INEQUALITIES FOR HILBERT OPERATOR AND ITS EXTENSIONS: THE PROBABILISTIC APPROACH

ADAM OSĘKOWSKI

Abstract. We present a probabilistic study of the Hilbert operator

$$Tf(x) = \frac{1}{\pi}\int_0^\infty \frac{f(y)\mathrm{d}y}{x+y}, \qquad x \geq 0,$$

defined on integrable functions f on the positive halfline. Using appropriate novel estimates for orthogonal martingales satisfying the differential subordination, we establish sharp moment, weak-type and Φ-inequalities for T. We also show similar estimates for higher dimensional analogues of the Hilbert operator and, by the further careful modification of martingale methods, we obtain related sharp localized inequalities for Hilbert and Riesz transforms.

1. Introduction

As evidenced in numerous papers (see e.g. [1], [2], [3], [6], [7], [15], [23], [26], [27]), martingale theory plays a fundamental role in obtaining various bounds for a wide class of singular integrals, Fourier multipliers and other important operators, in many cases producing optimal or almost-optimal constants. The problem of finding the exact values of various norms of such objects, most notably the Beurling-Ahlfors transform on the complex plane, has gained considerable interest in the recent literature and has been approached with the use of powerful probabilistic techniques developed by Burkholder [8], [10]. One of the motivations for this direction of research comes from the papers of Donaldson and Sullivan [13], and Iwaniec and Martin [21], [22], in which it was pointed out that good estimates for the L_p norm of the Riesz transforms on $\mathbb{R}^n$ and the Beurling-Ahlfors operator on $\mathbb{C}$ have important consequences in the study of quasiconformal mappings, related nonlinear geometric PDEs as well as in the L_p-Hodge decomposition theory.

The purpose of this paper is to illustrate further the fruitful connection between probability theory and the study of classical operators appearing in harmonic analysis. In particular, we will show how an appropriate "fine-tuning" of martingale methods can be used in the study of the so-called Hilbert operator (see below for the formal definition). This will lead us further to some interesting novel bounds for Hilbert and Riesz transforms.

To begin, let us describe the motivation, the results which interested many mathematicians at the beginning of the previous century. A celebrated inequality of Hilbert asserts that if $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ are two sequences of real numbers, then we

---

2000 *Mathematics Subject Classification.* Primary: 60G44. Secondary: 31B05.

Key words and phrases. Hilbert operator, martingale, differential subordination, best constants.

Partially supported by Polish Ministry of Science and Higher Education (MNiSW) grant IP2011 039571 'Iuventus Plus'.

have the sharp bound

$$\left|\sum_{m,n=1}^{\infty}\frac{a_mb_n}{m+n}\right|\leq\pi\left(\sum_{m=1}^{\infty}a_m^2\right)^{1/2}\left(\sum_{n=1}^{\infty}b_n^2\right)^{1/2}.$$

Actually, Hilbert proved this inequality with the constant 2π; the above sharp version is due to Schur [31]. This result was generalized by Hardy and Riesz (cf. [19]): for any $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ as above and any constant $1<p<\infty$,

$$\left|\sum_{m,n=1}^{\infty}\frac{a_mb_n}{m+n}\right|\leq\frac{\pi}{\sin(\pi/p)}\left(\sum_{m=1}^{\infty}|a_m|^p\right)^{1/p}\left(\sum_{n=1}^{\infty}|b_n|^q\right)^{1/q},$$

and constant $\pi/\sin(\pi/p)$ cannot be replaced in general by any smaller number. Here $q=p/(p-1)$ denotes the harmonic conjugate to p. See also the monograph [20] by Hardy, Littlewood and Polya, and the papers [24] by Oleszkiewicz and [33] by Ullrich for more on the subject. Clearly, the above inequalities imply that the operator S, acting on sequences $(a_n)_{n\geq 1}$ by the formula

$$Sa(n)=\frac{1}{\pi}\sum_{m=1}^{\infty}\frac{a_m}{m+n},\quad n=1,\,2,\,\ldots,$$

is bounded in ℓ_p, $1<p<\infty$, and $||S||_{\ell_p\to\ell_p}=\sin^{-1}(\pi/p)$ (note the normalization factor $1/\pi$ used in the definition of S).

We will be interested in the continuous version of S. For any locally integrable function f on $(0,\infty)$, define the Hilbert operator T by

$$Tf(x)=\frac{1}{\pi}\int_0^{\infty}\frac{f(y)\mathrm{d}y}{x+y},\qquad x>0.$$

This operator arises naturally in many settings; for instance, one easily checks that it is equal to the square of Laplace transform. By standard discretization arguments, one easily verifies that the norms of $||S||_{\ell_p\to\ell_p}$ and $||T||_{L_p(0,\infty)\to L_p(0,\infty)}$ coincide. So, for $1<p<\infty$ we have the identity

$$||T||_{L_p(0,\infty)\to L_p(0,\infty)}=\frac{1}{\sin(\pi/p)}.$$

An alternative proof of this fact, using Schur's lemma, can be found for example in Grafakos [16].

Our purpose is to develop a completely different approach to the study of various estimates for T, which rests on the theory of martingales. Not only will it allow us to give another proof of the above L_p bound, but it will also enable us to obtain certain weak-type and Φ-estimates for the operators S and T. Furthermore, it will also lead us to the study of higher-dimensional versions of Hilbert operator. To introduce these, pick a positive integer d, fix $j\in\{1,2,\,\ldots,\,d\}$ and let $\mathbb{R}^d_{j+}=\{x\in\mathbb{R}^d:x_j>0\}$. For any locally integrable function on $\mathbb{R}^d_{j+}$ and any $x\in\mathbb{R}^d_{j+}$, define

$$T_jf(x)=\frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}}\int_{\mathbb{R}^d_{j+}}\frac{f(y)(x_j+y_j)}{|x+y|^{d+1}}\mathrm{d}y.$$

Clearly, if $d=1$, then the family $\{T_1,T_2,\ldots,T_d\}$ contains only one element, the Hilbert operator on $(0,\infty)$.

We are ready to formulate the main results of this paper; we start with the continuous setting. Our first statement concerns the L_p-boundedness of $\{T_j\}_{j=1}^d$.

Theorem 1.1. *For any $d \geq 1$ and $j \in \{1, 2, \ldots, d\}$, we have*

$$||T_j||_{L_p(\mathbb{R}^d_{j+}) \to L_p(\mathbb{R}^d_{j+})} = \sin^{-1}(\pi/p). \tag{1.1}$$

The L_p inequality fails to hold for when $p = 1$ or $p = \infty$. However, in these boundary cases we will show certain weaker substitutes. Define the usual weak-L_1 quasinorm by $||f||_{L_{1,\infty}(\mathbb{R}^d_{j+})} = \sup_{\lambda>0} \left[\lambda|\{x \in \mathbb{R}^d_{j+} : |f(x)| \geq \lambda\}|\right]$. The first result is a weak-type version of (1.1) for $p = 1$; unfortunately, martingale methods allow us to establish this bound for $d = 1$ only.

Theorem 1.2. *We have*

$$||T||_{L_1(0,\infty) \to L_{1,\infty}(0,\infty)} = \pi^{-1}. \tag{1.2}$$

The problem for $d \geq 2$ arises from the fact that the passage from the martingale theory to the operators T_j involves the use of a kind of a conditional expectation, which is not a contraction on weak spaces; on the other hand, it is a contraction on L_p and hence Theorem 1.1 holds true. See Section 3.

The final result is a Φ-estimate for bounded and integrable functions. The reasoning will work for all dimensions.

Theorem 1.3. *Suppose that $\Phi : [0, \infty) \to [0, \infty)$ is a convex, strictly increasing function of class C^1, satisfying $\Phi(0) = \Phi'(0+) = 0$. Then for any $d \geq 1$, $j \in \{1, 2, \ldots, d\}$ and any integrable function f on $\mathbb{R}^d_{j+}$ satisfying $||f||_{L_\infty(\mathbb{R}^d_{j+})} \leq 1$, we have the sharp bound*

$$\int_{\mathbb{R}^d_{j+}} \Phi\big(|T_j f(x)|\big)\, dx \leq C_\Phi \int_{\mathbb{R}^d_{j+}} |f(x)|\, dx, \tag{1.3}$$

where

$$C_\Phi = \int_1^\infty \frac{\Phi\left(\frac{1}{\pi}\log s\right)}{(s-1)^2}\, ds. \tag{1.4}$$

Clearly, by straightforward scaling, the above result extends to general bounded integrable functions on $\mathbb{R}^d_{j+}$.

The above results have their versions in the discrete setting. The following statement is an immediate consequence of the above theorems, by standard approximation arguments.

Theorem 1.4. *Let $a = (a_n)_{n\geq 1}$ be an arbitrary sequence of real numbers.*

(i) We have

$$\#\{n : |Sa(n)| > 1\} \leq \frac{1}{\pi}\sum_{n=1}^{\infty} |a_n|.$$

(ii) If $||(a_n)_{n\geq 1}||_{\ell_\infty} \leq 1$, then for any function Φ as in the statement of Theorem 1.3 we have

$$\sum_{n=1}^{\infty} \Phi\left(|Sa(n)|\right) \leq C_\Phi \sum_{n=1}^{\infty} |a_n|.$$

The inequalities are sharp.

The remainder of the paper is divided into three sections. Section 2 contains our probabilistic contribution - appropriate martingale versions of (1.1), (1.2) and (1.3). In Section 3 we show how to exploit these estimates to deduce the bounds for the operators S and T. The final part is devoted to certain localized estimates for

Hilbert and Riesz transforms, which can be regarded as extensions of the theorems formulated above.

2. Martingale inequalities

This section is devoted to the probabilistic counterparts of the results formulated in the introduction. For the sake of clarity, we have decided to split the material into a few parts.

2.1. **Background and notation.** Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing family of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all the events of probability 0. Let X, Y be two adapted real-valued martingales with continuous trajectories. The maximal functions of X and Y will be given by $X^* = \sup_{t\geq 0}|X_t|$, $Y^* = \sup_{t\geq 0}|Y_t|$. If the martingales converge almost surely, their limits will be denoted by X_∞ and Y_∞, respectively. The symbol $[X, Y]$ will stand for the quadratic covariance process of X and Y: consult, for instance, Dellacherie and Meyer [12] for details. The martingales X, Y are said to be *orthogonal* if the process $[X, Y]$ is constant with probability 1. Following Bañuelos and Wang [3] and Wang [34], Y is said to be *differentially subordinate* to X, if the process $([X, X]_t - [Y, Y]_t)_{t\geq 0}$ is nonnegative and nondecreasing as a function of t.

The differential subordination (with or without orthogonality of martingales) implies plenty of interesting inequalities comparing the sizes of X and Y (e.g., L_p estimates, weak-type bounds, etc.). This type of problems was first studied by Burkholder in the eighties (see the seminal papers [8] and [10]) and, by now, the literature on the subject is quite extensive. It is impossible to give even a short review here and we refer the interested reader to the recent monograph [25] by Osękowski, which is devoted to the detailed exposition of this area. Here we only mention one result, due to Bañuelos and Wang [34], which will be needed in our further considerations. We use the notation $||X||_p = \sup_{t\geq 0}||X_t||_p$ for $1 \leq p \leq \infty$.

Theorem 2.1. *Suppose that X, Y are orthogonal martingales such that Y is differentially subordinate to X. Then for any $1 < p < \infty$,*

$$||Y||_p \leq \cot\left(\frac{\pi}{2p^*}\right)||X||_p,$$

where $p^* = \max\{p, p/(p-1)\}$. The constant is the best possible.

In what follows, we will also work with a slightly stronger condition than the differential subordination: in some estimates it will be necessary to assume that the process $[X, X] - [Y, Y]$ is constant and nonnegative with probability 1. With a lack of a better word, in such a case we will say that X and Y are *differentially equivalent.*

The martingale inequalities we plan to study are of very unusual form. Let us explain briefly their connection with the operators T and T_j. First, note that these operators are closely related to the Hilbert transform $\mathcal{H}$ and the Riesz transforms R_j, classical objects in harmonic analysis. Recall that the latter operators are given by the principal value integrals

$$\mathcal{H}f(x) = \frac{1}{\pi}\text{p.v.}\int_{\mathbb{R}} \frac{f(y)}{x-y}\text{d}y,$$

for sufficiently regular $f : \mathbb{R} \to \mathbb{R}$, and

$$R_j f(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}} \text{ p.v.} \int_{\mathbb{R}^d} \frac{f(y)(x_j - y_j)}{|x-y|^{d+1}} \mathrm{d}y, \qquad j = 1, 2, \ldots, d,$$

for sufficiently regular $f : \mathbb{R}^d \to \mathbb{R}$. There is a well-established connection between the estimates for these operators and the theory of orthogonal martingales satisfying the differential subordination. Actually, the two settings are essentially parallel and the optimal constants are the same: more precisely, the probabilistic analogue of the pair $(f, \mathcal{H}f)$ or $(f, R_j f)$ is the pair (X, Y). For instance, as the reader immediately notices, the sharp L_p bound of Theorem 2.1 is precisely the probabilistic version of the identity $||\mathcal{H}||_{L_p(\mathbb{R}) \to L_p(\mathbb{R})} = ||R_j||_{L_p(\mathbb{R}^d) \to L_p(\mathbb{R}^d)} = \cot(\pi/(2p^*))$, obtained by Bañuelos and Wang [3], Iwaniec and Martin [22] and Pichorides [29]. See the papers [3], [5], [23], [26] and the monograph [25] for the more detailed study of this phenomenon. This will also be made more clear in Section 3 below.

Coming back to T and T_j, we see that if f is a function on $[0, \infty)$, and we extend it to the whole real line by putting $f(x) = 0$ for $x \le 0$, then we have the identity

$$Tf(x) = -\mathcal{H}f(-x) \qquad \text{for } x > 0. \tag{2.1}$$

Similarly, if we take an arbitrary $f : \mathbb{R}^d_{j+} \to \mathbb{R}$ and extend it to $\mathbb{R}^d$ by $f(x) = 0$ for $x \notin \mathbb{R}^d_{j+}$, then

$$T_j f(x) = -R_j f(-x) \qquad \text{for } x \in \mathbb{R}^d_{j+}. \tag{2.2}$$

Consequently, any estimate for $\mathcal{H}$ or R_j (obtained, for instance, by probabilistic tools) immediately yields the corresponding bound for T or T_j, respectively. However, such an inequality is in general not sharp, and we can do better than that, by appropriate "fine-tuning" of the martingale methods. For instance, it is evident that the operators T and T_j are positive and hence in the proof of any reasonable estimate one may restrict oneself to the class of nonnegative functions. This, in the probabilistic setting, leads to the restriction to nonnegative X's, which in general improves the constants (in comparison to the case of general X).

Unfortunately, this is still not good enough. The problem is of the following type. Suppose we want to establish the L_p bound for T. If we use (2.1) and then apply the appropriate bound for $\mathcal{H}$ (i.e., we write $||Tf||_{L_p(0,\infty)} \le ||\mathcal{H}f||_{L_p(\mathbb{R})} \le c_p||f||_{L_p(0,\infty)}$), we do not discard the behavior of $\mathcal{H}f$ on the negative halfline, which can be substantial. A similar problem occurs in the higher dimensions when applying (2.2): then we do not control the contribution of $R_j f$ coming from the set $\mathbb{R}^d \setminus \mathbb{R}^d_{j+}$. Fortunately, this difficulty can be overcome, as we will see now. In analogy with the preceding reasoning, given a nonnegative function f on $\mathbb{R}^d_{j+}$, let us extend both f and $T_j f$ to the whole $\mathbb{R}^d$ by putting $f = T_j f = 0$ on $\mathbb{R}^d \setminus \mathbb{R}^d_{j+}$. Then we have the following crucial estimate:

$$T_j f(x) \le (-R_j f(-x))_+ 1_{\{f(-x)=0\}}, \qquad \text{for almost all } x \in \mathbb{R}^d_{j+}. \tag{2.3}$$

Indeed, if $x_j < 0$, then the left-hand side is zero, while the right-hand side is nonnegative; on the other hand, if $x_j > 0$, then $f(-x) = 0$ and

$$T_j f(x) = -\frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}} \int_{\mathbb{R}^d_{j+}} \frac{f(y)(-x_j - y_j)}{|-x-y|^{d+1}} \mathrm{d}y = -R_j f(-x),$$

so both sides are equal.

We turn to the formulation of our main probabilistic results.

Theorem 2.2. *Let $1 < p < \infty$ be fixed. Suppose that X, Y are orthogonal martingales such that Y is differentially subordinate to X. Assume further that X is nonnegative and bounded in L_p, and $Y_0 \equiv 0$. Then we have the sharp estimates*

$$\big|\big|(Y_\infty)_+ 1_{\{X_\infty=0\}}\big|\big|_p \leq \sin^{-1}(\pi/p)||X_\infty||_p \tag{2.4}$$

and

$$\big|\big|Y_\infty 1_{\{X_\infty=0\}}\big|\big|_p \leq C_p||X_\infty||_p, \tag{2.5}$$

where

$$C_p = \begin{cases} \sin^{-1/p}((p-1)\pi/2) & \text{if } 1 < p < 2, \\ \sin^{-1}(\pi/p) & \text{if } p \geq 2. \end{cases}$$

This statement is a version of moment inequalities for T_j. More precisely, the bound (2.4) will lead to (1.1), while the estimate (2.5) will correspond to appropriate L_p bound for the function $x \mapsto R_j f(-x) 1_{\{f(-x)=0\}}$ (which is of independent interest: see Section 4 below).

We should also briefly comment here on the existence of the pointwise limits X_∞ and Y_∞. Since X is nonnegative, the almost sure convergence of this process follows from the classical results of Doob [14]. On the other hand, Y is differentially subordinate to X, so the existence of Y_∞ is a consequence of the corresponding escape inequalities: see Lemma 4 in Wang [34]. The same reasoning will guarantee the existence of appropriate limits in the theorems below.

Next, we will establish the following weak-type inequality, which is a substitute for (2.4) and (2.5) in the case $p = 1$. Here we will assume the stronger condition of differential equivalence.

Theorem 2.3. *Suppose that X, Y are orthogonal, differentially equivalent martingales such that X is nonnegative and $Y_0 \equiv 0$. Then we have the sharp estimates*

$$\mathbb{P}(Y_\infty 1_{\{X_\infty=0\}} > 1) \leq \pi^{-1}||X_0||_1 \tag{2.6}$$

and

$$\mathbb{P}(|Y_\infty| 1_{\{X_\infty=0\}} > 1) \leq 2\pi^{-1}||X_0||_1. \tag{2.7}$$

Note that on the right we have the first moment of X_0; actually, neither of the inequalities holds if we replace X_0 by X_∞ there. Indeed, take (X, Y) to be the planar Brownian motion starting from $(1, 0)$, stopped at y-axis (clearly, this counterexample does not work for the preceding L_p estimates: X is not bounded in L_p).

Finally, we establish the following inequality in the bounded case, which is a martingale version of (1.3).

Theorem 2.4. *Let $\Phi : [0, \infty) \to [0, \infty)$ be a convex function of class C^1 satisfying $\Phi(0) = \Phi'(0+) = 0$. Assume further that X, Y are orthogonal martingales such that X takes values in the interval $[0, 1]$, $Y_0 \equiv 0$ and Y is differentially subordinate to X. Then we have the sharp estimates*

$$\mathbb{E}\Phi\left((Y_\infty)_+ 1_{\{X_\infty=0\}}\right) \leq C_\Phi||X_\infty||_1 \tag{2.8}$$

and

$$\mathbb{E}\Phi\left(|Y_\infty| 1_{\{X_\infty=0\}}\right) \leq 2C_\Phi||X_\infty||_1, \tag{2.9}$$

where C_Φ is given in (1.4) above.

Each of the above estimates will be established in a separate subsection below. However, as the arguments leading to the estimates share the same pattern, we have decided to explain first the approach in a general setting. A similar abstract description can be found in the work of Bañuelos and Wang [3], Wang [34] and Osękowski [25].

2.2. **On a method of proof.** Fix a Borel, locally bounded function $V : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ and suppose that our goal is to establish the estimate

$$\mathbb{E}V(X_\infty, Y_\infty) \leq 0 \tag{2.10}$$

for any pair (X, Y) of almost surely convergent, orthogonal martingales such that X is nonnegative, $Y_0 \equiv 0$ and such that one of the following conditions holds:

(A) Y is differentially subordinate to X.
(B) X and Y are differentially equivalent.

In order to study this problem, we use Burkholder's method, which, generally speaking, rests on the construction of an appropriate special function. Suppose that $U : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is a continuous function, which is of class C^1 in the interior of its domain. Assume further that there are pairwise disjoint open sets $D_1, D_2, \ldots, D_n \subset [0,\infty) \times \mathbb{R}$, satisfying $\overline{D_1} \cup \overline{D_2} \cup \ldots \cup \overline{D_n} = [0,\infty) \times \mathbb{R}$ and such that U is of class C^2 on each D_j. Finally, consider the following structural conditions the function U might satisfy:

$$U(x,0) \leq 0 \qquad \text{for all } x \geq 0, \tag{2.11}$$

$$U(x,y) \geq V(x,y) \qquad \text{for all } (x,y) \in [0,\infty) \times \mathbb{R}, \tag{2.12}$$

$$U \qquad \text{is superharmonic,} \tag{2.13}$$

$$U_{xx} \leq 0 \qquad \text{for all } (x,y) \in D_1 \cup D_2 \cup \ldots \cup D_n. \tag{2.14}$$

The existence of such a function brings us very close to the validity of (2.10). *Roughly speaking*, if we can find U satisfying all the above conditions, then (2.10) holds true for any orthogonal pair (X, Y) satisfying (A). Furthermore, the existence of U satisfying the above regularity and the conditions (2.11), (2.12) and (2.13) ((2.14) is not required) implies the validity of (2.10) under (B).

To see this, we first use the mollification trick and approximate U by a function which has appropriate regularity (this argument goes back to the works of Burkholder [9], Bañuelos and Wang [3] and Wang [34]). Fix a nonnegative function g of class C^∞, supported on the unit ball of $\mathbb{R}^2$, satisfying $\int_{\mathbb{R}^2} g = 1$. For a given $\delta > 0$, let U^δ stand for the convolution of U and g: that is,

$$U^\delta(x,y) = \int_{\mathbb{R}^2} U(x + \delta u, y + \delta v) g(u,v) \mathrm{d}u\mathrm{d}v,$$

where $x \geq \delta$ and $y \in \mathbb{R}$. Clearly, this function is of class C^∞, inherits the superharmonicity property and satisfies $U^\delta_{xx}(x,y) \leq 0$ for all $x > \delta$ and $y \in \mathbb{R}$.

Next, fix a pair (X, Y) of almost surely convergent, orthogonal martingales such that X is nonnegative, $Y_0 \equiv 0$ and such that Y is *differentially subordinate* to X (the case of differential equivalence will be studied later). Consider the stopping time

$$\tau_N = \inf\{t \geq 0 : |X_t| + |Y_t| \geq N\}, \tag{2.15}$$

where N is a given large positive integer. Next, pick $\varepsilon \geq \delta$. Since U^δ is of class C^∞, we may apply Itô's formula to this function and the pair $Z_t = (\varepsilon +$

$X^{\tau_N}1_{\{\tau_N>0\}}, Y^{\tau_N}1_{\{\tau_N>0\}})$, where $X^{\tau_N}=(X_{\tau_N\wedge t})_{t\geq 0}$ (we need to add ε to X^{τ_N} to make sure that the composition of U^δ and the pair makes sense). As the result, we obtain, for any $t\geq 0$,

$$U^\delta(Z_t)=I_0+I_1+I_2+I_3/2, \tag{2.16}$$

where

$$\begin{aligned}
I_0&=U^\delta(Z_0),\\
I_1&=\int_{0+}^{\tau_N\wedge t}U_x^\delta(Z_s)\mathrm{d}X_s+\int_{0+}^{\tau_N\wedge t}U_y^\delta(Z_s)\mathrm{d}Y_s,\\
I_2&=\int_{0+}^{\tau_N\wedge t}U_{xy}^\delta(Z_s)\mathrm{d}[X,Y]_s,\\
I_3&=\int_{0+}^{\tau_N\wedge t}U_{xx}^\delta(Z_s)\mathrm{d}[X,X]_s+\int_{0+}^{\tau_N\wedge t}U_{yy}^\delta(Z_s)\mathrm{d}[Y,Y]_s.
\end{aligned}$$

Let us now analyze the terms I_0 through I_3 separately. As we have assumed above, the martingale Y starts from the origin. Therefore, (2.11) gives that $I_0=U(\varepsilon+X_0 1_{\{\tau_N>0\}},0)\leq 0$. The term I_1 has zero expectation, by elementary properties of stochastic integrals. We have $I_2=0$, since the orthogonality implies $\mathrm{d}[X,Y]_s=0$. To handle I_3, we exploit the differential subordination and the inequality $U_{xx}^\delta\leq 0$. We get

$$\begin{aligned}
I_3&\leq\int_{0+}^{\tau_N\wedge t}U_{xx}^\delta(Z_s)\mathrm{d}[Y,Y]_s+\int_{0+}^{\tau_N\wedge t}U_{yy}^\delta(Z_s)\mathrm{d}[Y,Y]_s\\
&=\int_{0+}^{\tau_N\wedge t}\Delta U^\delta(Z_s)\mathrm{d}[Y,Y]_s\leq 0,
\end{aligned}$$

where in the last passage we have exploited the superharmonicity of U^δ. Plugging all these facts into (2.16) and taking expectation of both sides yields $\mathbb{E}U^\delta(Z_t)\leq 0$. Now let $\delta\to 0$; since the function U is continuous, we have the pointwise convergence $U^\delta(x,y)\to U(x,y)$ for all $(x,y)\in(0,\infty)$. But the process Z is bounded, so Lebesgue's dominated convergence theorem gives $\mathbb{E}U(Z_t)\leq 0$. Next, we let $t\to\infty$. By the continuity of the trajectories and the boundedness of the process Z, we get $\mathbb{E}U(\varepsilon+X_{\tau_N}1_{\{\tau_N>0\}},Y_{\tau_N}1_{\{\tau_N>0\}})\leq 0$. Finally, let $\varepsilon\to 0$ to get $\mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}},Y_{\tau_N}1_{\{\tau_N>0\}})\leq 0$, again by Lebesgue's dominated convergence theorem. Essentially, this is the farthest point the general method can take us. Our plan is to let $N\to\infty$ to obtain $\mathbb{E}U(X_\infty,Y_\infty)\leq 0$ and then apply (2.12) to get (2.10). However, we cannot do this passage to the limit without some further boundedness conditions on U, X and Y, which may be dependent on the specific inequality we study.

Now suppose we want to handle the more restrictive case in which X and Y are assumed to be differentially equivalent. Then it is enough to find U satisfying the above regularity and the conditions (2.11)-(2.13) (that is, we may remove the condition (2.14) from the list of the requirements). Indeed, the whole above analysis remains valid; the only change concerns the term I_3. If X and Y are differentially equivalent, then $\mathrm{d}[X,X]_t=\mathrm{d}[Y,Y]_t$ and hence

$$I_3=\int_{0+}^{t}\Delta U^\delta(\varepsilon+X_s,Y_s)\mathrm{d}[X,X]_s\leq 0,$$

by the superharmonicity of U^δ. So, as it was in the case of the differential subordination, we obtain the estimate $\mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}}, Y_{\tau_N}1_{\{\tau_N>0\}}) \leq 0$ and we need to carry out a limiting procedure basing on the boundedness conditions imposed on U, X and Y.

Remark 2.5. The method described above admits plenty of modifications. Let us mention here two of them, which will be useful in our further considerations. First, we do not have to assume that U is continuous on the whole halfplane $[0,\infty)\times\mathbb{R}$. For example, suppose we replace this condition by the following, weaker property: U is locally bounded, continuous on $(0,\infty)\times\mathbb{R}$, and for each y, the function $x \mapsto U(x,y)$ is continuous. Then one easily checks that the above proof works fine, as the passage $\lim_{\varepsilon\downarrow 0}\mathbb{E}U(\varepsilon + X_{\tau_N}1_{\{\tau_N>0\}}, Y_{\tau_N}1_{\{\tau_N>0\}}) = \mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}}, Y_{\tau_N}1_{\{\tau_N>0\}})$ is still valid.

Another extension, which will be needed below, is the following. Suppose that we want to establish (2.10) under (A) and the additional assumption that X takes values in the interval $[0,1]$. Then we do not have to construct U on the whole halfplane: it is enough to find $U : [0,1]\times\mathbb{R}\to\mathbb{R}$, satisfying (2.11), (2.12), (2.13) and (2.14) there. This is also very easy to see. For any $\delta\in(0,1/2)$, pick $\varepsilon\in(\delta,1/2)$ and let U^δ be a function given by the same convolution as above. This function is well-defined only on the strip $[\delta,1-\delta]\times\mathbb{R}$; therefore, if we want to apply Itô's formula, we need to make sure that the stochastic pair we use takes values in this set. This enforces us to use, for instance, the martingale $(\varepsilon+(1-2\varepsilon)X_t^{\tau_N},(1-2\varepsilon)Y_t^{\tau_N})$ (we do not need to include the indicator $1_{\{\tau_N>0\}}$: if N is large, this indicator is 1, since $X_0\in[0,1]$ and $Y_0\equiv 0$). The remaining arguments are the same: Itô's formula combined with the analysis of the corresponding terms I_1 to I_4 gives $\mathbb{E}U^\delta(\varepsilon+(1-2\varepsilon)X_t^{\tau_N},(1-2\varepsilon)Y_t^{\tau_N})\leq 0$, and then we let $\delta\to 0$, $t\to\infty$ and $\varepsilon\to 0$ to obtain $\mathbb{E}U(X_{\tau_N},Y_{\tau_N})\leq 0$. This puts us in the same position as in the preceding setting.

2.3. **Proof and sharpness of** (2.4)**.** Now we will show how the above methodology can be used to establish the moment estimate (2.4). This inequality is of the form (2.10), with $V(x,y)=(y_+)^p1_{\{x=0\}}-\sin^{-p}(\pi/p)x^p$. Consider $U:[0,\infty)\times\mathbb{R}\to\mathbb{R}$, given by

$$U(x,y)=\begin{cases}-\sin^{-p}(\pi/p)\cdot R^p\cos^p\theta & \text{if } \theta\in[-\pi/2,\pi/2-\pi/p],\\ -\sin^{-1}(\pi/p)\cdot R^p\sin\left(p\theta-\pi p/2-\pi/p\right) & \text{if } \theta\in(\pi/2-\pi/p,\pi/2],\end{cases}$$

where we have used the polar coordinates $x=R\cos\theta$, $y=R\sin\theta$ with $R\geq 0$ and $\theta\in[-\pi/2,\pi/2]$. Furthermore, put $D_1=\{(x,y):R>0,\,\theta\in(-\pi/2,\pi/2-\pi/p)\}$ and $D_2=\{(x,y):R>0,\,\theta\in(\pi/2-\pi/p,\pi/2)\}$. We will check that U has all the required properties.

Regularity. It is straightforward to check that U is continuous on $[0,\infty)\times\mathbb{R}$ and of class C^1 in the interior of its domain. It is also obvious that U is of class C^2 on D_1 and D_2.

The condition (2.11)*.* If $1<p<2$, then

$$U(x,0)=x^p\cdot\sin(\pi/p)\sin(\pi p/2+\pi/p)\leq 0,$$

since $\pi/p\in[0,\pi]$ and $\pi p/2+\pi/p\in[\pi,2\pi]$. On the other hand, if $p\geq 2$, then

$$U(x,0)=-x^p\sin^{-p}(\pi/p)\leq 0.$$

The majorization (2.12). If $\theta < \pi/2 - \pi/p$ or $\theta = \pi/2$, then both sides are equal; hence, we must prove the inequality for $\theta \in (\pi/2 - \pi/p, \pi/2)$. For these values of θ, the majorization can be rewritten as

$$-\frac{\sin(p\theta - \pi p/2 - \pi/p)}{\cos^p \theta} \geq -\sin^{1-p}(\pi/p).$$

As we have observed above, both sides are equal in the boundary case $\theta = \pi/2 - \pi/p$. Furthermore, if we denote the left-hand side by $F(\theta)$, then we easily compute that $F'(\theta) = -p\cos^{-p-1}\theta\cos((p-1)\theta - \pi p/2 - \pi/p)$. This is positive for $\theta \in (\pi/2 - \pi/p, \pi/2)$, since then the angle $(p-1)\theta - \pi p/2 - \pi/p$ lies between $-3\pi/2$ and $-\pi/2 - \pi/p$ (and hence the corresponding cosine function is negative).

Superharmonicity. It suffices to note that U is of class C^1 and satisfies $\Delta U = U_{xx} < 0$ on D_1 and $\Delta U = 0$ on D_2.

The condition (2.14). If $\theta < \pi/2 - \pi/p$, then $U(x,y) = -\sin^p(\pi/p)x^p$ and hence $U_{xx}(x,y) = -p(p-1)\sin^p(\pi/p)x^{p-2} < 0$. On the other hand, if $\theta > \pi/2 - \pi/p$, then some straightforward computations show that

$$U_{xx}(x,y) = -p(p-1)\sin(\pi/p)\cdot R^{p-2}\sin\big((p-2)\theta - \pi p/2 - \pi/p\big).$$

It is enough to observe that for the above values of θ, the angle $(p-2)\theta - \pi p/2 - \pi/p$ lies between $-\pi - \pi/p$ and $-2\pi + \pi/p$, and hence the sine function is positive; this implies $U_{xx} < 0$ and completes the analysis of (2.14).

Thus, the reasoning described in the preceding subsection implies that

$$\mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}}, Y_{\tau_N}1_{\{\tau_N>0\}}) \leq 0 \tag{2.17}$$

(recall that τ_N is given by (2.15)). Clearly, U enjoys the upper bound $|U(x,y)| \leq C(|x|^p + |y|^p)$ for some constant C depending only on p. Consequently,

$$|U(X_{\tau_N}, Y_{\tau_N})| \leq C((X^*)^p + (Y^*)^p).$$

Recall that we have assumed in the statement of the theorem that the martingale X is bounded in L_p. Hence, by Theorem 2.1, Y also has this property. Thus, by Doob's maximal estimate (cf. [14]), the right-hand side above is integrable. So, by Lebesgue's dominated convergence theorem, we are allowed to let $N \to \infty$ in (2.17). This gives $\mathbb{E}U(X_\infty, Y_\infty) \leq 0$ and the use of (2.12) establishes the desired inequality (2.4).

Sharpness. The fact that the constant $\sin^{-1}(\pi/p)$ is the best possible will follow from the arguments of Section 3. Therefore, for the reader's convenience, we will only describe an example which implies the optimality (and skip the verification of its properties). Fix $1 < p < \infty$, a number $\varphi < \pi/p$ and consider the sector $A_\varphi = \{(x,y) : \pi/2 - \varphi \leq \theta \leq \pi/2\}$. Let $B = (B^1, B^2)$ denote the two-dimensional Brownian motion starting at $(\frac{1}{2}\tan\varphi, 1)$ and stopped upon exiting A_φ. Finally, set $X = B^1$ and $Y = B^2 - 1$. Of course, then X, Y are orthogonal and differentially equivalent martingales; furthermore, X is nonnegative and $Y_0 = 0$, so all the requirements on the processes are met. One can show that

$$\lim_{\varphi\to\pi/p} \frac{||(Y_\infty)_+ 1_{\{X_\infty=0\}}||_p}{||X_\infty||_p} = \sin^{-1}(\pi/p).$$

2.4. **Proof and sharpness of** (2.5)**.** When $p \geq 2$, then the calculations are essentially the same as in the previous case. The inequality is of the form (2.10) with $V(x,y) = |y|^p 1_{\{x=0\}} - \sin^{-p}(\pi/p)x^p$. Let $U : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be given by

$$U(x,y) = \begin{cases} -\sin^{-p}(\pi/p) \cdot R^p \cos^p \theta & \text{if } |\theta| \leq \pi/2 - \pi/p, \\ -\sin^{-1}(\pi/p) \cdot R^p \sin\left(p|\theta| - \pi p/2 - \pi/p\right) & \text{if } |\theta| \in (\pi/2 - \pi/p, \pi/2]. \end{cases}$$

This is a "symmetrized" version of the function U from the preceding estimate in the sense that both objects agree on $[0,\infty) \times [0,\infty)$, and the function we have just introduced satisfies $U(x,y) = U(x,-y)$ for all $x \geq 0$ and $y \in \mathbb{R}$. Letting $D_1 = \{(x,y) : |\theta| \leq \pi/2 - \pi/p\}$ and $D_2 = \{(x,y) : |\theta| \in (\pi/2 - \pi/p, \pi/2]$, we repeat the above calculations and show that U has all the required properties. Thus, (2.17) holds true and the same limiting arguments as above yield (2.5).

In the case $1 < p < 2$, some new objects have to be introduced. We must take $V(x,y) = |y|^p 1_{\{x=0\}} - \sin^{-1}\left((p-1)\pi/2\right) x^p$ and define $U : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ by

$$U(x,y) = \frac{R^p \cos(p\theta)}{\cos(p\pi/2)}.$$

There is only one region $D_1 = (0,\infty) \times \mathbb{R}$. Let us check that all the requirements are met.

Regularity. This is trivial: U is of class C^∞ on D_1.

The condition (2.11). This is also evident, since $\cos(p\pi/2) < 0$.

The majorization (2.12). We may restrict ourselves to the case $\theta \geq 0$, since both U and V are symmetric with respect to the x-axis. If $\theta = 0$ or $\theta = \pi/2$, then both sides are equal. For the remaining θ's, note that the bound reads

$$\frac{R^p \cos(p\theta)}{\cos(p\pi/2)} \geq \frac{R^p \cos^p \theta}{\cos(p\pi/2)},$$

or, equivalently, $\cos(p\theta)/\cos^p \theta \leq 1$. Denoting the left-hand side by $F(\theta)$, we see that $F(0) = 1$ and

$$F'(\theta) = -\frac{\sin((p-1)\theta)}{\cos^{p+1}\theta} \leq 0 \qquad \text{for } |\theta| < \pi/2.$$

This yields (2.12).

Superharmonicity. This is trivial: U is harmonic on $(0,\infty) \times \mathbb{R}$.

The condition (2.14). A little calculation shows that

$$U_{xx}(x,y) = \frac{p(p-1)R^{p-2}\cos((p-2)\theta)}{\cos(p\pi/2)}.$$

Since $|\theta| < \pi/2$ and $1 < p < 2$, we see that the numerator is positive; on the other hand, the denominator is negative, so (2.14) holds true. So, the reasoning from §2.2 gives (2.17) and the same limiting argument as above yields the validity of (2.5).

Sharpness. When $p \geq 2$, the optimality of the constant C_p follows from the sharpness of (2.4) (which, in turn, will be a consequence of the reasoning of Section 3). So, let us assume that $1 < p < 2$. Let $c > 1$ be a large positive number and consider the set $A_c = ([0,\infty) \times \mathbb{R}) \setminus ([c,\infty) \times \{0\})$. Let (X,Y) be a two-dimensional Brownian motion starting at $(1,0)$ and stopped at the boundary of A_c. Then it is not difficult to see that both X and Y are L_p bounded, and $\lim_{c\to\infty} ||X||_p =$

$\lim_{c\to\infty}||Y||_p=\infty$. Since the above function U is harmonic on $(0,\infty)\times\mathbb{R}$, an application of Itô's formula yields

$$\mathbb{E}U(X_\infty,Y_\infty)=U(1,0).$$

On the other hand, the functions U and V coincide at the boundary of A_c, so the above equality implies

$$\mathbb{E}|Y_\infty|^p 1_{\{|X_\infty|=0\}}=\sin^{-1}((p-1)\pi/2)\mathbb{E}X_\infty^p+U(1,0).$$

Since $\mathbb{E}X_\infty^p$ explodes as $c\to\infty$, the constant $\sin^{-1}((p-1)\pi/2)$ cannot be replaced by a smaller number. This establishes the sharpness.

2.5. **Proof and sharpness of** (2.6) **and** (2.7)**.** Now we turn our attention to the weak-type inequality. Actually, it is enough to establish the first bound: indeed, having done this, we write

$$\mathbb{P}(|Y_\infty|1_{\{X_\infty=0\}}>1)=\mathbb{P}((Y_\infty)_+1_{\{X_\infty=0\}}>1)+\mathbb{P}((-Y_\infty)_+1_{\{X_\infty=0\}}>1) \tag{2.18}$$

and bound each probability on the right by $||X_0||_1/\pi$ (to see that this is permitted for the second term, note that the pair $(X,-Y)$ satisfies the assumptions).

Put $V(x,y)=1_{\{y1_{\{x=0\}}>1\}}-\frac{1}{\pi}x$. Then the inequality (2.6) is not exactly of the form (2.10), since in the former bound we have the appearance of the variable X_0. However, let us not worry about that and introduce the special function U by

$$U(x,y)=\begin{cases}\dfrac{1}{2}-\dfrac{1}{\pi}\arctan\left(\dfrac{1-y}{x}\right)-\dfrac{x}{\pi} & \text{if } x>0,\\ 1_{\{y>1\}}+\frac{1}{2}1_{\{y=1\}} & \text{if } x=0.\end{cases}$$

Here there is only one domain $D_1=(0,\infty)\times\mathbb{R}$. Let us verify that the function has the required properties.

Regularity. The first problem is that U is not continuous at the point $(0,1)$. However, it is of class C^2 on $(0,\infty)\times\mathbb{R}$ and for each $y\in\mathbb{R}$, the function $x\mapsto U(x,y)$ is continuous on $[0,\infty)$, which is sufficient for our purposes (see Remark 2.5 above).

The condition (2.11). For any $x\geq 0$, we have

$$U(x,0)=\frac{1}{2}-\frac{1}{\pi}\arctan\frac{1}{x}-\frac{x}{\pi}=\frac{1}{\pi}(\arctan x-x)\leq 0,$$

as needed.

The majorization (2.12). Observe that $\arctan((1-y)/x)\leq\pi/2$, and hence $U(x,y)\geq-\frac{x}{\pi}=V(x,y)$ when $y<1$ or $x>0$. Next, we have $U(0,1)=1/2>V(0,1)$. For remaining (x,y) (i.e., for $x=0$ and $y>1$), both sides of (2.12) are equal.

Superharmonicity. This is evident: as one easily verifies, U is actually harmonic on $(0,\infty)\times\mathbb{R}$.

Therefore, the reasoning of §2.2 gives us $\mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}},Y_{\tau_N}1_{\{\tau_N>0\}})\leq 0$ which, by the nonnegativity of X, can be rewritten as

$$\mathbb{E}\big[U(X_{\tau_N}1_{\{\tau_N>0\}},Y_{\tau_N}1_{\{\tau_N>0\}})+\pi^{-1}X_{\tau_N}1_{\{\tau_N>0\}}\big]\leq\pi^{-1}\mathbb{E}X_0 1_{\{\tau_N>0\}}.$$

The expression in the square brackets is bounded from below, so the use of Fatou's lemma and Lebesue's monotone convergence theorem yields

$$\mathbb{E}\big[U(X_\infty,Y_\infty)+\pi^{-1}X_\infty\big]\leq\pi^{-1}\mathbb{E}X_0.$$

But, by (2.12), the left-hand side is not smaller than $\mathbb{P}(Y_\infty 1_{\{X_\infty=0\}} > 1)$. This completes the proof of (2.6).

Sharpness. It is enough to show that the constant in (2.7) is optimal; by the argumentation in (2.18), this will also imply that the improvement of (2.6) is impossible. Fix a positive number c. Let (X, Y) be a two-dimensional Brownian motion, starting at a point $(c, 0)$ and stopped upon hitting the y-axis. Then $||X_0||_1 = c$, $X_\infty = 0$ and the law of Y_∞ is the Cauchy distribution with parameter c. Hence

$$\mathbb{P}(|Y_\infty|1_{\{X_\infty=0\}} > 1) = \mathbb{P}(|Y_\infty| > 1) = \int_{\mathbb{R}\setminus[-1,1]} \frac{c}{t^2+c^2}\mathrm{d}t = 1 - \frac{2}{\pi}\arctan\left(\frac{1}{c}\right)$$

and therefore

$$\lim_{c\to 0} \frac{\mathbb{P}(|Y_\infty|1_{\{X_\infty=0\}} > 1)}{||X_0||_1} = \frac{2}{\pi}.$$

This proves the desired sharpness.

2.6. **Proof and sharpness of** (2.8) **and** (2.9)**.** It is enough to establish the first estimate; arguing as in the preceding subsection, we see that then (2.9) follows. Here the reasoning will be a little more involved, as the special function is more complicated. Assume that $\Phi : [0,\infty) \to [0,\infty)$ is a fixed convex function satisfying $\Phi(0) = \Phi'(0+) = 0$. Clearly, we may assume in addition that $C_\Phi < \infty$, since otherwise there is nothing to prove. The inequality (2.8) is of the form (2.10), with $V(x, y) = \Phi(y_+ 1_{\{x=0\}}) - C_\Phi x$. As we have observed in Remark 2.5 above, it suffices to construct an appropriate special function on the strip $[0, 1] \times \mathbb{R}$. This will be done by the following two-step procedure. First define an auxiliary function $\mathcal{U} : \mathbb{R} \times (0,\infty) \to \mathbb{R}$ by the Poisson integral

$$\mathcal{U}(\alpha, \beta) = \frac{1}{\pi}\int_{-\infty}^{-1} \frac{\beta\Phi\left(\frac{1}{\pi}\log(-t)\right)}{(\alpha-t)^2+\beta^2}\mathrm{d}t.$$

Obviously, $\mathcal{U}$ is harmonic and satisfies

$$\lim_{(\alpha,\beta)\to(z,0)} \mathcal{U}(\alpha,\beta) = \begin{cases} \Phi\left(\frac{1}{\pi}\log(-z)\right) & \text{if } z < -1,\\ 0 & \text{if } z > -1.\end{cases} \tag{2.19}$$

Next, consider a conformal mapping $\varphi(z) = -e^{-i\pi z}$, or, in real coordinates,

$$\varphi(x, y) = \left(-e^{\pi y}\cos(\pi x), e^{\pi y}\sin(\pi x)\right).$$

One easily verifies that φ maps $(0, 1) \times \mathbb{R}$ onto the halfplane $\mathbb{R} \times (0,\infty)$. Now we are ready to define U on the strip $(0, 1) \times \mathbb{R}$: put

$$U(x, y) = \mathcal{U}(\varphi(x, y)) - C_\Phi x. \tag{2.20}$$

The function U is harmonic on $(0, 1) \times \mathbb{R}$, as a composition of a harmonic function with a conformal mapping. Furthermore, by (2.19), it can be extended to the continuous function on the whole strip $[0, 1]\times\mathbb{R}$ by $U(0, y) = \Phi(y_+)$, $U(1, y) = -C_\Phi$. As one easily verifies, U admits the following explicit formula in the interior of its domain:

$$U(x, y) = \frac{1}{\pi}\int_1^\infty \frac{\Phi\left(\left(\frac{1}{\pi}\log s + y\right)_+\right)\sin(\pi x)}{(s-\cos(\pi x))^2+\sin^2(\pi x)}\mathrm{d}s - C_\Phi x. \tag{2.21}$$

To complete the description of the setting, let us note that we consider only one domain $D_1 = (0,1) \times \mathbb{R}$. We will now verify that the function U enjoys all the required properties.

Regularity. This is clear: as we have already noted above, U is continuous on the strip and is obviously of class C^∞ in its interior.

Superharmonicity. This is trivial, since U is harmonic inside its domain.

The conditions (2.11) *and* (2.14). For any fixed $s \in \mathbb{R}$, the function $y \mapsto \Phi\left(\left(\frac{1}{\pi}\log s + y\right)_+\right)$ is convex and hence, by (2.21), the function U is "vertically convex", i.e., we have $U_{yy} \geq 0$ on $(0,1)\times\mathbb{R}$. This, by the harmonicity of U, implies $U_{xx}(x,0) \leq 0$ for $x \in (0,1)$. Thus, the condition (2.11) follows at once from the equalities $U(0,0) = 0$ and

$$\begin{aligned} U_x(0+,0) &= \lim_{x\mapsto 0} \frac{U(x,0)}{x} \\ &= \lim_{x\to 0}\left[\frac{\sin(\pi x)}{\pi x}\int_1^\infty \frac{\Phi\left(\frac{1}{\pi}\log s\right)\sin(\pi x)}{(s-\cos(\pi x))^2+\sin^2(\pi x)}\mathrm{d}s - C_\Phi\right] = 0. \end{aligned} \tag{2.22}$$

The majorization (2.12). We will study a stronger estimate

$$U(x,y) \geq \Phi(y_+)(1-x) - C_\Phi x. \tag{2.23}$$

Note that for a fixed $y \in \mathbb{R}$, the left-hand side is a concave function of x (we have shown above that $U_{xx} \leq 0$ on $(0,1)\times\mathbb{R}$), while the right-hand side is linear in x. Consequently, it is enough to prove the bound for $x = 0$ and $x = 1$; however, in both these cases, the estimate (2.23) becomes an equality.

Consequently, for any N we have $\mathbb{E}U(X_{\tau_N}1_{\{\tau_N>0\}}, Y_{\tau_N}1_{\{\tau_N>0\}}) \leq 0$ and hence, by the stronger majorization (2.23), we have the bound

$$\mathbb{E}\Phi((Y_{\tau_N})_+1_{\{\tau_N>0\}})(1-X_{\tau_N}1_{\{\tau_N>0\}}) \leq C_\Phi \mathbb{E}X_{\tau_N}1_{\{\tau_N>0\}}.$$

Now we let $N \to \infty$. The right hand side converges to $\mathbb{E}X_\infty$, by Lebesgue's dominated convergence theorem (X is bounded). The left-hand side is dealt with Fatou's lemma. As the result, we get

$$\mathbb{E}\Phi((Y_\infty)_+)(1-X_\infty) \leq C_\Phi \mathbb{E}X_\infty,$$

which is stronger than the inequality (2.8). The desired bound follows.

Sharpness. It suffices to prove that (2.9) is sharp. Fix a parameter $c \in (0,1)$ close to 0, and let (X,Y) be a planar Brownian motion, started at $(c,0)$ and stopped upon exiting the strip $[0,1]\times\mathbb{R}$. Then, of course, we have $X_\infty \in \{0,1\}$ almost surely and $\mathbb{E}X_\infty = c$. The distribution of Y_∞ on the set $\{X_\infty = 0\}$ is well-known and one can compute directly that $\mathbb{E}\Phi(|Y_\infty|1_{\{X_\infty=0\}})$; however, for reader's convenience, let us present here a quick derivation of this quantity, based on the above function U. Clearly, the distribution of Y_∞ is symmetric, so

$$\begin{aligned} \mathbb{E}\Phi(|Y_\infty|1_{\{X_\infty=0\}}) = 2\mathbb{E}\Phi((Y_\infty)_+1_{\{X_\infty=0\}}) &= 2\mathbb{E}\left[U(X_\infty,Y_\infty) + C_\Phi X_\infty\right] \\ &= 2U(c,0) + 2C_\Phi c. \end{aligned}$$

Consequently, by (2.22), we have

$$\lim_{c\downarrow 0}\frac{\mathbb{E}\Phi(|Y_\infty|1_{\{X_\infty=0\}})}{\mathbb{E}X_\infty} = 2C_\Phi + \lim_{c\downarrow 0}\frac{U(c,0)}{c} = 2C_\Phi,$$

which gives the claim.

3. Inequalities for S, T and T_j

Throughout this section, d is a fixed positive integer. The inequality (2.3) shows how to handle inequalities for the operator T_j by means of the corresponding bounds for the positive part of the Riesz transform. To handle the latter estimates, we will exploit the well-known probabilistic representation of Riesz transforms in terms of the so-called background radiation process, introduced by Gundy and Varopoulos in [18]. Let us briefly describe this connection. Suppose that X is a Brownian motion in $\mathbb{R}^d$ and let Y be an independent Brownian motion in $\mathbb{R}$ (both processes start from the appropriate origins). For any $y>0$, introduce the stopping time $\tau(y)=\inf\{t\geq 0: Y_t\in\{-y\}\}$. If f belongs to $\mathcal{S}(\mathbb{R}^d)$, the class of rapidly decreasing functions on $\mathbb{R}^d$, let $U_f:\mathbb{R}^d\times[0,\infty)\to\mathbb{R}$ stand for the Poisson extension of f to the upper half-space. That is,

$$U_f(x,y):=\mathbb{E}f\left(x+X_{\tau(y)}\right).$$

For any $(d+1)\times(d+1)$ matrix A we define the martingale transform $A{*}f$ by

$$A{*}f(x,y)=\int_{0+}^{\tau(y)} A\nabla U_f(x+X_s,y+Y_s)\cdot \mathrm{d}(X_s,Y_s).$$

Note that $A*f(x,y)$ is a random variable for each x, y. Now, for any $f\in C_0^\infty$, any $y>0$ and any matrix A as above, define $\mathcal{T}_A^y f:\mathbb{R}^d\to\mathbb{R}$ through the bilinear form

$$\int_{\mathbb{R}^d}\mathcal{T}_A^y f(x)g(x)\,\mathrm{d}x=\int_{\mathbb{R}^d}\mathbb{E}\big[A{*}f(x,y)g(x+X_{\tau(y)})\big]\mathrm{d}x, \tag{3.1}$$

where g runs over $C_0^\infty(\mathbb{R}^d)$. Less formally, $\mathcal{T}^y f$ is given as the following conditional expectation with respect to the measure $\tilde{\mathbb{P}}=\mathbb{P}\otimes\mathrm{d}x$ ($\mathrm{d}x$ denotes Lebesgue's measure on $\mathbb{R}^d$): for any $z\in\mathbb{R}^d$,

$$\mathcal{T}_A^y f(z)=\tilde{\mathbb{E}}\left[A{*}f(x,y)|x+X_{\tau(y)}=z\right].$$

See Gundy and Varopoulos [18] for the rigorous statement of this equality. The interplay between the operators $\mathcal{T}_A^y$ and Riesz transforms is explained in the following theorem, consult [18] or Gundy and Silverstein [17].

Theorem 3.1. *Let $A^j=[a_{\ell m}^j]$, $j=1,2,\ldots,d$ be the $(d+1)\times(d+1)$ matrices given by*

$$a_{\ell m}^j=\begin{cases}1 & \text{if } \ell=d+1,\ m=j,\\ -1 & \text{if } \ell=j,\ m=d+1,\\ 0 & \text{otherwise.}\end{cases}$$

Then $\mathcal{T}_{A^j}^y f\to R_j f$ almost everywhere as $y\to\infty$.

We will also require the following auxiliary fact (see [26], for instance). Namely, in the definition (3.1), g runs over the class $C_0^\infty(\mathbb{R}^d)$; however, having successfully defined the operator $\mathcal{T}_A^y$, we may extend the validity of (3.1) to a wider class of functions g.

Lemma 3.2. *Let $f\in C_0^\infty(\mathbb{R}^d)$ and $A=A^j$ for some j. Then for any $1<q<\infty$ and any $g\in L^q(\mathbb{R}^d)$, the equality (3.1) holds true.*

We are ready to establish the inequalities (1.1), (1.2) and (1.3).

Proof of $||T_j||_{L_p(\mathbb{R}^d_{j+})\to L_p(\mathbb{R}^d_{j+})} \leq \sin^{-1}(\pi/p)$. Fix $j\in\{1,2,\ldots,d\}$, $x\in\mathbb{R}$ and $y>0$. By a standard density argument, it suffices to establish the estimate (1.1) for $f\in C_0^\infty(\mathbb{R}^d)$. Furthermore, since T_j is a positive operator, we may assume that f is nonnegative. Consider the pair $\xi=(\xi_t)_{t\geq 0}$, $\zeta=(\zeta_t)_{t\geq 0}$ of martingales given by

$$\begin{aligned}\xi_t &= U_f(x+X_{\tau(y)\wedge t}, y+Y_{\tau(y)\wedge t})\\ &= U_f(x,y)+\int_{0+}^{\tau(y)\wedge t}\nabla U_f(x+X_s,y+Y_s)\cdot \mathrm{d}(X_s,Y_s)\end{aligned}$$

and

$$\zeta_t = -\int_{0+}^{\tau(y)\wedge t} A^j\nabla U_f(x+X_s,y+Y_s)\cdot\mathrm{d}(X_s,Y_s),$$

for $t\geq 0$. Then the martingale ζ is differentially subordinate to ξ, since

$$[\xi,\xi]_t-[\zeta,\zeta]_t = |U_f(x,y)|^2+\sum_{k\notin\{j,d+1\}}\int_{0+}^{\tau(y)\wedge t}\left|\frac{\partial U_f}{\partial x_k}(x+X_s,y+Y_s)\right|^2\mathrm{d}s$$

is nonnegative and nondecreasing as a function of t. Furthermore, ξ and ζ are orthogonal, which is a direct consequence of the equality $\langle A^jx,x\rangle=0$, valid for all $x\in\mathbb{R}^d$. Indeed,

$$[\xi,\zeta]_t = -\int_{0+}^{\tau(y)\wedge t}\langle A^j\nabla U_f(x+X_s,y+Y_s),\nabla U_f(x+X_s,y+Y_s)\rangle \mathrm{d}s = 0.$$

Finally, note that ξ is nonnegative (since so are f and U_f) and $\zeta_0\equiv 0$. By (2.4), we have $\mathbb{E}(\zeta_{\tau(y)})_+^p 1_{\{\xi_{\tau(y)}=0\}}\leq \sin^{-p}(\pi/p)\mathbb{E}\xi^p_{\tau(y)}$, or

$$\mathbb{E}(-A^j * f(x,y))_+^p 1_{\{f(x+X_{\tau(y)})=0\}}\leq \sin^{-p}(\pi/p)\mathbb{E}f(x+X_{\tau(y)})^p.$$

Integrating this estimate with respect to $x\in\mathbb{R}^d$ and using Fubini's theorem yields

$$\int_{\mathbb{R}^d}\mathbb{E}(-A^j * f(x,y))_+^p 1_{\{f(x+X_{\tau(y)})=0\}}\mathrm{d}x\leq \sin^{-p}(\pi/p)\int_{\mathbb{R}^d}(f(x))^p\mathrm{d}x.$$

Now take an arbitrary positive function $g\in L^q(\mathbb{R}^d)$. By the above estimate and Hölder's inequality, we get

$$\begin{aligned}\int_{\mathbb{R}^d}&\mathbb{E}\big[\big(-A^j * f(x,y)\big)1_{\{f(x+X_{\tau(y)})=0\}}g(x+X_{\tau(y)})\big]\mathrm{d}x\\ &\leq \sin^{-1}(\pi/p)||f||_{L_p(\mathbb{R}^d)}||g||_{L^q(\mathbb{R}^d)},\end{aligned}$$

or, by the definition of $T^y_{A^j}$ and Lemma 3.2 (applied to $x\mapsto 1_{\{f(x)=0\}}g(x)$),

$$\frac{\int_{\mathbb{R}^d}(-T^y_{A^j}f(x))1_{\{f(x)=0\}}g(x)\mathrm{d}x}{||g||_{L^q(\mathbb{R}^d)}}\leq \sin^{-1}(\pi/p)||f||_{L_p(\mathbb{R}^d)}.$$

Since g was an arbitrary positive function, the above inequality implies

$$||(-T^y_{A^j}f)_+1_{\{f=0\}}||_{L_p(\mathbb{R}^d)}\leq \sin^{-1}(\pi/p)||f||_{L_p(\mathbb{R}^d)}. \tag{3.2}$$

Letting $y\to\infty$ and combining this estimate with Lemma 3.1 and Fatou's lemma yields

$$||(-R_jf)_+1_{\{f=0\}}||_{L_p(\mathbb{R}^d)}\leq \sin^{-1}(\pi/p)||f||_{L_p(\mathbb{R}^d)}.$$

It remains to apply (2.3):

$$||T_jf||_{L_p(\mathbb{R}^{d,j}_+)}\leq ||(-R_jf)_+1_{\{f=0\}}||_{L_p(\mathbb{R}^d)}\leq \sin^{-1}(\pi/p)||f||_{L_p(\mathbb{R}^d)}. \qquad \square$$

Sharpness of (1.1) *in the case* $d = 1$. See Hardy [19]. □

Proof of (1.3). We use a similar reasoning as above. It is enough to handle non-negative f's only. An application of (2.8) to the martingales ξ and ζ gives

$$\mathbb{E}\Phi\big((-A{*}f(x,y))_+1_{\{f(x+X_{\tau(y)})=0\}}\big) \leq C_\Phi \mathbb{E} f(x+X_{\tau(y)}),$$

so integrating over $x \in \mathbb{R}^d$ yields

$$\int_{\mathbb{R}^d} \mathbb{E}\Phi\big((-A{*}f(x,y))_+1_{\{f(x+X_{\tau(y)})=0\}}\big) \leq C_\Phi \int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Pick $q \in (1,\infty)$ and a nonnegative $g \in L^q(\mathbb{R}^d)$. Let Ψ be the Legendre transform of Φ, i.e., the strictly increasing, C^1 convex function satisfying $\Psi(0) = \Psi'(0+) = 0$ such that Ψ' and Φ' are the inverses of each other. We obtain, by Young's inequality,

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathbb{E}\left[\big(-A{*}f(x,y)\big)1_{\{f(x+X_{\tau(y)})=0\}}\, g(x+X_{\tau(y)})\right]\mathrm{d}x \\
&\qquad \leq \int_{\mathbb{R}^d} \mathbb{E}\left[\big(-A{*}f(x,y)\big)_+1_{\{f(x+X_{\tau(y)})=0\}}\, g(x+X_{\tau(y)})\right]\mathrm{d}x \\
&\qquad \leq \int_{\mathbb{R}^d} \mathbb{E}\Phi\left(\big(-A{*}f(x,y)\big)_+1_{\{f(x+X_{\tau(y)})=0\}}\right)\mathrm{d}x + \int_{\mathbb{R}^d}\mathbb{E}\Psi\big(g(x+X_{\tau(y)})\big)\,\mathrm{d}x \\
&\qquad \leq C_\Phi \int_{\mathbb{R}^d} f(x)\mathrm{d}x + \int_{\mathbb{R}^d}\Psi(g(x))\,\mathrm{d}x.
\end{aligned}$$

Hence, (3.1) and Lemma 3.2 give

$$\int_{\mathbb{R}^d}\left[(-\mathcal{T}^y_{A^j}f(x))1_{\{f(x)=0\}}g(x) - \Psi(g(x))\right]\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Now fix $M > 0$. Then the above bound clearly yields

$$\int_{\mathbb{R}^d}\left[\min\{(-\mathcal{T}^y_{A^j}f(x))1_{\{f(x)=0\}}, M\}g(x) - \Psi(g(x))\right]\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x. \tag{3.3}$$

Apply this inequality with $g(x) = \Phi'\left(\min\left\{(-\mathcal{T}^y_{A^j}f(x))_+1_{\{f(x)=0\}}, M\right\}\right)$. It is easy to see that $g \leq c(-\mathcal{T}^y_{A^j}f)_+1_{\{f=0\}}$ for some positive $c = c(M,K)$ and hence $g \in L^q(\mathbb{R}^d)$, since the same is true for $(-T^y_{A^j}f)_+1_{\{f=0\}}$ (use (3.2) and the fact that $f \in C_0^\infty(\mathbb{R}^d) \subset L^q(\mathbb{R}^d)$). Since g vanishes on the set $\{\mathcal{T}^y_{A^j}f > 0\}$, the inequality (3.3) implies

$$\int_{\mathbb{R}^d}\left[\min\{(-\mathcal{T}^y_{A^j}f(x))_+1_{\{f(x)=0\}}, M\}g(x) - \Psi(g(x))\right]\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Therefore, since $\Psi(\Phi'(t)) + \Phi(t) = t\Phi'(t)$ for all t, we get

$$\int_{\mathbb{R}^d}\Phi\left(\min\{(-\mathcal{T}^y_{A^j}f(x))_+1_{\{f(x)=0\}}, M\}\right)\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Letting $M \to \infty$ and applying Fatou's lemma gives

$$\int_{\mathbb{R}^d}\Phi\left((-\mathcal{T}^y_{A^j}f(x))_+1_{\{f(x)=0\}}\right)\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Let $y \to \infty$ and combine Lemma 3.1 with Fatou's lemma to obtain

$$\int_{\mathbb{R}^d}\Phi\left((-R_jf(x))_+1_{\{f(x)=0\}}\right)\mathrm{d}x \leq C_\Phi\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Hence, by (2.3),

$$\int_{\mathbb{R}^d_{j+}} \Phi(T_j f(x))\mathrm{d}x \leq \int_{\mathbb{R}^d} \Phi\left((-R_j f(x))_+ 1_{\{f(x)=0\}}\right)\mathrm{d}x \leq C_\Phi \int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

The proof is complete. □

Sharpness of (1.3) *in the case* $d=1$. It is enough to show the optimality of the constant for the operator S. Suppose first that the constant C_Φ is finite. Fix a large integer N and consider the sequence $a_1 = a_2 = \ldots = a_N = 1$, $a_{N+1} = a_{N+2} = \ldots = 0$. Then $||(a_n)_{n\geq 1}||_{\ell_1} = N$ and for any $m = 1, 2, \ldots$ we have

$$\sum_{n=1}^\infty \frac{a_n}{m+n} = \sum_{n=1}^N \frac{1}{m+n} \geq \ln\left(1 + \frac{N}{m+1}\right),$$

where we have used the elementary bound $\ln(1+1/k) \geq 1/(k+1)$ several times. Consequently,

$$\frac{\sum_{m=1}^\infty \Phi\left(\frac{1}{\pi}\sum_{n=1}^\infty \frac{a_n}{m+n}\right)}{||(a_n)_{n\geq 1}||_{\ell_1}} \geq \frac{1}{N}\sum_{m=1}^\infty \Phi\left(\frac{1}{\pi}\ln\left(1+\frac{N}{m+1}\right)\right),$$

and the latter expression is a Riemann sum for the integral $\int_0^\infty \Phi(\frac{1}{\pi}\ln(1+t^{-1}))\mathrm{d}t = C_\Phi$. This proves the sharpness of the estimate (1.3). □

Proof of $||T||_{L_1(0,\infty)\to L_{1,\infty}(0,\infty)} \leq \pi^{-1}$. As previously, with no loss of generality we may assume that f is nonnegative. By homogeneity, we will be done if we prove that

$$|\{x\in(0,\infty) : Tf(x) > 1\}| \leq \pi^{-1}\int_{\mathbb{R}^d} f(x)\mathrm{d}x.$$

Recall that we study the estimate in the case $d=1$ only. In this case, the martingales ξ and ζ are differentially equivalent. Furthermore, by Itô's formula, we have

$$A{*}f(x,y) = \mathcal{H}f(x+X_{\tau(y)}) - U_{\mathcal{H}f}(x,y).$$

Hence (2.6) gives

$$\mathbb{P}\big((-\mathcal{H}f(x+X_{\tau(y)}) + U_{\mathcal{H}f}(x,y))_+ 1_{\{f(x+X_{\tau(y)})=0\}} > 1\big) \leq \pi^{-1}\mathbb{E}f(x+X_{\tau(y)}),$$

which, in turn, implies

$$\mathbb{P}\big((-\mathcal{H}f(x+X_{\tau(y)}) + \inf_{z\in\mathbb{R}} U_{\mathcal{H}f}(z,y))_+ 1_{\{f(x+X_{\tau(y)})=0\}} > 1\big) \leq \pi^{-1}\mathbb{E}f(x+X_{\tau(y)}).$$

Integrating over $x\in\mathbb{R}^d$ and using Fubini's theorem, we obtain

$$\left|\left\{x\in\mathbb{R} : \big(-\mathcal{H}f(x) + \inf_{z\in\mathbb{R}} U_{\mathcal{H}f}(z,y)\big)_+ 1_{\{f(x)=0\}} > 1\right\}\right| \leq \pi^{-1}\int_{\mathbb{R}} f(x)\mathrm{d}x.$$

However, $\inf_{z\in\mathbb{R}} U_{\mathcal{H}f}(z,y)$ converges to 0 as $y\to\infty$ (see e.g. [18]) and hence Fatou's lemma implies

$$\left|\left\{x\in\mathbb{R} : \big(-\mathcal{H}f(x)\big)_+ 1_{\{f(x)=0\}} > 1\right\}\right| \leq \pi^{-1}\int_{\mathbb{R}} f(x)\mathrm{d}x.$$

This yields the claim, in the light of (2.3). □

Proof of $||T||_{L_1(0,\infty)\to L_{1,\infty}(0,\infty)} \geq \pi^{-1}$. As in the Φ-estimate, it is enough to construct an appropriate example for the discrete operator S. Fix large integers K, N and put $a_1 = a_2 = \ldots = a_N = K$, $a_{N+1} = a_{N+2} = \ldots = 0$. Clearly, we have $||(a_n)_{n\geq 1}||_{\ell_1} = KN$ and, arguing as above,

$$\sum_{n=1}^{\infty} \frac{a_n}{m+n} \geq K \ln\left(1 + \frac{N}{m+1}\right).$$

Therefore,

$$\#\left\{m : \frac{1}{\pi}\sum_{n=1}^{\infty} \frac{a_n}{m+n} > \frac{1}{\pi}\right\} \geq \#\left\{m : K\ln\left(1+\frac{N}{m+1}\right) > 1\right\} = \left\lfloor \frac{N}{e^{1/K}-1}\right\rfloor - 1$$

and hence

$$\frac{\#\left\{m : Sa(m) > \frac{1}{\pi}\right\}}{||(a_n)_{n\geq 1}||_{\ell_1}} \geq \frac{\lfloor N/(e^{1/K}-1)\rfloor - 1}{KN}.$$

If we let $N \to \infty$ and then $K \to \infty$, then the right-hand side converges to 1. In other words, given $\varepsilon > 0$, we can take sufficiently large K and N, such that the above ratio is larger than $1 - \varepsilon$. Hence the weak-type constant is at least $1/\pi$. □

Sharpness of (1.1) *and* (1.3) *for* $d \geq 2$. This will follow from the well-known classical transference arguments (see e.g. [22]). Our reasoning will follow the arguments from [27]. We will focus on the estimate (1.1), the inequality (1.3) can be handled similarly. Of course, it suffices to deal with the operator T_1 only. Suppose that for a fixed $1 < p < \infty$ and some c_p we have

$$\int_{\mathbb{R}^d} (-R_1 f(x))_+^p 1_{\{f(x)=0\}} \mathrm{d}x \leq c_p^p \int_{\mathbb{R}^d} f^p(x)\mathrm{d}x \tag{3.4}$$

for all $f \in L_p(\mathbb{R}^d)$ satisfying $f > 0$ on $\mathbb{R}^d_{1+}$ and $f = 0$ on $\mathbb{R}^d \backslash \mathbb{R}^d_{1+}$. Clearly, if we show that $c_p \geq \sin^{-1}(\pi/p)$, this will yield the desired lower bound for $||T_1||_{L_p(\mathbb{R}^d)\to L_p(\mathbb{R}^d)}$. For $t > 0$, define the dilation operator δ_t as follows: for any function $g : \mathbb{R}\times\mathbb{R}^{d-1} \to \mathbb{R}$, we let $\delta_t g(\xi,\zeta) = g(\xi, t\zeta)$. By (3.4), the operator $\mathfrak{T}_t := \delta_t^{-1} \circ R_1 \circ \delta_t$ satisfies

$$\begin{aligned}
\int_{\mathbb{R}^d} (-\mathfrak{T}_t f(x))_+^p 1_{\{f(x)=0\}}\mathrm{d}x &= t^{d-1}\int_{\mathbb{R}^d} (-R_1 \circ \delta_t f(x))_+^p 1_{\{\delta_t f(x)=0\}}\mathrm{d}x \\
&\leq t^{d-1} c_p^p \int_{\mathbb{R}^d} (\delta_t f(x))^p \mathrm{d}x \\
&= c_p^p \int_{\mathbb{R}^d} f^p(x)\mathrm{d}x,
\end{aligned} \tag{3.5}$$

provided f is as above (that is, $f \in L_p(\mathbb{R}^d)$, $f > 0$ on $\mathbb{R}^d_{1+}$ and $f = 0$ on $\mathbb{R}^d \setminus \mathbb{R}^d_{1+}$). Now suppose that additionally f belongs to $L_2(\mathbb{R}^d)$. It is not difficult to check that the Fourier transform $\mathcal{F}$ satisfies the identity $\mathcal{F} = t^{d-1}\delta_t \circ \mathcal{F} \circ \delta_t$ and hence the operator $\mathfrak{T}_t$ has the property that

$$\widehat{\mathfrak{T}_t f}(\xi,\zeta) = -i\frac{\xi}{(\xi^2 + t^2|\zeta|^2)^{1/2}}\widehat{f}(\xi,\zeta), \qquad (\xi,\zeta) \in \mathbb{R}\times\mathbb{R}^{d-1}.$$

By Lebesgue's dominated convergence theorem, we have

$$\lim_{t\to 0} \widehat{\mathfrak{T}_t f}(\xi,\zeta) = \widehat{\mathfrak{T}_0 f}(\xi,\zeta)$$

in $L_2(\mathbb{R}^d)$, where $\widehat{\mathfrak{T}_0 f}(\xi,\zeta) = -i\,\mathrm{sgn}\,(\xi)\widehat{f}$. Combining this with Plancherel's theorem, we conclude that there is a sequence $(t_n)_{n\geq 1}$ decreasing to 0 such that $\mathfrak{T}_{t_n} f$ converges to $\mathfrak{T}_0 f$ almost everywhere. Using Fatou's lemma and (3.5), we obtain

$$\int_{\mathbb{R}^d} (-\mathfrak{T}_0 f(x))_+^p 1_{\{f(x)=0\}} \mathrm{d}x \leq c_p^p \int_{\mathbb{R}^d} f^p(x)\mathrm{d}x. \tag{3.6}$$

Note that $\mathfrak{T}_t$ are bounded on $L_p(\mathbb{R}^d)$ for $1 < p < \infty$ (in fact, $||\mathfrak{T}_t||_p = ||R_1||_p$), so $\mathfrak{T}_0$ also has this property and thus the above estimate holds true without the assumption $f \in L_2(\mathbb{R}^d)$. Next, fix $\kappa_p < \sin^{-1}(\pi/p)$ and let us use the sharpness in the case $d = 1$, which we have already established above. There is a function $h \in L_p(\mathbb{R})$, which vanishes on $(-\infty, 0)$ and is strictly positive on $[0,\infty)$, satisfying

$$\int_{\mathbb{R}} (-\mathcal{H}h(x))_+^p 1_{\{h(x)=0\}} \mathrm{d}x > \kappa_p^p \int_{\mathbb{R}} h^p(x)\,\mathrm{d}x. \tag{3.7}$$

Indeed, it suffices to take $h : [0,\infty) \to [0,\infty)$ for which $||Th||_{L_p(0,\infty)}/||h||_{L_p(0,\infty)} > \kappa_p$, modify it slightly so that h is strictly positive on $[0,\infty)$, and, finally, extend it to the whole $\mathbb{R}$ by setting $h = 0$ on $(-\infty, 0)$. Pick an arbitrary function $g : \mathbb{R}^{d-1} \to [0,\infty)$ satisfying $g > 0$ almost everywhere and $||g||_{L_p(\mathbb{R}^{d-1})} = 1$, and define $f : \mathbb{R} \times \mathbb{R}^{d-1} \to \mathbb{R}$ by $f(\xi,\zeta) = h(\xi)g(\zeta)$. Then f is a p-integrable nonnegative function supported on $\mathbb{R}^d_{1+}$, strictly positive on this set. Furthermore, $\mathfrak{T}_0 f(\xi,\zeta) = \mathcal{H}h(\xi)g(\zeta)$, which is due to the identity

$$\widehat{\mathfrak{T}_0 f}(\xi,\zeta) = -i\mathrm{sgn}\,(\xi)\,\widehat{h}(\xi)\widehat{g}(\zeta).$$

Plug this into (3.6). Clearly, we have $1_{\{f(\xi,\zeta)=0\}} \geq 1_{\{h(\xi)=0\}}$, so we obtain

$$\int_{\mathbb{R}} (-\mathcal{H}h(\xi))_+^p 1_{\{h(\xi)=0\}} \mathrm{d}\xi \leq c_p^p \int_{\mathbb{R}} h^p(\xi)\mathrm{d}\xi.$$

This implies $c_p > \sin^{-1}(\pi/p)$ by virtue of (3.7) and the fact that κ_p was an arbitrary number smaller than $\sin^{-1}(\pi/p)$. The proof is complete.

4. An inequality for Riesz transforms

The results obtained in the preceding sections motivate the following related problem, which is interesting on its own. Namely, suppose that D is an open, connected subset of $\mathbb{R}^d$ and let f be a function supported on D. What can be said about the size of Riesz transform $R_j f$ restricted to the compliment D^c of the set D? For example, if the size is measured by means of L_p norms, what is the best constant c_p in the estimate

$$||R_j f||_{L_p(D^c)} \leq c_p ||f||_{L_p(D)}\,?$$

Note that if $d = 1$ and D is the positive halfline, this leads precisely to the problem of bounding the Hilbert operator T. What can be said for arbitrary d and an arbitrary domain D?

We will study this question for convex domains only. The precise statement can be found in the three theorems below. Recall the constant C_p, introduced in Theorem 2.2.

Theorem 4.1. *Suppose that D is an open and convex subset of $\mathbb{R}^d$ and f is a function supported on D. Then for any $1 < p < \infty$ we have*

$$||R_j f||_{L_p(D^c)} \leq C_p ||f||_{L_p(D)}. \tag{4.1}$$

We will also establish the following weak-type and Φ-inequalities. The first result concerns, as in the preceding sections, the one-dimensional case only.

Theorem 4.2. *Suppose that D is an open interval contained in $\mathbb{R}$ and f is a function supported on D. Then we have*

$$||\mathcal{H}f||_{L_{1,\infty}(D^c)} \leq \frac{2}{\pi}||f||_{L_1(D)}. \tag{4.2}$$

Our final result works in all dimensions. Recall the constant C_Φ, defined in (1.4).

Theorem 4.3. *Suppose that D is an open and convex subset of $\mathbb{R}^d$ and f is an integrable function supported on D, satisfying $||f||_{L_\infty(D)} \leq 1$. Suppose that $\Phi : [0,\infty) \to [0,\infty)$ is a convex, strictly increasing function of class C^1, satisfying $\Phi(0) = \Phi'(0+) = 0$. Then we have*

$$\int_{D^c} \Phi\big(|R_j f(x)|\big)\, dx \leq C_\Phi ||f||_{L_1(D)}, \tag{4.3}$$

Thus, we see that in the one-dimensional case, we encounter a very interesting phenomenon: the constants change when we pass from halflines to intervals (except for the L_p bound for $p \geq 2$).

There are several more or less informal comments which are worth stating here. The first obstacle we face is that Riesz transforms are not positive operators, and hence we cannot - at least, not immediately - restrict ourselves to the class of nonnegative functions f. Without this step, we would be led to estimates for martingales without the assumption of nonnegativity for the dominated process, which would increase the constants. To overcome this difficulty, we will study first the one-dimensional case, in which the extremal functions still must be of constant sign. To see this, fix an open interval (a,b) and consider an integrable function f supported on (a,b). Then, for x outside (a,b) we have

$$\mathcal{H}f(x) = \frac{1}{\pi}\int_a^b \frac{f(y)}{x-y}\mathrm{d}y.$$

So, if we pass from f to $|f|$, then $|\mathcal{H}f|$ does not decrease. Furthermore, we see that $\mathcal{H}f$ does not change the sign on $(-\infty,a)$ and does not change the sign on (b,∞), the signs on these two halflines being different. Roughly speaking, this implies that martingale counterpart will have to involve $|Y|$; this explains why we have developed the estimates (2.5), (2.7) and (2.9) in Section 2. Now, in order to proceed to higher dimensions, we will need to exploit the method of rotations; we have been unable to find a probabilistic argument here.

Proof of (4.1)*,* (4.2) *and* (4.3)*.* We will focus on the L_p estimate; the remaining bounds can be shown in a similar manner. We start with the one-dimensional case; as we already know, we may assume that f is nonnegative. Basing on the inequality (2.5) and the reasoning of the preceding section, we establish the bound

$$||\mathcal{H}f||_{L_p(D^c)} \leq ||\mathcal{H}f1_{\{f=0\}}||_{L_p(\mathbb{R})} \leq C_p||f||_{L_p(\mathbb{R})} = C_p||f||_{L_p(D)},$$

as desired. Now we turn to the case $d \geq 2$. Given a vector θ belonging to the unit sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$, we define the directional Hilbert transform $\mathcal{H}_\theta$ by

$$\mathcal{H}_\theta(f)(x) = \frac{1}{\pi}\text{p.v.}\int_{\mathbb{R}} \frac{f(x-t\theta)}{t}\mathrm{d}t,$$

where f is a sufficiently regular real-valued function on $\mathbb{R}^d$ (for instance, belonging to the Schwarz class). For example, if e_1 denotes the vector $(1,0,0,\ldots,0)\in\mathbb{R}^d$, then $\mathcal{H}_{e_1}$ is obtained by applying the Hilbert transform in the first variable followed by the identity operator in the remaining variables. Thus, by the reasoning we have just presented, the inequality (4.1) holds true for $\mathcal{H}_{e_1}$ as well; here we use the fact that the set D is convex. Next, if A is an arbitrary orthogonal matrix, we have

$$\mathcal{H}_{Ae_1}(f)(x)=\mathcal{H}_{e_1}(f\circ A)(A^{-1}x),$$

which implies that (4.1) holds true if we replace R_j with $\mathcal{H}_\theta$ with an arbitrary θ (again, by the convexity of D). Repeating the classical arguments of Iwaniec and Martin [22] (a convenient reference is also Grafakos [16]), we get that

$$R_jf(x)=||\theta_j||^{-1}_{L_1(\mathbb{S}^{d-1})}\int_{\mathbb{S}^{d-1}}\theta_j\mathcal{H}_\theta(f)(x)\mathrm{d}\theta.$$

Consequently, if f is a function supported on D, then

$$\int_{D^c}|R_jf(x)|^p\mathrm{d}x\leq ||\theta_j||^{-1}_{L_1(\mathbb{S}^{d-1})}\int_{\mathbb{S}^{d-1}}|\theta_j|\int_{D^c}|\mathcal{H}_\theta(f)(x)|^p\mathrm{d}x\mathrm{d}\theta\leq C_p^p||f||^p_{L_p(D)},$$

which is the claim. □

We turn our attention to the sharpness of the above estimates. It is enough to prove the optimality of the constants in the case $d=1$ only. The passage to higher dimensions in (4.1) and (4.3) follows from a straightforward modification of the transference argument presented in the preceding section. We start with the weak-type and Φ-estimates.

Sharpness of (4.2) *and* (4.3). Let $D=[-1,1]$ and consider the function $f=\chi_{[-1,1]}$. Then $\mathcal{H}f(x)=\frac{1}{\pi}\log\left|\frac{x+1}{x-1}\right|$ and, for a given $\lambda>0$,

$$\begin{aligned}\frac{\lambda|\{x\in D^c:|\mathcal{H}f(x)|>\lambda\}|}{||f||_{L_1(D)}}&=\lambda|\{x\geq 1:\mathcal{H}f(x)>\lambda\}|\\&=\lambda|\{x\geq 1: x\leq (e^{\pi\lambda}+1)/(e^{\pi\lambda}-1)\}|\\&=\frac{2\lambda}{e^{\pi\lambda}-1}\xrightarrow{\lambda\to 0}\frac{2}{\pi},\end{aligned}$$

which yields the sharpness of (4.2). The inequality (4.3) is handled with the use of the same extremal function f. We have $||f||_{L_1(D)}=2$ and hence

$$\frac{\int_{D^c}\Phi\big(|\mathcal{H}f(x)|\big)\mathrm{d}x}{||f||_{L_1(D)}}=\int_1^\infty\Phi\big(\mathcal{H}f(x)\big)\mathrm{d}x=\int_1^\infty\Phi\left(\frac{1}{\pi}\log\frac{x+1}{x-1}\right)\mathrm{d}x=C_\Phi,$$

which is the desired claim. □

Sharpness of (4.1). When $p\geq 2$, the optimality of the constant C_p follows from (1.1); so, let us assume that $1<p<2$. Here the reasoning will be slightly more complicated. Let us look back at the almost-extremal example of Section 2, which was used to prove the sharpness of (2.5). So, fix $c>1$, let $A_c=([0,\infty)\times\mathbb{R})\setminus([c,\infty)\times\{0\})$ and let (X,Y) be a planar Brownian motion, started at $(1,0)$ and stopped upon leaving A_c. There exists a conformal mapping F_c which maps A_c onto K, the unit disc of $\mathbb{C}$, which

(i) sends $(1,0)$ to the origin,

(ii) sends the halfline $[c,\infty)\times\{0\}$ to a symmetric arc $\mathcal{A}_\alpha=\{e^{i\theta}:|\theta|\leq\alpha\}$, for some α depending on c.

Set $u_c=\mathfrak{Re}F_c^{-1}$ and $v_c=\mathfrak{Im}F_c^{-1}$. Then u_c and v_c are conjugate harmonic functions on K, u_c is nonnegative and supported on $\mathcal{A}_\alpha$. Furthermore, since the distribution of (X_∞,Y_∞) is the harmonic measure on ∂A_c with respect to $(1,0)$, we get

$$\frac{||v_c 1^c_{\mathcal{A}_\alpha}||_{L_p(\mathbb{T})}}{||u_c 1_{\mathcal{A}_\alpha}||_{L_p(\mathbb{T})}}=\frac{||v_c 1^c_{\mathcal{A}_\alpha}||_{L_p(\mathbb{T})}}{||u_c||_{L_p(\mathbb{T})}}=\frac{||Y_\infty 1_{\{X_\infty=0\}}||_p}{||X_\infty||_p}. \tag{4.4}$$

However, we have $v_c(0)=0$ and hence v_c is the periodic Hilbert transform of u_c. That is, if we identify the unit circle $\mathbb{T}$ with the interval $(-\pi,\pi]$, then

$$v_c(x)=\frac{1}{2\pi}\text{p.v.}\int_{-\pi}^{\pi}u_c(t)\cot\frac{x-t}{2}\mathrm{d}t.$$

But the function u_c is nonnegative and we have the elementary estimate $\cot u\leq 1/u$ for $|u|\leq\pi/2$. Consequently, if we take $x\notin\mathcal{A}_\alpha$ (more formally, $x\notin[-\alpha,\alpha]$) and we extend u_c to the whole $\mathbb{R}$ by setting $u_c=0$ outside $(-\pi,\pi]$, then

$$|\mathcal{H}u_c(x)|=\left|\frac{1}{2\pi}\int_{\mathbb{R}}\frac{2u_c(t)}{x-t}\mathrm{d}t\right|\geq\left|\frac{1}{2\pi}\int_{\mathbb{R}}u_c(t)\cot\left(\frac{x-t}{2}\right)\mathrm{d}t\right|=|v_c(x)|.$$

Combining this with (4.4), we obtain

$$\frac{||\mathcal{H}u_c||_{L_p(D^c)}}{||u_c||_{L_p(D)}}\geq\frac{||Y_\infty 1_{\{X_\infty=0\}}||_p}{||X_\infty||_p},$$

where $D=[-\alpha,\alpha]$. However, if c is sufficiently large, then the right-hand side can be made arbitrarily close to C_p. This shows that the constant C_p is indeed the best possible in (4.1). □

Acknowledgments

The author would like to thank an anonymous referee for the careful reading of the first version of the manuscript, for many helpful comments and remarks.

References

[1] R. Bañuelos, A. Bielaszewski and K. Bogdan, *Fourier multipliers for non-symmetric Lévy processes*, Marcinkiewicz centenary volume, Banach Center Publ. **95** (2011), 9-25, Polish Acad. Sci. Inst. Math., Warsaw.

[2] R. Bañuelos and P. J. Méndez-Hernandez, *Space-time Brownian motion and the Beurling-Ahlfors transform*, Indiana Univ. Math. J. **52** (2003), no. 4, 981–990.

[3] R. Banuelos, and G. Wang, *Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms*, Duke Math. J. **80** (1995), no. 3, 575–600.

[4] R. Bañuelos and G. Wang, *Orthogonal martingales under differential subordination and application to Riesz transforms*, Illinois J. Math. 40 (1996), 678–691.

[5] R. Bañuelos and G. Wang, *Davis's inequality for orthogonal martingales under differential subordination*, Michigan Math. J. **47** (2000), 109–124.

[6] A. Borichev, P. Janakiraman and A. Volberg, *Subordination by conformal martingales in L_p and zeros of Laguerre polynomials*, Duke Math. J. **162** (2013), no. 5, 889–924.

[7] A. Borichev, P. Janakiraman and A. Volberg, *On Burkholder function for orthogonal martingales and zeros of Legendre polynomials*, Amer. J. Math. **135** (2013), no. 1, 207–236.

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

[9] D. L. Burkholder, *A Sharp and Strict L_p-Inequality for Stochastic Integrals* , Ann. Probab. **15** (1987), 268–273.

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

[11] D. L. Burkholder, R. F. Gundy and M. L. Silverstein, *A maximal function characterization of the class* H^p, Trans. Amer. Math. Soc. **157** (1971), 137-153.

[12] C. Dellacherie and P. A. Meyer, *Probabilities and potential B*, North-Holland, Amsterdam, 1982.

[13] S. Donaldson and D. Sullivan, *Quasiconformal 4-manifolds*, Acta Math. **163** (1989), 181-252.

[14] J. L. Doob, Stochastic processes, Wiley, New York, 1953.

[15] E. Geiss, S. Mongomery-Smith, E. Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. **362** (2010), 553-575.

[16] L. Grafakos, *Classical Fourier analysis*, Second edition. Graduate Texts in Mathematics, 249. Springer, New York, 2008.

[17] R. F. Gundy and M. Silverstein, *On a probabilistic interpretation for Riesz transforms*, Lecture Notes in Math. 923, Springer, Berlin and New York, 1982.

[18] R.F. Gundy and N. Th. Varopoulos, *Les transformations de Riesz et les intgrales stochastiques*, C. R. Acad. Sci. Paris Sr. A-B 289 (1979), A13–A16.

[19] G. H. Hardy, *Note on a theorem of Hilbert concerning series of positive terms*, Proc. London Math. Soc. **23** (1925), Records of Proc. XLV–XLVI.

[20] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd edn. (Cambridge University Press, 1952).

[21] T. Iwaniec and G. Martin, *Quasiregular mappings in even dimensions*, Acta Math. **170** (1993), 29-81.

[22] T. Iwaniec and G. Martin, *Riesz transforms and related singular integrals*, J. Reine Angew. Math. **473** (1996), 25-57.

[23] P. Janakiraman, *Best weak-type* (p,p) *constants,* $1 < p < 2$*, for orthogonal harmonic functions and martingales*, Illinois J. Math. **48** (2004), 909–921.

[24] K. Oleszkiewicz, *An Elementary Proof of Hilbert's Inequality*, Amer. Math. Monthly **100** (1993), 276–280.

[25] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72**, Birkhäuser, 2012.

[26] A. Osękowski, *Sharp logarithmic inequalities for Riesz transforms*, J. Funct. Anal. **263** (2012), 89-108.

[27] A. Osękowski, *Logarithmic inequalities for Fourier multipliers*, Math. Z. **274** (2013), no. 1-2, 515–530.

[28] G. Peskir and A. Shiryaev, *Optimal stopping and free boundary problems*, Lect. Notes in Math. ETH, Zurich, 2005.

[29] S. K. Pichorides, *On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov*, Studia Math. **44** (1972), 165–179.

[30] M. Riesz, *Les fonctions conjugées et ler séries de Fourier*, C. R. Acad. Paris, **178** (1924), 1464–1467.

[31] I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, Journal f. Math. **140** (1911), 1–28.

[32] E. M. Stein, *Singular integrals and Differentiability Properties of Functions,* Princeton University Press, Princeton, 1970.

[33] D. C. Ullrich, *A simple elementary proof of Hilbert's inequality*, Amer. Math. Monthly **120** (2013), no. 2, 161–164.

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

[35] A. Zygmund, *Sur les fonctions conjugées*, Fund. Math. **13** (1929), 284–303.

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

E-mail address: ados@mimuw.edu.pl