-) (t(x_+ - x_-) + x_-) \exp(\dots).$$

If

$$\varphi'(t_*) = 0$$

then

$$\frac{x_+ - x_-}{y_+ - y_-} = x_- + t_*(x_+ - x_-),$$

so

$$t_* = \frac{1}{y_+ - y_-} - \frac{x_-}{(x_+ - x_-)},$$

or

$$t_*(y_+ - y_-) = 1 - (y_+ - y_-) \frac{x_-}{x_+ - x_-}.$$

Therefore,

$$\varphi(t_*) = \frac{x_+ - x_-}{y_+ - y_-} \exp\left(\left(y_+ - y_-\right) \frac{x_-}{x_+ - x_-} - 1 - y_-\right).$$

We now plug in  $x_{\pm}$  and  $y_{\pm}$ . First

$$x_{\pm} = 1 \pm r,$$

so  $x_+ - x_- = 2r$ . Also  $y_+ - y_- = \log \frac{1+r}{Q} - \log(1-r) = \log \frac{1+r}{Q(1-r)}$ . We notice that

$$Q(1-r) = Q\left(1 - \sqrt{1 - \frac{1}{Q}}\right) = Q - \sqrt{Q^2 - Q} = \frac{Q}{Q + \sqrt{Q^2 - Q}} \asymp 1.$$

Hence  $y_+ - y_- \asymp 1$ . As this proof involves no deep ideas, we finish it briefly. First, we are interested in large  $Q$ , because for bounded  $Q$  we can always find a uniform  $C$ . Hence,

$$r = \sqrt{1 - \frac{1}{Q}} \sim 1 - \frac{1}{2Q} \sim 1,$$

and  $x_- = 1 - r \sim 1/(2Q)$ ,  $y_- \sim \log(1/(2Q))$ . Consequently,

$$\varphi(t_*) \asymp 2 \left(1 - \frac{1}{2Q}\right) \exp \left( C \cdot \frac{1}{2Q} \frac{1}{2} - 1 - \log \frac{1}{2Q} \right) \asymp 2 \exp(\log(2Q)) \asymp Q.$$

This finishes the proof. □

## References

- [1] BEZNOSOVA, O. AND REZNIKOV, A.:  $L \log L$  and reverse Hölder property for  $A_\infty$  weights, and their applications. ArXiv:1107.1885, 2011.
- [2] BUCKLEY, S.: *Harmonic analysis on weighted spaces*. Ph.D. Thesis, University of Chicago, 1990.
- [3] BUCKLEY, S.: Estimates for operator norms on weighted spaces and reverse Jensen inequalities. *Trans. Amer. Math. Soc.* **340** (1993), no. 1, 253–272.
- [4] BUCKLEY, S.: Summation conditions on weights. *Michigan Math. J.* **40** (1993), no. 1, 153–170.
- [5] DINDOŠ, M. AND WALL, T.: The sharp  $A_p$  constant for weights in a reverse-Hölder class. *Rev. Mat. Iberoam* **25** (2009), no. 2, 559–594.
- [6] FEFFERMAN, R., KENIG, C. AND PIPHER, J.: The theory of weights and the Dirichlet problem for elliptic equations. *Ann. of Math. (2)* **134** (1991), no. 1, 65–124.
- [7] GARCÍA-CUERVA, J. AND RUBIO DE FRANCIA, J.: *Weighted norm inequalities and related topics*. North-Holland Mathematics Studies 116, North-Holland, Amsterdam, 1985.
- [8] GRAFAKOS, L.: *Classical and modern Fourier analysis*. Pearson Education, Upper Saddle River, NJ, 2004.
- [9] HYTÖNEN, T. AND LACEY, M.: The  $A_p$ - $A_\infty$  inequality for general Calderón–Zygmund operators. *Indiana Univ. Math. J.* **61** (2012), no. 6, 2041–2092.
- [10] HYTÖNEN, T. AND PÉREZ, C.: Sharp weighted bounds involving  $A_\infty$ . *Anal. PDE* **6** (2013), no. 4, 777–818.
- [11] HYTÖNEN, T., PÉREZ, C., TREIL, S. AND VOLBERG, A.: Sharp weighted estimates for dyadic shifts and the  $A_2$  conjecture. *J. Reine Angew. Math.* **687** (2014), 43–86.
- [12] IWANIEC, T. AND VERDE, A.: On the operator  $\mathcal{L}f = f \log |f|$ . *J. Funct. Anal.* **169** (1999), no. 2, 391–420.
- [13] NAZAROV, F., TREIL, S. AND VOLBERG, A.: Bellman function and two-weight inequality for Haar multipliers. *J. Amer. Math. Soc.* **12** (1999), no. 4, 909–928.
- [14] NAZAROV, F., TREIL, S. AND VOLBERG, A.: The  $Tb$  theorem on non-homogeneous spaces. *Acta Math.* **190** (2003), no. 2, 151–239.
- [15] PETERMICHLE, S. AND POTT, S.: An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.* **354** (2002), no. 4, 1699–1703.

- [16] REZNIKOV, A., VASYUNIN, V. AND VOLBERG, A.: An observation: cut-off of the weight  $w$  does not increase the  $A_{p_1, p_2}$ -“norm” of  $w$ . ArXiv:1008.3635, 2010.
- [17] REZNIKOV, A.: Sharp weak type estimates for weights in the class  $A_{p_1, p_2}$ . *Rev. Mat. Iberoam.* **29** (2013), no. 2, 433–478.
- [18] SLAVIN, L. AND VASYUNIN, V.: Sharp results in the integral-form John–Nirenberg inequality. *Trans. Amer. Math. Soc* **363** (2011), no. 8, 4135–4169.
- [19] STEIN, E. M.: Note on the class  $L \log L$ . *Studia Math.* **31** (1969), 305–310.
- [20] STEIN, E. M.: *Harmonic analysis: real-variable methods, orthogonality and oscillatory integrals*. Princeton Mathematical Series 43, Monographs in Harmonic Analysis III, Princeton University Press, Princeton, NJ, 1993.
- [21] VASYUNIN, V.: Mutual estimates of  $L^p$ -norms and the Bellman function. *J. Math. Sci. (N.Y.)* **156** (2009), no. 5, 766–798.
- [22] VASYUNIN, V.: Lecture Notes. Available at [http://homepages.uc.edu/~slavinld/TRS/cincinnati\\_bellman\\_lectures.pdf](http://homepages.uc.edu/~slavinld/TRS/cincinnati_bellman_lectures.pdf).
- [23] VASYUNIN, V. AND VOLBERG, A.: Monge–Ampere equation and Bellman optimization of Carleson embedding theorems. In *Linear and complex analysis*, 195–238. Amer. Math. Soc. Transl. Ser. 2, 226, Amer. Math. Soc., Providence, RI, 2009.
- [24] VASYUNIN, V. AND VOLBERG, A.: The Bellman function for the simplest two-weight inequality: an investigation of a particular case. *St. Petersburg Math. J* **18** (2007), no. 2, 201–222.
- [25] WILSON, M.: *Weighted Littlewood–Paley theory and exponential-square integrability*. Lecture Notes in Mathematics 1924, Springer, Berlin, 2008.
- [26] WITTWER, J.: A sharp estimate on the norm of the martingale transform. *Math. Res. Lett* **7** (2000), no. 1, 1–12.

Received September 24, 2012.

OLEKSANDRA BEZNOVA: Department of Mathematics, University of Alabama, Box 870350, Tuscaloosa, AL 35487-0350, USA.

E-mail: [ovbeznova@ua.edu](mailto:ovbeznova@ua.edu)

ALEXANDER REZNIKOV: Department of Mathematics, 1326 Stevenson Center, Vanderbilt University, Nashville, TN 36240, USA.

E-mail: [aleksandr.b.reznikov@vanderbilt.edu](mailto:aleksandr.b.reznikov@vanderbilt.edu)