Rend. Lincei Mat. Appl. (Online first)
DOI 10.4171/RLM/1023

Published by EMS Press

Mathematical Analysis. – *Two-weight dyadic Hardy inequalities*, by NICOLA ARCOZZI, NIKOLAOS CHALMOUKIS, MATTEO LEVI and PAVEL MOZOLYAKO, communicated on 12 May 2023.

ABSTRACT. – We present various results concerning the two-weight Hardy inequality on infinite trees. Our main aim is to survey known characterizations (and proofs) for trace measures, as well as to provide some new ones. Also for some of the known characterizations we provide here new proofs. In particular, we obtain a new characterization in terms of a reverse Hölder inequality for trace measures, and one based on the well-known Muckenhoupt–Wheeden–Wolff inequality, of which we here give a new probabilistic proof. We provide a new direct proof for the so-called isocapacitary characterization and a new simple proof, based on a monotonicity argument, for the so-called mass-energy characterization. Furthermore, we introduce a conformally invariant version of the two-weight Hardy inequality, characterize the compactness of the Hardy operator, provide a list of open problems, and suggest some possible lines of future research.

KEYWORDS. – Potential theory on trees, Carleson measures, Hardy operator, Dirichlet space, Besov spaces, maximal function, Bellman function, Muckenhoupt–Wheeden inequality, Wolff's inequality, reverse Hölder inequality.

2020 MATHEMATICS SUBJECT CLASSIFICATION. – Primary 31E05; Secondary 05C05, 05C63, 42B25, 42B35, 31A15, 30C85.

CONTENTS

1. Introduction

1.1. A (very) brief history of Hardy's inequality

Hardy's inequality [40] states that

$$\int_0^\infty \left(\frac{1}{x}\int_0^x f(y)dy\right)^p dx \le (p^*)^p \int_0^\infty f(x)^p dx,$$

for every positive measurable function f and every $p > 1$, where $p^* := p/(p-1)$, and the constant $(p^*)^p$ is optimal. Hardy himself, motivated by the goal of giving a simpler proof of "Hilbert's inequality for double series" [39], was actually primarily interested in the discrete analogue of the above inequality,

$$\text{(Hardy)} \qquad \sum_{n=1}^\infty \left(\frac{a_1+\cdots+a_n}{n}\right)^p \le (p^*)^p \sum_{n=1}^\infty a_n^p,$$

where a_n are positive real numbers. Indeed, for $p = 2$, the discrete version was the first one to be proved by Hardy in his earlier paper [38]. The discrete inequality for general p can either be deduced by the corresponding continuous one or proved directly, as communicated to Hardy by Landau [40]. Much more information in the fascinating history of the development of Hardy's original inequality can be found in the survey paper [45].

Despite its original purpose, it soon became clear that Hardy's inequality and its extensions lie in the heart of the developments in the broad area of harmonic analysis throughout the 20th century, up until modern days. From a general viewpoint the reason is that Hardy's inequality *is the prototype of a (weighted) norm inequality for an integration (averaging) operator between L^p spaces*. Averaging operators together with maximal and singular operators are the pillars of harmonic analysis. Under this light, a vast number of theorems can be considered as Hardy-type inequalities. Hence, it comes with no surprise that the original Hardy inequality was later generalized in many different directions.

Tomaselli in [68] and Talenti in [66] made the first steps towards a weighted Hardy inequality, i.e., an inequality of the form

$$\int_0^\infty \left(\int_0^x f(y)dy\right)^p U(x)dx \le C(p,U,V)\int_0^\infty f(x)^p V(x)dx, \tag{1.1}$$

in which V, U are positive measurable weights. Another way to state the same inequality is to say that the *Hardy operator*, which maps f to its primitive, is bounded from $L^p(\mathbb{R}_+, V(x)dx)$ to $L^p(\mathbb{R}_+, U(x)dx)$. The first to give a complete characterization of the weights such that the weighted Hardy inequality holds was Muckenhoupt in [54]. It is worth taking a closer look to Muckenhoupt's condition.

THEOREM 1 (Muckenhoupt). *Let* $1 \le p \le \infty$. *There exists a constant* C *such that* (1.1) *is true if and only if*

$$B := \sup_{r>0} \left(\int_r^\infty U(x)^p dx \right) \left(\int_0^r V(x)^{-p^*} dx \right)^{p-1} < \infty.$$

Furthermore, if C *is the smallest constant such that the inequality holds, then* $B \le C \le p(p^*)^{p-1} B$.

This theorem completes the picture of the weighted Hardy inequality on $\mathbb{R}_+$. In the meanwhile, several other extensions of Hardy's inequality to different spaces were considered: to higher dimensions, to different metric spaces, to fractional integral operators [55, 63, 65]. Also the weighted problem with different exponents, namely, the boundedness of the Hardy operator from $L^p(\mathbb{R}_+, V(x)dx)$ to $L^q(\mathbb{R}_+, U(x)dx)$, with p and q not necessarily coinciding, was extensively studied: we mention Bradley [22] for the case $1 < p \le q \le \infty$, Maz'ya [50] and Sawyer [60] for the case $p > q$.

Let us mention that there exists also a different stream of research which is devoted to finding weights which are optimal, in an appropriate sense, for weighted Hardy's inequalities, both in the continuous setting [18, 34] and in the setting of graphs [19, 43].

1.2. The two-weight Hardy inequality on trees

In the present paper, we focus on a generalized version of the discrete Hardy inequality (Hardy), the two-weight Hardy inequality on trees. A particular case of this inequality is the so-called two-weight dyadic Hardy inequality, presented in Section A. In that section, we will also make clear the intuitive fact that the classical inequality (Hardy) is a special case of the inequality on trees, and we will provide some examples of application of the dyadic inequality in complex analysis. In the paper, however, we are able to work in the generality of the two-weight Hardy inequality on trees presented in this section. Before stating such an inequality, we need to introduce some pieces of notation.

A *tree* $T = (V, E)$ is a simple connected graph with no cycles, where V is a finite or countable set of vertices and E is the set of edges. The number of edges sharing each vertex is assumed to be finite, but we do not require further restrictions. We will consistently use Greek letters for edges and Latin letters for vertices and boundary points, to be defined below. Henceforth, we will identify the tree T with its vertices V. We fix arbitrarily a *root* vertex o and we assume that there exists a *pre-root* vertex o^* which is connected to o and to no other vertex. We denote by ω the edge connecting o and o^* and we call it the *root edge*.

A fundamental property of trees is that for any couple of vertices $x, y \in T$ there exists a unique geodesic connecting the two, that is, a unique minimal sequence of pairwise connected vertices containing x and y. We write $[x, y]$, or equivalently $[y, x]$, for the (unique) set of edges connecting pairs of points in the geodesic joining x and y. The *confluent* of x and y is the vertex $x \wedge y$ such that $[o^*, x \wedge y] = [o^*, x] \cap [o^*, y]$. The edge-counting distance on T is given by $d(x, y) = \sharp[x, y]$. We use the same symbol for the vertex-counting distance on E, defined in the obvious way, and abbreviate $d(\alpha)$ for $d(\alpha, \omega)$. If we assign to each edge a weight $\sigma(\alpha) > 0$, we can define the associated distance d_σ by adding the weights on the edges of paths, instead of counting edges. Due to the elementary topology of a tree, the new metric has the same geodesics.

If α is an edge, we denote by $b(\alpha)$ its endpoint vertex which is closest to o^* and by $e(\alpha)$ the furthest. Observe that for any vertex x there exists a unique edge α with $e(\alpha) = x$, while there are possibly many edges with $b(\alpha) = x$. For each vertex x, we define the *predecessor* of x as its unique neighbor vertex $p(x)$ which is closer than x to o^*, and we denote by $s(x)$ the set of the remaining neighbors, the *children* on x. Predecessors and children may be defined also for edges in the very same way.

The choice of a root induces a partial order on T and E. For edges, we write $\alpha \subseteq \beta$ if $\beta \in [o^*, e(\alpha)]$, and for vertices we write $x \subseteq y$ if $[o^*, y] \subseteq [o^*, x]$. We will also write $x \subseteq \alpha \subseteq y$, comparing vertices and edges, with the obvious meaning. The notation $\subseteq$ is a reminder that trees often come from dyadic decompositions of metric spaces (see Section A).

The *boundary* of the rooted tree (T, ω), denoted by ∂T, is the set of the maximal geodesics emanating from o^*[1].

In the infinite tree case, we always assume that all maximal geodesics starting at o^* are infinite, so that $\partial T \cap T = \emptyset$. We set $\overline{T} := T \cup \partial T$.

The most important geometric objects we deal with are the *successor sets* $S(\alpha)$ of an edge α, $S(\alpha) = \{x \in \overline{T} : [o^*, x] \ni \alpha\}$. We also write $S(x)$ for $S(\alpha)$ when $x = e(\alpha)$. We remark that $\overline{T}$ is compact with respect to the topology generated by the family $\{S(\alpha)\}_{\alpha \in E}$ of successor sets and singletons $\{x\}_{x \in T}$, from which $\overline{T}$ is often referred to as the *standard compactification* of T.

Figure 1 below should help the reader to visualize and summarize some of the fundamental notation just introduced.

We are now ready to introduce our main object of study: the two-weight Hardy inequality on trees. Define the *Hardy operator* $\mathcal{I}$ as the operator mapping a function $\varphi : E \to \mathbb{R}_+$ to

$$\mathcal{I}\varphi(x) = \sum_{\alpha \in [o^*, x]} \varphi(\alpha), \quad x \in \overline{T},$$

(1) If a maximal geodesic is of finite length, i.e., it ends in a *leaf* vertex x, we identify such a geodesic with x, which is then a boundary point for T.

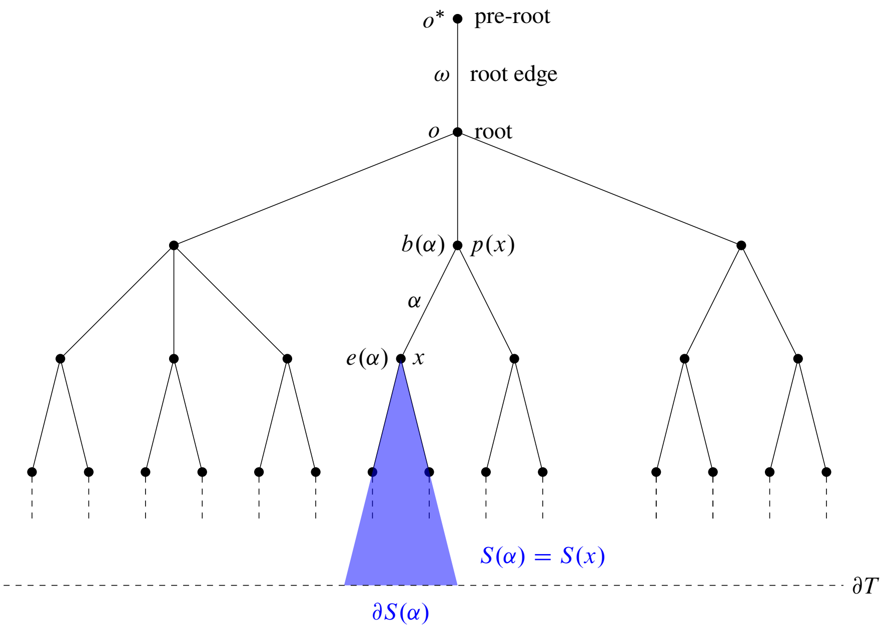

FIGURE 1. The infinite rooted tree.

provided that the sum converges. Let $\pi : E \to \mathbb{R}_+$ be some fixed edge weight, let $\mu \geq 0$ be a Borel measure on $\overline{T}$, and write $\ell^p(\pi) = \ell^p(E, \pi)$ and $L^p(\mu) = L^p(\overline{T}, \mu)$. The main problem which is discussed here is whether the following *two-weight Hardy inequality on T* holds:

$$\text{(H)} \qquad \int_{\overline{T}} \mathcal{I}\varphi(x)^p d\mu(x) \leq [\mu] \sum_{\alpha \in E} \varphi(\alpha)^p \pi(\alpha),$$

i.e., if $\mathcal{I} : \ell^p(\pi) \to L^p(\mu)$ is bounded, and what can be said about the constant $[\mu] = \|\mathcal{I}\|^p_{\ell^p(\pi)\to L^p(\mu)}$.

Any μ satisfying (H) (for some fixed π, p) with finite constant $[\mu]$ will be called a *trace measure*, and $[\mu]$ is the *Carleson measure norm*, or *trace norm*, of μ. We always consider π as a fixed, geometric object so that $[\mu]$ depends on π and p as well. Also, the quantity $[\cdot]$ is chosen so to be sublinear functional of μ:

$$[t\mu] = t[\mu] \text{ if } t \geq 0 \quad \text{and} \quad [\mu + \nu] \leq [\mu] + [\nu].$$

The inequality (H) has a natural corresponding dual inequality, which we will be of fundamental use in the sequel. For $\psi : V \to \mathbb{R}$, we set

$$\mathcal{I}^*_\mu \psi(\alpha) = \mathcal{I}^*(\psi d\mu)(\alpha) := \int_{S(\alpha)} \psi d\mu \quad \alpha \in E.$$

By obvious duality,

$$\langle \mathcal{I}\varphi, \psi\rangle_{L^2(\mu)} = \langle \varphi, \mathcal{I}^*(\psi d\mu)\rangle_{\ell^2}.$$

Thus, (H) is equivalent to

$$\text{(H*)} \qquad \sum_{\alpha} \pi(\alpha)^{1-p^*}\left(\mathcal{I}^*_\mu \psi(\alpha)\right)^{p^*} \le [\mu]^{p^*-1}\int_{\overline{T}} \psi^{p^*}\,d\mu,$$

which we will refer to as the *dual Hardy inequality*.

We reserve a specific symbol for a particular edge weight which will show up often in the paper; we write $|\alpha|$ for the unique edge weight satisfying the recursive formula

$$|\alpha| = \left|p(\alpha)\right|/q\big(e(\alpha)\big),$$

normalized such that $|\omega| = 1$. On the homogeneous tree where each vertex has $q+1$ neighbors, $|\alpha| = q^{-d(\alpha)}$. Observe that, if the tree arises from a dyadic decomposition of the unit interval (see Section A), then $|\alpha|$ is the Lebesgue measure of the corresponding subinterval in the decomposition. In full generality, we will then call Lebesgue measure on the boundary of T the measure dx defined by

$$\int_{\partial S(\alpha)} dx = |\alpha|, \quad \alpha \in E.$$

Characterizations of the triples (p, π, μ) for which (H) holds have been known since more than twenty years [6, 10, 26, 60]. The main goal of this expository paper is to survey old characterizations together with their proofs, provide some new proofs, and also present some completely new characterizations. We point out that even the known proofs are here adapted to work in our general framework, since they have all been originally given for the dyadic homogeneous tree only.

Our motivations are manifold. First, it is interesting and instructive to see the diverse machinery which can be employed in the solution of the problem. Second, we will see that the many conditions characterizing $\mu's$ for which the Hardy inequality holds are rather different one from the other, and their equivalence is a collection of interesting mathematical facts by itself. Third, as research moves to uncharted territories, such as multi-parameter dyadic Hardy inequalities, it is useful to have a place where different techniques are surveyed in a unified framework. We will mention some more recent areas of investigation, providing some results and mentioning a number of open problems.

1.3. Description of contents

We proceed now to a detailed description of the main theorems of this paper. The main goal of Sections 2 and 3 is to prove the following theorem.

THEOREM 2. *Let T be a locally finite connected tree, $1 < p < \infty$, and π a nonnegative weight on the edges. Then, for a positive Borel measure μ on $\overline{T}$ the following are equivalent.*

(i) *The two-weight Hardy inequality* (H) *holds with best constant $[\mu] < \infty$.*

(ii) *The* mass-energy condition *holds:*

$$\text{(ME)} \qquad \sup_{\alpha\in E} \mu\big(S(\alpha)\big)^{-1} \sum_{\beta\subseteq\alpha} \pi(\beta)^{1-p^*} \mu\big(S(\beta)\big)^{p^*} := [\![\mu]\!]^{p^*-1} < \infty.$$

(iii) *The isocapacitary condition holds:*

$$\text{(ISO)} \qquad \sup_{\alpha_1,\dots,\alpha_n\in E} \operatorname{Cap}_{p,\pi}\Big(\bigcup_{i=1}^{n} S(\alpha_i)\Big)^{-1} \sum_{i=1}^{n} \mu\big(S(\alpha_i)\big) := [\![\mu]\!]_c < \infty,$$

where the capacity $\operatorname{Cap}_{p,\pi}$ *is the one defined in Section 2.1.*

Furthermore, the following inequalities hold:

$$[\![\mu]\!]_c \le [\mu] \le 2^p [\![\mu]\!]_c, \quad [\![\mu]\!] \le [\mu] \le p^{p^*} [\![\mu]\!].$$

As we will show later, it is easy to prove that (ME) and (ISO) imply (H); the main issue in Theorem 2 is to prove the reverse implications. Section 2 is dedicated to developing a potential theory on the tree and proving the isocapacitary characterization, i.e., the equivalence of (i) and (iii) in Theorem 2. This equivalence was first proved in [13], though in an indirect way, passing through the mass-energy condition. Here we give a new direct proof that (iii) implies (i), which builds on ideas developed by Maz'ya in the continuous setting. The main tool is a *strong capacitary inequality* [1,51], which in tree language takes the form

$$\text{(Cap)} \qquad \sum_{k=-\infty}^{+\infty} 2^{pk} \operatorname{Cap}_{p,\pi}\big(\{x : \mathcal{I}\varphi(x) > 2^k\}\big) \le \frac{2^p}{2^p-1} \|\varphi\|^p_{\ell^p(E,\pi)}$$

and has an elementary proof.

In Section 3, we turn to the *mass-energy condition* introduced in [10]. We will give three proofs of its equivalence with (H): one based on maximal functions (originally proved for the homogeneous tree in [11]), one relying on a simple monotonicity argument, which is new and the simplest available at the moment, but only works for $p = 2$, and a very recent one using a Bellman function argument (originally proved for the homogeneous tree in [6, 26]).

The advantage of the isocapacitary condition (ISO) over the mass-energy condition (ME) is that the measure μ appears on the left-hand side only. It is obvious from it, for instance, that if $\nu \le \mu$, then $[\![\nu]\!]_c \le [\![\mu]\!]_c$. On the other hand, the mass-energy condition only has to be verified on single intervals, and not on arbitrary unions of them.

In some simple and particularly common cases of weights and trees, some further characterizations for trace measures can be proved to hold, lengthening the list of equivalent conditions summed up in Theorem 2. Sections 4 and 5 are dedicated to two different such additional characterizations.

More precisely, in Section 4, we restrict our attention to the case $\pi = 1$ and $p = 2$, and prove that in this case the Hardy inequality is equivalent to a one-parameter family of conditions. We prove the following theorem.

THEOREM 3. *The Hardy inequality*

$$\int_{\overline{T}} \mathcal{I}\varphi(x)^2 d\mu(x) \leq [\mu] \sum_{\alpha\in E} \varphi(\alpha)^2$$

is equivalent to

$$\text{(s-Testing)} \qquad \sup_{\alpha\in E} \fint_{S(\alpha)} \left(\int_{S(\alpha)} d(x\wedge y) d\mu(x) \right)^s d\mu(y) := [\![\mu]\!]_s < \infty,$$

for some (equivalently, for every) $s \geq 1$.

It can be readily verified that the s-testing condition is stronger than the mass-energy condition. The aim of the section is to prove that in fact they are all equivalent to the mass-energy condition which, in a sense, amounts to say that Carleson measures satisfy a reverse Hölder inequality; see Theorem 10. The results in this section are new and the techniques employed are partially inspired by the work of Tchoundja [67].

In Section 5, we provide another characterization of trace measures which holds (for any p) for a family of edge weights (depending on p) on homogeneous trees, and it can be generalized to trees having Ahlfors regular boundary [15, Section 3]. More precisely, we prove the following.

THEOREM 4. *Let T be a homogeneous tree, $1 < p < \infty$, $0 < s < 1$, and $\pi(\alpha) = |\alpha|^{\frac{1-p^* s}{1-p^*}}$. Then, the following conditions are both equivalent to the Hardy inequality* (H)*:*

(i) $\sup_{\alpha\in E} \mu(S(\alpha))^{-1} \int_{\partial T_\alpha} (\sum_{\beta\supset x} \frac{\mu(S(\beta))}{|\beta|^s})^{p^*} dx < \infty;$

(ii) $\sup_{\alpha\in E} \mu(S(\alpha))^{-1} \int_{\partial T} (\sup_{\beta\supset x} \frac{\mu(S(\beta))}{|\beta|^s})^{p^*} dx < \infty.$

Such a characterization is an immediate consequence of the Muckenhoupt–Wheeden inequality, which in this case reads as follows: for any $1 < p < \infty$, $0 < s < 1$ and for any measure μ on a homogeneous tree T,

$$\int_{\partial T} \left(\sum_{\alpha\supset x} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s} \right)^{p^*} dx \approx \sum_{\alpha} \frac{\mu\big(S(\alpha)\big)^{p^*}}{|\alpha|^{p^* s - 1}} \approx \int_{\partial T} \left(\sup_{\alpha\supset x} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s} \right)^{p^*} dx.$$

Indeed, it is easy to recognize in the middle term the form that the sum appearing in the mass-energy condition gets for the particular choice of

$$\pi(\alpha) = |\alpha|^{\frac{1-p^* s}{1-p^*}}.$$

Hence, for this family of weights on homogeneous trees we have an alternative characterization of Carleson measures. The above choice of π is of particular interest because of the connection with the theory of Bessel potentials in $\mathbb{R}^n$; see Section B. The inequality has also other interesting applications in potential theory; see for example [4].

Besides providing a different characterization of Carleson measures, the story of this inequality is itself interesting. That the term on the left is comparable to that on the right was proved in a non-dyadic language by Muckenhoupt and Wheeden in [55, Theorem 1]. They attribute the idea of the proof, which is a textbook example of a good lambda inequality, to Coifman and Fefferman [33]. Unaware of it, Wolff gave a wholly different proof [41, Theorem 1] that the term in the center is comparable to the term on the left. Independently of this, the first author, Rochberg, and Sawyer gave a different proof that the central term is bounded by the one on the right [10]. In this section, we will give a new proof of the "Wolff inequality" based on a probabilistic argument. The new proof has the advantage that it extends more easily to different settings. We will return on that in future work.

In Section 6, we discuss a conformally invariant version of (H) on the homogeneous tree. The left-hand side in (H) evidently depends on the arbitrary choice of a root in our tree. To remedy this, we propose the modified inequality

$$\int_{\overline{T}} \left| \mathcal{I} f(x) - \frac{1}{\mu(\overline{T})} \int_{\overline{T}} f d\mu \right|^2 d\mu(x) \leq [\mu]_{\mathrm{inv}} \sum_{\alpha} f(\alpha)^2.$$

We prove that it is "invariant"; i.e., if Ψ is a tree automorphism, then $[\Psi_* \mu]_{\mathrm{inv}} = [\mu]_{\mathrm{inv}}$, and that it is surprisingly equivalent to (H). We also provide a sharp estimate of the quantity $[\mu]_{\mathrm{inv}}$ in terms of capacity (Theorem 20). All results in this section appear here for the first time.

In Section 7, we collect some miscellaneous results on the topic. They are all new. First, we prove that a "vanishing" version of the mass-energy condition characterizes the compactness of the Hardy operator (Theorem 23). With a similar reasoning, one can obtain an equivalent isocapacitary-type vanishing condition (Theorem 24). We then discuss another very natural and easily determined necessary condition for (H) to hold, the *simple box-type condition*,

$$\text{(SB)} \qquad \sup_{\alpha \in E} \mu(S(\alpha)) \Big(\sum_{\beta \supseteq \alpha} \pi(\beta)^{1-p^*} \Big)^{p-1} := [\![\mu]\!]_{sc} < +\infty.$$

In contrast to the mass-energy and the isocapacitary conditions, however, for the many relevant weights and trees (SB) is not sufficient; see Example 25. We end the section by providing two easy examples of trees where the potential theory degenerates in two opposite ways.

In Section 8, we enlarge our horizon, including some dyadic structures variously related to the Hardy operator or to its applications. Most of this territory is uncharted, only a few results are known, investigation is still in its infancy, and there is a high potential for applications to harmonic analysis, holomorphic function theory, and more. This section is essentially a description of the few results that are known in the literature. We omit most of the proofs. In Section 8.1, we introduce the viewpoint of reproducing kernel Hilbert spaces, which provides a unified view of the preceding inequalities and is instrumental to state the problem of Hardy-type inequalities for quotient structures in Section 8.2. In Section 8.3, we briefly account on the topic of Hardy inequalities on poly-trees. This is a new area of research where very little is known. Recently, it has attracted a lot of interest because of the applications to function theory in the poly-disc.

We end the paper with an appendix, where we include the discussion on the model case of the purely dyadic Hardy inequality (Section A) and a comparison of the potential theory we use in the paper with that arising from Bessel's potentials (Section B).

In the text, we mention a number of open problems that we think are interesting and deserve further attention.

2. Potential theory on trees and the isocapacitary characterization

This section is dedicated to prove the equivalence of (i) and (iii) in Theorem 2. Section 2.1 introduces the potential theory which we need to define a p-capacity on the tree, while the actual proof is given in Section 2.2.

2.1. *Potential theory*

We define a potential theory following Adams and Hedberg's axiomatic approach [2]. Other approaches are also possible; see for example [62]. Consider the compact Hausdorff space $\overline{T}$ and make E into a measure space by endowing it with the measure associated to a weight $\sigma : E \to \mathbb{R}_+$. We introduce the kernel $k : \overline{T} \times E \to \mathbb{R}_+$, given by the characteristic function $k(x,\alpha) = \chi_{\{\alpha \supset x\}}(x,\alpha)$. Observe that $k(\cdot,\alpha)$ is continuous on ∂T, since $\partial S(\alpha)$ is open.

Given a function $\varphi : E \to \mathbb{R}_+$, we define the *potential* of φ, $\mathcal{I}_\sigma\varphi : \overline{T} \to \mathbb{R}_+ \cup \{+\infty\}$, by

$$\mathcal{I}_\sigma\varphi(x) = \sum_\alpha k(x,\alpha)\varphi(\alpha)\sigma(\alpha) = \sum_{E \ni \alpha \supset x} \varphi(\alpha)\sigma(\alpha).$$

The *co-potential* of a function $\psi : \overline{T} \to \mathbb{R}_+$ with respect to a positive Borel measure μ on $\overline{T}$ is defined as the edge function

$$\mathcal{I}^*_\mu \psi(\alpha) = \int_{\overline{T}} k(x,\alpha)\psi(x)d\mu(x) = \int_{S(\alpha)} \psi(x)d\mu(x), \quad \alpha \in E.$$

The co-potential of μ is intended to be $\mathcal{I}^*_\mu(\alpha) := \mathcal{I}^*_\mu 1(\alpha) = \mu(S(\alpha))$. Observe that, if $\psi \in L^1(\overline{T},\mu)$, by Fubini's theorem we have $\langle \mathcal{I}_\sigma \varphi, \psi\rangle_{L^2(\overline{T},\mu)} = \langle \varphi, \mathcal{I}^*_\mu \psi\rangle_{\ell^2(E,\sigma)}$.

For a Hölder dual pair of exponents p, p^*, we can further associate to the measure μ a *nonlinear Wolff potential*, $V^{\mu,\sigma}_p(x) = \mathcal{I}_\sigma(\mathcal{I}^*_\mu)^{p^*-1}(x)$, $x \in \overline{T}$. More explicitly, one has

$$\begin{aligned} V^{\mu,\sigma}_p(x) &= \sum_\alpha k(\alpha,x)\sigma(\alpha)\Big(\int_{\overline{T}} k(\alpha,y)d\mu(y)\Big)^{p^*-1} \\ &= \sum_{\alpha\in[o^*,x]} \sigma(\alpha)\mu\big(S(\alpha)\big)^{p^*-1}. \end{aligned}$$

In the linear case $p = 2$, the sum and the integral can be switched and the above potential is expressed as

$$V^{\mu,\sigma}_2(x) = \int_{\overline{T}} \sum_\alpha k(\alpha, x\wedge y)\sigma(\alpha)d\mu(y) = \int_{\overline{T}} \sum_{\alpha\supset x\wedge y} \sigma(\alpha)d\mu(y).$$

The *p-energy* of the charge distribution μ is given by

$$\mathcal{E}^\sigma_p(\mu) = \int_{\overline{T}} V^{\mu,\sigma}_p(x)d\mu(x).$$

If the energy is finite, by Fubini's theorem it holds that

$$\mathcal{E}^\sigma_p(\mu) = \|\mathcal{I}^*_\mu\|^{p^*}_{\ell^{p^*}(E,\sigma)} = \sum_\alpha \mu\big(S(\alpha)\big)^{p^*}\sigma(\alpha).$$

We define the capacity of a closed subset $A \subseteq \overline{T}$ as

$$\begin{aligned} \mathrm{Cap}^\sigma_p(A) &= \inf\big\{\|\varphi\|^p_{\ell^p(E,\sigma)} : \varphi \geq 0,\ \mathcal{I}_\sigma\varphi(x) \geq 1 \text{ on } A\big\} \\ &= \sup\left\{\frac{\mu(A)^p}{\mathcal{E}^\sigma_p(\mu)^{p-1}} : \mathrm{supp}(\mu) \subseteq A\right\}. \end{aligned}$$

The equality between the first and second line above is given by a classical theorem in potential theory that can be found, for instance, in [2]. The (p,σ)-equilibrium function for A is the unique function φ satisfying $\mathcal{I}_\sigma\varphi = 1$, Cap^σ_p-quasi everywhere on A, and

$$\mathrm{Cap}^\sigma_p(A) = \|\varphi\|^p_{\ell^p(E,\sigma)}.$$

Similarly, one defines the (p,σ)-equilibrium measure for A as the unique measure probability measure such that $\mathrm{Cap}_p^\sigma(A) = \mu(A)$. Two of the authors recently found a characterization for equilibrium measures on trees [7].

Observe that the boundedness of $\mathcal{I}_\sigma : \ell^p(\sigma) \to L^p(\mu)$ is equivalent to that of the Hardy operator

$$\mathcal{I} : \ell^p(\pi) \to L^p(\mu),$$

under the correspondence $\pi(\alpha) = \sigma(\alpha)^{1-p}$. Following this paradigm, we can translate the natural potential theoretic objects in the π-dictionary, which turns out to be more adjusted to our scopes:

$$\mathcal{E}_p^\sigma(\mu) = \mathcal{E}_{p,\pi}(\mu) := \sum_\alpha \mu\big(S(\alpha)\big)^{p^*}\pi(\alpha)^{1-p^*},$$

$$\mathrm{Cap}_p^\sigma(A) = \mathrm{Cap}_{p,\pi}(A) = \inf\left\{\|\varphi\|^p_{\ell^p(E,\pi)} : \varphi \geq 0,\ \mathcal{I}\varphi(x) \geq 1 \text{ on } A\right\}.$$

2.2. Strong capacitary inequality and the isocapacitary characterization

We are ready to prove the isocapacitary characterization for trace measures, that is, the equivalence of (i) and (iii) in Theorem 2. As it is common, we will prove it passing through the so-called capacitary strong inequality. There are various versions of such inequality. Here we naturally treat the case of a tree; a proof for a large class of kernels in the continuous case can be found in [2, Theorem 7.1.1].

THEOREM 5 (Capacitary strong inequality). *Let $1 < p < \infty$ and $\varphi : E \to \mathbb{R}_+$,*

$$\sum_{k=-\infty}^{+\infty} 2^{pk}\,\mathrm{Cap}_{p,\pi}\left(\{x : \mathcal{I}\varphi(x) > 2^k\}\right) \leq \frac{2^p}{2^p - 1}\|\varphi\|^p_{\ell^p(E,\pi)}.$$

PROOF. Set $\Omega_k = \{x \in T : \mathcal{I}\varphi(x) > 2^k\}$, $\partial\Omega_k = \{x \in \Omega_k : \mathcal{I}\varphi(p(x)) \leq 2^k\}$. For given k, let $\varphi_k(\alpha) = 2^{k-m(\alpha)}\varphi(\alpha)$ if $e(\alpha) \in \partial\Omega_k \cap \partial\Omega_{k+1}$, where $m(\alpha) = m$ is the largest integer for which $e(\alpha) \in \partial\Omega_{m+1}$, and $\varphi_k(\alpha) = \varphi(\alpha)\cdot\chi_{\Omega_k\setminus\Omega_{k+1}}(e(\alpha))$ otherwise. Let $x \in \partial\Omega_{k+1}$ and $y = [o,x] \cap \partial\Omega_k$. Then, if $x \neq y$, we have

$$\mathcal{I}\varphi_k(x) = \mathcal{I}\varphi(x) - \mathcal{I}\varphi\big(p(y)\big) > 2^{k+1} - 2^k = 2^k,$$

while if $x = y$ and α is the unique edge such that $x = e(\alpha)$, we have

$$\varphi(\alpha) = \mathcal{I}\varphi(x) - \mathcal{I}\varphi\big(p(x)\big) > 2^{m+1} - 2^m$$

and, therefore,

$$\mathcal{I}\varphi_k(x) = \varphi_k(\alpha) = 2^{k-m}\varphi(\alpha) > 2^k.$$

Hence, $2^{-k}\varphi_k$ is a testing function for $\mathrm{Cap}_{p,\pi}(x : \mathcal{I}\varphi(x) > 2^k)$. Summing and using that the supports of the $\varphi_k's$ can meet only at points belonging to multiple boundaries,

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}} 2^{pk}\mathrm{Cap}_{p,\pi}(\Omega_k) \\
&\quad\le \sum_{k\in\mathbb{Z}} 2^{pk} \sum_{\alpha} 2^{-pk} \left|\varphi_k(\alpha)\right|^p \pi(\alpha) \\
&\quad= \sum_{k\in\mathbb{Z}} \Big(\sum_{e(\alpha)\in\Omega_k\setminus\Omega_{k+1}} \left|\varphi(\alpha)\right|^p \pi(\alpha) + \sum_{e(\alpha)\in\partial\Omega_k\cap\partial\Omega_{k+1}} \frac{\left|\varphi(\alpha)\right|^p \pi(\alpha)}{2^{p(m(\alpha)-k)}} \Big) \\
&\quad= \sum_{\alpha} \pi(\alpha)\left|\varphi(\alpha)\right|^p + \sum_{\alpha} \pi(\alpha)\left|\varphi(\alpha)\right|^p \sum_{k<m(\alpha)} \frac{\chi_{\partial\Omega_k\cap\partial\Omega_{k+1}}\big(e(\alpha)\big)}{2^{p(m(\alpha)-k)}} \\
&\quad\le \frac{2^p}{2^p-1}\|\varphi\|^p_{\ell^p(E,\pi)}. \qquad \blacksquare
\end{aligned}$$

Proof of the equivalence of (i) and (iii) in Theorem 2. Suppose that as always φ is a positive function defined on the edges of the tree, then using the distribution function we write

$$\begin{aligned}
\int_{\overline{T}} \mathcal{I}\varphi^p d\mu &= \int_0^\infty \mu\big(\{x\in\overline{T} : \mathcal{I}\varphi(x) > \lambda\}\big) d\lambda^p \\
&\le (2^p-1)\sum_{k=-\infty}^{+\infty} 2^{kp}\mu\big(x\in\overline{T} : \mathcal{I}\varphi(x) > 2^k\big) \\
&\le (2^p-1)[\![\mu]\!]_c^p \sum_{k=-\infty}^{+\infty} 2^{kp}\,\mathrm{Cap}_{p,\pi}\left(\{x\in\partial T : \mathcal{I}\varphi(x) > 2^k\}\right) \\
&\le 2^p[\![\mu]\!]_c^p\|\varphi\|^p_{\ell^p(\pi)}.
\end{aligned}$$

This concludes the proof of sufficiency.

To prove necessity, let $\alpha_1,\dots,\alpha_n \in E$ and denote by φ the equilibrium function associated to the set $\bigcup_{i=1}^n S(\alpha_i)$. Then, $I\varphi \ge 1$, $\mathrm{Cap}_{p,\pi}$-quasi everywhere and in particular μ-a.e. on $\bigcup_{i=1}^n S(\alpha_i)$. It follows that

$$\begin{aligned}
[\mu]^p\,\mathrm{Cap}_{p,\pi}\left(\bigcup_{i=1}^n S(\alpha_i)\right) &= [\mu]^p\sum_{\alpha}\varphi(\alpha)^p\pi(\alpha) \ge \int_{\overline{T}}(\mathcal{I}\varphi)^p d\mu \\
&\ge \sum_{i=1}^n \int_{S(\alpha_i)}(\mathcal{I}\varphi)^p d\mu \ge \sum_{i=1}^n \mu\big(S(\alpha_i)\big). \qquad \blacksquare
\end{aligned}$$

Problem 1. Find the best constant in the inequality $[\mu] \le 2^p[\![\mu]\!]_c$ of Theorem 2.

3. Mass-energy characterization: three different proofs

In this section, we will prove the mass-energy characterization for trace measures, that is, the equivalence of (i) and (ii) in Theorem 2. We will give three different proofs, based on different techniques. The easiest proof only works for $p = 2$, while the other two work for any $1 < p < \infty$.

We recall that

$$[\![\mu]\!]^{p^*-1} = \sup_{\alpha\in E} \frac{\sum_{\beta\subseteq\alpha} \pi(\beta)^{1-p^*}\mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\alpha)\big)},$$

and we call it the *energy-mass ratio*.

3.1. Maximal function

This simple proof can be found, for instance, in [16]. It relies on the L^p inequality for a suitable maximal function. If $\mu, \sigma \geq 0$ are measures on $\overline{T}$ and f a function on $\overline{T}$, then

$$M_\mu(f d\sigma)(x) := \max_{\alpha\in[o^*,x]} \frac{1}{\mu\big(S(\alpha)\big)} \int_{S(\alpha)} |f| d\sigma.$$

We simplify the notation by setting $M_\mu f = M_\mu(f d\mu)$. We have the following weak-$(1, 1)$ estimate.

Theorem 6. *Suppose that $\mu, \sigma \geq 0$ are measures on $\overline{T}$ and ψ a positive function on $\overline{T}$. Then,*

(a) $\lambda\sigma(x : M_\mu\psi(x) > \lambda) \leq \int_{\overline{T}} \psi M_\mu(d\sigma)d\mu,$

(b) *for* $1 < p < \infty$, $\int_{\overline{T}} (M_\mu\psi)^p d\sigma \leq (p^*)^p \int_{\overline{T}} \psi^p M_\mu(d\sigma)d\mu.$

Proof. First we prove (a). Fix $\lambda > 0$ and set $E(\lambda) = \{x \in \overline{T} : M_\mu\psi(x) > \lambda\} = \sqcup_j S(\alpha_j)$, where $⨍_{S(\alpha_j)} \psi d\mu > \lambda$ and $⨍_{S(\beta)} \psi d\mu \leq \lambda$ if $\beta \supset \alpha_j$. Then,

$$\begin{aligned}
\sigma\big(E(\lambda)\big) &= \sum_j \sigma\big(S(\alpha_j)\big) = \sum_j \frac{\sigma\big(S(\alpha_j)\big)}{\mu\big(S(\alpha_j)\big)}\mu\big(S(\alpha_j)\big) \\
&\leq \frac{1}{\lambda}\sum_j \frac{\sigma\big(S(\alpha_j)\big)}{\mu\big(S(\alpha_j)\big)} \int_{S(\alpha_j)} \psi d\mu \leq \frac{1}{\lambda}\sum_j M_\mu(d\sigma)(x_j) \int_{S(\alpha_j)} \psi d\mu \\
&\leq \frac{1}{\lambda}\sum_j \int_{S(\alpha_j)} \psi M_\mu(d\sigma)d\mu \leq \frac{1}{\lambda}\int_{\overline{T}} \psi M_\mu(d\sigma)d\mu,
\end{aligned}$$

which proves the weak estimate. Then, a simple application of a variation of the classical Marcinkiewicz interpolation theorem [36, Exercise 1.3.3] gives us (b) from (a). ∎

PROOF OF THE EQUIVALENCE OF (I) AND (II) IN THEOREM 2. Assume that the mass-energy condition holds and define the measure σ on $\overline{T}$ by setting

$$\sigma\big(S(\alpha)\big) = \sum_{\beta\subseteq\alpha} \mu\big(S(\beta)\big)^{p^*} \pi(\beta)^{1-p^*}.$$

The mass-energy condition (ME) can be written as $\sigma(S(\alpha)) \le [\![\mu]\!]^{p^*} \mu(S(\alpha))$. Thus, for any positive function ψ on $\overline{T}$,

$$\begin{aligned}
\sum_{\alpha\in E} (I_\mu^*\psi)^{p^*}(\alpha)\pi(\alpha)^{1-p^*} &= \sum_{\alpha\in E} \left(⨍_{S(\alpha)} \psi d\mu \right)^{p^*} \sigma\big(e(\alpha)\big) \\
&\le \sum_{\alpha\in E} (M_\mu\psi)^{p^*} e\big((\alpha)\big)\sigma\big(e(\alpha)\big) \\
&\le \int_{\overline{T}} (M_\mu\psi)^{p^*}(x)\sigma(x) \\
&\le p^{p^*} \int_{\overline{T}} \psi^{p^*} M_\mu(d\sigma) d\mu \\
&\le p^{p^*} [\![\mu]\!]^{p^*-1} \int_{\overline{T}} \psi^{p^*} d\mu.
\end{aligned}$$

So the dual Hardy inequality (H*), equivalent to (H), is obtained.

Conversely, suppose that the dual Hardy inequality (H*) holds. By testing it on functions $\psi = \chi_{S(\alpha)}$, we obtain

$$\mu\big(S(\alpha)\big) \sum_{\beta\supset\alpha} \pi(\beta)^{1-p^*} + \sum_{\beta\subseteq\alpha} \pi(\beta)^{1-p^*} \mu\big(S(\beta)\big)^{p^*} \le [\mu]^{p^*-1} \mu\big(S(\alpha)\big),$$

which implies the mass-energy condition (ME). ■

3.2. *Monotone proof for $p = 2$*

This easiest proof only works for $p = 2$, because it uses the C^* identity for operators on Hilbert spaces.

LEMMA 7. *Let $T : H_1 \to H_2$ be a bounded and linear operator between Hilbert spaces. Then,*

$$\|T\| = \|T^*\| = \|TT^*\|^{1/2}.$$

LEMMA 8. *Let $\mu, \nu \ge 0$ be measures on $\overline{T}$, suppose that $\mu(S(\alpha)) \le \nu(S(\alpha))$ holds for all α in E, and suppose that $f : \overline{T} \to \mathbb{R}_+$ is monotone, $f(x) \ge f(y)$ if $x \subseteq y$. Then, $\int f d\mu \le \int f d\nu$.*

PROOF. It suffices to prove the inequality of the corresponding distribution functions. For $t > 0$,

$$\begin{aligned}\mu(x : f(x) > t) &= \sum_j \mu\big(S(\alpha_j)\big) \text{ since } f \text{ is monotone,}\\ &\le \sum_j \nu\big(S(\alpha_j)\big) = \nu\big(x : f(x) > t\big).\end{aligned}$$

■

PROOF OF THE EQUIVALENCE OF (I) AND (II) IN THEOREM 2 FOR $p = 2$. Suppose initially that our tree is finite, but arbitrarily large. This assures that all relevant operators are bounded. If we manage to estimate the norm of the Hardy operator independently of the length of the tree, then we can pass to the infinite case by a simple limiting argument.

For a $g : E \to \mathbb{R}_+$, we first compute

$$\begin{aligned}\mathcal{I}^*\mathcal{I}g(\alpha) &= \pi(\alpha)^{-1}\int_{S(\alpha)}\sum_{\beta\in[o^*,x]} g(\beta)d\mu(x)\\ &= \pi(\alpha)^{-1}\sum_\beta g(\beta)\mu\big(S(\beta)\cap S(\alpha)\big)\\ &= \pi(\alpha)^{-1}\sum_{\beta\subseteq\alpha} g(\beta)\mu\big(S(\beta)\big) + \pi(\alpha)^{-1}\mu\big(S(\alpha)\big)\sum_{\beta\supseteq\alpha} g(\beta)\\ &= \pi(\alpha)^{-1}\sum_{\beta\subseteq\alpha} g(\beta)\mu\big(S(\beta)\big) + \pi(\alpha)^{-1}\mu\big(S(\alpha)\big)\mathcal{I}g\big(e(\alpha)\big)\\ &= T_1g(\alpha) + T_2g(\alpha).\end{aligned}$$

Therefore,

$$\|\mathcal{I}^*\mathcal{I}\|_{\ell^2(\pi)} \le \|T_1\|_{\ell^2(\pi)} + \|T_2\|_{\ell^2(\pi)}.$$

The second norm can be computed in terms of the norm $\mathcal{I}$. Consider a measure ρ on T such that $\rho(e(\alpha)) = \pi^{-1}(\alpha)\mu(S(\alpha))$. The mass-energy condition allows us to apply Lemma 8 to the measures ρ and μ, obtaining, therefore,

$$\begin{aligned}\|T_2g\|^2_{\ell^2(\pi)} &= \sum_\alpha \pi(\alpha)^{-1}\mu\big(S(\alpha)\big)^2\big(\mathcal{I}g\big(e(\alpha)\big)\big)^2\\ &= \int_T(\mathcal{I}g)^2d\rho \le [\![\mu]\!]\int_T(\mathcal{I}g)^2d\mu\\ &= [\![\mu]\!]\sum_\alpha \mu(\alpha)\big(\mathcal{I}g\big(e(\alpha)\big)\big)^2\\ &\le [\![\mu]\!]\|g\|^2_{\ell^2(\pi)}\|\mathcal{I}\|^2_{\ell^2(\pi)\to L^2(\mu)}.\end{aligned}$$

A standard calculation shows that $T_2^* = T_1$ (with respect to the inner product in $\ell^2(\pi)$), hence we have for free the estimate on the norm of T_1. Putting everything together we get

$$\|\mathcal{I}\|^2_{\ell^2(\pi)\to L^2(\mu)} = \|\mathcal{I}\mathcal{I}^*\|_{\ell^2(\pi)} \le 2[\![\mu]\!]^{\frac{1}{2}}\|\mathcal{I}\|_{\ell^2(\pi)\to L^2(\mu)}.$$

Since $\mathcal{I}$ is bounded because our tree is finite, we can divide both sides of the inequality with its norm to get

$$\|\mathcal{I}\|^2_{\ell^2(\pi)\to L^2(\mu)} \le 4[\![\mu]\!].$$

Notice that this proof gives the best constant. ∎

3.3. *Bellman function*

In this section, we provide a different proof, based on a Bellman function approach, of the fact that (ME) implies (H). Let T be a general rooted tree and denote by $|\cdot|$, as usual, the *canonical* edge weight defined in Section 2.1. Set $k_\beta = |\beta|/|\alpha|$ when $\beta \in s(\alpha)$.

The following is a tree version of the weighted dyadic Carleson imbedding theorem by Nazarov, Treil, and Volberg [57]. With respect to the standard dyadic case, this tree analogue presents some extra difficulty, due to the fact that the objects into play are here not martingales but only supermartingales.

THEOREM 9 (Carleson imbedding theorem for trees). *Let σ be a nonnegative weight on E and λ a measure on $\overline{T}$ satisfying*

$$\sum_{\beta\subseteq\alpha}\sigma(\beta)\frac{I^*_\lambda(\beta)^p}{|\beta|^p} \le I^*_\lambda(\alpha), \quad \textit{for every } \alpha \in E. \tag{3.1}$$

Then,

$$\sum_{\alpha\in E}\sigma(\alpha)\frac{I^*_\lambda\varphi(\alpha)^p}{|\alpha|^p} \le (p^*)^p\|\varphi\|^p_{\ell^p(\lambda)}, \quad \textit{for every } \varphi : E \to \mathbb{R}_+. \tag{3.2}$$

Letting μ be the measure defined by $\mu(S(\alpha)) = I^*_\lambda(\alpha)/|\alpha|$, and switching to the π-dictionary under the usual correspondence $\pi(\alpha) = \sigma(\alpha)^{1-p}$, we see that (3.1) is equivalent to the mass-energy condition (ME) and (3.2) coincides with the dual Hardy inequality (H*).

On the dyadic tree, a proof of the above theorem relying on a Bellman function method was first given in [6] for $p = 2$, and later extended to every $1 < p < \infty$ in [26]. In his paper, it is proven that the Bellman function employed is *the* Bellman function of a problem in stochastic optimal control. We give here a slightly adapted proof which works on every tree T. We also remark that the result remains true substituting the canonical weight $|\cdot|$ with a general weight w fulfilling the so-called *flow condition*, that is, $\sum_{\beta\in s(\alpha)} w(\beta) = w(\alpha)$, for any $\alpha \in E$. Indeed, in the following proof the flow condition is the only property of the canonical weight which is used. We refer the reader to [30,47,48] for some recent result concerning trees endowed with flow measures and flow weights.

PROOF OF THEOREM 9. We begin by observing that it is enough to show the result for nonnegative functions. For any edge α and any quadruple of nonnegative real numbers F, f, A, v, define $\Omega_\alpha(F, f, A, v)$ to be the set of weights σ, measures λ, and functions φ such that

$$\frac{1}{|\alpha|} I^*_\lambda(\varphi^p)(\alpha) = F, \quad \frac{1}{|\alpha|} I^*_\lambda \varphi(\alpha) = f,$$
$$\frac{1}{|\alpha|} \sum_{\beta\subseteq\alpha} \sigma(\beta) \frac{I^*_\lambda(\beta)^p}{|\beta|^p} = A, \quad \frac{1}{|\alpha|} I^*_\lambda(\alpha) = v.$$

In order for $\Omega_\alpha(F, f, A, v)$ not to be empty, it must be $f^p \le F v^{p-1}$, the condition coming from Hölder's inequality. Moreover, (3.1) implies $A \le v$. We denote that by $\mathcal{D}$ the domain

$$\{f^p \le F v^{p-1},\ A \le v\} \subseteq \mathbb{R}^4,$$

which is clearly convex, being the intersection of the half plane $\{A \le v\}$ with the cylindroid having as a basis the convex set

$$\{f^p \le F v^{p-1}\} \subseteq \mathbb{R}^3.$$

Define the *Belmann function* $\mathbb{B} : \mathbb{R}^4 \to \mathbb{R}_+$,

$$\mathbb{B}(F, f, A, v) = \sup_{\Omega_\omega(F,f,A,v)} \frac{1}{|\alpha|} \sum_{\beta\subseteq\omega} \sigma(\beta) \frac{I^*_\lambda \varphi(\beta)^p}{|\beta|^p}.$$

We aim to prove that

$$\mathbb{B}(F, f, A, v) \le |\omega| (p^*)^p F, \quad \text{for all } (F, f, A, v) \in \mathcal{D}.$$

Let $x = (F, f, A, v)$ be a point in $\mathcal{D}$ and fix arbitrarily $(\sigma, \lambda, \varphi) \in \Omega_\omega(x)$. For each $\alpha \in E$, let $x_\alpha = (F_\alpha, f_\alpha, A_\alpha, v_\alpha) \in \mathbb{R}^4$ be the unique point such that $(\sigma, \lambda, \varphi) \in \Omega_\alpha(x_\alpha)$. In particular, $x_\omega = x$, and it is clear that $x_\alpha \in \mathcal{D}$ for each α. Moreover, the additivity of I^* gives the relations

$$F_\alpha = \frac{1}{|\alpha|}\varphi^p(\alpha)\lambda(\alpha) + \sum_{\beta\in s(\alpha)} k_\beta F_\beta, \quad f_\alpha = \frac{1}{|\alpha|}\varphi(\alpha)\lambda(\alpha) + \sum_{\beta\in s(\alpha)} k_\beta f_\beta,$$
$$A_\alpha = \sigma(\alpha)\frac{I^*_\lambda(\alpha)^p}{|\alpha|^{p+1}} + \sum_{\beta\in s(\alpha)} k_\beta A_\beta, \quad v_\alpha = \frac{\lambda(\alpha)}{|\alpha|} + \sum_{\beta\in s(\alpha)} k_\beta v_\beta.$$

By denoting $b_\alpha^p = |\alpha|^{-1}\varphi^p(\alpha)$, $c_\alpha = \sigma(\alpha)|\alpha|^{-(p+1)} I_\lambda^*(\alpha)^p$, and $a_\alpha^{p^*} = |\alpha|^{-1}\lambda(\alpha)$ and defining the point $y_\alpha = (b_\alpha^p, a_\alpha b_\alpha, c_\alpha, a_\alpha^{p^*})$, the above relations can be rewritten as

$$x_\alpha = y_\alpha + \sum_{\beta\in s(\alpha)} k_\beta x_\beta, \quad \alpha \in E. \tag{3.3}$$

Now, suppose that we can design a concrete function, $\mathcal{B} : \mathbb{R}^4 \to \mathbb{R}$ such that

(i) $\mathcal{B}(F, f, A, v) \le (p^*)^p F$ on $\mathcal{D}$, and satisfying

(ii) $|\alpha|\mathcal{B}(x_\alpha) - \sum_{\beta\in s(\alpha)} |\beta|\mathcal{B}(x_\beta) \ge \sigma(\alpha) f_\alpha^p$, for every $\alpha \in E$.

Then, summing over $\alpha \in E$ both sides of the inequality and exploiting the telescopic structure of the summand, we obtain

$$\sum_{\alpha\in E} \sigma(\alpha)\frac{1}{|\alpha|} I_\lambda^*(\varphi^p)(\alpha) = \sum_{\alpha\in E} \sigma(\alpha) f_\alpha^p \le |\omega|\mathcal{B}(x) \le |\omega|(p^*)^p F,$$

from which follows the thesis,

$$\mathbb{B}(F, f, A, v) \le |\omega|(p^*)^p F, \quad \text{for all } (F, f, A, v) \in \mathcal{D}.$$

We now claim that the function

$$\mathcal{B}(F, f, A, v) = (p^*)^p\left(F - \left(\frac{p-1}{A + (p-1)v}\right)^{p-1} f^p\right)$$

fulfills the desired properties. The construction of a Bellman function is a delicate matter. The interested reader can find more information and examples in [23, 24, 58]. It is immediate that (i) holds on $\mathcal{D}$. For any chosen $x = (F, f, A, v) \in \mathcal{D}$, let x_α be the associated family of points solving (3.3). Then, also the points $x_\alpha^* = x_\alpha - (0, 0, c_\alpha, 0)$ and $x_\alpha^{**} = x_\alpha - y_\alpha$ belong to the convex domain $\mathcal{D}$. Since $\mathcal{B}$ is clearly concave in the third variable, we have

$$\mathcal{B}(x_\alpha) - \mathcal{B}(x_\alpha^*) \ge c_\alpha \frac{\partial\mathcal{B}}{\partial A}(x_\alpha) = c_\alpha\left(\frac{p f_\alpha}{A_\alpha + (p-1)v_\alpha}\right)^p \ge c_\alpha\left(\frac{f_\alpha}{v_\alpha}\right)^p,$$

the last inequality following from the domain constraint $A_\alpha \le v_\alpha$. Indeed, $\mathcal{B}$ is also concave as a function of four variables on the convex set $\mathcal{D}$, as one can verify by checking that the Hessian matrix of

$$(F, f, A, v) \mapsto \mathcal{B}(F, f, A, v)$$

is positive semidefinite on $\mathcal{D}$. Hence, we have

$$\mathcal{B}(x_\alpha^*) - \mathcal{B}(x_\alpha^{**}) \ge \frac{\partial\mathcal{B}}{\partial F}(x_\alpha^*) b_\alpha + \frac{\partial\mathcal{B}}{\partial f}(x_\alpha^*) a_\alpha b_\alpha + \frac{\partial\mathcal{B}}{\partial v}(x_\alpha^*) a_\alpha^{p^*} \ge 0,$$

where the last inequality can be derived by direct calculations. Putting the pieces together we obtain

$$c_\alpha\left(\frac{f_\alpha}{v_\alpha}\right)^p \leq \mathcal{B}(x_\alpha)-\mathcal{B}(x_\alpha^*) = \mathcal{B}(x_\alpha)-\mathcal{B}(x_\alpha^{**})+\mathcal{B}(x_\alpha^{**})-\mathcal{B}(x_\alpha^*) \leq \mathcal{B}(x_\alpha)-\mathcal{B}(x_\alpha^{**}). \tag{3.4}$$

Exploiting the concavity of $\mathcal{B}$,

$$\mathcal{B}(x_\alpha^{**}) = \mathcal{B}\Big(\sum_{\beta\in s(\alpha)} k_\beta x_\beta\Big) \geq \sum_{\beta\in s(\alpha)} k_\beta \mathcal{B}(x_\beta),$$

which, by means of (3.4), yields to

$$c_\alpha\left(\frac{f_\alpha}{v_\alpha}\right)^p \leq \mathcal{B}(x_\alpha) - \sum_{\beta\in s(\alpha)} k_\beta \mathcal{B}(x_\beta).$$

It is easily seen that $c_\alpha = v_\alpha^p \sigma(\alpha)|\alpha|^{-1}$, which substituted above gives (ii). ■

It is clear that, a posteriori, the mass-energy and the isocapacitary conditions are equivalent, being both equivalent to (H). However, it is tempting to look for an a priori argument for the equivalence of these geometric conditions which does not pass through the boundedness of the Hardy operator. A direct proof that (ME) implies (ISO), for a family of weights including $\pi = 1$, is in [13], where it is also directly proven, for $p = 2$ and $\pi = 1$, that $\nu \leq \mu$ implies that $[\![\nu]\!] \leq 2[\![\mu]\!]$.

Problem 2. Find a proof of the equivalence between (ME) and (ISO), which works for every couple π, μ and does not require the boundedness of the Hardy operator.

4. A reverse Hölder inequality

In the particular case that $\pi \equiv 1$ and $p = 2$, the mass-energy condition can be rewritten in an interesting way as a consequence of the calculation

$$\begin{aligned}\sum_{\beta\subseteq\alpha}\mu\big(S(\beta)\big)^2 &= \sum_{\beta\subseteq\alpha}\int_{S(\beta)}\int_{S(\beta)} d\mu(x)d\mu(y)\\ &= \int_{S(\alpha)}\int_{S(\alpha)}\Big(\sum_{\beta\supseteq x\wedge y} 1\Big)d\mu(x)d\mu(y)\\ &= \int_{S(\alpha)}\int_{S(\alpha)} d(x\wedge y)d\mu(x)d\mu(y).\end{aligned}$$

Therefore, the mass-energy condition can be expressed as

$$\sup_{\alpha \in E} \fint_{S(\alpha)} \int_{S(\alpha)} d(x \wedge y) d\mu(x) d\mu(y) < +\infty.$$

Notice also that

$$\int_{S(\alpha)} d(x \wedge y) d\mu(x) = \mathcal{I}\mathcal{I}^*_\mu(\chi_{S(\alpha)}).$$

The following variation on the above condition:

$$\text{(s-Testing)} \qquad \sup_{\alpha \in E} \fint_{S(\alpha)} \left(\int_{S(\alpha)} d(x \wedge y) d\mu(x) \right)^s d\mu(y) := [\![\mu]\!]_s < \infty$$

is clearly stronger than the mass-energy condition for $s > 1$ due to Hölder's inequality. The surprising result is that in fact the conditions are equivalent, and the corresponding quantities are comparable. This result is in the spirit of the John–Nirenberg reverse Hölder inequality for BMO functions.

THEOREM 10. *For all measures μ, and $s > 1$,*

$$\sup_{\alpha \in E} \fint_{S(\alpha)} \left(\int_{S(\alpha)} d(x \wedge y) d\mu(x) \right)^s d\mu(y) \leq C_s [\![\mu]\!]^s.$$

In an implicit form, this result is contained in the work of Tchoundja [67]. The proof of the above theorem is based on a Calderón–Zygmund-type theorem for the operator $\mathcal{I}\mathcal{I}^*_\mu := T_\mu$. More precisely, the following theorem holds.

THEOREM 11. *Suppose that the operator*

$$T_\mu : L^2(\mu) \to L^2(\mu)$$

is bounded. Then, for any $s \in (1, +\infty)$ the operator

$$T_\mu : L^s(\mu) \to L^s(\mu)$$

is bounded. Furthermore, $\|T_\mu\|_{L^s(\mu)} \leq C_s \|T_\mu\|_{L^2(\mu)}$.

Since the underlying measure μ is not necessarily doubling, this theorem can be seen as a special case of [56, Theorem 1.1]. Here we will give a direct proof which also provides better quantitative estimates of the constants involved based on a good-λ inequality as in [67, 70].

LEMMA 12 (Good-λ inequality). *Let μ be a trace measure on $\overline{T}$. Then, for every $\eta > 0$, there exists $\gamma(\eta) > 0$ such that for any nonnegative function f on $\overline{T}$,*

$$\mu\{x \in \overline{T} : T_\mu f(x) > (1+\eta)\lambda \text{ and } M_\mu(f^2)(x) \leq \gamma^2\lambda^2\} \leq \frac{1}{2}\mu\{x \in \overline{T} : T_\mu f(x) > \lambda\}.$$

PROOF. Notice that the set $\{T_\mu f > \lambda\}$ is a *stopping time*. In other words, it can be written as a disjoint union of tent regions,

$$\{T_\mu f > \lambda\} = \bigcup_{i=1}^{\infty} S(\alpha_i).$$

It is, therefore, sufficient to prove that for all α_i we have

$$\mu\big\{x \in S(\alpha_i) : T_\mu f(x) > (1+\eta)\lambda \text{ and } M_\mu(f^2)(x) \le \gamma^2\lambda^2\big\} \le \frac{1}{2}\mu\big(S(\alpha_i)\big).$$

So for the rest of the proof we work on a fixed $S(\alpha_i)$ which we denote by $S(\alpha)$ to avoid an overload of notation. Let $f_1 = f\chi_{S(\alpha)}$ and $f_2 = f - f_1$. For $x \in S(\alpha)$,

$$\begin{aligned} T_\mu f_2(x) &= \int_{\bar{T}\setminus S(\alpha)} d(x \wedge y) f(y) d\mu(y) \\ &= \int_{\bar{T}\setminus S(\alpha)} d\big(b(\alpha) \wedge y\big) f(y) d\mu(y) \le T_\mu f\big(b(\alpha)\big) \le \lambda. \end{aligned}$$

Because $b(\alpha) \notin S(\alpha)$, therefore

$$T_\mu f(x) \le T_\mu f_1(x) + \lambda,$$

which implies that

$$\begin{aligned} &\mu\big\{x \in S(\alpha) : T_\mu f(x) > (1+\eta)\lambda \text{ and } M_\mu(f^2)(x) \le \gamma^2\lambda^2\big\} \\ &\quad\le \mu\big\{x \in S(\alpha) : T_\mu f_1(x) > \eta\lambda\big\} \\ &\quad\le \frac{1}{\eta^2\lambda^2}\int_{S(\alpha)} (T_\mu f_1)^2 d\mu \\ &\quad\le \frac{\|T_\mu\|^2_{L^2(\mu)}}{\eta^2\lambda^2}\int_{S(\alpha)} f^2 d\mu \\ &\quad\le \frac{\|T_\mu\|^2_{L^2(\mu)}\mu\big(S(\alpha)\big)}{\eta^2\lambda^2} M_\mu(f^2)\big(e(\alpha)\big) \\ &\quad\le \frac{\|T_\mu\|^2_{L^2(\mu)}\gamma^2}{\eta^2}\mu\big(S(\alpha)\big), \end{aligned}$$

where we assume without loss of generality that $M_\mu(f^2)(e(\alpha)) \le \gamma^2\lambda^2$; otherwise the left-hand side is zero. It suffices, therefore, to choose

$$\gamma = \frac{\eta}{\sqrt{2}\|T_\mu\|_{L^2(\mu)}}. \qquad \blacksquare$$

PROOF OF THEOREM 10. Since the operator T_μ is self-adjoint, it suffices to prove that $L^2(\mu)$ boundedness implies $L^s(\mu)$ boundedness for all $s > 2$. Let $s > 2$ and $f \in L^s(\overline{T}, \mu)$. Exploiting Lemma 12 and Theorem 6, we get

$$\begin{aligned}
\int_{\overline{T}} (T_\mu f)^s d\mu &= \int_0^\infty \mu(\{T_\mu f > \lambda\}) d\lambda^s \\
&= (1+\eta)^s \int_0^\infty \mu\big(\{T_\mu f > (1+\eta)\lambda\}\big) d\lambda^s \\
&\le (1+\eta)^s \int_0^\infty \mu\big(\{T_\mu f > (1+\eta)\lambda \text{ and } M_\mu(f^2) \le \gamma^2\eta^2\}\big) d\lambda^s \\
&\quad + (1+\eta)^s \int_0^\infty \mu\big(\{M_\mu(f^2) > \gamma^2\lambda^2\}\big) d\lambda^s \\
&\le \frac{(1+\eta)^s}{2} \int_0^\infty \mu(\{T_\mu f > \lambda\}) d\lambda^s + \frac{(1+\eta)^s}{\gamma^s} \int_{\overline{T}} M_\mu(f^2)^{\frac{s}{2}} d\mu \\
&\le \frac{(1+\eta)^s}{2} \int_{\overline{T}} (T_\mu f)^s d\mu + \frac{2^{s/2} s(1+\eta)^s}{(s-2)\gamma^s} \int_{\overline{T}} f^s d\mu,
\end{aligned}$$

which proves the thesis if η is chosen small. In particular,

$$\|T_\mu\|_{L^s(\mu)} \le C_s \|T_\mu\|_{L^2(\mu)}. \qquad \blacksquare$$

The reverse Hölder inequality is now a corollary of the above theorem.

PROOF OF THEOREM 11. Suppose that μ satisfies the mass-energy condition. Then,

$$\|T_\mu\|_{L^2(\mu)} = \|\mathcal{I}\mathcal{I}_\mu^*\|_{L^2(\mu)} = \|\mathcal{I}\|^2_{\ell^2 \to L^2(\mu)} \le 4[\![\mu]\!].$$

On the other hand,

$$\|T_\mu\|^s_{L^s(\mu)} \ge \frac{\|T_\mu(\chi_{S(\alpha)})\|^s_{L^s(\mu)}}{\mu(S(\alpha))} \ge \fint_{S(\alpha)} \mathcal{I}\mathcal{I}_\mu^*\big(\chi(S(a))\big)^s d\mu,$$

and the result follows from Theorem 10. $\blacksquare$

5. THE INEQUALITY OF MUCKENHOUPT AND WHEEDEN, AND WOLFF

In this section, we only consider only the case when T is a homogeneous tree. We recall that if each vertex of T has $q+1$ neighbors, then $|\alpha| = q^{-d(\alpha)}$, for each edge α. Let $0 < s < 1$ and $1 < p^* < \infty$. For any measure μ on $\overline{T}$, we trivially have

$$\left(\sup_{\alpha \supset x} \frac{\mu(S(\alpha))}{|\alpha|^s}\right)^{p^*} \le \sum_{\alpha \supset x} \left(\frac{\mu(S(\alpha))}{|\alpha|^s}\right)^{p^*} \le \left(\sum_{\alpha \supset x} \frac{\mu(S(\alpha))}{|\alpha|^s}\right)^{p^*} \quad \forall x \in \partial T.$$

The inequality of Muckenhoupt and Wheeden [55, Theorem 1], (MW) in the sequel, says that the chain of inequalities can be reversed, on average.[2]

THEOREM 13. *For any measure μ on ∂T, $p^* \geq 1$, and $0 < s < 1$, there is a constant $C = C(p^* s)$ such that*

$$\text{(MW)} \qquad \frac{1}{C}\int_{\partial T}\left(\sum_{\alpha\supset x}\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^{p^*}dx \leq \int_{\partial T}\sum_{\alpha\supset x}\left(\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^{p^*}dx$$
$$\leq C\int_{\partial T}\left(\sup_{\alpha\supset x}\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^{p^*}dx.$$

As usual, dx here is the Lebesgue measure for which $\int_{S(\alpha)} dx|\alpha|$. A first consequence of the (MW) inequality is that we have a one-parameter of seemingly different conditions characterizing μ's for which the Hardy inequality holds, provided that the weight π has the special form $\pi(\alpha) = |\alpha|^{\frac{p^* s-1}{1-p^*}}$. Indeed, the central term in (MW) can be written as an energy,

$$\int_{\partial T}\sum_{\alpha\supset x}\left(\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^{p^*}dx = \sum_{\alpha}\frac{\mu\big(S(\alpha)\big)^{p^*}}{|\alpha|^{-p^* s}}\int_{\partial S(\alpha)}dx$$
$$= \sum_{\alpha}\mu\big(S(\alpha)\big)^{p^*}|\alpha|^{(1-p^* s)}$$
$$= \mathcal{E}_{p,\pi}(\mu),$$

and the (MW) gives

$$\mathcal{E}_{p,\pi}(\mu) \approx \int_{\partial T}\left(\sum_{\alpha\supset x}\left(\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^q\right)^{p^*}dx, \quad \text{for all } q \geq 1.$$

PROPOSITION 14 (Wolff's inequality on the tree). *Let μ be a nonnegative Borel measure on ∂T. Then, for any $p^* \geq 1$ and $0 < s < 1$ one has*

$$\int_{\partial T}\left(\sum_{\alpha\supset x}\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\right)^{p^*}dx \lesssim \int_{\partial T}\sum_{\alpha\supset x}\frac{\big(\mu\big(S(\alpha)\big)\big)^{p^*}}{|\alpha|^{sp^*}}dx.$$

Since the particular choice of q plays no role, from now on, to keep the notation lighter, we fix the homogeneity of the tree T setting $q = 2$; i.e., we put ourselves back in the realm of the classical dyadic Hardy inequality (A.1). Given an edge $\alpha \in E$, we

(2) In fact, the full Muckenhoupt–Wheeden inequality in the Euclidean setting applies to more general situations when the underlying measure is only $\mathcal{A}_\infty$ equivalent to the Lebesgue measure dx.

denote by α^+ and α^- its two children edges. In this setting, Proposition 14 follows from the slightly more general statement below. The function $\varphi : T \to \mathbb{R}_+$ is a logarithmic supermartingale with the drift $d > 0$, if for every edge α one has

$$\frac{1}{2}\big(\log\varphi(\alpha^+) + \log\varphi(\alpha^-)\big) \le \log\varphi(\alpha) - d. \tag{5.2}$$

PROPOSITION 15. *Assume that φ is a logarithmic supermartingale with the drift $d > 0$, and that its jumps are bounded from above,*

$$\max\big(\varphi(\alpha^+), \varphi(\alpha^-)\big) \le C\varphi(\alpha), \quad \alpha \in E, \tag{5.3}$$

for some constant $C > 0$. Then, for any $p^* \ge 1$ one has

$$\begin{aligned}\int_{\partial T}\Big(\sum_{\alpha\supset x}\varphi(\alpha)\Big)^{p^*}dx &\le C_1(p^*, d)\int_{\partial T}\sum_{\alpha\supset x}\varphi^{p^*}(\alpha)\,dx \\ &\le C_2(p^*, d)\int_{\partial T}\sup_{\alpha\supset x}\varphi^{p^*}(\alpha)\,dx.\end{aligned} \tag{5.4}$$

PROOF OF PROPOSITION 14. One only needs to observe that $\varphi(\alpha) := \mu(S(\alpha))/|\alpha|^s$, $\alpha \in E$, defines a logarithmic supermartingale with the drift $\log 2\cdot(1-s)$ and bounded jumps. Indeed, given any edge α in T, we clearly have $|\alpha^\pm|^s = 2^{-s}|\alpha|^s$; hence, since $\mu(S(\alpha)) = \mu(S(\alpha^+)) + \mu(S(\alpha^-))$, we see that

$$\begin{aligned}\frac{\mu\big(S(\alpha^+)\big)}{|\alpha^+|^s}\cdot\frac{\mu\big(S(\alpha^-)\big)}{|\alpha^-|^s} &= 2^{2s}|\alpha|^{-2s}\mu\big(S(\alpha^+)\big)\mu\big(S(\alpha^-)\big) \\ &\le |\alpha|^{-2s}2^{2(s-1)}\big(\mu\big(S(\alpha^+)\big) + \mu\big(S(\alpha^-)\big)\big)^2 \\ &= 2^{2(s-1)}\bigg(\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}\bigg)^2.\end{aligned}$$

The logarithmic supermartingale property follows immediately. On the other hand,

$$\frac{\mu\big(S(\alpha^\pm)\big)}{|\alpha^\pm|^s} \le 2^s\frac{\mu\big(S(\alpha)\big)}{|\alpha|^s},$$

so the jumps of φ are clearly bounded from above. ■

In order to prove Proposition 15, we will need the following lemma, of which we postpone the proof.

LEMMA 16 (Wolff's lemma). *Fix $\delta > 0$ and $N > 1/2$. Let φ be a logarithmic supermartingale with positive drift $d > 0$ satisfying (5.3). Then, for any edge α_0 in T the following inequality holds:*

$$\sum_{\alpha\subseteq\alpha_0}\varphi^{\delta+N}(\alpha)|\alpha| \gtrsim \sum_{\alpha\subseteq\alpha_0}\varphi^{\delta}(\alpha)\sum_{\beta\subseteq\alpha}\varphi^N(\beta)|\beta|. \tag{5.5}$$

PROOF OF PROPOSITION 15. We only show the left inequality in (5.4). Let us introduce the following notations: write $[p^*]$ and $\{p^*\}$ for the integer and the decimal part of p and set

$$p^* = [p^*] + \{p^*\}, \quad r = \frac{1+\{p^*\}}{2}, \quad Q = [p^*] - 1.$$

Note that $p^* = Q + r + r$. Then, we have

$$\begin{aligned}
(5.6)\quad & \int_{\partial T} \Big(\sum_{\alpha \supset x} \varphi(\alpha)\Big)^q \, dx \\
& \leq \int_{\partial T} \Big(\sum_{\alpha \supset x} \varphi(\alpha)\Big)^Q \Big(\sum_{\alpha \supset x} \varphi^p(\alpha)\Big)\Big(\sum_{\alpha \supset x} \varphi^p(\alpha)\Big)\, dx \\
& = \int_{\partial T} \Big(\sum_{\alpha_1,\dots,\alpha_Q,\alpha_{Q+1},\alpha_{Q+2} \supset x} \varphi(\alpha_1)\dots\varphi(\alpha_Q)\varphi^p(\alpha_{Q+1})\varphi^p(\alpha_{Q+2})\Big)\, dx \\
& = \sum_{\pi \in S_{Q+2}} \sum_{\alpha_1 \supset \dots \supset \alpha_{Q+1} \supset \alpha_{Q+2}} \varphi^{p_{\pi(1)}}(\alpha_1)\dots\varphi^{p_{\pi(Q+2)}}(\alpha_{Q+2})|\alpha_{Q+2}|,
\end{aligned}$$

where S_{Q+2} is the symmetric group of all permutations of $\{1,\dots,Q+2\}$ and $p_j = 1$, $1 \leq j \leq Q$, $p_{Q+1} = p_{Q+2} = p$.

The next step is to use Wolff's lemma: given a permutation $\pi \in S_{Q+2}$, we apply (5.5) repeatedly to (5.6), obtaining

$$\begin{aligned}
& \sum_{\alpha_1 \supset \dots \supset \alpha_{Q+1} \supset \alpha_{Q+2}} \varphi^{p_{\pi(1)}}(\alpha_1)\dots\varphi^{p_{\pi(Q+2)}}(\alpha_{Q+2})|\alpha_{Q+2}| \\
& = \sum_{\alpha_1 \supset \dots \supset \alpha_Q} \varphi^{p_{\pi(1)}}(\alpha_1)\dots\varphi^{p_{\pi(Q)}}(\alpha_Q) \\
& \quad \times \Big(\sum_{\alpha_Q \supset \alpha_{Q+1} \supset \alpha_{Q+2}} \varphi^{p_{\pi(Q+1)}}(\alpha_{Q+1})\varphi^{p_{\pi(Q+2)}}(\alpha_{Q+2})|\alpha_{Q+2}|\Big) \\
& \lesssim \sum_{\alpha_1 \supset \dots \supset \alpha_Q} \varphi^{p_{\pi(1)}}(\alpha_1)\dots\varphi^{p_{\pi(Q)}}(\alpha_Q) \\
& \quad \times \Big(\sum_{\alpha_Q \supset \alpha_{Q+1}} \varphi^{p_{\pi(Q+1)}+p_{\pi(Q+2)}}(\alpha_{Q+1})|\alpha_{Q+1}|\Big) \\
& = \sum_{\alpha_1 \supset \dots \supset \alpha_{Q-1}} \varphi^{p_{\pi(1)}}(\alpha_1)\dots\varphi^{p_{\pi(Q-1)}}(\alpha_{Q-1}) \\
& \quad \times \Big(\sum_{\alpha_{Q-1} \supset \alpha_Q \supset \alpha_{Q+1}} \varphi^{p_{\pi(Q)}}(\alpha_Q)\varphi^{p_{\pi(Q+1)}+p_{\pi(Q+2)}}(\alpha_{Q+1})|\alpha_{Q+1}|\Big) \\
& \lesssim \dots \lesssim \sum_{\alpha_1} \varphi^{p_{\pi(1)}+\dots+p_{\pi Q+2}}(\alpha_1)|\alpha_1| = \int_{\partial T} \sum_{\alpha \supset x} \varphi^q(\alpha)\, dx.
\end{aligned}$$

Summing over all $\pi \in S_{Q+2}$, we obtain the first half of (5.4). ∎

We now prove Wolff's lemma. The proof is based on a careful analysis of slow and fast growing geodesics.

PROOF OF LEMMA 16. Without any loss of generality, we may assume that $N = 1$ (since the proof works all the same for every N) and $\alpha_0 = \omega$, the root edge.

What we are going to do next is to fix an edge α and look at the possible growth rate of $\varphi(\beta)$ for $\beta \subseteq \alpha$. The idea is that, if φ does not grow too fast in this region, then one could expect for the second sum on the right-hand side of (5.5) to be dominated by the value at the starting point,

$$\sum_{\beta\subseteq\alpha} \varphi(\beta)|\beta| \lesssim \varphi(\alpha)|\alpha|.$$

On the other hand, if φ grows very (exponentially) fast, then we write the right-hand side of (5.5) as $\sum_\beta \varphi(\beta)|\beta| \sum_{\alpha_0\supseteq\alpha\supseteq\beta} \varphi^\delta(\alpha)$, and expect the second sum to be estimated by the value at β.

Given $\alpha \in E$ and $k \geq 0$, let

$$A(\alpha, k) := \left\{\beta \subseteq \alpha : d(\beta, \alpha) = k,\ \frac{\varphi(\beta)}{\varphi(\alpha)} \leq (1 + r)^k\right\}$$

be the set of slowly growing successors (here $r = r(d, \delta) < 10^{-2}$ is some small constant to be chosen later), and let also $A(\alpha) = \bigcup_{k\geq 0} A(\alpha, k)$. We have

$$\begin{aligned}\sum_{\alpha\in E} \varphi^\delta(\alpha) \sum_{\beta\subseteq\alpha} \varphi(\beta)|\beta| &= \sum_{\alpha\in E} \varphi^\delta(\alpha) \sum_{\beta\in A(\alpha)} \varphi(\beta)|\beta| \\ &\quad + \sum_{\alpha\in E} \varphi^\delta(\alpha) \sum_{\beta\subseteq\alpha,\beta\notin A(\alpha)} \varphi(\beta)|\beta|.\end{aligned}$$

We start by estimating the second term,

$$\begin{aligned}\sum_{\alpha\in E} \varphi^\delta(\alpha) \sum_{\beta\subseteq\alpha,\beta\notin A(\alpha)} \varphi(\beta)|\beta| &= \sum_{\beta\in E} \varphi(\beta)|\beta| \sum_{\alpha\supseteq\beta,\beta\notin A(\alpha)} \varphi^\delta(\alpha) \\ &< \sum_{\beta\in E} \varphi(\beta)|\beta| \sum_{\alpha\supseteq\beta,\beta\notin A(\alpha)} \varphi^\delta(\beta)(1 + r)^{-\delta d(\alpha,\beta)} \\ &\leq C(r) \sum_{\beta\in E} \varphi^{\delta+1}(\beta)|\beta|.\end{aligned}$$

To deal with the first term we let

$$B(\alpha, k) := \left\{\beta \subseteq \alpha,\ d(\beta, \alpha) = k : \frac{\varphi(\beta)}{\varphi(\alpha)} \geq (1 - r)^k\right\}, \quad \alpha \in E,\ k \geq 0,$$

and, as before, $B(\alpha) = \bigcup_{k\geq 0} B(\alpha,k)$. The function φ decays exponentially outside of $B(\alpha)$, in particular,

$$\sum_{\alpha\in E}\varphi^{\delta}(\alpha)\sum_{\beta\subseteq\alpha,\,\beta\notin B(\alpha)}\varphi(\beta)|\beta| < \sum_{\alpha\in E}\sum_{\beta\subseteq\alpha,\,\beta\notin B(\alpha)}(1-r)^{d(\alpha,\beta)}\varphi^{1+\delta}(\alpha)|\beta| \le C(r)\sum_{\alpha\in E}\varphi^{1+\delta}(\alpha)|\alpha|.$$

So far, we took care of two types of behavior of φ: points of very fast growth (i.e., $\beta\notin A(\alpha)$), and points of very fast decay ($\beta\notin B(\alpha)$). Now we consider the points $\beta\subseteq\alpha$, where $\varphi(\beta)$ is roughly comparable to $\varphi(\alpha)$. It turns out that these points are very rare in the successor set of α. More precisely, we show that, for every $\alpha\in E$ and $k\geq 0$,

$$\big|B(\alpha,k)\big| = \sharp\big\{\beta\in B(\alpha,k)\big\} \le C(r)2^{\frac{k}{2}}. \tag{5.7}$$

The reason for this is that by the multiplicative property (5.2) the function φ decays exponentially on (geometric) average, and its pointwise growth rate is bounded from above.

Let $\Phi = \log\varphi$. By (5.3),

$$\max\left(\frac{\varphi(\beta^+)}{\varphi(\beta)},\frac{\varphi(\beta^-)}{\varphi(\beta)}\right)\lesssim 1,$$

hence

$$\max\Big(\Phi(\beta^+)-\Phi(\beta),\Phi(\beta^-)-\Phi(\beta)\Big)\lesssim 1.$$

On the other hand (remember that $r<10^{-2}$), if $\beta\in B(\alpha,k)$, then we get

$$\Phi(\beta)-\Phi(\alpha)\geq -kr = -d(\alpha,\beta)r.$$

Choose $r<\min(1/100,d/10)$. The inequality (5.7) now follows from the following lemma.

LEMMA 17. *Let $Y=\{Y_n\}$ be a dyadic supermartingale with drift $d>0$,*

$$\frac{1}{2}\big(Y(\beta^+)+Y(\beta^-)\big)\le Y(\beta)-d,$$

and its differences are bounded from above,

$$Y_n - Y_{n-1}\le C.$$

Then, for any $k\geq 0$ and $r\le d/10$ one has

$$\sharp\big\{\beta\in E: d(\beta,\omega)=k,\ Y(\beta)\geq -kr\big\}\le C(r)2^{\frac{k}{2}}.$$

Assume for a moment that we have Lemma 17, and hence (5.7). Then, we get

$$\begin{aligned}
&\sum_{\alpha\in E}\varphi^{\delta}(\alpha)\sum_{\beta\in B(\alpha)\bigcap A(\alpha)}\varphi(\beta)|\beta| \\
&\quad=\sum_{\alpha\in E}\varphi^{\delta}(\alpha)\sum_{k\geq 0}\sum_{\beta\in B(\alpha,k)\bigcap A(\alpha,k)}\varphi(\beta)|\beta| \\
&\quad\leq\sum_{\alpha\in E}\varphi^{\delta}(\alpha)\sum_{k\geq 0}\sum_{\beta\in B(\alpha,k)}(1+r)^k\varphi(\alpha)|\beta| \\
&\quad=\sum_{\alpha\in E}\varphi^{\delta+1}(\alpha)\sum_{k\geq 0}|B(\alpha,k)|(1+r)^k|\beta| \\
&\quad\leq\sum_{\alpha\in E}\varphi^{\delta+1}(\alpha)\sum_{k\geq 0}|B(\alpha,k)|2^{2rk}|\alpha|2^{-k} \\
&\quad\leq\sum_{\alpha\in E}\varphi^{\delta+1}(\alpha)|\alpha|\sum_{k\geq 0}|B(\alpha,k)|2^{-k(1-2r)} \\
&\quad\lesssim\sum_{\alpha\in E}\varphi^{\delta+1}(\alpha)|\alpha|\sum_{k\geq 0}2^{\frac{k}{2}}2^{-k(1-2r)} \\
&\quad\lesssim\sum_{\alpha\in E}\varphi^{\delta+1}(\alpha)|\alpha|,
\end{aligned}$$

and we are done. ■

Lemma 17 clearly follows from the following rescaled driftless version.

LEMMA 18. *Let $X=\{X_n\}$ be a dyadic supermartingale,*

$$\frac{1}{2}\big(X(\beta^+)+X(\beta^-)\big)\leq X(\beta), \tag{5.8}$$

and its differences are bounded from above

$$X_n-X_{n-1}\leq 1. \tag{5.9}$$

Then, for any $k\geq 0$ and $\eta>0$ one has

$$\sharp\big\{\beta\in E: d(\beta,\omega)=k,\ X(\beta)\geq k\eta\big\}\leq C(\eta)2^{\frac{k}{2}}.$$

This lemma is in turn a corollary of Azuma–Hoeffding inequality (essentially a good-λ argument for supermartingales).

PROPOSITION 19 (Azuma–Hoeffding inequality). *Let $Z=\{Z_n\}$ be a supermartingale with bounded differences,*

$$|Z_n-Z_{n-1}|\leq c_n,\quad n\in\mathbb{N}.$$

Then,

$$\mathbb{P}(Z_n - Z_0 \geq K) \leq e^{\frac{-K^2}{2\sum_{k=1}^{n} c_k^2}}, \quad K > 0,\ n \in \mathbb{N}.$$

While the proposition above requires the supermartingale differences to be bounded above and below, it is not really relevant here. Namely, assume that X satisfies the hypothesis of Lemma 18 and let $S = \{S_n\}$ be its differences,

$$S(\beta) = X(\beta) - X\big(P(\beta)\big), \quad \beta \in E,$$

where $P(\beta)$ is the parent of β. By (5.9), $S \leq 1$. Consider the set

$$F = \big\{\beta \in E : S(\beta) \leq -2\big\}$$

and define

$$\widetilde{S}(\beta) = \begin{cases} -2, & \beta \in F, \\ S(\beta) & \text{otherwise.} \end{cases}$$

Now let

$$Z(\beta) = \sum_{\alpha \supseteq \beta} \widetilde{S}(\alpha).$$

Clearly, Z is still a dyadic supermartingale, since $\widetilde{S}(\beta^-) + \widetilde{S}(\beta^+) \leq 0$ by (5.8). Also $\widetilde{S} \geq S$, hence $Z \geq X$. It is easily seen that Z has bounded differences and, therefore, satisfies Azuma–Hoeffding inequality.

6. Conformally invariant Hardy inequality

While the right-hand side of the Hardy inequality (H) does not depend upon the choice of the root vertex o, the Hardy operator contained in the left-hand side does, and consequently also the optimal constant $[\mu] = [\mu]_o$ depends on this choice. It is, therefore, natural to seek an alternative "conformal" invariant theory. The term "conformal invariant" should be interpreted in the sense that as (H) corresponds, as explained in the introduction, to a Carleson inequality for Besov spaces, in the same way the inequality we are going to introduce should correspond to a continuous inequality which remains invariant under the group of automorphisms of the unit disc.

We consider here the case $p = 2$, $\pi \equiv 1$. We also assume the tree is dyadic and not rooted: each vertex is the endpoint of three edges, and T is endowed with a rich group of automorphisms which, having the Poincaré distance in mind, play in T the role of conformal automorphisms. Such automorphisms are also isometries with respect to the distance d and act naturally also on the boundary ∂T (see [35] for a comprehensive exposition on the topic). Once we fix a root o, there are $3 \times 2^{n-1}$ vertices at distance n from it.

It is easily seen that the Hardy inequality (H), holding for functions $f : E \to \mathbb{R}$, is equivalent to

$$\int_{\overline{T}} \left|F(x)\right|^2 \mu(x) \leq [\mu]_o \Big(\left|F(o)\right|^2 + \sum_{\alpha \in E} \left|\nabla F(\alpha)\right|^2 \Big), \quad F : \overline{T} \to \mathbb{R}, \tag{6.1}$$

where $\nabla F(\alpha) = F(e(\alpha)) - F(b(\alpha))$ depends on the choice of the root, but $|\nabla F(\alpha)|^2$ does not. A first attempt to write down a "conformally invariant" formulation of the Hardy inequality is, assuming that $\mu(T) = 1$,

$$\int_{\overline{T}} \left|F(x) - \mu(F)\right|^2 d\mu(x) \leq [\mu]_{\mathrm{inv}} \sum_{\alpha} \left|\nabla F(\alpha)\right|^2, \tag{CH}$$

where $\mu(F) = \int_T F d\mu$ is the mean of F and $[\mu]_{\mathrm{inv}} \in [0, +\infty]$ the best constant in the inequality. The invariance is the following.

Let Ψ be an isometry of $\overline{T}$ and define $\Psi_* \mu(A) = \mu(\Psi^{-1}(A))$ and $\Psi^* F(x) = F(\Psi(x))$, $A \subseteq \overline{T}$. Then,

$$\begin{aligned} \int_{\overline{T}} \left|F(x) - \Psi_*\mu(F)\right|^2 d\Psi_*\mu(x) &= \int_{\overline{T}} \left|\Psi^* F(y) - \mu(\Psi^* F)\right|^2 d\mu(y) \\ &\leq [\mu]_{\mathrm{inv}} \sum_{\beta} \left|\nabla \Psi^* F(\beta)\right|^2 = [\mu]_{\mathrm{inv}} \sum_{\alpha} \left|\nabla F(\alpha)\right|^2, \end{aligned}$$

showing that $[\Psi_*\mu]_{\mathrm{inv}} = [\mu]_{\mathrm{inv}}$.

Observe that the finiteness of $[\mu]_o$ in (6.1) implies that μ has no atoms on ∂T. On the other hand, if μ is a Dirac delta measure supported on the boundary, the left-hand side of (CH) vanishes, while the average $\mu = \frac{\delta_x + \delta_y}{2}$ of two Dirac delta gives a true, nontrivial inequality.

We will show that if μ is not a boundary Dirac delta, then (CH) is equivalent to (H). We present two separate arguments, one for measures supported on the boundary of the tree, and another one for measures supported on the vertex set. The first case is proved by means of the isocapacitary characterization. Since also the capacity $\mathrm{Cap}(A) = \mathrm{Cap}_{2,\pi}(A)$ of a set $A \subseteq \overline{T}$ depends on the choice of the root o, in this section we will denote it by $\mathrm{Cap}_o(A)$, making explicit the dependence so far kept implicit. On the other hand, *condensers capacity* is invariant under Möbius trasformations of the unit disc and, on trees, under the action of automorphisms. Given two disjoint sets $A, B \subseteq \partial T$, each being a finite union of arcs[3], we define the capacity of the condenser

(3) A tent in a non-rooted dyadic tree is any rooted dyadic subtree, and an arc is the boundary of a tent. We consider, here and in the whole section, finite unions of tents only in order to avoid the complication of properties which, in the general case, only hold outside sets of capacity zero.

(A, B) as

$$\operatorname{Cap}(A, B) = \inf\left\{ \sum_{\alpha\in E} \left|\nabla F(\alpha)\right|^2 : F|_A = 1,\ F|_B = 0,\ \text{and } |\nabla F| \text{ has finite support}\right\}.$$

The next result shows that, for measures supported on the boundary, (6.1) and (CH) are equivalent, and relates the optimal constants to a capacitary expression.

THEOREM 20. *Let $\mu \geq 0$ be a Borel probability measure on $\overline{T}$, giving no mass to vertices, and not being a Dirac delta on the boundary. Then,*

$$[\![\mu]\!] \approx \inf_{o\in T} [\mu]_o \approx \sup_{A,B\subseteq \partial T} \frac{\mu(A)\mu(B)}{\operatorname{Cap}(A, B)},$$

where the supremum is over all couples of finite union of arcs.

The following lemma provides a recursive formula for calculating the capacity of a condenser of a special type. This kind of formulas arises often in the setting of discrete capacities.

LEMMA 21. *Let $o \in T$ be a vertex and let T_1, T_2, T_3 be the dyadic subtrees having it as pre-root. Let $A_j \subseteq \partial T_j$ be finite union of arcs. Then,*

$$\operatorname{Cap}(A_1, A_2 \cup A_3) = \frac{\operatorname{Cap}_o(A_1)\cdot\big(\operatorname{Cap}_o(A_2) + \operatorname{Cap}_o(A_3)\big)}{\operatorname{Cap}_o(A_1) + \operatorname{Cap}_o(A_2) + \operatorname{Cap}_o(A_3)}.$$

In particular, for $i \neq j \neq k$, $\operatorname{Cap}(A_i, \partial T_j \cup \partial T_k) \approx \operatorname{Cap}_o(A_i)$.

PROOF. Let $c_j = \operatorname{Cap}_o(A_j)$. As in [12, Proposition 1], it can be proved that there exists an extremal function $F \geq 0$ on T such that (a) $\lim_{T\ni x\to\zeta} F(x) = 0$ for $\zeta \in A_2 \cup A_3$, (b) $\lim_{T\ni x\to\xi} F(x) = 1$ for $\xi \in A_1$, (c) $\sum_{\alpha\in E} \nabla F(\alpha)^2 = \operatorname{Cap}(A_1, A_2 \cup A_3)$. Of course, such $|\nabla F|$ is not finitely supported. Similarly, there exist analogous extremal functions F_j for c_j, $j = 1, 2, 3$, with $F_j(a) = 0$; $\lim_{T\ni x\to\xi} F_j(x) = 1$ for $\xi \in A_j$; $\sum_{\alpha\in E} \nabla F_j(\alpha)^2 = c_j$.

By extremality, it is obvious that there are numbers $0 < s_j \leq 1$ such that

$$\left|\nabla F(\alpha)\right| = s_j \left|\nabla F_j(\alpha)\right|$$

on the edges α of T_j, and since $|\nabla F|$ adds to one along geodesics going from $A_1 \cup A_2$ to A_3, it must be $s_1 + s_2 = s_1 + s_3 = 1$. Again by minimality, we look for $t = s_1$ which minimizes

$$\begin{aligned} f(t) = \|\nabla F\|_{\ell^2}^2 &= (1-t)^2\big(\|\nabla F_2\|_{\ell^2}^2 + \|\nabla F_3\|_{\ell^2}^2\big) + t^2\|\nabla F_1\|_{\ell^2}^2 \\ &= (1-t)^2(c_2 + c_3) + t^2 c_1, \end{aligned}$$

which is minimal when $t = \frac{c_2+c_3}{c_1+c_2+c_3}$, and thus

$$\mathrm{Cap}(A_1, A_2 \cup A_3) = f\left(\frac{c_2+c_3}{c_1+c_2+c_3}\right) = \frac{c_1\cdot(c_2+c_3)}{c_1+c_2+c_3}.$$

Since clearly $\mathrm{Cap}_o(T_1) = \mathrm{Cap}_o(T_2) = \mathrm{Cap}_o(T_3) = 1$ and $\mathrm{Cap}_o(A_i) \le 1$,

$$\mathrm{Cap}(A_i, \partial T_j \cup \partial T_k) = \frac{2\,\mathrm{Cap}_o(A_i)}{2+\mathrm{Cap}_o(A_i)} \approx \mathrm{Cap}_o(A_i). \qquad \blacksquare$$

PROOF OF THEOREM 20. In one direction,

$$\begin{aligned}\int_{\partial T}(F-\mu(F))^2 d\mu &= \int_{\partial T}\big((F-F(o))-\mu(F-F(o))\big)^2 d\mu\\ &= \int_{\partial T}(F-F(o))^2 d\mu - \left(\int_{\partial T}(F-F(o))d\mu\right)^2\\ &\le [\mu]_o\cdot\sum_{\alpha\in E}\nabla F(\alpha)^2,\end{aligned}$$

hence, $[\![\mu]\!] \le \inf_{o\in T}[\mu]_o$.

In the other direction, consider closed subsets $A, B \subseteq \partial T$, which we might assume to be finite unions of arcs, and a function F with finitely supported $|\nabla F|$ such that $F = 1$ on A and $F = 0$ on B. Then,

$$\begin{aligned}[\![\mu]\!]\cdot\sum_{\alpha\in E}\nabla F(\alpha)^2 &\ge \int_{\partial T}\left|F-\mu(F)\right|^2 d\mu\\ &= \frac{1}{2}\int_{\partial T\times\partial T}\big(F(x)-F(y)\big)^2 d\mu(x)d\mu(y)\\ &\ge \mu(A)\mu(B).\end{aligned}$$

Passing to the infimum over all such F's, we obtain

$$\mu(A)\mu(B) \le [\![\mu]\!]\cdot\mathrm{Cap}(A,B),$$

giving

$$[\![\mu]\!] \ge \sup_{A,B\subseteq\partial T}\frac{\mu(A)\mu(B)}{\mathrm{Cap}(A,B)}.$$

It is not difficult to see that finiteness of $\sup_{A,B\subseteq\partial T}\frac{\mu(A)\mu(B)}{\mathrm{Cap}(A,B)}$ implies that μ does not have more than an atom on the boundary, which would then be a Dirac delta. Hence, if μ has boundary atoms, the statement holds.

Suppose now that μ is atomless. We claim that

(i) if μ is a probability measure on ∂T having no atoms, then there are disjoint arcs $I_1 \cup I_2 = \partial T$, such that $1/3 \le \mu(I_j) \le 2/3$.

With the claim given, let A be a finite union of arcs, let I_1, I_2 be as given by (i), and $A_1 = A \cap I_2$, $A_2 = A \cap I_1$. Let o_j be the pre-root of the dyadic tree T_1^j having boundary I_j and T_2^j, T_3^j the two other dyadic trees having it as a pre-root, so that, for $k \neq j$, $I_k = \partial T_2^j \cup \partial T_3^j$ and

$$[\![\mu]\!] \geq \sup_{A,B\subseteq\partial T} \frac{\mu(A)\mu(B)}{\mathrm{Cap}(A,B)} \geq \frac{\mu(A_j)\mu(I_j)}{\mathrm{Cap}(A_j,I_j)} \approx \frac{\mu(A_j)}{\mathrm{Cap}_{o_k}(A_j)}.$$

Observing that $\mathrm{Cap}_{o_k}(A_j) \leq \mathrm{Cap}_{o_j}(A_j)$, for $o \in \{o_1, o_2\}$ we have

$$\frac{\mu(A)}{\mathrm{Cap}_o(A)} \leq \frac{\mu(A_1)}{\mathrm{Cap}_{o_2}(A_1)} + \frac{\mu(A_2)}{\mathrm{Cap}_{o_1}(A_2)} \leq 2 \sup_{A,B\subseteq\partial T} \frac{\mu(A)\mu(B)}{\mathrm{Cap}(A,B)} \leq 2[\![\mu]\!].$$

By the isocapacitary condition (ISO) characterizing measures satisfying the Hardy inequality, $[\mu]_o = \sup\{\frac{\mu(A_j)}{\mathrm{Cap}_o(A_j)} : A \text{ finite union of arcs}\}$, and we have the promised statement.

We now come to the proof of the claim. Choose a vertex $x_0 \in T$ and for $j = 1, 2, 3$ let T_j^0 be the dyadic subtrees having pre-root at it and call x_j the neighborhood point of x_0 lying in T_j^0. For at least one j, $\mu(\partial T_j^0) \geq 1/3$; say $j = 1$. If $\mu(\partial T_1^0) \leq 2/3$, we set $I_1 = \partial T_1^0$ and we are done. Otherwise, let T_1^1, T_2^1 be the two infinite subtrees of T_1^0 with pre-root in x_1. For one $j \in \{1,2\}$, we have $\mu(\partial T_j^1) \geq 1/3$; let it again be $j = 1$. As before, set $I_1 = \partial T_1^1$ if $\mu(\partial T_1^1) \leq 2/3$; otherwise consider the two infinite subtrees of T_1^1 rooted at the neighborhoods of x_1, and iterate the reasoning. If there is no stop, we have a family of nested tents $\partial T_1^0 \supset \partial T_1^1 \supset \cdots$, whose intersection is a single boundary point x with $\mu(\{x\}) \geq 2/3$, contradicting the assumption. ∎

We come now to the case of measures supported on the vertex set of T. Since no extra work is required, we present a proof which holds in much higher generality.

Proposition 22. *Let X be a locally compact space and let B be a Banach space of functions on X such that for every compact K there is $C(K)$ with $\sup_{x\in K}|f(x)| \leq C(K)$ if $\|f\|_B \leq 1$. Then, the following are equivalent for a probability measure μ:*

(a) $\int |f - \mu(f)|^p d\mu \leq C_0 \|f\|_B^p$,

(b) $\int |f|^p d\mu \leq C_1 \|f\|_B^p$.

Proof. If (b) holds, then

$$\big\|f - \mu(f)\big\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)} + \big|\mu(f)\big| \leq 2C_1\|f\|_B.$$

Suppose that (a) holds and (b) does not. By (a),

$$\|f\|_{L^p(\mu)} \leq \big\|f - \mu(f)\big\|_{L^p(\mu)} + \big|\mu(f)\big| \leq C_0\|f\|_B + \big|\mu(f)\big|,$$

then there exists a sequence f_n in B with (i) $\|f_n\|_B = 1$, (ii) $\|f_n\|_{L^p(\mu)} \nearrow \infty$, (iii) $M_n := |\mu(f_n)| \geq 2^{-1}\|f_n\|_{L^p(\mu)}$.

We can find compact sets $K_n \nearrow X$ such that $\int_{K_n} |f_n| d\mu > \frac{M_n}{2}$. For any fixed compact S, we have that

$$\int_S |f_n| d\mu \leq C(S) \leq \frac{M_n}{4} \quad \text{for } n \geq n(S),$$

hence

$$\int_{K_n/S} |f_n| d\mu \geq \frac{M_n}{2} - \frac{M_n}{4} = \frac{M_n}{4},$$

for $n \geq n(S)$. Thus

$$\frac{M_n^p}{4^p} \leq \left(\int_{K_n/S} |f_n| d\mu\right)^p \leq \|f\|^p_{L^p(\mu)} \mu(X \setminus S)^{p/p^*};$$

i.e.,

$$\frac{|\mu(f_n)|}{\|f\|_{L^p(\mu)}} \leq 4\mu(X \setminus S)^{1/p^*},$$

which can be made as small as we wish, leading to a contradiction. ■

PROBLEM 3. A more general problem is that of characterizing the probability measures λ on $T \times T$ such that

(a) $$\int_T \int_T \left|F(x) - F(y)\right|^2 d\lambda(x, y) \leq \{\lambda\} \sum_\alpha \left|\nabla F(\alpha)\right|^2,$$

which is as well conformally invariant if we set

$$\Psi_* \lambda(A \times B) = \lambda\big(\Psi(a) \times \Psi(B)\big)$$

and reduces to the above for $\lambda = \mu \otimes \mu$:

$$\begin{aligned}
\int_T \left|F(x) - \int_T F(y) d\mu(y)\right|^2 d\mu(x) &= \int_T \left|\int_T \big(F(x) - F(y)\big) d\mu(y)\right|^2 d\mu(x) \\
&\leq \int_T \int_T \left|F(x) - F(y)\right|^2 d\mu(y) d\mu(x) \\
&= 2\left[\int_T F(x)^2 d\mu(x) - \left(\int_T F(x) d\mu(x)\right)^2\right] \\
&= 2\left[\int_T \left|F(x) - \int_T F(y) d\mu(y)\right|^2 d\mu(x)\right].
\end{aligned}$$

No characterization of the measures λ for which (a) holds is known, to the best of our knowledge.

PROBLEM 4. A natural generalization of the conformal Hardy inequality is to consider a p-version, or even a weighted version of it. Then, given a weight function $\pi : E \to \mathbb{R}_+$, an interesting problem would be that of characterizing all positive Borel probability measures such that there exists a constant $[\mu]_{\text{inv}}$, possibly depending on μ and π, for which

$$\int_{\overline{T}} \left|F(x) - \mu(F)\right|^p d\mu(x) \leq [\mu]_{\text{inv}} \sum_{\alpha} \left|\nabla F(\alpha)\right|^p \pi(\alpha),$$

for all $F : \overline{T} \to \mathbb{R}$.

There is a related interesting, conformally invariant interpolation problem.

PROBLEM 5. Set $\|F\|^2 = \sum_{\alpha \in E} \nabla F(\alpha)^2$. Then,

$$\left|F(x) - F(y)\right| \leq \|F\| \cdot d(x, y)^{1/2}.$$

Given a subset $Z \subset T$, we say that it is *universally interpolating* for the seminorm $\|\cdot\|$ if

(i) $\sum_{z \in T} \sum_{w \in T} \frac{|F(z)-F(w)|^2}{d(z,w)} \leq C\|F\|^2$,

(ii) for all sequences $\{a(z)\}_{z \in Z}$ such that

$$\sum_{z \in T} \sum_{w \in T} \frac{\left|a(z) - a(w)\right|^2}{d(z, w)} < \infty,$$

there exists F with $\|F\| < \infty$ such that $F(z) = a(z)$, $\forall z \in Z$.

Condition (i) says that the measure $\lambda = \sum_{z,w \in T} \frac{\delta_{(z,w)}}{d(z,w)}$ satisfies (a). We call the sequence *onto interpolating* if just (ii) holds. We think that it is an interesting problem finding characterizations of universally, or onto, interpolating sequences. For background on interpolating sequences for Dirichlet-type space see [27–29].

7. Miscellaneous results

7.1. *Compactness*

We will briefly discuss the compactness conditions for the Hardy operator. As it is natural to expect, the compactness of the Hardy operator corresponds to the "vanishing" versions of the conditions characterizing boundedness.

THEOREM 23. *The map $\mathcal{I} : \ell^p(\pi) \to L^p(\mu)$ is compact if and only if*

$$\lim_{n\to\infty} \sup_{d(\alpha)\geq n} \frac{\sum_{\beta \leq \alpha} \pi(\beta)^{1-p^*} \mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\alpha)\big)} = 0. \tag{7.1}$$

PROOF. Suppose that (7.1) holds, fix $n \geq 1$, and consider the finite rank truncation of $\mathcal{I}_\mu^*$, $\mathcal{I}_{\mu,n}^* g(\alpha) = \chi(d(\alpha) \leq n)\mathcal{I}_\mu^* g(\alpha)$. Then,

$$\begin{aligned}
&\|\mathcal{I}_\mu^* g - \mathcal{I}_{\mu,n}^* g\|_{\ell^{p^*}(\pi^{1-p^*})}^{p^*} \\
&\quad = \sum_{d(\alpha)=n} \sum_{\beta \subseteq \alpha} \left|\mathcal{I}_\mu^* g(\beta)\right|^{p^*} \pi(\beta)^{1-p^*} \\
&\quad \leq C \sum_{d(\alpha)=n} \sup_{\gamma \subseteq \alpha} \frac{\sum_{\beta \subseteq \gamma} \pi^{1-p^*}(\beta)\mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\gamma)\big)} \|g\chi_{S(\alpha)}\|_{\ell^{p^*}(\mu)}^{p^*} \\
&\quad \leq C \sup_{d(\alpha)\geq n} \frac{\sum_{\beta \subseteq \alpha} \pi^{1-p^*}(\beta)\mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\alpha)\big)} \|g\|_{\ell^{p^*}(\mu)}^{p^*} \to 0.
\end{aligned}$$

Therefore, $\mathcal{I}_\mu^*$ is compact as an operator norm limit of finite rank operators. By Schauder's theorem [46, Theorem 7, p. 243], $\mathcal{I}$ is also compact.

To see the necessity of the condition, we can again work with $\mathcal{I}_\mu^*$ courtesy of Schauder's theorem. Suppose that (7.1) does not hold. Then, there exists a $\varepsilon > 0$ and a sequence of edges $\{\alpha_k\}_k$ such that $\lim_k d(\alpha_k) = \infty$ and

$$\frac{\sum_{\beta \subseteq \alpha_k} \pi^{1-p^*}(\beta)\mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\alpha_k)\big)} \geq \varepsilon.$$

Then, consider the sequence of testing functions $g_k := \mu(S(\alpha_k))^{-1/p^*}\chi_{S(\alpha_k)}$, which converges weakly to 0 in the space $L^{p^*}(\mu)$. By compactness, we must have

$$0 = \lim_k \|\mathcal{I}_\mu^* g_k\|_{\ell^{p^*}(\pi^{1-p^*})}^{p^*} = \frac{\sum_{\beta \subseteq \alpha_k} \pi^{1-p^*}(\beta)\mu\big(S(\beta)\big)^{p^*}}{\mu\big(S(\alpha_k)\big)},$$

which contradicts the fact that the above quantity is bounded below by ε. ■

In a very similar way, one can characterize the compactness of the Hardy operator in terms of a *vanishing capacitary condition*. We state the theorem without a proof in order to avoid repetition.

THEOREM 24. *The map $\mathcal{I} : \ell^p \to L^p(\mu)$ is compact if and only if*

$$\lim_n \sup_{d(\alpha_i)\geq n} \mathrm{Cap}_{p,\pi}\left(\bigcup_{i=1}^k S(\alpha_i)\right)^{-1} \sum_{i=1}^k \mu\big(S(\alpha_i)\big) = 0.$$

PROBLEM 6. Find a simple characterization of measures μ such that the Hardy operator $\mathcal{I} : \ell^2(\pi) \to L^2(\mu)$ belongs to the p-Schatten ideal.

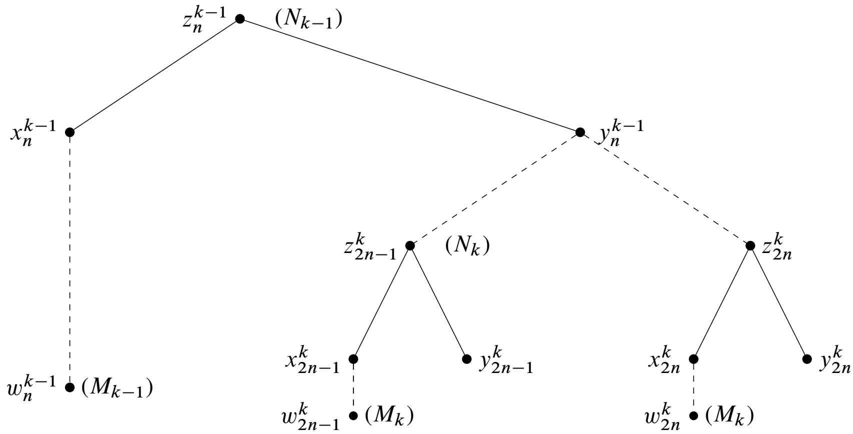

FIGURE 2. A snapshot from the defined family of points and their relations: continuous lines represent edges and dashed lines paths of (many) edges. Next to some vertices is specified, in parenthesis, their distance from the origin.

In [49], Luecking has studied trace ideal criteria for Toeplitz operators on weighted Bergman spaces. It is very possible that some of his results apply also to our case, although his results are not complete.

7.2. *No sufficiency of the simple box condition*

A necessary condition for the Hardy inequality to hold is the simple, one-box condition

$$\sup_{\alpha\in E} \mu\big(S(\alpha)\big) d_\pi\big(e(\alpha)\big)^{p-1} := [\![\mu]\!]_{sc} < +\infty,$$

where $d_\pi(x) = \sum_{\alpha\in[o^*,x]} \pi(\alpha)^{1-p^*}$. In fact, if $\varphi = \chi_{[o^*,x]}\pi^{1-p^*}$, then $\int_{\overline{T}} \mathcal{I}\varphi^p d\mu \geq \mu(S(\alpha))d_\pi(e(\alpha))^p$, while $\|\varphi\|^p_{L^p(\pi)} = d_\pi(e(\alpha))$.

In general, however, (SB) is not sufficient. We provide here a counterexample for $p = 2$, $\pi = 1$, and T the dyadic tree. Working a bit, the example can be modified to hold for $1 < p < \infty$.

EXAMPLE 25 ((SB) $\not\Longrightarrow$ (H)). Let T be a dyadic tree, $\pi \equiv 1$, and $p = 2$. At level $N_k = 2^k k$, choose 2^k vertices $\{z_n^k\}_{n=1}^{2^k}$, with children named x_n^k and y_n^k, such that $z_{2n-1}^k \wedge z_{2n}^k = y_n^{k-1}$. Choose points $w_n^k \in S(x_n^k)$ with $d(w_n^k) = M_k = 2^k k^2$. This configuration of points is described in Figure 2.

Let $\mu = \sum_{k,n} \frac{1}{M_k}\delta_{w_n^k}$. A simple reasoning shows that it suffices to verify the one-box condition at the nodes z_n^k:

$$\mu\big(S(z_n^k)\big) = \sum_{j=0}^{\infty} \frac{2^j}{M_{j+k}} = 2^{-k}\sum_{j=0}^{\infty} \frac{1}{(j+k)^2} \approx \frac{1}{N_k} = d(z_n^k)^{-1}.$$

On the other hand, the mass-energy condition (ME) fails. To see this, denote by Z the minimal subtree containing all the z_n^k points. Then,

$$\sum_{x\in T}\mu\big(S(x)\big)^2 \geq \sum_{x\in Z}\mu\big(S(x)\big)^2 = \sum_{k=0}^{\infty}(N_{k+1}-N_k)2^k\left(\sum_{j=0}^{\infty}\frac{1}{(j+k)^2}\right)^2$$
$$\approx \sum_{k=0}^{\infty}\left(\frac{N_{k+1}}{N_k^2}-\frac{1}{N_k}\right)2^k \approx \sum_{k=0}^{\infty}\frac{1}{k} = +\infty.$$

PROBLEM 7. Find a characterization of those couples (p,π) for which (SB) is not sufficient on the dyadic tree.

7.3. Two opposite examples

In the generality in which we have stated it, the dyadic Hardy inequality covers a variety of contexts; some of them very rich, some very poor. The richest context is in our opinion the unweighted case $\pi = 1$, corresponding to the classical Dirichlet space. We consider here two cases at the opposite extremes.

EXAMPLE 26 (Boundary having null capacity). Consider the dyadic, infinite tree with the weight $\pi(\alpha) = 2^{-d(\alpha)}$ and $p = 2$. The reader familiar with martingale theory can fruitfully think of π as the probability of a fair coin tossed $d(\alpha)$ times. Let us show that with this choice $\mathrm{Cap}_{2,\pi}(\partial T) = 0$. Let $g_N(\alpha) = \frac{1}{N}$ if $d(\alpha) \leq N$ and $g_N(\alpha) = 0$ elsewhere. Then,

$$\mathrm{Cap}_{2,\pi}(\partial T) = \inf\left\{\|f\|^2_{\ell^2(\pi)} : f \geq 0,\ \mathcal{I}f \geq 1 \text{ on } \partial T\right\} \leq \|g_N\|^2_{\ell^2(\pi)}$$
$$= \sum_{n=0}^{N} 2^n\frac{1}{N^2}2^{-n} = \frac{1}{N} \to 0, \quad \text{as } N \to +\infty.$$

It follows from the isocapacitary inequality (ISO) that ∂T does not support any Carleson measure $p = 2$ and the chosen weight π. In particular, the Lebesgue measure $\mu_0(\partial S(\alpha)) = 2^{-d(\alpha)}$ does not define a Carleson measure. This fact is best appreciated having in mind basic martingale theory.

Consider the filtration associated with the infinite tossing of a fair coin, where ∂T is the probability space and $\mu_0(\partial S(\alpha)) = 2^{-d(\alpha)}$ is the probability measure. A martingale for the filtration has the form $X_n(\zeta) = \mathcal{I}f(\alpha)$ for $\zeta \in \partial S(\alpha)$ and $d(\alpha) = n$. The martingale Hardy space $\mathcal{M}^2$ contains those martingales for which

$$\|X\|^2_{\mathcal{M}^2} = \sup_n \int_{\partial T} X_n^2 d\mu_0 = \int_{\partial T}\mathcal{I}f^2 d\mu_0 = \sum_\alpha f(\alpha)^2\pi(\alpha) < \infty;$$

i.e., $\mathcal{M}^2 = \ell^2(\pi) \cap \mathcal{M}$, where $\mathcal{M}$ is the space of all martingales. By definition, μ_0 is Carleson measure for $\mathcal{M}^2$, but it is not for $\ell^2(\pi)$. The underlying reason is that in $\mathcal{M}^2$ cancelations play a prominent role, and this much enlarges the class of the Carleson measures at the boundary.

But suppose we ask a measure μ to be Carleson for the variation of the martingales in $\mathcal{M}^2$,

$$\int_{\partial T} \Big(\sum_{\alpha \in P(\zeta)} \big| f(\alpha) \big| \Big)^2 d\mu(\zeta) \le C \sum_{\alpha} f(\alpha)^2 \pi(\alpha),$$

where the variation of $X = \mathcal{I} f$ is $V(X)(\zeta) = \sum_{\alpha \in P(\zeta)} |f(\alpha)|$. It is easy to see that this is the same as asking μ to be Carleson for $\ell^2(\pi)$, hence $\mu = 0$. This is reflected in the fact that functions in the classical Hardy space can have unbounded variation a.e. at the boundary of the unit disc [21, 59].

Let us note that, in contrast, for the weights $\pi_\lambda(\alpha) = 2^{\lambda d(\alpha)}$, $0 \le \lambda < 1$, the Carleson measures for $\ell^2(\pi_\lambda) \cap \mathcal{M}$ and $\ell^2(\pi_\lambda)$ are the same [9]. This is in much the spirit of Beurling's result on exceptional sets [20].

Example 27 (All boundary points have positive capacity). It is easy to see that for any tree T and any $\zeta \in \partial T$, $\mathrm{Cap}_p^\sigma(\{\zeta\}) = d_\sigma(\zeta)^{-1}$. In fact, for any function f which is admissible for ζ,

$$\begin{aligned} \|f\|_{\ell^p(\sigma)}^p &\ge \sum_{\alpha \supset \zeta} \big| f(\alpha) \big|^p \sigma(\alpha) \\ &\ge \Big(\sum_{\alpha \supset \zeta} \frac{\big| f(\alpha) \big|}{d_\sigma(\zeta)^{1/p^*}} \sigma(\alpha) \Big)^p \\ &= \frac{\big(I_\sigma f(\zeta) \big)^p}{d_\sigma(\zeta)^{p-1}} \ge d_\sigma(\zeta)^{1-p}, \end{aligned}$$

and the right-hand side is the $\ell^p(\sigma)$-norm of the admissible function taking constant value $d_\sigma(\zeta)^{-1}$ on the edges $\alpha \supset \zeta$ and zero elsewhere. Then, assuming that $\mathrm{Cap}_p^\sigma(\{\zeta\}) \ge c > 0$ for all boundary points is the same as saying that ∂T is bounded with respect to the distance d_σ. This is the case, for example, for the weights $\sigma(\alpha) = 2^{\lambda d(\alpha)}$ with $\lambda > 0$. Under this assumption, all functions $\mathcal{I} f$ with $f \in \ell^p(\pi)$ are bounded on ∂T:

$$\begin{aligned} \big| \mathcal{I} f(\zeta) \big| &\le \Big(\sum_{\alpha \in [o^*, \zeta]} \big| f(\alpha) \big|^p \pi(\alpha) \Big)^{1/p} \Big(\sum_{\alpha \in [o^*, \zeta]} \sigma(\alpha) \Big)^{1/p^*} \\ &\le \|f\|_{\ell^p(\pi)} \left(\frac{2}{c} \right)^{1/p^*}, \end{aligned}$$

from which follows that all bounded measures μ are Carleson for $\ell^p(\pi)$.

Under this assumption, all functions $\mathcal{I} f$ with $f \in \ell^p(\pi)$, where as usual $\pi = \sigma^{1-p}$, are continuous up to the boundary, with respect to the natural topology induced by the visual distance $d(\zeta, \xi) = e^{-d(\zeta \wedge \xi)}$:

$$\begin{aligned}
\left|\mathcal{I} f(\zeta) - \mathcal{I} f(\xi)\right| &\leq \sum_{\alpha \in [\zeta, \xi]} \left|f(\alpha)\right| \\
&\leq \Big(\sum_{\alpha \in S(\zeta \wedge \xi)} \left|f(\alpha)\right|^p \pi(\alpha)\Big)^{1/p} \Big(\sum_{\alpha \in [\zeta, \xi]} \sigma(\alpha)\Big)^{1/p^*} \\
&\leq \Big(\sum_{\alpha \in S(\zeta \wedge \xi)} \left|f(\alpha)\right|^p \pi(\alpha)\Big)^{1/p} \Big(\frac{2}{c}\Big)^{1/p^*},
\end{aligned}$$

which tends to zero as $\xi \to \zeta$ by dominated convergence. By Weierstrass theorem, all bounded measures μ are Carleson for $\ell^p(\pi)$.

8. Some variations on the structure

8.1. *The viewpoint of reproducing kernels*

Let us recall to the reader that a *reproducing kernel Hilbert space (RKHS)* is a Hilbert space $\mathcal{H}$ of functions defined on a set X such that the point evaluation functionals are continuous or, equivalently, such that for any $x \in X$ there exists an element $K_x \in \mathcal{H}$ which fulfills the reproducing property

$$f(x) = \langle f, K_x \rangle_{\mathcal{H}}, \quad \text{for all } f \in \mathcal{H}.$$

It is easy to see that the function $K : X \times X \to \mathbb{C}$ given by

$$K(y, x) := K_x(y) = \langle K_x, K_y \rangle_{\mathcal{H}}$$

is a kernel on $\mathcal{H}$; that is, it is positive semi-definite, nonzero on the diagonal and satisfies $K(x, y) = \overline{K(y, x)}$. See [17] or [3] for a systematic treatment of the topic.

Let H_K be a RKHS of functions on a locally compact space X with continuous reproducing kernel K. A simple and well-known "T^*T argument" shows that the imbedding $\mathrm{Id} : H_K \to L^2(\mu)$ (i.e., μ, a positive Borel measure on X, is a Carleson measure for H_K) can be rephrased various ways in terms of integral inequality on $L^2(\mu)$.

Lemma 28. *Given a RKHS H_K of functions on a locally compact space X with continuous reproducing kernel $K(x, y) = K_y(x)$, the following are equivalent for a Borel measure μ on X:*

(i) $\|f\|^2_{L^2(\mu)} \leq \{\mu\}_1 \|f\|^2_{H_K}$; *i.e.,* $\|\mathrm{Id}^* g\|^2_{H_K} \leq \{\mu\}_1 \|g\|^2_{L^2(\mu)}$,

(ii) $\langle \mathrm{Id}^* g, \mathrm{Id}^* g\rangle_{H_K} = \int_X \int_X g(x)\overline{g(y)} K_y(x) d\mu(x) d\mu(y) \leq \{\mu\}_1 \|g\|^2_{L^2(\mu)}$,

(iii) $\int_X \int_X g(x)\overline{g(y)} \operatorname{Re} K_y(x) d\mu(x) d\mu(y) \leq \{\mu\}_2 \|g\|^2_{L^2(\mu)}$,

(iv) $\int_X \int_X g(x) g(y) \operatorname{Re} K_y(x) d\mu(x) d\mu(y) \leq \{\mu\}_2 \|g\|^2_{L^2(\mu)}$, *for real (or even just positive) g,*

(v) $\|f\|^2_{L^2(\mu)} \leq \{\mu\}_1 \|F\|^2_{H_{\operatorname{Re} K}}$,

(vi) $|\langle \mathrm{Id}^* g, \psi\rangle_{L^2(\mu)}| = |\int_X \int_X g(x)\overline{\psi(y)} K_y(x) d\mu(x) d\mu(y)| \leq \{\mu\}_3 \|g\|_{L^2(\mu)} \|\psi\|_{L^2(\mu)}$,

(vii) $\|\mathrm{Id}^* g\|^2_{L^2(\mu)} \leq \{\mu\}_3 \|g\|^2_{L^2(\mu)}$,

where the statements are assumed to hold for all $f, g : X \to \mathbb{C}$, and the constants $\{\mu\}_1, \{\mu\}_2, \{\mu\}_3$, which are the best constants in the respective inequalities, might depend on μ and π but not on f, g. In particular, μ is Carleson measure for H_K if and only if it is a Carleson measure for $H_{\operatorname{Re} K}$.

The proof of Lemma 28 can be found with some variations in different sources: [16, Proposition 4.9], [13, Lemma 24], [14, pp. 9–10]. The proof itself is a short "soft" analysis argument, and we write it here for convenience of the reader. The statement we give is an elaboration of various statements in the literature.

Proof of Lemma 28. (i) says that $\mathrm{Id} : H_K \to L^2(\mu)$ is bounded with norm $\{\mu\}_1^{1/2}$, so $\mathrm{Id}^* : L^2(\mu) \to H_K$ is bounded with the same norm, where

$$\mathrm{Id}^* g(x) = \langle \mathrm{Id}^* g, K_x\rangle_{H_K} = \langle g, K_x\rangle_{L^2(\mu)} = \int_X g(y) K_y(x) d\mu(y),$$

where in the last equality we used $K_x(y) = \overline{K_y(x)}$. Hence,

$$\begin{aligned}\|g\|^2_{L^2(\mu)}\{\mu\}_1 &\geq \langle \mathrm{Id}^* g, \mathrm{Id}^* g\rangle_{H_K} = \langle g, \mathrm{Id}^* g\rangle_{L^2(\mu)} \\ &= \int_X \int_X g(x)\overline{g(y)} K_y(x) d\mu(x) d\mu(y),\end{aligned}$$

showing that (i) and (ii) are equivalent.

Testing (ii) on real g's, we have that (iv) holds with $\{\mu\}_2 = \{\mu\}_1$. If viceversa (iv) holds and $g = g_R + i g_I$ is the decomposition of g in real and imaginary parts, then

$$\begin{aligned}\|I^* g\|^2_{H_K} &\leq \left(\|I^* g_R\|_{H_K} + \|I^* g_I\|_{H_K}\right)^2 \\ &\leq 2\left(\|I^* g_R\|^2_{H_K} + \|I^* g_I\|^2_{H_K}\right) \\ &\leq 2\{\mu\}_2 \|g\|^2_{L^2(\mu)},\end{aligned}$$

and we obtain the dual form of (i) with $\{\mu\}_1 \leq 2\{\mu\}_2$. A similar reasoning works if we assume (iv) to hold for positive g's, but with a different constant.

It is clear that (vi) and (vii) are equivalent, and (vi) implies (ii) with $\{\mu\}_1 \leq \{\mu\}_3$,

$$\langle \mathrm{Id}^* g, \mathrm{Id}^* g\rangle_{H_K} = \langle g, \mathrm{Id}^* g\rangle_{L^2(\mu)} \leq \{\mu\}_3 \|g\|^2_{L^2(\mu)}.$$

In the other direction, set $Tg(x) = \int_X K_x(y)g(y)d\mu(y) = \mathrm{Id}^* g(x)$, where T is defined on $L^2(\mu)$, $T = T^*$ and T is positive, $\langle Tg, g\rangle_{L^2(\mu)} \geq 0$. Then, it has a positive square root $\sqrt{T}$ and by (ii)

$$\|\sqrt{T} g\|^2_{L^2(\mu)} = \langle Tg, g\rangle_{L^2(\mu)} \leq \{\mu\}_1 \|g\|^2_{L^2(\mu)},$$

hence (vii) holds with $\{\mu\}_3 = \{\mu\}_1$.

Since $\mathrm{Re}\, K$ is positive definite, hence a reproducing kernel, (iii), and (v) are equivalent for the same reason (i) and (ii) are. Finally, (iii) clearly implies (iv) and, on the other hand, if (iv) holds, then

$$\begin{aligned}
&\int_X \int_X g(x)\overline{g(y)}\, \mathrm{Re}\, K_y(x) d\mu(x) d\mu(y) \\
&\quad = \int_X \int_X g_R(x) g_R(y)\, \mathrm{Re}\, K_y(x) d\mu(x) d\mu(y) \\
&\qquad + \int_X \int_X g_I(x) g_I(y)\, \mathrm{Re}\, K_y(x) d\mu(x) d\mu(y) \\
&\quad \leq \{\mu\}_2 \big[\|g_R\|^2_{L^2(\mu)} + \|g_I\|^2_{L^2(\mu)}\big] = \{\mu\}_2 \|g\|^2_{L^2(\mu)}.
\end{aligned}$$

■

This lemma allows one to use methods from singular integral theory (where $\mathrm{Re}\, K$ is the kernel of the “singular” integral) on nonhomogeneous spaces.

This point of view, at least in dyadic theory, started in [13] in connection to the problem of Carleson measures for the Drury–Arveson space, which we will mention below, then it is taken up by Tchoundja [67] and Volberg–Wick [71] in order to study more general function spaces on the unit ball.

What we aim here, instead, is to give an interpretation of the conformal invariant inequality (CH) in terms of reproducing kernels. Indeed, we will provide such an interpretation for the even more general inequality (a) introduced in Problem 3. The idea is to read (a) as the imbedding inequality of an appropriate RKHS into $L^2(\mu \otimes \mu)$, where μ is a Borel measure on the rooted tree T. Lemma 28 would then provide various ways to reformulate inequality (a).

Let T be the rooted tree and π an arbitrary positive weight function. We first introduce the *Dirichlet space* $\mathcal{D}(\pi) := \{F = \mathcal{I} f : f \in \ell^2(\pi)\}$, endowed with the inner product $\langle F, G\rangle_{\mathcal{D}(\pi)} = \langle f, g\rangle_{\ell^2(\pi)} = \langle \nabla F, \nabla G\rangle_{\ell^2(\pi)}$, for $F = \mathcal{I} f$, $G = \mathcal{I} g \in \mathcal{D}(\pi)$. We claim that this (semi-)Hilbert space has a reproducing kernel,

$$K_x(y) = d_\pi(x \wedge y) := \sum_{\alpha \in [o^*, x \wedge y]} \pi(\alpha)^{-1}.$$

Indeed, for $f \in \ell^2(\pi)$ and $F = \mathcal{I} f$ we have

$$\begin{aligned} F(x) &= \sum_{\alpha\in[o^*,x]} f(\alpha)\pi^{-1}(\alpha)\pi(\alpha) \\ &= \sum_{\alpha\in E} f(\alpha)\chi_{[o^*,x]}(\alpha)\pi^{-1}(\alpha)\pi(\alpha) = \langle F, \chi_{[o^*,x]}\pi^{-1}\rangle_{\mathcal{D}(\pi)}, \end{aligned}$$

from which it follows that

$$K_x(y) = \mathcal{I}\chi_{[o^*,x]}\pi^{-1}(y) = \sum_{\alpha\in[o^*,y]} \pi^{-1}(\alpha)\chi\big(\alpha \in [o^*, x]\big) = d_\pi(x \wedge y).$$

It is then imprecise but harmless to say that $\mathcal{D}(\pi)$ is a RKHS. The inequality we are re-interpreting as a RKHS imbedding is (a) for $\lambda = \mu \otimes \mu$; i.e.,

$$\text{(a}'\text{)} \qquad \int_T\int_T \left|F(x) - F(y)\right|^2 d\mu(x)d\mu(y) \le \{\mu\} \sum_{\alpha\in E} \left|\nabla F(\alpha)\right|^2,$$

We are not quite done yet, since this inequality bounds the $L^2(\mu \otimes \mu)$ norm of the differences of a function with the ($\pi = 1$) Dirichlet (semi-)norm of the function itself. However, we argue here that, given a RKHS H_K on a set X such that $1 \in H_K$ (assume $\|1\|_{H_K} = 1$), there is a canonical way to construct the RKHS of its differences, having as elements the functions $(x, y) \mapsto F(x) - F(y) =: \nabla F(x, y)$, with $F \in H_K$. We define the kernel $\kappa : (X \times X) \times (X \times X) \to \mathbb{C}$ as

$$\text{(8.1)} \qquad \begin{aligned} \kappa_{(a,b)}(x, y) = \kappa\big((x, y), (a, b)\big) &:= \langle K_a - K_b, K_x - K_y\rangle \\ &= K(x, a) - K(y, a) - K(x, b) + K(y, b) \\ &= \nabla K_a(x, y) - \nabla K_b(x, y) \\ &= \nabla(K_a - K_b)(x, y). \end{aligned}$$

We show that κ reproduces the space H_∇ of the functions ∇F, endowed with the inner product

$$\text{(8.2)} \qquad \langle\nabla F, \nabla G\rangle_{H_\nabla} := \big\langle F - 1\langle F, 1\rangle_{H_K}, G - 1\langle G, 1\rangle_{H_K}\big\rangle_{H_K},$$

which is well defined since $\nabla F = 0$ if and only if F is constant, i.e., if and only if $F - 1\langle F, 1\rangle_{H_K} = 0$.

Lemma 29. *H_∇ with the inner product (8.2) is a RKHS with kernel κ given by (8.1).*

Proof. Since $\langle K_a - K_b, 1\rangle_{H_K} = 0$, we have

$$\begin{aligned} \langle\nabla F, \kappa_{(a,b)}\rangle_{H_\nabla} &= \big\langle F - 1\langle F, 1\rangle_{H_K}, K_a - K_b\big\rangle_{H_K} \\ &= F(a) - F(b) \\ &= \nabla F(a, b). \end{aligned}$$

■

Lemma 29, in particular, tells us that the space $\mathcal{D}_\nabla$ of differences of functions in the Dirichlet space $\mathcal{D} = \mathcal{D}(1)$ has reproducing kernel

$$k_{(a,b)}(x, y) = d(x \wedge a) - d(x \wedge b) - d(y \wedge a) + d(y \wedge b).$$

Inequality (a′) represents then the boundedness of the imbedding $\mathcal{D}_\nabla \to L^2(\mu \otimes \mu)$, which by means of Lemma 28 admits various re-writings.

A picture shows that the definition of the kernel of $\mathcal{D}_\nabla$ is independent of the choice of the root, as we know a priori by conformal invariance.

For many classical spaces of holomorphic functions, as far as it concerns their imbedding properties, one can substitute the reproducing kernel with its absolute value causing no losses. It is a natural question if the same applies here.

Problem 8. Is (a′) equivalent to the imbedding in $L^2(\mu \otimes \mu)$ of the space having $|k_{(a,b)}(x, y)|$ as kernel?

As a comment to the above problem, we observe that the kernel $k_{(a,b)}(x, y)$ seems to present important cancelations, which might be an indication that tools from singular integral theory are needed in the characterization of the conformally invariant Hardy inequality. Indeed, it is a simple exercise to check that for $(a, b), (x, y) \in T \times T$ and $[p, q] := [a, b] \cap [x, y]$, it holds that

$$k_{(a,b)}(x, y) = \begin{cases} +d(p, q) & \text{if } a \text{ and } x \text{ (hence, } b \text{ and } y\text{) can be joined in } T \setminus [p, q], \\ -d(p, q) & \text{if } a \text{ and } x \text{ (hence, } b \text{ and } y\text{) cannot be joined in } T \setminus [p, q]. \end{cases}$$

8.2. *Quotient structures*

Dyadic quotient structures appeared for the first time in [13], to the best of our knowledge, to deal with the problem of the Carleson measures for the Drury–Arveson space. Using the T^*T argument outlined in Section 8.1, the problem was shown to be equivalent to the immersion $\mathrm{Id} : H_K \to L^2(\mu)$ for a tree and a kernel which we are going to describe in a special case containing all essential information.

Consider a 4-adic, rooted tree T, whose vertices x at level $d(o^*, x) = n$ might be labeled as 4-adic rationals $x = 0.t_1 \dots t_n$, with $t_j \in \mathbb{Z}/4\mathbb{Z}$ and an edge joining the parent $0.t_1 \dots t_{n-1}$ with the child $0.t_1 \dots t_{n-1}t_n$. Define similarly the dyadic tree U and consider the surjective map $\Phi : T \to U$ induced by the map $[t]_{\mathrm{mod}\, 4} \mapsto [t]_{\mathrm{mod}\, 2}$, sending digits 0, 2 to binary digit 0, and digits 1, 3 to binary digit 1.

The map Φ is a root-preserving *tree epimorphism*: it is surjective, and $\Phi(x)$ and $\Phi(y)$ are joined by an edge in U if and only if x and y are joined by an edge in T. In other words, we have defined a quotient structure $U = T/\Phi$ on T.

We define a kernel $K_{\mathcal{g}}$ on T by

$$K_{\mathcal{g}}(x,y) = \frac{2^{-d([o^*],[x]\wedge_{\mathcal{g}}[y])}}{2^{-2d(o^*,x\wedge_{\mathcal{g}} y)}},$$

which can be proved to be positive definite, hence defining a RKHS $H_{K_{\mathcal{g}}}$. The wedge $\wedge_{\mathcal{g}}$ is a modified version of the wedge we have used so far. For the exact definition, the reader is referred to [13].

The following theorem is proved in [13, 67].

THEOREM 30. *The following are equivalent for a measure* $\mu \geq 0$ *on* T*:*

(i) *the map* $\mathrm{Id} : H_{K_{\mathcal{g}}} \to L^2(\mu)$ *is bounded,*

(ii) *we have both the simple condition* $\mu(S(\alpha)) \leq C_0 2^{-d(e(\alpha),o^*)}$ *and the inequality*

$$\int_{S(\alpha)} \left(\int_{S(\alpha)} K_{\mathcal{g}}(x,y)d\mu(y) \right)^s d\mu(x) \leq C_p \mu\big(S(\alpha)\big),$$

for one, or equivalently for all, $1 \leq s < \infty$.

It is not clear if one needs to introduce the modified wedge $\wedge_{\mathcal{g}}$ in order for the above theorem to hold. Thus the following problem remains open.

PROBLEM 9. Is it true that Theorem 30 remains true if we replace the kernel $K_{\mathcal{g}}$ with the kernel

$$K(x,y) := \frac{2^{-d([o^*],[x]\wedge[y])}}{2^{-2d(o^*,x\wedge y)}},$$

where $\wedge$ is the standard tree wedge?

The real part of the reproducing kernel of the Drury–Arveson space can be naturally written down as the quotient of two kernels which reflect this stratification. Passing to dyadic decompositions, this leads to the kernels $K_{\mathcal{g}}$ and K we have just described, and the Carleson measure problem for the Drury–Arveson space can be reduced to the theorem stated above.

We have seen that conditions similar to those in the theorem also provide alternative characterizations of the measures μ satisfying the Hardy inequality, at least when $p = 2$. We think that there are here some interesting questions for further investigation.

PROBLEM 10. Is it possible to have a characterization of the Carleson measures for H_K in terms of the potential theory associated with the kernel K?

8.3. Product structures: poly-trees

The dyadic tree T parametrizes the set of the dyadic subintervals of $[0, 1]$, and the corresponding product structure T^d is defined to parametrize Cartesian products

$R = I_1 \times \cdots \times I_d$ of such intervals: dyadic rectangles for $d = 2$, etcetera. Following the same lines of Section 2.1, a potential theory can now be defined on T^d by taking tensor products of everything on sight, as we will detail below. This leads to a natural extension of the Hardy inequality to the multi-parameter setting. In this situation, however, characterizing trace measures is a much more complicated problem. We remark that the poly-tree is not a tree, but a graph presenting cycles, and this creates new and major difficulties. So far, solutions to the problem are known for $\sigma \equiv 1$, $p = 2$ and for dimension $d = 2, 3$ only [5, 8, 53]. It is also known [52] that the techniques used in these works are not feasible to be extended to $d = 4$ and $p \neq 2$. Let us briefly expand on that.

We identify T with its vertex set, $T^d \ni x = (x_1, \ldots, x_d)$, and denote by E^d the edge set of T^d, $E^d \ni \alpha = (\alpha_1, \ldots, \alpha_d)$. Let $k : \overline{T} \times E \to \mathbb{R}_+$ be the kernel defined in Section 2.1. We define the kernel $\mathbf{k} : \overline{T}^d \times E^d \to \{0, 1\}$,

$$\mathbf{k}(x, \alpha) = \chi_{\{\alpha \supset x\}}(x, \alpha) = \Pi_{j=1}^d k(x_j, \alpha_j).$$

Let $\sigma : E^d \to \mathbb{R}_+$ be a positive weight. For a function $\varphi : E^d \to \mathbb{R}_+$ we set $\mathbf{I}_\sigma = \mathcal{I}_\sigma \otimes \cdots \otimes \mathcal{I}_\sigma$; i.e.,

$$\mathbf{I}_\sigma \varphi(x) = \sum_{\alpha \in E^d} \mathbf{k}(x, \alpha) \varphi(\alpha) \sigma(\alpha) = \sum_{E^d \ni \alpha \supset x} \varphi(\alpha) \sigma(\alpha),$$

and for $\mu \geqslant 0$ on $\overline{T}^d$,

$$\mathbf{I}^* \mu(\alpha) := \mu\big(S(\alpha)\big), \quad \alpha \in E^d.$$

The d-parameter, weighted Hardy inequality for such product structure[4] is

$$\int_{\overline{T}^d} \big(\mathbf{I}_\sigma \varphi(x)\big)^p d\mu(x) \leqslant [\mu]_\otimes \sum_{\alpha \in E^d} \varphi(\alpha)^p \sigma(\alpha), \quad \varphi \geqslant 0,$$

and the problem is characterizing μ's for which $[\mu]_\otimes < \infty$, or even better some geometric, sharp estimate of $[\mu]_\otimes$.

(^{4}) The prototype of the bi-parameter Hardy inequalities is Sawyer's result [61], where he considers, in much more generality, inequalities of the form

$$\sum_{m,n=0}^{\infty} \Big(\sum_{i=0}^{m} \sum_{j=0}^{n} f(i, j) \Big)^2 \mu(m, n) \leqslant [\mu]_{\mathrm{Saw}} \sum_{m,n=0}^{\infty} \big| f(m, n) \big|^2.$$

His very clever proof does not extend to the tri-parameter case. Moreover, his inequality is not dyadic and covers the facts here surveyed only in the case of the trivial homogeneous rooted tree $\mathbb{N}$.

Once we have the kernel $\mathbf{k}$, the general theory [2] provides us also with definitions of potentials and energies of measures, and capacities of compact set $K \subseteq \overline{T}^d$, as exposed in Section 2.1. We can hope at this point that the capacitary estimate does the job,

$$\mu\left(\bigcup_{j=1}^{n} \alpha^{(j)}\right) \leqslant [\mu]_{\otimes,c} \operatorname{Cap}_{\pi,p}\left(\bigcup_{j=1}^{n} \alpha^{(j)}\right), \quad \text{for all } \alpha^{(1)}, \dots, \alpha^{(n)} \in E^d.$$

Following Maz'ya's lead, this would follow from a (multi-parameter) *strong capacitary inequality* (see Section 2.2),

$$\text{(SCI)} \qquad \sum_{k=-\infty}^{+\infty} 2^{2k} \operatorname{Cap}_{\pi,p}\left(x : \mathbf{I}f(x) > 2^k\right) \leqslant A\|\varphi\|^p_{\ell^p_+(E^d,\pi)}.$$

Here a major difficulty appears: the standard proofs of (SCI) depend, more or less explicitly, on the *boundedness principle* for potentials of measures,

$$\sup\left\{\mathbf{V}_p^{\mu,\sigma}(x) : x \in \overline{T}^d\right\} \leqslant B \cdot \max\left\{\mathbf{V}_p^{\mu,\sigma}(x) : x \in \operatorname{supp}(\mu)\right\},$$

but in the multi-parameter situation such a principle miserably fails.

PROPOSITION 31 ([8]). *For $d \geqslant 2$, there exist measures μ^K which are equilibrium for a compact $K \subset (\partial T)^d$, hence automatically satisfy $\mathbf{V}^{\mu^K} := \mathbf{V}_2^{\mu^K,1} \leqslant 1$ on $\operatorname{supp}(\mu^K)$, such that $\mathbf{V}^{\mu^K}(x) = +\infty$ at some point $x \in (\partial T)^d$.*

The idea is to have a set K which is rarefied, but "curved" is such a way many "not too thin" rectangles join it to the point x, like rays focusing on it.

The way out of this difficulty, implemented in [8] for $d = 2$, $p = 2$, and $\sigma \equiv 1$, is proving a *distributional boundedness principle*.

THEOREM 32. *There is $C > 0$ such that for $\lambda > 1$ and for an equilibrium measure μ,*

$$\operatorname{Cap}\left(\{x : \mathbf{V}^\mu(x) > \lambda\}\right) \leqslant C\,\frac{\|\mathbf{I}^*\mu\|^2_{\ell^2}}{\lambda^{2+1}}.$$

The inequality would follows by rescaling if one has 2 instead of $2+1$. This weaker form of the boundedness principle suffices to produce a variation of a classical proof of the strong capacitary inequality,

$$[\mu]_\otimes \approx [\mu]_{\otimes,c}, \quad \text{for } d = 2.$$

This result was then extended to $d = 3$, but no higher, in [53]. With some major difficulty, the capacitary characterization of the measures for which the multi-parameter, dyadic Hardy inequality holds is true at least for $d = 2, 3$. What about the other characterizations and proofs?

It is proved in [5] that a mass-energy condition holds as well in $d = 2$, and in [53] this was extended to $d = 3$, always for $p = 2$. More precisely, for $d = 2, 3$ we have that $[\mu]_\otimes \approx [\![\mu]\!]_\otimes$, where $[\![\mu]\!]_\otimes$ is the best constant in

$$\text{(ME}\otimes\text{)} \qquad \sum_{E^d \ni \beta \subseteq \alpha} \mu\big(S(\beta)\big)^2 \leq [\![\mu]\!]_\otimes\, \mu\big(S(\alpha)\big) < \infty, \quad \alpha \in E^d.$$

This fact might surprise practitioners of the Hardy space on the bi-disc. It was proved in [25] that Carleson measures for the Hardy space on the bi-disc are not characterized by a "single-box condition" such as (ME⊗), and A. Chang proved in [31] that the characterization holds if one allows multiple boxes. One might expect that a multiple-box condition like

$$\sum_{\beta \subseteq \bigcup_{j=1}^n \alpha^{(j)}} \mu\big(S(\beta)\big)^2 \leq [\mu]_{\text{mult}} \mu\Bigg(\bigcup_{j=1}^n S(\alpha^{(j)})\Bigg) < \infty, \quad \text{for all } \alpha^{(1)}, \ldots, \alpha^{(n)} \in E^d,$$

might not be weakened, but in fact this is not the case.

The proofs we surveyed for the one-parameter Hardy operator seem not to work in the multi-parameter case. The simple maximal proof, for instance, does not work because, contrary to the usual dyadic, weighted maximal function, its several parameter versions,

$$\mathcal{M}_\mu f(x) = \sup_{E^d \ni \alpha \supset x} \frac{1}{\mu\big(S(\alpha)\big)} \int_{S(\alpha)} f d\mu,$$

are not necessarily weakly bounded on L^1, neither they are bounded on L^2. In the unweighted case, the L^2 boundedness of the multi-parameter maximal function was proved in [42], and a nice account of multi-parameter theory with applications to martingales and the Hardy space is in [37].

PROBLEM 11. It would be interesting to know whether, like in the one-parameter case, for $1 \leqslant s < \infty$,

$$\sup_{\alpha \in E^d} \frac{1}{\mu\big(S(\alpha)\big)} \int_{S(\alpha)} \bigg(\int_{S(\alpha)} \delta_d(x \wedge y) d\mu(y)\bigg)^s d\mu(x) \approx C_s [\![\mu]\!]^s,$$

where $\delta_d(x \wedge y) := \prod_{j=1}^d d(x_j \wedge y_j)$.

A. DYADIC DECOMPOSITION OF AHLFORS REGULAR METRIC SPACES

Let $\mathcal{D}$ be the sets of the dyadic intervals $I_{n,j} = [(j-1)/2^n, j/2^n)$ in $[0, 1)$, $1 \leq j \leq 2^n$, $n \geq 0$. Let also $\mu, \nu : \mathcal{D} \to \mathbb{R}_+$ be two weight functions. Then, we say that the *two-weight dyadic Hardy inequality* holds if there exists a positive constant $[\mu]$, possibly depending on π and p, such that for all $f : \mathcal{D} \to \mathbb{R}_+$,

$$\sum_{I\in\mathcal{D}}\Big(\sum_{J\supseteq I} f(J)\Big)^p \mu(I) \le [\mu]\sum_{I\in\mathcal{D}} f(I)^p\nu(I). \tag{A.1}$$

It is clear that the above is, in fact, a Hardy inequality on the homogeneous dyadic tree: interpret $\mathcal{D}$ as the vertex set of a tree, where two vertices are connected by an edge if and only if one of the corresponding intervals $I_{n,j}, I_{m,k}$ contains the other and $|n-m| = 1$, and set o to be the vertex corresponding to $[0, 1)$. Observe that here, as compared to the general formulation (H), we are in the simplest case when μ is supported on the tree $\mathcal{D}$ rather than on $\bar{\mathcal{D}}$. In the paper, we chose to work in higher generality and allow trees to be not necessarily homogeneous and the measure μ to give mass also to ∂T, the natural boundary of the tree. In this way, it becomes clear that the symmetric-space structure of the group of automorphisms of the homogeneous tree plays no role in (most of) this theory. Moreover, to extend the support of the measure up to boundary is justified by the problem of exceptional sets at the boundary, i.e., of trace measures.

We remark that $\mathcal{D}$ is just the prototype of decomposition of a metric space. In a more general context, if (X, μ, ρ) is a homogeneous metric measure space, then its Christ's decomposition [32, Theorem 11] provides a family of generalized "cubes" $\{Q_k^\alpha\}$ which can be readily checked to form a tree.

Let us show that (A.1) is a genuine generalization of the classical Hardy inequality (Hardy).

Suppose for example that $\{I_n\}$ is an infinite branch of dyadic intervals; i.e., $I_n \in \mathcal{D}$, $I_0 = [0, 1)$, $I_{n+1} \subseteq I_n$, and $2|I_{n+1}| = |I_n|$. Set also $\pi(I) = \mu(I) = 0$ if $I \in \mathcal{D}$ is not one of the I_n and $\mu(I_n) =: U_n$, $\pi(I_n) =: V_n$. Write φ_n for $\varphi(I_n)$. Then, the dyadic Hardy inequality takes the form

$$\sum_{n=0}^{\infty}\Big(\sum_{m=0}^{n}\varphi_m\Big)^p U_n \le [\mu]\sum_{n=0}^{\infty}\varphi_n^p V_n,$$

which is of course the discrete analogue of Muckenhoupt's two-weight Hardy inequality in $\mathbb{R}_+$. In particular, by choosing $U_n = n^{-p}$ and $V_n = 1$, one gets back to (Hardy).

An interesting example of an unexpected application of (A.1), coming from complex analysis, is the problem of characterizing Carleson measures for Besov spaces B_a^p. A Carleson measure for B_a^p is a positive Borel measure $\tilde{\mu}$ on the unit disc $\mathbb{D}$ of the complex plane for which there exists a constant $C(\tilde{\mu}) < \infty$ such that, for all f holomorphic on $\mathbb{D}$,

$$\int_{\mathbb{D}} |f|^p d\tilde{\mu} \le C(\tilde{\mu})\Big[\big|f(0)\big|^2 + \int_{\mathbb{D}}\big|f'(z)\big|^p\big(1-|z|^2\big)^{p+a-2}dxdy\Big], \tag{Besov}$$

with $0 \leq ap < 1$. Such problems appear in connection to the characterization of multipliers and of exceptional sets at the boundary for spaces of holomorphic functions, sequences of interpolation, and more. The first result is by Stegenga [64] who manages to characterize such measures for $p = 2$ in terms of a condition involving Riesz capacities of compact subsets of the unit disc. The root of Stegenga's work can be traced back to earlier works: Maz'ya [51] and Adams [1]. For the case $1 < p < \infty$, a similar characterization of Carleson measures in terms of nonlinear Riesz capacities was later obtained by Verbitskiĭ [69] and rediscovered by Wu [72].

More recently, it was proved in [10, 44] that the Carleson inequality for Besov spaces is equivalent to a dyadic Hardy inequality. More precisely, for a given dyadic interval $I = I_{n,j} \in \mathcal{D}$, let $Q(I)$ be the set of points $z = re^{i\theta}$ with $\theta/2\pi \in I$ and $1 - 2^{-n} \leq r \leq 1 - 2^{-n-1}$. Set also $\mu(I) := \widetilde{\mu}(Q(I))$ and $\pi(I) = 2^{-an}$. Then, $\widetilde{\mu}$ is a Carleson measure for B_a^p, $0 \leq ap < 1$, if and only if the triple p, μ, π satisfies the dyadic Hardy inequality (A.1). Motivated by this application, we will call Carleson measures (or trace measures) all measures $\mu : \mathcal{D} \to \mathbb{R}_+$ satisfying the Hardy inequality on trees (H).

The dyadic setting is much ductile. The same inequality (A.1), with different choices of the weight π, can be used to characterize Carleson measures for holomorphic spaces in several dimensions or for spaces of harmonic functions, trace inequalities for potential spaces, and more. Many such problems, in fact, can be proven to be equivalent to their dyadic counterparts, and often (A.1) is the form they assume.

B. Bessel potentials on the boundary of the dyadic tree

The content of this section is specific to the homogeneous tree. Out of simplicity, we consider only the dyadic case, but everything we say applies, mutatis mutandis, to homogeneous trees of any degree.

Our objective in this section is to introduce, using as usual the axiomatic theory of Adams and Hedberg, a seemingly different potential theory on the boundary of the dyadic tree which depends on two parameters p and s. Subsequently, we will use the inequality of Muckenhoupt and Wheeden in order to prove that it is "equivalent" to the potential theory introduced in Section 2.1 for the same parameter p and a particular choice of the weight π. "Equivalent" means that for compact sets that lie on the boundary of the dyadic tree, the capacities of the sets measured by means of the two different potential theories are comparable.

Let T be the dyadic tree and consider the compact Hausdorff space $\overline{T}$ and the measure space $(\partial T, dx)$. Fix some parameter $0 < s \leq 1/p^*$. We then define the s-Bessel kernel as

$$G_s(x, y) := |\alpha|^{-s}, \quad x, y \in \partial T,$$

where α is the (unique) edge such that $e(\alpha) = x \wedge y$.

Following, as always, [2], we define the Bessel potential of a function $\varphi : \partial T \to \mathbb{R}_+$ as

$$G_s\varphi(y) := \int_{\partial T} G_s(x,y)\varphi(x)dx, \quad y \in \overline{T}.$$

Also the Bessel co-potential of a measure μ is defined as

$$G_s^*\mu(x) := \int_{\overline{T}} G_s(x,y)d\mu(y), \quad x \in \partial T.$$

Notice that, if temporarily we use the notation $\widehat{\alpha}$ to denote the only child of α lying in $[o^*, x]$, for some fixed vertex x, we can estimate the Bessel co-potential as follows:

$$\begin{aligned} G_s^*\mu(x) &= \int_{\overline{T}} G_s(x,y)d\mu(y) \\ &= \sum_{\alpha\supset x} \int_{S(\alpha)\setminus S(\widehat{\alpha})} |\alpha|^{-s} d\mu(y) \\ &= \sum_{\alpha\supset x} \big(\mu\big(S(\alpha)\big) - \mu\big(S(\widehat{\alpha})\big)\big)|\alpha|^{-s}. \end{aligned}$$

Hence, we trivially have $G_s^*\mu(x) \le \sum_{\alpha\supset x} \mu(S(\alpha))|\alpha|^{-s}$. On the other hand,

$$\begin{aligned} G_s^*\mu(x) &= \sum_{\alpha\supset x} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s} - \sum_{\alpha\supset x} \frac{\mu\big(S(\widehat{\alpha})\big)}{2^s|\widehat{\alpha}|^s} \\ &= \mu(\partial T) + (1-2^{-s}) \sum_{\alpha\supset x, \alpha\neq\omega} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s} \\ &\ge (1-2^{-s}) \sum_{a\supset x} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s}. \end{aligned}$$

In other words,

$$G_s^*\mu(x) \approx \sum_{\alpha\supset x} \mu\big(S(\alpha)\big)|\alpha|^{-s}.$$

The associated Bessel energy is given by

$$\begin{aligned} E_p^s(\mu) &:= \int_{\overline{T}} \left(\int_{\partial T} G_s(x,y) G_s^*\mu(x)^{p^*-1} dx \right) d\mu(y) \\ &= \int_{\partial T} G_s^*\mu(x)^{p^*} dx \approx \int_{\partial T} \left(\sum_{\alpha\supset x} \frac{\mu\big(S(\alpha)\big)}{|\alpha|^s} \right)^{p^*} dx \\ &\approx \sum_{\alpha} \frac{\mu\big(S(\alpha)\big)^{p^*}}{|\alpha|^{sp^*-1}} = \mathcal{E}_{p,\pi}(\mu), \end{aligned}$$

where $\pi(\alpha) = |\alpha|^{\frac{1-sp^*}{1-p^*}}$. Notice that we have used the Muckenhoupt–Wheeden inequality (MW) in the last step.

Therefore, since the energies associated to a positive Borel measure via the two different potential theories are comparable, we can conclude that the $\mathrm{Cap}_{p,\pi}$ capacity of a compact subset of the boundary of the dyadic tree and its s-Bessel capacity are comparable. In particular, compact sets of zero $\mathrm{Cap}_{p,\pi}$ capacity coincide with those of zero s-Bessel capacity.

Funding. – N. Chalmoukis is partially supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers” (Project Number: 4662). M. Levi was partially supported by the Project “Harmonic Analysis on Continuous and Discrete Structures” (bando Trapezio Compagnia di San Paolo CUP E13C21000270007), the Project GNAMPA 2022 “Generalized Laplacians on Continuous and Discrete Structures” (CUP E55F22000270001), P. Mozolyako was supported by the Ministry of Science and Higher Education of the Russian Federation, agreement 075-15-2021-602.

References

[1] D. R. Adams, On the existence of capacitary strong type estimates in R^n. *Ark. Mat.* **14** (1976), no. 1, 125–140. Zbl 0325.31008 MR 417774

[2] D. R. Adams – L. I. Hedberg, *Function spaces and potential theory*. Grundlehren Math. Wiss. 314, Springer, Berlin, 1996. Zbl 0834.46021 MR 1411441

[3] J. Agler – J. E. McCarthy, *Pick interpolation and Hilbert function spaces*. Grad. Stud. Math. 44, American Mathematical Society, Providence, RI, 2002. Zbl 1010.47001 MR 1882259

[4] N. Arcozzi – N. Chalmoukis, Riesz capacities of a set due to Dobiński. *C. R. Math. Acad. Sci. Paris* **360** (2022), 679–685. Zbl 1495.30019 MR 4449868

[5] N. Arcozzi – I. Holmes – P. Mozolyako – A. Volberg, Bi-parameter embedding and measures with restricted energy conditions. *Math. Ann.* **377** (2020), no. 1-2, 643–674. Zbl 1439.42014 MR 4099629

[6] N. Arcozzi – I. Holmes – P. Mozolyako – A. Volberg, Bellman function sitting on a tree. *Int. Math. Res. Not. IMRN* **2021** (2021), no. 16, 12037–12053. Zbl 1486.42014 MR 4370368

[7] N. Arcozzi – M. Levi, Equilibrium measures on trees. *Collect. Math.* **74** (2023), no. 1, 61–79. Zbl 1509.31021 MR 4528947

[8] N. Arcozzi – P. Mozolyako – K.-M. Perfekt – G. Sarfatti, Bi-parameter Potential theory and Carleson measures for the Dirichlet space on the bidisc. 2018, arXiv:1811.04990, to appear in *Discrete Anal.*

[9] N. Arcozzi – R. Rochberg, Topics in dyadic Dirichlet spaces. *New York J. Math.* **10** (2004), 45–67. Zbl 1067.46038 MR 2052364

[10] N. Arcozzi – R. Rochberg – E. Sawyer, Carleson measures for analytic Besov spaces. *Rev. Mat. Iberoamericana* **18** (2002), no. 2, 443–510. Zbl 1059.30051 MR 1949836

[11] N. Arcozzi – R. Rochberg – E. Sawyer, The characterization of the Carleson measures for analytic Besov spaces: a simple proof. In *Complex and harmonic analysis*, pp. 167–177, DEStech Publishing, Lancaster, PA, 2007. Zbl 1148.30028 MR 2387287

[12] N. Arcozzi – R. Rochberg – E. Sawyer, The diameter space, a restriction of the Drury-Arveson-Hardy space. In *Function spaces*, pp. 21–42, Contemp. Math. 435, American Mathematical Society, Providence, RI, 2007. MR 2359415

[13] N. Arcozzi – R. Rochberg – E. Sawyer, Capacity, Carleson measures, boundary convergence, and exceptional sets. In *Perspectives in partial differential equations, harmonic analysis and applications*, pp. 1–20, Proc. Sympos. Pure Math. 79, American Mathematical Society, Providence, RI, 2008. Zbl 1182.30099 MR 2500486

[14] N. Arcozzi – R. Rochberg – E. T. Sawyer – B. D. Wick, The Dirichlet space: a survey. *New York J. Math.* **17A** (2011), 45–86. Zbl 1233.30017 MR 2782728

[15] N. Arcozzi – R. Rochberg – E. T. Sawyer – B. D. Wick, Potential theory on trees, graphs and Ahlfors-regular metric spaces. *Potential Anal.* **41** (2014), no. 2, 317–366. Zbl 1294.31004 MR 3232028

[16] N. Arcozzi – R. Rochberg – E. T. Sawyer – B. D. Wick, *The Dirichlet space and related function spaces*. Math. Surveys Monogr. 239, American Mathematical Society, Providence, RI, 2019. Zbl 1450.30002 MR 3969961

[17] N. Aronszajn, Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** (1950), 337–404. Zbl 0037.20701 MR 51437

[18] E. Berchio – D. Ganguly – G. Grillo – Y. Pinchover, An optimal improvement for the Hardy inequality on the hyperbolic space and related manifolds. *Proc. Roy. Soc. Edinburgh Sect. A* **150** (2020), no. 4, 1699–1736. Zbl 1457.46043 MR 4122432

[19] E. Berchio – F. Santagati – M. Vallarino, Poincaré and Hardy inequalities on homogeneous trees. In *Geometric properties for parabolic and elliptic PDEs*, pp. 1–22, Springer INdAM Ser. 47, Springer, Cham, 2021. Zbl 1473.35011 MR 4279612

[20] A. Beurling, Ensembles exceptionnels. *Acta Math.* **72** (1940), 1–13. Zbl 0023.14204 MR 1370

[21] J. Bourgain, On the radial variation of bounded analytic functions on the disc. *Duke Math. J.* **69** (1993), no. 3, 671–682. Zbl 0787.30020 MR 1208816

[22] J. S. Bradley, Hardy inequalities with mixed norms. *Canad. Math. Bull.* **21** (1978), no. 4, 405–408. Zbl 0402.26006 MR 523580

[23] D. L. Burkholder, Martingale inequalities. In *Martingales (Report Meeting, Oberwolfach, 1970)*, pp. 1–8, Lecture Notes in Math. 190, Springer, Berlin, 1971. MR 372985

[24] D. L. Burkholder, Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12** (1984), no. 3, 647–702. Zbl 0556.60021 MR 744226

[25] L. Carleson, A counterexample for measures bounded on H^p for the bi-disc. Report 7, Institut Mittag-Leffler, 1974.

[26] M. Cavina, Bellman function for Hardy's inequality over dyadic trees. 2020, arXiv:2002.07532.

[27] N. Chalmoukis, Onto interpolation for the Dirichlet space and for $H_1(\mathbb{D})$. *Adv. Math.* **381** (2021), article no. 107634. Zbl 1467.30022 MR 4215748

[28] N. Chalmoukis, A note on simply interpolating sequences for the Dirichlet space. *New York J. Math.* **29** (2023), 193–202. Zbl 1508.30075 MR 4551188

[29] N. Chalmoukis – A. Hartmann – K. Kellay – B. D. Wick, Random interpolating sequences in Dirichlet spaces. *Int. Math. Res. Not. IMRN* **2022** (2022), no. 17, 13629–13658. Zbl 07582361 MR 4475285

[30] N. Chalmoukis – M. Levi, Some remarks on the Dirichlet problem on infinite trees. *Concr. Oper.* **6** (2019), no. 1, 20–32. Zbl 1418.31012 MR 3938329

[31] S.-Y. A. Chang, Carleson measure on the bi-disc. *Ann. of Math. (2)* **109** (1979), no. 3, 613–620. Zbl 0401.28004 MR 534766

[32] M. Christ, A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.* **60/61** (1990), no. 2, 601–628. Zbl 0758.42009 MR 1096400

[33] R. R. Coifman – C. Fefferman, Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.* **51** (1974), 241–250. Zbl 0291.44007 MR 358205

[34] B. Devyver – M. Fraas – Y. Pinchover, Optimal Hardy weight for second-order elliptic operator: an answer to a problem of Agmon. *J. Funct. Anal.* **266** (2014), no. 7, 4422–4489. Zbl 1298.47057 MR 3170212

[35] A. Figà-Talamanca – C. Nebbia, *Harmonic analysis and representation theory for groups acting on homogeneous trees*. London Math. Soc. Lecture Note Ser. 162, Cambridge University Press, Cambridge, 1991. Zbl 1154.22301 MR 1152801

[36] L. Grafakos, *Classical Fourier analysis*. 2nd edn., Grad. Texts in Math. 249, Springer, New York, 2008. Zbl 1220.42001 MR 2445437

[37] R. F. Gundy, Inégalités pour martingales à un et deux indices: l'espace H^p. In *Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978)*, pp. 251–334, Lecture Notes in Math. 774, Springer, Berlin, 1980. MR 590628

[38] G. H. Hardy, Notes on some points in the integral calculus LI: On Hilbert's double-series theorem, and some connected theorems concerning the convergence of infinite series and integrals. *Messenger Math* **48** (1919), 107–112.

[39] G. H. Hardy, Note on a theorem of Hilbert. *Math. Z.* **6** (1920), no. 3-4, 314–317. Zbl 47.0207.01 MR 1544414

[40] G. H. Hardy, Notes on some points in the integral calculus LX: An inequality between integrals. *Messenger Math* **54** (1925), 150–156. Zbl 51.0192.01

[41] L. I. Hedberg – T. H. Wolff, Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* **33** (1983), no. 4, 161–187. Zbl 0508.31008 MR 727526

[42] B. Jessen – J. Marcinkiewicz – A. Zygmund, Note on the differentiability of multiple integrals. *Fund. Math.* **25** (1935), 217–234. Zbl 0012.05901

[43] M. Keller – Y. Pinchover – F. Pogorzelski, Optimal Hardy inequalities for Schrödinger operators on graphs. *Comm. Math. Phys.* **358** (2018), no. 2, 767–790. Zbl 1390.26042 MR 3774437

[44] R. Kerman – E. Sawyer, Carleson measures and multipliers of Dirichlet-type spaces. *Trans. Amer. Math. Soc.* **309** (1988), no. 1, 87–98. Zbl 0657.31019 MR 957062

[45] A. Kufner – L. Maligranda – L.-E. Persson, The prehistory of the Hardy inequality. *Amer. Math. Monthly* **113** (2006), no. 8, 715–732. Zbl 1153.01015 MR 2256532

[46] P. D. Lax, *Functional analysis*. Pure Appl. Math. (N.Y.), John Wiley & Sons, New York, 2002. Zbl 1009.47001 MR 1892228

[47] M. Levi – A. Martini – F. Santagati – A. Tabacco – M. Vallarino, Riesz transform for a flow Laplacian on homogeneous trees. *J. Fourier Anal. Appl.* **29** (2023), no. 2, article no. 15. Zbl 1512.05083 MR 4552579

[48] M. Levi – F. Santagati – A. Tabacco – M. Vallarino, Analysis on trees with nondoubling flow measures. *Potential Anal.* **58** (2023), no. 4, 731–759. Zbl 1511.05039 MR 4568880

[49] D. H. Luecking, Trace ideal criteria for Toeplitz operators. *J. Funct. Anal.* **73** (1987), no. 2, 345–368. Zbl 0618.47018 MR 899655

[50] V. Maz'ya, *Sobolev spaces with applications to elliptic partial differential equations*. augmented edn., Grundlehren Math. Wiss. 342, Springer, Heidelberg, 2011. Zbl 1217.46002 MR 2777530

[51] V. G. Maz'ya, On certain integral inequalities for functions of many variables. *J. Soviet Math.* **1** (1973), 205–234. Zbl 0284.26011

[52] P. Mozolyako – G. Psaromiligkos – A. Volberg, Counterexamples for multi-parameter weighted paraproducts. *C. R. Math. Acad. Sci. Paris* **358** (2020), no. 5, 529–534. MR 4149851

[53] P. Mozolyako – G. Psaromiligkos – A. Volberg – P. Zorin-Kranich, Carleson embedding on the tri-tree and on the tri-disc. *Rev. Mat. Iberoam.* **38** (2022), no. 7, 2069–2116. Zbl 1510.42032 MR 4526311

[54] B. Muckenhoupt, Hardy's inequality with weights. *Studia Math.* **44** (1972), 31–38. Zbl 0236.26015 MR 311856

[55] B. Muckenhoupt – R. Wheeden, Weighted norm inequalities for fractional integrals. *Trans. Amer. Math. Soc.* **192** (1974), 261–274. Zbl 0289.26010 MR 340523

[56] F. Nazarov – S. Treil – A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces. *Int. Math. Res. Not. IMRN* **1998** (1998), no. 9, 463–487. Zbl 0918.42009 MR 1626935

[57] F. Nazarov – S. Treil – A. Volberg, The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.* **12** (1999), no. 4, 909–928. Zbl 0951.42007 MR 1685781

[58] F. Nazarov – S. Treil – A. Volberg, Bellman function in stochastic control and harmonic analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000)*, pp. 393–423, Oper. Theory Adv. Appl. 129, Birkhäuser, Basel, 2001. Zbl 0999.60064 MR 1882704

[59] W. Rudin, The radial variation of analytic functions. *Duke Math. J.* **22** (1955), 235–242. Zbl 0064.31105 MR 79093

[60] E. Sawyer, Weighted Lebesgue and Lorentz norm inequalities for the Hardy operator. *Trans. Amer. Math. Soc.* **281** (1984), no. 1, 329–337. Zbl 0538.47020 MR 719673

[61] E. Sawyer, Weighted inequalities for the two-dimensional Hardy operator. *Studia Math.* **82** (1985), no. 1, 1–16. Zbl 0585.42020 MR 809769

[62] P. M. Soardi, *Potential theory on infinite networks*. Lecture Notes in Math. 1590, Springer, Berlin, 1994. Zbl 0818.31001 MR 1324344

[63] S. Sobolev, On a theorem in functional analysis. *Amer. Math. Soc. Transl. (2)* **34** (1963), 39–68.

[64] D. A. Stegenga, Multipliers of the Dirichlet space. *Illinois J. Math.* **24** (1980), no. 1, 113–139. Zbl 0432.30016 MR 550655

[65] E. M. Stein – G. Weiss, Fractional integrals on n-dimensional Euclidean space. *J. Math. Mech.* **7** (1958), 503–514. Zbl 0082.27201 MR 98285

[66] G. Talenti, Osservazioni sopra una classe di disuguaglianze. *Rend. Sem. Mat. Fis. Milano* **39** (1969), 171–185. Zbl 0218.26011 MR 280661

[67] E. Tchoundja, Carleson measures for the generalized Bergman spaces via a $T(1)$-type theorem. *Ark. Mat.* **46** (2008), no. 2, 377–406. Zbl 1159.32004 MR 2430733

[68] G. Tomaselli, A class of inequalities. *Boll. Un. Mat. Ital. (4)* **2** (1969), 622–631. Zbl 0188.12103 MR 255751

[69] I. È. Verbitskiĭ, Multipliers in spaces with "fractional" norms, and inner functions. *Sibirsk. Mat. Zh.* **26** (1985), no. 2, 51–72, 221. MR 788329

[70] J. Verdera, On the $T(1)$-theorem for the Cauchy integral. *Ark. Mat.* **38** (2000), no. 1, 183–199. Zbl 1039.42011 MR 1749365

[71] A. Volberg – B. D. Wick, Bergman-type singular integral operators and the characterization of Carleson measures for Besov-Sobolev spaces and the complex ball. *Amer. J. Math.* **134** (2012), no. 4, 949–992. Zbl 1252.42020 MR 2956255

[72] Z. Wu, Carleson measures and multipliers for Dirichlet spaces. *J. Funct. Anal.* **169** (1999), no. 1, 148–163. Zbl 0962.30032 MR 1726750

Received 9 February 2022,
and in revised form 6 May 2023

Nicola Arcozzi
Dipartimento di Matematica, Alma Mater Studorium Università di Bologna
Piazza di Porta San Donato 5, 40126 Bologna, Italy;
nicola.arcozzi@unibo.it

Nikolaos Chalmoukis
Dipartimento di Matematica e Applicazioni, Universitá degli studi di Milano Bicocca
Via Roberto Cozzi, 55, 20125 Milano, Italy;
nikolaos.chalmoukis@unimib.it

Matteo Levi
Facultat de Ciències, Universitat Autònoma de Barcelona
08193 Bellaterra;
Facultat de Matemàtiques i Informàtica, Universitat de Barcelona
Gran Via de les Corts Catalanes, 585, 08007 Barcelona, Spain;
matteo.levi@ub.edu

Pavel Mozolyako
Department of Mathematics and Computer Science, Saint Petersburg State University
199178 Saint Petersburg, Russia;
pmzlcroak@gmail.com