SAINT-PETERSBURG UNIVERSITY

MANUSCRIPT

PAVEL ALEXANDROVICH MOZOLYAKO

DISCRETE MODELS OF THE BOUNDARY BEHAVIOUR OF HARMONIC FUNCTIONS

Scientific specialty 1.1.1. Real, Complex and functional analysis

Doctoral Dissertation
in Physics and Mathematics

Translation from Russian

ST. PETERSBURG
2023

Contents

Introduction

The theory of Harmonic Functions holds a central place in Mathematical Analysis since for almost two hundred years. As of today it contains areas deemed classical and fundamental, especially if considered in parallel to its sibling, the Analytic function theory, like Hardy spaces or Subharmonic Functions, as well as the most recent and trending disciplines, such as Discrete Harmonic Functions, Harmonic Measure or inquiries into the growth and decay of Laplacian. It has innumerous connections to other domains – extensively and fruitfully exploited in both directions – Stochastic Processes and Random Fields, Potential Theory (which one might claim to be sired by Harmonic Function Theory), Analysis on Graphs and Networks, Statistical Physics, Fourier Analysis, Wavelet and Signal Analysis, and many many more.

Questions regarding the boundary behaviour of harmonic functions permeate the theory from its very foundations to the newest and freshly developed branches. One could argue that the very problem of existence – the Dirichlet problem – already contains most of issues usually studied by researchers in the field. How the function converges to its boundary values, or, in the other way around, how it grows and/or oscillates (which, mysteriously, is oftentimes easier to investigate – three-dimensional version of Privalov's theorem or tail Law of Iterated Logarithm for harmonic functions come to mind), depending on the manner we are approaching the boundary, infinity included, the zero sets and level sets, all of these and more was studied since pretty much the very beginning. Maximum Principle, Hadamard's three circles, Fatou theorems, angular derivative, regular points, boundary Harnack's inequality, propagation of smallness and the Law of Iterated Logarithm are just a few of those.

The relation between classical Harmonic Function Theory and its discrete models, which is so explicit in, for example, Stochastic Processes, came to the spotlight somewhat more recently. One could say that it was present in the background, hidden in, say Littlewood-Paley decomposition, stopping time arguments in the disc, Carleson tents and others. It did arrive in full power in early eighties with the famous results by Makarov on the support of harmonic measure. Then it gained further momentum with the emergence of wavelets, dyadic Singular Integral Theory and Probability on graphs, and by now discrete harmonic functions (of all flavors) serve both as a tool and as an independent object of study (in particular their dissimilarity with their continuous parents – discrete Liouville's theorem comes to mind among the recent developments).

The thesis is dedicated to studying three types of boundary behaviour problems in relation with discrete models.

The first topic considers Carleson measures for a certain class of analytic and harmonic spaces in the unit polydisc. Given $H(\Omega)$ – a Hilbert space of functions, usually harmonic or analytic, on a domain $\Omega \subset \mathbb{R}^d$ and a Radon measure μ on Ω (or sometimes on its closure $\overline{\Omega}$ – where it makes sense) we call μ a Carleson measure for $H(\Omega)$, if the inequality

$$\int_\Omega |f|^2\, d\mu \leq C_\mu \|f\|^2_{H(\Omega)}$$

holds for any $f \in H(\Omega)$ with some absolute constant C_μ.

The notion itself goes back to the famous paper of Carleson [21] about the corona problem, where, in particular, he gave a description of such measures for the Hardy space $H^2(\mathbb{D})$. Later on the extensions were given for other spaces of analytic and harmonic functions. We mention here the results of Stegenga [90] (Hardy-Sobolev scale on $\mathbb{D}$), Hastings [40] and Luecking [59] (Bergman spaces in several variables), [4] and [5] (Besov spaces on complex balls). For the Hardy-Sobolev scale on the polydisc the main references are [24, 25, 26], where the Hardy space (harmonic version) on the polydisc was considered. Carleson measures proved to be a central notion in the analysis of holomorphic spaces, as they intervene in the characterization of multipliers, interpolating sequences, and Hankel-type forms, in Corona-type problems, in the characterization of exceptional sets at the boundary, and more. Moreover in Potential Theory they come in the form of *trace measures* – measures that realize the bounded embeddings of Sobolev spaces into $L^p(\mu)$ called *trace inequalities*. This incarnation goes back to seminal works of V.G. Maz'ya which dealt with weak solutions of the Schrödinger equations, and later developments produced a stream of results about properties and descriptions of such measures.

Our aim is to characterize the Carleson measures for a certain scale of Hardy-Sobolev spaces on the polydisc. Analysis on polydisc is, in general, much more complicated compared to the classical one-dimensional case, and these difficulties will expose themselves in several places of the arguments. In particular we note that outside of a number of important papers by S.-Y. A. Chang and R. Fefferman in early eighties we have mentioned above no other results for Carleson measures were obtained in the multi-parametric setting, even on the bi-disc. The approach we used is the dyadic discrete one. Namely, following [6] we transfer our continuous problem into a discrete medium, reformulating the Carleson embedding as a *Hardy embedding* on a certain graph we call a d-tree. We then continue by solving the discrete problem, and show that the solution can be properly moved back to the polydisc. Moreover, it is the discrete embedding and related questions that actually come to the forefront of our investigations, as we consider it to be at least as interesting a problem as the description of Carleson measures due to connections to dyadic paraproducts, singular integral operators, multiparametric potentials and other notions. Loss of information when moving between continuous and discrete settings, especially in the context of analytic functions, is also mentioned.

The second topic considers the so-called growth spaces of harmonic functions. These consist of harmonic functions on a domain $\Omega \subset \mathbb{R}^d$ satisfying a growth condition near the boundary

$$|u(x)| \leq C_u w(\mathrm{dist}(x, \partial\Omega)), \quad x \in \Omega,$$

where the function w is called *the growth weight* and it is usually a doubling one.

Such spaces of analytic and harmonic functions in the unit disk were considered by A. Shields and D. Williams, while the Fourier series of such functions were studied by G. Bennett, D. Stegenga and R. Timoney, who showed that the growth of the function cannot be characterized by the growth of the partial sums of its Fourier series (which is directly related to our results). Another group of results is due to B. Korenblum, whose classical paper on the extension of Nevanlinna theory was a starting point for the research in the area, and follow-up papers by Yu. Lyubarskii, E. Malinnikova and P. Thomas. Moreover, these spaces on the unit disc were a subject of classical works by M. Cartwright who showed that one-sided polynomial growth condition actually implies the two-sided bound. They were extended to a large class of weights by A. Borichev.

Within this topic we show that functions in growth spaces can be described in terms of their wavelet decompositions. Wavelet series of the boundary values of a harmonic function is a convenient tool that replaces the Fourier series. For instance, the Haar wavelets provide the martingale representation. We chose smooth multiresolution analysis, the smoothness depending on the weight w. Next we prove that such a function in a Lipschitz domain should oscillate wildly near the boundary. It is the expected behaviour, expressed in a number of the Law of the Iterated Logarithm type scenarios, and we give a version adapted for our space.

Another result is the extension of M. Cartwright theorem to the unit ball in $\mathbb{R}^d$ and to a large class of regular weights.

Finally we prove a couple of oscillation-type results for divided differences of Hölder functions on the real line. On a first glance this is neither harmonic nor growth related, however such differences behave in a very similar way to the vertical, say, derivative of a harmonic function under growth condition, and the approach to handle them is very similar to the techniques in the results mentioned above, especially the dyadic martingale representation (actually they can be deduced from the wavelet decomposition representation).

The third topic is dedicated to the variation of harmonic functions near the boundary. For a given smooth domain $\Omega \subset \mathbb{R}^d$ and a positive harmonic function u on Ω we study the points $\xi \in \partial\Omega$ such that

$$\int_0^1 |\nabla u(p + t\vec{N}(\xi))| \, dt < +\infty,$$

where $\vec{N}(\xi)$ is the inward normal vector to $\partial\Omega$ at ξ.

The study of such an integral – *a normal variation of* u can be traced to a classical 1955 paper by W. Rudin, where he proved that it can be infinite at almost every boundary point for a bounded analytic function u in the unit disc. In 1993 Bourgain showed that the set of points of finite variation of bounded analytic (and later on, of poisitve harmonic) functions must nevertheless be quite large in the sense of Hausdorff dimension. This result was later extended to the unit ball by M. O'Neill. The normal variation turned up in many different contexts, we also mention works by J. Ortega-Cerda, D. Girela, P. Jones, P.F.X. Müller, D. Walsh and others, who used variants of this quantity to characterize spaces of analytic functions, or to deduce boundary properties of functions on such spaces. Our main result here is the extension of Bourgain's theorem to smooth domains in $\mathbb{R}^d$. We show that normal variation of a positive harmonic function is finite on a large set of boundary points. The important idea here is that we avoid using Fourier analysis (as done by Bourgain and O'Neill), and use the estimates of Green's functions instead. Hence the method is not a discrete one, nevertheless a proper discrete statement of the problem would be of interest.

As it happens in mathematics, the results are usually obtained by collaborative work and complex discussions. Nevertheless the author has chosen to present here those theorems and arguments that he can safely claim to put his name on, unless otherwise specified.

The thesis consists of Introduction, 8 Chapters and a Conclusion. It contains definitions, arguments, historical references, auxiliary statements and Theorems, which are **presented to defense**. Thesis results are published in 15 articles ([98]-[112]) in peer-reviewed journals. They were demonstrated many times in talks on international scientific conferences and research seminars. This provides their **reliability**.
Further in this chapter we will introduce the main results of the thesis, and formulate their statements as well as describe the setting. For reader's convenience the theorems are reintroduced and their proofs are given in the respective chapters. The numbering of displays in the main part of the thesis consists of three digits: chapter number, section number and the formula number within section. The numbering in the Introduction is consecutive, with a prefix **I.** attached.

The **relevance** of the thesis is justified by the choice of the topics. They are embedded into a wide mathematical context with numerous connections to Potential Theory, Stochastic Processes, Functional Analysis (in particular Hilbert Spaces of Analytic Functions), Analysis on Graphs, PDEs and many other areas of research. All of these are actively developing fields populated by prominent mathematicians and they attract a lot of interest from specialists. Recent activity around Nehari's theorem on the polydisc should be mentioned as well.

I consider the presented results and topics to be appropriately **developed** and ready for the defense. The included results contain general statements, their applications, examples and counterexamples. In addition there is a number of follow-up questions and a number of paths are opened for further investigations. Some of those are presented in the Conclusion.

The **purpose** of the work is three-fold. The first aim is to obtain new concrete mathematical results, the second is to develop new methods and techniques, and to refine already known ones, and the third is to open up directions for subsequent studies and expand our understanding of discussed problems.

Most of the results of the thesis are **new**, this is corroborated by publications in peer-reviewed journals. When the already known statement is discussed, the actual result is the application or presentation of a new approach. This is also of interest.

The work in thesis is of **theoretical** nature. The obtained results can be (and some of them already are) applied in the mathematical research in the areas mentioned above.

The **methods** of the thesis, both new and established, are related to: Functional Analysis, Measure Theory, Reproducing Kernel Hilbert Spaces, Potential Theory, Harmonic Functions on Graphs, Combinatorics – for Chapters 1-3 – and also to Probability (especially dyadic martingales), PDEs, Wavelet Analysis, Bloch Spaces, Harmonic Measure for the rest of the thesis. In particular, we mention the new techniques of *energy majorization* and *energy decay estimates* in multi-parametric setting, which, generally speaking, belong to the main achievements of the thesis.

Discrete setting

Here we introduce a discrete model which will serve as the working ground for Chapters 1 and 2. The result obtained in those chapters will be moved to a continuous medium – spaces of harmonic and analytic functions on the polydisc in Chapter 3. The L^2 weighted embedding problem which is the central object of study in these chapters also appears here first, and it gives rise to (a version of) Potential Theory. This model comes in two flavors: first, we do it for a finite graph, and then we consider a variant for a certain type of infinite graphs (d-trees). We discuss several possible ways to interpret this graph, like a discretization of (D) by Whitney cubes, a subset of $\mathbb{R}^n$ (as the natural embedding of a graph into Euclidean space) and representing of the vertices of a d-tree as dyadic rectangles in the unit cube $[0,1]^d$. We define the setting and give preliminary statements. Some more details (regarding trees mostly) about directed graphs, product graphs and potentials can be found in [6] or [64], and also in [97, 78].

The ***vertices*** of a graph we denote by greek letters α, β, γ etc., in particular by τ, ω we usually denote the *boundary points*. Sometimes – when it is more convenient to think of a graph as a collection of dyadic rectangles – we write vertices as Q, R, I, J instead. We *do not* make use of graph edges (though discrete gradients are usually defined on edges in the literature, and somehow it is a more natural way of thinking, we move all the objects we deal with to vertices). As a result we *identify a graph and its vertex set*, which is reflected in notation and definitions.
Also, throughout the text, we write $A \lesssim B$ for a couple of quantities A, B, if there exists a constant C such that $A \leq CB$. This constant is usually assumed to not depend on variable quantities, depending on the context. Also we write $A \approx B$, if $A \lesssim B$ and $B \lesssim A$.

A ***directed graph*** Γ ***without directed cycles*** is a partially ordered set such that for any $\alpha, \beta \in \Gamma$ there exists $\gamma \in \Gamma$ such that $\gamma \geq \alpha$, $\gamma \geq \beta$. The sets $\{\beta \in \Gamma : \beta \geq \alpha\}$ and $\{\beta \in \Gamma : \beta \leq \alpha\}$ are denoted by $\mathcal{P}_\Gamma(\alpha)$ and $\mathcal{S}_\Gamma(\alpha)$ respectively (we drop the subscript Γ wherever it is convenient). Similarly, given $E \subset \Gamma$ we put $\mathcal{P}(E) := \bigcup_{\alpha \in E} \mathcal{P}(\alpha)$ and $\mathcal{S}(E) := \bigcup_{\alpha \in E} \mathcal{S}(\alpha)$.
If $\mathcal{P}(\alpha)$ is totally ordered for any $\alpha \in \Gamma$, then we call Γ ***a tree***. A tree T is 2^n***-adic***, if any $\alpha \in T$ that is a non-least element (i.e. there is no $\beta \leq \alpha$) has exactly 2^n immediate children. A d***-tree*** T^d is a Cartesian product of d copies of T with the product order. For the sake of convenience we

always assume, unless it is mentioned specifically, that the trees we are dealing with are dyadic. Most of the results and properties can be easily moved to the 2^n-adic case anyway.

By attaching the usual edge-counting distance to a finite graph Γ we make it a compact metric space (while this is of no importance for finite graphs, we'll have to tread a bit more carefully later on).

Next we move to infinite graphs. We start with the infinite dyadic tree T with a single root (i.e. unique maximal element), which we denote by o, and we make it into a compact metric space (see [6] for details). First we consider the combinatorial boundary $\tilde{\partial} T$ of T which is just a collection of half-infinite geodesics $\omega = \{o, \omega_1, \omega_2, \dots\}$, $\omega_k \in T$, with respect to the edge-counting distance starting at the root. For any two points α, β in T their ***confluent*** or the ***least common ancestor*** $\alpha \wedge \beta$ is the minimal element in $\mathcal{P}(\alpha) \cap \mathcal{P}(\beta)$. Also, for a pair of different geodesics $\omega, \zeta \in \tilde{\partial} T$ ***the confluent*** $\omega \wedge \zeta$ (which is a point of T itself) is the minimal point in $\omega \cap \zeta$, and we put $\omega \wedge \omega := \omega$. Now given $\alpha, \beta \in T$ we put

$$\operatorname{dist}(\alpha, \beta) := 3^{-|\alpha\wedge\beta|} - \frac{1}{2}(3^{-|\alpha|} + 3^{-|\beta|}), \tag{I.1}$$

where $|\tau| := \#\mathcal{P}(\tau) - 1$, $\tau \in T$ is the usual distance to the root. This is a metric (ultrametric even) on T, and taking a metric completion of T we obtain a compact metric space $\overline{T}$. The ***metric boundary*** ∂T of T is $\overline{T} \setminus T$.

Clearly every $\omega \in \tilde{\partial} T$ is a Cauchy sequence with respect to dist, moreover, if we denote by $[\omega]$ the equivalence class of ω in ∂T, then the map $\omega \mapsto [\omega]$ is a homeomorphism of $\tilde{\partial} T$ onto ∂T, thus the metric and combinatorial boundaries are the same. As such the predecessor set $\mathcal{P}(\omega)$ of a boundary point ω is just the geodesic ω itself, and we write $\omega \in \mathcal{S}(\alpha)$, $\alpha \in T$, if α belongs to this geodesic, $\omega = (o, \omega_1, \dots, \alpha, \dots)$. The balls in ∂T are of the form $\mathcal{S}(\alpha) \cap \partial T$, $\alpha \in T$.

Next we write $\overline{T}^d$ for the product of d copies of $\overline{T}$, and extend the metric structure in the usual way

$$\operatorname{dist}(\alpha, \beta) = \sup_{k=1,\dots,d} \operatorname{dist}(\alpha_k, \beta_k), \quad \alpha_k, \beta_k \in \overline{T},$$

where $\alpha = (\alpha_1, \dots, \alpha_d)$, $\beta = (\beta_1, \dots, \beta_d)$. The ***distinguished boundary*** $(\partial T)^d$ is the product of d copies of ∂T. We also write $\partial\mathcal{S}(\alpha)$ to be the 'shadow' of α on the distinguished boundary, $\partial\mathcal{S}(\alpha) = \mathcal{S}(\alpha) \cap (\partial T)^d$. As before, the balls are of the form $\partial\mathcal{S}(\alpha)$, where all the coordinates of the vertex α have the same depth.

We also need the ***restricted tree*** T_N which consists of points $\alpha \in T$ such that $\#\mathcal{P}(\alpha) \le N$, $N \in \mathbb{N}$. The ***restricted*** d***-tree*** T_N^d is again a product of d copies of T_N. Clearly $T^d = \bigcup_{N\ge1} T_N^d$.

Discrete setting: graph potentials

Here we attach some new objects to our graph Γ (in what follows by Γ we understand either a finite graph Γ_f or an infinite d-tree $\overline{T}^d$) – ***measures, weights and functions***.
A real-valued non-negative Radon ***measure*** on a finite graph Γ or on $\overline{T}^d$ we usually denote by letters μ, ν, σ, ρ etc.

A Borel measurable ***function*** is a mapping from Γ_f or $\overline{T}^d$ to $\mathbb{R}_+$ (so all functions are assumed to be non-negative), we denote it by f, φ, ψ etc.
A ***weight*** is a mapping from Γ_f or T^d (we do not define it on the boundary ∂T^d of an infinite graph) to positive numbers. In particular, for a finite graph Γ_f all these are just collections of positive numbers attached to the vertices of Γ_f.

Given a function f and a weight w on Γ we define the ***weighted Hardy operator*** $\mathbf{I}_w$ as follows

$$\mathbf{I}_w f(\alpha) := \sum_{\gamma\geq\alpha} f(\gamma)w(\gamma) := \int_{\mathcal{P}(\alpha)} f\,dw. \tag{I.2}$$

In particular, if $w \equiv 1$, we write $\mathbf{I}$. Sometimes we also write $\mathbf{I}(wf)$ instead of $\mathbf{I}_w f$.
Similarly, given a function φ and a measure μ we define the 'adjoint' Hardy operator (we'll justify this particular notation later) by

$$\mathbf{I}^*_\mu \varphi(\alpha) := \int_{\mathcal{S}(\alpha)} \varphi\,d\mu. \tag{I.3}$$

For a finite graph Γ_f it turns into $\sum_{\gamma\leq\alpha}\varphi(\gamma)\mu(\gamma)$. Like before we write $\mathbf{I}^*$ for $\mu \equiv 1$.
The ***weighted potential*** of a measure μ for the weight w is

$$\mathbf{V}^\mu_w(\alpha) := \mathbf{I}_w(\mathbf{I}^*\mu)(\alpha) = \sum_{\gamma\geq\alpha}\int_{\mathcal{S}(\gamma)} d\mu\, w(\gamma) = \int_\Gamma w(\mathcal{P}(\alpha\wedge\tau))d\mu(\tau), \tag{I.4}$$

where $\mathcal{P}(\alpha\wedge\gamma) = \mathcal{P}(\alpha)\cap\mathcal{P}(\gamma)$ and $w(\mathcal{P}(\alpha\wedge\gamma)) = \sum_{\tau\in\mathcal{P}(\alpha\wedge\gamma)} w(\tau)$. The ***weighted energy*** of a measure μ is

$$\mathcal{E}_w[\mu] := \int_\Gamma \mathbf{V}^\mu_w\,d\mu, \tag{I.5}$$

and the ***mutual energy*** of two measures μ, ν is

$$\mathcal{E}_w[\mu,\nu] := \int_\Gamma \mathbf{V}^\mu_w\,d\nu. \tag{I.6}$$

Now we are ready to define the ***weighted capacity*** on Γ. Assume that $K \subset \Gamma$ is a compact set. We let

$$\mathrm{Cap}_w(K) := \inf\{\|f\|^2_{L^2(\Gamma,dw)} :\ \mathbf{I}_w f(\tau) \geq 1,\ \tau\in K\}, \tag{I.7}$$

where $\|f\|^2_{L^2(\Gamma,dw)} = \sum_{\gamma\in\Gamma} f^2(\gamma)w(\gamma)$. We assume that empty set has infinite capacity. As usual we say that a property holds ***quasi-everywhere*** or q.e. if it holds for all $\gamma\in\Gamma$ except on a set with zero capacity.

Remark. Due to our agreement for a weight to be a positive function, we see that the capacity of a singleton in the interior of Γ is always non-zero. See also the discussion in the beginning of Section 1.1.

Potential theory

In this Section we explain how to introduce Potential Theory on graphs, in particular on d-trees, and on polydisc $\mathbb{D}^d$. We adopt the general point of view presented by Adams and Hedberg in that we give a general construction (as described in [1, Chapter 2]) which we then adapt to the respective setting. Unfortunately, we are not able to directly quote [1], since they still use $\mathbb{R}^n$ within their general setting (so, the setting is not general enough for our purposes), and some additional work is needed to handle graphs. We avoid this issue by embedding our graphs into $\mathbb{R}^n$, so that the potential kernels we use are interpreted to be defined on euclidean spaces. We could also have just repeated arguments from [1, Chapter 2] almost verbatim to handle potentials on d-trees, but it would be just redundant.
In some form or another these arguments can be seen (though mostly on trees) in [6], [10], [62], [63], [79] and [64, Chapter 16]. Most of the basic capacitary properties for exponential weights (product capacities and maximum/domination principles discussion in Section 1.1, boundary projection estimates in Section 3.2) can be seen in [104].
Some more details and basic results about potentials and capacities are provided in Chapter 1.

General Potential Theory

Here we give a very short summary of results in [1, Chapter 2.3-2.5] about the definition of capacity, potentials and existence of the equilibrium measure.

Let (M,ν) be a measure space, and $g : M\times\mathbb{R}^n \to \mathbb{R}_+\cup\{+\infty\}$ be a ***kernel function*** that is (a) lower semicontinuous on $\mathbb{R}^n$ for each $y\in M$ and (b) measurable on M for each $x\in\mathbb{R}^n$. Given a positive Radon measure μ on $\mathbb{R}^n$ and a non-negative ν-measurable function f we define

$$\mathcal{G}f(x) := \textstyle\int_M g(x,y)f(y)\,d\nu(y),\quad x\in\mathbb{R}^n, \tag{I.8a}$$

$$\check{\mathcal{G}}\mu(x) := \textstyle\int_{\mathbb{R}^n} g(x,y)\,d\mu(x),\quad y\in M. \tag{I.8b}$$

For $E\subset\mathbb{R}^n$ let Ω_E be a set of ***admissible functions,***

$$\Omega_E := \{f\in L^2(M,\nu):\ f\geq 0,\, \mathcal{G}f\geq 1\,\text{on}\,E\},$$

and the g***-capacity***

$$\mathrm{Cap}_g(E) := \inf\left\{\int_M f^2\,d\nu:\ f\in\Omega_E\right\}.$$

Defined as such Cap_g is an *outer capacity* and Borel sets are capacitable. For a measure μ on $\mathbb{R}^n$ let the g***-potential*** be

$$\mathbb{V}_g^\mu := \mathcal{G}\check{\mathcal{G}}(\mu).$$

We say that a property happens g-quasi everywhere, g-q.e., if it happens everywhere, but on a set of zero g-capacity.

Theorem I.1 (Frostman theorem for Cap_g) *Let $E\subset\mathbb{R}^n$ be a compact set. Then there exists a* *measure* μ_E

$$\operatorname{Cap}_g(E)^{\frac{1}{2}} = \sup\left\{\mu(E):\ \mu \geq 0,\ \operatorname{supp}\mu \subset E,\ \|\check{\mathcal{G}}(\mu)\|_{L^2(M,\nu)} \leq 1\right\}, \tag{I.9a}$$

$$\operatorname{Cap}_g(E) = \mu_E(E) =: |\mu_E| = \textstyle\int_{\mathbb{R}^n} \mathbb{V}_g^{\mu_E}\, d\mu, \tag{I.9b}$$

$$\mathbb{V}_g^{\mu_E}(x) \leq 1, \quad x \in \operatorname{supp}\mu_E, \tag{I.9c}$$

$$\mathbb{V}_g^{\mu_E}(x) \geq 1, \quad g\text{-q.e. on } E. \tag{I.9d}$$

Potential theory on graphs

In order to apply the general construction of previous paragraph we have to make a few more steps. Essentially we want to have our kernel function g to be defined on $\Gamma \times \Gamma$, or, in other words, to let M be a graph (usually d-tree with a positive weight attached). However the other space above is $\mathbb{R}^n$, and to make the argument work we will consider Γ to be a subset of $\mathbb{R}^n$ for some appropriate n, i.e. embed Γ into $\mathbb{R}^n$.

Finite graphs

When Γ is a finite graph, we almost do not have to do anything, and the embedding is just mapping the vertices to some separated set of points, say in $\mathbb{R}^2$. The only thing we have to remember is the ordering.

Infinite d-trees

Now assume $\Gamma = \overline{T}^d$ is a d-tree. We embed each coordinate in $\mathbb{R}^2$ in a standard way, and then take the Cartesian product. The idea for this (mostly procedural – throughout the work we interpret trees in a different way) tree embedding can be found in [64, Chapter 1] or in [6].
We identify the vertices of T with the approximating intervals for the classical Cantor set on the unit interval. Namely, consider the ternary Cantor set $E = \bigcap_{j=0}^{\infty} E_j$, where $E_0 = [0,1]$, and E_j consists of 2^j closed intervals of length 3^{-j}. Then each point of T corresponds to a unique interval in E_j (or, more precisely, to its centerpoint), and, similarly, ∂T maps to E. In other words, if c_{jk} is the centerpoint of a k-th segment in E_j, and α is a k-th point (counted, say, from left to right) on the level j of T, $|\alpha| = j$, we set

$$\Phi(\alpha) := p_{kj} = (c_{kj}, 3^{-k}) \in \mathbb{R}^2.$$

Extend Φ to the boundary ∂T by continuity, so that $\Phi(\partial T) = E$. Hence we have an injective mapping $\Phi : \overline{T} \to \Phi(\overline{T}) \subset \mathbb{R}^2$, with euclidean distance of images comparable to the original tree metric defined in (I.1)

$$|\Phi(\alpha) - \Phi(\beta)| \approx \operatorname{dist}(\alpha, \beta), \quad \alpha, \beta \in \overline{T}.$$

In the same vein the points of T^d correspond to ternary rectangles (Cartesian products of centerpoints of intervals in E_j^i). In particular, $(\partial T)^d$ can be identified with E^d. The predecessor and successor sets on $\Phi\left(\overline{T}^d\right)$ are inherited from the graph.

Applying general theory

We do everything for $\overline{T}^d$, the case of finite graphs is done similarly.
We set $M := T^d$ and $\nu := w$ to be a discrete measure on M. Let $\tau_y \in M = T^d$ and $x \in \Phi(\overline{T}^d)$. Our kernel is

$$g(x,\tau_y) := \mathbb{1}_{\mathcal{S}(x)}(\tau_y) = \mathbb{1}_{\mathcal{P}(\tau_y)}(x), \quad x \in \Phi\left(\overline{T}^d\right),\ y \in T^d. \tag{I.10}$$

On the rest of $\mathbb{R}^{2d}$ we set g to be infinity,

$$g(x,\tau_y) := +\infty, \quad x \notin \Phi\left(\overline{T}^d\right). \tag{I.11}$$

It is easy to check that g is lower semicontinuous in the first variable and measurable (w.r.t. discrete measure w which is weight) in the second. Also, if $x = \Phi(\tau_x) \in \Phi\left(\overline{T}^d\right)$ and $\tau_y \in T^d$, then for any non-negative f on T^d and a measure $\tilde{\mu}$ supported on $F\left(\overline{T}^d\right)$ one has

$$\mathcal{G}f(x) = \int_M g(x,\tau_y)f(\tau_y)\,d\nu(\tau_y) = \sum\nolimits_{\tau_y \geq \tau_x} f(\tau_y)w(\tau_y) = \mathbf{I}_w f(\tau_x), \tag{I.12a}$$

$$\check{\mathcal{G}}\tilde{\mu}(\tau_y) := \int_{\mathbb{R}^n} g(x,\tau_y)\,d\tilde{\mu}(x) = \int_{\mathcal{S}(\tau_y)} d\mu = \mathbf{I}^*\mu(\tau_y), \tag{I.12b}$$

where μ is the Φ-preimage of $\tilde{\mu}$. For $x \notin \Phi\left(\overline{T}^d\right)$ or for a measure $\tilde{\mu}$ with mass outside of $\Phi\left(\overline{T}^d\right)$ we get infinity in the expressions above. It turns out that it makes sense to consider only the measures supported on the image of d-tree. In any case now we can run Adams-Hedberg machinery to obtain the following discrete version of Frostman theorem.

Theorem I.2 *Assume $E \subset \overline{T}^d$ is a compact set and $w : T^d \to \mathbb{R}_+$ is a weight with $w(o) > 0$. Then there exists a unique measure μ_E such that*

$$|\mu_E| = \int_{\overline{T}^d} \mathbf{V}_w^{\mu_E}\,d\mu_E = \mathrm{Cap}_w(E), \tag{I.13a}$$

$$\mathbf{V}_w^{\mu_E} \leq 1 \quad \text{on} \quad \operatorname{supp}\mu_E \tag{I.13b}$$

$$\mathbf{V}_w^{\mu_E} \geq 1 \quad \text{q.e. on} \quad E. \tag{I.13c}$$

Same holds for a finite graph Γ.

Hardy embeddings

Formulation of the main Theorem I.3

Our main object of study here is the ***discrete Hardy embedding*** on the d-tree in the linear case. As we have already mentioned before, it can be considered from different points of view: the multi-linear weighted paraproduct theory (see below), multi-parametric Potential theory, random walks on d-trees (so random walks with multi-parametric time in sense of [17]), combinatorics of dyadic rectangles, polydiscs (including probabilistic interpretation like in [38]) etc. We have chosen to explore it relation to the problem of describing Carleson embeddings of weighted Hardy-Sobolev spaces on the polydisc – the problem that originally motivated us to study analysis on d-trees.

In this Section we describe the discrete problem, and we explain how to move to the continuous case in the next Section. The potential-theoretic meaning, while not taking central place in our arguments, is almost always at least present in the background, and sometimes we use this language to formulate the results and study the nature of arising problems from this angle.

We also would like to mention that our results can be viewed as the two-weight paraproduct estimates, see [112, Section 1] for details and references. Two-weight estimates for singular integrals were considered by F. Nazarov, S. Treil, and A. Volberg regarding dyadic singular operators and by M. Lacey, C.-Y. Shen, E. Sawyer, and I. Uriarte-Tuero regarding the Hilbert transform, see [74], [75], [55], [53], and the references therein. Another example is a recent paper by A. Iosevich, B. Krause, E. Sawyer, K. Taylor, and I. Uriarte-Tuero [45] on the two weight problem for the spherical maximal operator. Classically, an estimate of paraproduct tri-linear forms [37] is based on $T1$ theorem of G. David and J.-L. Journé. The theory of Carleson measures (or classical BMO theory) is involved. It is well known [25, 26, 49, 48] that in the multi-parametric setting all these results and concepts of Carleson measure, BMO, John–Nirenberg inequality, Calderón–Zygmund decomposition are much more delicate. Paper [72] develops a completely new approach to prove natural tri-linear bi-parameter estimates on bi-parameter paraproducts, and in [73] a simplified method was used to address the multi-parameter paraproducts.

Assume that d is fixed. We ask when, for a given ***weight-measure pair*** (w,μ), the operator $\mathbf{I}_w$ is bounded acting from $L^2(T^d,w)$ to $L^2(\overline{T}^d,\mu)$, i.e.

$$\int_{\overline{T}^d}(\mathbf{I}_w f)^2\,d\mu \le C\int_{T^d} f^2\,dw,\quad f\in L^2(T^d,w), \tag{I.14}$$

or, if we consider the dual embedding,

$$\int_{T^d}(\mathbf{I}^*_\mu\varphi)^2\,dw \le C\int_{\overline{T}^d} \varphi^2\,d\mu,\quad \varphi\in L^2(\overline{T}^d,\mu). \tag{I.15}$$

The smallest constant C such that the inequalities above hold for any appropriate function we call the ***Carleson embedding constant*** and denote by $[w,\mu]_{CE}$.

A standard way to obtain the conditions on w and μ for $[w,\mu]_{CE}$ to be finite is to test the embedding on a certain subclass of functions (usually characteristic functions of some sets). In doing this we arrive a number of test embeddings. All sets are assumed to be Borel.

The ***subcapacitary*** constant is the smallest number $[w,\mu]_{SC}$ such that

$$\mu(E)\le [w,\mu]_{SC}\,\mathrm{Cap}_w(E),\quad \forall E\subset \overline{T}^d. \tag{I.16}$$

The ***hereditary Carleson constant*** (or the ***restricted energy condition constant or REC constant***) is the smallest number $[w,\mu]_{HC}$ such that

$$\mathcal{E}_w(\mu|_E)=\sum_{\alpha\in T^d} w(\alpha)(\mathbf{I}^*\mu|_E)^2(\alpha)\le [w,\mu]_{HC}\mu(E),\quad \forall E\subset \overline{T}^d. \tag{I.17}$$

The ***Carleson constant*** (or the ***multiple box constant***) is the smallest number $[w,\mu]_C$ such that

$$\sum_{\alpha\in\mathcal{S}(E)} w(\alpha)(\mathbf{I}^*\mu)^2(\alpha) \leq [w,\mu]_C\mu(E), \quad \forall E\subset \overline{T}^d. \tag{I.18}$$

The ***box constant*** is the smallest number $[w,\mu]_B$ such that

$$\sum_{\alpha\leq\beta} w(\alpha)(\mathbf{I}^*\mu)^2(\alpha) \leq [w,\mu]_B\mathbf{I}^*\mu(\beta) = [w,\mu]_B\,\mu(\mathcal{S}(\beta)), \quad \forall\beta\in T^d. \tag{I.19}$$

All of these are clearly generated by appropriate tests: the subcapacitary condition is the direct embedding tested on functions admissible for E, the hereditary Carleson is the dual embedding tested on characteristic functions of sets, the Carleson constant appears when we remove some extra parts from the left-hand side of (2.3b), and the box constant is the Carleson constant with sets restricted to be descendants of singletons in T^d (which may be considered as dyadic parallelepipeds in $[0,1]^d$ – hence the naming convention).

For positive numbers A,B we write $A\lesssim B$ if $A\leq CB$ for some absolute constant C.

The inequalities

$$\begin{aligned}&[w,\mu]_B\leq [w,\mu]_C\leq [w,\mu]_{HC}\leq [w,\mu]_{CE},\\ &[w,\mu]_{SC}\leq [w,\mu]_{CE}\end{aligned}$$

are obvious (also it is not hard to see that $[w,\mu]_{SC}\leq [w,\mu]_{HC}$). The converse inequalities for $d=1$ are known (see, for example, [6] for potential-theoretic proof, or [74] and [108] for the proof via Bellman function). The main result states that for product weights we have the converse inequalities also for $d=2$ and $d=3$ ([112, Theorem 1.4]).

Theorem I.3 *Assume that (w,μ) is a weight-measure pair on the d-tree with $d=2$ or $d=3$, and w has a product structure, $w(\alpha)=\prod_{k=1}^d w(\alpha_k)$ for any $\alpha=(\alpha_1,\dots,\alpha_d)\in T^d$. Then*

$$\begin{aligned}&[w,\mu]_B\gtrsim [w,\mu]_C\gtrsim [w,\mu]_{HC}\gtrsim [w,\mu]_{CE},\\ &[w,\mu]_{SC}\gtrsim [w,\mu]_{CE}.\end{aligned} \tag{I.20}$$

Motivation for the choice of a weight and medium

The reason to study this particular problem originated from the Carleson embedding problem on the polydisc, as it is explained in the next paragraph. It is because of that, in particular, we are considering d-trees and weights of product nature. The path we have chosen to obtain this result is also influenced by this motivation – it is potential-theoretic, at least in spirit, and it follows the train of thought presented in works of D. Stegenga ([90]), E. Sawyer ([86]), N. Arcozzi, R. Rochberg, E. Sawyer and B. Wick ([4], [5], [6]), to name a few, and which goes back to V.G. Maz'ya.

However, while the set of ideas we adopted eventually made it possible to lay down the road to the result, we had to reinvent most of the arguments along the way. The reason for this lies of course in the multi-parametric nature of our setting. This nature can be explained in a number of ways, like the d-dimensional time or presence of non-radial kernels, in particular we observe that the d-tree is not a tree – this directed graph has a lot of cycles. Hence the geodesics are not unique anymore, and the geometric structure becomes much more complicated. We also can not disintegrate the problem to d one-dimensional statements, since while the weight w is product, the measure μ is not. Still the product property of w helps immensely, and we will see later that a general two-weight problem (with general w and μ) is even more difficult, even on subgraphs of a d-tree.

These difficulties are also hidden in the formulation of equivalent test conditions. In considering the multi-parametric embeddings the rule of thumb is that the test should be more involved, compared to $d = 1$. Not only both the direct and the dual embedding should be tested, but also the test functions have to be more complicated. This, superficially, is not the case here, since the reverse inequalities (I.20) look exactly like what one expects from the case $d = 1$. However this is not completely true. First, we deal with product weights, so some one-dimensional structure has to be preserved. In a proper $2D$-setting, even for simple subgraphs of T^2, we have to have several single-box or multiple-box test, as evidenced in [86] (and nothing is known for T^3). Another reason is that in (I.20), specifically in the $[w, \mu]_B \gtrsim [w, \mu]_{CE}$ a small miracle happens – just one single-box test is enough. We attribute this to the tangible presence of Potential Theory in the structure of our problem, since, for instance, for the Carleson embedding on the bi-disc it is not true anymore (as is shown in [21], [24]). We can not emulate this embedding with a product weight – the capacity does not make sense then, and the proper discrete analogue has the roles of w and μ reversed.

We also want to mention that the restriction $d \leq 3$ here is important. We do not know if the Theorem I.3 is true in higher dimensions (we believe so, though), and the techniques for its proof should be reinvented once more.

Scheme of the proof

We prove Theorem I.3 in Chapter 2. The key argument in the proof of is the so-called ***Surrogate Maximum Principle*** which replaces the usual Maximum Principle which is absent in the multi-parametric setting. Actually we show that the main Theorem holds for every d modulo establishing said Principle.

Section 2.1

There we recall the formulation of the main Theorem.

Section 2.2

This Section is dedicated to the proof of the Surrogate Maximum Principle for $d \leq 3$. Let us elaborate.

In Section 2.2.1 we prove some auxiliary results on a usual dyadic tree.

In Sections 2.2.2 and 2.2.3 we extend these results to the 2-tree and 3-tree respectively. We do it separately, since we want to showcase that another jump of complexity happens when we move from $d = 2$ to $d = 3$, and yet another one when we consider higher dimensions. One would expect that all dimensions above 2 behave similarly, but this is not the case. Two key results are proven in these Sections. The first tells us how to rearrange the mass of a function on T^d in a more effective way, if it is supported on a set of small potential.

Lemma I.1 (**Small energy majorization** on T^d for $d = 2, 3$.) *Let $d = 2$ or $d = 3$, and $f : T^d \to \mathbb{R}_+$ be a superadditive (in each variable) function, and w be a product weight. Suppose that* $\operatorname{supp} f \subseteq \{\mathbf{I}(wf) \le \delta\}$*, and let $\lambda \ge 4\delta$. Then there exists an energy-effective redistribution $\varphi : T^d \to \mathbb{R}_+$ such that*

$$\mathbf{I}(w\varphi)(\alpha) \ge \mathbf{I}(wf)(\alpha), \quad \alpha \in \{\lambda \le \mathbf{I}(wf) \le 2\lambda\}, \tag{I.21a}$$

$$\textstyle\int_{T^d} \varphi^2 w \le C \left(\frac{\delta}{\lambda}\right)^{c(d)} \int_{T^d} f^2 w, \tag{I.21b}$$

where C is some absolute constant, and $c(2) = 2$, $c(3) = 1$.

The second key result considers the application of small energy majorization to estimate the capacity of exceptional sets – the sets where the equilibrium potential is very large (see [105, Theorem 1.11], [2, Lemma 3.1]).

Theorem I.4 *Let μ be a measure on T^d for $d = 2$ or $d = 3$, w a product weight and $\mathbf{V}^\mu_w \le 1$ on* $\operatorname{supp}\mu$*. Let $E_\lambda := \{\mathbf{V}^\mu_w \ge \lambda \ge 10\}$. Then*

$$\operatorname{Cap}_w E_\lambda \le \frac{C}{\lambda^{c(d)}} \mathcal{E}_w[\mu],$$

where C is an absolute constant and $c(2) = 4$, $c(3) = 3$.

Finally, in Section 2.2.4 we formulate the Surrogate Maximum principle for d-trees – it is a conjecture for $d \ge 4$ (see [107, Theorem 1.2]).

Theorem I.5 *We say that a weight w on a d-tree T^d satisfies **the surrogate maximum principle**, if for some $\kappa > 0$, $C < \infty$ and every positive measures $\mu, \rho : T^d \to [0, \infty)$ and $\delta > 0$ one has*

$$\int_{T^d} \mathbf{V}^\mu_{w,\delta}\, d\rho \le C\, (\delta|\rho|)^\kappa \left(\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho]\right)^{\frac{1-\kappa}{2}}. \tag{I.22}$$

For $d = 1, 2, 3$ every weight of product form satisfies this principle with $\kappa = \frac{1}{d}$ and C independent of w.

Conjecture 0.0.1 *Let w be of product type. Then w satisfies the surrogate maximum principle with $\kappa = \frac{1}{d}$ and $C = C(d)$ independent of w.*

Remark. As a corollary of the Surrogate Maximum Principle we also obtain the tail energy estimates for potentials on d-tree, we just substitute $\rho = \mu$ into SMP.

Section 2.3

In this Section we prove one of the test conditions in Theorem I.3 – the subcapacitary condition

$$[w,\mu]_{SC} \gtrsim [w,\mu]_{CE}.$$

In order to do this we show another important estimate – the ***Strong Capacitary Inequality*** on a d-tree.

Theorem I.6 *Let w on T^d satisfy the Surrogate Maximum Principle, and let $f : t^d \to \mathbb{R}_+$. Then*

$$\sum_{k\in\mathbb{Z}} 2^{2k}\,\mathrm{Cap}_w\left(\{\alpha \in T^d:\ \mathbf{I}_w f(\alpha) > 2^k\}\right) \lesssim \int_{T^d} f^2 w. \tag{I.23}$$

This is a multi-parametric version of the famous theorem by Maz'ya. Its (very simple) tree analogue is proven in Section 1.2.

Section 2.4

In this Section we prove another two test conditions in Theorem I.3 – the Carleson and Hereditary Carleson conditions (the second serves mostly as a procedural step for the first one)

$$[w,\mu]_{C} \gtrsim [w,\mu]_{HC} \gtrsim [w,\mu]_{CE}.$$

The left inequality above we show via the subcapacitary condition – it provides a nice short-cut.

Section 2.5

In this Section we prove the last remaining test condition in Theorem I.3 – the Box (or the Single Box Test) condition

$$[w,\mu]_{B} \gtrsim [w,\mu]_{CE}.$$

This is the most difficult and also the most surprising out of test conditions. As we have already mentioned above, single box test is not something one would expect to happen in several parameters, but still it does hold true.

Section 2.6

The last Section of Chapter 2 consists several examples and counterexamples that provide more information about our previous results, see [110] for a detailed exposition. Namely we explain:

- that it is enough to work on finite d-trees (which is what we were doing throughout the whole chapter) and how to pass to the limit in the d-tree depth;
- why the product structure of w is important for Theorem I.3 – one has to test more than just a direct or just a dual embedding (counterexamples from [106], Proposition 1.1);

- what problems arise when we try to pass to the case $d = 4$, and what kind of peculiar and unpleasant behaviour can have the tail energy of a measure even on T^2 with unit weight.

Carleson measures for weighted Hardy-Sobolev spaces

Hilbert spaces of analytic functions in polydisc

Let $d \in \mathbb{N}$ and $\vec{s} = (s_1, \ldots, s_d) \in \mathbb{R}^d$. We say that an analytic on the unit polydisc $\mathbb{D}^d$ function f belongs to the weighted Hardy-Sobolev space $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$, if

$$f(z) = \sum_{k \in (\mathbb{Z}_+)^d} c_k z^k, \quad z = (z_1, \ldots, z_d) \in \mathbb{D}^d,$$

and the Taylor coefficients $c_k = \hat{f}(k)$ of f satisfy

$$\sum_{k \in (\mathbb{Z}_+)^d} |c_k|^2 (k_1 + 1)^{s_1} \cdot \cdots \cdot (k_d + 1)^{s_d} < +\infty,$$

where $k = (k_1, \ldots, k_d),\ k_j \in \mathbb{Z}_+$ is the multi-index. The square root of the left-hand side of the expression above we call the $\vec{s}$-Hardy-Sobolev norm of f,

$$\|f\|^2_{\mathcal{H}_{\vec{s}}(\mathbb{D}^d)} := \sum_{k \in (\mathbb{Z}_+)^d} |c_k|^2 (k_1 + 1)^{s_1} \cdot \cdots \cdot (k_d + 1)^{s_d}.$$

Clearly,

$$\mathcal{H}_{\vec{s}}(\mathbb{D}^d) = \bigotimes_{j=1}^{d} \mathcal{H}_{s_j}(\mathbb{D}). \tag{I.24}$$

We also introduce the harmonic versions of such spaces, which we define through via tensor product for the sake of brevity.

$$\begin{aligned}
&\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d) = \bigotimes_{j=1}^{d} \mathcal{H}^h_{s_j}(\mathbb{D}),\\
&\mathcal{H}^h_{s_j}(\mathbb{D}) = \left\{ f(z) = \sum_{k \geq 0} \hat{f}(k) z^k + \sum_{k < 0} \hat{f}(k) \bar{z}^k : \sum_{k \in \mathbb{Z}} |hatf(k)|^2 (|k| + 1)^{s_j} < +\infty \right\}.
\end{aligned} \tag{I.25}$$

This is a kind of umbrella definition which includes a number of famous classical spaces. The first one to mention is of course the Hardy space on the unit disc:

$$\begin{aligned}
&H^2(\mathbb{D}) := \{f(z) = \sum_{k \geq 0} c_k z^k : \|f\|^2_{H^2} < +\infty\},\\
&\|f\|^2_{H^2} = \frac{1}{2\pi} \sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta,
\end{aligned}$$

which corresponds to the situation $d = 1,\ s = 0$. The next one is the (unweighted) Dirichlet space on $\mathbb{D}$

$$\mathcal{D}(\mathbb{D}) := \{f(z) = \sum_{k\geq 0} c_k z^k :\ \|f\|^2_{\mathcal{D}} < +\infty\},$$
$$\|f\|^2_{\mathcal{D}} = \|f\|^2_{H^2} + \frac{1}{\pi}\int_{\mathbb{D}} |f'(z)|^2 dA,$$

which is covered by $d = 1,\ s = 1$.

There is an immense amount of literature on Hardy space – it is probably the most well studied space in Complex Analysis. The Dirichlet space is slightly less popular, but still it is a subject of a large stream of books and papers (we would like to mention the recent book by Arcozzi, Rochberg, Sawyer and Wick [7]), also it enjoys being an analytic version of another classical Sobolev $W^{1,2}$ space. The polydisc versions are much less understood, with Hardy space again being the most well-researched.

In view of our definition we prefer to consider them as members of a scale of Hilbert spaces of analytic functions, which is obtained by varying the parameter $\vec{s}$. We restrict ourselves to $\vec{s} \in [0,1]^d$ (so Bergman spaces are not considered), moreover, the Hardy endpoint of the scale we only mention when discussing some counterexamples and open problems. The reason is that we intend to work with Potential Theory (at least in spirit), which does not survive taking $\vec{s}$ across zero.

Another interesting class of spaces is obtained when we take a diagonal version of $H_{\vec{s}}$ (especially $H_{\vec{1}}$ – see [102] for discussions regarding this space), i.e. we consider functions on the *unit disc* by taking all the variables z_k to be the same $z \in \mathbb{D}$.

Carleson measures: formulation of the main result

Our main object of study here are the so-called ***Carleson measures*** for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$.

Recall that a measure μ on $\mathbb{D}^d$ is called ***Carleson*** for $\mathcal{H}_{\vec{s}}$, if

$$\int_{\mathbb{D}^d} |f(z)|^2\, d\mu(z) \leq C_\mu \|f\|^2_{\mathcal{H}_{\vec{s}}(\mathbb{D}^d)}, \quad \forall f \in \mathcal{H}_{\vec{s}}(\mathbb{D}^d), \tag{I.26}$$

or, in other words, if the embedding $Id : \mathcal{H}_{\vec{s}}(\mathbb{D}^d) \mapsto L^2(\mathbb{D}^d, \mu)$ is bounded. We can extend this definition to measures on a closed polydisc by taking

$$\sup_{0<r<1} \int_{\mathbb{D}^d} |f(rz)|^2\, d\mu(z)$$

instead of the left-hand side of (I.26). There are numerous results that describe Carleson measures for various spaces of analytic and harmonic functions, we mention again the original paper [19] and the stream of literature that followed, [90], [40], [59], [24], [25], [26], [4], [5] and many others. Carleson embeddings can also be viewed in terms of trace inequalities for Sobolev spaces (see, say [50] and [68], the reference corps is quite vast here as well).

Generally speaking a typical description of a Carleson measure says that it should satisfy a testing condition on a Carleson tent erected over a boundary interval or a boundary set.
Our main result on Carleson measures follows this idea as well. We give a description of Carleson measures for the analytic space $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$ with $d = 1, 2, 3$ and $\vec{s}$ close to $\vec{1}$ (so our spaces are close to the unweighted Dirichlet space). We also describe Carleson measures for the harmonic version $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d)$ with the same d, but now with less restrictions on $\vec{s}$. This is covered in the following two theorems (the first one is from [107] and [112]).

Theorem I.7 *Let $\vec{s} = (s_1, \dots, s_d)$, $s_j \in (0,1], j = 1, \dots, d$, $1 \le d \le 3$, such that all s_i are sufficiently close to 1: $1 - s_j \le \varepsilon_d$, for a certain positive absolute $\varepsilon = \varepsilon(d)$ and $j = 1, \dots, d$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$, if and only if one of the following conditions holds true*

$$\nu(T(E)) \lesssim \mathrm{Cap}_{\vec{s}}(E), \quad E \subset \mathbb{T}^d \tag{I.27a}$$
$$\textstyle\sum_{R \subset E} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } E, \tag{I.27b}$$
$$\textstyle\sum_{R \subset Q} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } Q. \tag{I.27c}$$

Here Q, R are dyadic rectangles on the (poly) torus $\mathbb{T}^d$, and $T(Q)$ is the usual tent area above Q, while E is any finite union of such rectangles, and $T(E)$ is the union of respective tents.

Theorem I.8 *Let $\vec{s} = (s_1, \dots, s_d)$, $s_j \in (0,1], j = 1, \dots, d$, $1 \le d \le 3$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}^h_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d)$, if and only if one of the following conditions holds true*

$$\nu(T(E)) \lesssim \mathrm{Cap}_{\vec{s}}(E), \quad E \subset \mathbb{T}^d \tag{I.28a}$$
$$\textstyle\sum_{R \subset E} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } E, \tag{I.28b}$$
$$\textstyle\sum_{R \subset Q} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } Q. \tag{I.28c}$$

These Theorems are proven in Chapter 3.

Scheme of the proof

We deduce Theorems I.7 and I.8 from Theorem I.3 via discretization argument.

Discretization

In Section 3.1 we describe a method of discretizing our problem, which we then can attack with the help of the Hardy embedding. We note, that we *do not discretize the space $\mathcal{H}_{\vec{s}}$*, but rather we discretize the Carleson embedding (I.26) using the Reproducing Kernel property of our spaces. This turns out to be a less direct but a more convenient approach.
Remark. The discretization argument and the reproducing kernel trick are borrowed from [6].

In Section 3.1.1 we consider a discrete model of d-dimensional polydisc which turns out to be a d-tree. There we have to take some care regarding the difference between hyperbolic geometry on trees and in the continuous medium. We handle it by introducing some additional structure on T^d.

In Section 3.1.2 we use the *Reproducing Kernel* technique to relate the discrete weighted potential to the adjoint Carleson embedding. It is here that the analytic/harmonic nature of our continuous space becomes important, depending on the situation we have to reduce the available values of parameter $\vec{s}$ to be able to estimate the Reproducing Kernel. The main result of this section is the following theorem

Theorem I.9 *Let $\vec{s} = (s_1, \ldots, s_d)$, $s_j \in (0,1], j = 1, \ldots, d$, $d \geq 1$, such that all s_i are sufficiently close to 1: $1 - s_j \leq \varepsilon_d$, for a certain positive absolute $\varepsilon = \varepsilon(d)$ and $j = 1, \ldots, d$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$, if and only if $(w_{\vec{s}}, \tilde{\nu})$ is a trace weight-measure pair for $\overline{T}^d$,*

$$\sum_{\alpha \in T^d} (\mathbf{I}^* \psi \tilde{\nu})^2(\alpha) w_{\vec{s}}(\alpha) \leq C \int_{T^d} \psi^2 d\tilde{\nu}, \quad \forall \psi \in L^2(\overline{T}^d, \tilde{\nu}) \,. \tag{I.29}$$

Here $\tilde{\nu}$ is the discrete image of ν on $\overline{T}^d$.
The same holds true for the harmonic space $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d)$, only we do not require $\vec{s}$ to be close to $\vec{1}$.

Remark. It is obvious that for $d = 1$ the harmonic and analytic spaces are almost the same, at least in structure. Namely,

$$\mathcal{H}^h_s(\mathbb{D}) = \mathcal{H}_s(\mathbb{D}) \bigoplus \overline{\mathcal{H}_s(\mathbb{D})}.$$

Thus their Carleson measures are the same. In higher dimensions the analytic space is much smaller than the harmonic (the Fourier coefficients of its element lie only on a small part of $\mathbb{Z}^d$ – the positive octant $(\mathbb{Z}_+)^d$, and conjugation does not help any more). We mention this difference several times, see also the question in Section 8.7.

Capacitary condition

In Section 3.2 we concentrate on the equivalence between subcapacitary conditions for the Carleson embedding on $\mathbb{D}^d$ and Hardy embedding on T^d. There are ways to establish continuous versions of energy conditions as well (see, for example, [102, Theorem 2] for one-dimensional diagonal Dirichlet space), however the subcapacitary one is the most accessible in this case. In order to formulate such condition in terms of continuous capacities we have to establish the equivalence between d-tree and multi-parametric Bessel capacities on the distinguished boundary $\mathbb{T}^d$. This can be viewed as a variation on results and techniques from [6], [12], [62], [63], [64], where similar results were obtained for trees and metric spaces.

Let us define the ***Bessel multi-parametric capacity*** that we are referring to. We use again the [1]-scheme presented above. Assume that d and $\vec{s} = (s_1, \ldots, s_d)$ are fixed. Now we consider our kernel g to be defined on $\mathbb{T}^d \times \mathbb{T}^d$ (polytorus is naturally embedded into $\mathbb{R}^{2d}$). We set

$$
\begin{aligned}
g_{\vec{s}}(z,\zeta) &:= \prod_{k=1}^{d} g_{s_j}(z_j,\zeta_j), \quad z,\zeta \in \mathbb{T}^d, \\
g_{s_j}(z_j,\zeta_j) &:= \frac{1}{|z_j-\zeta_j|^{1-\frac{s_j}{2}}}, \quad z_j,\zeta_j \in \mathbb{T}.
\end{aligned}
\tag{I.30}
$$

The measure on $M = \mathbb{T}^d$ is just the Lebesgue measure. As a result, for $f : \mathbb{T}^d \to \mathbb{R}_+$ and Radon measure μ on $\mathbb{T}^d$ we have

$$\mathcal{G}f(z) = \textstyle\int_{\mathbb{T}^d} g_{\vec{s}}(z,\zeta) f(\zeta)\, d\nu(\zeta), \tag{I.31a}$$

$$\check{\mathcal{G}}\mu(\zeta) := \textstyle\int_{\mathbb{T}^d} g_{\vec{s}}(z,\zeta)\, d\mu(z). \tag{I.31b}$$

It is not hard to see (by comparing with classical Bessel kernels) that

$$\mathbf{U}^{\mu}_{\vec{s}}(z) := \mathcal{G}\check{\mathcal{G}}\mu(z) = \int_{\mathbb{T}^d\times\mathbb{T}^d} g(z,\zeta)g(\zeta,\tau)d\mu(\zeta)\, dm(\tau) \approx \int_{\mathbb{T}^d} \mathbb{K}_{\vec{s}}(z,\zeta)\, d\,\mu(\zeta), \tag{I.32}$$

where $dm(\tau)$ is just the Lebesgue measure on the polytorus, and

$$
\begin{aligned}
\mathbb{K}_{\vec{s}}(z,\zeta) &= \prod_{k=1}^{d} \mathbb{K}_{s_k}(z_k,\zeta_k), \\
\mathbb{K}_s(z_k,\zeta_k) &= \frac{1}{|z_k-\zeta_k|^{1-s}}, \quad s<1, \\
\mathbb{K}_1(z_k,\zeta_k) &= \log\left(\frac{1}{|z_k-\zeta_k|}\right).
\end{aligned}
\tag{I.33}
$$

Now for a compact set $F \subset \mathbb{T}^d$ we can define $\vec{s}$***-capacity*** by, say,

$$\mathrm{Cap}_{\vec{s}}(F) = \inf\left\{\int_{\mathbb{T}^d} \mathbf{U}^{\mu}_{\vec{s}}\, d\mu : \ \mathbf{U}^{\mu}_{\vec{s}} \geq 1 \text{ on } F\right\}.$$

A polytorus version of Frostman's theorem is as follows.

Theorem I.10 *Assume $F \subset \mathbb{T}^d$ is a compact set. Then there exists a unique measure μ_F such that*

$$\mathcal{E}_{\vec{s}}[\mu_F] := |\mu_F| = \textstyle\int_{\mathbb{T}^d} \mathbf{U}^{\mu_F}_{\vec{s}}\, d\mu_F = \mathrm{Cap}_{\vec{s}}(F), \tag{I.34a}$$

$$\mathbf{U}^{\mu_F}_{\vec{s}} \leq 1 \quad \text{on} \quad \operatorname{supp}\mu_F \tag{I.34b}$$

$$\mathbf{U}^{\mu_F}_{\vec{s}} \geq 1 \quad \text{q.e. on} \quad F. \tag{I.34c}$$

Now that we have defined the capacity on $\mathbb{T}^d$ we build the tools to establish the relation between continuous and discrete capacities. The first one says that for ***exponential product weights*** the capacity of a set and its boundary projection are similar. This allows us not to bother with estimating capacity in the interior, be it $\mathbb{D}^d$ or T^d. Let $\vec{s} \in (0,1]^d$ be as above, and $w_{\vec{s}}$ be a product weight on T^d

$$w_{\vec{s}}(\alpha) := 2^{\sum_{j=1}^{d}(1-s_j)|\alpha_j|}, \quad \alpha = (\alpha_1, \ldots, \alpha_j) \in T^d, \tag{I.35}$$

where $|\alpha_j| = \#\mathcal{P}(\alpha_j) - 1$.

Theorem I.11 *Let $E \subset \overline{T}^d$ and $w = w_{\vec{s}}$ be the product weight generated by $\mathcal{H}_{\vec{s}}$. Then the capacity of E and its boundary projection are equivalent,*

$$\mathrm{Cap}_{w_{\vec{s}}}(E) \approx \mathrm{Cap}_{w_{\vec{s}}}(\partial\mathcal{S}(E)),$$

where $\partial\mathcal{S}(E) = \{\omega \in (\partial T)^d : \mathcal{P}(\omega) \cap E \neq \emptyset\}$ and the constant in the inequality depends on $\vec{s}$ and d only (it blows up with $s_j \to 0$ for any of s_j).

After that we prove the equivalence theorem for boundary capacities.

Theorem I.12 *Let $w := w_{\vec{s}} : T^d \to \mathbb{R}_+$ be an exponential product weight, and assume that $E \subset (\partial T)^d$ is a compact set. Then the respective capacities of F and its polytorus image $F := \Lambda(E)$ are equivalent*

$$\mathrm{Cap}_{\vec{s}}(F) \approx \mathrm{Cap}_{w_{\vec{s}}}(E), \tag{I.36}$$

and the constant depends only on d and $\vec{s}$.

Growth spaces

We proceed to the second part of the thesis. Here we have a tonal change of sorts, since now we move from functional-analytic approach, where we consider discretization of the embedding (and of the function space $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$ itself in a sense), to the estimates and discretizations of individual functions. In addition, we are more focused on the discretizing process itself. Also, while in the previous arguments our efforts were more concentrated on the graph side of the discussion, where we mostly conducted our attack on the problem, here we are shifting to the representation of a harmonic function in a discrete way (be it wavelets or a martingales), and the discrete machinery is already mostly available.

In Chapter 4 we give a wavelet representation of a harmonic function in the growth class.
In Chapter 5 we continue studying growth functions, now on Lipschitz domains in $\mathbb{R}^d$ and prove a LIL-type result about their boundary oscillation. Its upper-half space version can be deduced from the wavelet decomposition result (and it was done in [99]), but we give a separate proof with a different discretization technique.
In Chapter 6 we consider divided differences of Hölder functions (which behave very much like the growth functions), and provide a couple of counterxamples.
Finally, in Chapter 7 we give a proof of M. Cartwright's theorem for the unit ball in $\mathbb{R}^d$.
The results and statements are formulated in the following sections.

Wavelet representation of functions from growth spaces

Now let us have an overview of our first result on the growth spaces. We still are in the context of several variables, $\mathbb{R}^{d+1}$, however it is not a true multi-parametric situation any more, so our discrete models are tree-like.

Let $w : \mathbb{R}_+ \to \mathbb{R}_+$ be a ***doubling weight***, i.e. a continuous decreasing function, $\lim_{t\to 0+} w(t) = +\infty$, $w(t) = 1$ when $t > 1$, that satisfies the doubling condition

$$w(t) \leq Dw(2t). \tag{I.37}$$

We consider harmonic functions in $\mathbb{R}^{d+1}_+$ with the following growth restriction

$$|u(x,t)| \leq Kw(t), \quad \text{where } (x,t) \in \mathbb{R}^{d+1}_+$$

The space of these functions – ***the growth space*** – is denoted by $h^\infty_v(\mathbb{R}^{d+1})$ and the least K for which the inequality above is satisfied is called the norm of u in h^∞_v, we denote it by $\|u\|_{v,\infty}$. We note that such a harmonic function is bounded in any half-space

$$\mathbb{R}^{d+1}_\delta = \{(x,t) \in \mathbb{R}^{d+1}, t \geq \delta > 0\} \tag{I.38}$$

and thus can be represented there by the Poisson integral of its values on the hyperplane $\{t = \delta\}$. We denote by h^0_v the subspace of h^∞_v consisting of functions u such that $u(x,t) = o(w(t))$ $(t \to 0)$ uniformly in $x \in \mathbb{R}^d$.

A brief history of such spaces includes works by A. Shields and D. Williams ([88] and [89]), G. Bennett, D. Stegenga and R. Timoney ([11]), B. Korenblum ([51]), W. Lusky ([61], [60]), A. Borichev, Yu. Lyubarskii, E. Malinnikova, P. Thomas ([14], [65]), K. Seip ([87]).

We intend to provide a description of functions in the growth spaces in terms of their boundary wavelet (multiresolutional) approximation. When the weight grows faster than t^{-a} for some a, our description is in terms of the wavelet coefficients; for slow growing weights we consider partial sums of the wavelet series.

Multiresolutional approximation: notation

We consider an ***r-regular multiresolution approximation*** $\{V_j\}$ of $L^2(\mathbb{R}^d)$, [69, ch 2.2], where $r \geq r_0(v)$ will be specified later. Then there exists $\phi \in V_0$ that satisfies

$$|\partial^\alpha \phi(x)| \leq C_N(1 + |x|)^{-N},$$

for any α such that $|\alpha| \leq r$ and every $N \in \mathbf{N}$, and $\{\phi(x-k), k \in \mathbb{Z}^d\}$ form an orthonormal basis for V_0. Further, there exists a collection of smooth (of class C^r) functions $\{\psi_p\}_{p=1}^q$ that form an orthonormal basis for $V_1 \ominus V_0$, decrease rapidly with all its derivatives of order up to r and satisfy the cancellation property. Then

$$\psi_{p,jk} = 2^{dj/2}\psi_p(2^j x - k), \quad j \in \mathbb{Z}, k \in \mathbb{Z}^d,\ p = 1, ..., q,$$

is an ***orthonormal wavelet basis*** in $L^2(\mathbb{R}^d)$, [69, ch 3.6]. We will use the orthonormal basis $\{\phi(x-k)\}_{k\in\mathbb{Z}^d} \cup \{\psi_{p,jk}\}_{1\le p\le q, j\ge 0, k\in\mathbb{Z}^d}$. For any function $f \in L^\infty(\mathbb{R}^d)$ we define ***the wavelet coefficients***

$$c_{p,jk}(f) = \int_{\mathbb{R}^d} f(x)\overline{\psi_{p,jk}(x)}dx, \quad j \ge 0, k \in \mathbb{Z}^d,\ p = 1, \dots, q,$$

and

$$b_k(f) = \int_{\mathbb{R}^d} f(x)\overline{\phi(x-k)}dx, \quad k \in \mathbb{Z}^d.$$

The N-th partial sum of the wavelet decomposition of f is

$$S_N(f)(x) = \sum_k b_k(f)\phi(x-k) + \sum_{p=1}^{q}\sum_{j=0}^{N}\sum_k c_{p,jk}\psi_{p,jk}(x).$$

Next, we define the *multiresolutional blocks.* Given a doubling weight v, we choose A large enough and define a sequence of integers $\{n_l\}$ such that $n_0 = 0$, $n_l > n_{l-1}$ and $w(2^{-n_l}) \in [A^l, A^{l+1})$. There exists m^* that depends on v only that satisfies

$$\frac{2^{-m^* n_l} w(2^{-n_l})}{2^{-m^* n_{l-1}} w(2^{-n_{l-1}})} < 1 - \varepsilon \tag{I.39}$$

for some positive ε. The idea is that instead of dealing with multiresolutional subspaces V_j separately we instead group them into blocks $V_{n_l} \setminus V_{n_{l-1}}$, and estimate block partial sums.

The main result is

Theorem I.13 *Let $u(x,t)$ be a harmonic function on $\mathbb{R}^{d+1}_+$ bounded on each half-space $\{(x,t) : t > t_0 > 0\}$. Then $u \in h^\infty_v$ if and only if there exists C such that*

$$M_N(u) = \sup_{t>0} \|S_N(u(\cdot,t))\|_{L^\infty(\mathbb{R}^d)} \le Cw(2^{-N}).$$

Similarly, $u \in h^0_v$ if and only if $\lim_{N\to\infty} M_N(u)(w(2^{-N}))^{-1} = 0$.

The proof of the theorem above combines standard tools of multiresolution analysis with a clever argument of J. Bourgain, [15], that allows one to squeeze a convolution with the appropriate Poisson kernel. This trick was also used in [98] (Theorem 1 and Corollary 3.1 there).

We also mention another result from [99]. Let g be a non-zero radial function in $\mathbb{R}^d$ such that $g \in C^r$, where r is large enough $r > r_0(v)$ (basically r_0 is the rate of growth of v). Assume also that g with all its partial derivatives of order up to r satisfies

$$|\partial^\beta g(x)| \le \frac{C}{(1+|x|^2)^{d+1}}.$$

For example g with compact support will work. Then $(1+|x|^{d+1})\partial^\beta g \in L^1(\mathbb{R}^d)$ when $|\beta| \le r$. We have

$$|\hat{g}(\tau)| \le \frac{C}{(1+|\tau|)^r}, \tag{I.40}$$

and similar estimates hold for partial derivatives of $\hat{g}$ up to order $d+1$.

Theorem I.14 *Let u be a harmonic function on $\mathbb{R}^{d+1}_+$ that is bounded in each half-space $\{(x,t) : t \ge t_0 > 0\}$ and let $g \in L^1(\mathbb{R}^d)$ be a radial function such that $\hat{g}$ has derivatives in $L^1(\mathbb{R}^d)$ up to order $d+1$, (I.40) holds for $\hat{g}$ and its derivatives, and $\hat{g}(0) \ne 0$. Then $u \in h^\infty_v$ if and only if there exists a constant C_u such that*

$$\left| \int_{\mathbb{R}^d} u(x,t) g\left(\frac{y-x}{a}\right) dx \right| \le C_u a^d w(a), \tag{I.41}$$

for all $t > 0, a > 0$ and $y \in \mathbb{R}^d$.

Remark. Basically this Theorem says that one can replace Poisson kernel in the convolution formula for u with something else, which is smooth enough. For example, a Steklov kernel convoluted with itself r times would work.

There is a general heuristics that goes more or less like this: if a harmonic function has a good enough growth estimate (like a Bloch function, or one investigated by Korenblum in [51]), then it has a natural dyadic martingale equivalent, which satisfies similar growth conditions. If we replace a martingale (so a Haar wavelet) by something more smooth, then we can obtain the same type of result with relaxed restrictions on growth. Sometimes it helps, since some of the properties are not readily transferable from dyadic martingales to the functions they model (and we will see such an example a bit later), on the other hand smooth wavelets have a certain downside – the supports of wavelets of the same rank must intersect a lot, so in order to use dyadic martingale machinery one has to do further work, this time with wavelet series.

Scheme of the proof

In Section 4.1 we collect some Lemmas and wavelet knowledge.

First, in Section 4.1.1 we prove some auxiliary statements, which are essentially a development of a trick used by Bourgain in [15, 16].

Following Y. Meyer [69] we then collect necessary results on multiresolution analysis in Sections 4.1.2 and 4.1.3.

In section 4.2 we prove Theorem I.13 – we work with blocks of wavelets of consecutive generations, where the size of the block depends on the weight function v. The proof is based on two following theorems that cover each direction of Theorem I.13. They are proved in Sections 4.2.1 and 4.2.2 respectively.

Theorem I.15 *For any $u \in h^\infty_v(\mathbb{R}^{d+1}_+)$ we define*

$$g_0(x,t) = \sum_{k \in \mathbb{Z}^d} \langle u(y,t), \phi(y-k) \rangle \phi(x-k), \quad \textit{and}$$

$$g_l(x,t) = \sum_{j=n_{l-1}+1}^{n_l} \sum_{p=1}^{q} \sum_{k\in\mathbb{Z}^d} \langle u(y,t), \psi_{p,jk}(y)\rangle \psi_{p,jk}(x), \quad l \geq 1.$$

Then

$$u(x,t) = \sum_{l=0}^{\infty} g_l(x,t), \quad g_l(\cdot,t) \in V_{n_l}(\infty) \quad \textit{and}$$

$$\|g_l(\cdot,t)\|_\infty \leq C\|u\|_{v,\infty} w(2^{-n_l}), \quad l \geq 0, \tag{I.42}$$

where C depends on ϕ and A only.

In other words, the wavelet decomposition block (of size depending on the weight itself) of a function in the growth space admits the same growth estimate. The converse is also true – if all such blocks have proper growth, then their sum, which is the function, belongs to the growth space.

Theorem I.16 *Let u be a harmonic function in $\mathbb{R}^{d+1}_+$ that is bounded on each half-space $\{(x,t) \in \mathbb{R}^{d+1}, t \geq t_0 > 0\}$. Suppose that for each $t > 0$*

$$u(x,t) = \sum_{l=0}^{\infty} g_l(x,t),$$

where the series converges uniformly on $\mathbb{R}^d$, $g_0(\cdot,t) \in V_0(\infty)$,

$$g_l(x,t) = \sum_{j=n_{l-1}+1}^{n_l} \sum_{p=1}^{q} \sum_{k\in\mathbb{Z}^d} a_p^{(jk)}(t)\psi_{p,jk}(x), \ \ l \geq 1$$

and there exists B such that

$$\|g_l(\cdot,t)\|_\infty \leq Bw(2^{-n_l}),$$

for any $t > 0$. Then $u \in h_v^\infty$ and $\|u\|_{v,\infty} \leq CB$, where C depends on A and ϕ only.

Growth classes on Lipschitz domains

In this Section we present an oscillation estimate for functions from growth classes in Lipschitz domains in $\mathbb{R}^d$. We describe results from [103].

Notation and statements

Statements

Let $w : \mathbb{R}_+ \to \mathbb{R}_+$ be a doubling weight as defined in (I.37). Given a Lipschitz function $\phi : \mathbb{R}^d \to \mathbb{R}$, we denote by Ω_ϕ the domain above the graph of ϕ,

$$\Omega_\phi = \{(x,y) : x \in \mathbb{R}^d, y > \phi(x)\}.$$

We consider harmonic functions in Ω_ϕ with the usual growth restriction

$$|u(x,y)| \le Cw(\mathrm{dist}((x,y),\partial\Omega_\phi)), \quad (x,y) \in \Omega_\phi.$$

The space of these functions is denoted by $h_w^\infty(\Omega_\phi)$ and the smallest C for which this inequality is satisfied is called the norm of u in $h_w^\infty(\Omega_\phi)$. We denote it by $\|u\|_{w,\infty}$.

Our main goal is to obtain an estimate in the spirit of the Law of the Iterated Logarithm for ***weighted averages*** of functions from $h_w^\infty(\Omega_\phi)$ ([103, Theorem 2]).

Theorem I.17 *Let ϕ be a Lipschitz function on $\mathbb{R}^d$ and let u be a function in $h_w^\infty(\Omega_\phi)$. For $x \in \mathbb{R}^d$ and $0 < \delta \le 1$ put*

$$I(x,\delta) = \int_\delta^1 u(x,\phi(x)+y)\, d\left(\frac{1}{w(y)}\right). \tag{I.43}$$

Then the following LIL holds

$$\limsup_{\delta\to 0} \frac{I(x,\delta)}{\sqrt{\log w(\delta) \log\log\log w(\delta)}} \le C\|u\|_{w,\infty}, \quad a.e.\, x \in \mathbb{R}^d, \tag{I.44}$$

where C depends only on the function ϕ, weight w and dimension d.

By a trivial estimate, if $u \in h_w^\infty(\mathbb{R}_+^{d+1})$, then immediately $I(x,\delta) \le C\log w(\delta)$. However, weighted average (I.43) enjoys some nice cancellation properties, so, with the help of the Law of Iterated Logarithm (LIL) techniques, a better estimate (I.44) is obtained. Since we are now working with Lipschitz domains, we can not use wavelet decomposition techniques any more. Now we employ ideas by J. Llorente and A. Nicolau, which, in short, work as follows. First we approximate I by a Bloch function H that also happens to belong to $h_{\log w}^\infty$. Recall that H is a Bloch function in Ω if it is harmonic there and $|\nabla H|(\xi) \le \frac{C}{\mathrm{dist}(\xi,\partial\Omega)}$. Here we would like to mention two of the LILs for harmonic functions, namely the Makarov-Llorente LIL for the Bloch functions, [66, Corollary **3.2**] [58, Theorem **1**], and the LIL of Bañuelos-Moore, [8, Theorem **3.04**]. Unfortunately, we could not use either of those directly, since the former does not provide the desired estimate for slow growing weights w, and the latter involves the Lusin area integral, which can not readily be estimated by the weight. Therefore we modify the ideas used in the proof of those LILs and we proceed by approximating H by a (super)dyadic martingale and estimating its quadratic function by w. Then it remains to apply the LIL for the martingales.

We would also like to note that Theorem I.17 remains true if we replace Ω_ϕ with some star-like Lipschitz domain.

As a corollary of Theorem I.17 we have the local version of Theorem **5** from [99] (which is [103, Theorem 3])

Theorem I.18 *Let u be a harmonic function in $\mathbb{R}_+^{d+1}$. Assume that there exists a set $\Sigma \subset \mathbb{R}^d$ of positive d-dimensional Lebesgue measure such that for every $x_0 \in \Sigma$*

$$|u(x,y)| \le Cw(y), \quad |x - x_0| \le My, \tag{I.45}$$

where M is some positive constant. Then

$$\limsup_{\delta\to 0}\frac{I(x,\delta)}{\sqrt{\log w(\delta)\log\log\log w(\delta)}}\leq C_1\cdot C,\quad a.e.\, x\in\Sigma, \tag{I.46}$$

where the constant C_1 depends only on M, w and d.

Note that the condition (I.45) restricts the *non-tangential* growth of u near the boundary. We do not know if this result remains true if we replace (I.45) by a *radial* growth condition.

Notation

By $|E|$ we denote the Lebesgue measure of a set $E\subset\mathbb{R}^d$ with dimension d depending on the context.
Given a point $x=(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$ and $r>0$ we denote by $Q(x,r)$ the cube of radius r centered at x

$$Q(x,r)=\prod_{i=1}^{d}(x_i-r,x_i+r],$$

we also put $Q(x,\frac{1}{2}):=Q(x)$. Further, given a cube Q we denote its center by x_Q, so that $Q=Q(x_Q,r)$ for some positive r.

Fix $x\in\mathbb{R}^d$. If $2^k x_i-\frac{1}{2}\in\mathbb{Z}$ for every $i=1\dots d$, and $r=2^{-k-1}$ for some $k\in\mathbb{Z}_+$, we call the cube $Q=Q(x,r)$ dyadic of rank k. For $x\in\mathbb{R}^n$ and $k\in\mathbb{Z}_+$ we denote by $\Delta_k(x)$ the collection of all (shifted) dyadic cubes of rank k in $Q(x)$,

$$\Delta_k(x)=\left\{\prod_{i=1}^{d}(x-\frac{1}{2}+m_i2^{-k},x-\frac{1}{2}+(m_i+1)2^{-k}),\quad m_i\in\mathbb{Z}_+,\ 0\leq m_i\leq 2^k-1\right\},$$

we also put $\Delta(x)=\bigcup_{k=0}^{\infty}\Delta_k(x)$. If $Q(x)=(0,1]^d$, we write Δ_k and Δ respectively. By $\mathcal{F}_k(x)$ we denote the (finite, obviously) sigma-algebra generated by dyadic cubes of rank k in $Q(x)$. Given a probability Borel measure μ on $Q(x)$ and an increasing sequence $\{n_k\}_{k=0}^{\infty}\subset\mathbb{Z}_+$ we can consider the ***(super-)dyadic martingales*** on $Q(x)$ with respect to the filtration $\{\mathcal{F}_{n_k}(x)\}_{k=0}^{\infty}$, they are usually denoted by $S=\{S_k,\mathcal{F}_{n_k}(x),\mu\}$. This means that S_k is a piecewise constant function on the (shifted) dyadic cubes of rank n_k, and if $\tilde{Q}$ is a dyadic cube in $\Delta_{n_{k-1}}(x)$, then

$$\frac{1}{\mu(\tilde{Q})}\int_{\tilde{Q}}S_k(t)\,d\mu(t)=S_{k-1}(x_{\tilde{Q}}).$$

In particular, if $n=1$ and μ is the Lebesgue measure on $Q\left(\frac{1}{2}\right)=(0,1]$ then S has a following *truncated wavelet representation*

$$S_k(t)=L+\sum_{j=0}^{n_k}\sum_{i=0}^{2^j-1}b_{ij}\psi_{ij}(t),\quad t\in(0,1], \tag{I.47}$$

where $L=\mathbb{E}S_k=\int_{Q_0}S_k(t)\,dt$, $b_{ij}=2^j\int_{Q_0}S_k(t)\psi_{ij}(t)\,dt$, $\psi_{ij}(t)=\psi(2^jt-i)$, $t\in\mathbb{R}$, and ψ is the

Haar wavelet, $\psi = \chi_{[0,1]} - 2\chi_{[0,\frac{1}{2}]}$ (instead of the usual L^2 scaling we use L^∞ one here, it is more convenient for our purposes). For any interval $I \subset (0,1]$ of length $2r$, $I = [x_I - r, x_I + r]$, we put

$$\psi_I(t) = \psi\left(\frac{t - x_I + r}{2r}\right), \quad t \in \mathbb{R}.$$

Then (I.47) can be written as follows

$$S_k(t) = L + \sum_{j=0}^{n_k} \sum_{I \in \Delta_j} b_I \psi_I(t), \quad t \in (0,1],$$

where $b_I = \frac{1}{|I|} \int_{\mathbb{R}} S_k(t) \psi_I(t)\, dt$.
By $\langle S \rangle_k$ we denote the ***quadratic function*** of S,

$$\langle S \rangle_k^2 = \sum_{j=1}^{k} \mathbb{E}[|S_j - S_{j-1}|^2 | \mathcal{F}_{n_{j-1}}].$$

If $n_k = k$ and we use the Haar representation of S, we can write the quadratic function in the following way

$$\langle S \rangle_k^2(t) = L^2 + \sum_{I \in \Delta: t \in I, |I| \geq 2^{-k}} b_I^2. \tag{I.48}$$

Let u be a harmonic function in Ω_ϕ. We say that u belongs to the ***Bloch class*** in Ω_ϕ, if there exists a constant $D > 0$ such that

$$|\nabla u(x,y)| \leq \frac{D}{\operatorname{dist}((x,y), \partial\Omega_\phi)}, \quad (x,y) \in \Omega_\phi.$$

We denote the space of such functions by $\mathcal{B}(\Omega_\phi)$ and the smallest D for which this inequality is satisfied by $\|u\|_{\mathcal{B}}$.
The connection between Bloch functions and dyadic martingales is well established, see, for example, [66] for the unit disc case and [58] for Lipschitz domains. Here, however, we use a superdyadic martingale, which is, essentially, a thinned dyadic martingale. It means that instead of the usual dyadic filtration $\mathcal{F}_k$ we use some subsequence of dyadic sigma-algebras $\mathcal{F}_{n_k}$ where n_k depends on the weight w (and is lacunary for slow growing w). The main reason for the transition from the dyadic to the superdyadic martingale approximation here is that the quadratic function of the superdyadic martingale is much easier to estimate (similar ideas were used in [65]).

Scheme of the proof

The proof is given in Section 5.1.2 and follows immediately once we have the following Lemma, which we prove in 5.1.3 and 5.1.4.

Lemma I.2 *Assume that $u \in h_w^\infty(\Omega_\phi)$. Then for every $x_0 \in \mathbb{R}^d$ there exists a probability measure μ on $Q(x_0)$ and a (super)dyadic martingale $S = \{S_k, \mathcal{F}_{n_k}, \mu\}_{k=0}^\infty$ on $Q(x_0)$ such that μ is absolutely* *continuous with respect to the Lebesgue measure on* $Q(x_0)$ *and for every* $k \in \mathbb{Z}_+$

$$|S_k(x) - I(x, s_k)| \lesssim 1, \tag{I.49a}$$
$$|S_k(x) - S_{k+1}(x)| \lesssim 1, \quad x \in Q(x_0), \tag{I.49b}$$

where $s_k = w^{-1}(2^k)$.

The local Theorem I.18 is proven in Section 5.1.5.

Lemma I.2 above relies on the approximation of I by **the** Bloch function in $h^\infty_{\log w}(\Omega_\phi)$, which satisfies the respective LIL. Now we wonder if **a** Bloch function in $h^\infty_{\log w}(\Omega_\phi)$ satisfies the LIL (I.44). In Section 5.2 we construct an example of a Bloch function that provides the negative answer to this question.

Growth classes: two counterexamples for divided differences

Here we discuss a couple of constructions that are related to the oscillatory behaviour of divided differences of Hölder functions. These differences can be seen as yet another example of growth behaviour of a harmonic function, only now instead of Poisson kernel, the extension to the upper half-plane is done via Steklov kernel. We restrict ourselves to the real line in this case. The results described here are from [109].

Statements and notation

Statements

For $0 < a < 1$ let $\mathrm{Hol}_a(\mathbb{R})$ be the Hölder class of functions $f : \mathbb{R} \to \mathbb{R}$ such that there exists a constant $C = C(f) > 0$ with $|f(x) - f(y)| \le C|x - y|^a$ for any $x, y \in \mathbb{R}$. The infimum of such constants C is denoted by $\|f\|_a$.
For a continuous real function f we define an ***a-divided difference*** as follows

$$\mathrm{D}_a(f)(x, h) = \frac{f(x+h) - f(x)}{|h|^a}. \tag{I.50}$$

Our goal here is to prove two theorems about oscillations of D_a for functions from Hol_a.

Theorem I.19 *Let* $0 < a < 1$. *Then there exists a function* $f \in \mathrm{Hol}_a(\mathbb{R})$ *such that at almost every* $x \in \mathbb{R}$ *one has*

$$\limsup_{h \to 0^+} \mathrm{D}_a(f)(x, h) > 0$$

and

$$\liminf_{h \to 0^+} \mathrm{D}_a(f)(x, h) = 0.$$

Theorem I.20 *Let* $0 < a < 1$. *Then there exists a function* $f \in \mathrm{Hol}_a(\mathbb{R})$ *and a constant* $C > 0$ *such that for any point* $x \in \mathbb{R}$ *there exist two sequences* $\{h_k\}_{k=1}^\infty$, $\{h'_k\}_{k=1}^\infty$ *of positive numbers,* *converging to zero, such that*

$$\begin{aligned}\limsup_{k\to\infty}\left|\frac{f(x+h'_k)-f(x)}{h'_k}\right| &\le 1,\\ \liminf_{k\to\infty}\frac{|f(x+h_k)-f(x)|}{|h_k|^\alpha} &> C.\end{aligned} \tag{I.51}$$

Notation and reasoning behind Theorems I.19, I.20

One might wonder why are we interested in such properties of D_a and the reason is as follows. For $b > 1$, G.H. Hardy proved in [39] that the Weierstrass function

$$f_{b,a}(x) = \sum_{n=0}^{\infty} b^{-na}\cos(b^n x), \quad x \in \mathbb{R}, \tag{I.52}$$

is in $\mathrm{Hol}_a(\mathbb{R})$ and exhibits the extreme behavior

$$\limsup_{h\to 0}\frac{|f_{b,a}(x+h)-f_{b,a}(x)|}{|h|^a} > 0$$

at any point $x \in \mathbb{R}$.
In [109] the following improvement that elaborates on this extreme behaviour was obtained.

Theorem *Let* $0 < a < 1$ *and* $f \in \mathrm{Hol}_a(\mathbb{R})$. *At almost every point* $x \in \mathbb{R}$ *such that there exists a constant* $\delta = \delta(x) > 0$ *with*

$$\limsup_{h\to 0^+}\frac{\sigma\{t\in[h,1] : \mathrm{D}_a(f)(x,t) > \delta\}}{\log h^{-1}} > 0, \tag{I.53}$$

there exists a constant $c = c(x) > 0$ *such that*

$$\limsup_{h\to 0^+}\frac{\sigma\{t\in[h,1] : \mathrm{D}_a(f)(x,t) < -c\}}{\log h^{-1}} > 0. \tag{I.54}$$

Here σ is the standard Haar measure of $(0,\infty)$ defined as

$$\sigma(E) = \int_E \frac{dh}{h}, \quad E \subset (0,\infty),$$

so that $\sigma[h,1] = \log h^{-1}$, $0 < h < 1$.
Remark. In particular, for any $b > 1$ and $0 < a < 1$, the Weierstrass function $f_{b,a}$ defined above in (I.52) satisfies condition (I.53) (and also (I.54)) at any point $x \in \mathbb{R}$ for certain uniform constants $\delta = \delta(b,a)$ and $c = c(b,a)$.

Observe however that in both the assumption and conclusion of the above theorem, one has to use σ (or some kind of averaging over scales) to measure the set of scales where $\mathrm{D}_a(f)(x,t)$ is not small. **This is necessitated exactly by Theorem I.19.**

The arguments in [109] relied heavily on dyadic martingale techniques. Essentially, the divided differences D_a were represented by the martingale

$$S_k(x) = 2^k(f(x_2) - f(x_1)), \quad x \in [x_1, x_2) \in \Delta_k, \tag{I.55}$$

where Δ_k is the collection of dyadic subintervals of $[0,1]$ of rank k. In particular, this martingale satisfies the *growth condition*

$$\sup_k 2^{-k(1-a)} \|S_k\|_\infty \leq \|f\|_a.$$

Then the continuous results were obtained through this discretization, and the martingale results are collected in the following Theorem ([109, Corollary 1]).

Theorem *Let $0 < \varepsilon < 1$ and let $\{T_n\}$ be a dyadic martingale with* $\sup_n 2^{-n\varepsilon}\|T_n - T_{n-1}\|_\infty \leq 1$.

(a) For $0 < \theta < 1$, consider the set $G(\theta)$ of points $x \in \mathbb{R}$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \#\{1 \leq k \leq N : 2^{-k\varepsilon} T_k(x) \geq \theta\} = 1. \tag{I.56}$$

Then $H^{\Phi(\theta(1-2^{-\varepsilon}))}(G(\theta)) \leq 1$ and consequently $\dim G(\theta) \leq \Phi(\theta(1 - 2^{-\varepsilon}))$.

(b) For $0 < \theta < 1$ consider the set $F(\theta)$ of points $x \in \mathbb{R}$ such that

$$\liminf_{k\to\infty} 2^{-k\varepsilon} T_k(x) \geq \theta.$$

Then $H^{\Phi(\theta(1-2^{-\beta}))}(F(\theta)) \leq 1$ and consequently $\dim F(\theta) \leq \Phi(\theta(1 - 2^{-\varepsilon}))$.

(c) At almost every point $x \in \mathbb{R}$ such that there exists a constant $\delta = \delta(x) > 0$ with

$$\limsup_{N\to\infty} \frac{1}{N} \#\{1 \leq k \leq N : 2^{-k\varepsilon} T_k(x) > \delta\} > 0,$$

there exists a constant $c = c(x) > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \#\{1 \leq k \leq N : 2^{-k\varepsilon} T_k(x) < -c\} > 0$$

Here H^s is the dyadic Hausdorff measure, dim is the Hausdorff dimension, and Φ is the ***entropy function***

$$\Phi(\eta) = \frac{1+\eta}{2} \log_2 \left(\frac{2}{1+\eta}\right) + \frac{1-\eta}{2} \log_2 \left(\frac{2}{1-\eta}\right), \quad 0 < \eta < 1.$$

Remark. We observe that the strategy of obtaining continuous results from their dyadic analogues has certain limitations. Fix $0 < \varepsilon < 1$ and let $\{T_n\}$ be a dyadic martingale such that $\sup_n 2^{-n\varepsilon}\|T_n\|_\infty < \infty$. J. Fernandez, J. Heinonen and J. Llorente, [35], proved the following $0-1$ Law: for any interval I either $\{T_n(x)\}$ converges at a set of points $x \in I$ of positive length or there exists a constant $C > 0$ such that

$$\mathrm{H}^{\infty}_{1-\varepsilon}(\{x \in I : \lim_{n\to\infty} T_n(x) = +\infty\}) > C|I|^{1-\varepsilon}.$$

Here $\mathrm{H}^{\infty}_{1-\varepsilon}$ denotes the $(1-\varepsilon)$-Hausdorff content. See [35]. However the continuous analogue of this result fails, moreover a Hölder continuous function may oscillate wildly around every point. **This is exactly the statement of Theorem I.20**.

Scheme of the proof

Section 6.1 contains the proof of Theorem I.19. Section 6.2 is devoted to the proof of Theorem I.20.

Cartwright's theorem for growth spaces in balls

For our next result we travel to the unit ball $\mathbb{B}$ of $\mathbb{R}^{d+1}$. There we show, following [100], that u is harmonic function subject to a *one-sided regular* growth condition (and now it does not need to be doubling), then it has the same estimate (possibly multiplied by a constant) from the other side as well.

Statements

A well-known theorem by M.L. Cartwright [22] states that if a function u, harmonic in the unit disk, $u(0) = 0$, satisfies the ***one-sided growth condition***

$$u(z) \le w(1 - |z|), \quad z \in \mathbb{D},$$

where $w(t) = \frac{1}{t^p}$ for some $p > 1$, then the reverse inequality holds

$$u(z) \ge -Cw(1 - |z|), \quad z \in \mathbb{D},$$

where C depends only on p. This result was later refined and extended to more general weights by Cartwright herself ([23]) and C.N. Linden ([56, 57]). The works by N. Nikolskii ([76]) and A. Borichev ([13]) should also be mentioned, the latter in particular, where a very nice estimate $u(z) \ge -(1 + o(1))w(1 - |z|)$ was obtained for sufficiently fast growing weights (see [13, section 1.3]); some estimates for the constant in the reverse inequality were also given earlier in [57]. The techniques used in all of the works mentioned above involve analytic functions and conformal mappings and are therefore limited to the complex plane. However, it is natural to ask if similar results hold for harmonic functions in higher dimensions, related problems in higher dimensions were studied by P.J. Rippon, K. Samotij and B. Korenblum, see [82, 84, 85, 52].

Let $w : \mathbb{R}_+ \to \mathbb{R}_+$ be a strictly decreasing function, such that $\lim_{y\to 0} w(y) = \infty$ and $w(1) = 1$. Furthermore we assume that $w \in C^2$ and the following *growth and regularity conditions* are satisfied

$$\lim_{y\to 0} \frac{w(y)}{w'(y)} = 0, \tag{I.57}$$

and

$$\left(\frac{w(y)}{w'(y)}\right)' \geq -\frac{1-\delta}{d}, \quad 0 < y < 1, \tag{I.58}$$

for some positive δ.

Remark. Observe that now we do not ask the weight w to be doubling anymore, rather we want it to be regular, as per estimates above.

Our main result is the following

Theorem I.21 *Let U be a harmonic function in the unit ball $\mathbb{B} = \{z \in \mathbb{R}^{d+1} : \ |z| < 1\}$ and $U(0) = 0$. Assume that U admits the growth condition*

$$U(z) \leq w(1-|z|), \quad z \in \mathbb{B}, \tag{I.59}$$

where the weight w satisfies (I.57) *and* (I.58) *above.*

Then the following two-sided estimate holds ([100, Theorem 1])

$$|U(z)| \leq Cw(1-|z|), \quad z \in \mathbb{B}, \tag{I.60}$$

where the constant $C = C(d, \delta)$ depends only on the parameter δ and the dimension d.

The conditions (I.57) and (I.58) assure that the weight w grows relatively fast as $|z| \to 1$ and is regular. The natural regularity for majorants of harmonic functions is *logarithmic convexity*, however it is shown in [13, Proposition 4.1] that some additional regularity of the weight is necessary for Theorem I.21 to hold.

For the rate of the growth, the weight $w_0(y) = y^{-d}$ is the natural threshold in this result. We see that $w(y) = y^{-p}$ satisfies (I.57) and (I.58) if and only if $p > d$. The result also fails when $p \leq d$, since one can easily see that the Poisson kernel for the ball $\mathbb{B}$ is strictly positive, but grows like w_0 near its singularity at the boundary. There is no upper bound on the growth of w.

Scheme of the proof

Section 7.1

Here we collect some notation and technical results.

Section 7.2

Here we restate our main Theorem a bit, formulating it in terms of averages over spherical caps.

Section 7.3

This Section considers with rewriting regularity conditions (I.57) and (I.58) into more convenient (with relation to our context) form.

Section 7.4

Here we discuss said regularity conditions and explain their inner workings.

Section 7.5

We get back to the proof, and we formulate the key technical result, from which we deduce Theorem I.21.

Theorem I.22 *Let $\tilde{k} : \mathbb{R}_+ \mapsto \mathbb{R}_+$ be a strictly decreasing absolutely continuous function such that*

$$\tilde{k}(0) < \infty, \tag{I.61a}$$

$$\int_0^1 \left(\frac{\tilde{k}(y)}{y}\right)^{\frac{1}{d+1}} dy \leq D, \tag{I.61b}$$

for some constant $1 < D < \infty$. Let $\tilde{u}$ be a harmonic function in $\mathbb{B}$, continuous up to the boundary, satisfying $\tilde{u}(0) = 0$ and $\tilde{u}(z) \leq \tilde{k}(|z|)$ for $z \in \mathbb{B}$ Then for any $x_0 \in \partial\mathbb{B}$ and $b \in \left[0, \frac{1}{2}\right]$ the following inequality holds

$$\int_{\{\phi(\zeta,x_0)\leq b\}} \tilde{u}(\zeta)\, d\sigma_d(\zeta) \geq -C\left(D^{d+1} + \tilde{k}(0)b^d\right). \tag{I.62}$$

where C depends only on the dimension n.

Here $\phi(\zeta, x_0)$ is the angle between two points $\zeta, x_0 \in \partial\mathbb{B}$ and σ_d is the normalized surface measure on a sphere.

Section 7.6

We have another technical lemma formulated and proven in this Section. The argument involves a construction of an auxiliary surface, which is a slight modification of one used in [52].

Section 7.7

We finish proving Theorem I.22 in this Section.

Mean variation

The last part of the thesis is devoted to a different kind of estimate of the boundary behaviour – the ***normal variation***. Under its maiden name, ***the radial variation***, it appeared first as a separate object probably in the very classical Rudin's paper [83] where he studied various spaces of analytic functions (disc algebra, Blaschke products, space $H^\infty(\mathbb{D})$ bounded analytic functions) in the unit disc. He proved that for any such space (and, rather trivially, for some others, like $H^2(\mathbb{D})$) there exists a function f such that its radial variation is infinite at *almost every* point on

$\mathbb{T}$. In other words,

$$\operatorname{var}(f)|_{[0,e^{i\theta}]} = \int_0^1 |f'(re^{i\theta})|\, dr = +\infty, \quad \text{for a.e. } \theta \in [0, 2\pi). \tag{I.63}$$

He also gave some examples of spaces where this quantity is finite everywhere or almost everywhere. Moreover, he actually showed that the set of θ such that the corresponding radial variation is finite is not only of Lebesgue measure zero, but also of first category. The only remaining thing to show was to get rid of such a set altogether. This, however, turned out to be not so easy, and in the following years became a rather interesting and known problem, at some point he actually offered a symbolic prize for solving it.

After quite a while, 38 years later to be precise, this conjecture was disproved by Bourgain in [15], where he showed that *every bounded analytic function* in the unit disc $\mathbb{D}$ has finite radial variation on a certain non-empty set of radii, moreover, this set, while it can be very small in some sense, it has to be very large in some other sense – it must have full Hausdorff dimension. In a followup paper, [16], he extended this result to *positive harmonic functions*, and, answering a question raised by V.P. Havin, also to extensions by some other types of sufficiently smooth approximate identities (like Fejer kernel, though not Steklov kernel).

Different forms of weighted integrals of derivatives of analytic functions are ubiquitous, they are usually used in definitions (Dirichlet or weighted Hardy-Sobolev spaces immediately come to mind) or characterizations of various spaces of functions, like in works of D. Girela on Besov spaces, or a very recent series of papers by A. Baranov and I. Kayumov on rational functions.
The estimates of radial variation itself were used, for instance, by P. Jones and P.F.X. Müller in the proof of Anderson's conjecture ([47], see also a recent paper [71] for multi-dimensional version).

The main result in [16] was later extended to the unit ball in $\mathbb{R}^d$ by M. O'Neill in [77]. A certain reformulation with different approximate identities was also considered in [16, 98]. In particular, for a lot of radii a stronger estimate holds

$$\int_0^1 \int_0^{2\pi} |f'(re^{i\tilde{\theta}})| P_{(r)}(\theta - \tilde{\theta})\, d\tilde{\theta} < +\infty, \tag{I.64}$$

where $P_{(r)}(\theta) = \dfrac{1 - r^2}{2\pi(1 + r^2 - 2r\cos\theta)}$ is the usual Poisson kernel for the disc. It is easy to see that the left-hand side majorizes the radial variation. On the other hand the additional convolution by a Poisson kernel makes this quantity more adjustable and stable (in a sense it is not unlike the relation between non-tangential and radial maximal functions – the latter is usually trickier to work with).

Our main goal is to extend Bourgain's result to positive functions on (smooth) domains in $\mathbb{R}^{d+1}$, $d \geq 1$. The result described here were obtained (in a slightly more general formulation) in [101].

Formulation of main results

Let $\Omega \subset \mathbb{R}^{d+1}$ be a domain and S be its boundary that we consider to be C^2-smooth. By $\vec{N}(\xi)$ we denote the inward normal vector to S at $\xi \in S$. The interval $\{ta + (1-t)b : 0 \le t < 1\}$, where $a, b \in \mathbb{R}^{d+1}$, is denoted by $(a, b]$. Given a point $\xi \in S$ let $t(\xi) > 0$ be a number such that $(\xi, +r(\xi)\vec{N}(\xi)] \subset \Omega$. Let u be some real-valued function on O. The ***normal variation of*** u ***at*** $\xi \in S$ is

$$(\mathrm{Nvar}\ u)(\xi) := \mathrm{var}\left(u|_{(\xi, p+t(\xi)\vec{N}(\xi)]}\right).$$

We are only concerned whether this quantity is finite, and u is assumed to be smooth on Ω. Therefore the explicit choice of $t(\xi)$ is of no importance.

Let $E \subset S_1 \subset S$. We say that the set E is ***ultradense*** in S_1, if for any $\xi \in S_1$ and $r > 0$ we have

$$\dim(E \bigcap \mathbb{B}(\xi, r)) = d,$$

here $\mathbb{B}(\xi, \rho)$ is the ball in $\mathbb{R}^{d+1}$ with radius r and center ξ, and dim is the Hausdorff dimension. Put

$$\mathcal{V}(u) := \{\xi \in S : (\mathrm{Nvar}\ u)(\xi) < +\infty\}.$$

Our main result here is the following Theorem ([101, Theorem 1]).

Theorem I.23 *If u is harmonic and positive on Ω, then $\mathcal{V}(u)$ is ultradense in S.*

It is a well-known fact that ([81], [43]) that any function u that is positive and harmonic on Ω has finite boundary values along *almost all* (with respect to d-dimensional Hausdorff measure) normal vectors $\vec{N}(\xi)$. Theorem I.23 states that for many ξ a stronger version of this property holds: the variation $(\mathrm{Nvar}\ u)(\xi)$ is finite.

Our theorem extends previously mentioned results of Bourgain and O'Neill. For $d = 1$ it can be easily deduced from [15] via conformal mappings. However, when studying the boundary behaviour of harmonic function on more or less arbitrary $(d+1)$-dimensional domains, the transition from $d = 1$ to $d \ge 2$ is usually quite complicated (see, for example, [20, pp. 48-49], [43], [44]). On top of that, the problem we are interested in is further complicated by the fact that the remarkable results in [15], [16] relied heavily upon the use of *Fourier transform*, which is well suited for the case when S is a group of some sort, but that can hardly be adapted to deal with the domains in Theorem I.23 (the paper [77] used spherical harmonics). While we still follow the general idea of [15] it was necessary to modify the arguments and to avoid the methods of harmonic analysis.

Let us look again at Theorem I.23. Under its hypothesis we see that

$$\mathrm{var}(u|(\xi, \xi + t(\xi)\vec{N}(\xi)]) = \int_0^{t(\xi)} \left| \frac{\partial}{\partial t}\left(u\left(\xi + t\vec{N}(\xi)\right)\right)\right| dt \tag{I.65}$$

and Theorem I.23 follows from Theorem I.24 ([101, Theorem 2]), where we replace the integral in (I.65) by an integral that considers the derivatives of u *along all directions* (and not just along the normal vector $\vec{\mathcal{N}}(\xi))$). Put

$$\mathcal{V}_{\text{grad}}(u) := \{\xi \in S : \int_0^{t(\xi)} \left|\nabla u(\xi + t\vec{N}(\xi))\right| \, dt < +\infty\}. \tag{I.66}$$

Theorem I.24 *Under hypothesis of Theorem I.23 the set $\mathcal{V}_{\text{grad}}(u)$ is ultradense in S.*

As we have already mentioned, the key obstacle to overcome in Theorems I.23 and I.24 is that S is not a group any more. Nevertheless we will demonstrate our argument on a very simple case of $\Omega = \mathbb{R}^2_+$ and u being a positive harmonic function on Ω with compactly supported boundary values. We will see that one can extend it to general positive functions and proper domains in $\mathbb{R}^{d+1}$ (it is done in [101] anyway), and we promise not to use the group structure of $\mathbb{R}$, convolution arguments and such – unless it is convenient for purposes of notation.
From now on we are acting under such assumptions.

Given a positive harmonic function u on $\mathbb{R}^2_+$ (which, as we have agreed, is a harmonic extension of some positive measure with compact support) and a number $y > 0$ let

$$u_y(x,t) := u(x, t+y), \quad x \in \mathbb{R}, \ t > 0.$$

Now let h^y_u be a ***least harmonic majorant*** of a subharmonic function ∇u_y (see ([89, Ch. 6, §4]) for details). Writing it in the convolution form, we get

$$h^y_u(x,s) := \int_{\mathbb{R}} |\nabla u(\xi, y)| P_{(s)}(x-\xi) \, d\xi, \quad x \in \mathbb{R}, \ y, s > 0,$$

where $P_{(s)}(\xi) = \dfrac{s}{\pi(\xi^2 + s^2)}$ is the Poisson kernel for the upper half-plane.
We define

$$\operatorname{Mvar} u(x) := \int_0^1 h^{2t}_u(x,t) \, dt.$$

This is the analogue of the averaged gradient from (I.64). A point $x \in \mathbb{R} = S$ is called ***a Bourgain point (B-point) of*** u, if $\operatorname{Mvar} u(x) < +\infty$. The set of such points we denote by $\mathcal{B}(u)$. It is easy to show that $\mathcal{V}_{\text{grad}}(u) \supset \mathcal{B}(u)$. We present the following theorem.

Theorem I.25 *The set $\mathcal{B}(u)$ is ultradense in S.*

Remark. It is of paramount importance that u is *bounded from below (or from above)*. We can not drop this condition, since there exists a harmonic on $\mathbb{D}$ function u with boundary values in BMO such that $\mathcal{V}(u)$ (on the unit circle) is empty (see [46]).

Discussion of the proof of Theorem I.25

In Chapter 8 we will actually be proving only Theorem I.25. It allows us to significantly simplify the exposition, and the argument is given in full generality in [101].

Differential equation for h^{2t}

Given a positive harmonic function φ on $\mathbb{R}^2_+$ such that $\varphi(\infty) = 0$ we write

$$\varphi_{[t]}(x) := \varphi(x,t), \quad x \in \mathbb{R},\ t > 0.$$

In other words we restrict our harmonic function ϕ on the line $\{(x,t) :\ x \in \mathbb{R}\}$.
Now imagine for a moment that we are gifted with a *regular family of integral operators* B_t with kernels $b_t : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ that act on $C_0^\infty(\mathbb{R})$, and on a specific choice of the argument – $u_{[t]}$ – it gives

$$B_t\left[u_{[t]}\right] := \left(h^{2t}_u\right)_{[t]}, \quad t > 0. \tag{I.67}$$

Assume also, while we are at it, that for any $\varepsilon > 0$ we have another family $\Psi_t = \Psi_{t,\varepsilon}$ (we usually drop the second subscript) with *positive kernels* ψ_t that satisfies the following properties

$$\Psi_\theta\left[\varphi_{[t]}\right] \le C\Psi_t\left[\varphi_{[t]}\right], \quad \forall\varphi \text{ and } 0 < \theta \le t \le 1, \tag{I.68a}$$
$$\tfrac{\partial}{\partial t}\left(\Psi_t\left[\varphi_{[t]}\right]\right) = \varepsilon\Psi_t\left[B_t\left[\varphi_{[t]}\right]\right], \quad 0 < t < 1. \tag{I.68b}$$

Now let $g_t := (h^{2t}_u)_{[t]}$ and observe that due to positivity of u, Ψ_t, B_t for any $0 < \theta \le \delta \le 1$ we have

$$\begin{aligned}\Psi_1\left[u_{[1]}\right] &\ge \Psi_1\left[u_{[1]}\right] - \Psi_\delta\left[u_{[\delta]}\right] = \varepsilon\int_\delta^1 \frac{\partial}{\partial t}\left(\Psi_t\left[u_{[t]}\right]\right)dt = \varepsilon\int_\delta^1 \Psi_t\left[B_t\left[u_{[t]}\right]\right]dt = \\ \varepsilon&\int_\delta^1 \Psi_t[g_t]\,dt \ge \frac{\varepsilon}{C}\int_\delta^1 \Psi_\theta[g_t]\,dt = \frac{\varepsilon}{C}\Psi_\theta\left[\int_\delta^1 g_t\,dt\right].\end{aligned} \tag{I.69}$$

Broadly speaking this computation already suggests that $\operatorname{Mvar} u = \int_0^1 g_t\,dt$ is finite at a lot of points. We need to go deeper, though, so we ask even more from Ψ_t. Namely, we want the kernel ψ_t to have unit mass,

$$\int_{\mathbb{R}} \psi_t(x,\xi)\,d\xi \equiv 1,$$

or, in other words, that Ψ_t^* maps the set of probability measures on $\mathbb{R}$ to itself. Moreover, let us pick any smooth probability measure ν on $\mathbb{R}$ – so a Gaussian measure would just suffice. Next we set $\nu_t := \Psi_t^*[\nu]$, $0 < t \le 1$ (remember that ν_t do depend on ε) and claim that

a. *there exist some positive constants c_1, c_2 such that for any $r > 0$*

$$\nu_t(I) \le c_1 r^{1-c_2\varepsilon} \tag{I.70}$$

for any interval $I \subset \mathbb{R}$ with $|I| = r$;

b. *For an interval I the measure ν_t has positive mass on I, if $\varepsilon < \varepsilon_I$ is small enough.*

Having secured all of the above it remains to integrate (I.69) with respect to ν to obtain

$$\int_{\mathbb{R}} \Psi_1\left[u_{[1]}\right]d\nu \ge \int_{\mathbb{R}} \frac{\varepsilon}{C}\Psi_\theta\left[\int_\delta^1 g_t\,dt\right]d\nu = \frac{\varepsilon}{C}\int_{\mathbb{R}}\left(\int_\delta^1 g_t\,dt\right)\Psi_\theta^*[\nu].$$

Then pass to the limit in δ to arrive at

$$\int_{\mathbb{R}} \mathrm{Mvar}_u \, d\nu_0 \le \frac{C}{\varepsilon},$$

since $u_{[1]}$ is bounded and ν is some reasonable probability measure. Here ν_0 is the weak limit of ν_θ. The dimension estimate of $\operatorname{supp} \nu_0$ follows immediately from a and b above.

How to construct such operators

It might seem that B_t and Ψ_t appear out of thin air, it is not the case though. We write down B_t explicitly, and then solve the differential equation I.68b, accruing the rest of properties along the way. Unfortunately the operators B_t do not really commute, so we can not just write down the solution as an exponent (like in [27]). Hence we have to construct the solution by hand, basically as a Riemann sum, or more precisely a Riemann product – this is a version of the multiplicative integral.

Essentially we have to check three main things: that the solution Ψ_t even exists, that it is a positive operator, and the 'focusing' property (I.68a). All of these are in a nutshell given by an 'extra' Poisson convolution in the definition of $\mathrm{Mvar}\, u$ and generous application of Harnack inequality. We stress again, that this not only a half-space result, and our choice of Ω is dictated by reasons of convenience.

Further comments

Unlike previous topics, here we use 'continuous' analysis, one can even say that we decided to move away from discrete techniques of wavelets and martingales that were implicitly present in [15] (there is also a wavelet description of $\mathcal{B}(u)$). Besides, if we try to do a direct reformulation of the problem in the martingale language, it will turn out to be rather trivial (see [18] for further discussions). Nevertheless, it would be very interesting to have a proper look at the equation (I.68a). So far it seems to be rather elusive to understand properly, and a discrete models usually help in such things.

Such an attempt was made in [111] where the set of Bourgain points was studied for a special choice of u – *harmonic measure of a Cantor-type set* E on the boundary. The description of these points came in the form of dyadic encoding of E, [111, Theorem 1].

Theorem I.26 *Let $\{q_j\} \in \ell^1$ be a sequence of positive numbers, and $E \subset [0,1]$ be a Cantor-type set generated by $\{q_j\}$. In other words, on the step k of construction of E we throw away q_k-th proportion of the segment of previous generation (summability of $\{q_j\}$ guarantees that $|E| > 0$). Every point $x \in E$ can be encoded in a natural way as a sequence of* 0 *and* 1*, denote such a sequence by $\kappa(x) = (\kappa_1(x), \kappa_2(x), \dots)$. Then $x_0 \in E$ is a Bourgain point for the harmonic extension of $\mathbb{1}_E$ to the upper half-plane if and only if*

$$\sum_{k \ge 1} 2^{n_{k+1}(x_0) - n_k(x_0)} q_{n_k(x_0)-1} < +\infty,$$

where $n_k(x_0)$ *are the times where the trajectory* $\kappa(x_0)$ *changes, from* 0 *to* 1 *or vice versa.*

This Theorem sheds some light on the geometric distribution of Bourgain points – they must lie deep in E. For instance, the endpoints of E, i.e. such points where trajectory stabilizes after some number, are definitely not Bourgain, and this can be checked directly. On the other hand, if $n_k(x_0) + 1 = n_{k+1}(x_0)$, then x_0 is a B-point.

Reference list

For the convenience of the reader here we have collected references for the presented results.

Theorems I.3, I.4, I.5, I.6, and Lemma I.1 are presented in [107] and (mostly) in [112]. Paper [112] contains results that comprise Theorems I.7, I.8 and I.9. One-dimensional version of Theorem I.3 is proven in [108], Carleson constant estimate for the bitree is done in [105]. Theorem I.11 is proven in [104], while the discrete-continuous capacity equivalence (Theorem I.12) is shown in [2]. Various examples and counterexamples are collected in papers [106] and [110].

Theorems I.13, I.14, I.15, I.16 are presented in [99] (in particular, Theorem I.14 can be deduced from the estimates in [98]).

Theorems I.17, I.18 are proven in [103], and Theorems I.19, I.20 in [109].

Multidimensional versions of M. Cartwright's theorem (Theorems I.21, I.22) are done in Finally, the results related to the Bourgain points, which are Theorems I.23, I.24, I.25, I.26, are contained in [101] and [111].

Chapter 1 Discrete model

1.1 Some basic facts

In this section we collect a few results about the basic behaviour of capacities on (mostly) a d-tree and the structure of the respective equilibrium measures.

Remark. In what follows the weight w is always assumed to be positive and finite. This is needed to properly run the Potential Theory on graphs, especially the definition and properties of weighted capacity. Sometimes in the counterexamples we allow it to take zero values at some points for our convenience, but we do not invoke any kind of capacity there, they happen on finite graphs, and the weights can be modified to be made positive (but very small) anyway while still maintaining the counterexample. See, for example, Section 2.6.4.

The capacity of a given set is usually very hard to compute explicitly, even in the most classical case $\Gamma = \overline{T}$ and $w \equiv 1$ (which corresponds to the classical harmonic capacity). We give a couple of examples for the most basic situations.

Proposition 1.1.1 *Let w be a weight on a graph Γ. Then the capacity of a singleton is*

$$\operatorname{Cap}_w(\{\alpha\}) = \frac{1}{w(\mathcal{P}(\alpha))}, \quad \alpha \in \Gamma. \tag{1.1}$$

Proof. We want to construct a minimizer for (I.7). Clearly the extremal function $f_{\{\alpha\}}$ should be supported only on $\mathcal{P}(\alpha)$. The elementary solution is immediately given by $f(\gamma) = \frac{1}{w(\mathcal{P}(\alpha))} \cdot \mathbb{1}_{(\mathcal{P}(\alpha))}(\gamma)$. $\square$

The next proposition considers the structure of the equilibrium measure for subsets of a weighted tree.

Proposition 1.1.2 *Assume T is a tree equipped with a weight w, and $E \subset \overline{T}$ is a compact set. Then the capacities of E and $\mathcal{S}(E)$ coincide, moreover, if μ_E is the equilibrium measure of E, then*

$$\mathbf{V}_w^{\mu_E} = 1 \quad \text{q.e. on } \mathcal{S}(E). \tag{1.2}$$

Remark. This equilibrium Frostman property (1.2) definitely does not hold a d-tree with $d \geq 2$, actually most of the times (unless it is a product set, see below) the equilibrium potential is strictly greater than 1 on the set in question. The equality of capacities though holds on any graph.

Proof. The proof is almost immediate. Consider the set $\tilde{E}$ of maximal points of E (in the directed graph structure), i.e. $\alpha \in \tilde{E}$, if $\alpha \in E$ and there is no $\beta \in E$, $\beta > \alpha$. Then $\mathcal{S}(\tilde{E}) = \mathcal{S}(E)$ and, by monotonicity, $\mathbf{I}_w f(\alpha) \geq \mathbf{I}_w f(\beta)$ for $\tilde{E} \ni \alpha \geq \beta$. It turns out that the 'important' points from the point of view of the potential are exactly those in $\tilde{E}$, hence if f is admissible for $\tilde{E}$, it is also admissible for $\mathcal{S}(E)$. On the other hand, $\tilde{E} \subset \mathcal{S}(E)$, therefore the capacity of the second set must be greater or equal than that of the first set.
For (1.2) we may assume that E is a down-set, $E = \mathcal{S}(E)$ by the argument above. Below, in Proposition1.1.4, we show that Maximum Principle holds with constant 1 on a tree. Hence the general Frostman Theorem I.2 implies that $\mathbf{V}_w^{\mu_E} \leq 1$ outside of $\operatorname{supp} \mu_E$. On the other hand by the same Theorem $\mathbf{V}_w^{\mu_E} \geq 1$ q.e. on E.

□

While the next estimate seems to be of a rather specific nature, it will be quite useful later on – d-trees with exponential product weights serve as a natural model for Hardy-Sobolev spaces on the polydiscs

Proposition 1.1.3 *Assume Γ is a d-tree $\overline{T}^d$, and let w be an exponentially increasing product weight*

$$\begin{aligned} w(\alpha) &= w_1(\alpha_1) \cdot \dots \cdot w_d(\alpha_d), \quad \alpha_k \in T \\ w_k(\alpha_k) &:= 2^{(1-s_k)|\alpha_k|}, \quad 0 < s_k \leq 1. \end{aligned} \tag{1.3}$$

Then the distinguished boundary $(\partial T)^d$ has positive capacity. If, however, one of s_k is less or equal than zero, then the capacity of the distinguished boundary is zero.
Also the capacity of a product set is a product of capacities,

$$\operatorname{Cap}_w(E) = \prod_{k=1}^d \operatorname{Cap}_{w_k}(E_k), \quad E = \prod_{k=1}^d E_k,\ \overline{T} \supset E_k \text{ a compact set.} \tag{1.4}$$

Proof. For the first part of the statement we refer to Theorem I.11, which says that the capacity of a set (so, say, a root) and its (distinguished) boundary projection (the total distinguished boundary $(\partial T)^d$ in this case) are comparable. The capacity of a root is obviously non-zero.
For fast growing weights, i.e. when one of s_k's is ≤ 0, we prove the product estimate first.
If f_{E_k} is the extremizer for the k-th coordinate set, then, clearly, $\prod_{k=1}^d f_{E_k}$ is admissible for E, i.e.

$$\sum_{\alpha \geq \omega} w(\alpha) \prod_{k=1}^d f_{E_k}(\alpha_k) = \sum_{k=1}^d \sum_{\alpha_k \geq \omega_k} \prod_{k=1}^d w_k(\alpha_k) f_{E_k}(\alpha_k) \geq 1, \quad \omega \in E.$$

On the other hand, as we have already seen, the one-dimensional equilibrium potential is exactly 1 q.e. not only the support of the equilibrium measure, but q.e. on the whole set E_k itself. Hence, if μ is a product of equilibrium measures for coordinate sets E_k, its d-weighted potential $\mathbf{V}_w^\mu$ is also 1 q.e. on $E = \prod_{k=1}^d E_k$. Hence it can not be but equilibrium for E.
Now assume that a weight w grows too fast on a tree T, i.e. $w(\alpha) \geq 2^{|\alpha|}$ for $\alpha \in T$. If μ is any

non-zero measure on ∂T, then $\mathbf{I}^*(o) = |\mu| > 0$, and

$$\mathbf{I}^*\mu(\alpha) = \mathbf{I}^*\mu(\alpha^1) + \mathbf{I}^*\mu(\alpha^2),$$

where α^i are the two children of α. But then, for two children of the first rank, α_1^1, α_1^2 we have

$$(\mathbf{I}^*\mu(\alpha_1^1))^2 w(\alpha_1^1) + (\mathbf{I}^*\mu(\alpha_1^1))^2 w(\alpha_1^2) \geq 2\frac{1}{2}\left(\mathbf{I}^*\mu(\alpha_1^1) + \mathbf{I}^*\mu(\alpha_2^1)\right)^2 = (\mathbf{I}^*(o))^2.$$

By iteration we see that for any level-set of depth $k+1$ we have

$$\sum_{j=1}^{2^{k+1}} (\mathbf{I}^*\mu(\alpha_{k+1}^j))^2 w(\alpha_{k+1}^j) \geq 2^{k+1} \sum_{j=1}^{2^{k+1}} (\mathbf{I}^*\mu(\alpha_{k+1}^j))^2 \geq$$
$$2^k \sum_{j=1}^{2^k} (\mathbf{I}^*\mu(\alpha_k^j))^2 \geq \cdots \geq |\mu|^2.$$

Hence the energy of μ is infinite,

$$\sum_{\alpha \in T} (\mathbf{I}^*\mu(\alpha))^2 w(\alpha) = \infty.$$

To estimate the capacity of $(\partial T)^d$ we use the product property from above. We are done. $\square$

Remark. The growth rate $2^{|\alpha|}$ is a critical one, and it corresponds to the case where the Potential Theory stops being useful. As we will see later, it is parallel to the Hardy-Sobolev scale of spaces in the disc, where the end-point of the scale, the unweighted Hardy space $H^2(\mathbb{D})$, is exactly when the capacitary conditions do not work any more. The Carleson condition for $H^2(\mathbb{D})$ is realized with the Lebesgue measure on $\mathbb{T}$, which we can still interpret as a Bessel capacity (again, at the end-point of the corresponding scale), however most of the arguments break down here, as they should.

The Potential Theory on the tree is similar enough to the classical theory on the plane (and the classical Bessel potentials can be modeled by the discrete version with an appropriate choice of the weight w). However, the bi-parametric version, even in the case $w \equiv 1$, is quite different in many ways. The next result provides some details.

Proposition 1.1.4 *Let w be a weight on a (dyadic) tree T and μ be a measure on $\overline{T}$ with finite w-energy. Then the supremum of its weighted potential is achieved on its support,*

$$\sup_{\tau \in \operatorname{supp} \mu} \mathbf{V}_w^\mu(\tau) = \sup_{\tau \in \overline{T}} \mathbf{V}_w^\mu(\tau). \tag{1.5}$$

Also, if $f : T \to \mathbb{R}_+$ satisfies $f(\alpha) \geq f(\alpha^-) + f(\alpha^+)$ for any point α and its two immediate children $\alpha^\pm$, and ν is a measure with finite energy such that $\mathbf{I}_w f \geq \mathbf{V}_w^\nu$ on the support of ν, then the same *inequality holds true everywhere*

$$\mathbf{I}_w f(\tau) \geq \mathbf{V}_w^\nu(\tau), \quad \tau \in \overline{T}. \tag{1.6}$$

On the other hand, if $w \equiv 1$ is a weight on T^2, then both the Maximum and Domination principles do not hold in general, even for equilibrium measures. In other words, for any N there exists a set E such that the supremum of the potential of its equilibrium measure μ_E is greater than N, while $\mathbf{V}^{\mu_E} = 1$ on $\operatorname{supp}\mu_E$.

Proof. **Maximum Principle**

It is enough to check (1.5) inside the tree (i.e. for $\tau \in T$), since for any $\beta \in \partial T$ we have $\mathbf{V}_w^\mu(\beta) = \sup_{\tau > \beta} \mathbf{V}_w^\mu(\alpha)$. Now assume that there exists a point $\beta \in T \setminus \operatorname{supp}\mu$ such that

$$\mathbf{V}_w^\mu(\beta) > \mathbf{V}_w^\mu(\alpha), \quad \alpha \in \operatorname{supp}\mu.$$

We see immediately that $\mathcal{S}(\beta) \bigcap \operatorname{supp}\mu = \emptyset$, hence there exists a unique point $\tau_\beta > \beta$ such that $\mathcal{S}(\tau_\beta) \bigcap \operatorname{supp}\mu \neq \emptyset$, but $\mathcal{S}(\tau) \bigcap \operatorname{supp}\mu = \emptyset$ for every $\tau < \tau_\beta$. Then $(\mathbf{I}^*\mu)(\tau) = 0$ for such points τ, and

$$\mathbf{V}_w^\mu(\beta) = \mathbf{V}_w^\mu(\tau_\beta) + \sum_{\tau_\beta > \tau \geq \beta} (\mathbf{I}^*\mu)(\beta) w(\beta) = \mathbf{V}_w^\mu(\tau_\beta).$$

Monotonicity (with respect to natural order on T) of $\mathbf{V}_w^\mu$ implies that

$$\mathbf{V}_w^\mu(\tau_\beta) < \mathbf{V}_w^\mu(\alpha), \quad \text{for any } \alpha \in \mathcal{S}(\tau_\beta) \bigcap \operatorname{supp}\mu,$$

and we have a contradiction.

Domination Principle

As before, it is enough to show (1.6) only for points inside T. Now suppose there exists $\alpha^0 \in T$ such that

$$(\mathbf{I}_w f)(\alpha^0) \leq \mathbf{V}_w^\nu(\alpha^0),$$

clearly we may also assume that

$$(\mathbf{I}_w f)(\tau) > \mathbf{V}_w^\nu(\tau), \quad \tau > \alpha^0.$$

It follows immediately that $f(\alpha^0) \leq (\mathbf{I}^*\nu)(\alpha^0)$, hence one of the sons of α^0 (which we denote by α_1) satisfies $f(\alpha^1) \leq (\mathbf{I}^*\nu)(\alpha^1)$ and $(\mathbf{I}^*\nu)(\alpha^1) > 0$. Repeating this argument we obtain a sequence $\{\alpha^k\}_0^\infty$ of nested points such that $f(\alpha^k) \leq (\mathbf{I}^*\nu)(\alpha^k)$, $k = 0, 1, \ldots$ and $(\mathbf{I}^*\nu)(\alpha^k) > 0$. Denote the endpoint of this geodesic by $\omega = \bigcap_k \mathcal{S}(\alpha^k)$. Clearly $\omega \in \operatorname{supp}\nu$. It follows that

$$(\mathbf{I}_w f)(\omega) = (\mathbf{I}_w f)(\alpha^0) + \sum_{k=1}^\infty f(\alpha^k) w(\alpha^k) \leq \mathbf{V}_w^\nu(\alpha^0) + \sum_{k=1}^\infty (\mathbf{I}^*\nu)(\alpha^k) w(\alpha^k) = \mathbf{V}_w^\nu(\omega),$$

and we have a contradiction.

Why the Maximum and Domination principles fail on T^2

See Section 2.6.4 for Maximum Principle, it fails even for an equilibrium measure on a collection of predecessors of one boundary point. The lack of Domination principle also follows immediately – we can take ν to be the equilibrium measure without Maximum Principle and f to be a point mass at the root. Since $\mathbf{I}wf \equiv 1$ in this case, we have domination on the support of ν, but obviously not on the set of exceptional potential. $\square$

1.2 Strong capacitary inequality in one-parametric case

Below we show a simple version of the so-called Strong Capacitary Inequality (for a 2- or 3-parametric version see Section 2.3). It is a discrete version of the a collection of results that consider potentials with radial kernels. Here the main point is that on a tree one has unique geodesics (so no cycles unlike d-tree) which allows us to make a spherical change of variables of sorts.

Proposition 1.2.1 *Let w be a weight on a tree T, and f be a (non-negative) function there. Then there exists a constant $C = C(w)$ such that*

$$\sum_{k\in\mathbb{Z}} 2^{2k}\operatorname{Cap}_w(\{\tau \in \overline{T}:\ \mathbf{I}_w f(\tau) \ge 2^k\}) \le C\|f\|^2_{L^2(T,dw)}. \tag{1.7}$$

Proof. For every $k \in \mathbb{Z}$ we define a stopping time E_k which consists of maximal (in tree order) points α such that $\mathbf{I}wf(\alpha) \ge 2^k$. In other words, if $\alpha \in E_k$, then $\mathbf{I}_w f(\alpha) \ge 2^k$, but $\mathbf{I}_w f(p(\alpha)) < 2^k$, where $p(\alpha)$ is the unique parent of α. Clearly,

$$\{\tau \in \overline{T}:\ \mathbf{I}_w f(\tau) \ge 2^k\} = \{\tau : \tau \in \mathcal{S}(E_k)\} = \{\tau :\ \exists\alpha \in E_k :\ \alpha \ge \tau\}.$$

If for some geodesic we never have $\mathbf{I}_w(\alpha) \ge 2^k$, i.e. the stopping time is infinite, we just do not have points from E_k on that geodesic. The sets E_k do not contain any proper descendants of their points, i.e. if $\alpha \in E_k$, then there is no $\beta < \alpha$ such that $\beta \in E_k$.
Next, clearly, we can have $F_k := \{\tau \in \overline{T}:\ 2^{k+1} > \mathbf{I}_w f(\tau) \ge 2^k\}$ on the left-hand side of (1.7) instead, we just increase the constant a bit. Our goal is to construct an almost partition Ω_k of T, such that $\bigcup_k \Omega_k = T$, and Ω_k are almost disjoint – their union covers any point at most thrice, and $2^{-k+2} f|_{\Omega_k}$ is admissible for F_k. Such a (almost) partition would provide the desired estimate (1.7), since then we can write

$$\cap_w(F_k) \le 16 \cdot 2^{-2k} \cdot \|f|_{\Omega_k}\|^2_{L^2T,dw},$$

so that

$$\sum_{k\in\mathbb{Z}} 2^{2k}\,\mathrm{Cap}_w(\{\tau\in\overline{T}:\ \mathbf{I}_w f(\tau)\geq 2^k\})\leq 16\sum_k \|f|_{\Omega_k}\|^2_{L^2T,dw}\leq C\cdot\|f\|^2_{L^2T,dw}.$$

The existence of such partition is also almost immediate. Indeed, consider any individual geodesic $\mathcal{P}(\omega)=(o,\tau_1(\omega),\tau_2(\omega),\dots,\omega)$ where ω is some point on ∂T, and let n_k be the stopping time indices on that particular geodesic, i.e. $\tau_{n_k}(\omega)\in E_k$. Due to discrete nature of the setting, it can happen that some of these times coincide, $n_k=n_{k+1}=\dots=n_{k+i}$ for some k,i – this happens when the jump in $\mathbf{I}_w f$ is too large. Nevertheless the sets $\{\tau_j(\omega): n_k\leq j<n_{k+1}\}$ are disjoint, hence are the sets

$$\bigcup_{\omega\in\partial T}\{\tau_j(\omega): n_k\leq j<n_{k+1}\}=F_k.$$

Now, given $\omega\in\partial T$, we put

$$E_k(\omega):=\{\tau_{n_k-1}(\omega),\tau_{n_k}(\omega),\dots,\tau_{n_{k+1}}(\omega)\},\quad \text{if }\ n_k(\omega)<n_{k+1}(\omega),$$

and

$$E_k(\omega):=\emptyset\quad\text{otherwise.}$$

Let

$$E_k:=\bigcup_{\omega\in\partial T}E_k(\omega).$$

We claim that these sets provide the promised almost partition. Indeed, one can also see that

$$E_k=F_k\bigcup\{\text{ immediate parents of points in }F_k\}\bigcup\{\text{ immediate children of points in }F_k\}.$$

It follows that any point $\alpha\in E_k$ cannot be a parent or a children of other F_j's more than once, since $\bigcup_k E_k(\omega)$ covers $\mathcal{P}(\omega)$ at most three times. Finally, if $E_k(\omega)=\emptyset$, so is $F_k(\omega)=F_k\cap\mathcal{P}(\omega)$. Otherwise $\tau_{n_{k+1}}(\omega)$ is the maximal point of $F_{k+1}(\omega)$ – by definition $\mathbf{I}_w f(\tau_{n_{k+1}-1}(\omega))<2^{k+1}$. Clearly,

$$\mathbf{I}_w f|_{E_k(\omega)}(\tau_{n_{k+1}}(\omega))=\mathbf{I}_w f(\tau_{n_{k+1}}(\omega))-\mathbf{I}_w f(\tau_{n_k-1}(\omega))\geq 2^{k+1}-2^k=2^k.$$

Hence $2^{-k}f|_{E_k}$ is admissible for F_{k+1}, and we are done. $\square$

Chapter 2 Hardy embeddings on finite d-trees

2.1 Main results

Let us recall the definitions of the main objects of study and the main statement.
The Hardy embedding inequality is

$$\int_{\overline{T}^d} (\mathbf{I}_w f)^2 \, d\mu \le [w,\mu]_{CE} \int_{T^d} f^2 \, dw, \quad f \in L^2(T^d, w), \tag{2.1}$$

and its dual version is

$$\int_{T^d} (\mathbf{I}^*_\mu \varphi)^2 \, dw \le [w,\mu]_{CE} \int_{\overline{T}^d} \varphi^2 \, d\mu, \quad \varphi \in L^2(\overline{T}^d, \mu). \tag{2.2}$$

The *subcapacitary constant* $[w,\mu]_{SC}$, *hereditary Carleson constant* $[w,\mu]_{HC}$, *Carleson constant* $[w,\mu]_C$, *box constant* $[w,\mu]_B$ are the smallest numbers that realize the respective inequalities below

$$\mu(E) \le [w,\mu]_{SC} \operatorname{Cap}_w(E), \quad \forall E \subset \overline{T}^d, \tag{2.3a}$$

$$\mathcal{E}_w(\mu|_E) = \textstyle\sum_{\alpha\in T^d} w(\alpha)(\mathbf{I}^*\mu|_E)^2(\alpha) \le [w,\mu]_{HC}\mu(E), \quad \forall E \subset \overline{T}^d, \tag{2.3b}$$

$$\textstyle\sum_{\alpha\in\mathcal{S}(E)} w(\alpha)(\mathbf{I}^*\mu)^2(\alpha) \le [w,\mu]_C \mu(E), \quad \forall E \subset \overline{T}^d, \tag{2.3c}$$

$$\textstyle\sum_{\alpha\le\beta} w(\alpha)(\mathbf{I}^*\mu)^2(\alpha) \le [w,\mu]_B \mathbf{I}^*\mu(\beta) = [w,\mu]_B\, \mu(\mathcal{S}(\beta)), \quad \forall \beta \in T^d. \tag{2.3d}$$

The main result of this chapter states that for product weights we have the converse inequalities also for $d = 2$ and $d = 3$.

Theorem 2.1.1 *Assume that (w,μ) is a weight-measure pair on the d-tree with $d = 2$ or $d = 3$, and w has a product structure, $w(\alpha) = \prod_{k=1}^d w(\alpha_k)$ for any $\alpha = (\alpha_1, \dots, \alpha_d) \in T^d$. Then*

$$\begin{aligned} &[w,\mu]_B \gtrsim [w,\mu]_C \gtrsim [w,\mu]_{HC} \gtrsim [w,\mu]_{CE}, \\ &[w,\mu]_{SC} \gtrsim [w,\mu]_{CE}. \end{aligned} \tag{2.4}$$

Remark. It is important that the weight w is of product structure – later on we will provide some counterexamples to (2.4) with non-product weight even on the T^2 (these counterexamples can actually be traced to [86]). The measure μ however is not assumed to be of product structure (otherwise Theorem 2.1.1 would just immediately follow from 1-dimensional result by disintegration), and this complicates things greatly. Also it is not known what happens for $d \ge 4$, we will discuss it later in more detail.

In what follows we will be working on truncated d-tress T^d_N, i.e. on product of d copies of trees cut at the level N. The reason is that it is much more convenient to consider objects on finite graphs (so that, for instance, all the capacities are non-zero), and to obtain the estimate on the infinite d-tree we then just pass to the limit with some care. The details are in Section 2.6.1.
As such in what happens below we do not specify the depth N of our truncated d-tree, we only take care that in any estimate the constants do not depend on N, and we drop the subscript writing T^d instead of T^d_N.

Here we mostly use the arguments from [112]. In dimension 2 the chain $[w,\mu]_C \gtrsim [w,\mu]_{HC} \gtrsim [w,\mu]_{CE}$ was proven in [105] (Theorems 1.8, 1.9) in a slightly different fashion.

2.2 Surrogate Maximum Principle

In this Section we build our main instrument with which we handle various estimates in (2.4) – the so-called *Surrogate Maximum Principle.* As we have seen above, in Proposition 1.1.4 the standard Maximum Principle usually does not hold on T^d, even for $w \equiv 1$. It turns out, however, that we still can salvage some information by estimating the size of the set where it fails. In doing this we also obtain the tail energy estimates for measures on T^d, $d = 2$ or $d = 3$. Within this Section we write $\mathbf{I}$ as the Hardy operator on a tree, bi-tree or a tri-tree, depending on the context. We also make use of the coordinate projections $\mathbf{I_k}$, where

$$\mathbf{I}_k f(\alpha) = \sum_{\beta_k \geq \alpha_k} f(\alpha_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_d), \quad \alpha = (\alpha_1, \dots, \alpha_d) \in T^d,$$

and f is a function on T^d (again, the actual value of the dimension will depend on the context). Projections $\mathbf{I}^*_k$ are defined in the same way,

$$\mathbf{I}^*_k f(\alpha) = \sum_{\beta_k \leq \alpha_k} f(\alpha_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_d), \quad \alpha = (\alpha_1, \dots, \alpha_d) \in T^d.$$

We observe that these projections commute, i.e. for any $A \subset \{1, 2, \dots, d\}$

$$\prod_{k \in A} \mathbf{I}_k = \prod_{k \in A} \mathbf{I}_{\sigma(k)}$$

for any permutation σ of A. Same goes for $\mathbf{I}^*_k$.

2.2.1 Estimates on a tree

We start with a few tree lemmas.

Lemma 2.2.1 *Let $f, g : T \to \mathbb{R}_+$ be any functions. Then*

$$\mathbf{I}f \cdot \mathbf{I}g \leq \mathbf{I}(\mathbf{I}f \cdot g + f \cdot \mathbf{I}g).$$

Proof. Essentially this is just discrete integration by parts. Indeed, given $\alpha \in T$ we have

$$\begin{aligned}
\mathbf{I}f(\alpha)\mathbf{I}g(\alpha) \leq \mathbf{I}f(\alpha)\mathbf{I}g(\alpha) + \mathbf{I}(f \cdot g)(\alpha) = \\
\sum_{\alpha' \geq \alpha,\ \alpha'' \geq \alpha} f(\alpha')g(\alpha'') + \sum_{\alpha' \geq \alpha} f(\alpha')g(\alpha') = \\
\sum_{\alpha' \geq \alpha'' \geq \alpha} f(\alpha')g(\alpha'') + \sum_{\alpha'' \geq \alpha' \geq \alpha} f(\alpha')g(\alpha'') = \\
\sum_{\alpha'' \geq \alpha} \mathbf{I}f(\alpha'')g(\alpha'') + \sum_{\alpha' \geq \alpha} f(\alpha')\mathbf{I}g(\alpha') = \\
\mathbf{I}(\mathbf{I}f \cdot g)(\alpha) + \mathbf{I}(f \cdot \mathbf{I}g)(\alpha).
\end{aligned}$$

□

For the next Lemma we introduce another bit of notation. Given a (finite) tree T the set of children of a vertex $\beta \in T$ consists of the maximal elements of T that are strictly smaller than β,

$$\operatorname{ch}\beta := \max\{\beta' \in T :\ \beta' < \beta\},$$

in particular, if $\beta \in \partial T$, then $\operatorname{ch}\beta = \emptyset$.

A function $g : T \to \mathbb{R}$ is called *superadditive*, if for every $\beta \in T$ we have

$$g(\beta) \geq \sum_{\beta' \in \operatorname{ch}\beta} g(\beta'),$$

and *additive*, if one has the equality in the expression above. The difference operator is defined by

$$\Delta g(\beta) := g(\beta) - \sum_{\beta' \in \operatorname{ch}\beta} g(\beta').$$

This choice of notation is somewhat motivated by the fact that if $G = \mathbf{I}g$, then Δg is just the graph (non-normalized) laplacian of G.

Lemma 2.2.2 *Let T be a finite tree. Then for any pair of functions $f, g : T \to \mathbb{R}$ we have*

$$\sum_{\alpha \in T} f(\alpha)g(\alpha) = \sum_{\alpha' \in T} \Delta f(\alpha')\mathbf{I}g(\alpha'). \tag{2.5}$$

Proof. Summing down the levels of the tree one can easily see that

$$f(\alpha) = \sum_{\alpha' \leq \alpha} \Delta f(\alpha').$$

Hence

$$\begin{aligned}
\sum_{\alpha} f(\alpha)g(\alpha) = \sum_{\alpha \geq \alpha'} \Delta f(\alpha')g(\alpha) = \\
\sum_{\alpha' \in T} \Delta f(\alpha') \sum_{\alpha' \leq \alpha} g(\alpha) = \sum_{\alpha' \in T} \Delta f(\alpha')\mathbf{I}g(\alpha').
\end{aligned}$$

□

We need another 'integration by parts' result, now with $\mathbf{I}^*$.

Lemma 2.2.3 *Let $f, g : T \to \mathbb{R}$. Then*

$$\mathbf{I}^*(fg) = \mathbf{I}^*(\Delta f \cdot Ig) - f(\mathbf{I}g - g)$$

Proof. For a given β the set $\mathcal{S}(\beta)$ is a (sub-)tree itself, with the root β. Applying (2.5) on that tree, we have

$$\mathbf{I}(fg)(\beta) = \int_{\mathcal{S}(\beta)} fg = \int_{\mathcal{S}(\beta)} \Delta f \cdot \mathcal{I}(g\mathbb{1}_{\mathcal{S}(\beta)}).$$

Now, if $\alpha \in \mathcal{S}(\beta)$, then

$$\mathcal{I}(g\mathbb{1}_{\mathcal{S}(\beta)})(\alpha) = \sum_{\alpha\le\gamma\le\beta} g(\gamma) = \sum_{\alpha\le\gamma} g(\gamma) - \sum_{\beta\le\gamma} g(\gamma) + g(\beta) =$$
$$\mathbf{I}g(\alpha) - \mathbf{I}g(\beta) + g(\beta),$$

hence

$$\mathbf{I}^*(fg)(\beta) = \int_{\mathcal{S}(\beta)} \Delta f \cdot (\mathbf{I}g(\alpha) - \mathbf{I}g(\beta) + g(\beta)) =$$
$$\int_{\mathcal{S}(\beta)} \Delta f \cdot \mathcal{I}g - (\mathcal{I}g(\beta) - g(\beta)) \int_{\mathcal{S}(\beta)} \Delta f =$$
$$\mathcal{I}^*(\Delta f \cdot \mathcal{I}g)(\beta) - (\mathcal{I}g(\beta) - g(\beta))f(\beta).$$

□

Corollary 2.2.1 *Given a pair of non-negative functions $f, g : T \to \mathbb{R}_+$ one has*

$$\mathbf{I}^*(fg) \le \mathbf{I}^*(\Delta f \cdot g).$$

2.2.2 Estimates on a bi-tree

We continue by moving up one dimension, and we prove the biparametric analogues of the results above. Our first Lemma is the analogue of Lemma 2.2.1.

Lemma 2.2.4 *Let $f, g : T^2 \to \mathbb{R}_+$. Then*

$$\mathbf{I}f \cdot \mathbf{I}g \le \mathbf{I}(\mathbf{I}f \cdot g + \mathbf{I}_1 f \cdot \mathbf{I}_2 g + \mathbf{I}_2 f \cdot \mathbf{I}_1 g + f\mathbf{I}g).$$

Proof. Since $\mathbf{I}_1$, $\mathbf{I}_2$ commute, we can apply Lemma 2.2.1 obtaining

$$\mathbf{I}f \cdot \mathbf{I}g = (\mathbf{I}_1\mathbf{I}_2 f)(\mathbf{I}_1\mathbf{I}_2 g) \le \mathbf{I}_1((\mathbf{I}_1\mathbf{I}_2 f) \cdot \mathbf{I}_2 g + \mathbf{I}_2 f \cdot (\mathbf{I}_1\mathbf{I}_2 g)),$$

and

$$((\mathbf{I}_1\mathbf{I}_2 f)\cdot \mathbf{I}_2 g + \mathbf{I}_2 f\cdot(\mathbf{I}_1\mathbf{I}_2 g)) \le$$
$$\mathbf{I}_2((\mathbf{I}_2\mathbf{I}_1 f)\cdot g + \mathbf{I}_1 f\cdot \mathbf{I}_2 g) + \mathbf{I}_2(\mathbf{I}_2 f\cdot \mathbf{I}_1 g + f\cdot(\mathbf{I}_1\mathbf{I}_2 g)).$$

Hence

$$\mathbf{I}f\cdot\mathbf{I}g \le \mathbf{I}_1(\mathbf{I}_2((\mathbf{I}_2\mathbf{I}_1 f)\cdot g + \mathbf{I}_1 f\cdot\mathbf{I}_2 g) + \mathbf{I}_2(\mathbf{I}_2 f\cdot\mathbf{I}_1 g + f\cdot(\mathbf{I}_1\mathbf{I}_2 g))) =$$
$$\mathbf{I}(\mathbf{I}f\cdot g + \mathbf{I}_1 f\cdot\mathbf{I}_2 g + \mathbf{I}_2 f\cdot\mathbf{I}_1 g + f\mathbf{I}g).$$

□

Next three results consider the *energy estimates* on the bi-tree. Our end goal is to learn how to estimate the capacities of sets of large potential for measures supported on sets of small potential.

Lemma 2.2.5 *Let $f : T^2 \to \mathbb{R}_+$ be a function which is superadditive separately in each variable, and let $w_1w_2 = w : T^2 \to \mathbb{R}_+$ be a product weight. Suppose also that* $\operatorname{supp} f \subset \{\mathbf{I}(wf) \le \delta\}$. *Then*

$$\int_{T^2} wf\cdot\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f)\cdot\mathbf{I}(wf) \le \delta^2\int_{T^2} f^2w.$$

Proof. Observe that (due to the product structure) the operator $\mathbf{I}_1$ 'does not see' w_2, that is $\mathbf{I}_1(wf) = w_2\mathbf{I}_1(w_1f)$. We then apply the hypothesis removing one of the terms in the left-hand side,

$$\begin{aligned}&\int_{T^2} wf\cdot\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f)\cdot\mathbf{I}(wf) \le \delta\int_{T^2} wf\cdot\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f) =\\ &\delta\int_{T^2} w_1f\cdot\mathbf{I}_1(wf)\cdot\mathbf{I}_2(w_2f) = \delta\int_{T^2} wf\cdot\mathbf{I}_1^*\left(w_1f\cdot\mathbf{I}_2(w_2f)\right) =\\ &\delta\int_{T^2} wf\cdot\mathbf{I}_1^*\left(f\cdot\mathbf{I}_2(wf)\right).\end{aligned} \tag{2.6}$$

By Corollary 2.2.1 we have

$$\mathbf{I}_1^*(f\cdot\mathbf{I}_2(wf)) \le \mathbf{I}_1^*\left(\Delta_1 f\cdot\mathbf{I}_1\mathbf{I}_2(wf)\right).$$

The support of f is an *up-set*, i.e. it contains all the ancestors of its elements, hence $\Delta_1 f$ is also supported there, and is positive, $\Delta_1 f \ge 0$ by superadditivity of f. Hence,

$$\mathbf{I}_1^*\left(f\cdot\mathbf{I}_2(wf)\right) \le \mathbf{I}_1^*\left(\Delta_1 f\cdot\mathbf{I}(wf)\right) \le \mathbf{I}_1^*(\Delta_1 f\cdot\delta) = \delta f. \tag{2.7}$$

It remains to plug (2.7) into (2.6). □

Lemma 2.2.6 *Under the conditions of previous Lemma we also have*

$$\int_{T^2} w(\mathbf{I}_1(w_1f))^2(\mathbf{I}_2(w_2f))^2 \le 4\delta^2\int_{T^2} f^2w.$$

Proof. By Lemma 2.2.1 and commutativity of $\mathbf{I}_1$, $\mathbf{I}_2$ we have

$$\int_{T^2} w(\mathbf{I}_1(w_1 f))^2(\mathbf{I}_2(w_2 f))^2 \leq 4\int_{T^2} w\mathbf{I}_1\left(w_1 f\cdot \mathbf{I}_1(w_1 f)\right)\cdot \mathbf{I}_2\left(w_2 f\cdot \mathbf{I}_2(w_2 f)\right) =$$
$$4\int_{T^2} \mathbf{I}_1\left(w_1 f\cdot \mathbf{I}_1(wf)\right)\cdot \mathbf{I}_2\left(w_2 f\cdot \mathbf{I}_2(wf)\right) =$$
$$4\int_{T^2} \mathbf{I}_2^*\left(w_1 f\cdot \mathbf{I}_1(wf)\right)\cdot \mathbf{I}_1^*\left(w_2 f\cdot \mathbf{I}_2(wf)\right) =$$
$$4\int_{T^2} w\mathbf{I}_2^*\left(f\cdot \mathbf{I}_1(wf)\right)\cdot \mathbf{I}_1^*\left(f\cdot \mathbf{I}_2(wf)\right).$$

It remains to use (2.7) □

We are ready to prove the energy majorization Lemma. Namely, we show that if f is supported on the set of small ($\leq \delta$) $\mathbb{I}_w$-potential (w.r.t. the function f itself), then $\mathbf{I}_w f$, generally speaking, is not the most effective way, in the sense of energy, to obtain its own values on the set of large ($\geq \lambda \geq 4\delta$) potential. Let us formulate this in a more precise way.

Lemma 2.2.7 *Assume that $f: T^2 \to \mathbb{R}_+$ satisfies the conditions of previous two Lemmas, and assume $\lambda \geq 4\delta$. Then there exists the energy-efficient redistribution $\varphi: T^2 \to \mathbb{R}_+$ of f such that*

$$\mathbf{I}(w\varphi)(\alpha) \geq \mathbf{I}(wf)(\alpha), \quad \alpha \in \{\lambda \leq \mathbf{I}(wf) \leq 2\lambda\}, \tag{2.8a}$$
$$\int_{T^2} \varphi^2 w \leq C\frac{\delta^2}{\lambda^2}\int_{T^2} f^2 w, \tag{2.8b}$$

where C is some absolute constant.

Proof. Since $2\delta\lambda^{-1} \leq \frac{1}{2}$, we have

$$(\mathbf{I}(wf))\mathbb{1}_{\{\mathbf{I}(wf)\geq\lambda\}} \leq 4\lambda^{-1}\mathbf{I}\left(\mathbf{I}_1(w_1 f)\cdot \mathbf{I}_2(w_2 f)\right),$$

hence

$$(\mathbf{I}(wf))\mathbb{1}_{\{\lambda\leq\mathbf{I}(wf)\leq2\lambda\}} \leq 4\lambda^{-1}\mathbf{I}\left(\mathbf{I}_1(w_1 f)\cdot \mathbf{I}_2(w_2 f)\right)\mathbb{1}_{\{\lambda\leq\mathbf{I}(wf)\leq2\lambda\}} \leq$$
$$4\lambda^{-1}\mathbf{I}\left(\mathbf{I}_1(w_1 f)\cdot \mathbf{I}_2(w_2 f)\cdot \mathbb{1}_{\{\lambda\leq\mathbf{I}(wf)\leq2\lambda\}}\right).$$

Let

$$\varphi := 4\lambda^{-1}\mathbf{I}\left(\mathbf{I}_1(w_1 f)\cdot \mathbf{I}_2(w_2 f)\cdot \mathbb{1}_{\{\lambda\leq\mathbf{I}(wf)\leq2\lambda\}}\right).$$

Clearly, φ satisfies the condition (2.8a) of the statement. To get (2.8b) we just aaply Lemma 2.2.6. □

Now the reason why we do care about this Lemma is that it allows us to estimate the size of 'exceptional' sets. In other words, if μ is a measure on T^2 supported on the set $\{\mathbf{V}_w^\mu \leq 1\}$ of 'small' potential, then its set of $E_\lambda = \{\mathbf{V}_w^\mu \geq \lambda \geq 10\}$ 'large' potential generally should not be empty. It is possible (for product weights) to estimate its capacity, or, to be more precise, to improve on the standard weak estimate.

Theorem 2.2.1 *Let μ be as above. Then*

$$\operatorname{Cap}_w E_\lambda \leq \frac{C}{\lambda^4}\mathcal{E}_w[\mu],$$

where C is an absolute constant.

Remark. Observe that $\frac{1}{\lambda}\mathbf{I}^*\mu$ is admissible (i.e. $\mathbf{I}\left(\frac{1}{\lambda}\mathbf{I}^*\mu\right) \geq 1$ on E_λ). By definition of capacity it follows immediately that

$$\operatorname{Cap}_w E_\lambda \leq \frac{1}{\lambda^2}\mathcal{E}_w[\mu].$$

This is a trivial estimate. It turns out, however, that we can say a bit more – that the capacity of E_λ decays faster than expected as $\lambda \to \infty$.

Proof. Put $f := \mathbf{I}^*\mu$ and $\delta := 1$. If $f(\alpha) \neq 0$ then there exists a point $\beta \leq \alpha$ such that $\beta \in \operatorname{supp}\mu$. By our assumption it means that $\mathbf{I}_w f = \mathbf{V}_w^\mu(\beta) \leq 1$, so, by motonicity of $\mathbf{I}$ it follows that $\mathbf{I}_w f(\alpha) \leq 1$. Hence

$$\operatorname{supp} f \subset \{\mathbf{I}_w f \leq \delta = 1\},$$

and we fall squarely into assumptions of Lemma 2.2.7. We apply it with data $(f; \delta = 1; \lambda_m = 2^m\lambda)$ to get functions φ_m, $m = 0, 1, \ldots$ such that

$$\mathbf{I}_w\varphi_m \geq \mathbf{I}_w f = \mathbf{V}_w^\mu, \quad \text{where } \mathbf{V}_w^\mu \in [2^m\lambda, 2^{m+1}\lambda],$$

which means that

$$2^{-m}\lambda^{-1}\mathbf{I}_w\varphi_m \geq 1, \quad \text{where } \mathbf{V}_2^\mu \in [2^m\lambda, 2^{m+1}\lambda].$$

Let us sum everything up – let $\varphi := \sum_{m\geq 0} 2^{-m}\lambda^{-1}\varphi_m$, we first obtain

$$\mathbf{I}_w\varphi \geq 1, \quad \text{where } \mathbf{V}_w^\mu \in [\lambda, \infty),$$

and also

$$\begin{aligned}
&\int_{T^2}\varphi^2 w \leq \left(\lambda^{-1}\sum_{m\geq 0} 2^{-m}\left(\int_{T^2}\varphi_m^2 w\right)^{\frac{1}{2}}\right)^2 \leq \\
&C\left(\lambda^{-1}\sum_{m\geq 0}\lambda^{-1}2^{-2m}\left(\int_{T^2} f^2 w\right)^{\frac{1}{2}}\right)^2 \leq \\
&C'\lambda^{-2}\int_{T^2} f^2 w.
\end{aligned}$$

Since $f = \mathbf{I}^*\mu$,

$$\int_{T^2} f^2 w = \int_{T^2}(\mathbf{I}^*\mu)^2 w = \int_{T^2}\mathbf{V}_2^\mu\, d\mu = \mathcal{E}_w[\mu],$$

and we are done. $\square$

Remark. The precise decay rate λ^{-4} is probably not the best possible here (we believe we can actually improve it to $e^{-c\sqrt{\lambda}}$ at least for $w \equiv 1$), however we do know that the lower bound is indeed exponential, that is there exists a measure μ such that $\mathbf{V}^\mu \leq 1$ on $\operatorname{supp}\mu$ but with absolute positive constant c the following holds

$$\operatorname{Cap}(\{\mathbf{V}^\mu > \lambda\}) \geq ce^{-2\lambda},$$

see the lack of Maximum Principle (Proposition 1.1.4).

Finally we need yet another energy estimate, this time about mixed energy behaviour. We recall that the *delta*-truncated potential and energy are

$$\mathbf{V}^\mu_{w,\delta}(\alpha) = \sum_{\beta\geq\alpha:\ \mathbf{V}^\mu_w(\beta)\leq\delta} w(\beta)\mathbf{I}^*\mu(\beta)$$
$$\mathcal{E}_{w,\delta}[\mu] = \int_{T^2} \mathbf{V}^\mu_{w,\delta} d\mu.$$

Lemma 2.2.8 *Let μ, ρ be positive measures on T^2 and $\delta > 0$. Let $w : T^2 \to \mathbb{R}_+$ be a product weight. Then*

$$\left(\int_{T^2} \mathbf{V}^\mu_{w,\delta}\, d\rho\right)^4 \leq 28\cdot\delta^2\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho]|\rho|^2.$$

Proof. Let $f := \mathbb{1}_{\mathbf{V}^\mu_w\leq\delta}\mathbf{I}^*\mu$. Then

$$\int_{T^2} \mathbf{V}^\mu_{w,\delta}\, d\rho = \int_{T^2} \mathbf{I}(fw)\, d\rho \leq |\rho|^{\frac12}\left(\int_{T^2} (\mathbf{I}(wf))^2\, d\rho\right)^{\frac12} \leq$$

by Lemma 2.2.4

$$|\rho|^{\frac12}\left(2\int_{T^2} \mathbf{I}\left(\mathbf{I}_1(wf)\cdot\mathbf{I}_2(wf) + (wf)\cdot\mathbf{I}(wf)\right)\, d\rho\right)^{\frac12} =$$
$$2^{\frac12}|\rho|^{\frac12}\left(\int_{T^2} w\left(\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f) + (wf)\cdot\mathbf{I}(wf)\right)\mathbf{I}^*\rho\right)^{\frac12} \leq$$
$$2^{\frac12}|\rho|^{\frac12}\mathcal{E}_w^{\frac14}[\rho]\left(\int_{T^2} w\left(\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f) + (wf)\cdot\mathbf{I}(wf)\right)^2\right)^{\frac14}.$$

Now we expand the square while applying Lemmas 2.2.5 and 2.2.6, and obtain

$$2^{\frac12}|\rho|^{\frac12}\mathcal{E}_w^{\frac14}[\rho]\left(\int_{T^2} w\left(\mathbf{I}_1(w_1f)\cdot\mathbf{I}_2(w_2f) + (wf)\cdot\mathbf{I}(wf)\right)^2\right)^{\frac14} \leq$$
$$2^{\frac12}|\rho|^{\frac12}\mathcal{E}_w^{\frac14}[\rho]\left(7\delta^2\int_{T^2} f^2 w\right)^{\frac14} =$$
$$28^{\frac14}|\rho|^{\frac12}\mathcal{E}_w^{\frac14}[\rho]\delta^{\frac12}\mathcal{E}_{w,\delta}^{\frac14}[\mu].$$

We are done. $\square$

2.2.3 Estimates on a tri-tree

Now we repeat most of the results from the previous Section on T^3 (and the Hardy operator is now three-dimensional one). Similarly to Lemma 2.2.4 we obtain the following result for tri-trees.

Lemma 2.2.9 *Let $f, g : T^3 \to \mathbb{R}_+$. Then*

$$\mathbf{I}f \cdot \mathbf{I}g \leq \mathbf{I}\left(\sum_{A \subseteq \{1,2,3\}} \mathbf{I}_A f \cdot \mathbf{I}_{A^c} g\right),$$

where $\mathbf{I}_A = \prod_{j \in A} \mathbf{I}_j$ is the composition of the respective coordinate Hardy operators.

Proof. It goes exactly like in Lemma 2.2.4 with appropriate changes. □

Corollary 2.2.2 *Let $0 < \delta < \frac{\lambda}{4}$ and let $f : T^3 \to \mathbb{R}_+$ with* $\operatorname{supp} f \subset \{\mathbf{I}f \leq \delta\}$. *Then*

$$(\mathbf{I}f)\mathbb{1}_{\lambda \leq \mathbf{I}f \leq 2\lambda} \leq 4\lambda^{-1}\left(\sum_{j \in \{1,2,3\}} \mathbf{I}_j f \cdot \mathbf{I}_{\{j\}^c} f \cdot \mathbb{1}_{\mathbf{I}f \leq 2\lambda}\right),$$

where $\mathbf{I}_{\{j\}^c} = \prod_{k \neq j} \mathbf{I}_k$.

Proof. Substituting $f = g$ Lemma 2.2.9 implies that

$$(\mathbf{I}f)^2 \leq \mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j f \cdot \mathbf{I}_{\{j\}^c} f + 2f \cdot \mathbf{I}f\right).$$

Now we apply the support condition, obtaining

$$(\mathbf{I}f)\mathbb{1}_{\lambda \leq \mathbf{I}f \leq 2\lambda} \leq \lambda^{-1}\mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j \cdot \mathbf{I}_{\{j\}^c} + 2\delta f\right) \leq$$
$$\lambda^{-1}\mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j f \cdot \mathbf{I}_{\{j\}^c} f\right) + 2\delta\lambda^{-1}\mathbf{I}f.$$

Since $2\delta\lambda^{-1} \leq \frac{1}{2}$, we see that

$$(\mathbf{I}f)\mathbb{1}_{\lambda \leq \mathbf{I}f \leq 2\lambda} \leq 2\lambda^{-1}\mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j f \cdot \mathbf{I}_{\{j\}^c} f\right) \cdot \mathbb{1}_{\lambda \leq \mathbf{I}f \leq 2\lambda} \leq$$
$$2\lambda^{-1}\left(2\sum_{j=1,2,3} \mathbf{I}_j f \cdot \mathbf{I}_{\{j\}^c} f \cdot \mathbb{1}_{\mathbf{I}f \leq 2\lambda}\right).$$

□

Lemma 2.2.10 (Energy bound on T^3) *Let $f : T^3 \to \mathbb{R}_+$ be superadditive in each variable. Let $w : T^3 \to \mathbb{R}_+$ be a product weight, and suppose that* $\operatorname{supp} f \subseteq \{\mathbf{I}_w f \leq \delta\}$. *Then for any $j = 1, 2, 3$*

$$\int_{T^3} w \cdot \left(\mathbf{I}_j(w_j f) \cdot \mathbb{I}_{\{j\}^c}(w_{\{j\}^c} f)\right)^2 \mathbb{1}_{\mathbf{I}(wf) \leq \lambda} \leq 2\delta\lambda \int_{T^3} f^2 w,$$

where $w_{\{j\}^c} = \prod_{k\neq j} w_k$.

Proof. Without any loss of generality we may assume that $j=1$. Then by Lemma 2.2.1 we have

$$\begin{aligned}
&\int_{T^3} w\cdot(\mathbf{I}_1(w_1 f)\cdot \mathbf{I}_2\mathbf{I}_3(w_2w_3f))^2\, \mathbf{I}\mathrm{f} \le \lambda \le \\
&2\int_{T^3} w\mathbf{I_1}\left((w_1 f\cdot \mathbf{I}_1(w_1 f))\right)\cdot (\mathbf{I}_2\mathbf{I}_3(w_2w_3f))^2\cdot \mathbb{1}_{\mathbf{I}(wf)\le\lambda} = \\
&2\int_{T^3} \mathbf{I}_1\left(w_1 f\cdot \mathbf{I}_1(wf)\right)\cdot(\mathbf{I}_2\mathbf{I}_3(w_2w_3f))\cdot(\mathbf{I}_2\mathbf{I}_3(wf))\cdot \mathbb{1}_{\mathbf{I}(wf)\le\lambda} = \\
&2\int_{T^3} w_1 f\cdot \mathbf{I}_1(wf)\cdot \mathbf{I}_1^*\left((\mathbf{I}_2\mathbf{I}_3(w_2w_3f))\cdot(\mathbf{I}_2\mathbf{I}_3(wf))\cdot \mathbb{1}_{\mathbf{I}(wf)\le\lambda}\right).
\end{aligned} \tag{2.9}$$

By Corollary 2.2.1 we have

$$\begin{aligned}
&\mathbf{I}_1^*\left((\mathbf{I}_2\mathbf{I}_3(w_2w_3f))\cdot(\mathbf{I}_2\mathbf{I}_3(wf))\cdot \mathbb{1}_{\mathbf{I}(wf)\le\lambda}\right) \le \\
&\mathbf{I}_1^*\left(\Delta_1\left(\mathbb{1}_{\mathbf{I}(wf)\le\lambda}\mathbf{I}_2\mathbf{I}_3(w_2w_3f)\right)\cdot \mathbf{I}_1\left(\mathbf{I}_2\mathbf{I}_3(wf)\right)\right).
\end{aligned}$$

Since $\{\mathbf{I}(wf)\le\lambda\}$ is an up-set and f is superadditive, we have $\Delta_1\left(\mathbb{1}_{\mathbf{I}(wf)\le\lambda}\mathbf{I}_2\mathbf{I}_3(w_2w_3f)\right)\ge 0$ and $\mathbf{I}_1\left(\mathbf{I}_2\mathbf{I}_3(wf)\right) = \mathbf{I}(wf)\le\lambda$ on the support of the former function. Hence

$$\begin{aligned}
&\mathbf{I}_1^*\left((\mathbf{I}_2\mathbf{I}_3(w_2w_3f))\cdot(\mathbf{I}_2\mathbf{I}_3(wf))\cdot \mathbb{1}_{\mathbf{I}(wf)\le\lambda}\right) \le \\
&\mathbf{I}_1^*\left(\Delta_1\left(\mathbb{1}_{\mathbf{I}(wf)\le\lambda}\mathbf{I}_2\mathbf{I}_3(w_2w_3f)\right)\cdot\lambda\right) = \\
&\lambda\mathbb{1}_{\mathbf{I}(wf)\le\lambda}\cdot \mathbf{I}_2\mathbf{I}_3(w_2w_3f).
\end{aligned}$$

This bound implies

$$\begin{aligned}
(2.9)\le & 2\lambda\int_{T^3} w_1 f\cdot \mathbf{I}_1(wf)\cdot \mathbf{I}_2\mathbf{I}_3(w_2w_3f) = \\
& 2\lambda\int_{T^3} f\cdot \mathbf{I}_1(wf)\cdot \mathbf{I}_2\mathbf{I}_3(wf) = 2\lambda\int_{T^3} wf\cdot \mathbf{I}_1^*(f\cdot \mathbf{I}_2\mathbf{I}_3(wf)).
\end{aligned}$$

Similar to (2.7) we see that

$$\mathbf{I}_1^*(f\cdot \mathbf{I}_2\mathbf{I}_3(wf))\le \delta f.$$

We are done. □

Next Lemma is the three-dimensional version of Lemma 2.2.7 (observe a slightly worse decay rate here).

Lemma 2.2.11 (Small energy majorization on T^3.) *Let $f: T^3\to\mathbb{R}_+$ be a superadditive (in each variable) function, and $w: T^3\to\mathbb{R}_+$ be a product weight. Suppose that* $\operatorname{supp} f\subseteq\{\mathbf{I}(wf)\le\delta\}$, *and let $\lambda\ge 4\delta$. Then there exists an energy-effective redistribution $\varphi: T^3\to\mathbb{R}_+$ such that*

$$\mathbf{I}(w\varphi)(\alpha)\ge \mathbf{I}(wf)(\alpha),\quad \alpha\in\{\lambda\le \mathbf{I}(wf)\le 2\lambda\}, \tag{2.10a}$$

$$\textstyle\int_{T^3}\varphi^2 w\le C\frac{\delta}{\lambda}\int_{T^3} f^2 w, \tag{2.10b}$$

where C is some absolute constant.

Proof. Since $2\delta\lambda^{-1} \le \frac{1}{2}$, we have

$$(\mathbf{I}(wf)) \cdot \mathbb{1}_{\lambda \le \mathbf{I}(wf) \le 2\lambda} \le 2\lambda^{-1}\mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j(w_j f) \cdot \mathbf{I}_{\{j\}^c}(w_{\{j\}^c} f)\right) \cdot \mathbb{1}_{\lambda \le \mathbf{I}(wf) \le 2\lambda}$$
$$2\lambda^{-1}\mathbf{I}\left(2\sum_{j=1,2,3} \mathbf{I}_j(w_j f) \cdot \mathbf{I}_{\{j\}^c}(w_{\{j\}^c} f) \cdot \mathbb{1}_{\lambda \le \mathbf{I}(wf) \le 2\lambda}\right).$$

We, therefore, get (2.10a), and to obtain (2.10b) just apply Lemma 2.2.10 □

The exceptional set estimate follows as well, albeit again with a worse decay rate.

Theorem 2.2.2 *Let μ be a measure on T^3, $w : T^3 \to \mathbb{R}_+$ a product weight and $\mathbf{V}_w^\mu \le 1$ on $\operatorname{supp}\mu$. Let $E_\lambda := \{\mathbf{V}_w^\mu \ge \lambda \ge 10\}$. Then*

$$\operatorname{Cap}_w E_\lambda \le \frac{C}{\lambda^3}\mathcal{E}_w[\mu],$$

where C is an absolute constant.

Proof. We basically repeat the proof of Theorem 2.2.1, only instead we use small majorization on tri-tree, Lemma 2.2.11 instead of Lemma 2.2.7. □

Remark. Again we do not know how precise the rate λ^{-3} is. It obviously should be worse than the one on the bi-tree, but we do not know anything beyond that.

We continue to estimate the mixed energy on T^3.

Lemma 2.2.12 *Let μ, ρ be positive measures on T^3 and $\delta > 0$. Let $w : T^3 \to \mathbb{R}_+$ be a product weight. Then*

$$\left(\int_{T^3} \mathbf{V}_{w,\delta}^\mu \, d\rho\right)^3 \le C \cdot \delta\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho]|\rho|. \tag{2.11}$$

Proof. Without loss of generality $\mathcal{E}_{w,\delta}[\mu] \ne 0$ and $\rho \not\equiv 0$. Let $\lambda > 0$ be chosen later. Let $f := \mathbf{I}^*\mu \cdot \mathbb{1}_{\mathbf{V}_w^\mu \le \delta}$, this function is clearly superadditive (in each variable). Also, $\mathbf{I}(wf) = \mathbf{V}_{w,\delta}^\mu \le \mathbf{V}_w^\mu \le \delta$ on $\operatorname{supp} f$, and $\mathcal{E}_{w,\delta}[\mu] = \int_{T^3} f^2 w$.

For $m = 1, 0, \ldots$ let

$$\varphi_m := 4(2^m\lambda)^{-1}\left(\sum_{j=1,2,3} \mathbf{I}_j(w_j f) \cdot \mathbf{I}_{\{j\}^c}(w_{\{j\}^c} f) \cdot \mathbb{1}_{\mathbf{I}(wf) \le 2^{m+1}\lambda}\right).$$

Then by Corollary 2.2.2 with wf in place of f we have

$$\mathbf{I}(wf) \cdot \mathbb{1}_{2^m\lambda < \mathbf{I}(wf) < 2^{m+1}\lambda} \le \mathbf{I}(w\varphi_m),$$

and, by Lemma 2.2.10, we have

$$\int_{T^3} \varphi_m^2 w \le C\frac{\delta}{2^m\lambda}\int_{T^3} f^2 w.$$

Hence

$$\int_{T^3} \mathbf{V}^\mu_{w,\delta}\,d\rho = \int_{\{\mathbf{V}^\mu_{w,\delta}\le\lambda\}} \mathbf{V}^\mu_{w,\delta}\,d\rho + \sum_{m\ge0}\int_{\{2^m\lambda<\mathbf{V}^\mu_{w,\delta}\le 2^{m+1}\lambda\}} \mathbf{V}^\mu_{w,\delta}\,d\rho \le$$
$$\lambda|\rho| + \sum_{m\ge0}\int_{T^3}\mathbf{I}(w\varphi_m)\,d\rho = \lambda|\rho| + \sum_{m\ge0} w\varphi_m\mathbf{I}^*\rho \le$$
$$\lambda|\rho| + \sum_{m\ge0}\left(\int_{T^3}\varphi_m^2 w\right)^{\frac12}\mathcal{E}_w^{\frac12}[\rho] \le$$
$$\lambda|\rho| + \sum_{m\ge0} c\left(\frac{\delta}{2^m\lambda}\right)^{\frac12}\mathcal{E}_{w,\delta}^{\frac12}[\mu]\mathcal{E}_w^{\frac12}[\rho] \le$$
$$\lambda|\rho| + C\left(\frac{\delta}{\lambda}\right)^{\frac12}\mathcal{E}_{w,\delta}^{\frac12}[\mu]\mathcal{E}_w^{\frac12}[\rho].$$

Substituting $\lambda = (\delta\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho])^{\frac13}|\rho|^{-\frac23}$ we obtain (2.11). □

Corollary 2.2.3 *Let μ,ρ be positive measures on T^3 and $\delta>0$. Then*

$$\int_{T^3}\mathbf{V}^\mu_{w,\delta}\,d\rho \le C_{(2.11)}^{\frac12}\delta^{\frac12}\mathcal{E}_w^{\frac16}[\mu]|\mu|^{\frac16}\mathcal{E}_w^{\frac13}[\rho]|\rho|^{\frac13}.$$

Proof. By Lemma 2.2.12 (applied twice) we have

$$\left(\int_{T^3}\mathbf{V}^\mu_{w,\delta}\,d\rho\right)^{\frac13} \le C_{(2.11)}\delta\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho]|rho| \le$$
$$C_{(2.11)}\delta\left(C_{(2.11)}\delta\mathcal{E}_w[\mu]|\mu|\right)^{\frac12}\mathcal{E}_w[\rho]|\rho|,$$

and we are done. □

2.2.4 Estimates on d-trees (conditional)

We say that a weight w on a d-tree T^d satisfies *the surrogate maximum principle*, if for some $\kappa>0$, $C<\infty$ and every positive measures $\mu,\rho: T^d\to[0,\infty)$ and $\delta>0$ one has

$$\int_{T^d}\mathbf{V}^\mu_{w,\delta}\,d\rho \le C\,(\delta|\rho|)^\kappa\,(\mathcal{E}_{w,\delta}[\mu]\mathcal{E}_w[\rho])^{\frac{1-\kappa}{2}}\,. \tag{2.12}$$

For $d=1,2,3$ every weight of product form satisfies this principle – as we have shown above, with $\kappa=\frac1d$ and C independent of w. This leads us to the following conjecture

Conjecture 2.2.1 (Surrogate maximum principle for d-trees) *Let w be of product type. Then w satisfies the surrogate maximum principle with $\kappa=\frac1d$ and $C=C(d)$ independent of w.*

It is very important that w is of product form here, since if it is not, one can construct a counterexample even on T^2 and w taking only values 0 and 1 (and monotone as well). While it seems very believable that the surrogate maximum principle holds for higher dimensions, there are some issues here that prohibit just repeating the proof (and, strangely enough, they happen not in the usual '$d = 1$ to $d = 2$' transition).

In what follows we can actually work conditionally on the surrogate maximum principle (SMP). All implicit constants are allowed to depend on κ, C in (2.12) but not on w. As we have already mentioned, our results hold unconditionally for $d = 1, 2, 3$.

Taking $\rho = \mu$ in (2.12) we obtain the following Lemma.

Lemma 2.2.13 *Let $w : T^d \to \mathbb{R}_+$ be such that the SMP holds. Let μ be a positive measure on T^d and $\delta > 0$. Then*

$$\int_{T^d} \mathbf{V}^{\mu}_{w,\delta}\, d\mu \le C_{(2.12)}^{\frac{2}{1+\kappa}} (\delta|\mu|)^{\frac{2\kappa}{1+\kappa}} \mathcal{E}_w^{\frac{1-\kappa}{1+\kappa}}[\mu]. \tag{2.13}$$

Conjecture 2.2.2 *For all positive integers d one has*

$$\int_{T^d} \mathbf{V}^{\mu}_{w,\delta}\, d\mu \le C_n (\delta|\mu|)^{\frac{2}{d+1}} \mathcal{E}_w^{\frac{d-1}{d+1}}[\mu].$$

2.3 Subcapacitary condition

Now that we have done the preliminary work estimating the various mixed energies on d-trees we are ready to handle our first condition – the subcapacitary one. Namely, we are going to prove the following proposition.

Proposition 2.3.1 *Assume that SMP holds for a weight w, and let μ be a non-negative measure on T^d satisfying*

$$\mu(E) \le C_\mu \operatorname{Cap}_w(E), \quad \forall E \subset T^d. \tag{2.14}$$

Then for any function $f : T^d \to \mathbb{R}_+$ one has

$$\int_{T^d} (\mathbf{I}_w f)^2\, d\mu \le C' \int_{T^d} f^2 w, \tag{2.15}$$

where C' depends only on C_μ and $C_{(2.12)}$. Or, in other words, $[w,\mu]_{SC} \gtrsim [w,\mu]_{CE}$.

We prove this in several steps, following mostly [1, Chapter 7]. First, we modify the left-hand side of (2.15) by using the subcapacitary condition to arrive to the so-called *Strong Capacitary Inequality*, which we then proceed to prove. Next, we separate the ℓ^2-norm of f, reducing (2.15) to estimates of the level sets of $\mathbf{I}_w f$. We then invoke (2.12) to verify that the energy scalar product of two equilibrium measures can be estimated by the capacities of the respective sets. This is the key point of the argument. We finish the proof by showing that the mixed energy of the level sets (energy scalar product of their equilibrium measures) is concentrated on the diagonal.

2.3.1 Proof of Proposition 2.3.1: Strong Capacitary Inequality

Let w, μ be like in the statement of Proposition, and let f be any (non-negative) function in $L^2(T^d, w)$. Given $k \in \mathbb{Z}$ let

$$E_k := \{\alpha \in T^d : \mathbf{I}_w f(\alpha) > 2^k\}.$$

Writing the distribution of the left-hand side of (2.15) (or using 'spherical' coordinates of the level sets of $\mathbf{I}_w f$) and applying the subcapacitary condition (2.14) we obtain

$$\int_{T^d} (\mathbf{I}_w f)^2 \, d\mu \lesssim \sum_{k\in\mathbb{Z}} 2^{2k}\mu(E_k) \lesssim \sum_{k\in\mathbb{Z}} 2^{2k} \operatorname{Cap}_w(E_k).$$

Our job is done then, if we manage to prove the following estimate – Strong Capacitary Inequality

$$\sum_{k\in\mathbb{Z}} 2^{2k} \operatorname{Cap}_w(E_k) \lesssim \int_{T^d} f^2 w. \tag{2.16}$$

The rest of the proof is dedicated to this estimate.

Remark. Observe that the weak capacitary inequality,

$$2^{2k} \operatorname{Cap}_w(E_k) \leq \int_{T^d} f^2 w$$

is absolutely trivial, holds for all weights and d-trees, and follows immediately from the definition of capacity.

2.3.2 Getting rid of f

Let μ_k be the equilibrium measure for E_k. Essentially we need the following properties of equilibrium measures

$$\begin{aligned} &\operatorname{Cap}_w(E_k) = |\mu_k| = \mathcal{E}_w[\mu_k] \\ &\mathbf{V}_w^{\mu_k} = 1 \ \text{ on } \operatorname{supp}\mu_k \\ &\mathbf{V}_w^{\mu_k} \geq 1 \ \text{ on } E_k. \end{aligned} \tag{2.17}$$

The left-hand side of (2.16) is

$$\begin{aligned} \sum_{k\in\mathbb{Z}} 2^{2k} \operatorname{Cap}_w(E_k) = \sum_{k\in\mathbb{Z}} 2^k \int_{T^d} 2^k \, d\mu_k \leq \sum_{k\in\mathbb{Z}} 2^k \int_{T^d} \mathbf{I}_w f \, d\mu_k = \\ \sum_{k\in\mathbb{Z}} 2^k \int_{T^d} f w \mathbf{I}^*\mu_k = \int_{T^d} f w \left(\sum_{k\in\mathbb{Z}} 2^k \mathbf{I}^*\mu_k\right) \leq \\ \left(\int_{T^d} f^2 w\right)^{\frac{1}{2}} \left(\int_{T^d} \left(\sum_{k\in\mathbb{Z}} 2^k \mathbf{I}^*\mu_k\right)^2 w\right)^{\frac{1}{2}} \end{aligned}$$

We deduce that (2.16) would follow from

$$\int_{T^d} \left(\sum_{k\in\mathbb{Z}} 2^k \mathbf{I}^*\mu_k\right)^2 w \lesssim \sum_{k\in\mathbb{Z}} 2^{2k} \operatorname{Cap}_w(E_k) = \int_{T^d} \sum_{k\in\mathbb{Z}} 2^{2k} \left(\mathbf{I}^*\mu_k\right)^2 w, \tag{2.18}$$

or, in other words, that the main term in the left-hand side of (2.18) is the diagonal one.

2.3.3 The diagonal term estimate

Expanding the left-hand side of (2.18) we get

$$\sum_{k,j\in\mathbb{Z}} 2^{k+j}\int_{T^d}\mathbf{I}^*\mu_k\cdot\mathbf{I}^*\mu_j\cdot w=\sum_{k,j\in\mathbb{Z}}2^{k+j}\int_{T^d}\mathbf{V}_w^{\mu_k}\,d\mu_j\leq 2\sum_{j\leq k\in\mathbb{Z}}2^{k+j}\int_{T^d}\mathbf{V}_w^{\mu_k}\,d\mu_j$$

By the Surrogate Maximum Principle (2.12) with data $\delta=1$, $\mu=\mu_j$ and $\rho=\mu_k$we have

$$\int_{T^d}\mathbf{V}_w^{\mu_k}\,d\mu_j\lesssim\mathcal{E}_w^{\frac{1-\kappa}{2}}[\mu_j]\mathcal{E}_w^{\frac{1+\kappa}{2}}[\mu_k]=|\mu_k|^{\frac{1+\kappa}{2}}\cdot|\mu_j|^{\frac{1-\kappa}{2}},$$

since both measures are equilibrium ones. Plugging it back into the inequalities above and applying Hölder inequality twice we get

$$\sum_{j\leq k\in\mathbb{Z}}2^{k+j}\int_{T^d}\mathbf{V}_w^{\mu_k}\,d\mu_j\lesssim\sum_{j\leq k\in\mathbb{Z}}2^{k+j}|\mu_k|^{\frac{1+\kappa}{2}}\cdot|\mu_j|^{\frac{1-\kappa}{2}}=\sum_{k\in\mathbb{Z}}2^{k(1+\kappa)}|\mu_k|^{\frac{1+\kappa}{2}}\cdot 2^{-k\kappa}\sum_{j\leq k}2^j|\mu_j|^{\frac{1-\kappa}{2}}\leq$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{k\in\mathbb{Z}}2^{-k\frac{2\kappa}{1-\kappa}}\left(\sum_{j\leq k}2^j|\mu_j|^{\frac{1-\kappa}{2}}\right)^{\frac{2}{1-\kappa}}\right)^{\frac{1-\kappa}{2}}=$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{k\in\mathbb{Z}}2^{-k\frac{2\kappa}{1-\kappa}}\left(\sum_{j\leq k}2^{j\frac{\kappa}{2}}2^{j(1-\frac{\kappa}{2})}|\mu_j|^{\frac{1-\kappa}{2}}\right)^{\frac{2}{1-\kappa}}\right)^{\frac{1-\kappa}{2}}\leq$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{k\in\mathbb{Z}}2^{-k\frac{2\kappa}{1-\kappa}}\left(\sum_{j\leq k}2^{j\frac{\kappa}{1+\kappa}}\right)^{\frac{1+\kappa}{1-\kappa}}\sum_{j\leq k}2^{j\frac{2-\kappa}{1-\kappa}}|\mu_j|\right)^{\frac{1-\kappa}{2}}\lesssim$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{k\in\mathbb{Z}}2^{-k\frac{2\kappa}{1-\kappa}}2^{k\frac{\kappa}{1-\kappa}}\sum_{j\leq k}2^{j\frac{2-\kappa}{1-\kappa}}|\mu_j|\right)^{\frac{1-\kappa}{2}}=$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{j\in\mathbb{Z}}2^{j\frac{2-\kappa}{1-\kappa}}|\mu_j|\sum_{k\geq j}2^{-k\frac{\kappa}{1-\kappa}}\right)^{\frac{1-\kappa}{2}}\lesssim$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{j\in\mathbb{Z}}2^{j\frac{2-\kappa}{1-\kappa}}|\mu_j|\cdot 2^{-j\frac{\kappa}{1-\kappa}}\right)^{\frac{1-\kappa}{2}}=$$

$$\left(\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|\right)^{\frac{1+\kappa}{2}}\left(\sum_{j\in\mathbb{Z}}2^{2j}|\mu_j|\right)^{\frac{1-\kappa}{2}}=\sum_{k\in\mathbb{Z}}2^{2k}|\mu_k|.$$

But $|\mu_k|$ is just $\mathcal{E}_w[\mu_k]=\int_{T^d}(\mathbf{I}^*\mu_k)^2w$, hence

$$\sum_{k,j\in\mathbb{Z}}2^{k+j}\int_{T^d}\mathbf{I}^*\mu_k\cdot\mathbf{I}^*\mu_j\cdot w\lesssim\sum_{k\in\mathbb{Z}}2^{2k}\int_{T^d}(\mathbf{I}^*\mu_k)^2w,$$

and we are done.

Remark. The Strong Capacitary Inequality (which essentially goes back to Maz'ya, [67]) usually does not hold – along with the Surrogate Maximum Principle – for non-product weights. Actually the same counterexample for SMP messes up SCI as well.

2.4 From the Carleson condition to the embedding

Here we show two of the converse inequalities in (2.4), namely that

$$[w,\mu]_{CE} \lesssim [w,\mu]_{HC} \lesssim [w,\mu]_C.$$

We start with an auxiliary statement. Given $E \subset T^d$ let

$$\mathcal{E}_{w,E}[\mu] := \int_E (\mathbf{I}^*\mu)^2 w.$$

Lemma 2.4.1 *Let $w: T^d \to \mathbb{R}_+$ be such that SMP (2.12) holds. Let ν be a measure on T^d and*

$$E := \left\{\mathbf{V}_w^\nu > (2C_{(2.13)})^{-\frac{1}{\kappa}}]\frac{\mathcal{E}_w[\nu]}{|\nu|}\right\}] \subset T^d.$$

Then

$$\mathcal{E}_{w,E}[\nu] = \sum_{\alpha\in E}(\mathbf{I}^*\nu)^2(\alpha)w(\alpha) \geq \frac{1}{2}\mathcal{E}_w[\nu].$$

Proof. Put $\delta := (2C_{(2.13)})^{-\frac{1}{\kappa}}\dfrac{\mathcal{E}_w[\nu]}{|\nu|}$. By Lemma 2.2.13 we have

$$\mathcal{E}_{w,E}[\nu] = \mathcal{E}_w[\nu] - \mathcal{E}_{w,\delta}[\nu] \geq \mathcal{E}_w[\nu] - C_{(2.13)}(\delta|\nu|)^\kappa \mathcal{E}_w^{1-\kappa}[\nu] = \frac{1}{2}\mathcal{E}_w[\nu].$$

□

Remark. Essentially this Lemma tells us that if we truncate the energy of ν 'high enough', i.e. on a set of small – compared to the average value of $\mathbf{V}_w^\nu$ – potential, then we are left with only a small portion of the full energy $\mathcal{E}_w[\nu]$.

We are ready to prove the main estimate. The idea here goes roughly as follows. If $[w,\mu]_{HC}$ is way larger than $[w,\mu]_C$ (which we may assume to be 1 say), it would mean that there exists a set E such that μ restricted to this set has extremely large (compared to its mass) energy. But then, if we go 'up' from this set (i.e. consider enough of its immediate ancestors), then the energy of $\mu\mathbb{1}_E$ is mostly supported on this ancestor set F, this is due to Lemma 2.4.1 (and, essentially, due to the tail energy estimate 2.2.13). On the other hand, this set F is not much larger than E in capacitary sense, so in order to keep $[w,\mu]_C$ down, we have to rebalance μ on it, but in doing this we obtain even worse ratio between $[w,\mu]_{HC}$ and $[w,\mu]_C$, and we can repeat the argument. We have the following theorem.

Theorem 2.4.1 *Let $w: T^d \to \mathbb{R}_+$ be such that SMP (2.12) holds. Then for every non-negative measure μ on T^d we have*

$$[w,\mu]_{HC} \lesssim [w,\mu]_C.$$

Proof. Without any loss of generality we may assume that $[w,\mu]_C = 1$. Let

$$A := [w,\mu]_{HC} = \sup_{E\subset T^d,\, \mu(E)\neq 0} \frac{\mathcal{E}_w[\mu\cdot \mathbb{1}_E]}{\mu(E)}. \tag{2.19}$$

Since T^d is finite, the constant A is finite as well, and the maximizer E for (2.19) does exist. Let $\nu := \mu\mathbb{1}_E$ and

$$F := \{\mathbf{V}_w^\nu > cA\}$$

for some absolute constant c. Then by Lemma 2.4.1, if c is small enough (as in previous Lemma), we have

$$\mathcal{E}_{w,F}[\nu] \geq \frac{1}{2}\mathcal{E}_w[\nu].$$

Hence $0 < \mathcal{E}_w[\nu] \leq 2\mathcal{E}_{w,F}[\nu] \leq 2\mathcal{E}_{w,F}[\mu] \leq 2\mu(F)$, in particular $\mu(F)\neq 0$. Since $\mathbf{V}_w^\nu > cA$ on F, we have

$$cA\mu(F) \leq \int_F \mathbf{V}_w^\nu\, d\mu \leq \mathcal{E}_w^{\frac{1}{2}}[\nu]\mathcal{E}_w^{\frac{1}{2}}[\mu\mathbb{1}_F] \leq (2\mu(F))^{\frac{1}{2}}(A\mu(F))^{\frac{1}{2}}.$$

It follows immediately that $A \lesssim 1$. $\square$

2.4.1 From the Hereditary Carleson condition to the embedding

The next inequality to reverse is the one between $[w,\mu]_{HC}$ and the embedding constant $[w,\mu]_{CE}$. The Hereditary Carleson/Restricted Energy condition is not very useful by itself, since it involves testing on arbitrary sets, and not in a good way at that. It serves as a step, however, between more appropriate tests and the boundedness of the original embedding.

Thankfully the inequality $[w,\mu]_{CE} \lesssim [w,\mu]_{HC}$ we can get practically for free, via the results in Section 2.3.

Proposition 2.4.1 *Assume that $w : T^d \to \mathbb{R}_+$ is a weight that satisfies SMP* (2.12)*, and μ is a non-negative measure on T^d. Then for any $E\subset T^d$ one has*

$$\mu(E) \leq [w,\mu]_{HC}\,\mathrm{Cap}_w(E).$$

Proof. Fix any set $E\subset T^d$, and let f be an admissible function for E, that is $\mathbf{I}_w f \geq 1$ on E, and let $A = [w,\mu]_{HC}$. Clearly, the energy is positive, so, expanding, we have

$$\begin{aligned}0 \leq \int_{T^d} (\mathbf{I}^*\mu|_E - Af)^2 w = \sum_{\alpha\in T^d} (\mathbf{I}^*\mu|_E)^2(\alpha)w(\alpha) - \\ - 2A\sum_{\alpha\in T^d} \mathbf{I}^*\mu|_E(\alpha)\cdot f(\alpha)w(\alpha) + A^2\sum_{\alpha\in T^d} f^2(\alpha)w(\alpha).\end{aligned}$$

which means, in turn,

$$
\begin{aligned}
0 \leq & \left(\sum_{\alpha \in T^d} (\mathbf{I}^* \mu|_E)^2(\alpha) w(\alpha) - A \sum_{\alpha \in T^d} \mathbf{I}^* \mu|_E(\alpha) \cdot f(\alpha) w(\alpha) \right) - \\
& - A \left(\sum_{\alpha \in T^d} \mathbf{I}^* \mu|_E(\alpha) \cdot f(\alpha) w(\alpha) - A \sum_{\alpha \in T^d} f^2(\alpha) w(\alpha) \right) = \\
& \left(\sum_{\alpha \in T^d} (\mathbf{I}^* \mu|_E)^2(\alpha) w(\alpha) - A \int_{T^d} \mathbf{I}_w f \, d\mu|_E \right) - \\
& - A \left(\int_{T^d} \mathbf{I}_w f \, d\mu|_E - A \sum_{\alpha \in T^d} f^2(\alpha) w(\alpha) \right).
\end{aligned} \tag{2.20}
$$

Now, since f is admissible for E and $\mu|_E$ is supported on E, we see that $\int_{T^d} \mathbf{I}_w f \, d\mu|_E \geq |\mu|_E|$. By definition of $[w, \mu]_{HC} = A$ it follows that

$$
\sum_{\alpha \in T^d} (\mathbf{I}^* \mu|_E)^2(\alpha) w(\alpha) - A \int_{T^d} \mathbf{I}_w f \, d\mu|_E \leq 0.
$$

However than the second term on the right-hand side (2.20) must be non-negative,

$$
|\mu|_E| \leq A \sum_{\alpha \in T^d} f^2(\alpha) w(\alpha).
$$

Minimizing over all admissible functions f we arrive at

$$
|\mu|_E| \leq A \, \mathrm{Cap}_w(E),
$$

and we are done. $\square$

It remains to see, that we have already proven in Section 2.3 that subcapacitary measures realize the embedding, $[w, \mu]_{CE} \lesssim [w, \mu]_{SC}$, hence it follows immediately that

$$
[w, \mu]_{CE} \lesssim [w, \mu]_{HC}.
$$

2.5 Single box test

We are left with the last of the converse inequalities (2.4), the one that tells it is enough to test the embedding on the successor sets of singletons. The general idea of the proof is somewhat similar to the idea behind $[w, \mu]_{CE} \lesssim [w, \mu]_C$, however since we now only have estimates on boxes and not on arbitrary sets, we must deal with additional complications. Also, we would like to mention again the unlikely nature of such a single box test which seems to go against the usual course of events in multi-parametric settings.

2.5.1 Main estimate

We start with some additional notation. Within this section we always assume that a weight w on T^d is already fixed, so we will drop it from subscripts. Given a non-negative measure ν on T^d define

$$\mathbf{V}^{\nu}_{\tau}(\omega) := \textstyle\sum_{\omega\le\beta\le\tau} \mathbf{I}^*\nu(\beta)w(\beta) \tag{2.21a}$$

$$\mathbf{V}_{\varepsilon',\,good}(\omega) := \textstyle\sum_{\beta\ge\omega:\ \mathbf{V}^{\nu}_{\omega}>\varepsilon'} \mathbf{I}^*\nu(\beta)w(\beta). \tag{2.21b}$$

Next Lemma employs a staircase-like construction inspired by [86] to estimate the size of truncated 'good' potentials.

Lemma 2.5.1 *Let $d\ge 2$ and μ be a non-negative measure on T^d. Let $w: T^d\to\mathbb{R}_+$ be a weight that satisfies SMP* (2.12)*. Assume that $\mathcal{E}[\mu]\le|\mu|$ and*

$$\mathbf{V}^{\mu} \ge \frac{1}{3} \quad \text{on } \operatorname{supp}\mu. \tag{2.22}$$

Then, if ε' is small enough, we have

$$\int_{T^d} \mathbf{V}^{\mu}_{\varepsilon',\,good}\,d\mu \gtrsim |\mu|.$$

Proof. It suffices to show that for some ε' and ε_{d-1} we have

$$\mu\left\{\omega\in T^d:\ \mathbf{V}^{\mu}_{\varepsilon',\,good}(\omega)\ge\varepsilon_{d-1}\right\}\ge\frac{|\mu|}{2}.$$

Let $\varepsilon>0$ be chosen later, and define

$$\varepsilon_1 := \varepsilon,\ \ \varepsilon_2 := \varepsilon\cdot\varepsilon_1^{\frac{1}{\kappa}},\ \ \varepsilon_3 := \varepsilon\cdot\varepsilon_2^{\frac{1}{\kappa}},\dots.$$

By Lemma 2.2.13 we have

$$\int_{\mathbb{T}^d} \mathbf{V}^{\mu}_{\varepsilon_j}\,d\mu \lesssim \varepsilon_j^{\kappa}|\mu|^{\kappa}\mathcal{E}^{1-\kappa}[\mu] \lesssim \varepsilon_j^{\kappa}\int_{T^d} d\mu$$

for some $\kappa>0$. We recall that $\mathbf{V}^{\mu}_{\varepsilon_j}(\omega) = \sum\limits_{\omega\le\alpha\in T^d:\ \mathbf{V}^{\mu}\le\varepsilon_j} \mathbf{I}^*\mu(\alpha)w(\alpha)$. By Chebyshev's inequality it follows that

$$\mathbf{V}^{\mu}_{\varepsilon_j}(\omega) \le \frac{1}{10}\left(\frac{\varepsilon_j}{\varepsilon}\right)^{\kappa} \tag{2.23}$$

on at least $(1-C\varepsilon^{\kappa})$-portion (w.r.t. μ) of points ω. With that in mind we now only consider such points ω that (2.23) holds for every $j=1,2,\dots,d-1$, and also such that $\mathbf{V}^{\mu}(\omega)\lesssim 1$. Let

$$\varepsilon' := \varepsilon\cdot\varepsilon_1\cdot\dots\cdot\varepsilon_{d-1}.$$

Given a point ω let

$$\mathcal{U} := \{\tau \geq \omega :\ \mathbf{V}^{\mu}_{\tau}(\omega) > \varepsilon'\} \tag{2.24}$$

and

$$\mathcal{W}_j := \{\tau \geq \omega :\ \mathbf{V}^{\mu}(\tau) \leq \varepsilon_j\}, \quad 1 \leq j \leq d-1. \tag{2.25}$$

We again recall the definitions of successor and predecessor sets, $\mathcal{P}(\alpha) = \{\beta \in T^d :\ \beta \geq \alpha\}$ and $\mathcal{S}(\beta) = \{\alpha \in T^d :\ \alpha \leq \beta\}$.

Now if $\mathcal{U} \not\subset \mathcal{W}_{d-1}$, then there exists a point $\gamma \notin \mathcal{W}_{d-1}$ such that $\mathcal{P}(\gamma) \subset \mathcal{U}$. Hence

$$\mathbf{V}^{\mu}_{\varepsilon', good}(\omega) \geq \sum_{\gamma' \geq \gamma} \mathbf{I}^*\mu(\gamma')w(\gamma') = \mathbf{V}^{\mu}(\gamma) \geq \varepsilon_{d-1},$$

and in this case the potential at ω is large enough. Assume next, that $\mathcal{U} \subseteq \mathcal{W}_{d-1}$. We are going to cover $\mathcal{P}(\omega) \setminus \mathcal{W}_1$ by boundedly many sets of the form $\mathcal{S}(\beta)$ with $\beta \in \mathcal{P}(\omega) \setminus \mathcal{U}$. This will lead to a contradiction with (2.22), since by (2.23) and (2.24) the integral of $f := \mathbf{I}^*\mu \cdot w$ is small on $\mathcal{W}_1$ and also on each such a set $\mathcal{S}(\beta)$. Let us elaborate.

For a set of coordinates $J \subset \{1, 2, \ldots, d\}$ and a point $\gamma \in T^d$ let

$$\mathcal{P}_J(\gamma) := \{\beta \in T^d :\ \beta_j \geq \gamma_j,\ j \in J,\ \beta_j = \gamma_j,\ j \notin J\}.$$

Given a non-empty $J \subset \{1, 2, \ldots, d\}$ and $\gamma \in T^d$ we define a set $\mathcal{Q}_J(\gamma) \subset T^d$ as follows. If $|J| = 1$, then $\mathcal{Q}_J(\gamma)$ consists of the unique maximal element of $\mathcal{P}_J(\gamma) \setminus \mathcal{U}$, if the latter set is non-empty, and is defined to be empty otherwise. If $|J| \geq 2$, then $\mathcal{Q}_J(\gamma)$ is a maximal set of maximal elements of $\mathcal{P}_J(\gamma) \setminus \mathcal{W}_{d-|J|+1}$ such the sets $\mathcal{P}_J(\beta) \setminus \mathcal{W}_{d-|J|+2}$ are pairwise disjoint for $\beta \in \mathcal{Q}_J(\gamma)$.

Yet another set we need is $\mathcal{R}_J$ which we define recursively, $\mathcal{R}_{\emptyset}(\gamma) := \{\gamma\}$, and

$$\mathcal{R}_J(\gamma) := \bigcup_{J' \subset J} \bigcup_{\gamma' \in \mathcal{Q}_J(\gamma)} \mathcal{R}_{J'}(\gamma),$$

where we run over all sets $J' \subset J$ with cardinality $|J'| = |J| - 1$.

We claim that for every $\gamma \in \mathcal{P}(\omega)$ and every $J \subset \{1, 2, \ldots, n\}$ with $J \neq \emptyset$ we have

$$\bigcup_{\gamma' \in \mathcal{R}_J(\gamma)} \mathcal{S}(\gamma') \supseteq \mathcal{P}_J(\gamma) \setminus \mathcal{W}_{d-|J|+1}, \tag{2.26}$$

where we set $\mathcal{W}_d := \mathcal{U}$ to close the notation. We prove (2.26) by induction on size of J. For $|J| = 1$ the claim is obvious. Now, let J, $|J| \geq 2$, be given and suppose we have (2.26) for all proper subsets of J. Let

$$\mathcal{F} := \bigcup_{\gamma' \in \mathcal{R}_J} \mathcal{S}(\gamma'), \quad \mathcal{G} := \mathcal{P}_J(\gamma) \setminus \mathcal{W}_{d-|J|+1}.$$

By hypothesis

$$\mathcal{F} \supseteq \mathcal{P}_{J'}(\gamma') \setminus \mathcal{W}_{d-|J|+2} \tag{2.27}$$

for every $\gamma' \in \mathcal{R}_J(\gamma)$ and every $J' \subsetneq J$. Suppose that

$$\mathcal{F} \not\supseteq \mathcal{G}. \tag{2.28}$$

Chose a maximal element $\beta \in \mathcal{F} \setminus \mathcal{G}$. Since $\mathcal{F}$ is a down-set (so it contains all successors of its elements), we see that β is also a maximal element of $\mathcal{G}$. We claim

$$\left(\mathcal{P}_J(\beta) \bigcap \mathcal{P}_J(\gamma)\right) \setminus \mathcal{W}_{d-|J|+2} = \emptyset \quad \text{fro all } \gamma' \in \mathcal{Q}(\gamma). \tag{2.29}$$

Indeed, suppose for a contradiction that there exists $\beta' \in \left(\mathcal{P}_J(\beta) \bigcap \mathcal{P}_J(\gamma)\right) \setminus \mathcal{W}_{d-|J|+2}$ and let β' be minimal with such a property. Since $\mathcal{W}_{d-|J|+2}$ is an up-set, β' is also a minimal element of $\mathcal{P}_J(\beta) \bigcap \mathcal{P}_J(\gamma)$. Also, since $\beta, \gamma' \in \mathcal{P}_J(\gamma)$, we have that β' is in fact the coordinate-wise maximum of β, γ'. Finally, since β and γ' are two distinct maximal elements de $\mathcal{P}$, we deduce that β' coincides with γ' in at least one coordinate, so $\beta' \in \mathcal{P}_{J'}\gamma'$ for some $J' \subsetneq H$. Now (2.27) implies that $\beta' \in \mathcal{F}$ which is a down-set, but $\gamma' \geq \gamma$, therefore $\gamma \in \mathcal{F}$, and we have a contradiction.

It follows that (2.29) holds true. But this, in turn, contradicts the maximality of $\mathcal{Q}_J(\gamma)$. Thus the assumption (2.28) is false, and we finally get (2.26). Fix a point $\gamma \geq \omega$. For $2 \leq |J| \leq d$ we have

$$1 \gtrsim \mathbf{V}^{\mu}(\omega) \geq \mathbf{V}^{\mu}(\gamma) \geq \sum_{\beta \in \mathcal{Q}_J(\gamma)} \int_{\mathcal{P}_J(\beta) \setminus \mathcal{W}_{d-|J|+2}} fw$$
$$\geq \sum_{\beta \in \mathcal{Q}_J(\gamma)} \left(\mathbf{I}f(\beta) - \mathbf{I}(f \mathbb{1}_{\mathcal{W}_{d-|J|+2}})(\omega)\right)$$

by definition of $\beta \in \mathcal{W}_{d-|J|+1}$ and by (2.23)

$$\sum_{\beta \in \mathcal{Q}_J(\gamma)} \left(\varepsilon_{d-|J|+1} - \frac{\left(\frac{\varepsilon_{d-|J|+2}}{\varepsilon}\right)^{\frac{1}{2}}}{10}\right) \gtrsim |\mathcal{Q}_J(\gamma)| \varepsilon_{d-|J|+1}.$$

It follows that

$$\varepsilon_1 \cdot \dots \cdot \varepsilon_{d-1} |\mathcal{R}_{1,2,\dots,d}(\omega)| \lesssim 1.$$

Hence by (2.26)

$$\begin{aligned}\mathbf{V}^{\mu}(\omega) - \mathbf{V}^{\mu}_{\varepsilon} = \int_{\mathcal{P}(\omega) \setminus \mathcal{W}_1} fw \leq \sum_{\gamma' \in \mathcal{R}_{1,2,\dots,d}(\omega)} \int_{\mathcal{S}(\gamma')} fw = \\ \sum_{\gamma' \in \mathcal{R}_{1,2,\dots,d}(\omega)} \mathbf{V}^{\mu}_{\gamma'}(\omega) \leq \\ \varepsilon' |\mathcal{R}_{1,2,\dots,d}(\omega)| \lesssim \\ \frac{\varepsilon'}{\varepsilon_1 \cdot \dots \varepsilon_{d-1}} = \varepsilon.\end{aligned}$$

By the power of (2.23)

$$\frac{1}{3} \le \mathbf{V}^{\mu}(\omega) = (\mathbf{V}^{\mu}(\omega) - \mathbf{V}^{\mu}_{\varepsilon}(\omega)) + \mathbf{V}^{\mu}_{\varepsilon}(\omega) \le C\varepsilon + \frac{1}{10}.$$

With ε taken to be small enough the inequality above clearly becomes false, and we get the contradiction to our assumption that $\mathcal{U} \subset \mathcal{W}_{d-1}$. $\square$

2.5.2 Single box implies Carleson

Before we proceed we need another auxiliary result (of rather general nature).

Lemma 2.5.2 (Balancing Lemma) *Let w be a weight (of any kind) and ν be a measure on T^d, such that*

$$\mathcal{E}[\nu] \ge A|\nu|,$$

for some $A > 0$. Then there exists a down-set $\tilde{E} \subset T^d$ such that the measure $\tilde{\nu} := \nu \cdot \mathbb{1}_{\tilde{E}}$ satisfies

$$\mathbf{V}^{\tilde{\nu}} \ge \frac{A}{3} \quad \text{on } \tilde{E},$$
$$\mathcal{E}[\tilde{\nu}] \ge \frac{1}{3}\mathcal{E}[\nu].$$

Proof. The argument is straightforward – we consecutively throw away parts of ν with small potential and then check that we are left with enough energy. To elaborate, let us replace ν by $\frac{3\nu}{A}$, so that we may assume that $A = 3$. Let $E_0 = T^d$ and $\nu_0 := \nu\mathbb{1}_{E_0}$. We proceed by induction

$$E_{k+1} := E_k \setminus \{\mathbf{V}^{\nu_k} \le 1\} \quad \nu_{k+1} := \nu \cdot \mathbb{1}_{E_{k+1}}.$$

The sequence $\{E_k\}$ consists of down-sets and is decreasing, and since T^d is finite, it must stabilize at some point, $E_m = E_{m+1},\ m \ge M$ for some number M. Let $\tilde{\nu} := \nu_M$ and $\tilde{E} := E_M$. By construction we already have the first estimate, since

$$\mathbf{V}^{\tilde{\nu}} \ge 1 = \frac{A}{3} \quad \text{on } \tilde{E}.$$

To estimate what remains of energy we put $\sigma_k := \nu_k - \nu_{k+1} = \nu_k \cdot \mathbb{1}_{E_k \setminus E_{k+1}}$. Then we have

$$\begin{aligned}\mathcal{E}[\nu] = \int_{T^d} \mathbf{V}^{\nu_0}\, d\nu_0 = \int_{T^d} \mathbf{V}^{\nu_1}\, d\nu_1 + \int_{T^d} \mathbf{V}^{\nu_0}\, d\sigma_0 + \int_{T^d} \mathbf{V}^{\sigma_0}\, d\nu_0 \le \\ \int_{T^d} \mathbf{V}^{\nu_1}\, d\nu_1 + 2\int_{T^d} \mathbf{V}^{\nu_0}\, d\sigma_0 \le \mathcal{E}[\nu_1] + 2|\sigma_0| \le \\ \cdots \le \mathcal{E}[\nu_M] + 2|\sigma_0| + 2|\sigma_1| + \cdots + 2|\sigma_M| \le \mathcal{E}[\tilde{\nu}] + 2|\nu|,\end{aligned}$$

since $\mathbf{V}^{\nu_k} \le 1$ on $\operatorname{supp} \sigma_k$. But $\mathcal{E}[\nu] \ge 3|\nu|$ by assumption, so we arrive at

$$\mathcal{E}[\tilde{\nu}] \ge \frac{1}{3}\mathcal{E}[\nu],$$

and we are done. $\square$

We are ready to attack the last remaining inequality from (2.4).

Theorem 2.5.1 *Let $d \geq 2$ and $w : T^d \to \mathbb{R}_+$ be such that the SMP (2.12) holds. Then for any non-negative measure ν on T^d one has*

$$[w,\nu]_{HC} \lesssim [w,\nu]_B.$$

Proof. By rescaling we may assume that $[w,\nu]_B = 1$. Let $A := [w,\mu]_{HC}$, and let $E \subset T^d$ be a set such that $\mu := \nu \mathbb{1}_E \neq 0$ and $\mathcal{E}[\mu] = A|\mu|$ (such a subset does exist, since, say, T^d is finite). By Lemma 2.5.2 there exists a further refinement $\tilde{E}$ of E such that $\tilde{\mu} = \mu \mathbb{1}_{\tilde{E}}$ satisfies

$$\mathbf{V}^{\tilde{\mu}} \geq \frac{A}{3} \quad \text{on } \tilde{E},$$

and $\tilde{\mu} \neq 0$. Thus, replacing μ by $\tilde{\mu}$ we may assume that $\mathbf{V}^{\mu} \geq \dfrac{A}{3}$ on $\operatorname{supp}\mu$.

By Lemma 2.5.1 applied with $\frac{\mu}{A}$ in place of μ for sufficiently small $\varepsilon, \theta > 0$ we have

$$\int_{T^d} \mathbf{V}^{\mu}_{\varepsilon A,\, good}\, d\mu \geq 2\theta \mathcal{E}[\mu].$$

We claim that with these values of ε and θ we actually have

$$\mathcal{E}[\mu] \leq \frac{\theta}{1-\theta} \sum_{\alpha:\, \theta\varepsilon A \mathbb{I}^*\mu(\alpha) \leq \mathcal{E}_\alpha[\mu]} (\mathbf{I}^*\mu(\alpha))^2 w(\alpha). \tag{2.30}$$

Indeed, suppose that α is such that

$$\theta\varepsilon A \mathbb{I}^*\mu(\alpha) > \mathcal{E}_\alpha[\mu] = \sum_{\omega \leq \alpha} \mathbf{V}^{\mu}_{\alpha}(\omega)\mu(\omega),$$

where $\mathbf{V}^{\mu}_{\alpha}(\omega) = \displaystyle\sum_{\beta:\, \omega \leq \beta \leq \alpha} \mathbf{I}^*\mu(\beta) \cdot w(\beta)$. Then we have

$$\begin{aligned} \sum_{\omega \leq \alpha:\, \mathbf{V}^{\mu}_{\alpha}(\omega) \leq \varepsilon A} \mu(\omega) = \mathbf{I}^*\mu(\alpha) - \sum_{\omega \leq \alpha:\, \mathbf{V}^{\mu}_{\alpha}(\omega) > \varepsilon A} \mu(\omega) \geq \\ \mathbf{I}^*\mu(\alpha) - \frac{1}{\varepsilon A} \sum_{\omega \leq \alpha} \mathbf{V}^{\mu}_{\alpha}(\omega)\mu(\omega) \geq \\ (1-\theta)\mathbf{I}^*\mu(\alpha). \end{aligned}$$

It follows that

$$\sum_{\alpha:\,\theta\varepsilon A\mathbf{I}^*\mu(\alpha)>\mathcal{E}_\alpha[\mu]}(\mathbf{I}^*\mu(\alpha))^2w(\alpha)\le\sum_\alpha w(\alpha)\mathbf{I}^*\mu(\alpha)\frac{1}{1-\theta}\sum_{\omega\le\alpha:\,\mathbf{V}^\mu_\alpha(\omega)\le\varepsilon A}\mu(\omega)=$$
$$\frac{1}{1-\theta}\sum_\omega\mu(\omega)\sum_{\alpha\ge\omega:\,\mathbf{V}^\mu_\alpha(\omega)\le\varepsilon A}w(\alpha)\mathbf{I}^*\mu(\alpha)=$$
$$\frac{1}{1-\theta}\sum_\omega\mu(\omega)\left(\mathbf{V}^\mu(\omega)-\mathbf{V}^\mu_{\varepsilon A,\,good}(\omega)\right)\le$$
$$\frac{1-2\theta}{1-\theta}\mathcal{E}[\mu].$$

Now we have the claim (2.30).

Next, by Lemma 2.2.13 and the fact that $\mathbf{V}^\mu\ge\dfrac{A}{4}$ on $\operatorname{supp}\mu$ we have

$$\mathcal{E}_{c'A}[\mu]\lesssim(c'A)^\kappa|\mu|^\kappa\mathcal{E}^{1-\kappa}[\mu]\lesssim(c')^\kappa\mathcal{E}[\mu]. \tag{2.31}$$

Taking c' to be sufficiently small and combining (2.30) with (2.31) we obtain

$$\mathcal{E}[\mu]\lesssim\sum_{\alpha\in\mathcal{R}}(\mathbf{I}^*\mu(\alpha))^2w(\alpha),\quad\text{where}$$
$$\mathcal{R}:=\left\{\alpha\in T^d:\ \theta\varepsilon A\mathbf{I}^*\mu(\alpha)\le\mathcal{E}_\alpha[\mu],\ \mathbf{V}^\mu(\alpha)\ge c'A\right\}.$$

For each $\alpha\in\mathcal{R}$ we have

$$\theta\varepsilon A\mathbf{I}^*\mu(\alpha)\le\mathcal{E}_\alpha[\mu]\le\mathcal{E}_\alpha[\nu]\le[w,\nu]_B\mathbf{I}^*\nu(\alpha)=\mathbf{I}^*\sigma(\alpha),$$

where $\sigma L=\nu\mathbb{1}_F$ and $F:=\{\beta\in T^d:\ \exists\alpha\in\mathcal{R},\ \alpha\le\beta\}$. It follows that

$$A^2\mathcal{E}[\mu]\lesssim\mathcal{E}[\sigma]. \tag{2.32}$$

On the other hand, by definition of A, the fact that $\mathbf{V}^\mu\gtrsim A$ on $\operatorname{supp}\sigma$ and Cauchy-Schwartz, we obtain

$$\mathcal{E}[\sigma]\le A|\sigma|\lesssim\int_{T^d}\mathbf{V}^\mu\,d\sigma\le\mathcal{E}^{\frac12}[\mu]\mathcal{E}^{\frac12}[\sigma]. \tag{2.33}$$

From (2.33) we obtain $\mathcal{E}[\sigma]\lesssim\mathcal{E}[\mu]$, and inserting this into (2.32), we see that $A\lesssim1$.

We are done. $\square$

2.6 Comments, examples and counterexamples

In this Section, which contains results from [106] and [110], we explore a number of results that shed some light on the behaviour of weighted potentials. First we justify the fact that we were working with finite trees in this Chapter. Next, we show that if the weight w *does not have a product structure*, then none of the statements of Theorem 2.1.1 holds true anymore, even if w takes only values 0 or 1. We will do this by constructing N-coarse measures μ and weights w

on finite bi-trees T^2 of depth N such that the discrepancies between box, Carleson, REC, and embedding constants grow with N. These counterexamples are intimately related to those in [86], basically they are their discrete analogues. Also we prove a more general version of small energy majorization statement on a 1-tree, and show that its 2-tree version is false.

2.6.1 From finite d-tree T_N^d to an infinite d-tree $\overline{T}^d$

We have proven Theorem 2.1.1 for any finite d-tree T_N^d (or, equivalently, for pairs (w, μ) restricted to T_N^d), in particular there is no dependence on N in (2.4). We claim that it implies the embedding Theorem on the full d-tree.
Indeed, we have proven Theorem 2.1.1 for any finite d-tree T_N^d and the constants that govern the relations between $[w, \mu]_{CE}, [w, \mu]_{HC}, [w, \mu]_C, [w, \mu]_B$ and $[w, \mu]_{SC}$ do not depend on the depth N. Now let us fix a measure μ on the full d-tree $\overline{T}^d$ and a product weight w on its interior T^d. We claim that the reverse inequalities (2.4) hold true.
In order to show this we fix any $N \in \mathbb{N}$ and consider truncations of w and μ to T_N^d. For $\alpha = (\alpha_1, \ldots, \alpha_d) \in T^d$ define $|\alpha| := \max_{1\le k\le d} |\alpha_k|$ with $|\alpha_k| := \#\mathcal{P}(\alpha_k) - 1$ being the depth of α_k in the respective coordinate tree. The truncation of the weight is straightforward

$$w_N := w \cdot \mathbb{1}_{T_N^d}.$$

The measure μ has to be moved to the truncated tree a bit more carefully

$$\mu_N := \mu \cdot \mathbb{1}_{T_N^d} + \sum_{\alpha:|\alpha|=N+1} \mathbb{1}_\alpha \cdot \mathbf{I}^*\mu(\alpha).$$

In other words, we leave the portion of μ that lies on T_N^d as it is, but also 'rise up' the rest of μ to the lower-most level of T_N^d (which consists of points of depth N – exactly those with $|\alpha| = N+1$) by taking $\mathbf{I}^*\mu$.
It is easily seen that

$$\mathbf{I}^*\mu_N(\alpha) \le \mathbf{I}^*\mu(\alpha), \quad \alpha \in T^d. \tag{2.34}$$

Indeed, for any two distinct points $\alpha, \beta \in (\partial T_N)^d$ (i.e. for $|\alpha| = |\beta| = N+1$) the successor sets of α and β *do not intersect*, $\mathcal{S}(\alpha) \cap \mathcal{S}(\beta) = \emptyset$. Hence

$$\int_{\mathcal{S}(\gamma)} \mathbb{1}_{|\alpha|\ge N} d\mu = \sum_{\alpha\le\gamma,\, |\alpha|=N+1} \int_{\mathcal{S}(\alpha)} d\mu,$$

so for $\gamma \in T_N^d$ we have equality in (2.34), and for other γ's the left-hand side is zero (we do not care about these points here anyway).
Similarly, for any collection of singleton successors

$$E := \bigcup_j \mathcal{S}(\alpha^j), \quad \alpha^j \in T^d, \tag{2.35}$$

we have

$$\mu_N(E) = \mu\left(E \cap T_N^d\right)$$

for the same reason. We deduce that if the pair (w, μ) satisfies any of the test conditions (2.3a) – (2.3d), be it Heredetary Carleson, Carleson, Box or Subcapacitary on $\overline{T}^d$, then the restricted pairs (w_N, μ_N) do the same on T_N^d with same constants.
Next we observe that for any $f \in L^2(T^d, dw)$ we have by dominated convergence

$$\lim_{N\to\infty} \int_{T^d} f^2(\tau)\, dw_N(\tau) = \lim_{N\to\infty} \int_{T^d} (f \cdot \mathbb{1}_{T_N^d})^2(\tau)\, dw(\tau) = \int_{T^d} f^2(\tau)\, dw(\tau).$$

Finally,

$$\lim_{N\to\infty} \int_{T^d} (\mathbf{I}_{w_N} f)^2\, d\mu_N = \lim_{N\to\infty} \int_{T^d} (\mathbf{I}_w f)^2\, d\mu_N = \int_{\overline{T}^d} (\mathbf{I}_w f)^2\, d\mu,$$

since $\mathbf{I}_w f$ is monotone in the d-tree order and I_w-potentials converge pointwise to boundary values

$$\lim_{|\alpha|\to\infty,\, \alpha\geq\omega} \mathbf{I}_w f(\alpha) = \mathbf{I}_w f(\omega), \quad \omega \in \partial T^d.$$

This finishes the reduction argument.

2.6.2 General setting for counterexamples

Within this Section we assume $d = 2$ and $T^2 = T_N^2$ be a finite (but very deep) dyadic bi-tree – since we are constructing counterexamples we are considering the simplest possible situation. Also, here we will make use of the rectangle interpretation – we identify T^2 with the set of all dyadic rectangles in the square $Q_0 = [0,1]^2$ with side lengths at least 2^{-N} ordered by inclusion.

We denote the set of minimal elements of this bi-tree, that is, the small squares of size $2^{-N} \times 2^{-N}$, by $(\partial T)^2$, and elements of this set will be denoted by ω. We denote sets of ω's by $E \subset (\partial T)^2$ and identify them with their union, so we will write $Q \subset E$ if Q is covered by the elements of E.

In all examples the measure μ will be supported on the square Q_0. We identify it with a function on T^2 by setting $\tilde{\mu}(\omega) := \mu(\omega)$ for $\omega \in (\partial T)^2$ and $\tilde{\mu}(Q) := 0$ for $Q \notin (\partial T)^2$. Then $\mathbf{I}^*\tilde{\mu}(Q) = \mu(Q)$. With this convention the box condition (2.3d) for the measure μ and weight $w = \{w_Q\}$ becomes

$$\sum_{Q\in T^2,\, Q\subset R} \mu^2(Q) w_Q \leq C\mu(R), \quad \text{for any } R \in T^2. \tag{2.36}$$

The Carleson condition (2.3c) becomes

$$\sum_{Q\in T^2,\, Q\subset E} \mu^2(Q) w_Q \leq C\mu(E), \quad \text{for any } E \subset (\partial T)^2, \tag{2.37}$$

the hereditary Carleson (or Restricted Energy) condition (2.3b) becomes

$$\sum_{Q\in T^2} \mu^2(Q\cap E)w_Q \le C\mu(E), \quad \text{for any } E\subset(\partial T)^2, \tag{2.38}$$

and the (dual version of) Carleson embedding (2.2) becomes

$$\sum_{Q\in T^2}\left(\int_Q \varphi\, \mathrm{d}\mu\right)^2 w_Q \le C\int_{Q_0}\varphi^2 d\mu \quad \text{for any } \varphi\in L^2(Q_0,d\mu). \tag{2.39}$$

2.6.3 Box condition does not imply Carleson condition

In [21] Carleson constructed families $\mathcal{R}$ of dyadic sub-rectangles of $Q=[0,1]^2$ having the following two properties:

$$\forall R_0\in T^2, \qquad \sum_{R\subset R_0, R\in\mathcal{R}} m_2(R)\le C_0 m_2(R_0)\,, \tag{2.40}$$

but

$$\sum_{R\in\mathcal{R}} m_2(R) > C_1 m_2(\cup_{R\in\mathcal{R}}R)\,, \tag{2.41}$$

with arbitrarily large ratios C_1/C_0, where m_2 is the planar Lebesgue measure. Choosing $\mu=m_2$ and

$$w_R := \begin{cases} \frac{1}{m_2(R)}, & R\in\mathcal{R},\\ 0, & \text{otherwise}\end{cases}$$

we can identify the left-hand sides of (2.40) and (2.41) with the left-hand sides of (2.36) and (2.37), respectively. Hence the box condition (2.36) holds with constant C_0, while the Carleson condition (2.37) can only hold with constant $\ge C_1$.

The weight w is rather wild here. But there is also a counterexample with $w_R\in\{0,1\}$ for all R.

2.6.4 Carleson condition does not imply REC

Our aim here is to show that for general w,μ the Carleson condition (2.37) is no longer sufficient for the embedding (2.39) or even the Restricted Energy Condition (2.38). Namely we prove the following statement.

Proposition 2.6.1 *For any $\delta>0$ there exists a number $N\in\mathbb{N}$, a weight $w: T_N^2\to\{0,1\}$, and a measure μ on ∂T^2 such that μ satisfies the Carleson condition* (2.37) *with the constant $C_\mu=\delta$:*

$$\sum_{Q\subset E}\mu^2(Q)w_Q\le\delta\mu(E), \quad \textit{for any } E\subset(\partial T)^2, \tag{2.42}$$

but there exists a set $F\subset Q_0$ such that

$$\sum_{Q\in\mathcal{D}}\mu^2(Q\cap F)w_Q > \mu(F), \tag{2.43}$$

hence the constant in (2.38) *is at least* 1.

We intend to give two examples of this kind. The first example is quite simple and is inspired by the counterexample for L^2-boundedness of the biparameter maximal function. In this example the weight w is supported on a very small subset of the bi-tree, which differs greatly from the original graph. The second example is somewhat more involved, but the weight w is supported on a much larger portion of the bi-tree; in fact it has the monotonicity property $w_R \geq w_Q$ for $R \supseteq Q$. Nevertheless there are not enough rectangles in the support of w to have the Carleson-REC equivalence.

We introduce some additional notation. We denote by $\omega_0 := [0, 2^{-N}]^2$ the left lower corner of the unit square. Given a dyadic rectangle $R = [a, b] \times [c, d]$ let

$$\begin{aligned} R^{+\circ} &:= [(a+b)/2, b] \times [c, d], \\ R^{\circ+} &:= [a, b] \times [(c+d)/2, d], \\ R^{++} &:= [(a+b)/2, b] \times [(c+d)/2, d], \end{aligned}$$

be its right half, upper half, and upper right quadrant, respectively. Again we fix the weight w at the beginning of our arguments, so we will drop it from the subscript in $\mathbf{V}^{\mu}_{w}$, $\mathbf{I}_w$ etc.

A simple example

Let $Q_i = [0, 2^{-i+1}] \times [0, 2^{-N+i}]$ for $j = i, \ldots, N$. Let measure μ have mass 1 on ω_0 and each of Q_i^{++}, and mass 0 everywhere else. Let

$$w_R := \begin{cases} 1 & \text{if } R \in \{\omega_0, Q_1, Q_2, \ldots, Q_N\}, \\ 0 & \text{else.} \end{cases}$$

So we have $N+1$ weights w_R equal to 1. For the set $E = \omega_0$ we have

$$\mathcal{E}[\mu|E] = \mu(\omega_0)^2 + \sum_{i=1}^{N} \mu(\omega_0 \cap Q_i)^2 = (N+1) = (N+1)\mu(E).$$

So the REC constant (2.38) is $\geq N+1$.

Denoting $Q_0 := \omega_0$, for an arbitrary $E \subseteq \partial T^2$ we have

$$\mathcal{E}_E[\mu] = \sum_{R \subset E,\, w_R \neq 0} \mu(R)^2 = \sum_{j: Q_j \subset E} \mu(Q_j)^2.$$

Then since $Q_i^{++} \cap Q_j = \emptyset$ unless $i \in \{0, j\}$, we have

$$\mathcal{E}_E[\mu] \leq \sum_{j: Q_j \subset E} 2^2 \leq 4\mu(E)$$

So the Carleson condition (2.37) holds with constant 4.

The lack of maximal principle matters

Now we construct a more complicated example in which the Carleson condition holds, but the restricted energy condition fails. The weight w still has values either 0 or 1, but the support $\mathcal{R}$ of w is an *up-set*, that is, it contains every ancestor of every rectangle in $\mathcal{R}$.

The example is based on the fact that potentials on bi-tree may not satisfy the maximal principle. So we start with constructing an N-coarse μ such that we have

$$\mathbf{V}^\mu \lesssim 1 \quad \text{on } \operatorname{supp}\mu, \tag{2.44}$$

but

$$\max \mathbf{V}^\mu \geq \mathbf{V}^\mu(\omega_0) \gtrsim \log N. \tag{2.45}$$

We define a collection of rectangles

$$Q_j := [0, 2^{-2^j}] \times [0, 2^{-2^{-j}N}], \quad j = 1, \dots, M \approx \log N. \tag{2.46}$$

Now we put

$$\begin{aligned} \mathcal{R} &:= \{R :\ Q_j \subset R \text{ for some } j = 1 \dots M\} \\ w_Q &:= \mathbb{1}_{\mathcal{R}}(Q) \\ \mu(\omega) &:= \frac{1}{N}\sum_{j=1}^{M} \frac{1}{|Q_j^{++}|}\mathbb{1}_{Q_j^{++}}(\omega). \end{aligned} \tag{2.47}$$

here $|Q|$ denotes the total amount of points $\omega \in (\partial T)^2 \cap Q$, i.e. the amount of the smallest possible rectangles (of size $2^{-N} \times 2^{-N}$) in Q.

Observe that on Q_j the measure is basically a uniform distribution of the mass $\frac{1}{N}$ over the upper right quarter Q_j^{++} of the rectangle Q_j (and these quadrants are disjoint).

To prove (2.44) we fix $\omega \in Q_j^{++}$ and split

$$\mathbf{V}^\mu(\omega) = \mathbf{V}^\mu_{Q_j^{++}}(\omega) + \mu(Q_j^{\circ +}) + \mu(Q_j^{+\circ}) + \mathbf{V}^\mu(Q_j),$$

where the first term sums up $\mu(Q)$ (recall that in our setting $\mathbf{I}^*\mu(Q) = \mu(Q)$ due to identification of measures on $[0,1]^2$ and $(\partial T)^2$) for Q between ω and Q_j^{++}. It is easy to see that $\mathbf{V}^\mu_{Q_j^{++}}(\omega) \lesssim \frac{1}{N}$ (the left-hand side is a double geometric sum). Trivially $\mu(Q_j^{\circ +}) + \mu(Q_j^{+\circ}) \leq \frac{2}{N}$. The non-trivial part is the estimate

$$\mathbf{V}^\mu(Q_j) \lesssim 1\,. \tag{2.48}$$

For each dyadic rectangle $R \supseteq \omega_0$ and each j' we have

$$\text{either } Q_{j'} \subseteq R, \text{ or } Q_{j'}^{++} \cap R = \emptyset. \tag{2.49}$$

Moreover, since the sides of rectangles Q_j are nested, the set $\{j' : Q_{j'} \subseteq R\}$ is an interval that contains j. For an interval of integers $[m, m+k]$ let

$$C^{[m,m+k]} := \{R \supseteq \omega_0 : \{j' : Q_{j'} \subseteq R\} = [m, m+k]\}.$$

Since each rectangle in $C^{[m,m+k]}$ contains $[0, 2^{-2^m}] \times [0, 2^{-2^{-m-k}N}]$, we have

$$\#C^{[m,m+k]} \leq (2^m + 1)(2^{-m-k}N + 1) \lesssim 2^{-k}N. \tag{2.50}$$

It follows that

$$\mathbf{V}^\mu(Q_j) = \sum_{[m,m+k]\ni j} (\#C^{[m,m+k]})(k+1)\frac{1}{N} \lesssim \sum_{k\geq 0}(k+1)^2 2^{-k} N \frac{1}{N} \lesssim 1. \tag{2.51}$$

This shows (2.48), and hence (2.44) is also proved.

Now we estimate $\mathbf{V}^\mu(\omega_0)$ from below. To this end we need a more careful *lower* bound on $\#C^{[m,m+k]}$. The set $C^{\{j\}}$ contains all rectangles R that contain Q_j and are contained in $[0, 2^{-2^{j-1}-1}] \times [0, 2^{-2^{-j-1}N-1}]$, so

$$\#C^{\{j\}} \geq 2^{j-1} \cdot 2^{-j-1}N \gtrsim N. \tag{2.52}$$

Hence

$$\mathbf{V}^\mu(\omega_0) \geq \sum_{j=1}^{M} (\#C^{\{j\}})\frac{1}{N} \gtrsim M. \tag{2.53}$$

This shows (2.45) as $M \asymp \log N$.

Next we construct the second example of ν and w such that the Carleson condition holds, but the REC (restricted energy condition) fails. The weight w is chosen as in (2.47), so this time it is the indicator function of an up-set. With the measure μ that we have just constructed we put

$$\nu := \mu + \nu|\omega_0,$$

where $\nu|\omega_0$ is the uniformly distributed over ω_0 measure of total mass $\frac{1}{N}$.

REC constant is large

Let us first give a lower bound for the REC constant. Consider $F = \omega_0$. Then by (2.52) we have

$$\mathcal{E}[\nu|F] \geq \sum_{j=1}^{M} (\#C^{\{j\}})\nu(\omega_0)^2 \gtrsim MN \cdot \nu(\omega_0)^2.$$

This shows that $[w, \nu]_{HC} \gtrsim \nu(\omega_0) \cdot NM = M$.

Carleson constant is small

Next we will verify that the Carleson condition (2.37) holds with a small constant. We may remove from the sum on the left-hand side of (2.37) all rectangles $Q \notin \mathcal{R}$. Then we can replace E by the union of remaining Q's without changing the left-hand side and decreasing the right-hand side. Hence we may reduce to the case when E is a union of members of $\mathcal{R}$. By (2.49) it follows that for each j we have either $Q_j \subseteq E$ or $Q_j^{++} \cap E = \emptyset$. Let $\mathcal{J} := \{j : Q_j \subseteq E\}$. Then we obtain

$$LHS(2.37) \leq \sum_{[m,m+k]\subseteq\mathcal{J}} \sum_{Q\in C^{[m,m+k]}} \big((k+1)/N + 1/N\big)^2. \tag{2.54}$$

Using (2.50) this implies

$$\begin{aligned} LHS(2.37) &\lesssim \sum_{[m,m+k]\subseteq\mathcal{J}} 2^{-k} N(k+2)^2(1/N)^2 \\ &\lesssim (\#\mathcal{J})/N \\ &\leq \mu(E) \\ &\leq \nu(E), \end{aligned}$$

so that $[w,\nu]_C \lesssim 1$.

2.6.5 REC does not imply embedding

In this section we emulate the previous construction, we start with $\{Q_j\}$ and measure μ but instead of adding ω_0 we will add a more sophisticated piece of measure.

We define $Q_j, \mu, \mathcal{R}, w$ as in the previous section. We continue with denoting

$$Q_{0,j} := Q_j, \quad \mu_0 := \mu \text{ from the previous section}\,.$$

Next we continue with defining a sequence of collections $\mathcal{Q}_k,\ k = 0,\dots,K \approx \log M$ of dyadic rectangles as follows

$$\mathcal{Q}_k := \left\{ Q_{k,j} := \bigcap_{i=j}^{j+2^k-1} Q_{0,i},\ j = 1,\dots,M-2^k \right\},\ k = 1,\dots,K. \tag{2.55}$$

In other words, $\mathcal{Q}_k$ consists of the intersections of 2^k consecutive elements of the basic collection $\mathcal{Q}_0$. The total amount of rectangles in $\mathcal{Q}_k$ is denoted by $M_k = M - 2^k + 1$.

For $k = 1,\dots,K$ let

$$\mu_k(\omega) := \frac{2^{-2k}}{N} \sum_{j=1}^{M_k} \frac{1}{|Q_{k,j}^{++}|} \mathbb{1}_{Q_{k,j}^{++}}(\omega), \quad \omega \in (\partial T)^2,$$

and define

$$\mu := \mu_0 + \sum_{k=1}^{K} \mu_k.$$

Embedding constant is large

Let us recall the direct version of Carleson embedding for our bi-tree

$$\int_{(\partial T)^2} (\mathbf{I}(fw))^2 \, \mathrm{d}\mu \le [w,\mu]_{CE} \int_{T^2} f^2 \cdot w. \tag{2.56}$$

We test the inequality (2.56) with the function

$$f(R) := \mu_0(R) = \mathbf{I}^*\mu_0(R).$$

Using (2.44) we obtain

$$\int_{T^2} f^2 \cdot w = \int_{T^2} \mathbf{V}^{\mu_0} \, \mathrm{d}\mu_0 \lesssim |\mu_0| = \frac{M}{N}. \tag{2.57}$$

On the other hand, by definition (2.55) and replacing M by 2^k in (2.53) we obtain

$$\mathbf{V}^{\mu_0}(Q_{k,j}) \gtrsim 2^k N \cdot \frac{1}{N} = 2^k. \tag{2.58}$$

It follows that

$$\int_{T^2} (\mathbf{I}(fw))^2 \, \mathrm{d}\mu = \int_{(\partial T)^2} (\mathbf{V}^{\mu_0})^2 \, \mathrm{d}\mu = \sum_{k=1}^{K} \int_{(\partial T)^2} (\mathbf{V}^{\mu_0})^2 \, \mathrm{d}\mu_k \gtrsim \sum_{k=1}^{K} 2^{2k} |\mu_k| \sim \frac{M}{N} \log M. \tag{2.59}$$

Substituting (2.57) and (2.59) in (2.56) we obtain $[w,\mu]_{CE} \gtrsim \log M$.

REC constant is small

We claim that $[w,\mu]_{HC} \lesssim 1$. This means that for any collection $\mathcal{A}$ of dyadic rectangles, setting $A := \cup_{R\in\mathcal{A}} R$, we have

$$\mathcal{E}[\mu|_A] \lesssim \mu(A). \tag{2.60}$$

To show (2.60) let $\nu_k := \mu_k|_A$, $k = 0, \ldots, K$. Then

$$\mathcal{E}[\mu|_A] = \sum_{n,k} \int_{(\partial T)^2} \mathbf{V}^{\nu_n} \, \mathrm{d}\nu_k \le 2 \sum_{n\ge k} \int (\partial T)^2 \mathbf{V}^{\nu_n} \, \mathrm{d}\nu_k \le 2 \sum_{n\ge k} \int (\partial T)^2 \mathbf{V}^{\mu_n} \, \mathrm{d}\nu_k.$$

Since $\operatorname{supp} \nu_k \subseteq \operatorname{supp} \mu_k$ it suffices to show

$$\sum_{n\ge k} \mathbf{V}^{\mu_n} \lesssim 1 \quad \text{on} \quad \operatorname{supp} \mu_k. \tag{2.61}$$

The claim (2.61) has the advantage that it does not depend on $\mathcal{A}$ any more.

For every $R\in\mathcal{R}$ we have

$$\begin{aligned}\mu_n(R) &= 2^{-2n}\#\{Q_{n,j}\subseteq R\}\le 2^{-2n}(\#\{Q_{0,j}\subseteq R\}+2^n)\\ &\le 2^{-n}(\#\{Q_{0,j}\subseteq R\}+1)\le 2\cdot 2^{-n}\mu_0(R).\end{aligned}$$

It follows that

$$\mathbf{V}^{\mu_n}(Q_{k,j})\lesssim 2^{-n}\mathbf{V}^{\mu_0}(Q_{k,j})\le 2^{-n}\sum_{i=j}^{j+2^k-1}\mathbf{V}^{\mu_0}(Q_{0,i})\lesssim 2^{k-n},$$

where the last inequality follows from (2.44). This implies (2.61) and therefore (2.60).

2.6.6 Two-functions small energy lemma

The key element in the proof of Theorem 2.1.1 was the so-called *energy majorization* technique, which we stated (and proved) separately on T^2 and T^3, see Lemmas 2.2.7 and 2.2.11 respectively. There we showed, essentially, that if the measure μ on T^2 is concentrated on a set of 'small' μ-potential (say $\mathbf{V}^\mu\le 1$) then $\mathbf{I}^*\mu$ is a rather ineffective, in the sense of energy, way to provide the same potential on the set of 'large' μ-potential, i.e. where $\mathbf{V}^\mu\ge\lambda\ge 10$. In a sense it is another way to say that these 'small potential' and 'large potential' sets are far from each other. It turns out (and that was how the arguments in [2] and [3] ran) that on a tree there is even a more powerful statement, where the 'distance' between these two sets is measured with one measure, and the actual potential is realized via another measure. In particular the energy majorization on T^2 follows from this extended version pretty much immediately. A natural suggestion was that in order to move to higher dimensions one should obtain T^d version of this two-measure majorization for $d\ge 2$. However, it turns out that it fails to hold already on T^2, so that our energy majorization argument on T^3 uses a workaround (which itself stops working on T^d with $d\ge 4$).

In this Section we first give a proof of two-measure (two-function actually) Lemma on a tree, and then we show a counterexample that disproves it on T^2.

Lemma 2.6.1 *Let $f,g:T\to\mathbb{R}_+$ be functions on a 1-tree T such that* $\operatorname{supp} f\subset\{\mathbf{I}g\le\delta\}$, *and g is superadditive. Then there exists $\varphi:T\to\mathbb{R}_+$ such that*

$$a)\ \mathbf{I}\varphi(\omega)\ge If(\omega)\quad\forall\omega\in\partial T\colon \mathbf{I}g(\omega)\in[\lambda,2\lambda] \tag{2.62}$$

$$b)\ \int_T\varphi^2\le C\frac{\delta}{\lambda}\int_T f^2. \tag{2.63}$$

The proof of Lemma 2.6.1

We start with lemma that holds regardless of operator and medium.

Lemma 2.6.2 *Let K be an integral operator with a positive kernel and f,g positive functions. Then*

$$\int(Kf)^2g\le\Big(\sup_{\operatorname{supp} g}KK^*g\Big)\int f^2.$$

Proof. Without loss of generality f is positive. By duality we have

$$\int (Kf)^2 g = \int fK^*(Kf \cdot g) \leq \|f\|_2 \|K^*(Kf \cdot g)\|_2.$$

We call the operator and its kernel by the same letter K. By the hypothesis $Kh(x) = \int K(x,y)h(y)$ with a positive kernel K. Hence

$$\begin{aligned}
\|K^*(Kf \cdot g)\|_2^2 &= \int K^*(Kf \cdot g)K^*(Kf \cdot g) \\
&= \int K(x,y)((If)(x)g(x))K(x',y)((Kf)(x')g(x'))\,\mathrm{d}(x,x',y) \\
&\leq \int \frac{1}{2}(Kf(x)^2 + Kf(x')^2)K(x,y)(g(x))K(x',y)(g(x'))\,\mathrm{d}(x,x',y) \\
&= \frac{1}{2}\int K^*((Kf)^2 \cdot g)K^*(g) + \int K^*(g)K^*((Kf)^2 \cdot g) \\
&= \int (KK^*g) \cdot (Kf)^2 \cdot g \leq \Big(\sup_{\operatorname{supp} g} KK^*g\Big) \int (Kf)^2 \cdot g.
\end{aligned}$$

Substituting the second displayed estimate into the first we obtain

$$\int (Kf)^2 g \leq \|f\|_2 \Big(\sup_{\operatorname{supp} g} KK^*g\Big) \Big(\int (Kf)^2 \cdot g\Big)^{1/2}.$$

The conclusion follows. $\square$

In the preceding lemma operator K could have been either $\mathbf{I}$ on T or $\mathbf{I}$ on T^d, this did not matter. But in the next lemma, it matters whether we are on T or T^2.

Lemma 2.6.3 *Let T be a finite tree and $g, h : T \to \mathbb{R}_+$. Assume that g is superadditive and $\lambda = \|Ih\|_{L^\infty(\operatorname{supp} g)}$. Then for every $\beta \in T$ we have*

$$\mathbf{I}(gh)(\beta) = \sum_{\alpha \leq \beta} g(\alpha)h(\alpha) \leq \lambda g(\beta).$$

Proof. Without loss of generality we may think that β is the unique maximal element of T and $T = \operatorname{supp} g$. We induct on the depth of the tree. Let T be given and suppose that the claim is known for all its branches. Then by the inductive hypothesis and superadditivity of g we have

$$\begin{aligned}
\sum_{\alpha \leq \beta} g(\alpha)h(\alpha) &= g(\beta)h(\beta) + \sum_{\beta' \in \operatorname{ch}(\beta)} \sum_{\alpha \leq \beta'} g(\alpha)h(\alpha) \\
&\leq g(\beta)h(\beta) + \sum_{\beta' \in \operatorname{ch}(\beta)} g(\beta') \sup_{\alpha \leq \beta'} \sum_{\alpha \leq \alpha' \leq \beta'} h(\alpha') \\
&\leq g(\beta)h(\beta) + \sum_{\beta' \in \operatorname{ch}(\beta)} g(\beta') \sup_{\alpha < \beta} \sum_{\alpha \leq \alpha' < \beta} h(\alpha') \\
&\leq^{key} g(\beta)h(\beta) + g(\beta) \sup_{\alpha < \beta} \sum_{\alpha \leq \alpha' < \beta} h(\alpha') = g(\beta) \sup_{\alpha \leq \beta} \sum_{\alpha \leq \alpha' \leq \beta} h(\alpha'). \qquad \square
\end{aligned}$$

Now we present the proof of Lemma 2.6.1 by means of Lemma 2.6.2 and Lemma 2.6.3. Let $\varphi = 2\lambda^{-1}\mathbf{I}f \cdot g \cdot \mathbb{1}_{\mathbf{I}g\leq 4\lambda}$. Let ω be such that $\mathbf{I}g(\omega) \geq \lambda$. Then $f(\omega) = 0$ and $f(\gamma) = 0$ for all ancestors of ω up to the first γ' such that $\mathbf{I}g(\gamma') \leq \delta$. Hence, on such ω

$$\sum_{\gamma\geq\omega} \mathbf{I}f \cdot g \cdot \mathbb{1}_{\mathbf{I}g\leq 4\lambda} = \sum_{\gamma\geq\omega} \mathbf{I}f \cdot g = \mathbf{I}f(\omega)(\lambda - \delta) \geq \frac{\lambda}{2}\mathbf{I}f(\omega).$$

We checked (2.62) of Lemma 2.6.1.

To check 2.63 we first apply Lemma 2.6.2 with

$$K := I \circ \mathbb{1}_{\mathbf{I}g\leq\delta},$$

which a composition of multiplication operator and $\mathbf{I}$. Then

$$\int_T \varphi^2 = \frac{4}{\lambda^2}\int_T (\mathbf{I}f)^2 (g\mathbb{1}_{\mathbf{I}g\leq 4\lambda})^2 \leq \frac{4}{\lambda^2} \sup_{\operatorname{supp} g} KK^*(g^2\mathbb{1}_{\mathbf{I}g\leq 4\lambda}) \int_T f^2\,.$$

To understand $\sup_{\operatorname{supp} g} KK^*(g^2\mathbb{1}_{\mathbf{I}g\leq 4\lambda})$ we use Lemma 2.6.3. By this lemma for any node α

$$K^*(g^2\mathbb{1}_{\mathbf{I}g\leq 4\lambda})(\alpha) \leq \mathbf{I}^*(g^2\mathbb{1}_{\mathbf{I}g\leq 4\lambda})(\alpha) \leq 4\lambda g(\alpha)\,.$$

Now we are left to estimate $Kg = \mathbf{I}(\mathbb{1}_{\mathbf{I}g\leq\delta}g)$. But just by definition of $\mathbf{I}$ we have

$$\mathbf{I}(\mathbb{1}_{\mathbf{I}g\leq\delta}g) \leq \delta. \tag{2.64}$$

So $\sup_{\operatorname{supp} g} KK^*(g^2\mathbb{1}_{\mathbf{I}g\leq 4\lambda}) \leq 4\delta\lambda$ and we get

$$\int_T \varphi^2 \leq \frac{16\delta}{\lambda}\int f^2\,.$$

How it could have worked

As we will see in a short while, Lemma 2.6.1 fails on the bi-tree. However, imagine just for a moment, that we do not know it yet and try to see what happens if we make an attempt of emulating its proof.

A believable choice of φ would be

$$\begin{aligned}\varphi = \lambda^{-1}(&\mathbf{I}_1 f\cdot\mathbf{I}_{23}g + \mathbf{I}_2 f\cdot\mathbf{I}_{13}g + \mathbf{I}_3 f\cdot\mathbf{I}_{12}g +\\ &\mathbf{I}_1 gf\cdot\mathbf{I}_{23}f + \mathbf{I}_2 g\cdot\mathbf{I}_{13}f + \mathbf{I}_3 g\cdot\mathbf{I}_{12}f + g\mathbf{I}f).\end{aligned} \tag{2.65}$$

This function satisfies

$$\mathbf{I}(\cdot\mathbb{1}_{\mathbf{I}f\leq 2\lambda}\cdot\varphi) \geq \mathbf{I}f, \quad \text{where } \mathbf{I}g \in [\lambda, 2\lambda], \tag{2.66}$$

which the analogue of a) of Lemma 2.6.1. However, we cannot (and it is eventually not possible) prove the analogue of b) of Lemma 2.6.1 for this function. This misfortune happens since we have no good estimate of $\int_{T^d} (\mathbf{I}f)^2 g^2$ via $\int_{T^d} f^2$ for g that is *separately superadditive* in all variables.

Notice that this hurdle is removed if $f = g$ because then

$$\mathbf{I}(\lambda^{-1} g \mathbf{I}f) = \mathbf{I}(\lambda^{-1} f \mathbf{I}f) \leq \frac{\delta}{\lambda} \mathbf{I}f \leq \frac{1}{10} \mathbf{I}f,$$

and in place of φ from (2.65), we have another φ for majorization:

$$\tilde{\varphi} := c\lambda^{-1}(2\mathbf{I}_1 f \cdot \mathbf{I}_{23} f + 2\mathbf{I}_2 f \cdot \mathbf{I}_{13} f + 2\mathbf{I}_3 f \cdot \mathbf{I}_{12} f),$$

where $c = \frac{10}{9}$. In fact from (2.66) it now follows that

$$\mathbf{I}(\mathbb{1}_{\mathbf{I}f \leq 2\lambda} \cdot \tilde{\varphi}) \geq \mathbf{I}f, \quad \text{where } \mathbf{I}f \in [\lambda, 2\lambda]\,. \tag{2.67}$$

The analogue of inequality b) of Lemma 2.6.1 $\equiv$ (2.63) on tri-tree now follows from Lemma 2.2.10.

But What goes wrong on 4-tree?

So, we do not know how to estimate $\int_{T^3} (\mathbf{I}f)^2 g^2$ via $\int_{T^3} f^2$, but we know how to remove this hurdle in the case $f = g$. May be one can also remove this hurdle for $f = g$ on d-tree, $d \geq 4$?

Unfortunately, we can see now that the trick does not work for $d \geq 4$. Let us notice that by the analogy with (2.65) we can construct φ for 4-tree:

$$\begin{aligned}
&\varphi = \lambda^{-1}(\mathbf{I}_1 f \cdot \mathbf{I}_{234} g + \mathbf{I}_2 f \cdot \mathbf{I}_{134} g + \mathbf{I}_3 f \cdot \mathbf{I}_{124} g + \mathbf{I}_4 f \cdot \mathbf{I}_{123} g + \\
&\mathbf{I}_1 g \cdot \mathbf{I}_{234} f + \mathbf{I}_2 g \cdot \mathbf{I}_{134} f + \mathbf{I}_3 g \cdot \mathbf{I}_{124} f + \mathbf{I}_4 g \cdot \mathbf{I}_{123} f + \\
&\mathbf{I}_{12} g \cdot \mathbf{I}_{34} f + \mathbf{I}_{23} g \cdot \mathbf{I}_{14} f + \mathbf{I}_{34} g \cdot \mathbf{I}_{12} f + \mathbf{I}_{12} f \cdot \mathbf{I}_{34} g + \mathbf{I}_{23} f \cdot \mathbf{I}_{14} g + \mathbf{I}_{34} f \cdot \mathbf{I}_{12} g + \\
&g\mathbf{I}f).
\end{aligned} \tag{2.68}$$

Here $\mathbf{I}$ means summation in all 4 variables, the Hardy operator on T^4. Let us consider what happens for the case $g = f$. We again can absorb the last term $g\mathbf{I}f = f\mathbf{I}f \leq \delta f$ into the left hand side because $\operatorname{supp} f \subset \{\mathbf{I}f \leq \delta\}$.

But to prove the analogue of b) of Lemma 2.6.1 we would need to know how to estimate e. g.

$$\int_{T^4} (\mathbf{I}_{12} f \cdot \mathbf{I}_{34} f)^2 \leq C \int_{T^4} f^2\,.$$

We do not know how to achieve such an estimate. In a sense, everything works as long as we have a 1-dimensional factor $\mathbf{I}_j$ present everywhere in the formula. However, when a term with only 2-dimensional factors, like above, appears, our argument does not work any more.

Lemma 2.6.1 does not hold on T^2

Proposition 2.6.2 *Lemma 2.6.1 does not hold on T^d. Namely, let h be a function on $\mathbb{R}$ such that* $\lim_{t\to 0} h(t) = 0$ *and $w : T^2 \to \mathbb{R}_+$ be a unit weight, $w \equiv 1$. There exists a pair of numbers $\lambda \geq 10\delta > 0$ and a pair functions $f, g : T^2 \to \mathbb{R}_+$ such that* $\operatorname{supp} f \subset \{\mathbf{I}f \leq \delta\}$ *and g is* *superadditive in each variable, and such that for any* $\varphi : T^2 \to \mathbb{R}_+$ *such that*

$$\mathbf{I}\varphi \geq \mathbf{I}f \quad \text{on } \{2\lambda \leq \mathbf{I}g \leq 4\lambda\}$$

the estimate

$$\int_{T^2} \varphi^2 \leq h\left(\frac{\delta}{\lambda}\right) \int_{T^2} f^2 \tag{2.69}$$

does not hold for the bitree of sufficiently large depth (recall that we work with finite d-trees T_N^d *here).*

Below f, g have special form, namely

$$f = \mathbf{I}^*\mu,\ g = \mathbf{I}^*\nu,$$

with certain positive measures on T^2, where the measure μ is trivial — it is just a unit mass at the root o of T^2. In particular, $f(o) = 1, f(v) = 0, \forall v \neq o$. Clearly $\mathbf{I}f \equiv 1$ on T^2.

The choice of ν is more sophisticated. First we choose a large number M. Consider now another number $n = 2^{2^s} >> M$ for some natural s, its value is defined in a few lines. In the unit square Q^0 consider dyadic sub-squares $Q_1, \dots, Q_{2^M}$, which are South-West to North-East diagonal squares of sidelength 2^{-M}.

In each Q_j choose ω_j, the South-West corner dyadic square of sidelength 2^{-n-M}. Now let ν be a sum of identical masses at ω_j and let n and these masses satisfy the following relation

$$\begin{aligned} &\nu(\omega) := \frac{1}{n^2}, \quad j = 1, \dots, 2^M \\ &2^M = \frac{n}{\log n}. \end{aligned} \tag{2.70}$$

We have immediately

$$g(o) = \mathbf{I}^*\nu(o) = |\nu| = \frac{1}{n^2} \cdot \frac{n}{\log n} = \frac{1}{n \log n} =: \delta.$$

Clearly we have chosen f, g satisfying $\operatorname{supp} f = \{o\} \subset \{\mathbf{I}g \leq \delta\}$ with g being sub-additive in both variables on T^2: it is true for any function of the form $\mathbf{I}^*\nu$.

Now what is λ, and what is the set $\{2\lambda \leq \mathbf{I}g \leq 4\lambda\}$? For Q_1 and ω_1 consider the family $\mathcal{F}_1$ of dyadic rectangles containing ω_1 and contained in Q_1 of the following sort:

$$[0, 2^{-n}2^{-M}] \times [0, 2^{-M}], [0, 2^{-n/2}2^{-M}] \times [0, 2^{-2}2^{-M}], \dots, [0, 2^{-n/2^k}2^{-M}] \times [0, 2^{-2^k}2^{-M}],$$

there is $\frac{\log n}{\log 2}$ of them, and they are called $q_{10}, q_{11}, \dots, q_{1k}$, $k \asymp \log n$. We do the same for each ω_j, Q_j and we get $q_{j0}, q_{j1}, \dots, q_{jk}$.

As we have already seen in the 'REC-not Carleson embedding', one has

$$\mathbf{I}g(q_{ji}) \asymp \frac{1}{n} \quad \forall j, i. \tag{2.71}$$

Let

$$F := \bigcup_{ik} q_{ik} \,. \tag{2.72}$$

So we choose $\lambda = \frac{c}{n}$ with an appropriate c. Then $F \subset \{2\lambda \le \mathbf{I}g \le 4\lambda\}$. Since $\mathbf{I}f \ge 1$, then if majorizing function φ as in 2.69 would exist, we would have $\mathbf{I}\varphi \ge 1$ on F and

$$\int_{T^2} \varphi^2 \le \frac{C}{\log n} \int_{T^2} f^2 = \frac{C}{\log n} \,.$$

By the definition of capacity this would mean that

$$\mathrm{Cap}(F) \le \frac{C}{\log n} \,.$$

Below we show that $\mathrm{Cap}(F) \asymp 1$. Hence there is no such majorizing function.

Let ρ on F be an equilibrium measure of F, and let μ be a measure charging $\frac{1}{n}$ on each q_{jk} with $\frac{\log n}{4} \le k \le \frac{3\log n}{4}$, and zero otherwise. Clearly $|\mu| = \sum_{j=1}^{\frac{n}{\log n}} \sum_{k=\frac{\log n}{4}}^{\frac{3\log n}{4}} \mu(q_j k) \asymp 1$. We claim that

$$\mathbf{V}^\mu \asymp 1 \quad \text{on } \operatorname{supp} \mu. \tag{2.73}$$

Assuming for a moment that this estimate holds, we write for $\varepsilon > 0$

$$0 \le \mathcal{E}[\rho - \varepsilon\mu] = \left(\int_{T^2} \mathbf{V}^\rho \, d\rho - \varepsilon \int_{T^2} \mathbf{V}^\rho \, d\mu\right) + \varepsilon \left(\varepsilon \int_{T^2} \mathbf{V}^\mu \, d\mu - \int_{T^2} \mathbf{V}^\rho \, d\mu\right). \tag{2.74}$$

Since ρ is capacitary for $F \supset \operatorname{supp}\mu$ and T^2 is finite (i.e. every singleton has positive capacity), we have $\mathbb{V}^\rho \ge 1$ on $\operatorname{supp}\mu$, and $\int_{T^2} \mathbf{V}^\rho \, d\mu \ge |\mu|$. By (2.73) there is some absolute ε such that $\varepsilon \int_{T^2} \mathbf{V}^\mu \, d\mu \le |\mu|$, so that the second term in (2.74) must be negative. But then the first term is positive, which means

$$\mathrm{Cap}\, F = \int_{T^2} \mathbf{V}^\rho \, d\rho \ge \varepsilon \int_{T^2} \mathbf{V}^\rho \, d\mu \ge \varepsilon |\mu| \asymp 1.$$

It remains to prove (2.73). By symmetry it is enough to estimate the potential at q_{1k}. For that we split $\mathbf{V}^\mu$ to $\mathbb{V}_1$, this is the contribution of rectangles containing Q_1, to $\mathbb{V}_2$, the contribution of rectangles containing q_{1k} and contained in Q_1, and $\mathbb{V}_3$, the contribution of rectangles containing q_{1k} that strictly intersect Q_1 and that are "vertical", meaning that there vertical side contains vertical side of Q_1 (there is $\mathbb{V}_4$ totally symmetric to $\mathbb{V}_3$).

Two of these are easy, $\mathbb{V}_1$ 'almost' consists of 'diagonal squares containing Q_1'. Not quite, but other rectangles are also easy to take care of. Denote

$$r = |\mu|, \quad M = \log \frac{n}{\log n} \,.$$

Then we write the diagonal part first and then the rest:

$$\mathbb{V}_1 = r + \frac{r}{2} + \frac{r}{4} + \dots \frac{r}{2^M} + \frac{r}{2} + \frac{r}{2} + 2\frac{r}{4} + 2\frac{r}{4} + \dots k\frac{r}{2^k} + 2\frac{r}{2^k} + \dots \asymp 1$$

To estimate $\mathbb{V}_2$ notice that there are at most cn rectangles containing q_{1k} and contained in Q_1 that do not contain any other q, there are $\frac{cn}{2}$ of rectangles containing q_{1k} and one of its siblings (and lie in Q_1), there are $\frac{cn}{4}$ of rectangles containing q_{1k} and two of its siblings (and lie in Q_1), et cetera.

Hence,

$$\mathbb{V}_2 \le Cn\frac{1}{n} + \frac{Cn}{2}\frac{2}{n} + \frac{Cn}{4}\frac{3}{n} + \dots \lesssim 1$$

Now consider $\mathbb{V}_3$. The horizontal size of q_{1k} is $2^{-M} \cdot 2^{-n2^{-k}}$. Its vertical size is $2^{-M} \cdot 2^{-2^k}$. So the rectangles of the third type that do not contain the siblings: their number is at most (we are using that $k \ge \frac{1}{4}\log n$)

$$n2^{-k}(2^k + M) \le n + n^{\frac{3}{4}} \log n\,.$$

Regarding those that contain q_{1k} and one sibling, their number is at most

$$n2^{-k}(2^{k-1} + M) \le \frac{n}{2} + n^{\frac{3}{4}} \log n\,.$$

We continue, and arrive at

$$\mathbb{V}_3 \le n\frac{1}{n} + \frac{n}{2}\frac{2}{n} + \frac{n}{4}\frac{3}{n} + \dots + n^{\frac{3}{4}} \log n \frac{\log^2 n}{n} \lesssim 1\,.$$

We deal with $\mathbb{V}_4$ in exactly the same way, only now we use that $k \le \frac{3}{4}\log n$. Finally after adding all $\mathbb{V}_i$ we get

$$\mathbb{V}_1 + \mathbb{V}_2 + \mathbb{V}_3 + \mathbb{V}_4 \le C_1 + C_2 + C_3\frac{\log^3 n}{n^{\frac{1}{4}}}.$$

Since the inverse estimate is already given by $\mathbb{V}_1$, we obtain (2.73). We are done.

The shape of the graph of function $x \to$ cap$(\mathbb{V}^\nu \ge x)$ for $w \equiv 1$

Let E be a subset of T or T^2 and ν be equilibrium measure for E,

$$\mathrm{Cap}(E) = |\nu|,\ \mathbf{V}^\nu = 1 \text{ on } \operatorname{supp}\nu,\ f := \mathbf{I}^*\nu = \{f : \int f^2 \to \min \text{ for } \mathbf{I}f \ge 1 \text{ on } E\}\,.$$

First consider the case of T. Let $x \in [|\nu|, 1]$ and we study the set

$$D_x := \{\alpha \in T : \mathbf{V}^\nu(\alpha) \ge x\}\,.$$

We want to understand a bit the shape of the graph of

$$C(x) := \mathrm{Cap}(D_x)\,.$$

We start with $x = |\nu| = \mathbf{V}(E)$. Notice that o, the root of T, is obviously such that $\mathbb{V}^\nu(o) = |\nu|$, so $0 \in D_{|\nu|}$. But $\mathrm{cap}(o) \asymp \mathrm{cap}(T) = 1$. Thus

$$C(|\nu|) = 1\,.$$

Now consider $x = 1$. On E we have $\mathbf{V}^\nu = 1$ and maximum principle (we are on T, so it exists) says that $E = \{\alpha : \mathbf{V}^\nu \geq 1\}$. Therefore,

$$C(1) = \mathrm{cap}(E) = |\nu|\,.$$

Now let $|\nu| < x < 1$. We know (again this is maximum principle) that

$$\int_T \mathbb{1}_{\mathbf{I}g \leq x} \cdot g^2 = \int_T \mathbf{V}^\nu_x \, d\nu \leq x|\nu|\,. \tag{2.75}$$

Notice that if $\mathbf{I}g(\alpha) \leq x$ and $\mathbf{I}g(\mathrm{son}\,\alpha) > x$ then $\mathbb{I}g(\alpha) \geq x/2$ just because $g = \mathbf{I}^*\nu$ is monotonically increasing on T. But this means that

$$\mathbf{I}(\mathbb{1}_{\mathbf{I}g \leq x} \cdot g) \geq x/2, \quad \text{on } D_x = \{\mathbf{I}g = \mathbf{V}^\nu \geq x\}\,. \tag{2.76}$$

The definition of capacity and relationships (2.75), (2.76) show the following:

Proposition 2.6.3 *On a simple tree T the capacity of the level set $D_x = \{\alpha \in T : \mathbf{V}^\nu(\alpha) \geq x\}$ for any equilibrium measure ν of a set E satisfies the following inequality*

$$C(x) = \mathrm{cap}(\{\alpha \in T : \mathbf{V}^\nu(\alpha) \geq x\}) \leq \frac{4\,\mathrm{cap}(E)}{x} = \frac{4|\nu|}{x},\ \mathrm{cap}(E) \leq x \leq 1\,.$$

This is absolutely not the case for T^2. The capacity of level set of equilibrium potentials on T^2 behave in a much stranger and wild way. We saw it in the proof of Proposition 2.6.2. In fact, our measure ν there is (after multiplying by a constant) an equilibrium measure,

$$|\nu| = \frac{1}{n\log n}\,.$$

We put

$$x = \frac{c}{n}\,.$$

But we saw above that if the absolute constant c is chosen correctly, then

$$\mathrm{cap}((\alpha_1, \alpha_2) \in T^2 : \mathbf{V}^\nu(\alpha_1, \alpha_2) \geq \frac{c}{n}) \asymp 1 >> \frac{|\nu|}{x}\,. \tag{2.77}$$

This means that Proposition 2.6.3 is false for T^2 because if it were true, that we would have $\mathrm{cap}((\alpha_1, \alpha_2) \in T^2 : \mathbf{V}^\nu(\alpha_1, \alpha_2) \geq \frac{c}{n}) \lesssim \frac{1}{\log n}$.

The main reason behind this is that on T^2 we do not have a proper Maximum Principle, which is (2.75) above. Instead we have only its surrogate version that makes the estimate of capacity much faster blowing up than in Proposition 2.6.3. In fact, SMP means that

$$\operatorname{cap}(\{\mathbf{V}^{\nu} \geq x\}) \leq \frac{C_{\tau}\operatorname{cap}(E)}{x^{1+\tau}}\,.$$

and we saw that τ is indispensable. Of course the capacity of any subset of T^2 is bounded by 1, so we have

$$\operatorname{cap}(\{\mathbf{V}^{\nu} \geq x\}) \leq \max\left(1, \frac{C_{\tau}\operatorname{cap}(E)}{x^{1+\tau}}\right).$$

This explains a flat piece of graph $C(x) \asymp 1$, when x is between $\frac{1}{n\log n}$ and $\frac{1}{n}$.

Chapter 3 From the d-tree to the polydisc: Carleson measures

In this Chapter we will make the transition from the discrete medium we worked with in the previous Chapter to the continuous one – polydisc $\mathbb{D}^d$ of dimension $d = 1, 2, 3$. The space $L^2(T^d, w)$ will be considered as a discrete model of the weighted Hardy-Sobolev space $\mathcal{H}_s$ of harmonic (and sometimes analytic!) functions on the polydisc, and the discrete Hardy embedding (2.1) (after a suitable choice of weight w) will transform into the Carleson embedding for $\mathcal{H}_s$ (with respective changes in the testing conditions – especially subcapacitory). Observe that, strictly speaking, we do not discretize the spaces $\mathcal{H}_s(\mathbb{D})$, but rather the polydisc $\mathbb{D}^d$ and the Carleson embedding inequality. This allows us to evade some extra technicalities regarding moving the values of functions from continuous to discrete setting (though, naturally the arising L^2 spaces on T^d serve perfectly well as discrete models of $\mathcal{H}_s(\mathbb{D})$).

3.1 Discretization procedure: polydisc and embedding

3.1.1 Discretizing the polydisc $\mathbb{D}^d$

We start with making a decomposition of the unit disc into dyadic Carleson boxes. For integer $j \geq 0$ and $1 \leq l \leq 2^j$ let $z_{jl} = (1-2^{-j})e^{\frac{2\pi i(2l-1)}{2^{j+1}}}$, and for $z = re^{it} \in \mathbb{D}$ let $J(z) = \{e^{is} : t-(1-r)\pi \leq s < t + (1-r)\pi\}$, $S(z) = \{\rho e^{is} : e^{is} \in J(z); r \leq \rho \leq 1\}$, and let $Q(z) = \{\rho e^{is} \in S(z) : \frac{1-r}{2} < 1 - \rho \leq 1 - r\}$ be the 'upper half' of $S(z)$. We write $Q_{jl} := Q(z_{jl})$. Now we see that there is one-to-one map between points (vertices) of T and dyadic Carleson half-cubes Q_{jl}; Q_{00} corresponds to the root o, Q_{11} and Q_{12} to its two children etc. In other words, for every $\alpha \in T$ there exists a unique half-cube Q_α, and vice versa, for every half-cube Q_{jl} there is exactly one point $\alpha^{jl} \in T$. The collection $\{Q_\alpha\}_{\alpha \in T}$ forms a covering of the unit disc. Note also that given a point $z \in \mathbb{D}$ it is possible to pick the half-box $Q_\alpha \ni z$ in a unique way.

Next we introduce an auxiliary graph structure on T by setting

$$\mathcal{P}_{\mathfrak{G}}(\tau) := \bigcup_{\beta \in \mathcal{P}(\tau)} \{\gamma \in T : \operatorname{cl} Q_\gamma \bigcap \operatorname{cl} Q_\beta \neq \emptyset,\ |Q_\gamma| = |Q_\beta|\}. \tag{3.1}$$

Essentially, for a point $\tau \in T$ we define the $\mathfrak{G}$-*extended predecessor set* by taking the usual predecessor set $\mathcal{P}(\tau)$ and adding all the 'euclidean-adjacent' vertices γ, i.e. to each $\beta \in \mathcal{P}(\tau)$ we attach two extra vertices γ of the same rank (depth in T) and such that their respective Carleson half-boxes intersect with Q_β. This type of construction is often used when dealing with dyadic structures – basically it the same as to take a dyadic interval and consider also two adjacent ones of the same length. The $\mathfrak{G}$-*extended successor set* is defined accordingly,

$$\tau \in \mathcal{S}_{\mathfrak{G}}(\tau) \iff \gamma \in \mathcal{P}_{\mathfrak{G}}(\tau). \tag{3.2}$$

Clearly this definition can be extended to the completed tree $\overline{T}$.

Since by definition $\mathcal{P}(\alpha) \subset \mathcal{P}_{\mathfrak{G}}(\alpha)$ for any $\alpha \in \overline{T}$, we have the same inclusion for the successor sets, $\mathcal{S}(\alpha) \subset \mathcal{S}_{\mathfrak{G}}(\alpha)$, and this inclusion is proper unless the vertex in question is the root o. On the other hand, the successor sets are 'comparable on average'. To elaborate, for any point α there exists a set $N(\alpha)$ which consists of exactly three points – α and its two 'euclidean neighbors' – such that

$$\mathcal{S}_{\mathfrak{G}}(\alpha) \subset \bigcup_{\beta \in N(\alpha)} \mathcal{S}(\beta), \tag{3.3}$$

also $\bigcup_{\alpha \in T} N(\alpha)$ covers each point at most 3 times.

The correspondence between ∂T and $\mathbb{T}$ is a bit more complicated – there is no one-to-one correspondence anymore, since dyadic-rational points are counted twice. On the other hand, we do not really need to consider the respective boundaries, and the exceptional points form a countable hence polar set (for any related capacity).

Given a boundary point $\omega \in \partial T$ which is also a geodesic $\{o, \omega^1, \dots\}$ one can consider its image $S(\omega) := \bigcap_{k \geq 0} S_{\omega^k}$. This mapping is not injective though, since a point on a circle can be represented by two different geodesics on the tree.

Now we do the same for the polydisc $\mathbb{D}^d$. The next step is more or less automatic – we consider the partition

$$\begin{aligned} \mathbb{D}^d &= \bigcup_{\alpha \in T^d} Q_\alpha, \\ Q_\alpha &= Q_{(\alpha_1, \dots, \alpha_d)} = \prod_{k=1}^{d} Q_{\alpha_k}, \quad \alpha_k \in T. \end{aligned}$$

Similarly

$$\begin{aligned} \mathcal{P}_{\mathfrak{G}}(\alpha) &= \prod_{k=1}^{d} \mathcal{P}(\alpha_k), \\ \mathcal{S}_{\mathfrak{G}}(\alpha) &= \prod_{k=1}^{d} \mathcal{S}(\alpha_k), \quad \alpha = (\alpha_1, \dots, \alpha_d) \in T^d. \end{aligned}$$

Again, the usual predecessor/successor sets and the $\mathfrak{G}$-extended ones on T^d are not very different, in the same sense

$$\mathcal{S}_{\mathfrak{G}}(\alpha) \subset \bigcup_{\beta \in N(\alpha)} \mathcal{S}(\beta), \quad \alpha \in T^d, \tag{3.4}$$

with $\#N(\alpha) = 3^d$ and $\bigcup_{\alpha \in T} N(\alpha)$ covering each point in T^d at most 3^d times. We also consider the (non-injective) mapping

$$\Lambda : (\partial T)^d \to (\partial \mathbb{D})^d = (\Lambda(\omega_1), \dots, \Lambda(\omega_d))\,. \tag{3.5}$$

Remark. The main reason to introduce this auxiliary graph structure $\mathfrak{G}$ is that the geometry of the tree T does not completely agree with the geometry of the unit disc $\mathbb{D}$. For instance, one can easily find a pair of points $z, w \in \mathbb{D}$, very close to each other, while the tree distance between α and β corresponding to these points (i.e. $z \in Q_\alpha,\ w \in Q_\beta$) is very large. It is a well-known (if somewhat minor) obstacle, and there are several ways to overcome it. We have chosen what we think is the simplest one, especially since we do not care about precise values of arising constants.

3.1.2 Carleson embedding on T^d is equivalent to the polydisc embedding

Setting and basic definitions

Let us recall the definitions of the objects we are working with.
Given an integer $d \geq 1$ and $\vec{s} = (s_1, \dots, s_d) \in [0,1]^d$ we consider a Hilbert space $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$ of analytic functions on the polydisc $\mathbb{D}^d$ with the norm

$$\|f\|^2_{\mathcal{H}_{\vec{s}}(\mathbb{D}^d)} := \sum_{n_1,\dots,n_d \geq 0} |\widehat{f}(n_1, \dots, n_d)|^2 (n_1+1)^{s_1} \cdot \dots \cdot (n_d+1)^{s_d},$$

where

$$f(z) = \sum_{n_1,\dots,n_d \geq 0} \widehat{f}(n_1, \dots, n_d) z_1^{n_1} \cdot \dots \cdot z_d^{n_d}, \quad z = (z_1, \dots, z_d) \in \mathbb{D}^d.$$

Also,

$$\mathcal{H}_{\vec{s}}(\mathbb{D}^d) = \bigotimes_{j=1}^{d} \mathcal{H}_{s_j}(\mathbb{D}). \tag{3.6}$$

Recall that (w, μ) is a trace pair for the weighted Hardy inequality on $\overline{T}^d$, if $[w, \mu]_{CE} < +\infty$, i.e.

$$\textstyle\int_{\overline{T}^d} (\mathbf{I}_w f)^2\, d\mu \leq [w, \mu]_{CE} \int_{T^2} f^2\, dw, \quad f \in L^2(T^d, w), \tag{3.7a}$$

$$\textstyle\int_{T^d} (\mathbf{I}^*_\mu \varphi)^2\, dw \leq [w, \mu] \int_{\overline{T}^d} \varphi^2\, d\mu, \quad \varphi \in L^2(\overline{T}^d, \mu). \tag{3.7b}$$

Recall also that a measure ν on $\overline{\mathbb{D}}^d$ is called a Carleson measure for $\mathcal{H}_{\vec{s}}$, if there exists a constant C_ν such that

$$\sup_{r<1} \int_{\overline{\mathbb{D}}^d} |f(rz)|^2\, d\nu(z) \leq C_\nu \|f\|^2_{\mathcal{H}_s(\mathbb{D}^d)}, \tag{3.8}$$

or, in other words, the embedding $Id:\ \mathcal{H}_s(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, d\nu)$ is bounded. We are going to give a description of the relation between trace pairs and Carleson measures for $\mathcal{H}_{\vec{s}}$. We start by assuming that $\operatorname{supp} \nu \subset r\mathbb{D}^d$ for some $r < 1$ (the latter is just a convenience assumption, no estimate below will depend on r, or on the depth of the graph and later we will se how to deal with boundary measures).

It is well known that $\mathcal{H}_{s_j}(\mathbb{D})$, $1 \leq j \leq d$, is a *reproducing kernel Hilbert space* (RKHS) with kernel K_{s_j} satisfying (possibly after a suitable change of norm)

$$\begin{aligned} |K_{s_j}|(z_j,\zeta_j) &\asymp |1-z_j\bar{\zeta}_j|^{s_j-1}, \quad 0<s_j<1 \\ |K_{s_j}|(z_j,\zeta_j) &\asymp \log|1-z_j\bar{\zeta}_j|^{-1}, \quad s_j=1. \end{aligned} \tag{3.9}$$

Moreover it is not hard to verify that

$$\Re K_s \asymp |K_s|, 0<s\le 1\,. \tag{3.10}$$

However, the case $s=0$ is a special case as

$$\text{Poisson kernel is not equivalent to the absolute value of Cauchy kernel}\,. \tag{3.11}$$

It follows immediately that $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$ is a reproducing kernel Hilbert space as well, and

$$K_{\vec{s}}(z,\zeta)=\prod_{j=1}^d K_{s_j}(z_j,\zeta_j), \quad z,\zeta\in\mathbb{D}^d.$$

Going back to the Carleson embedding we see that $Id:\mathcal{H}_{\vec{s}}(\mathbb{D}^d)\to L^2(\mathbb{D}^d,d\nu)$ is bounded if and only if its adjoint Θ is bounded as well. Let us compute its action on a function $g\in L^2(\mathbb{D}^d,d\nu)$

$$(\Theta g)(z)=\langle\Theta g,K_{\vec{s}}(z,\cdot)\rangle_{\mathcal{H}_{\vec{s}}(\mathbb{D})}=\langle g,K_{\vec{s}}(z,\cdot)\rangle_{L^2()\mathbb{D}^d,d\nu}=\int_{\mathbb{D}^d} g(\zeta)\overline{K_{\vec{s}}(z,\zeta)}\,d\nu(\zeta).$$

Hence, for Θ to be bounded it must satisfy

$$\|g\|^2_{L^2(\mathbb{D}^d,d\nu)}\gtrsim\|\Theta g\|_{\mathcal{H}_{\vec{s}}(\mathbb{D}^d)}=\langle g,\Theta g\rangle_{L^2(\mathbb{D}^d,d\nu)}=\int_{\mathbb{D}^{2d}} g(z)\overline{g(\zeta)}K_{\vec{s}}(z,\zeta)\,d\nu(z)\,d\nu(\zeta). \tag{3.12}$$

If inequality (3.12) holds then trivially the following holds:

$$\|g\|^2_{L^2(\mathbb{D}^d,d\nu)}\gtrsim\int_{\mathbb{D}^{2d}} g(z)g(\zeta)K_{\vec{s}}(z,\zeta)\,d\nu(z)\,d\nu(\zeta), \quad g\ge 0\,. \tag{3.13}$$

If we would know that the real part of the coordinate reproducing kernel is comparable to its absolute value, we deduce that Θ is bounded, if and only if

$$\int_{\mathbb{D}^{2d}} g(z)g(\zeta)|K_{\vec{s}}(z,\zeta)|\,d\nu(z)\,d\nu(\zeta)\lesssim\|g\|^2_{L^2(\mathbb{D}^d,d\nu)} \tag{3.14}$$

for any positive g on $\mathbb{D}^d$.

In fact, (3.12) implies (3.13), and we can take the real part of both sides of (3.13), putting real part on kernel. Now if to know that

$$\Re K_{\vec{s}}(z,\zeta)=\Re\prod_{j=1}^d K_{s_j}(z_j,\zeta_j)\asymp|\prod_{j=1}^d K_{s_j}(z_j,\zeta_j)|=|K_{\vec{s}}(z,\zeta)|, \quad z,\zeta\in\mathbb{D}^d, \tag{3.15}$$

we would deduce (3.12)⇒(3.14). The only thing we need for this implication is the above pointwise equivalence (3.15). On the other hand, the implication (3.14)⇒(3.12) obviously always holds.

We conclude that in the presence of pointwise equivalence (3.15) we have (3.12)≡(3.14).

However, equivalence (3.15) – ultimately important for us to prove equivalence of dyadic and analytic embeddings (see below) – has limitations. First of all it is false even for $1D$ case $d = 1$ if $s = 0$, see (3.11). That makes the case $s = 0$ quite special. It is well known that for $1D$ case embedding measures for Poisson and Cauchy kernels on $L^2(\mathbb{T})$ are the same. This is rather simple, but should be consider as "a miracle". Already in $2D$ situation it is not known whether the embedding measures for Poisson $P_{z_1}P_{z_2}$ and Cauchy $K_{\vec{0}}(z, \zeta) = (1 - z_1\bar{\zeta}_1)^{-1}(1 - z_2\bar{\zeta}_2)^{-1}$ kernels on $L^2(\mathbb{T}^2)$ are the same, and the attempts to prove it lead to recent developments around [34].

Another interesting distinction of the case $s = 0$ is again about (3.11). For $s > 0$ characterize the embedding can be characterized in terms of simple box (rectangular) test. As it is well known from the works of Chang, Fefferman and Carleson [24], [31], [21], [93], such characterization is not possible for Poisson embedding of $L^2(\mathbb{T}^d)$ if $d \geq 2$.

Unweighted Dirichlet space

We first consider the case when all $s_j = 1$. For brevity we assume $d = 2$ – for unweighted Dirichlet space this is not a restriction of generality.

The reproducing kernel $K_{\vec{1}}(z, \zeta) = \log(1 - z_1\bar{\zeta}_1)\log(1 - z_2\bar{\zeta}_2) = K_1(z_1, \zeta_1)K_1(z_2, \zeta_2)$. The first idea is to see that our inequality (3.12) (equivalent to embedding):

$$\int_{\mathbb{D}^2} g(z)\overline{g(\zeta)}K_{\vec{1}}(z, \zeta)\, d\nu(z)\, d\nu(\zeta) \leq A\|g\|^2_{L^2(\mathbb{D}^2, d\nu)} \tag{3.16}$$

implies that for every $C \geq 0$ we have

$$\int_{\mathbb{D}^2} g(z)\overline{g(\zeta)}(C + K_1(z_1, \zeta_1))(C + K_1(z_2, \zeta_2))\, d\nu(z)\, d\nu(\zeta) \leq B(C)\|g\|^2_{L^2(\mathbb{D}^2, d\nu)} \tag{3.17}$$

To deduce the latter inequality from (3.16) one should open the brackets and consider 4 terms in the LHS. The term with $K_1(z_1, \zeta_1))K_1(z_2, \zeta_2)$ is $\lesssim \|g\|^2_{L^2(\mathbb{D}^2, d\nu)}$ by (3.16). The term with $C^2 \int_{\mathbb{D}^2} g(z)\overline{g(\zeta)}d\nu(z)\, d\nu(\zeta)$ obviously is $\lesssim \|g\|^2_{L^2(\mathbb{D}^2, d\nu)}$ by Hölder inequality. Consider one of mixed terms (they are treated symmetrically):

$$C\int_{\mathbb{D}^2} g(z)\overline{g(\zeta)}K_1(z_1, \zeta_1)\, d\nu(z)\, d\nu(\zeta) =: CI,$$

skip C, and, using disintegration theorem and pushing forward of ν to the first coordinate (we call that push forward ν_1), we write I as follows

$$I = \int_{\mathbb{D}} G(z_1)\overline{G(\zeta_1)}K_1(z_1, \zeta_1))\, d\nu_1(z_1)\, d\nu_1(\zeta_1),$$

where $G(w) := \int g(w, u)d\nu_w(u)$ and $d\nu_w(u)$ are slicing measures: $\nu(E) = \int \nu_w(E)d\nu_1(w)$.

Push forward measure ν_1 on $\mathbb{D}$ is obviously a Carleson measure for $1D$ Dirichlet space, if ν is a Carleson measure for Dirichlet space in $2D$. Therefore,

$$\int_{\mathbb{D}} G(z_1)\overline{G(\zeta_1)}K_1(z_1,\zeta_1))\,d\nu_1(z_1)\,d\nu_1(\zeta_1) \le B\int_{\mathbb{D}} |G_1(z_1)|^2\,d\nu_1(z_1) \le$$
$$B\int_{\mathbb{D}}\Big(\int_{\mathbb{D}} |g(z_1,z_2)|\,d\nu_{z_1}(z_2)\Big)^2\,d\nu_1(z_1) \le B'\int_{\mathbb{D}^2} |g(z_1,z_2)|^2\,d\nu_{z_1}(z_2)d\nu_1(z_1) \le$$
$$B'\int_{\mathbb{D}^2} |g(z_1,z_2)|^2\,d\nu(z)\,.$$

We deduced (3.17) from (3.16) by the use of the disintegration theorem and slicing measures. Notice that the nature of the kernel did not play any role. We could have done this with any dimension d and any kernel $K_{\vec{s}}$ instead of $K_{\vec{1}}$.

The fact that we worked with precisely $K_{\vec{1}}$ is crucial. In fact, values of $1-z\bar{\zeta}, z,\zeta\in\mathbb{D}$ K_1 are obviously in the right half-plane. Hence, as $\Im K_1$ is the argument of $\log\frac{1}{1-z\bar{\zeta}}$, we have

$$|\Im K_1(z,\zeta)| \le \pi/2. \tag{3.18}$$

Hence, by adding sufficiently large constant $C>0$ to $K_1(z,\zeta)$ we achieve a) $|\Re(C+K_1)| >> |\Im(C+K_1)|$, b) $|\Re(C^d+K_{\vec{1}}(z,\zeta))| \ge c\Re(\Pi_{j=1}^d(C+K_1)(z_j,\zeta_j)))$ for any dimension d, it is enough to choose $C=C(d)$ large positive number. The latter inequality implies that

$$\Re\Pi_{j=1}^d(C+K_1(z_j,\zeta_j)) \asymp |\Pi_{j=1}^d(C+K_1(z_j,\zeta_j))|\,. \tag{3.19}$$

Therefore, for $\vec{s}=\vec{1}$ by modifying the kernel we can achieve (3.15) without changing the class of Carleson measures. This is shown by (3.17). This means that without changing the set of embedding measures we can equivalently replace inequality (3.12) by (3.14). This reasoning works for $\vec{s}=\vec{1}$ and any dimension d.

Weighted Hardy-Sobolev spaces

Now $\vec{s}=(s_j)_{j=1}^d, 0<s_j\le 1$, but $\vec{s}\ne\vec{1}$. We are unable to repeat the trick that was successful in the previous section. In fact, for $K_s=(1-z\bar{\zeta})^{s-1}$ with $0<s<1$ (3.18) does not hold, the imaginary part will not be bounded, and so the previous reasoning with adding a large constant to each kernel of each variable does not work.

However, to reduce the analytic embedding (3.12) to dyadic embedding on multi-tree we seem to really need to show that (3.12) implies (3.14) (the converse implication being always trivial).

Here we have only partial results, namely for the case when

$$1-\epsilon(d) \le s_j \le 1 \tag{3.20}$$

for $\epsilon(d)$ sufficiently close to 0.

We just notice that $1-z\bar{\zeta}$ lies in the right half-plane if $z,\zeta\in\mathbb{D}$, and so $(1-z\bar{\zeta})^{\epsilon}$ lies in the cone $C_\epsilon = \{u+iv, u\ge 0, |v|\le u\cdot\tan\pi\epsilon\}$. Therefore, for every $s_j\in(1-\epsilon,1)$,

$$|\Im K_{s_j}(z_j,\zeta_j)|\le \tan\pi\epsilon\cdot\Re K_{s_j}(z_j,\zeta_j)\,.$$

This implies that if ϵ is sufficiently small (depending on the dimension d) then (3.15) holds, which, as we have already explained gives us the equivalence of (3.12) and (3.14).

From (3.14) we will now proceed to conclude that dyadic embedding holds. Then we will explain why dyadic embedding implies (3.14), thus closing the circular argument.

Remark. The argument of Section 3.1.2 fails for a number of reasons, if even one of the parameters s_j becomes zero. For example, if some s_j vanish, we have 'a phase transition' in the kernel, and (3.15) stops to be true in general. This explains the special role of Hardy spaces on the polydisc. For the classical Hardy space on the polydisc one can still make a connection between Carleson embedding and Hardy inequality, only now the direct embedding (3.7a) should be used in place of the dual (3.7b), and the roles of μ and w are reversed.
However Chang–Fefferman theory gives the characterization of embedding measures in the d-*harmonic* space
$h^2(\mathbb{D}^d)=H^h_0(\mathbb{D}^d)$. As, obviously, the Hardy space of holomorphic functions in the polydisc is such that $H^2(\mathbb{D}^d)\subset h^2(\mathbb{D}^d)$, the Chang–Fefferman theory gives the sufficient condition for measure to be an embedding measure for the Hardy class, but whether it is a necessary condition (we believe it is) is not known outside the classical case $d=1$. If the influential paper [34] were correct, then its proof can be modified to give this necessity, but unfortunately the note [95] indicated a counterexample to the reasoning (but not to the result) of [34].

Transition to the dyadic setting

We define a canonical map $\Lambda_*: Meas^+(\mathbb{D}^d)\to Meas^+(T^d)$ given by

$$\Lambda^*\nu(\alpha)=\tilde\nu(\alpha):=\nu(Q_\alpha). \tag{3.21}$$

For boundary measures we also define the natural push-forward from $Meas^+(\partial T)^d$ to $Meas^+(\mathbb{T}^d)$,

$$\Lambda_*\mu(F):=\mu(\Lambda^{-1}(F)),\quad F\subset\mathbb{T}^d. \tag{3.22}$$

Similarly, given a function $g\in L^2(\mathbb{D}^d,d\nu)$ we write

$$\tilde g(\alpha):=\frac{1}{\mu(Q_\alpha)}\int_{Q_\alpha} g(z)\,d\mu(z),\quad \alpha\in T^d, \tag{3.23}$$

and we put $\tilde g(\alpha):=0$, if $\nu(Q_\alpha)=0$.

Define

$$
\begin{aligned}
w_{s_j}(\alpha_j) &:= 2^{(1-s_j)(|\alpha_j|-1)}, \quad \alpha_j \in T,\\
w_{\vec{s}}(\alpha) &:= \prod_{j=1}^{d} w_{s_j}(\alpha_j), \quad \alpha = (\alpha_1, \dots, \alpha_d) \in T^d,
\end{aligned}
\tag{3.24}
$$

where $|\alpha_j| = \#\mathcal{P}(\alpha_j)$ is the depth of α_j in T (the root is assumed to have depth 1).

In order to proceed we need the following Lemma.

Lemma 3.1.1 *For any $\alpha, \beta \in T^d$ we have*

$$
|K_{\vec{s}}(z,\zeta)| \approx \tilde{k}_{\vec{s}}(\alpha,\beta) := w_{\vec{s}}\left(\mathcal{P}_{\mathfrak{G}}(\alpha) \bigcap \mathcal{P}_{\mathfrak{G}}(\beta)\right) = \sum_{\gamma \in \mathcal{P}_{\mathfrak{G}}(\alpha) \bigcap \mathcal{P}_{\mathfrak{G}}(\beta)} w_{\vec{s}}(\gamma), \quad z \in Q_\alpha,\ \zeta \in Q_\beta.
$$

Same equivalence holds on the boundaries as well

$$
|K_{\vec{s}}(z,\zeta)| \approx \tilde{k}_{\vec{s}}(\tau,\omega) := \sum_{\gamma \in \mathcal{P}_{\mathfrak{G}}(\tau) \bigcap \mathcal{P}_{\mathfrak{G}}(\omega)} w_{\vec{s}}(\gamma), \quad z = \Lambda(\tau),\ \zeta = \Lambda(\omega),\ \ \tau, \omega \in (\partial T)^d.
$$

Proof. Since $w_{\vec{s}}$ is a product weight, it is enough to consider points on a tree $\overline{T}$ and unit disc $\overline{\mathbb{D}}$ respectively. Recall that the $\mathfrak{G}$-structure adds two additional elements of each rank to the usual predecessor set $\mathcal{P}(\tau)$. Now fix any two points $\tau, \omega \in \overline{T}$ (boundary or inner, it does not really matter), we always may assume they are of the same depth in $\overline{T}$. Indeed, let, say $|\omega| > |\tau|$, consider then $\omega' \in \mathcal{P}(\omega)$ with $|\omega| = |\omega'|$. It is elementary to verify that

$$
\mathcal{P}_{\mathfrak{G}}(\tau) \bigcap \mathcal{P}_{\mathfrak{G}}(\omega) = \mathcal{P}_{\mathfrak{G}}(\tau) \bigcap \mathcal{P}_{\mathfrak{G}}(\omega').
$$

Same happens and on the continuous side, where

$$
|K_s(z,\zeta)| \approx \left| K_s\left(z, \zeta\frac{|z|}{|\zeta|}\right)\right|, \quad \frac{1}{2} \le |z| \le |\zeta|.
$$

Now, for a point $z \in \mathbb{T}$, one has

$$
\begin{aligned}
1 + \log\frac{1}{|z-1|} &\approx \sum_{j=1}^{2+\log_2\frac{1}{|z-1|}} |\tilde{I}_j(z)|^0,\\
1 + \frac{1}{|z-1|^{1-s}} &\approx \sum_{j=1}^{2+\log_2\frac{1}{|z-1|}} |\tilde{I}_j(z)|^{1-s}, \quad 0 < s < 1,
\end{aligned}
$$

where $\tilde{I}_j(z) = \{e^{2\pi i\theta} : \ \theta \in [-2^{-j}, 2^{-j}+1)\}$ is the union of three dyadic subintervals of $\mathbb{T}$ rank j – the one that contains 0 and its two immediate neighbours. Such an interval also contains z as well, and it does not when j becomes larger than $2 - \log_2|z-1|$ (we are being a bit imprecise here). It follows immediately that

$$1+\log\frac{1}{|z-\zeta|}\approx\sum_{j=1}^{\log_2\frac{1}{|z|}}|\tilde{I}_j(z,\zeta)|^0,$$
$$1+\frac{1}{|z-\zeta|^{1-s}}\approx\sum_{j=1}^{\log_2\frac{1}{|z|}}|\tilde{I}_j(z,\zeta)|^{1-s},\quad 0<s<1,$$

with $I_j(z,\zeta)$ being the union of three sequential dyadic intervals of the same rank that contain both ζ and z. Since $w_s(\alpha)=2^{|\alpha|(1-s)}$, and $J_\alpha=2^{-|\alpha|}$ (we recall that J_α is the base of the Carleson square corresponding to α), from this one can deduce that for $|\tau|=|\omega|$

$$w_1(\mathcal{P}_{\mathfrak{G}}(\tau)\bigcap\mathcal{P}_{\mathfrak{G}}(\omega))=|\tau|\approx\log\frac{1}{|z_\tau-z_\omega|},$$
$$w_s(\mathcal{P}_{\mathfrak{G}}(\tau)\bigcap\mathcal{P}_{\mathfrak{G}}(\omega))=\sum_{j=1}^{|\tau|}2^{(1-s)j}\approx 2^{(1-s)|\tau|}\approx\frac{1}{|z_\tau-z_\omega|^{1-s}},\quad 0<s<1,$$

where z_τ is the center of the Carleson box Q_τ for $\tau\in T$, or its circle image $\Lambda(\tau)$, if $\tau\in\partial T$. It remains to compare it with $|K_s(z,\zeta)|$ for $z\in Q_\tau$ and $\zeta\in Q_\omega$ (or some boundary images thereof) and we are done. $\square$

Applying Lemma 3.1.1 to the left-hand side of (3.14) we get

$$\begin{aligned}&\int_{\mathbb{D}}\int_{\mathbb{D}}g(z)g(\zeta)|K_{\vec{s}}(z,\zeta)|\,d\nu(z)d\nu(\zeta)=\\&\sum_{\alpha\in T^d}\sum_{\beta\in T^d}\int_{Q_\alpha}\int_{Q_\beta}g(z)g(\zeta)|K_{\vec{s}}(z,\zeta)|\,d\nu(z)\,d\nu(\zeta)\approx\\&\sum_{\alpha\in T^d}\sum_{\beta\in T^d}\int_{Q_\alpha}\int_{Q_\beta}g(z)g(\zeta)k_{\vec{s}}(\alpha,\beta)\,d\mu(z)\,d\mu(\zeta)=\\&\sum_{\alpha\in T^d}\sum_{\beta\in T^d}\tilde{g}(\alpha)\tilde{g}(\beta)k_{\vec{s}}(\alpha,\beta)\tilde{\nu}(\alpha)\tilde{\nu}(\beta)\end{aligned}$$

We attack the calculation from the end, letting $\sigma(\alpha)=\tilde{g}(\alpha)\tilde{\nu}(\alpha),\ \alpha\in T^d$ (recall that measures and functions on T^d are the same):

$$\begin{aligned}&\sum_{\gamma\in T^d}w_{\vec{s}}(\gamma)\left(\sum_{\alpha\in\mathcal{S}_{\mathfrak{G}}(\gamma)}\sigma(\alpha)\right)^2=\sum_{\gamma\in T^d}\sum_{\alpha,\beta\in\mathcal{S}_{\mathfrak{G}}(\gamma)}\sigma(\alpha)\sigma(\beta)w_{\vec{s}}(\gamma)=\\&\sum_{\alpha,\beta\in T^d}\sigma(\alpha)\sigma(\beta)\sum_{\gamma\in\mathcal{P}_{\mathfrak{G}}(\alpha)\bigcap\mathcal{P}_{\mathfrak{G}}(\beta)}w_{\vec{s}}(\gamma)=\\&\sum_{\alpha\in T^d}\sum_{\beta\in T^d}w_{\vec{s}}\left(\mathcal{P}_{\mathfrak{G}}(\alpha)\bigcap\mathcal{P}_{\mathfrak{G}}(\beta)\right)\cdot\sigma(\alpha)\sigma(\beta)=\\&\sum_{\alpha\in T^d}\sum_{\beta\in T^d}\tilde{g}(\alpha)\tilde{g}(\beta)\tilde{k}_{\vec{s}}(\alpha,\beta)\tilde{\nu}(\alpha)\tilde{\nu}(\beta).\end{aligned}$$

Define

$$k_{\vec{s}}(\alpha,\beta) := w(\vec{s})(\mathcal{P}(\alpha)\bigcap\mathcal{P}(\beta)) = \sum_{\gamma\geq\alpha,\beta} w(\gamma), \quad \alpha,\beta\in T^d,$$

and repeat the calculation above but with $\mathcal{P}$ instead of $\mathcal{P}_{\mathfrak{G}}$. We obtain

$$\sum_{\gamma\in T^d} w_{\vec{s}}(\gamma)\left(\sum_{\alpha\in\mathcal{S}(\gamma)}\sigma(\alpha)\right)^2 = \sum_{\alpha\in T^d}\sum_{\beta\in T^d}\tilde{g}(\alpha)\tilde{g}(\beta)k_{\vec{s}}(\alpha,\beta)\tilde{\nu}(\alpha)\tilde{\nu}(\beta).$$

The successor set formula (3.4)– the one that tells that $\mathcal{S}_{\mathfrak{G}}$ can be covered by boundedly many sets $\mathcal{S}$ and that each point would be used a bounded amount of times – implies that

$$\sum_{\gamma\in T^d} w_{\vec{s}}(\gamma)\left(\sum_{\alpha\in\mathcal{S}_{\mathfrak{G}}(\gamma)}\sigma(\alpha)\right)^2 = \sum_{\gamma\in T^d}\sigma\left(\mathcal{S}_{\mathfrak{G}}(\gamma)\right)^2 w_{\vec{s}}(\gamma) \approx$$

$$\sum_{\gamma\in T^d}\sigma\left(\mathcal{S}(\gamma)\right)^2 w_{\vec{s}}(\gamma) = \sum_{\gamma\in T^d} w_{\vec{s}}(\gamma)\left(\sum_{\alpha\in\mathcal{S}(\gamma)}\sigma(\alpha)\right)^2.$$

Combining the estimates above we see that (3.14) is equivalent to

$$\sum_{\alpha\in T^d}\int_{Q_\alpha} g^2(z)\,d\nu(z) \gtrsim \sum_{\alpha\in T^d}\sum_{\beta\in T^2}\tilde{g}(\alpha)\tilde{g}(\beta)k_{\vec{s}}(\alpha,\beta)\tilde{\nu}(\alpha)\tilde{\nu}(\beta),$$

where $\tilde{g}$ and $\tilde{\nu}$ are defined as in (3.21), (3.23). We see that if g is constant on the boxes Q_α and ν is Carleson measure for $\mathcal{H}_{\vec{s}}$, then

$$\|\tilde{g}\|^2_{L^2(T^d,d\tilde{\nu})} \gtrsim \sum_{\alpha,\beta\in T^d}\tilde{g}(\alpha)\tilde{g}(\beta)k_{\vec{s}}(\alpha,\beta)\tilde{\mu}(\alpha)\tilde{\mu}(\beta) = \sum_{\gamma\in T^d} w_{\vec{s}}(\gamma)\left(\sum_{\beta\leq\gamma}\tilde{g}(\beta)\tilde{\nu}(\beta)\right)^2. \tag{3.25}$$

On the other hand, by Jensen's inequality,

$$\sum_{\alpha\in T^d}\int_{Q_\alpha} g^2(z)\,d\nu(z) \geq \sum_{\alpha\in T^d}\tilde{g}^2(\alpha)\tilde{\nu}(\alpha),$$

so if (3.25) holds for any non-negative $\tilde{g}$ in $L^2(T^2, d\tilde{\nu})$, then ν is Carleson.

But, clearly, what we have above is just (3.7b) with $w = w_{\vec{s}}$ and $\mu = \tilde{\nu}$ and $\varphi = \tilde{g}$. The only difference is that (3.7b) has a closed d-tree $\overline{T}^d$ in it, but this is not really important, as we will see in a few lines.

Proposition 3.1.1 *Assume that ν is a measure on $\overline{\mathbb{D}}^d$ – a measure on $\mathbb{C}^d$ with* $\operatorname{supp}\nu\subset\overline{\mathbb{D}}^d$*. For $0<r<1$ let ν_r be a dilated measure supported on $r\overline{\mathbb{D}}^d$,*

$$d\nu_r(z) := d\nu(r^{-1}z), \quad z\in\mathbb{C}^d.$$

Then ν is Carleson for $\mathcal{H}_{\vec{s}}$ if and only if the measure ν_r is Carleson and the constant C_{ν_r} in (3.8) *does not depend on r.*

Proof. It is almost immediate. Since

$$\int_{\overline{\mathbb{D}}^d} |f(r\zeta)|^2\, d\nu(\zeta) = \int_{r\overline{\mathbb{D}}^d} |f(z)|^2\, d\nu_r(z) = \int_{\overline{\mathbb{D}}^d} |f(z)|^2\, d\nu_r(z),$$

we see that, if ν is Carleson, then, clearly ν_r is with C_{ν_r} independent of r. On the other hand, if $C_{\nu_r} < C < \infty$, then $C_\nu = \sup_{r<1} C_{\nu_r} < C$ and also $\lim_{r\to 1} \int_{\overline{\mathbb{D}}^d} |f(r\zeta)|^2\, d\nu(\zeta)$ clearly exists as well. $\square$

Remark. We stress that all of our Carleson conditions do not depend on r or on the depth of a d-tree.

We, therefore, obtain the following Theorem (Theorem I.9)

Theorem 3.1.1 *Let $\vec{s} = (s_1, \dots, s_d)$, $s_j \in (0,1], j = 1, \dots, d$, $d \geq 1$, such that all s_i are sufficiently close to 1: $1 - s_j \leq \varepsilon_d$, for a certain positive absolute $\varepsilon = \varepsilon(d)$ and $j = 1, \dots, d$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$, if and only if $(w_{\vec{s}}, \tilde{\nu})$ is a trace weight-measure pair for $\overline{T}^d$,*

$$\sum_{\alpha\in T^d} (\mathbf{I}^*\psi\tilde{\nu})^2(\alpha) w_{\vec{s}}(\alpha) \leq C \int_{T^d} \psi^2 d\tilde{\nu}, \quad \forall \psi \in L^2(\overline{T}^d, \tilde{\nu})\,. \tag{3.26}$$

Here $\tilde{\nu} = \Lambda^*\nu$ is the discrete image of ν on $\overline{T}^d$.

3.2 Weighted capacity on T^d and Bessel multiparametric capacity on $\mathbb{D}^d$

To get to the Carleson embedding conditions for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$ with $d = 1, 2, 3$, it remains to use the discrete Theorem 2.1.1. However it is convenient to consider the continuous reformulation of the conditions. For this we have chosen the subcapacitary one (2.3a). In this section we will show the equivalence between T^d-subcapacitary and d-parametric Bessel subcapacitary conditions, and we will finally deduce Theorem I.7.

Recall that for a given measure on $\mathbb{T}^d = (\partial\mathbb{D})^d$ and $\vec{s} \in (0,1]^d$ we can define the multi-parametric $\vec{s}$-Bessel potential

$$\mathbf{U}^{\mu}_{\vec{s}}(z) = \int_{\mathbb{T}^d} \mathbb{K}_{\vec{s}}(z,\zeta)\, d\mu(\zeta), \quad z \in \mathbb{T}^d, \tag{3.27}$$

where

$$\begin{aligned}
\mathbb{K}_{\vec{s}}(z,\zeta) &= \prod_{k=1}^{d} \mathbb{K}_{s_k}(z_k,\zeta_k), \\
\mathbb{K}_{s}(z_k,\zeta_k) &= \frac{1}{|z_k - \zeta_k|^{1-s}}, \quad s<1, \\
\mathbb{K}_{1}(z_k,\zeta_k) &= \log\left(\frac{1}{|z_k - \zeta_k|}\right).
\end{aligned} \tag{3.28}$$

We aim to prove that capacities on the polytorus and the boundary of d-tree are comparable.

Theorem 3.2.1 *Let $w := w_{\vec{s}} : T^d \to \mathbb{R}_+$ be an exponential product weight defined as in (3.24), and assume that $E \subset (\partial T)^d$ is a compact set. Then the respective capacities of F and its polytorus image $F := \Lambda(E)$ are equivalent*

$$\mathrm{Cap}_{\vec{s}}(F) \approx \mathrm{Cap}_{w_{\vec{s}}}(E), \tag{3.29}$$

and the constant depends only on d and $\vec{s}$.

Proof. Our argument runs as follows.
First we show that for any (Radon non-negative) measure μ on $(\partial T)^d$ its ($w_{\vec{s}}$) potential is comparable to the ($\vec{s}$-Bessel) potential of its continuous image $\Lambda_* \mu =: \nu$. This immediately leads to the equivalence of their respective energies. Therefore, the equilibrium measure for E is almost equilibrium for F. On the other hand, if ν is equilibrium for F, then its pull-back $\Lambda^* \nu$ again still has comparable energy to ν, hence $\Lambda^* \nu$ is almost equilibrium for E. Let us elaborate.

The equivalence of potentials is already given by Lemma 3.1.1. Indeed, for $\zeta, z \in \mathbb{T}^d$ the potential kernel $\mathbb{K}_{\vec{s}}$ and the reproducing kernel $K_{\vec{s}}$ are the same. Let μ be a measure supported on E (equilibrium measures are just like that). Similar to the argument in the previous section we can write

$$\begin{aligned}
&\mathcal{E}_{w_{\vec{s}}}[\mu] = \int_E \mathbf{V}^{\mu}_{w_{\vec{s}}}\, d\mu = \sum_{\gamma \in T^d} w_{\vec{s}}(\gamma) \left(\int_{\partial \mathcal{S}(\gamma)} d\mu \right)^2 \\
&\approx \sum_{\gamma \in T^d} w_{\vec{s}}(\gamma) \left(\int_{\partial \mathcal{S}_{\mathfrak{G}}(\gamma)} d\mu \right)^2 = \int_{(\partial T)^2 d} \left(\sum_{\gamma \in \mathcal{P}_{\mathfrak{G}}(\tau) \cap \mathcal{P}_{\mathfrak{G}}(\omega)} w_{\vec{s}}(\gamma) \right) d\mu(\tau)\, d\mu(\omega) = \\
&\int_{E^2} k_{\vec{s}}(\tau, \omega)\, d\mu(\tau) d\mu(\omega).
\end{aligned}$$

Hence instead of the usual discrete potential we can look at $\mathfrak{G}$-extended ones. Given $F \ni \zeta = \Lambda(\omega),\ \omega \in E$, we have

$$\{\tau \in E : k_{\vec{s}}(\tau, \omega) \geq Ct\} \subset E \cap \Lambda^{-1}\left(\{z \in F : K_{\vec{s}}(z, \zeta) \geq t\}\right) \subset \left\{\tau \in E : k_{\vec{s}}(\tau, \omega) \geq \frac{T}{C}\right\}$$

for some absolute constant C, since by Lemma 3.1.1

$$k_{\vec{s}}(\tau, \omega) \approx K_{\vec{s}}(\Lambda(\tau), \Lambda(\omega)).$$

Hence, due to $\operatorname{supp} \mu \subset E$,

$$\begin{aligned}
&\int_E k_{\vec{s}}(\tau, \omega)\, d\mu(\tau) = \int_0^\infty \mu\left(E \cap \{k_{\vec{s}}(\tau, \omega) \geq t\}\right) dt \approx \\
&\int_0^\infty \nu\left(\{K_{\vec{s}}(z, \zeta) \geq t\}\right) dt = \int_F \mathbb{K}_{\vec{s}}(z, \zeta)\, d\nu(z).
\end{aligned}$$

The equivalence of energies of μ and $\nu = \Lambda_*\nu$ follows immediately, and we see that

$$\mathrm{Cap}_{W_{\vec{s}}}(E) \lesssim \mathrm{Cap}_{\vec{S}}(\Lambda(E)),$$

since $\Lambda_*\mu$ is admissible for F after multiplication by a suitable constant.

To get the reverse estimate we need to move measures in the other direction, from $\mathbb{T}^d$ to $(\partial T)^d$. In order to do so, for ν on $\mathbb{T}^d$ and Borel $\tilde{E} \subset (\partial T)^d$ we write

$$\Lambda^*\nu(\tilde{E}) := \int_{\mathbb{T}^d} \frac{\#\{\tilde{E} \cap \{\Lambda^{-1}(z)\}\}}{\#\{\Lambda^{-1}(z)\}}\, d\nu(z). \tag{3.30}$$

This is a well-defined mapping, for details about the correctness, measurability etc. see [2] (or [6] for 1D argument). Since the integrand in (3.30) is bounded from above and below, we see that, $\Lambda^*\Lambda_*\mu \approx \mu$ and $\Lambda_*\Lambda^*\nu \approx \nu$. Actually, it is easy (but rather tedious) to show that these are equalities for measures with finite energies, which are exactly those that we consider. However even the two-sided estimate is good enough, since we do not keep track of constants meticulously. Now repeating the previous argument with ν – equilibrium measure for F, we see that $\Lambda^*\nu$ is (after multiplication by a constant) admissible for E, and they have comparable energies. This proves the lemma. $\square$

To completely move our discrete subcapacitary condition to the polytorus we have to make one more observation (see also [104, Section 3]).

Theorem 3.2.2 *Let $E \subset \overline{T}^d$ and $w = w_{\vec{s}}$ be the product weight generated by $\mathcal{H}_{\vec{s}}$. Then the capacity of E and its boundary projection are equivalent,*

$$\mathrm{Cap}_w(E) \approx \mathrm{Cap}_w(\partial\mathcal{S}(E)),$$

where $\partial\mathcal{S}(E) = \{\omega \in (\partial T)^d : \mathcal{P}(\omega) \cap E \neq \emptyset\}$.

Remark. This is a more precise version of Proposition 1.1.3.

Proof. One direction is trivial, since by monotonicity of $\mathbf{I}_w$ one has $\mathbf{V}_w^\mu(\omega) \geq \mathbf{V}_w^\mu(\beta)$, where $\omega \in \partial\mathcal{S}(E)$ is a descendant of $\beta \in E$ for any measure μ. To obtain the reverse estimate we aim to show that the potentials of a measure μ and its 'boundary projection' are comparable *everywhere on* $\overline{T}^d$. This would imply that admissible measures for E and $\partial\mathcal{S}(E)$ are the same, after multiplication by a constant, and that their energies are comparable. S before, that settles the equivalence of capacities.

We start with defining a boundary projection of a measure. For this we need to introduce the Lebesgue measure on $(\partial T)^d$ – we just put it to be the image of the normalized Lebesgue measure on the polytorus.

$$m_d := \frac{1}{(2\pi)^d}\Lambda^*(dx_1 \dots dx_d).$$

In other words, for $\partial\mathcal{S}(\alpha) \subset (\partial T)^d$ we set $m_d(\partial\mathcal{S}(\alpha)) := 2^{-|\alpha|}$, where $|\alpha| = \prod_{k=1}^d |\alpha_k|$ and $|\alpha_k| = \#\mathcal{P}(\alpha) - 1$. Then we extend it to Borel subsets of $(\partial T)^d$ in the usual way.

Now assume that $\mu = \mu_\alpha$ is a mass at a singleton $\operatorname{supp}\mu_\alpha = \alpha \in T^d$. Then its projection $(\mu_\alpha)_b$ is defined as follows

$$(\mu_\alpha)_b(E) = |\mu_\alpha| m_d \left(E \bigcap \partial\mathcal{S}(\alpha)\right).$$

Essentially we take the mass μ_α and distribute it in a uniform way over $\partial\mathcal{S}(\alpha)$ – the shadow of α over (∂T^d). Next, if μ is a measure on T^d (so, sans boundary, just a collection of singleton masses), we write

$$\mu_b = \sum_{\alpha\in T^d} (\mu_\alpha)_b, \quad \mu_\alpha := \mu \cdot \mathbb{1}_\alpha.$$

Finally, if μ is supported on some portion of ∂T^d, say on $(\partial T)^{k_0} \times T^{d-k_0}$, then we do the same thing by leaving the $(\partial T)^{k_0}$ coordinate variables intact and dumping the rest down to $(\partial T)^{d-k_0}$ singleton by singleton.

$$d\mu_b(\omega_1,\dots,\omega_{k_0},\omega_{k_0+1},\dots,\omega_d) = \sum_{\alpha\in T^{d-k_0}} \left(d\mu(\omega_1,\dots,\omega_{k_0},\cdot,\dots,\cdot)\cdot \mathbb{1}_{(\partial T)^{k_0}\times\{\alpha\}}\right)_b (\omega_{k_0}+1,\dots,\omega_d).$$

If μ is supported on $(\partial T)^d$, we leave it as it is, $\mu_b = \mu$.
As a result, since every μ on $\overline{T}^d$ can be decomposed into the respective inner/boundary/distinguished boundary parts, we obtain a measure μ_b supported on $(\partial T)^d$ with the following important property

$$\mu(\overline{T}^d) = \mu_b((\partial T)^d).$$

Next step is to compare $w_{\vec{s}}$ potentials of μ and μ_b. Due to linearity it is enough to make estimates only for singleton masses $\mu_\alpha = \mu\cdot\mathbb{1}_\alpha$, since μ – and hence its potential – is just a sum of such terms. A singleton $\{\alpha\}$, however is a product set, as is its boundary projection $\partial\mathcal{S}(\alpha)$, so we have

$$\mathbf{V}_{w_{\vec{s}}}^{\mu_\alpha} = \prod_{k=1}^d \mathbf{V}_{w_{s_k}}^{\mu_{\alpha_k}}, \quad \mathbf{V}_{w_{\vec{s}}}^{(\mu_\alpha)_b} = \prod_{k=1}^d \mathbf{V}_{w_{s_k}}^{(\mu_{\alpha_k})_b},$$

so it is enough to do the estimate on T. So, fix any $\gamma \in T$ and a singleton mass σ at γ, and $0 < s \le 1$, and also a point $\tau \in \overline{T}$. There are two options, depending on whether τ is a descendant of γ. If it is not, then the potentials coincide,

$$\mathbf{V}_{w_s}^{\sigma}(\tau) = \sum_{\beta\ge\tau} \mathbf{I}^*\sigma(\beta) w_s(\beta) = \sum_{\beta\ge\tau} \mathbf{I}^*\sigma_b(\beta) w_s(\beta) = \mathbf{V}_{w_s}^{\sigma_b}(\tau),$$

since, clearly, $\mathbf{I}^*\sigma = \mathbf{I}^*\sigma_b$ above γ and both are zero between $\gamma\wedge\tau$ and τ. On the other hand, if $\tau \le \gamma$, then the uniform distribution of the boundary projection mass implies

$$\begin{aligned}
&\mathbf{V}_{w_s}^{\sigma_b}(\tau) = \sum_{\beta\ge\tau} \mathbf{I}^*\sigma_b(\beta) w_s(\beta) = \sum_{\beta\ge\gamma} \mathbf{I}^*\sigma_b(\beta) w_s(\beta) + \sum_{\gamma>\beta\ge\tau} \mathbf{I}^*\sigma_b(\beta) w_s(\beta) = \\
&\mathbf{V}_{w_s}^{\sigma}(\tau) + \sum_{\gamma>\beta\ge\tau} \mathbf{I}^*\sigma_b(\gamma)\cdot 2^{|\gamma|-|\beta|} w_s(\beta) \le \mathbf{V}_{w_s}^{\sigma}(\tau) + C\mathbf{I}^*\sigma_b(\gamma) w_s(\gamma) \le (C+1)\mathbf{V}_{w_s}^{\sigma}(\tau),
\end{aligned}$$

since $\mathbf{V}^{\sigma}_{w_s}(\tau) = \mathbf{V}^{\sigma}_{w_s}(\gamma)$ and

$$\sum_{\beta \le \gamma} 2^{|\gamma|-|\beta|} w_s(\beta) \le \sum_{k \ge 0} 2^{-k+(1-s)k} w_s(\gamma) \le C w_s(\gamma).$$

We see that

$$\mathbf{V}^{\mu}_{w_{\vec{s}}} \approx \mathbf{V}^{\mu_b}_{w_{\vec{s}}} \quad \text{and} \quad \mathcal{E}_{w_{\vec{s}}}[\mu] \approx \mathcal{E}_{w_{\vec{s}}}[\mu_b],$$

since $|\mu| = |\mu_b|$ by construction. Consider now the equilibrium measures of E and $\partial\mathcal{S}(E)$, which we denote by μ_E and ν respectively, and let $\mu_b := (\mu_E)_b$. Since $\mathbf{V}^{\nu}_{w_{\vec{s}}} \ge 1$ on the support of μ_b, we see that their mutual energy dominates $|\mu_b|$ which is actually proportional to $\mathcal{E}_{w_{\vec{s}}}[\mu_b]$ by equivalence of potentials. Hence for any C one has

$$\begin{aligned} 0 &\le \mathcal{E}_{w_{\vec{s}}}[\mu_b] - 2C\mathcal{E}_{w_{\vec{s}}}[\mu_b, \nu] + C^2\mathcal{E}_{w_{\vec{s}}}[\nu] \le \\ &\mathcal{E}_{w_{\vec{s}}}[\mu_b] - C|\mu_b| + C(C\mathcal{E}_{w_{\vec{s}}}[\nu] - |\mu_b|). \end{aligned}$$

For large enough C the first term above becomes negative, hence the second is positive, which means that

$$\cap_{w_{\vec{s}}}(\partial\mathcal{S}(E)) = \mathcal{E}_{w_{\vec{s}}}[\nu] \ge \frac{1}{C}\mathcal{E}_{w_{\vec{s}}}[\mu_b] \approx \cap_{w_{\vec{s}}}(E).$$

We are done. $\square$

Now we are ready to finish the proof of Theorem I.7. Let us recall its statement.

Theorem 3.2.3 *Let $\vec{s} = (s_1, \ldots, s_d)$, $s_j \in (0,1], j = 1, \ldots, d$, $1 \le d \le 3$, such that all s_i are sufficiently close to 1: $1 - s_j \le \varepsilon_d$, for a certain positive absolute $\varepsilon = \varepsilon(d)$ and $j = 1, \ldots, d$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}_{\vec{s}}(\mathbb{D}^d)$, if and only if one of the following conditions holds true*

$$\nu(T(E)) \lesssim \mathrm{Cap}_{\vec{s}}(E), \quad E \subset \mathbb{T}^d \tag{3.31a}$$

$$\textstyle\sum_{R \subset E} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } E, \tag{3.31b}$$

$$\textstyle\sum_{R \subset Q} \nu^2(T(R)) w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } Q. \tag{3.31c}$$

Here Q, R are dyadic rectangles on the (poly) torus $\mathbb{T}^d$, and $T(Q)$ is the usual tent area above Q, while E is any finite union of such rectangles, and $T(E)$ is the union of respective tents.

Proof. By invoking Theorem 3.1.1 we see that ν is Carleson if and only if its discrete image $\tilde{\nu} := \Lambda^*\nu$ combined with the weight $w_{\vec{s}}$ forms a trace weight-measure pair for T^d. Theorem 2.1.1 provides testing conditions for such a pair. In particular, (3.31b) follows from $[w_{\vec{s}}, \tilde{\nu}]_C \gtrsim [w_{\vec{s}}, \tilde{\nu}]_{CE}$ and (3.31c) from $[w_{\vec{s}}, \tilde{\nu}]_B \gtrsim [w_{\vec{s}}, \tilde{\nu}]_{CE}$. The subcapacitary condition (3.31a) is obtained by $[w_{\vec{s}}, \tilde{\nu}]_{SC} \gtrsim [w_{\vec{s}}, \tilde{\nu}]_{CE}$, Theorems 3.2.1 and 3.2.2. Indeed, writing $T(E)$ as a union of Q_α's and some boundary fragments, we see that $\nu(T(E)) \approx \tilde{nu}(\tilde{E})$, where $\tilde{E} \subset \overline{T}^d$ is the union of those α's and Λ-preimages of boundary fragments. Since $\nu \approx \Lambda^*\Lambda_*\nu$ (again, they are actually equal, but we do not need that), we have that due to Theorem 3.2.1 the condition (3.31a) is equivalent to

$$\tilde{\nu}(\tilde{E}) \lesssim \mathrm{Cap}_{w_{\vec{s}}}\, \partial\mathcal{S}(\tilde{E}), \quad \forall \tilde{E} = \mathcal{S}(\cup_{k=1}^{M}\{\alpha_k\}),\ \alpha_k \in T^d,$$

but the right-hand side of the above is equivalent to $\mathrm{Cap}_{w(\vec{s})}(\tilde{E})$ by Theorem 3.2.2, and we get the discrete subcapacitary condition, which is equivalent to the embedding. □

If we consider the *harmonic* spaces $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d)$ for $d = 1, 2, 3$, then we do need $\vec{s}$ to be close to $\vec{1}$.

Theorem 3.2.4 *Let $\vec{s} = (s_1, \dots, s_d)$, $s_j \in (0, 1], j = 1, \dots, d$, $1 \le d \le 3$. Let ν be a non-negative measure in $\overline{\mathbb{D}}^d$. Then embedding operator $id : \mathcal{H}^h_{\vec{s}}(\mathbb{D}^d) \to L^2(\overline{\mathbb{D}}^d, \nu)$ is bounded, i.e. ν is Carleson for $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d)$, if and only if one of the following conditions holds true*

$$\nu(T(E)) \lesssim \mathrm{Cap}_{\vec{s}}(E), \quad E \subset \mathbb{T}^d \tag{3.32a}$$

$$\textstyle\sum_{R\subset E} \nu^2(T(R))w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } E, \tag{3.32b}$$

$$\textstyle\sum_{R\subset Q} \nu^2(T(R))w_{\vec{s}}(R) \lesssim \nu(T(Q)), \quad \textit{for any } Q. \tag{3.32c}$$

Proof. It is exactly the same as in the previous Theorem, only now we observe that we do no need to care about real and imaginary parts of the reproducing kernels. Indeed, the harmonic reproducing kernel $K^h_{\vec{s}}(z, \zeta)$ has the following nice estimate

$$K^h_{\vec{s}}(z, \zeta) \approx |K_{\vec{s}}(z, \zeta)|,$$

since $H^h_{s_j}(\mathbb{D}) = H_{s_j}(\mathbb{D}) \bigoplus \overline{H_{s_j}(\mathbb{D})}$ and $\mathcal{H}^h_{\vec{s}}(\mathbb{D}^d) = \bigotimes_{j=1}^d H^h_{s_j}(\mathbb{D})$. Since the kernel is real now, we do not need the restrictions on s_j, and the rest of the proof runs verbatim. □

Chapter 4 Growth classes of harmonic functions: wavelet decomposition

4.1 Auxiliary results and facts about MRA

4.1.1 Main lemma

In this subsection we prove some lemmas. The first one is an elementary estimate.

Lemma 4.1.1 *Let $g \in L^1(\mathbb{R}^d)$ and $|x|^{2M} g(x) \in L^1(\mathbb{R}^d)$ for some integer $M > \frac{d}{2}$, if $\|\hat{g}\|_1 = C_1$ and $\|\Delta^M(\hat{g})\|_1 \le C_2$, then*

$$\|g\|_1 \le c(d, M) \left(C_1^{2M-d} C_2^d\right)^{1/(2M)}.$$

Proof. Since $g, \hat{g} \in L^1(\mathbb{R}^d)$ the inversion formula implies that $|g(x)| \le C_1$. Similarly, $|x|^{2M} g(x), \widehat{|x|^{2M} g(x)} \in L^1$ and $|g(x)| \le C_2 |x|^{-2M}$. Thus

$$\|g\|_1 \le c_d \left(C_1 \frac{R^d}{d} + C_2 \frac{R^{d-2M}}{2M - d} \right),$$

for any $R > 0$. Choosing R such that $R^{2M} = C_2 C_1^{-1}$ we obtain the required estimate. □

Let $P(x) = c_d (1+|x|^2)^{-(d+1)/2}$ be the standard Poisson kernel for the upper half-space, $P_{(s)}(x) = s^{-d} P(\frac{x}{s})$ as usual. The constant c_d is chosen such that $\hat{P}(\tau) = e^{-2\pi|\tau|}$. The next lemma (which is a special case of [98, Theorem 1]) will be used to divide by the Poisson kernel in the Fourier transforms and is inspired by [15]. We give a separate proof for the convenience of the reader.

Lemma 4.1.2 *There exists $\Phi \in L^1(\mathbb{R}^d)$ such that $\hat{\Phi}(\tau) = e^{2\pi|\tau|}$ when $|\tau| \le 1$.*

Proof. We note that $2\cosh(2\pi|\tau|)$ is a smooth function in $\mathbb{R}^d$ and we can find a smooth function Θ with compact support such that $\Theta(\tau) = 2\cosh(2\pi|\tau|)$ when $|\tau| \le 1$. Let Ξ be the inverse Fourier transform of Θ. Clearly, $\Xi \in L^1(\mathbb{R}^d)$. Finally we let $\Phi = \Xi - P$, then $\Phi \in L^1(\mathbb{R}^d)$ and $\hat{\Phi}(\tau) = \Theta(\tau) - e^{-2\pi|\tau|}$. For $|\tau| \le 1$ we have $\hat{\Phi}(\tau) = e^{2\pi|\tau|}$. □

Now we can give a preliminary estimate for a part of $u(\cdot, t) \in h_v^\infty$ with bounded frequencies. The next result is our Main lemma.

Lemma 4.1.3 *Let u be a bounded harmonic function in $\mathbb{R}_+^{d+1}$, and let $\sigma \in L^1(\mathbb{R}^d)$ be such that* supp $\hat{\sigma} \subset B_{\delta^{-1}}$. *Then*

$$\left| \int_{\mathbb{R}^d} u(x, t) \sigma(x) dx \right| \le C_d \|u\|_{L^\infty(\mathbb{R}_\delta^{d+1})} \|\sigma\|_1,$$

where C_d is a constant that depends on d only and $\mathbb{R}^{d+1}_\delta = \{(x,t) \in \mathbb{R}^{d+1} : \ t \geq \delta > 0\}$.

Proof. First, in the sense of distributions, we have

$$\int_{\mathbb{R}^d} u(x,t)\sigma(x)dx = \int_{\mathbb{R}^n} \widehat{u(\cdot,t)}(\tau)\hat{\sigma}(\tau)d\tau.$$

Now let Φ be the function in Lemma 4.1.2 and let $\Phi_\delta(x) = \delta^{-d}\Phi(\delta^{-1}x)$, then $\widehat{\Phi_\delta}(\tau) = \hat{\Phi}(\delta\tau)$ and $\|\Phi_\delta\|_1 = \|\Phi\|_1$. Since $\hat{\sigma}$ vanishes outside the ball $B_{1/\delta}$, we have

$$\hat{\sigma}(\tau) = \widehat{\Phi_\delta}(\tau)e^{-2\pi|\tau|\delta}\hat{\sigma}(\tau) = \widehat{\Phi_\delta}(\tau)\hat{P}_\delta(\tau)\hat{\sigma}(\tau).$$

Then

$$\left|\int_{\mathbb{R}^d} u(x,t)\sigma(x)dx\right| = \left|\int_{\mathbb{R}^d} u\widehat{(\cdot,t+\delta)}(\tau)\widehat{\sigma * \Phi_\delta}(\tau)d\tau\right| =$$

$$\left|\int_{\mathbb{R}^d} u(x,t+\delta)(\sigma * \Phi_\delta)(x)dx\right| \leq \|u\|_{L^\infty(\mathbb{R}^{d+1}_\delta)}\|\sigma\|_1\|\Phi_\delta\|_1$$

and the required estimate follows. □

4.1.2 Basic facts about smooth multiresolution analysis

We consider smooth (of order r) multiresolution approximation (MRA) in $\mathbb{R}^d$. Our main references here are the classical books by I. Daubechies [28] and by Y. Meyer [69]. We denote by $K(x,y)$ the kernel of the orthogonal projection onto V_0 and assume that

$$K(x,y) = \sum_{k\in\mathbb{Z}^d} \phi(x-k)\overline{\phi(y-k)},$$

where $\phi(x)$ and all its derivatives of order up to r decay faster at infinity than any power of x. We assume further that

$$\sum_{k\in\mathbb{Z}^d} \phi(x-k) = 1,$$

see [69, ch 2.10].

Let $K_j(x,y) = 2^{jn}K(2^jx, 2^jy)$. Further, let $D(x,y) = K_1(x,y) - K(x,y)$ and

$$D_j(x,y) = 2^{jd}D(2^jx, 2^jy) = K_{j+1}(x,y) - K_j(x,y).$$

Since we work with r-smooth MRA, we have

$$D(x,y) = \sum_{|\beta|=r} \partial_y^\beta D_\beta(x,y), \tag{4.1}$$

where D_β are Schwartz functions that satisfy $|D_\beta(x,y)| \leq C_m(1+|x-y|)^{-m}$, see [69, ch 2.8].

Further for any $f \in L^2(\mathbb{R}^d)$ we define

$$K_0 f(x) = \int_{\mathbb{R}^d} K(x,y) f(y) dy \quad \text{and} \quad D_j f(x) = \int_{\mathbb{R}^d} D_j(x,y) f(y) dy.$$

As usual $V_j = \{f(x) : f(2^{-j}x) \in V_0\}$. We will also need L^∞-version of these spaces,

$$V_0(\infty) = \{f(x) = \sum_{k \in \mathbb{Z}^d} a(k)\phi(x-k), \quad \{a(k)\} \in l^\infty(\mathbb{Z}^d)\},$$

and $V_j(\infty) = \{f(x) : f(2^{-j}x) \in V_0(\infty)\}$, $V_j \subset L^\infty(\mathbb{R}^d)$.
The following Bernstein's inequality holds, [69, ch 2.5]. There exists $C = C(\phi)$ such that

$$\|\partial^\beta f\|_\infty \leq C 2^{|\beta| j} \|f\|_\infty \tag{4.2}$$

for any $f \in V_j(\infty)$ and any multi-index β such that $|\beta| \leq r$.

4.1.3 Multiresolution approximation of the Poisson kernel

We need two estimates for smooth multiresolution approximation of Poisson kernels.

Lemma 4.1.4 *There exists C such that*

$$\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} (P_{(s)}(\xi - y) - P_{(s)}(\xi - x)) K(x,y) dx \right| dy \leq C s^{-r}$$

for any $\xi \in \mathbb{R}^d$.

Proof. We denote $f_\xi(x) = P_{(s)}(\xi - x)$, then

$$\int_{\mathbb{R}^d} (P_{(s)}(\xi - y) - P_{(s)}(\xi - x)) K(x,y) dx =$$
$$P_{(s)}(\xi - y) - \int_{\mathbb{R}^d} f_\xi(x) E(y,x) dx = f_\xi(y) - (K_0 f_\xi)(y).$$

We have

$$|f_\xi(x) - K_0 f_\xi(x)| \leq \sum_{j=0}^\infty \left| \int_{\mathbb{R}^d} D_j(x,y) f_\xi(y) dy \right|.$$

Integrating (4.1) we obtain

$$\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} D(x,y) f_\xi(y) dy \right| dx \leq C \sum_{|\beta| = r} \|\partial^\beta f_\xi\|_1 \leq C s^{-r}.$$

By rescaling we have also $D_j f(x) = (D_0 f_j)(2^j x)$, where $f_j(y) = f(y 2^{-j})$. Then

$$\|D_j f\|_1 = 2^{-dj} \|D f_j\|_1 \leq C 2^{-rj} \sum_{|\beta| = r} \|\partial^\beta f\|_1.$$

We apply this estimate for every $j = 0, 1, ...$ and $f = f_\xi$ and get

$$\int_{\mathbb{R}^d} |f_\xi(x) - K_0 f_\xi(x)| \leq \sum_{j=0}^{\infty} \|D_j f_\xi\|_1 \leq C \sum_{j=0}^{\infty} 2^{-rj} \sum_{|\beta|=r} \|\partial^\beta f_\xi\|_1 \leq C s^{-r}.$$

□

Lemma 4.1.5 *There exists C such that*

$$\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} (K_J(x,y) - K(x,y)) P_{(s)}(y-\xi)\, dy \right| d\xi \leq C s^{-r}$$

for any $x \in \mathbb{R}^d$ and any number $J \geq 1$.

Proof. In the notation of the last lemma we have

$$\int_{\mathbb{R}^d} (K_J(x,y) - K(x,y)) P_{(s)}(y-\xi)\, dy = \sum_{j=0}^{J-1} D_j f_\xi(x)$$

and by (4.1)

$$\int_{\mathbb{R}^d} |D_0 f_\xi(x)|\, d\xi \leq C \sum_{|\beta|=r} \|\partial^\beta P_{(s)}\|_1 \leq C s^{-r}.$$

Then similarly for $j \geq 1$

$$\int_{\mathbb{R}^d} |D_j f_\xi(x)|\, d\xi \leq C 2^{-rj} s^{-r}.$$

That concludes the proof of the lemma. □

4.2 Multiresolution analysis in growth spaces

4.2.1 Decomposition into blocks and direct estimates

Given a weight w that satisfies the doubling condition, we choose A large enough and define a sequence of integers $\{n_l\}_l$ such that $n_0 = 0$, $n_l > n_{l-1}$ and $w(2^{-n_l}) \in [A^l, A^{l+1})$. There exists m^* that depends on w only that satisfies

$$\frac{2^{-m^* n_l} w(2^{-n_l})}{2^{-m^* n_{l-1}} w(2^{-n_{l-1}})} < 1 - \varepsilon. \tag{4.3}$$

For weights w with some regularity we can satisfy the last inequality by choosing m^* such that $t^{m^*-1} w(t)$ is increasing.

We consider sufficiently smooth multiresolution approximation in $\mathbb{R}^d$, more precisely we choose r such that $r > m^* + d$, where m^* was chosen above. Instead of the usual dyadic partition $L^2(\mathbb{R}^d) = V_0 \cup \bigcup_{j\geq 1} V_j \setminus V_{j-1}$ we work with blocks adjusted to the weight v,

$$L^2(\mathbb{R}^d) = V_0 \cup \bigcup_{l\geq 1} V_{n_l} \setminus V_{n_{l-1}}.$$

We take wavelet series and combine all terms in generations $n_{l-1}+1, ..., n_l$ into one block; we work with bounded functions that do not belong to $L^2(\mathbb{R}^d)$ in general. Let $\{\psi_p\}_{p=1}^q$ be a collection of r-smooth rapidly decreasing functions such that $\{\psi_p(x-k), 1 \le p \le q, k \in \mathbb{Z}^d\}$ form an orthogonal basis for $V_1 \setminus V_0$, see [69, ch 3.1, 3.6] for details. For each $j \in \mathbb{Z}_+$ and $k \in \mathbb{Z}^d$ we have

$$\psi_{p,jk} = 2^{dj/2}\psi_p(2^j x - k).$$

In what follows we write

$$\langle f(y), g(y)\rangle = \int_{\mathbb{R}^d} f(y)\overline{g(y)}\, dy.$$

when the integral converges.
Let us recall Theorem I.15.

Theorem 4.2.1 *For any $u \in h_v^\infty(\mathbb{R}_+^{d+1})$ we define*

$$g_0(x,t) = \sum_{k\in\mathbb{Z}^d} \langle u(y,t), \phi(y-k)\rangle \phi(x-k), \quad \textit{and}$$

$$g_l(x,t) = \sum_{j=n_{l-1}+1}^{n_l} \sum_{p=1}^{q} \sum_{k\in\mathbb{Z}^d} \langle u(y,t), \psi_{p,jk}(y)\rangle \psi_{p,jk}(x), \quad l \ge 1.$$

Then

$$u(x,t) = \sum_{l=0}^{\infty} g_l(x,t), \quad g_l(\cdot,t) \in V_{n_l}(\infty) \quad \textit{and}$$

$$\|g_l(\cdot,t)\|_\infty \le C\|u\|_{v,\infty} w(2^{-n_l}), \quad l \ge 0, \tag{4.4}$$

where C depends on ϕ and A only.

Proof. Let as usual $K_j(x,y) = 2^{jd}K(2^j x, 2^j y)$, then

$$g_0(x,t) = \int_{\mathbb{R}^d} K(x,y)u(y,t)\, dt \quad \text{and}$$
$$g_l(x,t) = \int_{\mathbb{R}^d} (K_{n_l}(x,y) - K_{n_{l-1}}(x,y))u(y,t)\, dy.$$

Clearly, $g_l(\cdot,t) \in V_{n_l}(\infty)$. Moreover, for each t the function $u(\cdot,t)$ is uniformly continuous, thus the series $\sum_l g_l(x,t)$ converges to $u(x,t)$ uniformly on $\mathbb{R}^d$.

We take the Fourier transform of $K(x,\cdot)$ in second variable and denote it by $\hat{K}(x,\tau)$. (We never use the Fourier transform in the whole $\mathbb{R}^{2d}$.) We note that

$$|\hat{K}(x,\tau)| \le C(1+|\tau|)^{-m^*-d}$$

uniformly in x since $r > m^* + d$, the same inequality holds for all derivatives of $\hat{K}(x,\tau)$ in τ. Let further $\eta : \mathbb{R} \to \mathbb{R}$ be a smooth function with support in $[-1,1]$ that is equal to 1 on $[-1/2,1/2]$, $0 \le \eta \le 1$.

Then the Fourier transform of $K_N(x,\cdot)$, where $N \le n_l$, has the following partition

$$\widehat{K_N}(x,\tau) = \hat{K}(2^N x, 2^{-N}\tau) = \hat{K}(2^N x, 2^{-N}\tau)\eta(2^{-n_l}|\tau|) + \sum_{i=l+1}^{\infty} \hat{K}(2^N x, 2^{-N}\tau)(\eta(2^{-n_i}|\tau|) - \eta(2^{-n_{i-1}}|\tau|)) =: \zeta^0_{N,l}(x,\tau) + \sum_{i=l+1}^{\infty} \zeta_{N,i}(x,\tau).$$

First, we have

$$\|\zeta^0_{N,l}(x,\cdot)\|_1 \le \|\hat{K}(2^N x, 2^{-N}\cdot)\|_1 \le 2^{dN}\|\hat{K}(2^N x,\cdot)\|_1.$$

And also, since $N \le n_l$, we obtain $\|\Delta^M_\tau \zeta^0_{N,l}(x,\cdot)\|_1 \le C2^{(d-2M)N}$. Let further, $\sigma^0_{N,l} = \mathcal{F}^{-1}(\zeta^0_{N,l})$. Then Lemma 4.1.1 implies $\|\sigma^0_{N,l}(x,\cdot)\|_1 \le C$.

Next, using the estimates for the decay of $\hat{K}(x,\tau)$, we obtain

$$\|\zeta_{N,i}(x,\cdot)\|_1 = \int_{\mathbb{R}^d} |\hat{K}(2^N x, 2^{-N}\tau)(\eta(2^{-n_i}|\tau|) - \eta(2^{-n_{i-1}}|\tau|))|\, d\tau \le 2^{dN}\int_{2^{n_{i-1}-N-1} \le |\xi| \le 2^{n_i - N}} |\hat{K}(2^N x,\xi)|\, d\xi \le C2^{dN+m^*(N-n_{i-1})}.$$

Similarly, since $n_i > N$ and the derivatives of $\hat{K}(x,\tau)$ satisfy the same decay estimates, we get

$$\|\Delta^M_\tau \zeta_{N,i}(x,\cdot)\|_1 = \int_{\mathbb{R}^d} \left|\Delta^M_\tau \left(\hat{K}(2^{n_l}x, 2^{-n_l}\tau)(\eta(2^{-n_i}|\tau|) - \eta(2^{-n_{i-1}}|\tau|))\right)\right| d\tau \le C2^{(d-2M)N+m^*(N-n_{i-1})},$$

for any $M \ge 1$.

Further, we define $\sigma_{N,i}(x,y) = \mathcal{F}^{-1}(\zeta_{N,i}(x,\cdot)),\ i > l$. Then, applying Lemma 4.1.1 once again, we have $\|\sigma_{N,i}\|_1 \le C2^{m^*(N-n_{i-1})}$. Finally, applying Lemma 4.1.3 and (4.3), we obtain

$$\left|\int_{\mathbb{R}^d} K_N(x,y)u(y,t)\,dy\right| \le C\|u\|_{v,\infty}\left(w(2^{-n_l}) + 2^{m^*N}\sum_{i=l+1}^{\infty} 2^{-m^*n_{i-1}}w(2^{-n_i})\right) \le C\|u\|_{v,\infty}A^l, \tag{4.5}$$

for any $x \in \mathbb{R}^d$ and $N \le n_l$. Then (4.4) follows by taking $N = n_l$ and $N = n_{l-1}$. □

Corollary 4.2.1 *Let $\{\psi_{p,jk}\}$ be an orthogonal smooth wavelet basis as above. Then there exist C such that for any $u \in h^\infty_v$*

$$|c_{p,jk}(u(\cdot,t))| \le C2^{-dj/2}\|u\|_{v,\infty}w(2^{-j}), \tag{4.6}$$

when $t > 0,\ j \in \mathbb{Z}_+,\ k \in \mathbb{Z}^d$.

Proof. Clearly,

$$|\langle u(x,t), \psi_{p,jk}(x)\rangle| = |\langle g_l(x,t), \psi_{p,jk}(x)\rangle| \le \|g_l(x,t)\|_\infty \|\psi_{p,jk}\|_1,$$

where $j \in (n_{l-1}, n_l]$. Then (4.4) implies (4.6). $\square$

4.2.2 Converse estimates and coefficient characterization

The converse of Theorem I.15 is also true. We remind its statement as well.

Theorem 4.2.2 *Let u be a harmonic function in $\mathbb{R}^{d+1}_+$ that is bounded on each half-space $\{(x,t) \in \mathbb{R}^{d+1}, t \geq t_0 > 0\}$. Suppose that for each $t > 0$*

$$u(x,t) = \sum_{l=0}^{\infty} g_l(x,t),$$

where the series converges uniformly on $\mathbb{R}^d$, $g_0(\cdot,t) \in V_0(\infty)$,

$$g_l(x,t) = \sum_{j=n_{l-1}+1}^{n_l} \sum_{p=1}^{q} \sum_{k\in\mathbb{Z}^d} a_p^{(jk)}(t)\psi_{p,jk}(x),\ \ l \geq 1$$

and there exists B such that

$$\|g_l(\cdot,t)\|_\infty \leq Bw(2^{-n_l}),$$

for any $t > 0$. Then $u \in h^\infty_v$ and $\|u\|_{v,\infty} \leq CB$, where C depends on A and ϕ only.

Proof. We fix $s \in (0,1]$ and take L such that $s \in [2^{-n_{L+1}}, 2^{-n_L})$. Then

$$\begin{aligned} u(x,t+s) = \left(u(\cdot,t) * P_{(s)}\right)(x) = &\sum_{l=0}^{\infty}(g_l(\cdot,t) * P_{(s)})(x) = \\ &\sum_{l=0}^{L+1} \left(g_l(\cdot,t) * P_{(s)}\right)(x) + \sum_{l=L+2}^{\infty} \left(g_l(\cdot,t) * P_{(s)}\right)(x). \end{aligned}$$

For each $l \leq L+1$ we have

$$|g_l(\cdot,t) * P_{(s)}| \leq \|g_l(\cdot,t)\|_\infty \leq Bw(2^{-n_l}).$$

Since $\langle g_l(y,t), K_{n_{l-1}}(x,y)\rangle = 0$, for $l > L+1$ we get

$$\begin{aligned} &\left(g_l(\cdot,t) * P_{(s)}\right)(x) = \\ &\quad \int_{\mathbb{R}^d} g_l(y,t)P_{(s)}(x-y)\,dy - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g_l(y,t)K_{n_{l-1}}(\xi,y)P_{(s)}(x-w)\,d\xi dy = \\ &\quad\quad \int_{\mathbb{R}^d} g_l(y,t)\int_{\mathbb{R}^d}(P_{(s)}(x-y) - P_{(s)}(x-\xi))K_{n_{l-1}}(\xi,y)\,d\xi\,dy. \end{aligned}$$

Then by rescaling and applying Lemma 4.1.4, we obtain

$$|g_l(\cdot,t) * P_{(s)}| \leq CBw(2^{-n_l})(2^{n_{l-1}}s)^{-r}.$$

Now we remark that $2^{n_{l-1}}s \geq 1$ and $r > m^*$, where m^* is chosen such that (4.3) holds. Then

$$|g_l(\cdot,t) * P_{(s)}| \leq CBA^2 s^{-m^*} w(2^{-n_{l-1}}) 2^{-m^* n_{l-1}}.$$

Finally we add up the estimates and take into account (4.3) to get

$$|u(x,t+s)| \leq CBw(s),$$

for any $t > 0$.

□

When the weight v grows sufficiently fast we can reformulate the result in terms of the wavelet coefficients. For general weights such characterization is not possible (see also Example below).

Definition 4.2.1. *We say that a weight v is of power-type growth if the doubling condition (I.37) is fulfilled and there exists κ such that the sequence n_j defined in 4.2.1 satisfies $n_{j+1} - n_j \leq \kappa$ for any $j \geq 0$.*

Typical examples of weights of power-type growth are $w(t) = t^{-a}$, $a > 0$. Normal weights in the terminology of Shields and Williams, [88], are of power-type growth. When v is a weight of power-type growth, harmonic functions in h_v^∞ can be characterized by their wavelet *coefficients* if one combines Corollary 4.2.1 with the one below.

Corollary 4.2.2 *Let v be a weight of power-type growth, and let u be harmonic in $\mathbb{R}^{d+1}_+$ and bounded on each half-space $\{(x,t) \in \mathbb{R}^{d+1}, t \geq t_0 > 0\}$. Suppose there exists B such that*

$$|b_k(u(\cdot,t))| \leq B, \quad \textit{and} \quad |c_{p,jk}(u(\cdot,t))| \leq 2^{-dj/2} Bw(2^{-j}), \tag{4.7}$$

for any $t > 0$, $j \in \mathbb{Z}_+$, $k \in \mathbb{Z}^d$. Then $u \in h_v^\infty$ and $\|u\|_{v,\infty} \leq CB$, where C does not depend on u.

Proof. Let g_l be defined as in Theorem 4.2.1. We want to show that $\|g_l(x,t)\| \leq Bw(2^{-n_l})$. Since we have only finitely many dyadic generations between n_{l-1} and n_l it suffices to estimate

$$\sum_{k \in \mathbb{Z}^d} \langle u(y,t), \psi_{p,jk}(y) \rangle \psi_{p,jk}(x)$$

for each j. Applying (4.7) and the inequality from [69, ch 3.1],

$$\max_x \sum_{k \in \mathbb{Z}^d} |\psi_{p,jk}(x)| \leq C 2^{dj/2},$$

we get the required estimate. □

4.2.3 Wavelet characterization

Now we prove Theorem I.13, we recall its formulation for reader's convenience.

Theorem 4.2.3 *Let $u(x,t)$ be a harmonic function on $\mathbb{R}^{d+1}_+$ bounded on each half-space $\{(x,t): t > t_0 > 0\}$. Then $u \in h^\infty_v$ if and only if there exists C such that*

$$M_N(u) = \sup_{t>0} \|S_N(u(\cdot,t))\|_{L^\infty(\mathbb{R}^d)} \le Cw(2^{-N}).$$

Similarly, $u \in h^0_v$ if and only if $\lim_{N\to\infty} M_N(u)(w(2^{-N}))^{-1} = 0$.

While it follows readily from the proofs of Theorem 4.2.1 and 4.2.2, we provide an argument below just for the sake of completeness. We choose to reformulate the statement using blocks of wavelet decomposition.

Suppose that $u \in h^\infty_v$ and $n_{l-1} \le N < n_l$. We have

$$S_N(u(x,t)) = \int_{\mathbb{R}^d} K_N(x,y)u(y,t)\,dy,$$

and $w(2^{-N}) \ge cA^l$. Then, applying (4.5) for this l, we obtain $|S_N(u(x,t))| \le C\|u\|_{v,\infty} w(2^{-N})$. If in addition $u \in h^0_v$, then similarly we have

$$M_N(u) = \sup_t |S_N(u(x,t))| = o(w(2^{-N})), \quad N \to \infty.$$

To prove the converse, assume that $M_N(u) \le \varepsilon_l w(2^{-N})$ when $N \ge n_{l-1}$. Then clearly

$$|g_l(x,t)| = |S_{n_l}(x,t) - S_{n_{l-1}}(x,t)| \le 2\varepsilon_l w(2^{-n_l}).$$

Theorem 4.2.2 implies that $u \in h^\infty_v$ when ε_l are bounded. When $\varepsilon_l \to 0$ as $l \to \infty$, we get

$$|g_l(\cdot,t) * P_{(s)}| \le 2\varepsilon_l w(2^{-n_l})$$

for any $t > 0$. Moreover, as in the proof of Theorem 4.2.2,

$$|g_l(\cdot,t) * P_{(s)}| \le C\varepsilon_l w(2^{-n_l})(2^{n_{l-1}}s)^{-r}.$$

Then we choose L such that $s \in [2^{-n_{L+1}}, 2^{-n_L})$ and write

$$|u(x,t+s)| \le \sum_{l=0}^{L+1} 2\varepsilon_l w(2^{-n_l}) + C\varepsilon_L \sum_{l=L+2}^{\infty} w(2^{-n_l})(2^{n_{l-1}}s)^{-m} \le$$
$$\sum_{l=0}^{L+1} 2\varepsilon_l w(2^{-n_l}) + C_1\varepsilon_L(2^{n_{L+1}}s)^{-m} w(2^{-n_{L+1}}) \le c_L w(s),$$

where c_L goes to zero as L goes to infinity. This concludes the proof of Theorem 4.2.3.

We will also reformulate the result in terms of the boundary values of harmonic functions. Let $h^{-a} = h^\infty_{t^{-a}}$ for $a > 0$ and $h^{-\infty} = \cup_a h^{-a}$. The doubling condition on the weight v implies that $h^\infty_v \subset h^{-a}$ for some $a > 0$. Harmonic functions in $h^{-\infty}$ admit boundary values in the sense of distributions of finite order, see for example [92]. Thus when we take sufficiently smooth multiresolution approximation and choose compactly supported wavelets, we can define the wavelet coefficients of the boundary values of $u \in h_v^\infty$. Then we can reformulate the main result in the following way.

Suppose that $u \in h^{-a}$ and let U be the boundary values of u in the sense of distributions. Let further $b_k(U)$ and $c_{p,jk}(U)$ be the wavelet coefficients of U with respect to a sufficiently smooth compactly supported wavelet basis. We define

$$S_N(U)(x) = \sum_{p=1}^{q} \sum_{j=0}^{N} \sum_{k \in \mathbb{Z}^d} c_{p,jk}(U)\psi_{p,jk}(x) + \sum_{k \in \mathbb{Z}^d} b_k(U)\phi(x-k).$$

Corollary 4.2.3 *Let $h_v^\infty \subset h^{-a}$ and let $u \in h^{-a}$. Then $u \in h_v^\infty$ if and only if there exists $K > 0$ such that*

$$\|S_N(U)\|_\infty \le K w(2^{-N})$$

for any N.

Proof. Clearly, $c_{p,jk}(U) = \lim_{t\to 0} c_{p,jk}(u(\cdot,t))$ and then

$$S_N(U)(x) = \lim_{t\to 0} S_N(u(\cdot,t))(x), \quad x \in \mathbb{R}^d.$$

Thus if $u \in h_v^\infty$ the required estimate holds.

We want to prove the converse. Consider the sequence $u_N(\cdot,y) = S_N(U) * P_{(y)}$ of harmonic functions in the upper half-space. By repeating the estimates from the proof of Theorem 4.2.2, we conclude that $u_N \in h_v^\infty$ and $\|u_N\|_{v,\infty} \le CK$. Thus $\{u_N\}$ form a normal family in the upper half-space and we can choose a convergent subsequence $\{u_{N_j}\}$ that converges to $u_0 \in h_v^\infty$. Further, let U_0 be the boundary values of u_0. We have $c_{p,jk}(U_0) = c_{p,jk}(U)$ and $b_k(U_0) = b_k(U)$. This implies $U_0 = U$ and since u_0 and u are bounded in $\{(x,t), t \ge 1\}$ we conclude that $u = u_0 \in h_v^\infty$. □

Chapter 5 Growth classes on Lipschitz domains: boundary oscillation

5.1 Proof of Theorems I.17 and I.18

5.1.1 Main approximation lemma

Given two functions f and g, we say that $f \lesssim g$ if there is a positive constant $C = C(w, d, \|\phi'\|_\infty, \|u\|_{w,\infty})$ such that $f \le Cg$. We write $f \approx g$ if $f \lesssim g$ and $g \lesssim f$ simultaneously. Consider a positive decreasing sequence $\{s_k\}_{k=0}^\infty$ such that

$$w(s_k) = 2^k, \quad k \in \mathbb{Z}_+,$$

and put

$$n_0 = 0,\ n_k = -\left[\frac{\log s_k}{\log 2}\right], \quad k \in \mathbb{N}$$

(we do not mind if $n_k = n_{k+1}$ for some $k \in \mathbb{N}$, it can happen for fast growing weights).
It follows from the doubling property (I.37) that $w(2^{-n_k}) \approx 2^k$. Consider $I(x,\delta)$ defined in (I.43). The approximation of $I(x,\delta)$ by martingales is provided by Lemma I.2 which we recall here.

Lemma 5.1.1 *Assume that $u \in h_w^\infty(\Omega_\phi)$. Then for every $x_0 \in \mathbb{R}^d$ there exists a probability measure μ on $Q(x_0)$ and a (super)dyadic martingale $S = \{S_k, \mathcal{F}_{n_k}, \mu\}_{k=0}^\infty$ on $Q(x_0)$ such that μ is absolutely continuous with respect to the Lebesgue measure on $Q(x_0)$ and for every $k \in \mathbb{Z}_+$*

$$|S_k(x) - I(x, s_k)| \lesssim 1, \tag{5.1a}$$

$$|S_k(x) - S_{k+1}(x)| \lesssim 1, \quad x \in Q(x_0). \tag{5.1b}$$

5.1.2 How to deduce Theorem I.17

Assuming that Lemma 5.1.1 holds, we proceed by the standard argument. Fix any $x_0 \in \mathbb{R}^d$ and put

$$E = \{x \in Q(x_0) : \lim_{m\to\infty} |\langle S\rangle_m|(x) < \infty\}.$$

The inequality (5.1b) implies that $\langle S\rangle_m^2 \lesssim m,\ m \ge 1$. Applying the Law of the Iterated Logarithm for martingales to S (see, for example, Theorem **3.0.2** in [8]), we get

$$\limsup_{m\to\infty} \frac{|S_m|(x)}{\sqrt{m \log\log m}} \lesssim 1 \qquad \mu\ a.e.\ x \in Q(x_0) \setminus E.$$

It is well known that for μ almost every $x \in E$ the sequence $\{S_m(x)\}$ is bounded, so (5.1a) implies that the sequence $\{I(x, s_m)\}$ is bounded μ *a.e.* on E as well. It follows that

$$\limsup_{m\to\infty} \frac{|I(x,s_m)|}{\sqrt{m \log\log m}} \lesssim 1 \quad \mu \ a.e.\ x \in Q(x_0). \tag{5.2}$$

Now for $s_m \le \delta \le s_{m-1}$ and $x \in Q(x_0)$ we have

$$\begin{aligned} |I(x,s_m) - I(x,\delta)| &\le \int_{s_m}^{\delta} |u(x,\phi(x)+y)|\, d\left(\frac{1}{w(y)}\right) \\ &\lesssim \log w(s_m) - \log w(s_{m-1}) = 1, \end{aligned}$$

also, clearly, $w(\delta) \ge w(s_{m-1}) = \frac{1}{2}w(s_m)$. Combined with (5.2) and the fact that μ is absolutely continuous with respect to the Lebesgue measure, it gives us

$$\limsup_{\delta\to\infty} \frac{|I(x,\delta)|}{\sqrt{\log w(\delta) \log\log\log w(\delta)}} \lesssim 1, \quad a.e.\ x \in Q(x_0).$$

The inequality (I.44) follows immediately.

5.1.3 Proof of Lemma 5.1.1: auxiliary function H

The approximation of $I(x,\theta)$ by a Bloch function is covered by the following lemma

Lemma 5.1.2 *Assume that $u \in h_w^\infty(\Omega_\phi)$. Put*

$$H(x,t) = \int_0^1 u(x,t+y)\, d\left(\frac{1}{w(y)}\right), \quad (x,t) \in \Omega_\phi. \tag{5.3}$$

Then H belongs to $\mathcal{B}(\Omega_\phi)$ and $\|H\|_{\mathcal{B}} \lesssim 1$. Moreover

$$|H(x,\phi(x)+\theta) - I(x,\theta)| \lesssim 1, \quad x \in \mathbb{R}^d, 0 < \theta \le 1. \tag{5.4}$$

Proof. First we note that H is harmonic in Ω_ϕ (it is the average of harmonic functions). We proceed by proving the following inequality

$$|\nabla u|(x,\phi(x)+\theta) \lesssim \frac{w(\theta)}{\theta}, \quad x \in \mathbb{R}^d,\ \theta > 0. \tag{5.5}$$

Fix any positive θ. Since ϕ is a Lipschitz function, we see that $\mathrm{dist}((x,\theta+\phi(x)), \partial\Omega_\phi) \approx \theta$ for any $x \in \mathbb{R}^n$ and positive θ. It follows from (I.37) that for $y \ge \frac{\theta}{2}$ we have

$$|u(x,\phi(x)+y)| \le w(\mathrm{dist}(x, y+\phi(x)), \partial\Omega_\phi)) \lesssim w\left(\frac{\theta}{2}\right), \tag{5.6}$$

so there exists a constant $C = C(d,\phi,w)$ such that

$$0 \le u(x,y) + Cw\left(\frac{\theta}{2}\right) \le (C+1)w\left(\frac{\theta}{2}\right), \quad (x,y) \in \Omega_{\phi+\frac{\theta}{2}}.$$

Then the estimate (5.5) follows from (I.37) and the Harnack inequality. For $(x,\theta) \in \Omega_\phi$ (5.5) implies

$$
\begin{aligned}
|\nabla H|\,(x,\phi(x)+\theta) &\le \int_0^1 |\nabla u|(x,\phi(x)+\theta+y)\,d\left(\frac{1}{w(y)}\right)\\
&\lesssim \int_0^1 \frac{w(\theta+y)}{\theta+y}\,d\left(\frac{1}{w(y)}\right) = \int_0^\theta \frac{w(\theta+y)}{\theta+y}\,d\left(\frac{1}{w(y)}\right) + \int_\theta^1 \frac{w(\theta+y)}{\theta+y}\,d\left(\frac{1}{w(y)}\right).
\end{aligned}
$$

Since the function $\frac{w(y)}{y}$ is decreasing, we have

$$
\int_0^\theta \frac{w(\theta+y)}{\theta+y}\,d\left(\frac{1}{w(y)}\right) \le \int_0^\theta \frac{w(\theta)}{\theta}\,d\left(\frac{1}{w(y)}\right) = \frac{1}{\theta}.
$$

On the other hand,

$$
\begin{aligned}
\int_\theta^1 \frac{w(\theta+y)}{\theta+y}\,d\left(\frac{1}{w(y)}\right) &\le \int_\theta^1 \frac{w(y)}{y}\,d\left(\frac{1}{w(y)}\right)\\
&\le \sum_{k=0}^{[\log\frac{1}{\theta}]} \int_{2^k\theta}^{2^{k+1}\theta} \frac{1}{y} d\log\frac{1}{w(y)} \le \sum_{k=0}^{[\log\frac{1}{\theta}]} \frac{1}{2^k\theta} \int_{2^k\theta}^{2^{k+1}\theta} d\log\frac{1}{w(y)}\\
&\le \frac{1}{\theta}\sum_{k=0}^{[\log\frac{1}{\theta}]} 2^{-k}(\log w(2^k\theta) - \log w(2^{k+1}\theta))\\
&\le \frac{1}{\theta}\sum_{k=0}^{[\log\frac{1}{\theta}]} 2^{-k}(\log\left(Dw(2^{k+1}\theta)\right) - \log w(2^{k+1}\theta))\\
&\le \frac{\log D}{\theta}\sum_{k=0}^{[\log\frac{1}{\theta}]} 2^{-k} \lesssim \frac{1}{\theta}.
\end{aligned}
\tag{5.7}
$$

Gathering the estimates, we arrive at

$$
|\nabla H|\,(x,\phi(x)+\theta) \lesssim \frac{1}{\theta} \approx \frac{1}{\operatorname{dist}((x,\phi(x)+\theta),\partial\Omega_\phi)},
$$

and we get the first part of the lemma.

To prove (5.4) we write

$$
\begin{aligned}
&H(x,\phi(x)+\theta) - I(x,\theta)\\
&\quad= \int_0^1 u(x,\phi(x)+\theta+y)\,d\left(\frac{1}{w(y)}\right) - \int_\theta^1 u(x,y)\,d\left(\frac{1}{w(y)}\right)\\
&\quad= \int_0^\theta u(x,\phi(x)+\theta+y)\,d\left(\frac{1}{w(y)}\right)\\
&\qquad+ \int_\theta^1 (u(x,\phi(x)+\theta+y) - u(x,\phi(x+y)))\;d\left(\frac{1}{w(y)}\right).
\end{aligned}
$$

Following the same reasoning as above, we see that

$$\left|\int_0^\theta u(x,\phi(x)+\theta+y)\,d\left(\frac{1}{w(y)}\right)\right| \le w(\theta)\int_0^\theta d\left(\frac{1}{w(y)}\right) = 1.$$

Again, (5.5) implies that

$$\begin{aligned}\left|\int_\theta^1 (u(x,\phi(x)+\theta+y)-u(x,\phi(x+y)))\,d\left(\frac{1}{w(y)}\right)\right|&\\ \le \int_\theta^1\int_y^{y+\theta} |\nabla u|(x,\phi(x)+s)\,ds\,d\left(\frac{1}{w(y)}\right)&\\ \le \int_\theta^1\int_y^{y+\theta}\frac{w(s)}{s}\,ds\,d\left(\frac{1}{w(y)}\right) \le \int_\theta^1 w(y)\frac{\theta}{y}\,d\left(\frac{1}{w(y)}\right) \lesssim 1,&\end{aligned}$$

just like in (5.7). Combining these two inequalities we get (5.4). □

5.1.4 Proof of Lemma 5.1.1: dyadic martingale

Once we obtained the intermediate approximation of I by a Bloch function, we can proceed to martingales. It is well known (see, for example, [66]) that the Bloch functions in the unit disc can (up to a constant error) be viewed as dyadic martingales. The case of Lipschitz domains was considered by Llorente, Corollary **2** in [58] is the main instrument in the following argument.

Fix any point $x_0 \in \mathbb{R}^d$ and let

$$\begin{aligned}&A = \|\phi'\|_\infty\sqrt{d},\ \lambda = 8+\frac{1}{A},\\ &\Omega_1 = \{(x,y) : x\in\lambda Q(x_0) : \phi(x)\le y\le\phi(x)+\lambda A\}.\end{aligned}$$

The following proposition holds

Proposition 5.1.1 (Corollary **2**, [58]) *If $v\in\mathcal{B}(\Omega_1)$ then there exists a dyadic martingale $\mathcal{M} = \{M_k,\mathcal{F}_k(x_0),\omega\}$ in $Q(x_0)$ and a positive constant $C = C(\phi,d)$ such that ω is absolutely continuous with respect to the Lebesgue measure on $Q(x_0)$, and for every $k\in\mathbb{N}$ if $A2^{-(k+1)}\le t\le A2^{-k}$, then for every $x\in Q(x_0)$*

$$|M_k(x)-v(x,\phi(x)+t)| \le C\|v\|_{\mathcal{B}}, \tag{5.8a}$$

$$|M_{k+1}(x)-M_k(x)| \le C\|v\|_{\mathcal{B}}. \tag{5.8b}$$

We apply this proposition to H and put

$$S = \{S_k,\mathcal{F}_{n_k}(x_0),\omega\} := \{M_{n_k},\mathcal{F}_{n_k}(x_0),\omega\}.$$

Now we prove (5.1a) and (5.1b). It follows from (5.8a) and (5.4) that

$$\begin{aligned}|S_k(x) - I(x, s_k)| &= |M_{n_k}(x) - I(x, s_k)| \\ &\le |M_{n_k}(x) - H(x, \phi(x) + A2^{-n_k})| + |H(x, \phi(x) + A2^{-n_k}) - H(x, \phi(x) + s_k)| \\ &\quad + |H(x, \phi(x) + s_k) - I(x, s_k)| \lesssim 1 + \int_{s_k}^{A2^{-n_k}} |\nabla H(x, \phi(x) + y)| \, dy \\ &\lesssim 1, \quad x \in Q(x_0),\end{aligned}$$

since $\|H\|_{\mathcal{B}} \lesssim 1$, and we get (5.1a). To obtain (5.1b) we note that

$$\begin{aligned}|S_k(x) - S_{k+1}(x)| &= |M_{n_k}(x) - M_{n_{k+1}}(x)| \le |M_{n_k}(x) - I(x, s_k)| \\ &\quad + |I(x, s_k) - I(x, s_{k+1})| + |I(x, s_{k+1}) - M_{n_{k+1}}(x)| \\ &\lesssim 1 + |I(x, s_k) - I(x, s_{k+1})|.\end{aligned}$$

Clearly,

$$|I(x, s_k) - I(x, s_{k+1})| \le w(s_k) \int_{s_{k+1}}^{s_k} d\left(\frac{1}{w(y)}\right) = 2^k(2^{-k} - 2^{-k-1}) = \frac{1}{2},$$

and the inequality (5.1b) follows.

5.1.5 Proof of Theorem I.18

The proof is standard. We apply the usual ice-cream cone construction to Σ, i.e. consider the domain

$$\Omega = \bigcup_{x \in \Sigma} \Gamma(x, M),$$

where $\Gamma(x, M)$ is the cone with vertex x and aperture M, $\Gamma(x, M) = \{(\tilde{x}, y) \in \mathbb{R}_+^{d+1} : |\tilde{x} - x| \le My\}$. Clearly Ω is the area above the graph of some Lipschitz function ϕ with $\|\phi'\|_\infty = \frac{1}{M}$, so that $\Omega = \Omega_\phi$. The condition (I.45) then implies that

$$|u(x, y)| \le Kw(y) \lesssim w(\mathrm{dist}((x, y), \partial\Omega)), \quad (x, y) \in \Omega,$$

and we can apply Theorem I.17 to obtain

$$\limsup_{\delta \to 0} \frac{I(x, \delta)}{\sqrt{\log w(\delta) \log\log\log w(\delta)}} \le C, \quad a.e.\ x \in \mathbb{R}^d.$$

Theorem I.18 follows immediately.

5.2 An example

In the proof of Theorem I.17 we introduced the harmonic function H which is shown to be a Bloch function. In addition, the estimate (5.4) implies that $H \in h^\infty_{\log w}(\Omega_\phi)$, and that the LIL in (I.44) holds for H as well,

$$\limsup_{\delta\to 0} \frac{H(x,\delta)}{\sqrt{\log w(\delta)\log\log\log w(\delta)}} \lesssim 1, \quad a.e.\ x\in\mathbb{R}^d. \tag{5.9}$$

To obtain this estimate we used the special nature of H, namely that it was constructed on $u\in h_w^\infty(\Omega_\phi)$. It is then natural to ask if an *arbitrary* function $v\in h_{w_0}^\infty(\Omega_\phi)\bigcap\mathcal{B}(\Omega_\phi)$ satisfies the following LIL

$$\limsup_{\delta\to 0} \frac{v(x,\phi(x)+\delta)}{\sqrt{w_0(\delta)\log\log w_0(\delta)}} \lesssim 1, \quad a.e.\ x\in\mathbb{R}^n. \tag{5.10}$$

The answer to this question is negative as provided by the following proposition.

Proposition 5.2.1 *Let* $w_0(y)=\log\log\frac{e}{y}+1,\ y\in(0,1]$. *There exists a constant* $A>0$, *a function* $v\in h_{w_0}^\infty(\mathbb{R}^2_+)\bigcap\mathcal{B}(\mathbb{R}^2_+)$, *a number* $k_0\in\mathbb{N}$ *and a sequence* $\{y_k\}_{k=k_0}^\infty\to 0$ *such that*

$$\left|\left\{t\in[0,1]:\ |v(t,y_k)|\geq\frac{w_0(y_k)}{A}\right\}\right|\geq\frac{1}{10}. \tag{5.11}$$

It is known that a function in $h_w^\infty(\mathbb{R}^2_+)$ can grow as fast as w only on small part of vertical rays $\{x+iy\}$, $y\in\mathbb{R}_+$, however it can attain the maximal growth on the subsets of those rays for a.e. $x\in\mathbb{R}$ (see [65], [14]). Unfortunately, we can not use the example provided there, since it is constructed as a lacunary trigonometric series, for which, as it can be shown, (5.9) holds if it belongs to the Bloch class.

Proof. Given two real-valued functions $f,g\in L^2(\mathbb{R}^n)$ we denote the scalar product $\int_{\mathbb{R}^n} f(t)g(t)dt$ by $\langle f,g\rangle$. Consider a function $\varphi:\mathbb{R}\mapsto\mathbb{R}$ such that $\operatorname{supp}\varphi\subset[0,1]$, $\varphi\in C^{10}$, $\|\varphi\|_\infty\leq 1$. We also require that $\int_{\mathbb{R}}\varphi(t)\,dt=0$ and $\langle\varphi,\psi\rangle\neq 0$ (where ψ is a Haar wavelet mentioned earlier). For example we can take the suitable renormalization of the Daubechies wavelet (or any other smooth wavelet with compact support that satisfies our conditions). By $P_{(y)}$ we denote the Poisson kernel for the halfplane, $P_{(y)}(t)=\frac{y}{\pi(y^2+t^2)}$, $y>0$, $t\in\mathbb{R}$.

The idea is to obtain a functional series of the form

$$\sum_{j=0}^{k}\sum_{I\in\Delta_j} c_I\varphi_I(t) := \Phi_k(t), \quad t\in\mathbb{R}, \tag{5.12}$$

that satisfies properties similar to those in the statement, and then prove that the corresponding Bloch function provides the required example. To elaborate we first construct Φ_k and an increasing sequence $\{b_j\}_{j=1}^\infty\subset\mathbb{Z}_+$ in such a way that we have

$$\|\Phi_k-\Phi_{k-1}\|_\infty\leq 1, \tag{5.13a}$$
$$\|\Phi_k\|_\infty\leq w_0(2^{-k})+2, \tag{5.13b}$$
$$\left|\left\{t\in(0,1]:|\Phi_{b_k}|(t)\geq\tfrac{w_0(2^{-b_k})}{4}\right\}\right|\geq\tfrac{1}{10}, \tag{5.13c}$$

for any integer $k\geq k_0$.

The property (5.13a) is an analogue of the Bloch condition, (5.13b) is the growth restriction, and (5.13c) corresponds to (5.11) (so that there is no LIL for Φ_k with w_0).

5.2.1 Construction of $\{b_j\}$ and Φ_k

First we chose $a \in \mathbb{N}$ such that $2^{-a+1}\|\varphi'\|_\infty \le \frac{1}{4}\,|\langle\varphi,\psi\rangle|$. Now chose a natural $j_0 \ge 4\|\varphi'\|_\infty + 4$ and an increasing sequence $b_j \in \mathbb{N}$ in such a way that

$$\begin{aligned} &b_1 = 0,\ \frac{b_j}{a} \in \mathbb{N},\\ &j-1 \le w_0(2^{-b_j}) \le j,\\ &\left(\frac{b_j - b_{j-1}}{a} - 1\right)\langle\varphi,\psi\rangle^2 \ge 4j^2, \quad j \ge j_0. \end{aligned} \tag{5.14}$$

It is not hard to verify that such choice is possible (we remind that $w_0(y) = \log\log\frac{e}{y} + 1$).

The functions Φ_k are constructed via double induction, first on j, and then on m between b_j and b_{j+1}. Put $\Phi_0(t) = \varphi(t)$. Assume now that we obtained Φ_{b_j} for some $j \in \mathbb{N}$. Consider all the intervals $I \in \Delta_{b_j}$ such that $\sup_{t\in I}|\Phi_{b_j}|(t) > j$, we denote the set of these intervals by $\mathcal{E}_j^{b_j}$. Now suppose that we have constructed Φ_{m-1} and $\mathcal{E}_j^{m-1}$ for some m, $b_j + 1 \le m \le b_{j+1}$. If $\frac{m}{a} \in \mathbb{Z}$, then for $I \in \Delta_m$ and $t \in I$ let

$$\Phi_m(t) = \Phi_{m-1}(t), \quad t \in \textstyle\bigcup_{J\in\mathcal{E}_j^{m-1}} J, \tag{5.15a}$$

$$\Phi_m(t) = \Phi_{m-1}(t) + \varphi_I(t), \quad t \notin \textstyle\bigcup_{J\in\mathcal{E}_j^{m-1}} J, \tag{5.15b}$$

$$\mathcal{E}_j^m = \mathcal{E}_j^{m-1} \textstyle\bigcup \{J \in \Delta_m : \sup_{t\in J}|\Phi_m(t)| > j\}. \tag{5.15c}$$

Otherwise we put

$$\Phi_m(t) = \Phi_{m-1}(t), \quad t \in (0,1], \tag{5.16a}$$

$$\mathcal{E}_j^m = \mathcal{E}_j^{m-1}. \tag{5.16b}$$

Finally, put

$$\mathcal{E}_j = \bigcup_{m=b_j}^{b_{j+1}-1} \mathcal{E}_j^m = \mathcal{E}_j^{b_{j+1}-1}.$$

What we do here is, essentially, a stopping time procedure applied (instead of martingales as usual) to the functional series of the form like in (5.12). We see that if $I \in \mathcal{E}_j$, then the construction is stopped at this interval, and $\Phi_{b_{j+1}}(t) = \Phi_m(t)$, $t \in I$, $m = \operatorname{rank} I$, where rank of an interval is just its depth in the respective tree, or $\operatorname{rank} I = |\log_2|I||$. If, on the other hand, $t \in (0,1] \setminus \bigcup_{J\in\mathcal{E}_j} J$, then the construction happens on every step (divisible by a) up until b_{j+1}. The set $(0,1] \setminus \bigcup_{J\in\mathcal{E}_j} J$ can be decomposed into a disjoint union of intervals from $\Delta_{b_{j+1}}$, we denote the set of these intervals by $\mathcal{G}_j$.

Clearly Φ_m is of the form like in (5.12), moreover,

$$\Phi_{b_{j+1}}(t) = \Phi_{b_j}(t) + \sum_{m=b_j+1}^{b_{j+1}} \sum_{J\in\Delta_m} c_J\varphi_J(t), \quad t\in(0,1],$$

where $c_J = 1$ only if $\frac{\operatorname{rank} J}{a} \in \mathbb{Z}$ *and* there is no interval $I \in \mathcal{E}_j$ such that $I \supset J$, $c_J = 0$ otherwise. We also see that

$$\sup_{t\in I} |\Phi_{b_{j+1}}|(t) > j, \quad I \in \mathcal{E}_j;$$
$$|\Phi_{b_{j+1}}| \le j+1, \quad j \in \mathbb{N}.$$

We are left to check (5.13a)-(5.13c). The condition (5.13a) follows straight from (5.15), since $\|\varphi_I\|_\infty = 1$ for any interval I. For any $k\in\mathbb{N}$ there exists $j_k\in\mathbb{N}$ such that $b_{j_k} \le k \le b_{j_k+1}-1$. We therefore have

$$\|\Phi_k\|_\infty \le j_k + 1 \le w_0(2^{-b_{j_k}}) + 2 \le w_0(2^{-k}) + 2,$$

and we obtain (5.13b).

5.2.2 Proof of (5.13c): martingale decomposition

Pick any $j \ge j_0$ (we remind that j_0 was defined in (5.14)). Since $j_0 \ge 4\|\varphi'\|_\infty + 4$, we see that $\frac{j}{2} - \|\varphi'\|_\infty \ge \frac{j}{2} - \frac{j_0}{4} \ge \frac{j}{4}$, and, due to (5.14) we have $\frac{j}{2} - \|\varphi'\|_\infty \ge \frac{w_0(2^{-b_j})}{4}$. It follows that to obtain (5.13c) it is enough to prove

$$\left|\left\{t\in(0,1] : |\Phi_{b_{j+1}}(t)| \ge \frac{j}{2} - \|\varphi'\|_\infty\right\}\right| \ge \frac{1}{10}. \tag{5.17}$$

The first step is to prove that $|\Phi_{b_{j+1}}|$ is "sufficiently large" on the intervals from $\mathcal{E}_j$, namely that for any $I\in\mathcal{E}_j$ we have

$$|\Phi_{b_{j+1}}|(t) \ge j - 2\|\varphi'\|_\infty, \quad t\in I. \tag{5.18}$$

Indeed, for $m = \operatorname{rank} I$ we have

$$|\Phi'_m|(t) \le \sum_{J\in\Delta\colon t\in J,\, \operatorname{rank} J\le m} c_J\|\varphi'_J\|_\infty = \|\varphi'\|_\infty \sum_{J\in\Delta\colon t\in J,\, \operatorname{rank} J\le m} \frac{c_J}{|J|} \le 2^{m+1}\|\varphi'\|_\infty, \quad t\in(0,1],$$

since $|c_J| \le 1$ for any $J\in\Delta$. Again we see that $\Phi_{\beta_{j+1}}(t) = \Phi_m(t)$ on I, therefore $|\sup_{t\in I}\Phi_m(t) - \inf_{t\in I}\Phi_m(t)| \le \int_I |\Phi'_m(t)|\,dt \le 2\|\varphi'\|_\infty$, and we get (5.18).

Now we show that

$$\left|\bigcup_{J\in\mathcal{G}_j} J\right| \le \frac{3}{4}, \tag{5.19}$$

combined with (5.18) it implies (5.17). In order to do this consider the Haar decomposition of

$\Phi_{b_{j+1}}$,

$$\Phi_{b_{j+1}} = \sum_{m=0}^{\infty} \sum_{I \in \Delta_m} b_I \psi_I, \tag{5.20}$$

where $b_I = 2^{\operatorname{rank} I} \langle \Phi_{b_{j+1}}, \psi_I \rangle = 2^{\operatorname{rank} I} \sum_{k=0}^{b_{j+1}} \sum_{J \in \Delta_k} c_J \langle \varphi_J, \psi_I \rangle$, and c_J is either 0 or 1. Here we sum from $m = 0$, since $\operatorname{supp} \Phi_k \subset [0,1]$ and $\int_0^1 \Phi_k(t)\, dt = 0$ for any $k \in \mathbb{Z}_+$. If we put

$$\tilde{S}_k = \sum_{m=0}^{k} \sum_{I \in \Delta_m} b_I \psi_I,$$

we see that $\{\tilde{S}_k, \mathcal{F}_k, |\cdot|\}$ is a dyadic martingale on $(0,1]$. Since $\Phi_{b_{j+1}} \in C^{10}(\mathbb{R})$, the sum on the right-hand side in (5.20) converges to $\Phi_{b_{j+1}}$ uniformly on $\mathbb{R}$ as $k \to \infty$. It follows immediately that $\langle \tilde{S} \rangle_k$ converges uniformly to a bounded limit which we denote by $\langle \tilde{S} \rangle_\infty$.

Our goal here is to prove that the quadratic function of $\tilde{S}$ is "big" on the intervals from $\mathcal{G}_j$, so that we can use the standard dyadic martingale methods to estimate the size of $\bigcup_{J \in \mathcal{G}_j} J$. For any $k \in \mathbb{Z}_+$, the following equality holds

$$\int_0^1 \langle \tilde{S} \rangle_k^2(t)\, dt = \int_0^1 \tilde{S}_k^2(t)\, dt. \tag{5.21}$$

Assume for a moment that we know that

$$\langle \tilde{S} \rangle_\infty^2(t) \geq 4j^2, \quad t \in \bigcup_{J \in \mathcal{G}_j} J. \tag{5.22}$$

Then (5.21) implies that

$$\begin{aligned} 4j^2 \cdot \left| \bigcup_{J \in \mathcal{G}_j} J \right| &\leq \int_{\bigcup_{J \in \mathcal{G}_j} J} \langle \tilde{S} \rangle_\infty^2(t)\, dt \leq \int_0^1 \langle \tilde{S} \rangle_\infty^2(t)\, dt \\ &= \int_0^1 \tilde{S}_\infty^2(t)\, dt = \int_0^1 \Phi_{b_{j+1}}^2(t)\, dt \leq (j+1)^2, \end{aligned}$$

and (5.19) follows immediately. It remains to prove the estimate (5.22).

5.2.3 Proof of (5.13c): inequality (5.22)

First we show that if $c_I = 1$ for some $I \in \Delta_m$, $b_j \leq m \leq b_{j+1} - 1$, then

$$|b_I| \geq \frac{1}{2} \left| \langle \varphi, \psi \rangle \right|. \tag{5.23}$$

Fix such an interval I. For any $J \in \Delta_k,\ k \le m$, the standard calculation gives

$$\begin{aligned}
&|\langle \varphi_J, \psi_I \rangle| \\
&\quad = \left| \int_{\mathbb{R}} \varphi(2^k t - x_J)\psi(2^m t - x_I)\, dt \right| = \left| \int_{\mathbb{R}} \varphi(2^k(t - 2^{-m}x_I) - x_J)\psi(2^m t)\, dt \right| \\
&\quad = 2^{-k} \left| \int_{\mathbb{R}} \varphi(t - 2^{k-m}x_I - x_J)\psi(2^{m-k}t)\, dt \right| \\
&\quad = 2^{-k} \left| \int_{\mathbb{R}} \left(\varphi(t - 2^{k-m}x_I - x_J) - \varphi(-2^{k-m}x_I - x_J) \right) \psi(2^{m-k}t)\, dt \right| \\
&\quad = 2^{-k} \left| \int_{\mathbb{R}} \int_{-2^{k-m}x_I - x_J}^{t-2^{k-m}x_I - x_J} \varphi'(s)\, ds\, \psi(2^{m-k}t)\, dt \right| \le 2^{-k} \|\varphi'\|_\infty \int_{\mathbb{R}} \left| t\psi(2^{m-k}t) \right| \, dt \\
&\quad = 2^{k-2m} \|\varphi'\|_\infty \int_{\mathbb{R}} |t\psi(t)|\, dt \le 2^{k-2m} \|\varphi'\|_\infty.
\end{aligned}$$

Now we see that if $k > m$, then $\langle \varphi_J, \psi_I \rangle = 0$ for any $J \in \Delta_k$, and if $k \le m$, then there exists at most one $J \in \Delta_k$ such that $\langle \varphi_J, \psi_I \rangle \ne 0$. We therefore have

$$\begin{aligned}
|b_I| &= 2^m \left| \sum_{k=0}^{b_{j+1}} \sum_{J \in \Delta_k} c_J \langle \varphi_J, \psi_I \rangle \right| = 2^m \left| \sum_{k \le m,\, J \in \Delta_k,\, J \supset I} c_J \langle \varphi_J, \psi_I \rangle \right| \\
&\ge 2^m \left| \langle \varphi_I, \psi_I \rangle \right| - 2^m \sum_{k \le m-1,\, J \in \Delta_k,\, J \supset I} |c_J| |\langle \varphi_J, \psi_I \rangle| \\
&\ge \langle \varphi, \psi \rangle - 2^m \sum_{k \le m-1,\, J \in \Delta_k,\, J \supset I} |c_J| \|\varphi'\|_\infty 2^{k-2m} \\
&\ge \langle \varphi, \psi \rangle - \|\varphi'\|_\infty \sum_{k \le m-1,\, J \in \Delta_k,\, J \supset I} 2^{k-m} |c_J|.
\end{aligned}$$

It follows from (5.15b) that if $c_J = 1$ then $c_J = 0$ for $J \in \Delta_k$, $m - a < k \le m-1$ (the decomposition of $\Phi_{b_{j+1}}$ has very sparse coefficients). Combined with the choice of a, it gives

$$|\langle \varphi, \psi \rangle| - \|\varphi'\|_\infty \sum_{k \le m-1,\, J \in \Delta_k,\, J \supset I} 2^{k-m} |c_J| \ge \langle \varphi, \psi \rangle - \|\varphi'\|_\infty 2^{-a} \ge \frac{1}{2} |\langle \psi, \varphi \rangle|,$$

and we have (5.23).

Fix any $I \in \mathcal{G}_j$. Again we note that $c_J = 1$ for any $J \in \Delta_m$ such that $\frac{m}{a} \in \mathbb{Z}$, $J \supset I$ and $b_j \le m \le b_{j+1} - 1$. Therefore (5.23) implies that $|b_J| \ge \frac{1}{2} |\langle \varphi, \psi \rangle|$ for such intervals J, and due to (5.14) we have

$$\langle \tilde{S} \rangle_\infty^2(t) \ge \sum_{b_j \le m \le b_{j+1} - 1,\, \frac{m}{a} \in \mathbb{Z},\, J \in \Delta_m,\, t \in J} |b_J|^2 \ge \frac{1}{4} \left(\frac{b_{j+1} - b_j}{a} - 1 \right) |\langle \varphi, \psi \rangle|^2 \ge 100 j^2,$$

for $t \in I$, and we get (5.22).

5.2.4 How to create a Bloch function from Φ_j

Let

$$v_k(x,y) = \left(\Phi_k * P_{(y)}\right)(x), \quad x \in \mathbb{R},\ k \geq k_0.$$

First we show that $v_k \to v$ as $k \to \infty$, where v is a harmonic function.

Fix any $y > 0$. Since c_I is either 0 or 1, we have for natural $m \leq n$

$$\begin{aligned}
&|v_m(x,y) - v_n(x,y)| \\
&\qquad = \left| \sum_{j=m+1}^{n} \sum_{I \in \Delta_j} c_I \varphi_I * P_{(y)} \right|(x) \leq \sum_{j=m+1}^{n} \sum_{I \in \Delta_j} |c_I| \left| \int_{\mathbb{R}} \varphi_I(t) P_{(y)}(x-t)\, dt \right| \\
&\qquad \leq \sum_{j=m+1}^{n} \sum_{i=0}^{2^j-1} \left| \int_{\mathbb{R}} \varphi(2^j t) P_{(y)} \left(x - t - \left(i + \frac{1}{2}\right) 2^{-j} \right) dt \right|.
\end{aligned}$$

A standard calculation gives

$$\begin{aligned}
&\sum_{i=0}^{2^j-1} \left| \int_{\mathbb{R}} \varphi(2^j t) P_{(y)} \left(x - t - \left(i + \frac{1}{2}\right) 2^{-j} \right) dt \right| \\
&= \sum_{i=0}^{2^j-1} \left| \int_{\mathbb{R}} \varphi(2^j t) \left(P_{(y)} \left(x - t - \left(i + \frac{1}{2}\right) 2^{-j} \right) - P_{(y)} \left(x - \left(i + \frac{3}{2}\right) 2^{-j} \right) \right) dt \right| \\
&\leq \sum_{i=0}^{2^j-1} \left| \int_0^{\frac{2^{-j}}{y}} \varphi(2^j y t) \left(P \left(\frac{x}{y} - t - \frac{i+\frac{1}{2}}{y} 2^{-j} \right) - P \left(\frac{x}{y} - \frac{i+\frac{3}{2}}{y} 2^{-j} \right) \right) dt \right| \\
&\leq \sum_{i=0}^{2^j-1} \left| \int_0^{\frac{2^{-j}}{y}} \varphi(2^j y t) \int_{\frac{x}{y} - \frac{i+\frac{3}{2}}{y} 2^{-j}}^{\frac{x}{y} - t - \frac{i+\frac{1}{2}}{y} 2^{-j}} |P'|(s)\, ds\, dt \right| \\
&\leq \int_0^{\frac{2^{-j}}{y}} |\varphi(2^j y t)| \sum_{i=0}^{2^j-1} \int_{\frac{x}{y} - \frac{i+\frac{3}{2}}{y} 2^{-j}}^{\frac{x}{y} - t - \frac{i+\frac{1}{2}}{y} 2^{-j}} |P'|(s)\, ds\, dt \\
&\leq \int_0^{\frac{2^{-j}}{y}} |\varphi(2^j y t)| \int_{\mathbb{R}} |P'|(s)\, ds\, dt \leq C \frac{2^{-j}}{y}, \quad x \in \mathbb{R},\ y > 0,\ j \in \mathbb{N}.
\end{aligned} \tag{5.24}$$

We therefore have

$$|v_m(x,y) - v_n(x,y)| \leq C \sum_{j=m+1}^{n} \frac{2^{-j}}{y}, \quad x \in \mathbb{R},\ y > 0, \tag{5.25}$$

and the uniform convergence follows immediately.

Next we show that v satisfies the $h^\infty_{w_0}$ growth condition. For $y \geq 2^{-k}$ (5.25) implies that

$$|v(x,y) - v_k(x,y)| \leq C, \quad x \in \mathbb{R}.$$

Combined with (5.13b) and definition of w_0 this implies that

$$|v(x,y)| \leq C + |v_k(x,y)| = C + \left|\Phi_k * P_{(y)}\right|(x) \\ \leq C + w_0(2^{-k}) \leq Cw_0(y), \quad x \in \mathbb{R},\ 2^{-k+1} \geq\ y \geq 2^{-k},\ k \in \mathbb{N},$$

and, therefore, $v \in h^\infty_{w_0}$.

Now we prove that $v \in \mathcal{B}(\mathbb{R}^2_+)$. Fix any positive $y \leq 1$ and $m \in \mathbb{Z}$ such that $2^{-m+1} \geq y \geq 2^{-m}$. We have

$$|\nabla v(x,y)| \leq |\nabla v(x,y) - \nabla v_m(x,y)| + |\nabla v_m(x,y)|, \quad x \in \mathbb{R}.$$

Repeating the estimate in (5.24) verbatim we get for any $x \in \mathbb{R}$,

$$\begin{aligned} &|\nabla v(x,y) - \nabla v_m(x,y)| \\ &\qquad \leq \left|\frac{\partial}{\partial y} v(x,y) - \frac{\partial}{\partial y} v_m(x,y)\right| + \left|\frac{\partial}{\partial x} v(x,y) - \frac{\partial}{\partial x} v_m(x,y)\right| \\ &\qquad = \left|\sum_{j=m+1}^{\infty} \sum_{I\in\Delta_j} c_I \varphi_I * \left(\frac{\partial}{\partial y} P_{(y)}\right)\right|(x) + \left|\sum_{j=m+1}^{\infty} \sum_{I\in\Delta_j} c_I \varphi_I * \left(\frac{\partial}{\partial x} P_{(y)}\right)\right|(x) \\ &\qquad \leq \frac{C2^{-m}}{y^2} \leq \frac{C}{y}. \end{aligned} \tag{5.26}$$

Recall that $\varphi \in C^{10}(\mathbb{R})$ and that for any two different intervals $I, J \in \Delta_j$ the supports of φ_I and φ_J are disjoint. A simple rescaling gives

$$\left|\sum_{I\in\Delta_j} c_I \varphi_I * \left(\frac{\partial}{\partial x} P_{(y)}\right)\right|(x) + \left|\sum_{I\in\Delta_j} c_I \varphi_I * \left(\frac{\partial}{\partial y} P_{(y)}\right)\right|(x) \leq C2^j, \quad x \in \mathbb{R},\ y > 0,\ j \in \mathbb{N}.$$

It follows that

$$\left|\frac{\partial}{\partial x} v_m(x,y)\right| + \left|\frac{\partial}{\partial y} v_m(x,y)\right| = \left|\Phi_m * \left(\frac{\partial}{\partial x} P_{(y)}\right)\right|(x) + \left|\Phi_m * \left(\frac{\partial}{\partial y} P_{(y)}\right)\right|(x) \\ \leq C\sum_{j=0}^{m} 2^j = C2^{m+1} \leq \frac{C}{y}, \quad x \in \mathbb{R}.$$

This estimate and (5.26) imply that $v \in \mathcal{B}(\mathbb{R}^2_+)$.

It remains to prove (5.11). Fix any $k \geq k_0$. Since $\|\Phi'_{b_k}\|_\infty \leq 2^{b_k+1}\|\varphi'\|_\infty$, we see that for any $x \in (0,1]$ such that $|\Phi_{b_k}(x)| \geq \frac{k-1}{2}$, there exists an interval $I_x = [x - \rho_k, x + \rho_k]$, $\rho_k = \frac{2^{-b_k-2}}{\|\varphi'\|_\infty}$, such that $|\Phi_{b_k}(t)| \geq \frac{k-3}{2}$, $t \in I_x$. Clearly then

$$\int_{\mathbb{R}\setminus I_x} P_y(x-t)\,dt \leq \frac{1}{4},$$

for $0 < y \le \frac{\rho_k}{10}$. Now if we fix such an x and put $y_k = \frac{2^{-b_k-2}}{10\|\varphi'\|_\infty}$, we have

$$\begin{aligned}
|v_{b_k}(x,y)| &= \left|\Phi_{b_k} * P_{(y)}\right|(x) \\
&\ge \left|\int_{I_x} \Phi_{b_k}(t)P_{(y)}(x-t)\,dt\right| - \left|\int_{\mathbb{R}\setminus I_x} \Phi_{b_k}(t)P_{(y)}(x-t)\,dt\right| \\
&\ge \frac{k-3}{2} - \|\Phi_{b_k}\|_\infty \int_{\mathbb{R}\setminus I_x} P_{(y)}(x-t)\,dt \ge \frac{k-3}{2} - \frac{k+2}{4} \\
&\qquad\qquad = \frac{k-4}{2} \ge \frac{w_0(2^{-b_k})-4}{2},
\end{aligned}$$

so that

$$\left|\left\{x \in (0,1] : |v_{b_k}|(x,y) \ge \frac{w_0(2^{-b_k})-4}{2}\right\}\right| \ge \frac{1}{10}, \quad 0 < y \le y_k.$$

Again, following (5.24), we obtain

$$|v_{\beta_k}(x,y_k) - v(x,y_k)| \le C_0\|\varphi'\|_\infty. \tag{5.27}$$

The doubling property of w_0 implies

$$w_0(y_k) = w_0\left(\frac{2^{-\beta_k-2}}{10\|\varphi'\|_\infty}\right) \le C_1\frac{w_0(2^{-\beta_k})-4}{2} - C_0\|\varphi'\|_\infty \le C\frac{w_0(2^{-\beta_k})-4}{2}$$

for k large enough. We therefore have

$$\left|\left\{x \in (0,1] : |v|(x,y_k) \ge \frac{w_0(y_k)}{C}\right\}\right| \ge \frac{1}{10}, \quad y_k = \frac{2^{-b_k-2}}{10\|\varphi'\|_\infty},$$

which is (5.11). □

The way we did the construction of v is, probably, not the most effective one. Unfortunately we could not use here the dyadic martingale methods, as described, for example, in [36]. Instead we decided to employ the wavelet-like series for Bloch functions (see [69] for the description of $\mathcal{B}(\mathbb{R}^2_+)$ in terms of wavelet representation).

Note that the averaging process $u(x,\delta) \to H(x,\delta) = \int_0^1 u(x,y+\delta)\,d\frac{1}{w(y)}$ can be viewed as an application of some multiplier M to the boundary values of u,

$$\widehat{Mf}(\tau) = \widehat{f}(\tau)\int_0^1 e^{-2\pi y|\tau|}d\frac{1}{w(y)}, \quad \tau \in \mathbb{R},$$

where $f = u(\cdot,0)$ (these boundary values exist in some sense, at least as a distribution). The doubling condition (I.37) implies that

$$\int_0^1 e^{-2\pi y|\tau|}d\frac{1}{w(y)} \approx \frac{1}{w\left(\frac{1}{|\tau|}\right)}, \quad |\tau| > 0,$$

so we basically divide the Fourier transform of $u(\cdot,0)$ by w. It would be interesting to find out the image of M, we see at least that $Mu \in h_w^\infty(\mathbb{R}_+^{n+1}) \bigcap \mathcal{B}(\mathbb{R}_+^{n+1})$ if $u \in h_w^\infty(\mathbb{R}_+^{n+1})$. The example in Proposition 5.2.1 shows that the image of M can (in the case of slowly growing weights) be a proper subset of $h_w^\infty(\mathbb{R}_+^{n+1}) \bigcap \mathcal{B}(\mathbb{R}_+^{n+1})$. For more information about the multipliers on the growth spaces see [30].

Chapter 6 Growth classes: divided differences

6.1 Proof of Theorem 6.1.1

The proof consists of two parts. First we show the dyadic martingale version of Theorem I.19. Let us recall its statement.

Theorem 6.1.1 *Let $0 < a < 1$. Then there exists a function $f \in \mathrm{Hol}_a(\mathbb{R})$ such that at almost every $x \in \mathbb{R}$ one has*

$$\limsup_{h\to 0^+} \mathrm{D}_a(f)(x,h) > 0$$

and

$$\liminf_{h\to 0^+} \mathrm{D}_a(f)(x,h) = 0.$$

Then we approximate the a-divided differences by their discrete versions arriving at the continuous statement.

Lemma 6.1.1 *Let $0 < \varepsilon < 1$. Then there exists a dyadic martingale $\{S_n\}$ with $\sup_n 2^{-n\varepsilon}\|S_n\|_\infty < \infty$, such that*

$$\limsup_{n\to\infty} 2^{-n\varepsilon} S_n(x) > 0,$$

and

$$\liminf_{n\to\infty} 2^{-n\varepsilon} S_n(x) \geq 0$$

for almost every $x \in \mathbb{R}$. Actually the following uniform version of the last inequality holds: for any $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$2^{-n\varepsilon} S_n(x) \geq -\delta \quad \text{for any} \;\; n \geq n_0, \;\; x \in \mathbb{R}.$$

Proof. It is enough to define $\{S_n\}$ on the unit interval $[0,1)$. It will be constructed via a double induction argument. More precisely, we define a pair of increasing sequences $\{k_{jn}\}_{1\leq n\leq n_j}$ and M_j, $n \in \mathbb{Z}_+$, of natural numbers satisfying

$$\begin{aligned} &k_{00} + M_0 \leq k_{01} \leq k_{01} + M_0 \leq \cdots \leq k_{0n_0} + M_0 \leq k_{10} \leq k_{10} + M_1 \cdots \leq k_{1n_1} + M_1 \leq \ldots \\ &\leq k_{20} \leq \cdots \leq k_{(j-1)n_{j-1}} + M_{j-1} \leq k_{j0} \leq \cdots \leq k_{jn_j} + M_j \leq \ldots, \end{aligned}$$

and a martingale $\{S_m\}$ such that: *(a)* for any $n \geq 0$ there exists $j \geq 0$ such that $2^{-m\varepsilon} S_m \geq -2^{-n}$ for $m \geq k_{jn_j} + M_j$, and *(b)* $2^{-m\varepsilon} S_m \geq \frac{1}{3}$ for at least one number m between k_{j0} and $k_{jn_j} + M_j$ on a large portion of $[0,1)$. We start describing the building block of our construction.

Block construction

Given a dyadic interval J with $|J| = 2^{-K}$ and a number $\delta > 0$ we define a *building block* $W(\delta, J)$ as follows.

Consider a nested sequence of dyadic subintervals of J that shrinks to its left end-point. In other words, let $J_0 := J$, and, given J_{k-1} define $J_k := J_{k-1}^-$, $k \geq 1$ (where I^- is the left half-interval of I). Let $M = M(\delta) := \left[\frac{\log \frac{1}{2\delta}}{(1-\varepsilon)\log 2}\right] + 1$, so that

$$\frac{1}{2} \leq 2^{M(1-\varepsilon)} \cdot \delta \leq \frac{1}{2^\varepsilon}.$$

Now let h_I be a (slightly renormalized) Haar function corresponding to a dyadic interval I, $h_I(x) = 2\chi_{I^-} - \chi_I$, and define

$$s_{k,J}(x) := \delta \cdot 2^{K\varepsilon} \cdot 2^k h_{J_k}(x), \quad 0 \leq k \leq M.$$

Since $|J_k| = 2^{-K-k}$, then, clearly, $s_{k,J}$ is a martingale difference of rank $K + k$, and

$$\|2^{-(K+k)\varepsilon} \sum_{m=0}^{k} s_{m,J}\|_\infty \leq 2\delta 2^{k(1-\varepsilon)} \leq 2^{1-\varepsilon}, \quad 0 \leq k \leq M. \tag{6.1}$$

On the other hand

$$2^{-(K+k)\varepsilon} \sum_{m=0}^{k} s_{m,J} \geq -2^{-k\varepsilon} \cdot \delta \geq -\delta. \tag{6.2}$$

Define

$$W(\delta, J) := \sum_{k=0}^{M} s_{k,J},$$

and observe that

$$2^{-(M+K)\varepsilon} \|W(\delta, J)\|_\infty \leq 2^{1-\varepsilon}, \tag{6.3}$$

and

$$2^{-(M+K)\varepsilon} W(\delta, J)(x) \geq \frac{1}{2}\left(1 - 2^{-M}\right), \quad x \in J_M. \tag{6.4}$$

In particular, $|J_M| = 2^{-M}|J|$. To summarize, we have constructed a step function $W(\delta, J)$ supported on J whose values are $-\delta 2^{K\varepsilon}$ on $J \setminus J_M$, and $\delta 2^{K\varepsilon}(2^M - 1)$ on J_M. Since $2^{M(1-\varepsilon)}\delta \approx 1$, we have $\delta 2^{K\varepsilon} \approx |J|^{-\varepsilon} 2^{M(\varepsilon-1)}$ and therefore $\delta 2^{K\varepsilon}(2^M - 1) \approx |J_M|^{-\varepsilon}$.

Arranging the blocks, first step

Let $\delta_j := 2^{-j-2}$, $j \in \mathbb{Z}_+$. We define a (very lacunary) sequence k_{mn} of numbers in the following way. Put $k_{00} := 0$, $J = [0, 1)$, and

$$S_{M(\delta_0)} := W(\delta_0, J).$$

Now let k_{01} be such that

$$2^{-k_{01}\varepsilon} \|S_{M(\delta_0)}\|_\infty \leq \frac{\delta_0}{2}.$$

Then we let

$$
\begin{aligned}
S_i &:= S_{M(\delta_0)}, \quad M(\delta_0) \le i \le k_{01}, \\
S_{k_{01}+M(\delta_0)} &:= S_{M(\delta_0)} + \sum_{J \in \Delta_{k_{01}}} W(\delta_0, J),
\end{aligned}
$$

We remind that Δ_i is the collection of dyadic intervals of rank i.
We continue iterating the procedure. To elaborate, assume we have defined numbers k_{0n} and the martingale S_i with $0 \le i \le k_{0n} + M(\delta_0)$. Then we pick $k_{0(n+1)}$ in such a way that

$$
2^{-k_{0(n+1)}\varepsilon} \|S_{k_{0n}+M(\delta_0)}\|_\infty \le \frac{\delta_0}{2},
$$

and

$$
\begin{aligned}
S_i &:= S_{k_{0n}+M(\delta_0)}, \quad k_{0n} + M(\delta_0) \le i < k_{0(n+1)}, \\
S_{k_{0(n+1)}+M(\delta_0)} &:= S_{k_{0n}+M(\delta_0)} + \sum_{J \in \Delta_{k_{0(n+1)}}} W(\delta_0, J).
\end{aligned}
$$

We repeat the construction until we have $n = n_0 := \left[\frac{\log(1-2^{-M(\delta_0)})}{\log \delta_0}\right] + 1$.

Arranging the blocks, second step

We continue to iterate, now also with respect to the parameter j. Assume that we have defined a sequence of numbers $\{k_{mn}\}_{m=0}^{j-1} = \{\{k_{0n}\}_{n=0}^{n_0}, \dots, \{k_{(j-1)n}\}_{n=0}^{n_{j-1}}\}$ and a sequence of partial sums $\{S_i\},\ i = 0, \dots, k_{(j-1)n_{j-1}} + M(\delta_{j-1})$. We apply the procedure from the previous step, now using δ_j in place of δ_0. In other words, we fix a number $k_{j0} \ge k_{(j-1)n_{j-1}}$ such that

$$
2^{-k_{j0}\varepsilon} \|S_{k_{(j-1)n_{j-1}}+M(\delta_{j-1})}\| \le \frac{\delta_j}{2},
$$

and define S_i for $k_{(j-1)n_{j-1}} + M(\delta_{j-1}) \le i \le k_{j0} + M(\delta_j)$ as above. Then we proceed to k_{j1} and so on, until we have $n = n_j = \left[\frac{\log(1-2^{-M(\delta_j)})}{\log \delta_j}\right] + 1$ (by our assumptions $m_j = M(\delta_j) \approx j$, and $n_j \approx j2^j$).

Behaviour of $\{S_m\}$

First we claim that S_i satisfies the growth condition, that is

$$
\sup_i 2^{-i\varepsilon} \|S_i\|_\infty \le 2^{1-\varepsilon}.
$$

Indeed, fix a number i and consider the largest k_{jn} such that $k_{jn} \le i$. We have two options: *(a)* $k_{jn} + M(\delta_j) < i$, and *(b)* $k_{jn} + M(\delta_j) \ge i$. For the option (a) the martingale just stops until we hit the next number $k_{j(n+1)}$ or $k_{(j+1)0}$, in any case, clearly, $S_i = S_{k_{jn}+M(\delta_j)}$, and we have

$$2^{-i\varepsilon}\|S_i\|_\infty = 2^{-i\varepsilon}\|S_{k_{jn}+M(\delta_j)}\|_\infty \leq 2^{-(k_{jn}+M(\delta_j))\varepsilon}\|S_{k_{jn}+M(\delta_j)}\|_\infty \leq$$
$$2^{-(k_{jn}+M(\delta_j))\varepsilon}\left(\|S_m\|_\infty + \|S_{k_{jn}+M(\delta_j)} - S_m\|_\infty\right),$$

where $m = m(n,j)$ is either $k_{j(n-1)} + M(\delta_j)$, if $n \geq 1$, or $k_{(j-1)n_{j-1}} + M(\delta_{j-1})$, if $n = 0$. In both cases k_{jn} was chosen in such a way that

$$2^{-(k_{jn}+M(\delta_j))\varepsilon}\|S_m\|_\infty \leq 2^{-k_{jn}\varepsilon}\|S_m\|_\infty \leq \frac{\delta_j}{2}.$$

On the other hand, by construction we have

$$S_{k_{jn}+M(\delta_j)} - S_m = \sum_{J\in\Delta_{k_{jn}}} W(\delta_j, J),$$

hence $\|S_{k_{jn}+M(\delta_j)} - S_m\|_\infty = \|W(\delta_j,J)\|_\infty$ for any $J \in \Delta_{k_{nj}}$. By our choice of $W(\delta_j, J)$ (see (6.3)) we have

$$2^{-(k_{nj}+M(\delta_j))\varepsilon}\|W(\delta_j,J)\|_\infty \leq 2^{1-\varepsilon}.$$

Option (b) is handled in the same way, only now we use estimate (6.1) instead.
Next we aim to show that

$$\liminf_{i\to\infty} 2^{-i\varepsilon}S_i(x) \geq 0, \quad x \in [0,1).$$

Again, it follows from our construction, since the martingale S_i consists of very sparse and independent pieces, and by the choice of k_{jn} we can always consider only the tail end of it. In particular, if $i \geq k_{jn} + M(\delta_j)$ for some j, n, then by (6.2) we have $2^{-i\varepsilon}W(\delta_j, J) \geq -2\delta_j$ for any $J \in \Delta_{k_{jn}}$, hence using the previous argument we get $2^{-i\varepsilon}S_i \geq -3\delta_j$, which proves the estimate, as well as the last part of the statement.

Finally we want to estimate the size of the set E of points $x \in \mathbb{R}$ where $\limsup_{i\to\infty} 2^{-i\varepsilon}S_i(x) \geq \frac{1}{5}$. Fix a pair of numbers $j \in \mathbb{Z}_+$ and $0 \leq n \leq n_j - 1$. Since, as before,

$$2^{-(k_{jn}+M(\delta_j))\varepsilon}\|S_{k_{jn}+M(\delta_j)}\|_\infty \geq 2^{-(k_{jn}+M(\delta_j))\varepsilon}\|W(\delta_j,J)\|_\infty - \frac{\delta_j}{2}$$

for any $J \in \Delta_{k_{jn}}$, we can only consider the respective building block $W(\delta_j, J)$. Now, if $|J| = 2^{-k_{jn}}$, we have seen in (6.4) that $2^{-(k_{jn}+M(\delta_j))\varepsilon}W(\delta_j, J) \geq \frac{1}{4}$ on the interval $J_{M(\delta_j)}$ with $|J_{M(\delta_j)}| = 2^{-M(\delta_j)}|J|$. On the other hand, if I is the dyadic interval of the next construction step in J, that is $|I| = 2^{-k_{j(n+1)}}$, $I \subset J$, again by (6.4) we have $2^{-(k_{j(n+1)}+M(\delta_j))\varepsilon}W(\delta_j, I) \geq \frac{1}{4}$ on $I_{M(\delta_j)}$. Denote by $\mathcal{F}(J)$ the set of all such intervals, that is,

$$\mathcal{F}(J) = \{I_{M(\delta_j)} \subset I : I \in \Delta_{k_{j(n+1)}}(J)\},$$

where $\Delta_m(J)$ is the collection of dyadic intervals of rank m that lie inside J. The intervals in $\mathcal{F}(J)$ are disjoint, and they are uniformly distributed over J (for any $I \in \Delta_{k_{j(n+1)}}(J)$ recall that $I_{M(\delta_j)}$

is a leftmost dyadic subinterval of I of rank $k_{j(n+1)} + M(\delta_j)$). It follows that

$$\left|\left(\bigcup_{\mathcal{F}(J)} I_{M(\delta_j)}\right) \setminus J_{M(\delta_j)}\right| = \sum_{I_{M(\delta_j)} \in \mathcal{F}(J),\ I_{M(\delta_j)} \subset J \setminus J_{M(\delta_j)}} |I_{M(\delta_j)}| = \tag{6.5}$$
$$(2 - 2^{-M(\delta_j)})2^{-M(\delta_j)}|J \setminus J_{M(\delta_j)}|.$$

An interval I' is called δ_j-*special*, if there exists a number $0 \le n \le n_j$ and an interval $J \in \Delta_{k_{jn}}$ such that $I' = J_{M(\delta_j)}$, that is I' is the left-most dyadic subinterval of J of rank $k_{jn} + M(\delta_j)$. The collection of δ_j-special intervals is denoted by $\mathcal{F}_j$. As before, $|I'|^\varepsilon W(\delta_j, J) \ge \frac{1}{4}$ on I', and therefore $|I'|^\varepsilon S(I') \ge \frac{1}{5}$ (where $S(I) := S_i(x)$ with $x \in I$ and $i = \log_2 |I|^{-1}$). It follows from (6.5) that

$$\left|\bigcup_{I' \in \mathcal{F}_j} I'\right| \ge 1 - (1 - 2^{-M(\delta_j)})^{n_j}.$$

Therefore the set F_j of points $x \in [0, 1)$ where

$$2^{-(k_{jn} + M(\delta_j))\varepsilon} \sum_{J \in \Delta_{k_{jn}}} W(\delta_j, J)(x) \le \frac{1}{4}$$

for all $n = 0, \ldots, n_j$, has small Lebesgue measure, namely

$$|F_j| \le (1 - 2^{-M(\delta_j)})^{n_j} \lesssim \delta_j$$

by our choice of n_j. Hence

$$\left|\left\{x : 2^{-i\varepsilon} S_i(x) \le \frac{1}{5},\ k_{j0} \le i \le k_{jn_j}\right\}\right| \lesssim \delta_j.$$

Since $\sum_j \delta_j \le 1$, we see immediately that

$$\left|\left\{x :\ \limsup_{i \to \infty} 2^{-i\varepsilon} S_i(x) \le \frac{1}{5}\right\}\right| = 0.$$

We make another observation which will be useful later. Given a δ_j-special interval I' consider the dyadic interval $\tilde{I}$ of the same length that lies immediately on the left of I', in other words, if $I' = [i2^{-m}, (i+1)2^{-m})$, then $\tilde{I} := [(i-1)2^{-m}, i2^{-m})$ (if the left end-point of I' is 0, we put $\tilde{I} := \emptyset$, so the intervals that fall out of $[0, 1)$ are discarded). These intervals are called *left-δ_j-special*, and their collection is denoted by $\tilde{\mathcal{F}}_j$. Arguing as above we see that

$$\left|[0, 1) \setminus \left(\bigcup_{\tilde{I} \in \tilde{\mathcal{F}}_j} \tilde{I}\right)\right| \le 2(1 - 2^{-M(\delta_j)})^{n_j} \lesssim \delta_j, \tag{6.6}$$

so that almost every point $x \in [0,1)$ lies in $\bigcup_{\tilde{I}\in\tilde{\mathcal{F}}_j} \tilde{I}$ for infinitely many $j \in \mathbb{Z}_+$. $\square$

Now we are ready to prove Theorem 6.1.1.

Proof of Theorem 6.1.1. Fix $\varepsilon := 1 - a$. Consider the martingale $\{S_n\}$ constructed in Lemma 6.1.1. We can assume $S_0 = 0$. We will define a function f defined in the real line as follows. Let $f(0) = 0$. The relation $f(b_n) - f(a_n) := 2^{-n} S_n(I)$ for any $I = [a_n, b_n) \in \Delta_n$, $n \geq 0$, defines f on the dyadic points of $[0,1]$ and we extend f to non-dyadic points of $[0,1]$ by continuity. Observe that since $S_0 = 0$ we have $f(0) = f(1) = 0$. Finally we extend f from $[0,1]$ to the whole real line by periodicity. Let us prove that $f \in \mathrm{Hol}_\alpha$. Fix a point $x \in \mathbb{R}$ and a number $0 < h \leq 1$. We aim to show that $|f(x+h) - f(x)| \leq Ch^a$ for some absolute constant $C > 0$. There exists an increasing sequence of dyadic-rational points $\{a_k\}_{k\in\mathbb{Z}}$ such that $[a_{k-1}, a_k) \in \Delta$, $\lim_{k\to-\infty} a_k = x$, $\lim_{k\to+\infty} a_k = x+h$, and for any $n \in \mathbb{N}$ there exists at most 4 dyadic intervals of rank n of the form $[a_{k-1}, a_k)$. In other words, we consider a Whitney decomposition of the interval $[x, x+h)$ with $\{a_k\}$ being the endpoints of the corresponding dyadic intervals. Given $k \in \mathbb{Z}$ denote by r_k the length of the interval $[a_{k-1}, a_k)$, that is $r_k = a_k - a_{k-1}$. Clearly,

$$|f(x+h) - f(x)| = \left|\sum_{k\in\mathbb{Z}} (f(a_k) - f(a_{k-1}))\right| \leq \sum_{k\in\mathbb{Z}} |f(a_k) - f(a_{k-1})| =$$
$$= \sum_{k\in\mathbb{Z}} r_k |S([a_{k-1}, a_k))|.$$

Since by construction $\sup_n 2^{-n(1-a)} \|S_n\|_\infty < \infty$, and the amount of points a_k generating the dyadic intervals of rank n is bounded, there exists a constant $C = C(a) > 0$ such that

$$|f(x+h) - f(x)| \leq C \sum_{n \geq \log_2 \frac{1}{h}} 2^{-n} 2^{n(1-a)} \leq \frac{C}{1 - 2^{-a}} h^a,$$

so f belongs to the corresponding Hölder class Hol_a. Next we show that

$$\liminf_{h\to 0^+} \frac{f(x+h) - f(x)}{h^a} = 0, \quad x \in [0,1). \tag{6.7}$$

Fix any $x \in [0,1)$ and an arbitrarily small $\delta > 0$. By the last part of Lemma 6.1.1 there exists a number N such that $2^{-n(1-a)} S_n(t) \geq -\delta$ for any $n \geq N$ and $t \in [0,1)$. Now fix any $0 < h \leq 2^{-N}$, and consider the Whitney decomposition of $[x, x+h)$ as before. Clearly, $r_k \leq 2^{-N}$ for all $k \in \mathbb{Z}$, therefore we have

$$f(x+h) - f(x) = \sum_{k\in\mathbb{Z}} (f(a_k) - f(a_{k-1})) = \sum_{k\in\mathbb{Z}} r_k \frac{f(a_k) - f(a_{k-1})}{r_k} =$$
$$\sum_{k\in\mathbb{Z}} r_k S([a_{k-1}, a_k)) = \sum_{k\in\mathbb{Z}} 2^{-n_k} S([a_{k-1}, a_k)) \geq -\delta \sum_{k\in\mathbb{Z}} 2^{-n_k} 2^{n_k(1-a)},$$

where $2^{-n_k} = r_k$ and $\sup_k r_k \leq h$. Since the numbers n_k do not accumulate (we recall that for any

n there are at most four numbers $n_k = n$), it follows that

$$\sum_{k\in\mathbb{Z}} 2^{-n_k} 2^{n_k(1-a)} \le Ch^a$$

for some absolute constant $C > 0$, and (6.7) follows immediately.

It remains to show that for almost every $x \in [0,1]$ we have

$$\limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h^a} > \frac{1}{20}. \tag{6.8}$$

Fix a point $x \in [0,1)$ and a number N such that $2^{-n(1-a)}S_n(t) \ge -\frac{1-a}{40}$ for any $n \ge N$ and $t \in [0,1)$. It follows from (6.6) that almost every x belongs to infinitely many left-δ_j-special intervals, in particular there is an increasing sequence $\{j_m(x)\}_{m=0}^{\infty}$ such that $x \in \tilde{I}_m(x) \in \tilde{\mathcal{F}}_{j_m}$ and $|\tilde{I}_m| \le 2^{-N}$. Now for any j_m we define h_m in such a way that $x + h_m$ is the right end-point of the δ_{j_m}-special interval I_m corresponding to $\tilde{I}_m$. In other words, if $\tilde{I}_m = [(i-1)|\tilde{I}_m|, i|\tilde{I}_m|)$ for some $i \in \mathbb{Z}_+$, then $h_m := (i+1)|\tilde{I}_m| - x$. Since I_m is δ_{j_m}-special, we have $|I_m|^{1-a}S(I_m) \ge \frac{1}{5}$. Consider a Whitney-type decomposition of $[x, x+h_m)$ generated by $\{a_k\}_{k\in\mathbb{Z}}$ as above. In this case, since $x + h_m$ is dyadic-rational, we assume $a_0 = a_1 = \cdots = x + h_m$, also, clearly, $[a_{-1}, a_0) = I_m$ and $a_k - a_{k-1} = r_k \le |I_m| \le 2^{-N}$ for any $k \le 0$. In particular, $r_k S([a_{k-1}, a_k)) \ge -r_k^a \frac{1-a}{40}$, $k < 0$. We therefore have

$$\begin{aligned} f(x+h_m) - f(x) &= \sum_{k\le 0} (f(a_k) - f(a_{k-1})) = \\ &= r_0 \frac{f(a_0) - f(a_{-1})}{r_0} + \sum_{k<0} r_k \frac{f(a_k) - f(a_{k-1})}{r_k} = \\ &= |I_m| S(I_m) + \sum_{k<0} r_k S([a_{k-1}, a_k)) \ge \frac{1}{5}|I_m|^a - \frac{1-a}{40} \sum_{k<0} r_k^a. \end{aligned}$$

Since for any given rank there are at most 4 dyadic intervals of this rank of the form $[a_{k-1}, a_k)$, we have

$$\sum_{k<0} r_k^a \le 4 \sum_{n \ge \log_2 |I_m|^{-1}} 2^{-na} \le \frac{4}{1-a} |I_m|^a.$$

Hence

$$f(x+h_m) - f(x) \ge \frac{1}{5}|I_m|^a - \frac{1}{10}|I_m|^a = \frac{1}{10}|I_m|^a \ge \frac{1}{20} h_m^a,$$

because $h_m \le |I_m| + |\tilde{I}_m| = 2|I_m|$. This finishes the proof of Theorem 6.1.1.

□

6.2 Proof of Theorem 6.2.1

Let us recall the formulation of the Theorem.

Theorem 6.2.1 *Let $0 < a < 1$. Then there exists a function $f \in \mathrm{Hol}_a(\mathbb{R})$ and a constant $C > 0$ such that for any point $x \in \mathbb{R}$ there exist two sequences $\{h_k\}_{k=1}^{\infty}$, $\{h'_k\}_{k=1}^{\infty}$ of positive numbers,* *converging to zero, such that*

$$\limsup_{k\to\infty}\left|\frac{f(x+h'_k)-f(x)}{h'_k}\right|\le 1,$$
$$\liminf_{k\to\infty}\frac{|f(x+h_k)-f(x)|}{|h_k|^\alpha}>C. \tag{6.9}$$

We will construct the function f via a rarefied (with respect to space variable) and lacunary (with respect to frequency scale variable) wavelet series. In fact it will be an analogue of the classical Weierstrass functions which admits better control over the individual atoms. We start by defining the base wavelet $\varphi\in C^\infty(\mathbb{R})$ that satisfies the following conditions

$$\operatorname{supp}\varphi=\left[-\frac12,\frac12\right],\quad \varphi\equiv 1 \text{ on } \left[-\frac{1}{16},\frac{1}{16}\right],\quad \varphi\equiv -1 \text{ on } \left[-\frac{7}{16},-\frac38\right]\cup\left[\frac38,\frac{7}{16}\right],$$
$$\int_{\mathbb{R}} x^n\varphi(x)\,dx=0,\ 0\le n\le 2.$$

It is easy to verify (see e.g. [42]) that for any sequence $\{c_{jk}\},\ j\in\mathbb{Z},k\in\mathbb{Z}_+$, satisfying $|c_{jk}|\le 2^{-ka}$, $k\in\mathbb{N}$, the function

$$f:=\sum_{j\in\mathbb{Z},k\in\mathbb{N}} c_{jk}\varphi_{jk},\quad \text{where } \varphi_{jk}(t):=\varphi(2^kt-j),$$

belongs to Hol_a.

We consider a *superlacunary* sequence k_n of positive integers that will be defined by induction. We put $k_1:=1$. We set $k_n\ge k_{n-1}+4$ to satisfy a certain condition (6.10) that we announce in a few lines. Next we put $c_{jk}:=2^{-ka}$, if $k=k_n$ for some $n\in\mathbb{N}$, and $c_{jk}\equiv 0$ otherwise, and we let

$$f:=\sum_{n=1}^{\infty}\sum_{j\in\mathbb{Z}} c_{jk_n}\varphi_{jk_n}.$$

For any $m\ge 2$ we define $S_m:=\sum_{n=1}^{m-1}\sum_{j\in\mathbb{Z}}c_{jk_n}\varphi_{jk_n}$ and $R_m:=\sum_{n=m}^{\infty}\sum_{j\in\mathbb{Z}}c_{jk_n}\varphi_{jk_n}$ to be the main part and the tail of the series representing f.

Assume we have defined k_n for $n=1\ldots m-1$ (and therefore S_m) for some $m\ge 2$. We pick k_m to satisfy the following conditions:

$$2^{-k_m}\cdot\|S'_m\|_\infty\le\tilde\varepsilon 2^{-k_ma}$$
$$\sup_{|\theta|\le 10\cdot 2^{-k_m}}|S'_m(t_0+\theta)|\le\varepsilon,\quad \text{for every } t_0 \text{ such that } S'_m(t_0)=0 \tag{6.10}$$

for some very small absolute constant $\tilde\varepsilon>0$ to be chosen later. Observe that for any m the functions φ_{jk_m} have disjoint supports, and there are nested sequences $\{J_n^\pm\}$ of intervals of length $\frac18 2^{-k_n}$ such that $\varphi_{j_n^\pm k_n}\equiv\pm1$ on $J_n^\pm$ for some $j_n^\pm$ with $n\le m$.

Fix any point $x\in\mathbb{R}$. Given $m\in\mathbb{N}$ there exist four numbers $r_m^\pm=r_m^\pm(x)$, $\rho_m^\pm=\rho_m^\pm(x)$, such that $R_m(x+r_m^+)=\sup_{2^{-k_m}<t\le 2^{-k_m+1}}R_m(x+t)$, $R_m(x+r_m^-)=\inf_{2^{-k_m}<t\le 2^{-k_m+1}}R_m(x+t)$, and $R_m(x-\rho_m^+)=\sup_{2^{-k_m}<t\le 2^{-k_m+1}}R_m(x-t)$, $R_m(x-\rho_m^-)=\inf_{2^{-k_m}<t\le 2^{-k_m+1}}R_m(x-t)$. In other

words, $x + r_m^{\pm}$ is the maximum/minimum point of the 2^{-k_m}-periodic function R_m on the interval $[x + 2^{-k_m}, x + 2^{-k_m+1}]$ (and $x - \rho_m^{\pm}$ on the interval $[x - 2^{-k_m+1}, x - 2^{-k_m}]$ respectively). Clearly $2^{-k_m} \le |r_m^{\pm}|, |\rho_m^{\pm}| \le 2^{-k_m+1}$, and

$$R_m(x - \rho_m^{\pm}) = R_m(x + r_m^{\pm}) = \pm \sum_{n=m}^{\infty} 2^{-k_n a}.$$

Clearly

$$\frac{f(x + r_m^{\pm}) - f(x)}{r_m^{\pm}} = \frac{S_m(x + r_m^{\pm}) - S_m(x)}{r_m^{\pm}} + \frac{R_m(x + r_m^{\pm}) - R_m(x)}{r_m^{\pm}} := (I^{\pm}) + (II^{\pm}),$$

and we have $(II^+) \ge 0$, $(II^-) \le 0$ by the definition of $r_m^{\pm}$. Consider the following possible situations:

(i) For one of the numbers $r_m^{\pm}$ we have

$$\left| \frac{f(x + r_m^{\pm}) - f(x)}{r_m^{\pm}} \right| \le 1. \tag{6.11}$$

(ii) We have

$$\frac{f(x + r_m^{+}) - f(x)}{r_m^{+}} > 1, \qquad \frac{f(x + r_m^{-}) - f(x)}{r_m^{-}} < -1,$$

or

$$\frac{f(x + r_m^{+}) - f(x)}{r_m^{+}} < -1, \qquad \frac{f(x + r_m^{-}) - f(x)}{r_m^{-}} > 1.$$

(iii) For both $r_m^{\pm}$

$$\frac{f(x + r_m^{\pm}) - f(x)}{r_m^{\pm}} > 1 \tag{6.12}$$

or

$$\frac{f(x + r_m^{\pm}) - f(x)}{r_m^{\pm}} < -1.$$

Case (i). Assume that the inequality holds, say, for r_m^+. We claim that in this case

$$\frac{f(x + r_m^{-}) - f(x)}{r_m^{-}} \le -\frac{1}{2} 2^{k_m(1-a)}. \tag{6.13}$$

Indeed, by (6.10) the sequence $\{k_n\}$ is chosen in such a way that for $0 \le \theta \le 2^{-k_m+1}$ we have

$$|S_m(t + \theta) - S_m(t)| \le \int_0^{\theta} |S'(t + s)|\, ds \le 2\tilde{\varepsilon} \cdot 2^{-k_m a}, \quad t \in \mathbb{R}.$$

On the other hand, clearly,

$$\sup_{t \in \mathbb{R}} |R_m(t)| = \sum_{n=m}^{\infty} 2^{-k_n a} \ge 2^{-k_m a}.$$

Take $\tilde{\varepsilon} \leq \frac{1}{200}$. It follows immediately that $|S_m(x+r_m^+) - S_m(x+r_m^-)| \leq \frac{1}{50} 2^{-k_m a}$, therefore

$$\frac{f(x+r_m^-) - f(x)}{r_m^-} = \frac{r_m^+}{r_m^-} \frac{f(x+r_m^+) - f(x)}{r_m^+} + \frac{f(x+r_m^-) - f(x+r_m^+)}{r_m^-} =$$
$$= \frac{r_m^+}{r_m^-} \frac{f(x+r_m^+) - f(x)}{r_m^+} + \frac{S_m(x+r_m^-) - S_m(x+r_m^+)}{r_m^-} + \frac{R_m(x+r_m^-) - R_m(x+r_m^+)}{r_m^-}.$$

Since $\frac{r_m^+}{r_m^-} \leq 2$, we obtain

$$\left| \frac{r_m^+}{r_m^-} \frac{f(x+r_m^+) - f(x)}{r_m^+} + \frac{S_m(x+r_m^-) - S_m(x+r_m^+)}{r_m^-} \right| \leq \frac{1}{10} 2^{k_m(1-a)}$$

On the other hand, $R_m(x+r_m^-) - R_m(x+r_m^+) = -2\sum_{n=m}^{\infty} 2^{-k_n a} \leq -2^{-k_m a+1}$, therefore we deduce

$$\frac{f(x+r_m^-) - f(x)}{r_m^-} \leq \frac{R_m(x+r_m^-) - R_m(x+r_m^+)}{r_m^-} + \frac{1}{10} 2^{k_m(1-a)} \leq -\frac{1}{2} 2^{k_m(1-a)}.$$

This proves (6.13) and we put $h_m := r_m^-$ and $h'_m := r_m^+$. If (6.11) is attained at r_m^-, we repeat the argument above exchanging r_m^+ and r_m^-.

Case (ii). Clearly there must exist a point $x + \tilde{r}_m$ between $x + r_m^+$ and $x + r_m^-$ such that

$$f(x+\tilde{r}_m) - f(x) = 0,$$

we immediately put $h'_m := \tilde{r}_m$. On the other hand

$$\max\left\{|R_m(x+\tilde{r}_m) - R_m(x+r_m^+)|, |R_m(x+\tilde{r}_m) - R_m(x+r_m^-)|\right\} \geq \sup_{t\in\mathbb{R}} |R_m(t)| \geq 2^{-k_m a}. \quad (6.14)$$

Assume that the maximum is attained at r_m^+. Then

$$f(x+r_m^+) - f(x) = f(x+r_m^+) - f(x+\tilde{r}_m) = S_m(x+r_m^+) - S_m(x+\tilde{r}_m) + R_m(x+r_m^+) - R_m(x+\tilde{r}_m),$$

and arguing as in the case *(i)* we have

$$\left| \frac{f(x+r_m^+) - f(x)}{r_m^+} \right| \geq \frac{1}{2} 2^{k_m(1-a)}.$$

We then put $h_m := r_m^+$. If the maximum in (6.14) is attained at r_m^-, we argue similarly.

Case (iii). Assume we have (6.12) (the other option is dealt with exactly the same way). Since $R_m(x+r_m^-) - R_m(x) \leq 0$, the arguments above imply that $S_m(x+r_m^-) - S_m(x) \geq r_m^-$. We now show that $R_m(x) \leq 0$. Indeed, by our choice of $\{k_n\}$ satisfying (6.10) the difference $|S_m(x+r_m^-) - S_m(x)|$ is dominated by $2^{-k_m a} \leq -R_m(x+r_m^-)$. Hence the condition $R_m(x) \geq 0$ would immediately imply that $f(x+r_m^-) - f(x) \leq 0$ which contradicts our assumption.

Now we look at the minimum/maximum on the left of x. First we claim that both $S_m(x) - S_m(x-\rho_m^+)$ and $S_m(x) - S_m(x-\rho_m^-)$ are positive. Assume it is not the case, say, for ρ_m^+, that is $S_m(x) - S_m(x-\rho_m^+) \leq 0$. Then S'_m should vanish at some point $x+\theta \in [x-\rho_m^+, x+r_m^-]$. By our choice of k_n, see (6.10), it follows immediately that

$$\sup_{\theta\approx 2^{-k_n}} |S'_m(x+\theta)| \leq \frac{1}{10}.$$

Therefore

$$\left|\frac{S_m(x+r_m^-)-S_m(x)}{r_m^-}\right| \leq \frac{1}{10},$$

and

$$\frac{f(x+r_m^-)-f(x)}{r_m^-} \leq \frac{1}{10} + \frac{R_m(x+r_m^-)-R_m(x)}{r_m^-} \leq \frac{1}{10},$$

so we have a contradiction. This proves that $S_m(x)-S_m(x-\rho_m^+)\geq 0$. A similar argument shows that $S_m(x)-S_m(x-\rho_m^-)\geq 0$.

Since $R_m(x)\geq R_m(x-\rho_m^-)$, we obtain

$$\frac{f(x)-f(x-\rho_m^-)}{\rho_m^-} = \frac{S_m(x)-S_m(x-\rho_m^-)}{\rho_m^-} + \frac{R_m(x)-R_m(x-\rho_m^-)}{\rho_m^-} \geq 0.$$

On the other hand, since $R_m(x)\leq 0$ we have $R_m(x)-R_m(x-\rho_m^+)\leq -R_m(x-\rho_m^+)$. Hence, as in the previous cases,

$$\frac{f(x)-f(x-\rho_m^+)}{\rho_m^+} \leq -\frac{1}{2}2^{k_m(1-a)}.$$

in particular there exists a point $x-\tilde{\rho}_m$ such that $f(x)-f(x-\tilde{\rho}_m)=0$. We define $h_m := -\rho_m^+$, and $h'_m := -\tilde{\rho}_m$.

Remark. We have constructed a function $f\in \mathrm{Hol}_a$ such that for every $x\in\mathbb{R}$ there exists a couple of sequences h_m, h'_m that satisfy

$$\left|\frac{f(x+h'_m)-f(x)}{h'_m}\right| \leq 1$$
$$\frac{|f(x+h_m)-f(x)|}{|h_m|^a} \gtrsim 1.$$

It follows from the construction that these two sequences can be chosen in such a way that they both lie on the same side of x (right or left, but it *depends* on the point x), but it is not immediately clear that we can fix the side beforehands, i.e. that we can pick such a function f that both h_m and h'_m are, say, positive numbers. One therefore could ask, if for every function $f\in\mathrm{Hol}_a$ there exists at least one point x such that either

$$\liminf_{\theta\downarrow 0}\frac{f(x+\theta)-f(x)}{\theta} = +\infty,$$

or there exists a finite right derivative of f at x.

Chapter 7 Growth classes in the Ball: Cartwright theorem revisited

7.1 Notation

Given two functions f and g defined in $\mathbb{R}^{d+1}$, we say that $f \lesssim g$ if there is a positive constant C, depending only on the dimension n, such that $f \leq Cg$. We write $f \sim g$ if $f \lesssim g$ and $g \lesssim f$ simultaneously. A point z in the unit ball $\mathbb{B}$ in $\mathbb{R}^{d+1}$ will be denoted by (x, y), where $x \in S = \partial\mathbb{B}$, $x = \frac{z}{|z|}$ and $y = 1 - |z| > 0$. Then y is the distance from z to the unit sphere and x is the closest point to z on the sphere; this notation turns out to be convenient for our problem. Despite the inconsistency we will sometimes write $u(z)$ and sometimes $u(x, y)$. By $P_y(x, \xi)$ we denote the Poisson kernel for $\mathbb{B}$,

$$P_y(x, \xi) = \frac{y(2-y)}{|(1-y)x - \xi|^{d+1}} = \frac{1-|z|^2}{|z-\xi|^{d+1}}, \quad x, \xi \in S,\ y \in [0,1],\ z = (1-y)\cdot x.$$

Let also $\phi(z, \zeta) \in [0, \pi]$ be the angle between z and ζ where $z, \zeta \in \mathbb{R}^{d+1} \setminus \{0\}$,

$$\phi(z, \zeta) = \cos^{-1}\left(\frac{\langle z, \zeta\rangle}{|z||\zeta|}\right).$$

Let η be the south pole of $\mathbb{B}$, $\eta = (0, \ldots, 0, -1)$, we fix this notation for the rest of the chapter. Given $0 \leq t \leq \pi$ and $0 \leq y \leq 1$ we denote by $A(y, t)$ the "antarctic" cap

$$A(y, t) = \{z \in \mathbb{B} :\ |z| = 1 - y,\ \phi(z, \eta) \leq t\},$$

and we also put $S(y, t) = \partial A(y, t)$. Following [52], we consider the averaged Poisson kernel,

$$\Phi(x, y, t) = \frac{1}{\sigma_{d-1}(S(0,t))} \int_{S(0,t)} P_y(x, \xi)\, d\sigma_{d-1}(\xi), \quad x \in S,\ 0 < y \leq 1,\ 0 \leq t \leq \pi, \tag{7.1}$$

where σ_{d-1} is the $(d-1)$-dimensional surface measure on $S(0, t)$, $\sigma_{d-1}(S(0, t)) = C(d) \sin^{d-1} t$. Note that $\Phi(x, 1, t) = 1$, for $x \in S$, $0 \leq t \leq \pi$.

We need the following estimate (Lemma 1 from [52])

Lemma 7.1.1 *For any $x \in S$ and $y \in (0, 1]$, $t \leq \phi(x, \eta)$ we have*

$$\Phi(x, y, t) \sim \frac{y}{q^2(q^{d-1} + \sin^{d-1}\phi(x, \eta))},$$

where $q^2 = 1 + (1-y)^2 - 2(1-y)\cos(\phi(x,\eta) - t) = \operatorname{dist}^2(S(0,t), S(y,\phi(x,\eta)))$.

This averaged Poisson kernel will be useful later on (section 7.5), when we deal with axially symmetric harmonic functions. We call the function $\tilde{u}$ on the unit ball axially symmetric if $u(z)$ depends only on $|z|$ and the angle $\phi(z,\eta)$ between z and $\eta = (0,\dots,0,-1)$. If such a function has boundary values $\tilde{u}(x,0) = \varphi(t),\ t = \phi(x,\eta)$, we rewrite the Poisson representation formula in the following way

$$\begin{aligned}\tilde{u}(x,y) = \int_S \tilde{u}(\xi,0)P_y(x,\xi)\,d\sigma_d(\xi) &= \int_0^\pi \int_{S(0,t)} \tilde{u}(\xi,0)P_y(x,\xi)\,d\sigma_{d-1}(\xi)\,dt \\ &= C(d)\int_0^\pi \varphi(t)\Phi(x,y,t)\sin^{d-1}t\,dt, \quad (7.2)\end{aligned}$$

where σ_d is normalized surface measure on S and $C(d)$ is the surface measure of the $(d-1)$-dimensional unit sphere.

7.2 Averaging theorem

In this section we show that in order to deduce (I.60) from (I.59) it is sufficient to obtain an estimate for averages of U over certain spherical caps $A(\theta,\alpha)$ for some $\alpha = \alpha(\theta)$, where $0 < \theta < \frac{1}{2}$. We also make some preliminary estimates to deduce inequalities for the averages from (I.59) and the regularity conditions (I.57) and (I.58).

First we prove the following theorem.

Theorem 7.2.1 *Let U be a harmonic function in $\mathbb{B}$, continuous up to the boundary, satisfying*

$$\begin{aligned}&U(0) = 0,\\ &U(x,y) \le w(y), \quad x \in S,\ 0 \le y \le 1,\end{aligned}$$

where w is a strictly decreasing function. Assume that for some positive $\theta < \frac{1}{2}$ there exists a positive $\alpha = \alpha(\theta) \le \frac{\theta}{4}$ such that

$$w(\theta - 2\alpha) \le C_1 w(\theta), \tag{7.3a}$$

$$\frac{1}{\alpha^d}\left|\int_{A(\theta,\alpha)} U(z)\,d\sigma_d(z)\right| \le C_2 w(\theta), \tag{7.3b}$$

for some positive constants C_1, C_2. Then

$$U(\eta,\theta) \ge -C_3 w(\theta), \tag{7.4}$$

where $C_3 = C_3(C_1, C_2, d)$.

Proof. Consider the ball $\mathbb{B}'$ with center (η,θ) and radius 2α. The condition (7.3a) implies that for any $z \in \mathbb{B}'$ we have

$$w(1 - |z|) \le C_1 w(\theta),$$

and therefore

$$-U(z) + C_1 w(\theta) \geq 0, \quad z \in \mathbb{B}'.$$

Now we can use the Harnack inequality to obtain

$$-U(\eta,\theta) + C_1 w(\theta) \leq C(d)\left(-U(z) + C_1 w(\theta)\right), \quad |z| = 1-\theta,\ \phi(z,\eta) \leq \alpha.$$

All that remains is to take the average over $\{z : |z| = 1-\theta,\ \phi(z,\eta) \leq \alpha\}$,

$$-U(\eta,\theta) + C_1 w(\theta) \leq \tilde{C}(d)\frac{1}{\alpha^d}\int_{\{z:|z|=1-\theta,\ \phi(z,\eta)\leq\alpha\}} \left(-U(z) + C_1 w(\theta)\right)\, d\sigma_d(z),$$

which, combined with (7.3b), implies (7.4). □

7.3 Two lemmas

Now we want to show that the regularity conditions (I.57) and (I.58) imply (7.3a) and (7.3b) for an appropriately chosen $\alpha = \alpha(\theta)$. It turns out that a natural way to define α (at least for somewhat smooth weights) is

$$\alpha(\theta) := -\frac{w(\theta)}{10w'(\theta)}, \quad 0 < \theta < 1. \tag{7.5}$$

We refer the reader to Section 7.4 for a further discussion. The validity of our choice is provided by the following lemma.

Lemma 7.3.1 *If the weight w satisfies* (I.57) *and* (I.58)*, and $\alpha(\theta)$ is given by* (7.5)*, then $0 \leq \alpha(\theta) \leq \frac{\theta}{4}$ and*

$$w(\theta - 2\alpha(\theta)) \leq 2w(\theta), \quad 0 < \theta \leq \frac{1}{2}.$$

Now we need to see if the α we have chosen in (7.5) satisfies (7.3b). This is a much more complicated task than verifying (7.3a), and the first step is the statement below.

Lemma 7.3.2 *If the weight w satisfies* (I.57) *and* (I.58) *for some $\delta > 0$ and $\alpha(\theta)$ is defined by* (7.5)*, then for $0 < \theta \leq \frac{1}{2}$*

$$\int_0^1 \left(\frac{w(y(1-\theta)+\theta)}{y}\right)^{\frac{1}{d+1}} dy \leq \left(\frac{d+1}{d} + \frac{40(d+1)}{\delta}\right) w^{\frac{1}{d+1}}(\theta)\alpha^{\frac{d}{d+1}}(\theta). \tag{7.6}$$

7.3.1 Proof of Lemma 7.3.1

For any $0 < \theta < \frac{1}{2}$ there exists $\theta_1 \in [\theta - 2\alpha(\theta), \theta]$ such that

$$w(\theta) = w(\theta - 2\alpha(\theta)) + 2\alpha(\theta)w'(\theta_1). \tag{7.7}$$

The regularity condition (I.58) implies that

$$\alpha'(\theta) \le \frac{1-\delta}{10d}. \tag{7.8}$$

Hence $\alpha(\theta_1) \ge \alpha(\theta) - \frac{\theta-\theta_1}{10}$, and, on the other hand, $\theta - \theta_1 \le 2\alpha(\theta)$, so

$$\alpha(\theta_1) \ge \alpha(\theta) - \frac{\alpha(\theta)}{5} = \frac{4}{5}\alpha(\theta).$$

We see that

$$-\alpha(\theta)w'(\theta_1) = \frac{\alpha(\theta)w(\theta_1)}{10\alpha(\theta_1)} \le \frac{w(\theta_1)}{8}.$$

Plugging this inequality into (7.7), we obtain

$$w(\theta) \ge w(\theta - 2\alpha(\theta)) - \frac{w(\theta_1)}{4} \ge \frac{w(\theta - 2\alpha(\theta))}{2}$$

and the lemma follows.

7.3.2 Proof of Lemma 7.3.2

We split the integral in (7.6) into two parts

$$\int_0^1 \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy$$
$$= \int_0^{\alpha(\theta)} \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy + \int_{\alpha(\theta)}^1 \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy = I_1 + I_2.$$

To estimate the first integral we just note that for $\theta \le 1$

$$I_1 = \int_0^{\alpha} \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy \le w^{\frac{1}{d+1}}(\theta) \int_0^{\alpha} \left(\frac{1}{y}\right)^{\frac{1}{d+1}} dy \le \frac{d+1}{d} w^{\frac{1}{d+1}}(\theta)\alpha^{\frac{d}{d+1}}(\theta).$$

To deal with the second integral we let $\kappa(y) := w^{\frac{1}{d+1}}(y),\ y > 0$. So it suffices to verify that

$$I_2 = \int_{\alpha}^1 \kappa((1-\theta)y+\theta)y^{-\frac{1}{d+1}}\, dy \le \tilde{C}\kappa(\theta)\alpha^{\frac{d}{d+1}}(\theta), \tag{7.9}$$

where $\tilde{C} = \frac{40(d+1)}{\delta}$. It follows from the definitions of $\alpha = \alpha(y)$ and $\kappa = \kappa(y)$ that

$$\alpha \cdot \frac{\kappa'}{\kappa} = -\frac{1}{10(d+1)},$$

and, using (I.58), we obtain

$$\left(\alpha^{\frac{d}{d+1}}\kappa\right)' = \alpha^{-\frac{1}{d+1}}\kappa\left(\frac{\kappa'}{\kappa}\alpha + \frac{d}{d+1}\alpha'\right)$$
$$= \alpha^{-\frac{1}{d+1}}\kappa\left(-\frac{1}{10(d+1)} + \frac{d}{d+1}\alpha'\right) \le -\frac{1}{10}\alpha^{-\frac{1}{d+1}}\kappa\left(\frac{1}{(d+1)} - \frac{1-\delta}{d+1}\right)$$
$$= -\frac{\delta}{10(d+1)}\alpha^{-\frac{1}{d+1}}\kappa. \quad (7.10)$$

Let $c = \frac{\delta}{10(d+1)}$. Then integrating (7.10) over $[\theta, 1]$ for $\theta < \frac{1}{2}$, we see that

$$\kappa(\theta)\alpha^{\frac{d}{d+1}}(\theta) \ge c\int_\theta^1 \kappa(y)\alpha^{-\frac{1}{d+1}}(y)\,dy$$
$$= c(1-\theta)\int_0^1 \kappa\left(y(1-\theta)+\theta\right)\alpha^{-\frac{1}{d+1}}\left(y(1-\theta)+\theta\right)\,dy.$$

Now, for $y \ge \alpha(\theta)$ by (7.8) we have $\alpha(\theta + (1-\theta)y) \le \alpha(\theta) + (1-\theta)y \le 2y$. Therefore

$$\int_0^1 \kappa\left(y(1-\theta)+\theta\right)\alpha^{-\frac{1}{d+1}}\left(y(1-\theta)+\theta\right)\,dy \ge 2^{-\frac{1}{d+1}}\int_{\alpha(\theta)}^1 \kappa\left((1-\theta)y+\theta\right)y^{-\frac{1}{d+1}}\,dy.$$

This gives (7.9) with $\tilde{C} = \frac{2^{\frac{1}{d+1}}10(d+1)}{\delta(1-\theta)}$. Combining the estimates for both integrals, we get

$$\int_0^1 \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy \le \left(\frac{d+1}{d} + \frac{40(d+1)}{\delta}\right) w^{\frac{1}{d+1}}(\theta)\alpha^{\frac{d}{d+1}}(\theta),$$

and we are done.

7.4 Intermezzo: some comments about regularity

We have seen that in order to prove the main theorem we need the conditions (7.3a) and (7.3b). They are rather independent: the proof of the first one is self-contained, and the second one, as it will be shown later, follows from Lemma 7.3.2, where we do not use any information on the doubling property of α. Combining them, we see that for every fixed θ we essentially need to find some $\alpha = \alpha(\theta)$ such that

$$w(\theta-\alpha) \le 2w(\theta), \quad (7.11a)$$
$$\int_\alpha^1 \left(\frac{w(y+\theta-y\theta)}{y}\right)^{\frac{1}{d+1}} dy \le C(w,n)w^{\frac{1}{d+1}}(\theta)\alpha^{\frac{d}{d+1}}(\theta). \quad (7.11b)$$

These two inequalities are actually “fighting” with each other. Indeed, if we put α to be very small, then the first condition is immediately satisfied, but the second one fails miserably. On the other hand α cannot be large (compared to θ), because of the first condition: the faster the weight w grows the smaller must α be. If we try to unify these two inequalities, we (albeit probably with some loss of information) would arrive at the “regularity” of the weight w as stated in (I.57) and

(I.58).

It should be noted that this is the only place that we need the regularity conditions, so if we have a weight w that satisfies (7.11a) and (7.11b) with some α (not necessarily defined as in (7.5)), then Theorem I.21 still holds. One important example is the weight w of polynomial growth. Assume that $w \in C^1$ and

$$-\frac{N}{y} \le \frac{w'(y)}{w(y)} \le -\frac{d+\varepsilon}{y}, \; y \in (0,1], \tag{7.12}$$

for some positive ε and $N \ge d+\varepsilon$. Put $\alpha(\theta) = \frac{\theta}{2N}$. Clearly $w(\theta - \alpha) \le 2w(\theta), \; \theta \in (0,1]$, so that we have (7.11a). Furthermore

$$\left(y^{\frac{d}{d+1}} w^{\frac{1}{d+1}}(y)\right)' = \frac{y^{-\frac{1}{d+1}} w^{\frac{1}{d+1}}(y)}{d+1} \left(\frac{w'(y)}{w(y)} y + d\right),$$

so that

$$\left(\alpha^{\frac{d}{d+1}}(y) w^{\frac{1}{d+1}}(y)\right)' \le -C\alpha^{-\frac{1}{d+1}}(y) w^{\frac{1}{d+1}}(y),$$

which is basically (7.10). Following the proof of Lemma 7.3.2 closely, we see that (7.11b) also holds. Note that in this case the weight w can be a little less smooth than required by the regularity condition (I.58).

Note also that in order to bound U from below we do not need w to be regular on the entire interval $(0,1]$. Assume that $w \in C^2$, w is decreasing, and (I.57) and (I.58) hold only for $0 < y \le y_0$ (or $w \in C^1$ and (7.12) holds only for $0 < y \le y_0$) for some $y_0 < 1$. We can still prove a version of Theorem I.21 replacing (I.60) with

$$|U(z)| \le C_1 + C_2 w(1 - |z|), \quad z \in \mathbb{B}. \tag{7.13}$$

Indeed, it is easy to show that there exists a C^2 function $\tilde{w}$ that satisfies (I.57) and (I.58) for $y \in (0,1]$ and such that

$$\begin{aligned} \tilde{w}(y) &\ge w(y), \quad y_1 \le y \le 1, \\ \tilde{w}(y) &= A w(y), \quad 0 < y \le y_1. \end{aligned}$$

For example, one may choose $\tilde{w}(y) = c(y+b)^s$ for $y \ge y_1$ and some $y_1 \le y_0$ such that $\left(\frac{w}{w'}\right)'(y_1) < 0$. Since Theorem I.21 holds for $\tilde{w}$, we immediately have (7.13). A similar argument works for $w \in C^1$ satisfying (7.12).

7.5 Main technical theorem

7.5.1 Statement

The next theorem allows us to estimate from above the absolute values of some averages of the harmonic function.

Theorem 7.5.1 *Let $\tilde{k} : \mathbb{R}_+ \mapsto \mathbb{R}_+$ be a strictly decreasing absolutely continuous function such that*

$$\tilde{k}(0) < \infty, \tag{7.14a}$$

$$\int_0^1 \left(\frac{\tilde{k}(y)}{y}\right)^{\frac{1}{d+1}} dy \le D, \tag{7.14b}$$

for some constant $1 < D < \infty$. Let $\tilde{u}$ be a harmonic function in $\mathbb{B}$, continuous up to the boundary, satisfying $\tilde{u}(0) = 0$ and $\tilde{u}(x, y) \le \tilde{k}(y)$ for $x \in S$, $0 \le y \le 1$. Then for any $x_0 \in S$ and $\beta \in [0, \frac{1}{2}]$ the following inequality holds

$$\int_{\{\phi(x,x_0)\le\beta\}} \tilde{u}(x, 0)\, d\sigma_d(x) \ge -C \left(D^{d+1} + \tilde{k}(0)\beta^d\right). \tag{7.15}$$

where C depends only on the dimension n.

7.5.2 Theorems 7.5.1 and 7.2.1 imply Theorem I.21

Fix any positive $\theta \le \frac{1}{2}$. Let U and w be as in Theorem I.21, and α be defined as in (7.5). The weight we are going to use in Theorem 7.5.1 is defined as follows

$$\tilde{k}(y) := \frac{w(y + \theta - y\theta)}{w(\theta)\alpha(\theta)^d}, \quad 0 \le y \le 1.$$

Indeed, if we apply Lemma 7.3.2, we obtain

$$\int_0^1 \left(\frac{\tilde{k}(y)}{y}\right)^{\frac{1}{d+1}} dy = \int_0^1 \left(\frac{w(y + \theta - y\theta)}{w(\theta)\alpha^d(\theta)y}\right)^{\frac{1}{d+1}} dy \le \left(\frac{d+1}{d} + \frac{40(d+1)}{\delta}\right),$$

so we have the condition (7.14b) with $D = \frac{d+1}{d} + \frac{40(d+1)}{\delta}$. Now put $\beta = \alpha(\theta)$, so $0 \le \beta \le \frac{1}{2}$ by (7.8). Clearly, $\tilde{k}(0) < \infty$, and, by (I.59), the function

$$\tilde{u}(z) := \frac{U(z(1-\theta))}{w(\theta)\alpha^d(\theta)}, \quad |z| \le 1,$$

can be estimated from above by $\tilde{k}(y)$. Theorem 7.5.1 therefore implies that, for any $x_0 \in S$,

$$\int_{\{\phi(x,x_0)\le\alpha(\theta)\}} \tilde{u}(x, 0)\, dx \ge -C(d) \left(D^{d+1} + \tilde{k}(0)\alpha^d(\theta)\right) \ge -C(d) \left(D^{d+1} + 1\right).$$

Since $\tilde{u}$ is bounded from above by $\tilde{k}$, the last inequality implies

$$\frac{1}{\alpha^d(\theta)} \left| \int_{A(\theta,\alpha(\theta))} U(z)\, d\sigma_d(z) \right| \lesssim D^{d+1} w(\theta),$$

and we get (7.3b). The condition (7.3a) will follow from Lemma 7.3.1. We obtain the following inequality

$$\frac{1}{\alpha^d(\theta)}\left|\int_{A(\theta,\alpha(\theta))} U(z)\,d\sigma_d(z)\right| \lesssim \left(\frac{d+1}{d}+\frac{40(d+1)}{\delta}\right)^{d+1} w(\theta),$$

which combined with Theorem 7.2.1 proves Theorem I.21.

7.6 The weight lemma

The aim of the rest of this Chapter is to prove Theorem 7.5.1. Before proceeding further we need to introduce some additional notation. Fix any $\beta \in [0, \frac{\pi}{2}]$ and let

$$A = A(0,\beta) = \{x \in S : \phi(x,\eta) \le \beta\},\ A' = S \setminus A.$$

Recall that $\Phi(x,y,t)$ is the averaged Poisson kernel, defined in (7.1). The main ingredient in the proof of Theorem 7.5.1 is the following lemma.

Lemma 7.6.1 *Let $k : \mathbb{R}_+ \mapsto \mathbb{R}_+$ be a strictly decreasing absolutely continuous function such that*

$$k(0) \le \tfrac{\lambda}{\beta^d}, \tag{7.16a}$$

$$\textstyle\int_0^1 \left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}} dy \le \lambda^{\frac{1}{d+1}}, \tag{7.16b}$$

for some positive $\lambda \le \frac{1}{\pi}$. There exist a domain $\Omega \subset \mathbb{B}$ and a positive function v_A, harmonic in Ω, such that

$$A \subset \partial\Omega,\ 0 \in \Omega, \tag{7.17a}$$

$$v_A(0) \le C(d)\lambda^{\frac{1}{d+1}}, \tag{7.17b}$$

$$v_A(x,y) \gtrsim k(y) \gtrsim \Phi(x,y,\beta), \quad (x,y) \in \partial\Omega \setminus A, \tag{7.17c}$$

where the constants depend only on the dimension n.

The proof of this lemma uses a modification of the argument presented in Lemma **4** in [52]. Basically it allows us to estimate the average of the weight k on $\partial\Omega \setminus A$ with respect to the harmonic measure of Ω at zero. The key point here is the second inequality in (7.17c) which will be used later (in 7.7) to obtain the lower bound in (7.15).

7.6.1 Proof of Lemma 7.6.1: auxiliary surface Γ_A

To obtain Ω we construct its boundary $\partial\Omega = \Gamma_A \bigcup A$. The surface Γ_A is defined below, in such a way that the second inequality in (7.17c) holds on Γ_A and moreover $k(y) \approx \Phi(x,y,\beta)$ there.

Formally, consider the function $\frac{y}{k(y)}$, which is strictly increasing. Let $s = s(\beta)$ be the solution of the following equation

$$\frac{y}{k(y)} = \beta^{d+1}.$$

Since k is decreasing, (7.16a) implies

$$0 < s = k(s)\beta^{d+1} \leq k(0)\beta^{d+1} \leq \lambda\beta. \tag{7.18}$$

Further, (7.16b) and the monotonicity of k imply that

$$k^{\frac{1}{d+1}}(1) \int_0^1 \left(\frac{1}{y}\right)^{\frac{1}{d+1}} dy \leq \int_0^1 \left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}} dy \leq \lambda^{\frac{1}{d+1}},$$

so $k(1) \leq \lambda \leq \frac{1}{\pi}$. Now if we let $\rho = \rho(\beta)$ be the solution of

$$\frac{y}{k(y)} = (\pi - \beta)^{d+1},$$

we see that $s < \rho < 1$. Further let

$$\begin{aligned} \gamma(y) &= \beta + \left(\frac{y}{k(y)}\right)^{\frac{1}{d+1}}, \quad s \leq y \leq \rho, \\ \gamma(y) &= \beta + \left(\frac{y}{k(y)\beta^{d-1}}\right)^{\frac{1}{2}}, \quad 0 \leq y \leq s, \end{aligned} \tag{7.19}$$

and note that $\gamma(0) = \beta$, $\gamma(s) = 2\beta$ and $\gamma(\rho) = \pi$. The surface Γ_A is defined as follows

$$\Gamma_A := \{(x, y) : \phi(x, \eta) = \gamma(y),\ x \in A' = S \setminus A,\ y \in [0, \rho]\},$$

and we define Ω as the domain bounded by $A \bigcup \Gamma_A$, so that Ω satisfies (7.17a).

7.6.2 Proof of Lemma 7.6.1: auxiliary function v_A

We define v_A on the unit sphere by

$$\begin{aligned} v_A(x, 0) &= k(y), \quad (x, y) \in \Gamma_A, \\ v_A(x, 0) &= 0, \quad x \in A, \end{aligned} \tag{7.20}$$

and let v_A be the harmonic continuation of $v_A(\cdot, 0)$ to the ball. Note that the function v_A is axially symmetric. It remains to verify (7.17b) and (7.17c).

In what follows the letter C denotes a constant, depending only on n, whose value can change from line to line. The proof of the first inequality is straightforward if somewhat cumbersome. We have $\gamma^{-1}(\beta) = 0$, $\gamma^{-1}(2\beta) = s$ and it follows from (7.2) that

$$\begin{aligned} v_A(0) &= C \int_\beta^\pi \int_{\phi(x,\eta)=t} v_A(x, 0)\, d\sigma_{d-1}(x)\, dt = C \int_\beta^\pi k(y(\gamma)) \sin^{d-1}\gamma\, d\gamma \\ &= C \int_\beta^{2\beta} k(y(\gamma)) \sin^{d-1}\gamma\, d\gamma + C \int_{2\beta}^\pi k(y(\gamma)) \sin^{d-1}\gamma\, d\gamma \\ &= C \int_0^s k(y) \sin^{d-1}(\gamma(y))\gamma'(y)\, dy + C \int_s^\rho k(y) \sin^{d-1}(\gamma(y))\gamma'(y)\, dy. \end{aligned}$$

These two integrals are dealt with more or less in the same way. For the first one we have $\gamma(y) \le \gamma(s) = 2\beta$ so that $\sin\gamma(y) \le 2\beta$. Then

$$\begin{aligned}
\int_0^s k(y)\sin^{d-1}(\gamma(y))\gamma'(y)\,dy &\le C\beta^{d-1}\int_0^s k(y)\gamma'(y)\,dy \\
&= C\beta^{\frac{d-1}{2}}\int_0^s y^{-\frac{1}{2}}\left(k^{\frac{1}{2}}(y) - yk'(y)k^{-\frac{1}{2}}(y)\right)\,dy \\
&\le C\beta^{\frac{d-1}{2}}\left(\int_0^s \sqrt{\frac{k(y)}{y}}\,dy + \int_0^s y^{\frac{1}{2}}\,dk^{\frac{1}{2}}(y)\right) \\
&\qquad \le C\beta^{\frac{d-1}{2}}\sqrt{k(0)s} \le C\beta^{\frac{d-1}{2}}\sqrt{\lambda^2\cdot\beta^{1-n}} \le C\lambda,
\end{aligned}$$

the next to last inequality follows from (7.16a) and (7.18). Analogously, for the second integral we have $y \in [s,\rho]$ and

$$\gamma'(y) = \frac{1}{d+1}\left(\frac{y}{k(y)}\right)^{-\frac{d}{d+1}}\frac{k(y)-yk'(y)}{k^2(y)},$$

also

$$\gamma(y) \le 2\left(\frac{y}{k(y)}\right)^{\frac{1}{d+1}}.$$

We get

$$\begin{aligned}
\int_s^\rho k(y)\sin^{d-1}(\gamma(y))\gamma'(y)\,dy &\le \int_s^\rho k(y)(\gamma(y))^{d-1}\gamma'(y)\,dy \\
\le 2^{d-1}\int_0^1 k(y)\frac{k^{-\frac{d-1}{d+1}}(y)y^{\frac{d-1}{d+1}}}{d+1}&\left(y^{-\frac{d}{d+1}}k^{-\frac{1}{d+1}}(y) - y^{\frac{1}{d+1}}k^{-\frac{1}{d+1}-1}(y)k'(y)\right)\,dy \\
\le \frac{2^{d-1}}{d+1}&\left(\int_0^1\left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}}\,dy - (d+1)\int_0^1 y^{\frac{d}{d+1}}\,d(k^{\frac{1}{d+1}}(y))\right) \\
= \frac{2^{d-1}}{d+1}\left(\int_0^1\left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}}\,dy + n\int_0^1\left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}}\right.&\left.dy - (d+1)k^{\frac{1}{d+1}}(1)\right) \\
&\le C\int_0^1\left(\frac{k(y)}{y}\right)^{\frac{1}{d+1}}\,dy \le C\lambda^{\frac{1}{d+1}}.
\end{aligned}$$

Combining these two estimates, we obtain that $v_A(0) \le C(\lambda + \lambda^{\frac{1}{d+1}}) \lesssim \lambda^{\frac{1}{d+1}}$, since $\lambda \le 1$.

The second part of (7.17c), i.e. the inequality $k(y) \gtrsim \Phi(x,y,\beta)$, for $(x,y)\in\Gamma_A$ follows directly from Lemma 7.1.1. Indeed, Lemma 7.1.1 implies that for $t \le \phi(x,\eta) \le \frac{\pi}{2}$

$$\Phi(x,y,t) \sim \frac{y}{((\phi(x,\eta)-t)^2+y^2)\left(((\phi(x,\eta)-t)^2+y^2)^{\frac{d-1}{2}} + \phi^{d-1}(x,\eta)\right)}. \tag{7.21}$$

If $\beta \le \phi(x,\eta) \le 2\beta$ then, clearly, $((\phi(x,\eta)-\beta)^2+y^2)^{\frac{d-1}{2}} + \phi^{d-1}(x,\eta) \ge \beta^{d-1}$. Further $(\phi(x,\eta) -$

$\beta)^2 = \frac{y}{k(y)\beta^{d-1}}$ for $(x,y) \in \Gamma_A$ by(7.19). Therefore we get

$$\Phi(x,y,\beta) \le \frac{Cy}{((\phi(x,\eta)-\beta)^2+y^2)\beta^{d-1}} \le \frac{Cy}{(\phi(x,\eta)-\beta)^2\beta^{d-1}}$$
$$\le \frac{Cyk(y)\beta^{d-1}}{y\beta^{d-1}} \le Ck(y), \quad (x,y) \in \Gamma_A.$$

It follows from (7.21) that $\Phi(x,y,\beta) \le \frac{cy}{(\phi(x,\eta)-\beta)^{d+1}}$. For $2\beta \le \phi(x,\eta) \le \pi$ we have $4(\phi(x,\eta)-\beta)^2 \ge \phi^2(x,\eta)$, and therefore, by (7.19),

$$\Phi(x,y,\beta) \le \frac{Cy}{(\phi(x,\eta)-\beta)^{d+1}} = \frac{Cy}{(\gamma(y)-\beta)^{d+1}} = \frac{Ck(y)y}{y} = Ck(y), \quad (x,y) \in \Gamma_A.$$

To obtain the first part of (7.17c), we first show that

$$y \le \phi(x,\eta) - \beta \tag{7.22}$$

for $(x,y) \in \Gamma_A$. Indeed, for $2\beta \le \phi(x,\eta)$ it follows from (7.16b) and (7.19) that

$$\frac{y}{\phi(x,\eta)-\beta} = k^{\frac{1}{d+1}}(y)y^{1-\frac{1}{d+1}} \le \int_0^y \left(\frac{k(\tau)}{\tau}\right)^{\frac{1}{d+1}} d\tau \le \lambda^{\frac{1}{d+1}},$$

since $\frac{k(y)}{y}$ is decreasing. If $\beta \le \phi(x,\eta) \le 2\beta$, then $y \le s \le \lambda\beta$ and (7.19) gives

$$\frac{y}{\phi(x,\eta)-\beta} = k^{\frac{1}{2}}(y)\beta^{\frac{d-1}{2}}y^{\frac{1}{2}} \le k^{\frac{1}{2}}(0)\beta^{\frac{d-1}{2}}s^{\frac{1}{2}} \le (\lambda\beta^{-n})^{\frac{1}{2}}\beta^{\frac{d-1}{2}}(\lambda\beta)^{\frac{1}{2}} \le \lambda \le 1.$$

Put $E(x,y) = \{\xi \in S : \phi(\xi,\eta) \le \phi(x,\eta),\ \phi(x,\xi) \le y\}$. It follows from (7.22) that $\sigma_d(E(x,y)) \sim y^d$ for $(x,y) \in \Gamma_A$ and $E(x,y) \subset a$. Since for $\xi \in E(x,y)$ we have $P_y(\xi,x) \gtrsim \frac{1}{y^d}$, the function $v_A(x,0)$ is axially symmetric and by definition strictly decreasing with respect to $\phi(x,\eta)$ for $\phi(x,\eta) \ge \beta$, we get

$$\begin{aligned} v_A(x,y) &= \int_S v_A(\xi,0)P_y(\xi,x)\,d\sigma_d(\xi) \\ &\ge \int_{\{\xi\in a:\phi(\xi,\eta)\le\phi(x,\eta)\}} v_A(\xi,0)P_y(\xi,x)\,d\sigma_d(\xi) \gtrsim \int_{E(x,y)} v_A(\xi,0)\frac{1}{y^d}\,d\sigma_d(\xi) \\ &\ge \frac{1}{y^d}\int_{E(x,y)} v_A(x,0)\,d\sigma_d(\xi) \ge \frac{1}{y^d}v_A(x,0)\sigma_d(E(x,y)) \gtrsim Ck(y). \end{aligned}$$

This completes the proof of Lemma 7.6.1.

7.7 Proof of Theorem 7.5.1

First we renormalize the weight $\tilde{k}$ and the function $\tilde{u}$ as

$$\begin{aligned} k(y) &= \frac{\lambda}{D^{d+1} + \tilde{k}(0)\beta^d} \tilde{k}(y), \\ u(z) &= \frac{\lambda}{D^{d+1} + \tilde{k}(0)\beta^d} \tilde{u}(z), \end{aligned} \tag{7.23}$$

where $\lambda = \lambda(d) \leq \frac{1}{\pi}$ is a small positive constant to be chosen later. We may assume that $x_0 = \eta$ and that $\tilde{u}(\cdot, 0)$ (and therefore $u(\cdot, 0)$) is a axially symmetric function,

$$u(x, 0) = \varphi(|\phi(x, \eta)|), \quad x \in S.$$

By u_A and $u_{A'}$ we denote the harmonic continuation to $\mathbb{B}$ of the functions $u(\cdot, 0) \cdot \chi_A$ and $u(\cdot, 0) \cdot \chi_{A'}$ correspondingly.

Clearly $0 = u(0) = u_{A'}(0) + u_A(0)$. Let

$$K = -u_A(0) = -\int_A u(x, 0)\, d\sigma_d(x),$$

and assume, as we may, that $K \geq 0$ (otherwise (7.15) is trivial). We see that (7.14a) and (7.14b) imply that the weight k satisfies the conditions (7.16a) and (7.16b). Let Γ_A and v_A be as in Lemma 7.6.1.

Our first aim is to prove the following inequality

$$u_a(x, y) \leq C(1 + K) v_A(x, y), \quad (x, y) \in \Gamma_A. \tag{7.24}$$

Since $u_a(\cdot, 0)$ is just the part of the boundary values of u that lies in a, we have

$$u_a(x, y) = u(x, y) - u_A(x, y) \leq k(y) - u_A(x, y), \quad (x, y) \in \Gamma_A,$$

so to get an upper estimate on u_a we actually need to bound u_A from below on Γ_A. Again, (7.2) provides us with

$$u_A(x, y) = C(d) \int_0^\beta \varphi(t) \Phi(x, y, t) \sin^{d-1} t\, dt,$$

so, in particular, we have

$$u_A(0) = C(d) \int_0^\beta \varphi(t) \sin^{d-1} t\, dt. \tag{7.25}$$

Clearly $\varphi(t) - k(0) \leq 0$, so using the mean value theorem (the first one, unlike in [52]) we see that there exists $t_0 \in [0,\beta]$ such that

$$\begin{aligned}
&\int_0^\beta \varphi(t)\Phi(x,y,t)\sin^{d-1} t\,dt \\
&\qquad = \int_0^\beta (\varphi(t)-k(0))\,\Phi(x,y,t)\sin^{d-1} t\,dt + \int_0^\beta k(0)\Phi(x,y,t)\sin^{d-1} t\,dt \\
&\qquad = \Phi(x,y,t_0)\int_0^\beta (\varphi(t)-k(0))\sin^{d-1} t\,dt + k(0)\int_0^\beta \Phi(x,y,t)\sin^{d-1} t\,dt \\
&\qquad \geq \Phi(x,y,t_0)\int_0^\beta \varphi(t)\sin^{d-1} t\,dt - \Phi(x,y,t_0)k(0)\int_0^\beta \sin^{d-1} t\,dt \\
&\qquad = \Phi(x,y,t_0)\frac{u_A(0)}{C(d)} - \Phi(x,y,t_0)k(0)\int_0^\beta \sin^{d-1} t\,dt,
\end{aligned}$$

the last equality following from (7.25). Now (7.23) implies that $k(0)\beta^d \leq \lambda \leq 1$, and we also have $\int_0^\beta \sin^{d-1} t\,dt \approx \beta^d$. We continue the estimate, obtaining

$$\begin{aligned}
\int_0^\beta \varphi(t)\Phi(x,y,t)\sin^{d-1} t\,dt &\geq \Phi(x,y,t_0)\left(\frac{u_A(0)}{C(d)} - k(0)\int_0^\beta \sin^{d-1} t\,dt\right) \\
&\geq \Phi(x,y,t_0)\left(-\frac{K}{C(d)} - C(d,\beta)\lambda\right),
\end{aligned}$$

where $C(d,\beta) \sim 1$. It follows from (7.21) that $\sup_{0\leq t\leq\beta}\Phi(x,y,t) \sim \Phi(x,y,\beta)$, when $\phi(x,\eta) > \beta$. We therefore have

$$\begin{aligned}
u_A(x,y) &\geq C(d)\Phi(x,y,t_0)\left(-\frac{K}{C(d)} - C(d,\beta)\lambda\right) \\
&\geq C(d)\Phi(x,y,\beta)\left(-\frac{K}{C(d)} - C(d,\beta)\lambda\right) \geq -C(d)(K+1)\Phi(x,y,\beta).
\end{aligned}$$

Gathering all the estimates and applying (7.17c), we get

$$u_{A'}(x,y) \leq k(y) - u_A(x,y) \leq k(y) + C(d)(K+1)\Phi(x,y,\beta) \lesssim v_A(x,y)(K+1)$$

for $(x,y) \in \Gamma_A$, and we obtain (7.24).

Once we have this estimate it is quite easy to finish the proof. Indeed, it follows from (7.24),(7.17b) and the maximum principle that

$$K = u_{A'}(0) \leq C(1+K)v_A(0) \leq C_0\lambda^{\frac{1}{d+1}}(1+K) \leq \frac{1+K}{3},$$

for sufficiently small λ. Therefore we have $K \leq \frac{1}{2}$, which means that

$$\int_{A(0,\beta)} \tilde{u}(x,0)\,d\sigma_d(x) \geq -\frac{1}{2\lambda}\left(D^{d+1} + \tilde{k}(0)\beta^d\right),$$

and we are done.

Chapter 8 Normal variation of positive harmonic functions

In this Chapter we prove Theorem I.25. Let us recall its statement.

Theorem 8.0.1 *Assume that u is a positive harmonic function on $\mathbb{R}^2_+$ such that its boundary measure has compact support, so, in particular, $\lim_{|(x,t)|\to\infty} u(x,t) = 0$. Then the set of Bourgain points, i.e. points where the* ***mean variation*** *of u is finite*

$$\operatorname{Mvar} u(x) = \int_0^1 h_u^{2t}(x,t)\,dt < \infty,$$

is ultradense in $\mathbb{R}$, that is its intersection with ***any interval*** *$I \subset \mathbb{R}$ has full Hausdorff dimension. Here h_u^{2t} is the least harmonic majorant of the subharmonic function $|\nabla u|(\cdot, \cdot + 2t)$ which can be written as follows*

$$h_u^{2t}(x,s) := \int_{\mathbb{R}} |\nabla u(\xi, 2t)| P_{(s)}(x-\xi)\,d\xi, \quad x \in \mathbb{R},\ t,s > 0.$$

The scheme of the proof was briefly outlined in the Introduction. Before we have a closer look let us introduce some notation and establish some preliminary facts.

8.1 Operators B_y

8.1.1 Integral operators and their kernels

Since we aim to maintain the idea that our arguments are not restricted to the half-spaces, we try to abstain from the usual $\mathbb{R}^{d+1}_+$ type of notation where it is not completely inconvenient. In particular we write Ω instead of $\mathbb{R}^2_+$ and $S = \partial\Omega$ instead of $\mathbb{R} = \partial\mathbb{R}^2_+$. In what follows by a *kernel* we call a function defined on $S \times S$. The kernels are denoted by lowercase Latin or Greek letters, and the corresponding integral operators by respective capital letters. For example, the operator K is generated by the kernel k,

$$K[\varphi](x) := \int_S k(x,\xi)\varphi(\xi)\,d\xi, \quad x \in S. \tag{8.1}$$

The unfortunate exception is the Poisson kernel $P_{(y)}$ which, besides being written with capital P, is also defined on $\Omega \times S$ or on S, depending on what variables we fix. Because of this we will use it in a slightly different notation. Also we put $x_{[t]} := (x,t),\ x \in S,\ t > 0$.

The composition $k_1 \circ k_2$ of the kernels k_1 and k_2 is defined as follows

$$(k_1 \circ k_2)(x,\xi) := \int_S k_1(x,\eta)k_2(\eta,\xi)\,d\eta, \quad x,\xi \in S. \tag{8.2}$$

Here we assume that the integrand in (8.2) is summable with respect to s for any $x, \xi \in S$. We take special care for the equality $K_1(K_2(\varphi)) = K(\varphi)$ to hold (here K is the integral operator with the kernel $k := k_1 \circ k_2$). We also need the compositions $k_1 \circ \cdots \circ k_n$ that are defined in the same way for any natural n.

By k^* we denote the kernel adjoint to k:

$$k^*(x,\xi) := k(\xi,x), \quad x,\xi \in S, \tag{8.3}$$

and K^* is the respective integral operator, that also gives the following formula

$$\int_S K^*(\varphi)\cdot\psi\,ds = \int_S \varphi\cdot K(\psi)\,ds.$$

8.1.2 Kernels p_y, c_y, b_y

The first of these kernels is just another way to write Poisson kernel for Ω, the other two also depend on u.

Kernel p_t

Put

$$p_t(x,\xi) := P_{(t)}(x-\xi) = \frac{t}{\pi(t^2+(x-\xi)^2)}, \quad x,\xi \in S,\ t>0. \tag{8.4}$$

Given $x \in S,\ t>0$, the measure $p_t(x,\cdot)\,ds$ on S is the harmonic measure in Ω with the pole at $x_{[t]} := (x,t) \in \Omega$. In particular, $\int_S p_t(x,\xi)\,d\xi = 1$.
We will need the *semigroup property* of p_t,

$$p_{t_1+t_2} = p_{t_1} \circ p_{t_2}, \quad t_1,t_2 > 0. \tag{8.5}$$

We also need the following property of p_t – *everywhere* on $S \times S$ we have

$$\frac{p_{t_2}}{p_{t_1}} \le c(S)\left(\frac{t_2}{t_1}\right), \quad 0 < t_1 \le t_2 \le 1. \tag{8.6}$$

Kernel c_t

Given $\vec{a} \in \mathbb{R}^d$ we put

$$\operatorname{sgn}\vec{a} := \begin{cases} \frac{\vec{a}}{|\vec{a}|}, & \text{if } \vec{a} \neq 0 \\ 0, & \text{if } \vec{a} = 0. \end{cases}$$

The scalar product of $\vec{a}$ and $\vec{b}$ is denoted by $\langle \vec{a}, \vec{b} \rangle$. Let $\phi(\xi) := \operatorname{sgn}(\nabla u)(\xi),\ \xi \in \Omega$. The vector field ϕ vanishes on the zero set $\mathcal{Z}$ of ∇u, is smooth and unitary outside it.

Let

$$c_t(x,\xi) := \left\langle \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x} \right) \circ p_t(x,\xi), \phi(x_{2t}) \right\rangle, \quad x, \xi \in S,\ t > 0, \tag{8.7}$$

or, in other words, we take the partial derivative of the Poisson kernel w.r.t. ϕ. Both ϕ and c_t depend on u, we drop the index though.

By Harnack's inequality one has

$$|c_t(x,\xi)| \leq \frac{c(S) p_t(x,\xi)}{t}, \quad x, \xi \in S,\ t > 0. \tag{8.8}$$

If we differentiate $\int_S p_t(x,\xi) ds(\xi) \equiv 1,\ (x,t) \in \Omega$ over $\phi(x_{2t})$, we get

$$C_t[1] = 0 \text{ on } S, \quad t > 0. \tag{8.9}$$

Kernel b_t

Now we define the kernel as follows $b_t := p_t \circ c_t,\ t > 0$, the corresponding integral clearly converges. Next we observe that for $x \in S,\ t > 0$ and $\phi = (\phi_1, \phi 2)$ we have

$$\begin{aligned}
&C_t \left[u_{[t]} \right](x) = \int_S c_t(x,\xi) u_{[t]}(\xi)\, d\xi = \\
&= \int_S \left(\frac{\partial}{\partial x} p_t(x,\xi) \cdot \phi_1(x,2t) + \frac{\partial}{\partial t} p_t(x,\xi) \cdot \phi_2(x,2t) \right) u(\xi,t)\, d\xi = \\
&\phi_1(x,2t) \int_S \frac{\partial}{\partial x} p_t(x,\xi) u(\xi,t)\, d\xi + \phi_2(x,2t) \int_S \frac{\partial}{\partial t} p_t(x,\xi) u(\xi,t)\, d\xi = \\
&\phi_1(x,2t) \frac{\partial}{\partial x} u(x,2t) + \phi_2(x,2t) \frac{\partial}{\partial t} u(x,2t) = |\nabla u|(x,2t)
\end{aligned} \tag{8.10}$$

The least harmonic majorant h^{2t} for $|\nabla u_{[2t]}|$ is therefore the Poisson extension of the above,

$$h^{2t}\left(x_{[t]}\right) = (h^{2t})_{[t]}(x) = B_t\left(u_{[t]}\right)(x), \quad x \in S, \tag{8.11}$$

and

$$\operatorname{Mvar} u(x) = \int_0^1 B_t\left(u_{[t]}\right)(x)\, dt, \quad x \in S. \tag{8.12}$$

Note that $B_t\left[u_{[t]}\right] > 0$.

We need b_t to satisfy the following properties: *for any* $t \in (0,1)$

$$|b_t| \leq \tfrac{c}{t} p_t, \tag{8.13a}$$

$$B_t(1) \equiv 0 \text{ on } S. \tag{8.13b}$$

Indeed, by (8.8) and Harnack's inequality we have

$$\begin{aligned} |b_t| &\le p_t \circ |c_t| \le c\frac{p_t \circ p_t}{t} \\ &\le c'(S)\frac{p_t}{t}, \end{aligned}$$

while (8.13b) follows from (8.9).
In Section 8.7 we show that the function $(x, \xi, t) \mapsto b_t(x, \xi)$ is continuous on $S \times S \times (0, +\infty)$.

8.1.3 Outline of the proof of Theorem 8.0.1, measures $\nu^{\varepsilon,u}$

Assume that there exists a Borel measure ν supported on S such that $\int_S \mathrm{Mvar}\, u\, d\nu < +\infty$ and $\nu(I) > 0$ for any interval $I \subset S$. Then the set $\mathcal{B}(u)$ of Bourgain points of u is *dense* in S. However, we can not yet say that $\mathcal{B}(u)$ is *ultradense* in S – such a measure ν can still be supported on some countable subset of S.
To prove the *ultradensity* of $\mathcal{B}(u)$ in S, we construct the family $\{\nu^\varepsilon\}$ (with $0 < \varepsilon < \varepsilon_0$ for some ε_0 that ultimately depends on S) of measures supported on S such that

a. $\int_S V\, d\nu^\varepsilon < +\infty$;

b. there exist some positive constants c_1, c_2 *such that*

$$\nu^\varepsilon(I) \le c_1 |I|^{1-c_2\varepsilon} \tag{8.14}$$

for any interval $I \subset S$;

c. for any interval $I \subset S$ *there exists such an* $\varepsilon_I > 0$ *that* $\nu^\varepsilon(I) > 0$, *if* $0 < \varepsilon < \varepsilon_I$.

Let us verify that *the existence of such a family of measures* (ν^ε) *guarantees the ultradensity of* $\mathcal{B}(u)$ *in* S.

The following argument is classical, see, for instance, [12, Lemma 1.2.8]. Let $I \subset S$ be an arbitrary interval in S. Put $\Sigma := \mathcal{B}(u) \bigcap \mathbb{B}$, so that $\nu_\varepsilon(\Sigma) = \nu^\varepsilon(I)$ due to (a). For any covering $\{I_j\}_{j=1}^\infty$ of the set Σ we have

$$0 < \nu^\varepsilon(I) = \nu^\varepsilon(\Sigma) \le \sum_j \nu^\varepsilon(I_j) \le c_1 \sum_j |I_j|^{1-c_2\varepsilon},$$

if $0 < \varepsilon < \min\left\{\varepsilon_0, \varepsilon(I), \frac{1}{c_2}\right\}$. For such an ε we see that $(1-c_2\varepsilon)$-Hausdorff measure of Σ is positive, hence $\dim \Sigma = 1$.

The plan of the construction of ν^ε is laid down in Section 8.2. The construction itself (and the proof of (a)-(c)) is done in Sections 8.3 - 8.6. We note that the measures ν^ε are probability measures (i.e. $\nu_\varepsilon(S) = 1$).

We call the measures ν^ε $(= \nu^{\varepsilon,u})$ the Bourgain measures (B-measures) of the function u. The idea to use these measures to prove the ultradensity of B-points is borrowed from [15], [16]. The main difference here lies in the construction of B-measures. Our argument works not only for the case when Ω is the ball or the upper halfspace, but also when Ω is the star-like domain.

8.2 Construction of measures ν^ε: outline

8.2.1 The plan

It remains to construct a family of B-measures that satisfy conditions (a), (b), (c) (see Section 8.1.3). It is done in Sections 8.3-8.6 according to the plan we introduce below.

By $M_+(S)$ we denote the set of finite (positive) Radon measures on $\mathbb{R}^{d+1}$ supported on S. Below we define a family of mappings $(\mathcal{W}_{\varepsilon,u})_{0<\varepsilon<\varepsilon_0}$ of the set $M_+(S)$ into itself, such that $\nu^\varepsilon := \mathcal{W}_{\varepsilon,u}(\nu)$ satisfies (a), (b), (c), for any nonzero $\nu \in M_+(S)$. This proves Theorem 8.0.1.

Remark. The mappings $\mathcal{W}_{\varepsilon,u}$ are actually restrictions on $M_+(S)$ of the linear operators that map the set $M(S)$ of finite Borel charges on $\mathbb{R}^d$ supported on S into itself, and, moreover, $\mathcal{W}_{\varepsilon,u}(\nu)(S) = \nu(S)$ for any $\nu \in M(S)$.

8.2.2 Kernels $\psi_{t,u,\varepsilon}$

Let $\varepsilon \in (0, \varepsilon_0)$ be sufficiently small, and from now on we stop mentioning ε_0. The measure $\mathcal{W}_{\varepsilon,u}[\nu]$ is obtained via a continuous transformation of the measure $\nu \in M_+(S)$, that depends on the parameter t. For $t \in (0,1)$ we define a positive kernel $\psi_t (= \psi_{t,\varepsilon,u}) \in C(\bar{S} \times \bar{S})$ (and, consequently, an integral operator Ψ_t), such that

$$\int_S \psi_t(x,\xi)\, d\xi = \Psi_t[1](x) \equiv 1$$

for any $t \in (0,1)$. Given a probability measure $\nu \in M_+(S)$ and a Borel set $E \subset S$ let

$$\begin{aligned} \nu_t(E) :&= \int_S \int_S \mathbb{1}_E(x)\psi_t(\xi,x)\, d\nu(\xi)\, dx \\ &=: \Psi_t^*[\nu](E),\ t \in (0,1). \end{aligned} \tag{8.15}$$

Clearly ν_t is a probability measure supported on S. The measure $\nu^\varepsilon := \mathcal{W}[\nu]\ (= \mathcal{W}_{\varepsilon,u}[\nu))$ is defined as the weak limit $\lim_{t\downarrow 0} \nu_t s$ (its existence is proven in Section 8.3.4). The operators Ψ_t tend to the identity as $t \uparrow 1$ (see Section 8.3.3). More precisely, $\lim_{t\uparrow 1} \Psi_t[\varphi] = \varphi$ uniformly on S for any $\varphi \in C(\bar{S})$. From now on Ψ_1 is the identity operator on $C(\bar{S})$, and Ψ_1^* is the identity mapping of $M(\bar{S})$ into itself.

We thus see that for $t = 1$ we start with a unit mass ν on S, and then we gradually redistribute it as t tends to zero. Every $t \in (0,1)$ corresponds to the distribution $\Psi_t^*[\nu] = \nu_t$. Passing to the limit we acquire the desired distribution $\nu^\varepsilon = \mathcal{W}[\nu]$, that is 'adjusted' to u, that is $\int_S \operatorname{Mvar} u\, d\nu^\varepsilon$ is finite. Hence the *existence* of Bourgain points of u is provided. Their *ultradensity* is implied by the fact that the family $\{\nu^\varepsilon\}$ satisfies (b) and (c). These conditions are proven in Section 8.5. Condition (a) is deduced at once from some properties of the kernels ψ_t.

8.2.3 Two key facts about ψ_t

These facts are shown in Sections 8.3-8.4.

Let φ be a positive harmonic function on Ω that has a finite limit at infinity. As before, we put $\varphi_{[t]}(x) := \varphi(x,t),\ x \in S,\ t > 0$, so that $\varphi_{[t]} \in C(\bar{S})$. Let $0 < \theta < t \leq 1$. *Then*

i. $\Psi_\theta \left[\varphi_{[t]}\right] \leq C\Psi_t \left[\varphi_{[t]}\right]$;

ii. if $\lim_\infty \varphi = 0$, then for any $x \in S$ *the function* $f^x : t \mapsto \Psi_t \left[\varphi_{[t]}\right]$ *is continuously differentiable on* $(0,1]$, and
$$\left(\frac{\partial}{\partial t} f^x\right)(t) = \varepsilon \Psi_t \left[B_t \left[\varphi_{[t]}\right]\right](x), \quad x \in S,\ t \in (0,1] \tag{8.16}$$

The statement (i) is proven in Section 8.3.3, statement (ii) and the continuity of $B_t \left[\varphi_{[t]}\right]$ is shown in Section 8.4.

8.2.4 Mean variation is finite

Now we show that (i) and (ii) imply that $\int_S \operatorname{Mvar} u \, d\nu^\varepsilon$ is finite, i.e. the property (a) of the measure ν^ε.

Given $t \in (0,1]$ we let
$$g_t := (h^{2t})_{[t]} = B_t \left[u_{[t]}\right], \tag{8.17}$$
see (8.10) and (8.11). We note that for any $t \in (0,1]$ the function g_t coincides on S with some positive and harmonic function in Ω, that vanishes at infinity. To prove (a) it is enough to show that $\displaystyle\int_S \int_\delta^1 g_t \, dt \, d\nu^\varepsilon$ is uniformly bounded for $\delta \in (0,1)$. For such a δ the function $x \mapsto \int_\delta^1 g_t(x) dt,\ x \in S$, coincides on S with some positive harmonic function on Ω, that vanishes at infinity (see Section 8.7). Therefore, due to (i) we have

$$\begin{aligned}
&\lim_{\theta \to 0} \int_S \left(\int_\delta^1 g_t \, dt\right) d\nu_\theta = \lim_{\theta \to 0} \int_S \left(\Psi_\theta \left[\int_\delta^1 g_t \, dt\right]\right) d\nu \\
&= \lim_{\theta \to 0} \int_S \left(\int_\delta^1 \Psi_\theta[g_t] \, dt\right) d\nu \leq C_{(\mathrm{i})} \int_S \left(\int_\delta^1 \Psi_t(g_t) \, dt\right) d\nu.
\end{aligned}$$

But it follows from (ii), (8.16) and (8.17), that the last integral is

$$\begin{aligned}
&\frac{C_{(\mathrm{i})}}{\varepsilon} \int_S \left(\int_\delta^1 \frac{\partial}{\partial t} \Psi_t \left[u_t\right] \, dt\right) d\nu = \frac{C_{(\mathrm{i})}}{\varepsilon} \int_S \left(\Psi_1 \left[u_{[1]}\right] - \Psi_\delta \left[u_{[\delta]}\right]\right) d\nu \\
&\leq \frac{C_{(\mathrm{i})}}{\varepsilon} \int_S \Psi_1 \left[u_{[1]}\right] d\nu \leq \frac{C_{(\mathrm{i})}}{\varepsilon} \sup_S u_{[1]}
\end{aligned}$$

(it is precisely in the penultimate inequality that we used the positivity of u). Recall that u vanishes at infinity so u_1 is bounded on S. We are done.

8.2.5 Differential equations (8.16)

We have to solve these equations (8.16), and $t \mapsto \Psi_t$, is the unknown operator-function variable. In order to do this we use a (version of) a well known method of solving linear differential equations in a vector space ([27]). We construct an operator function $J \mapsto \Psi_J$ that maps a compact interval $J \subset (0,1]$ to an integral operator Ψ_J with positive kernel $\psi_J \in C(S \times S)$. This function – '*multiplicative integral*' – satisfies the following condition:

$$0 < a < b < c \Rightarrow \psi_{[a,c]} = \psi_{[b,c]} \circ \omega_{[a,b]}. \tag{8.18}$$

The kernels ψ_t are defined as follows $\psi_t := \psi_{[t,1]}$, $0 < t < 1$. In Sections 8.3-8.6 we make sure that this choice of kernels provides all the necessary properties of ν^ε.

8.3 Kernels ψ_J, weak convergence of ν_t, condition (i)

Here and in the next few Sections we will make proclamations about this or that properties of ψ_t, and prove the properties we desired above modulo these proclamations. All will converge in due sense in Section 8.6. Also we will assume the value of ε to be fixed and small (essentially we will choose it at the very end of our chain of arguments), so we do not use it in the notation.

8.3.1 Kernels b_J, $\tilde{\psi}_J, \psi_J$

By segm_+ we denote the set of all non-degenerate compact intervals (segments) in $(0, +\infty)$. Given $J = [a,b] \in \mathrm{segm}_+$ let

$$\begin{aligned} &m(J) = a,\ M(J) = b,\ |J| = b - a, \\ &b_J(x,\xi) = \int_J b_\theta(x,\xi)\, d\theta, \quad x, \xi \in S. \end{aligned} \tag{8.19}$$

It follows from (8.13a) and (8.6) that

$$|b_J| \le c \frac{p_{m(J)}}{m(J)} |J|, \tag{8.20}$$

for $M(J) \le 1$. We call the segment $J \in \mathrm{segm}_+$ *short*, if $|J| \le m(J)$. Given a short segment $J \subset (0,1]$ we have

$$|b_J| \le c' p_{|J|}, \tag{8.21}$$

again, due to (8.6). Note that for any $J \in \mathrm{segm}_+$

$$B_J[1] = \int_J B_\theta[1]\, d\theta = 0.$$

Let

$$\tilde{\psi}_J := p_{|J|} - \varepsilon b_J, \quad J \in \mathrm{segm}_+.$$

Clearly, $\tilde{\Psi}_J[1] = 1$. The kernel $\tilde{\psi}_J$ *is positive*, if J is short and ε is small enough (see (8.21), we stress again that the choice of maximal ε_0 is postponed until the very end). Moreover under these conditions we have

$$(1 - c\varepsilon)p_{|J|} \leq \tilde{\psi}_J \leq (1 + c\varepsilon)p_{|J|} \tag{8.22}$$

We proceed by accumulating properties of kernels ψ_J postponing their construction until later. For sufficiently small $|J|$ the kernel ψ_J can be considered as a small perturbation of $\tilde{\psi}_J$. Namely, if we put $r_J := \psi_J - \tilde{\psi}_J$, we shall see that

$$|r_J| \leq c\varepsilon^2 \frac{|J|^2}{m(J)^2} p_{m(J)}. \tag{8.23}$$

Then (8.6) implies

$$|r_J| \leq c'\varepsilon^2 p_{|J|}, \tag{8.24}$$

for short J. *The kernel ψ_J is therefore positive.* Indeed, then we have by (8.22)

$$\psi_J = \tilde{\psi}_J + r_J \geq (1 - c''\varepsilon)\, p_{|J|},$$

for short J. Since an arbitrary J is a disjoint union $\bigsqcup_{n=1}^N J_n$, of short intervals J_n, we see that $\psi_J = \psi_{J_N} \circ \cdots \circ \psi_{J_1} > 0$ (see (8.18)).

We also note that $R_J[1] = 0$, hence

$$\Psi_J[1] = 1.$$

8.3.2 Behaviour of ψ_t for small t

Here we estimate kernels $\psi_{[\theta,1]}$ for small positive θ. First we observe that $[t, 2t]$ is short, therefore

$$\psi_{[t,2t]} \leq \tilde{\psi}_{[t,2t]} + c\varepsilon p_t \leq p_t + c'\varepsilon \int_t^{2t} \frac{p_y}{y}\, dy + c\varepsilon p_t,$$

where c and c' are absolute positive constants (see (8.13), (8.21)). Further, by (8.6)

$$\int_t^{2t} \frac{p_y}{y}\, dy \leq \int_t^{2t} c' \frac{p_t}{t} \cdot \frac{y}{t}\, d\theta = c' p_t.$$

Hence

$$\psi_{[t,2t]} \leq C p_t, \quad C = 1 + \varepsilon c' > 0,\ 0 < t \leq 1, \tag{8.25}$$

we assume that $\varepsilon < 1$. The estimate of $\psi_{[\theta,1]}$ follows from (8.25): *if* $0 < \theta < \frac{1}{2}$, *then*

$$\psi_{[\theta,1]} \leq c \cdot \frac{1}{\theta^{c\varepsilon}} \cdot p_{1-\theta}. \tag{8.26}$$

Indeed, by writing $\psi_{[\theta,1]}$ as a composition of $\psi_{\left[2^k\theta,2^{k+1}\theta\right]}$ and $\psi_{[2^K\theta,1]}$ for some $K \approx \log\frac{1}{\theta}$, and applying (8.25) to each term, we get

$$\psi_{[\theta,1]} \le (1+c\varepsilon)^{K+1} p_{1-2^K\theta} \circ p_{2^{K-1}\theta} \circ \cdots \circ p_\theta = (1+c\varepsilon)^{K+1} p_{(1-2^K\theta)+(2^{K-1}+\cdots+1)\theta}$$
$$= (1+c\varepsilon)^{K+1} p_{1-\theta} \le= \frac{c}{\theta^{c'\varepsilon}} p_{1-\theta}.$$

Analogously, under the same conditions we deduce from (8.25) the following estimate

$$\psi_{[\theta,1]} \ge c''\theta^{c''\varepsilon} \cdot p_{1-\theta}, \quad c'' > 0. \tag{8.27}$$

Also, the last term $p_{1-\theta}$ looks like a constant function for small values of θ, so the estimates above are uniform in a way.

8.3.3 Focusing property of the operator Ω_Δ

Given a short segment J the operator Ψ_J behaves like the Poisson operator $P_{|J|}$: for a class of functions ψ defined on S, the function $\Psi_J[\varphi]$ converges to φ as $|J| \to 0$; whereupon $\xi \mapsto \psi_J(x,\xi)$, $\xi \in S$ 'is focusing onto $x \in S$', becoming similar to δ_x. In what follows we use this 'focusing property' repeatedly.

Lemma 8.3.1 *Let $t \in (0,1)$ and φ be a positive harmonic function on Ω. Then for any $J \in$ segm$_+$, $J \subset (0,t]$ we have*

$$\left|\Psi_J\left[\varphi_{[t]}\right] - \varphi_{[t]}\right| \le c\frac{|J|}{t}\varphi_{[t]} \tag{8.28}$$

everywhere on S (and the constants do not depend on ε).

Proof. Let $I \in$ segm$_+$, $I \subset J$, and $f := \varphi_{[t]} = \varphi(\cdot,t)$. Then

$$|\Psi_I[f] - f| \le |P_{|I|}[f] - f| + |B_I[f]| + |R_I[f]| =: I + II + III. \tag{8.29}$$

The first term is very simple, since by Harnack's inequality

$$\left|P_{|I|}[f] - f(x)\right| \le \int_t^{t+|I|} |\nabla\varphi|(x,\theta)\,d\theta \le C|I||\nabla\varphi|(x,t) \le C\frac{|I|}{t}\varphi(x,t).$$

We deal with II next. Given $\theta \in I$, $x \in S$ we have

$$|C_\theta[f](x)| \le |\nabla\left(P_\theta[f]\right)(x)| = \left|\nabla\varphi_{[t+\theta]}(x)\right| \le c_3\frac{\varphi(x,t)}{t},$$

therefore

$$II \le \int_I P_\theta\left[|C_\theta[f]|\right]\,d\theta \le \frac{c_3}{t}\int_I \varphi(x,t+\theta)\,d\theta \le \frac{c_4|I|}{t}\varphi(x,t).$$

Finally (8.23) implies that for $x \in S$

$$\begin{aligned} III(x) &\le c_5 \int_S \frac{|I|^2}{m(I)^2} p_{m(I)}(x,\xi) f(\xi)\, d\xi \\ &\le c_5 \frac{|I||J|}{(m(J))^2} \varphi(x, t+m(I)) \le c_6 f(x) \frac{|I||J|}{t(m(J))^2}. \end{aligned} \tag{8.30}$$

It follows from (8.29) – (8.30) that

$$(1-\theta_I) f \le \Psi_I[f] \le (1+\theta_J) f, \tag{8.31}$$

where $0 < \theta_I \le c_7 \frac{|I|}{t}\left(1 + \frac{|I||J|}{(m(J))^2}\right)$.

Now we decompose J into N nonoverlapping segments $I_1, I_2, \ldots, I_N$ of the same length. Let $N = N(J,t)$ be large enough, so that

$$\theta_{I_n}(:= \theta_n) \le 2c_7 \frac{|J|}{Nt} < \frac{1}{2}, \quad k = 1, \ldots, K.$$

Then due to (8.31) and (8.17),

$$\begin{aligned} &\Psi_J[f] = \prod_{n=N}^{1} \Psi_{I_n}[f] \le \prod_{n=N}^{2} \Psi_{I_n}[(1+\theta_1) f) \le \cdots \le \\ &\left(1 + 2c_7 \frac{|J|}{Nt}\right)^N f \le \left(1 + c_8 \frac{|J|}{t}\right) f \end{aligned} \tag{8.32}$$

(we recall that $\frac{|J|}{t} \le 1$, and the kernel of Ψ_{I_n} is positive). Further

$$\Psi_J[f] \ge \left(1 - 2c_7 \frac{|J|}{Nt}\right)^N f \ge \left(1 - c_9 \frac{|J|}{t}\right) f. \tag{8.33}$$

Now (8.32) and (8.33) imply (8.28). $\square$

As an immediate corollary of Lemma 8.3.1 we get (i) from Section 8.2.3. To prove this we first put $J := [\theta, t]$.

By Lemma 8.3.1 we have:

$$\Psi_J\left[\varphi_{[t]}\right] \le (1+c)\varphi_{[t]}.$$

Since $[\theta, 1] = J \bigcup [t, 1]$, we get

$$\Psi_\theta\left[\varphi_{[t]}\right] = \Psi_t(\Psi_J\left[\varphi_{[t]}\right] \le (1+c)\Psi_t\left[\varphi_{[t]}\right]$$

and we are done.

8.3.4 Weak convergence of measures $\gamma_y s$

Here we show that as $t \downarrow 0$ *the measures* $\nu_t s$ *converge weakly on* $\bar{S}$ *to some measure* $\nu_0\,(=\nu^\varepsilon$ *supported on* S, *such that* $\nu_0(S)=1$.

Proof. Since all measures ν_t are probability measures, we can find a decreasing sequence $\{t_k\}\downarrow 0$ in $(0,1)$, such that ν_{t_k} converge weakly on $\bar{S}$ to some measure:

$$\lim_{k\to\infty}\int_S f d\nu_{t_k} = \int_{\bar{S}} f\, d\nu_0, \tag{8.34}$$

for any $f\in C(\bar{S})$. Let us verify that $\nu_0(\{\infty\})=0$, so that $\nu_0(S)=1$, and the mass of ν_0 does not run away to infinity. Consider a large interval $I\subset S$ with $|I|=L$ that is centered at zero. By $\omega_L(x,t)$ we denote the harmonic measure of I in Ω taken at (x,t). Now choose a large enough number M in such a way that

$$\omega_L(x,t)\le \frac{1}{2},\quad x\in S\setminus M\cdot I,\quad 0<t\le 1,$$

here $M\cdot I$ is the interval of length $M|I|$ concentric to I. Clearly $\omega_L(x,t)\in C(\bar{S})$ for a given t. Further

$$\begin{aligned}
&\nu_0(\{\infty\})\le \nu_0(\bar{S}\setminus M\cdot I)\le 2\int_{\bar{S}}(1-\omega_L)(x,t)\,d\nu_0(x)=\lim_{k\to\infty}2\int_S(1-\omega_L)(x,t)\,d\nu_{t_k}(x)\\
&=\lim_{k\to\infty}\int_S \Psi_{t_k}\left[(1-\omega_L)_{[t]}\right]\,d\nu\le \frac{c(S)}{t}\int_S(1-\omega_L)_{[t]}\,d\nu(x),
\end{aligned}$$

here we used the focusing property of Ψ_{y_k} and the harmonicity of $1-\omega_L$. Since ν is a probability measure on S, and $\omega_L(x,t)\to 1$ as $L\to\infty$, we see that the last integral vanishes.

It remains to show that

$$\lim_{t\downarrow 0}\int_S f\,d\nu_t=\int_S f\,d\nu_0$$

for any $f\in C(\bar{S})$. Clearly it is enough to consider traces of positive harmonic functions, so we assume f to be such a trace, $f(x)=\varphi(x,\theta)$, $x\in S$ for some positive harmonic φ and $t>0$. Using Lemma 8.3.1, multiplicative property of ψ_J and the fact that $\Psi_t[1]=1,\ \nu(S)=1$, we obtain:

$$\begin{aligned}
&\left|\int_S f\,d\nu_t-\int_S f\,d\nu_{t_k}\right|=\left|\int_S \Psi_{t_k}\left[\Psi_{[t,t_k]}[f]-f\right]\,d\nu\right|\\
&\le \|\Psi_{[t_k,t]}[f]-f\|_\infty\le c(S)\frac{t_k}{\theta}\|f\|_\infty,
\end{aligned}$$

and we arrive to the desired conclusion. □

8.4 Identity (ii) for operator functions $t\mapsto\Psi_t$

In this section we (still taking for granted the existence of ψ_J and their properties) demonstrate the equations (ii) that have already been used in Section 8.2.4.

Let φ be a positive harmonic function on Ω that vanishes at infinity. Let

$$f^x(t) := \Psi_t\left[\varphi_{[t]}\right](x), \quad t \in (0,1],\ x \in S,$$

where, as before, $\Psi_t := \Psi_{[t,1]}$ and Ψ_1 is the identity operator. To compute the derivative of f^x we first prove that it is Lipschitz on any segment $[t_0, 1]$, $0 < t_0 < 1$. Following that (Section 8.4.1) we compute the *left* derivative $(f^x)'_-$. As we show later it exists everywhere on $(0,1]$ and is continuous (the right derivative $(f^x)'_+$ is trickier to handle since the kernels (8.16) are not necessarily commutative). The Lipschitz property of $f|_{[t_0,1]}$, $t_0 \in (0,1]$ implies that

$$f^x(y) = f^x(1) - \int_t^1 (f^x)'_-(\theta)\, d\theta, \quad t \in (0,1].$$

We therefore have $(f^x)'_-(t) = (f^x)'_+(t)$, $t \in (0,1]$, and, consequently, $f^x \in C^1((0,1])$.

Now we show that $f^x|_{[t_0,1]}$ is Lipschitz. Let $t \in (0,1]$, $\delta > 0$, $t - \delta \geq t_0$, $J := [t,1]$. Considering (8.17) we have

$$f^x(t) - f^x(t-\delta) = \left(\Psi_t\left[I + II\right]\right)(x) + III(x), \tag{8.35}$$

where

$$\begin{aligned} I &:= \varphi_{[t]} - \Psi_J\left[\varphi_{[t]}\right], \\ II &:= \varphi_{[t]} - \varphi_{[t-\delta]}, \\ III &:= (\Psi_{t-\delta} - \Psi_t)[II]. \end{aligned}$$

Setting $N := \sup_{t_0 \leq t \leq 1} \|\varphi_{[t]}\|_\infty$, we see that focusing property of ψ_J (Lemma 8.3.1) implies:

$$\|I\|_\infty \leq c_1(S) N \frac{\delta}{t_0}.$$

By Harnack's inequality for any $x \in S$ there exists $\theta = \theta(x) \in (t-\delta, t)$ such that

$$|II|(x) \leq |\nabla\varphi(x,\theta)|\delta \leq c_2(S) N \frac{\delta}{t_0}.$$

Hence

$$\|\Psi_t[I + II]\|_\infty \leq \|I\|_\infty + \|II\|_\infty \leq c_3(S) N \frac{\delta}{t_0}.$$

Finally,

$$\|III\|_\infty \leq 2\|II\|_\infty \leq 2c_3(S) N \frac{\delta}{t_0},$$

so that

$$|f^x(t) - f^x(t-\delta)| \leq c_4(S) N \frac{\delta}{t_0}, \quad x \in S,\ 0 < t_0 \leq t - \delta < t \leq 1.$$

We are done.

8.4.1 Computing the derivative $(f^x)'$

It follows from (8.35) that for $0 < \delta \le \dfrac{t}{2}$

$$\frac{f^x(t-\delta) - f^x(t)}{-\delta} = \Psi_t[A_1](x) + A_2(x), \quad x \in S,\ 0 < t \le 1 \tag{8.36}$$

where $A_1 := \dfrac{I + II}{2}$, $A_2 := \dfrac{III}{\delta}$ (I, II, III are the same as in (8.35)).
Let us verify that $\lim_{\delta \downarrow 0} A_2 = 0$ on S. Indeed, $|\nabla \varphi(x, \theta)|$ is uniformly bounded for $\frac{t}{2} \le \theta \le 2t$ by Harnack's inequality, also $\Psi_\theta[1] = 1$, hence

$$\frac{1}{\delta}\left(\Psi_{t-\delta} - \Psi_t\right)\left[\int_{t-\delta}^{t} \frac{\partial}{\partial \theta}\varphi_{[\theta]}\, d\theta\right] = \frac{1}{\delta}\int_{t-\delta}^{t}\left(\Psi_{t-\delta} - \Psi_t\right)\left[\frac{\partial}{\partial \theta}\varphi_{[\theta]}\right] d\theta$$

and the integrand converges to zero.
It remains to prove that the following equality holds on S

$$\lim_{\delta \downarrow 0} \frac{I + II}{\delta} = \varepsilon B_t\left[\varphi_{[t]}\right], \tag{8.37}$$

since then

$$\lim_{\delta \downarrow 0} \Psi_t[A_1] = \varepsilon \Psi_t\left[B_t\left[\varphi_{[t]}\right]\right]$$

would follow immediately by dominated convergence. We have

$$\lim_{\delta \downarrow 0} \frac{I + II}{\delta} = \lim_{\delta \to 0}\left(\frac{\varepsilon}{\delta}\int_{t-\delta}^{t} B_\theta\left[\varphi_{[t]}\right]\, d\theta + \frac{1}{\delta} R_{[t-\delta,\delta]}\left[\varphi_{[t]}\right]\right) = \varepsilon B_t\left[\varphi_{[t]}\right] + \lim_{\delta \downarrow 0}\frac{1}{\delta} R_{[t-\delta,t]}\left[\varphi_{[t]}\right] \tag{8.38}$$

since $\theta \mapsto B_\theta\left[\varphi_{[t]}\right](x)$ is continuous on $[t - \delta, t]$ for any $x \in S$, see 8.7. The second term in the expression above clearly vanishes due to (8.23).

8.5 Properties (b) and (c) of measures ν^ε

8.5.1 Positivity of $\nu^\varepsilon(I)$

Lemma 8.5.1 *Given an interval $I \subset S$ there exists ε_I such that for any $\varepsilon \le \varepsilon_I$ we have $\nu^\varepsilon(I) > c$, where $c = c(S, I, \nu) > 0$.*

Proof. Let φ be a harmonic extension of some partition of unity element attached to I, that is the boundary data $\varphi_{[0]}$ is a smooth function between 0 and 1, also $\varphi_{[0]} \equiv 1$ on $\frac{1}{2}I$, and $|\nabla \varphi_{[0]}| \le \frac{2}{|I|}$. It is enough to show that for ε small enough

$$\int_S \varphi_{[t]}\, d\nu^\varepsilon \ge c(S, I) > 0. \tag{8.39}$$

This, in turn, follows from

$$\int_S \varphi_{[t]}\, d\Psi_\delta^*(\nu) \geq c(S, I, \nu) > 0, \text{ where } 0 < \delta < y < |I|. \tag{8.40}$$

We always can assume that $|I| \leq 1$. The estimate (8.40) is equivalent to

$$\int_S \Psi_{[\delta,t]}\left[\varphi_{[t]}\right]\, d\Psi_t^*(\nu) \geq c > 0,$$

that, due to Lemma 8.3.1 follows from

$$\int_S \varphi_{[t]}\, d\Psi_t^*(\nu) \left(= \int_S \Psi_t\left[\varphi_t\right]\, d\nu \right) \geq c_1 > 0 \tag{8.41}$$

If $t \in (0, |I|)$, then

$$\begin{aligned} \Psi_t\left[\varphi_{[t]}\right] &= \Psi_{|I|}\left[\varphi_{[|I|]}\right] - \int_t^{|I|} \frac{\partial}{\partial\theta}\left(\Psi_\theta\left[\varphi_{[\theta]}\right]\right)\, d\theta \\ &= \Psi_{|I|}\left[\varphi_{[|I|]}\right] - \varepsilon \int_t^{|I|} \Psi_\theta\left[B_\theta\left[\varphi_{[\theta]}\right]\right]\, d\theta \end{aligned} \tag{8.42}$$

By (8.27)

$$\Psi_{|I|}\left[\varphi_{[|I|]}\right] \geq c_2(S)|I|^\varepsilon P_{1-|I|}\left[\varphi_{[|I|]}\right] = c_2(S)|I|^\varepsilon \varphi_1, \tag{8.43}$$

hence

$$\int_S \Psi_{|I|}\left[\varphi_{[|I|]}\right]\, d\nu \geq c_3|I|^\varepsilon.$$

The second term in (8.42) is given by

$$\varepsilon \int_t^{|I|} \Psi_\theta\left[B_\theta\left[\varphi_{[\theta]}\right]\right]\, d\theta \leq \varepsilon \sup_{\theta\in(0,|I|)} \|B_\theta\left[\varphi_{[\theta]}\right]\|_\infty \cdot |I|. \tag{8.44}$$

On the other hand

$$\|B_\theta\left[\varphi_{[\theta]}\right]\|_\infty \leq c_4(S) \sup_\Omega |\nabla\varphi| =: \frac{c_5}{|I|} \tag{8.45}$$

(see the definition of b_t in 8.1.2 and the first of inequalities (8.8)). We arrive at

$$\int_S \Psi_t\left[\varphi_{[t]}\right]\, d\nu \geq c_3|I|^\varepsilon - c_5\varepsilon \geq \frac{c_3}{2},$$

for small values of ε. $\square$

8.5.2 Mass distribution of ν_ε

We move to (b). Let $I \subset S$ be any finite interval, we may assume $|I| \le \frac{1}{2}$, and let $g(x)$ be its harmonic measure in Ω taken at $(x, |I|) \in \Omega$. Since, clearly $g(x) \ge \frac{1}{10}$ on I, we have

$$\nu^\varepsilon(I) \le C \int_S g\, d\nu^\varepsilon = C \lim_{\delta \to 0} \int_S g\, d\Psi^*_\delta[\nu].$$

Taking $\delta < |I|$ and $J = [|I|, 1]$ we get

$$\int_S g\, d\Psi^*_\delta[\nu] = \int_S \Psi_{[\delta,|I|]}[g]\, d\Psi^*_J[\nu] \le C \int_S g\, d\Psi^*_J[\nu],$$

due to Lemma 8.3.1 and g being a trace of positive harmonic function on level $|I|$. Next we use (8.26) to see that

$$\int_S g\, d\Psi^*_J[\nu] \le C \cdot |I|^{-C\varepsilon} \int_S g(x) \int_S p_{1-|I|}(x, \xi)\, d\nu(\xi)\, dx \le$$
$$C_1 \cdot |I|^{-C\varepsilon} \int_S g(x)\, dx \le C_2 |I|^{1-C\varepsilon},$$

since $\int_S p_{1-|I|}(x,\xi)\, d\nu(\xi)$ is clearly uniformly bounded, and L^1-norm of g is just $|I|$.

8.6 Kernels ψ_J : existence, properties

8.6.1 Some extra notation

Consider $J \subset (0,1]$ and let $\Lambda \subset \mathrm{segm}_+$ be a finite set of nonoverlapping intervals such that $J = \bigcup_{I \in \Lambda} I$. We call such a set a *partition* of $J \in \mathrm{segm}_+$. Sometimes we understand Λ as a*family* $(I_k)_{k=1}^N$ of segments with positive (and nondecreasing) left endpoints: $0 < m(J) = m(I_1) < \cdots < m(I_N) < M(I_N) = M(J)$. The number $\delta_\Lambda := \max_{I \in \Lambda} |I|$ is called the *mesh* of the partition Λ; we write $\Lambda_2 \succ \Lambda_1$ if any element of Λ_2 lies in some element of Λ_1. Given a partition Λ of J we define the kernel ψ^Λ,

$$\psi^\Lambda := \tilde{\psi}_{I_N} \circ \tilde{\psi}_{I_{N-1}} \circ \ldots \tilde{\psi}_{N_1}$$

(see related definitions in Section 8.3.1). The kernel ψ_J is defined (in Section 8.6.5) as the limit of the sequence of kernels $(\psi^{\Lambda_n})_{n=1}^\infty$, where Λ_n is some sequence of partitions of J, with $\delta_{\Lambda_n} \to 0$.

8.6.2 Decomposition of ψ^Λ

By $\mathbf{N}_q$ we denote the set of the subsets of $\{1, 2, \ldots, N\}$ of cardinality q. We start with the following identity

$$\psi^\Lambda = (p_{|I_1|} - \varepsilon b_{I_1}) \circ (p_{|I_2|} - \varepsilon b_{I_2}) \circ \cdots \circ (p_{|I_N|} - \varepsilon b_{I_N}) =$$
$$p_{|J|} + \sum_{q=1}^N \sum_{\Sigma \in \mathbf{N}_q} \pi_\Sigma, \tag{8.46}$$

where $\pi_\Sigma := r^\Sigma_N \circ r^\Sigma_{N-1} \circ \cdots \circ r^\sigma_1$, and

$$r^\Sigma_j = \begin{cases} -\varepsilon b_{I_j}, & \text{if } j \in \Sigma, \\ p_{|I_j|}, & \text{if } j \notin \Sigma. \end{cases}$$

Consider the sum in (8.46) that corresponds to $q = 1$,taking into account that $b_J = \sum_{q=1}^N b_{I_q}$, while

$$\sum_{\Sigma \in \mathbf{N}_1} \pi_\Sigma = -\varepsilon \sum_{j=1}^N p_{|I_j^+|} \circ b_{I_j} \circ p_{|I_j^-|},$$

where $I_j^\pm$ is $[M(I_j), M(J)]$ or $[m(J), m(I_j)]$ respectively (if one of them degenerates to a point, then $p_{|I_q^\pm|}$ is understood as the composition identity) . It follows from (8.46) that

$$\psi^\Lambda = \tilde{\psi}_J + \varepsilon \sum_{j=1}^N v_j + \rho_\Lambda, \tag{8.47}$$

where $v_j := b_{I_j} - p_{|I_j^+|} \circ b_{I_j} \circ p_{|I_j^-|}$ and $\rho_\Lambda := \sum_{q=2}^N \sum_{\Sigma \in \mathbf{N}_q} \pi_\Sigma$.

8.6.3 Estimating $\psi^\Lambda - \tilde{\psi}_J$

We need to estimate the remainder terms v_j and ρ_Λ.

Kernels v_j

Note that for any θ, $\lambda > 0$ we have

$$|b_\theta| + |c_\theta| \le c(S)\tfrac{p_\theta}{\theta} \tag{8.48a}$$

$$|p_{\theta+\lambda} - p_\theta| \le c(S)\tfrac{\lambda}{\theta} p_\theta, \tag{8.48b}$$

if $\theta + \lambda < 1$. Now, (8.48a) follows from (8.8) and (8.13), and the second is fairly obvious, say, by Harnack's inequality again.

Let $I \in \mathrm{segm}_+,\ I \subset J$, and set $L(I) := \dfrac{|I|}{m(I)}$. Put $v_I := b_I - p_{|I^+|} \circ b_I \circ p_{|I^-|}$. We have

$$|v_I| \le |b_I - p_{|I^+|} \circ b_I| + |p_{|I^+|} \circ b_I - p_{|I^+|} \circ b_I \circ p_{|I^-|}|. \tag{8.49}$$

Using (8.48b) and (8.8) we obtain

$$\begin{aligned} &|b_I - p_{|I^+|} \circ b_I| \le \int_I |p_\theta c_\theta - p_{|I^+|+\theta} \circ c_\theta|\, d\theta \le c(S,J)|I^+| \int_I \frac{p_\theta}{\theta} \circ \frac{p_\theta}{\theta}\, d\theta \\ &= c(S,J)|I^+| \int_I \frac{p_{2\theta}}{\theta^2}\, d\theta \le c'(S,J)|I^+| \int_I \frac{p_{m(J)}}{\theta^2} \cdot \frac{2\theta}{m(J)}\, d\theta \\ &\le L(J)\frac{|I|}{m(J)} p_{m(J)}. \end{aligned} \tag{8.50}$$

As we move to deal with the second term in (8.49), let us recall that one can shave one half of the Poisson kernel off c_θ. In other words, if we look at (8.7) and use the semi-group property of p_θ, we get

$$c_\theta = \tilde{c}_\theta \circ p_{\frac{\theta}{2}}, \quad \theta > 0,$$

where $\tilde{c}_\theta$ is derivative of $p_{\frac{\theta}{2}}$ w.r.t. ϕ, and it satisfies the same estimates as c_θ itself. Hence

$$\begin{aligned}
&|p_{|I^+|} \circ b_I - p_{|I^+|} \circ b_I \circ p_{|I^-|}| \leq \\
&\int_I |p_{|I^+|+\theta} \circ (c_\theta - c_\theta \circ p_{|I^-|})| \, d\theta = \int_I \left| p_{|I^+|+\theta} \circ \tilde{c}_\theta \circ (p_{\frac{\theta}{2}} - p_{\frac{\theta}{2}+|I^-|}) \right| \, d\theta \\
&\leq c(S) \int_I \frac{1}{\theta} \frac{|I^-|}{\theta} p_{|I^+|+\frac{3}{2}\theta} \circ p_{\frac{\theta}{2}} \, d\theta \leq c''(S) \int_I \frac{|I^-|}{\theta^2} \frac{|I^+| + 2\theta}{m(J)} p_{m(J)} \, d\theta \\
&\leq c''(S,J) L(J)(3L(J)+2) \frac{|I|}{m(J)} p_{m(J)} \leq 3c''(S,J) \frac{M(J)}{m(J)} L(J) \frac{|I|}{m(J)} p_{m(J)}
\end{aligned} \tag{8.51}$$

we recall that $L(J) = \frac{M(J)}{m(J)} - 1$. It follows from (8.50), (8.51) and (8.49) that

$$|v_I| \leq c(S,J) \frac{M(J)}{m(J)} L(J) \frac{|I|}{m(J)} p_{m(J)},$$

and $c(S,J)$ increases with J. Returning to the partition Λ of the segment J (see (8.47)) we get

$$\sum_{j=1}^{N} |v_j| \leq c(S,J) \frac{M(J)}{m(J)} L(J) \sum_{j=1}^{N} \frac{|I|_k}{m(J)} p_{m(J)} = c(S,J) \frac{M(J)}{m(J)} (L(J))^2 p_{m(J)}. \tag{8.52}$$

Estimate of ρ_Λ

Going back to (8.46) we have

$$\rho_\Lambda = \sum_{q=2}^{N} \sum_{\Sigma \in \mathbf{N}_q} \pi_\Sigma,$$

If $j \in \Sigma$, then by (8.20)

$$|r_j^\Sigma| = \varepsilon |b_{I_j}| \leq c(S) \varepsilon L(I_j) p_{m(I_j)}$$

Assume that Λ is a *regular* partition, that is $\delta(\Lambda) \leq \frac{2|J|}{N}$. Then, putting $h := 2c(S)\varepsilon L(J)$, we obtain

$$|r_j^\Sigma| \leq \frac{h}{N} p_{m(I_j)}, \quad j \in \Sigma.$$

On the other hand if $j \notin \Sigma$, then $r_j^\Sigma = p_{|I_j|}$. It means that for $\Sigma \in \mathbf{N}_q$

$$|\pi_\Sigma| \leq h^q N^{-q} p_{a(\Sigma)},$$

where

$$a(\Sigma) := \sum_{j \in \Sigma} m(I_j) + \sum_{j \notin \Sigma} |I_j| \leq qM(J) + |J|$$

Taking into account that $\#\mathbf{N}_q = C_N^q \le \frac{N^q}{q!}$, we arrive at

$$
\begin{aligned}
|\rho_\Lambda| &\le \sum_{q=2}^{N} \frac{N^q}{q!} h^q N^{-q} c(S) \frac{qM(J)+|J|}{m(J)} p_{m(J)} \le Ch^2 \sum_{q=2}^{\infty} \frac{h^{q-2}}{(q-1)!} \frac{M(J)+|J|}{m(J)} p_{m(J)} \\
&\le C(2L(J)+1)h^2 e^h p_{m(J)}.
\end{aligned}
$$

Now we are ready to work with $\psi^\Lambda - \tilde{\psi}_J$. Combining the estimate above with (8.47) and (8.52) we deduce

$$
|\psi^\Lambda - \tilde{\psi}_J| \le c(S, L(J))(L(J))^2 p_{m(J)}, \tag{8.53}
$$

where the function $s \mapsto c(S,s)$ is increasing. Here we assume Λ to be a regular partition.

The estimate of the kernel ψ^Λ follows from (8.53) as well

$$
\begin{aligned}
&|\psi^\Lambda| \le |\tilde{\psi}(J)| + c(S, L(J))(L(J))^2 p_{m(J)} \le p_{|J|} + \\
&+ (c(S)\varepsilon L(J) + c(S, L(J))(L(J))^2 p_{m(J)}) = p_{|J|} + A \cdot (L(J))^2 p_{m(J)},
\end{aligned} \tag{8.54}
$$

here $A = A(L(J), S)$, and the function $x \mapsto A(x, S)$ is increasing.

8.6.4 Estimate of $\psi^{\tilde{\Lambda}} - \psi^\Lambda$, $\Lambda \succ \tilde{\Lambda}$

This is the main estimate of this Section, after we obtain it the construction of ψ_J is almost done.

Lemma 8.6.1 *Let $\tilde{\Lambda}$ be a partition of $J \in \mathrm{segm}_+$ and $\Lambda \succ \tilde{\Lambda}$. Then*

$$
|\psi^{\tilde{\Lambda}} - \psi^\Lambda| \le C(S, J)\delta(\tilde{\Lambda}) p_{m(J)} \tag{8.55}
$$

where $\delta(\tilde{\Lambda})$ is the mesh of partition (see Section 8.6.1).

Proof. Let $\tilde{\Lambda} = \{J_1, J_2, \dots, J_N\}$, $m(J_1) < m(J_2) \dots < m(J_N)$. Put $\Lambda_k := \{I \in \Lambda : I \subset J_k\}$ so that Λ_k is the partition of J_k, $\Lambda = \bigcup_{k=1}^N \Lambda_k$. For $i = 2, 3, \dots, N$ by Λ_i^- we denote the part of Λ that lies in J_i^-: $\Lambda_i^- = \bigcup_{1 \le q < i} \Lambda_q$; $\Lambda_1^- := \emptyset$. For $i = 1, \dots, N$ by $\tilde{\Lambda}_i^+$ we denote the part of $\tilde{\Lambda}$ that lies in J_i^+ : so $\tilde{\Lambda}_i^+ = \bigcup_{i < q \le N} J_q$; $\tilde{\Lambda}_{N+1}^+ := \emptyset$. Finally, let $\tilde{\Lambda}_i := \Lambda_i^- \bigcup \{J_i\} \bigcup \tilde{\Lambda}_i^+$, $1 \le i \le N$; $\tilde{\Lambda}_{N+1} := \Lambda$. Here the kernel $\psi^{\tilde{\Lambda}_i}$ is written as ψ_i, $i = 1, \dots, N+1$. In particular, $\psi_1 = \psi^{\tilde{\Lambda}}$, $\psi_{N+1} = \psi^\Lambda$,

$$
\psi^{\tilde{\Lambda}} - \psi^\Lambda = \sum_{i=1}^{K} (\psi_i - \psi_{i+1}).
$$

If $i \ne 1$, N, then

$$
\psi_i - \psi_{i+1} = \psi^{\tilde{\Lambda}_i^+} \circ (\tilde{\psi}_{J_i} - \psi^{\Lambda_i}) \circ \psi^{\Lambda_i^-}. \tag{8.56}
$$

This equality also holds for $i = 1$, N, if $\psi^\emptyset$ is understood as a identity convolution operator. It follows from (8.56), (8.53), (8.54), and the inequality $|J_i|^2 \le \delta(\tilde{\Lambda})|J_i|$ that

$$
|\psi_i - \psi_{i+1}| \le (p_{|J_i^+|} + Cp_{m(J_i^+)}) \circ \delta(\tilde{\Lambda})|J_i| p_{m(J)} \circ (p_{|J_i^-|} + Cp_{m(J_i^-)}), \tag{8.57}
$$

here $C = C(J,S)$. The right-hand side in (8.57) does not exceed

$$\frac{A(M(J)+|J|)}{m(J)}\delta(\tilde\Lambda)|J_i|p_{m(J)} = A(1+2L(J))\delta(\tilde\Lambda)|J_i|p_{m(J)},$$

where $A = A(\Delta, S)$. Therefore

$$|\psi^{\tilde\Lambda} - \psi^{\Lambda}| \le \sum_{i=1}^{N} |\psi_i - \psi_{i+1}| \le A(1+2L(J))|J|\delta(\tilde\Lambda)p_{m(J)}.$$

□

8.6.5 Dyadic partitions of ω_Δ

Let $n \in \mathbb{Z}_+,\ J \in \mathrm{segm}_+$. By $\tilde\Lambda_n(J)$ we denote the partition of J that consists of all intersections of the form $J \bigcap \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right],\ j \in \mathbb{Z}_+$. We call such a partition *dyadic of the rank* n. Clearly this is a regular partition, $\tilde\Lambda_{n+1}(J) \succ \tilde\Lambda_n(J),\ \delta(\tilde\Lambda_n(J)) \le \frac{|J|}{2^n}$. Put $\psi_n := \psi^{\tilde\Lambda_n(J)}$. Lemma 8.6.1 implies that $|\psi_n - \psi_{n+1}| \le C(J,S)\frac{|J|}{2^n}p_{m(J)}$. Therefore the series

$$\psi_J := \lim_{n\to\infty} \psi_n = \psi_1 + (\psi_2 - \psi_1) + (\psi_3 - \psi_2) + \dots \tag{8.58}$$

converges uniformly on $S \times S$ and defines the kernel ψ_J that satisfies

1. $\psi_J \in C(S \times S)$;

2. $\psi_J > 0$, if $\varepsilon \in (0, \varepsilon_0)$;

3. $\int_S \psi_J(x,\xi)\, d\xi = 1$ for any $x \in S$;

4. if $0 < a < b < c$, then $\psi_{[a,c]} = \psi_{[b,c]} \circ \psi_{[a,b]}$;

5. $|\psi_J - \tilde\psi_J| \le C(S)\varepsilon^2(L(J))^2 p_{m(J)}$, if J is short (i.e. if $L(J) \le 1$).

Proof. (1) follows from continuity of b_J (see Section 8.7) and from the fact that series in (8.58) converges uniformly on $S \times S$.
The statement (2) follows from the positivity of $\tilde\psi_J$ for small values of ε (see (8.21)).
To obtain (3) we note that $\tilde\psi_J(x,\xi)\, d\xi$ is a probability measure on S, therefore $\int_S \psi_n(x,\xi)\, d\xi \equiv 1$, and we can pass to the limit in the integral since by (8.55)

$$|\psi_n - \psi_J| \le C|J|2^{-n}p_{m(J)}.$$

Now we show (4). Let $J := [a, c]$, $J^- := [a, b]$, $J^+ := [b, c]$, $\tilde{\Lambda}_n'(J) := \tilde{\Lambda}_n(J^-) \bigcup \tilde{\Lambda}_n(J^+)$. Clearly, $\tilde{\Lambda}_n'(J) \succ \tilde{\Lambda}_n(J)$, hence Lemma 8.6.1 implies

$$|\psi^{\tilde{\Lambda}_n'} - \psi^{\tilde{\Lambda}_n}| \leq c(S, J) \frac{|J|}{2^n} p_{m(J)}.$$

We thus have $\lim_{n\to\infty} \psi^{\tilde{\Lambda}_n'(J)} = \psi_J$ everywhere on $S \times S$. On the other hand we see that

$$\lim_{n\to\infty} \psi^{\tilde{\Lambda}_n(J^+)} \circ \psi^{\tilde{\Lambda}_n(J^-)} = \psi_{J^+} \circ \psi_{J^-}.$$

We can pass to the limit on account of the estimates

$$|\psi^{\tilde{\Lambda}_n(J^\pm)}| \leq \tilde{\psi}_{J^\pm} + c p_{m(J)}$$

that follow from (8.53). It remains to note that $\psi^{\tilde{\Lambda}_n'(J)} \equiv \psi^{\tilde{\Lambda}_n(J^+)} \circ \psi^{\tilde{\Lambda}_n(J^-)}$.
We are left to prove (5). It follows from (8.53) that for $\Lambda := \tilde{\Lambda}_n(J)$ we have

$$|\psi^{\tilde{\Lambda}_n(J)} - \tilde{\psi}(J)| \leq c(S, L(J)) \varepsilon^2 (L(J))^2 p_{m(J)}, \quad n = 1, 2, \ldots,$$

where the function $x \mapsto c(S, x)$ is increasing. If J is a short segment (i.e. $L(J) \leq 1$), then passing to the limit we obtain

$$|\psi_J - \tilde{\psi}_J| \leq c(S, 1) \varepsilon^2 \left(\frac{|J|}{m(J)} \right) p_{m(J)}.$$

□

We are almost done – the only thing left to check is the continuity of b_t.

8.7 Kernels b_t are continuous

Here we consider the kernels b_t defined in 8.1.2, and show that they are continuous in all three variables. While it takes a bit more work to establish for Lipschitz domains in $\mathbb{R}^{d+1}$, it is almost immediate for $\Omega = \mathbb{R}^2_+$. Indeed, the zero sets of gradients of harmonic functions are discrete, and the boundary geometry is somewhat easier to deal with.

So, let us consider the kernel b_t written in the convolution – here is one of a select few places where we make our life easier by actually using a simpler structure of $\Omega = \mathbb{R}^2_+$ and $S = \mathbb{R}$. We have

$$\begin{gathered} b_t(x,\xi) = (p_t \circ c_t)(x,\xi) = \int_S p_t(x,\eta) c_t(\eta,\xi)\, d\eta = \\ \int_S p_t(x,\eta) \cdot \left(\frac{\partial}{\partial \eta} p_t(\eta,\xi) \cdot \phi_1(\eta, 2t) + \frac{\partial}{\partial t} p_t(\eta,\xi) \cdot \phi_2(\eta, 2t) \right) d\eta = \\ \frac{1}{\pi^2} \int_{\mathbb{R}} \frac{t}{t^2 + (x-\eta)^2} \left(\frac{2t(\eta - \xi)}{(t^2 + (\xi-\eta)^2)^2} \cdot \phi_1(\eta, 2t) + \frac{t^2 - (\xi-\eta)^2}{(t^2+(\xi-\eta)^2)^2} \cdot \phi_2(\eta, 2t) \right) d\eta, \end{gathered} \tag{8.59}$$

where $(\phi_1, \phi_2)(\eta, 2t) = \phi(\eta, 2t) = \dfrac{\nabla u(\eta, 2t)}{|\nabla u(\eta, 2t)|}$, if $|\nabla u(\eta, 2t)| \neq 0$, and 0 otherwise. Since $|\nabla u|$ is bounded on any half-plane $\{(x, \theta) : x \in \mathbb{R},\ \theta \geq \theta_0 > 0\}$ by Harnack's inequality, we see that the formula in (8.59) guarantees that $b_x(x, \xi)$ is continuous on $\mathbb{R} \times \mathbb{R} \times (0, +\infty)$, and we are done.

Conclusion

We wrap up the thesis with a couple of observations. First, the results presented here, especially those related to Potential Theory on d-trees can be considered as a notable advancement in the studies of the spaces on the polydisc. Moreover, we believe that it is the methods developed that constitute the main achievement, the energy majorization and extremal capacity estimates scan be successfully implemented in other models. A discretization of Bessel capacity seems also be of use. Second, there are numerous pathways open to continue the investigations. A greater understanding of the situation in higher dimensions and further analysis of Bourgain measures seem to be especially interesting.
We finish the thesis with a number of open questions and list of possible problems for the follow-up research.

Weighted Potential Theory on a graph

Non-linear theory

A natural question considering the weighted Hardy embedding (2.1) concerns its non-linear version. Namely we ask what are the conditions on w, μ that guarantee the boundedness of the embedding

$$\mathbf{I}_w : \ L^p(T^d, w) \to L^q(\overline{T}^d, \mu),$$

or, in other words, when the inequality

$$\left(\int_{\overline{T}^d} (\mathbf{I}_w f)^p \, d\mu\right)^{\frac{1}{p}} \leq C(w, \mu, p, q) \left(\int_{T^d} f^q \, dw\right)^{\frac{1}{q}} \tag{C.1}$$

holds true for any $f \in L^q(T^d, w)$? The one-dimensional case, $d = 1$, is rather well studied, see for example [6], [74] or [108]. No results are known so far for $d \geq 2$, even for $w \equiv 1$ and $p = q$. One of the (numerous) obstacles is that if one defines the non-linear potential on a 2-tree analogously to the linear case, it is not p-harmonic any more (it is p-harmonic for $d = 1$, while the linear version is still harmonic in each variable). Also it would be interesting to consider the questions raised in [91] on a bi-tree.

Higher dimensions

Another question is about what happens for product weights w and $d \geq 4$. Before we have shown that our arguments, strictly speaking, already do not work very well for $d = 3$, and we have to use the work-around (see the counterexample to two-function energy majorization, Proposition 2.6.2). The case of T^4 is therefore out of our reach for now. One of the possible approaches would be to modify the two-function Lemma 2.6.1 so it would work for very specific functions – slices of 3-dimensional measures and potentials.

Two-weight problem

Sawyer's result for d-trees

A careful reading of [86] provides the statement of Theorem 8.7.1 below. The setting in Sawyer's paper [86] can be reduced to the following situation. Assume that the weight $w : T^2 \to \mathbb{R}_+$ has a very special structure: it lives on ancestors of one small square. Namely, there exists a point $\omega_0 \in (\partial T)^2$ such that

$$\operatorname{supp} w \subset \{\alpha : \ \alpha \geq \omega_0\}.$$

If such w is identically 1 on ancestors of one boundary point ω_0, then it is a product weight, and we know the answer: condition (C.2c) below is necessary and sufficient for embedding. But for general weight sitting on ancestors of ω_0, Sawyer proved the following result.

Theorem 8.7.1 *Let w live on ancestors of $\omega_0 \in (\partial T)^2$.*

$$\sup\nolimits_{\beta \geq \omega_0} (\mathbf{I}^* \mu(\beta) \mathbb{I} w(\beta)) \leq A^2 < \infty, \tag{C.2a}$$

$$\textstyle\sum_{\alpha \geq \beta \geq \omega_0} \mu(\alpha)(\mathbf{I} w(\alpha))^2 \leq A^2 \sum_{\alpha \geq \beta} w(\alpha), \tag{C.2b}$$

$$\textstyle\sum_{\omega_0 \leq \alpha \leq \beta} w(\alpha)(\mathbf{I}^* \mu(\alpha))^2 \leq A^2 \sum_{\alpha \leq \beta} \mu(\alpha) \tag{C.2c}$$

for any $\beta \in T^2$. All of these three conditions are necessary.

Remark. The last condition (C.2c) is just the single-box test. In other words, if we restrict the weight to be supported only on the hooked rectangles, but drop the requirement that it has a product structure, we see that the single box test $[w, \mu]_B < \infty$ is getting replaced by three tests: one has to have three single box tests for the pair (μ, w).

Remark. This particular version of 2-tree has an unusual property – assume that w is also of a product nature, i.e. $w = w_1 \cdot w_2$, where w_j is supported on a single tree geodesic. Then, as it is shown in [96], any of these three conditions would suffice, even (C.2a). But this is just the sub-capacitary condition for singletons,

$$\mu(\{\alpha\}) \leq A \operatorname{Cap}_w(\{\alpha\}), \quad \forall \alpha \in T^2.$$

On the other hand, it is well known that on a tree T, even with $w \equiv 1$ single box subcapacitary condition is not sufficient, mostly due to lack of additivity of capacity (unless it is Lebesgue measure, of course). This difference can be attributed to a more 'connected' structure of T^2 restricted to $\mathcal{P}(\omega_0)$.

Another way to look at these conditions is to rewrite them in a capacitary language. A bi-tree T^2 with such a weight w attached actually has a sort of symmetric structure, and the weight w and a measure μ become interchangeable. To elaborate, for a measure μ that is supported on ancestors of a single point we define yet another version of capacity (we consider finite graphs for simplicity) as follows

$$\mathbf{W}^w_\mu(\alpha) := \sum_{\gamma\leq\alpha}(\mathbf{I}w)(\gamma)\mu(\gamma),$$

$$\mathrm{Pac}_\mu(F) := \inf_w \left\{ \sum_{\beta\in T^2}(\mathbf{I}w)^2(\beta)\mu(\alpha) :\ \mathbf{W}^w_\mu \geq 1 \text{ on } F \right\}.$$

Essentially this is a symmetric version of $\mathbf{V}$-potential and w-capacity – we reverse the ordering of our graph, and replace $\mathbf{I}^*$ with $\mathbf{I}$ and w with μ. It turns out that Sawyer's conditions (C.2) are also equivalent to

$$\mu(E) \leq A\,\mathrm{Cap}_w(E), \quad \forall E\subset T^2, \tag{C.3a}$$

$$w(F) \leq A\,\mathrm{Pac}_\mu(F), \quad \forall F\subset T^2. \tag{C.3b}$$

The 3-dimensional version of Sawyer's result is still wide open. The construction he employs depends on the fact that $d = 2$ in a crucial way. However, we hope that machinery developed in Section 2.5 combined with capacitary-pacacitary approach would allow to obtain results for higher dimensions.

General weights on 2-tree

It is very natural then to ask what are the conditions for a proper two-weight (i.e. for non-product weights w) problem on a d-tree, even in the case $d = 2$. The obvious candidates are three single-box tests (C.2) and potential-theoretic tests (C.3). An important thing to note, however, is that this problem would also include the discrete Carleson embedding for $H^2(\mathbb{D}^2)$ in the sense of Chang-Fefferman. If it is indeed can be covered by three single boxes, it would be very surprising. This question is also wide open, again since Sawyer's construction is quite rigid and can not be properly modified, at least so far.

Relation between analytic and harmonic embeddings

As we have seen above, the analytic version of the Carleson embedding theorem 3.2.3 even for 2-tree requires the space $\mathcal{H}_{\vec{s}}$ to be not far from the unweighted Dirichlet space $\mathcal{H}_{\vec{1}}$. This is not the case for harmonic spaces, and it is interesting to know, if this difference is the artefact of the proof, or if it has some inherent nature. In the latter case one would expect some kind of a critical curve in $\vec{s}$ to exist, such that below this curve the analytic and harmonic Carleson measures are different.

Also the question is quite intriguing already for $s = \vec{0}$ (though, technically, we did not consider such values of parameters). It turned out that the equivalence of harmonic and analytic embeddings is not known, and establishing (or disproving) such equivalence could shed new light on the multi-parameter Nehari theorem of [34].

Variation near the boundary

In our opinion the differential equality (I.68b) that defines the family Ψ_t and the Bourgain measures ν_θ generated by it are extremely interesting. In particular, it seems that these measures somehow contain information about the geometric structure of a zero set of derivatives of positive harmonic functions near the boundary. Hence the true nature of these measures is not yet properly illuminated.

A proper (i.e. not immediately trivial) discrete formulation of the radial variation problem for harmonic functions would also be very useful, by itself and also as a way to better understand the continuous situation.

Bibliography

[1] R. Adams, L. Hedberg, *Function Spaces and Potential Theory*, Springer 1999.

[2] N. Arcozzi, P. Mozolyako, K.-M. Perfekt, G. Sarfatti, *Carleson measures for the Dirichlet space on the bidisc*, arXiv:1811.04990, (2018), 1–44.

[3] N. Arcozzi, P. Mozolyako, G. Psaromiligkos, A. Volberg, P. Zorin-Kranich, *Bi-parameter Carleson embeddings with product weights*, arXiv:1906.11150 (2020), 1–24.

[4] N. Arcozzi, R. Rochberg, E. Sawyer, *Carleson Measures for Analytic Besov Spaces.* Rev. Mat. Iberoamericana **18** (2002), no. 2, 443–510.

[5] N. Arcozzi, R. Rochberg, E. Sawyer, *Carleson Measures for the Drury-Arveson Hardy Space and other Besov-Sobolev Spaces on Complex Balls.* Adv. Math. **218** (2008), no. 4, 1107–1180.

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

[7] N. Arcozzi, R. Rochberg, E. Sawyer, B.D. Wick, *The Dirichlet space and related function spaces.* American Mathematical Society, 2019.

[8] R. Bañuelos, C. N. Moore, *Probabilistic behavior of harmonic functions.* Birkhäuser, Basel-Boston-Berlin (1999).

[9] A. Barron, J. Pipher, *Sparse domination for bi-parameter operators using square functions*, arXiv:1709.05009, 1–22.

[10] I. Benjamini, Y. Peres, *Random walks on a tree and capacity in the interval*, Ann. Inst. H. Poincare Probab. Statist. **28** (1992), no. 4, 557-592.

[11] G. Bennett, D. A. Stegenga, R. M. Timoney, *Coefficients of Bloch and Lipschitz functions*, Ill. J. Math. **25** (1981), 520–531.

[12] C. Bishop, Y. Peres *Fractals in probability and analysis*, Cambridge University Press, 2017.

[13] A. Borichev, *On the minimum of harmonic functions*, J. Anal. Math. **89** (2003), 199–212.

[14] A. Borichev, Yu. Lyubarskii, E. Malinnikova, P. Thomas, *Radial growth of functions in the Korenblum space*, St. Petersburg Math. J. **21** (2010), no. 6, 877–891.

[15] J. Bourgain, *On the radial variation of bounded analytic functions on the disk*, Duke Math. J. **69** (1993), no. 3, 671–682.

[16] J. Bourgain, *Bounded variation of measure convolutions*, Math. Zametki **54** (1993), no. 4, 25–34 (transl. from Russian in J. Math Notes **54** (1993)).

[17] R. Cairoli, J.B. Walsh, *Stochastic integrals in the plane*, Acta math. **134** (1975), 11–183.

[18] A. Cantón, J.L. Fernández, D. Pestana, J.M. Rodríguez, *On harmonic functions on trees*, Potential Analysis, **15** (2001), 199-244.

[19] L. Carleson, *Interpolations by bounded analytic functions and the corona problem*, Ann. of Math. **76** (1962), no. 3, 547–559

[20] L. Carleson, *Selected Problems on Exceptional Sets*, Van Nostrand, 1967.

[21] L. Carleson, *A counter example for measures bounded on H^p for the bi-disc*, Mittag-Leffler Report No. **7** (1974).

[22] M.L. Cartwright, *On analytic functions regular in the unit circle, I*, Quart. J. Math. Oxford **4** (1933), 246–257.

[23] M.L. Cartwright, *On analytic functions regular in the unit circle, II*, Quart. J. Math. Oxford **6** (1935), 94–105.

[24] Sun-Yung A. Chang, *Carleson measure on the bi-disc*, Ann. of Math. **109** (1979), no. 3, 613–620.

[25] Sun-Yung A. Chang, R. Fefferman, *A continuous version of duality of H1 with BMO on the bidisc*, Ann. of Math. **112** (1980), no. 1, 179–201.

[26] Sun-Yung A. Chang, R. Fefferman, *Some recent developments in Fourier analysis and H^p-theory on product domains*, Bull. Amer. Math. Soc. **12** (1985), no. 1, 1–43.

[27] J.L. Daleckii, M.G. Krein, *Stability of solutions of differential equations in Banach spaces*, Translations of Mathematical Monographs, v. 43, AMS, Providence, Rhode-Island, 1974.

[28] I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, PA, 1992.

[29] L.E. Dor, *On projections in L^1*, Ann. of Math. **102** (1975), no. 3, 463–474.

[30] K.S. Eikrem, E. Malinnikova, *Coefficient multipliers of growth spaces of harmonic functions*, Integr. Equ. Oper. Theory **82** (2015), 555–573.

[31] R. Fefferman, *Harmonic analysis on product spaces*, Ann. of Math. **126** (1987), no. 1, 109–130.

[32] R. Fefferman, *Calderón-Zygmund theory for product domains: H^p spaces*, Proc. Nat. Acad. Sci. U.S.A. **83** (1986), no. 4, 840–843.

[33] R. Fefferman, *Some recent developments in Fourier analysis and H^p theory on product domains. II*, Function spaces and applications (Lund, 1986), 44–51, Lecture Notes in Math., 1302, Springer, Berlin, 1988.

[34] S. Ferguson, M. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), 143–160.

[35] J. L. Fernández, J. Heinonen, J. G. Llorente, *Asymptotic values of subharmonic functions*, Proc. London Math. Soc. **73.2** (1996), no. 3, 404–430.

[36] J.B. Garnett, D.E. Marshall, *Harmonic measure*, Cambridge University Press, Cambridge (2005).

[37] L. Grafakos, R. Torres, *Multilinear Calderón–Zygmund theory*, Adv. Math. **165**, 2002, 124–164.

[38] R.F. Gundy, *Inégalités pour martingales a un et deux indices: L'espace H^p*, in Ecole d'Eté de Probabibilités de Saint-Flour VIII-1978, Springer (1980).

[39] G. H. Hardy, *Weierstrass's non-differentiable function*, Trans. Amer. Math. Soc. **17** (1916), no. 3, 301–325.

[40] W. Hastings, *A Carleson Measure Theorem for Bergman Spaces*, Proceedings of the AMS **2** (1975), no. 1, 237–241.

[41] W.K. Hayman, P.B. Kennedy, *Subharmonic functions*, Vol. 1. Acad. Press, London, 1976.

[42] M. Holschneider, P. Tchamitchian, *Pointwise analysis of Riemann's "nondifferentiable" function*, Inv. Math. **105** (1991), no. 1 157–175.

[43] R. Hunt, R. Wheeden, *Positive harmonic functions on Lipschitz domains*, Trans. Amer. Math. Soc. **147** (1970), 507–527.

[44] R. Hunt, R. Wheeden, *On the boundary values of harmonic functions*, Trans. Amer. Math. Soc. **132** (1968), no. 2,307–322.

[45] A. Iosevich, B. Krause, E. Sawyer, K. Taylor, I. Uriarte-Tuero, *Maximal operators: scales, curvature and the fractal dimension*, Anal. Math. **45** (2019), 63–86.

[46] P. Jones, *A complete bounded submanifold in C^3*, Proc. Amer. Math. Soc. **76** (1979), no. 2 305–306.

[47] P. Jones, P.F.X. Müller, *Radial variation of Bloch functions.*, Math. Res. Lett. **4** (1997), no. 3, 395–400.

[48] J.-L. Journé, *Calderoón–Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), 55–91.

[49] J.-L. Journé, *Two problems of Calderón-Zygmund theory on product-spaces*, Ann. Inst. Fourier (Grenoble) **38** (1988), no. 1, 111–132.

[50] R. Kerman, E. Sawyer, *Carleson Measures and Multipliers of Dirichlet-type Spaces*, Trans. Amer. Math. Soc. **309** (1988), no. 1, 87–98.

[51] B. Korenblum, *An extension of the Nevanlinna theory*, Acta Math. **135** (1975), no. 3-4, 187–219.

[52] B. Korenblum, P.J. Rippon, K. Samotij, *On integrals of harmonic functions over annuli*, Ann. Acad. Sci. Fenn. **20** (1995), 3–26.

[53] M. T. Lacey, *Two-weight inequality for the Hilbert transform: a real variable characterization II*, Duke Math. J. **163** (2014), no. 15, 2821–2840.

[54] M. Lacey, *Lectures on Nehari's Theorem on the Polydisk*, Contemp. Math. Volume **444**, 2007, pp. 185–213.

[55] M. T. Lacey, E. T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, *Two-weight inequality for the Hilbert transform: a real variable characterization I*, Duke Math. J. **163** (2014), no. 15, 2795–2820.

[56] C.N. Linden, *Functions regular in the unit circle*, Quart. J. Math. Oxford **52** (1956), 196–216.

[57] C.N. Linden, *Inequalities for functions regular in the unit circle*, Quart. J. Math. Oxford **58** (1962), 26–37.

[58] J.G. Llorente, *Boundary values of harmonic Bloch functions in Lipschitz domains: a martingale approach*, Potential Analysis **9** (1998), 229–260.

[59] D. Luecking, *A Technique for Characterizing Carleson Measures on Bergman Spaces*, Proceedings of the AMS **87** (1983), no. 4, 656–660.

[60] W. Lusky, *On weighted spaces of harmonic and holomorphic functions*, J. London Math. Soc. **51** (1995), no. 2, 309–320.

[61] W. Lusky, *On the isomorphism classes of weighted spaces of harmonic and holomorphic functions*, Studia Math. **175** (2006), no. 1, 19–45.

[62] R. Lyons, *Random walks and percolation on trees*, Ann. Probab. **18** (1990), no. 3 931–958.

[63] R. Lyons, *Random walks, capacity and percolation on trees*, Ann. Probab. **20** (1992), no. 4, 2043–2088.

[64] R. Lyons, Y. Peres *Probability on trees and networks*, Cambridge, (2016).

[65] Yu. Lyubarskii, E. Malinnikova, *Radial oscillation of harmonic functions in the Korenblum class*, Bull. London Math. Soc. **44** (2012), no. 1, 68–84.

[66] N.G. Makarov, *Probability methods in the theory of conformal mapping*, Algebra i Analiz **1** (1989), no. 1, 3–59, (Engl. transl. Leningrad Math. J., **1**, (1990)).

[67] V.G. Maz'ya, *On the theory of n-dimensional Schrödinger operator*, Izv. Akad. Nauk SSSR, Ser. Mat. **28** (1964), 1145–1172.

[68] V.G. Maz'ya, *Imbedding Theorems and their Applications*, Baku Sympos. (1966) Nauka, (1970), 142–159 (in Russian).

[69] Y. Meyer, *Wavelets and operators*, Cambridge studies in advanced mathematics, **37**, Cambridge University Press, 1992.

[70] P. Mozolyako, G. Psaromiligkos, A. Volberg, P. Zorin-Kranich, *Improved surrogate bi-parameter maximum principle*, arXiv:2101.01094v2, (2021), 1–15.

[71] P.F.X. Müller, K. Riegler, *Radial variation of Bloch functions on the unit ball of* $\mathbb{R}^d$, Ark. Mat. 58(1), (2020) 161–178.

[72] C. Muscalu, J. Pipher, T. Tao, C. Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), 269–296.

[73] C. Muscalu, J. Pipher, T. Tao, C. Thiele, *Multi-parameter paraproducts*, Rev. Mat. Iberoamericana **22** (2006), no. 3, 963–976.

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

[75] F. Nazarov, S. Treil, A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583–597.

[76] N.K. Nikolski, *Selected Problems of weighted approximation and spectral analysis*, Proc. of the Steklov Institute of Math, Vol. **120**, Amer. Math. Soc., Providence, 1976.

[77] M. D. O'Neill, *Vertical variation of harmonic functions in upper half-spaces*, Colloq. Math. **87** (2001), 1–12.

[78] M. Paolo, P. M. Soardi *Potential Theory on Infinite Networks*, Springer, 1994.

[79] R. Pemantle, Y. Peres, *Galton-Watson trees with the same mean have the same polar sets*, Ann. Probab. **23** (1995), 1102–1124.

[80] J. Pipher, *Journé's covering lemma and its extension to higher dimensions*, Duke Journal of Math **53** (1986), no. 3, 683–690.

[81] I.I. Privalov, P.I. Kuznetsov, *On boundary problems and various classes of harmonic and subharmonic functions on an arbitrary domain*, Mat. Sb. **6(48)** 1939, no. 3, 345–375 (in Russian).

[82] P.J. Rippon *A boundary estimate for harmonic functions*, Mathematical Proceedings of the Cambridge Philosophical Society **91** (1982), 79–90.

[83] W. Rudin *The radial variation of analytic functions*, Duke Math. J. **22** (1955), 235–242.

[84] K. Samotij *A representation theorem for harmonic functions in the ball in $\mathbb{R}^n$*, Ann. Acad. Sci. Fenn. **11** (1986), no. 1, 29–37.

[85] K. Samotij *A critical growth rate for harmonic and subharmonic functions in the open ball in $\mathbb{R}^n$*, Colloq. Math. **52** (1987), no. 1, 145–158.

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

[87] K. Seip *Interpolation and sampling in small Bergman spaces*, Collect. Math. **64** (2013), 61–72.

[88] A. L. Shields, D. L. Williams *Bounded projections, duality, and multipliers in spaces of analytic functions*, Trans. Amer. Math. Soc. **162** (1971), 287–302.

[89] A. L. Shields, D. L. Williams *Bounded projections, duality, and multipliers in spaces of harmonic functions*, J. Reine Angew. **299/300** (1978), 256–279.

[90] D.A. Stegenga *Multipliers of the Dirichlet Space*, Illinois J. Math. **24** (1980), no. 1, 113–139.

[91] V. Stepanov, E. Ushakova, *On weighted Hardy inequality with two-dimensional rectangular operator – extension of the E. Saywer theorem*, Math. Ineq. & Appl. **24** (2021), no. 3, 617–634.

[92] E. J. Straube, *Harmonic and analytic functions admitting a distribution boundary value*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **11** (1984), no. 4, 559-591.

[93] Terence Tao, *Dyadic product H^1, BMO, and Carleson's counterexample*, unpublished note, (1999), available at http://www.math.ucla.edu/ tao/preprints/Expository/product.dvi.

[94] I. E. Verbitsky, *Embedding and multiplier theorems for discrete Littlewood-Paley spaces.* Pacific J. Math. **176** (1996), no. 2, 529–556.

[95] A. Volberg, *Non-symmetry for Fourier transform of positive functions of two variables*, preprint, Dec. 2019, 2 pages.

[96] A. Wedestig, *Weighted inequalities for the Sawyer two-dimensional Hardy operator and its limiting geometric mean operator*, J. Inequal. Appl. **4** (2005), 387–394.

[97] W. Woess, *Random walks on infinite graphs and groups*, Cambridge, 2000.

Papers included in the thesis.

[98] P. Mozolyako, *On the definition of the B-points of Borel charge on the real line*, Zap. Nauch Sem. POMI **389** (2011), 191–205 (English transl. in J. Math. Sci. **182** (2012), no. 5, 690-698).

[99] K.S. Eikrem, E. Malinnikova, P. Mozolyako, *Wavelet characterization of growth spaces of harmonic functions*, Journal d'Analyze Mathématique **122** (2014), no. 1, 87–111.

[100] A. Logunov, E. Malinnikova, P. Mozolyako, *On a theorem of Cartwright in higher dimensions*, Journal of London Mathematical Society **93** (2016), no. 1, 65–82.

[101] V.P. Havin, P.A. Mozolyako, *Boundedness of variation of a positive harmonic function along the normals to the boundary*, Algebra and Analysis **28** (2016), no. 3, 67–110 (English transl. in St. Petersburg Mathematical Journal, **28** (2017) no. 3, 345–375).

[102] N. Arcozzi, P. Mozolyako, K.-M. Perfekt, S. Richter, G. Sarfatti, *Some Hilbert Spaces related with the the Dirichlet space*, Concrete Operators **3** (2016), 94–101.

[103] P. Mozolyako, *Boundary oscillations of harmonic functions in Lipschitz domains*, Collectanea Mathematica **68** (2017), no. 3, 359–376.

[104] N. Arcozzi, P. Mozolyako, K.-M. Perfekt, *Some properties related to trace inequalities for the multi-parameter Hardy operators on poly-trees*, Anal. Math. Phys. **9**, (2019), 937–954.

[105] N. Arcozzi, I. Holmes, P. Mozolyako, A. Volberg, *Bi-parameter embedding and measures with restriction energy condition*, Math. Ann. **377** (2020), 643–674.

[106] P. Mozolyako, G. Psaromiligkos, A. Volberg, *Counterexamples for multi-parameter weighted paraproducts*, Comptes Rendus Mathématique **358** (2020), no. 5, 529–534.

[107] P. Mozolyako, G. Psaromiligkos, A. Volberg, P. Zorin-Kranich, *Combinatorial property of all positive measures in dimensions 2 and 3*, Comptes Rendus Mathématique, **358** (2020), no. 6, 721–725.

[108] N. Arcozzi, I. Holmes, P. Mozolyako, A. Volberg, *Bellman function sitting on a tree*, Int. Math. Res. Notes **2021** (2021), no. 16, 12037–12053.

[109] P. Mozolyako, A. Nicolau, *Oscillation of Functions in the Hölder class*, Potential Analysis **55** (2021), 53–74.

[110] P. Mozolyako, A. Volberg, *Essential differences of potential theories on a tree and on a bi-tree*, Comptes Rendus Mathématique **360** (2022), 1039–1048.

[111] P. Mozolyako, *B-points of the Cantor-type set*, Zap. Nauch Sem. POMI **512** (2022), 148–172 (in Russian).

[112] P. Mozolyako, G. Psaromiligkos, A. Volberg, P. Zorin-Kranich, *Carleson embedding on the tri-tree and on the tri-disc*, Revista Matemática Iberoamericana **38** (2022), 2069–2116.