Association for Women in Mathematics Series

María Cristina Pereyra
Stefania Marcantognini
Alexander M. Stokolos
Wilfredo Urbina *Editors*

Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)

Celebrating Cora Sadosky's Life

AWM
ASSOCIATION FOR WOMEN IN MATHEMATICS

Springer

Association for Women in Mathematics Series

Volume 5

More information about this series at http://www.springer.com/series/13764

Association for Women in Mathematics Series

Focusing on the groundbreaking work of women in mathematics past, present, and future, Springer's Association for Women in Mathematics Series presents the latest research and proceedings of conferences worldwide organized by the Association for Women in Mathematics (AWM). All works are peer-reviewed to meet the highest standards of scientific literature, while presenting topics at the cutting edge of pure and applied mathematics. Since its inception in 1971, The Association for Women in Mathematics has been a non-profit organization designed to help encourage women and girls to study and pursue active careers in mathematics and the mathematical sciences and to promote equal opportunity and equal treatment of women and girls in the mathematical sciences. Currently, the organization represents more than 3000 members and 200 institutions constituting a broad spectrum of the mathematical community, in the United States and around the world.

María Cristina Pereyra • Stefania Marcantognini
Alexander M. Stokolos • Wilfredo Urbina
Editors

Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)

Celebrating Cora Sadosky's Life

Springer

Editors
María Cristina Pereyra
Department of Mathematics and Statistics
University of New Mexico
Albuquerque, NM, USA

Stefania Marcantognini
Department of Mathematics
Venezuelan Institute for Scientific Research
Caracas, Venezuela

Alexander M. Stokolos
Department of Mathematical Sciences
Georgia Southern University
Statesboro, GA, USA

Wilfredo Urbina
Department of Mathematical and Actuarial Sciences
Roosevelt University Chicago
Chicago, IL, USA

ISSN 2364-5733 ISSN 2364-5741 (electronic)
Association for Women in Mathematics Series
ISBN 978-3-319-51591-5 ISBN 978-3-319-51593-9 (eBook)
DOI 10.1007/978-3-319-51593-9

Library of Congress Control Number: 2016946423

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Photograph by Margaret Randall, February 2, 2004

Preface for Volume 2

On April 4, 2014, we celebrated Cora Sadosky's life with an *afternoon in her honor*, preceded by the 13th New Mexico Analysis Seminar[1] on April 3–4, 2014, and followed by the Western Sectional Meeting of the AMS on April 5–6, 2014, all held in Albuquerque, New Mexico, USA. It was a mathematical feast, gathering more than a hundred analysts – fledgling, junior and senior – from all over the USA and the world such as Canada, India, Mexico, Sweden, the UK, South Korea, Brazil, Israel, Hungary, Finland, Australia, Venezuela, and Spain, to remember her outspokenness, her uncompromising ways, her sharp sense of humor, her erudition, and above all her profound love for mathematics.

Many speakers talked about how their mathematical lives were influenced by Cora's magnetic personality and her mentoring early in their careers and as they grew into independent mathematicians. Particularly felt was her influence among young Argentinian and Venezuelan mathematicians. Rodolfo Torres, in a splendid lecture about Cora and her mathematics, transported us through the years from Buenos Aires to Chicago back to Buenos Aires, from Caracas to the USA back to Buenos Aires, and from Washington D.C. to California. He reminded us of Cora always standing up for human rights, Cora president of the Association for Women in Mathematics (AWM), and Cora always encouraging and fighting for what she thought was right.

[1]The 13th New Mexico Analysis Seminar and *An Afternoon in Honor of Cora Sadosky* were sponsored by National Science Foundation (NSF) Grant DMS-140042, the Simons Foundation, and the Efroymson Foundation, and the events were done *in cooperation* with the Association for Women in Mathematics (AWM). See the conferences websites:

www.math.unm.edu/conferences/13thAnalysis

people.math.umass.edu/~nahmod/CoraSadosky.html

An Afternoon in Honor of Cora Sadosky was organized by Andrea Nahmod, Cristina Pereyra, and Wilfredo Urbina. The 13th New Mexico Analysis Seminar organizers were Matt Blair, Cristina Pereyra, Anna Skripka, and Maxim Zinchenko from the University of New Mexico and Nick Michalowski from New Mexico State University.

Cora was born in Buenos Aires, Argentina, on May 23, 1940, and died on December 3, 2010, in Long Beach, CA. Cora got her PhD in 1965 at the University of Chicago under the supervision of both Alberto Calderón and Anthoni Zygmund, the grandparents of the now-known Calderón-Zygmund School. Shortly after her return from Chicago, she married Daniel J. Goldstein, her lifelong companion who sadly passed away on March 13, 2014, a few weeks before the Albuquerque gathering. Daniel and Cora had a daughter, Cora Sol, who is now a political science professor at California State University in Long Beach, and a granddaughter, Sasha Malena, who brightened their last years. During her life, Cora wrote more than 50 research papers, a graduate textbook (*Interpolation of Operators and Singular Integrals: An Introduction to Harmonic Analysis*, Marcel Dekker 1979), and she edited two volumes: one celebrating Mischa Cotlar's 70th birthday (*Analysis and Partial Differential Equations: A Collection of Papers Dedicated to Mischa Cotlar*, CRC Press, 1989) and one celebrating Alberto Calderón's 75th birthday (*Harmonic Analysis and Partial Differential Equations: Essays in Honor of Alberto Calderón*, edited with M. Christ and C. Kenig, The University of Chicago Press, 1999). In the first volume, we have included a list as complete as possible of her scholarly work. Notable are her contributions to harmonic analysis and operator theory, in particular her lifelong and very fruitful collaboration with Mischa Cotlar.

When news of Cora's passing spread like wildfire in December 2010, many people were struck. The mathematical community quickly reacted. The AWM organized an impromptu memorial at the 2011 Joint Mathematical Meeting (JMM), as reported by Jill Pipher, at the time AWM president:

> Many people wrote to express their sadness and to send remembrances. The AWM business meeting on Thursday, January 6 at the 2011 JMM was largely devoted to a remembrance of Cora.

This appeared in the March-April issue of the AWM Newsletter[2] which was entirely dedicated to the memory of Cora Sadosky.

An obituary by Allyn Jackson for Cora Sadosky appeared in Notices of the American Mathematical Society in April 2011.[3]

In June 2011, Cathy O'Neal wrote in her blog mathbabe[4] a beautiful remembrance for Cora:

> [...] Cora, whom I met when I was 21, was the person that made me realize there is a community of women mathematicians, and that I was also welcome to that world. [...] And I felt honored to have met Cora, whose obvious passion for mathematics was absolutely awe-inspiring. She was the person who first explained to me that, as women mathematicians, we will keep growing, keep writing, and keep getting better at math as we grow older [...]. When I googled her this morning, I found out she'd died about 6 months ago. You can read

---

[2] *President's Report*, AWM Newsletter, Vol. 41, No. 2, March-April 2011, p. 1. This issue was dedicated to the memory of Cora Sadosky and it was partially reproduced in Volume 1.

[3] Notices AMS, Vol. 58, Number 4, April 2011, pp. 613–614.

[4] http://mathbabe.org/2011/06/29/cora-sadosky/

> about her difficult and inspiring mathematical career in this biography.[5] It made me cry and made me think about how much the world needs role models like Cora.

In 2013, the Association for Women in Mathematics established the biennial AWM-Sadosky Prize in Analysis,[6] to be awarded every other year starting in 2014. The purpose of the award is to highlight exceptional research in analysis by a woman early in her career. Svitlana Mayboroda was the first recipient of the AWM-Sadosky Research Prize in Analysis awarded in January 2014. Mayboroda contributed a survey paper joint with Ariel Barton to the first of this series of two volumes. As the first volume went into press, the second recipient of the award, the 2016 AWM-Sadosky Prize, was announced: Daniela de Silva, from Columbia University. The award was presented to her in the January 2016 Joint Mathematical Meeting.

In 2015, Kristin Lauter, president of the AWM, started her report in the May-June issue of the AWM Newsletter,[7] with a couple of paragraphs remembering Cora:

> I remember very clearly the day I met Cora Sadosky at an AWM event shortly after I got my PhD, and, knowing very little about me, she said unabashedly that she didn't see any reason that I should not be a professor at Harvard someday. I remember being shocked by this idea, and pleased that anyone would express such confidence in my potential, and impressed at the audacity of her ideas and confidence of her convictions.
>
> Now I know how she felt: when I see the incredibly talented and passionate young female researchers in my field of mathematics, I think to myself that there is no reason on this earth that some of them should not be professors at Harvard. But we are not there yet ... and there still remain many barriers to the advancement and equal treatment of women in our profession and much work to be done.

In these two volumes, friends, colleagues, and/or mentees have contributed research papers, surveys, and/or short remembrances about Cora. The remembrances were sometimes weaved into the article submitted (either at the beginning or the end), and we have respected the format each author chose. Many of the authors gave talks in *the 13th New Mexico Analysis Seminar*, in *An Afternoon in Honor of Cora Sadosky*, and/or in the Special Sessions of the AMS; others could not attend these events but did not think twice when given the opportunity to contribute to this homage.

The mathematical contributions naturally align with Cora's mathematical interests: harmonic analysis and PDEs, weighted norm inequalities, Banach spaces and BMO, operator theory, complex analysis, and classical Fourier theory.

Volume 1 contains articles about Cora and her mathematics and mentorship, remembrances by colleagues and friends, her bibliography according to MathSciNet, and survey and research articles on harmonic analysis and partial differential equations, BMO, Banach and metric spaces, and complex and classical Fourier analysis.

---

[5] Biographies of Women in Mathematics: Cora Sadosky http://www.agnesscott.edu/lriddle/women/corasadosky.htm

[6] More details in the AWM-Sadosky Research Prize in Analysis webpage: https://sites.google.com/site/awmmath/programs/sadosky-prize.

[7] President's Report. AWM Newsletter, Vol. 45, No. 3, May-June, p. 1 (2015).

The year 2014 saw the resolution of the two-weight problem for the Hilbert transform à la Muckenhoupt by Michael Lacey, Erik Sawyer, Chun-Yen Shen, and Ignacio Uriarte-Tuero, a problem that had been open for 40 years. This problem was solved à la Helson-Szegö by Cora Sadosky and Mischa Cotlar in the early 1980s using complex analysis and operator theory methods. In the last 15 years, a number of techniques have been developed and refined to yield this result, including stopping time arguments, Bellman functions, Lerner's median approach, and bumped approach.

Volume 2 contains more remembrances and survey and research articles on weighted norm inequalities, operator theory, complex analysis, dynamical systems, and dyadic harmonic analysis. The articles illustrate surprising connections to Tauberian functions, number theory, and wavelet systems. A survey of the two-weight problem for the Hilbert transform by Michael Lacey is featured.

Before describing in detail the contents of the second volume in this series, we would like to end with some words by Nikolai Nikolski[8] regarding Cora Sadosky's joint work with Mischa Cotlar on the two-weight problem:

> My next impression on Mischa's mathematics is dated about 10 years later when his great series of papers with Cora Sadosky on Generalized Toeplitz Kernels (GTK) started to appear. On the age when all people involved in "weighted analysis" were excited with the Muckenhoupt-type approach (which is efficient for real variable applications), the Cotlar-Sadosky's idea to develop Helson-Szegö classical techniques was revolutionary. They immediately obtained important applications of the GTK theory in a variety of domains where complex analysis language is more appropriate than the real analysis one (scattering theory, Hankel and Toeplitz operators, dilation theory…, but also singular integrals for so important problems as the famous two-weighted estimates). This Cotlar-Sadosky series appeared almost simultaneously with the well-known Krein's school achievements (Adamyan-Arov-Krein) and the Lax-Phillips approach to scattering theory, but the GTK theory showed several advantages (as, for example, an important - and growing with time! - efficiency in several complex variables.)

Contents of Volume 2

We now describe in more detail the contents of the second volume. Volume 2 includes remembrances, photos, two survey articles, and research articles by an array of mathematicians representing themes at the heart of Cora's mathematical interests: weighted inequalities, complex analysis, and operator theory.

In the chapter "Remembering Corita", author and poet Margaret Randall, longtime friend of Corita and her parents, Manuel and Cora, shares her memories of the family, and in chapter "Remembering Cora" Neil Hindman, Howard University Cora's former colleague, shares a few memories.

[8]This is a paragraph in a remembrance for Mischa Cotlar that can be found at www.math.unm.edu/conferences/10thAnalysis/resources/cotlar/nikolski.pdf.

In the chapter "The Two-Weight Inequality for the Hilbert Transform: A Primer," Michael Lacey surveys the resolution of the two-weight problem for the Hilbert transform. This is a valuable and insightful survey on the recent advances on the two-weight inequality for the Hilbert transform, where the author has been a leading contributor. With the fast development of the field in the last few years, hundreds of pages have been published. While much of that has quickly become outdated with the arrival of more powerful and more efficient approaches, there are still important parts in the non-latest papers that have not been redone or surpassed by the newest developments. In this survey article, there is a detailed presentation of the unconditional characterization of the two-weight inequality by Lacey, Sawyer, Shen, and Uriarte-Tuero, incorporating at the core a refinement by Hytönen. An earlier conditional result of Nazarov, Treil, and Volberg, under a "pivotal" condition, is reworked in the style of the more recent papers for the sake of both simplification and easier comparison. Some important counterexamples and a thorough discussion of motivation and applications are presented. Michael Lacey reflects in his article about Cora Sadosky and how his research was strongly influenced by her passion and interests, he gave a talk titled *Cora Sadosky Influence on my Work* in an AMS Special Session on "Harmonic Analysis and Operator Theory (In Memory of Cora Sadosky)" co-organized by two of the editors of this volume, Stokolos and Urbina, in Albuquerque on April 2014.

In the chapter "Singular Integrals, Rank-One Perturbations, and Clark Model in General Situation," Constance Liaw and Sergei Treil present a survey and discuss generalizations of the Clark model to the case of non-singular measures and applications to the study of rank-one perturbations for unitary and self-adjoint operators. This survey summarizes several well-known papers in that direction written previously by the authors and gives some new ideas on the construction of a similar model for dissipative operators. Rank-one perturbations play an important role in operator theory and mathematical physics. One of the principal attractions of rank-one perturbations is that for such operators almost everything can be explicitly computed, and then advanced techniques of harmonic analysis, like the study of fine properties of Cauchy-type integrals or advanced theory of singular integral operators, can be applied. These lecture notes give an account of the mini-course delivered by the authors in the 13th New Mexico Analysis Seminar in April 2014 on *Perturbations, Two-Weight Estimates, and Clark Model.* Sergei Treil also gave an invited lecture on *Two-Weight Estimates Following Arocena-Cotlar-Sadosky* during "An Afternoon in Honor to Cora Sadosky" held in Albuquerque, NM, in April 2014.

In the chapter "On Two-Weight Estimates for Dyadic Operators," Oleksandra Beznosova, Daewon Chung, Jean Moraes, and María Cristina Pereyra discuss quantitative two-weight estimates for dyadic operators in a nonhomogeneous setting. They review the prior known estimates for the maximal dyadic function, dyadic square function, martingale transform, and the dyadic paraproduct. They compare their results for the dyadic paraproduct to results of Holmes, Lacey, and Wick on a homogeneous setting where the two weights are assumed to be in the Muckenhoupt A_p class and the Bloom *BMO property* is necessary and sufficient for boundedness of the dyadic paraproduct and its adjoint. In this paper, the weights are not necessarily doubling, they satisfy a joint A_2 condition, and the dyadic square function is assumed to be two-weight bounded.

In the chapter “Potential Operators with Mixed Homogeneity,” Calixto Calderón and Wilfredo Urbina present a *fitting tribute to Corita* (in the reviewer’s words). In 1966 Cora Sadosky discussed a quasi-homogeneous version of Sobolev’s immersion theorem. Later the first author and T. Kwembe proved a similar result for potential operators with kernels having mixed homogeneity very much in the spirit of Sadosky’s result. The aim of this paper is to extend Calderón-Kwembe’s theorem in two directions: first by establishing a corresponding result in terms of mixed norms in the Benedek-Panzone’s sense and second by obtaining results for the case of unbounded characteristics.

In the chapter “Elementary Proofs of One-Weight Norm Inequalities for Fractional Integral Operators and Commutators,” David Cruz-Uribe presents new proofs of some recent results concerning weighted inequalities for the fractional integral operator and its commutator with *BMO* functions. The author reduces the problem to obtaining estimates for a sparse fractional operator which majorizes the fractional integral operators. As pointed out by the author, the advantage of this approach is its simplicity: it avoids extrapolation, good-λ inequalities, and comparisons to the fractional maximal operator; however the proofs do not give sharp dependence on the $A_{p,q}$ characteristic of the weights; nevertheless this dependence is carefully tracked. This is a nice summary and introduction into the modern dyadic techniques in weighted inequalities. This chapter should be read in conjunction with the chapter about two-weight inequalities for fractional integral operators by Sawyer, Shen, and Uriarte-Tuero in this volume. Cruz-Uribe ends with a very touching personal story about Cora Sadosky.

In the chapter “Finding Cycles in Nonlinear Autonomous Discrete Dynamical Systems,” Dmitriy Dmitrishin, Anna Khamitova, Alex Stokolos, and Michai Tohaneanu provide an exposition of their recent results concerning cycle localization and stabilization in nonlinear dynamical systems. Both the general theory and numerical applications to well-known dynamical systems are presented. Following on the footsteps of the pioneering work of Grebogi, Ott, Pyragas, York, et al., the authors consider associating to a given map f on $\mathbb{R}^n$ and for each N another map F_N on $(\mathbb{R}^n)^N$ that will stabilize certain unstable orbits common with the initial map f for appropriately chosen parameters. The authors show for a fixed N not all unstable orbits may be stabilized by the given control and indicate explicit bounds on the multiplier of a T-cycle of f that enable control of the above form to stabilize the orbit. This paper will be of primary interest to those with prior experience in dynamical systems; fortunately the authors provide a very suitable list of references for those who wish to study the prerequisites necessary for a good understanding of this paper. Alex Stokolos gave an invited talk on *Complex and Harmonic Analysis in Nonlinear Dynamics* during “An Afternoon in Honor of Cora Sadosky” held in Albuquerque, NM, in April 2014. Stokolos ends their article with a heartfelt remembrance of Cora Sadosky.

In the chapter “Smooth Analytic Functions and Model Subspaces,” Konstantin Dyakonov surveys the canonical Riesz-Nevanlinna factorization in various classes of analytic functions on the disk that are smooth up to its boundary and model subspaces (i.e., invariant subspaces of the backward shift) in the Hardy spaces H^p and in *BMOA*. It is the interrelationship and a peculiar cross-fertilization between these two topics that the author wishes to highlight. This article deals with the canonical factorization in classes of "smooth" analytic functions on the unit disk on one hand and the so-called model subspaces on the other hand. The author gives a (almost) self-contained presentation (with proofs) of several deep and beautiful results which are related in a natural way to the theory of Hankel and Toeplitz operators, one of Cora Sadosky's favorite themes.

In the chapter "Rational Inner Functions on a Square-Matrix Polyball," Annatoli Grinshpam, Dmitry Kaliuzhnyi-Verbovetskyi, Victor Vinnikov, and Hugo Woerdeman establish among other results the existence of a finite-dimensional unitary realization for every matrix-valued rational inner function from the Schur-Agler class on a unit square-matrix polyball. It is well known that for polydisks, the Schur-Agler class and the Schur class only coincide in dimensions 1 and 2. Furthermore, in the other cases, though one might naively expect otherwise, not all rational inner functions are in the Schur-Agler class. As a consequence, it is of interest to characterize those which are. The authors describe those matrix-valued rational inner functions in the Schur-Agler class over square-matrix polyballs, which includes the polydisks. This is done in terms of the transfer function representation, which, as might be expected, coincides with the unitary colligation being finite dimensional. The last section of the paper attempts in the scalar case to more directly describe the polynomials appearing the Korányi and Vagi description when dealing with Schur-Agler class rational inner functions. This connects with much interesting recent work, including by the authors, on determinantal representations. Finally, a number of thought-provoking open questions are posed.

In the chapter "A Note on Local Hölder Continuity of Weighted Tauberian Functions," Paul Hagelstein and Ioanis Parissis discuss how Solyanik estimates may be used to establish local Hölder continuity estimates for the Tauberian functions associated to the Hardy-Littlewood and strong maximal operators in the context of Muckenhoupt weights. The Tauberian condition for the geometric maximal operators was introduced by A. Córdoba and R. Fefferman in 1977. In this nice and clearly written article, the authors introduce the Tauberian function as a weighted generalization of the halo function, which is a classical object in the theory of differentiation of integrals, and they establish belonging of the function to the local Hölder classes, with the Hölder exponent proportional to the reciprocal value of the A_∞ norm of the weight. The main tool used is Solyanik estimates – a series of results initiated by A. Solyanik in 1995 and developed further by the authors of the current article. In particular, they proved that Solyanik estimates imply the continuity of halo function for the most important case of density bases. The main result of the article provides the quantitative version of continuity of the Tauberian function for bases of all cubes and all rectangles with sides parallel to the coordinate axis. The novelty of the article is in the remarkable connection of the A_∞ norm of the weights with the smoothness of the Tauberian functions.

In the chapter "Three Observations on Commutators of Singular Integral Operators with *BMO* Functions," Carlos Pérez and Ismael Rivera present interesting observations concerning commutators of singular integral operators with *BMO* functions. Namely, they discuss sharpness of the sub-exponential local decay, sparse domination result for commutators, and the failure of an endpoint estimate motivated by the conjugation method. This paper is very useful not only for the researcher who is carefully studying the commutator but also for the beginners who want to be acquainted with the theory of commutators. Carlos Pérez gave an invited talk on *Optimality of Exponents and Yano's Condition in Weighted Estimates and Endpoint Estimates* during "An Afternoon in Honor of Cora Sadosky" held in Albuquerque, NM, in April 2014.

In the chapter "A Two-Weight Fractional Singular Integral Theorem with Side Conditions, Energy, and k-Energy Dispersed," Erik Sawyer, Chun-Yen Shen, and Ignacio Uriarte-Tuero present a follow-up to a paper of theirs, where the two-weight inequality for fractional singular integral operators is studied under the assumption that the pair of weights does not have common point masses. In this chapter, the authors allow for common point masses. Under appropriate Muckenhoupt (joint $\mathcal{A}_\alpha^2$ and punctured $\mathcal{A}_\alpha^2$ conditions) and α-quasi-energy side conditions, the authors show that a fractional singular integral operator, T_α, is bounded from one weighted space to another if certain quasicube testing conditions (involving a globally bi-Lipschitz map $\Omega : \mathbb{R}^n \to \mathbb{R}^n$) hold for T_α and its dual and if the quasiweak boundedness property holds for T_α. Conversely, if T_α is bounded, then the quasitesting conditions hold, and the quasiweak boundedness condition holds. It is unknown whether the quasi-energy conditions are necessary in higher dimensions in general. This is a highly technical paper and the authors do a very good job placing road maps and diagrams and sometimes iterating ideas so the reader does not get lost.

In the chapter "A Partition Function Connected with the Göllnitz-Gordon Identities," Nicolas A. Smoot presents a Rademacher-type formula for the partition of a positive integer into parts of special type, associated to the Göllnitz-Gordon identities. The subject of this paper constitutes a beautiful application of complex analysis to number theory. The proof of the main result follows in the spirit of the Rademacher approach and involves its basic components – generating functions, Rademacher contours, the Hardy-Ramanujan circle method, Bessel functions. The argument is quite involved but very well written and illustrated by nice sketches. The paper is almost self-contained and is accessible to nonexpert researchers and graduate students.

In the chapter "On Toeplitz Operators with Quasi-radial and Pseudo-homogeneous Symbols," Nikolai Vassilevskii explores a new wide class of symbols that generate commutative Banach algebras on each weighted Bergman space on the unit ball in $\mathbb{C}^n$. These symbols are a natural extension of the previously studied quasi-radial quasi-homogeneous symbols and contain them as a very special particular case. The commutative C^*-algebras of Toeplitz operators on Bergman spaces of the unit ball $\mathbb{B}^n$ of $\mathbb{C}^n$ are fairly well understood nowadays. Some examples of commutative Banach (not C^*) algebras of such operators are known. The current paper exhibits a new wide class of functions that generate such algebras on each of the standard weighted Bergman spaces on $\mathbb{B}^n$. The results are of interest and contribute toward the (as yet fairly incomplete) understanding of the commutative subalgebras of Toeplitz operators on the ball in $\mathbb{C}^n$, in dimension larger than one.

In the chapter "A Bump Theorem for Weighted Embedding and Maximal Operator: The Bellman Function Approach," Sasha Volberg gives a simple Bellman function proof of Carlos Pérez's "bump theorem" for the two-weight estimates of maximal operators. The author shows how the Bellman function method proves a certain "discrete" inequality (i.e., a bound on discrete operator), which, in its turn, implies a bound of a certain continuous operator. The main difference from a, by now, classical procedure is that here one needs a Bellman "functional." Since the operator under consideration (maximal function) is, in some sense, easy, the paper is intended to give a clear and not very technical presentation of the method.

In the chapter "The Necessity of A_∞ for Translation and Scale Invariant Almost-Orthogonality," Mike Wilson continues his study of general wavelet systems in the context of weights. For a weight μ in the Muckenhoupt class A_∞, the author has shown previously that for sufficiently nice mother wavelets and arbitrary T-systems, the resulting system is almost orthogonal in $L^2(\mu)$. This paper concerns the converse result. In the reviewer's own words: *The current version of the converse seems now fully satisfactory. It has the nice feature then that if you can obtain the almost orthogonality in $L^2(d\mu)$ for one choice of a wavelet meeting the minimum requirements and all T-systems, then one can conclude that $\mu \in A_\infty$, and then that one has the almost orthogonality in $L^2(\mu)$ for all reasonable wavelets systems and all T-systems, by the other direction of the theorem. This "prove it for one, get it for all" feature, although not uncommon in this general area, is quite pleasing.*

Alex Stokolos, Sergei Treil, and Carlos Pérez were invited speakers to "An Afternoon in Honor of Cora Sadosky." Beznosova, Cruz-Uribe, and Pereyra and Stokolos and Urbina were co-organizers of AMS Special Sessions on "Weighted Norm Inequalities and Related Topics" and on "Harmonic Analysis and Operator Theory (In Memory of Cora Sadosky)," respectively, on April 5–6, 2014, held in Albuquerque, NM. Michael Lacey, Oleksandra Beznosova, Daewon Chung, Jean Moraes, Paul Hagelstein, Dmitry S. Kaliuzhnyi-Verbovetskyi, Constance Liaw, Nikolai Vassilevskii, and Mike Wilson gave talks in the AMS meeting in Albuquerque in April and honored there as they are doing here the life and work of Cora Sadosky. Vinnikov is Cora Sadosky's coauthor; in fact they were both co-authors of the abstract for the AMS talk delivered by Kaliuzhnyi-Verbovetskyi. Other authors could not make it to the conference but were more than happy to contribute to this volume.

Acknowledgments

These volumes would not have been possible without the contributions from all the authors. We are grateful for the time and care you placed into crafting your manuscripts. Thank you!

All the articles were peer reviewed, and we are indebted to our dedicated referees, who timely and often enthusiastically pitched in to help make these volumes a reality. We used your well-placed comments and words to describe in the preface the articles in these volumes. Thank you!

We would like to thank Cora Sol Goldstein, Cora's daughter, who blessed the project and gave us a selection of beautiful photos for us to choose and use. When she first heard about the volumes, she said, *"My father would had been so happy to know about these books."*

Andrea Nahmod spotted a photo with Cora at the Mathematical Sciences Research Institute (MSRI) at Berkeley, CA, and facilitated communication with Hélène Barcelo (MSRI deputy director), Christine Marshall (MSRI program manager), and David Eisenbud (MSRI director), who kindly gave us permission and sent us a high-resolution copy of the photo which is in display at MSRI that we are reproducing in this second volume. Margaret Randall, writer and longtime friend of Corita and her parents, shared her memories and provided the opening photo for both volumes after a serendipitous encounter in Albuquerque. Thank you!

Our editor at Springer, Jay Popham, was very accommodating and patient, so was editor Marc Strauss who came on board later as the editor of the AWM-Springer Series.

We thank Kristin Lauter, AWM president, for embracing this project and the AWM staff for helping us in the final laps, in particular Anne Leggett Macdonald, AWM Newsletter editor, who provided guidance and moral support.

We cannot help but to think that Cora's spirit was around helping us finish this project. Cora's legacy is strong and will continue inspiring many more mathematicians!

Albuquerque, New Mexico, USA — María Cristina Pereyra
Caracas, Venezuela — Stefania Marcantognini
Statesboro, Georgia, USA — Alexander M. Stokolos
Chicago, Illinois, USA — Wilfredo Urbina

Contents Volume 2

Contents Volume 1

Part I
More Remembrances and photos

Remembering Corita . . .

Margaret Randall

Corita Sadosky—her mother's name was Cora, so we called her Corita—was an Argentinean mathematician who belonged to her country's "lost generation." What that meant was that she came of age during the 1970s, dark years of dictatorship when 30,000 young people were murdered by neo-fascist forces much like those once again threatening us in so many countries today. Some because they were actively working against the dictatorships that had taken control in the Southern Cone of South America. Some because they were simply young, and youth alone was considered a crime. Most for some combination of both.

I first met Corita and her parents, Cora and Manuel, when we had all taken refuge in Cuba. Corita and her family, including her husband Daniel Goldstein, had come to the Caribbean island by way of Venezuela. Eventually they had a daughter, Corasol. My family and I had come from Mexico. Those years saw interwoven webs of exiles moving from country to country, often just a few steps ahead of death threats or worse. Corita and I became instant friends.

In the summer of 1973 it seemed our liberated territories had expanded; Chile, under Salvador Allende was trying to show the world that freedom could come not only by means of armed struggle but also through democratic elections. People throughout the Americas were hopeful. In Peru the Velasco Alvarado government was developing progressive initiatives, especially for the country's large Quechua population. Corita was invited to head a program there, to improve women's lives. The program involved an oral history project, in which they would ask women themselves what they needed.

Looking back, I can't help but smile. Cora had been invited because she was a woman, without consideration of her life as a mathematician. She promptly told

M. Randall (✉)
Albuquerque, NM, USA
e-mail: mrandall36@gmail.com

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_1

the Peruvian organization issuing the invitation that she was not the best person for the job. She recommended me—already a feminist and oral historian—and that is how I happened to travel to Peru in the fall of 1973, just after Pinochet's bloody coup toppled the Chilean experiment. Hundreds of Chileans were making their way across their country's northern border into Peru. I would spend 3 months working for the United Nations' International Labor Office in what turned out to be a fascinating project. Unfortunately, it too would be doomed.

From then on, Corita's and my friendship grew. We spent time together in Cuba and later in Washington, D.C., where she eventually ended up, teaching at Howard University. Our families, too, remained interwoven—in that strong fabric made up of Latin American revolutionaries forced to migrate from country to country. I remember Cora's sudden death, and a visit to Manuel and his new wife in Buenos Aires. My son, Gregory, at the time vice president of the University of the Republic in Uruguay, bestowed an honorary doctorate on Manuel just months before the latter died. Corita made a life for herself in the United States. Despite having been displaced from the country of her birth, she adapted to Washington, taught several generations of upcoming mathematicians, and continued with the research that had become her life's passion. My last memories of Corita are from a visit Barbara and I made to her in our nation's capital.

When Corita died, it was a shock. We hadn't been in touch for a few years, but I had no doubt she was there—in that "there" we shared for so many years although its physical location might vary. Like many, she died much too young. But, like some, she had had the strength to overcome a violent displacement that has defeated so many.

It is wonderful that she is being honored in this beautiful book. She was a good friend. But she was also an example of the humanity, brilliance and fortitude needed again today to keep on fighting the good fight against the powers of disrespect and arrogance so obvious on today's political scene.

October 2016

Margaret Randall (New York, 1936) is a poet, essayist, oral historian, translator, photographer and social activist. She lived in Latin America for 23 years (in Mexico, Cuba, and Nicaragua). Randall's most recent titles include "She becomes time", "Che on my mind", "Haydée Santamaría, Cuban revolutionary: she led by transgression", and "Exporting revolution: Cuba's global solidarity". Randall has also devoted herself to translation, producing "Only the road/Solo el camino", an anthology of eight decades of Cuban poetry, among other books. She lives in New Mexico with her partner (now wife) of more than 30 years, the painter Barbara Byers, and travels extensively to read, lecture and teach.

Remembering Cora

Neil Hindman

I spent almost three decades in an office adjacent to Cora's office. We were in a building separate from the department office, where most of the Mathematics Department faculty had their offices. We were fortunate, since we had individual offices, while those in the other building had two or three people to an office. We were unfortunate because the university doesn't tend to turn on the heat in our building until mid November. We both had electric space heaters, and when we ran them at the same time, we would blow a circuit breaker.

My main interaction with Cora was the fact that we shared my coffee maker. Every morning when she came in to the office, she would come and get a cup of coffee—we both drank it black. And she would occasionally bring in a can of ground coffee.

Beyond that, the main thing I remember is that Cora did not like to carry the heavy calculus text book to class. So she would tear out the relevant pages for the day's lesson and just take those to class.

N. Hindman (✉)
Howard University, Washington, DC, USA
e-mail: nhindman@howard.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_2

Dinner with MSRI's Human Resources Advisory Committee, November 10, 2004, at the house of Director David Eisenbud. Photo by Monika Eisenbud

1996 University of Chicago Conference in honor of Alberto Calderón's 75th Birthday: Front row, seated (left to right): M. Christ, C. Sadosky, A.P. Calderon, M.A. Muschietti. First row, standing (left to right): C.E. Kenig, J. Alvarez Alonso, C. Gutierrez, E. Berkson, J. Neuwirth. Second row, standing (left to right): A. Torchinsky, J. Polking, S. Vagi, R.R. Reitano, A.E. Gatto, R. Seeley.

Photo courtesy of Cora Sol Goldstein, photographer unknown

Part II
Survey Articles

The Two Weight Inequality for the Hilbert Transform: A Primer

Michael T. Lacey

In memory of my father, H. Elton Lacey

Abstract Given a pair of weights w, σ, the two weight inequality for the Hilbert transform is of the form $\|H(\sigma f)\|_{L^2(w)} \lesssim \|f\|_{L^2(\sigma)}$. Recent work of Lacey-Sawyer-Shen-Uriarte-Tuero and Lacey have established a conjecture of Nazarov-Treil-Volberg, giving a real-variable characterization of which pairs of weights this inequality holds, provided the pair of weights do not share a common point mass. In this paper, the characterization is proved, collecting details from across several papers; counterexamples are detailed; and areas of application are indicated.

Introduction

By a *weight* we mean a non-negative Borel locally finite measure, typically on $\mathbb{R}$. We consider the *two weight inequality for the Hilbert transform* for a pair of weights w, σ on $\mathbb{R}$:

$$\|H(f \cdot \sigma)\|_{L^2(w)} \leq \mathcal{N}\|f\|_{L^2(\sigma)}. \tag{1}$$

Here, $\mathcal{N}$ denotes the best constant in the inequality. And $H\nu(x)$ is the Hilbert transform of ν

Research supported in part by grant NSF-DMS 0968499, a grant from the Simons Foundation (#229596 to Michael Lacey), and the Australian Research Council through grant ARC-DP120100399. The author benefited from two research programs, first 'Operator Related Function Theory and Time-Frequency Analysis' at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during 2012–2013, and second 'Interactions between Analysis and Geometry' program at IPAM, UCLA, 2013.

M.T. Lacey (✉)
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
e-mail: lacey@math.gatech.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_3

$$H\nu(x) := \int \frac{\nu(dy)}{y-x}. \tag{2}$$

We do not insist on the existence of the principal value, a point addressed in section "Principal Values".

The central question is then a real-variable characterization of the inequality (1). In the special case that the pair of weights σ and w do not share a common point mass, this was supplied in three papers, one of Lacey-Sawyer-Shen-Uriarte-Tuero [23] with the refinement of Hytönen [15], and another of the present author [17], answering a beautiful conjecture of Nazarov-Treil-Volberg [56].

Theorem 1.1 *Define two positive constants $\mathcal{A}_2$ and $\mathcal{T}$ as the best constants in the inequalities below, uniform over intervals I, and with respect to interchanging the roles of σ and w.*

$$\frac{\sigma(I)}{|I|} \cdot P(w\mathbf{1}_{\mathbb{R}\setminus I}, I) \le \mathcal{A}_2, \tag{3}$$

$$\int_I H(\sigma \mathbf{1}_I)^2 \, dw \le \mathcal{T}^2 \sigma(I). \tag{4}$$

There holds $\mathcal{N} \simeq \mathcal{H} := \mathcal{A}_2^{1/2} + \mathcal{T}$.

The first condition is an extension of the Muckenhoupt A_2 condition to a 'half Poisson condition with a hole.' The exact Poisson extension of σ to the upper half-plane is not needed, rather we use the approximation below, which is roughly the Poisson extension evaluated at the center of I, and up into the half-plane the length of I, see Fig. 1.

$$P(\sigma, I) := \int_{\mathbb{R}} \frac{|I|}{(|I| + \mathrm{dist}(x, I))^2} \, \sigma(dx). \tag{5}$$

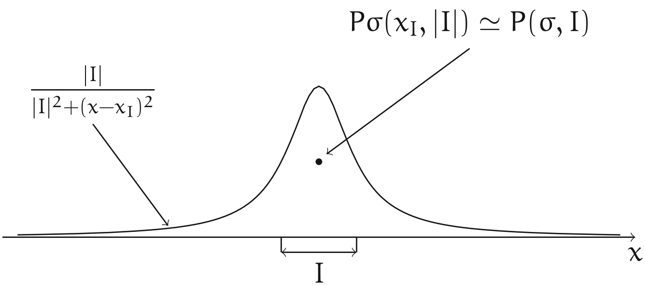


Fig. 1 The value of $P(\sigma, I)$ is approximately the Poisson extension of σ evaluated at point in the upper half-plane given by the center of I, and the length of I

The remaining conditions are referred to as the Sawyer-type *testing conditions*, as Eric Sawyer first introduced these conditions into the two weight setting in his fundamental papers on the maximal function [52], and later the fractional and Poisson integral operators [53]. It is well-known that the A_2 condition (3) is necessary for the two weight inequality, and it is obvious that the testing conditions are necessary. Thus, the substance of the Theorem above concerns the sufficiency of the A_2 and testing inequalities for the norm inequality.

This Theorem is a central result in the non-homogeneous harmonic analysis, as founded in a sequence of influential papers of Nazarov-Treil-Volberg [35–37]. The proof of the theorem is involved, encompassing arguments and points of view that were spread across several papers [17, 21, 23, 38]. Finally, the interest in the two weight inequality is well-motivated by applications to operator theory, model spaces, and spectral theory, themselves spread across additional papers.

The point of this paper is to

(a) state and prove the Theorem, in all detail.
(b) give the proof under the influential *pivotal condition*, which serves to highlight where the difficulties arise in the general case;
(c) collect relevant, explicit, counterexamples;
(d) give complements and extensions of the theorem, and the proof techniques;
(e) and point to areas of applications.

Sections proceed directly towards proofs, but many conclude with some context and discussion. The proof is entirely elementary, assuming only the well known facts about martingale differences.

An Overview of the Proof

The result is an *individual two weight inequality.* It characterizes the boundedness of the Hilbert transform, and no other operator. Therefore, particular properties of this transform must guide the proof. The elementary examples of these are the *monotonicity principle,* Lemma 3.2, valid for all pairs of weights, and then the *energy inequality*, Lemma 3.3, valid under the assumption of interval testing and the A_2 condition. These properties are a last vestige of positivity: The kernel $\frac{1}{y}$ is monotone increasing on $\mathbb{R}\setminus\{0\}$. This feature will deliver to us the *energy inequality*; finding it, and unlocking its secrets is the key to the proof.

The main line of the argument begins with the bilinear form $\langle H_\sigma f, g\rangle_w$. It's decomposition is made to 'regularize' all four quantities in the expression, the two functions f and g, as well as the 'irregularities' of the pair of weights, as expressed by the energy inequality. Only half of the decomposition needs to be specified, due to the self-dual nature of the question, and some of these considerations are familiar to experts in both the $T1$ and the Tb theorems. But the underlying difficulties do not have any classical analog.

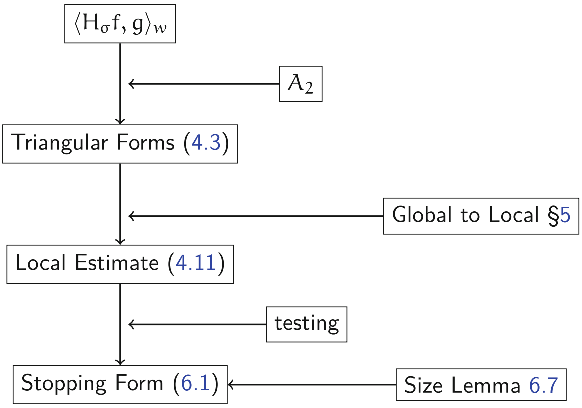


Fig. 2 A schematic tree of the proof of the main theorem

The proof strategy is outlined in Fig. 2. The passage to the 'triangular forms' in Lemma 4.1 is a rather standard step in many $T1$-type theorems. The *Calderón-Zygmund stopping data* defined in section "Global to Local Reduction" is the foundational tool. It (a) controls the values of certain telescoping sums of martingale differences; (b) regularizes the weights, from the point of view of the energy inequality; and (c) allows the use of the *quasi-orthogonality* argument, an important simplification. The triangular forms are of a 'local' and a 'global' form, and have dual forms as well. There are two steps in the analysis, a 'global to local' reduction in section "Global to Local Reduction", and an analysis of the 'stopping form' in the section "The Stopping Form".

The stopping data is essential to the 'global to local reduction' in Theorem 4.4. A simple appeal to the testing condition, allows an application of the monotonicity principle to rephrase the inequality in this Theorem as a certain two-weight inequality for the Poisson integral. In this inequality, the Poisson integral maps functions on $\mathbb{R}$ to those on $\mathbb{R}^2_+$. The weight on $\mathbb{R}$ is, say, σ. The weight on $\mathbb{R}^2_+$ is then derived from w in a specific fashion from the stopping data, and hence depend upon f and the pair of weights. But the Poisson operator is a positive operator, and one has a quite adequate understanding of their two weight inequalities. We directly implement this understanding, without proving any more general result.

The local term is then dominated by the analysis of the *stopping form* (53). This is again a familiar object, to experts in $T1$ theorem, addressed by *ad hoc* off-diagonal estimates, which absolutely do not apply in the current context. Control of the irregularities of the weights is now the main point, complicated by the fact that the stopping form is not intrinsically defined. A notion of 'size' is introduced—it serves as an approximate of the operator norm of the stopping form, and again is most naturally defined in terms of a measure on $\mathbb{R}^2_+$, derived from the two given weights. The *size lemma*, Lemma 6.3, decomposes a stopping form into constituent parts. Those of large size have a simpler form, which allows one to estimate their operator norm by size. What is left has smaller size, and so one can recurse. This argument relies heavily on the Hilbertian structure of the question.

Some readers will have noticed that a very common set of objects, Carleson measures, are not mentioned, and indeed, they do not appear in the proof at all. The wide spread prevalence of Carleson measures in $T1$ theorems can be traced to two facts, first that associated paraproducts operators are the principle obstacle to a simple proof, and second, the paraproduct operators have an essentially canonical form. In this theorem, neither of these facts hold, and so we have abandoned the notions of Carleson measures and paraproducts.

Carleson measures are also used to, indirectly, control the sums of martingale differences. Rather than this, we use the simpler method of stopping data, as described in section "Global to Local Reduction".

The A_2 Theory

The classical case of an A_2 weight corresponds to the case of $w(dx) = w(x)dx$, and $w(x) > 0$ a.e. Moreover, the weight σ also has density given by $\sigma(x) := w(x)^{-1}$. It is assumed that both w and σ are locally integrable, so that they are both weights. See Fig. 3. Note that $w(x) \cdot \sigma(x) \equiv 1$. The Muckenhoupt A_2 condition asserts that this same equality approximately holds, uniformly over location and scale.

$$[w]_{A_2} := \sup_I \frac{w(I)}{|I|} \cdot \frac{\sigma(I)}{|I|} < \infty.$$

These are 'simple' averages. This condition is equivalent to the uniform norm bound on $L^2(w)$ for the class of simple averaging operators

$$f \mapsto \frac{1}{|I|} \int_I f \, dx \cdot \mathbf{1}_I, \qquad I \text{ is an interval.}$$

From this condition flows a rich theory, including the boundedness of all Calderón-Zygmund operators. The classical result of Hunt-Muckenhoupt-Wheeden [11] states

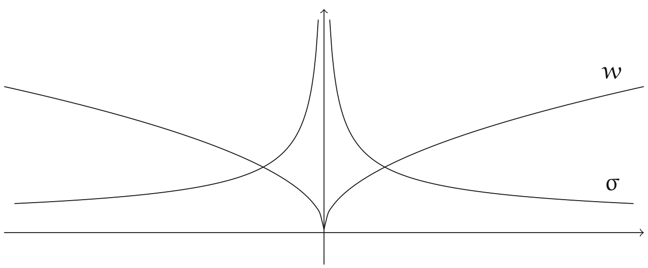


Fig. 3 For $0 < \epsilon < 1$, the function $w(x) = |x|^{1-\epsilon}$ is an A_2 weight. It and the dual weight $\sigma(x) = |x|^{\epsilon-1}$ are graphed above. One can check that $[w]_{A_2} \simeq \epsilon^{-1}$

that w in A_2 if and only if the Hilbert transform maps $L^2(w)$ to $L^2(w)$. By a basic change of variables argument, first noted by Sawyer [52], this is equivalent to H_σ mapping $L^2(\sigma)$ to $L^2(w)$. Stefanie Petermichl [44] quantified the Hunt-Muckenhoupt-Wheeden theorem as follows.

Theorem A *A weight $w \in A_2$ if and only if H is bounded from $L^2(w)$ to $L^2(w)$, and moreover the constant $\mathcal{N}$ in (1) satisfies $\mathcal{N} \simeq [w]_{A_2}$.*

To place this result in the context of our main result, it is classical and easy to see that the Poisson A_2 characteristic satisfies $\mathcal{A}_2 \lesssim [w]_{A_2}^2$. And, using the remarkable Haar shift representation of the Hilbert transform due to Petermichl [43], one can check that the testing condition satisfies $\mathcal{T} \lesssim [w]_{A_2}$. This is what Petermichl's original proof did. All existing proofs of Petermichl's Theorem (see [13, 18, 29]) depend ultimately on known Lebesgue measure estimates for the Hilbert transform, or closely related operators. For instance, [16, 29] use the weak-$L^1(dx)$ bound for sparse shift operators. Estimates of these type are irrelevant for the two weight theorem.

It is perhaps worth emphasizing that the powerful Haar shift technique of Petermichl, even with its impressive extension by Hytönen [13], seems to be of little use in the general two weight problem. There are two obstacles: Firstly, in order to use it, one must essentially have control on a Haar shift operator, independently of how the grid defining the shift is defined. The resulting condition on the pair of weights is more subtle than the two weight inequality for the Hilbert transform. Secondly, one should recover the energy inequality of Lemma 3.3. But, the energy of any fixed Haar shift is zero, and indeed, the two weight inequality for Haar shift operators [39] has just a few difficulties in its proof.

By the A_2 *Theorem*, it is meant the linear in A_2 bound for all Calderón-Zygmund operators. This result, pursued by many, and established by Hytönen [13], has many points of contact with the subject of this note. But, we refer the reader to [14] and references there in for more information, and see [16] for what is arguably the most elementary proof.

In the A_2 theory, it is essential that $w(x) > 0$ a.e. Suppose one relaxes this condition to $w(x)$ is positive on a measurable set $E \subsetneq \mathbb{R}$, and define $\sigma(x)$ to be supported on E, and equal to $w(x)^{-1}$. One can then ask if the Hilbert transform is bounded for this pair of weights, and Theorem 1.1 applies here. This question is an instance of the non-homogeneous A_2 theory advocated by A. Volberg. One can hope that specificity in the way the weights are prescribed could introduce some additional simplifications in the characterization of the two weight inequality in this setting. But, none has yet been found.

The Individual Two Weight Problem

Given an operator T, the *individual L^p two weight inequality for T* is the inequality

$$\|T_\sigma f\|_{L^p(w)} \leq \mathcal{N}_T \|f\|_{L^p(\sigma)}. \tag{6}$$

Here and throughout we use the notation $T_\sigma f := T(\sigma f)$. We understand that T applied to a signed measure $\sigma \cdot f$ should make sense. And, the inequality above is the preferred form of the inequality as duality is expressed in the natural way: The inequality (6) is equivalent to

$$\|T_w^* g\|_{L^{p'}(\sigma)} \le \mathcal{N}_T \|g\|_{L^{p'}(w)}.$$

The question is then to characterize the pairs of weights for which (6) holds.

This specificity of the question is of interest for a few canonical operators, ones for which the corresponding two weight inequality will naturally present itself. The leading examples of this are, for positive operators, the Hardy operator by Muckenhoupt [32], the maximal function, Sawyer's Theorem of 1981 [52] and Sawyer's 1988 theorem for the fractional integrals [53]. It is noteworthy that the two weight inequalities for the Hardy and the Poisson integral are used in the proof of our main theorem, as are various purely dyadic variants of these Theorems.

It is interesting to that this is not only a chronological list, but it also reflects the depth of the results as well. The Hardy operator is easiest, characterized by an 'A_2-type condition,' as recalled in Theorem F. It was Sawyer's insight, however, that the maximal function characterization requires a testing condition. The fractional integrals are harder still. For the sake of comparison, let us state a special case of the result for the fractional integrals in one dimension. Besides Sawyer's results, one should also consult Casscante-Ortega-Verbitsky [6], as well as those of Vuorinen [57]. Both results give a characterization in terms of testing conditions. And, while we state just one case of the general result, one should note that there is no Sobolev condition imposed on the L^p indices.

Theorem B *For two weights w, σ, and $0 < \alpha < 1$, the operator $R_\sigma f(x) := \int f(x-y) \frac{\sigma(dy)}{|y|^\alpha}$ maps $L^2(\sigma)$ to $L^2(w)$ if and only if the testing inequalities below hold.*

$$\int_I R_\sigma(\mathbf{1}_I)^2 \, dw \le \mathcal{T}^2 \sigma(I), \qquad \int_I R_w(\mathbf{1}_I)^2 \, d\sigma \le \mathcal{T}^2 w(I).$$

Moreover the norm of the operator is equivalent to $\mathcal{T}$, the best constant in the inequalities above.

The analysis of the individual two weight inequality for positive operators is much simpler, as is the case of dyadic operators. For certain non-positive dyadic operators, see the result of Nazarov-Treil-Volberg [39], and the much more recent works of Vuorinen [57, 58]. These results have found significant interest, due to the Haar shift operators of Petermichl [43], the remarkable median inequality of Lerner [28] and its extension in [16], and the Hytönen representation theorem [13].

The Hilbert transform is the first non-positive continuous operator for which the individual two weight problem has been solved. And, one would only ever expect that the solution would be of interest (or even possible) for a few canonical choices of operators, such as Hilbert, Cauchy and Riesz transforms. Foundational to the solution for the Hilbert transform is the monotonicity of the kernel. No other canonical choice will satisfy such a simple condition. For a special case of the Cauchy transform [24] one can make progress. But the case of Riesz transforms is much harder [26, 54].

The individual two weight question makes sense for any $1 < p < \infty$, and there are characterizations in this, and other off-diagonal cases for positive operators. For dyadic analogs of singular integrals Vuorinen [57] has shown that these inequalities can be characterized by *quadratic* testing conditions. Also see [27]. The extension of this characterization to the setting of the Hilbert transform is challenging.

The Hilbert Transform

The two weight inequality for the Hilbert transform was addressed as early as 1976 by Muckenhoupt and Wheeden [33].[1] But, it received much wider recognition as an important problem with the 1988 work of Sarason [50]. The latter was part of important sequence of investigations that identified de Branges spaces as an essential tool in operator theory. His question concerning the composition of Toeplitz operators, see section "Sarason's Question on Toeplitz Operators", was raised therein, and advertised again in [51]. This question related the individual two weight problem for the Hilbert transform to a profound question from operator theory.

While not stated in the language of the Hilbert transform, Sarason wrote that it was 'tempting' to conjecture that the full Poisson A_2 condition would be sufficient for the two weight inequality. In an important development, F. Nazarov [34] showed that this was not the case. The two weight problem was seen to be important to Model spaces, namely certain embedding questions for Model spaces can be realized as a two weight inequality for the Hilbert transform. In particular, a more delicate counterexample was developed by Nazarov-Volberg [40] to disprove a conjectured characterization of the Carleson measures for a model space. The Nazarov counterexample was also used by Nikol'skiĭ-Treil [42], in the context of spectral theory.

The Nazarov counterexample is by way of a Bellman function approach. In section "Example Weights", we give an explicit example. It is worth noting that in Sarason's question, the weights have a density $|f|^2$, for analytic f, and the subharmonicity could be an important part of the problem. But, in the context of model spaces, completely singular arbitrary measures can arise. In section "Example Weights", one of the weights is uniform measure on a Cantor set.

[1] In particular, they noted that the simple A_2 condition was not sufficient for the boundedness of the Hilbert transform, and conjectured that half-Poisson A_2 conditions would be sufficient, an indication of the powerful sway held by the Muckenhoupt A_2 condition in the early years of the weighted theory.

Nazarov-Treil-Volberg were creating the field of non-homogeneous Harmonic Analysis, in a series of ground-breaking papers [35–37]. Their work, and a revitalization of the perspective of Eric Sawyer from the 1980's, lead them to conjecture the characterization proved in this paper. Moreover, their influential proof strategy, devised in [38, 56], lead to a verification of the conjecture in the case that both weights were doubling. This paper uses their strategy, with several additional features. At the same time, their approach is generic, in that it applies to general Calderón-Zygmund operators. Specific properties of the Hilbert transform had to be used in the characterization. These properties were identified in [17, 20, 21, 23], and the more precise description of what was accomplished at each stage is spread out throughout the paper.

The Circle

The two weight inequality has an equivalent formulation on the circle, which we formulate now. Given two weights w, σ on the circle group $\mathbb{T} \equiv \mathbb{R}/2\pi\mathbb{Z}$, we consider the norm inequality

$$\int_{\mathbb{T}} \left| \int f(y) \cdot \cot\left(\frac{x-y}{2}\right) \sigma(dy) \right|^2 dw \leq \mathcal{N}^2 \|f\|^2_{L^2(\mathbb{T},\sigma)}. \tag{7}$$

This is abbreviated to $\|H^{\mathbb{T}}_{\sigma} f\|_{L^2(w)} \leq \mathcal{N}\|f\|_{L^2(\sigma)}$.

Theorem 1.3 *The inequality (7) holds if and only if the pair of weights below satisfy the conditions below and their duals. For all intervals $I \subset \mathbb{T}$, with $|I| \leq 1$, there are finite constants $\mathcal{A}_2$ and $\mathcal{T}$, such that*

$$\frac{\sigma(I)}{|I|} \cdot P^{\mathbb{T}}(w\mathbf{1}_{\mathbb{T}\setminus I})(x_I, 1-|I|) < \mathcal{A}_2, \tag{8}$$

$$\int_I |H^{\mathbb{T}}_{\sigma} \mathbf{1}_I|^2 \, dw \leq \mathcal{T}^2 \sigma(I). \tag{9}$$

Moreover, letting $\mathcal{A}_2$ and $\mathcal{T}$ be the best constants in these inequalities and their duals, there holds $\mathcal{N} \simeq \mathcal{A}_2^{1/2} + \mathcal{T}$.

In (8), the term $P^{\mathbb{T}} w(x_I, r)$ is the standard Poisson operator on the disk, evaluated at a point in the unit disk given by the center of the of the interval x_I, and the radial factor r.

Let us indicate how to prove the theorem above from Theorem 1.1. Fix σ and w be two weights on $\mathbb{T}$. Embed the weight w into $[0, 1]$ in the natural way, and call the resulting measure w'. Place three copies of σ on the intervals $[-1, 0]$, $(0, 1]$ and $(1, 2]$, and call the resulting measure σ'. Thus, σ' and w' are two weights on $\mathbb{R}$. It is clear that σ' and w' satisfy the Poisson A_2 condition with holes on $\mathbb{R}$.

For a function $f \in L^2(\mathbb{T};\sigma)$, let f' be three copies of f on the intervals $[-1,0]$, $(0,1]$ and $(1,2]$. Viewing $\mathbb{T}$ as $[0,1]$, there is a subtle difference between $H_\sigma^{\mathbb{T}} f(x)$ and $H_{\sigma'} f'(x)$, the former computed on $\mathbb{T}$, and the latter on $\mathbb{R}$. Namely

$$|H_\sigma^{\mathbb{T}} - \tfrac{1}{2} H_{\sigma'} f'(x)| \lesssim \int_{-3}^{3} |f'(y)| \cdot |x-y|^2 \, \sigma(dy).$$

It is easy to see that the A_2 condition implies that the operator on the right is bounded. Hence, the testing conditions on $\mathbb{T}$ imply those for w' and σ'. Hence $H_{\sigma'}$ maps $L^2(\sigma')$ to $L^2(w')$. From that, we deduce the boundedness of $H_\sigma^{\mathbb{T}}$.

Cora Sadosky. Cora Sadosky and I met only a couple of times, which is a pity, since my research has been so strongly influenced by her passions and interests. Her work with Cotlar on the L^p variant of the Helson-Szegő theorem is a beautiful complex variable result well beyond the reach of the current real-variable techniques. Her interest in Hankel forms on two and more complex variables has been my own for several years. And, in a number of small ways, I work to support more diversity in the profession, again following her lead.

Cora Sadosky's family came up in 2005, during a three month stay in Argentina, in a antiquarian bookstore just a few steps from the Casa Rosada in Buenos Aires. The proprietor, upon hearing I was a mathematician, remembered his own youth and a compelling Professor Manuel Sadosky. He remembered that the Professor had a daughter and asked after her. This was the third or fourth conversation of this type I had in that lovely city! It is a privilege to work on the beautiful subject of mathematics. Even more so to have passion, and insights that others will carry forward.

Preliminaries

Principal Values

We make no assertion about principal values of the Hilbert transform, and do not expect them to exist in the generality in which we are considering. One can then be concerned about how the definition is made. There are a couple of different options. One can impose some sort of truncation on the integrals, and the statements of the theorems are then understood to be uniform over all truncations. Many of the different possible truncations will be equivalent, since the A_2 condition will hold, see [31] for a general discussion of this issue. Alternatively, one can formally define

$$\langle H_\sigma f, g\rangle_w := \int\int f(y)g(x)\frac{dy\,dx}{y-x}$$

for all f, g which have closed supports that are a positive distance apart, and extend H linearly from there.

In our proof, all of the essential difficulties in the proof arise when f and g have widely separated supports. The definition of H_σ in this case is of course by the formula above.

Dyadic Grids and Haar Functions

A *grid* is a collection $\mathcal{D}$ of left closed, right open intervals so that for all $I, J \in \mathcal{D}$, $I \cap J = \emptyset, I, J$. Further say that $\mathcal{D}$ is a *dyadic grid* if for all integers n, the collection $\{I \in \mathcal{D} \,:\, |I| = 2^n\}$ partitions $\mathbb{R}$, aside from the endpoints of the intervals.

For a sub collection $\mathcal{F}$ of a dyadic grid $\mathcal{D}$, set $\pi_{\mathcal{F}} I$ to be the minimal element of $\mathcal{F}$ that contains I; I need not be a member of F. Set $\pi^1_{\mathcal{F}} I$ to be the minimal member of $\mathcal{F}$ that strictly contains I, inductively define $\pi^{t+1}_{\mathcal{F}} I = \pi^1_{\mathcal{F}}(\pi^t_{\mathcal{F}} I)$.

Say that the collection $\mathcal{D}$ is *admissible for weight* σ if σ does not have a point mass at any endpoint of an interval $I \in \mathcal{D}$.

Haar Functions

Let $\mathcal{D}$ be admissible for σ be a weight on $\mathbb{R}$. If $I \in \mathcal{D}$ is such that σ assigns non-zero weight to both children of I, the associated Haar function is chosen to have a non-negative inner product with the independent variable, $\langle x, h^\sigma_I(x) \rangle_\sigma \geq 0$, a convenient choice due to the central role of the energy inequality, (19).

$$h^\sigma_I(x) := \sqrt{\frac{\sigma(I_-)\sigma(I_+)}{\sigma(I)}} \left(\frac{I_+(x)}{\sigma(I_+)} - \frac{I_-(x)}{\sigma(I_-)} \right). \tag{10}$$

In this definition, we are identifying an interval with its indicator function, and we will do so throughout the remainder of the paper. This is an $L^2(\sigma)$-normalized function, and has σ-integral zero. If σ is supported only on one child of I, then we set $h^\sigma_I \equiv 0$.

For any dyadic interval I_0 with $\sigma(I_0) > 0$, the non-zero functions among $\{\sigma(I_0)^{-1/2} I_0\} \cup \{h^\sigma_I \,:\, I \in \mathcal{D}, I \subset I_0\}$ form an orthonormal basis for $L^2(I_0, \sigma)$. We will use the notation $L^2_0(I_0, \sigma)$ for the subspace of $L^2(I_0, \sigma)$ of functions with mean zero. It has orthonormal basis consisting of the non-zero functions in $\{h^\sigma_I \,:\, I \in \mathcal{D}, I \subset I_0\}$. These are familiar properties. But, another familiar property, that the positive and negative values of h^σ_I are comparable in absolute value, fails in a dramatic fashion for non-doubling measures. See Fig. 4.

We will use the notations $\mathbb{E}^\sigma_I f = \sigma(I)^{-1} \int_I f \; d\sigma, \hat{f}(I) = \langle f, h^\sigma_I \rangle_\sigma$, as well as the equality below, holding for those I with $h^\sigma_I \not\equiv 0$.

$$\Delta^\sigma_I f = \langle f, h^\sigma_I \rangle_\sigma h^\sigma_I = I_+ \mathbb{E}^\sigma_{I_+} f + I_- \mathbb{E}^\sigma_{I_-} f - I \mathbb{E}^\sigma_I f. \tag{11}$$

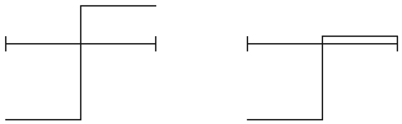


Fig. 4 Two Haar functions. For the *left* function, the weight is nearly equally distributed between the two halves of the interval, in sharp contrast to the function on the *right*, in which the weight on the *right* half is much larger than on the *left*

This is the familiar martingale difference equality, and so we will refer to $\Delta_I^\sigma f$ as a martingale difference. It implies the familiar telescoping identity $\mathbb{E}_J^\sigma f = \sum_{I\,:\,I \supsetneq J} \mathbb{E}_J^\sigma \Delta_I^\sigma f$.

The *Haar support* of a function $f \in L^2(\sigma)$ is the collection $\{I \,:\, \hat{f}(I) \neq 0\}$.

Random Dyadic Grids

Let $\widehat{\mathcal{D}}$ be the standard dyadic grid in $\mathbb{R}$, thus all intervals $[0, 2^n]$ for $n \in \mathbb{N}$ are in $\widehat{\mathcal{D}}$. A *random dyadic grid* $\mathcal{D}$ is specified by $\omega = \{\omega_n\} \in \{0, 1\}^{\mathbb{Z}}$, and the elements are

$$I = \hat{I} \dot{+} \omega := \hat{I} + \sum_{n\,:\,2^{-n} < |I|} 2^{-n}\omega_n, \qquad \hat{I} \in \widehat{\mathcal{D}}.$$

The natural uniform probability measure $\mathbb{P}$ is placed upon $\{0, 1\}^{\mathbb{Z}}$.

Fix $0 < \varepsilon < 1$ and $r \in \mathbb{N}$. An interval $I \in \mathcal{D}$ is said to (ε, r)-*good* if for all intervals $J \in \mathcal{D}$ with $|J| \geq 2^{r-1}|I|$, the distance from ∂J and *either child of* I is at least $|I|^\varepsilon |J|^{1-\varepsilon}$. Otherwise I is said to be (ε, r)-*bad*. These are the basic properties of this definition.

Proposition 2.1 *These three properties hold.*

(1) The property of $I = \hat{I} \dot{+} \omega$ *being* (ε, r)*-good only depends upon* ω *and* $|I|$*.*
(2) $\mathbf{p}_{good} := \mathbb{P}(I \text{ is } (\varepsilon, r)\text{-good})$ *is independent of* I*.*
(3) $\mathbf{p}_{bad} := 1 - \mathbf{p}_{good} \lesssim \varepsilon^{-1} 2^{-\varepsilon r}$*.*

Proof An interval $I = \hat{I} \dot{+} \omega$ is equally likely to be the left or right half of its parent $\pi_{\mathcal{D}}^1 I$, depending only on ω_n, where $|I| = 2^n$. Similarly, I is equally likely to be any one of the 2^t potential positions in $\pi_{\mathcal{D}}^t I$, and its exact position is determined by $\{\omega_n, \ldots, \omega_{n+t-1}\}$. This proves the first two claims.

For the last, if I is bad, then for some $t > r$, there holds dist $(I,\ \partial \pi_{\mathcal{D}}^t I) \leq 2^{(1-\varepsilon)t}|I|$. For this to happen, it is necessary that the numbers $\{\omega_s \,:\, n + \lceil (1-\varepsilon)t \rceil < u \leq n + t - 1\}$ all be equal, and hence are either all 0 or all 1. This clearly proves that

$$\mathbf{p}_{\text{bad}} \leq \sum_{t=r+1}^{\infty} 2^{1-(t-\lceil(1-\varepsilon)t\rceil)} \lesssim \varepsilon^{-1}2^{-\varepsilon r}.$$

□

This elementary proposition is used in the following fundamental way. Fix two weights w, σ. With probability one, a random $\mathcal{D}$ is admissible for both w and σ. Indeed, the collection of points that are point masses for one of the two weights is a fixed countable collection of points. And any fixed point has probability zero of being an endpoint of an interval in $\mathcal{D}$. Hence, we can, with probability one, define the Haar basis adapted to these two weights. Write the identity operator on $L^2(\sigma)$ by

$$P^{\sigma}_{\text{good}}f + P^{\sigma}_{\text{bad}}f \qquad \text{where} \quad P^{\sigma}_{\text{good}}f := \sum_{I\in\mathcal{D}\,:\, I \text{ is } (\varepsilon,r)\text{-good}} \langle f, h^{\sigma}_I\rangle_{\sigma} h^{\sigma}_I .$$

Use the same notation for the weight w.

Proposition 2.2 *There holds*

$$\mathbb{E}\|P^{\sigma}_{bad}f\|^2_{\sigma} \lesssim \varepsilon^{-1}2^{-\varepsilon r}\|f\|^2_{\sigma}.$$

Proof The location of I and the property of I being bad are independent, hence

$$\mathbb{E}\|P^{\sigma}_{\text{bad}}f\|_{\sigma} = \mathbb{E}\sum_{I\in\mathcal{D}} \mathbf{1}_{I \text{ is bad}}\,\hat{f}(I)^2 = \mathbf{p}_{\text{bad}}\mathbb{E}\sum_{I\in\mathcal{D}}\hat{f}(I)^2 = \mathbf{p}_{\text{bad}}\|f\|^2_{\sigma}$$

and then the proposition follows. □

Lemma 2.3 *For any constant* $1 \leq C < \infty$*,* $0 < \varepsilon < 1$*, there is a choice of* $r \in \mathbb{N}$ *sufficiently large so that this holds. Let* w, σ *be a pair of weights for which the constant* $\mathcal{H}$ *and the constant* $\mathcal{N}$ *in* (1) *are finite. Suppose there holds uniformly over admissible dyadic grids* $\mathcal{D}$*,*

$$|\langle H_{\sigma}P^{\sigma}_{good}f, P^{w}_{good}g\rangle_w| \leq C\mathcal{H}\|f\|_{\sigma}\|g\|_w, \tag{12}$$

then, $\mathcal{N} \leq 2C\mathcal{H}$*.*

Proof Use Proposition 2.2 on the good and bad projections, as written and the same version for $L^2(w)$.

$$\begin{aligned}|\langle H_{\sigma}f, g\rangle_w| \leq \mathbb{E}\big\{&|\langle H_{\sigma}P^{\sigma}_{\text{good}}f, P^{w}_{\text{good}}g\rangle_w| + |\langle H_{\sigma}P^{\sigma}_{\text{good}}f, P^{w}_{\text{bad}}g\rangle_w| \\ &+ |\langle H_{\sigma}P^{\sigma}_{\text{bad}}f, P^{w}_{\text{good}}g\rangle_w| + |\langle H_{\sigma}P^{\sigma}_{\text{bad}}f, P^{w}_{\text{bad}}g\rangle_w|\big\}.\end{aligned}$$

The first term is controlled by the assumption (12), and the remaining terms are controlled by the finiteness of $\mathcal{N}$ and average-norm estimate on the bad projection. By appropriate selection of $f \in L^2(\sigma)$ and $g \in L^2(w)$, there holds

$$\mathcal{N}_{\tau_0} \le C\mathcal{H} + C'\varepsilon^{-1}2^{-\varepsilon r/2}\mathcal{N}_{\tau_0}.$$

For any fixed ε, we can take $r \gtrsim \varepsilon^{-1}\log\varepsilon^{-1}$, so that the second term can be absorbed into the left hand side. □

Context and Discussion

The random grid method was pioneered in [36], and is a critical tool in non-homogeneous analysis [56], where the weights need not be doubling. It has a broader set of uses, as witnessed by a powerful representation of a general Calderón-Zygmund operator as a rapidly convergent sum of dyadic operators due to Hytönen [13].

The parameterization of the grids used here follows Hytönen [12], but the statistics of this parameterization are those of the random shift in Nazarov-Treil-Volberg [35, 36].

Necessary Conditions

Herein, we take up the necessity of the A_2 condition from the norm inequality. Following that is the monotonicity property, an essential property of the Hilbert transform, and then showing the necessity of the energy inequality from the A_2 and interval testing condition. The energy inequality is foundational to the proof.

The A_2 Condition

The A_2 condition has different forms, and so we clarify the language associated with the A_2 condition here. The *simple* A_2 condition is

$$\sup_I \frac{\sigma(I)}{|I|}\cdot\frac{w(I)}{|I|},$$

the supremum formed over all intervals I. This reduces to the classical Muckenhoupt condition if $w(dx) = w(x)dx$, where $w(x) > 0$ a.e., and $\sigma(dx) = w(x)^{-1}dx$. Next, are the *half-Poisson* conditions:

$$\sup_I P(\sigma, I)\frac{w(I)}{|I|} < \infty.$$

Finally there is the full Poisson A_2 condition

$$\sup_I P(\sigma, I) \cdot P(w, I) < \infty \tag{13}$$

and of course, we only use the Poisson condition with holes, of Hytönen [15]. We verify that the Poisson A_2 condition (3) is necessary for the two weight inequality (1).

Proposition 3.1 *Assume that the pair of weights do not share a common point mass, and that the norm inequality* (1) *holds. Then, the* A_2 *condition* (3) *holds.*

Proof Fix the interval $I = (a, b)$ as in (3), and let $a \in I$. We will estimate half of the Poisson integral of w using the notation

$$p_I(x)^2 := \frac{|I|}{(|I| + \operatorname{dist}(x, I))^2} \mathbf{1}_{[b,\infty)} \tag{14}$$

so that $P(w \cdot [b, \infty), I) = \|p_I \cdot [b, \infty)\|^2_{L^2(w)}$. Below, we estimate the right half of the Poisson integral of w.

$$\begin{aligned} \frac{\sigma(I)}{|I|^{1/2}} \cdot P(w \cdot [b, \infty), I) &\le \int_I \int_b^\infty \frac{1}{|I| + \operatorname{dist}(x, I)} \cdot \frac{1}{y - x}\, w(dx)\, \sigma(dy) \\ &= \langle H_\sigma(I), p_I \rangle_w \lesssim \mathcal{N} \sigma(I)^{1/2} \|p_I\|_w . \end{aligned}$$

Rearranging,

$$\frac{\sigma(I')}{|I|} \cdot P(w \cdot (a, \infty), I) \lesssim \mathcal{N}^2. \tag{15}$$

Clearly, the same inequality holds for $(-\infty, a]$. □

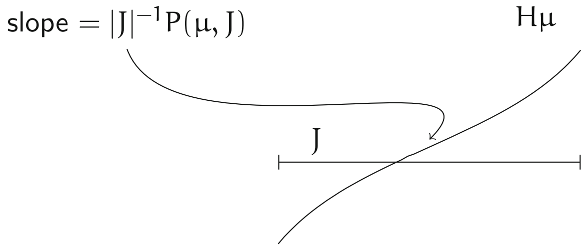
The Monotonicity Principle

Certain kinds of off-diagonal estimates for the Hilbert transform have concrete estimates in terms of the Poisson integral. This estimate makes this precise, and shows moreover that we need not be that careful about exactly which function appears in the Poisson integral. It is at the core of the entire proof.

Lemma 3.2 (Monotonicity Principle) *Suppose that the two weights* σ *and* w *satisfy the* A_2 *bound, and neither has a point mass at an endpoint of* I. *Let* $J \subset I$. *There holds for any* $g \in L^2(J, w)$, *with* w*-integral zero,*

$$P(\sigma(\mathbb{R} - I), I) \Big\langle \frac{x}{|I|}, \overline{g} \Big\rangle_w \lesssim \langle H(\sigma(\mathbb{R} - I)), \overline{g} \rangle_w . \tag{16}$$

Fig. 5 An illustration of the monotonicity principle



Here, $\overline{g} = \sum_{J'} |\widehat{g}(J')| h^w_{J'}$, *is a Haar multiplier applied to g. Suppose that* $J \subset I$ *is good, with* $2^r|J| \le |I|$. *Then for any two compactly supported weights* $|\nu| \le \mu$ *supported off of the interval I, there holds*

$$|\langle H\nu, g\rangle_w| \lesssim \langle H\mu, \overline{g}\rangle_w \simeq P(\mu, J)\Big\langle \frac{x}{|J|}, \overline{g}\Big\rangle_w. \tag{17}$$

Note that in the first estimate, the Poisson term is always estimated above by an inner product involving the Hilbert transform. In the second, note that the inner product can always be made larger by making the weight positive. Moreover, under moderate assumptions on the support of the weight, the first inequality can be reversed. See Fig. 5. In that figure, the function μ is outside of $2^{r(1-\epsilon)}J$, so that $H\mu$ is a smooth increasing function on J. Moreover, the derivative of $H\mu$ is approximately $|J|^{-1}P(\mu, J)$. So, if we form an inner product with the Haar function h^w_J, we only need to be concerned with the linear approximation to $H\mu$. However, the conditions to get the reversal are particular, and this drives the case analysis in different sections of the proof.

Proof We consider the first estimate. By linearity, it suffices to consider the case of $g(x) = h^w_J(x)$, for $J \subset I$, and indeed we can take $J = I$. We need to separate the two weights involved. The A_2 condition is the only condition needed for the weak-boundedness principle, Proposition 7.4. Applying it in this setting, notice that it shows that for $\lambda > 1$,

$$|\langle H_\sigma(\lambda I - I), h^w_I\rangle| \lesssim \mathcal{A}_2^{1/2}\sqrt{\sigma(\lambda I - I)}.$$

The assumption that σ does not have mass at the endpoints of I implies that $\sigma(\lambda I - I)$ can be made arbitrarily small, as $\lambda \downarrow 1$. Therefore, it suffices to consider $H_\sigma(\mathbb{R} - \lambda I)$, for some fixed $\lambda > 1$.

Then estimate

$$
\begin{aligned}
\langle H^c(\sigma\cdot(\mathbb{R}-\lambda I)), h_I^w\rangle_w &= \int_{\mathbb{R}-\lambda I}\int_J \frac{1}{y-x} h_J^w(x)\; w(dx)\sigma(dy) \\
&= \int_{\mathbb{R}-\lambda I}\int_I \Big[\frac{1}{y-x} - \frac{1}{y-x_J}\Big] h_I^w(x)\; w(dx)\,\sigma(dy) \\
&= \int_{\mathbb{R}-\lambda I}\int_I \frac{x-x_J}{(y-x)(y-x_J)} h_I^w(x)\; w(dx)\,\sigma(dy) \\
&\gtrsim \int_{\mathbb{R}-I}\int_I \frac{|I|}{(|I|+\operatorname{dist}(y,I))^2}\cdot\frac{x-x_J}{|I|} h_I^w(x)\; w(dx)\,\sigma(dy) \\
&= P(\sigma\cdot(\mathbb{R}-\lambda I), I)\Big\langle \frac{x}{|I|}, h_I^w\Big\rangle_w .
\end{aligned}
$$

Here, x_J is the center of J, and it can be inserted for the usual reason that h_J^w has w-integral zero. Then, use the fact that $(x-x_J)h_J^w \ge 0$, and that $(y-x)(y-x_J) > 0$. So (16) holds.

The second inequality (17) comes with the assumption that $J \subset I$, $2^r|J| < |I|$, whence $\operatorname{dist}(J, I) > |J|^\epsilon |I|^{1-\epsilon} \ge 2^{r(1-\epsilon)}|J|$. Namely, the support of h_J^w and that of μ are separated. Then, inserting a constant as we can since the Haar function has integral zero,

$$
\begin{aligned}
\langle H\nu, h_J^w\rangle_w &= \int_{\mathbb{R}-I}\int_J \{\frac{1}{y-x} - \frac{1}{y-x_J}\} h_J^w(x)\; \nu(dy)\, w(dx) \\
&= \int_{\mathbb{R}-I}\int_J \frac{x-x_J}{(y-x)(y-x_J)} h_J^w(x)\; \nu(dy)\, w(dx)
\end{aligned}
$$

Notice that the integrand is non-negative, hence we can make the integral bigger in absolute value by replacing $|\nu|$ by μ. This is the first inequality in (17). For the second equivalence, by the separation in supports, we have $\frac{1}{(y-x)(y-x_J)} \simeq \frac{1}{(y-x_J)^2}$ in the range of integration. And this finishes the proof. □

The Energy Inequality

The energy inequality is phrased in terms of the quantity

$$
E(w,I)^2 := |I|^{-2}\mathbb{E}_I^w\big|x\cdot I - \mathbb{E}_I^w x\big|^2 = |I|^{-2}\sum_{J\,:\,J\subset I}\langle x, h_J^w\rangle_w^2. \tag{18}
$$

Lemma 3.3 (The Energy Inequality) *For any interval I_0 and any partition $\mathcal{P}$ of I_0 into intervals such that neither σ nor w have point masses at the endpoints, there holds*

$$\sum_{I \in \mathcal{P}} P(\sigma(I_0 \setminus I), I)^2 E(w, I)^2 w(I) \le C_0 \mathcal{H}^2 \sigma(I_0). \tag{19}$$

Here, C_0 is an absolute constant.

Proof It follows from (16), viewed in dual fashion, that

$$\begin{aligned} P(\sigma(I_0 \setminus I), I)^2 E(w, I)^2 w(I) &\lesssim \|H(\sigma(I_0 \setminus I)) \cdot I\|_w^2 \\ &\lesssim \|H(\sigma \cdot I_0) \cdot I\|_w^2 + \|H(\sigma \cdot I) \cdot I\|_w^2 \\ &\lesssim \|H(\sigma \cdot I_0) \cdot I\|_w^2 + \mathcal{T}^2 \sigma(I). \end{aligned}$$

Above, we have appealed to the testing assumption (4). Summing over $I \in \mathcal{P}$, the second term above is clearly no more than $\mathcal{T}^2 \sigma(I_0)$. And the second term is no more than

$$\|H(\sigma \cdot I_0) \cdot I_0\|_w^2 \le \mathcal{T}^2 \sigma(I_0).$$

□

Context and Discussion

In the absence of common point masses, the necessity of the *full* $\mathcal{A}_2$ condition, namely (13), was easily available, with an argument of Sergei Treil already pointed out by Sarason in his note [51]. This argument, based upon complex variables, has close analogs in [38, 56]. A real variable proof is in [20], it is essentially an elaboration of the argument in the early paper of Muckenhoupt and Wheeden [33]. Despite the necessity, only the half Poisson A_2 condition is used, together with testing, in the proof of sufficiency, in the case of no common point masses.

Higher dimensional extensions of the A_2 which are not straight forward, are discussed in [19]. There are notable distinctions important to higher dimensions. First, the necessary Poisson type condition only comes in its 'half' form. Second, the power on the Poisson kernel comes as the square of the dimension of the kernel involved, a feature familiar from the analysis of reproducing kernel spaces. Third, the degree of the Poisson kernel matches the important derivative Poisson decay, important to energy considerations, only when the dimension of the kernel is one.

The energy inequality was influenced by the following assumption placed upon the pair of weights in [38, 56]. Assume that there is a finite constant $\mathcal{P}$ so that for all intervals I_0, and all partitions $\mathcal{P}$ of I_0 into intervals,

$$\sum_{I \in \mathcal{P}} P(\sigma \cdot I_0, I)^2 w(I) \le \mathcal{P}^2 \sigma(I_0). \tag{20}$$

Also assume that the dual inequality holds. In the language of Nazarov-Treil-Volberg, this is the *pivotal condition*. They proved

Theorem C *Assume that* w *and* σ *do not share a common point mass. Then, there holds* $\mathcal{N} \lesssim \mathcal{A}_2^{1/2} + \mathcal{T} + \mathcal{P}$.

This is a very strong Theorem, with an important proof. It decisively used the tools of non-homogeneous harmonic analysis, namely random grids, and good-bad projections. The pivotal condition controlled certain degeneracies in the pair of weights, compare to Definition 4.3. To illustrate the difficulties in the general case, we prove this theorem in section "Proof Under the Pivotal Assumption".

The pivotal condition holds if the pair of maximal function estimates hold, namely $M_\sigma : L^2(\sigma) \mapsto L^2(w)$ and $M_w : L^2(w) \mapsto L^2(\sigma)$. This is easy to see. From (20),

$$\begin{aligned}\sum_{I \in \mathcal{P}} P(\sigma \cdot I_0, I)^2 w(I) &\leq \sum_{I \in \mathcal{P}} \inf_{x \in I} M(\sigma \cdot I_0)(x)^2 w(I) \\ &\leq \int_{I_0} M(\sigma \cdot I_0)^2 \, dw \lesssim \sigma(I_0),\end{aligned}$$

by the assumed norm bound on the maximal function. One sees that Theorem C offered a complete characterization of the two weight inequality for the triple of operators $(H_\sigma, M_\sigma, M_w)$. If the pair of weights are doubling, then the boundedness of the maximal functions is a consequence of the A_2 condition.[2] The full characterization of the boundedness of the Hilbert transform was thus known for doubling measures. See [56].

The pivotal condition is generic in the following sense. Assuming the pivotal condition, the Hilbert transform can be replaced by a generic Calderón-Zygmund operator with one derivative on its kernel. This, and its extension to operators with a rougher kernel, was fundamental in the paper [45], whose main result was an important intermediate one in the solution of the A_2 conjecture [13].

Nazarov-Treil-Volberg, in language reminiscent of Sarason, wrote that 'perhaps the pivotal condition is necessary' for the boundedness of the Hilbert transform. This turned out to have a strong measure of truth, in that using the specific structure of the Hilbert transform, the *energy inequality* was shown necessary in [20]. Note that one can formally obtain the pivotal condition (20) from the energy inequality (19) by raising the energy term $E(w, I)$ to the zero power, rather than the necessary power 2. The paper [20] then adapted the approach of [38, 56], essentially imposing a new weaker condition on the pair of weights in which one raised the energy to a power intermediate between 0 and 2. In addition, that paper provided an explicit example,

[2] Alternatively, under the assumption of w being doubling, check that the energy satisfies $E(w, I) \gtrsim 1$, with the implied constant depending upon the doubling constant. Thus, the necessary energy inequality implies the pivotal condition.

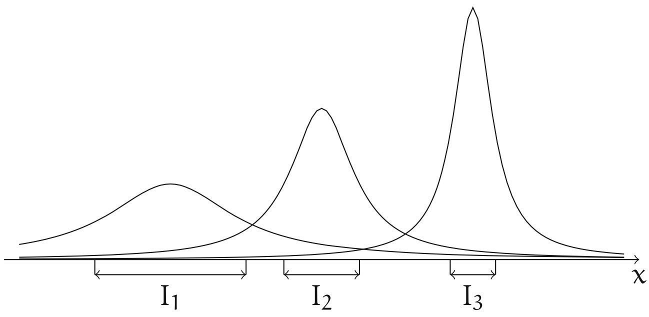


Fig. 6 The function $\frac{|I|}{|I|^2+|x-x_I|^2}$ are graphed for three separate intervals

recounted in section "Example Weights", that showed that the pivotal condition (20) is *not necessary* for the boundedness of the Hilbert transform.

The energy inequality is rather subtle. The Poisson term $P(\sigma, I)$ can be much larger than the simple average, but this is compensated for with the terms $E(w,I)^2 w(I)$. The Fig. 6 is offered to provide some insight into the 'long tails' that the Poisson term can have.

Another indication of this subtlety is the observation that the energy inequality will not follow from just the A_2 condition. Given interval I_0, and partition $\mathcal{P}$ of I_0, one can write

$$\sum_{I\in\mathcal{P}} P(\sigma,I)^2 E(w,I)^2 w(I) \le \mathcal{A}_2 \sum_{I\in\mathcal{P}} |I|\cdot P(\sigma,I)^2$$

$$= \mathcal{A}_2 \int_{I_0} \sum_{I\in\mathcal{P}} \frac{|I|^2}{(|I|+\operatorname{dist}(x,I))^2}\,\sigma(dx).$$

To finish, one would have to know that the function inside the integral is bounded. But, this is not true in general. Though a very tame *BMO* function, this fact does not help, since σ is a general measure, and need not satisfy any A_∞ type condition. Indeed, the proof of the main theorem would be more or less classical if the weights satisfy a A_∞ type conditions.

The monotonicity principle, Lemma 3.2, was noted in [21]. It, with the energy inequality, are essential aspects of the proof.

Global to Local Reduction

Our aim is to prove the estimate (12),

$$\sup |\langle H_\sigma P^\sigma_{\text{good}} f, P^w_{\text{good}} g\rangle_w| \lesssim \mathcal{H}\|f\|_\sigma \|g\|_w.$$

That is, the bilinear form only needs to be controlled for (ε, r)*-good functions* $f = P^{\sigma}_{\text{good}}f$ *and similarly for g, goodness being defined with respect to a fixed dyadic grid.* Suppressing the notation, we write 'good' for '(ε, r)-good,' and it is always assumed that the dyadic grid $\mathcal{D}$ is fixed, and only good intervals are in the Haar support of f and g. We clearly remark on goodness when the property is used; any value of $0 < \varepsilon \leq \frac{1}{4}$ is sufficient for our purposes. The symbol ε is kept throughout, as a guide to the appearance of the good property of intervals.

The inequality above is reduced to the local estimate, (32), at the end of this section. It is sufficient to assume that f and g are supported on an interval I^0; by trivial use of the interval testing condition, we can further assume that f and g are of integral zero in their respective spaces. Thus, f is in the linear span of (good) Haar functions h^{σ}_I for $I \subset I^0$, and similarly for g.

The distinction between $J \subset I$ and $J \Subset I$ ($J \subset I$ and $2^r|J| \leq |I|$) forces some case analysis. This is further simplified by this assumption on the Haar supports of f, g. There are two integers s_f, s_g such that

$$f = \sum_{\substack{I\,:\,I\subset I^0 \\ \log_2|I|\in s_f + r\mathbb{Z}}} \Delta^{\sigma}_I f \tag{21}$$

and similarly for g. Thus, the lengths of the (good) intervals I are restricted to an equivalence class mod r, which is to say that the *scales of f are separated by r*, and the same for g. This will be a convenience at a few technical points below. Set $\mathcal{D}_f := \{I \;:\; \log_2|I| \in s_f - 1 + r\mathbb{Z}\}$, so these are the children of the intervals that appear in (21). Due to the probabilistic way in which the grids are constructed, we can further assume that $I^0 \in \mathcal{D}_f$. Also set $\mathcal{D}_g := \{I \;:\; \log_2|I| \in s_g + r\mathbb{Z}$.

We are to control the bilinear form

$$\langle H_\sigma f, g\rangle_w = \sum_{I,J\,:\,I,J\subset I^0} \langle H_\sigma \Delta^{\sigma}_I f, \Delta^{w}_J g\rangle_w. \tag{22}$$

The sum is broken into many summands, as is typical in these arguments, but the manner in which it is done has some important points below. The most important of these are the two 'triangular' forms

$$B^{\text{above}}(f, g) := \sum_{I\,:\,I\subset I^0} \sum_{J\,:\,J\Subset I_J} \mathbb{E}^{\sigma}_{I_J} \Delta^{\sigma}_I f \cdot \langle H_\sigma I_J, \Delta^{w}_J g\rangle_w \tag{23}$$

and the dual form, $B^{\text{below}}(f, g)$. Here, $J \Subset I$ means that $J \subset I$ and $2^r|J| \leq |I|$, in words 'J is strongly contained in I'. And the interval I_J is the child of I that contains J. Goodness of J justifies the use of this condition. A basic fact, proved in section "Elementary Estimates", is

Lemma 4.1 *There holds*

$$\left|\langle H_\sigma f, g\rangle_w - B^{above}(f, g) - B^{below}(f, g)\right| \lesssim \mathcal{H}\|f\|_\sigma \|g\|_w.$$

Thus, the main technical result is as below; it immediately supplies our main theorem.

Theorem 4.2 *There holds*

$$|B^{above}(f,g)| \lesssim \mathcal{H}\|f\|_\sigma \|g\|_w.$$

The same inequality holds for the dual form $B^{below}(f,g)$.

The remainder of this section is devoted to a reduction of the global Theorem 4.2 to a local estimate described in section "The Stopping Form". In the local estimate, the function f is more structured in that it has bounded averages on a fixed interval, and the pair of functions f, g are more structured in that their Haar supports avoid intervals that strongly violate the energy inequality. Still the argument to control this term requires a subtle recursion.

We construct stopping data, which accomplishes two ends, in that it will control certain telescoping sums of martingale differences of f, and that it controls certain degeneracies in an energy estimate on the weights.

Definition 4.3 Define $\mathcal{F}$, the stopping intervals, recursively by initializing $I^0 \in \mathcal{F}$, and in the recursive step, if $F \in \mathcal{F}$ is minimal, add to $\mathcal{F}$ the maximal subintervals $F' \subset F$, with $F' \in \mathcal{D}_f$, either

***f* stopping** $\quad \mathbb{E}^\sigma_{F'}|f| > C\alpha_f(F) := \mathbb{E}^\sigma_F|f|.$
Energy Stopping $\quad \|H_\sigma(F \setminus F') \cdot F'\|^2_w \geq C\mathcal{H}^2\sigma(F').$

That is, we stop if either the average of f becomes too large, or, essentially, the energy condition becomes too large.

For appropriate constant C, it follows that $\mathcal{F}$ is σ-Carleson, namely

$$\sum_{F'\in\mathcal{F}\,:\,F'\subset F} \sigma(F') \leq \tfrac{1}{10}\sigma(F), \qquad F \in \mathcal{F}. \tag{24}$$

Many properties of the σ-Carleson property are used below. But, also note the following property:

$$|\mathbb{E}^\sigma_I f| \leq C\alpha_f(\pi_{\mathcal{F}} I) \tag{25}$$

We will use the notation

$$P^\sigma_F f := \sum_{I\in\mathcal{D}\,:\,\pi_{\mathcal{F}}I=F} \Delta^\sigma_I f, \qquad F \in \mathcal{F}. \tag{26}$$

and a dual projection $Q^w_F g$, is defined similarly, but importantly, we replace $\pi_{\mathcal{F}} J = F$ by $\dot{\pi}_{\mathcal{F}} J = F$, meaning that F is the smallest interval in $\mathcal{F}$ such that $J \Subset F$. (Note that both are projections, but $P^\sigma_F f$ is a structured function, while $Q^w_F g$ is not.) The σ-Carleson property allows us to estimate

$$\sum_{F\in\mathcal{F}}\{\alpha_f(F)\sigma(F)^{1/2}+\|P^\sigma_F f\|_\sigma\}\|Q^w_F g\|_w$$
$$\leq\left[\sum_{F\in\mathcal{F}}\{\alpha_f(F)^2\sigma(F)+\|P^\sigma_F f\|^2_\sigma\}\times\sum_{F\in\mathcal{F}}\|Q^w_F g\|^2_w\right]^{1/2}\lesssim\|f\|_\sigma\|g\|_w. \tag{27}$$

We will refer to as the *quasi-orthogonality* argument. It holds only under the assumption that the projections Q^w_F are pairwise orthogonal. It is very useful.

We henceforth concentrate on the 'above' forms, with all considerations applying in their dual formulation to control the 'below' forms. Return to the double sum (22), and define

$$B^{\mathrm{above}}_{\mathcal{F},\mathrm{loc}}(f,g):=\sum_{F\in\mathcal{F}}B^{\mathrm{above}}(P^\sigma_F f,Q^w_F g). \tag{28}$$

The global to local reduction is:

Corollary 4.4 (Global to Local Reduction) *There holds*

$$\left|B^{above}(f,g)-B^{above}_{\mathcal{F},loc}(f,g)\right|\lesssim\mathcal{H}\|f\|_\sigma\|g\|_w.$$

Proof Observe that $B^{\mathrm{above}}(f,g)$ is a sum over pairs of intervals (I,J) with $J\Subset I_J$, whence $\dot\pi_{\mathcal{F}}J\subset\pi_{\mathcal{F}}I$. Now, the case of $\dot\pi_{\mathcal{F}}J=\pi_{\mathcal{F}}I$ is contained in the form $B^{\mathrm{above}}_{\mathcal{F},\mathrm{loc}}(f,g)$, hence we need only concern ourselves with the case of $\dot\pi_{\mathcal{F}}J\subsetneq\pi_{\mathcal{F}}I$, that is, we need only bound

$$\sum_{F\in\mathcal{F}}\sum_{\substack{F'\in\mathcal{F}\\F'\subsetneq F}}B^{\mathrm{above}}_{\mathcal{F}}(P^\sigma_F f,Q^w_{F'}g).$$

Set $g_F:=Q^w_F g$. The sum in question is

$$\sum_{F\in\mathcal{F}}\sum_{I\,:\,I\supsetneq F}\mathbb{E}^\sigma_{I_J}\Delta^\sigma_I f\cdot\langle H_\sigma I_F,g_F\rangle_w. \tag{29}$$

We invoke, for the first time, the *Hilbert-Poisson exchange argument*: (a) Replace the argument of the Hilbert transform by a stopping interval. (b) Invoke the stopping tree construction to control the sum of martingale differences of f. (c) Apply interval testing, on the stopping interval. (d) Use the monotonicity principle to dominate the complementary term in terms of a Poisson integral. (e) Analyze the Poisson term. (f) Use quasi-orthogonality, as needed.

The argument of the Hilbert transform is I_F, the child of I that contains F. Write $I_F = F + (I_F - F)$, and use linearity of H_σ. Note that by the standard martingale difference identity and the construction of stopping data,

$$\Big| \sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_F} \Delta^\sigma_I f \Big| \lesssim \alpha_f(F), \qquad F \in \mathcal{F}.$$

Hence, invoking interval testing,

$$\begin{aligned}\Big| \sum_{F\in\mathcal{F}} \sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_F} \Delta^\sigma_I f \cdot \langle H_\sigma F, g_F\rangle_w \Big| &\lesssim \sum_{F\in\mathcal{F}} \alpha_f(F) \big| \langle H_\sigma F, g_F\rangle_w \big| \\ &\lesssim \mathcal{H} \sum_{F\in\mathcal{F}} \alpha_f(F)\sigma(F)^{1/2} \|g_F\|_w.\end{aligned}$$

Quasi-orthogonality bounds this last expression.

For the second expression, when the argument of the Hilbert transform is $I_F - F$, is the objective of section "The Remaining Part of the Global Estimate". We have proved

$$\Big| \sum_{F\in\mathcal{F}} \sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f \cdot \langle H_\sigma I_F, g_F\rangle_w \Big| \lesssim \mathcal{H} \|f\|_\sigma \|g\|_w. \tag{30}$$

This completes the Hilbert-Poisson exchange argument.

$$\Big| \sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_F} \Delta^\sigma_I f \cdot (I_F - F) \Big| \lesssim \Phi := \sum_{F'\in\mathcal{F}} \alpha_f(F') \cdot F', \qquad F \in \mathcal{F}.$$

Therefore, the monotonicity property (17) applies, and yields

$$\Big| \sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_F} \Delta^\sigma_I f \cdot \langle H_\sigma (I_F - F), g_F\rangle_w \Big| \lesssim \sum_{J\in\mathcal{J}^*(F)} P_\sigma(\Phi\cdot F^c, J) \Big\langle \frac{x}{|J|}, J\overline{g}_F \Big\rangle_w, \qquad F \in \mathcal{F}. \tag{31}$$

Here $\overline{g}_F := \sum_{J\in\mathcal{J}(F)\,:\,J \Subset F} |\hat{g}(J)| \cdot h^w_J$, so that every term has a positive inner product with x, and $\mathcal{J}^*(F)$ are the maximal good intervals $J \Subset F$, and $J \in \mathcal{D}_g$. (If $J \notin \mathcal{D}_g$, then $\langle g, h^w_J\rangle_w = 0$, by choice of g at the beginning of the proof.)

The control of the sum over $F \in \mathcal{F}$ of (31) □

It remains to control $B^{\text{above}}_{\mathcal{F},\text{loc}}(f, g)$. Keeping the quasi-orthogonality argument in mind, appropriate control on the individual summands is enough to control it. To describe what has been done, one must note that the functions $P^\sigma_F f$ *need not be bounded*. But, we are only concerned with averages over intervals where the average will be bounded. In addition this function and $Q^w_F g$ are well-adapted to the pair of weights w, σ. The next lemma, combined with the quasi-orthogonality estimate clearly completes the proof of the Theorem.

Lemma 4.5 (The Local Estimate) *For each $F \in \mathcal{F}$, there holds*

$$|B^{above}(P^{\sigma}_{F}f, Q^{w}_{F}g)| \le \mathcal{H}\{\alpha_f(F)\sigma(F)^{1/2} + \|P^{\sigma}_{F}f\|_{\sigma}\}\|Q^{w}_{F}g\|_{w}. \tag{32}$$

The first step in the proof of the Lemma above is to invoke the Hilbert-Poisson exchange argument again, but we will arrive at a Poisson term which falls outside the immediate scope of the energy inequality. Focusing on the argument of the Hilbert transform in (32), we write $I_J = F - (F - I_J)$. When the interval is F, and J is in the Haar support of $Q^{w}_{F}g$, notice that the scalar

$$\alpha_f(F)\varepsilon_J := \sum_{I\,:\,J \Subset I \subset F} \mathbb{E}^{\sigma}_{J} \Delta^{\sigma}_{I} f$$

is bounded by an absolute constant, by construction of the stopping intervals. Indeed, by the telescoping identity for martingale differences,

$$\alpha_f(F)\varepsilon_J = \sum_{I\,:\,I^{-} \subsetneq I \subset F} \mathbb{E}^{\sigma}_{I^{-}} \Delta^{\sigma}_{I} f = \mathbb{E}^{\sigma}_{I_J} f,$$

which is at most $C\alpha_f(F)$, since $\dot{\pi}_{\mathcal{F}} J = F$. Therefore, we can write

$$\Bigl| \sum_{I\,:\,I \subset F} \sum_{J\,:\,J \Subset I} \mathbb{E}^{\sigma}_{J} \Delta^{\sigma}_{I} f \cdot \langle H_{\sigma} F, \Delta^{w}_{J} g\rangle \Bigr| \le \alpha_f(F) \Bigl| \Bigl\langle H_{\sigma} F, \sum_{J\,:\,J \Subset F} \varepsilon_J \Delta^{w}_{J} g \Bigr\rangle_{w} \Bigr| \tag{33}$$

$$\le \mathcal{T}\alpha_f(F)\sigma(F)^{1/2} \Bigl\| \sum_{J\,:\,J \Subset F} \varepsilon_J \Delta^{w}_{J} g \Bigr\|_{w} \le \mathcal{T}\sigma(F)^{1/2}\|g\|_{w}. \tag{34}$$

This uses only interval testing and orthogonality of the martingale differences, and it matches the first half of the right hand side of (32).

When the argument of the Hilbert transform is $F - I_J$, this is the *stopping form*, the last component of the local part of the problem. It requires a subtle recursion, described in section "The Stopping Form".

Context and Discussion

Many $T1$ theorems have arguments, sometimes subtle ones, about telescoping sums which collapse. These arguments are systematically handled herein with the stopping data, as opposed to more intricate Carleson measure arguments.

The use of the energy stopping intervals is motivated by the use of the corresponding intervals, under the pivotal condition (20), in [38, 56]. However, the pivotal condition is not necessary for the two weight inequality, while the energy inequality is necessary from the A_2 and interval testing conditions.

Initial arguments had largely ignored the structure of the pair of functions f, g in the inner product $\langle H_\sigma f, g\rangle_w$, instead concentrating on proving an intricate series of Carleson measure type estimates. This changed with the argument of [21], which introduced Calderón-Zygmund stopping intervals, and the quasi-orthogonality argument into the subject. It was only then that the role of the global to local step was identified, but not proved. Stopping data also allows us to avoid the subtle problem of *absence of canonical paraproducts*. Attempts to introduce them induce *ad hoc* elements into the proof.

This section begins with the elementary and familiar Lemma 4.1, and then argues that the control of the triangular form $B^{\text{above}}(f, g)$ splits into the 'global to local' and the 'local' part. The authors of [23] only had the first reduction. And, using the techniques of that paper, could prove

Theorem D ([23]) *There holds* $|B^{above}(f, g)| \lesssim \{\mathcal{H} + \mathcal{B}_\infty\}\|f\|_\sigma \|g\|_w$, *where* $\mathcal{H} = \mathcal{A}_2^{1/2} + \mathcal{T}$, *and the remaining constant is the best constant in*

$$|B^{above}(f, g)| \lesssim \mathcal{B}_\infty \sigma(I_0)^{1/2} \|g\|_w,$$

where $|f| \le \mathbf{1}_{I_0}$, *and* I_0 *is any interval. The corresponding estimate holds for the dual from* $B^{below}(f, g)$.

This is a powerful Theorem, strongly suggesting that the A_2 condition and testing the Hilbert transform over bounded functions is sufficient for the L^2 boundedness of H_σ. But, there is no obvious way to deduce such a result from the Theorem above. Phrasing things differently, it can be very difficult to translate partial information about the triangular form $B^{\text{above}}(f, g)$ to information about $\langle H_\sigma f, g\rangle_w$, a potentially serious obstacle if a richer theory of two weight inequalities for singular integrals is to be developed.

The *parallel corona* was introduced in [22] to surmount this obstacle. With it, the result that could be proved the first real variable characterization of the two weight inequality for any continuous singular integral.

Theorem E (Lacey Sawyer Shen Uriarte-Tuero [22]) *There holds* $\mathcal{N} \simeq \mathcal{A}_2^{1/2} + \mathcal{T}_\infty$, *where the latter constant is the best constant in the inequalities below, uniform over all intervals I, and Borel subsets* $E \subset I$.

$$\int_I |H_\sigma \mathbf{1}_E|^2 \, dw \le \mathcal{T}_\infty^2 \sigma(I), \qquad \int_I |H_w \mathbf{1}_E|^2 \, d\sigma \le \mathcal{T}_\infty^2 w(I).$$

(One tests the Hilbert transform on $\mathbf{1}_E$*, but only the weight of the interval I appears on the right.)*

The parallel corona delays the application of Lemma 4.1, this feature combined with a special function theory specific to Haar expansions for non-doubling measures, were the critical ingredients.

The parallel corona has been used to give short transparent proofs of two weight inequalities for singular integrals. See the last page of Hytönen's survey [14] and the article of Tanaka [55].

It is natural to wonder if there are any L^p analogs of the main Theorem. We have some clues as to how this might work, in the more complicated testing conditions of Vuorinen [57, 58]. One could see that the global to local reduction would work under variants of these more complicated testing conditions. The control of the local term is however a heavily Hilbertian argument, and so potentially very difficult to extend to an L^p-setting.

The Remaining Part of the Global Estimate

The last part of the global-to-local part of the arugment is this Lemma.

Lemma 5.1 *Using the notation of section "Global to Local Reduction", there holds*

$$\Bigl\lvert \sum_{F\in\mathcal{F}} \sum_{I\,:\,I\supsetneq F} \mathbb{E}^{\sigma}_{I_F} \Delta^{\sigma}_{I} f \cdot \langle H_\sigma (I_F\setminus F), g_F\rangle_w \Bigr\rvert \lesssim \mathcal{H}\lVert f\rVert_\sigma \lVert g\rVert_w. \tag{35}$$

Our method of proof has these elements. (a) Use monotonicity to pass to a positive operator. (b) Identify the inequality needed as an instance of a two weight inequality, but not for general functions, only one fixed function, and a derived weight $\mu = \mu_{\sigma,w,f}$ that is well-adapted to the function; (c) Invoke the *parallel corona* method to prove the desired two weight inequality. Along the way, we will identify simplifications of the general case of a two weight inequality for a positive operator.

Begin the proof by observing that

$$\Bigl\lvert \sum_{I\,:\,I\supsetneq F} \mathbb{E}^{\sigma}_{I_F} \Delta^{\sigma}_{I} f \cdot (I_F - F)\Bigr\rvert \lesssim \Phi \coloneqq \sum_{\substack{F'\in\mathcal{F}\\ F'\supsetneq F}} \alpha_f(F')\cdot \overline{F}'_F, \qquad F\in\mathcal{F}.$$

where $\overline{F}'_F = F'\setminus F''$, with F'' being the $\mathcal{F}$-child of F' that contains F. Also, by monotonicity, the left-side of (35) is at most

$$\sum_{F\in\mathcal{F}} \sum_{J^*\in\mathcal{J}^*(F)} P\Bigl(\sum_{\substack{F'\in\mathcal{F}\\ F'\supsetneq F}} \alpha_f(F')\cdot \overline{F}'_F, J^*\Bigr) \sum_{\substack{J\,:\,J\subset J^*\\ \dot{\pi}_{\mathcal{F}} J = J^*}} \Bigl\langle \frac{x}{\lvert J^*\rvert}, h^{\sigma}_J \Bigr\rangle_\sigma \lvert\langle g, h^w_J\rangle_w\rvert.$$

The desired estimate is a consequence of new L^2-estimate for the modified Poisson operator

$$\tilde{P}f(x,t) = \int \frac{f(y)}{t^2+(x-y)^2}\, dy \tag{36}$$

which is extended to $\tilde{P}f(I) = Pf(x_I, |I|)$. The relevant measure on the upper half-plane is given by

$$\mu := \sum_{F \in \mathcal{F}} \sum_{J^* \in \mathcal{J}^*(F)} \delta_{x_{J^*}, |J^*|} \sum_{\substack{J\,:\,J \subset J^* \\ \dot{\pi}_{\mathcal{F}} J = J^*}} \langle x, h_J^w \rangle_w^2 . \tag{37}$$

Finally, the estimate we need is as below, in which we have eliminated the sum of $J \in \mathcal{J}^*$.

$$\Bigg\| \sum_{\substack{F' \in \mathcal{F} \\ F \in \mathrm{Ch}_{\mathcal{F}}(F')}} \alpha_f(F') \sum_{J \in \mathcal{J}^*(F)} \tilde{P}(\overline{F}'_F, J) \cdot Q_J \Bigg\|_{\mu} \lesssim \mathcal{H} \|f\|_{\sigma} . \tag{38}$$

Here, $\mathrm{Ch}_{\mathcal{F}}(F')$ is the collection of $\mathcal{F}$-children of F', and $Q_J = J \times [0, |K|]$ is the Carleson box over interval J.

This last inequality is in fact *universal*, in that we could fix the measuer μ, replace f by an arbitrary function, and the inequality is still true. But this fact is not needed. And, we can use the fact that f and the measure μ are related through the stopping data, to simplify the proof of (38).

Our knowledge of two weight estimates suggest that the inequality (38) is easiest to prove by duality, and using the joint stopping data on f and the dual function $\gamma \in L^2(\mathbb{R}^2_+, \mu)$, a technique refered to as the *parallel corona.* We will reduce the inequality (38) to two testing inequalities. One will be a reformulation of the energy inequality and the other will be a consequence of the A_2 condition.

By duality, the inequality we establish is

$$\sum_{\substack{F' \in \mathcal{F} \\ F \in \mathrm{Ch}_{\mathcal{F}}(F')}} \alpha_f(F') \sum_{J \in \mathcal{J}^*(F)} \tilde{P}(\overline{F}'_F, J) \int_{Q_J} \gamma \, d\mu \lesssim \mathcal{H} \|f\|_{\sigma} \|\gamma\|_{\mu} . \tag{39}$$

Here, γ is a non-negative function, supported on a Carleson cube Q_{J_0}, where $J_0 \in \mathcal{J}^* := \bigcup_{F \in \mathcal{F}} \mathcal{J}^*(F)$. We construct stopping intervals $\mathcal{G}$ for γ, by initializing $\mathcal{G} = \{J_0\}$, and setting $\alpha_{J_0}(g) = \mathbb{E}^{\mu}_{Q_{J_0}} \gamma$. In the recursive step, for minimal $J \in \mathcal{G}$, we add to $\mathcal{G}$ the maximal subintervals $J' \subsetneq J$ with $J' \in \mathcal{J}^*$ such that $\alpha_g(J') := \mathbb{E}^{\mu}_{Q_{J'}} \gamma > 10 \alpha_g(J)$. We let $\pi_{\mathcal{G}} I$ be the minimal element of G that contains I.

Now, in the sum (39), a given interval J that occurs satisfies either $\pi_{\mathcal{G}} J \subset F'$ or $F' \subsetneq \pi_{\mathcal{G}} J$. (Keep in mind that there could be many intervals $G \in \mathcal{G}$ that lie between J and F'.) This division splits the sum into two terms, the first is the sum over $F' \in \mathcal{F}$ of

$$\alpha_f(F') \sum_{F \in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J \in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J \subset F'}} \tilde{P}(\overline{F}'_F, J) \int_{Q_J} \gamma \, d\mu . \tag{40}$$

And the second is sum over $G \in \mathcal{G}$ of

$$\sum_{\substack{F' \in \mathcal{F} \\ \pi_{\mathcal{G}} F' = G \\ F' \neq G}} \alpha_f(F') \sum_{F \in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J \in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J = G}} \tilde{P}(\overline{F}'_F, J) \int_{Q_J} \gamma \, d\mu. \tag{41}$$

The first testing inequality is this inequality, uniform over $F' \in \mathcal{F}$.

$$(40) \lesssim \mathcal{H} \alpha_f(F') \sigma(F')^{1/2} \Bigg[\sum_{\substack{G \in \mathcal{G} \\ \pi_{\mathcal{F}} G = F'}} \alpha_\gamma(G)^2 \mu(Q_G) \Bigg]^{1/2}. \tag{42}$$

That this completes the bound of (40) is an immediate consequence of quasi-orthogonality. By Cauchy-Schwarz applied to the right of (42), note that

$$\sum_{F \in \mathcal{F}} \alpha_f(F')^2 \sigma(F') \lesssim \|f\|_\sigma^2,$$

and as well, by the construction of the stopping data for γ,

$$\sum_{F \in \mathcal{F}} \sum_{\substack{G \in \mathcal{G} \\ \pi_{\mathcal{F}} G = F'}} \alpha_\gamma(G)^2 \mu(Q_G) \lesssim \|\gamma\|_\mu^2.$$

This completes half of the proof of (39). The other half follows from the second testing inequality: Uniformly in $G \in \mathcal{G}$, there holds

$$(41) \lesssim \mathcal{H} \alpha_\gamma(G) \mu(Q_G)^{1/2} \Bigg[\sum_{\substack{F' \in \mathcal{F} \\ \pi_{\mathcal{G}} F' = G}} \alpha_f(F')^2 \sigma(F') \Bigg]^{1/2}. \tag{43}$$

It is bounded again by quasi-orthogonality. It remains to prove the two testing inequalities (42) and (43).

Proof of (42) This is just the energy inequality. By construction

$$(40) \lesssim \alpha_f(F') \sum_{F \in \mathrm{Ch}_{\mathcal{F}}(F)} \sum_{\substack{J \in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J \subset F'}} \alpha_\gamma(\pi_{\mathcal{G}} J) \tilde{P}_\sigma(F' \setminus F, J) \mu(\{(x_J, |J|)\}).$$

Of course we use Cauchy-Schwarz on the right above. Recall the definition of μ, to see that this inequality

$$\sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F)} \sum_{\substack{J\in \mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J\subset F'}} \tilde{P}_\sigma(F'\setminus F,J)^2\mu(\{(x_J,|J|)\})$$

$$\leq \sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F)} \sum_{\substack{J\in \mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J\subset F'}} \tilde{P}_\sigma(F'\setminus F,J)^2 \sum_{J'\,:\,J'\subset J} \langle x, h^w_{J'}\rangle_w^2 \lesssim \mathcal{H}^2\sigma(F')$$

is simply a reformulation of the energy inequality (19).

The other part of the application of Cauchy-Schwarz is

$$\sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F)} \sum_{\substack{J\in \mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J\subset F'}} \alpha_\gamma(\pi_{\mathcal{G}}J)^2\mu(\{(x_J,|J|)\}) \lesssim \sum_{\substack{G\in\mathcal{G}\\ \pi_{\mathcal{F}}G\in\{F'\}\cup \mathrm{Ch}_{\mathcal{F}}(F')}} \alpha_\gamma(\pi_{\mathcal{G}}J)^2\mu(Q_G).$$

This completes the proof of (42). □

Proof of (43) In (41), we dominate $\int_{Q_J}\gamma \;\; d\mu \;\leq\; 10\mathbb{E}^\mu_G\gamma\cdot\mu(Q_J)$, and then express (41) using the dual to the operator $\tilde{P}$ defined in (36). We have

$$(41)\lesssim \mathbb{E}^\mu_G\gamma\times \sum_{\substack{F'\in\mathcal{F}\\ \pi_{\mathcal{G}}F'=G\\ F'\neq G}} \alpha_f(F') \sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in\mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J=G}} \tilde{P}(\overline{F}'_F,J)\mu(Q_J) \tag{44}$$

$$=\mathbb{E}^\mu_G\gamma\times\int_G \sum_{\substack{F'\in\mathcal{F}\\ \pi_{\mathcal{G}}F'=G\\ F'\neq G}} \sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \alpha_f(F')\cdot F'_F \sum_{\substack{J\in\mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J=G}} \tilde{P}^*_\mu(Q_J)\,d\sigma \tag{45}$$

Apply Cauchy-Schwartz in the variable F', and $L^2(\sigma)$. One of the terms that result is

$$\sum_{\substack{F'\in\mathcal{F}\\ \pi_{\mathcal{G}}F'=G}} \alpha_f(F')^2\sigma(F').$$

Compare to the right side of (43). The other term is the following inequality, holding uniformly in $G\in\mathcal{G}$:

$$\int_G \sum_{\substack{F'\in\mathcal{F}\\ \pi_{\mathcal{G}}F'=G\\ F'\neq G}} \Bigg[\sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in\mathcal{J}^*(F)\\ \pi_{\mathcal{G}}J=G}} F'_F\cdot\tilde{P}^*_\mu(Q_J)\Bigg]^2 d\sigma \lesssim \mathcal{A}_2\mu(Q_G). \tag{46}$$

As the inequlaity shows, this follows from the A_2 condition.

An obstacle to a proof is that the sets Q_J overlap. This is addressed with the definition $W^j_J = J\times(2^{-j-1},2^{-j}]$, for $j\geq 0$. These sets are disjoint in j and J. We will then show that for each $F'\in\mathcal{F}$,

$$\int_{\mathbb{R}} \Bigg[\sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J = G}} F'_F \cdot \tilde{P}^*_{\mu}(W^j_J) \Bigg]^2 \, d\sigma \lesssim 2^{-j} \mathcal{A}_2 \sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J = G}} \mu(W^j_J). \tag{47}$$

This easily implies (46).

Two additional summing variables are convenient. For integers $k \geq r$, we restrict the sum to $J \in \mathcal{J}^*(F)$ with $2^k|J| = |F|$. And, for integers $\ell \geq k(1-\epsilon)$, we further require that

$$2^{\ell-1}|J| \leq \mathrm{dist}(J, \partial F) < 2^{\ell}|J|. \tag{48}$$

By goodness, $\ell \geq k(1-\epsilon)$, but there is in general no other condition that we have here. Then, we prove this estimate, which is (48), with these two additional restrictions on J. Uniformly in $F' \in \mathcal{F}$,

$$\int_{\mathbb{R}} \Bigg[\sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J = G \\ 2^k|J|=|F|,\ (48)\text{ holds}}} F'_F \cdot \tilde{P}^*_{\mu}(W^j_J) \Bigg]^2 \, d\sigma \lesssim 2^{-j-k-\ell} \mathcal{A}_2 \sum_{F\in \mathrm{Ch}_{\mathcal{F}}(F')} \sum_{\substack{J\in \mathcal{J}^*(F) \\ \pi_{\mathcal{G}} J = G}} \mu(W^j_J). \tag{49}$$

In (49), there are at most 2^{ℓ} intervals J. We can therefore pass the square inside the sum, at cost of a factor of 2^{ℓ}. But,

$$\int_{F'_F} \tilde{P}^*_{\mu}(W^j_J)(x)^2 \, d\sigma(x) \lesssim \mu(W^j_J) \int_{F'_F} \int_{W^j_J} \frac{1}{[y_2^2 + |x-y_1|^2]^2} \, d\mu(y)\, d\sigma(x) \tag{50}$$

$$\lesssim \frac{\mu(W^j_J)^2}{2^{2\ell}|J|^2} \tilde{P}_{\sigma}(F'_F, J) \tag{51}$$

$$\lesssim 2^{-2\ell-2j} \mu(W^j_J) w(J) \tilde{P}_{\sigma}(F'_F, J) \lesssim 2^{-2\ell-2j} \mathcal{A}_2 \mu(W^j_J). \tag{52}$$

Here, we have used Cauchy-Schwartz, followed by the estimate below, which holds for $x \in F'_F$,

$$\int_{W^j_J} \frac{1}{[y_2^2 + |x-y_1|^2]^2} \, d\mu(y) \lesssim \frac{\mu(W^j_J)}{2^{2\ell}|J|^2(|J|^2 + |x-x_J|^2)}$$

Then, besides disjointness, the sets W_J^j enjoy the estimate $\mu(W_J^j) \lesssim 2^{-2j}|J|^2 w(J)$, which follows from the definition of μ in (37), and the estimate $|x, h_J^w|^2 \le |J|^2 w(J)$. Finally, we just appeal to the $\mathcal{A}_2$ condition. The bound in (52) is multiplied by 2^ℓ, to prove (49). This finishes the proof. □

Context and Discussion

The inequality (38) is universal. This was first proved in [23], in the case that the weights did not share a common point mass. It was down by appealing to the Sawyer theorem [53] on two weight inequalities for the Poisson operator. This technique does not allow common point masses, however. Addressing this, Hytönen [15] found a clever way to use dyadic approximates to the 'Poisson operator with holes,' by using dyadic approximates to an arbitrary interval, and prinving a novel dyadic two weight inequality.

The proof herein does not attempt to prove the *universal* form of (38). Indeed, this inequality is not needed. Indeed, the close relationship between the function f, and the derived measure μ in (37) permits a short self-contained proof.

The Stopping Form

The last step in the proof of Theorem 4.2, hence in the proof of the main theorem, is to show that the local inequality (32) holds. Using the discussion at the end of the previous section, this amounts to controlling the *stopping form*. Given an interval $F \in \mathcal{F}$, the stopping form is

$$B_F^{\text{stop}}(f,g) := \sum_{I\,:\,\pi_{\mathcal{F}} I = F} \ \sum_{J\,:\,J \Subset I_J, \dot{\pi}_{\mathcal{F}} J = F} \mathbb{E}_{I_J}^{\sigma} \Delta_I^{\sigma} f \cdot \langle H_\sigma(I_0 - I_J), \Delta_J^w g\rangle_w. \tag{53}$$

Lemma 6.1 *There holds for each $F \in \mathcal{F}$,*

$$|B_{I_0}^{stop}(f,g)| \lesssim \mathcal{H} \|P_F^\sigma f\|_\sigma \|Q_F^w g\|_w. \tag{54}$$

The stopping form arises naturally in any proof of a $T1$ theorem using Haar or other bases. In the non-homogeneous case, or in the Tb setting, where (adapted) Haar functions are important tools, it frequently appears in more or less this form. Regardless of how it arises, the stopping form is treated as a error, in that it is bounded by some simple geometric series, obtaining decay as e. g. the ratio $|J|/|I|$ is held fixed. (See for instance [38] (7.16).)

These sorts of arguments, however, implicitly require some additional hypotheses, such as the weights being mutually A_∞. Of course, the two weights above can be mutually singular. There is no *a priori* control of the stopping form in terms of simple parameters like $|J|/|I|$, even supplemented by additional pigeonholing of various parameters.

Our method is inspired by proofs of Carleson's Theorem on Fourier series [5, 10, 25], and has one particular precedent in the current setting, a much simpler bound for the stopping form in [22].

Admissible Pairs

We can assume that $f = P^{\sigma}_{F} f$ and $g = Q^{w}_{F} g$. For all pair of intervals $J \Subset I \subset F$ that we need to consider, we have $\dot{\pi}_{\mathcal{F}} J = F$, and hence by the Energy Stopping condition, there holds

$$P(\sigma(F - I_J), I_J) E(w, I_J)^2 w(I_J) \le C\sigma(I_J). \tag{55}$$

For if not, by monotonicity (16), we would have that the interval I_J would be an energy stopping interval, hence $I_J \in \mathcal{F}$, and $\dot{\pi}_{\mathcal{F}} J = I_J$. It is this condition that is our starting point for the recursion.

A range of decompositions of the stopping form necessitate a somewhat heavy notation that we introduce here. The individual summands in the stopping form involve four distinct intervals, namely F, I, I_J, and J. The interval F will not change in this argument, and the pair (I, J) determine I_J. Subsequent decompositions are easiest to phrase as actions on collections $\mathcal{Q}$ of pairs of intervals $Q = (Q_1, Q_2)$ with $F \supset Q_1 \Supset Q_2$. (The letter P is already taken for the Poisson integral.) And we consider the bilinear forms

$$B_{\mathcal{Q}}(f, g) := \sum_{Q \in \mathcal{Q}} \mathbb{E}^{\sigma}_{(Q_1)_{Q_2}} \Delta^{\sigma}_{Q_1} f \cdot \langle H_\sigma(F - (Q_1)_{Q_2}), \Delta^{w}_{Q_2} g \rangle_w .$$

We will have the standing assumption that for all collections $\mathcal{Q}$ that we consider are *admissible*.

Definition 6.2 A collection of pairs $\mathcal{Q}$ is *admissible* if it meets these criteria. For any $Q = (Q_1, Q_2) \in \mathcal{Q}$,

(1) $Q_2 \Subset Q_1 \subset F$, and both Q_1 and Q_2 are good.
(2) (convexity in Q_1) If $Q'' \in \mathcal{Q}$ with $Q''_2 = Q_2$ and $Q''_1 \subset I \subset Q_1$, with I good, then there is a $Q' \in \mathcal{Q}$ with $Q'_1 = I$ and $Q'_2 = Q_2$.

The first property is self-explanatory. The second property is convexity in Q_1, subject to goodness, holding Q_2 fixed, which is used in the estimates on the stopping form which conclude the argument. A third property is described below.

We exclusively use the notation $\mathcal{Q}_k$, $k = 1, 2$ for the collection of intervals $\bigcup\{Q_k : Q \in \mathcal{Q}\}$, not counting multiplicity. Similarly, set $\tilde{\mathcal{Q}}_1 := \{(Q_1)_{Q_2} : Q \in \mathcal{Q}\}$, and $\tilde{Q}_1 := (Q_1)_{Q_2}$.

(3) Every interval $Q_2 \in \mathcal{Q}_2$ satisfies $\dot{\pi}_{\mathcal{F}} Q_2 = F$ (And so, every $\tilde{\mathcal{Q}}_1$ has $\mathcal{F}$-parent F.)

The last requirement comes from the assumption that the functions f and g be adapted to $\mathcal{F}_{\text{energy}}(F)$. We will be appealing to different Hilbertian arguments below, so we prefer to make this an assumption about the pairs rather than the functions f, g. The Hilbert space will be the space of good functions in $L^2(\sigma)$ and $L^2(w)$.

Typically, one only ever needs goodness of the *small* interval, in this case Q_2. We will use the term size$(\mathcal{Q})$ below, in which it will be apparent that goodness of the intervals Q_1 will be helpful. Namely, at this point goodness is used to as in the monotonicity principle, to estimate off-diagonal inner products involving the Hilbert transform by Poisson averages, and to regularize Poisson averages. Both are made more explicit in section "Upper Bounds on the Stopping Form".

The stopping form is obtained with the admissible collection of pairs given by

$$\mathcal{Q}_0 = \{(I, J) : J \Subset I \subset F, I \text{ and } J \text{ are good}, \ \dot{\pi}_{\mathcal{F}} J = F\}. \tag{56}$$

There holds $B_F^{\text{stop}}(f, g) = B_{\mathcal{Q}_0}(f, g)$.

There is a very important notion of the size of $\mathcal{Q}$.

$$\operatorname{size}(\mathcal{Q})^2 := \sup_{K \in \tilde{\mathcal{Q}}_1 \cup \mathcal{Q}_2} \frac{\mathsf{P}(\sigma(F - K), K)^2}{\sigma(K)|K|^2} \sum_{J \in \mathcal{Q}_2 : J \subset K} \langle x, h_J^w \rangle_w^2. \tag{57}$$

For admissible $\mathcal{Q}$, there holds size$(\mathcal{Q}) \lesssim \mathcal{H}$, as follows (55).

More definitions follow. Set the norm $\mathbf{B}_{\mathcal{Q}}$ of the bilinear form $\mathcal{Q}$ to be the best constant in the inequality

$$|B_{\mathcal{Q}}(f, g)| \le \mathbf{B}_{\mathcal{Q}} \|f\|_\sigma \|g\|_w.$$

Thus, our goal is show that $\mathbf{B}_{\mathcal{Q}} \lesssim$ size$(\mathcal{Q})$ for admissible $\mathcal{Q}$, but we will only be able to do this directly in the case that the pairs (Q_1, Q_2) are weakly decoupled in a collection $\mathcal{Q}$. The relevant decoupling is precisely described in section "Upper Bounds on the Stopping Form".

Say that collections of pairs $\mathcal{Q}^j$, for $j \in \mathbb{N}$, are *mutually orthogonal* if on the one hand, the collections $(\mathcal{Q}^j)_2$, of second coordinates of the pairs, are pairwise disjoint, and on the other, that the collections $\widetilde{(\mathcal{Q}^j)}_1$ are pairwise disjoint. The concept has to be different in the first and second coordinates of the pairs, due to the different role of the intervals $\tilde{Q}_1$ and Q_2, which comes up again in the next paragraph.

The meaning of mutual orthogonality is best expressed through the norm of the associated bilinear forms. Under the assumption that $B_{\mathcal{Q}} = \sum_{j \in \mathbb{N}} B_{\mathcal{Q}^j}$, and that the $\{\mathcal{Q}^j : j \in \mathbb{N}\}$ are mutually orthogonal, the following essential inequality holds.

$$\mathbf{B}_{\mathcal{Q}} \leq \sqrt{2} \sup_{j\in\mathbb{N}} \mathbf{B}_{\mathcal{Q}^j}. \tag{58}$$

Indeed, for $j \in \mathbb{N}$, let Π_j^w be the projection onto the linear span of the Haar functions $\{h_J^w : J \in \mathcal{Q}_2^j\}$, and use a similar notation for Π_j^σ. We then have the two inequalities

$$\sum_{j\in\mathbb{N}} \|\Pi_j^w g\|_w^2 \leq \|g\|_w^2, \qquad \sum_{j\in\mathbb{N}} \|\Pi_j^\sigma f\|_\sigma^2 \leq 2\|f\|_\sigma^2.$$

Since a given interval I can be in two collections $\mathcal{Q}_1^j$, we have the factor of 2 in the second inequality. Therefore, we have

$$\begin{aligned} |B_{\mathcal{Q}}(f,g)| &\leq \sum_{j\in\mathbb{N}} |B_{\mathcal{Q}^j}(f,g)| \\ &= \sum_{j\in\mathbb{N}} |B_{\mathcal{Q}^j}(\Pi_j^\sigma f, \Pi_j^w g)| \\ &\leq \sum_{j\in\mathbb{N}} \mathbf{B}_{\mathcal{Q}^j} \|\Pi_j^\sigma f\|_\sigma \|\Pi_j^w g\|_w \leq \sqrt{2} \sup_{j\in\mathbb{N}} \mathbf{B}_{\mathcal{Q}^j} \cdot \|f\|_\sigma \|g\|_w. \end{aligned}$$

This proves (58).

The Recursive Argument

This is the essence of the matter.

Lemma 6.3 (Size Lemma) *An admissible collection of pairs $\mathcal{Q}$ can be partitioned into collections $\mathcal{Q}^{large}$ and admissible $\mathcal{Q}_t^{small}$, for $t \in \mathbb{N}$ such that*

$$\mathbf{B}_{\mathcal{Q}} \leq C size(\mathcal{Q}) + (1+\sqrt{2}) \sup_t \mathbf{B}_{\mathcal{Q}_t^{small}}, \tag{59}$$

$$\textit{and} \quad \sup_{t\in\mathbb{N}} size(\mathcal{Q}_t^{small}) \leq \tfrac{1}{4} size(\mathcal{Q}). \tag{60}$$

Here, $C > 0$ is an absolute constant.

The point of the lemma is that all of the constituent parts are better in some way, and that the right hand side of (59) involves a favorable supremum. We can quickly prove the main result of this section.

Proof of Lemma 6.1 The stopping form of this Lemma is of the form $B_{\mathcal{Q}}(f,g)$ for admissible choice of $\mathcal{Q}$, with size$(\mathcal{Q}) \leq C\mathcal{H}$, as we have noted in (56). Define

$$\zeta(\lambda) := \sup\{\mathbf{B}_{\mathcal{Q}} : \text{size}(\mathcal{Q}) \leq C\lambda\mathcal{H}\}, \qquad 0 < \lambda \leq 1,$$

where $C > 0$ is a sufficiently large, but absolute constant, and the supremum is over admissible choices of $\mathcal{Q}$. We are free to assume that $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are further constrained to be in some fixed, but large, collection of intervals $\mathcal{I}$. Then, it is clear that $\zeta(\lambda)$ is finite, for all $0 < \lambda \leq 1$. Because of the way the constant $\mathcal{H}$ enters into the definition, it remains to show that $\zeta(1)$ admits an absolute upper bound, independent of how $\mathcal{I}$ is chosen.

It is the consequence of Lemma 6.3 that there holds

$$\zeta(\lambda) \leq C\lambda + (1+\sqrt{2})\zeta(\lambda/4), \qquad 0 < \lambda \leq 1. \tag{61}$$

Iterating this inequality beginning at $\lambda = 1$ gives us

$$\zeta(1) \leq C + (1+\sqrt{2})\zeta(1/4) \leq \cdots \leq C \sum_{t=0}^{\infty} \left[\tfrac{1+\sqrt{2}}{4}\right]^t \leq 4C.$$

So we have established an absolute upper bound on $\zeta(1)$. □

Proof of Lemma 6.3

We restate the conclusion of Lemma 6.3 to more closely follow the line of argument to follow. The collection $\mathcal{Q}$ can be partitioned into two collections $\mathcal{Q}^{\mathrm{large}}$ and $\mathcal{Q}^{\mathrm{small}}$ such that

(1) $\mathbf{B}_{\mathcal{Q}^{\mathrm{large}}} \lesssim \tau$, where $\tau := \mathrm{size}(\mathcal{Q})$.
(2) $\mathcal{Q}^{\mathrm{small}} = \mathcal{Q}_1^{\mathrm{small}} \cup \mathcal{Q}_2^{\mathrm{small}}$.
(3) The collection $\mathcal{Q}_1^{\mathrm{small}}$ is admissible, and $\mathrm{size}(\mathcal{Q}_1^{\mathrm{small}}) \leq \frac{\tau}{4}$.
(4) For a collection of dyadic intervals $\mathcal{L}$, the collection $\mathcal{Q}_2^{\mathrm{small}}$ is the union of mutually orthogonal admissible collections $\mathcal{Q}_{2,L}^{\mathrm{small}}$, for $L \in \mathcal{L}$, with

$$\mathrm{size}(\mathcal{Q}_{2,L}^{\mathrm{small}}) \leq \tfrac{\tau}{4}, \qquad L \in \mathcal{L}.$$

Thus, we have by inequality (58) for mutually orthogonal collections,

$$\begin{aligned}\mathbf{B}_{\mathcal{Q}} &\leq \mathbf{B}_{\mathcal{Q}^{\mathrm{large}}} + \mathbf{B}_{\mathcal{Q}_1^{\mathrm{small}} \cup \mathcal{Q}_2^{\mathrm{small}}} \\ &\leq \mathbf{B}_{\mathcal{Q}^{\mathrm{large}}} + \mathbf{B}_{\mathcal{Q}_1^{\mathrm{small}}} + \mathbf{B}_{\mathcal{Q}_2^{\mathrm{small}}} \\ &\leq C\tau + (1+\sqrt{2}) \max\left\{\mathbf{B}_{\mathcal{Q}_1^{\mathrm{small}}}, \sup_{L \in \mathcal{L}} \mathbf{B}_{\mathcal{Q}_{2,L}^{\mathrm{small}}}\right\}.\end{aligned}$$

This, with the properties of size listed above prove Lemma 6.3 as stated, after a trivial re-indexing.

In a manner similar to the argument of section "Global to Local Reduction", there is an induced measure on the upper half-plane that is relevant to our considerations.

This time it is given by

$$\mu_{\mathcal{Q}} = \mu := \sum_{J \in \mathcal{Q}_2 \,:\, J \subset F} \langle x, h_J^w \rangle_w^2 \delta_{(x_J, |J|)}, \qquad x_J \text{ is the center of } J.$$

The tent over L is the triangular region $T_L := \{(x, y) \ : \ |x - x_L| \leq |L| - y\}$, so that

$$\mu(T_L) = \sum_{J \in \mathcal{Q}_2 \,:\, J \subset L} \langle x, h_J^w \rangle_w^2.$$

Observe that

$$\text{size}(\mathcal{Q})^2 = \sup_{K \in \tilde{\mathcal{Q}}_1 \cup \mathcal{Q}_2} \frac{\mathsf{P}(\sigma(F - K), K)^2}{\sigma(K)|K|^2} \mu(T_K).$$

All else flows from this construction of a subset $\mathcal{L}$ of dyadic subintervals of F. The initial intervals in $\mathcal{L}$ are the minimal intervals $L \in \tilde{\mathcal{Q}}_1 \cup \mathcal{Q}_2$ such that

$$\frac{\mathsf{P}(\sigma(F - L), L)^2}{|L|^2} \mu(T_L) \geq \frac{\tau^2}{16} \sigma(L). \tag{62}$$

Since $\text{size}(\mathcal{Q}) = \tau$, there are such intervals L.

Initialize $\mathcal{S}$ (for ‘stock’ or ‘supply’) to be all the dyadic intervals in $\tilde{\mathcal{Q}}_1 \cup \mathcal{Q}_2$ which strictly contain some interval in $\mathcal{L}$. In the recursive step, let $\mathcal{L}'$ be the minimal elements $S \in \mathcal{S}$ such that

$$\mu(T_S) \geq \rho \sum_{\substack{L \in \mathcal{L} \,:\, L \subset S \\ L \text{ is maximal}}} \mu(T_L), \qquad \rho = \tfrac{17}{16}. \tag{63}$$

(The inequality would be trivial if $\rho = 1$.) If $\mathcal{L}'$ is empty the recursion stops. Otherwise, update $\mathcal{L} \leftarrow \mathcal{L} \cup \mathcal{L}'$, and $\mathcal{S} \leftarrow \{K \in \mathcal{S} \ : \ K \not\subset L \ \forall L \in \mathcal{L}\}$. See Fig. 7.

Once the recursion stops, report the collection $\mathcal{L}$. It has this crucial property: For $L \in \mathcal{L}$, and integers $t \geq 1$,

$$\sum_{L' \,:\, \pi_{\mathcal{L}}^t L' = L} \mu(T_{L'}) \leq \rho^{-t} \mu(T_L). \tag{64}$$

Indeed, in the case of $t = 1$, is a criteria for membership in $\mathcal{L}$, and a simple induction proves the statement for all $t \geq 1$.

The decomposition of $\mathcal{Q}$ is based upon the relation of the pairs to the collection $\mathcal{L}$, namely a pair $\tilde{Q}_1, Q_2$ can (a) both have the same parent in $\mathcal{L}$; (b) have distinct parents in $\mathcal{L}$; (c) Q_2 can have a parent in $\mathcal{L}$, but not $\tilde{Q}_1$; and (d) Q_2 does not have a parent in $\mathcal{L}$.

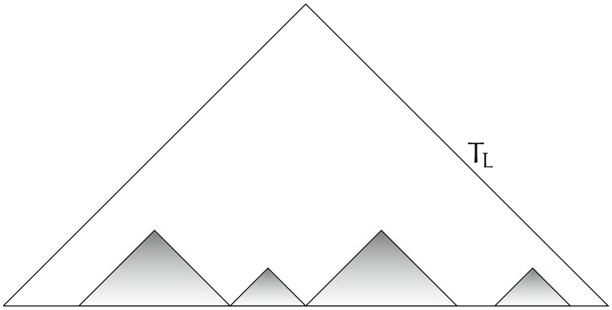


Fig. 7 The shaded smaller tents have been selected, and T_L is the minimal tent with $\mu(T_L)$ larger than ρ times the μ-measure of the shaded tents

A particularly vexing aspect of the stopping form is the linkage between the martingale difference on g, which is given by J, and the argument of the Hilbert transform, $F - I_J$. The 'large' collections constructed below will, in a certain way, decouple the J and the $F - I_J$, enough so that norm of the associated bilinear form can be estimated by the size of $\mathcal{Q}$.

In the 'small' collections, there is however no decoupling, but critically, the size of the collections is smaller, and we only have to estimate the maximal operator norm among the small collections.

Pairs Comparable to $\mathcal{L}$

Define

$$\mathcal{Q}_{L,t} := \{Q \in \mathcal{Q} \;:\; \pi_{\mathcal{L}} \tilde{Q}_1 = \pi^t_{\mathcal{L}} Q_2 = L\}, \qquad L \in \mathcal{L},\ t \in \mathbb{N}.$$

These are admissible collections, as the convexity property in Q_1, holding Q_2 constant, is clearly inherited from $\mathcal{Q}$. Now, observe that for each $t \in \mathbb{N}$, the collections $\{\mathcal{Q}_{L,t} \;:\; L \in \mathcal{L}\}$ are mutually orthogonal. The collection of intervals $(\mathcal{Q}_{L,t})_2$ are obviously disjoint in $L \in \mathcal{L}$, with $t \in \mathbb{N}$ held fixed. And, since membership in these collections is determined in the first coordinate by the interval $\tilde{Q}_1$, and the two children of Q_1 can have two different parents in $\mathcal{L}$, a given interval I can appear in at most two collections $(\widetilde{\mathcal{Q}_{L,t}})_1$, as $L \in \mathcal{L}$ varies, and $t \in \mathbb{N}$ held fixed.

Define $\mathcal{Q}_1^{\text{small}}$ to be the union over $L \in \mathcal{L}$ of the collections

$$\mathcal{Q}_{L,1}^{\text{small}} := \{Q \in \mathcal{Q}_{L,1} \;:\; \tilde{Q}_1 \neq L\}.$$

Note in particular that we have only allowed $t = 1$ above, and $\tilde{Q}_1 = L$ is not allowed. For these collections, we need only verify that

Lemma 6.4 *There holds*

$$\mathit{size}(\mathcal{Q}_{L,1}^{\mathit{small}}) \le \sqrt{(\rho-1)}\cdot\tau = \frac{\tau}{4}, \qquad L\in\mathcal{L},\ t\in\mathbb{N}. \tag{65}$$

Proof An interval $K \in \widetilde{(\mathcal{Q}_{L,1}^{\mathrm{small}})}_1 \cup \mathcal{Q}_2$ is not in $\mathcal{L}$, by construction. Suppose that K does not contain any interval in $\mathcal{L}$. By the selection of the initial intervals in $\mathcal{L}$, the minimal intervals in $\tilde{\mathcal{Q}}_1 \cup \mathcal{Q}_2$ which satisfy (62), it follows that the interval K must fail (62). And so we are done.

Thus, K contains some element of $\mathcal{L}$, whence the inequality (63) must fail. Namely, rearranging that inequality, and using the measure μ associated with $\mathcal{Q}_{L,1}^{\mathrm{small}}$,

$$\mu_{\mathcal{Q}_{L,1}^{\mathrm{small}}}(T_L) \le (\rho-1) \sum_{\substack{L'\in\mathcal{L}\,:\,L'\subset K\\ L' \text{ is maximal}}} \mu(T_L) \tag{66}$$

$$\le \frac{1}{16}\mu(T_L) \le \frac{\tau^2}{16}\cdot\frac{|K|^2\cdot\sigma(K)}{\mathsf{P}(\sigma(L-K),K)^2}. \tag{67}$$

Here, note that we begin with the measure $\mu_{\mathcal{Q}_{L,1}^{\mathrm{small}}}$; use $\rho = 1+\frac{1}{16}$; and the last inequality follows from the definition of size. This finishes the proof of (65). □

The collections below are the first contribution to $\mathcal{Q}^{\mathrm{large}}$. Take $\mathcal{Q}_1^{\mathrm{large}} := \bigcup\{\mathcal{Q}_{L,1}^{\mathrm{large}} : L\in\mathcal{L}\}$, where

$$\mathcal{Q}_{L,1}^{\mathrm{large}} := \{Q\in\mathcal{Q}_{L,1} \;:\; \tilde{Q}_1 = L\}.$$

Note that Lemma 6.8 applies to this Lemma, take the collection $\mathcal{S}$ of that Lemma to be $\{L\}$, and the quantity η in (79) satisfies $\eta \lesssim \tau = \mathrm{size}(\mathcal{Q})$, by (82). From the mutual orthogonality (58), we then have

$$\mathbf{B}_{\mathcal{Q}_1^{\mathrm{large}}} \le \sqrt{2}\sup_{L\in\mathcal{L}} \mathbf{B}_{\mathcal{Q}_{L,1}^{\mathrm{large}}} \lesssim \tau.$$

The collections $\mathcal{Q}_{L,t}$, for $L\in\mathcal{L}$, and $t\ge 2$ are the second contribution to $\mathcal{Q}^{\mathrm{large}}$, namely

$$\mathcal{Q}_2^{\mathrm{large}} := \bigcup_{L\in\mathcal{L}}\bigcup_{t\ge 2}\mathcal{Q}_{L,t}.$$

For them, we need to estimate $\mathbf{B}_{\mathcal{Q}_{L,t}}$.

Lemma 6.5 *There holds* $\mathbf{B}_{\mathcal{Q}_{L,t}} \lesssim \rho^{-t/2}\tau$.

From this, we can conclude from (58) that

$$\begin{aligned}\mathbf{B}_{\mathcal{Q}_2^{\text{large}}} &\le \sum_{t\ge 2} \mathbf{B}_{\bigcup\{\mathcal{Q}_{L,t} \,:\, L\in\mathcal{L}\}} \\ &\le \sqrt{2}\sum_{t\ge 2} \sup_{L\in\mathcal{L}} \mathbf{B}_{\mathcal{Q}_{L,t}} \lesssim \tau \sum_{t\ge 2} \rho^{-t/2} \lesssim \tau.\end{aligned}$$

Proof For $L \in \mathcal{L}$, let $\mathcal{S}_L$, the $\mathcal{L}$-children of L. For each $Q \in \mathcal{Q}_{L,t}$, we must have $Q_2 \subset \pi_{\mathcal{S}_L} Q_2 \subset \tilde{Q}_1$. Then, divide the collection $\mathcal{Q}_{L,t}$ into three collections $\mathcal{Q}^{\ell}_{L,t}$, $\ell = 1, 2, 3$, where

$$\begin{aligned}\mathcal{Q}^1_{L,t} &\coloneqq \{Q \in \mathcal{Q}_{L,t} \ : \ Q_2 \Subset \pi_{\mathcal{S}_L} Q_2\}, \\ \mathcal{Q}^2_{L,t} &\coloneqq \{Q \in \mathcal{Q}_{L,t} \ : \ Q_2 \not\Subset \pi_{\mathcal{S}_L} Q_2 \Subset \tilde{Q}_1\},\end{aligned}$$

and $\mathcal{Q}^3_{L,t} \coloneqq \mathcal{Q}_{L,t} - (\mathcal{Q}^1_{L,t} \cup \mathcal{Q}^2_{L,t})$ is the complementary collection. Notice that $\mathcal{Q}^1_{L,t}$ equals the whole collection $\mathcal{Q}_{L,t}$ for $t > r+1$.

We treat them in turn. The collections $\mathcal{Q}^1_{L,t}$ fit the hypotheses of Lemma 6.8, just take the collection of intervals $\mathcal{S}$ of that Lemma to be $\mathcal{S}_L$. It follows that $\mathbf{B}_{\mathcal{Q}^1_{L,t}} \lesssim \boldsymbol{\beta}(t)$, where the latter is the best constant in the inequality

$$\sum_{J\in(\mathcal{Q}_{L,t})_2 \,:\, J\Subset K} \mathsf{P}(\sigma(F-K),J)^2 \Big\langle \frac{x}{|J|}, h^w_J \Big\rangle_w^2 \le \boldsymbol{\beta}(t)^2 \sigma(K), \qquad K \in \mathcal{S}_L,\ L \in \mathcal{L},\ t \ge 2. \tag{68}$$

We will prove the estimate below, which is clearly summable in $t \in \mathbb{N}$ to the estimate we want.

Lemma 6.6 *There holds* $\boldsymbol{\beta}(t) \lesssim \rho^{-t/2}\tau$.

Proof We have the estimate without decay in t, $\boldsymbol{\beta}(t) \lesssim \text{size}(\mathcal{Q})$, as follows from (82). Use this estimate for $1 \le t \le r+3$, say. In the case of $t > r+3$, the essential property is (64). The left hand side of (68) is dominated by the sum below. Note that we index the sum first over L', which are $r+1$-fold $\mathcal{L}$-children of K, whence $L' \Subset K$, followed by $t-r-2$-fold $\mathcal{L}$-children of L'.

$$\sum_{\substack{L'\in\mathcal{L} \\ \pi^{r+1}_{\mathcal{L}} L' = K}} \sum_{\substack{L''\in\mathcal{L} \\ \pi^{t-r-2}_{\mathcal{L}} L'' = L'}} \sum_{J\in\mathcal{Q}_2 \,:\, J\subset L''} \mathsf{P}(\sigma(F-K),J)^2 \Big\langle \frac{x}{|J|}, h^w_J \Big\rangle_w^2 \tag{69}$$

$$\overset{(77)}{\lesssim} \sum_{\substack{L'\in\mathcal{L} \\ \pi^{r+1}_{\mathcal{L}} L' = K}} \frac{\mathsf{P}(\sigma(F-K),L')^2}{|L'|^2} \sum_{\substack{L''\in\mathcal{L} \\ \pi^{t-r-2}_{\mathcal{L}} L'' = L'}} \mu(T_{L''}) \tag{70}$$

$$\overset{(64)}{\lesssim} \rho^{-t+r+2} \sum_{\substack{L' \in \mathcal{L} \\ \pi_{\mathcal{L}}^{r+1} L' = K}} \frac{\mathsf{P}(\sigma(F-K), L')^2}{|L'|^2} \mu(T_{L'}) \tag{71}$$

$$\lesssim \rho^{-t} \tau^2 \sum_{\substack{L' \in \mathcal{L} \\ \pi_{\mathcal{L}}^{r+1} L' = K}} \sigma(L') \lesssim \tau^2 \rho^{-t} \sigma(K). \tag{72}$$

We have also used (77), and then the central property (64) following from the construction of $\mathcal{L}$, finally appealing to the definition of size. Hence, $\boldsymbol{\beta}(t) \lesssim \tau \rho^{-t/2}$. This completes the analysis of $\mathcal{Q}^1_{L,t}$. □

We need only consider the collections $\mathcal{Q}^2_{L,t}$ for $1 \le t \le r+1$, and they fall under the scope of Lemma 6.9. A variant of (82) shows that $\mathbf{B}_{\mathcal{Q}^2_{L,t}} \lesssim \tau$. Similarly, we need only consider the collections $\mathcal{Q}^3_{L,t}$ for $1 \le t \le r+1$. It follows that we must have $2^r \le |Q_1|/|Q_2| \le 2^{2r+2}$. Namely, this ratio can take only one of a finite number of values, implying that Lemma 6.10 applies easily to this case to complete the proof. □

Pairs Not Strictly Comparable to $\mathcal{L}$

It remains to consider the pairs $Q \in \mathcal{Q}$ such that $\tilde{Q}_1$ does not have a parent in $\mathcal{L}$. The collection $\mathcal{Q}^{\mathrm{small}}_2$ is taken to be the (much smaller) collection

$$\mathcal{Q}^{\mathrm{small}}_2 := \{Q \in \mathcal{Q} \,:\, Q_2 \text{ does not have a parent in } \mathcal{L}\}.$$

Observe that $\mathrm{size}(\mathcal{Q}^{\mathrm{small}}_2) \le \sqrt{(\rho-1)}\tau \le \frac{\tau}{4}$. This is as required for this collection. (The collections $\mathcal{Q}^{\mathrm{small}}_1$ and $\mathcal{Q}^{\mathrm{small}}_2$ are also mutually orthogonal, but this fact is not needed for our proof.)

Proof Suppose $\eta < \mathrm{size}(\mathcal{Q}^{\mathrm{small}}_2)$. Then, there is an interval $K \in \widetilde{(\mathcal{Q}^{\mathrm{small}}_1)}_1 \cup (\mathcal{Q}^{\mathrm{small}}_2)_2$ so that

$$\eta^2 \sigma(K) \le \frac{P(\sigma(F-K), K)^2}{|K|^2} \mu_{\mathcal{Q}^{\mathrm{small}}_2}(T_K).$$

Suppose that K does not contain any interval in $\mathcal{L}$. It follows from the initial intervals added to $\mathcal{L}$, see (62), that we must have $\eta \le \frac{\tau}{4}$.

Thus, K contains an interval in $\mathcal{L}$. This means that K must fail the inequality (63). Therefore, we have

$$\eta^2 \sigma(K) \le (\rho - 1) \frac{P(\sigma(F-K), K)^2}{|K|^2} \mu(T_K) \le \frac{\tau^2}{16} \sigma(K).$$

This relies upon the definition of size, and proves our claim. □

For the pairs not yet in one of our collections, it must be that Q_2 has a parent in $\mathcal{L}$, but not $\tilde{Q}_1$. Using $\mathcal{L}^*$, the maximal intervals in $\mathcal{L}$, divide them into the three collections

$$\mathcal{Q}_3^{\text{large}} := \{Q \in \mathcal{Q} \,:\, Q_2 \Subset \pi_{\mathcal{L}^*} Q_2 \subset \tilde{Q}_1\}, \tag{73}$$

$$\mathcal{Q}_4^{\text{large}} := \{Q \in \mathcal{Q} \,:\, Q_2 \not\Subset \pi_{\mathcal{L}^*} Q_2 \Subset \tilde{Q}_1\}, \tag{74}$$

$$\mathcal{Q}_5^{\text{large}} := \{Q \in \mathcal{Q} \,:\, Q_2 \not\Subset \pi_{\mathcal{L}^*} Q_2 \subsetneq \tilde{Q}_1, \text{and } \pi_{\mathcal{L}^*} Q_2 \not\Subset \tilde{Q}_1\}. \tag{75}$$

Observe that Lemma 6.8, with (82), gives

$$\mathbf{B}_{\mathcal{Q}_3^{\text{large}}} \lesssim \tau. \tag{76}$$

Take the collection $\mathcal{S}$ of Lemma 6.8 to be $\mathcal{L}^*$.

Observe that Lemma 6.9 applies to show that the estimate (76) holds for $\mathcal{Q}_4^{\text{large}}$. Take $\mathcal{S}$ of that Lemma to be $\mathcal{L}^*$. The estimate from Lemma 6.9 is given in terms of η, as defined in (94). But, is at most τ.

In the last collection, $\mathcal{Q}_5^{\text{large}}$, notice that the conditions placed upon the pair implies that $|Q_1| \le 2^{2r+2}|Q_2|$, for all $Q \in \mathcal{Q}_5^{\text{large}}$. It therefore follows from a straight forward application of Lemma 6.10, that (76) holds for this collection as well.

Upper Bounds on the Stopping Form

We prove upper bounds on the norm of the stopping form in a situation in which there is some decoupling between the martingale difference on g, and the argument of the Hilbert transform. First, an elementary observation.

Proposition 6.7 *For intervals $J \subset L \Subset K$, with L either good, or the child of a good interval,*

$$\frac{P(\sigma(F-K), J)}{|J|} \simeq \frac{P(\sigma(F-K), L)}{|L|}. \tag{77}$$

Proof The property of interval I being good, says that if $I \subset \tilde{I}$, and $2^{r-1}|I| \le |\tilde{I}|$, then the distance of either child of I to the boundary of $\tilde{I}$ is at least $|I|^{\epsilon}|\tilde{I}|^{1-\epsilon}$. Thus, in the case that L is the child of a good interval, the parent $\hat{L}$ of L is contained in K, and $2^{r-1}|\hat{L}| \le |K|$, so by the definition of goodness,

$$\begin{aligned}\operatorname{dist}(J, F-K) &\ge \operatorname{dist}(L, F-K) \\ &\ge |L|^{\epsilon}|K|^{1-\epsilon} \ge 2^{r(1-\epsilon)}|L|.\end{aligned}$$

The same inequality holds if L is good. Then, one has the equivalence above, by inspection of the Poisson integrals. □

Lemma 6.8 *Let $\mathcal{S}$ be a collection of pairwise disjoint intervals in F. Let $\mathcal{Q}$ be admissible such that for each $Q \in \mathcal{Q}$, there is an $S \in \mathcal{S}$ with $Q_2 \Subset S \subset \tilde{Q}_1$. Then, there holds*

$$|B_{\mathcal{Q}}(f,g)| \lesssim \eta \|f\|_\sigma \|g\|_w, \tag{78}$$

$$\textit{where} \quad \eta^2 := \sup_{S\in\mathcal{S}} \frac{1}{\sigma(S)} \sum_{J\in\mathcal{Q}_2 \,:\, J\Subset S} P(\sigma(F-S),J)^2 \langle \frac{x}{|J|}, h_J^w \rangle_w^2. \tag{79}$$

It is useful to note that η is always smaller than the size: For $S \in \mathcal{S}$, let $\mathcal{J}^*$ be the maximal intervals $J \in \mathcal{Q}_2$ with $J \Subset S$, and note that (77) applies to see that

$$\sum_{J\in\mathcal{Q}_2 \,:\, J\Subset S} P(\sigma(F-S),J)^2 \langle \frac{x}{|J|}, h_J^w \rangle_w^2 = \sum_{J^*\in\mathcal{J}^*} \sum_{J\in\mathcal{Q} \,:\, J\subset J^*} P(\sigma(F-S),J)^2 \langle \frac{x}{|J|}, h_J^w \rangle_w^2 \tag{80}$$

$$\lesssim \sum_{J^*\in\mathcal{J}^*} \frac{P(\sigma(F-S),J^*)^2}{|J^*|^2} \sum_{J\in\mathcal{Q} \,:\, J\subset J^*} \langle x, h_J^w \rangle_w^2 \tag{81}$$

$$\lesssim \sum_{J^*\in\mathcal{J}^*} \sigma(J^*) \lesssim \operatorname{size}(\mathcal{Q})\sigma(S). \tag{82}$$

Proof An interesting part of the proof is that it depends very much on cancellative properties of the martingale differences of f. (Absolute values must be taken *outside* the sum defining the stopping form!) This argument will invoke the stopping data, and part of the Hilbert-Poisson exchange argument.

Assume, as we can, that the Haar support of f is contained in $\mathcal{Q}_1$. Take $\mathcal{F}$ and $\alpha_f(\cdot)$ to be stopping data defined in this way: First, add to $\mathcal{F}$ the interval F, and set $\alpha_f(F) := \mathbb{E}_F^\sigma |f|$. Inductively, if $F \in \mathcal{F}$ is minimal, add to $\mathcal{F}$ the maximal children F' such that $\alpha_f(F') := \mathbb{E}_{F'}^\sigma |f| > 4\alpha_f(F)$. This is a simple form of the stopping data construction in section "Global to Local Reduction". In particular quasi-orthogonality (27) holds.

Write the bilinear form as

$$B_{\mathcal{Q}}(f,g) = \sum_J \langle H_\sigma \varphi_J, \Delta_J^w g \rangle_w \tag{83}$$

$$\text{where} \quad \varphi_J := \sum_{Q\in\mathcal{Q} \,:\, Q_2=J} \mathbb{E}_J^\sigma \Delta_{Q_1}^\sigma f \cdot (F - \tilde{Q}_1). \tag{84}$$

The function φ_J is well-behaved, as we now explain. At each point x with $\varphi_J(x) \neq 0$, the sum above is over pairs Q such that $Q_2 = J$ and $x \in F - \tilde{Q}_1$. By the convexity property of admissible collections, the sum is over consecutive (good) martingale differences of f. The basic telescoping property of these differences shows that the sum is bounded by the stopping value $\alpha_f(\pi_{\mathcal{F}} J)$. Let I^* be the maximal interval of the form $\tilde{Q}_1$ with $x \in F - \tilde{Q}_1$, and let I_* be the child of the minimal such interval which contains J. Then,

$$\begin{aligned} |\varphi_J(x)| &= \Big| \sum_{\substack{Q \in \mathcal{Q} : Q_2 = J \\ x \in I - \tilde{Q}_1}} \mathbb{E}_J^\sigma \Delta_{Q_1}^\sigma f(x) \Big| \\ &= \big| \mathbb{E}_{I^*}^\sigma f - \mathbb{E}_{I_*}^\sigma f \big| \lesssim \alpha_f(\pi_{\mathcal{F}} J)(F - S), \end{aligned} \tag{85}$$

where S is the $\mathcal{S}$-parent of J.

We can estimate as below, for $F \in \mathcal{F}$:

$$\Xi(F) := \Big| \sum_{Q \in \mathcal{Q} : \pi_{\mathcal{F}} Q_2 = F} \mathbb{E}_{Q_2} \Delta_{Q_1}^\sigma f \cdot \langle H_\sigma(F - \tilde{Q}_1), \Delta_J^w g \rangle_w \Big| \tag{86}$$

$$\overset{(84)}{=} \Big| \sum_{J \in \mathcal{Q}_2 : \pi_{\mathcal{F}} J = F} \langle H_\sigma \varphi_J, \Delta_J^w g \rangle_w \Big| \tag{87}$$

$$\overset{(85)}{\lesssim} \alpha_f(F) \sum_{\substack{S \in \mathcal{S} \\ \pi_{\mathcal{F}} S = F}} \sum_{\substack{J \in \mathcal{Q}_2 \\ J \subset S}} P(\sigma(F - S), J) \Big| \Big\langle \frac{x}{|J|}, \Delta_J^w g \Big\rangle_w \Big| \tag{88}$$

$$\lesssim \alpha_f(F) \Big[\sum_{\substack{S \in \mathcal{S} \\ \pi_{\mathcal{F}} S = F}} \sum_{\substack{J \in \mathcal{Q}_2 \\ J \subset S}} P(\sigma(F - S), J)^2 \Big\langle \frac{x}{|J|}, h_J^w \Big\rangle_w^2 \times \sum_{\substack{J \in \mathcal{Q}_2 \\ \pi_{\mathcal{F}} J = F}} \hat{g}(J)^2 \Big]^{1/2} \tag{89}$$

$$\overset{(79)}{\lesssim} \operatorname{size}(\mathcal{Q}) \alpha_f(F) \Big[\sum_{\substack{S \in \mathcal{S} \\ \pi_{\mathcal{F}} S = F}} \sigma(S) \times \sum_{\substack{J \in \mathcal{Q}_2 \\ \pi_{\mathcal{F}} J = F}} \hat{g}(J)^2 \Big]^{1/2} \tag{90}$$

$$\lesssim \operatorname{size}(\mathcal{Q}) \alpha_f(F) \sigma(F)^{1/2} \Big[\sum_{J \in \mathcal{Q}_2 : \pi_{\mathcal{F}} J = F} \hat{g}(J)^2 \Big]^{1/2}. \tag{91}$$

The top line follows from (84). In the second, we appeal to (85) and monotonicity principle, the latter being available to us since $J \subset S$ implies $J \Subset S$, by hypothesis. We also take advantage of the strong assumptions on the intervals in $\mathcal{Q}_2$: If $J \in \mathcal{Q}_2$, we must have $\pi_{\mathcal{F}} J = \pi_{\mathcal{F}}(\pi_{\mathcal{S}} J)$. The third line is Cauchy–Schwarz, followed by the appeal to the hypothesis (79), while the last line uses the fact that the intervals in $\mathcal{S}$ are pairwise disjoint.

The quasi-orthogonality argument (27) completes the proof, namely we have

$$\sum_{F \in \mathcal{F}} \Xi(F) \lesssim \operatorname{size}(\mathcal{Q}) \|f\|_\sigma \|g\|_w. \tag{92}$$

□

Lemma 6.9 *Let $\mathcal{S}$ be a collection of pairwise disjoint intervals in F. Let $\mathcal{Q}$ be admissible such that for each $Q \in \mathcal{Q}$, there is an $S \in \mathcal{S}$ with $Q_2 \subset S \Subset \tilde{Q}_1$. Then, there holds*

$$|B_{\mathcal{Q}}(f,g)| \lesssim \eta \|f\|_\sigma \|g\|_w, \tag{93}$$

$$\textit{where} \quad \eta^2 := \sup_{S \in \mathcal{S}} \frac{\mathsf{P}(\sigma(Q_1 - \pi_{\tilde{Q}_1} S), S)^2}{\sigma(S)|S|^2} \sum_{J \in \mathcal{Q}_2 \,:\, J \subset S} \langle x, h_J^w \rangle_w^2. \tag{94}$$

Proof Construct stopping data $\mathcal{F}$ and $\alpha_f(\cdot)$ as in the proof of Lemma 6.8. The fundamental inequality (85) is again used. Then, by the monotonicity principle (17), there holds for $F \in \mathcal{F}$,

$$\begin{aligned}
\Xi(F) &:= \Big| \sum_{Q \in \mathcal{Q} \,:\, \pi_{\mathcal{F}} Q_2 = F} \mathbb{E}_{Q_2} \Delta_{Q_1}^\sigma f \cdot \langle H_\sigma(F - \tilde{Q}_1), \Delta_{Q_2}^w g \rangle_w \Big| \\
&\lesssim \alpha_f(F) \sum_{S \in \mathcal{S} \,:\, \pi_{\mathcal{F}} S = F} \mathsf{P}(\sigma(F - \pi_{\tilde{Q}_1} S), S) \sum_{J \in \mathcal{Q}_2 \,:\, J \subset S} \Big\langle \frac{x}{|S|}, h_J^w \Big\rangle_w \cdot |\hat{g}(J)| \\
&\lesssim \alpha_f(F) \Big[\sum_{S \in \mathcal{S} \,:\, \pi_{\mathcal{F}} S = F} \mathsf{P}(\sigma(F - \pi_{\tilde{Q}_1} S), S)^2 \sum_{J \in \mathcal{Q}_2 \,:\, J \subset S} \Big\langle \frac{x}{|S|}, h_J^w \Big\rangle_w^2 \times \sum_{J \in \mathcal{Q}_2 \,:\, J \subset S} \hat{g}(J)^2 \Big]^{1/2} \\
&\lesssim \eta \alpha_f(F) \Big[\sum_{S \in \mathcal{S} \,:\, \pi_{\mathcal{F}} S = F} \sigma(S) \times \sum_{J \in \mathcal{Q}_2 \,:\, J \subset S} \hat{g}(J)^2 \Big]^{1/2} \\
&\lesssim \eta \alpha_f(F) \sigma(F)^{1/2} \Big[\sum_{J \in \mathcal{Q}_2 \,:\, \pi_{\mathcal{F}} J = F} \hat{g}(J)^2 \Big]^{1/2}.
\end{aligned}$$

After the monotonicity principle (17), we have used Cauchy-Schwarz, and the definition of η. The quasi-orthogonality argument (27) then completes the analysis of this term, see (92). □

The last Lemma that we need is elementary, and is contained in the methods of [38].

Lemma 6.10 *Let $u \geq r + 1$ be an integer, and $\mathcal{Q}$ be an admissible collection of pairs such that $|Q_1| = 2^u |Q_2|$ for all $Q \in \mathcal{Q}$. There holds*

$$|B_{\mathcal{Q}}(f,g)| \lesssim \textit{size}(\mathcal{Q}) \|f\|_\sigma \|g\|_w.$$

Proof Recall the form of the stopping form in (53). Observe, from inspection of the definition of the Haar function (10), that

$$|\mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f| \leq \frac{|\hat{f}(I)|}{\sigma(I_J)^{1/2}}.$$

Then, an elementary application of the monotonicity principle gives us

$$
\begin{aligned}
|B_{\mathcal{Q}}(f,g)| &\le \sum_{I\in\mathcal{Q}_1} |\hat{f}(I)| \sum_{J\,:\,(I,J)\in\mathcal{Q}} \sigma(I_J)^{-1/2}\mathsf{P}(\sigma(F-I_J),J)\Big\langle \frac{x}{|J|}, h_J^w\Big\rangle_w |\hat{g}(J)| \\
&\le \|f\|_\sigma \Bigg[\sum_{I\in\mathcal{Q}_1}\Bigg[\sum_{J\,:\,(I,J)\in\mathcal{Q}} \frac{1}{\sigma(I_J)}\mathsf{P}(\sigma(F-I_J),J)\Big\langle \frac{x}{|J|}, h_J^w\Big\rangle_w |\hat{g}(J)|\Bigg]^2\Bigg]^{1/2} \\
&\le \operatorname{size}(\mathcal{Q})\|f\|_\sigma\|g\|_w
\end{aligned}
$$

This follows immediately from Cauchy-Schwarz, and the fact that for each $J \in \mathcal{Q}_2$, there is a unique $I \in \mathcal{Q}_1$ such that the pair (I, J) contribute to the sum above. □

Context and Discussion

The proof herein succeeds because the notion of size approximates the operator norm of the stopping form. Moreover, the 'large' portions of the stopping form, there is a decoupling that takes place.

It is very interesting that one can prove unconditional results about the two weight Hilbert transform, following the techniques in [23], without solving the local problem.

Elementary Estimates

This section is devoted to the proof of Lemma 4.1. The estimates fall into many subcases, and are of a more classical nature, albeit the A_2 assumption is critical. (In fact, all the estimates in this section depend only on the half-Poisson A_2 hypothesis, but this is not systematically tracked in the notation.) In addition, all estimates should be interpreted as uniform over all smooth truncations. Some of these are off-diagonal estimates, for which the smooth truncations are important. The uniformity over truncations is however suppressed in notation.

First some basic estimates are collected. This is property of good intervals, which can be effectively used in non-critical situations.

Lemma 7.1 *For three intervals $J, I, I' \in \mathcal{D}$ with $J \subset I \subset I'$, $|J| = 2^{-s}|I|$, with $s \ge r$ and J good, then*

$$P(\sigma\cdot(I'-I),J) \le 2^{-(1-\varepsilon)s}P(\sigma\cdot I',I). \tag{95}$$

Proof Note that for $x \in I' - I$ we have

$$\operatorname{dist}(x, J) \geq |I|^{1-\varepsilon} |J|^{\varepsilon} = 2^{s(1-\varepsilon)} |J|.$$

Using this in the definition of the Poisson integral, we get

$$\begin{aligned}
P(\sigma \cdot (I' - I), J) &\leq 2 \int_{I'-I} \frac{|J|}{\operatorname{dist}(x, J)^2} \, \sigma(dx) \\
&\lesssim \frac{|J|}{|I|} \int_{I'-I} \frac{|I|}{(|J| + \operatorname{dist}(x, J))^2} \, \sigma(dx) \\
&\lesssim 2^{-s(1-2\varepsilon)} \int_{I'-I} \frac{|I|}{(|I| + \operatorname{dist}(x, I))^2} \, \sigma(dx) = 2^{-s(1-2\varepsilon)} P(\sigma(I' - I), I).
\end{aligned}$$

□

Proposition 7.2 *Suppose that two intervals $I, J \in \mathcal{D}$ satisfy $|I| \geq |J|$, and $3I \cap J = \emptyset$, then*

$$\sup_{0<\alpha<\beta} |\langle H(\sigma I), h_J^w \rangle_w| \lesssim \sigma(I) \sqrt{w(J)} \frac{|J|}{(|J| + dist(I, J))^2} \tag{96}$$

Proof Since h_J^w has w-integral zero, estimate as below, where x_J is the center of J.

$$\begin{aligned}
|\langle HI, h_J^w \rangle_w| &= \Big| \int_I \int_J K_{\alpha,\beta}(y - x) \cdot h_J^w(x) \; w(dx) \sigma(dy) \Big| \\
&= \Big| \int_I \int_J \{K_{\alpha,\beta}(y - x) - K_{\alpha,\beta}(y - x_J)\} h_J^w(x) \; w(dx) \sigma(dy) \Big| \\
&\lesssim \int_I \int_J \frac{|J|}{(|J| + \operatorname{dist}(I, J))^2} |h_J^w(x)| \; w(dx) \sigma(dy).
\end{aligned}$$

The Lemma follows by inspection. □

Proposition 7.3 *Suppose that two intervals $I, J \in \mathcal{D}$ satisfy $2^s |J| = |I|$, where $s > r$, the interval J is good, and $J \subset 3I \setminus I$, then*

$$\sup_{0<\alpha<\beta} |\langle H(\sigma I), h_J^w \rangle_w| \lesssim 2^{-(1-2\varepsilon)s} \sigma(I) \sqrt{w(J)} |I|^{-1} \tag{97}$$

Proof Under the assumption of the Lemma, the proof of Proposition 7.2 holds, supplying the estimate estimate of that Lemma. But, the extra assumption that J is good implies that $\operatorname{dist}(J, I) > 2^{s(1-\varepsilon)} |J|$, and then the estimate follows by inspection.

□

The Weak Boundedness Inequality

The following inequality is a weak-boundedness inequality, a consequence of the A_2 inequality. Here, we look at the Hilbert transform inequality on two disjoint intervals.

Proposition 7.4 *There holds for all disjoint intervals I, J with no point masses at their endpoints,*

$$\sup_{0<\alpha<\beta} |\langle H(\sigma f \cdot I), g \cdot J \rangle_w| \lesssim \mathcal{A}_2^{1/2} \|f\|_\sigma \|g\|_w. \tag{98}$$

The constant on the right can in fact be taken as follows. For a point a that separates the interiors of I and J, with I to the left of a,

$$\sup_{r>0} P(\sigma \mathbf{1}_{(-\infty,a)}, (a, a+r)) \frac{w(a, a+r)}{r} + P(w \mathbf{1}_{(a,\infty)}, (a, a+r)) \frac{\sigma(a-r, a)}{r}. \tag{99}$$

In particular, for arbitrary intervals I and J with no point masses at the endpoints,

$$|\langle H_\sigma I, J \rangle_w| \lesssim \mathcal{A}_2^{1/2} [\sigma(I) w(J)]^{1/2} \tag{100}$$

It is useful to note that the *global* integrability of indicators is then a consequence of the A_2 and interval testing conditions.

Since the intervals are disjoint, there is no possibility of cancellation in the estimate, and it therefore is closely relate to the Hardy inequality. In the two weight setting, this has been characterized by Muckenhoupt [32].

Theorem F *For weights $\widehat{w}$ and σ supported on $\mathbb{R}_+$.*

$$\left\| \int_{(0,x)} f \, \sigma(dy) \right\|_{\hat{w}} \leq \mathcal{B} \|f\|_\sigma, \tag{101}$$

$$\textit{where} \quad \mathcal{B}^2 \simeq \sup_{0<r<\infty} \int_{(r,\infty)} \widehat{w}(dx) \times \int_{(0,r)} \sigma(dy). \tag{102}$$

For the sake of completeness, we recall Muckenhoupt's proof of this result. This preparation is proved by integration by parts.

Proposition 7.7 *Let ϕ be an increasing function on $(0, \infty)$, with $\phi(0) = 0$ and ϕ strictly positive on $(0, \infty)$. Then,*

$$\int_{(0,x]} \phi(t)^{-1/2} d\phi(t) \leq 2\phi(x)^{1/2}, \tag{103}$$

with equality if ϕ is continuous.

Proof of Theorem F We are free to assume that the function $\phi(x) = \sigma((0,x))$ is strictly positive on $(0,\infty)$. Then, multiply and divide by $\phi(x)^{1/4}$, and use Cauchy–Schwarz to see that

$$\Big\| \int_{(0,x)} f \, \sigma(dy) \Big\|_{\hat{w}}^2 \le \int_{(0,\infty)} \int_{(0,x)} f(y)\phi(y)^{1/2} \, \sigma(dy) \cdot \int_{(0,x)} \phi(y)^{-1/2} \, \sigma(dy) \; \hat{w}(dx)$$
$$\le 2 \int_{(0,\infty)} \int_{(0,x)} f(y)\phi(y)^{1/2} \, \sigma(dy) \cdot \phi(x)^{1/2} \; \hat{w}(dx)$$
$$= 2 \int_{(0,\infty)} f(y)\phi(y)^{1/2} \int_{(y,\infty)} \phi(x)^{1/2} \; \hat{w}(dx) \; \sigma(dy)$$

Above, we have used (103), and then Fubini. Concentrate on the inner integral. Our definition of $\mathcal{B}$ and Proposition 7.7 gives us

$$\mathcal{B} \int_{(y,\infty)} \Big[\int_{(x,\infty)} \hat{w}(dt) \Big]^{-1/2} \hat{w}(dx) \le 2\mathcal{B} \Big[\int_{(y,\infty)} \hat{w}(dt) \Big]^{1/2}$$

And, now we can estimate

$$\Big\| \int_{(0,x)} f \, \sigma(dy) \Big\|_{\hat{w}}^2 \le 4\mathcal{B} \int_{(0,\infty)} f(y)\phi(y)^{1/2} \Big[\int_{(y,\infty)} \hat{w}(dt) \Big]^{1/2} \sigma(dy) \le 4\mathcal{B}^2 \|f\|_{\sigma}^2 .$$

The proof is complete. □

Proof of Proposition 7.4 Interval testing and (98) prove the estimate (100), so we turn to the proof of (98).

After a translation, we can assume that 0 separates the interiors of I and J. Let us assume that I is to the left of zero. We change the problem. Set $\tilde{\sigma}(dx) = \sigma(-dx)$ for $x \ge 0$, and for $f \in L^2(I,\sigma)$, set $\phi(x) = f(-x)$. Then,

$$\langle H_\sigma f, g \rangle_w = \int_{(-\infty,0)} \int_{(0,\infty)} \frac{f(y)g(x)}{y-x} \, \sigma(dy) w(dx) \tag{104}$$
$$= -\int_{(0,\infty)} \int_{(0,\infty)} \frac{\phi(y)g(x)}{x+y} \, \tilde{\sigma}(dy) w(dx). \tag{105}$$

The double integral is split into dual terms, one of which is

$$\int_{(0,\infty)} \int_{(0,x)} \frac{\phi(y)g(x)}{x+y} \, \tilde{\sigma}(dy) \; w(dx). \tag{106}$$

We analyze this bilinear form.

Note that $x + y \simeq x$ in (106). Thus, it suffices to estimate

$$\int_{(0,\infty)} \Big| \int_{(0,x)} \frac{\phi(y)}{x} \, \tilde{\sigma}(dy) \Big|^2 \, w(dx) = \int_{(0,\infty)} \Big| \int_{(0,x)} \frac{\phi(y)}{x} \, \tilde{\sigma}(dy) \Big|^2 \, \frac{w(dx)}{x^2} \le \mathcal{B}^2 \|\phi\|_{\tilde{\sigma}}^2.$$

where $\mathcal{B}$ is as in (102), and $\hat{w}(dx) = \frac{w(dx)}{x^2}$ and $\sigma = \tilde{\sigma}$

It remains to estimate the constant $\mathcal{B}$, For any $0 < r < \infty$,

$$\int_{(0,r)} \tilde{\sigma}(dy) \int_{(r,\infty)} d\widehat{w} = \frac{\sigma(-r,0)}{r} \int_{(r,\infty)} \frac{r}{x^2} \, w(dx) \lesssim \mathcal{A}_2.$$

The more precise conclusion (99) can be read off from this inequality. Recall that (105) is split into two bilinear forms, and we have only considered one of them. This explains the symmetric form of (99). □

The Different Subcases of Lemma 4.1

Lemma 4.1 follows from appropriate bounds on these bilinear forms, and their duals.

$$B^{\text{nearby}}(f,g) \coloneqq \sum_{\substack{I,J \,:\, 2^{-r-1}|I| \le |J| \le |I| \\ 3I \cap J \ne \emptyset}} |\langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w|, \tag{107}$$

$$B^{\text{far}}(f,g) \coloneqq \sum_{I,J \,:\, 3I \cap 3J = \emptyset} |\langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w|, \tag{108}$$

$$B^{\text{close}}(f,g) \coloneqq \sum_{\substack{I,J \,:\, 2^r|J| \le |I| \\ J \subset 3I \setminus I}} |\langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w|, \tag{109}$$

$$B^{\text{adjacent}}(f,g) \coloneqq \sum_{I,J \,:\, J \Subset I_J} |\mathbb{E}_{I-I_J}^\sigma \Delta_I^\sigma f \langle H_\sigma (I - I_J), \Delta_J^w \phi \rangle_w|. \tag{110}$$

Lemma 7.8 *For* $\star \in \{nearby, \ far, \ close, \ adjacent\}$, *there holds*

$$B^\star(f,g) \lesssim \mathcal{A}_2^{1/2} \|f\|_\sigma \|g\|_w.$$

The Nearby Term

One can check directly that for each interval I, with child I', there holds $|\mathbb{E}^{\sigma}_{I'} h^{\sigma}_I| \le \sigma(I')^{-1/2}$. It then follows from (98) that $|\langle H_\sigma h^\sigma_I, h^w_J\rangle_w| \lesssim \mathcal{H}$. And then,

$$B^{\text{nearby}}(f,g) \lesssim \mathcal{H} \sum_{\substack{I,J\,:\,2^{-r-1}|I|\le|J|\le|I| \\ 3I\cap J\neq\emptyset}} |\hat{f}(I)\hat{g}(J)| \lesssim \mathcal{H}\|f\|_\sigma \|g\|_w .$$

The last line follows from the fact that for each I, there are only a bounded number of J occurring in the sum.

Here, and below, we will be using the notation $\hat{f}(I) = \langle f, h^\sigma_I\rangle_\sigma$.

The Far Term

We consider the case of $|J| \le |I|$, and $3I \cap 3J \neq \emptyset$. It follows that $J \subset 3^{s+1}I \setminus 3^s I$ for some integer $s \ge 1$. For an interval K, integer $s \ge r$ and $t \ge 0$, consider the two projections

$$\Pi^\sigma_{K,s,t} f := \sum_{\substack{I\,:\,I\subset 3^{t+2}K\setminus 3^{t+1}K \\ |I|=|K|}} \Delta^\sigma_I f$$

$$\Pi^w_{K,s,t} g := \sum_{\substack{J\,:\,J\subset 3^t K \\ 2^s|J|=|K|}} \Delta^w_J g .$$

These projections satisfy, for fixed s,t,

$$\sum_K \|\Pi^\sigma_{K,s,t} f\|^2_\sigma \le \|f\|^2_\sigma , \tag{111}$$

with a similar bound for $\Pi^w_{K,s,t} g$. Also, we need to bound

$$\sum_{s\ge r}\sum_{t\ge 0} \left|\langle H_\sigma \Pi^\sigma_{K,s,t} f, \Pi^w_{K,s,t} g\rangle_w\right| . \tag{112}$$

But, using the fact that $\Delta^w_J g$ has mean zero, and the distance between the support of $\Pi^\sigma_{K,s,t} f$ and $\Pi^w_{K,s,t} g$ is approximately $3^t|K|$, we have

$$\left|\langle H_\sigma \Pi^\sigma_{K,s,t} f, \Pi^w_{K,s,t} g\rangle_w\right| \lesssim \frac{2^{-s}|K|}{3^{2t}|K|^2} \|\Pi^\sigma_{K,s,t} f\|_{L^1(\sigma)} \|\Pi^w_{K,s,t} g\|_{L^1(w)}$$

$$\lesssim \frac{\sqrt{\sigma(3^{t+2}K)w(3^tK)}}{2^s 3^{2t}|K|} \| \Pi^{\sigma}_{K,s,t} f\|_{\sigma} \| \Pi^{w}_{K,s,t} g\|_{w}$$

$$\lesssim 2^{-s}3^{-t}\mathcal{A}_2^{1/2} \| \Pi^{\sigma}_{K,s,t} f\|_{\sigma} \| \Pi^{w}_{K,s,t} g\|_{w}.$$

Since we have gained geometric decay in s, and t, and we have the inequality (111), we can easily complete the proof of (112).

The Close Term

For integers $s \geq r$, the sum below a relative length of J with respect to I. Applying (97),

$$\sum_{\substack{I,J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} |\langle H_\sigma \Delta^\sigma_I f, \Delta^w_J \phi\rangle_w| \lesssim 2^{(1-2\varepsilon)s} \sum_{\substack{I,J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} |\hat f(I)\hat g(J)| \frac{\sqrt{\sigma(I)w(J)}}{|I|}$$

$$\lesssim 2^{(1-2\varepsilon)s} \sum_I |\hat f(I)| \frac{\sqrt{\sigma(I)}}{|I|} \sum_{\substack{J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} |\hat g(J)| \sqrt{w(J)}$$

We have the geometric decay in s. Apply Cauchy–Schwarz, one term is $\|f\|_\sigma$. The other term, squared, is

$$\sum_I \frac{\sigma(I)}{|I|^2} \sum_{\substack{J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} \hat g(J)^2 \times \sum_{\substack{J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} w(J) \lesssim \mathcal{A}_2 \sum_I \sum_{\substack{J\,:\,2^s|J|=|I| \\ J\subset 3I\setminus I}} \hat g(J)^2 \lesssim \mathcal{A}_2 \|g\|_w^2.$$

This completes the estimate.

The Adjacent Term

We argue as in the previous case. It is easy to see that $|\mathbb{E}^\sigma_{I-I_J} \Delta^\sigma_I f| \lesssim |\hat f(I)| \sigma(I - I_J)^{-1/2}$.

For $\theta \neq \theta' \in \{\pm\}$, and consider the sum below, where s plays the same role as before.

$$\sum_{\substack{I,J\,:\,2^s|J|=|I| \\ J\subset I+(\theta'|I|)}} \left| \mathbb{E}^\sigma_{I_\theta} \Delta^\sigma_I f \cdot \langle H_\sigma I_\theta, \Delta^w_J g\rangle_w \right| \lesssim 2^{-(1-2\varepsilon)s} \sum_{\substack{I,J\,:\,2^s|J|=|I| \\ J\subset I+(\theta'|I|)}} |\hat f(I)\hat g(J)| \frac{\sqrt{\sigma(I_\theta)w(J)}}{|I|}$$

$$\lesssim 2^{-(1-2\varepsilon)s} \mathcal{A}_2^{1/2} \|f\|_\sigma \|g\|_w.$$

The details are suppressed.

Context and Discussion

The techniques of this section are all drawn from the work of Nazarov-Treil-Volberg [38, 56], aside from the use of the two weight Hardy inequality, which is drawn from [20].

Proof Under the Pivotal Assumption

We prove an upper bound for a two weight inequality assuming a pivotal condition on a pair of weights. The setup is as follows. Let $K(y)$ satisfy the size and gradient condition

$$|x-y| \cdot |\nabla K(x,y)| + |K(x,y)| \leq |x-y|^{-1}.$$

We will consider the operator Tf given formally by p.v. $\int K(x,y)f(y)\,dy$. In the two weight setting, no principal value need exist, so given two weights σ, w, we consider the constant $\mathcal{N}_T$, which is be the best constant in the inequality

$$\Big\| \int K(x,y)f(y)\ \sigma(dy) \Big\|_w \leq \mathcal{N}_T \|f\|_\sigma .$$

Let $\mathcal{P}$ be the best constant in the *pivotal inequality*, defined as follows. For any interval I_0 and any partition $\mathcal{P}$ of I_0 into intervals such that neither σ nor w have point masses at the endpoints, there holds

$$\sum_{I \in \mathcal{P}} P(\sigma(I_0 \setminus I), I)^2 w(I) \leq \mathcal{P}^2 \sigma(I_0). \tag{113}$$

We also require that the dual inequality, with the roles of w and σ reversed, holds. One can note that this inequality will hold if the maximal function satisfies the two weight inequality $\|M_\sigma f\|_w \lesssim \|f\|_\sigma$, and its dual.

Theorem 8.1 (Nazarov-Treil-Volberg [56]) *Assume that the pair of weights w, σ satisfy the A_2 condition (3), and the pivotal conditions hold, namely $\mathcal{P} < \infty$. Then, there holds $\mathcal{N}_T \lesssim \mathcal{T}_T + \mathcal{A}_2^{1/2} + \mathcal{P}$, where $\mathcal{T}$ is the best constant in the inequalities*

$$\int_I |T_\sigma I|^2\ w(dx) \leq \mathcal{T}_T^2 \sigma(I), \qquad \int_I |T_w I|^2\ \sigma(dx) \leq \mathcal{T}_T^2 w(I).$$

We give the proof, with the goal of highlighting some of the difficulties that one must face in the general case. In addition, a quantitative higher dimensional version of this Theorem was key to [45]. We will use Calderón-Zygmund stopping data, to facilitate comparisons to the general case. This will also give an easier proof than is in [45, 56].

Off-Diagonal Estimates

We need a typical off-diagonal estimate, one that is far less refined than the monotonicity principle.

Lemma 8.2 *For all* $0 < \alpha < \beta$*, good intervals* $J \Subset I$*, and function* f *is supported off of* I*, there holds*

$$|\langle T\sigma f, g\rangle| \lesssim P(\sigma|f| \cdot I^c, I) w(J)^{1/2} \|g\|_w. \tag{114}$$

for any function $g \in L^2(w)$*, supported on* J *and with integral zero.*

Proof Use the standard subtraction argument to see that

$$\begin{aligned} |\langle T\sigma f, g(x)\rangle| &= \Big| \int_J \int_{\mathbb{R}\setminus I} \{K(x,y) - K(x_J, y)\} f(y) g(x)\, \sigma(dy)\, w(dx) \Big| \\ &\lesssim \int_J \int_{\mathbb{R}\setminus I} \frac{|x - x_J|}{(x_J - y)^2} \cdot |f(y) g(x)|\, \sigma(dy)\, w(dx). \end{aligned}$$

The bound follows by Cauchy–Schwarz and inspection. □

The Global to Local Reduction

One need only prove that

$$|\langle T_\sigma P^\sigma_{\text{good}} f, P^w_{\text{good}} g\rangle_w| \lesssim \mathcal{T} \|f\|_\sigma \|g\|_w,$$

where $\mathcal{T} := \mathcal{T}_T + \mathcal{A}_2^{1/2} + \mathcal{P}$. The set up is much like section "Global to Local Reduction". We will understand that the functions f and g can be assumed to be good functions. In fact, f has the 'thin' Haar expansion in (21), and similarly for g, in order to reduce some case analysis below.

In analogy to (23), define

$$B^{\text{above}}(f, g) := \sum_{I : I \subset I_0} \sum_{J : J \Subset I} \mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f \cdot \langle T_\sigma I_J, \Delta^w_J g\rangle_w, \tag{115}$$

and define $B^{\text{below}}(f, g)$ similarly. Since Lemma 4.1 depends only on the A_2 assumption, we have

Lemma 8.3 *There holds*

$$\left|\langle T_\sigma f, g\rangle_w - B^{above}(f,g) - B^{below}(f,g)\right| \lesssim \mathcal{A}_2^{1/2}\|f\|_\sigma \|g\|_w.$$

Thus, the main technical result is

Lemma 8.4 *There holds*

$$|B^{above}(f,g)| \lesssim \mathfrak{T}\|f\|_\sigma \|g\|_w. \tag{116}$$

The same inequality holds for $B^{below}(f,g)$.

The stopping intervals are defined similarly.

Definition 8.5 Define $\mathcal{F}$, the stopping intervals, recursively by initializing $I^0 \in \mathcal{F}$, and in the recursive step, if $F \in \mathcal{F}$ is minimal, add to $\mathcal{F}$ the maximal subintervals $F' \subset F$, with $F' \in \mathcal{D}_f$, so that meet either of these conditions:

f **stopping** $\mathbb{E}^\sigma_{F'}|f| > C\alpha_f(F) := \mathbb{E}^\sigma_F|f|$.
Pivotal Stopping $\mathsf{P}(\sigma \cdot I_0, I)^2 w(I) > 10\mathcal{P}^2\sigma(I)$.

That is, we stop if either the average of f becomes too large, or, essentially, the pivotal quantity becomes too large.

We use the same notation as in section "Global to Local Reduction", and in analogy to Corollary 4.4, there holds

Lemma 8.6 (The Global to Local Reduction) *There holds*

$$|B^{above}_{\mathcal{F},glob}(f,g)| \lesssim \mathfrak{T}\|f\|_\sigma \|g\|_w, \tag{117}$$

$$\textit{where} \quad B^{above}_{\mathcal{F},glob}(f,g) := \sum_{\substack{I,J\,:\,\dot{\pi}_{\mathcal{F}}J \subsetneq I \\ J \Subset I}} \mathbb{E}^\sigma_{I_J}\Delta^\sigma_I f \cdot \langle T_\sigma I_J, \Delta^w_J g\rangle_w. \tag{118}$$

Proof This variant of the 'Hilbert-Poisson exchange' argument is needed. Holding $F \in \mathcal{F}$ fixed, we sum over J with $\dot{\pi}_{\mathcal{F}}J = F$ and I with $F \subsetneq I$. Then, the argument of T_σ is I_F which is written as $I_F = F + (I_F \setminus F)$. Defining ε_F by

$$\sum_{I\,:\,I \supsetneq F} \mathbb{E}^\sigma_{I_J}\Delta^\sigma_I f := \varepsilon_F \alpha_f(F),$$

these constants are bounded by a constant: $|\varepsilon_F| \lesssim 1$. Then,

$$\begin{aligned}\Phi(F) &:= \left|\sum_{I\,:\,I \supsetneq F}\ \sum_{J\,:\,\dot{\pi}_{\mathcal{F}}J = F} \mathbb{E}^\sigma_{I_F}\Delta^\sigma_I f \cdot \langle T_\sigma F, \Delta^w_J g\rangle_w\right| \\ &= \left|\left\langle T_\sigma F, \sum_{J\,:\,\dot{\pi}_{\mathcal{F}}J = F} \varepsilon_J \Delta^w_J g\right\rangle_w\right| \le \mathfrak{T}\alpha_f(F)\sigma(F)^{1/2}\|Q^w_F g\|_w\end{aligned}$$

This depends upon the testing assumption on T_σ applied to intervals. The operator $Q^w_F g$ is the Haar projection defined at (26). Quasi-orthogonality as in (27) finishes the sum over $F \in \mathcal{F}$.

The complementary case is that of the global-to-local reduction. But, under the pivotal condition there is a geometric decay along the stopping tree. For $F \in \mathcal{F}$, and integer j, let $\mathrm{ch}_j(F)$ be the j-fold descendants of F in the collection $\mathcal{F}$. That is, $\mathrm{ch}_0(F) = \{F\}$, and $F' \in F' \in \mathrm{ch}_{j+1}(F)$ iff F' is the child of some interval $F'' \in \mathrm{ch}_j(F)$.

We will index by $F \in \mathcal{F}$, $F' \in \mathrm{ch}_1(F)$, and $F'' \in \mathrm{ch}_j(F')$, where $j \geq 0$. Using (114) and critically, Lemma 7.1, we have

$$\Big| \sum_{I\,:\,\pi_{\mathcal{F}} I = F} \mathbb{E}^\sigma_{I_{F'}} f \langle T_\sigma (I_{F'} \setminus F'), Q^w_{F''} g\rangle_w \Big| \lesssim \alpha_f(F)\mathbb{P}(\sigma \cdot (F \setminus F'), F'') w(F'') \| Q^w_{F''} g\|_w$$

$$\lesssim \alpha_f(F) 2^{(1-\epsilon)j} \mathbb{P}(\sigma \cdot (F \setminus F''), F'') w(F'')^{1/2} \| Q^w_{F''} g\|_w .$$

We have geometric decay in j above. Moreover, summing over F' and F'', we can appeal to the pivotal condition (20) to see that

$$\sum_{F'' \in \mathrm{ch}_{j+1}(F)} \mathbb{P}(\sigma \cdot (F \setminus F''), F'') w(F'')^{1/2} \| Q^w_{F''} g\|_w$$

$$\leq \Bigg[\sum_{F'' \in \mathrm{ch}_{j+1}(F)} \mathbb{P}(\sigma \cdot (F \setminus F''), F'')^2 w(F'') \times \sum_{F'' \in \mathrm{ch}_{j+1}(F)} \| Q^w_{F''} g\|_w^2 \Bigg]^{1/2}$$

$$\lesssim \mathcal{P} \Bigg[\sigma(F) \sum_{F'' \in \mathrm{ch}_{j+1}(F)} \| Q^w_{F''} g\|_w^2 \Bigg]^{1/2} .$$

Then, quasi-orthogonality is used to estimate

$$\sum_{F \in \mathcal{F}} \alpha_f(F) \Bigg[\sigma(F) \sum_{F'' \in \mathrm{ch}_{j+1}(F)} \| Q^w_{F''} g\|_w^2 \Bigg]^{1/2} \lesssim \|f\|_\sigma \|g\|_w .$$

This completes the global to local reduction. □

The Local Estimate

It remains to prove the following *local estimate*:

$$|B^{\mathrm{above}}(P^\sigma_F f, g)| \lesssim \mathcal{T}\big\{ \alpha_f(F)\sigma(F)^{1/2} + \|P^\sigma_F f\|_\sigma \big\} \|g\|_w , \qquad Q^w_F g = g,$$

for then quasi-orthogonality will complete the bound on $B^{\mathrm{above}}_{\mathcal{F}}(f, g)$.

In the bilinear form above, the argument of T_σ is, for a pair of intervals $J \Subset I$, $I_J = (F - I_J) + F$. Using linearity, and focusing on the argument of T_σ being F, we can repeat the argument of (34), which depends upon the fact that the averages of f are controlled. Below, there is an requirement that I_J has $\mathcal{F}$-parent F, which we are free to add since $Q^w_F g = g$.

$$\Big| \sum_{I\,:\,\pi_{\mathcal{F}} I = F} \sum_{J\,:\,\dot{\pi}_{\mathcal{F}} I_J = F} \mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f \cdot \langle T_\sigma F, \Delta^w_J g \rangle_w \Big| \lesssim \mathfrak{T} \alpha_f(F) \sigma(F)^{1/2} \|g\|_w .$$

This bound follows the argument of (34), and we suppress the details.

It therefore remains to consider the *stopping form*

$$B^{\text{stop}}_F(f,g) := \sum_{I\,:\,\pi_{\mathcal{F}} I = F} \sum_{J\,:\,\dot{\pi}_{\mathcal{F}} I_J = F} \mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f \cdot \langle T_\sigma(I_0 - I_J), \Delta^w_J g \rangle_w .$$

Lemma 8.7 *For all $F \in \mathcal{F}$, there holds*

$$|B^{stop}_F(f,g)| \lesssim \mathcal{P} \|f\|_\sigma \|g\|_w .$$

Proof This depends very much on the selection of stopping intervals. In fact there is geometric decay, holding the relative lengths of I and J fixed. Estimate for integers $s \geq r$,

$$\begin{aligned}
&\sum_{I\,:\,\pi_{\mathcal{F}} I = F} \sum_{\substack{J\,:\,J \Subset I_J, \pi_{\mathcal{F}} I_J = F \\ |I| = 2^s |J|}} |\mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f \cdot \langle T_\sigma(I_0 - I_J), \Delta^w_J g \rangle_w| \\
&\qquad \leq \sum_{I\,:\,\pi_{\mathcal{F}} I = F} \sum_{\theta \in \{\pm\}} \frac{|\hat{f}(I)|}{\sigma(I_\theta)^{1/2}} \sum_{\substack{J\,:\,J \Subset I_\theta, \pi_{\mathcal{F}} I_J = F \\ |I| = 2^s |J|}} P(\sigma(F - I_\theta), J) \langle \tfrac{x}{|J|}, h^w_J \rangle_w |\hat{g}(J)| \\
&\qquad \lesssim M_s \Big[\sum_{I\,:\,\pi_{\mathcal{F}} I = F} \hat{f}(I)^2 \Big]^{1/2} \times \Big[\sum_{J\,:\,J \subset F} \hat{g}(I)^2 \Big]^{1/2} \\
&\qquad \text{where} \quad M_s^2 := \max_{\theta \in \{\pm\}} \sup_{I\,:\,\pi_{\mathcal{F}} I_\theta = F} \frac{1}{\sigma(I_\theta)} \sum_{\substack{J\,:\,J \Subset I_\theta, \pi_{\mathcal{F}} I_J = F \\ |I| = 2^s |J|}} P(\sigma(F - I_\theta), J)^2 w(J).
\end{aligned}$$

Here, we have used (a) used the bound $|\mathbb{E}^\sigma_{I_J} \Delta^\sigma_I f| \leq \frac{|\hat{f}(I)|}{\sigma(I_\theta)^{1/2}}$; (b) appealed to (114); (c) used Cauchy–Schwarz, together with the fact that for $J \Subset F$, there is a unique I containing it, with length $2^s |J|$.

It remains to bound M_s, gaining a geometric decay in s, and appealing to the pivotal condition. Return to the inequality (95), to gain the geometric decay,

$$\sum_{\substack{J\,:\,J\Subset I_\theta,\pi_{\mathcal{F}}I_J=F\\ |I|=2^s|J|}} P(\sigma(F-I_\theta),J)^2 w(J) \lesssim 2^{-(1-\varepsilon)s} P(\sigma\cdot F, I_\theta)^2 w(I_\theta) \lesssim 2^{-(1-\varepsilon)s}\mathcal{P}^2\sigma(I_\theta),$$

where the decisive point is that I_θ has $\mathcal{F}$-parent F, hence it must *fail* the pivotal stopping condition. □

Example Weights

The sharpness of the different conditions in the main theorem is the subject of the this section.

Theorem 9.1 *There are pairs of weights σ, w, with no common point masses, that satisfy any one of these conditions.*

(1) The pair of weights satisfies the full Poisson A_2 condition, but the norm inequality for the Hilbert transform (1) *does not hold.*
(2) The pair of weights satisfies the full Poisson A_2 condition, and the testing inequality (4)*, but the norm inequality for the Hilbert transform* (1) *does not hold.*
(3) The pair of weights satisfy the two weight norm inequality (1)*, but not the pivotal condition* (20)*.*

Point (1) is a counterexample to Sarason's Conjecture, first disproved by Nazarov [34]. In contrast to his argument, an explicit pair of weights are exhibited.

The Initial Steps in the Main Construction

Let $C = \bigcap_{n=0}^{\infty} C_n$ be the standard middle third Cantor set in the unit interval. Thus, $C_0 = [0,1]$, $C_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$, and more generally

$$C_n = \bigcup \Big\{[x, x+3^{-n}] \,:\, x = \sum_{j=1}^{n} \epsilon_j 3^{-j},\ \epsilon_j \in \{0,2\}\Big\}.$$

Let w be the standard uniform measure on C. Thus $w(I) = 2^{-n}$ on each component of C_n, $n \in \mathbb{N}_0$. This is phrased slightly differently. Let $\mathcal{K}$ be the collection of components of all the sets C_n. Then, for each $K \in \mathcal{K}$, there holds $w(K) = |K|^{\frac{\ln 2}{\ln 3}}$.

The weight σ will be a sum of point masses selected from the intervals in $\mathcal{G}$, taken to be the components of the open set $[0,1] - C$. (G is for 'gap'. See Fig. 8.) Consider the Hw restricted an interval $G \in \mathcal{G}$. This is a smooth, monotone function, hence it has a unique zero z_G (Fig. 9). Then, the weight σ is

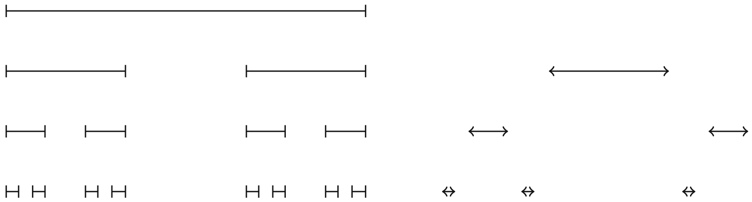


Fig. 8 The approximates to the Cantor set C on the *left*, and on the *right*, the gaps, namely the components of $[0, 1] - C$. The intervals on the *left* are in $\mathcal{K}$, and those on the *right* are in $\mathcal{G}$

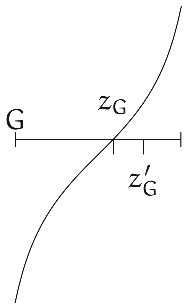


Fig. 9 The selection of the points z_G and z'_G for a gap interval G. The function Hw, restricted to G is monotone increasing, hence has a unique zero, the point z_G. The second point z'_G will be to the *right*, but a distance to the boundary of G that is at least $c|G|$, for absolute constant $0 < c < \frac{1}{2}$

$$\sigma := \sum_{G \in \mathcal{G}} s_G \cdot \delta_{z_G}, \tag{119}$$

where $s_G > 0$ will be chosen momentarily, consistent with the A_2 condition. A second measure is given by $\sigma' := \sum_{G \in \mathcal{G}} s_G \cdot \delta_{z'_G}$, where z'_G is the unique point in G at which $Hw(z'_G) = |G|^{-1+\frac{\ln 2}{\ln 3}}$. See Fig. 9.

The constants s_G are be specified by the simple A_2 ratio

$$\frac{w(3G)}{|G|} \cdot \frac{\sigma(G)}{|G|} = 2, \qquad \text{that is} \quad s_G = 2|G|^{2-\frac{\ln 2}{\ln 3}}. \tag{120}$$

To see this, note that

$$w(3G) = w(G - |G|) + w(G + |G|) = 2|G|^{\frac{\ln 2}{\ln 3}},$$

since $G \pm |G|$ are components of some C_n. With this definition, the basic facts about the w and σ come from the geometry of the Cantor set and the relations below,

$$\begin{aligned} w(I) &\lesssim |I|^{\frac{\ln 2}{\ln 3}}, && I \text{ is triadic,} \\ \sigma(I) &\lesssim |I|^{2-\frac{\ln 2}{\ln 3}} && I \text{ is triadic, } I \text{ not strictly contained in any } G \in \mathcal{G}. \end{aligned} \tag{121}$$

On the other hand, if $I \in \mathcal{G} \cup \mathcal{K}$, the inequalities above can be reversed, namely

$$w(3I) \simeq |I|^{\frac{\ln 2}{\ln 3}}, \quad \sigma(I) \simeq |I|^{2-\frac{\ln 2}{\ln 3}}, \qquad I \in \mathcal{G} \cup \mathcal{K}. \tag{122}$$

The properties of these measures that we are establishing are as follows.

Lemma 9.2 *For the measures just defined, there holds*

(1) The Hilbert transform H_σ is bounded from $L^2(\sigma)$ to $L^2(w)$.
(2) The Hilbert transform $H_{\sigma'}$ is unbounded *from $L^2(\sigma')$ to $L^2(w)$, but the pair of weights satisfy the A_2 condition, and the testing conditions*

$$\sup_{I \text{ an interval}} \sigma'(I)^{-1} \int_I |H_{\sigma'} I|^2 \, dw < \infty.$$

Concerning point 2, the unboundedness of H_w is direct from the construction of σ'.

$$\int (Hw)^2 \, d\sigma' = \sum_{G \in \mathcal{G}} Hw(z'_G)^2 \sigma'(\{z'_G\}) \tag{123}$$

$$= \sum_{G \in \mathcal{G}} |G|^{2-\frac{\ln 2}{\ln 3} - 2(1-\frac{\ln 2}{\ln 3})} = \sum_{G \in \mathcal{G}} |G|^{+\frac{\ln 2}{\ln 3}} = \infty. \tag{124}$$

There are exactly 2^{n-1} elements of $\mathcal{G}$ of length 3^{-n}, proving the sum is infinite.

The Poisson A_2 Condition

Lemma 9.3 *For either weight $\mu \in \{\sigma, \sigma'\}$, the pair of weights w, μ satisfy the A_2 condition.*

Proof It suffices to check the A_2 condition on the triadic intervals in the unit interval. Let us begin by showing that for any triadic interval $I \in \mathcal{K} \cup \mathcal{G}$,

$$P(\sigma, I) \lesssim \frac{\sigma(I)}{|I|}, \quad \text{and} \quad P(w, I) \lesssim \frac{w(3I)}{|I|}. \tag{125}$$

For then, the control of the simple A_2 ratio will imply the control of the full A_2 ratio. (For the inequality on w, the triple of the interval appears on the right, since $w(I)$ can be zero if $I \in \mathcal{G}$.) Now, it will be clear that this argument is insensitive to the location of the points z_G and z'_G, so the same argument for σ will work equally well for σ'.

Let us consider σ. Using (122), there holds

$$P(\sigma, I) \le \frac{\sigma(I)}{|I|} + \sum_{k=1}^{\infty} \int_{3^k I \setminus 3^{k-1} I} \frac{|I|}{(|I|^2 + \mathrm{dist}(x,I))^2}\, \sigma(dx)$$
$$\lesssim \frac{\sigma(I)}{|I|} + \sum_{k=1}^{\infty} \frac{\sigma(3^k I)}{3^k |3^k I|}$$
$$\lesssim \frac{\sigma(I)}{|I|} + \sum_{k=1}^{\infty} 3^{-k} |3^k I|^{1-\frac{\ln 2}{\ln 3}} \lesssim \frac{\sigma(I)}{|I|} \sum_{k=0}^{\infty} 3^{-k\frac{\ln 2}{\ln 3}} \lesssim \frac{\sigma(I)}{|I|}.$$

Turning to the weight w, one has

$$P(w, I) \le \frac{w(3I)}{|I|} + \sum_{k=2}^{\infty} \int_{3^k I \setminus 3^{k-1} I} \frac{|I|}{(|I|^2 + \mathrm{dist}(x,I))^2}\, w(dx)$$
$$\lesssim \frac{w(3I)}{|I|} + \sum_{k=2}^{\infty} \frac{w(3^k I)}{3^k |3^k I|}$$
$$\lesssim \frac{w(3I)}{|I|} + \sum_{k=2}^{\infty} 3^{-k} |3^k I|^{-1+\frac{\ln 2}{\ln 3}} \lesssim \frac{w(3I)}{|I|} \sum_{k=1}^{\infty} 3^{-k(2-\frac{\ln 2}{\ln 3})} \lesssim \frac{w(3I)}{|I|}.$$

The A_2 product $P(\sigma, I) \cdot P(w, I)$ has been bounded for $I \in \mathcal{K} \cup \mathcal{G}$. Suppose that I is a triadic interval that is not in these two collections. Then, I must be strictly contained in some gap $G \in \mathcal{G}$. Writing $I^{(k)} = G$, where, $I^{(k)}$ denotes the k-fold parent of I in the triadic grid, we have $w(G) = 0$. Hence,

$$P(w, I) = \int_{[0,1]\setminus G} \frac{|I|}{(|I| + \mathrm{dist}(x,I))^2}\, w(dx) \simeq 3^{-k} P(w, G).$$

First, consider σ restricted to the gap G:

$$P(w, I) P(\sigma \cdot G, I) \lesssim 3^{-k} P(w, G) \frac{\sigma(G)}{|I|} \simeq P(w, G) \frac{\sigma(G)}{|G|} \lesssim 1.$$

Now, we have to consider the Poisson average of σ off of the gap G, in which case we have

$$P(\sigma \cdot ([0,1] \setminus G), I) \simeq 3^{-k} P(\sigma, G),$$

and so the estimate follows. □

The Testing Conditions

We turn to the testing conditions, using in an essential way the precise definition of the weight σ: it gives a huge cancellation, which simplifies things considerably.

Lemma 9.4 *For any interval I, there holds*

$$\int_I |H_w I|^2 \, d\sigma \lesssim w(I). \tag{126}$$

Proof By construction of σ, there are two reductions. The first is simple, namely that the two endpoints of the interval I can be taken to be an endpoint of an interval in $\mathcal{G}$. The second comes from the construction of σ: $Hw \equiv 0$, relative to $d\sigma$ measure. Hence,

$$\int_I |H_w I|^2 \, d\sigma = \int_I H_w([0,1]-I)^2 \, d\sigma,$$

namely the *complement of I* is the argument of the Hilbert transform on the right.

Then, one abandons all further cancellations. Let us show that for all intervals $K \in \mathcal{K}$ (the components of the sets C_n which generate the Cantor set),

$$\int_K |H_w K_{\mathrm{rt}}|^2 \, d\sigma \lesssim w(K), \tag{127}$$

where K_{rt} is the right component of $[0,1] \setminus K$. The same estimate holds for the left component, and this completes the proof. For, if we set I_{rt} to be the right component of $[0,1] \setminus I$, and take $K^1, K^2, \dots,$ to be the maximal intervals in K contained in I, there holds

$$\begin{aligned}\int_I (H_w I_{\mathrm{rt}})^2 \, d\sigma &\le \sum_{n=1}^{\infty} \int_{K^n} (H_w K^n_{\mathrm{rt}})^2 \, d\sigma \\ &\lesssim \sum_{n=1}^{\infty} w(K^n) \lesssim w(I).\end{aligned}$$

Now, for $K \in \mathcal{K}$, let $K_1, K_2, \dots,$ be the maximal intervals in $\mathcal{K}$ that lie to the right of K. Arranging them in increasing length, note that the length of K_1 is either $|K|$ or $3|K|$. For $n \ge 2$, the length of K_n increases by a factor of 3, and $\operatorname{dist}(K, K_n) \gtrsim |K_n|$, and hence there are at most $1 - \log_3 |K|$ such intervals in $\mathcal{K}$. Here is an illustration:

Then, one has the estimate below, where the sum is of a decreasing geometric series, estimated by its first term.

$$|H_w K_{\rm rt}| \lesssim \sum_{n=1}^{\infty} \frac{w(K_n)}{|K_n|} \simeq \frac{w(K)}{|K|}.$$

Hence, (127) follows from the control of the A_2 ratio. □

An important part of the remaining arguments is that points z_G, and z'_G cannot cluster close to the boundary of G.

Lemma 9.5 *There is a constant* $0 < c < \frac{1}{2}$ *such that*

$$|z_G - z'_G| \leq c|G|.$$

Proof Estimate Hw at the midpoint z''_G of a component G. By symmetry of the Hilbert transform, and the Cantor set, it always holds that $H(w\mathbf{1}_{3G})(z'_G) = 0$, so that appealing to (121),

$$\begin{aligned} |Hw(z''_G)| &= |H(w\mathbf{1}_{(3G)^c})(z''_G)| \\ &\lesssim \sum_{k=2}^{n} \frac{w(3^k G)}{|3^k G|} \\ &\lesssim \sum_{k=2}^{n} |3^k G|^{-1+\frac{\ln 2}{\ln 3}} \lesssim |G|^{-1+\frac{\ln 2}{\ln 3}} \end{aligned}$$

Next, we turn to a derivative calculation. The function Hw, restricted to G is a smooth function, one that diverges at the end points of G at a rate that reflect the fractal dimension of G. For any $x \in G$ note that

$$\begin{aligned} \frac{d}{dx} Hw(x) &= \int_C \frac{w(dy)}{(y-x)^2} \\ &\gtrsim \frac{w(3G)}{|G|^2} \simeq |G|^{-2+\frac{\ln 2}{\ln 3}}. \end{aligned}$$

This is a uniform lower bound, and in fact the lower bound is very poor at the boundaries of G. Indeed,

$$\frac{d}{dx} Hw(x) \gtrsim \operatorname{dist}(x, \partial G)^{-2+\frac{\ln 2}{\ln 3}}.$$

It follows that we have to have $|z_G - z'_G| < c|G|$, for some $0 < c < \frac{1}{2}$. That is, one need only move at fixed small multiple of $|G|$, passing from the location of the zero z_G to the point z'_G. □

The second half of the testing intervals inequalities is as follows.

Lemma 9.6 *For $\mu \in \{\sigma, \sigma'\}$, and any interval I,*

$$\int_I |H_\mu I|^2 \, dw \lesssim \mu(I). \tag{128}$$

Proof For the sake of specificity, let $\mu = \sigma$. Indeed, by Lemma 9.5, the same argument will work for σ'. To fix ideas, let us assume that $I \in \mathcal{K}$. Write the left, middle and right thirds of I as I_{-1}, I_0, I_1, respectively. Then, note that

$$\int_I H_\sigma(I)^2 \, dw = \int_{I_{-1}\cup I_1} H_\sigma(I)^2 \, dw \tag{129}$$

$$\lesssim \int_{I_{-1}\cup I_1} H_\sigma(I_0)^2 \, dw + \int_{I_{-1}} H_\sigma(I_0+I_1)^2 \, dw + \int_{I_1} H_\sigma(I_{-1}+I_0)^2 \, dw \tag{130}$$

$$+ \int_{I_{-1}} H_\sigma(I_{-1})^2 \, dw + \int_{I_1} H_\sigma(I_1)^2 \, dw. \tag{131}$$

The first term on the right is simple. On the interval I_0, σ is a point mass, at a point that is at distance $\geq c|I|$ from $I_{\pm 1}$. Thus, by (122),

$$\int_{I_{-1}\cup I_1} H_\sigma(I_0)^2 \, dw \lesssim \frac{|I|^{4-2\frac{\ln 2}{\ln 3}}}{|I|^2} |I|^{\frac{\ln 2}{\ln 3}} \simeq \sigma(I).$$

That completes the first integral. The remaining two integrals in (130) are handled by a similar argument.

Concerning the two integrals in (131), one should note that $I_{\pm 1} \in \mathcal{K}$ and that $\sigma(I_{\pm 1}) \leq 3^{-2+2\frac{\ln 2}{\ln 3}} \sigma(I)$. This geometric factor is smaller than $\frac{1}{2}$, therefore one can recurse on (130) and (131) to see that

$$\int_K H_\sigma(K)^2 \, dw \lesssim \sigma(K), \qquad K \in \mathcal{K}. \tag{132}$$

For a general interval I, since σ is a sum of Dirac masses, we can assume that the interval I is in a canonical form. Namely, each endpoint of I can be assumed to be an endpoint of an interval in $\mathcal{G}$. The basic inequality is

$$\sum_{K\in\mathcal{K}_I} \int_K |H_\sigma(I-K)|^2 \, dw \lesssim \sigma(I), \tag{133}$$

where $\mathcal{K}_I$ is the maximal elements of $\mathcal{K}$ contained in I. The integration is over K, and the argument of the Hilbert transform is $I - K$.

To see that (133) implies the Lemma, note that by (132),

$$\begin{aligned}\int_I H_\sigma(I)^2\,dw &= \sum_{K\in\mathcal{K}_I}\int_K H_\sigma(I)^2\,dw\\ &\lesssim \sum_{K\in\mathcal{K}_I}\int_K H_\sigma(I-K)^2\,dw + \sum_{K\in\mathcal{K}_I}\int_K H^\sigma(K)^2\,dw\\ &\lesssim \sigma(I) + \sum_{K\in\mathcal{K}_I}\sigma(K) \lesssim \sigma(I).\end{aligned}$$

In fact, (133) follows from

$$\int_K |H_\sigma(I-K)|^2\,dw \lesssim \frac{\sigma(I)^2}{|I|^2}w(K), \qquad K\in\mathcal{K}_I. \tag{134}$$

For this is summed over $K\in\mathcal{K}_I$, and then one uses the A_2 property.

To prove (134), all hope of cancellation is abandoned. For an interval $K\in\mathcal{K}_I$, let us consider component $I_{\rm rt}$ of $I-K$ which lies to the right of K. It has a Whitney like decomposition into a finite sequence of intervals $J_1,\dots,J_t$ that we construct now. These intervals will have the property that they are (a) pairwise disjoint, (b) their union is $I_{\rm rt}$, (c) and $\operatorname{dist}(K,\operatorname{supp}(\sigma J_s)) \gtrsim |J_s| \gtrsim 3^{\frac{s}{2}}|K|$, for all $1\le s\le t$.

Now, $J_1 = K+|K| \in \mathcal{G}$. If this interval is not contained in I, it follows that K contains the right hand endpoint of I, and there is nothing to prove. Assuming that $J_1\subset I$, the inductive step is this. Given $J_1,\dots,J_s$, as above, whose union is not $I_{\rm rt}$

(1) If $J_s\in\mathcal{G}$, then $J_s+|J_s|\in\mathcal{K}$. If this interval is contained in $I_{\rm rt}$, then we take $J_{s+1}=J_s+|J_s|\in\mathcal{K}$, and repeat the recursion. Otherwise, we update $J_s := I_{\rm rt}-\bigcup_{u=1}^{s-1}J_t$, and the recursion stops.
(2) If $J_s\in\mathcal{K}$, then it follows that $J_{s-1}\in\mathcal{G}$, and the element of $\mathcal{G}$ immediately to the right of J_s is $3(J_s+6|J_s|)$. If this interval is contained in $I_{\rm rt}$, then we take $J_{s+1}=3(J_s+6|J_s|)\in\mathcal{G}$, and repeat the recursion. Otherwise, we update $J_s := I_{\rm rt}-\bigcup_{u=1}^{s-1}J_t$, and the recursion stops.

With this construction, it follows that

$$|H_\sigma(I_{\rm rt})\cdot K| \lesssim \sum_{u=1}^{t}\frac{\sigma(J_s)}{|J_s|} \lesssim \sum_{n=1}^{\infty}|J_s|^{1-\frac{\ln 2}{\ln 3}} \lesssim \frac{\sigma(I)}{|I|}.$$

This proves the 'right half' of (134), that is, when the argument of the Hilbert transform is $I_{\rm rt}$. The 'left half' is the same, so the proof is complete. □

At this point, we have proven that the pair of weights (w,σ') satisfy the full Poisson A_2 condition, and the testing condition (128). But, $\|Hw\|_{L^2(\sigma')}$ is infinite, by (124). Hence, points (1) and (2) of Theorem 9.1 are shown.

We have also shown that the pair of weights (w, σ) satisfy the full Poisson A_2 condition, and both sets of testing conditions. Hence, by our main theorem, H_w is bounded from $L^2(w)$ to $L^2(\sigma)$. This pair of weights also fail the pivotal condition (20) of Nazarov-Treil-Volberg [38]. This is verified by observing that the collection $\mathcal{G}$ of gaps is a partition of [0, 1], and

$$\begin{aligned}\sum_{G\in\mathcal{G}} P(w,G)^2 w(G) &\simeq \sum_{G\in\mathcal{G}} \frac{w(3G)^2}{|G|^2}\sigma(I) \\ &\simeq \sum_{G\in\mathcal{G}} w(3G) \simeq \sum_{G\in\mathcal{G}} |G|^{\frac{\ln 2}{\ln 3}} = \infty\end{aligned}$$

since $\mathcal{G}$ contains 2^n intervals of length 3^{-n}, for all integers n. Here, we have used (125), followed by (121). Since $\inf_{x\in G} Mw(x) \gtrsim P(w,G)$, this also shows that the maximal function M is not bounded from $L^2(w)$ to $L^2(\sigma)$.

Notice in contrast that the energy inequality (19) for the partition $\mathcal{G}$ is trivial, since σ restricted to any interval G is a point mass, hence $E(\sigma, G) = 0$, for all $G \in \mathcal{G}$.

Context and Discussion

Counterexamples were an important source of inspiration on these questions. The early paper of Muckenhoupt and Wheeden [33] includes an example of the fact that the simple A_2 condition is not sufficient for the two weight inequality. For instance, the boundedness of the simple A_2 ratio is simple to check for the pair $w = \delta_0$, and $\sigma(dx) = x\mathbf{1}_{[0,\infty)}dx$. Then, one sees that for $f = \frac{1}{x}\mathbf{1}_{[1,L]}$,

$$\sqrt{\log L} \simeq \|f\|_\sigma \ll \log L \simeq \|H_\sigma f\|_w, \qquad L > 1.$$

Thus, the Hilbert transform is unbounded. And, one can directly see that the half-Poisson A_2 condition fails.

Much harder, is the fact that the Poisson A_2 condition is not sufficient. This was the contribution of Nazarov [34]. This example lead to the conjecture of Nazarov-Treil-Volberg [56] proved herein. A more delicate example, of a pair of weights which satisfied the Poisson A_2 condition, and one set of testing conditions, say (4), but not the norm inequality was that of Nazarov-Volberg [40]. Also see Nikol'skiĭ-Treil [42], for a related example to disprove a conjecture about similarity to a normal operator. Both of these latter examples were based upon Nazarov's indirect example.

The example given here is directly inspired by a Cantor set type example in Sawyer's two weight maximal function paper [52]. It is drawn from [20], with the purpose to show that the *pivotal condition* of Nazarov-Treil-Volberg [38, 56] was *not necessary* for the two weight inequality to hold. This was an explicit example, and also pointed to the primary role of the notion of energy. It is very interesting and delicate, in that the point masses have to be placed on the zeros of the Hilbert transform, in order to obtain the boundedness of the transform. It is also humbling in that it still does not reveal how delicate the proof of the sufficiency in the main theorem needs to be.

It is subtle example of Maria Carmen Reguera [47] and Reguera-Thiele [49] that proves this, as is pointed out by Reguera-Scurry [48].

Theorem G *There is a pair of weights for which the maximal function* M_σ *is bounded from* $L^2(\sigma) \to L^2(w)$ *and* M_w *is bounded from* $L^2(w) \to L^2(\sigma)$*, but norm inequality for the Hilbert transform* (1) *does not hold.*

This is quite a bit more intricate than the examples we have presented. It had been suggested, in the early days of the weighted theory, that the boundedness of the maximal functions would be sufficient for the norm boundedness of the Hilbert transform. On the other hand, if one considers ‘off-diagonal’ estimates, then boundedness of the maximal function is sufficient for norm inequalities for singular integrals [8].

Applications of the Main Inequality

The interest in the two weight problem stems from a range of potential applications arising in sophisticated arenas of complex function and spectral theory. The motivations for these questions are complicated, and based upon subtle theories. The connections to the two weight Hilbert transform are not always immediate, and the properties of interest are frequently more intricate than those of mere boundedness of a transform. Nevertheless, the acknowledged experts Belov-Mengestie-Seip in [4] write “…we have found it both useful and conceptually appealing to transform the subject into a study of the mapping properties of discrete Hilbert transforms. We have learned to appreciate that the essential difficulties thus seem to appear in a more succinct form.” A brief guide to the subjects, and some of the ‘essential difficulties’ follow.

Sarason’s Question on Toeplitz Operators

This question arose from Sarason’s work on exposed points of H^1 [50]. Indeed, this was part of an influential body of work that pointed to the distinguished role of de Branges spaces in the subject. This paper contains examples of pairs of functions f, g, for which the individual Toeplitz operators where unbounded, but the composition bounded.

Question 1 (Sarason [51]) *Characterize those pairs of outer functions $g, h \in H^2$ for which the composition of Toeplitz operators $T_g T_{\overline{h}}$ is bounded on H^2.*

Following [51], for a function $h \in L^2(\mathbb{T})$, the Toeplitz operator T_h can be thought of as taking $f \in H^2$ to the space of analytic functions by the definition

$$T_h f(z) := \frac{1}{2\pi} \int_{\partial\mathbb{D}} f(\mathrm{e}^{i\theta}) h(\mathrm{e}^{i\theta}) \overline{k_z(\mathrm{e}^{i\theta})}\, d\theta,$$

where $k_w(z) := \frac{(1-|w|^2)^{1/2}}{1-\overline{w}z}$ is the reproducing kernel.

Also in [51] is an argument of S. Treil that a Poisson A_2 condition is necessary condition for the boundedness of the composition:

$$\sup_{z\in\mathbb{D}} P|f|^2(z) P|g|^2(z) < \infty, \tag{135}$$

where P denotes the Poisson extension to the unit disk. Sarason wrote that *'It is tempting to conjecture that the last condition is also sufficient for the boundedness of $T_g T_{\overline{h}}$.'* This statement, widely referred to as the Sarason Conjecture, is of interest in both the Hardy and Bergman space settings.(Aleman-Pott-Reguera [2] have resolved the conjecture in the negative in a Bergman space setting. A striking argument in which they prove the boundedness of the Bergman projection is *equivalent to* the boundedness of the positive part of the Bergman projection. This allows a much simpler counterexample to be identified.)

The connection with the two weight problem for the Hilbert transform is indicated by the diagram from [7] §5, see Fig. 10. In the diagram, $M_{\overline{h}}$ is multiplication by $\overline{h}$ and P_+ is the Riesz projection from L^2 to H^2. The boundedness is equivalent to

$$M_g P_+ M_{\overline{f}} \;:\; H^2 \mapsto H^2.$$

The structure of outer functions leads to these simplifications. Since the product of analytic is analytic, the second H^2 above can be replaced by L^2, and then, the outside multiplication M_g can then be replaced by $M_{|g|}$. Thus, we are considering

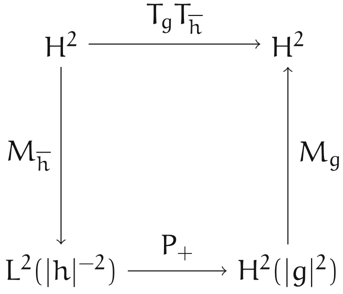


Fig. 10 Sarason's Question concerns the *top line* of the diagram, which is equivalent to the *lower part* of the diagram. The operator $M_{\overline{h}}$ on the *left* is an isometry onto its range, while M_g, the operator on the *right* is an isometry between the two spaces

$M_{|g|}P_+M_{\bar{f}} \;:\; H^2 \mapsto L^2$. Now, $\bar{f}$ is anti-analytic, so we can replace H^2 above by L^2. Moreover, the multiplication operator $M_{f/|f|}$ is unitary, since an outer function can be equal to zero on $\mathbb{T}$ only on a set of measure zero. Thus, it is equivalent to consider

$$M_{|g|}P_+M_{|f|} \;:\; L^2 \mapsto L^2.$$

This is a two weight inequality for P_+. (Sergei Treil helped us with the history of this question.)

The Riesz projection is a linear combination of the identity and the Hilbert transform, and our main theorem will apply to it. Note that the inequality

$$\|P_+(|f|\phi)\|_{L^2(|g|^2dx)} \lesssim \|\phi\|_{L^2(dx)}$$

is equivalent to

$$\|P_+(|f|^2\psi)\|_{L^2(|g|^2dx)} \lesssim \|\psi\|_{L^2(|f|^2dx)}.$$

Recall that $P_+ = I - \frac{\pi}{i}H$, according to how we defined the Hilbert transform, where I represents the identity operator. In the two weight setting, we interpret the norm inequality $\|P_+(\sigma f)\|_w \lesssim \|f\|_\sigma$, as uniform over all truncations $0 < \tau < 1$ defined by

$$P_{+,\tau}(\sigma f) := \sigma f + \frac{i}{\pi}\int_{\tau<|x-y|<\tau^{-1}} f(y)\frac{\sigma(dy)}{y-x}$$

Theorem 10.1 *For pairs of weights w, σ that absolutely continuous with respect to Lebesgue measure, the norm inequality $\|P_+(\sigma f)\|_w \lesssim \|f\|_\sigma$ holds if and only if the pair of weights satisfy the Poisson A_2 condition (3), and these testing inequalities hold, uniformly over all intervals I, for a finite positive constant $\mathcal{P}$,*

$$\int_I |P_+(\sigma\mathbf{1}_I)|^2\, w(dx) \le \mathcal{P}^2\sigma(I), \qquad \int_I |P_+(w\mathbf{1}_I)|^2\, \sigma(dx) \le \mathcal{P}^2 w(I).$$

One must be sure that the A_2 inequality is necessary from the norm inequality. As it suffices to test real-valued functions, the real-variable proof given here will suffice. This in particular shows that for the densities of the weights, $\sigma(x)\cdot w(x) \le \mathcal{A}_2$, for a.e.$x$. Thus, the identity part of the norm, and testing, inequalities are trivial. The remaining parts just concern the Hilbert transform, so one can use the main result.

If one is interested in the Sarason question for functions f, g that are not outer, there is no simple reduction to the two weight inequality for the Hilbert transform, and the problem is quite subtle, as the role of the multiplier $P_+M_{\bar{f}}$ is more involved than that of just a weight.

Model Spaces

For a probability measure σ on $\mathbb{T}$, define a holomorphic function θ on $\mathbb{D}$ by the Poisson integral

$$\frac{1}{1-\theta(z)} \coloneqq \int_{\mathbb{T}} \frac{1}{1-z\bar{\zeta}}\,\sigma(d\zeta).$$

This is an inner function: A holomorphic map of $\mathbb{D}$ to itself which is unimodular a.e. on $\mathbb{T}$. Also, $\theta(0) = 1$. (The measure σ is a Clark measure for θ, frequently written as σ_1.)

The shift operator $Sf(z) = zf(z)$ on H^2 has invariant subspace $\theta H^2 = \{\theta f : f \in H^2\}$, whence $K_\theta \coloneqq H^2 \ominus \theta H^2$ is invariant for S^*. Beurling's theorem states that every invariant subspace for S^* is of this form. The model operator is $S_\theta \coloneqq P_\theta S$, where P_θ is the orthogonal projection from H^2 onto K_θ. Remarkably, subject to mild conditions, every contractive operator on a Hilbert space is unitarily equivalent to a properly chosen S_θ. For this, and other reasons, properties of the K_θ spaces have broad significance.

The spaces K_θ and $L^2(\sigma)$ are unitarily equivalent, with the unitary map from $f \in L^2(\sigma)$ to $F \in K_\theta$ given by

$$F(z) = (1-\theta(z)) \int_{\mathbb{T}} \frac{f(\zeta)}{1-z\bar{\zeta}}\sigma(d\zeta).$$

One is interested in those measures μ on $\mathbb{T}$ for which the natural embedding operator is bounded from K_θ to $L^2(\mu)$, namely, is it the case that $\|F\|_\mu \lesssim \|F\|_{K_\theta}$. We see that this bound is equivalent to

$$\int_{\mathbb{T}} \left| \int_{\mathbb{T}} \frac{f(\zeta)}{1-z\bar{\zeta}}\sigma(d\zeta) \right|^2 |1-\theta(z)|^2 \mu(dz) \lesssim \|f\|_\sigma^2.$$

That is, the question is equivalent to a two weight inequality for the Hilbert transform on $\mathbb{T}$.

From this perspective, one can lift counterexamples concerning the two weight Hilbert transform to those for embedding operators, which is the tactic of [40], from which we have taken this condensed presentation. A characterization of the embedding question can be read off from our main theorem.

But note that Clark measure is on $\mathbb{T}$, by definition, and the second measure μ is constrained to be supported on $\mathbb{T}$, whereas the disk would be the natural assumption. In the case where μ is supported on the disk, and one seeks an *isometric* embedding, the question has a remarkable answer, found by Aleksandrov [1]. The general question is resolved in [24], which gives a characterization of a two weight inequality for the Cauchy transform, under these restrictions on the supports of the weights. The method of the proof is similar to that of the Hilbert transform, with some additional complications.

The model spaces are also important to spectral theory, and the subject of rank one perturbations of a unitary operator. In spectral theory, it is important to understand the structure of the unitary operator that sends the Hilbert space to into L^2 of the spectral measure. Weighted Hilbert transforms arise therein. See for instance [42], which uses the example of Nazarov showing that the A_2 condition is not sufficient for the boundedness of the Hilbert transform. Also see [30].

We point the interested readers to [41, 46], and the many citations therein for more information about these subjects.

de Branges Spaces

We recall the setting of [3, 4]. For a sequence of distinct points $\Gamma = \{\gamma_n\} \subset \mathbb{C}$ and a sequence of positive numbers $v = \{v_n\}$ consider the Cauchy transform

$$H_{(\Gamma,v)} \ : \ a = \{a_n\} \mapsto \sum_{n\,:\,z\neq\gamma_n} \frac{a_n v_n}{z-\gamma_n}$$

This is well defined for $a \in \ell^2_v$ and $z \in \Omega$, defined by

$$\Omega := \Big\{ z \in \mathbb{C} \ : \ \sum_{n\,:\,z\neq\gamma_n} \frac{v_n}{|z-\gamma_n|^2} < \infty \Big\}.$$

Call $\mathcal{H}(\Gamma, v)$ the space of functions analytic on Ω given by the image of ℓ^2_v under $H_{(\Gamma,v)}$. For appropriate choices of (Γ, v), these Hilbert spaces have deep connections to analytic function spaces. For instance, the reproducing kernels of $\mathcal{H}(\Gamma, v)$ are

$$k_z(\zeta) := \sum_n \frac{v_n}{(\overline{z-\gamma_n})(\zeta-\gamma_n)}, \qquad z \in \Omega\,.$$

And, many natural questions, such as the structure of frames of reproducing kernels for $\mathcal{H}(\Gamma, v)$, require knowledge about the two weight inequality for the Cauchy transform. For instance, the main real-variable result in [4] is a characterization of a two weight inequality, but under the requirement that both measures be a sum of point masses on sparse collections of points. This yields interesting results in the setting of de Branges spaces.

The definition of $\mathcal{H}(\Gamma, v)$ provides just one possible representation of a de Branges space, a class of Hilbert spaces with remarkable properties. The standard reference for them is [9]. Beginning from the works of Sarason [50], they have become an essential part of subject of analytic function spaces.

References

1. A.B. Aleksandrov, Isometric embeddings of co-invariant subspaces of the shift operator. Zap. Nauchn. Sem. S. Peterburg Otdel. Mat. Inst. Steklov. (POMI) **232** (1996), no. Issled. po Linein. Oper. i Teor. Funktsii. **24**, 5–15, 213 (Russian, with English and Russian summaries); English transl., J. Math. Sci. (New York) **92**(1), 3543–3549 (1998)
2. A. Aleman, S. Pott, M.C. Reguera, Sarason conjecture on the Bergman space, Available at http://www.arxiv.org/abs/1304.1750
3. Y. Belov, T.Y. Mengestie, K. Seip, Unitary discrete Hilbert transforms. J. Anal. Math. **112**, 383–393 (2010)
4. Y. Belov, T.Y. Mengestie, K. Seip, Discrete Hilbert transforms on sparse sequences. Proc. Lond. Math. Soc. (3) **103**(1), 73–105 (2011)
5. L. Carleson, On convergence and growth of partial sums of Fourier series. Acta Math. **116**, 135–157 (1966)
6. C. Cascante, J.M. Ortega, I.E. Verbitsky, On L^p-L^q trace inequalities. J. Lond. Math. Soc. (2) **74**(2), 497–511 (2006)
7. D. Cruz-Uribe, The invertibility of the product of unbounded Toeplitz operators. Integr. Equ. Oper. Theory. **20**(2), 231–237 (1994)
8. D. Cruz-Uribe, J. Martell, C. Pérez, A note on the off-diagonal Muckenhoupt-Wheeden conjecture, 2012, Available at http://www.arxiv.org/abs/1203.5906
9. L. de Branges (ed.), *Hilbert Spaces of Entire Functions* (Prentice-Hall Inc., Englewood Cliffs, 1968)
10. C. Fefferman, Pointwise convergence of Fourier series. Ann. Math. (2) **98**, 551–571 (1973)
11. R. Hunt, B. Muckenhoupt, R. Wheeden, Weighted norm inequalities for the conjugate function and Hilbert transform. Trans. Am. Math. Soc. **176**, 227–251 (1973)
12. T.P. Hytönen, On Petermichl's dyadic shift and the Hilbert transform. C. R. Math. Acad. Sci. Paris **346**(21–22), 1133–1136 (2008). (English, with English and French summaries)
13. T.P. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators. Ann. Math. (2) **175**(3), 1473–1506 (2012)
14. T.P. Hytönen, The A_2 theorem: remarks and complements, in *Harmonic Analysis and Partial Differential Equations*. Contemporary Mathematics, vol. 612 (American Mathematical Society, Providence, 2014), pp 91–106. doi:10.1090/conm/612/12226. MR3204859
15. T.P. Hytönen, The two weight inequality for the Hilbert transform with general measures, 2013, Available at http://www.arxiv.org/abs/1312.0843
16. M.T. Lacey, An elementary proof of the A_2 Bound. Isr. J. Math. **217**, 181 (2017). doi:10.1007/s11856-017-1442-x
17. M.T. Lacey, Two-weight inequality for the Hilbert transform: a real variable characterization, II. Duke Math. J. **163**(15), 2821–2840 (2014)
18. M.T. Lacey, S. Petermichl, M.C. Reguera, Sharp A_2 inequality for Haar shift operators. Math. Ann. **348**(1), 127–141 (2010)
19. M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, A characterization of two weight norm inequalities for maximal singular integrals with one doubling measure. Anal. PDE **5**(1), 1–60 (2012)
20. M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, A two weight inequality for the Hilbert transform assuming an energy hypothesis. J Funct. Anal. **263**, 305–363 (2012)
21. M.T. Lacey, E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, The two weight inequality for the Hilbert transform, coronas, and energy conditions, (2011), Available at http://www.arXiv.org/abs/118.2319
22. M.T. Lacey, E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, Two weight inequality for the Hilbert transform: a real variable characterization. https://arxiv.org/abs/1201.4319v6
23. 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**(15), 2795–2820 (2014)
24. M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, C.-Y. Shen, B. Wick, Two weight inequalities for the Cauchy transform from **R** to **C**+, 2013, Available at http://arxiv.org/abs/1310.4820

25. M.T. Lacey, C. Thiele, A proof of boundedness of the Carleson operator. Math. Res. Lett. **7**, 361–370 (2000)
26. M.T. Lacey, B. Wick, Two weight inequalities for Riesz transforms: uniformly full dimension weights, 2013, Available at http://arxiv.org/abs/1312.6163
27. J. Lai, S. Treil, Two weight L^p estimates for paraproducts in non-homogeneous settings, Available at http://arxiv.org/abs/1507.05570
28. A.K. Lerner, A pointwise estimate for the local sharp maximal function with applications to singular integrals. Bull. Lond. Math. Soc. **42**(5), 843–856 (2010)
29. A.K. Lerner, A simple proof of the A_2 conjecture. Int. Math. Res. Not. IMRN **14**, 3159–3170 (2013). MR3085756
30. C. Liaw, S. Treil, Rank one perturbations and singular integral operators. J. Funct. Anal. **257**(6), 1947–1975 (2009)
31. C. Liaw, S. Treil, Regularizations of general singular integral operators. Rev. Mat. Iberoam. **29**(1), 53–74 (2013)
32. B. Muckenhoupt, Hardy's inequality with weights. Studia Math. **44**, 31–38 (1972). Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I
33. B. Muckenhoupt, R.L. Wheeden, Two weight function norm inequalities for the Hardy-Littlewood maximal function and the Hilbert transform. Studia Math. **55**(3), 279–294 (1976)
34. F. Nazarov, A counterexample to Sarason's conjecture. Preprint, MSU, 1997, Available at http://wwwmathmsuedu/~fedja/prepr.html
35. F. Nazarov, S. Treil, A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces. Int. Math. Res. Notices **15**, 703–726 (1997)
36. F. Nazarov, S. Treil, A. Volberg, Accretive system Tb-theorems on nonhomogeneous spaces. Duke Math. J. **113**(2), 259–312 (2002)
37. F. Nazarov, S. Treil, A. Volberg, The Tb-theorem on non-homogeneous spaces. Acta Math. **190**(2), 151–239 (2003)
38. F. Nazarov, S. Treil, A. Volberg, Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures, 2004, Available at http://arxiv.org/abs/1003.1596
39. F. Nazarov, S. Treil, A. Volberg, Two weight inequalities for individual Haar multipliers and other well localized operators. Math. Res. Lett. **15**(3), 583–597 (2008)
40. F. Nazarov, A. Volberg, The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces $K\theta$. J. Anal. Math. **87**, 385–414 (2002). Dedicated to the memory of Thomas H. Wolff
41. N. K. Nikol'skiĭ, in *Treatise on the Shift Operator*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 273 (Springer, Berlin, 1986). Spectral function theory; With an appendix by S.V. Hruščev [S.V. Khrushchëv], V.V. Peller; Translated from the Russian by Jaak Peetre
42. N. Nikolski, S. Treil, Linear resolvent growth of rank one perturbation of a unitary operator does not imply its similarity to a normal operator. J. Anal. Math. **87**, 415–431 (2002). Dedicated to the memory of Thomas H. Wolff
43. S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol. C. R. Acad. Sci. Paris Sér. I Math. **330**(6), 455–460 (2000). (English, with English and French summaries)
44. S. Petermichl, The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. Am. J. Math. **129**(5), 1355–1375 (2007)
45. C. Pérez, S. Treil, A. Volberg, OnA_2 conjecture and Corona decomposition of weights, 2010, Available at http://arxiv.org/abs/1006.2630
46. A. Poltoratski, D. Sarason, Aleksandrov-Clark measures, in *Recent Advances in Operator-Related Function Theory*. Contemporary Mathematics, vol. 393 (American Mathematical Society, Providence, 2006), pp. 1–14
47. M.C. Reguera, On Muckenhoupt-Wheeden conjecture. Adv. Math. **227**(4), 1436–1450 (2011)
48. M.C. Reguera, J. Scurry, On joint estimates for maximal functions and singular integrals on weighted spaces. Proc. Am. Math. Soc. **141**(5), 1705–1717 (2013)

49. M.C. Reguera, C. Thiele, The Hilbert transform does not map $L^1(Mw)$ to $L^{1\infty}(w)$. Math. Res. Lett. **19**(1), 1–7 (2012)
50. D. Sarason, Exposed Points in H^1. I. The Gohberg anniversary collection, vol. II (Calgary, AB, 1988), Operator Theory Advances and Applications, vol. 41 (Birkhäuser, Basel, 1989), pp. 485–496
51. D. Sarason, in *Products of Toeplitz Operators*, Linear and Complex Analysis. Problem Book 3. Part I, ed. by V.P. Havin, N.K. Nikolski. Lecture Notes in Mathematics, vol. 1573 (Springer, Berlin, 1994), pp. 318–319
52. E.T. Sawyer, A characterization of a two-weight norm inequality for maximal operators. Studia Math. **75**(1), 1–11 (1982)
53. E.T. Sawyer, A characterization of two weight norm inequalities for fractional and Poisson integrals. Trans. Am. Math. Soc. **308**(2), 533–545 (1988)
54. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A note on failure of energy reversal for classical fractional singular integrals. Int. Math. Res. Not. IMRN **2015**(19), 9888–9920
55. H. Tanaka, A characterization of two-weight trace inequalities for positive dyadic operators in the upper triangle case. Potential Anal. **41**(2), 487–499 (2014). doi:10.1007/s11118-013-9379-0. MR3232035
56. A. Volberg, *Calderón-Zygmund Capacities and Operators on Nonhomogeneous Spaces*. CBMS Regional Conference Series in Mathematics, vol. 100 (Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2003)
57. E. Vuorinen, $L^p(\mu) \to L^q(\nu)$ characterization for well localized operators. J. Fourier Anal. Appl. **22**(5), 1059–1075 (2016)
58. E. Vuorinen, Two weight L^p-inequalities for dyadic shifts and the dyadic square function. Studia Math. **237**(1), 25–56 (2017). 42B20 (42B25)

Singular Integrals, Rank One Perturbations and Clark Model in General Situation

Constanze Liaw and Sergei Treil

To the memory of Cora Sadosky

Abstract We start with considering rank one self-adjoint perturbations $A_\alpha = A + \alpha(\cdot, \varphi)\varphi$ with cyclic vector $\varphi \in \mathcal{H}$ on a separable Hilbert space $\mathcal{H}$. The spectral representation of the perturbed operator A_α is realized by a (unitary) operator of a special type: the Hilbert transform in the two-weight setting, the weights being spectral measures of the operators A and A_α.

Similar results will be presented for unitary rank one perturbations of unitary operators, leading to singular integral operators on the circle.

This motivates the study of abstract singular integral operators, in particular the regularization of such operator in very general settings.

Further, starting with contractive rank one perturbations we present the Clark theory for arbitrary spectral measures (i.e. for arbitrary, possibly not inner characteristic functions). We present a description of the Clark operator and its adjoint in the general settings. Singular integral operators, in particular the so-called normalized Cauchy transform again plays a prominent role.

Finally, we present a possible way to construct the Clark theory for dissipative rank one perturbations of self-adjoint operators.

The work of Constanze Liaw is supported by the Simons Foundation Collaboration Grant for Mathematicians #426258. Work of S. Treil is supported by the National Science Foundation under the grant DMS-1301579.

C. Liaw (✉)
CASPER and Department of Mathematics, Baylor University, One Bear Place #97328, Waco, TX 76798, USA
e-mail: Constanze_Liaw@baylor.edu
http://sites.baylor.edu/constanze_liaw/

S. Treil
Department of Mathematics, Brown University, 151 Thayer Str./Box 1917, Providence, RI 02912, USA
e-mail: treil@math.brown.edu
http://www.math.brown.edu/~treil

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_4

These lecture notes give an account of the mini-course delivered by the authors, which was centered around (Liaw and Treil, J Funct Anal 257(6):1947–1975, 2009; Rev Mat Iberoam 29(1):53–74, 2013; J Anal Math). Unpublished results are restricted to the last part of this manuscript.

2010 Mathematics Subject Classification. 42B20, 44A15, 47A10, 47A20, 47A55.

Introduction

Rank one perturbations play an important role in operator theory and mathematical physics. One of the principal attractions of rank one perturbations is that for such operators almost everything can be explicitly computed, and then the advanced technique of Harmonic Analysis, like the study of fine properties of Cauchy type integrals, or advanced theory of singular integral operators can be applied.

Main Players

Rank One Perturbations

Self-adjoint rank one perturbations occurred naturally in mathematical physics [45]. For example, a change in the boundary condition of a limit-point half-line Schrödinger operator from Dirichlet to Neumann, or to mixed conditions, can be reformulated as adding a rank one perturbation (see for example [40]).

The technique of rank one perturbations was used in some results on orthogonal polynomials and Jacobi matrices, and there are some interesting applications to free probability (see e.g. [9, 10]). They also turned out to be useful in the investigation of certain random Hamiltonian systems called Anderson models and the longstanding Anderson localization conjecture [8]. Many specializations of this conjecture were studied in literature and the field is still very active (see e.g. [2, 17, 19–21], also see [28, 46] for a recent account of parts of the field). Rank one perturbations play a role in [1, 27, 30, 41]. Recent studies of closely related unitary Anderson models as well as accessible explanations of the physical relevance of these models can be found, e.g. in [23–25, 42]. The additive perturbation is replaced by a multiplicative one and the dynamical localization behavior is known to be quite similar to its self-adjoint analogue.

Singular Integral Operators

Singular integrals is a classical and actively studied field in Harmonic Analysis, and rank one perturbations serves as a source of very interesting problems. Many results for rank one perturbations are obtained by investigating fine properties of singular integrals. For example, investigation of the boundary behavior of the Cauchy transform of measures lead via the so-called Aronszajn–Krein formula, see (4), to the famous Aronszajn–Donoghue theorem stating that singular parts of the spectral measures of the family of rank one perturbation by a cyclic vector are mutually singular.

As for a different example, basic facts about Cauchy transform of a measure coupled with the Aronszajn–Krein formula (4) give a proof of the famous Kato–Rosenblum theorem about preservation of the absolutely continuous spectrum for rank one (and automatically for finite rank) perturbations. While the proof for the trace class perturbations using the technique of wave operator is probably more elegant, the approach of singular integrals gives some helpful insights.

A deep relation between singular integral operators and rank one perturbations is based on the fact that a unitary operator realizing the spectral representation of a rank one perturbation is given by a singular integral operator acting $L^2(\mu) \to L^2(\mu_\alpha)$, where μ_α is the spectral measure of the perturbed operator, see [32]; we explain this connection in the present notes. We should mention here that the spectral measures μ and μ_α can be extremely bad, without any reasonable smoothness, so the above operator gives a natural example of a two weight estimates for Cauchy type operators with extremely "pathological" measures.

Clark Measures and Clark Model

In the paper [12] that started what is now called the "Clark theory" D. Clark considered all unitary rank one perturbations of a special contractive operator (the so-called *model operator* with scalar inner characteristic function). He also described the spectral measures and the spectral representations of the perturbed unitary operators.

The spectral measures of these unitary rank one perturbations were later called the *Clark measures*. Note, that if we fix one such rank one unitary perturbation, then the other unitary rank one perturbations are the rank one perturbations of the fixed one. In the original paper [12] all the spectral measures were purely singular, but very often the term *Clark measures* (or Clark family of measures) was used for spectral measures of unitary rank one perturbations of a unitary operator, or for the spectral measures of self-adjoint rank one perturbations of a self-adjoint operator.

Later many deep function-theoretic results about Clark measures were proved by A. Aleksandrov, (see [3–7], or see [37] for a survey), so sometimes people refer to *Aleksandrov–Clark theory*, or *Aleksandrov–Clark measures*. Extremely significant contributions to the theory were made then by A. Poltoratskii, who, in particular, proved the a.e. existence with respect to the singular part of the measure of the non-tangential boundary values of the so-called *normalized Cauchy transform*, see [38].

We also mention an important book [39] by D. Sarason where many aspects of Clark theory were treated from the point of view of function space theory. In particular, a description of the Clark operator was obtained in the case when the characteristic function θ is an extreme point of the unit ball in the Hardy space H^∞.

Within classical analysis many fruitful connections of Clark measures with holomorphic composition operators, rigid functions and the Nehari interpolation problem have been discovered and studied, see for example, [37]. Some problems in the theory of Hardy spaces, and more generally of other spaces of analytic functions are closely related to Clark theory. Thus, recently M. Jury [26] computed the asymptotic symbols of a certain class of weakly asymptotic Toeplitz operators in terms of the Aleksandrov–Clark measures which occur in the context of rank one perturbations.

Plan of the Notes

These lecture notes give an account of the mini-course delivered by the authors, which was centered around [31–33]. Unpublished results are restricted to the last part of this manuscript.

Self-Adjoint and Unitary Rank One Perturbations

We begin section "Self-Adjoint and Unitary Rank One Perturbations" with an introduction of self-adjoint rank one perturbations. We then find a unitary operator V_α giving the spectral representation of the perturbed operator, see Theorem 2.1 below; this operator looks like a singular integral operator with Cauchy type kernel $(s-t)^{-1}$, although the formula of the operator looks quite different from the classical singular integral operators of Cauchy type.

In particular, the so-called *Rigidity Theorem*, see Theorem 2.2 below, holds for such operators: it essentially says that if the formula from Theorem 2.1 gives a bounded operator with trivial kernel, then, after probably a renormalization (multiplication by a non-vanishing weight) of the measure in the target space, we get exactly the unitary operator from the perturbation theory, given by Theorem 2.1.

We then give a different representation of the operator V_α that looks more in line with the traditional formulas for singular integral operators. Regularizations of singular kernels, treated later in section "Singular Integral Operators", play an important role in getting this alternative representation.

We then present similar results for the unitary rank one perturbations of unitary operators. Everything works out similarly to the self-adjoint case; some formulas for the unitary case might not look as transparent as the ones in the self-adjoint case, but in the unitary case we avoid technical difficulties related to dealing with unbounded operators.

Regularizations of Singular Integral Operators

Section "Singular Integral Operators" is devoted to the theory of regularization of singular kernels, which we believe have applications far beyond the perturbation theory. We show that under very general assumptions about a singular kernel, its so-called *restricted boundedness* implies the uniform boundedness of all "reasonable" regularizations of the corresponding formal singular integral operator.

The restricted boundedness of the kernel is the weakest boundedness property of the corresponding singular integral operator. Usually, it is assumed in the theory of singular integral operators that a singular kernel K blows up on the diagonal $x = y$, so the formal integral representation $Tf(x) = \int K(x,y)f(y)d\mu(y)$ is not well defined.

However, even if we only start out with a kernel K (without assuming the we are given an operator) for bounded functions f and g with separated compact supports the expression

$$\langle Tf, g\rangle = \int K(x,y)f(y)g(x)d\mu(y)d\nu(x)$$

is well defined, and if the "correct" estimate $|\langle Tf, g\rangle| \le C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\nu)}$, $1/p + 1/p' = 1$ holds for all such pairs, we say that K is $L^p(\mu) \to L^p(\nu)$ restrictedly bounded. And we show in section "Singular Integral Operators" that if the measures μ and ν do not have common atoms and the kernel K is restrictedly bounded, then for any "reasonable" regularization K_ε of the kernel the corresponding regularized operators T_ε are uniformly (in ε) bounded. This result gives us a way to define for each restrictedly bounded kernel a corresponding singular integral operator.

Clark Model for Contractive Perturbations of Unitary Operators

Section "Clark Theory for Rank One Perturbations of Unitary Operators" is devoted to the Clark theory in full generality. We start with unitary rank one perturbations of a unitary operator U by a $*$-cyclic vector. All such perturbations can be parametrized by a scalar parameter $\gamma \in \mathbb{T}$; if one takes $\gamma \in \mathbb{D}$ the resulting operator will be a completely non-unitary (c.n.u.) contraction with defect indices 1-1. For such a contraction a so-called *functional model*, cf. [43] can be constructed; in fact functional models are the canonical way of investigating non-normal contractions.

Thus, the perturbed operator U_γ, $\gamma \in \mathbb{D}$ has two unitarily equivalent representations: one in the spectral representation of U and the other one in the model space for the functional model. The Clark operator is a unitary operator intertwining these representations. In Clark's original paper [12] this operator was constructed for the case of the operator U having purely singular spectrum. In [12] the starting point was a c.n.u. contraction with inner characteristic function, which — after translation to our language — means that the unitary operator U (and thus all its rank one unitary perturbations U_γ) has a purely singular spectral measure.

In the general case (general spectral measure, or equivalently, a general scalar characteristic function) our approach of starting with perturbations of unitary operators looks more natural; in particular, it allowed us to describe the Clark operator. Of course, now when we know all the formulas, it is possible to go in the opposite direction and start with a c.n.u. contraction; but using this approach without knowing the formulas in advance we would have a hard time getting the results. It could well be just our personal preference, but deducing the formulas in our setup starting from a unitary operator was a natural and a straightforward process.

The main problem with the general case of Clark theory is that for general scalar characteristic function the model is vector-valued, i.e. the model space consists of vector-valued functions (with values in $\mathbb{C}^2$). Earlier approaches based on function spaces theory, see for example [39], dealt with spaces of scalar-valued functions. Some parts of the Clark operator were obtained using such methods, but for the full operator one had to honestly write down a complete model space and do all the computations.

The adjoint of the Clark operator is described using singular integral operators of Cauchy type. The so-called *normalized Cauchy transform* investigated by A. Poltoratskii, see [38], plays a prominent role there. The Clark operator itself then can be represented via boundary values of the analytic functions.

In the model theory we adapt the point of view of *coordinate-free model* by N. Nikolski and V. Vasyunin, cf. [35, 36], where by picking different spectral representations of the minimal unitary dilation one gets different *transcriptions* of the model. We present a "universal" representation formula, valid in any transcription, as well as formulas adapted to two popular transcriptions, the Sz.-Nagy–Foiaş transcription and the de Branges–Rovnyak one.

Clark Model for Dissipative Perturbations of Self-Adjoint Operators

The last part, section "Few Remarks About Clark Theory for the Dissipative Case" is devoted to the Clark model for the dissipative perturbations of a self-adjoint operator. We adapt a common approach that the model space for a dissipative operator is the model space of its Cayley transform (which is a contraction), with one detail: since the original operator lives in $L^2(\mathbb{R}, \mu)$, we, using the standard conformal map between the upper half-plane $\mathbb{C}_+$ and the unit disc $\mathbb{D}$, move the model space to the real line (half-plane). The results in this section were not presented before.

Note, that the formulas in this section do not look as elegant as in the case of perturbations of unitary operators. Probably, a different approach to the model of dissipative operators would be more appropriate, but we do not know a serious contender yet. As a pure speculation, the de Branges spaces $\mathfrak{L}(\varphi)$ could serve as appropriate model spaces for the dissipative perturbations. These spaces were introduced in the first chapter of [13], but were not much investigated, unlike the spaces $\mathcal{H}(E)$ which were investigated in details in [13] and were subject of extensive research by many authors.

Self-Adjoint and Unitary Rank One Perturbations

Self-Adjoint Rank One Perturbations

For a self-adjoint (possibly unbounded) operator A on a separable Hilbert space $\mathcal{H}$ let us consider the family of rank-one perturbations A_α, $\alpha \in \mathbb{R}$, given by

$$A_\alpha := A + \alpha(\,\cdot\,, \varphi)_{\mathcal{H}}\varphi \qquad \text{on } \mathcal{H}. \tag{1}$$

Here, if the operator A is bounded, then φ is a vector in $\mathcal{H}$. For unbounded A, we can consider the wider class of "singular form-bounded" perturbations where we assume $\varphi \in \mathcal{H}_{-1}(A) \supset \mathcal{H}$, where $\mathcal{H}_{-r}(A)$, $r \in \mathbb{N}$, is the completion of $\mathcal{H}$ with respect to the norm $\|\cdot\|_{\mathcal{H}_{-r}(A)}$, $\|f\|_{\mathcal{H}_{-r}(A)} = \|(I + |A|)^{-r/2} f\|_{\mathcal{H}}$. In particular, the perturbation $\alpha(\,\cdot\,, \varphi)\varphi$ can be unbounded (see [29, 32] and the references within for further details). If $\mathcal{H} = L^2(\mathbb{R}, \mu)$ and $A = M_t$ is the multiplication by the independent variable,

$$M_t f(t) = tf(t), \qquad \forall t \in \mathbb{R},$$

then $\mathcal{H}_{-1}(A)$ is exactly the collection of measurable functions such that

$$\int_{\mathbb{R}} \frac{|f(t)|^2}{1 + |t|} d\mu(t) < \infty.$$

For $r \geq 2$ the formal expression (1) does not uniquely determine a self-adjoint operator: For fixed α there is a family of self-adjoint operators corresponding to (1). For this reason we do not consider this case, but rather assume that $r < 2$.

Without loss of generality we can assume that φ is cyclic for A, that is,

$$\mathcal{H} = \operatorname{clos\,span}\{(A - \lambda\mathbf{I})^{-1}\varphi : \lambda \in \mathbb{C} \setminus \mathbb{R}\}.$$

Otherwise, i.e. if $\widetilde{\mathcal{H}} = \operatorname{clos\,span}\{(A - \lambda\mathbf{I})^{-1}\varphi : \lambda \in \mathbb{C} \setminus \mathbb{R}\} \subsetneq \mathcal{H}$, then we restrict our attention to the action on $\widetilde{\mathcal{H}}$ as the perturbation is trivial (does nothing) on $\mathcal{H} \ominus \widetilde{\mathcal{H}}$.

Then according to the Spectral Theorem the operator A is unitarily equivalent to the multiplication M_t by the independent variable in a space $L^2(\mu)$ where μ is a spectral measure of the operator A. Spectral measure is of course not unique, multiplying a spectral measure by a non-vanishing weight (i.e. by a function $w \in L^1_{\text{loc}}(\mu)$, $w > 0$ μ-a.e.) we get a different spectral measure.

It is customary in the operator theory and mathematical physics to consider the *canonical* spectral measure to be the spectral measure associated with the "vector" φ, i.e. the unique measure μ such that

$$F(\lambda) := ((A - \lambda\mathbf{I})^{-1}\varphi, \varphi) = \int_{\mathbb{R}} \frac{d\mu(x)}{x - \lambda}.$$

In this case φ is represented by the function $\mathbf{1}$, and the assumption that $\varphi \in \mathcal{H}_{-1}(A)$ means simply that $\int_{\mathbb{R}}(1+|x|)^{-1}d\mu(x) < 1$.

Knowing function F one can say a lot about the spectral measure μ: since the imaginary part of F is (up to the factor π) the Poisson integral of μ we can immediately conclude that the density of the absolutely continuous part of μ is given by the non-tangential boundary values of $\pi^{-1}\operatorname{Im} F(z)$ (such values exist a.e. by classical results). It is also not hard to show that the singular part of μ is supported on a set where the (non-tangential) boundary values of $\operatorname{Im} F$ are infinite.

In the heart of the theory of rank one perturbations lies the simple fact that there is a simple relation between the function F and the corresponding functions F_α for the perturbed operators.

Namely, the following simple formula for the inverse of the rank one perturbation of the identity is well known

$$\Big(\mathbf{I} - (\,\cdot\,, b)a\Big)^{-1} = \mathbf{I} + \frac{1}{d}\,(\,\cdot\,, b)a; \tag{2}$$

here $d = 1 - (b, a)$ is the so-called perturbation determinant, and the operator is invertible if and only if $d \neq 0$. The proof of this formula is an easy exercise, we leave it to the reader.

Using the above formula (2) one can easily compute the resolvent of the perturbed operator A_α,

$$(A_\alpha - \lambda\mathbf{I})^{-1}f = (A - \lambda\mathbf{I})^{-1}f - \frac{\alpha\left((A-\lambda\mathbf{I})^{-1}f, \varphi\right)}{1+\alpha\left((A-\lambda\mathbf{I})^{-1}\varphi, \varphi\right)}(A-\lambda\mathbf{I})^{-1}\varphi \tag{3}$$

which immediately implies the relation between the function F and the corresponding functions F_α, $F_\alpha(\lambda) := ((A_\alpha - \lambda\mathbf{I})^{-1}\varphi, \varphi)$, commonly known as the Aronszajn–Krein formula:

$$F_\alpha = \frac{F}{1+\alpha F}\,. \tag{4}$$

If μ_α denotes the spectral measure of the perturbed operator A_α associated with φ, then

$$F_\alpha(\lambda) = \int_{\mathbb{R}} \frac{d\mu_\alpha(x)}{x-\lambda}.$$

(Note that it is not hard to show that if φ is cyclic for A, then φ is also cyclic for A_α and therefore A_α is unitarily equivalent to the multiplication operator M_t in $L^2(\mu_\alpha)$).

Many classical results in perturbation theory can be obtained from the Aronszajn–Krein formula (4) and classical results about boundary values of the Cauchy transform.

For example, it is not hard to show that all the absolutely continuous parts of the measures μ_α are equivalent (i.e. mutually absolutely continuous), which is just the Kato–Rosenblum theorem for rank one perturbations. Also, the analysis of the singular parts of the measures μ_α yields the famous Aronszajn–Donoghue theorem, stating that the singular parts of μ_α are mutually singular.

Rank One Perturbations and Singular Integral Operators

We find a sufficient condition on the absence of singular spectrum by studying the spectral representation, which comes in the form of a two weight Hilbert transform. Part of this material can be understood as a first example for section "Singular Integral Operators".

Consider a family of rank one perturbations given by $A_\alpha := A + \alpha(\cdot, \varphi)\varphi$, see (1), where $\varphi \in \mathcal{H}_{-1}(A)$ is cyclic for A. Let μ denote the spectral measure of operator A with respect to φ, so A is unitarily equivalent to the multiplication operator M_t in $L^2(\mu)$. Let us consider the operator A in its spectral representation, i.e. let us assume that A *is* the multiplication operator M_t in $L^2(\mu)$. As we discussed before, the assumption that $\varphi \in \mathcal{H}_{-1}(A)$ means simply that $\int_{\mathbb{R}}(1 + |x|)^{-1} d\mu(x) < \infty$,

Then the operator A_α is defined by

$$A_\alpha = A + \alpha(\cdot, \varphi)\varphi = M_t + \alpha(\,\cdot\,, \mathbf{1})_{L^2(\mu)}\mathbf{1}$$

on $L^2(\mu)$. On the other hand, the operator A_α is unitarily equivalent to the multiplication M_s by the independent variable s in $L^2(\mu_\alpha)$ (we use a different letter for the independent variable here to distinguish between the multiplication operators in $L^2(\mu)$ and $L^2(\mu_\alpha)$).

We want to find a unitary operator giving the spectral representation of the operator A_α, i.e. a unitary operator

$$V_\alpha : L^2(\mu) \to L^2(\mu_\alpha)$$

such that

$$V_\alpha A_\alpha = M_s V_\alpha.$$

We also want φ to be represented by $\mathbf{1}$ in both representations, which translates to additional condition $V_\alpha \varphi = \mathbf{1}$.

Theorem 2.1 of [32] gives the representation of V_α as the Hilbert transform type singular integral.

Theorem 2.1 (Representation Theorem) *Under the above assumptions the spectral representation $V_\alpha : L^2(\mu) \to L^2(\mu_\alpha)$ of A_α is given by*

$$V_\alpha f(s) = f(s) - \alpha \int \frac{f(s) - f(t)}{s - t}\, d\mu(t) \tag{5}$$

for all compactly supported C^1 functions f.

Without going into the details of the proof, we indicate the proof strategy for bounded operators A. The intertwining condition

$$M_s V_\alpha = V_\alpha A_\alpha = V_\alpha (M_t + \alpha(\,\cdot\,, \varphi)\varphi) \tag{6}$$

can be rewritten as

$$V_\alpha M_t = M_s V_\alpha - \alpha(\,\cdot\,, \varphi) V_\alpha \varphi.$$

Using induction we get

$$V_\alpha M_t^n = M_s^n V_\alpha - \alpha \sum_{k=0}^{n-1} (\,\cdot\,, M_t^k \varphi) M_s^{n-k-1} V_\alpha \varphi.$$

Recalling that $\varphi \equiv \mathbf{1}$, $V_\alpha \varphi \equiv \mathbf{1}$, we get by applying the above identity to φ and denoting $f_n(t) := t^n$ we get

$$V_\alpha f_n(s) = s^n - \alpha \int_{\mathbb{R}} \left(\sum_{k=0}^{n-1} t^k s^{n-k-1} \right) d\mu(t).$$

Summing the geometric progression under the integral we get the representation formula (5) for $f = f_n$, $f_n(t) = t^n$. Linearity of (5) implies that it holds for all polynomials, and rather standard approximation reasoning allows to extend this formula to the case of compactly supported C^1 functions.

This reasoning, of course, works only for bounded operators A (i.e. when the measure μ is compactly supported). In the case of unbounded operators the resolvent identity (3) is used instead of (6), see [32] for the details.

Aside we mention that integral operators represented by formula (5) are very interesting objects, probably deserving more careful investigation. Without proof we mention one property (see Theorem 2.2 of [32]), which can be understood as a converse to the last Representation Theorem.

Theorem 2.2 (Rigidity Theorem) *Let measure μ on $\mathbb{R}$ be supported on at least two distinct points and satisfy $\int (1 + |t|)^{-1}\, d\mu(t) < \infty$. Let V be defined on compactly supported C^1 functions f by formula (5).*

Assume V extends to a bounded operator from $L^2(\mu)$ to $L^2(\nu)$. Assume $\operatorname{Ker} V = \{0\}$.

Then there exists a function h such that $1/h \in L^\infty(\nu)$, and $M_h V$ is a unitary operator from $L^2(\mu) \to L^2(\nu)$ (equivalently, that $V : L^2(d\mu) \to L^2(|h|^2\, d\nu)$ is unitary).

Moreover, the unitary operator $U := M_h V$ gives the spectral representation of the operator $A_\alpha := M_t + \alpha(\,\cdot\,, \varphi)\varphi$, $\varphi \equiv \mathbf{1}$, in $L^2(\mu)$, namely $U A_\alpha = M_s U$, where M_s is the multiplication by the independent variable s in $L^2(\nu)$.

The integral in the representation formula looks like a singular integral operator, but not exactly in the traditional sense. The attempt to understand the precise connection with the theory of classical singular integral operators lead us to the theory of regularizations. We describe these results in more detail in section "Singular Integral Operators" on general abstract singular integral operators.

But now, let us first notice that

$$(V_\alpha f, g)_{L^2(\mu_\alpha)} = -\alpha \iint \frac{f(t)\overline{g(s)}}{s-t}\, d\mu(t)\, d\mu_\alpha(s)$$

for all $f \in L^2(\mu)$ and $g \in L^2(\mu_\alpha)$ with separated compact supports. This equality is trivial for compactly supported C^1 function f and g (with separated compact supports) and can be extended to the general case by a standard approximation argument.

Since $V_\alpha : L^2(\mu) \to L^2(\mu_\alpha)$ is further unitary, the kernel $K(s,t) = 1/(s-t)$ is what we call *restrictedly bounded* kernel, see Definition 3.1 below.

An application of Theorem 3.2 and Remark 3.3 shows the following result.

Theorem 2.3 *For the measures μ, μ_α as above, the operators $T_\varepsilon : L^2(\mu) \to L^2(\mu_\alpha)$,*

$$T_\varepsilon f(s) := \int_{\mathbb{R}} \frac{f(t)}{s-t+i\varepsilon} d\mu(t),$$

are uniformly (in ε) bounded.

Uniform boundedness of the operators T_ε implies that there exists a w.o.t. limit point of T_ε, as $\varepsilon \to 0$. In fact, it can be shown that this limit point is unique if $\varepsilon \to 0^+$ or $\varepsilon \to 0^-$, so we can say that there exist a w.o.t.-limits $T_\pm = \text{w.o.t.-}\lim_{\varepsilon\to 0^\pm} T_\varepsilon$.

The existence of w.o.t. limits follows, for example, from the lemma below the fact that for $\operatorname{Im} z > 0$ and for $\operatorname{Im} z < 0$ the non-tangential boundary values of $Rf\mu(z) := \int_{\mathbb{R}} \frac{f d\mu(t)}{t-z}$ exist μ_α-a.e.

Lemma 2.4 *For any $f \in L^2(\mu)$ the non-tangential boundary values of $Rf\mu(z) = \int_{\mathbb{R}} \frac{f d\mu(t)}{t-z}$, $z \in \mathbb{C}_+$ or $z \in \mathbb{C}_-$ exist μ_α-a.e.*

Proof The a.e. convergence with respect to Lebesgue measure (and so with respect to the absolutely continuous part of μ_α) follows from classical facts about boundary values of functions from Hardy spaces: for $f \geq 0$ the function $Rf\mu(z)$ has non-positive imaginary part, so composing it with a conformal mapping from the lower half-plane $\mathbb{C}_-$ we get a bounded analytic function, which has non-tangential limits on $\mathbb{R}$ a.e. with respect to Lebesgue measure. Representing arbitrary complex-valued function as linear combination of 4 non-negative ones we get the a.e. existence (with respect to Lebesgue measure) in general case.

To prove the convergence with respect to the singular part $(\mu_\alpha)_s$ of μ_α we get by applying functional φ to the resolvent formula (3) and denoting $f_\alpha = V_\alpha f$ that

$$Rf_\alpha \mu_\alpha = \frac{Rf\mu}{1+\alpha R\mu} .$$

By Poltoratkii's theorem, see [38, Theorem 2.7], the non-tangential boundary values of $Rf_\alpha\mu_\alpha/R\mu_\alpha$ exist (and coincide with f_α) $(\mu_\alpha)_{\mathrm{s}}$-a.e. Combining the above identity with the Aronszajn–Krein formula (4) we get

$$\frac{Rf_\alpha\mu_\alpha}{R\mu_\alpha} = \frac{Rf\mu}{R\mu}. \tag{7}$$

But it follows from the Aronszajn–Krein formula (4) that the non-tangential boundary values of $F = R\mu$ exist (and equal to $-1/\alpha$) $(\mu_\alpha)_{\mathrm{s}}$-a.e. Indeed

$$\operatorname{Im} F_\alpha = \frac{\operatorname{Im} F}{|1+\alpha F|^2}$$

and $\operatorname{Im} F_\alpha$ function is the Poisson extension (up to the factor π) of the measure μ_α. Therefore, since the singular part of the measure μ_α is supported on the subset of $\mathbb{R}$ where non-tangential boundary values of $\operatorname{Im} F_\alpha$ equal $+\infty$, we can conclude that the non-tangential boundary values of F equal $-1/\alpha$ $(\mu_\alpha)_{\mathrm{s}}$-a.e.

Since the non-tangential boundary values in (7) exist $(\mu_\alpha)_{\mathrm{s}}$-a.e., we conclude that same for $Rf\mu$. □

The above Lemma 2.4 implies the w.o.t. convergence of T_ε as $\varepsilon \to 0^+$ or $\varepsilon \to 0^-$. Indeed, Lemma 2.4 implies the μ_α-a.e. convergence, which, in turn implies that any weakly convergent subsequence of $T_\varepsilon f$ converges to the same function (the a.e. limit). And this, as one can easily see, means that $T_\varepsilon f$ has a weak limit as $\varepsilon \to 0^+$ or $\varepsilon \to 0^-$.

So, we can define the operators $T_\pm$ either as w.o.t. limits of T_ε as $\varepsilon \to 0^\pm$ or define $T_\pm f$ as the non-tangential boundary values of $Rf\mu(z)$, $z \in \mathbb{C}_\pm$.

Using the operators $T_\pm$ we obtain an alternative representation formula, see Theorem 3.2 of [32]:

Theorem 2.5 *Let μ and μ_α be the spectral measures of A and A_α, and let $T_\pm$ be as defined above.*

Then V_α can be written as

$$V_\alpha f(s) = f(s)(\mathbf{1} - \alpha\, T_\pm \mathbf{1}) + \alpha\, T_\pm f, \qquad \forall f \in L^2(\mu). \tag{8}$$

Proof Consider operators $V_\alpha^\varepsilon : L^2(\mu) \to L^2(\mu_\alpha)$,

$$V_\alpha^\varepsilon f(s) = f(s) - \alpha \int \frac{f(s)-f(t)}{s-t+i\varepsilon}\, d\mu(t) = f(s)(1-\alpha T_\varepsilon \mathbf{1}(s)) + \alpha T_\varepsilon f(s)$$

and notice that for compactly supported C^1 function f the functions $V_\alpha^\varepsilon f(s)$ converge uniformly and in $L^2(\mu_\alpha)$ to V_f as $\varepsilon \to 0$. Together with uniform bounds on T_ε this immediately implies that V_α^ε converges in the strong operator topology to V_α.

Taking w.o.t. limits we arrive to the representations (8). □

Remark Note that for the existence of the w.o.t. limits of T_ε it is sufficient to have μ_α-a.e. convergence on a dense set. As we just discussed above, for compactly supported $f \in C^1$ the functions $V_\alpha^\varepsilon f$ converge uniformly to $V_\alpha f$. It was also shown in the proof of Lemma 2.4 that $T_\varepsilon \mathbf{1}(s) = -F(s + i\varepsilon)$ converges μ_α-a.e., which immediately implies μ_α-a.e. convergence of $T_\varepsilon f$ for compactly supported C^1 functions.

This approach was used in [32].

Unitary Rank One Perturbations

In this section we present the analogues of the Representation Theorem 2.1 and the Rigidity Theorem 2.2 for the case of unitary rank one perturbations of unitary operators, that were proved in [31, Section 8].

We should mention, that these results cannot be obtained just by taking the Cayley transform of the self-adjoint case, we will explain this in section "Few Remarks About Clark Theory for the Dissipative Case".

In the contrast with the self-adjoint case the description of all unitary rank one perturbations of a unitary operator is not immediately self-evident, but with a little effort one could see that all unitary rank one perturbations of a unitary operator U can be parametrized as

$$U_{b,\alpha} = U + (\alpha - 1)(\,\cdot\,, U^*b)b \qquad b \in \mathcal{H},\ \|b\| = 1, \quad \alpha \in \mathbb{T}. \tag{9}$$

The fact that this formula indeed gives us the parametrization of the unitary rank one perturbations can be easily seen in the case $U = \mathbf{I}$; the general case then is obtained by right multiplying the formula for the perturbation of $\mathbf{I}$ by U.

In what follows we assume that the vector b is fixed and use the notation U_α for $U_{b,\alpha}$, so our perturbations will be parametrized by the scalar parameter $\alpha \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$.

Since the action of perturbation $(\,\cdot\,, U^*b)b$ is trivial (zero) on $(\operatorname{span}\{U^n b : n \in \mathbb{Z}\})^\perp$, we can ignore what is going on there and assume without loss of generality that b is $*$-cyclic vector for U, meaning that $\operatorname{span}\{U^n b : n \in \mathbb{Z}\} = \mathcal{H}$.

Then by the Spectral Theorem U is unitarily equivalent to the multiplication by the independent variable ξ in $L^2(\mu) = L^2(\mathbb{T}, \mu)$, where μ is a spectral measure of U. As in the self-adjoint case we fix a spectral measure μ to be the spectral measure corresponding to the vector b, so μ is a probability measure and the vector b in the spectral representation is given by the function $\mathbf{1}$.

So, as before let us assume that U is not just unitarily equivalent, but *is* a multiplication operator M_ξ by the independent variable ξ in $L^2(\mu) = L^2(\mathbb{T}, \mu)$, $\mu(\mathbb{T}) = 1$ and the rank unitary perturbations U_α are given by (9) with $b = \mathbf{1}$.

It is not hard to show that if b is $*$-cyclic for U then it is also $*$-cyclic for $U_\alpha = U_{b,\alpha}$, so U_α is unitarily equivalent to the multiplication M_z by the independent variable z in $L^2(\mu_\alpha)$. We take for μ_α the spectral measure corresponding to the vector b, so $b = \mathbf{1}$ in the spectral representation of U_α in $L^2(\mu_\alpha)$.

Under these assumptions we want to describe the unitary operator giving the unitary equivalence between U_α and its spectral representation of, i.e. the unitary operator $\mathcal{V}_\alpha : L^2(\mu) \to L^2(\mu_\alpha)$ such that $\mathcal{V}_\alpha \mathbf{1} = \mathbf{1}$ and

$$\mathcal{V}_\alpha U_\alpha = M_z \mathcal{V}_\alpha. \tag{10}$$

In Theorem 8.1 of [31] we proved:

Theorem 2.6 (Representation Theorem) *Let $\mathcal{V}_\alpha : L^2(\mu) \to L^2(\mu_\alpha)$ be a unitary operator satisfying* (10) *and such that $\mathcal{V}_\alpha \mathbf{1} = \mathbf{1}$ (which means that μ_α is the spectral measure of U_α corresponding to the cyclic vector b, $b(\xi) \equiv \mathbf{1}$). Then*

$$\mathcal{V}_\alpha f(z) = f(z) + (1-\alpha) \int_{\mathbb{T}} \frac{f(\xi) - f(z)}{1 - \bar{\xi} z}\, d\mu(\xi) \qquad \textit{for all } f \in C^1(\mathbb{T}). \tag{11}$$

Proof The proof goes similarly to the proof of the self-adjoint case (Theorem 2.1 above) for the bounded perturbations sketched above. Namely, using "linear algebra" notation, i.e. identifying $b \in \mathcal{H}$ with the operator $b : \mathbb{C} \to \mathcal{H}$, $b(\alpha) = \alpha b$ and denoting by b^* its adjoint $b^* : \mathcal{H} \to \mathbb{C}$, $b^*(x) = (x, b)_{\mathcal{H}}$ we can write

$$U_\alpha = U + (\alpha - 1) b b_1^* = M_\xi + (\alpha - 1) b b_1^*,$$

where $b_1 := U^* b$. Then the intertwining relationship (10) gives us

$$\mathcal{V}_\alpha U = M_z \mathcal{V}_\alpha + (1-\alpha)(\mathcal{V}_\alpha b) b_1^*. \tag{12}$$

Inductively one can show that for $n \geq 0$

$$\mathcal{V}_\alpha U^n = M_z^n \mathcal{V}_\alpha + (1-\alpha) \sum_{k=1}^{n} M_z^{k-1} (\mathcal{V}_\alpha b) \left((U^*)^{n-k} b_1\right)^*.$$

Applying this formula to the function $b \equiv \mathbf{1} \in L^2(\mu)$ and recalling that $(U^n b)(\xi) = \xi^n$, $\mathcal{V}_\alpha b = \mathbf{1}$, $b_1(\xi) \equiv \xi$, $(U_1^*)^{n-k} b_1 \equiv \xi^{n-k+1}$ we obtain summing the geometric series

$$(\mathcal{V}_\alpha \xi^n)(z) = z^n + (1-\alpha) \int_{\mathbb{T}} \frac{\xi^n - z^n}{1 - \bar{\xi} z}\, d\mu(\xi). \tag{13}$$

The action of $\mathcal{V}_\alpha$ on $\bar{\xi}^n$, $n \geq 0$ is proved similarly. Namely, taking the adjoint of the intertwining formula $\mathcal{V}_\alpha U_\alpha = M_z \mathcal{V}_\alpha$ and right and left multiplying by $\mathcal{V}_\alpha$ we get that $\mathcal{V}_\alpha U_\alpha^* = M_{\bar{z}} \mathcal{V}_\alpha$, so

$$\mathcal{V}_\alpha U^* = M_{\bar{z}} \mathcal{V}_\alpha + (1 - \overline{\alpha})(\mathcal{V}_\alpha b_1) b^*.$$

But that is exactly the intertwining relationship (12) with $U^* = U^{-1}$ instead of U and $M_{\bar{z}} = M_{z^{-1}}$ instead of M_z. So applying the same reasoning as above we get that (13) holds also for $n \leq 0$, and therefore for all trigonometric polynomials.

A standard approximation argument concludes the proof. □

A converse of the Representation Theorem is also true in the unitary setting. Under mild conditions bounded injective operators $\mathcal{V} : L^2(\mu) \to L^2(\nu)$ that are given by (11) induce a Clark family. More precisely, we quote Theorem 8.4 of [31].

Theorem 2.7 (Rigidity Theorem) *Let a probability measure μ on $\mathbb{T}$ be supported on at least two distinct points. Let $\alpha \in \mathbb{T}\setminus\{1\}$, and let $\mathcal{V}f$ be defined for C^1 functions f by the right hand side of (5).*

Assume $\mathcal{V}$ extends to a bounded operator from $L^2(\mu)$ to $L^2(\nu)$ and assume $\operatorname{Ker} \mathcal{V} = \{0\}$.

Then there exists a function h such that $1/h \in L^\infty(\nu)$, and $M_h \mathcal{V}$ is a unitary operator from $L^2(\mu) \to L^2(\nu)$ (equivalently, that $\mathcal{V} : L^2(d\mu) \to L^2(|h|^2\, d\nu)$ is unitary).

Moreover, the measure $|h|^2 \nu$ is exactly the Clark measure μ_α defined as above, and $\mathcal{V}$ treated as the operator $L^2(\mu) \to L^2(\mu_\alpha)$ is exactly the operator $\mathcal{V}_\alpha$ from Theorem 2.6.

As in the self-adjoint setting, the Representation Theorem reminds us of singular integral operators. Acting as in the self-adjoint case we show that the kernel $K(z, \xi) = 1/(1 - \bar{\xi} z)$ is restrictedly bounded (see Definition 3.1 below). Again, Theorem 3.2 and Remark 3.3 show the uniform boundedness of the regularization of the singular integral operator.

Theorem 2.8 *For the Clark measures μ and μ_α, the operators $T_r : L^2(\mu) \to L^2(\mu_\alpha)$ given by*

$$T_r f(z) := \int_{\mathbb{T}} \frac{f(\xi) d\mu(\xi)}{1 - r\bar{\xi} z}$$

are uniformly (in $r \in \mathbb{R}_+ \setminus \{1\}$) bounded.

An analog of Lemma 2.4 holds for the unit circle with essentially the same proof (for a different proof, see [31, Proposition 8.2]), so the limits $\lim_{r \to 1\mp} T_r f(z)$ exist μ_α-a.e. on $\mathbb{T}$. So we can define operators $T_\pm$ as the μ_α-a.e. limits

$$T_\pm f(z) := \lim_{r \to 1\mp} T_r f(z), \qquad z \in \mathbb{T},$$

or, equivalently, as w.o.t. limits

$$T_{\pm} f := \text{w.o.t.-} \lim_{r \to 1^{\mp}} T_r f.$$

Replacing the kernel in (11) by $1/(1 - r\overline{\xi} z)$ and taking the limit as $r \to 1^{\mp}$, we get an alternative formula for $\mathcal{V}_\alpha$.

Theorem 2.9 *Let μ and μ_α be the spectral measures of U and U_α respectively, and let $T_{\pm} =$ w.o.t.-$\lim_{r \to 1^{\mp}} T_r$ (the existence of the limit was just discussed). Then $\mathcal{V}_\alpha$ has the alternative representation*

$$\mathcal{V}_\alpha f = [\mathbf{1} - (1-\alpha) T_{\pm} \mathbf{1}] f + (1-\alpha) T_{\pm} f \qquad \forall f \in L^2(\mu).$$

How Unstable Can the Singular Spectrum Become?

By the Kato-Rosenblum theorem we know that the absolutely continuous spectrum remains invariant under rank one perturbations. But under a rank one perturbation by a cyclic vector, the singular perturbation can change type, as was shown by an example by Donoghue. So the question becomes: To which extent may the spectral properties of the measures μ_α vary as we change α? Much work has been done and many interesting examples were discovered, several are included in [40].

First of all notice that in the context of rank one perturbations for pure point and the singular continuous spectrum can behave quite different. For example, it is possible for A_α to have purely singular continuous spectrum on the interval $[0, 1]$ for all α. But the same behavior is not possible for pure point spectrum. In fact, the perturbations A_α have pure point spectrum for all α if and only if the spectrum is countable without accumulation points.

Another question concerns the type of parameter sets that allow dense singular embedded (in absolutely continuous) spectrum. For several years, all examples exhibited dense singular embedded spectrum only for a Lebesgue measure zero set of parameters α. It came as a surprise when Del Rio, Fuentes and Poltoratskii [15] proved the existence of a family of rank one perturbations with dense absolutely continuous spectrum and dense singular spectrum for almost every parameter α in an arbitrary (previously given) set $B \subset \mathbb{R}$ and with purely absolutely continuous spectrum for almost every $\alpha \in \mathbb{R} \backslash B$. Their proof uses Clark theory. Via a complicated construction they show the existence of a characteristic function for which the corresponding family of rank one unitary perturbations has the desired properties. In fact, it is possible to produce most any type of singular spectrum in this setting, see [16]. In the last reference, it was remarked that replacing the words 'almost every' by 'every' in their statement would be a non-trivial improvement, requiring a rather different approach. Namely, they suggested the following open problem: Fix an interval $I \subset \mathbb{R}$ and a measurable subset $B \subset \mathbb{R}$. Can one find a family of measures μ_β so that dense singular spectrum on I occurs precisely when $\beta \in B$ (and the corresponding operator has purely absolutely continuous spectrum on I for all $\beta \in \mathbb{R}\backslash B$)?

A class of examples is concerned with the question of how unstable the spectral type may be, if we do not have absolutely continuous part. A result of Del Rio, Makarov and Simon [17] which was independently proved by Gordon [22] states the following. Consider $I \subset \operatorname{supp} \mu$ closed and not a singleton. If $\mu_\alpha|_I$ is singular, then the set of α's for which $\mu_\alpha|_I$ is purely singular continuous is a dense G_δ set.

A converse to this result was presented by C. Sundberg [44]: For any closed subinterval I which is not a singleton and any G_δ subset of $\mathbb{R}$, there exists a family of measures (corresponding to a family of rank one perturbations) such that $\operatorname{supp} \mu \subset I$, μ_α is purely singular continuous for $\alpha \in G$ and μ_α is pure point for $\alpha \in \mathbb{R}\backslash G$. In the proof, Sundberg applies Clark theory. He constructs the characteristic function by defining a function on a Riemann surface $\mathcal{R}$ over the disk $\mathbb{D}$, and then applies the projection from $\mathcal{R}$ to $\mathbb{D}$.

Behavior of the Singular Continuous Spectrum

To this day, a characterization of the singular continuous part of the perturbed operator's spectral measure in terms of the unperturbed operator remains an open problem. Several sufficient conditions for the absence of singular continuous spectrum are known (see, for example, [11, 40]). Within the realm of our methods, an application of Theorem 2.3 empowers us with control over singular spectrum of the perturbed operator.

Lemma 2.10 (Lemma 4.4 of [32]) *Operators A_α, $\alpha \in \mathbb{R} \setminus \{0\}$, have a pure absolutely continuous spectrum on a closed interval I, if*

$$\int_0^\varepsilon x^{-2} w_I^* dx = \infty.$$

Here $d\mu = wdx + d\mu_s$ ($w \in L^1(dx)$) is the Lebesgue decomposition, and w_I^* denotes the increasing rearrangement of w on I.

This result allows a construction of unperturbed operators A with arbitrary embedded singular spectrum and for which all of the perturbed operators A_α, $\alpha \neq 0$ have no embedded singular spectrum.

Singular Integral Operators

Preliminaries

The Hilbert transform T

$$Tf(x) = \int_{\mathbb{R}} \frac{f(y)dy}{x-y}$$

is an example of what is usually called a *singular integral operator*. “Singular” here means that the kernel $K(x, y)$ of the operator is not integrable in y near the diagonal, so in the formal expression $Tf(x) = \int K(x, y)f(y)dy$ the integral is not well defined.

In the case of Hilbert transform it is very easy to show that the integral in the sense of principal value is well defined for C^1 compactly supported functions, so the operator is defined on a dense set in L^2 (and L^p, $1 < p < \infty$). It also can be shown that it can be extended to a bounded operator there.

Moreover, it can be shown that the integral in the sense of principal value exists a.e. for all $f \in L^p$, $1 \le p < \infty$; the proof is not as easy as for the C^1 functions, and is, in fact, quite involved.

A part of the operator V_α from Theorem 2.1 looks like the Hilbert transform, with the difference that the integration there is with respect to a general Radon measure μ. And what makes things even more complicated, is that the target space is $L^2(\mu_\alpha)$ with μ_α being a new measure.

In the theory of singular integral operators, there are several ways to define such an operator rigorously. One of the accepted ways, is what one would call the *axiomatic* approach. Namely, to define a singular integer operator $T : L^p(\mu) \to L^p(\nu)$ with kernel K we assume that we are given its bilinear form, defined on a dense subset of $L^p(\mu) \to L^{p'}(\nu)$, $1/p + 1/p' = 1$. The fact that T is an integral operator with kernel K means simply that

$$\langle Tf, g\rangle_\nu = \int K(x, y)f(y)g(x)d\mu(y)d\nu(x) \tag{14}$$

for all (say bounded) f and g with separated compact supports. Since the kernel K blows up only on the diagonal $x = y$, the integral above is well defined. Note, that according to this definition the multiplication operator M_φ, $M_\varphi f = \varphi f$ is an operator with kernel $K(x, y) \equiv 0$.

Moreover, it can be shown that any bounded singular integral operator with kernel $K \equiv 0$, where kernel is understood in the sense of (14), is a multiplication operator. So, according to the axiomatic approach, any two bounded singular integral operators that differ by a multiplication operator are identified as equal.

Another way to define the singular integral operator with kernel K is to consider the truncated operators T_ε,

$$T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} K(x,y)f(y)dy$$

which under usual assumptions about kernel K are well defined for bounded functions f with compact support. And we say that the integral operator with kernel K is bounded if all operators T_ε are uniformly bounded. If the operators T_ε are uniformly bounded, we can take w.o.t. limit of T_ε as $\varepsilon \to 0^+$, so in this case K is indeed a kernel of a bounded singular integral operator in the sense of the axiomatic approach.

Moreover, in all known examples if an axiomatically defined operator T is uniformly bounded then the operators T_ε are uniformly bounded. And as it turns out, this is not a coincidence, but a corollary of a very general fact.

Setup

In this paper we assume that μ and ν are Radon measures on $\mathbb{R}^d$ and that K belongs to $L^2_{\text{loc}}(\mu \times \nu)$ off the diagonal $x = y$, meaning that for any $x_0 \neq y_0$ there exists a neighborhood G of $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{R}^d$ such that $K\mathbf{1}_G \in L^2(\mu \times \nu)$. Note, that these assumptions are weaker than what is usually assumed about the kernels of singular integral operators.

The main results are also true for (at least some) locally compact abelian groups, in particular for tori $\mathbb{T}^d$. Also, since everything is local, the results can be modified to hold on smooth manifolds.

Definition 3.1 Let $K \in L^2_{\text{loc}}(\mu \times \nu)$ off the diagonal $x = y$. We say that K is $L^p(\mu) \to L^p(\nu)$ restrictedly bounded if for all $f \in L^\infty(\mu)$, $g \in L^\infty(\nu)$ with separated compact supports

$$\left| \int K(x,y)f(y)g(x)d\mu(y)d\nu(x) \right| \le C\|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)}. \tag{15}$$

The best constant C in (15) is called the $L^p(\mu) \to L^p(\nu)$ restricted bound of K, and denoted by $[K]^{\text{r}}_{L^p(\mu)\to L^p(\nu)}$.

If the exponent p and the measures μ, ν are fixed, we will skip $L^p(\mu) \to L^p(\nu)$ and simply say *restrictedly bounded.*

Going back, we can see that the operator V_α from Theorem 2.1 is a singular integral operator (in the sense of axiomatic approach) with kernel $K(s,t) = \alpha/(s-t)$. Since V_α is a unitary operator $L^2(\mu) \to L^2(\mu_\alpha)$ its norm is 1 and therefore the kernel $\alpha/(s-t)$ is restrictedly bounded with the $L^2(\mu) \to L^2(\mu_\alpha)$ restricted norm at most 1. Equivalently, one can say that the $L^2(\mu) \to L^2(\mu_\alpha)$ restricted norm of the kernel $1/(s-t)$ is at most $1/|\alpha|$.

Similarly, the operator $\mathcal{V}_\alpha$ from Theorem 2.6 is a singular integral operator with kernel $K(z,\xi) = (1-\alpha)/(1-\overline{\xi}z)$, $z,\xi \in \mathbb{T}$, and the $L^2(\mu) \to L^2(\mu_\alpha)$ restricted norm of the kernel $1/(1-\overline{\xi}z)$ is at most $1/|1-\alpha|$.

Regularizations of Singular Kernels

Let $m : \mathbb{R}^d \to \mathbb{R}$ be a *regularizer*, i.e. a bounded function which is 0 in a neighborhood of 0 and 1 in a neighborhood of ∞. Define the *regularized kernel* K_ε by $K_\varepsilon(x,y) = K(x,y)m((x-y)/\varepsilon)$. The regularized kernels K_ε are in $L^2_{\text{loc}}(\mu\times\nu)$ so the regularized integral operators T_ε,

$$T_\varepsilon f(x) := \int K_\varepsilon(x,y)f(y)d\mu(y)$$

are well defined for bounded compactly supported f. In particular, if $m(x) = \mathbf{1}_{(1,\infty)}(|x|)$ then we get the classical truncation

$$T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} K(x,y)f(y)d\mu(y). \tag{16}$$

If for $1 < p < \infty$ operators $T_\varepsilon : L^p(\mu) \to L^p(\nu)$ are uniformly bounded, then by taking w.o.t. limit point as $\varepsilon \to 0^+$ we conclude that K is a kernel of a singular integral operator (in the sense of the axiomatic approach) with kernel K, acting $L^p(\mu) \to L^p(\nu)$.

It turns out that the converse statement is true, even in a stronger sense, if we assume that the measures μ and ν do not have common atoms. Namely, the following theorem holds, see [33, Proposition 2.12].

Theorem 3.2 *Let a kernel K be $L^p(\mu) \to L^p(\nu)$ restrictedly bounded, and assume that μ and ν do not have common atoms. Then for any regularizer $m \in C^\infty$ the regularized operators T_ε are uniformly (in ε) bounded, $\|T_\varepsilon\|_{L^p(\mu)\to L^p(\nu)} \le C(m) < \infty$.*

Moreover, for all "interesting" kernels the $L^p(\mu) \to L^p(\nu)$ restricted boundedness implies the uniform boundedness of the classical truncations (16).

Without going into details, we just mention that the "interesting" kernels include kernel $1/(x-y)$, $x,y \in \mathbb{R}$ of the Hilbert transform, the kernel $(x-y)/|x-y|^{\alpha+1}$, $\alpha > 0$, $x,y \in \mathbb{R}^d$ of the generalized Riesz transform R_α in $\mathbb{R}^d$, the kernel $1/(z-w)$, $z,w \in \mathbb{C}$ of the Cauchy transform, the kernel $1/(z-w)^2$, $z,w \in \mathbb{C}$ of the Beurling–Ahlfors transform and many others. For more information please refer to [33].

Regularizations with smooth functions m seem to be a more logical and convenient choice, than the classical one; for example if one starts with a Calderón–Zygmund kernel then after smooth regularizations the resulting kernel will still be a Calderón–Zygmund one with uniform estimates of the constants. However, the classical truncations are used most.

Remark To define truncation of a kernel on the unit circle $\mathbb{T}$ we take the function m on the line, $m \equiv 0$ in a neighborhood of 0 and $m \equiv 1$ in a neighborhood of ∞, and define functions $\widetilde{m}_\varepsilon$ on $\mathbb{T}$ by

$$\widetilde{m}_\varepsilon(e^{it}) = m(t/\varepsilon), \qquad -\pi < t \leq \pi.$$

Then the regularized kernel K_ε is defined as

$$K_\varepsilon(z, \xi) = K(z, \xi) m_\varepsilon(z/\xi), \qquad z, \xi \in \mathbb{T}.$$

The regularized kernels on $\mathbb{T}^d$ are defined similarly, and the same results as in $\mathbb{R}^d$ holds in $\mathbb{T}^d$.

Remark 3.3 For singular integrals related to complex analysis there is another type of natural regularization. Namely for the kernel $K(x, y) = 1/(x - y)$ on $\mathbb{R}$ one can consider kernels

$$K_{\pm\varepsilon}(x, y) = 1/(x - y \pm i\varepsilon). \tag{17}$$

Similarly, for the kernel $K(z, \xi) = 1/(1 - \overline{\xi} z)$ on $\mathbb{T}$ define the regularized kernel

$$K_r(z, \xi) = 1/(1 - r\overline{\xi} z), \qquad 0 \leq r < \infty \quad r \neq 1. \tag{18}$$

For these kernels Theorem 3.2 holds as well.

Now let us discuss the main ideas of the proofs.

First Step: Schur Multipliers

The first idea is very simple: we want to multiply a restrictedly bounded kernel by a function M such that the resulting kernel is still restrictedly bounded.

Definition 3.4 We call a function $M(\,\cdot\,,\,\cdot\,)$ an $L^p(\mu) \to L^p(\nu)$ Schur multiplier if for any $L^p(\mu) \to L^p(\nu)$ restrictedly bounded kernel K the kernel KM is also $L^p(\mu) \to L^p(\nu)$ restrictedly bounded and

$$[KM]^{\mathrm{r}}_{L^p(\mu) \to L^p(\nu)} \leq C[K]^{\mathrm{r}}_{L^p(\mu) \to L^p(\nu)}.$$

The best constant C in the above inequality is called the Schur norm of M.

Traditionally, Schur multipliers are defined with respect to the operator norm of the corresponding integral operators, or with respect to the Schatten–von-Neumann norm, but our definition is very close in spirit, so we use the same term.

Formally, our definition depends on μ, ν and p, but we will construct "universal" multipliers, that work for all μ, ν and p with the same estimate on the Schur norm. They also are Schur multipliers with respect to the operator norm, as well as with respect to the Schatten–von-Neumann norms.

Thus, in what follows we will omit $L^p(\mu) \to L^p(\nu)$ and simply say Schur multiplier.

Constructing Schur Multipliers via Fourier Transform

We start with an elementary observation: the function M_a, $M_a(x,y) := e^{-ia\cdot x}e^{ia\cdot y}$, $a, x, y \in \mathbb{R}^d$ is a Schur multiplier with the Schur norm 1 (as a product of two unimodular functions of one variable).

Averaging in a we get that if σ is a complex-valued measure of bounded variation and $m = \widehat{\sigma}$ is its Fourier transform,

$$\widehat{\sigma}(s) := \int_{\mathbb{R}^d} e^{-is\cdot t} d\sigma(t)$$

then the function $M(x,y) = m(x-y)$ is a Schur multiplier with the Schur norm at most $\operatorname{var}\sigma$.

Note also that for $m_\varepsilon(s) = m(s/\varepsilon)$ and the measure σ_ε defined by $\sigma_\varepsilon(E) = \sigma(\varepsilon E)$ we have $m_\varepsilon = \widehat{\sigma}_\varepsilon$. Since $\operatorname{var}\sigma_\varepsilon = \operatorname{var}\sigma$ we get that all the functions M_ε

$$M_\varepsilon(x,y) = m_\varepsilon(x-y) = m((x-y)/\varepsilon)$$

are Schur multipliers with the Schur norm estimated by $\operatorname{var}\sigma$.

Since a compactly supported C^∞ function is a Fourier transform of an L^1 function (it is a Fourier transform of a Schwartz class function), and 1 is trivially a Schur multiplier, we can conclude that functions M_ε, $M_\varepsilon(x,y) = m((x-y)/\varepsilon)$ where m is the C^∞ regularizer defined in section "Regularizations of Singular Kernels" are Schur multipliers.

So we see that the regularized kernels K_ε obtained using smooth regularizers m are restrictedly bounded with the uniform (in ε) estimate on the restricted norm.

To get the corresponding result for the torus $\mathbb{T}^d$ we just need to restrict the regularizers m_ε to the cube $(-\pi, \pi]^d$ and then map the cube to the torus via the standard map.

Cauchy Type Regularizations

Let us now discuss the Cauchy type regularizations (17) and (18). For $\rho(x) = \mathbf{1}_{[0,\infty)}e^{-x}$ define

$$m(s) = 1 - \widehat{\rho}(s) = \frac{s}{s-i}.$$

Then $m_\varepsilon(s) = m(s/e) = s/(s - i\varepsilon)$, and the functions $M_\varepsilon(x, y) = m_\varepsilon(x - y)$ are Schur multipliers with Schur norm at most 2. Computing the regularized kernel we get

$$K_\varepsilon(x, y) = \frac{1}{x-y}\frac{x-y}{x-y-i\varepsilon} = \frac{1}{x-(y+i\varepsilon)},$$

so the kernels $K_{+\varepsilon}$ from (17) are uniformly restrictedly bounded.

Repeating the same reasoning with $\rho(x) = \mathbf{1}_{(-\infty,0]}e^x$ we get the conclusion for $K_{-\varepsilon}$.

For the kernel (18) on $\mathbb{T}$ we use the Fourier transform on $\mathbb{Z}$. Namely, it is easy to show that if $a \in \ell^1(\mathbb{Z})$ and $m(z) := \sum_{k\in\mathbb{Z}} a_k z^k$, $z \in \mathbb{T}$, then the function M

$$M(z, \xi) = m(z/\xi) \qquad z, \xi \in \mathbb{T}$$

is a Schur multiplier with Schur bound at most $\|a\|_{\ell^1}$.

Then for $0 \le r < 1$ multiplying $K(z, \xi) = 1/(1 - \overline{\xi}z)$ by

$$m(z/\xi) = 1 + \sum_{n=1}^{\infty}(r^n - r^{n-1})(\overline{\xi}z)^n = \frac{1-\overline{\xi}z}{1-r\overline{\xi}z}$$

we at most double the restricted norm (because $1 + \sum_{n=1}^{\infty} |r^n - r^{n-1}| = 1 + r \le 2$). So, for the kernel

$$K(z, \xi)\cdot\frac{1-\overline{\xi}z}{1-r\overline{\xi}z} = \frac{1}{1-r\overline{\xi}z} = K_r(z, \xi), \qquad r < 1$$

we get for $r < 1$

$$[K_r]_{L^p(\mu)\to L^p(\nu)} \le 2[K]_{L^p(\mu)\to L^p(\nu)}. \tag{19}$$

For $r > 1$ we can write

$$m(z/\xi) = \frac{1-\overline{\xi}z}{1-r\overline{\xi}z} = 1 - \sum_{n=1}^{\infty}(r^{-n} - r^{-(n+1)})(\overline{\xi}z)^{-n}.$$

Noticing that $1 + \sum_{n=1}^{\infty} |r^{-n} - r^{-(n+1)}| = 1 + r^{-1} \le 2$ we see that in the case $r > 1$ (19) holds as well.

Final Step: Boundedness of the Regularized Operators

Theorem 3.5 *Let μ and ν be Radon measures in $\mathbb{R}^d$ without common atoms. Assume that a kernel $K \in L^2_{\text{loc}}(\mu \times \nu)$ is $L^p(\mu) \to L^p(\nu)$ restrictedly bounded, with the restricted norm C. Then the integral operator with T kernel K is a bounded operator $L^p(\mu) \to L^p(\nu)$ with the norm at most $2C$.*

Restricting the kernels to compact subsets exhausting $\mathbb{R}^d \times \mathbb{R}^d$ one can easily reduce the proof to the case $K \in L^2(\mu \times \nu)$ (globally, not locally). Then the idea of the proof is very simple. Taking bounded compactly supported functions f and g we can write

$$\langle Tf, g\rangle_\nu = \int K(x,y)f(y)g(x)d\mu(y)d\nu(x).$$

The main idea of the proof is to construct bounded functions f_n, g_n with separated compact supports such that $f_n \rightharpoonup \frac{1}{2}f$ weakly in $L^2(\mu)$, $g_n \rightharpoonup \frac{1}{2}g$ weakly in $L^2(\nu)$ and such that

$$\limsup_{n\to\infty} \|f_n\|_{L^p(\mu)} \le 2^{-1/p}\|f\|_{L^p(\mu)}, \qquad \limsup_{n\to\infty} \|g_n\|_{L^{p'}(\mu)} \le 2^{-1/p'}\|g\|_{L^{p'}(\nu)}. \quad (20)$$

Since the operator T is Hilbert–Schmidt, and so compact (as an operator $L^2(\mu) \to L^2(\nu)$) the weak convergence implies that

$$\langle Tf_n, g_n\rangle_\nu \to \frac{1}{4}\langle Tf, g\rangle_\nu.$$

Therefore, using (20) we get

$$|\langle Tf, g\rangle| \le \limsup_{n\to\infty} 4|\langle Tf_n, g_n\rangle| \le 2C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\nu)}.$$

The main idea of the construction of the functions f_n and g_n is quite simple, at least for the absolutely continuous piece: we define $f_n := \mathbf{1}_{E_n}$, $g_n := \mathbf{1}_{F_n}$ where E_n and F_n are separated "mesh like" subsets, that are well mixed, meaning that for all dyadic cubes Q of size at least 2^{-n} the Lebesgue measure of the sets $Q \cap E_n$ and $Q \cap F_n$ is almost half (with relative error of say 2^{-n}) of the measure of Q. Construction of such sets in for the Lebesgue measure is rather trivial and can be left as an exercise for the reader.

For the measures μ and ν without atoms the construction is almost the same, only the "well mixed" property is with respect to the measure $\sigma = \mu + \nu$, meaning that for any dyadic cube Q of size at least 2^{-n} the measures $\sigma(Q \cap E_n)$, $\sigma(Q \cap F_n)$ are almost half of $\sigma(Q)$ with relative error 2^{-n}. It might not be immediately obvious how to construct such sets E_n, F_n, but the construction is relatively simple and straightforward, see [33] for details.

The construction in the general case is just a bit more complicated. Namely, we first construct the sets E_n and F_n with respect to the continuous parts μ_c, ν_c of the measures (making sure that the sets do not contain any atoms). Then we define f_n and g_n by adding to $f\mathbf{1}_{E_n}$ and $g\mathbf{1}_{F_n}$ the functions

$$\frac{1}{2}\sum_{k=1}^{n} f(a_k)\delta_{a_k}, \qquad \frac{1}{2}\sum_{k=1}^{n} g(b_k)\delta_{b_k}$$

respectively, where a_k, b_k are atoms of μ and ν respectively. To make sure that the functions f_n and g_n have separated supports, we then just need to "shrink" the sets E_n F_n by removing small discs around atoms. Again, the reader is referred to [33] for the details.

This idea of using "well mixed" set was exploited in [34] in the case of Lebesgue measure. It was later used in [32], where some of the result in this section were proved under the assumption that the singular parts of μ and ν are mutually singular.

The results in full generality were proved in [33], the reader should look there for full details.

Clark Theory for Rank One Perturbations of Unitary Operators

Plan of the Game

As we discussed above in section "Unitary Rank One Perturbations", rank one unitary perturbations of a unitary operator U are parametrized by the formula (9). If in (9) we take $|\alpha| < 1$ (instead of $|\alpha| = 1$) the resulting operator U_α will be not a unitary, but only a contractive ($\|U_\alpha\| \le 1$) operator.

If, as in section "Unitary Rank One Perturbations" we assume by ignoring the trivial part that b is $*$-cyclic vector for U, then for $|\gamma| < 1$ the operator $U_\gamma = U + (\gamma - 1)bb_1^*$, $b_1 = U^*b$ is a *completely non-unitary* (c.n.u.) contraction. The term completely non-unitary means that there is no reducing (i.e. invariant for U_γ and U_γ^*) subspace on which U_γ acts unitarily.

A completely non-unitary contraction T is up to unitary equivalence determined by its so-called *characteristic function* $\theta = \theta_T$, see the definition below. Namely, T is unitarily equivalent to its *model* $\mathcal{M} = \mathcal{M}_\theta$, where $\mathcal{M}_\theta$ is a *compression* of the multiplication operator M_z,

$$\mathcal{M}_\theta = P_\theta M_z \big|_{\mathcal{K}_\theta};$$

here $\mathcal{K}_\theta$ is a subspace of a generally vector-valued, and possibly weighted L^2 space on the unit circle, $P_\theta = P_{\mathcal{K}_\theta}$ is the orthogonal projection onto $\mathcal{K}_\theta$, and M_z is the multiplication by the independent variable z, $M_z f(z) = zf(z)$, $z \in \mathbb{T}$.

So, we have two unitarily equivalent representations of the operator U_γ, $|\gamma| < 1$: the representation

$$U_\gamma = M_\xi + (\gamma - 1)bb_1^*, \qquad b = \mathbf{1}, \quad b_1 = M_\xi^* \mathbf{1}$$

in the spectral representation of U in $L^2(\mu)$, where μ is the spectral measure of U corresponding to the vector b, and the representation as the model operator $\mathcal{M}_{\theta_\gamma}$ in the model subspace $\mathcal{K}_{\theta_\gamma}$.

The Clark theory describes the unitary operator providing this unitary equivalence, i.e. a unitary operator $\Phi_\gamma : \mathcal{K}_{\theta_\gamma} \to L^2(\mu)$ such that

$$\Phi_\gamma \mathcal{M}_{\theta_\gamma} = U_\gamma \Phi_\gamma.$$

D. Clark in his original paper [12] described such operators for the particular case when θ_γ is an inner function. He started with the model operator (unitarily equivalent to U_γ, $|\gamma| < 1$ in our notation) in a particular case of inner characteristic function, described all its unitary rank one perturbations (U_α, $|\alpha| = 1$ in our notation) and described the unitary operator between the model operator $\mathcal{M}_\theta$ and the spectral representation of U_α, $|\alpha| = 1$.

Translated to our language the fact that the characteristic function θ is inner means that the operator U (and so all U_α, $|\alpha| = 1$) have purely singular spectrum.

A Functional Model for a c.n.u. Contraction

Let us recall the definition related to the functional model. For a c.n.u. contraction T acting in a separable Hilbert space we define the defect operators

$$D_T := (\mathbf{I} - T^*T)^{1/2}, \qquad D_{T^*} := (\mathbf{I} - TT^*)^{1/2},$$

and the defect subspaces

$$\mathfrak{D} = \mathfrak{D}_T := \operatorname{clos}\operatorname{Ran} D_T, \qquad \mathfrak{D}_* = \mathfrak{D}_{T^*} := \operatorname{clos}\operatorname{Ran} D_{T^*}.$$

The characteristic function $\theta = \theta_T$ of the operator T is an analytic function $\theta = \theta_T \in H^\infty_{\mathfrak{D}\to\mathfrak{D}_*}$ whose values are bounded operators (in fact, contractions) acting from $\mathfrak{D}$ to $\mathfrak{D}_*$ defined by the equation

$$\theta_T(z) = \left(-T + zD_{T^*}(\mathbf{I} - zT^*)^{-1}D_T\right)\Big|_{\mathfrak{D}}, \qquad z \in \mathbb{D}. \tag{21}$$

Note that $T\mathfrak{D} \subset \mathfrak{D}_*$, so for $z \in D$ the above expression indeed can be interpreted as an operator from $\mathfrak{D}$ to $\mathfrak{D}_*$.

It is customary to assume that the characteristic function is defined up to constant unitary factors on the right and on the left, i.e. one considers the whole equivalence class consisting of functions $U\theta V$, where $U : \mathfrak{D}_* \to E_*$ and $V : E \to \mathfrak{D}$ are unitary operators and E_*, E are Hilbert spaces of appropriate dimensions. The advantage of this point of view is that we are not restricted to using the defect spaces of T, but can work with arbitrary Hilbert spaces of appropriate dimensions.

Note, that the characteristic function (defined up to constant unitary factors) is a unitary invariant of a completely non-unitary contraction: any two such contractions with the same characteristic function are unitarily equivalent.

Note also, that given a characteristic function, any representative gives us a model, and there is a standard unitary equivalence between the model for different representatives.

Remark Another way to look at a choice of a representative of a characteristic function is to pick orthonormal bases in the defect spaces and treat the characteristic function as a matrix-valued function (possibly of infinite size). The choice of the orthonormal bases is equivalent to the choice of the constant unitary factors.

In this paper by a *functional model* associated to an operator-valued function $\theta \in H^\infty_{E \to E_*}$ we understand the following: a model space $\mathcal{K}_\theta$ is an appropriately constructed subspace of a (possibly) weighted space $L^2(E_* \oplus E, W)$ on the unit circle $\mathbb{T}$ with the operator-valued weight W. The model operator $\mathcal{M}_\theta$ is a compression of the multiplication operator M_z onto $\mathcal{K}_\theta$,

$$\mathcal{M}_\theta = P_\theta M_z \big|_{\mathcal{K}_\theta}; \tag{22}$$

where $P_\theta = P_{\mathcal{K}_\theta}$ is the orthogonal projection onto $\mathcal{K}_\theta$.

All the functional models for the same θ are unitarily equivalent, so sometimes people interpret them as different *transcriptions* of one object.

As we already mentioned above, a completely non-unitary contraction with characteristic function θ is unitarily equivalent to its model $\mathcal{M}_\theta$.

On the other hand, for any purely contractive $\theta \in H^\infty_{E \to E_*}$, $\|\theta\|_\infty \le 1$ the model operator $\mathcal{M}_\theta$ is a completely non-unitary contraction, with θ being its characteristic function. Thus, any such θ is a characteristic function of a completely non-unitary contraction.

Sz.-Nagy–Foiaş Transcription

The Sz.-Nagy–Foiaş model (transcription) is probably the most used.

The model space $\mathcal{K}_\theta$ is defined as a subspace of $L^2(E_* \oplus E)$ (non-weighted, $W(z) \equiv \mathbf{I}$),

$$\mathcal{K}_\theta = \begin{pmatrix} H^2_{E_*} \\ \operatorname{clos} \Delta L^2_E \end{pmatrix} \ominus \begin{pmatrix} \theta \\ \Delta \end{pmatrix} H^2_E \tag{23}$$

where the defect Δ is given by

$$\Delta(z) := (1 - \theta(z)^* \theta(z))^{1/2}, \qquad z \in \mathbb{T}. \tag{24}$$

If the characteristic function θ is *inner*, meaning that its boundary values are isometries a.e. on $\mathbb{T}$, then $\Delta \equiv 0$, so the lower "floor" of $\mathcal{K}_\theta$ collapses and we get a simpler, "one-story" model subspace,

$$\mathcal{K}_\theta = H^2(E_*) \ominus \theta H^2(E).$$

This subspace is probably much more familiar to analysts, especially when θ is a scalar-valued function.

The model operator $\mathcal{M}_\theta$ is defined by (22) as the compression of the multiplication operator M_z (also known as forward shift operator) onto $\mathcal{K}_\theta$, and the multiplication operator M_z is understood as the entry-wise multiplication by the independent variable z,

$$M_z \begin{pmatrix} g \\ h \end{pmatrix} = \begin{pmatrix} zg \\ zh \end{pmatrix}.$$

As we discussed above, the characteristic function θ is defined up to constant unitary factors on the right and on the left. But one has to be a bit careful here, because if $\widetilde{\theta}(z) = U\theta(z)V$, where U and V are constant unitary operators, then the spaces $\mathcal{K}_\theta$ and $\mathcal{K}_{\widetilde{\theta}}$ are different.

However, the map $\mathcal{U}$

$$\mathcal{U} \begin{pmatrix} g \\ h \end{pmatrix} = \begin{pmatrix} Ug \\ V^* h \end{pmatrix}$$

is the canonical unitary map transferring the model from one space to the other.

Namely, it is easy to see that $\mathcal{U}$ is a unitary map from $H^2(E_*) \oplus \operatorname{clos} \Delta L^2(E)$ onto $H^2(UE_*) \oplus \operatorname{clos} \widetilde{\Delta} L^2(V^* E)$, where $\widetilde{\Delta} = \Delta_{\widetilde{\theta}} = V^* \Delta V$. Moreover, it is not difficult to see that $\mathcal{U}\mathcal{K}_\theta = \mathcal{K}_{\widetilde{\theta}}$ and that $\mathcal{U}$ commutes with the multiplication by z, so $\mathcal{U}_\theta := \mathcal{U}\, \big|_{\mathcal{K}_\theta}$ intertwines the model operators,

$$\mathcal{U}_\theta \mathcal{M}_\theta = \mathcal{M}_{\widetilde{\theta}} \mathcal{U}_\theta.$$

de Branges–Rovnyak Transcription

Let us present this transcription as it is described in [35]. Since the ambient space in this transcription is a weighted L^2 space with an operator-valued weight, let us recall that if W is an operator-valued weight on the circle, i.e. a function whose values are self-adjoint non-negative operators in a Hilbert space E, then the norm in the space $L^2(W)$ is defined as

$$\|f\|_{L^2(W)} = \int_{\mathbb{T}} (W(z)f(z), f(z))_E \frac{|dz|}{2\pi}.$$

There are some delicate details here in defining the above integral if we allow the values $W(z)$ to be unbounded operators, but we will not discuss it here. In our case when the characteristic function is scalar-valued the values $W(z)$ are bounded self-adjoint operators on $\mathbb{C}^2$, and the definition of the integral is straightforward.

Let

$$W_\theta(z) = \begin{pmatrix} \mathbf{I} & \theta(z) \\ \theta(z)^* & \mathbf{I} \end{pmatrix}.$$

The weight in the ambient space will be given by $W = W_\theta^{[-1]}$, $W_\theta^{[-1]}(z) = (W_\theta(z))^{[-1]}$ where $A^{[-1]}$ stands for the Moore–Penrose inverse of the operator A. If $A = A^*$ then $A^{[-1]}$ is $\mathbf{O}$ on $\operatorname{Ker} A$ and is equal to the left inverse of A on $\operatorname{Ran} A$. The model space $\mathcal{K}_\theta$ is defined as

$$\mathcal{K}_\theta = \left\{ \begin{pmatrix} g_+ \\ g_- \end{pmatrix} : g_+ \in H^2(E_*),\ g_- \in H^2_-(E),\ g_- - \theta^* g_+ \in \Delta L^2(E) \right\}. \quad (25)$$

Remark 4.1 The original de Branges–Rovnyak model was initially described in [14] using completely different terms. To give the definition from [14] we need to recall the notion of a Toeplitz operator. For $\varphi \in L^\infty_{E\to E_*}$ the Toeplitz operator $T_\varphi : H^2(E) \to H^2(E_*)$ with symbol φ is defined by

$$T_\varphi f := P_+(\varphi f), \qquad f \in H^2(E).$$

The (preliminary) space $\mathcal{H}(\theta) \subset H^2(E_*)$ is defined as a range $(\mathbf{I} - T_\theta T_{\theta^*})^{1/2} H^2(E)$ endowed with the *range norm* (the minimal norm of the preimage).

Let the involution operator J on $L^2(\mathbb{T})$ be defined as

$$Jf(z) = \bar{z} f(\bar{z}).$$

Following de Branges–Rovnyak [14] define the *model space* $\mathcal{D}(\theta)$ as the set of vectors

$$\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} : g_1 \in \mathcal{H}(\theta),\ g_2 \in H^2(E),\ \text{such that } z^n g_1 - \theta P_+(z^n J g_2) \in \mathcal{H}(\theta)\ \forall n \ge 0,$$

and such that

$$\left\| \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\|^2_{\mathcal{D}(\theta)} := \lim_{n\to\infty} \left(\|z^n g_1 - \theta P_+(z^n J g_2)\|^2_{\mathcal{H}(\theta)} + \|P_+(z^n J g_2)\|_2^2 \right) < \infty.$$

It might look surprising, but it was proved in [36] that the operator $\begin{pmatrix} g_+ \\ g_- \end{pmatrix} \mapsto \begin{pmatrix} g_+ \\ J g_- \end{pmatrix}$ is a unitary operator between the described above model space $\mathcal{K}_\theta$ in the de Branges–Rovnyak transcription and the model space $\mathcal{D}(\theta)$.

Model for the Operator U_γ

For the perturbations U_γ, $|\gamma| < 1$ the functional model can be computed explicitly.

The defect operators are computed to be

$$D_{U_\gamma} = \left(\mathbf{I} - U_\gamma{}^* U_\gamma\right)^{1/2} = \left(1 - |\gamma|^2\right)^{1/2} b_1 b_1^*,$$
$$D_{U_\gamma{}^*} = \left(\mathbf{I} - U_\gamma U_\gamma{}^*\right)^{1/2} = \left(1 - |\gamma|^2\right)^{1/2} b b^*$$

and the defect spaces are

$$\mathfrak{D} = \mathfrak{D}_{U_\gamma} = \operatorname{span}\{b_1\} \qquad \text{and} \qquad \mathfrak{D}_* = \mathfrak{D}_{U_\gamma{}^*} = \operatorname{span}\{b\}.$$

Note that the defect spaces are one-dimensional, so the characteristic function $\theta = \theta_\gamma$ is a scalar-valued function. We already mentioned above that $\theta \in H^\infty$, $\|\theta\|_\infty \le 1$. Note also that the defect spaces do not depend on γ.

The characteristic function θ_γ of U_γ can be computed in terms of Cauchy type transforms. For a (possibly complex-valued) measure τ on $\mathbb{T}$ and $\lambda \notin \mathbb{T}$ define the Cauchy type transforms R, R_1 and R_2 by

$$R\tau(\lambda) := \int_{\mathbb{T}} \frac{d\tau(\xi)}{1 - \overline{\xi}\lambda}, \quad R_1\tau(\lambda) := \int_{\mathbb{T}} \frac{\overline{\xi}\lambda d\tau(\xi)}{1 - \overline{\xi}\lambda}, \quad R_2\tau(\lambda) := \int_{\mathbb{T}} \frac{1 + \overline{\xi}\lambda}{1 - \overline{\xi}\lambda} d\tau(\xi). \tag{26}$$

If we pick b_1 and b to be the basis vectors in the corresponding defect spaces, then the characteristic function θ_γ of the operator U_γ, $|\gamma| < 1$ is given by

$$\theta_\gamma(\lambda) = -\gamma + \frac{(1-|\gamma|^2)R_1\mu(\lambda)}{1+(1-\overline{\gamma})R_1\mu(\lambda)} = \frac{(1-\gamma)R_2\mu(\lambda)-(1+\gamma)}{(1-\overline{\gamma})R_2\mu(\lambda)+(1+\overline{\gamma})}, \qquad \lambda \in \mathbb{D}. \tag{27}$$

Note that the formulas for θ_0 ($\gamma = 0$) are especially simple. And θ_0 is related to θ_γ by a fractional transformation:

$$\theta_\gamma = \frac{\theta_0 - \gamma}{1-\overline{\gamma}\theta_0} \qquad \text{or equivalently} \qquad \theta_0 = \frac{\theta_\gamma + \gamma}{1+\overline{\gamma}\theta_\gamma}. \tag{28}$$

To compute the characteristic function one can use the definition (21) of the characteristic function with U_γ instead of T and the inversion formula (2). Namely, writing

$$\mathbf{I} - zU_\gamma^* = (\mathbf{I} - zU^*)\left(I - z(\overline{\gamma}-1)(\mathbf{I}-zU^*)^{-1}b_1b^*\right)$$

and applying the inversion formula (2) we get denoting $\beta = \gamma - 1$

$$(\mathbf{I} - zU_\gamma^*)^{-1} = \left(\mathbf{I} + \frac{1}{\left(z\overline{\beta}(\mathbf{I}-zU^*)^{-1}b_1, b\right)_{\mathcal{H}}} z\overline{\beta}(\mathbf{I}-zU^*)^{-1}b_1b^*\right)(\mathbf{I}-zU^*)^{-1}.$$

In the spectral representation of U in $L^2(\mu)$ the operator $(I - zU^*)^{-1}$ is the multiplication by the function $1/(1-\overline{\xi}z)$, $b \equiv \mathbf{1}$, $b_1(\xi) \equiv \xi$, so the above inverse can be explicitly computed. Then standard algebraic manipulations lead to the formulas (27) for the resolvent.

A different way of computing the characteristic function for finite rank perturbations can be found in [18].

We point out that if the measure μ is purely singular (with respect to the Lebesgue measure), then the functions θ_γ are inner ($|\theta_\gamma| = 1$ a.e. on $\mathbb{T}$). In this case the model is especially simple, the model space consists of scalar functions, and that is the case treated by the original Clark theory.

However, in our case, μ is an arbitrary probability measure, so the characteristic functions can be non inner, and the model is more complicated: the model space consists of vector-valued functions (with values in $\mathbb{C}^2$).

Preliminaries About Clark Operator

Recall that our goal is to describe a *Clark operator*, i.e. a unitary operator (non-uniqueness is discussed in the next paragraph) that realizes unitary equivalence between U_γ and $\mathcal{M}_{\theta_\gamma}$. Namely, we want to find a unitary operator $\Phi_\gamma : \mathcal{K}_{\theta_\gamma} \to L^2(\mu)$ such that

$$\Phi_\gamma \mathcal{M}_{\theta_\gamma} = U_\gamma \Phi_\gamma. \tag{29}$$

Let us discuss what freedom do we have in choosing such an operator. Clearly, Φ_γ maps defect spaces of $\mathcal{M}_{\theta_\gamma}$ to the corresponding defect spaces of U_γ. Therefore, $\Phi_\gamma^* b$ and $\Phi_\gamma^* b_1$ must be unit vectors in $\mathfrak{D}_{\mathcal{M}_{\theta_\gamma}^*}$ and $\mathfrak{D}_{\mathcal{M}_{\theta_\gamma}}$ respectively.

We say that the unit vectors $c \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}^*}$ and $c_1 \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}}$ *agree* if there exists a unitary map $\Phi_\gamma : \mathcal{K}_{\theta_\gamma} \to L^2(\mu)$ satisfying (29) such that

$$\Phi_\gamma^* b = c, \qquad \Phi_\gamma^* b_1 = c_1.$$

If $\gamma = 0$ and μ is the Lebesgue measure, then it is not hard to see that $\theta_\gamma \equiv 0$. It is also easy to see that in this case, any two unit vectors $c \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}^*}$ and $c_1 \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}}$ agree.

Otherwise, if either $\gamma \neq 0$ or μ differs from the Lebesgue measure, then for any unit vector $c \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}^*}$ there exist a unique vector $c_1 \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}}$ which agrees with c; for details see Proposition 2.9 of [31]. That means the operator Φ_γ is unique up to a multiplicative unimodular constant $\alpha \in \mathbb{T}$; in particular, if we fix a unit vector $c \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}^*}$ then the condition $\Phi_\gamma c = b$ uniquely determines the Clark operator Φ_γ.

In the trivial case when μ is the normalized Lebesgue measure and $\gamma = 0$ the Clark operator Φ_γ can be easily constructed via elementary means, so in what follows we will ignore this case.

A "Universal" Representation Formula for the Adjoint of the Clark Operator

An explicit computation of the defect spaces of the compressed shift operator $\mathcal{M}_\theta$ yields that in the Sz.-Nagy–Foiaş transcription

$$\mathfrak{D}_{\mathcal{M}_\theta^*} = \operatorname{span}\{c\}, \qquad \mathfrak{D}_{\mathcal{M}_\theta} = \operatorname{span}\{c_1\},$$

where

$$c(z) := \left(1 - |\theta(0)|^2\right)^{-1/2} \begin{pmatrix} 1 - \overline{\theta(0)}\theta(z) \\ -\overline{\theta(0)}\Delta(z) \end{pmatrix}, \tag{30}$$

$$c_1(z) := \left(1 - |\theta(0)|^2\right)^{-1/2} \begin{pmatrix} z^{-1}\left(\theta(z) - \theta(0)\right) \\ z^{-1}\Delta(z) \end{pmatrix}, \tag{31}$$

where $\Delta := (1 - |\theta|^2)^{1/2}$.

Moreover, the vectors c and c_1 are of unit length and agree. Note also that it follows from (27) that $\theta_\gamma(0) = -\gamma$, so the above formulas can be further simplified.

The following theorem describes the adjoint Φ_γ^* of the Clark operator. Note that the intertwining relation (29) can be rewritten as

$$\Phi_\gamma^* U_\gamma = \mathcal{M}_{\theta_\gamma} \Phi_\gamma^*.$$

Theorem 4.2 (A "universal" representation formula; Theorem 3.1 of [31]) *Let θ_γ be a characteristic function (one representative) of U_γ, $|\gamma| < 1$, and let $\mathcal{K}_{\theta_\gamma}$ and $\mathcal{M}_\gamma = \mathcal{M}_{\theta_\gamma}$ be the model subspace and the model operator respectively. Assume that the unit vectors $c = c^\gamma \in \mathfrak{D}_{\mathcal{M}^*_{\theta_\gamma}}$, $c_1 = c_1^\gamma \in \mathfrak{D}_{\mathcal{M}_{\theta_\gamma}}$ agree. Let $\Phi_\gamma^* : L^2(\mu) \to \mathcal{K}_{\theta_\gamma}$ be the unitary operator satisfying*

$$\Phi_\gamma^* U_\gamma = \mathcal{M}_{\theta_\gamma} \Phi_\gamma^*,$$

and such that $\Phi_\gamma^* b = c^\gamma$, $\Phi_\gamma^* b_1 = c_1^\gamma$.

Then for all $f \in C^1(\mathbb{T})$

$$\Phi_\gamma^* f(z) = A_\gamma(z) f(z) + B_\gamma(z) \int \frac{f(\xi) - f(z)}{1 - \bar{\xi} z} \, d\mu(\xi) \tag{32}$$

where $A_\gamma(z) = c^\gamma(z)$, $B_\gamma(z) = c^\gamma(z) - z c_1^\gamma(z)$.

Idea of the proof To some extent we mirror the proof of Theorem 2.6. However, several miracles occur (beyond the fact that we are now dealing with the vector-valued setting of the model space makes the computations are more cumbersome):

Again, we begin with the intertwining relation $\Phi_\gamma^* U_\gamma = \mathcal{M}_{\theta_\gamma} \Phi_\gamma^*$ and evaluate the projection of the model operator

$$\mathcal{M}_{\theta_\gamma} = M_z - z c_1^\gamma (c_1^\gamma)^* - \theta_\gamma(0) c^\gamma (c_1^\gamma)^* = M_z + (\gamma c^\gamma - z c_1^\gamma)(c_1^\gamma)^*. \tag{33}$$

We notice that the model operator $\mathcal{M}_{\theta_\gamma}$ on $\mathcal{K}_\theta$ is a rank one perturbation of the unitary M_z, and the operator U_γ on $L^2(\mu)$ is a rank one perturbation of the unitary U_1 (multiplication by the independent variable). So we expect that the commutator $\Phi_\gamma^* U_1 - M_z \Phi_\gamma^*$ is at most of rank 2. But in fact, it turns out to be of rank one!

Indeed, the intertwining relation $\Phi_\gamma^* U_\gamma = \mathcal{M}_{\theta_\gamma} \Phi_\gamma^*$ can be rewritten as

$$\Phi_\gamma^* U_1 + (\gamma - 1) c^\gamma b_1^* = M_z \Phi_\gamma^* + (\gamma c^\gamma - c_2^\gamma) b_1^*$$

(here we used that $\Phi_\gamma^* b = c^\gamma$ and $(c_1^\gamma)^* \Phi_\gamma^* = (\Phi_\gamma c_1^\gamma)^* = b_1^*$), and therefore

$$\Phi_\gamma^* U_1 = M_z \Phi_\gamma^* + (c^\gamma - z c_1^\gamma) b_1^*. \tag{34}$$

From here, we proceed in analogy to the proof of Theorem 2.6 to obtain a formula for $\Phi_\gamma^* \xi^n$.

The formula for $\Phi_\gamma^* \bar{\xi}^n$ cannot be computed by simply taking the formal adjoint of the commutation relation (34). This is due to the fact that in general $zc_1^\gamma \notin \mathcal{K}_\theta$. Instead we compute the adjoint of the model operator in analogy to (33)

$$\mathcal{M}_{\theta_\gamma}^* = M_{\bar{z}} - M_{\bar{z}} c^\gamma (c^\gamma)^* - \overline{\theta(0)} c_1^\gamma (c^\gamma)^* = M_{\bar{z}} + (\overline{\gamma} c_1^\gamma - M_{\bar{z}} c^\gamma)(c^\gamma)^*.$$

We find ourselves in the lucky situation that the formulas for $\Phi_\gamma^* \xi^n$ and $\Phi_\gamma^* \bar{\xi}^n$ turn out to be the same. □

In the Sz.-Nagy–Foiaş transcription we derive concrete formulas. For $\gamma = 0$ we have $\theta_0(0) = 0$ and by (28) we obtain $\theta_\gamma(0) = -\gamma$. With this, the vector-valued functions $A_\gamma(z)$ and $B_\gamma(z)$ in the universal representation formula (32) evaluate to

$$A_\gamma(z) = c^\gamma(z) = (1 - |\gamma|^2)^{-1/2} \begin{pmatrix} 1 + \overline{\gamma}\theta_\gamma(z) \\ \overline{\gamma}\Delta_\gamma(z) \end{pmatrix} = \begin{pmatrix} \frac{(1-|\gamma|^2)^{1/2}}{1-\overline{\gamma}\theta_0(z)} \\ \frac{\overline{\gamma}\Delta_0(z)}{|1-\overline{\gamma}\theta_0(z)|} \end{pmatrix}, \tag{35}$$

$$B_\gamma(z) = c^\gamma(z) - zc_1^\gamma(z) = (1 - |\gamma|^2)^{-1/2} \begin{pmatrix} 1 + (\overline{\gamma} - 1)\theta_\gamma(z) - \gamma \\ (\overline{\gamma} - 1)\Delta_\gamma(z) \end{pmatrix} \tag{36}$$

$$= \begin{pmatrix} (1 - |\gamma|^2)^{1/2}(1 - \theta_0(z))/(1 - \overline{\gamma}\theta_0(z)) \\ (\overline{\gamma} - 1)\Delta_0(z)/|1 - \overline{\gamma}\theta_0(z)| \end{pmatrix},$$

where $\Delta_\gamma = (1 - |\theta_\gamma|^2)^{1/2}$.

Singular Integral Operators and a Representation for Φ_γ^* in the Sz.-Nagy–Foiaş Transcription

In this section we get a representation of Φ_γ^* adapted to the Sz.-Nagy–Foiaş transcription, similar to the representations given in Theorem 2.5.

We first note that for $v(\xi) = |B_\gamma(\xi)|^2$ the kernel $K(z, \xi) = 1/(1 - \bar{\xi}z)$ is an $L^2(\mu) \to L^2(v)$ restrictedly bounded kernel, see Definition 3.1. Indeed, taking C^1 functions f and g with separated compact supports we get that

$$(\mathcal{V}_\gamma f, g) = \int_{\mathbb{T}} \frac{\big(B(z)f(\xi), g(z)\big)_{\mathcal{H}}}{1 - \bar{\xi}z} d\mu(\xi) \frac{|dz|}{2\pi},$$

and standard approximation reasoning extend this formula to all bounded functions with separated supports. But that means that the vector-valued kernel[1] $B_\gamma(z)/(1-\overline{\xi}z)$ is a kernel of a singular integral operator $L^2(\mu) \to L^2$ with norm 1, and so it is $L^2(\mu) \to L^2$ restrictedly bounded (with restricted norm at most 1).

A standard renormalization argument then implies the $L^2(\mu) \to L^2(v)$ restricted boundedness of the scalar kernel $1/(1-\overline{\xi}z)$.

Therefore, as we discussed in section "Singular Integral Operators", see Theorem 3.2 and Remark 3.3, the regularized operators T_r with kernel $K_r(z,\xi) = 1/(1-r\overline{\xi}z)$ are uniformly bounded operators $L^2(\mu) \to L^2(v)$, so the operators $B_\gamma T_r$ are uniformly bounded $L^2(\mu) \to L^2$.

On the other hand, the boundary values of the Cauchy transform R (defined in (26)) exist a.e. with respect to Lebesgue measure by the classical theory of Hardy spaces; it is easier than for the operators $\mathcal{V}_\alpha$, since we do not need a.e. convergence with respect to a singular measure here.

In combination with the uniform bounds we can see the existence of weak operator topology limit

$$T_\pm := \text{w.o.t.-}\lim_{r\to 1^\mp} T_r.$$

Note also that $T_\pm$ can be defined as a.e. limits, $T_\pm f = \lim_{r\to 1^\mp} T_r f$.

Theorem 4.3 *Operator Φ_γ^* can be represented in the Sz.-Nagy–Foiaş transcription as*

$$(1-|\gamma|^2)^{1/2}\Phi_\gamma^* f = \begin{pmatrix} 0 \\ (\overline{\gamma}-(\overline{\gamma}-1)T_+\mathbf{1})\Delta_\gamma \end{pmatrix} f + \begin{pmatrix} (1+\overline{\gamma}\theta_\gamma)/T_+\mathbf{1} \\ (\overline{\gamma}-1)\Delta_\gamma \end{pmatrix} T_+ f$$

$$= \begin{pmatrix} 0 \\ \frac{1-\overline{\gamma}\theta_0}{|1-\overline{\gamma}\theta_0|} T_+\mathbf{1}\cdot\Delta_0 \end{pmatrix} f + \begin{pmatrix} \frac{1-|\gamma|^2}{1-\overline{\gamma}\theta_0}\cdot\frac{1}{T_+\mathbf{1}} \\ (\overline{\gamma}-1)\frac{(1-|\gamma|^2)^{1/2}}{|1-\overline{\gamma}\theta_0|}\Delta_0 \end{pmatrix} T_+ f$$

for $f \in L^2(\mu)$.

As expected this formula reduces to the normalized Cauchy transform for $\gamma = 0$ and inner functions θ. To see this, we notice that the second component collapses as $\Delta(z) = (1-|\theta(z)|^2)^{1/2} = 0$ Lebesgue a.e. on $\mathbb{T}$, and that $T_+ f/T_+\mathbf{1}$ is equal to the normalized Cauchy transform.

Idea of the proof For smooth functions f we replace the term $1-\overline{\xi}z$ in the denominator of (32) by $1-r\overline{\xi}z$ and take the limit as $r \to 1^-$. We obtain the same formula (32). Since we also have weak convergence of the operators we have

[1]We did not discuss singular integral operators with vector-valued kernels, but the extension of the theory presented in section "Singular Integral Operators" to the case of kernels with values in $\mathbb{R}^d$ or $\mathbb{C}^d$ is trivial and we omit it.

$$(T_+f)(z) - f(z)(T_+\mathbf{1})(z) = \int_{\mathbb{T}} \frac{f(\xi) - f(z)}{1 - \bar{\xi} z} d\mu(\xi), \qquad z \in \mathbb{T}.$$

We extend the operator by continuity to all of $L^2(\mu)$ and derive

$$T_+\mathbf{1} = 1/(1 - \theta_0) \tag{37}$$

from (26) through (28). Technical computations then yield the desired formula. □

Interestingly, similar arguments show that

$$T_-\mathbf{1} = -\bar{\theta}_0/(1 - \bar{\theta}_0). \tag{38}$$

Representing Φ^* in the de Branges–Rovnyak Transcription

We translate the formula in the last Theorem 4.3 from the Sz.-Nagy–Foiaş transcription to the de Branges–Rovnyak transcription, rather than starting from the universal representation formula in Theorem 4.2. This strategy seemed less cumbersome as we circumvent having to re-do much of the subtle work of regularizing singular integral operators. Also, we found it refreshing to understand the connection between the transcriptions.

By virtue of the definition of the Sz.-Nagy–Foiaş model space $\mathcal{K}_\theta$, see (23), a function

$$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \in \begin{pmatrix} H^2 \\ \operatorname{clos} \Delta L^2 \end{pmatrix}$$

is in $\mathcal{K}_\theta$ if and only if

$$g_- := \overline{\theta} g_1 + \Delta g_2 \in H^2_- := L^2(\mathbb{T}) \ominus H^2. \tag{39}$$

Note, that knowing g_1 and g_- one can restore g_2 on $\mathbb{T}$:

$$g_2 \Delta = g_- - g_1 \overline{\theta}.$$

The equality (39) means that the pair $g_+ = g_1$ and g_- belongs to the de Branges–Rovnyak space, see (25). It is also not hard to check that the norm of the pair (g_1, g_-) in the Branges–Rovnyak space (i.e. in the weighted space $L^2(W)$, $W = W_\theta^{[-1]}$, see subsection "de Branges–Rovnyak Transcription") coincides with the norm of the pair (g_1, g_2) in the Sz.-Nagy–Foiaş space (i.e. in non-weighted L^2). Indeed, we have

$$\begin{pmatrix} g_1 \\ g_- \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \overline{\theta} & \Delta \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}.$$

Let B be a "Borel support" of Δ, i.e. the set where one of the representative from the equivalence class of Δ is different from 0. A direct computation shows that for

$$W_\theta = \begin{pmatrix} 1 & \theta \\ \overline{\theta} & 1 \end{pmatrix}$$

we have a.e. on $\mathbb{T}$

$$\begin{pmatrix} 1 & \theta \\ 0 & \Delta \end{pmatrix} W_\theta^{[-1]} \begin{pmatrix} 1 & 0 \\ \overline{\theta} & \Delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{1}_B \end{pmatrix},$$

which gives the desired equality of the norms. (Here $\mathbf{1}_B$ denotes the characteristic function of B.)

Note that functions in H^2_- admit analytic continuation to the exterior of the unit disc, so a function in $\mathcal{K}_\theta$ is determined by the boundary values of two functions g_1 and g_- analytic in $\mathbb{D}$ and $\operatorname{ext}(\overline{\mathbb{D}})$ respectively.

Since the first component of a function in the Branges–Rovnyak space is the same as in the Sz.-Nagy–Foiaş space and by virtue of Theorem 4.3 we immediately know

$$g_+(z) = g_1(z) = \frac{(1-|\gamma|^2)^{1/2}}{1-\overline{\gamma}\theta_0} \frac{T_+ f}{T_+\mathbf{1}}, \tag{40}$$

where T_+ was defined in the paragraph prior to Theorem 4.3.

The second component $g_- = g_-^\gamma$ is analytic on $\operatorname{ext}(\overline{\mathbb{D}})$. Therefore, we do need to return to the universal representation formula. After some reformulation we observe.

Theorem 4.4 (Theorem 5.5 of [31]) *Let μ be not the Lebesgue measure. Then the function $g_- = g_-^\gamma$ is given by*

$$g_-^\gamma = (1-|\gamma|^2)^{-1/2}\left(\overline{\theta}_\gamma + \overline{\gamma}\right)\frac{T_- f}{T_-\mathbf{1}} = \frac{(1-|\gamma|^2)^{1/2}\overline{\theta}_0}{1-\gamma\overline{\theta}_0}\cdot\frac{T_- f}{T_-\mathbf{1}}. \tag{41}$$

Formulas for $\boldsymbol{\Phi_\gamma}$

A representation of the Clark operator Φ_γ is given in terms of the components g_+ and g_- of a vector in the de Branges–Rovnyak transcription. This formula is given piecewise. For a function $f \in L^2(\mu)$ we denote by f_a and f_s its "absolutely continuous" and "singular" parts, respectively. Formally, f_s and f_a can be defined as Radon–Nikodym derivatives $f_\mathrm{s} = d(f\mu)_\mathrm{s}/d\mu_\mathrm{s}$, $f_\mathrm{a} = d(f\mu)_\mathrm{a}/d\mu_\mathrm{a}$.

Let w denote the density of the absolutely continuous part of $d\mu$, i.e. $w = d\mu/dx \in L^1$.

Theorem 4.5 *Let* $g = \begin{pmatrix} g_+ \\ g_- \end{pmatrix} \in \mathcal{K}_{\theta_\gamma}$ *(in the de Branges–Rovnyak transcription) and let* $f \in L^2(\mu)$, $f = \Phi_\gamma g$. *Then*

(1) *the non-tangential boundary values of the function*

$$z \mapsto \frac{1-\overline{\gamma}}{(1-|\gamma|^2)^{1/2}} g_+(z), \qquad z \in \mathbb{D}$$

exist and coincide with f_{s} μ_{s}*-a.e. on* $\mathbb{T}$.

(2) *for the "absolutely continuous" part* f_{a} *of* f

$$(1-|\gamma|^2)^{1/2} w f_{\mathrm{a}} = \frac{1-\overline{\gamma}\theta_0}{1-\theta_0} g_+ + \frac{1-\gamma\overline{\theta}_0}{1-\overline{\theta}_0} g_-$$

a.e. on $\mathbb{T}$.

We provide the idea of the proof. First consider statement (2). By taking the limit as $r \to 1$ in $T_r f - T_{1/r} f$ we prove the Fatou type result [31, Lemma 5.6]:

$$T_+ f - T_- f = wf \quad \text{a.e. on } \mathbb{T}$$

(with respect to the Lebesgue measure) for all $f \in L^2(\mu)$. Together with (38) and (37) we can use the representations (40) and (41) for g_+ and g_- to see the desired result for the absolutely continuous part.

Statement (1) uses Poltoratskii's theorem [38, Theorem 2.7].

Clark Operator for Other α*,* $|\alpha| = 1$

Consider the Clark operator $\Phi_{\alpha,\gamma} : \mathcal{K}_{\theta_\gamma} \to L^2(\mu_\alpha)$, where μ_α, $|\alpha| = 1$ is the spectral measure corresponding to the cyclic vector b of the unitary operator U_α. Operator $\Phi_{\alpha,\gamma}$ is a unitary operator, which intertwines the model operator $\mathcal{M}_{\theta_\gamma}$ and the c.n.u. contraction $(U_\gamma)_\alpha$ which is the operator U_γ in the spectral representation of the operator U_α.

We deduce everything from the results we already obtained. First, let us write the c.n.u. contraction U_γ, $|\gamma| < 1$ as a rank one perturbation of the unitary operator U_α, $|\alpha| = 1$:

$$U_\gamma = U + (\gamma - 1) b b_1^* = U + (\alpha - 1) b b_1^* + (\gamma - \alpha) b b_1^* = U_\alpha + (\gamma/\alpha - 1) b \widetilde{b}_1^*,$$

where $\widetilde{b}_1 = \overline{\alpha} b_1$.

From this we can just read off the results for $\alpha \in \mathbb{T}$ from the results we already proved (for $\alpha = 1$).

But to be consistent, we need the operators $\Phi^*_{\alpha,\gamma}$ to agree. First we need them to the same model spaces, so let us fix the model spaces to be the ones we got for the case $\alpha = 1$. Second, we want them to be consistent with respect to the operators $\mathcal{V}_\alpha$ from section "Unitary Rank One Perturbations":

$$\Phi^*_{\alpha,\gamma} = \Phi^*_\gamma \mathcal{V}^*_\alpha \,. \tag{42}$$

Then an appropriately interpreted "universal" representation formula (Theorem 4.2) gives us a formula for $\Phi^*_{\alpha,\gamma}$.

Namely, in the spectral representation of U_α the c.n.u. contraction U_γ is given by

$$M_\xi + (\gamma/\alpha - 1) b^\alpha (b_1^\alpha)^*, \tag{43}$$

where $b^\alpha = \mathcal{V}_\alpha b$, $b_1^\alpha = \mathcal{V}_\alpha \widetilde{b_1} = \overline{\alpha}\mathcal{V}_\alpha b_1$, which yields $b^\alpha = \mathbf{1}$, $b_1^\alpha(\xi) \equiv \overline{\xi}$, $\xi \in \mathbb{T}$. Notice that

$$c^{\alpha,\gamma} = \Phi^*_{\alpha,\gamma} b^\alpha = \Phi^*_\gamma \mathcal{V}^*_\alpha b^\alpha = \Phi^*_\gamma b = c^\gamma,$$

and that

$$c_1^{\alpha,\gamma} = \Phi^*_{\alpha,\gamma} b_1^\alpha = \Phi^*_\gamma \mathcal{V}^*_\alpha b_1^\alpha = \overline{\alpha}\Phi^*_\gamma b_1 = \overline{\alpha} c_1^\gamma.$$

Therefore, to get the formula for $\Phi^*_{\alpha,\gamma}$ with $\Phi^*_{\alpha,\gamma} b^\alpha = c^\gamma$ (i.e. such that $\Phi^*_{\alpha,\gamma}\mathbf{1} = c^\gamma$) one just has to replace in (32) μ by μ_α, and c_1^γ by $\overline{\alpha} c_1^\gamma$ (c^γ remains the same). Note, that as long as c^γ and c_1^γ are computed, the parameter γ does not appear in (32).

Now let us get the representations in the Sz.-Nagy–Foiaş and de Branges–Rovnyak transcriptions. One of the ways to get the formula for $\Phi^*_{\alpha,\gamma}$ would be to take the "universal formula" above and then repeat the proofs of Theorem 4.3 and of Theorem 4.4.

But the there is a simpler (in our opinion) way, that allows us to get the result with almost no computations: one just have to "translate" Theorems 4.3, 4.4 to the spectral representation of U_α.

In both these theorems the characteristic function and the parameter γ are included explicitly, so we need to see how they change when we move to the spectral representation of U_α.

If we want to apply know formulas (27), they give us the characteristic function $\theta^\alpha_{\gamma/\alpha}$ of the operator (43) with b_1^α and b^α taken for the basis vectors in the corresponding defect subspaces.

So, by replacing μ with μ_α and γ with γ/α in (27) and (24) we get the characteristic function and the defect given by

$$\theta^\alpha_{\gamma/\alpha} = \overline{\alpha}\theta_\gamma \,, \qquad \text{and} \qquad \Delta_{\gamma,\alpha} = \Delta_\gamma \,.$$

Substituting these functions to (35) and replacing γ there by γ/α we get a representation formula for the adjoint of the Clark operator mapping $L^2(\mu_\alpha) \to \mathcal{K}_{\bar\alpha\theta_\gamma}$ in Sz.-Nagy–Foiaş transcription,

$$(1-|\gamma|^2)^{1/2}\widetilde{\Phi}^*_{\alpha,\gamma} f = \begin{pmatrix} 0 \\ (\overline{\gamma}/\overline{\alpha} - (\overline{\gamma}/\overline{\alpha}-1)T^\alpha_+\mathbf{1})\Delta_\gamma \end{pmatrix} f + \begin{pmatrix} (1+\overline{\gamma}\theta_\gamma)/T^\alpha_+ 1 \\ (\overline{\gamma}/\overline{\alpha}-1)\Delta_\gamma \end{pmatrix} T^\alpha_+ f \tag{44}$$

$$= \begin{pmatrix} 0 \\ \frac{1-\overline{\gamma}\theta_0}{|1-\overline{\gamma}\theta_0|} T^\alpha_+\mathbf{1}\cdot\Delta_0 \end{pmatrix} f + \begin{pmatrix} \frac{1-|\gamma|^2}{1-\overline{\gamma}\theta_0}\cdot\frac{1}{T^\alpha_+\mathbf{1}} \\ (\overline{\gamma}/\overline{\alpha}-1)\frac{(1-|\gamma|^2)^{1/2}}{|1-\overline{\gamma}\theta_0|}\Delta_0 \end{pmatrix} T^\alpha_+ f\,,$$

where we let $T^\alpha_+ f$ denote the non-tangential boundary values of $Rf\mu_\alpha(z)$, $z \in \mathbb{D}$.

But the above formula is not yet the formula we are looking for! To get it we applied Theorem 4.3 with μ_α instead of μ and $\theta^\alpha_{\gamma/\alpha} = \overline{\alpha}\theta_\gamma$ instead of θ_γ. But that means that the result in the right hand side there belongs to $\mathcal{K}_{\overline{\alpha}\theta}$. So the above expression is an absolutely correct formula giving the representation of the operator $\Phi^*_{\alpha,\gamma}$ in the model space $\mathcal{K}_{\overline{\alpha}\theta_\gamma}$; that is why we used $\widetilde{\Phi}^*_{\alpha,\gamma}$ and not $\Phi^*_{\alpha,\gamma}$ there.

To get the representation with the model space $\mathcal{K}_{\theta_\gamma}$ we notice that the map

$$\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \mapsto \begin{pmatrix} g_1 \\ \overline{\alpha} g_2 \end{pmatrix}$$

is a unitary map from $\mathcal{K}_{\overline{\alpha}\theta_\gamma}$ onto $\mathcal{K}_{\theta_\gamma}$. Moreover, it maps the defect vector c given by equation (30) for the space $\mathcal{K}_{\overline{\alpha}\theta_\gamma}$ to the corresponding defect vector c for the space $\mathcal{K}_{\theta_\gamma}$. Therefore, to obtain the representation formula for $\Phi^*_{\alpha,\gamma}$ we need to multiply the bottom entries in (44) by $\overline{\alpha}$, which gives us the following representation.

Theorem 4.6 *Operator $\Phi^*_{\alpha,\gamma}$ can be represented in the Sz.-Nagy–Foiaş transcription as*

$$(1-|\gamma|^2)^{1/2}\Phi^*_{\alpha,\gamma} f = \begin{pmatrix} 0 \\ (\overline{\gamma} - (\overline{\gamma}-\overline{\alpha})T^\alpha_+\mathbf{1})\Delta_\gamma \end{pmatrix} f + \begin{pmatrix} (1+\overline{\gamma}\theta_\gamma)/T^\alpha_+ 1 \\ (\overline{\gamma}-\overline{\alpha})\Delta_\gamma \end{pmatrix} T^\alpha_+ f$$

$$= \begin{pmatrix} 0 \\ \overline{\alpha}\frac{1-\overline{\gamma}\theta_0}{|1-\overline{\gamma}\theta_0|} T^\alpha_+\mathbf{1}\cdot\Delta_0 \end{pmatrix} f + \begin{pmatrix} \frac{1-|\gamma|^2}{1-\overline{\gamma}\theta_0}\cdot\frac{1}{T^\alpha_+\mathbf{1}} \\ (\overline{\gamma}-\overline{\alpha})\frac{(1-|\gamma|^2)^{1/2}}{|1-\overline{\gamma}\theta_0|}\Delta_0 \end{pmatrix} T^\alpha_+ f\,.$$

Few Remarks About Clark Theory for the Dissipative Case

Consider a family of rank one perturbations similar to section "Self-Adjoint and Unitary Rank One Perturbations", but with perturbation parameter $\alpha \in \mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. In other words, in the spectral representation of A (with respect to the cyclic vector φ and spectral measure $\mu = \mu^\varphi$) we study the family of perturbations given by

$$A_\alpha = M_t + \alpha(\,\cdot\,, \mathbf{1})_{L^2(\mu)}\mathbf{1} \qquad \text{on } L^2(\mu) \text{ with } \alpha \in \mathbb{C}_+.$$

Recall that we consider the extended class of form bounded rank one perturbations. In the spectral representation this condition is equivalent to

$$\int_{\mathbb{R}} \frac{d\mu(t)}{1+|t|} < \infty.$$

Without going into details on the definition of the perturbation in this case, we just say that one of the ways is to use the resolvent formula (3).

While there is no "canonical" model for the dissipative operator, a widely accepted way is to construct the model for the Cayley transform $\widetilde{T}_\alpha = (A_\alpha - i\mathbf{I})(A_\alpha + i\mathbf{I})^{-1}$.

So, let us compute $\widetilde{T}_\alpha$, introducing some notation along the way.

Denote by $\widetilde{U}$ the Cayley transform of $A = A_0$, $\widetilde{U} = (A - i\mathbf{I})(A + i\mathbf{I})^{-1}$. Using the resolvent formula (3) and denoting

$$\widetilde{b} := \|(A+i\mathbf{I})^{-1}\varphi\|^{-1}(A+i\mathbf{I})^{-1}\varphi, \qquad \widetilde{b}_1 := \|(A-i\mathbf{I})^{-1}\varphi\|^{-1}(A-i\mathbf{I})^{-1}\varphi$$

we can write

$$T_\alpha = \widetilde{U}_\gamma = \widetilde{U} + (\gamma - 1)\widetilde{b}(\widetilde{b}_1)^*,$$

where

$$\gamma = \gamma(\alpha) = \frac{1+\alpha\overline{Q}}{1+\alpha Q}, \qquad Q = ((A+i\mathbf{I})^{-1}\varphi, \varphi) = \int_{\mathbb{R}} \frac{d\mu(s)}{s+i}. \tag{45}$$

If we denote

$$F(z) := \int_{\mathbb{R}} \frac{d\mu(s)}{s-z},$$

we get that $Q = F(-i) = \overline{F(i)}$.

Note also that $\|\widetilde{b}\| = \|\widetilde{b}_1\| = 1$ and $\widetilde{b}_1 = \widetilde{U}^*\widetilde{b}$. It is obvious that $\gamma(\alpha) \in \mathbb{T}$ for $\alpha \in \mathbb{R}$. Since $\operatorname{Im} Q < 0$, we conclude that $\gamma(\alpha) \in \mathbb{D}$ for $\operatorname{Im}\alpha > 0$. Thus T_α is a contractive rank one perturbation of the unitary operator $\widetilde{U}$. Under our assumptions about cyclicity of φ, one can easily see that $\widetilde{b}$ is a $*$-cyclic vector for $\widetilde{U}$, so $\widetilde{U}$ is unitarily equivalent to the multiplication $U = M_\xi$ by the independent variable ξ in $L^2(\mu_{\mathbb{T}})$, where $\mu_{\mathbb{T}}$ is the spectral measure of $\widetilde{U}$ corresponding to the vector $\widetilde{b}$.

Let us fix some notation: for $\gamma = \gamma(\alpha)$ given by (45) we denote $\widetilde{U}_\gamma = \widetilde{T}_\alpha$, and by $U_\gamma = T_\alpha$ we denote the representation of the same operator in $L^2(\mu_{\mathbb{T}})$. In other words, we use T in conjunction with the parameter $\alpha \in \mathbb{C}_+$ and U in conjunction with the parameter $\gamma = \gamma(\alpha) \in \mathbb{D}$; also we use $\widetilde{\ }$ for the operators in $L^2(\mu)$, and T and U act in $L^2(\mu_{\mathbb{T}})$.

The spectral measure $\mu_{\mathbb{T}}$ of $\widetilde{U}$ is easily computed. Namely, if ω denotes the standard conformal map from $\mathbb{C}_+$ to $\mathbb{D}$ (and from $\mathbb{R}$ to $\mathbb{T}$),

$$\omega(z) := \frac{z-i}{z+i}, \qquad \omega^{-1}(\xi) = i\frac{1+\xi}{1-\xi},$$

then one can easily see that

$$\mu_{\mathbb{T}} := \widetilde{\mu} \circ \omega^{-1}, \qquad \text{where } d\widetilde{\mu}(x) = \frac{1}{P} \cdot \frac{d\mu(x)}{1+x^2};$$

here by $\widetilde{\mu} \circ \omega^{-1}$ we mean that $\widetilde{\mu} \circ \omega^{-1}(E) = \widetilde{\mu}(\omega^{-1}(E))$, $E \subset \mathbb{T}$, and $P := \int_{\mathbb{R}} \frac{d\mu(t)}{1+t^2}$.

What Is the Model for the Dissipative Case?

As we mentioned above, it is customary for dissipative operator to consider the model for its Cayley transform. Using formulas (27) with $\mu_{\mathbb{T}}$ for μ, and the above description of $\mu_{\mathbb{T}}$ we can write the characteristic function θ_γ, $\gamma = \gamma(\alpha)$.

However, since our original objects live on the real line (in $L^2(\mu)$), it is natural to consider the model also to be a space of functions on the real line. The standard unitary mapping $\Omega : L^2(\mathbb{T}) \to L^2(\mathbb{R})$,

$$\Omega f(x) := \frac{1}{\sqrt{\pi}(x+i)} f \circ \omega(x)$$

maps $H^2(\mathbb{D})$ onto $H^2(\mathbb{C}_+)$ and so $H^2_-(\mathbb{D})$ onto $H^2_-(\mathbb{C}_+)$. So if we use Ω^{-1} to transfer the model space $\mathcal{K}_\theta$ to the space of functions on $\mathbb{R}$, the model space on the real line in Sz.-Nagy–Foiaş and the de Branges–Rovnyak transcriptions will be defined exactly the same way as the model space on the circle.

The multiplication M_ω by the function ω on $\mathbb{R}$ corresponds to the multiplication by ξ on $\mathbb{T}$.

Note also that $(\pi/P)^{1/2}\Omega$ is a unitary operator $L^2(\mu_{\mathbb{T}}) \to L^2(\mu)$ and that the map $f \mapsto f \circ \omega$ maps $L^2(\mu_{\mathbb{T}}) \to L^2(\widetilde{\mu})$ unitarily.

Characteristic Function in the Half-Plane

Let us now compute the characteristic function for A_α. For $\gamma = \gamma(\alpha)$ defined by (45) let $\underset{\sim}{\theta_\gamma}$ be the characteristic function of T_α computed with respect to the vectors $\widetilde{b}_1$ and $\widetilde{b}$. Let

$$\Theta_\alpha = \widetilde{\theta_\gamma} := \theta_\gamma \circ \omega$$

be the transfer of θ_γ from the disc to the half-plane; note that we use capital Θ in conjunction with the parameter $\alpha \in \mathbb{C}_+$.

Let us now transfer the Cauchy type integrals (26) to $\mathbb{C}_+$. For $w \in \mathbb{C}_+$ let $\lambda = \omega(w)$. Then

Lemma 5.1 *We have*

$$R\mu_{\mathbb{T}}(\lambda) = \int_{\mathbb{T}} \frac{d\mu_{\mathbb{T}}(\xi)}{1-\overline{\xi}\lambda} \qquad = \frac{1}{2iP}\int_{\mathbb{R}} \left[\frac{1}{x-w} - \frac{1}{x+i}\right] d\mu_{\mathbb{R}}(x) =: \widetilde{R}\mu(w), \tag{46}$$

$$R_1\mu_{\mathbb{T}}(\lambda) = \int_{\mathbb{T}} \frac{\overline{\xi}\lambda d\mu_{\mathbb{T}}(\xi)}{1-\overline{\xi}\lambda} \qquad = \frac{1}{2iP}\int_{\mathbb{R}} \left[\frac{1}{x-w} - \frac{1}{x-i}\right] d\mu_{\mathbb{R}}(x) =: \widetilde{R}_1\mu(w), \tag{47}$$

$$R_2\mu_{\mathbb{T}}(\lambda) = \int_{\mathbb{T}} \frac{1+\overline{\xi}\lambda}{1-\overline{\xi}\lambda} d\mu_{\mathbb{T}}(\xi) \quad = \frac{1}{iP}\int_{\mathbb{R}} \left[\frac{1}{x-w} - \frac{x}{x^2+1}\right] d\mu_{\mathbb{R}}(x) =: \widetilde{R}_2\mu(w). \tag{48}$$

Using formulas (27) for the disc we can write the characteristic functions as

$$\widetilde{\theta}_\gamma(w) = -\gamma + \frac{(1-|\gamma|^2)\widetilde{R}_1\mu(w)}{1+(1-\overline{\gamma})\widetilde{R}_1\mu(w)} = \frac{(1-\gamma)\widetilde{R}_2\mu(w) - (1+\gamma)}{(1-\overline{\gamma})\widetilde{R}_2\mu(w) + (1+\overline{\gamma})}, \qquad w \in \mathbb{C}_+. \tag{49}$$

Note that the formulas for $\widetilde{\theta}_0$ ($\gamma = 0$, equivalently $\alpha = -1/\overline{Q} = -1/F(i)$) are especially simple. And $\widetilde{\theta}_0$ is related to $\widetilde{\theta}_\gamma$ by a fractional transformation:

$$\widetilde{\theta}_\gamma = \frac{\widetilde{\theta}_0 - \gamma}{1-\overline{\gamma}\widetilde{\theta}_0} \qquad \text{or equivalently} \qquad \widetilde{\theta}_0 = \frac{\widetilde{\theta}_\gamma + \gamma}{1+\overline{\gamma}\widetilde{\theta}_\gamma}.$$

Model and Defect Vectors in the Half-Plane

Recall that the model operator $\mathcal{M}_{\widetilde{\theta}_\gamma}$ is the compression of the multiplication operator M_ω by the function ω,

$$\mathcal{M}_{\widetilde{\theta}_\gamma} f := P_{\mathcal{K}_{\widetilde{\theta}_\gamma}} M_\omega f, \qquad f \in \mathcal{K}_{\widetilde{\theta}_\gamma}.$$

Let us compute defect subspaces of $\mathcal{M}_{\widetilde{\theta}_\gamma}$.

Considering vectors $c = c^\gamma$ and $c_1 = c_1^\gamma$ defined by (30), (31) with $\mu_{\mathbb{T}}$ instead of μ, define $\widetilde{c} := c \circ \omega$, $\widetilde{c}_1 := c_1 \circ \omega$, $\widetilde{\Delta} := \Delta \circ \omega = 1 - |\widetilde{\theta}_0|^2$. Computing we get in the Sz.-Nagy–Foiaş transcription

$$\widetilde{c}(z) := \left(1 - |\widetilde{\theta}_\gamma(i)|^2\right)^{-1/2} \begin{pmatrix} 1 - \overline{\widetilde{\theta}_\gamma(i)}\widetilde{\theta}_\gamma(z) \\ -\overline{\widetilde{\theta}_\gamma(i)}\widetilde{\Delta}(z) \end{pmatrix},$$

$$\widetilde{c}_1(z) := \left(1 - |\widetilde{\theta}_\gamma(i)|^2\right)^{-1/2} \begin{pmatrix} \omega(z)^{-1}\left(\widetilde{\theta}_\gamma(z) - \widetilde{\theta}_\gamma(i)\right) \\ \omega(z)^{-1}\widetilde{\Delta}(z) \end{pmatrix}.$$

Then the defect subspaces $\mathfrak{D}_{\mathcal{M}^*_{\widetilde{\theta}_\gamma}}$ and $\mathfrak{D}_{\mathcal{M}_{\widetilde{\theta}_\gamma}}$ of $\mathcal{M}_{\widetilde{\theta}_\gamma}$ are spanned by the vectors

$$\frac{\widetilde{c}(z)}{\sqrt{\pi}(z+i)}, \qquad \frac{\widetilde{c}_1(z)}{\sqrt{\pi}(z+i)}$$

and these vectors agree.

Representations of the Adjoint Clark Operator in the Half-Plane

Using these formulas we can transfer the universal representation formula given by Theorem 4.2 from the unit circle $\mathbb{T}$ to the real line $\mathbb{R}$. For a function f on the real line $\mathbb{R}$ define $\widetilde{f}$ by

$$\widetilde{f}(x) := (x+i)\cdot f(x).$$

and let $f_{\mathbb{T}} := \widetilde{f} \circ \omega^{-1}$. Then we can easily transfer Theorem 4.2 from the disc $\mathbb{D}$ to the half-plane $\mathbb{C}_+$.

To simplify the notation let us assume that the measure μ is Poisson normalized, i.e. that

$$P := \int_{\mathbb{R}} \frac{d\mu(x)}{x^2+1} = 1.$$

Formulas for the general case $P \neq 1$ can be then obtained if one notice that the map $f \mapsto P^{1/2} f$ is a unitary map $L^2(\mu) \to L^2(\mu/P)$.

A Universal Representation Formula

Theorem 5.2 (A "universal" representation formula for dissipative perturbations) *Let the measure μ be Poisson normalized ($P = 1$). Let $\widetilde{\theta}_\gamma$ be the characteristic function of $T_\alpha = \widetilde{U}_\gamma$, $|\gamma| < 1$, computed with respect to the vectors $\widetilde{b}_1$*

and $\widetilde{b}$ *(note that* $\widetilde{\theta}_\gamma$ *is given by* (49)*). Let* $\mathcal{K}_{\widetilde{\theta}_\gamma}$ *and* $\mathcal{M}_\gamma = \mathcal{M}_{\widetilde{\theta}_\gamma}$ *be the model subspace and the model operator respectively. Let* $\widetilde{\Phi}^*_\gamma : L^2(\mu) \to \mathcal{K}_{\widetilde{\theta}_\gamma}$ *be the unitary operator satisfying*

$$\widetilde{\Phi}^*_\gamma \widetilde{U}_\gamma = \mathcal{M}_{\widetilde{\theta}_\gamma} \widetilde{\Phi}^*_\gamma,$$

and such that $\widetilde{\Phi}^*_\gamma \widetilde{b}(z) = \widetilde{c}^\gamma(z)/(\sqrt{\pi}(z+i))$, $\widetilde{\Phi}^*_\gamma \widetilde{b}_1(z) = \widetilde{c}^\gamma_1(z)/(\sqrt{\pi}(z+i))$.

Then for all compactly supported $f \in C^1(\mathbb{R})$

$$\sqrt{\pi}(z+i)\widetilde{\Phi}^*_\gamma f(z) =$$

$$= \widetilde{A}_\gamma(z)\widetilde{f}(z) + \widetilde{B}_\gamma(z) \int \left(\widetilde{f}(s) - \widetilde{f}(z)\right) \frac{1}{2i} \left[\frac{1}{s-z} - \frac{1}{s+i}\right] d\mu(s)$$

where $\widetilde{A}_\gamma(z) = \widetilde{c}^\gamma(z)$, $\widetilde{B}_\gamma(z) = \widetilde{c}^\gamma(z) - \omega(z)\widetilde{c}^\gamma_1(z)$.

A Representation Formula in the Sz.-Nagy–Foiaş Transcription

For a measure μ on the real line define $T_+ f$ to be the non-tangential boundary values of the function

$$\widetilde{R} f\mu(z) = \frac{1}{2iP} \int_{\mathbb{R}} f(s) \left[\frac{1}{s-z} - \frac{1}{s+i}\right] d\mu(s), \qquad \operatorname{Im} z > 0,$$

and let $T^1_+ f$ be the non-tangential boundary values of

$$\widetilde{R}^1 f\mu(z) := \frac{1}{2iP} \int_{\mathbb{R}} \frac{f(s)d\mu(s)}{s-z}, \qquad \operatorname{Im} z > 0;$$

the non-tangential boundary values exist a.e. with respect to the Lebesgue measure by classical result about boundary values of the functions in the Hardy spaces H^p.

Theorem 5.3 *Let* μ *be Poisson normalized,* $P = 1$*. The operator* $\widetilde{\Phi}^*_\gamma$ *can be represented in the Sz.-Nagy–Foiaş transcription as*

$$\sqrt{\pi}(1-|\gamma|^2)^{1/2}\widetilde{\Phi}^*_\gamma f = \begin{pmatrix} 0 \\ (\overline{\gamma} - (\overline{\gamma}-1)T_+\mathbf{1})\widetilde{\Delta}_\gamma \end{pmatrix} f + \begin{pmatrix} (1+\overline{\gamma}\widetilde{\theta}_\gamma)/T_+\mathbf{1} \\ (\overline{\gamma}-1)\widetilde{\Delta}_\gamma \end{pmatrix} T^1_+ f$$

$$= \begin{pmatrix} 0 \\ \frac{1-\overline{\gamma}\widetilde{\theta}_0}{|1-\overline{\gamma}\widetilde{\theta}_0|} T_+\mathbf{1} \cdot \widetilde{\Delta}_0 \end{pmatrix} f + \begin{pmatrix} \frac{1-|\gamma|^2}{1-\overline{\gamma}\widetilde{\theta}_0} \cdot \frac{1}{T_+\mathbf{1}} \\ (\overline{\gamma}-1)\frac{(1-|\gamma|^2)^{1/2}}{|1-\overline{\gamma}\widetilde{\theta}_0|}\widetilde{\Delta}_0 \end{pmatrix} T^1_+ f$$

for $f \in L^2(\mu)$.

The same recipe as above gives the representation in the de Branges–Rovnyak transcription, and a formula for the Clark operator Φ_γ.

References

1. E. Abakumov, C. Liaw, A. Poltoratskiĭ, Cyclic vectors for rank one perturbations and Anderson-type Hamiltonians. J. Lond. Math. Soc. **88**(2), 523–537 (2013)
2. M. Aizenman, S. Molchanov, Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**(2), 245–278 (1993)
3. A.B. Aleksandrov, Multiplicity of boundary values of inner functions. Izv. Akad. Nauk Armyan. SSR Ser. Math. **22**(5), 490–503, 515 (1987)
4. A.B. Aleksandrov, Inner functions and related spaces of pseudocontinuable functions. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI), **170** (1989) Issled. Linein Oper. Teorii Funktsii. 17, 7–33, 321; translation in J. Soviet Math. **63**(2), 115–159 (1993)
5. A.B. Aleksandrov, On the existence of angular boundary values for pseudocontinuable functions. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **222** (1995), Issled. po Linein Oper. i Teor. Funktsii. **23**, 5–17, 307; translation in J. Math. Soc. New York **87**(5), 3781–3787 (1997)
6. A.B. Aleksandrov, Isometric embeddings of coinvariant subspaces of the shift operator. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **232** (1996), Issled. po Linein Oper. i Teor. Funktsii. **24**, 5–15, 213; translation in J. Math. Sci. (New York) **92**(1), 3543–3549 (1998)
7. A.B. Aleksandrov, On the maximum principle for pseudocontinuable functions. (Russian) Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **217** (1994), Issled. po Linein. Oper. i Teor. Funktsii. **22**, 16–25, 218; translation in J. Math. Sci. (New York) **85**(2), 1767–1772 (1997)
8. P.W. Anderson, Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492–1505 (1958)
9. M. Anshelevich, Bochner-Pearson-type characterization of the free Meixner class. Adv. in Appl. Math. **46**, 25–45 (2011). (Special issue in honor of Dennis Stanton)
10. M. Anshelevich, Two-state free Brownian motions. J. Funct. Anal. **260**, 541–565 (2011)
11. J.A. Cima, A.L. Matheson, W.T. Ross, in *The Cauchy Transform*. Mathematical Surveys and Monographs, vol. 125 (American Mathematical Society, Providence, 2006)
12. D.N. Clark, One dimensional perturbations of restricted shifts. J. Anal. Math. **25**, 169–191 (1972)
13. L. de Branges, in *Hilbert Spaces of Entire Functions* (Prentice-Hall, Inc., Englewood Cliffs, 1968)
14. L. de Branges, J. Rovnyak, Canonical models in quantum scattering theory, in *Perturbation Theory and its Applications in Quantum Mechanics*. Proceeding of Advance Seminar Mathematics Research Center, U.S. Army, Theoretical Chemistry Institute at the University of Wisconsin, Madison, 1965 (Wiley, New York, 1966), pp. 295–392
15. R. Del Rio, S. Fuentes, A. Poltoratskii, Coexistence of spectra in rank-one perturbation problems. Bol. Soc. Mat. Mexicana (3) **8**(1), 49–61 (2002)
16. R. Del Rio, S. Fuentes, A. Poltoratskii, Families of spectral measures with mixed types, Operator methods in ordinary and partial differential equations (Stockholm, 2000). Oper. Theory Adv. Appl. **132**, 131–140 (2002). Birkhäuser Basel
17. R. Del Rio, N. Makarov, B. Simon, Operators with singular continuous spectrum. II. Rank one operators. Commun. Math. Phys. **165**(1), 59–67 (1994)
18. R.G. Douglas, C. Liaw, A geometric approach to finite rank unitary perturbations. Indiana Univ. Math. J. **62**(1), 333–354 (2013)
19. J. Fröhlich, T. Spencer, Absence of diffusion in the tight binding model for large disorder of low energy. Commun. Math. Phys. **88**, 151–184 (1983)
20. F. Germinet, A. Klein, J.H. Schenker, Dynamical delocalization in random Landau Hamiltonians. Ann. Math. (2) **166**(1), 215–244 (2007)

21. F. Ghribi, P.D. Hislop, F. Klopp, Localization for Schrödinger operators with random vector potentials. Contemporary Mathematics, **447**, 123–138 (2007)
22. A.Ya. Gordon, Instability of dense point spectrum under finite rank perturbations. Commun. Math. Phys. **187**(3), 583–595 (1997)
23. E. Hamza, A. Joye, G. Stolz, Localization for random unitary operators. Lett. Math. Phys. **75**, 255–272 (2006)
24. E. Hamza, A. Joye, G. Stolz, Dynamical localization for unitary Anderson models. Math. Phys. Anal. Geom. **12**, 381–444 (2009)
25. A. Joye, Dynamical localization for d-dimensional random quantum walks. Quantum Inf. Process **11**(5), 1251–1269 (2012). doi:10.1007/s11128-012-0406-7
26. M.T. Jury, Completely positive maps induced by composition operators (2009, preprint) http://people.clas.ufl.edu/mjury/files/cpcomp.pdf
27. W. King, R.C. Kirby, C. Liaw, Delocalization for 3-D discrete random Schroedinger operator at weak disorder. J. Phys. A: Math. Theor. **47**, 305202 (2014). doi:10.1088/1751-8113/47/30/305202
28. W. Kirsch, An invitation to Random Schrödinger operators. arXiv:0709.3707
29. P. Kurasov, Singular and supersingular perturbations: Hilbert space methods, in *Spectral Theory of Schrödinger Operators* (Universidad Nacional AutÓnoma de México, México, 2004)
30. C. Liaw, Approach to the extended states conjecture. J. Stat. Phys. **153**(6), 1022–1038 (2013)
31. C. Liaw, S. Treil, Clark model in general situation. Journal d'Analyse Mathématiques **130**, 287–328 (2016). doi:10.1007/s11854-016-0038-4
32. C. Liaw, S. Treil, Rank one perturbations and singular integral operators. J. Funct. Anal. **257**(6), 1947–1975 (2009)
33. C. Liaw, S. Treil, Regularizations of general singular integral operators. Rev. Mat. Iberoam. **29**(1), 53–74 (2013)
34. N. Nikolski, S. Treil, Linear resolvent growth of rank one perturbation of a unitary operator does not imply its similarity to a normal operator. J. Anal. Math. **87**, 415–431 (2002). Dedicated to the memory of Thomas H. Wolff
35. N. Nikolski, V. Vasyunin, in *Elements of Spectral Theory in Terms of the Free Function Model. I. Basic Constructions*. Holomorphic spaces (Berkeley, 1995), Mathematical Sciences Research Institute Publication, vol. 33 (Cambridge University Press, Cambridge, 1998), pp. 211–302
36. N. Nikolskiĭ, V. Vasyunin, The Gohberg anniversary collection, *A Unified Approach to Function Models, and the Transcription Problem*, vol. II (Calgary, 1988). Operator Theory: Advances and Applications, vol. 41 (Birkhäuser, Basel, 1989), pp. 405–434
37. A. Poltoratskiĭ, D. Sarason, Aleksandrov-Clark measures, in *Recent Advances in Operator-Related Function Theory*. Contemporary Mathematics, vol. 393 (American Mathematical Society, Providence, 2006), pp. 1–14
38. A.G. Poltoratskiĭ, Boundary behavior of pseudocontinuable functions. Algebra i Analiz **5**(2), 189–210 (1993), engl. translation in St. Petersburg Math. J. 5(2),389–406 (1994)
39. D. Sarason, in *Sub-Hardy Hilbert Spaces in the Unit Disk*. University of Arkansas Lecture Notes in the Mathematical Sciences, vol. 10 (Wiley, New York, 1994), A Wiley-Interscience Publication
40. B. Simon, Spectral analysis of rank one perturbations and applications, in *Mathematical Quantum Theory. II. Schrödinger Operators*, (Vancouver, 1993), in CRM Proceedings & Lecture Notes, vol. 8 (American Mathematical Society, Providence, 1995), pp. 109–149
41. B. Simon, T. Wolff, Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math. **39**(1), 75–90 (1986)
42. M. Stoiciu, in *Poisson Statistics for Eigenvalues: From Random Schrdinger Operators to Random CMV Matrices*. CRM Proceedings & Lecture Notes, vol. 42 (American Mathematical Society, Providence, 2007), pp. 465–475

43. B. Sz.-Nagy, C. Foiaş, H. Bercovici, L. Kérchy, *Harmonic analysis of operators on Hilbert space*, Universitext, 2nd ed., (Springer, New York, 2010). Original edition: B. Sz.-Nagy, C. Foiaş, *Analyse harmonique des opérateurs de l'espace de Hilbert* (Masson et Cie, Paris, 1967). Translated from the French and revised, North-Holland Publishing Co., Amsterdam, 1970.
44. C. Sundberg, Private communication and presentation at the conference "Waves and Spectra" in January 2011
45. H. Weyl, Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen. Math. Ann. **68**(2), 220–269 (1910)
46. J. Bellissard, P. Hislop, A. Klein, G. Stolz, *Random Schrödinger operators: Universal Localization, Correlations, and Interactions*, Conference report (for the conference held in April 2009 at the Banff International Research Station)

Part III
Research Articles

On Two Weight Estimates for Dyadic Operators

Oleksandra Beznosova, Daewon Chung, Jean Carlo Moraes, and María Cristina Pereyra

In memory of our good friend and mentor Cora Sadosky

Abstract We provide a quantitative two weight estimate for the dyadic paraproduct π_b under certain conditions on a pair of weights (u, v) and b in $Carl_{u,v}$, a new class of functions that we show coincides with BMO when $u = v \in A_2^d$. We obtain quantitative two weight estimates for the dyadic square function and the martingale transforms under the assumption that the maximal function is bounded from $L^2(u)$ into $L^2(v)$ and $v \in RH_1^d$. Finally we obtain a quantitative two weight estimate from $L^2(u)$ into $L^2(v)$ for the dyadic square function under the assumption that the pair

The first author was supported by the University of Alabama RGC grant.

The second author was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(2015R1C1A1A02037331).

The third author was supported by the AVG program, 459895/2013-3, funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq.

O. Beznosova (✉)
Department of Mathematics, University of Alabama, 345 Gordon Palmer Hall, Tuscaloosa, AL 35487, USA
e-mail: ovbeznosova@ua.edu

D. Chung
Faculty of Basic Sciences, Mathematics, Keimyung University, 1095 Dalgubeol-daero, Daegu 42601, Korea
e-mail: dwchung@kmu.ac.kr

J.C. Moraes
Instituto de Matemática, Universidade Federal do Rio Grande do Sul, Av. Bento Goncalves, 9500 – Prédio 43-111, Agronomia, Caixa Postal 15080, 91509-900 Porto Alegre, RS, Brazil
e-mail: jean.moraes@ufrgs.br

M.C. Pereyra
Department of Mathematics and Statistics, University of New Mexico, 311 Terrace NE MSC01 1115, Albuquerque, NM 87131, USA
e-mail: crisp@math.unm.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_5

(u, v) is in joint $\mathcal{A}_2^d$ and $u^{-1} \in RH_1^d$, this is sharp in the sense that when $u = v$ the conditions reduce to $u \in A_2^d$ and the estimate is the known linear mixed estimate.

2010 *Mathematics Subject Classification.* Primary 42B20, 42B25; Secondary 47B38

Introduction

We study quantitative two weight inequalities for some dyadic operators. More precisely, we study conditions on pairs of locally integrable a.e. positive functions (u, v) so that a linear or sublinear dyadic operator T is bounded from $L^2(u)$ into $L^2(v)$, that is there exists a constant $C_{T,u,v} > 0$ such that for all $f \in L^2(u)$,

$$\|Tf\|_{L^2(v)} \leq C_{T,u,v} \|f\|_{L^2(u)},$$

with estimates on $C_{T,u,v}$ involving the constants that appear in the conditions imposed on the weights and/or the operator.

There are two current schools of thought regarding the two weight problem. First, given one operator find necessary and sufficient conditions on the weights to ensure boundedness of the operator on the appropriate spaces. Second, given a family of operators find necessary and sufficient conditions on the weights to ensure boundedness of the family of operators. In the first case, the conditions are usually "testing conditions" obtained from checking boundedness of the given operator on a collection of test functions. In the second case, the conditions are more "geometric", meaning to only involve the weights and not the operators, such as Carleson conditions or bilinear embedding conditions, Muckenhoupt A_2 type conditions or bumped conditions. Operators of interest are the maximal function [43, 54, 55, 59], fractional and Poisson integrals [11, 56], the Hilbert transform [9, 10, 30, 32, 33, 36, 49, 50] and general Calderón–Zygmund singular integral operators and their commutators [12–14, 47], the square functions [7, 26, 34, 35], paraproducts and their dyadic counterparts [3, 20, 21, 28, 44]. Necessary and sufficient conditions are only known for the maximal function, fractional and Poisson integrals [55], square functions [34] and the Hilbert transform [33, 36], and among the dyadic operators for the martingale transform, the dyadic square functions, positive and well localized dyadic operators [17, 18, 25, 37, 38, 49, 51, 57, 58, 60–62]. If the weights u and v are assumed to be in A_2^d, then necessary and sufficient conditions for boundedness of dyadic paraproducts and commutators in terms of Bloom's *BMO* are known [20, 21]. The assumption that a weight is in dyadic A_p^d is a strong assumption, it implies, for example, that the weight is dyadic doubling. On the other hand if the paraproduct is adapted to the weights u and v, then necessary and sufficient conditions for its boundedness from $L^p(u)$ into $L^p(v)$ are known [39] even in the non-homogeneous case, interestingly enough the conditions are different depending on whether $1 \leq p < 2$ or $p > 2$.

In this paper we obtain a quantitative two weight estimate for π_b, the dyadic paraproduct associated to b, where $b \in Carl_{u,v}$ a new class of functions that we show coincides with BMO^d when $u = v \in A_2^d$. The sufficient conditions on the pair of weights (u, v) required in our theorem are half of the conditions required for the boundedness of the martingale transform, namely (i) $(u, v) \in \mathcal{A}_2^d$ (joint dyadic A_2 condition) and (ii) a Carleson condition on the weights, or equivalently, the conditions required for the boundedness of the dyadic square function from $L^2(v^{-1})$ into $L^2(u^{-1})$.

In what follows $\mathcal{D}$ denotes the dyadic intervals, $\mathcal{D}(J)$ denotes the dyadic subintervals of an interval J, $|J|$ denotes the length of the interval J, $\{h_I\}_{I\in\mathcal{D}}$ denotes the Haar functions, $m_I f := \frac{1}{|I|}\int_I f$ denotes the integral average of f over the interval I with respect to Lebesgue measure, and $\langle f, g\rangle := \int f\overline{g}$ denotes the inner product on $L^2(\mathbb{R})$. We prove the following theorem.

Theorem 1.1 *Let (u, v) be a pair of measurable functions on $\mathbb{R}$ such that v and u^{-1}, the reciprocal of u, are weights on $\mathbb{R}$, and such that*

(i) $(u, v) \in \mathcal{A}_2^d$, *that is* $[u, v]_{\mathcal{A}_2^d} := \sup_{I\in\mathcal{D}} m_I(u^{-1})\, m_I v < \infty$.
(ii) *there is a constant $\mathcal{D}_{u,v} > 0$ such that*

$$\sum_{I\in\mathcal{D}(J)} |\Delta_I v|^2 |I|\, m_I(u^{-1}) \le \mathcal{D}_{u,v}\, v(J) \qquad \text{for all } J \in \mathcal{D},$$

where $\Delta_I v := m_{I_+} v - m_{I_-} v$, and $I_\pm$ are the right and left children of I.

Assume that $b \in Carl_{u,v}$, that is $b \in L^1_{loc}(\mathbb{R})$ and there is a constant $\mathcal{B}_{u,v} > 0$ such that

$$\sum_{I\in\mathcal{D}(J)} \frac{|\langle b, h_I\rangle|^2}{m_I v} \le \mathcal{B}_{u,v}\, u^{-1}(J) \qquad \text{for all } J \in \mathcal{D}.$$

Then π_b, the dyadic paraproduct associated to b, is bounded from $L^2(u)$ into $L^2(v)$. Moreover, there exists a constant $C > 0$ such that for all $f \in L^2(u)$

$$\|\pi_b f\|_{L^2(v)} \le C\sqrt{[u, v]_{\mathcal{A}_2^d}\mathcal{B}_{u,v}}\left(\sqrt{[u, v]_{\mathcal{A}_2^d}} + \sqrt{\mathcal{D}_{u,v}}\right)\|f\|_{L^2(u)},$$

where $\pi_b f := \sum_{I\in\mathcal{D}} m_I f\, \langle b, h_I\rangle\, h_I$.

When $u = v = w$ the conditions in Theorem 1.1 reduce to $w \in A_2^d$ and $b \in BMO^d$, but we do not recover the first author's linear bound for the dyadic paraproduct [1], we are off by a factor of $[w]_{RH_1^d}^{1/2}$. In [44, 45] similar methods yield the linear bound in the one weight case, but there is a step in that argument that can not be taken in the two weight setting. More precisely, in the one weight case, $u = v = w$, we have $ww^{-1} = 1$ a.e. and $1 \le m_I w\, m_I(w^{-1})$; in the two weight case we can no longer bound vu^{-1} nor can we bound $m_I v\, m_I(u^{-1})$ positively away from zero.

We compare the known two weight results for the martingale transform, the dyadic square function, and the dyadic maximal function. Assuming the maximal operator is bounded from $L^2(u)$ into $L^2(v)$, and under the additional condition that v is in the RH_1^d class, we conclude the other operators are bounded with quantitative estimates involving the operator norm of the maximal function and the RH_1^d constant. Notice that the boundedness of the maximal function implies that the weights (u, v) obey the joint $\mathcal{A}_2^d$ condition, but this is not sufficient for boundedness neither of the martingale transform nor the dyadic square function. Finally we obtain quantitative two weight estimates for the dyadic square function when $(u, v) \in \mathcal{A}_2^d$ and u^{-1} is in RH_1^d. This extends work of the first author [2] where similar quantitative two weight bounds were obtained under the stronger assumption that $u^{-1} \in A_q^d$ for some $q > 1$ (in other words, $u^{-1} \in A_\infty^d$).

Theorem 1.2 *Let (u, v) be a pair of measurable functions such that $(u, v) \in \mathcal{A}_2^d$ and $u^{-1} \in RH_1^d$. Then there is a constant such that*

$$\|S^d\|_{L^2(u)\to L^2(v)} \leq C[u, v]_{\mathcal{A}_2^d}^{1/2}\left(1 + [u^{-1}]_{RH_1^d}^{1/2}\right).$$

When the two weights equal w the conditions in Theorem 1.2 reduce to $w \in A_2^d$ and we improve the sharp linear estimates of Hukovic et al. [22] to a mixed linear estimate. Compare to one weight mixed type estimates of Lerner [41], and two weight strong and weak estimates in [7, 26, 34] where similar estimates are obtained for the g-function and Wilson's intrinsic square function [63]. In the aformentioned papers, both weights are assumed to be in A_∞.

The one weight problem, corresponding to $u = v = w$ is well understood. In 1960, Helson and Szegö [19] presented the first necessary and sufficient conditions on w for the boundedness of the Hilbert transform on $L^2(w)$ in the context of prediction theory. They used methods involving analytic functions and operator theory. The two weight characterization for the Hilbert transform in this direction was completely solved by Cotlar and Sadosky in [9] and [10]. The class of A_p weights was introduced in 1972 by Muckenhoupt [46], these are the weights w for which the Hardy-Littlewood maximal function maps $L^p(w)$ into itself. We say the positive almost everywhere and locally integrable function w satisfies the A_p condition if and only if

$$[w]_{A_p} := \sup_I \left(\frac{1}{|I|}\int_I w(x)dx\right)\left(\frac{1}{|I|}\int_I w^{-\frac{1}{p-1}}(x)dx\right)^{p-1} < \infty,$$

where $[w]_{A_p}$ denotes the A_p characteristic (often called A_p constant or norm) of the weight. In 1973, Hunt, Muckenhoupt, and Wheeden [23] showed that the Hilbert transform is bounded on $L^p(w)$ if and only if $w \in A_p$. Also, in 1973, Coiffman and Fefferman [8] extended this result to the classical Calderón-Zygmund operators. When $u = v = w$ the joint $\mathcal{A}_2$ condition coincides with A_2. The joint $\mathcal{A}_2$ condition is necessary and sufficient for the two weight weak (1,1) boundedness of the maximal function but is not enough for the strong boundedness [55]. In 1982 Sawyer found necessary and sufficient conditions on pairs of weights for the boundedness of the maximal function, namely joint $\mathcal{A}_2$ and the testing conditions [55]. In the 1990s the interest shifted toward the study, in the one weight case, of the sharp dependence of A_p characteristic for a general Calderón-Zygmund operator on weighted Lebesgue spaces $L^p(w)$. In 2012 Hytönen proved the A_2-conjecture (now theorem), see [24, 31]: Let T be a Calderón-Zgmund operator and w be an A_2 weight then

$$\|Tf\|_{L^2(w)} \le C\,[w]_{A_2}\|f\|_{L^2(w)}\,,$$

where the constant C depends only on the dimension d, the growth and smoothness of the kernel of T, and its norm in the non-weighted L^2. From sharp extrapolation [15] one deduces that for $1 < p < \infty$, and $w \in A_p$,

$$\|Tf\|_{L^p(w)} \le C_{d,T,p}\,[w]_{A_p}^{\max\{1,1/(p-1)\}}\|f\|_{L^p(w)}.$$

After these groundbreaking results, improvements were found in the form of mixed type estimates such as the following $L^2(w)$ estimate

$$\|Tf\|_{L^2(w)} \le C\,[w]_{A_2}^{1/2}([w]_{A^d_\infty}^{1/2} + [w^{-1}]_{A^d_\infty}^{1/2})\|f\|_{L^2(w)},$$

where $A^d_\infty = \cup_{p>1}A^d_p$, and $[w]_{A^d_\infty}$ is the Hruščev constant or is replaced by the smaller $[w]_{RH^d_1}$ as we do in this paper, see [25, 27] and [42, 54] for other variations. Currently a lot of effort has been put into finding two weight analogues of these estimates as described at the beginning of this introduction. In this paper we present two weight quantitative and mixed type estimates for the dyadic paraproduct, martingale transform, and the dyadic square function.

In this paper we work in $\mathbb{R}$ but the results should hold in $\mathbb{R}^d$ and in spaces of homogeneous type.

Preliminary definitions and results are collected in section "Definitions and Frequently Used Theorems", including joint $\mathcal{A}^d_2$, regular and weighted Haar functions, w-Carleson sequences, the class $Carl_{u,v}$, the class RH^d_1 and its quantitative relation to A^d_∞, weighted Carleson's and Buckley's Lemmas. The main dyadic operators are introduced in section "Dyadic Operators and Known Two Weight Results": dyadic maximal function, dyadic square function, martingale transform and the dyadic paraproduct, we record the known two weight results for these operators. In section "The Dyadic Paraproduct, Bump Conditions, and BMO vs $Carl_{u,v}$" we prove our quantitative two weight result for the dyadic paraproduct, we also show that when $u = v \in A^d_2$ then $Carl_{u,u} = BMO^d$. We compare our conditions to bumped conditions and argue that neither result implies the other, we also compare $Carl_{u,v}$ to the Bloom BMO and related conditions. In section "The Maximal and the Square Functions" we obtain some quantitative two weight estimates for the dyadic square function and the martingale transforms under the assumptions that the maximal function is bounded and the additional assumption v is a weight in RH_1^d. In section "The Sharp Quantitative Estimate for the Dyadic Square Function" we obtain a sharp two weight estimate for the dyadic square function under the assumptions that $(u, v) \in \mathcal{A}_2$ and $u^{-1} \in RH_1^d$.

The authors would like to thank the referee for thoughtful comments, and for enticing us to explore in more depth the Bloom BMO condition and compare it to $Carl_{u,v}$. The authors would also like to thank Jethro van Ekeren, a friend of the third author and a native English speaker, who proofread the article.

Definitions and Frequently Used Theorems

Throughout the proofs a constant C will be a numerical constant that may change from line to line. The symbol $A_n \lesssim B_n$ means there is a constant $c > 0$ independent of n such that $A_n \leq cB_n$, and $A_n \approx B_n$ means that $A_n \lesssim B_n$ and $B_n \lesssim A_n$. Given a measurable set E in $\mathbb{R}$, $|E|$ will denote its Lebesgue measure. We say that a function $v : \mathbb{R} \to \mathbb{R}$ is a weight if v is an almost everywhere positive locally integrable function. For a given weight v, the v-measure of a measurable set E, denoted by $v(E)$, is $v(E) = \int_E v(x)dx$. We say that a weight v is a regular weight if $v((-\infty, 0)) = v((0, \infty)) = \infty$. Let us denote $\mathcal{D}$ the collection of all dyadic intervals, and let us denote $\mathcal{D}(J)$ the collection of all dyadic subintervals of J.

We say that a pair of weights (u, v) satisfies the joint $\mathcal{A}_2^d$ condition if and only if both v and u^{-1}, the reciprocal of u, are weights, and

$$[u, v]_{\mathcal{A}_2^d} := \sup_{I \in \mathcal{D}} m_I(u^{-1})\, m_I v < \infty, \tag{1}$$

where $m_I v$ stands for the integral average of a weight v over the interval I. Note that $(u, v) \in \mathcal{A}_2^d$ is equivalent to $(v^{-1}, u^{-1}) \in \mathcal{A}_2^d$ and the corresponding constants are equal. Similarly a pair of weights (u, v) satisfies the joint $\mathcal{A}_p^d$ condition iff

$$[u, v]_{\mathcal{A}_p^d} := \sup_{I \in \mathcal{D}} m_I(u^{\frac{-1}{p-1}})^{p-1} m_I v < \infty.$$

Note also that $(v, v) \in \mathcal{A}_p^d$ coincides with the usual one weight definition of $v \in A_p^d$.

Haar Bases

For any interval $I \in \mathcal{D}$, there is a Haar function defined by

$$h_I(x) = \frac{1}{\sqrt{|I|}}\Big(\mathbb{1}_{I_+}(x) - \mathbb{1}_{I_-}(x)\Big),$$

where $\mathbb{1}_I$ denotes the characteristic function of the interval I, and I_+, I_- denote the right and left child of I respectively. For a given weight v and an interval I define the weighted Haar function as

$$h_I^v(x) = \frac{1}{\sqrt{v(I)}}\left(\sqrt{\frac{v(I_-)}{v(I_+)}}\mathbb{1}_{I_+(x)} - \sqrt{\frac{v(I_+)}{v(I_-)}}\mathbb{1}_{I_-(x)}\right).$$

The space $L^2(v)$ is the collection of square integrable complex valued functions with respect to the measure $d\mu = vdx$, it is a Hilbert space with the weighted inner product defined by $\langle f, g\rangle_v = \int f\overline{g}vdx$. It is a well known fact that the Haar systems $\{h_I\}_{I\in\mathcal{D}}$ and $\{h_I^v\}_{I\in\mathcal{D}}$ are orthonormal systems in $L^2(\mathbb{R})$ and $L^2(v)$ respectively. Therefore, for any weight v, by Bessel's inequality we have the following:

$$\sum_{I\in\mathcal{D}} |\langle f, h_I^v\rangle_v|^2 \le \|f\|^2_{L^2(v)}.$$

Furthermore, if v is a regular weight, then every function $f \in L^2(v)$ can be written as

$$f = \sum_{I\in\mathcal{D}} \langle f, h_I^v\rangle_v h_I^v ,$$

where the sum converges a.e. in $L^2(v)$, hence the family $\{h_I^v\}_{I\in\mathcal{D}}$ is a complete orthonormal system. Note that if v is not a regular weight so that $v((-\infty, 0))$, $v((0,\infty))$, or both are finite, then either $\mathbb{1}_{(-\infty,0)}$, $\mathbb{1}_{(0,\infty)}$, or both are in $L^2(v)$ and are orthogonal to h_I^v for every dyadic interval I.

The weighted and unweighted Haar functions are related linearly as follows:

Proposition 2.1 ([49]) *For any weight v and every $I \in \mathcal{D}$, there are numbers α_I^v, β_I^v such that*

$$h_I(x) = \alpha_I^v\, h_I^v(x) + \beta_I^v\, \frac{\mathbb{1}_I(x)}{\sqrt{|I|}}$$

where (i) $|\alpha_I^v| \le \sqrt{m_I v}$, (ii) $|\beta_I^v| \le \frac{|\Delta_I v|}{m_I v}$, *and* $\Delta_I v := m_{I_+} v - m_{I_-} v$.

Dyadic BMO

A locally integrable function b is in the space of dyadic bounded mean oscillation (BMO^d) if and only if there is a constant $C > 0$ such that for all $I \in \mathcal{D}$ one has

$$\int_I |b(x) - m_I b|\, dx \le C|I|.$$

The smallest constant C is the BMO^d-norm of b. The celebrated John-Nirenberg Theorem (see [53]) implies that for each $1 \leq p < \infty$, $b \in BMO^d$ iff

$$\|b\|^p_{BMO^d_p} := \sup_{I\in\mathcal{D}} \frac{1}{|I|}\int_I |b(x) - m_I b|^p dx < \infty.$$

Furthermore $\|b\|_{BMO^d_p}$ is comparable to the BMO-norm of b.

In this paper we will mostly be concerned with $p = 2$ and we will declare

$$\|b\|_{BMO^d} := \|b\|_{BMO^d_2} = \sup_{I\in\mathcal{D}} \left(\frac{1}{|I|}\int_I |b(x) - m_I b|^2 dx\right)^{1/2}.$$

Lemma 2.2 *If $b \in BMO^d$ then*

$$\|b\|^2_{BMO^d} = \sup_{I\in\mathcal{D}} \frac{1}{|I|} \sum_{J\in\mathcal{D}(I)} |\langle b, h_J\rangle|^2.$$

Proof The family $\{h_J\}_{J\in\mathcal{D}(I)}$ is an orthonormal basis of the space $L^2_0(I) := \{f \in L^2(I) : \int_I f = 0\}$. The function $(b - m_I b)\mathbb{1}_I \in L^2_0(I)$, hence by Plancherel

$$\int_I |b(x) - m_I b|^2 dx = \sum_{J\in\mathcal{D}(I)} |\langle b, h_J\rangle|^2.$$

This proves the lemma. □

In other words, $b \in BMO^d$ if and only if there is a constant $C > 0$ such that for all $I \in \mathcal{D}$

$$\sum_{J\in\mathcal{D}(I)} |\langle b, h_J\rangle|^2 \leq C|I|.$$

Carleson Sequences

A positive sequence $\{\lambda_I\}_{I\in\mathcal{D}}$ is a v-Carleson sequence if there is a constant $C > 0$ such that for all dyadic intervals J

$$\sum_{I\in\mathcal{D}(J)} \lambda_I \leq Cv(J). \tag{2}$$

When $v = 1$ almost everywhere we say that the sequence is a Carleson sequence or a dx-Carleson sequence. The infimum among all C's that satisfy the inequality (2) is called the intensity of the v-Carleson sequence $\{\lambda_I\}_{I\in\mathcal{D}}$. For instance, $b \in BMO^d$

if and only if $\{|\langle b, h_I\rangle|^2\}_{I\in\mathcal{D}}$ is a Carleson sequence with intensity $\|b\|^2_{BMO^d}$. The following lemma gives a relationship between unweighted and weighted Carleson sequences.

Lemma 2.3 (Little Lemma, [1]) *Let v be a weight, such that v^{-1} is also a weight, and let $\{\alpha_I\}_{I\in\mathcal{D}}$ be a Carleson sequence with intensity B then $\{\alpha_I/m_I(v^{-1})\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensity at most $4B$, that is for all $J\in\mathcal{D}$,*

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)}\frac{\alpha_I}{m_I(v^{-1})}\le 4Bm_Jv.$$

We also need to define a class of objects that will take the place of the BMO^d class in the two weighted case, we will call this class the two weight Carleson class.

Definition 2.4 Given a pair of functions (u, v) such that v and u^{-1} are weights, we say that a locally integrable function b belongs to the two weight Carleson class, $Carl_{u,v}$, if $\big\{|b_I|^2/m_Iv\}_{I\in\mathcal{D}}$ is a u^{-1}- Carleson sequence where $b_I=\langle b, h_I\rangle$.

Note that if $u=v$, then we have that $b\in Carl_{v,v}$ iff $\{|b_I|^2/m_Iv\}_{I\in\mathcal{D}}$ is a v^{-1}-Carleson sequence. The later statement is true if $\{|b_I|^2\}_{I\in\mathcal{D}}$ is a Carleson sequence (by Lemma 2.3), which in turn is equivalent to saying that $b\in BMO^d$. Therefore for any weight v , such that v^{-1} is also a weight, we have that

$$BMO^d\subset Carl_{v,v}.$$

Moreover, if $\mathcal{B}_{v,v}$ is the intensity of the v^{-1}-Carleson sequence $\{|b_I|^2/m_Iv\}_{I\in\mathcal{D}}$ then $\mathcal{B}_{v,v}\le 4\|b\|^2_{BMO^d}$. In section "The Dyadic Paraproduct, Bump Conditions, and *BMO* vs $Carl_{u,v}$" we will show that if $v\in A_2^d$ then $BMO^d=Carl_{v,v}\cap L^2_{loc}(\mathbb{R})$ (see Corollary 4.6).

We now introduce some useful lemmas which will be used frequently throughout this paper. You can find proofs in [45]. The following lemma was stated first in [49].

Lemma 2.5 (Weighted Carleson Lemma) *Let v be a weight, then $\{\alpha_I\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensity $\mathcal{B}$ if and only if for all non-negative v-measurable functions F on the line,*

$$\sum_{I\in\mathcal{D}}(\inf_{x\in I}F(x))\alpha_I\le\mathcal{B}\int_{\mathbb{R}}F(x)\,v(x)\,dx. \tag{3}$$

In relation to Carleson sequences we consider another class of weights which is called the Reverse Hölder class with index 1 and is defined as follows.

Definition 2.6 A weight v belongs to the dyadic Reverse Hölder class RH_1^d whenever its characteristic $[v]_{RH_1^d}$ is finite, where

$$[v]_{RH_1^d}:=\sup_{I\in\mathcal{D}}m_I\left(\frac{v}{m_Iv}\log\frac{v}{m_Iv}\right)<\infty.$$

It is well known that $v \in A_\infty$ if and only if $v \in RH_1$. In the dyadic case, $v \in RH_1^d$ does not imply that v is dyadic doubling, however $v \in A_\infty^d$ does. See [53] for more details. Recently, the first author and Reznikov obtained, in [4], the sharp comparability of the A_∞^d and RH_1^d characteristics. The RH_1^d characteristic is also known as the Fujii-Wilson A_∞^d characteristic, see for example [54, Equation (2.5)] and references therein.

Theorem 2.7 ([4]) *If a weight v belongs to the A_∞^d class, then $v \in RH_1^d$. Moreover,*

$$[v]_{RH_1^d} \le \log(16)[v]_{A_\infty^d}.$$

The constant $\log(16)$ *is the best possible.*

We would also like to note here that results of Iwaniek and Verde [29] show that $[w]_{RH_1^d} \approx \sup_{I\in D} \frac{\|w\|_{L\log L,I}}{\|w\|_{L,I}}$, where $\|\cdot\|_{\Phi(L),I}$ stands for the $\Phi(L)$-Luxemburg norm (for more details see [4]). In the same paper you can find the following characterization of the $L\log L$-norm (Part (a)) and a sharp version of Buckley's theorem (Part (b)).

Theorem 2.8 (a) [4, Theorem II.6(2)] *There exist real positive constants c and C, independent of the weight v, such that for every weight v and every interval J we have*

$$c\, m_J\left(v\log\left(\frac{v}{m_J v}\right)\right) \leqslant \frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} \frac{|\Delta_I v|^2}{m_I v}|I| \leqslant C\, m_J\left(v\log\left(\frac{v}{m_J v}\right)\right) \quad (4)$$

and as a consequence $\|v\|_{L\log L,J} \approx \frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} \frac{|\Delta_I v|^2}{m_I v}|I|$.

(b) *Let v be a weight such that $v \in RH_1^d$. Then $\{|\Delta_I v|^2|I|/m_I v\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensity comparable to $[v]_{RH_1^d}$. That is, there is a constant $C > 0$ such that for any $J \in \mathcal{D}$,*

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} \frac{|\Delta_I v|^2}{m_I v}|I| \le C[v]_{RH_1^d} m_J v.$$

Dyadic Operators and Known Two Weight Results

We now introduce several dyadic operators which will be considered in this paper, and record known two weight results for them.

Dyadic Weighted Maximal Function

First we recall the dyadic weighted maximal function.

Definition 3.1 We define the dyadic weighted maximal function M_v^d as follows

$$M_v^d f(x) := \sup_{\substack{I \ni x \\ I \in \mathcal{D}}} \frac{1}{v(I)} \int_I |f(y)|\, v(y) dy$$

The weighted maximal function M_v is defined analogously by taking the supremum over all intervals not just dyadic intervals. A very important fact about the weighted maximal function is that the $L^p(v)$ norm of M_v^d only depends on $p' = p/(p-1)$ not on the weight v.

Theorem 3.2 *Let v be a locally integrable function such that $v > 0$ a.e. Then for all $1 < p < \infty$, M_v^d is bounded in $L^p(v)$. Moreover, for all $f \in L^p(v)$*

$$\|M_v^d g\|_{L^p(v)} \leq C_p \|f\|_{L^p(v)}.$$

This result follows by the Marcinkiewicz interpolation theorem, with constant $C_p = 2(p')^{\frac{1}{p}}$, using the facts that M_v^d is bounded on $L^\infty(v)$ with constant 1 and it is weak-type $(1, 1)$ also with constant 1. Note that as $p \to 1$, $C_p/p' \to 2$ and $C_2 = 2\sqrt{2}$.

When $v = 1$, M_1 is the maximal function that we will denote M. In [6], Buckley showed that the $L^p(w)$ norm of M behaves like $[w]_{A_p}^{\frac{1}{p-1}}$, in particular the $L^2(w)$ norm of M depends linearly on the A_2 charateristic of the weight. The next theorem is Sawyer's celebrated two weight result for the maximal function M in the case $p = 2$.

Theorem 3.3 ([55]) *The maximal function M is bounded from $L^2(u)$ into $L^2(v)$ if and only if there is a constant $C_{u,v} > 0$ such that*

$$\int_I \big(M(\mathbb{1}_I u^{-1})(x)\big)^2 v(x)\, dx \leq C_{u,v} u^{-1}(I), \qquad \textit{for all intervals} \quad I. \tag{5}$$

A quantitative version of this result was given by Moen, he showed in [43] that the operator norm of M from $L^2(u)$ into $L^2(v)$ is comparable to $2C_{u,v}$. Note that Sawyer's test condition (5) implies $(u, v) \in \mathcal{A}_2$, moreover $[u, v]_{\mathcal{A}_2} \leq C_{u,v}$.

A quantitative two weight result for the maximal function not involving Sawyer's test conditions, instead involving joint $\mathcal{A}_2$ and RH_1 constant of u^{-1}, has been recently found by Hyönen and Përez [27], see also Pérez and Rela [54].

Theorem 3.4 ([54, Corollary 1.4])[1] *Let u and v be weights such that $(u, v) \in \mathcal{A}_2$ and $u^{-1} \in RH_1$ then*

[1]Note that in [54, Corollary 1.4] the statement is not exactly this one. The authors are using a well-known change of variables that we are not using in this paper: their w is our v, their σ is our u^{-1}, their $[w, \sigma]_{A_2^d}$ corresponds to our $[\sigma^{-1}, w]_{A_2^d}$ and hence equals to our $[u, v]_{A_2^d}$. Finally in their case $M(\cdot\sigma)$ acts on $g \in L^2(\sigma)$, in our case M acts on $f \in L^2(u)$, and clearly $g \in L^2(\sigma) = L^2(u^{-1})$ if and only if $f = g\sigma = gu^{-1} \in L^2(u)$ with equal norms.

$$\|M\|_{L^2(u)\to L^2(v)} \le C([u,v]_{\mathcal{A}_2}[u^{-1}]_{RH_1})^{1/2}.$$

This result is valid in certain spaces of homogeneous type, see [54]. In fact they prove a result valid in L^p replacing joint $\mathcal{A}_2$ by joint $\mathcal{A}_p$ and the power $1/2$ by $1/p$. More precisely they show

$$\|M\|_{L^p(u)\to L^p(v)} \le Cp'([u,v]_{\mathcal{A}_p}[u^{-1}]_{RH_1})^{1/p},$$

where $p' = p/(p-1)$ is the dual exponent to p.

Dyadic Square Function

Second, we introduce the dyadic square function.

Definition 3.5 We define the dyadic square function as follows

$$S^d f(x) := \left(\sum_{I\in\mathcal{D}} |m_I f - m_{\hat{I}} f|^2 \mathbb{1}_I(x)\right)^{1/2},$$

where $\hat{I}$ denotes the dyadic parent of I.

In [22], Hukovic, Treil and Volberg showed that the $L^2(v)$ norm of S^d depends linearly on the A_2 characteristic of the weight. Cruz-Uribe, Martell, and Pérez [13] showed that the $L^3(v)$ norm of S^d depends on $[v]_{A_3}^{1/2}$. One concludes that $\|S^d f\|_{L^p(v)} \le C[v]_{A_p}^{\max\{\frac{1}{2},\frac{1}{p-1}\}}\|f\|_{L^p(v)}$ by sharp extrapolation [15], this bound is optimal. Lerner [40] has shown that this holds for Wilson's intrinsic square function [63].

The following two weight characterization was introduced by Wilson, see also [49]

Theorem 3.6 ([62]) *The dyadic square function S^d is bounded from $L^2(u)$ into $L^2(v)$ if and only if*

(i) $(u,v) \in \mathcal{A}_2^d$
(ii) $\{|I|\,|\Delta_I u^{-1}|^2 m_I v\}_{I\in\mathcal{D}}$ *is a u^{-1}-Carleson sequence with intensity $\mathcal{C}_{u,v}$.*

Condition (ii) can be viewed as a localized testing condition on the test functions $u^{-1}\mathbb{1}_J$ for $J \in \mathcal{D}$. Thus, $\mathcal{C}_{u,v} \le \|S^d\|^2_{L^2(u)\to L^2(v)}$.

Recently Lacey and Li [34] showed a continuous quantitative analogue of this theorem and they claim the dyadic version is "a direct analog of their theorem", their estimate would read

$$\|S^d\|_{L^2(u)\to L^2(v)} \lesssim ([u,v]_{\mathcal{A}_2^d} + \mathcal{C}_{u,v})^{1/2}. \tag{6}$$

We will present a proof of this estimate in section "The Sharp Quantitative Estimate for the Dyadic Square Function". We will get quantitative two weight estimates for the dyadic square function involving either the two weight norm of the maximal operator and $[v]_{RH_1^d}^{1/2}$, or $[u,v]_{\mathcal{A}_2^d}^{1/2}$, $[u^{-1}]_{RH_1^d}^{1/2}$, and $[v]_{RH_1^d}^{1/2}$, under appropriate assumptions in each case.

Theorem 2.8(b) implies that if $u^{-1} \in RH_1^d$ and $(u,v) \in \mathcal{A}_2^d$ then condition (ii) in Theorem 3.6 holds with $\mathcal{C}_{u,v} \lesssim [u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d}$. As a corollary of (6) we get that if $u^{-1} \in RH_1^d$ and $(u,v) \in \mathcal{A}_2^d$ then

$$\|S^d\|_{L^2(u)\to L^2(v)} \leq C([u,v]_{\mathcal{A}_2^d} + [u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d})^{1/2}. \tag{7}$$

This improves [2, Theorem 4.1] where the stronger assumption $u^{-1} \in A_q^d$ for some $q > 1$ was made and a similar quantitative two weight estimate was obtained with $[u^{-1}]_{A_q^d}$ replacing $[u^{-1}]_{RH_1^d}$ and the constant C depending on q. Her results are proved in a setting where the underlying Lebesgue measure is replaced by a doubling measure σ on $\mathbb{R}$ (a space of homogeneous type), introducing a dependence on the doubling constant of σ which is tracked in the aformentioned theorem. We will prove (7) without relying on (6) in section "The Sharp Quantitative Estimate for the Dyadic Square Function". A closer look shows that the same argument will allow us to recover (6). When $u = v = w \in A_2^d$ this improves Hukovic's linear bound to a mixed bound:

$$\|S^d\|_{L^2(w)} \leq C([w]_{A_2^d}[w^{-1}]_{RH_1^d})^{1/2}.$$

Martingale Transform

Third, we introduce the martingale transforms.

Definition 3.7 Let r be a function from $\mathcal{D}$ into $\{-1, 1\}$ so that $r(I) = r_I$, then we define the martingale transform T_r associated to r, acting on functions $f \in L^2(\mathbb{R})$, by

$$T_r f(x) := \sum_{I\in\mathcal{D}} r_I \langle f, h_I\rangle h_I(x).$$

In [64], Wittwer showed that the $L^2(w)$ norm of T_r depends linearly on the A_2 characteristic of the weight w. The next theorem is from [49] and it gives necessary and sufficient conditions for the martingale transforms T_r to be uniformly bounded from $L^2(u)$ into $L^2(v)$. Before we state the theorem, let us define the positive operator

$$T_0 f(x) := \sum_{I\in\mathcal{D}} \frac{\alpha_I}{|I|} m_I f \, \mathbb{1}_I(x)\,,$$

where $\alpha_I = \frac{|\Delta_I v|}{m_I v}\frac{|\Delta_I(u^{-1})|}{m_I(u^{-1})}|I|$.

Theorem 3.8 ([49]) *The martingale transforms T_r are uniformly bounded from $L^2(u)$ to $L^2(v)$ if and only if the following four assertions hold simultaneously:*

(i) $(u, v) \in \mathcal{A}_2$
(ii) $\{|I|\,|\Delta_I u^{-1}|^2 m_I v\}_{I\in\mathcal{D}}$ *is a u^{-1}-Carleson sequence.*
(iii) $\{|I|\,|\Delta_I v|^2 m_I(u^{-1})\}_{I\in\mathcal{D}}$ *is a v-Carleson sequence.*
(iv) *The positive operator T_0 is bounded from $L^2(u)$ into $L^2(v)$.*

As a corollary of the previous results in this section we can rewrite Theorem 3.8 as follows,

Corollary 3.9 *The martingale transforms T_r are uniformly bounded from $L^2(u)$ to $L^2(v)$ if and only if the following three assertions hold simultaneously:*

(i) *S^d is bounded from $L^2(u)$ into $L^2(v)$.*
(ii) *S^d is bounded from $L^2(v^{-1})$ into $L^2(u^{-1})$.*
(iii) *The positive operator T_0 is bounded from $L^2(u)$ into $L^2(v)$.*

Dyadic Paraproduct

Finally we recall the definition of the dyadic paraproduct.

Definition 3.10 We formally define the dyadic paraproduct π_b associated to $b \in L^1_{loc}(\mathbb{R})$ as follows for functions f which are at least locally integrable:

$$\pi_b f(x) := \sum_{I\in\mathcal{D}} m_I f\, \langle b, h_I\rangle h_I(x).$$

It is a well know fact that the dyadic paraproduct is bounded not only on $L^p(dx)$ but also on $L^p(v)$ when $b \in BMO^d$ and $v \in A^d_p$. Beznosova proved in [1] that the $L^2(v)$ norm of the dyadic paraproduct depends linearly on both $[v]_{A^d_2}$ and $\|b\|_{BMO^d}$. Sharp extrapolation [15] then shows

$$\|\pi_b f\|_{L^p(w)} \le C\|b\|_{BMO^d}[w]_{A^d_p}^{\max\{1,\frac{1}{p-1}\}}\|f\|_{L^p(w)}.$$

When both weights $u, v \in A^d_p$ then it is known that the boundedness of the dyadic paraproduct $\pi_b : L^p(u) \to L^p(v)$ is equivalent to b being in a weighted $BMO^d(\mu)$ where $\mu = u^{1/p}v^{-1/p}$, that is,

$$\|b\|_{BMO^d(\mu)} := \sup_{I\in\mathcal{D}} \frac{1}{\mu(I)}\int_I |b(x) - m_I b|\, dx < \infty. \tag{8}$$

This space is known as Bloom's *BMO* [5]. In fact there are a number of conditions equivalent to (8) (see [20]) one of them being the boundedness of the adjoint of the dyadic paraproduct $\pi_b^* : L^p(u) \to L^p(v)$. By duality the last result is equivalent to the boundedness of the dyadic paraproduct $\pi_b : L^{p'}(v') \to L^{p'}(u')$, where p, p' are dual exponents, $\frac{1}{p} + \frac{1}{p'} = 1$, and u', v' are dual weights, namely $u' = u^{\frac{-1}{p-1}} = u^{-p'/p}$. Not surprisingly $\mu' = (v')^{1/p'}(u')^{-1/p'} = \mu$, so that $BMO(\mu') = BMO(\mu)$. The assumption that both weights are in A_p^d is very symmetric and forces boundedness of the paraproduct and its adjoint to occur simultaneously. This is the appropriate setting when dealing with two-weight inequalities for commutators which very naturally can be separated into commutators with a paraproduct, its adjoint, and other terms which will all be bounded from $L^2(u)$ into $L^2(v)$ provided $u, v \in A_p^d$ and b is in Bloom's $BMO(\mu)$. Assuming both $u, v \in A_p^d$ allows one to use Littlewood-Paley theory for the dyadic square function S^d, specifically, the $L^p(w)$ norm of $S^d g$ is comparable to the $L^p(w)$ norm of g whenever $w \in A_p^d$. In particular $\|\pi_b f\|_{L^2(v)}^2$ is comparable to $\|S^d(\pi_b f)\|_{L^2(v)}^2 = \sum_{I\in\mathcal{D}} |m_I f|^2 b_I^2 m_I(v)$, and from here boundedness from $L^2(u)$ into $L^2(v)$ of the dyadic paraproduct is reduced to verifying the following estimate

$$\sum_{I\in\mathcal{D}} |m_I f|^2 b_I^2 m_I(v) \le C_{u,v,b}\|f\|_{L^2(u)}^2.$$

This inequality holds by the weighted Carleson lemma (Lemma 2.5) and the boundedness of the maximal function in $L^2(u)$ when $u \in A_2^d$, provided the sequence $\{b_I^2 m_I(v)\}_{I\in\mathcal{D}}$ is a u-Carleson sequence, namely

$$\sum_{I\in\mathcal{D}(J)} b_I^2 m_I(v) \le Cu(J). \tag{9}$$

Another use of the Littlewood-Paley theory ($v \in A_2$) allows us to compare the left-hand-side to $\int_J |b(x) - m_J b|^2 v(x)\,dx$ yielding what turns out is an equivalent condition for the boundedness of the paraproduct from $L^2(u) \to L^2(v)$ when $u, v \in A_2^d$ (see [20])

$$\sup_{J\in\mathcal{D}} \frac{1}{u(J)} \int_J |b(x) - m_I b|^2 v(x)\,dx < \infty. \tag{10}$$

In [21, Theorem 3.1] the authors present an equivalent condition for the boundedness of the paraproduct from $L^2(u) \to L^2(v)$ when only $v \in A_2^d$, namely

$$\mathcal{B}_2(u,v) := \sup_{J\in\mathcal{D}} \frac{1}{u^{-1}(J)} \sum_{I\in\mathcal{D}(J)} b_I^2 (m_I u^{-1})^2 m_I(v) < \infty. \tag{11}$$

Conditions (10) and (11) are testing conditions for the test functions $u^{-1}\mathbb{1}_J$.

It should be noted that the dyadic paraproduct is a well-localized operator (and for trivial reasons) in the sense of Nazarov, Treil and Volberg [51]. Therefore the known necessary and sufficient testing conditions for well-localized operators apply in this case, these testing conditions involve the two weights and the function b (B. Wick, Personal communication. April 2016).

In section "The Dyadic Paraproduct, Bump Conditions, and *BMO* vs $Carl_{u,v}$" we provide sufficient conditions on a pair of weights (u, v) for the two weight boundedness of the dyadic paraproduct operator from $L^2(u)$ into $L^2(v)$ when $b \in Carl_{u,v}$, together with a quantitative estimate. The conditions we consider are less symmetric, we assume a priori that $(u, v) \in \mathcal{A}_2^d$ (which is equivalent to $(v^{-1}, u^{-1}) \in \mathcal{A}_2^d$), and an assymetric weighted Carleson condition, or equivalently we assume the dyadic square S^d function is bounded from $L^2(v^{-1}) \to L^2(u^{-1})$. Under these conditions we show that if $b \in Carl_{u,v}$ then π_b is bounded from $L^2(u)$ into $L^2(v)$. We would have liked to show that $b \in Carl_{u,v}$ is not only a sufficient condition but also a necessary condition for the boundedness of the dyadic paraproduct under the a priori assumptions on the pair of weights, but we have not been able to identify the appropriate testing functions that will yield this result. If we wish to show that both the paraproduct and its adjoint are bounded from $L^2(u)$ into $L^2(v)$ then we need to assume a priori joint $\mathcal{A}_2$ and two mixed Carleson conditions on the weights, and we need to assume $b \in Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$. It will be interesting to compare these conditions, for example can one show that if $u, v \in A_2^d$ then Bloom's *BMO* coincides with $b \in Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$? Can we conclude that when $(u, v) \in \mathcal{A}_2^d$ then $Carl_{u,v}$ is equivalent to $\mathcal{B}_2(u, v) < \infty$? or that when $v \in A_2^d$ then $Carl_{v,v}$ is equivalent to $\mathcal{B}_2(v, v) < \infty$? We record some results comparing these conditions in section "$Carl_{u,v}$ vs Bloom's *BMO*".

The Dyadic Paraproduct, Bump Conditions, and *BMO* vs $Carl_{u,v}$

In this section we will state and prove our main two weight result about the dyadic paraproduct (Theorem 1.1 in the introduction, called Theorem 4.1 in this section). We will also compare our result to known two weight bump conditions, compare the class $Carl_{v,v}$ with BMO^d when $v \in A_2^d$, and compare the class $Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$ with Bloom's *BMO* when both $u, v \in A_2^d$.

Two Weight Estimate for the Dyadic Paraproduct

In this section we obtain quantitative two-weight estimates for the dyadic paraproduct π_b when $b \in Carl_{u,v}$ and (u, v) are two weights with some additional conditions. Note that by definition b is a locally integrable function, thus $b_I = \langle b, h_I\rangle$ is well defined. The next theorem is a literal restatement of Theorem 1.1 which we provide for ease of the reader.

Theorem 4.1 *Let (u, v) be a pair of functions such that v and u^{-1} are weights, $(u, v) \in \mathcal{A}_2^d$, and $\{|\Delta_I v|^2 |I| m_I(u^{-1})\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensity $\mathcal{D}_{u,v}$. Then π_b is bounded from $L^2(u)$ into $L^2(v)$ if $b \in Carl_{u,v}$. Moreover, if $\mathcal{B}_{u,v}$ is the intensity of the u^{-1}-Carleson sequence $\{|b_I|^2/m_I v\}_{I\in\mathcal{D}}$ then there exists $C > 0$ such that for all $f \in L^2(u)$*

$$\|\pi_b f\|_{L^2(v)} \leq C\sqrt{[u,v]_{\mathcal{A}_2^d}\mathcal{B}_{u,v}}\left(\sqrt{[u,v]_{\mathcal{A}_2^d}} + \sqrt{\mathcal{D}_{u,v}}\right)\|f\|_{L^2(u)}.$$

Proof Fix $f \in L^2(u^{-1})$ and $g \in L^2(v)$, then $fu^{-1} \in L^2(u)$, $\|fu^{-1}\|_{L^2(u)} = \|f\|_{L^2(u^{-1})}$, $gv \in L^2(v^{-1})$ and $\|gv\|_{L^2(v^{-1})} = \|g\|_{L^2(v)}$, $\pi_b(fu^{-1})$ is expected to be in $L^2(v)$, then $gv \in L^2(v^{-1})$ is in the right space for the pairing. Thus, by duality, suffices to prove:

$$|\langle \pi_b(fu^{-1}), gv\rangle| \leq C\sqrt{[u,v]_{\mathcal{A}_2^d}\mathcal{B}_{u,v}}\left(\sqrt{[u,v]_{\mathcal{A}_2^d}} + \sqrt{\mathcal{D}_{u,v}}\right)\|f\|_{L^2(u^{-1})}\|g\|_{L^2(v)}. \quad (12)$$

Replace h_I by $\alpha_I h_I^v + \beta_I \frac{\mathbb{1}_I}{\sqrt{|I|}}$ where $\alpha_I = \alpha_I^v$ and $\beta_I = \beta_I^v$ as described in Proposition 2.1, to get

$$|\langle \pi_b(fu^{-1}), gv\rangle| \leq \sum_{I\in\mathcal{D}} |b_I| m_I(|f|u^{-1}) \left|\left\langle gv, \alpha_I h_I^v + \beta_I \frac{\mathbb{1}_I}{\sqrt{|I|}}\right\rangle\right|. \quad (13)$$

Use the triangle inequality to separate the sum in (13) into two summands

$$|\langle \pi_b(fu^{-1}), gv\rangle| \leq \sum_{I\in\mathcal{D}} |b_I||\alpha_I| m_I(|f|u^{-1})|\langle gv, h_I^v\rangle| + \sum_{I\in\mathcal{D}} |b_I|\frac{|\beta_I|}{\sqrt{|I|}} m_I(|f|u^{-1})|\langle gv, \mathbb{1}_I\rangle|.$$

Using the estimates $|\alpha_I| \leq \sqrt{m_I v}$ and $|\beta_I| \leq \frac{|\Delta_I v|}{m_I v}$ in Proposition 2.1, we have that,

$$|\langle \pi_b(fu^{-1}), gv\rangle| \leq \Sigma_1 + \Sigma_2,$$

where

$$\Sigma_1 := \sum_{I\in\mathcal{D}} |b_I| m_I(|f|u^{-1})|\langle gv, h_I^v\rangle|\sqrt{m_I v}$$

$$\Sigma_2 := \sum_{I\in\mathcal{D}} |b_I| m_I(|f|u^{-1})|\langle gv, \mathbb{1}_I\rangle|\frac{|\Delta_I v|}{m_I v}\frac{1}{\sqrt{|I|}}.$$

Estimating Σ_1: We have

$$\Sigma_1 \leq \sum_{I\in\mathcal{D}} \frac{|b_I|}{\sqrt{m_I v}} m_I^{u^{-1}}(|f|)|\langle g, h_I^v\rangle_v| m_I(u^{-1}) m_I v$$

$$\leq [u,v]_{\mathcal{A}_2^d} \sum_{I\in\mathcal{D}} \frac{|b_I|}{\sqrt{m_I v}} \inf_{x\in I} M_{u^{-1}} f(x) |\langle g, h_I^v\rangle_v|$$

$$\leq [u,v]_{\mathcal{A}_2^d} \left(\sum_{I\in\mathcal{D}} \frac{|b_I|^2}{m_I v} \inf_{x\in I} M^2_{u^{-1}} f(x) \right)^{1/2} \left(\sum_{I\in\mathcal{D}} |\langle g, h^v\rangle_v|^2 \right)^{1/2}.$$

Here in the first line we use that $\langle gv, f\rangle = \langle g, f\rangle_v$, in the second line we use that $m_I^{u^{-1}}|f| := \frac{m_I(|f|u^{-1})}{m_I(u^{-1})} \leq M_{u^{-1}} f(x)$ for all $x \in I$, and that $m_I(u^{-1}) m_I v \leq [u,v]_{\mathcal{A}_2^d}$ and in the third line we use the Cauchy-Schwarz inequality.

Using the fact that $\{h_I^v\}_{I\in\mathcal{D}}$ is an orthonormal system in $L^2(v)$ and the Weighted Carleson Lemma 2.5, with $F(x) = M^2_{u^{-1}} f(x)$, and $\alpha_I = |b_I|^2/m_I v$, which is a u^{-1}-Carleson sequence with intensity $\mathcal{B}_{u,v}$, by assumption, we get

$$\Sigma_1 \leq [u,v]_{\mathcal{A}_2^d} \sqrt{\mathcal{B}_{u,v}} \left(\int_{\mathbb{R}} M^2_{u^{-1}} f(x) u^{-1}(x) dx \right)^{1/2} \|g\|_{L^2(v)}$$

$$\leq 2\sqrt{2} [u,v]_{\mathcal{A}_2^d} \sqrt{\mathcal{B}_{u,v}} \|f\|_{L^2(u^{-1})} \|g\|_{L^2(v)}. \tag{14}$$

In the second inequality we used Theorem 3.2.

Estimating Σ_2: Using similar arguments as the ones used for Σ_1 , we conclude that,

$$\Sigma_2 \leq \sum_{I\in\mathcal{D}} |b_I| m_I^{u^{-1}}(|f|) m_I^v(|g|) \frac{|\Delta_I v|}{m_I v} \sqrt{|I|} m_I(u^{-1}) m_I v$$

$$= \sum_{I\in\mathcal{D}} \frac{|b_I|}{\sqrt{m_I v}} m_I^{u^{-1}}(|f|) m_I^v(|g|) |\Delta_I v| \sqrt{|I|} m_I(u^{-1}) \sqrt{m_I v}$$

$$\leq [u,v]^{1/2}_{\mathcal{A}_2^d} \sum_{I\in\mathcal{D}} \frac{|b_I|}{\sqrt{m_I v}} |\Delta_I v| \sqrt{|I|} \sqrt{m_I u^{-1}} \inf_{x\in I} M_{u^{-1}} f(x) \inf_{x\in I} M_v g(x)$$

$$\leq [u,v]^{1/2}_{\mathcal{A}_2^d} \left(\sum_{I\in\mathcal{D}} \frac{|b_I|^2}{m_I v} \inf_{x\in I} M^2_{u^{-1}} f(x) \right)^{1/2} \left(\sum_{I\in\mathcal{D}} |\Delta_I v|^2 m_I(u^{-1}) |I| \inf_{x\in I} M_v^2 g(x) \right)^{1/2}.$$

By hypothesis $\{|b_I|^2/m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence and $\{|\Delta_I v|\ |I| m_I(u^{-1})\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensities $\mathcal{B}_{u,v}$ and $\mathcal{D}_{u,v}$ respectively. By Lemma (2.5),

$$\Sigma_2 \leq \sqrt{[u,v]_{\mathcal{A}_2^d} \mathcal{B}_{u,v} \mathcal{D}_{u,v}} \left(\int_{\mathbb{R}} M^2_{u^{-1}} f(x) u^{-1}(x) dx \right)^{1/2} \left(\int_{\mathbb{R}} M_v^2 g(x) v(x) dx \right)^{1/2}$$

$$\leq \sqrt{[u,v]_{\mathcal{A}_2^d} \mathcal{B}_{u,v} \mathcal{D}_{u,v}} \|M_{u^{-1}} f\|_{L^2(u^{-1})} \|M_v g\|_{L^2(v)}$$

$$\leq 8 \sqrt{[u,v]_{\mathcal{A}_2^d} \mathcal{B}_{u,v} \mathcal{D}_{u,v}} \|f\|_{L^2(u^{-1})} \|g\|_{L^2(v)}.$$

This estimate, together with estimate (14), gives (12). □

We can replace the conditions on the pair (u, v) by boundedness of the dyadic square function to deduce boundedness of the dyadic paraproduct when $b \in Carl_{u,v}$.

Corollary 4.2 *Let $b \in L^1_{loc}(\mathbb{R})$ and (u, v) be a pair of functions such that v and u^{-1} are weights and $\{|b_I|^2/m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence ($b \in Carl_{u,v}$) with intensity $\mathcal{B}_{u,v}$. If the dyadic square function S^d is bounded from $L^2(v^{-1})$ into $L^2(u^{-1})$ then the paraproduct π_b is bounded from $L^2(u)$ into $L^2(v)$. Moreover*

$$\|\pi_b f\|_{L^2(v)} \leq C\sqrt{[u, v]_{\mathcal{A}_2^d}\mathcal{B}_{u,v}}\Big(\sqrt{[u, v]_{\mathcal{A}_2^d}} + \|S^d\|_{L^2(v^{-1})\to L^2(u^{-1})}\Big)\|f\|_{L^2(u)}.$$

Proof Assume S^d is bounded from $L^2(v^{-1})$ into $L^2(u^{-1})$. Theorem 3.6 implies that $(u, v) \in \mathcal{A}_2$ and $\{|\Delta_I v|^2|I|m_I(u^{-1})\}_{I\in\mathcal{D}}$ is v-Carleson sequence with intensity $\mathcal{C}_{v^{-1},u^{-1}}$. Moreover, $\mathcal{C}_{v^{-1},u^{-1}} \leq \|S^d\|^2_{L^2(v^{-1})\to L^2(u^{-1})}$. These two facts together with the hypothesis that $\{|b_I|^2/m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence imply, by Theorem 1.1, that π_b is bounded from $L^2(u)$ to $L^2(v)$. The claimed estimate holds. □

If we especialize to the one weight case $u = v = w \in A_2^d$ then $\|S^d\|_{L^2(w^{-1})} \leq C[w^{-1}]_{A_2^d} = C[w]_{A_2^d}$. Moreover, $b \in Carl_{w,w} \cap L^2_{loc}$ is equivalent to $b \in BMO^d$ and $\mathcal{B}_{w,w} \leq C\|b\|^2_{BMO^d}$, we show this in Corollary 4.6. The previous Corollary would give us that

$$\|\pi_b\|_{L^2(w)\to L^2(w)} \leq C\|b\|_{BMO^d}[w]^{\frac{3}{2}}_{A_2^d}.$$

Thus, we do not recover Beznosova's linear bound, we are off by $[w]^{\frac{1}{2}}_{A_2^d}$.

Comparison to One-Sided Bump Theorems

The dyadic paraproduct is especially interesting because it allows us to estimate Calderón-Zygmund singular integral operators (CZSIO). The general approach to the two weight estimates for the CZSIO as a class is a bump-approach. We refer the reader to [48] for the precise definitions and statements, the interested reader can also consult [59] in this volume.

Theorem 4.3 ([48, Theorem 3.2]) *Suppose Φ satisfies several conditions.*[2] *Suppose that there exists a constant C such that for all $I \in \mathcal{D}$*

$$\|u^{-1}\|_{L,I}\|v\|_{\Phi(L),I} \leqslant C. \tag{15}$$

[2]The conditions on the function Φ are satisfied by the functions $\Phi(L) = L\log^{1+\sigma} L$ and $L\log L \log\log^{1+\sigma} L$ (for sufficiently large $\sigma > 0$), but not by $\Phi(L) = L\log L$.

Then any Calderón-Zygmund singular integral operator T is weakly bounded from $L^2(u)$ into $L^{2,\infty}(v)$, i.e.,

$$v\{x \in \mathbb{R} : |Tf(x)| \geq \lambda\} \leqslant \left(\frac{C\|f\|_{L^2(u)}}{\lambda}\right)^2. \tag{16}$$

Let us assume that u and v are such that

$$\|u^{-1}\|_{L,J}\|v\|_{L\log L,J} \leqslant C,$$

which is a weaker condition than the condition in Theorem 4.3. Then by Theorem 2.8 we have that, for every $J \in \mathcal{D}$,

$$\|v\|_{L\log L,J} \approx \frac{1}{|J|}\sum_{I\in\mathcal{D}(J)}\frac{|\Delta_I v|^2}{m_I v}|I| \leqslant \frac{C}{m_J(u^{-1})}. \tag{17}$$

The condition we have for the paraproduct is

$$\frac{1}{|J|}\sum_{I\in D(J)}|\Delta_I v|^2 m_I(u^{-1})|I| \leqslant Cm_J(v) \tag{18}$$

Note that if $(u, v) \in \mathcal{A}_2^d$ we have that

$$\frac{1}{|J|}\sum_{I\in D(J)}|\Delta_I v|^2 m_I(u^{-1})|I| \leqslant [u, v]_{\mathcal{A}_2^d}\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)}\frac{|\Delta_I v|^2}{m_I v}|I|$$

while $m_J v \leqslant \frac{[u,v]_{A_2^d}}{m_J(u^{-1})}$. Therefore we cannot compare bump conditions to the conditions in our results without the additional assumption that there is a constant $q > 0$ such that $m_J(u^{-1})m_J v \geqslant q$ for all $J \in \mathcal{D}$. If $q \leqslant m_J(u^{-1})m_J v \leqslant Q$ for all $J \in \mathcal{D}$; the two conditions (17) and (18) become equivalent, but this assumption essentially reduces the problem to the one weight case [44, Proposition 7.4].

BMO vs $\mathcal{C}_{v,v}$

Formally the dyadic paraproduct is a bilinear operator for the locally integrable functions b and f. After we fix b in BMO^d, we consider π_b as a linear operator acting on f. In the following proposition, we try to answer the question: *if π_b is bounded on (weighted) Lebesgue spaces, then in which space does the locally square integrable function b lie?*

Proposition 4.4 (A necessary condition for boundedness of π_b) *Let u and v be weights and, for $1 < p < \infty$, $b \in L^2_{loc}(\mathbb{R})$. Assume $\pi_b : L^p(u) \to L^p(v)$ is a bounded operator then there is a constant $C_p > 0$ such that for any $I \in \mathcal{D}$,*

$$\int_I |b(x) - m_I b|^p v(x)dx \le C_p u(\hat{I}), \tag{19}$$

where $\hat{I}$ is the dyadic parent of I. The constant $C_p^{1/p}$ is the operator norm $\|\pi_b\|_{L^p(u)\to L^p(v)}$.

Proof Let us choose $f = h_J$ for some dyadic interval J. Then, by assumption, there exists a constant $C_p = \|\pi_b\|^p_{L^p(u)\to L^p(v)}$ such that

$$\int_{\mathbb{R}} |\pi_b(h_J)(x)|^p v(x)dx \le C_p \int_{\mathbb{R}} |h_J(x)|^p u(x)dx = C_p \frac{u(J)}{|J|^{p/2}}. \tag{20}$$

On the other hand,

$$\begin{aligned}
\pi_b(h_J)(x) &= \sum_{I\in\mathcal{D}} m_I(h_J)\langle b, h_I\rangle h_I(x) \\
&= \sum_{I\in\mathcal{D}(J_+)} \frac{1}{\sqrt{|J|}}\langle b, h_I\rangle h_I(x) - \sum_{I\in\mathcal{D}(J_-)} \frac{1}{\sqrt{|J|}}\langle b, h_I\rangle h_I(x) \\
&= \frac{1}{\sqrt{|J|}}\Big[(b(x) - m_{J_+}b)\mathbb{1}_{J_+}(x) - (b(x) - m_{J_-}b)\mathbb{1}_{J_-}(x)\Big],
\end{aligned}$$

where the last equality is due to the fact that $(b - m_J b)\mathbb{1}_J = \sum_{I\in D(J)}\langle b, h_I\rangle h_I$. Therefore we can write

$$\begin{aligned}
\int_{\mathbb{R}} |\pi_b(h_J)(x)|^p u(x)dx &= \frac{1}{|J|^{p/2}}\int_{\mathbb{R}} \left|(b(x) - m_{J_+}b)\mathbb{1}_{J_+}(x) - (b(x) - m_{J_-}b)\mathbb{1}_{J_-}(x)\right|^p v(x)dx \\
&= \frac{1}{|J|^{p/2}}\left(\int_{J_+} |b(x) - m_{J_+}b|^p v(x)dx + \int_{J_-} |b(x) - m_{J_-}b|^p v(x)dx\right).
\end{aligned}$$

Thus we can conclude that there is a constant C_p such that for all $I \in \mathcal{D}$

$$\int_I |b(x) - m_I b|^p v(x)dx \le C_p u(\hat{I}).$$

□

The condition (19) can be considered as a testing condition for the boundedness of the dyadic paraproduct from $L^p(u)$ into $L^p(v)$. When $u, v \in A^d_p$ both weights are doubling weights, in particular $u(\hat{I}) \le D(u)\, u(I)$ (where $D(u) := \sup_{I\in\mathcal{D}} u(\hat{I})/u(I) < \infty$ is the dyadic doubling constant of u). In this case, (19) becomes

$$\int_I |b(x) - m_I b|^p v(x)\, dx \leq C_p u(I)$$

which is equivalent to the boundedness of the paraproduct and its adjoint [20, Theorem 4.1] from $L^p(u)$ into $L^p(v)$ when $u, v \in A_p^d$. When $u = v$ this necessary condition was known in the more general matrix A_p context [28].

One can immediately conclude that the inequality (19) implies that b is in BMO^d for $u = v = 1$ (Lebesgue space). Thus, one can view the condition $b \in BMO^d$ as a testing condition for the boundedness of the paraproduct on $L^2(\mathbb{R})$, in the same way that the conditions $T1, T^*1 \in BMO$ in the celebrated $T1$ Theorem are testing conditions.

For the weighted Lebesgue space, we have the following corollary.

Corollary 4.5 *For $1 < p < \infty$, $b \in L^2_{loc}(\mathbb{R})$, if π_b is bounded from $L^p(v)$ into itself and v is an A_p^d weight, then b belongs to BMO^d. Moreover, $\|b\|_{BMO^d} \leq 2\|\pi_b\|_{L^p(v)\to L^p(v)}[v]_{A_p^d}^{1/p}$.*

Proof For any $I \in \mathcal{D}$, we have

$$\begin{aligned}
\int_I |b(x) - m_I b|\, dx &= \int_I |b(x) - m_I b| v^{\frac{1}{p}}(x) v^{-\frac{1}{p}}(x)\, dx \\
&\leq \left(\int_I |b(x) - m_I b|^p v(x)\, dx \right)^{\frac{1}{p}} \left(\int_I v^{-\frac{p'}{p}}(x)\, dx \right)^{\frac{1}{p'}} \\
&\leq C_p^{1/p} \left(\int_{\hat{I}} v(x)\, dx \right)^{\frac{1}{p}} \left(\int_{\hat{I}} v^{-\frac{p'}{p}}(x)\, dx \right)^{\frac{1}{p'}} \qquad (21) \\
&= C_p^{\frac{1}{p}} |\hat{I}| \left(\frac{1}{|\hat{I}|} \int_{\hat{I}} v(x)\, dx \right)^{\frac{1}{p}} \left(\frac{1}{|\hat{I}|} \int_{\hat{I}} v^{-\frac{1}{p-1}}(x)\, dx \right)^{\frac{p-1}{p}} \qquad (22) \\
&\leq 2\|\pi_b\|_{L^p(v)\to L^p(v)} [v]_{A_p^d}^{\frac{1}{p}} |I|.
\end{aligned}$$

Here the inequality (21) holds due to (19) with $v = u$. □

Notice that $b \in BMO$ implies that $b \in L^p_{loc}(\mathbb{R})$ for all $1 \leq p < \infty$ by the John-Nirenberg inequality.

For the two weight case, in order to show that (22) is bounded, we need $(v, u) \in \mathcal{A}_p$ which is totally different from $(u, v) \in \mathcal{A}_p$. Thus, we cannot conclude anything more than (19) for the two weight situation.

To finish this section, we give a relation between BMO^d and $Carl_{v,v}$.

Corollary 4.6 *If $v \in A_2^d$ then*

$$BMO^d = Carl_{v,v} \cap L^2_{loc}(\mathbb{R}).$$

Proof In section "Carleson Sequences", we observed that $BMO^d \subset Carl_{v,v}$ for any weight v such that v^{-1} is also a weight. Also recall that by the John-Nirenberg theorem if $b \in BMO$ then $b \in L^2_{loc}(\mathbb{R})$. Thus, to complete the proof, we need to show that if $v \in A_2^d$ and $b \in Carl_{v,v} \cap L^2_{loc}(\mathbb{R})$ then $b \in BMO^d$. If $v \in A_2^d$ then in particular $v \in RH_1^d$. By Theorem 2.8, it follows that, for every dyadic interval J, we have

$$\frac{1}{|J|}\sum_{I\in D(J)} |\Delta_I v|^2 m_I(v^{-1})|I| \leqslant [v]_{A_2^d}\frac{1}{|J|}\sum_{I\in D(J)} \frac{|\Delta_I v|^2}{m_I v}|I| \leqslant C[v]_{A_2^d}[v]_{RH_1^d} m_J v \ . \quad (23)$$

Since $v \in A_2^d$ and $b \in Carl_{v,v}$, all conditions of Theorem 1.1 are satisfied, and we know that the dyadic paraproduct, π_b, is bounded from $L^2(v)$ into $L^2(v)$. Thus, by Corollary 4.5, b must belong to BMO^d. □

Similar one weight results are shown by Isralowitz, Kwon, and Pott [28] in the much more general matrix A_p context.

$Carl_{u,v}$ vs Bloom's BMO

There are other weighted bounded mean oscillation spaces in the literature. The dyadic weighted BMO space for a weight μ in $\mathbb{R}^d$, denoted $BMO^d(\mu)$ in [20, Section 2.6], consists of all locally integrable functions b such that

$$\|b\|_{BMO^d(\mu)} := \sup_Q \frac{1}{\mu(Q)}\int_Q |b(x) - m_Q b|\,dx < \infty,$$

where the supremum is taken over all dyadic cubes with sides parallel to the axes. In that paper, it is pointed out that when the weight is in A_∞ (hence, in particular, is a doubling weight), one can replace the L^1 with L^p norm provided the integration with respect to the Lebesgue measure is replaced by $\sigma\,dx$ where $\sigma = \mu^{\frac{-1}{p-1}}$ is the conjugate weight.

When $u, v \in A_2^d$, let $\mu := u^{1/2}v^{-1/2}$, the corresponding weighted $BMO^d(\mu)$ is Bloom's BMO [5]. In [20, Theorem 4.1] it is shown that the following are equivalent conditions.

(i) $b \in BMO^d(\mu)$.
(ii) $b \in BMO_2^d(\mu)$ meaning $\sup_{I\in\mathcal{D}} \frac{1}{\mu(I)}\int_I |b(x) - m_I b|^2 \mu^{-1}(x)\,dx < \infty$.
(iii) $\sup_{I\in\mathcal{D}} \frac{1}{u(I)}\int_I |b(x) - m_I b|^2 v(x)\,dx < \infty$.
(iv) $\sup_{I\in\mathcal{D}} \frac{1}{v^{-1}(I)}\int_I |b(x) - m_I b|^2 u^{-1}(x)\,dx < \infty$.
(v) π_b is bounded from $L^2(u)$ into $L^2(v)$.
(vi) π_b^* from $L^2(u)$ into $L^2(v)$.

Theorem 4.7 *Assume $u, v \in A_2^d$ and let $\mu = u^{1/2}v^{-1/2}$. Then $b \in Carl_{u,v}$ if and only if $b \in Carl_{v^{-1},u^{-1}}$ if and only if $b \in BMO^d(\mu^{-1})$.*

Proof First we will show that $Carl_{u,v} \cup Carl_{v^{-1},u^{-1}} \subset BMO^d(\mu^{-1})$. Assume $b \in Carl_{u,v} \cup Carl_{v^{-1},u^{-1}}$. By assumption there is $C > 0$ such that for all $J \in \mathcal{D}$

$$(a) \sum_{I\in\mathcal{D}(J)} \frac{b_I^2}{m_I v} \le Cu^{-1}(J), \quad \text{or} \quad (b) \sum_{I\in\mathcal{D}(J)} \frac{b_I^2}{m_I u^{-1}} \le Cv(J).$$

When $w \in A_2^d$ the dyadic square function S^d obeys an inverse estimate $\|f\|_{L^2(w)} \le C[w]_{A_2^d}^{1/2}\|S^d f\|_{L^2(w)}$. In case (a), since $v^{-1} \in A_2^d$ we can use the inverse estimate for S^d in $L^2(v^{-1})$ and get, for all $J \in \mathcal{D}$, the estimate

$$\begin{aligned}\|(b - m_J b)\mathbb{1}_J\|_{L^2(v^{-1})}^2 &\le C[v]_{A_2^d}\|S^d\big((b - m_J b)\mathbb{1}_J\big)\|_{L^2(v^{-1})}^2 \\ &= C[v]_{A_2^d} \sum_{I\in\mathcal{D}(J)} b_I^2 m_I v^{-1} \\ &\le C[v]_{A_2^d}^2 \sum_{I\in\mathcal{D}(J)} \frac{b_I^2}{m_I v} \\ &\le C[v]_{A_2^d}^2 u^{-1}(J)\end{aligned}$$

Hence we conclude that $\sup_{I\in\mathcal{D}} \frac{1}{u^{-1}(I)} \int_I |b(x) - m_I b|^2 v^{-1}(x)\, dx < \infty$.

Similarly if we assume (b), we will conclude $\sup_{I\in\mathcal{D}} \frac{1}{v(I)} \int_I |b(x) - m_I b|^2 u(x)\, dx < \infty$, using this time that $u \in A_2^d$. These integral conditions are each separately equivalent to $b \in BMO(\mu^{-1})$ when $u, v \in A_2^d$ by the results in [20, Theorem 4.1].

Assume now that $b \in BMO(\mu^{-1})$ and $u, v \in A_2^d$. We will show that $b \in Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$. The assumption implies that

$$\|(b - m_J b)\mathbb{1}_J\|_{L^2(v^{-1})}^2 \le Cu^{-1}(J) \ \text{ and } \ \|(b - m_J b)\mathbb{1}_J\|_{L^2(u)}^2 \le Cv(J).$$

Both $u, v \in A_2^d$ so are $u^{-1}, v^{-1} \in A_2^d$, also $1 \le m_I v\, m_I v^{-1}$, and the dyadic square function is bounded in $L^2(w)$ for $w \in A_2^d$, moreover $\|S^d\big((b - b_J)\mathbb{1}_J\big)\|_{L^2(w)}^2 = \sum_{I\in\mathcal{D}(J)} |b_I|^2 m_I w$. We therefore conclude that

$$\sum_{I\in\mathcal{D}(J)} \frac{|b_I|^2}{m_I v} \le \sum_{I\in\mathcal{D}(J)} |b_I|^2 m_I v^{-1} \le C[v]_{A_2^d}^2 \|(b - m_J b)\mathbb{1}_J\|_{L^2(v^{-1})}^2 \le Cu^{-1}(J),$$

$$\sum_{I\in\mathcal{D}(J)} \frac{|b_I|^2}{m_I u^{-1}} \le \sum_{I\in\mathcal{D}(J)} |b_I|^2 m_I u \le C[u]_{A_2^d}^2 \|(b - m_J b)\mathbb{1}_J\|_{L^2(u)}^2 \le Cv(J).$$

Hence $b \in Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$.

All together we have shown $Carl_{u,v} \cup Carl_{v^{-1},u^{-1}} \subset BMO(\mu^{-1}) \subset Carl_{u,v} \cap Carl_{v^{-1},u^{-1}}$ which implies that $Carl_{u,v} = BMO(\mu^{-1}) = Carl_{v^{-1},u^{-1}}$ when $u, v \in A_2^d$. □

We just showed that when $u, v \in A_2^d$ and the dyadic paraproduct π_b is bounded from $L^2(u)$ into $L^2(v)$ then $b \in Carl_{v,u}$. Compare to Corollary 4.9 where only $v \in A_2^d$ and the pair (u, v) is in joint $\mathcal{A}_2$, but we assume $b \in Carl_{u,v}$ (note that the roles of u and v have been interchanged, and in general $Carl_{u,v} \neq Carl_{v,u}$).

When we assume only $v \in A_2^d$ then π_b is bounded from $L^2(u)$ into $L^2(v)$ iff (11), that is

$$B_2(u, v) := \sup_{I\in\mathcal{D}} \frac{1}{u^{-1}(I)} \sum_{J\in\mathcal{D}(I)} b_J^2 m_J(u^{-1})^2 m_J v < \infty.$$

Lemma 4.8 *If $(u, v) \in \mathcal{A}_2^d$ and $b \in Carl_{u,v}$ with intensity $\mathcal{B}_{u,v}$ then $B_2(u, v) < \infty$. Moreover*

$$B_2(u, v) \leq [u, v]^2_{\mathcal{A}_2^d} \mathcal{B}_{u,v}.$$

Proof The result follows immediately using first the joint $\mathcal{A}_2$ condition and then the $Carl_{u,v}$ condition,

$$\sum_{J\in\mathcal{D}(I)} b_J^2 m_J(u^{-1})^2 m_J v \leq [u, v]^2_{\mathcal{A}_2^d} \sum_{J\in\mathcal{D}(I)} \frac{b_J^2}{m_J v} \leq [u, v]^2_{\mathcal{A}_2^d} \mathcal{B}_{u,v} u^{-1}(I).$$

This implies $B_2(u, v) \leq [u, v]^2_{\mathcal{A}_2^d} \mathcal{B}_{u,v} < \infty$ as required. □

Using the results in [21] we will conclude that

Corollary 4.9 *If $(u, v) \in \mathcal{A}_2^d$, $v \in A_2^d$, and $b \in Carl_{u,v}$ then π_b is bounded from $L^2(u)$ into $L^2(v)$.*

As observed in [44] if $(u, v) \in \mathcal{A}_2^d$, $v \in A_2^d$ (or $u \in A_2^d$), and $b \in BMO^d$ then the boundedness of the paraproduct reduces to one weight boundedness on $L^2(v)$ (or on $L^2(u)$). The observation being that joint $\mathcal{A}_2$ implies, by the Lebesgue Differentiation Theorem, that $v(x) \leq [u, v]_{\mathcal{A}_2^d} u(x)$ for a.e. x, and therefore $\|g\|_{L^2(v)} \leq [u, v]^{1/2}_{\mathcal{A}_2^d} \|g\|_{L^2(u)}$. If $v \in A_2^d$ then by Beznosova's one weight linear bound for the paraproduct in $L^2(v)$ [1] one has

$$\|\pi_b f\|_{L^2(v)} \leq C[b]_{BMO^d}[v]_{A_2^d}\|f\|_{L^2(v)} \leq C[b]_{BMO^d}[v]_{A_2^d}[u, v]^{1/2}_{\mathcal{A}_2^d}\|f\|_{L^2(u)}.$$

Likewise if $u \in A_2^d$, then

$$\|\pi_b f\|_{L^2(v)} \le [u,v]_{\mathcal{A}_2^d}^{1/2} \|\pi_b f\|_{L^2(u)} \le C[b]_{BMO^d}[u]_{A_2^d}[u,v]_{\mathcal{A}_2^d}^{1/2} \|f\|_{L^2(u)},$$

where we used Beznosova's result in the last inequality. Using this observation we can deduce Corollary 4.9 without using the machinery of [21] if we can prove that $(u,v) \in \mathcal{A}_2^d$, $v \in A_2^d$, and $b \in Carl_{u,v}$ imply $b \in BMO^d$.

Lemma 4.10 *If $(u,v) \in \mathcal{A}_2^d$, $v \in A_2^d$, and $b \in Carl_{u,v} \cap L^2_{loc}(\mathbb{R})$ then $b \in BMO^d$.*

Proof Suffices to show that $b \in Carl_{v,v}$. Notice that the Cauchy-Schwarz inequality and the joint $\mathcal{A}_2$ condition imply $\frac{1}{v^{-1}(J)} \le v(J) \le \frac{[u,v]_{\mathcal{A}_2^d}}{u^{-1}(J)}$, therefore

$$\frac{1}{v^{-1}(J)} \sum_{I \in \mathcal{D}(J)} \frac{b_I^2}{m_I v} \le \frac{[u,v]_{\mathcal{A}_2^d}}{u^{-1}(J)} \sum_{I \in \mathcal{D}(J)} \frac{b_I^2}{m_I v} \le [u,v]_{\mathcal{A}_2^d} \mathcal{B}_{u,v}.$$

We conclude that $b \in Carl_{v,v} \cap L^2_{loc}(\mathbb{R}) = BMO^d$ by Corollary 4.6. □

It may be worth to point out that when $u = v \in A_2$, the condition $\mathcal{B}_2(v,v) < \infty$ coincides with $b \in Carl_{v,v}$. The reason being that now we do have the lower bound as well as the upper bound $1 \le m_I v\, m_I(v^{-1}) \le [v]_{A_2}$.

Lemma 4.11 *if $w \in A_2$ then $b \in Carl_{w,w}$ if and only if $\mathcal{B}_2(w) < \infty$, where*

$$\mathcal{B}_2(w) := \mathcal{B}_2(w,w) = \sup_{J \in \mathcal{D}} \frac{1}{w^{-1}(J)} \sum_{I \in \mathcal{D}(J)} m_I^2(w^{-1})|b_I|^2 m_I w.$$

The Maximal and the Square Functions

In this section we relate the boundedness of the square function with the boundedness of the Maximal function from $L^2(u)$ into $L^2(v)$. The main result of this section states that if the weight v is in RH_1^d and the Maximal function is bounded then the square function is also bounded. This result is an adaptation of Buckley's proof [6], for the fact that if $w \in A_2^d$ then S^d is bounded on $L^2(w)$. The last author proved a similar result, in [52], for the weighted maximal function and the weighted square function in $L^q(\mathbb{R})$ and $1 < q < \infty$.

Theorem 5.1 *Let (u,v) be a pair of weights such that $v \in RH_1^d$ and the Maximal function M is bounded from $L^2(u)$ into $L^2(v)$ with bound $\mathcal{M}_{u,v}$ then there exist $C > 0$ such that*

$$\|S^d f\|_{L^2(v)} \le C\mathcal{M}_{u,v}(1 + [v]_{RH_1^d}^{1/2})\|f\|_{L^2(u)}.$$

As an immediate Corollary of Theorem 5.1 and Theorem 3.4 we get,

Corollary 5.2 *Assume $(u,v)\in\mathcal{A}_2^d$, $u^{-1}\in RH_1^d$, and $v\in RH_1^d$, then*

$$\|S^d f\|_{L^2(v)}\le C\big([u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d}\big)^{1/2}(1+[v]_{RH_1^d}^{1/2})\|f\|_{L^2(u)}.$$

Note that this estimate does not recover the linear estimate in the one weight case $u=v\in A_2$, it is off by a factor of the form $[v]_{RH_1}^{1/2}$, unlike the estimate we will present in Theorem 6.1.

Proof (Proof of Theorem 5.1)
Given real-valued $f\in L^2(u)$ we have

$$\begin{aligned}\|S^d f\|_{L^2(v)}^2 &=\sum_{I\in\mathcal{D}}|\langle f,h_I\rangle|^2 m_I v=\frac{1}{2}\sum_{I\in\mathcal{D}}|m_I f-m_{\hat{I}}f|^2 v(\hat{I})\\ &=\frac{1}{2}\sum_{I\in\mathcal{D}}\big(m_I^2 f-m_{\hat{I}}^2 f\big)v(\hat{I}):=\Sigma_1.\end{aligned}$$

Adding and subtracting $2v(I)m_I^2 f$ in the sum and rearranging

$$\Sigma_1=\sum_{I\in\mathcal{D}}\big(2v(I)m_I^2 f-v(\hat{I})m_{\hat{I}}^2 f\big)+\sum_{I\in\mathcal{D}}\big(v(\hat{I})-2v(I)\big)m_I^2 f=:\Sigma_2+\Sigma_3.$$

Therefore, it is enough to check that for all $f\in L^2(u)$:

$$|\Sigma_i|\le C\mathcal{M}_{u,v}^2(1+[v]_{RH_1^d}^{1/2})^2\|f\|_{L^2(u)}^2\qquad\text{for}\qquad i=2,3.$$

Estimating Σ_2: First, let $a_m:=\sum_{I\in\mathcal{D}_m}2v(I)m_I^2 f=2\int(E_m f(x))^2 v(x)dx$ where $E_m f(x):=m_I f$ for $x\in I\in\mathcal{D}_m$ and $\mathcal{D}_m$ is the collection of all dyadic intervals with length 2^{-m}. Then

$$\Sigma_2:=\sum_{I\in\mathcal{D}}\big(2v(I)m_I^2 f-v(\hat{I})m_{\hat{I}}^2 f\big)=\sum_{m=-\infty}^{\infty}(a_m-a_{m-1}).$$

Using the fact that $E_m f(x)\le Mf(x)$ for all $x\in\mathbb{R}$ we can bound each a_m by

$$|a_m|\le 2\int_{\mathbb{R}}|Mf(x)|^2 v(x)dx=2\|Mf\|_{L^2(v)}^2\le C\mathcal{M}_{u,v}^2\|f\|_{L^2(u)}^2.$$

The last inequality follows since M is assumed to be bounded from $L^2(u)$ to $L^2(v)$. Let $s_n:=\sum_{|m|\le n}(a_m-a_{m-1})$, the partial sum sequence of Σ_2. Since this is a telescoping sum we have $s_n=(a_n-a_{-n-1})$ for all $n\in\mathbb{N}$. Therefore $|s_n|\le 2C\mathcal{M}_{u,v}^2\|f\|_{L^2(u)}^2$ for all $n\in\mathbb{N}$ which leads us to the better than desired estimate

$$|\Sigma_2|\le C\mathcal{M}_{u,v}^2\|f\|_{L^2(u)}^2.$$

Estimating Σ_3: Since every interval has two children, switching the sum over I to a sum over the parents $J = \hat{I}$ we have the following cancellation,

$$\sum_{I\in\mathcal{D}} \big(v(\hat{I}) - 2v(I)\big)m_{\hat{I}}^2 f = \sum_{J\in\mathcal{D}} \big(v(J) - 2v(J_+) + v(J) - 2v(J_-)\big)m_J^2 f = 0.$$

Hence we can write

$$\Sigma_3 = \sum_{I\in\mathcal{D}} \big(v(\hat{I}) - 2v(I)\big)\big(m_I^2 f - m_{\hat{I}}^2 f\big).$$

Applying the Cauchy-Schwarz inequality,

$$\begin{aligned}|\Sigma_3| &\le \Bigg(\sum_{I\in\mathcal{D}} \frac{\big(v(\hat{I}) - 2v(I)\big)^2}{v(\hat{I})}(m_I f + m_{\hat{I}} f)^2\Bigg)^{1/2}\Bigg(\sum_{I\in\mathcal{D}} v(\hat{I})(m_I f - m_{\hat{I}} f)^2\Bigg)^{1/2} \\ &= 2^{1/2}\sqrt{\Sigma_4\Sigma_1} \le 2^{-1/2}(\Sigma_4 + \Sigma_1)\,,\end{aligned}$$

where $\Sigma_4 := \sum_{I\in\mathcal{D}} \frac{\left(v(\hat{I})-2v(I)\right)^2}{v(\hat{I})}(m_I f + m_{\hat{I}} f)^2$. Thus,

$$\Sigma_1 \le |\Sigma_2| + |\Sigma_3| \le C\mathcal{M}_{u,v}^2 \|f\|_{L^2(u)}^2 + 2^{-1/2}(\Sigma_4 + \Sigma_1).$$

Subtracting $2^{-1/2}\Sigma_1$ from both sides of this inequality and multiplying by $(2+\sqrt{2})$ give us

$$\Sigma_1 \lesssim \mathcal{M}_{u,v}^2 \|f\|_{L^2(u)}^2 + \Sigma_4. \tag{24}$$

Estimating Σ_4: Note that $m_I|f| \le 2m_{\hat{I}}|f|$, hence $(m_I f + m_{\hat{I}} f)^2 \le (3m_{\hat{I}}|f|)^2$, and also note that $|v(\hat{I}) - 2v(I)| = 2|\hat{I}|^{-1}|\Delta_{\hat{I}} v|$. Switching the sum over I to a sum over the parents $J = \hat{I}$ gives

$$\Sigma_4 \lesssim \sum_{J\in\mathcal{D}} \frac{|\Delta_J v|^2}{m_J v}\,|J|\, m_J^2|f|.$$

Thus

$$\begin{aligned}\Sigma_4 &\lesssim \sum_{I\in\mathcal{D}} \frac{|\Delta_I v|^2}{m_I v}|I| m_I^2|f| \lesssim \sum_{I\in\mathcal{D}} \frac{|\Delta_I v|^2}{m_I v}|I| \inf_{x\in I} M^2 f(x) \\ &\lesssim [v]_{RH_1^d}\int_{\mathbb{R}} M^2 f(x)v(x)dx = [v]_{RH_1^d}\|Mf\|_{L^2(v)}^2 \le [v]_{RH_1^d}\mathcal{M}_{u,v}^2\|f\|_{L^2(u)}^2.\end{aligned}$$

Note that in the third inequality we use the fact that if $v \in RH_1^d$ then, by Theorem 2.8, $\{|\Delta_I v|^2|I|/m_I v\}_{I\in\mathcal{D}}$ is a v-Carleson sequence with intensity $[v]_{RH_1^d}$. This estimate together with (24) give us the desired estimate for real-valued functions. Using this estimate for the real and complex parts of $f \in L^2(v)$ we will conclude that

$$\|S^d f\|_{L^2(v)} \leq C\mathcal{M}_{u,v}(1 + [v]_{RH_1^d}^{1/2})\|f\|_{L^2(u)}.$$

□

Even though not explicitly we are still assuming that $(u, v) \in \mathcal{A}_2^d$, since we assumed that $M : L^2(u) \to L^2(v)$ which implies $(u, v) \in \mathcal{A}_d^2$, see [16].

Remark 5.3 In the last theorem we are providing a connection between the boundedness of the square function and the boundedness of the Maximal function. Another novelty of this result is that we have an estimate on how the norm of the square function depends on $[v]_{RH_1^d}$ and the norm of the Maximal function.

As a consequence of Theorems 5.1 and 3.6, we can show that the boundeness of the Maximal function from $L^2(u)$ into $L^2(v)$ together with the assumption that $v \in RH_1^d$ will imply the boundedness of the martingale transform.

Theorem 5.4 *Let (u, v) be a pair of weights such that $v \in RH_1^d$ and the Maximal function M is bounded from $L^2(u)$ into $L^2(v)$ then the martingale transforms T_r are uniformly bounded from $L^2(u)$ into $L^2(v)$.*

Proof Let us consider a pair of weights (u, v) satisfying the assumptions. By Theorem 5.1 the dyadic square function is bounded, and by Theorem 3.8 the pair of weights (u, v) satisfies

(i) $(u, v) \in \mathcal{A}_2^d$
(ii) $\{|\Delta_I u^{-1}|^2 m_I v|I|\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence.

Let us denote the intensity of the u^{-1}-Carleson sequence in (ii) by $\mathcal{D}_{u,v}$. To prove the boundedness of the martingale transform T_r, we need to show that (u, v) also satisfies the last two conditions in Theorem 3.6

(iii) $\{|\Delta_I v|^2 m_I(u^{-1})|I|\}_{I\in\mathcal{D}}$ is a v-Carleson sequence.
(iv) The operator T_0 is bounded from $L^2(u)$ into $L^2(v)$.

For condition (iii), we use the assumption $v \in RH_1^d$, Theorem 2.8(b), and $(u, v) \in \mathcal{A}_2^d$. More precisely, for any $J \in \mathcal{D}$,

$$\sum_{I\in\mathcal{D}} |\Delta_I v|^2 m_I(u^{-1})|I| = \sum_{I\in\mathcal{D}(J)} \frac{|\Delta_I v|^2}{m_I v} m_I v\, m_I(u^{-1})|I|$$

$$\leq [u, v]_{\mathcal{A}_2^d} \sum_{I\in\mathcal{D}(J)} \frac{|\Delta_I v|^2}{m_I v}|I| \leq C[u, v]_{\mathcal{A}_2^d}[v]_{RH_1^d} v(J).$$

We now need to check condition (iv), which for any positive $f \in L^2(u^{-1})$ and $g \in L^2(v)$ is equivalent to

$$|\langle T_0(fu^{-1}), gv\rangle| \leq C\|f\|_{L^2(u^{-1})}\|g\|_{L^2(v)}\,.$$

Thus, it suffices to verify the estimate

$$\sum_{I\in\mathcal{D}} m_I(|f|u^{-1})m_I(|g|v)\frac{|\Delta_I v|}{m_I v}\frac{|\Delta_I u^{-1}|}{m_I(u^{-1})}|I| \leq C\|f\|_{L^2(u^{-1})}\|g\|_{L^2(v)}. \qquad (25)$$

To see that (25) holds, we use the Cauchy-Schwarz inequality:

$$\begin{aligned}
&\sum_{I\in\mathcal{D}} m_I(|f|u^{-1})m_I(|g|v)\frac{|\Delta_I v|}{m_I v}\frac{|\Delta_I u^{-1}|}{m_I(u^{-1})}|I| \\
&\quad\leq \left(\sum_{I\in\mathcal{D}}\left(\frac{m_I(|f|u^{-1})}{m_I(u^{-1})}\right)^2 |\Delta_I u^{-1}|^2 m_I v|I|\right)^{1/2}\left(\sum_{I\in\mathcal{D}}\left(\frac{m_I(|g|v)}{m_I v}\right)^2\frac{|\Delta_I v|^2}{m_I v}|I|\right)^{1/2} \\
&\quad\leq \left(\sum_{I\in\mathcal{D}}|\Delta_I u^{-1}|^2 m_I v|I| \inf_{x\in I} M^2_{u^{-1}}f(x)\right)^{1/2}\left(\sum_{I\in\mathcal{D}}\frac{|\Delta_I v|^2}{m_I v}|I| \inf_{x\in I} M^2_v g(x)\right)^{1/2}.
\end{aligned}$$

Since $|\Delta_I u^{-1}|^2 m_I v|I|$ is a u^{-1}-Carleson sequence with intensity $\mathcal{D}_{u,v}$ and $\frac{|\Delta_I v|^2}{m_I v}|I|$ is a v-Carleson sequence with intensity $[v]_{RH_1^d}$, by condition (ii) and Theorem 2.8(b) respectively, we have that

$$\begin{aligned}
\sum_{I\in\mathcal{D}} m_I(|f|u^{-1})m_I(|g|v)\frac{|\Delta_I v|}{m_I v}\frac{|\Delta_I u^{-1}|}{m_I(u^{-1})}|I| &\leq \sqrt{\mathcal{D}_{u,v}[v]_{RH_1^d}}\,\|M_{u^{-1}}f\|_{L^2(u^{-1})}\|M_v g\|_{L^2(v)} \\
&\leq 8\sqrt{\mathcal{D}_{u,v}[v]_{RH_1^d}}\,\|f\|_{L^2(u^{-1})}\|g\|_{L^2(v)}\,,
\end{aligned}$$

the last inequality by Theorem 3.2. □

As an immediate consequence of the proof of Theorem 5.4 and Corollary 5.2 we get the following corollary.

Corollary 5.5 *If* $(u, v) \in \mathcal{A}_2^d$, $u^{-1} \in RH_1^d$ *and* $v \in RH_1^d$ *then the martingale transforms* T_r *are uniformly bounded* $L^2(u)$ *into* $L^2(v)$.

The Sharp Quantitative Estimate for the Dyadic Square Function

Our last theorem provides the dependence of the operator norm $\|S^d\|_{L^2(u)\to L^2(v)}$ on the joint $\mathcal{A}_2$ characteristic of the weights and $[u^{-1}]_{RH_1^d}$. This extends results of

Beznosova [2], and we follow the template of her original proof. We could have used instead Proposition 2.1 as remarked by our careful referee, we leave to the interested reader to verify that this can indeed be done.

Theorem 6.1 *Let (u, v) be a pair of weights such that $(u, v) \in \mathcal{A}_2^d$ and $u^{-1} \in RH_1^d$. Then there is a constant such that*

$$\|S^d\|_{L^2(u)\to L^2(v)} \le C[u, v]_{\mathcal{A}_2^d}^{1/2}\big(1 + [u^{-1}]_{RH_1^d}^{1/2}\big).$$

Proof We can write the square of the norm $\|S^d f\|_{L^2(v)}$ as:

$$\|S^d f\|_{L^2(v)}^2 = \int \sum_{I\in\mathcal{D}} |\langle f, h_I\rangle|^2 \frac{\mathbb{1}_I(x)}{|I|} v(x)dx = \sum_{I\in\mathcal{D}} |\langle f, h_I\rangle|^2 m_I v.$$

We decompose h_I in a slightly different way. For any weight u^{-1}, we can write h_I as

$$h_I(x) = \frac{1}{\sqrt{|I|}}\Big(H_I^{u^{-1}}(x) + A_I^{u^{-1}}\mathbb{1}_I(x)\Big) \quad \text{where} \quad A_I^{u^{-1}} = \frac{\Delta_I u^{-1}}{2m_I(u^{-1})}.$$

The family $\{u^{-1/2}H_I^{u^{-1}}\}_{I\in\mathcal{D}}$ is orthogonal in $L^2(dx)$ with norms satisfying the inequality

$$\|u^{-1/2}H_I^{u^{-1}}\|_{L^2(\mathbb{R})} \le \sqrt{|I| m_I(u^{-1})}.$$

Hence by Bessel's inequality we have that for all $f \in L^2(u)$ (recall that $f \in L^2(u)$ if and only if $fu^{1/2} \in L^2(\mathbb{R})$),

$$\sum_{I\in\mathcal{D}} \frac{|\langle f, H_I^{u^{-1}}\rangle|^2}{|I| m_I(u^{-1})} \le \|fu^{1/2}\|_{L^2(\mathbb{R})}^2 = \|f\|_{L^2(u)}^2.$$

Since $m_I v \le [u, v]_{\mathcal{A}_2^d}/m_I(u^{-1})$ we conclude that for all $f \in L^2(u)$,

$$\sum_{I\in\mathcal{D}} \left|\left\langle f, \frac{H_I^{u^{-1}}}{\sqrt{|I|}}\right\rangle\right|^2 m_I v \le [u, v]_{\mathcal{A}_2^d} \|f\|_{L^2(u)}^2 . \tag{26}$$

We claim that

$$\sum_{I\in\mathcal{D}} \left|\left\langle f, \frac{A_I^{u^{-1}}\mathbb{1}_I}{\sqrt{|I|}}\right\rangle\right|^2 m_I v \le C[u, v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_d^1}\|f\|_{L^2(u)}^2. \tag{27}$$

Using estimates (26) and (27) and the Cauchy-Schwarz inequality we conclude that

$$\sum_{I\in\mathcal{D}} |\langle f, h_I\rangle|^2\, m_I v \le C\big([u,v]_{\mathcal{A}_2^d} + 2[u,v]_{A_d^2}[u^{-1}]_{RH_1^d}^{1/2} + [u,v]_{A_d^2}[u^{-1}]_{RH_1^d}\big)\|f\|^2_{L^2(u)}\,,$$

which completes the proof.

Let us return to our claim. The left hand side of (27) can be written as

$$\sum_{I\in\mathcal{D}} \left|\left\langle f, \frac{A_I^{u^{-1}}\mathbb{1}_I}{\sqrt{|I|}}\right\rangle\right|^2 m_I v = \frac{1}{4}\sum_{I\in\mathcal{D}} |m_I f|^2 \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I v.$$

By our assumptions: $(u,v)\in\mathcal{A}_2^d$ and $u^{-1}\in RH_1^d$, for any $J\in\mathcal{D}$, we have

$$\begin{aligned}\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I^2(u^{-1}) m_I v &\le \frac{[u,v]_{A_2^d}}{|J|}\sum_{I\in\mathcal{D}(J)} \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I(u^{-1}) \\ &\le [u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d} m_J(u^{-1})\,, \qquad (28)\end{aligned}$$

The last inequality (28) is an application of Lemma 2.8(b). Therefore the sequence $\{\alpha_I := \big(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\big)^2 |I| m_I^2(u^{-1}) m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence with intensity $[u,v]_{A_2^d}[u^{-1}]_{RH_1^d}$.

We now can prove the claimed estimate (27),

$$\begin{aligned}\sum_{I\in\mathcal{D}} |m_I f|^2 \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I v &= \sum_{I\in\mathcal{D}} \left(\frac{|m_I f|}{m_I(u^{-1})}\right)^2 \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I^2(u^{-1}) m_I v \\ &\le \sum_{I\in\mathcal{D}} \left(m_I^{u^{-1}}(|f|u)\right)^2 \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I^2(u^{-1}) m_I v \\ &\le \sum_{I\in\mathcal{D}} \inf_{x\in I} M^2_{u^{-1}}(fu)(x) \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I^2(u^{-1}) m_I v.\end{aligned}$$

Finally using Lemma 2.5 with $F(x) = M^2_{u^{-1}}(fu)(x)$ and the u^{-1}-Carleson sequence $\{\alpha_I\}_{I\in\mathcal{D}}$ with intensity $[u,v]_{A_2^d}[u^{-1}]_{RH_1^d}$, will give us that

$$\begin{aligned}\sum_{I\in\mathcal{D}} |m_I f|^2 \left(\frac{\Delta_I u^{-1}}{m_I(u^{-1})}\right)^2 |I| m_I v &\le C[u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d} \|M_{u^{-1}}(fu)\|^2_{L^2(u^{-1})} \\ &\le C[u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d} \|fu\|^2_{L^2(u^{-1})} \\ &= C[u,v]_{\mathcal{A}_2^d}[u^{-1}]_{RH_1^d} \|f\|^2_{L^2(u)}.\end{aligned}$$

□

Analyzing carefully the proof above we realize that if instead of assuming $u^{-1} \in RH_1^d$ we assume that $\{|\Delta_I u^{-1}|^2 m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence with intensity $\mathcal{C}_{u,v}$ the argument will go through and we will recover the Lacey-Li estimate.

Theorem 6.2 *Let (u, v) be a pair of weights such that $(u, v) \in \mathcal{A}_2^d$ and $\{|\Delta_I u^{-1}|^2 m_I v\}_{I\in\mathcal{D}}$ is a u^{-1}-Carleson sequence with intensity $\mathcal{C}_{u,v}$. Then there is a constant $C > 0$ such that*

$$\|S^d\|_{L^2(u)\to L^2(v)} \le C\big([u, v]_{\mathcal{A}_2^d} + \mathcal{C}_{u,v}\big)^{1/2}.$$

We leave the details of the proof to the reader.

References

1. O. Beznosova, Linear bound for the dyadic paraproduct on weighted Lebesgue space $L^2(w)$. J. Func. Anal. **255**, 994–1007 (2008)
2. O. Beznosova, Bellman functions, paraproducts, Haar multipliers, and weighted inequalities. Ph.D. dissertation, University of New Mexico, 2008
3. O. Beznosova, Perfect dyadic operators: weighted $T(1)$ theorem and two weight estimates. J. Math. Anal. Appl. **439**(2), 813–831 (2016)
4. O. Beznosova, A. Reznikov, Equivalent definitions of dyadic Muckenhoupt and Reverse Holder classes in terms of Carleson sequences, weak classes, and comparability of dyadic $L \log L$ and A_∞ constants. Rev. Mat. Iberoam. **30**(4), 1191–1190 (2014)
5. S. Bloom, A commutator theorem and weighted *BMO*. Trans. Am. Math. Soc. **292**(1), 103–122 (1985)
6. S.M. Buckley, Summation condition on weights. Mich. Math. J. **40**, 153–170 (1993)
7. M. Cao, K. Li, Q. Xue, A characterization of two weight norm inequality for Littlewood-Paley g_λ^*-function (2015, Preprint). Available at arXiv:1504.07850
8. R. Coiffman, C. Fefferman, Weighted norm inequalities for maximal functions and singular integrals. Stud. Math. **51**, 241–250 (1974)
9. M. Cotlar, C. Sadosky, On the Helson-Szegö theorem and a related class of modified Toeplitz kernels, in *Harmonic Analysis in Euclidean Spaces*, ed. by G.Weiss, S. Wainger, vol. 35. Proceedings of Symposia in Pure Mathematics (American Mathematical Society, Providence, 1979), pp. 383–407
10. M. Cotlar, C. Sadosky, On some L^p versions of the Helson-Szegö theorem, in *Conference on Harmonic Analysis in Honor of Antoni Zygmund*, vols. I, II, Chicago, 1981. Wadsworth Mathematics Series (Belmont, Wadsworth, 1983), pp. 306–317
11. D. Cruz-Uribe, Two weight norm inequalities for fractional integral operators and commutators (2014, Preprint). Available at arXiv:1412.4157v1
12. D. Cruz-Uribe, K. Moen, Sharp norm inequalities for commutators of classical operators. Publ. Mat. **56**, 147–190 (2012)
13. D. Cruz-Uribe, J. Martell, C. Pérez, Sharp weighted estimates for classical operators. Adv. Math. **229**, 408–441 (2012)
14. D. Cruz-Uribe, A. Reznikov, A. Volberg, Logarithmic bump conditions and the two weight boundedness of Calderón-Zygmund operators. Adv. Math. **255**, 706–729 (2014)
15. O. Dragičevič, L. Grafakos, M.C. Pereyra, S. Petermichl, Extrapolation and sharp norm estimates for classical operators in weighted Lebesgue spaces. Publ. Mat. **49**, 73–91 (2005)
16. J. García-Cuerva, J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*. North-Holland Mathematics Studies, vol. 116 (North-Holland, Amsterdam, 1981)

17. T. Hänninen, Two weight inequality for vector-valued positive dyadic operators by parallel stopping cubes. To appear in Isr. J. Math.
18. T. Hänninen, T. Hytönen, K. Li, Two-weight $L^p - L^q$ bounds for positive dyadic operators: unified approach to $p \leq q$ and $p > q$. Potential Anal. **45**(3), 579–608 (2016)
19. H. Helson, G. Szegö, A problem in prediction theory. Ann. Math. Pura. Appl. **51**, 107–138 (1960)
20. I. Holmes, M. Lacey, B. Wick, Commutators in the two-weight setting. Math. Ann. (2016). doi:10.1007/s00208-016-1378-1
21. I. Holmes, M. Lacey, B. Wick, Bloom's inequality: commutators in a two-weight setting. Arch. Math. (Basel) **106**(1), 53–63 (2016)
22. S. Hukovic, S. Treil, A. Volberg, The Bellman function and sharp weighted inequalities for square functions, in *Complex Analysis, Operators and Related Topics*. Operator Theory, Advances and Applications, vol 113 (Birkaüser, Basel, 2000), pp. 97–113
23. R. Hunt, B. Muckenhoupt, R. Wheeden, Weighted norm inequalities for the conjugate function and the Hilbert transform. Trans. Am. Math. Soc. **176**, 227–252 (1973)
24. T. Hytönen, The sharp weighted bound for general Calderón-Sygmund Operators. Ann. Math. (2) **175**(3), 1473–1506 (2012)
25. T. Hytönen, M. Lacey, The $A_p - A_\infty$ inequality for general Calderón-Zygmund operators. Indiana Univ. Math. J. **61**, 2041–2052 (2012)
26. T. Hytönen, K. Li, Weak and strong $A_p - A_\infty$ estimates for square functions and related operators (2015, Preprint). Available at arXiv:1509.00273
27. T. Hytönen, C. Pérez, Sharp weighted bounds involving A_∞. Anal. PDE **6**(4), 777–818 (2013)
28. J. Isralowitz, H.K. Kwon, S. Pott, Matrix weighted norm inequalities for commutators and paraproducts with matrix symbols (2015, Preprint). Available at arXiv:1507.04032
29. T. Iwaniec, A. Verde, On the operator $\mathfrak{L}f = f \log |f|$. J. Funct. Anal. **169**(2), 391–420 (1999)
30. N.H. Katz, M.C. Pereyra, On the two weight problem for the Hilbert transform. Rev. Mat. Iberoam. **13**(01), 211–242 (1997)
31. M. Lacey, *The Linear Bound in A_2 for Calderón-Zygmund Operators: A Survey*. Marcinkiewicz Centenary Volume. Banach Center Publications, vol. 95 (Institute of Mathematics of the Polish Academy of Sciences, Warsaw, 2011), pp. 97—114
32. M. Lacey, The two weight inequality for the Hilbert transform: a primer, in *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, ed. by M.C. Pereyra, et al. Association for Women in Mathematics Series, vol. 5 (Springer, 2017)
33. M. Lacey, Two weight inequality for the Hilbert transform: a real variable characterization, II. Duke Math. J. **163**(15), 2821–2840 (2014)
34. M. Lacey, K. Li, Two weight norm inequalities for the g function. Math. Res. Lett. **21**(03), 521–536 (2014)
35. M. Lacey, K. Li, On $A_p - A_\infty$ estimates for square functions (2015, Preprint). Math. Z. **284**, 1211–1222 (2016)
36. M. Lacey, E. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, The two weight inequality for the Hilbert transform, coronas and energy conditions. Duke Math. J. **163**(15), 2795–2820 (2014)
37. M. Lacey, E. Sawyer, I. Uriarte-Tuero, Two weight inequalities for discrete positive operators (2009, Preprint). Available at arXiv:0911.3437
38. J. Lai, A new 2-weight estimate for a vector-valued positive operator (2015, Preprint). Available at arXiv:1503.06778
39. J. Lai, S. Treil, Two weight L^p estimates for paraproducts in non-homogeneous setting (2015, Preprint). Available at arXiv:1507.05570
40. A. Lerner, Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals. Adv. Math. **226**, 3912–3926 (2011)
41. A. Lerner, Mixed $A_p - A_r$ inequalities for classical singular integrals and Littlewood-Paley operators. J. Geom. Anal. **23**, 1343–1354 (2013)
42. A. Lerner, K. Moen, Mixed $A_p - A_\infty$ estimates with one supremum. Stud. Math. **219**(3), 247–267 (2013)

43. K. Moen, Sharp one-weight and two-weight bounds for maximal operators. Stud. Math. **194**(2), 163–180 (2009)
44. J.C. Moraes, Weighted estimates for dyadic operators with complexity. Ph.D. dissertation, University of New Mexico, 2011
45. J.C. Moraes, M.C. Pereyra, Weighted estimates for dyadic Paraproducts and t-Haar multiplies with complexity (m, n). Publ. Mat. **57**, 265–294 (2013)
46. B. Muckenhoupt, Weighted norm inequalities for the Hardy–Littlewood maximal function. Trans. Am. Math. Soc. **165**, 207–226 (1972)
47. F. Nazarov, A. Reznikov, S. Treil, A. Volberg, A Bellman function proof of the L^2 bump conjecture. J. Anal. Math. **121**, 255–277 (2013)
48. F. Nazarov, A. Reznikov, A. Volberg, Bellman approach to the one-sided bumping for weighted estimates of Calderón–Zygmund operators (2013, Preprint). Available at arXiv:1306.2653
49. F. Nazarov, S. Treil, A. Volberg, The Bellman functions and the two-weight inequalities for Haar Multipliers. J. AMS **12**, 909–928 (1999)
50. F. Nazarov, S. Treil, A. Volberg, Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures (2004, Preprint). Available at arXiv:1003.11596
51. F. Nazarov, S. Treil, A. Volberg, Two weight inequalities for individual Haar multipliers and other well localized operators. Math. Res. Lett. **15**(3), 583–597 (2008)
52. M.C. Pereyra, On the resolvents of dyadic paraproducts. Rev. Mat. Iberoam. **10**(3), 627–664 (1994)
53. M.C. Pereyra, Lecture notes on dyadic harmonic analysis. Contemp. Math. **289**, 1–60 (2001)
54. C. Pérez, E. Rela, A new quantitative two weight theorem for the Hardy-Littlewood maximal operator. Proc. Am. Math. Soc. **143**, 641–655 (2015)
55. E. Sawyer, A characterization of a two weight norm inequality for maximal functions. Stud. Math. **75**(1), 1–11 (1982)
56. E. Sawyer, A characterization of two weight norm inequalities for fractional and Poisson integrals. Trans. Am. Math. Soc. **308**(2), 533–545 (1988)
57. H. Tanaka, A characterization of two weight trace inequalities for positive dyadic operators in the upper triangular case. Potential Anal. **41**(2), 487–499 (2014)
58. S. Treil, A remark on two weight estimates for positive dyadic operators, in *Operator-Related Function Theory and Time-Frequency Analysis Abel Symposia*, vol. 9 (Springer, Cham, 2015), pp. 185–195
59. A. Volberg, A bump theorem for weighted embeddings and maximal operator: the Bellman function approach, in *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, ed. by M.C. Pereyra, et al. Association for Women in Mathematics Series, vol. 5 (Springer, 2017)
60. E. Vuorinen, $L^p(\mu) \to L^q(\nu)$ characterization for well localized operators. J. Fourier Anal. Appl. **22**, 1059 (2016). doi:10.1007/s00041-015-9453-7
61. E. Vuorinen, Two weight L^p-inequalities for dyadic shifts and the dyadic square function (2015, Preprint). Available at arXiv:1504.05759
62. M. Wilson, Weighted inequalities for the dyadic square function without dyadic A_∞. Duke Math. J. **55**, 19–49 (1987)
63. M. Wilson, *Weighted Littlewood-Paley Theory and Exponential-Square Integrability*. Lecture Notes in Mathematics, vol. 1924 (Springer, Berlin, 2008)
64. J. Wittwer, A sharp estimate on the norm of the martingale transform. Math. Res. Lett. **7**, 1–12 (2000)

Potential Operators with Mixed Homogeneity

Calixto P. Calderón and Wilfredo Urbina

Dedicated to the memory of Cora Sadosky

Abstract In 1966 Cora Sadosky introduced a number of results in a remarkable paper "A note on Parabolic Fractional and Singular Integrals", see Sadosky (Studia Math 26:295–302, 1966), in particular, a quasi homogeneous version of Sobolev's immersion theorem was discussed in the paper. Later, C. P. Calderón and T. Kwembe, following those ideas and incorporating the context of Fabes-Riviere homogeneity (Fabes and Riviere, Studia Math 27:19–38, 1966), proved a similar results for potential operators with kernels having mixed homogeneity. Calderón-Kwembe's (Dispersal models. X Latin American School of Mathematics (Tanti, 1991). Rev Un Mat Argent 37(3–4):212–229, 1991/1992) basic theorem was very much in the spirit of Sadosky's result. The natural extension of Sadosky's paper is nevertheless the joint paper by C. Sadosky and M. Cotlar (On quasi-homogeneous Bessel potential operators. In: Singular integrals. Proceedings of symposia in pure mathematics, Chicago, 1966. American Mathematical Society, Providence, 1967, pp 275–287) which constitutes a true tour de force through, what is now considered, local properties of solutions of parabolic partial differential equations. The tools are the introduction of "Parabolic Bessel Potentials" combined with mixed homogeneity local smoothness estimates.

The aim of this paper is to extend Calderón-Kwembe's theorem in two directions: (a) establish a corresponding result in terms of mixed norms in the Benedek-Panzone's sense, see Benedek and Panzone (Duke Math J 28:3–21, 1961). (b) establish results for the case of unbounded characteristics (integrable to the power r on the unit sphere). Calderón-Kwembe's theorem can also be estated in the frame

C.P. Calderón (✉)
Department of Mathematics, University of Illinois at Chicago, Chicago, IL 60607, USA
e-mail: cpc@uic.edu

W. Urbina
Department of Mathematical and Actuarial Sciences, Roosevelt University Chicago, Chicago, IL 60605, USA
e-mail: wurbinaromero@roosevelt.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_6

of generalized homogeneity but that case will not be consider here, see N. Riviere (Arkiv för Math 9(2):243–278, 1971) and A. P. Calderón and A. Torchinsky (Adv Math 16:1–64, 1975).

2010 *Mathematics Subject Classification.* Primary 47G05; Secondary 45P05.

Introduction

In 1966 Cora Sadosky introduced a number of results in a remarkable paper "A note on Parabolic Fractional and Singular Integrals", see [8], in particular, a quasi homogeneous version of Sobolev's immersion theorem was discussed in the paper. Later C. P. Calderón and T. Kwembe, following those ideas and incorporating the context of Fabes-Riviere homogeneity [5], proved a similar results for potential operators with kernels having mixed homogeneity. Calderón-Kwembe's basic theorem was very much in the spirit of Sadosky's result. The natural extension of Sadosky's paper is nevertheless the joint paper by C. Sadosky and M. Cotlar [9] which constitutes a true tour de force through, what is now considered, local properties of solutions of parabolic partial differential equations. The tools are the introduction of "Parabolic Bessel Potentials" combined with mixed homogeneity local smoothness estimates.

The aim of this paper is to extend Calderón-Kwembe's theorem in two directions: (a) establish a corresponding result in terms of mixed norms in the Benedek-Panzone's sense, see [1]. (b) establish results for the case of unbounded characteristics (integrable to the power r on the unit sphere). Calderón-Kwembe's theorem can also be estated in the frame of generalized homogeneity but that case will not be consider here, see N. Riviere [7] and A.P. Calderón and A. Torchinsky [2].

In what follows C is a positive real constant that may change from line to line.

A Change of Variable of Polar Type

Let $a_1, \cdots, a_n$ (fixed) real numbers, $a_j \geq 1$ and let $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, consider the function

$$F(x, \rho) = \sum_{j=1}^{n} \frac{x_j^2}{\rho^{2a_j}}, \tag{1}$$

for a fixed x. $F(x, \rho)$ is a continuous decreasing function of ρ ($\rho > 0$) and therefore there is a unique solution of $F(x, \rho) = 1$ that will be denoted as $\rho(x)$. It defines a distance in $\mathbb{R}^n$, see [9] which is called *parabolic distance*,

$$\rho(x, y) = \rho(x - y).$$

Then, we have,

$$(\frac{x_1}{\rho^{a_1}(x)}, \cdots, \frac{x_n}{\rho^{a_n}(x)}) \in \Sigma_n = \{y \in \mathbb{R}^n : |y| = 1\}.$$

The point x can be written (change of variables of polar type) as

$$\begin{aligned}
x_1 &= \rho^{a_1} \cos\phi_1 \cos\phi_2 \cdots \cos\phi_{n-1} = \rho^{a_1} \prod_{i=1}^{n-1} \cos\phi_i = \rho^{a_1} \Psi_1(\phi) \\
x_2 &= \rho^{a_2} \sin\phi_1 \cos\phi_2 \cdots \cos\phi_{n-1} = \rho^{a_2} \sin\phi_1 \prod_{i=2}^{n-1} \cos\phi_i = \rho^{a_2} \Psi_2(\phi) \\
x_3 &= \rho^{a_3} \sin\phi_2 \cos\phi_3 \cdots \cos\phi_{n-1} = \rho^{a_3} \sin\phi_2 \prod_{i=3}^{n-1} \cos\phi_i = \rho^{a_3} \Psi_3(\phi) \\
\cdots & \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_{n-1} &= \rho^{a_{n-1}} \sin\phi_{n-2} \cos\phi_{n-1} = \rho^{a_{n-1}} \Psi_{n-1}(\phi) \\
x_n &= \rho^{a_n} \sin\phi_{n-1} = \rho^{a_n} \Psi_n(\phi).
\end{aligned}$$

From the construction, it follows

$$\sum_{i=1}^{n} \Psi_i^2(\phi) = 1.$$

Then, the element of volume is

$$dx = \rho^{|a|-1} J(\phi_1, \cdots, \phi_{n-1}) d\rho d\sigma, \tag{2}$$

where $a = (a_1, \cdots, a_n)$ and $|a| = a_1 + \cdots + a_n$, see [5].

Notice that $|x_i| \leq \rho^{a_i}$ and therefore $\rho \geq |x_i|^{1/a_i}$. On the other hand, since $a_i \geq 1$, we have

$$\begin{aligned}
|x_i|^{1/a_i} &= \rho |\sin\phi_{i-1} \cos\phi_i \cdots \cos\phi_{n-1}|^{1/a_i} \\
&\geq \rho |\sin\phi_{i-1} \cos\phi_i \cdots \cos\phi_{n-1}|.
\end{aligned}$$

This will be used below.

We shall use in this paper the following distance,

$$\overline{\rho}(x) = \sum_{i=1}^{n} |x_i|^{1/a_i}.$$

The fact that $\overline{\rho}$ is actually a distance, follows immediately. The triangle inequality is obtained since $a_i \geq 1$ and hence $\frac{1}{a_i} \leq 1$, consequently,

$$\overline{\rho}(x+y) = \sum_{i=1}^{n} |x_i + y_i|^{1/a_i} \leq \sum_{i=1}^{n} |x_i|^{1/a_i} + \sum_{i=1}^{n} |y_i|^{1/a_i} = \overline{\rho}(x) + \overline{\rho}(y).$$

From the fact that $\mathbb{R}^n$ is a finite dimensional vector space, we know, that ρ and $\overline{\rho}$ are equivalent. For the sake of completeness, the equivalence will be given here explicitly,

$$\frac{|x_i|}{\rho^{a_i}} \leq 1, \quad \text{i.e.} \quad \rho^{a_i} \geq |x_i|,$$

then $\rho \geq |x_i|^{1/a_i}$, and therefore, $\rho \geq \frac{1}{n} \sum_{i=1}^{n} |x_i|^{1/a_i}$.

On the other hand,

$$|x_i|^{1/a_i} = \rho |\Psi_i(\phi)|^{1/a_i} \geq \rho |\Psi_i(\phi)|,$$

hence,

$$\sum_{i=1}^{n} |x_i|^{2/a_i} \geq \rho^2,$$

and therefore,

$$\rho \leq \sum_{i=1}^{n} |x_i|^{1/a_i}.$$

An equivalent metric was also considered in C. Sadosky [8] and C. Sadosky & M. Cotlar [9] , when $a_1, \cdots, a_n$ are rational numbers,

$$r(x) = (\sum_{i=1}^{n} |x_i|^{m/a_i})^{1/m},$$

where m is the smallest integer that is divisible by $2a_i,\ i = 1, 2, \cdots, n$.

Additionally, P. Krée [6] considered another equivalent distance,

$$\bar{r}(x) = \max_{1\le i\le n} |x_i|^{1/a_i}.$$

Main Results

Given an operator

$$Tf(x) = \int_{\mathbb{R}^n} K(x-y)f(y)dy, \tag{3}$$

We want to study the conditions on the kernel K such that it makes the operator T to be $L^p - L^q$ bounded for suitable values of p and q.

Potential Operators of Mixed Homogeneity

Given kernel K, we say it is a potential kernel with *mixed homogeneity* if

$$K(\lambda^{a_1}x_1, \lambda^{a_2}x_2, \cdots, \lambda^{a_n}x_n) = \lambda^{\beta-|a|}K(x), \tag{4}$$

for $0 < \beta < |a| = \sum_{i=1}^n a_i,\ a_i \ge 1$.

If x is expressed in polar type coordinates, $x = \rho^a\Psi(\phi)$, as above, then

$$K(x_1, x_2, \cdots, x_n) = \rho^{\beta-|a|}K(\phi). \tag{5}$$

For the case when K can be written as

$$K(x_1, x_2, \cdots, x_n) = \frac{\Omega(\phi)}{\rho^{|a|-\beta}}, \text{ with } \|\Omega\|_{L^\infty(\sigma_n)} < M, \tag{6}$$

C. P. Calderón and T. Kwembe [3], proved that

$$\|Tf\|^q \le C\|f\|_p, \quad \text{for} \quad \frac{1}{q} = \frac{1}{p} - \frac{\beta}{|a|}. \tag{7}$$

As we have already said, the aim of this paper is to extend Calderón-Kwembe's theorem in two directions:

(i) Establish a corresponding result in terms of mixed norms in the Benedek-Panzone's sense, [1].

(ii) Establish results for the case of unbounded characteristics (integrable to the power r on the unit sphere).

As it was mentioned in the introduction, Calderón-Kwembe's theorem can also be extended to potential operators with generalized homogeneity, but that case will not be consider here, see N. Riviere [7] and A.P. Calderón and A. Torchinsky [2].

Estimates of Benedek-Panzone [1], Du Plessis [4]

If the kernel K satisfies $\|K(\phi)\|_\infty < M$, then

$$|K(x_1, x_2, \cdots, x_n)| \leq C\rho^{-|a|+\beta}\|K(\phi)\|_\infty, \tag{8}$$

for $\sum_{j=1}^n \beta_j = \beta,\ \beta_j = \theta_j a_j,\ 0 < \theta_j < 1$. Then, one has the estimate,

$$\begin{aligned}\frac{1}{(\sum_{i=1}^n |x_i|^{1/a_i})^{|a|-\beta}} &= \frac{1}{\prod_{j=1}^n (\sum_{i=1}^n |x_i|^{1/a_i})^{a_j-\beta_j}} \\ &\leq \prod_{j=1}^n \frac{1}{|x_j|^{1/a_j(a_j-\beta_j)}} = \prod_{j=1}^n \frac{1}{|x_j|^{(1-\theta_j)}}.\end{aligned}$$

For each j we have a regular (one) dimensional potential operator, that maps L^{p_j} into L^{q_j} with $\frac{1}{q_j} = \frac{1}{p_j} - \theta_j$.

We will denote $L_{p_1,p_2,\cdots,p_n}$ the space of measurable functions in $\mathbb{R}^n$, such that

$$\|f\|_{p_1,p_2,\cdots,p_n} = (\int_{-\infty}^{\infty}(\int_{-\infty}^{\infty}\cdots(\int_{-\infty}^{\infty}|f(x_1,x_2,\cdots,x_n)|^{p_1}dx_1)^{p_2/p_1}dx_2)^{p_3/p_2}\cdots)^{p_n/p_{n-1}}dx_n)^{1/p_n} < \infty,$$

see Benedek-Panzone [1], page 301.

Theorem 2.1 *The operator defined as $Tf = K * f$, where K satisfies (4); maps continuously $L_{p_1,p_2,\cdots,p_n}$ into $L_{q_1,q_2,\cdots,q_n}$, with q_i given by $\frac{1}{q_i} = \frac{1}{p_i} - \theta_i$ and $\beta = \sum_{j=1}^n \theta_j a_j$.*

Proof It will suffice to show it for the case $n = 2$, the general case follows by induction. Let $f \in L_{p_1,p_2}$ and $g \in L_{q_1^*,q_2^*}$, q_i^* being the dual exponents of q_i. By the potential theorem in dimension one applied twice,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{|x_1-y_1|^{1-\theta_1}}\frac{1}{|x_2-y_2|^{1-\theta_2}}f(x_1,x_2)g(y_1,y_2)dx_1dx_2dy_1dy_2$$

$$= \int_{-\infty}^{\infty} \frac{1}{|x_2 - y_2|^{1-\theta_2}} \Big(\int_{-\infty}^{\infty} \frac{1}{|x_1 - y_1|^{1-\theta_1}} f(x_1, x_2) g(y_1, y_2) dx_1 dy_1 \Big) dx_2 dy_2$$
$$= \int_{-\infty}^{\infty} \frac{1}{|x_2 - y_2|^{1-\theta_2}} \|f(\cdot, x_2)\|_{p_1} \|g(\cdot, y_2)\|_{q_1^*} dx_2 dy_2$$
$$\le \|f\|_{p_1, p_2} \|g\|_{q_1^*, q_2^*},$$

using the duality in space with mixed L^p-norms, we get the thesis, see Benedek-Panzone [1], page 303. □

It should be noted that the above proof follows closely the corresponding one in Benedek-Panzone [1], pages 321–322.

Potential Operators with Not Bounded Characteristic

Consider kernels of the form

$$K(x_1, x_2, \cdots, x_n) = \rho^{-|a|+\beta} K(\phi), \tag{9}$$

with $|a| = a_1 + \cdots + a_n;\ a_i \ge 1;\ 0 < \beta < |a|$. $K(\phi)$ is considered no longer bounded on $\Sigma = \{x : \rho(x) = 1\}$ but instead we will assume that on Σ we have

$$\int_{\Sigma} |K(\phi)|^r d\sigma < \infty, \tag{10}$$

for some r; $1 < r < \infty$. Here $d\sigma$ stands for the element of "surface" on the unit sphere of $\mathbb{R}^n$.

Theorem 2.2 *Let the kernel K be such that (10) holds with $0 < \frac{1}{r} + \frac{\beta}{|a|} \le 1$. If $f \in L^p(\mathbb{R}^n)$ and $\frac{1}{r} = \frac{1}{p} - \frac{\beta}{|a|}$, then calling $T(f) = (K * f)$ we have*

(i)

$$|\{x : |Tf(x)| > \lambda\}| < \frac{C}{\lambda^r} \int_{\Sigma} |\Psi(\alpha)|^r d\sigma \|f\|_p^r, \tag{11}$$

here $\Psi(\alpha) = K(\alpha)$.

(ii) For values of $p \ge 1$ such that $\frac{1}{p} - \frac{\beta}{|a|} > \frac{1}{r_0}$ we have

$$\|K * f\|_r \le C\|f\|_p,$$

for $\frac{1}{r} = \frac{1}{p} - \frac{\beta}{|a|}$, whenever $\int_{\Sigma} |K(\phi)|^{r_0} d\sigma < \infty$.

Proof (ii) follows from (i) by application of Marcinkiewicz interpolation theorem.

The part (i) of the thesis will follow after few steps. We decompose the kernel $K = K_1 + K_2$, where $K_2 = K$ if $x \in A$ where,

$$A = \{x = (\rho, \phi) : \rho(x) > (\frac{\Psi(\phi)}{\lambda})^{\delta}\},$$

(for an appropriated value of δ to be determine later) and $K_2 = 0$ otherwise; and $K_1 = K - K_2$. For this selection $\delta = \frac{p}{|a|-\beta p}$.

Estimates for $K * f$. Let $p > 1$ and q its conjugate exponent; $\frac{1}{p} + \frac{1}{q} = 1$. We will use the point estimate,

$$|K_2 * f| \leq (\int |K_2|^q dx)^{1/q} \cdot \|f\|_p \tag{12}$$

We will consider $\|f\|_p = 1$ and move to the evaluation of $(\int |K_2|^q dx)^{1/q}$. Using Fubini's theorem, we set

$$\Big(\int_{\Sigma} |\Psi(\alpha)|^q d\sigma \int_{\{\rho(x) > (\frac{\Psi(\phi)}{\lambda})^{\delta}} \frac{1}{\rho^{(|a|-\beta)q}} \rho^{|a|-1} d\rho\Big)^{1/q}. \tag{13}$$

By using the value of δ from above we have

$$\delta[-(|a| - \beta)q + |a|] = -q.$$

Thus the inner integral above is

$$\int_{\frac{\Psi(\phi)}{\lambda}}^{\infty} \frac{1}{\rho^{(|a|-\beta)q}} \rho^{|a|-1} d\rho = C\lambda^q \Psi(\phi)^{-q} \tag{14}$$

then, (13) is immediately seen to be equal to

$$C\lambda^q \int_{\Sigma} |\Psi|^q |\Psi|^{-q} d\sigma, \tag{15}$$

taking now the q root, we get

$$|K_2 * f| \leq C\lambda.$$

We pass now to the estimate of $|K_1 * f|$. Using Young's inequality for the convolution, we obtain

$$|\{x : |K_1 * f|(x) > \lambda\}| \leq \frac{C}{\lambda^p} \|K_1\|_1^p \|f\|_p^p. \tag{16}$$

Using “polar” coordinates and Fubini’s theorem, we get for (16)

$$C\int_{\Sigma}|\frac{\Psi(\alpha)}{\lambda}|(|\frac{\Psi(\alpha)}{\lambda}|)^{\delta(|a|-(|a|-\beta)}\,d\sigma \tag{17}$$

Recall that for this case $\delta = \frac{p}{|a|-\beta p}$, so the exponent for the integrand above is

$$1+\frac{p\beta}{|a|-p\beta}=\frac{|a|}{|a|-p\beta}$$

Thus, the above integral is dominated by

$$C\frac{1}{\lambda^{p|a|/(|a|-p\beta)}}(\int_{\Sigma}|\Psi(\alpha)|^{|a|/(|a|-p\beta)}d\sigma)^{p}. \tag{18}$$

Using Jensen’s inequality for the integral above and remembering that

$$\frac{1}{r}=\frac{1}{p}-\frac{\beta}{|a|}=\frac{|a|-p\beta}{p|a|},$$

we get

$$C\frac{1}{\lambda^{r}}(\int_{\Sigma}|\Psi(\alpha)|^{r}d\sigma)^{p}=C\frac{1}{\lambda^{r}}\|\Psi\|_{r}^{r}. \tag{19}$$

From the constructions above and keeping in mind that $\|f\|_p = 1$, we have for a large but fixed C,

$$\{x:|K*f|(x)>C\lambda\}\subset\{x:|K_1*f|(x)>C\lambda/2\}\cup\{x:|K_2*f|(x)>C\lambda/2\}$$

With the appropriate selection of C we get

$$\{x:K_2*f(x)>C\lambda/2\}=\emptyset.$$

On the other hand,

$$|\{x:|K_1*f|(x)>\lambda\}|\le\frac{C}{\lambda^{r}}\|\Psi\|_{r}^{r}. \tag{20}$$

An homogeneity argument gives (i) for the case $p > 1$.

The limit case $p = 1$. Consider as above the decomposition $K = K_1 + K_2$. In this case,

$$\frac{1}{r}=1-\frac{\beta}{|a|},$$

the choice of δ is

$$\delta = \frac{1}{|a| - \beta}.$$

Consider now $f(x-y) = \bar{f}(y)$. A pointwise estimate gives,

$$|K_2 * \bar{f}| \leq \int_\Sigma |\psi(\alpha)| d\sigma \int_{\{\rho(y) > (\frac{\Psi(\alpha)}{\lambda})^\delta\}} \frac{1}{\rho^{|a|-\beta}} \bar{f}(\rho,\phi) \rho^{|a|-1} d\rho.$$

The above expression does not exceed

$$\int_\Sigma |\psi(\alpha)| d\sigma (\frac{\Psi(\alpha)}{\lambda})^{\delta(|a|-\beta)} \int_0^\infty \bar{f}(\rho,\phi)\rho^{|a|-1} d\rho \leq C\lambda.$$

Evaluation of $|K_1 * f|$. Remember that $\|f\|_1 = 1$, using Young's Theorem

$$|\{x : (K_1 * f)(x) > \lambda\}| \leq \frac{C}{\lambda} \|K_1\|_1 \|f\|_1 = \frac{C}{\lambda} \|K_1\|_1. \tag{21}$$

The evaluation of $\frac{1}{\lambda}\|K_1\|_1$ is obtained using Fubini's theorem

$$\begin{aligned} \frac{C}{\lambda} \int_\Sigma \Psi(\alpha) d\sigma \int_0^{(\frac{\Psi(\alpha)}{\lambda})^\delta} \frac{1}{\rho^{|a|-\beta}} \rho^{|a|-1} d\rho &= \frac{1}{\lambda^{1+\beta\delta}} \int_\Sigma \Psi(\alpha)^{1+\beta\delta} d\sigma \\ &= \frac{C}{\lambda^r} \int_\Sigma |\Psi(\alpha)|^r d\sigma. \end{aligned}$$

The proof now follows from the pattern from the previous case. □

References

1. A. Benedek, R. Panzone, The space L^p, with mixed norm. Duke Math. J. **28**, 3–21 (1961)
2. A.P. Calderón, A. Torchinsky, Parabolic maximal functions associated with a distribution. Adv. Math. **16**, 1–64 (1975)
3. C.P. Calderón, T. Kwembe, Dispersal models. X Latin American School of Mathematics (Tanti, 1991). Rev. Un. Mat. Argent. **37**(3–4), 212–229 (1991/1992)
4. N. Du Plessis, Some theorems about Riesz fractional integrals. Trans. Am. Math. **80**, 124–134 (1955)
5. E.B. Fabes, N.M. Riviere, Singular integrals with mixed homogeneity. Studia Math. **27**, 19–38 (1966)
6. P. Krée, Distributions quasihomogènes. C. R. Acad. Sci. Paris **261**, 2560–2563 (1965)
7. N.M. Riviere, Singular integrals and multiplier operators. Arkiv för Math. **9**(2), 243–278 (1971)

8. C. Sadosky, A note on parabolic fractional and singular integrals. Studia Math. **26**, 295–302 (1966)
9. C. Sadosky, M. Cotlar, On quasi-homogeneous Bessel potential operators, in *Singular Integrals*. Proceedings of Symposia in Pure Mathematics, Chicago, 1966 (American Mathematical Society, Providence, 1967), pp. 275–287

Elementary Proofs of One Weight Norm Inequalities for Fractional Integral Operators and Commutators

David Cruz-Uribe, OFS

Dedicated to the memory of Professor Cora Sadosky

Abstract We give new and elementary proofs of one weight norm inequalities for fractional integral operators and commutators. Our proofs are based on the machinery of dyadic grids and sparse operators used in the proof of the A_2 conjecture.

2010 *Mathematics Subject Classification.* 42B25, 42B30, 42B35

Introduction

The fractional integral operators, also called the Riesz potentials, are the convolution operators

$$I_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy, \qquad 0 < \alpha < n.$$

These operators are classical and for $1 < p < \frac{n}{\alpha}$ and q defined by $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$, satisfy $I_\alpha : L^p \to L^q$. When $p = 1$ they satisfy the endpoint estimate $I_\alpha : L^1 \to L^{q,\infty}$. (Cf. Stein [24].) One weight norm inequalities for these operators were first considered by Muckenhoupt and Wheeden [19], who introduced the governing class of weights, $A_{p,q}$. For $1 < p < \frac{n}{\alpha}$ and q such that $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$, a weight (i.e., a non-negative, locally integrable function) w is in $A_{p,q}$ if

The author is supported by NSF grant 1362425 and the Dean of the College of Arts & Sciences at UA. As this paper was being written he was also supported by the Stewart-Dorwart faculty development fund at Trinity College.

D. Cruz-Uribe (✉)
Department of Mathematics, University of Alabama, Tuscaloosa, AL 35487-0350, USA
e-mail: dcruzuribe@ua.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_7

$$[w]_{A_{p,q}} = \sup_Q \left(\fint_Q w^q \, dx \right)^{\frac{1}{q}} \left(\fint_Q w^{-p'} \, dx \right)^{\frac{1}{p'}} < \infty,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$. When $p = 1$, $q = \frac{n}{n-\alpha}$, we say $w \in A_{1,q}$ if

$$[w]_{A_{1,q}} = \sup_Q \operatorname*{ess\,sup}_{x \in Q} \left(\fint_Q w^q \, dx \right)^{\frac{1}{q}} w(x)^{-1} < \infty.$$

Muckenhoupt and Wheeden showed that when $p > 1$, $I_\alpha : L^p(w^p) \to L^q(w^q)$ if and only if $w \in A_{p,q}$, and when $p = 1$, $I_\alpha : L^1(w) \to L^{q,\infty}(w^q)$ when $w \in A_{1,q}$. Their proof used a good-λ inequality relating I_α and the fractional maximal operator,

$$M_\alpha f(x) = \sup_Q |Q|^{\frac{\alpha}{n}} \fint_Q |f(y)| \, dy \cdot \chi_Q(x).$$

Weighted norm inequalities for M_α were proved by generalizing the earlier results for the Hardy-Littlewood maximal operator. A different proof of the strong type inequality was given in [7]: there they used Rubio de Francia extrapolation to prove a norm inequality relating I_α and M_α.

Given $b \in BMO$ we define the commutator of a fractional integral by

$$[b, I_\alpha]f(x) = b(x)I_\alpha f(x) - I_\alpha(bf)(x) = \int_{\mathbb{R}^n} \big(b(x) - b(y)\big) \frac{f(y)}{|x-y|^{n-\alpha}} \, dy.$$

These commutators were introduced by Chanillo [2], who proved that with p and q defined as above, $[b, I_\alpha] : L^p \to L^q$. He also proved that when p is an even integer, $b \in BMO$ is a necessary condition. (The necessity for the full range of p was recently shown by Chaffee [1].) Commutators are more singular than the fractional integral operator: this can be seen by the fact that when $p = 1$, the do not map L^1 into $L^{q,\infty}$. For a counter-example and substitute endpoint estimate, see [6]. In this paper it was also shown that the strong type inequality is governed by $A_{p,q}$ weights: if $1 < p < \frac{n}{\alpha}$ and $w \in A_{p,q}$, then $[b, I_\alpha] : L^p(w^p) \to L^q(w^q)$. This proof relied on a sharp maximal function estimate relating the commutator, I_α, and M_α. A different proof using extrapolation to relate the commutator to an Orlicz fractional maximal operator was given in [7]. Yet another proof, one that gave the sharp constant in terms of the $[w]_{A_{p,q}}$ characteristic, was given in [9]. This proof used a Cauchy integral formula argument due to Chung et al. [4].

In this paper we give new and elementary proofs of the one weight inequalities for fractional integral operators and commutators. More precisely, we prove the following three theorems.

Theorem 1.1 *Given* $0 < \alpha < n$, *let* $q = \frac{n}{n-\alpha}$. *If* $w \in A_{1,q}$, *then for all* $f \in L^1(w)$,

$$\sup_{t>0} t\, w^q(\{x \in \mathbb{R}^n : |I_\alpha f(x)| > t\})^{\frac{1}{q}} \leq C(n,\alpha)[w]_{A_{1,q}} \int_{\mathbb{R}^n} |f(x)|w(x)\,dx.$$

Theorem 1.2 *Given* $0 < \alpha < n$, $1 < p < \frac{n}{\alpha}$, *let* q *be such that* $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$. *If* $w \in A_{p,q}$, *then for all* $f \in L^p(w^p)$,

$$\left(\int_{\mathbb{R}^n} |I_\alpha f(x)|^q w(x)^q\,dx\right)^{\frac{1}{q}} \leq C(n,p,\alpha)[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \left(\int_{\mathbb{R}^n} |f(x)|^p w(x)^p\,dx\right)^{\frac{1}{p}}.$$

Theorem 1.3 *Given* $0 < \alpha < n$, $1 < p < \frac{n}{\alpha}$, *let* q *be such that* $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$. *If* $w \in A_{p,q}$, *then for all* $f \in L^p(w^p)$ *and* $b \in BMO$,

$$\left(\int_{\mathbb{R}^n} |[b,I_\alpha]f(x)|^q w(x)^q\,dx\right)^{\frac{1}{q}}$$
$$\leq C(n,p,\alpha)[w]_{A_{p,q}}^{\max(p',q)+1+\frac{q}{p'}+\frac{p'}{p}} \|b\|_{BMO} \left(\int_{\mathbb{R}^n} |f(x)|^p w(x)^p\,dx\right)^{\frac{1}{p}}.$$

We prove Theorems 1.1, 1.2, and 1.3 using the machinery of dyadic grids and sparse operators. Dyadic fractional integral operators date back to the work of Sawyer and Wheeden [23]. More recently, using the machinery developed as part of the proof of the A_2 conjecture for singular integral operators (see [15, 18] and the references they contain) dyadic fractional integral operators were further developed and applied to commutators in [9–11]. The advantage of this approach is its simplicity: it avoids extrapolation, good-λ inequalities and comparisons to the fractional maximal operator. One weakness of our proofs is that they do not give sharp dependence on the $A_{p,q}$ characteristic of the weights in Theorems 1.2 and 1.3: this is to be expected since we freely use their properties to simplify the proofs, whereas any sharp constant proof must be arranged to use their properties as few times as possible. We record the precise constant we obtain to highlight where we use their properties. Sharp constants for the fractional integral operator are given in [17], and for commutators in [9].

The remainder of this paper is organized as follows. In section "Preliminary Results" we give some preliminary results about dyadic grids, sparse operators, weighted fractional maximal operators, and weights. In section "Proof of Theorem 1.1" we prove Theorem 1.1. Our proof adapts an argument that seems to have been part of the folklore of harmonic analysis. We want to thank the anonymous referee for suggesting this approach; it is much simpler than our original proof, which adapted Sawyer's proof of two weight weak (p,q) inequalities for fractional integrals [22]. In section "Proof of Theorem 1.2" we prove Theorem 1.2. Our proof uses ideas of Pérez [20] from his proof of two weight inequalities for fractional integrals, and from the elementary proof of one weight inequalities for the Hardy-Littlewood maximal operator due to Christ and Fefferman [3]. In section "Proof of Theorem 1.3" we prove Theorem 1.3. Our proof uses some ideas from the proof of two weight results in [9] to reduce the problem to an estimate essentially the same as the one for the fractional integral in the previous section. Finally, in section "Tres Recuerdos de Cora Sadosky" we give some personal recollections about the late Cora Sadosky.

Throughout this paper notation is standard or will be defined as needed. The constant n will always denote the dimension. We will denote constants by C, c, etc. and their value may change at each appearance. Unless otherwise specified, we will assume that constants can depend on p, α and n but we will keep track of the dependence on the $A_{p,q}$ characteristic explicitly.

Preliminary Results

Dyadic Grids and Operators

We begin by defining dyadic grids and the dyadic fractional integral operators. Unless otherwise noted, the results given here are taken from [5, 9–11].

Definition 2.1 A collection of cubes $\mathcal{D}$ in $\mathbb{R}^n$ is a dyadic grid if provided that:

(1) If $Q \in \mathcal{D}$, then $\ell(Q) = 2^k$ for some $k \in \mathbb{Z}$.
(2) If $P, Q \in \mathcal{D}$, then $P \cap Q \in \{P, Q, \emptyset\}$.
(3) For every $k \in \mathbb{Z}$, the cubes $\mathcal{D}_k = \{Q \in \mathcal{D} : \ell(Q) = 2^k\}$ form a partition of $\mathbb{R}^n$.

Definition 2.2 Given a dyadic grid $\mathcal{D}$, a set $\mathcal{S} \subset \mathcal{D}$ is sparse if for every $Q \in S$,

$$\Bigg| \bigcup_{\substack{P \in S \\ P \subsetneq Q}} P \Bigg| \leq \frac{1}{2}|Q|.$$

Equivalently, if we define

$$E(Q) = Q \setminus \bigcup_{\substack{P \in S \\ P \subsetneq Q}} P,$$

then the sets $E(Q)$ are pairwise disjoint and $|E(Q)| \geq \frac{1}{2}|Q|$.

The classic example of a dyadic grid and sparse families are the standard dyadic grid on $\mathbb{R}^n$ and the Calderón-Zygmund cubes associated with an L^1 function. See [8, Appendix A].

We now define a dyadic version of the fractional integral operator and show that it can be used to bound I_α pointwise. For $f \in L^1_{loc}$ and a cube Q, let

$$\langle f \rangle_Q = \fint_Q f(x)\,dx = \frac{1}{|Q|}\int_Q f(x)\,dx.$$

Given $0 < \alpha < n$ and a dyadic grid $\mathcal{D}$, define

$$I_\alpha^{\mathcal{D}} f(x) = \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \langle f \rangle_Q \cdot \chi_Q(x).$$

Similarly, given a sparse subset $\mathcal{S} \subset \mathcal{D}$, define

$$I_\alpha^{\mathcal{S}} f(x) = \sum_{Q\in\mathcal{S}} |Q|^{\frac{\alpha}{n}} \langle f \rangle_Q \cdot \chi_Q(x).$$

Lemma 2.3 *There exists a collection* $\{\mathcal{D}_i\}_{i=1}^N$ *of dyadic grids such that for* $0 < \alpha < n$ *and every non-negative function* f*,*

$$I_\alpha f(x) \le C \sup_i I_\alpha^{\mathcal{D}_i} f(x).$$

Moreover, given any dyadic grid $\mathcal{D}$ *and a non-negative function* $f \in L_c^\infty$*, there exists a sparse set* $\mathcal{S} \subset \mathcal{D}$ *such that*

$$I^{\mathcal{D}} f(x) \le C I^{\mathcal{S}} f(x).$$

Given $b \in BMO$ and $0 < \alpha < n$, for any dyadic grid $\mathcal{D}$, define the dyadic commutator

$$C_b^{\mathcal{D}} f(x) = \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \fint_Q |b(x) - b(y)| f(y)\,dy \cdot \chi_Q(x).$$

Lemma 2.4 *There exists a collection* $\{\mathcal{D}_i\}_{i=1}^N$ *of dyadic grids such that for* $0 < \alpha < n$*, every non-negative function* f*, and every* $b \in BMO$*,*

$$|[b, I_\alpha] f(x)| \le C \sup_i C_b^{\mathcal{D}^i} f(x).$$

Remark 2.5 It follows at once from Lemmas 2.3 and 2.4 that to prove norm inequalities for I_α and $[b, I_\alpha]$ it will suffice to prove them for their dyadic counterparts. Moreover, since these dyadic integral operators are positive, we may assume that f is non-negative in our proofs. Finally, by Fatou's lemma it will suffice to prove our results for functions $f \in L_c^\infty$. In particular, this will let us pass to sparse operators.

Weighted Fractional Maximal Operators

We begin with some basic facts about Orlicz spaces. Some of these will also be needed in section "Proof of Theorem 1.3" below. For further information on these spaces, see [21]; for their use in weighted norm inequalities, see [8]. Given a weight σ, let $d\sigma = \sigma\, dx$. We define averages with respect to the measure $d\sigma$:

$$\langle f\rangle_{Q,\sigma} = \fint_Q f(x)\, d\sigma = \frac{1}{\sigma(Q)} \int_Q f(x)\, d\sigma.$$

Given a Young function Φ and a cube Q, define the normalized Luxemburg norm with respect to Φ and $d\sigma$ by

$$\|f\|_{\Phi,Q,\sigma} = \inf\left\{\lambda > 0 : \fint_Q \Phi\left(\frac{|f(x)|}{\lambda}\right) d\sigma \leq 1\right\}.$$

If we let $\Phi(t) = t^p$, $1 \leq p < \infty$, then

$$\|f\|_{\Phi,Q,\sigma} = \left(\fint_Q |f(x)|^p\, d\sigma\right)^{\frac{1}{p}} = \|f\|_{p,Q,\sigma}.$$

Associated to any Young function is its associate function $\bar{\Phi}$. We have the generalized Hölder's inequality: for any cube Q,

$$\fint_Q |f(x)g(x)|\, d\sigma \leq C\|f\|_{\Phi,Q,\sigma}\|g\|_{\bar{\Phi},Q,\sigma};$$

the constant depends only on Φ.

Hereafter, we will let $\Phi(t) = t\log(e+t)$; then it can be shown that $\bar{\Phi}(t) \approx e^t - 1$. It follows for this choice of Φ that for $1 < p < \infty$,

$$\|f\|_{1,Q,\sigma} \leq \|f\|_{\Phi,Q,\sigma} \leq C(p)\|f\|_{p,Q,\sigma}.$$

We now define a weighted dyadic fractional maximal operator. Given a dyadic grid $\mathcal{D}$ and a weight σ, for $0 \leq \alpha < n$ define

$$M^{\mathcal{D}}_{\Phi,\sigma,\alpha} f(x) = \sup_{Q\in\mathcal{D}} \sigma(Q)^{\frac{\alpha}{n}} \|f\|_{\Phi,Q,\sigma}.$$

If $\Phi(t) = t$ we write $M^{\mathcal{D}}_{\sigma,\alpha}$. If $\alpha = 0$, then we simply write $M^{\mathcal{D}}_{\Phi,\sigma}$ or $M^{\mathcal{D}}_{\sigma}$ if $\Phi(t) = t$.

Lemma 2.6 *Let $\Phi(t) = t\log(e+t)$. Given $1 < p < \infty$ and $0 \leq \alpha < n$, define q by $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$. Given a weight σ and a dyadic grid $\mathcal{D}$, $M^{\mathcal{D}}_{\Phi,\sigma,\alpha} : L^p(\sigma) \to L^q(\sigma)$. The same inequality holds for $M^{\mathcal{D}}_{\sigma,\alpha}$.*

Proof This result is well-known when $\Phi(t) = t$; the proof is essentially the same for $\Phi(t) = t\log(e+t)$ and we sketch the details. By off-diagonal Marcinkiewicz interpolation (see [25]) it will suffice to prove the corresponding weak (p, q) inequality:

$$\sigma(\{x \in \mathbb{R}^n : M^{\mathcal{D}}_{\Phi,\sigma,\alpha}f(x) > t\})^{\frac{1}{q}} \leq \frac{C}{t}\left(\int_{\mathbb{R}^n} |f(x)|^p\, d\sigma\right)^{\frac{1}{p}}.$$

Fix $t > 0$; then we can decompose the level set as the union of disjoint cubes $Q \in \mathcal{Q}_t \subset \mathcal{D}$ that satisfy

$$\sigma(Q)^{\frac{\alpha}{n}} \|f\|_{\Phi,Q,\sigma} > t.$$

Therefore, since the cubes in $\mathcal{Q}_t$ are disjoint and $q/p \geq 1$, we have that

$$\begin{aligned}
\sigma(\{x \in \mathbb{R}^n : M^{\mathcal{D}}_{\Phi,\sigma,\alpha}f(x) > t\}) &= \sum_{Q\in\mathcal{Q}_t} \sigma(Q) \\
&\leq t^{-q} \sum_{Q\in\mathcal{Q}_t} \sigma(Q)^{1+q\frac{\alpha}{n}} \|f\|^q_{\Phi,Q,\sigma} \\
&\leq \frac{C}{t^q} \sum_{Q\in\mathcal{Q}_t} \sigma(Q)^{1+q\frac{\alpha}{n}} \left(\fint_Q |f|^p\, d\sigma\right)^{\frac{q}{p}} \\
&\leq \frac{C}{t^q} \sum_{Q\in\mathcal{Q}_t} \left(\int_Q |f|^p\, d\sigma\right)^{\frac{q}{p}} \\
&\leq \frac{C}{t^q} \left(\sum_{Q\in\mathcal{Q}_t} \int_Q |f|^p\, d\sigma\right)^{\frac{q}{p}} \\
&\leq \frac{C}{t^q} \left(\int_{\mathbb{R}^n} |f(x)|^p\, d\sigma\right)^{\frac{q}{p}}.
\end{aligned}$$

□

Properties of $A_{p,q}$ Weights

In this section we gather a few basic facts about the $A_{p,q}$ weights and the closely related Muckenhoupt A_p weights; for further information see [12]. For $1 < p < \infty$, $w \in A_p$ if

$$[w]_{A_p} = \sup_Q \fint_Q w(x)\,dx \left(\fint_Q w(x)^{1-p'}\,dx\right)^{p-1} < \infty.$$

For $p = 1$, $w \in A_1$ if

$$[w]_{A_1} = \sup_Q \operatorname*{ess\,sup}_{x\in Q} \left(\fint_Q w(x)\,dx\right) w(x)^{-1} < \infty.$$

It follows at once from the definition that for all p and q such that $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$, then $w \in A_{p,q}$ if and only if $w^q \in A_r$, $r = 1 + \frac{q}{p'}$ and $[w^q]_{A_r} = [w]^q_{A_{p,q}}$. (When $p = 1$, we interpret $\frac{q}{p'}$ as 0.) By the duality of A_p weights, $w^{-p'} \in A_{r'}$ and $[w^{-p'}]_{A_{r'}} = [w]^{p'}_{A_{p,q}}$.

As a consequence, we have that both w^q and $w^{-p'}$ are in A_∞. Below we will need to use two properties of A_∞ weights; for both we give the sharp constant version. There are multiple definitions of the A_∞ characteristic (cf. [13]) but for our purposes we will simply use the fact that $[w]_{A_\infty} \leq C(n)[w]_{A_p}$.

Lemma 2.7 *Given a weight $\sigma \in A_p \subset A_\infty$, then for any cube Q and set $E \subset Q$:*

(1) $\left(\frac{|E|}{|Q|}\right)^p \leq [\sigma]_{A_p} \frac{\sigma(E)}{\sigma(Q)}$;

(2) $\frac{\sigma(E)}{\sigma(Q)} \leq 2\left(\frac{|E|}{|Q|}\right)^{\frac{1}{s'}}$, *where* $s' = c(n)[\sigma]_{A_\infty}$.

Proof The first inequality follows from the definition of A_p: see [12]. The second follows from the sharp form of the reverse Hölder inequality due to Hytönen and Pérez [16]: for every cube Q,

$$\left(\fint_Q \sigma(x)^s\,dx\right)^{\frac{1}{s}} \leq 2\fint_Q \sigma(x)\,dx,$$

where $s = 1 + \frac{1}{c(n)[\sigma]_{A_\infty}}$. The desired inequality follows if we apply Hölder's inequality with exponent s to $\fint_Q \sigma(x)\chi_E(x)\,dx$. □

Proof of Theorem 1.1

By Remark 2.5 it will suffice to prove the weak $(1, q)$ inequality for the dyadic operator $I^{\mathcal{D}}_\alpha$, where $\mathcal{D}$ is an arbitrary dyadic grid, and for f a non-negative function in L^∞_c.

Since $q > 1$, $\|\cdot\|_{L^{q,\infty}(w^q)}$ is a norm, so by Minkowski's inequality,

$$\|I_\alpha^{\mathcal{D}} f\|_{L^{q,\infty}(w^q)} = \left\| \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}-1} \int_{\mathbb{R}^n} f(y)\chi_Q(y)\,dy\ \chi_Q(\cdot) \right\|_{L^{q,\infty}(w^q)}$$

$$\leq \int_{\mathbb{R}^n} f(y) \left\| \sum_{Q\in\mathcal{D}} |Q|^{-\frac{1}{q}} \chi_Q(y)\chi_Q(\cdot) \right\|_{L^{q,\infty}(w^q)} dy.$$

To complete the proof we need to show that for almost every y,

$$\left\| \sum_{Q\in\mathcal{D}} |Q|^{-\frac{1}{q}} \chi_Q(y)\chi_Q(\cdot) \right\|_{L^{q,\infty}(w^q)} \leq CM(w^q)(y)^{\frac{1}{q}} \leq C[w^q]_{A_{1,q}} w(y).$$

The second inequality is just the fact that $w \in A_{1,q}$, so it remains to prove the first.

Let $K = \frac{1}{1-2^{-\frac{n}{q}}}$. Fix $y \in \mathbb{R}^n$ and $t > 0$, and let Q_t be the largest cube in $\mathcal{D}$ containing y such that $K|Q_t|^{-\frac{1}{q}} > t$. Now fix $x \neq y$ such that

$$\sum_{Q\in\mathcal{D}} |Q|^{-\frac{1}{q}} \chi_Q(y)\chi_Q(x) > t,$$

and let Q_x be the smallest cube in $\mathcal{D}$ containing x and y. Then

$$\sum_{Q\in\mathcal{D}} |Q|^{-\frac{1}{q}} \chi_Q(y)\chi_Q(x) = \sum_{k=0}^{\infty} |2^k Q_x|^{-\frac{1}{q}} = K|Q_x|^{-\frac{1}{q}},$$

and so by maximality, $x \in Q_t$. Since this is true for all $t > 0$, we have that

$$\left\| \sum_{Q\in\mathcal{D}} |Q|^{-\frac{1}{q}} \chi_Q(y)\chi_Q(\cdot) \right\|_{L^{q,\infty}(w^q)} = \sup_{t>0} t w^q(\{x \in \mathbb{R}^n : K|Q_x|^{-\frac{1}{q}} > t\})^{\frac{1}{q}}$$

$$\leq K \sup_{t>0} |Q_t|^{-\frac{1}{q}} w^q(Q_t)^{\frac{1}{q}} \leq KM(w^q)(y)^{\frac{1}{q}}.$$

Proof of Theorem 1.2

By Remark 2.5, it will suffice to prove the strong (p, q) inequality for f non-negative and in L_c^∞. We may also replace I_α by the sparse operator $I_\alpha^{\mathcal{S}}$, where $\mathcal{S}$ is any sparse subset of a dyadic grid $\mathcal{D}$.

Let $v = w^q$ and $\sigma = w^{-p'}$ and estimate as follows: there exists $g \in L^{q'}(w^{-q'})$, $\|gw^{-1}\|_{q'} = 1$, such that

$$\|(I_\alpha^{\mathcal{S}} f)w\|_q = \int_{\mathbb{R}^n} I_\alpha f(x)g(x)\,dx = \sum_{Q\in\mathcal{S}} |Q|^{\frac{\alpha}{n}} \langle f\rangle_Q \int_Q g(x)\,dx$$
$$= \sum_{Q\in\mathcal{S}} |Q|^{\frac{\alpha}{n}-1}\sigma(Q)v(Q)^{1-\frac{\alpha}{n}} \langle f\sigma^{-1}\rangle_{Q,\sigma} v(Q)^{\frac{\alpha}{n}} \langle gv^{-1}\rangle_{Q,v}.$$

Since $1-\frac{\alpha}{n} = \frac{1}{p'}+\frac{1}{q}$, by the definition of the $A_{p,q}$ condition and Lemma 2.7 (applied to both v and σ) we have that

$$|Q|^{\frac{\alpha}{n}-1}\sigma(Q)v(Q)^{1-\frac{\alpha}{n}} \leq [w]_{A_{p,q}}\sigma(Q)^{\frac{1}{p}}v(Q)^{\frac{1}{p'}} \leq [w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}}\sigma(E(Q))^{\frac{1}{p}}v(E(Q))^{\frac{1}{p'}}. \tag{1}$$

If we combine these two estimates, then by Hölder's inequality and Lemma 2.6 we get that

$$\begin{aligned}
&\|(I_\alpha^{\mathcal{S}} f)w\|_q \\
&\leq C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \sum_{Q\in\mathcal{S}} \langle f\sigma^{-1}\rangle_{Q,\sigma}\sigma(E(Q))^{\frac{1}{p}} v(Q)^{\frac{\alpha}{n}}\langle gv^{-1}\rangle_{Q,v} v(E(Q))^{\frac{1}{p'}} \\
&\leq C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \bigg(\sum_{Q\in\mathcal{S}} \langle f\sigma^{-1}\rangle_{Q,\sigma}^p \sigma(E(Q))\bigg)^{\frac{1}{p}} \bigg(\sum_{Q\in\mathcal{S}} [v(Q)^{\frac{\alpha}{n}}\langle gv^{-1}\rangle_{Q,v}]^{p'} v(E(Q))\bigg)^{\frac{1}{p'}} \\
&\leq C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \bigg(\sum_{Q\in\mathcal{S}} \int_{E(Q)} M_\sigma^{\mathcal{D}}(f\sigma^{-1})(x)^p\,d\sigma\bigg)^{\frac{1}{p}} \bigg(\sum_{Q\in\mathcal{S}} \int_{E(Q)} M_{v,\alpha}^{\mathcal{D}}(gv^{-1})(x)^{p'}\,dv\bigg)^{\frac{1}{p'}} \\
&\leq C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \bigg(\int_{\mathbb{R}^n} M_\sigma^{\mathcal{D}}(f\sigma^{-1})(x)^p\,d\sigma\bigg)^{\frac{1}{p}} \bigg(\int_{\mathbb{R}^n} M_{v,\alpha}^{\mathcal{D}}(gv^{-1})(x)^{p'}\,dv\bigg)^{\frac{1}{p'}} \\
&\leq C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \bigg(\int_{\mathbb{R}^n} (f(x)\sigma(x)^{-1})^p\,d\sigma\bigg)^{\frac{1}{p}} \bigg(\int_{\mathbb{R}^n} (g(x)v(x)^{-1})^{q'}\,dv\bigg)^{\frac{1}{q'}} \\
&= C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \|fw\|_p\|gw^{-1}\|_{q'} \\
&= C[w]_{A_{p,q}}^{1+\frac{q}{p'}+\frac{p'}{p}} \|fw\|_p.
\end{aligned}$$

Proof of Theorem 1.3

For our proof we need two lemmas. The first is a weighted estimate for functions in *BMO*; our proof adapts ideas from Ho [14].

Lemma 5.1 *Let* $\Phi(t) = t\log(e+t)$. *Given a weight* $\sigma \in A_\infty$, *then for any* $b \in BMO$ *and any cube* Q,

$$\|b - \langle b\rangle_Q\|_{\bar{\Phi},Q,\sigma} \leq C[\sigma]_{A_\infty}\|b\|_{BMO}.$$

Proof By the John-Nirenberg inequality, there exist constants C_1, C_2 such that for every cube Q and λ, $t > 0$,

$$|\{x \in Q : |b(x) - \langle b\rangle_Q| > \lambda t\}| \leq C_1|Q| \exp\left(-\frac{C_2\lambda t}{\|b\|_{BMO}}\right).$$

Since $\sigma \in A_\infty$, by Lemma 2.7 we have that

$$\sigma(\{x \in Q : |b(x) - \langle b\rangle_Q| > \lambda t\}) \leq 2C_1^{\frac{1}{s'}}\sigma(Q)\exp\left(-\frac{C_2\lambda t}{\|b\|_{BMO}s'}\right),$$

where $s' = c(n)[\sigma]_{A_\infty}$. Let

$$\lambda = \frac{(1 + 2kC_1^{\frac{1}{s'}})\|b\|_{BMO}s'}{C_2} = C[\sigma]_{A_\infty}\|b\|_{BMO},$$

where $\bar{\Phi}(t) \leq ke^t$. Then we have that

$$\begin{aligned}
\fint_Q \bar{\Phi}\left(\frac{|b(x) - \langle b\rangle_Q|}{\lambda}\right) d\sigma &\leq k\sigma(Q)^{-1}\int_0^\infty e^t\sigma(\{x \in Q : |b(x) - \langle b\rangle_Q| > \lambda t\})\, dt \\
&\leq 2kC_1^{\frac{1}{s'}}\int_0^\infty e^t \exp\left(-\frac{C_2\lambda t}{\|b\|_{BMO}s'}\right) dt \\
&\leq 2kC_1^{\frac{1}{s'}}\int_0^\infty e^{-2kC_1^{\frac{1}{s'}}t}\, dt \\
&= 1.
\end{aligned}$$

Therefore, by the definition of the Luxemburg norm, we get the desired inequality. □

The second lemma is a weighted variant of an estimate from [7]; when $\Phi(t) = t$ the unweighted estimate is originally due to Sawyer and Wheeden [23].

Lemma 5.2 *Fix $0 < \alpha < n$, a dyadic grid $\mathcal{D}$, a weight σ and a Young function Φ. Then for any $P \in \mathcal{D}$ and any function f,*

$$\sum_{\substack{Q\in\mathcal{D}\\ Q\subset P}} |Q|^{\frac{\alpha}{n}}\sigma(Q)|\|f\|_{\Phi,Q,\sigma} \leq C(\alpha)|P|^{\frac{\alpha}{n}}\sigma(P)|\|f\|_{\Phi,P,\sigma}.$$

Proof To prove this we need to replace the Luxemburg norm with the equivalent Amemiya norm [21, Section 3.3]:

$$\|f\|_{\Phi,P,\sigma} \leq \inf_{\lambda>0}\left\{\lambda \fint_P 1 + \Phi\left(\frac{|f(x)|}{\lambda}\right)\,d\sigma\right\} \leq 2\|f\|_{\Phi,P,\sigma}.$$

By the second inequality, we can fix $\lambda_0 > 0$ such that the middle quantity is less than $3\|f\|_{\Phi,P,\sigma}$. Then by the first inequality,

$$\begin{aligned}
\sum_{\substack{Q\in\mathcal{D}\\ Q\subset P}} |Q|^{\frac{\alpha}{n}}|Q|\|f\|_{\Phi,Q,\sigma} &= \sum_{k=0}^{\infty}\sum_{\substack{Q\subset P\\ \ell(Q)=2^{-k}\ell(P)}} |Q|^{\frac{\alpha}{n}}|Q|\|f\|_{\Phi,Q,\sigma} \\
&\leq |P|^{\frac{\alpha}{n}}\sum_{k=0}^{\infty}2^{-k\alpha}\sum_{\substack{Q\subset P\\ \ell(Q)=2^{-k}\ell(P)}} \lambda_0\int_Q 1 + \Phi\left(\frac{|f(x)|}{\lambda_0}\right)\,d\sigma \\
&= C|P|^{\frac{\alpha}{n}}\lambda_0\int_P 1 + \Phi\left(\frac{|f(x)|}{\lambda_0}\right)\,d\sigma \\
&\leq C|P|^{\frac{\alpha}{n}}\sigma(P)\|f\|_{\Phi,P,\sigma}.
\end{aligned}$$

□

Proof of Theorem 1.3 Again by Remark 2.5, it will suffice to prove that given any dyadic grid $\mathcal{D}$,

$$\|C_b^{\mathcal{D}}f\|_{L^q(w^q)} \leq C\|f\|_{L^p(w^p)},$$

where f is non-negative and in L^∞_c. By duality there exists a non-negative function $g \in L^{q'}(w^{-q})$, $\|g\|_{L^{q'}(w^{-q})} = 1$, such that

$$\begin{aligned}
\|C_b^{\mathcal{D}}f\|_{L^q(w^q)} &= \int_{\mathbb{R}^n} C_b^{\mathcal{D}}f(x)g(x)\,dx \\
&= \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}}\int_Q\fint_Q |b(x)-b(y)|f(y)g(x)\,dy\,dx
\end{aligned}$$

$$\leq \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \fint_Q |b(y) - \langle b\rangle_Q| f(y)\,dy \int_Q g(x)\,dx$$

$$+ \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \fint_Q |b(x) - \langle b\rangle_Q| g(x)\,dx \int_Q f(y)\,dy$$

$$= I_1 + I_2.$$

We will first estimate I_1. Let $v = w^q$ and $\sigma = w^{-p'}$, and let $\Phi(t) = t\log(e+t)$. Then by Lemma 5.1, since $[\sigma]_{A_\infty} \leq [w]_{A_{p,q}}^{p'}$,

$$I_1 = \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \sigma(Q) \fint_Q |b(y) - \langle b\rangle_Q| f(y)\sigma(y)^{-1}\,d\sigma \fint_Q g(x)\,dx$$

$$\leq C \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \sigma(Q) \|f\sigma^{-1}\|_{\Phi,Q,\sigma} \|b - \langle b\rangle_Q\|_{\bar{\Phi},Q,\sigma} \fint_Q g(x)\,dx$$

$$\leq C[w]_{A_{p,q}}^{p'} \|b\|_{BMO} \sum_{Q\in\mathcal{D}} |Q|^{\frac{\alpha}{n}} \sigma(Q) \|f\sigma^{-1}\|_{\Phi,Q,\sigma} \fint_Q g(x)\,dx.$$

We now want to show that we can replace the summation over cubes in $\mathcal{D}$ by a summation over a sparse subset $\mathcal{S}$ of $\mathcal{D}$. We do this using an argument from [7]; see also [9]. Fix $a = 2^{n+1}$ and define the sets

$$\Omega_k = \{x \in \mathbb{R}^n : M^{\mathcal{D}} g(x) > a^k\}.$$

Then arguing exactly as in the construction of the Calderón-Zygmund cubes (see [8, Appendix A]), each set Ω_k is the union of a collection $\mathcal{S}_k$ of maximal, disjoint cubes in $\mathcal{D}$ that have the property that $a^k < \langle g\rangle_Q \leq 2^n a^k$. Moreover, the set $\mathcal{S} = \bigcup_k \mathcal{S}_k$ is sparse.

Now let

$$\mathcal{C}_k = \{Q \in \mathcal{D} : a^k < \langle g\rangle_Q \leq a^{k+1}\}.$$

Then by the maximality of the cubes in $\mathcal{S}_k$, every cube $P \in \mathcal{C}_k$ is contained in a unique cube in $\mathcal{S}_k$. Therefore, we can continue the above estimate and apply Lemma 5.2 to get

$$I_1 \leq C[w]_{A_{p,q}}^{p'} \|b\|_{BMO} \sum_k \sum_{Q\in\mathcal{C}_k} |Q|^{\frac{\alpha}{n}} \sigma(Q) \|f\sigma^{-1}\|_{\Phi,Q,\sigma} \fint_Q g(x)\,dx$$

$$\leq C[w]_{A_{p,q}}^{p'} \|b\|_{BMO} \sum_k a^k \sum_{P\in\mathcal{S}_k} \sum_{Q\subset P} |Q|^{\frac{\alpha}{n}} \sigma(Q) \|f\sigma^{-1}\|_{\Phi,Q,\sigma}$$

$$\leq C[w]_{A_{p,q}}^{p'} \|b\|_{BMO} \sum_k \sum_{P \in \mathcal{S}_k} |P|^{\frac{\alpha}{n}} \sigma(P) \|f\sigma^{-1}\|_{\Phi,P,\sigma} \fint_P g(x)\,dx$$

$$= C[w]_{A_{p,q}}^{p'} \|b\|_{BMO} \sum_{P \in \mathcal{S}} |P|^{\frac{\alpha}{n}-1} \sigma(P) v(P)^{1-\frac{\alpha}{n}} \|f\sigma^{-1}\|_{\Phi,Q,\sigma} v(P)^{\frac{\alpha}{n}} \fint_Q g v^{-1}\,dv.$$

We can now argue exactly as in the proof of Theorem 1.2, beginning with the estimate (1) and applying Lemma 2.6 to complete the estimate of I_1 with a constant $[w]_{A_{p,q}}^{p'+1+\frac{q}{p'}+\frac{p'}{p}}$.

The estimate for I_2 is essentially the same, exchanging the roles of f and g and σ and v. This yields the above estimate except that the constant is now $[w]_{A_{p,q}}^{q+1+\frac{q}{p'}+\frac{p'}{p}}$. This completes the proof. □

Tres Recuerdos de Cora Sadosky

I first met Cora at an AMS sectional meeting in Burlington, Vermont, in 1995. It began inauspiciously: at the reception on the first night a determined looking woman came up to me, waved her finger under my nose and said, "I have a bone to pick with you. We will talk later," and then marched off. She found me again about 30 min later and proceeded to explain. The year before she had been asked by an NSF reviewer for her opinion of my proposal which mentioned the two weight problem for the Hilbert transform. She had told the reviewer to refer me to a paper by her and Mischa Cotlar where they gave the first (and for a long time the only) characterization of these pairs of weights. He did not, however, share this reference, and when I published the paper based on this work I did not cite it. Cora assumed that I had simply disregarded this advice and was understandably annoyed. However, once I explained that I had never received this information she immediately became much friendlier and invited me to visit her in Washington, D.C.

For the next few years she took an interest in my career. Her first major intervention on my behalf came in 1996, when she applied her forceful personality, first to convince me that I must attend the *El Escorial* conference in 1996 (despite moving, changing jobs, and having two small children and a pregnant wife), and then to strong-arm funding from a colleague to pay for my trip. It was at this meeting that I met, among others, Carlos Pérez, and began a collaboration that has continued to the present day.

Two years later, in 1998, we met again at an AMS sectional meeting in Albuquerque. At this meeting she picked up on a point that my Spanish colleagues were also making: given my name and my ancestry, I really ought to be able to speak Spanish. Her solution was that I should "read a good math book in Spanish." She strongly recommended that I read Javier Duoandikoetxea's book, *Análisis de Fourier*, telling me that I would see some good mathematics as well as "learn

Spanish." For the next year I worked through the text line by line, in the process writing a complete translation. I approached Javi with an offer to complete the translation and update the notes, and together we produced an English edition. In the process I did in fact learn a great deal of harmonic analysis, but unfortunately, Cora's original goal was not achieved: my spoken Spanish did not improve appreciably. Moreover, this translation had the unintended consequence of convincing large numbers of mathematicians from Spain, Argentina and elsewhere that I did in fact speak Spanish.

Cora and I never collaborated on a paper. She suggested several projects, but my interests were moving away from hers and nothing came to fruition. At the time I never really quite understood or appreciated the support she provided at these points in my career, and it is only in looking back that I realize how much I owe her. So belatedly I say, *muchísimas gracias, Cora.*

References

1. L. Chaffee, Characterizations of BMO through commutators of bilinear singular integral operators. Proc. R. Soc. Edinb. Sec. A. **146**(6), 1159–1166 (2016). arXiv:1410.4587
2. S. Chanillo, A note on commutators. Indiana Univ. Math. J. **31**(1), 7–16 (1982)
3. M. Christ, R. Fefferman, A note on weighted norm inequalities for the Hardy-Littlewood maximal operator. Proc. Am. Math. Soc. **87**(3), 447–448 (1983)
4. D. Chung, C. Pereyra, C. Pérez, Sharp bounds for general commutators on weighted Lebesgue spaces. Trans. Am. Math. Soc. **364**(3), 1163–1177 (2012)
5. D. Cruz-Uribe, Two weight norm inequalities for fractional integral operators and commutators. (2015, preprint). arXiv:1412.4157
6. D. Cruz-Uribe, A. Fiorenza, Endpoint estimates and weighted norm inequalities for commutators of fractional integrals. Publ. Mat. **47**(1), 103–131 (2003)
7. D. Cruz-Uribe, J.M. Martell, C. Pérez, Extensions of Rubio de Francia's extrapolation theorem. Collect. Math. (Vol. Extra), 195–231 (2006)
8. D. Cruz-Uribe, J.M. Martell, C. Pérez, *Weights, Extrapolation and the Theory of Rubio de Francia*. Operator Theory: Advances and Applications, vol. 215 (Birkhäuser/Springer, Basel, 2011)
9. D. Cruz-Uribe, K. Moen, Sharp norm inequalities for commutators of classical operators. Publ. Mat. **56**, 147–190 (2012)
10. D. Cruz-Uribe, K. Moen, A fractional Muckenhoupt-Wheeden theorem and its consequences. Integral Equ. Oper. Theory **76**(3), 421–446 (2013)
11. D. Cruz-Uribe, K. Moen, One and two weight norm inequalities for Riesz potentials. Ill. J. Math. **57**(1), 295–323 (2013)
12. J. Duoandikoetxea, *Fourier Analysis*. Graduate Studies in Mathematics, vol. 29 (American Mathematical Society, Providence, 2001)
13. J. Duoandikoetxea, F.J. Martín-Reyes, S. Ombrosi, On the A_∞ conditions for general bases. Math Z. **282**(3), 955–972 (2016)
14. K.-P. Ho, Characterizations of BMO by A_p weights and p-convexity. Hiroshima Math. J. **41**(2), 153–165 (2011)
15. T. Hytönen, The A_2 theorem: Remarks and complements (2012, preprint). arXiv:1212.3840
16. T. Hytönen, C. Pérez, Sharp weighted bounds involving A_∞. Anal. PDE **6**(4), 777–818 (2013)
17. M.T. Lacey, K. Moen, C. Pérez, R.H. Torres, Sharp weighted bounds for fractional integral operators. J. Funct. Anal. **259**(5), 1073–1097 (2010)

18. A.K. Lerner, A simple proof of the A_2 conjecture. Int. Math. Res. Not. IMRN (14), 3159–3170 (2013)
19. B. Muckenhoupt, R.L. Wheeden, Weighted norm inequalities for fractional integrals. Trans. Am. Math. Soc. **192**, 261–274 (1974)
20. C. Pérez, Two weighted inequalities for potential and fractional type maximal operators. Indiana Univ. Math. J. **43**(2), 663–683 (1994)
21. M.M. Rao, Z.D. Ren, *Theory of Orlicz Spaces*. Monographs and Textbooks in Pure and Applied Mathematics, vol. 146 (Marcel Dekker Inc., New York, 1991)
22. E.T. Sawyer, A two weight weak type inequality for fractional integrals. Trans. Am. Math. Soc. **281**(1), 339–345 (1984)
23. E.T. Sawyer, R.L. Wheeden, Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces. Am. J. Math. **114**(4), 813–874 (1992)
24. E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, vol. 30 (Princeton University Press, Princeton, 1970)
25. E.M. Stein, G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton Mathematical Series, vol. 32 (Princeton University Press, Princeton, 1971)

Finding Cycles in Nonlinear Autonomous Discrete Dynamical Systems

Dmitriy Dmitrishin, Anna Khamitova, Alexander M. Stokolos, and Mihai Tohaneanu

Dedicated to Alexey Solyanik on his 55th birthday

Abstract The goal of this paper is to provide an exposition of recent results of the authors concerning cycle localization and stabilization in nonlinear dynamical systems. Both the general theory and numerical applications to well-known dynamical systems are presented. This paper is a continuation of Dmitrishin et al. (Fejér polynomials and chaos. Springer proceedings in mathematics and statistics, vol 108, pp. 49–75, 2014).

Introduction

The problem of cycle detection is fundamental in mathematics. In this paper we will be mainly concerned with the problem of detecting cycles of large length in an autonomous discrete system $x_{n+1} = f(x_n)$. The standard approach is to consider the composition map $f_T(x) := f(\ldots f(x)\ldots)$ and then solve the equation $f_T(x) = x$. However, this approach does not work well even in some basic cases. For example, in the model case of the logistic map $f(x) = 4x(1-x)$ it leads to a polynomial equation of degree 2^T, and thus a relatively small cycle length T can give rise to very serious computational difficulties.

The goal of this article is to suggest an alternative approach to the problem of cycle localization. In full generality the problem is very difficult and it is hard to believe that a universal technique could be developed. We thus start with a model

D. Dmitrishin (✉)
Odessa National Polytechnic University, 1 Shevchenko Ave., Odessa 65044, Ukraine
e-mail: dmitrishin@opu.ua

A. Khamitova • A.M. Stokolos
Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30458, USA
e-mail: anna_khamitova@georgiasouthern.edu; astokolos@georgiasouthern.edu

M. Tohaneanu
University of Kentucky, Lexington, KY 40506, USA
e-mail: mihai.tohaneanu@uky.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_8

case of non-linear autonomous discrete dynamical systems. The simplicity of the setting allows us to make some progress and to develop a feasible plan for further developments of the method. Since one fundamental tool of dynamics, often used for analyzing continuous time systems, is the reduction of continuous time flow to its Poincaré section, which is a discrete system, understanding the case of discrete systems is of great help in studying continuous systems also.

The core of the suggested method is the stabilization of the solutions by delayed feedback control (DFC) of a special type. We will briefly discuss a way to linearly stabilize the system in section"Linear Control"; however, it turns out that the linear DFC method has some obvious limitations regardless of the number of prehistory terms involved. In contrast, in the subsequent sections we will show that a certain nonlinear DFC schedule allows one to robustly stabilize chaotic solutions for any admissible range of parameters.

The methods we developed can be considered as chaos stabilization, and we believe they are of interest in other disciplines. Chaos theory is a part of modern Physics and the majority of the publications on chaos are in Physics literature; for instance, problems of stability have been discussed in [1–5, 12, 14–18, 21–27]. There are many specialists in chaos theory who are physicists, among whom we mention P. Cvetanović, C. Grebogi, E. Ott, K. Pyragas, J.A. Yorke etc. On the other hand, many biological systems exhibit chaotic behavior as well. A fundamental monograph of I.D. Murray [20] contains deep and advanced discussions and applications of the non-linear dynamical systems to models of population growth.

Settings

Consider the discrete dynamical system

$$x_{n+1} = f(x_n), \qquad f : A \to A,\ A \subset \mathbb{R}^m. \tag{1}$$

where A is a convex set that is invariant under f. Let us assume that the system has an unstable T-cycle $(x_1^*, \ldots, x_T^*)$. We define the cycle multipliers $\mu_1, \ldots, \mu_m$ as the zeros of the characteristic polynomial

$$\det\left(\mu I - \prod_{j=1}^{T} Df(x_j^*)\right) = 0. \tag{2}$$

We will assume that the multipliers are located in a region $M \subset \mathbb{C}$.

Stability Analysis

A standard approach (cf. [19]) to investigate stability with no delays is to construct a new map that has points of the cycle as equilibriums and then linearize about the equilibriums. Let us consider a system with time delay in a general form

$$x_{n+1} = F(x_n, x_{n-1}, \ldots, x_{n-\tau}), \qquad F : \mathbb{R}^m \to \mathbb{R}^m,\ \tau \in \mathbb{Z}_+. \tag{3}$$

We will study the local stability of a cycle $\{\eta_1, \ldots, \eta_T\}$ where $\eta_j \in \mathbb{R}^m$. In other words for all $n \geq \tau + 1$ the following equations are valid

$$\eta_{(n+1) \mod T} = F(\eta_{n \mod T}, \eta_{(n-1) \mod T}, \ldots, \eta_{(n-\tau) \mod T}).$$

where, slightly abusing notation, we assume that $T \mod T = T$.

We can now consider an auxiliary system with respect to the vector

$$z_n = \begin{pmatrix} z_n^{(1)} \\ z_n^{(2)} \\ \vdots \\ z_n^{(\tau+1)} \end{pmatrix} = \begin{pmatrix} x_{n-\tau} \\ x_{n-\tau+1} \\ \vdots \\ x_n \end{pmatrix}.$$

of size $m(\tau + 1)$:

$$z_{n+1} = \begin{pmatrix} z_{n+1}^{(1)} \\ z_{n+1}^{(2)} \\ \vdots \\ z_{n+1}^{(\tau+1)} \end{pmatrix} = \begin{pmatrix} z_n^{(2)} \\ z_n^{(3)} \\ \vdots \\ f(z_n^{(\tau+1)}, z_n^{(\tau)}, \ldots, z_n^{(1)}) \end{pmatrix}.$$

We can now rewrite (3) in the form

$$z_{n+1} = \mathcal{F}(z_n)$$

with $\mathcal{F} : \mathbb{R}^{m(\tau+1)} \to \mathbb{R}^{m(\tau+1)}$.

Let $\Psi(z) := \mathcal{F}(\ldots \mathcal{F}(z) \ldots)$ be $\mathcal{F}$ composed with itself T-times. We can now analyze the system

$$y_{n+1} = \Psi(y_n). \tag{4}$$

Let us periodically repeat the elements of the cycle:

$\{\eta_1, \eta_2, \ldots, \eta_T, \eta_1, \eta_2, \ldots, \eta_T, \ldots\}$. The first $\tau + 1$ elements of this sequence form a vector

$$y_1^* = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \end{pmatrix}.$$

In the same way we define the vectors

$$y_2^* = \begin{pmatrix} \eta_2 \\ \eta_3 \\ \vdots \end{pmatrix}, \ \ldots \ , \ y_T^* = \begin{pmatrix} \eta_T \\ \eta_1 \\ \vdots \end{pmatrix}.$$

It is clear that the vectors $y_1^*, \ldots, y_T^*$ are equilibria of the system (4).

The cycle $\{\eta_1, \ldots, \eta_T\}$ of the system (3) is asymptotically locally stable if and only if all equilibriums $y_1^*, \ldots, y_T^*$ of the system (4) are asymptotically locally stable.

For the equilibrium point y_1^* of the system (4) the Jacobi matrix is defined by the formula

$$D\Psi(y_1^*) = \prod_{j=1}^{T} D\mathcal{F}(y_j^*) \tag{5}$$

where the matrix $D\mathcal{F}(y_j^*)$ has size $m(\tau + 1) \times m(\tau + 1)$ and equals

$$D\mathcal{F}(y_j^*) = \begin{pmatrix} O & I & O & \ldots & O \\ O & O & I & \ldots & O \\ \ldots. & & & & \\ O & O & O & \ldots & I \\ Q_1^{(j)} & Q_2^{(j)} & Q_3^{(j)} & \ldots & Q_{\tau+1}^{(j)} \end{pmatrix}. \tag{6}$$

Here the matrices O and I are the $m \times m$ zero and identity matrices. Further,

$$Q_r^{(j)} = \left.\frac{\partial f}{\partial z^{(r)}}\right|_{y_j^*}, \quad r = 1, \ldots, \tau + 1; \ j = 1, \ldots, T,$$

i.e. the value of the derivative evaluated at the point y_j^*.

For all other equilibria y_j^* the Jacobi matrices $D\mathcal{F}(y_j^*)$ can be computed in the same manner, and

$$D\Psi(y_j^*) = D\mathcal{F}(y^*_{(T+j) \mod T}) \cdots D\mathcal{F}(y_j^*),$$

which can be obtained from (5) by a cyclic permutation of the factors. The eigenvalues of $D\Psi(y_j^*)$ thus coincide for all $j = 1, \ldots, T$ (see for example [13] for more details.)

If all eigenvalues of the matrix $D\Psi(y_j^*)$, which are the roots of the polynomial

$$\det\left(\lambda I - D\Psi(y_j^*)\right)$$

are less than one in absolute values then the cycle of the system (3) is locally asymptotically stable.

Note that in the scalar case $m = 1$ the matrices $D\mathcal{F}(y_j)$ are in Frobenius form, and the matrix (6) is a generalized form of the companion matrix. If the system has the special form (10) below, then the characteristic equation can be found explicitly by means of induction, see [6].

Linear Control

There is a common belief that a generalized linear control

$$u = -\sum_{j=1}^{N-1} \varepsilon_j \left(x_{n-j} - x_{n-j+1}\right) \tag{7}$$

can stabilize the equilibrium for the whole range of the admissible multipliers of the system (1). In this case $F(x_n, x_{n-1}, \ldots, x_{n-(N-1)T}) = f(x_n) + u$ and the characteristic equation for the system closed by the control (7) is $\chi_\mu(\lambda) = \prod_{j=1}^{m} \chi_{\mu_j}(\lambda)$, where $\chi_\mu(\lambda) := \lambda^N - \mu\lambda^{N-1} + p(\lambda)$ and

$$p(\lambda) = a_1\lambda^{N-1} + a_2\lambda^{N-2} + \cdots + a_N. \tag{8}$$

Here μ_j are cycle multipliers, i.e. the roots of the characteristic equation (2) of the open loop system, while the coefficients a_i and the gain ε_j are related by the bijection

$$\varepsilon_j = \sum_{k=j+1}^{N} a_k, \; j = 1, \ldots, N-1. \tag{9}$$

The proof can be done by the methods considered in the section "Stability Analysis" above. It is not trivial but similar to a non-linear scalar case [6].

In the case of real multipliers μ a careful application of Vieta's theorem implies that a necessary condition for the polynomials $\chi_\mu(\lambda)$ to be Schur stable is $1 - 2^N < \mu < 1$. It turns out [25] that for any fixed μ in this range there are coefficients that guarantee the stability of the polynomial $\chi_\mu(\lambda)$ for this given μ.

At this point, a natural question to ask is how robust the selected control can be, i.e. assuming that we are given the Schur-stable polynomials $\chi_\mu(\lambda)$, how much can we perturb the multipliers μ so that $\chi_\mu(\lambda)$ remain Schur-stable? More rigorously, the inquiry is the following:

What is the maximum length of a connected component of M?

In [10] we discovered a remarkable fact - the answer to the above question is 4 regardless of how large N is. Below is an idea of the proof.

Solyanik Visualisation

The polynomial $\chi_m(\lambda)$ is quite complicated and difficult to study directly. Alexey Solyanik (Personal communication) suggested a remarkable way to visualize the situation. Namely, $\chi_\mu(\lambda)$ is a stable polynomial if and only if $\chi_\mu(\lambda) = \lambda^N - \mu\lambda^{N-1} + p(\lambda) \neq 0$ for $|\lambda| \geq 1$ or

$$\frac{1}{\mu} \neq \frac{\frac{1}{\lambda}}{1 + \frac{p(\lambda)}{\lambda^N}} \Rightarrow \frac{1}{\mu} \neq \frac{z}{1 + q(z)} =: \Phi(z),\ |z| \leq 1.$$

where $z = \frac{1}{\lambda}$ and $q(z) = \frac{p(\lambda)}{\lambda^N}$. Therefore χ_μ is Schur stable if and only if $1/\mu \notin \Phi(\bar{\mathbb{D}})$, where $\mathbb{D} = \{z : |z| < 1\}$. This can be rewritten as $\mu \in (\bar{\mathbb{C}} \backslash \Phi(\bar{\mathbb{D}})^*$, where $z^* := 1/\bar{z}$ is an inversion. The above formula reduces the problem of stability to the problem of verifying whether μ is in the above set, which is still difficult, but more manageable.

Köbe Quarter Theorem Application

Now, let us expand $\Phi(z)$ in a power series $\Phi(z) = z + a_2 z^2 + a_3 z^3 + \ldots$ in $\mathbb{D}$. If Φ is univalent, then by the Köbe Quarter Theorem $\frac{1}{4}\mathbb{D} \subset \Phi(\mathbb{D})$ and therefore

$$\left|\frac{1}{\mu}\right| > \frac{1}{4} \Rightarrow |\mu| < 4.$$

We were able to get in [10] a generalization of the Köbe Quarter Theorem which allowed us to obtain the result mentioned above. We also remark that the inequality above explains the value 4 mentioned in the previous section.

Finally, it was proved in [10] that *if the diameter of the set of multipliers is larger than 16, or the diameter of any of its connected component is larger than 4, then for any N there is no control* (7) *that stabilizes equilibria of the system* (1) *for all admissible parameters of the system.*

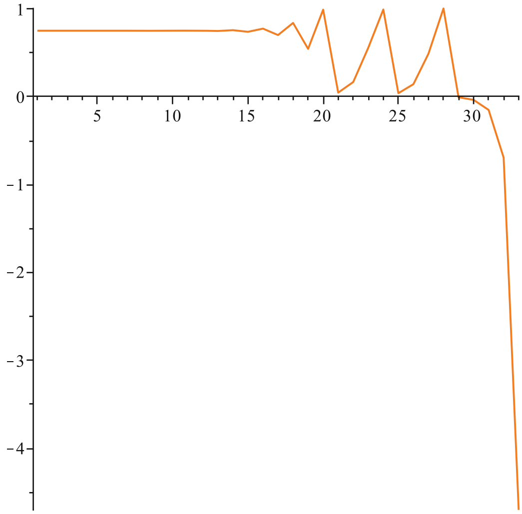


Fig. 1 Logistic close-loop system

The other obvious problem with the linear control is that the close-loop system can have solutions that are outside the domain of the map, i.e. the convex invariant set for the open-loop system is not necessarily an invariant set for the closed-loop system. In Fig. 1 above the solution to the logistic system closed by the control $u = -0.01(x_n - x_{n+1})$ with $x_0 = 0.7501$ is displayed. Note that $x_0 = 0.75$ is an equilibrium for a logistic map while a little perturbation produces a solution that blows up after 30 iterations.

These two basic obstacles - the range for of the close-loop system and the limited rage for the connected component for the multiplier - justify the introduction of the non-linear controls.

Average Non-linear Control

Typically, an arbitrarily chosen initial value x_0 produces a chaotic solution, i.e. we observe strong oscillations. An efficient way to kill oscillations is averaging, as can be readily seen in the summability of many trigonometric series. So, we decided to consider a new system

$$x_{n+1} = \sum_{k=1}^{N} a_k f(x_{n-kT+T}), \qquad \sum_{k=1}^{N} a_k = 1. \tag{10}$$

It is useful to rewrite the system as $x_{n+1} = f(x_n) + u_n$, where

$$u_n = -\sum_{j=1}^{N-1} \varepsilon_j \left(f\left(x_{n-jT+T}\right) - f\left(x_{n-jT}\right)\right), \tag{11}$$

where a_k and ε_k are in one-to-one correspondence as in (9).

We can now see that (10) in fact is the system (1) closed by the control (11). In this case the convex invariant set for the open-loop system is also invariant for the closed-loop system. We also remark that $u_n = 0$ for cycle points of period T, which implies that the closed-loop system $x_{n+1} = f(x_n) + u_n$ preserves the T-cycles of the initial one, which is very important for us.

Stability Analysis

The characteristic equation for the system (10) can be written in a remarkably useful form

$$\prod_{j=1}^{m}\left[\lambda^{T(N-1)+1} - \mu_j\left(\sum_{k=1}^{N} a_k\lambda^{N-k}\right)^T\right] = 0, \quad \mu_j \in M,\ j = 1, \ldots, m.$$

Here μ_j are the multipliers of the open loop system (1). The proof for the scalar case $m = 1$ can be found in [6], and the vector case can be done in a similar way. Denote

$$\phi(\lambda) := \lambda^{T(N-1)+1} - \mu\left(\sum_{k=1}^{N} a_k\lambda^{N-k}\right)^T,$$

$$q(z) = \sum_{j=1}^{N} a_j z^{j-1} \quad \text{and} \quad z = \frac{1}{\lambda}.$$

Then $\phi(\lambda) = 0$ is equivalent to $\mu^{-1} = z\,(q(z))^T$. So, if

$$F_T(z) := z\,(q(z))^T \tag{12}$$

then the inclusion

$$M \subset (\bar{\mathbb{C}} \backslash F_T(\bar{\mathbb{D}})^* \tag{13}$$

guarantees the local asymptotic stability of the cycle with multipliers in the set M. The inclusion (13) is the Solyanik visualization in this setting.

We are thus left with the following problem in geometric complex function theory: *given a set M containing all the multipliers, find a properly normalized polynomial map $z \to F_T(z)$ such that $M \subset \left(\bar{\mathbb{C}} \backslash F_T(\bar{\mathbb{D}})\right)^*$.*

Since in the simplest case $x_{n+1} = \mu x_n$ the solution $x_n = C\mu^n$ is exponentially blowing up for $\mu > 1$ there is no way to stabilize the case when one of the multipliers satisfies $\mu > 1$ by only using the small gain (11). Therefore we will consider only negative real multipliers, and in the complex case we will assume that M is disjoint from $[1, \infty)$. We also remark that the cycle is already stable if the multipliers are in $(-1, 1)$ or more generally in the unit disc $\mathbb{D}$ of the complex plane $\mathbb{C}$.

It is natural to expect that the stability will be getting worse if the set M is close to the half axis $[1, \infty)$. Below we will see several examples supporting this thesis.

Since $((0, \infty)\backslash(0, 1])^* = (0, 1)$, and (0,1) is the largest admissible range for real positive multipliers, the best case scenario for us is a polynomial whose image of the unit disc looks like a narrow neighborhood of $(0, 1]$. Solyanik (Personal communication) suggested as an example of such an object a famous Alexander polynomial - the polynomial with the coefficients $a_k = 1/k$, properly normalized [1]. The corresponding set $(\bar{\mathbb{C}}\backslash F_T(\bar{\mathbb{D}})^*$ has a cardioid type shape and for a large N can cover any given point except for the real numbers $z \geq 1$. We refer the reader to Fig. 2, where the image of the unit disc under the Alexander polynomial map is the interior of the inner (orange) curve, and the set $(\bar{\mathbb{C}}\backslash F_T(\bar{\mathbb{D}})^*$ is the interior of the outer (green) curve.

We thus know that theoretically there is a way to stabilize equilibria with any multipliers outside $[1, \infty)$ by the control (11). However, the length of the prehistory is growing exponentially with the magnitude of the multiplier. In Fig. 2 the largest multiplier which can be covered has a magnitude of 15, while the length of the prehistory is 20,000. We are thus led to the problem of finding the optimal coefficients a_j so that for a given set of multipliers M the prehistory N is as small as possible.

It turns out that in the case of multipliers with negative real part the number N can be shown to be much smaller, of only polynomial growth with respect to the size of the multipliers, which is of practical use. This will be explored in the subsequent sections.

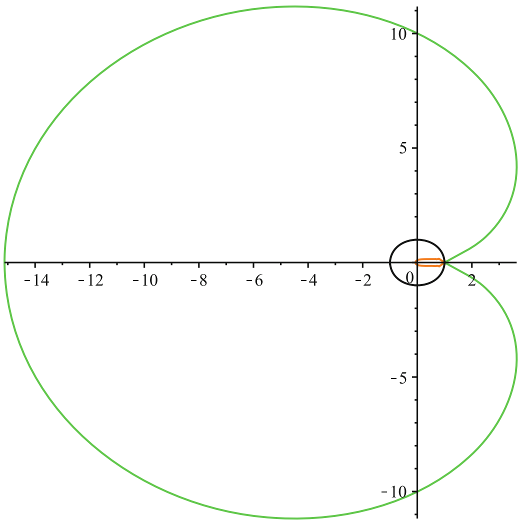


Fig. 2 Image of $F(\mathbb{D})$ with the Alexander polynomial $F(z)$, $N = 20{,}000$

Optimization Problem

Let us look first at the case when the multipliers lie on the half-axis $(-\infty, 1)$. In this case the problem of stabilization can be reduced to the following optimization problem: find

$$I_N^{(T)} = \sup_{\sum_{j=1}^N a_j = 1} \min_{t\in[0,\pi]} \left\{ \Re\left(F_T(e^{it})\right) : \Im\left(F_T(e^{it})\right) = 0 \right\}.$$

Technically speaking, this leads to a disconnected set of multipliers, as the boundary of $(\bar{\mathbb{C}} \backslash F_T(\bar{\mathbb{D}})^*$ will be tangent to the real axis (see, for example, Fig. 4 below). However, there is an easy trick to get rid of the tangent points: given $\epsilon > 0$ the polynomial $F_T^\epsilon(z) = (1+\epsilon)^{-1}(F_T(z) + \epsilon z)$ satisfies

$$\min_{t\in[0,\pi]} \left\{ \Re\left(F_T^\epsilon(e^{it})\right) : \Im\left(F_T^\epsilon(e^{it})\right) = 0 \right\} > I_N^{(T)} - \epsilon.$$

and does not intersect the real axis except for $t = 0$ and $t = \pi$. Since $\lim_{\epsilon\to 0} F_T^\epsilon(z) = F_T(z)$, one can use the coefficients of the polynomials $F_T(z)$ instead of $F_T^\epsilon(z)$ in computer simulations. In particular, it is done below.

It can be shown that for the closed-loop system a robust stabilization (i.e. by the same control for all $\mu \in (-\mu^*, 1)$) of any T-cycle is possible if

$$(\mu^*) \cdot |I_N^{(T)}| \le 1. \tag{14}$$

By duality, for any $\mu^* \ge 1$ a robust stabilization of T-cycle in the closed-loop system is possible if $N \ge N^*$, where N^* is the minimal integer N such that (14) holds. Formula (14) provides a practical criterion for the choice of N given μ^* and T.

Real Multipliers, Optimal Polynomials for T = 1

It was proved in [7] (see also [8]) that given N the largest μ such that

$$(-\mu, 1) \subset (\bar{\mathbb{C}} \backslash F_1(\bar{\mathbb{D}})^*$$

can be achieved if the coefficients of $q(z)$ in (12) are the coefficients of a polynomial related to the well-known Fejér polynomial, namely

$$a_j = 2\tan\frac{\pi}{2(N+1)}\left(1 - \frac{j}{N+1}\right)\sin\frac{\pi j}{N+1}, \quad j = 1, \dots, N. \tag{15}$$

Moreover, any N such that

$$\mu \cdot \tan^2\frac{\pi}{2(N+1)} \le 1$$

allows stabilization and *this inequality is sharp*.

Fig. 3 $F_1(\mathbb{D})$, N=12

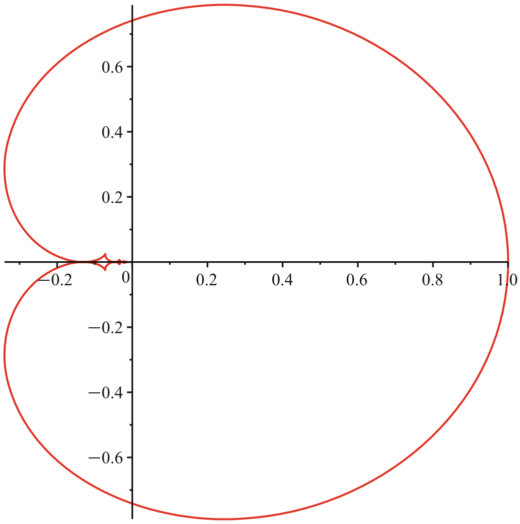
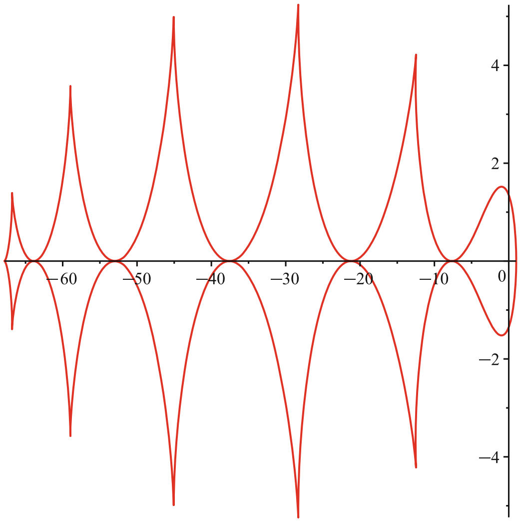


Fig. 4 $M = (\bar{\mathbb{C}} \backslash F_1(\bar{\mathbb{D}})^*$



For the choice $N = 12$ we display the image $F_1(\mathbb{D})$ and the maximal multiplier set M that allows for stability in Figs. 3 and 4.

Note that for $N = 12$ the boundary for the multiplier is

$$\mu \cdot \tan^2(\pi/26) \leq 1.$$

That implies that $\mu \leq 68$ which is easy to see in Fig. 4.

Real Multipliers, Optimal Polynomials for T = 2

It was proved in [9] (see also [8]) that given N the largest μ such that

$$(-\mu, 1) \subset (\bar{\mathbb{C}} \backslash F_2(\bar{\mathbb{D}})^*$$

can be achieved if the coefficients of $q(z)$ in (12) are coefficients related to the Fejér kernel of order $2N$

$$a_j = \frac{2}{N}\left(1 - \frac{2j-1}{2N}\right), \qquad j = 1, \ldots, N. \tag{16}$$

Moreover, any N such that

$$\mu \cdot \frac{1}{N^2} \leq 1$$

allows stabilization of a 2-cycle and *this inequality is sharp*.

For the choice $N = 12$ we display the image $F_1(\mathbb{D})$ and the maximal multiplier set M that allows for stability in Figs. 5 and 6.

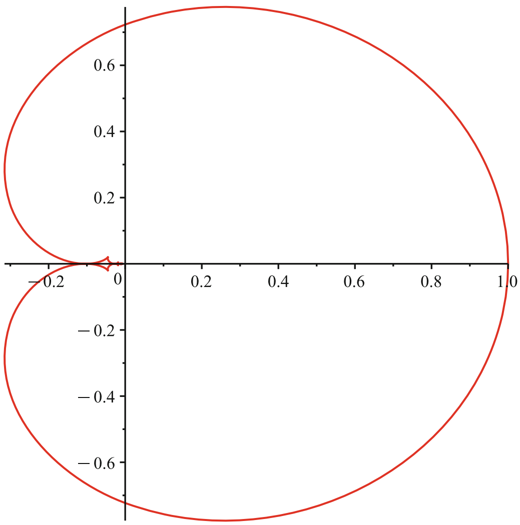
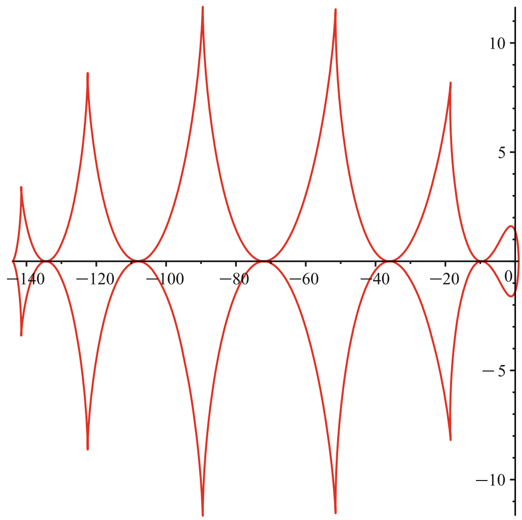


Fig. 5 $F_2(\mathbb{D})$, N = 12

Fig. 6 $M = (\bar{\mathbb{C}} \backslash F_2(\bar{\mathbb{D}})^*$



Note that for $N = 12$ the boundary for the multiplier is $\mu \cdot \frac{1}{12^2} \leq 1$. That implies that $\mu \leq 144$ which is easy to see in Fig. 6.

Real Multipliers, Quasi-optimal Polynomials, $T \geq 3$

The case $T \geq 3$ is much more difficult compare to the cases $T = 1, 2$. We were unable to employ harmonic analysis technique and had to use complex analysis methods. Remarkably enough, we were able to construct a family of polynomials that are optimal for $T = 1, 2$ and that produce the expected estimate for the multiplier range if $T \geq 3$.

Define the set of points

$$t_j = \frac{\pi(\sigma + T(2j-1))}{\tau + (N-1)T}, \quad j = 1, .., \frac{N}{2} \text{ (N-even)}, \left(\frac{N-1}{2} \text{ (N-odd)}\right)$$

and the generating polynomials

$$\eta_N(z) = z(z+1) \prod_{j=1}^{\frac{N-2}{2}} (z - e^{it_j})(z - e^{-it_j}), \quad \text{N-even;}$$

$$\eta_N(z) = z \prod_{j=1}^{\frac{N-1}{2}} (z - e^{it_j})(z - e^{-it_j}), \quad \text{N-odd.}$$

Writing $\eta_N(z)$ in a standard form

$$\eta_N(z) = z\sum_{j=1}^{N} c_j z^{j-1}$$

we can define the following three-parameter family of polynomials

$$q(z, T, \sigma, \tau) = K\sum_{j=1}^{N}\left(1 - \frac{1+(j-1)T}{2+(N-1)T}\right)c_j z^{j-1}, \tag{17}$$

where K is a normalization factor that makes $q(1, T, \sigma, \tau) = 1$. In the particular case $\sigma = \tau = 2$ K is given by

$$\frac{1}{K} = 2^{\frac{N-2}{2}}\prod_{j=1}^{\frac{N-2}{2}}(1-\cos t_j), \quad N \text{ even},$$

and

$$\frac{1}{K} = 2^{\frac{N-3}{2}}\prod_{j=1}^{\frac{N-1}{2}}(1-\cos t_j), \quad N \text{ odd}.$$

The polynomials (17) are substitutes for $q(z)$ in (12) and play the same role in the $T \geq 3$ scenario as Fejér polynomials do in the cases $T = 1, 2$. Because of that we call them quasi-optimal.

For any T and N, by choosing $\sigma = \tau = 2$ the relation (13) is valid for

$$\mu\left(\frac{T}{2+(N-1)T}\prod_{j=1}^{\frac{N-2}{2}}\cot^2\frac{t_j}{2}\right)^T < 1, \text{ N-even},$$

and

$$\mu\left(\prod_{j=1}^{\frac{N-1}{2}}\cot^2\frac{t_j}{2}\right)^T < 1, \text{ N-odd}.$$

Moreover, for large N the left hand side in the above inequalities is approximately

$$\frac{\mu}{N^2}\left(\pi^{\frac{2-T}{T}}\left(\Gamma\left(\frac{T+2}{2T}\right)\right)^2\right)^T \sim \pi^2\frac{\mu}{N^2}, \qquad T \to \infty.$$

The proof is work in preparation by the authors [11].

We conjecture that the coefficients we have found are actually optimal.

Conjecture A *Assume that N and T are given. Then the largest* μ *such that*

$$(-\mu, 1) \subset (\bar{\mathbb{C}} \backslash F_T(\bar{\mathbb{D}})^*$$

has a magnitude proportional to N^2 *and is achieved by picking* $q(z)$ *in* (12) *to be* $q(z, T, 2, 2)$.

In favor of this conjecture are numeric simulations and the fact that for $T = 1, 2$ the new family coincides with the polynomials that are optimal. Moreover, Figs. 7, 9 and 8, 10 are remarkably similar to Figs. 3, 5 and 4, 6 which correspond to the cases $T = 1, 2$.

Examples

Example of a Quasi-optimal Polynomial for $T = 3, N = 5$

Let us consider a numeric example $T = 3, N = 5$. In this case $t_1 = 5\pi/14$ and $t_2 = 11\pi/14$. The generating polynomial is

$$\eta(z) = z+2(\cos\frac{3\pi}{14}-\sin\frac{\pi}{7})z^2+2(1-\cos\frac{\pi}{7}+\sin\frac{\pi}{14})z^3+2(\cos\frac{3\pi}{14}-\sin\frac{\pi}{7})z^4+z^5.$$

The normalized factor is

$$K = \frac{1}{2}\left(1 - \cos\frac{5\pi}{14}\right)^{-1}\left(1 - \cos\frac{11\pi}{14}\right)^{-1} = 0.496\ldots$$

To get the region for the locations of μ one needs first to build a covering polynomial

$$F_3(z) = z(q(z, 3, 2, 2))^3 = K^3 z\left(\frac{13}{14} + \frac{10}{7}(\cos\frac{3\pi}{14} - \sin\frac{\pi}{7})z+ (1 - \cos\frac{\pi}{7} + \sin\frac{\pi}{14})z^2 + \frac{4}{7}(\cos\frac{3\pi}{14} - \sin\frac{\pi}{7})z^3 + \frac{1}{14}z^4\right)^3.$$

See Fig. 7. Then take the inverse and get the region displayed in Fig. 8, where the range for the multiplier is $\mu \in (-33, 1)$. Using the above polynomials one can get the estimates $I_5^{(3)} \geq -0.03$. We conjecture that these values are optimal.

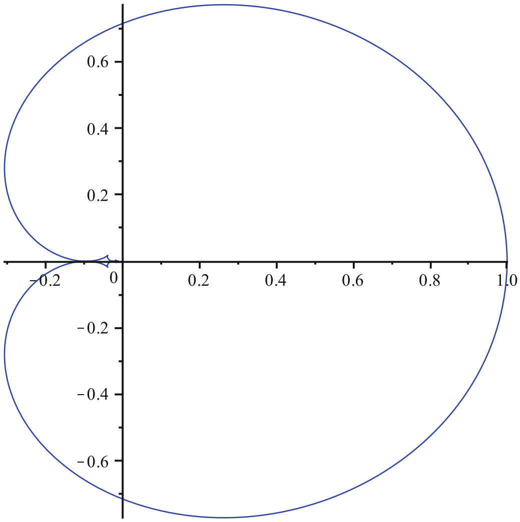


Fig. 7 $F_3(\mathbb{D})$, N=5

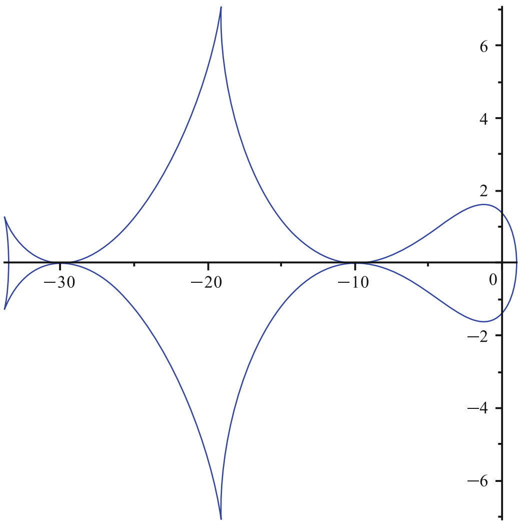


Fig. 8 $(\bar{\mathbb{C}}\backslash F_3(\bar{\mathbb{D}})^*$

Example of 8-Cycle in Logistic Equation

As an example of an application of the above method let us consider the logistic equation

$$x_{n+1} = 4x_n(1 - x_n). \tag{18}$$

It is well known that it has cycles of any length and that the cycles are unstable.

Consider the problem of *finding cycles of length 8*. Figure 9 displays the polynomial images of the unit disc $F_8(\mathbb{D})$ with the quasi-optimal polynomial of degree 12. Figure 10 displays the inverse image $(\bar{\mathbb{C}}\backslash F_8(\bar{\mathbb{D}})^*$.

Fig. 9 $F_8(\mathbb{D})$, N=12

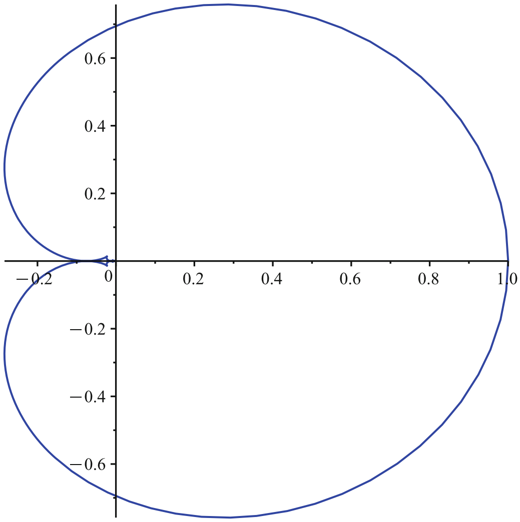
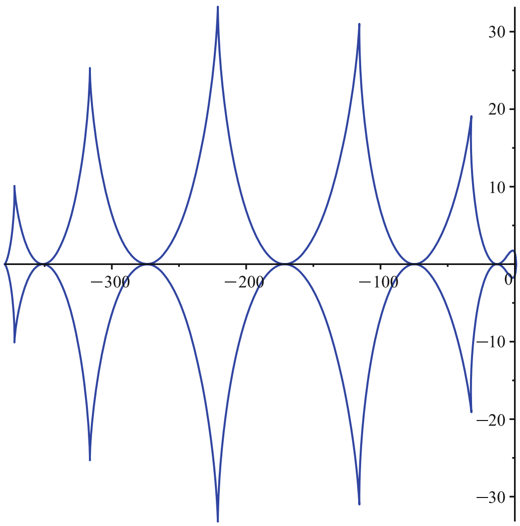


Fig. 10 $(\bar{\mathbb{C}}\backslash F_8(\bar{\mathbb{D}})^*$



To do that let us consider the system (18) and close it by the control

$$u_n = -\sum_{j=1}^{N-1} \varepsilon_j (f(x_{n-8j+8}) - f(x_{n-8j})).$$

We provide numeric simulation with the 8-cycle control above applied to the standard logistic equation. Figure 11 below reveals the existence of two 8-cycles. Moreover, the size of the control u_n goes to 0 as $n \to \infty$, so it provides an increasingly better approximation to the initial system. For $n \geq 9000$, for example,

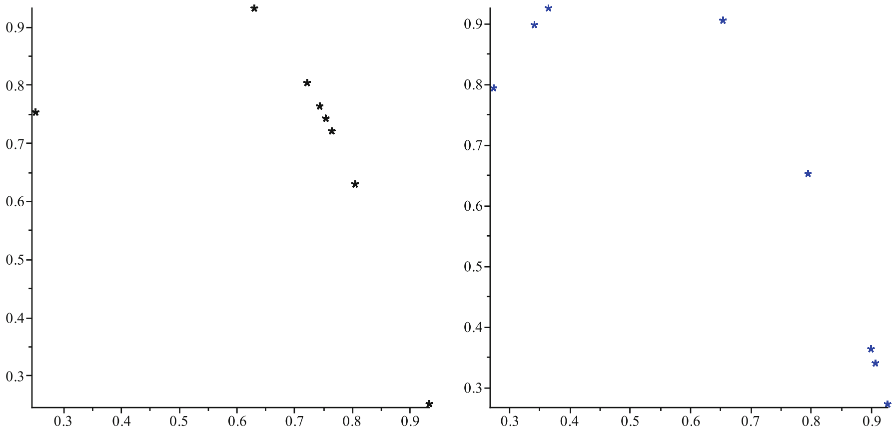


Fig. 11 Two 8-cycles in logistic equation on (x_{n-1}, x_n) plane

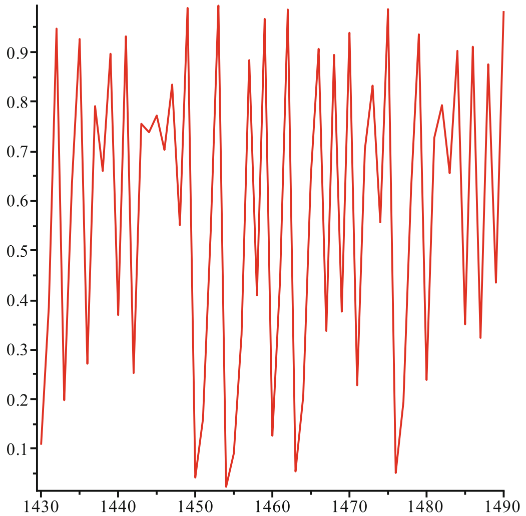


Fig. 12 Dynamics of logistic equation with $x_0 = 0.2518$

the control $|u_n| \leq 0.00002$, so we can obtain the value of the points of the two 8-cycles from Fig. 11 up to the fifth decimal:

{0.2518;0.7535;0.7429;0.7640; 0.7213;0.8042;0.6299;0.9325}, and

{0.3408;0.8987;0.3642;0.9262;0.2733;0.7944;0.6533;0.9059}.

By increasing the number of iterations n one can get more digits in the cycle.

The subtlety of the situation is well illustrated by the fact that knowledge of a point on a cycle does not guarantee that the whole cycle can be found numerically by the iterative procedure (18) because of chaotic behavior of the solutions to (18). Figure 12 demonstrates what happens when we plug in $x_0 = 0.2518$ in (18). Since the system is chaotic, the numerical simulations do not reveal the existence of the 8-cycle. On the other hand, after adding an 8-cycle control the solutions exhibit 8-periodicity, see Fig. 13.

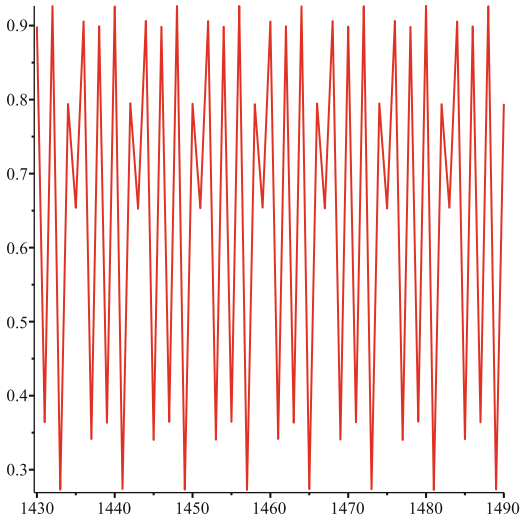


Fig. 13 Dynamics of the solution in the closed-loop system with $x_0 = 0.2518$

Let us note that the standard approach would be to search for equilibria of the 8-folded composition of the logistic map. However this new map is a polynomial of degree 2^8 and therefore one should consider 512 roots on the interval [0,1]. Identifying those roots is a serious practical problem.

Complex Multipliers, $\Re(\mu) < 0$.

Complex Multipliers, Case of Equilibrium

Let us assume that $\mu \in \{\Re(z) < 0\} \cup \{|z| < 1\}$. Note the we consider the unit disk because if eigenvalues are in the unit disk, we have stability without any control added.

This domain may be considered as a union of the domains $M_R := \{|z + R| < R\} \cup \{|z| < 1\}$. If $R = N/2$ then choosing the polynomial map

$$F(z) = \frac{2}{N} \sum_{j=1}^{N} (1 - \frac{j}{N+1}) z^j$$

we can guarantee that the image of the unit disc will be to the right of the line $\Re(z) = -1/N$. Therefore, $M_{N/2}$ will be included in $\left(\bar{\mathbb{C}} \backslash F(\bar{\mathbb{D}})\right)^*$. We illustrate this in Figs. 14 and 15.

Fig. 14 $F(\mathbb{D})$, N=12

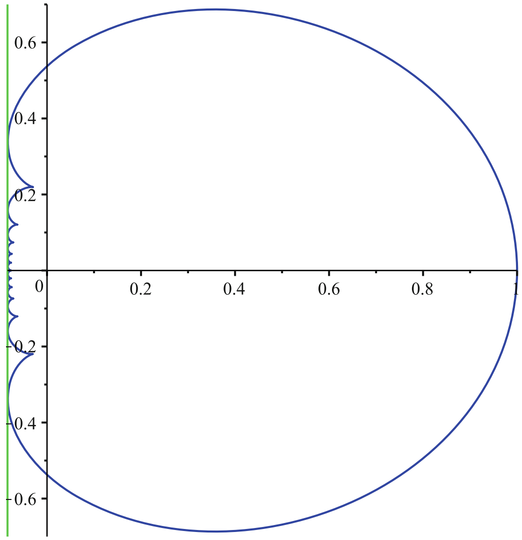
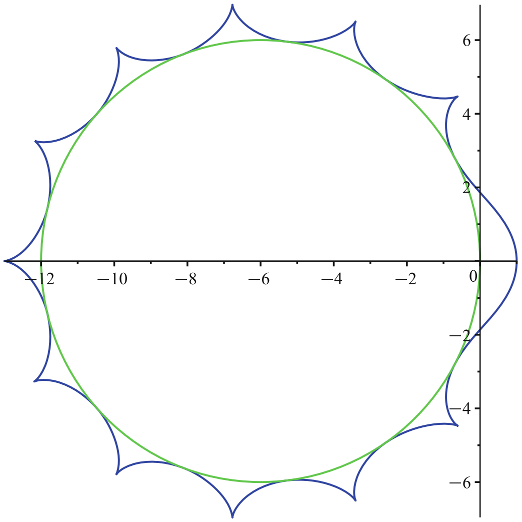


Fig. 15 $\left(\bar{\mathbb{C}} \backslash F(\bar{\mathbb{D}})\right)^*$



Burgers Map

As an example of an application let us consider the well-known Burgers map

$$\begin{cases} x_{n+1} = & 0.75x_n - y_n^2, \\ y_{n+1} = & 1.75y_n - x_n y_n. \end{cases}$$

Here different colors correspond to different initial values. The plot in Fig. 16 as well as several plots below are in the (x_n, x_{n+1}) coordinate plane. We can see that after adding the nonlinear control an equilibrium point is clearly revealed in Fig. 17.

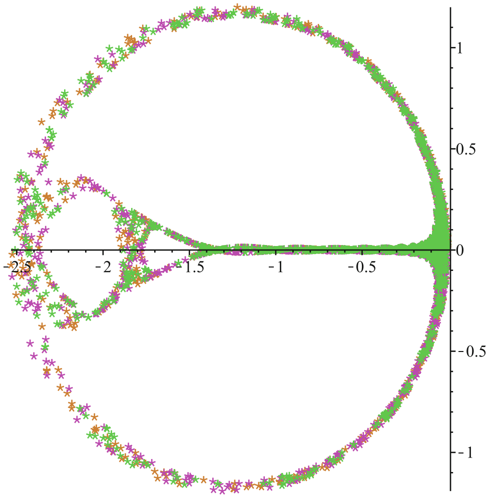


Fig. 16 Chaos

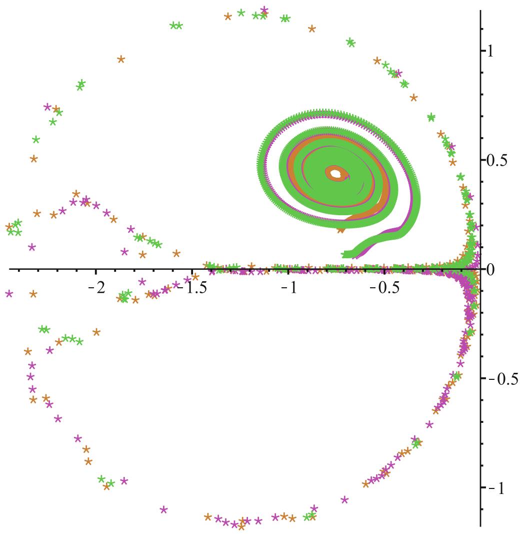


Fig. 17 Equilibrium, N = 55

Arnold's Cat Map

As another example, let us consider the famous Arnold's Cat map, Fig. 18

$$\begin{cases} x_{n+1} = & (x_n + y_n) \mod 1, \\ y_{n+1} = & (x_n + 2y_n) \mod 1. \end{cases}$$

Fig. 18 Chaos

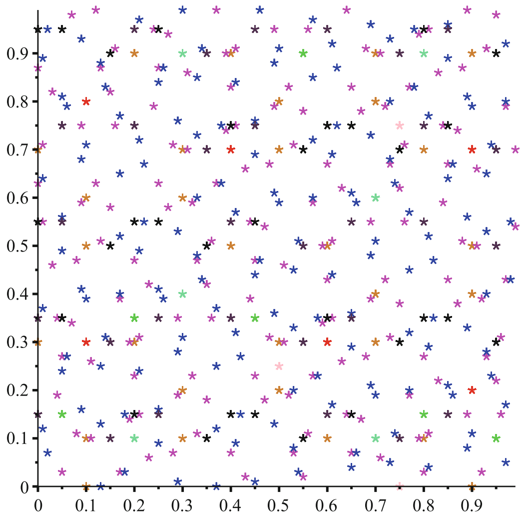
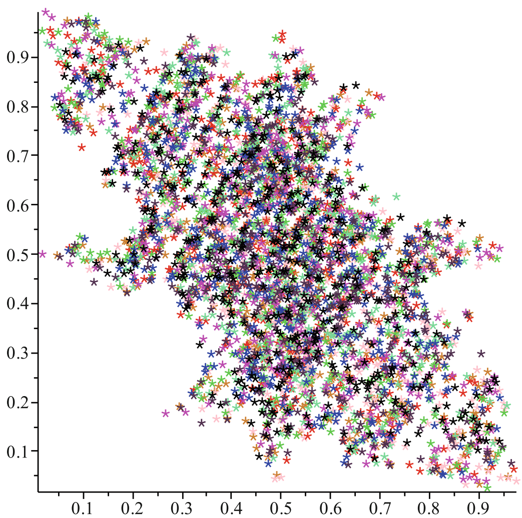


Fig. 19 $T = 1, N = 3$



Its multipliers are real numbers, one greater than 1 so it is unlikely that our method will stabilize the equilibrium, see Fig. 19. However, it still has a regularizing effect on the dynamics, as can be seen in the pictures below. It would be interesting to understand why different orbits, corresponding to different colors, end up being separated by the nonlinear control, see Fig. 20.

Fig. 20 T = 1, N = 50

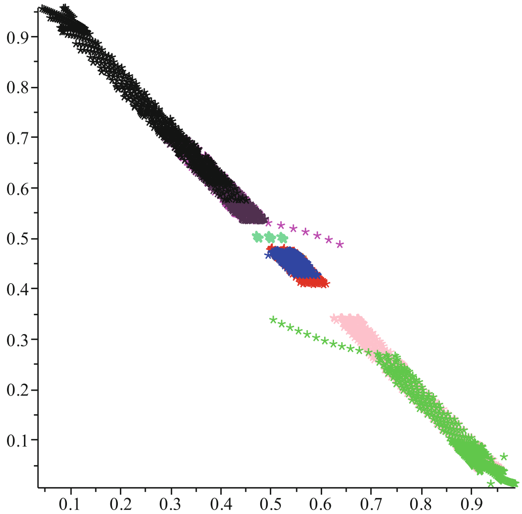
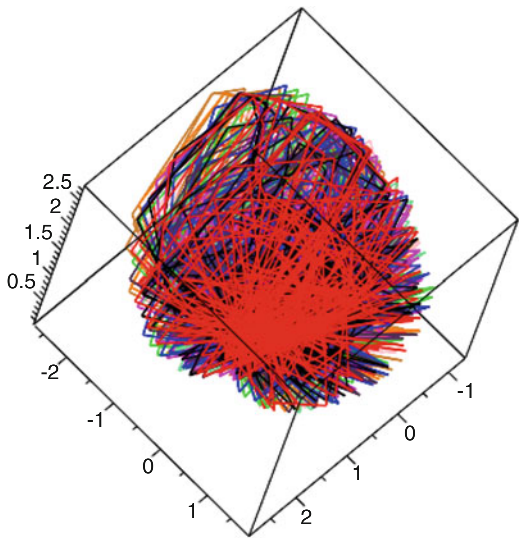


Fig. 21 Chaos



Ikeda 3D Map

Let us also look at the 3D Ikeda map, Fig. 21

$$\begin{cases} x_{n+1} = & 1 + 0.9\left(x_n \cos\left(0.4 - \frac{6}{1+z_n^2}\right) - y_n \sin\left(0.4 - \frac{6}{1+z_n^2}\right)\right), \\ y_{n+1} = & 0.9\left(x_n \sin\left(0.4 - \frac{6}{1+z_n^2}\right) + y_n \cos\left(0.4 - \frac{6}{1+z_n^2}\right)\right), \\ z_{n+1} = & \sqrt{(x_n-1)^2 + y_n^2}. \end{cases}$$

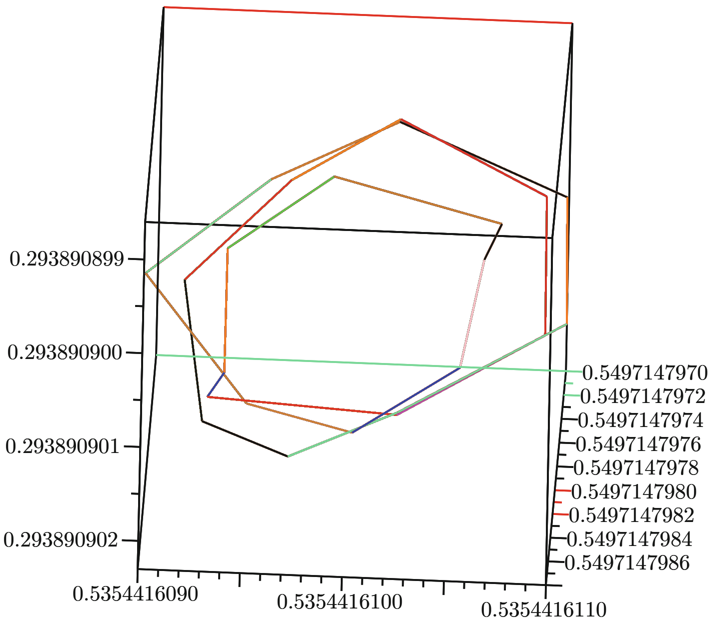


Fig. 22 Equilibrium, N = 3

One again notices an equilibrium, whose first few digits in the decimal expansion are $(0.5354416, 0.5497147, 0.2938909)$, see Fig. 22.

Complex Multipliers, $\Re(z) < 0,\ T \geq 2$

We illustrate the case of cycles of length 8, i.e. N = 12, T = 8, in Figs. 23 and 24. Here we use $F_T(z) = z\,(q(z, T, 1, 1))^T$.

In the following few subsections we illustrate how 4-cycles become visible after adding the nonlinear control.

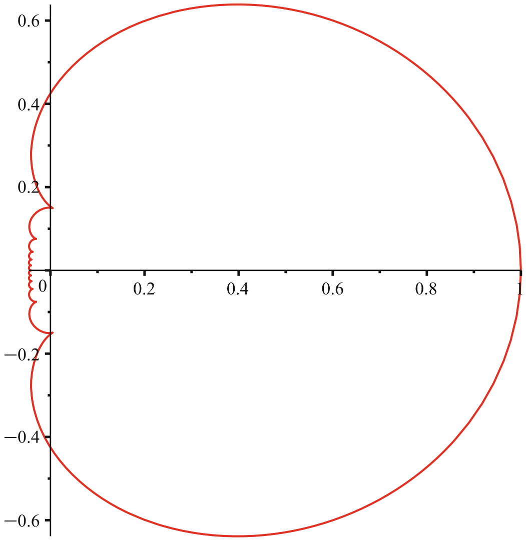


Fig. 23 $F_8(\mathbb{D})$, N = 12

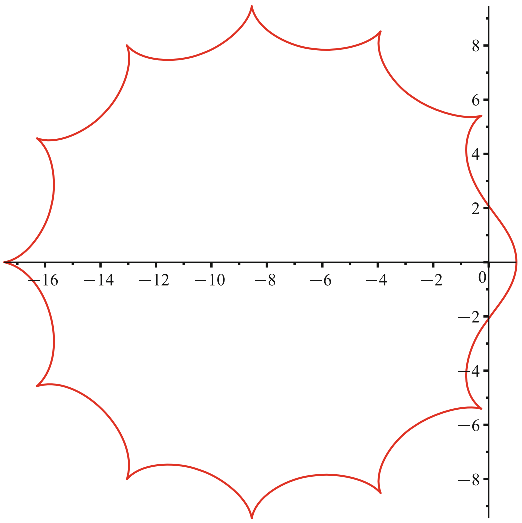


Fig. 24 $(\bar{\mathbb{C}} \backslash F_8(\bar{\mathbb{D}})^*$

Hennon Map, T = 4

Let us consider the Hennon map, Fig. 25

$$\begin{cases} x_{n+1} = & 1 - 1.4x_n^2 + y_n, \\ y_{n+1} = & 0.3x_n. \end{cases}$$

Note the appearance of 4-cycles after adding a nonlinear control in Fig. 26.

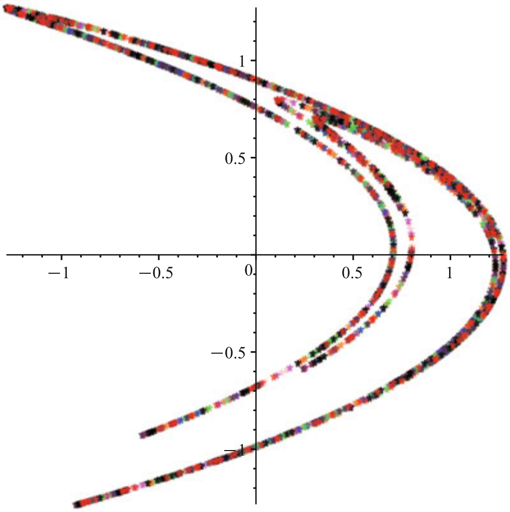


Fig. 25 Chaos

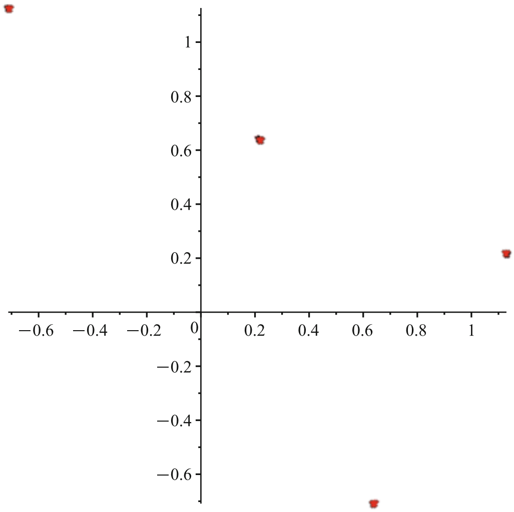


Fig. 26 4-cycle, N = 10

Ikeda Map, T = 4

The 2D Ikeda map is given by the system below and is displayed in Fig. 27 and the 4 cycle in Fig. 28

$$\begin{cases} x_{n+1} = & 1 + 0.9\left(x_n \cos\left(0.4 - \frac{6}{1+x_n^2+y_n^2}\right) - y_n \sin\left(0.4 - \frac{6}{1+x_n^2+y_n^2}\right)\right), \\ y_{n+1} = & 0.9\left(x_n \sin\left(0.4 - \frac{6}{1+x_n^2+y_n^2}\right) + y_n \cos\left(0.4 - \frac{6}{1+x_n^2+y_n^2}\right)\right), \end{cases}$$

Fig. 27 Chaos

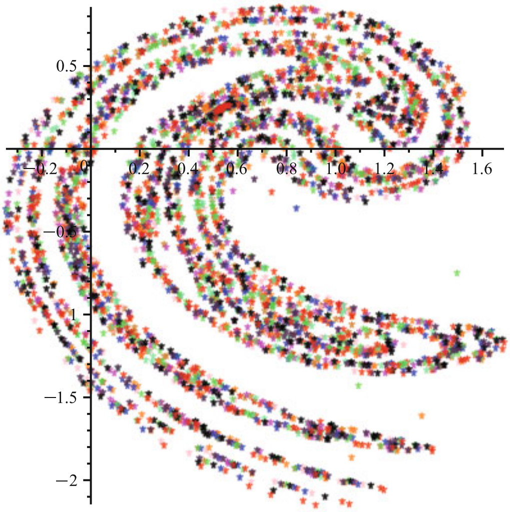
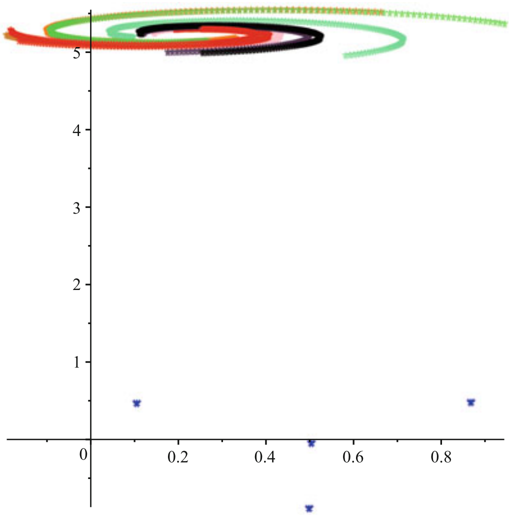


Fig. 28 4-cycle, N = 6



Lozi Map, T = 4

The Lozi map is defined by the system below and is displayed in Fig. 29 and the 4 cycle is in Fig. 30.

$$\begin{cases} x_{n+1} = 1 - 1.7|x| + 0.5y, \\ y_{n+1} = x_n. \end{cases}$$

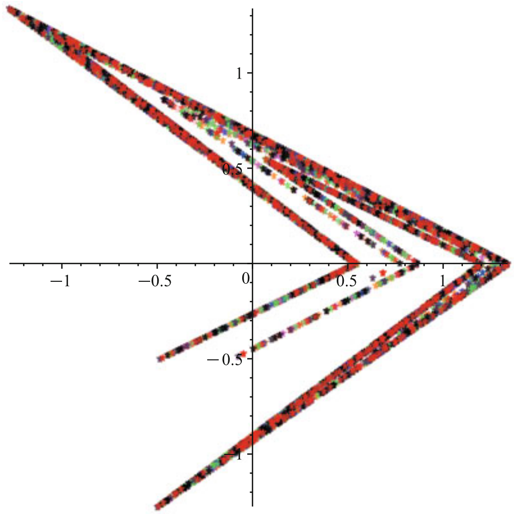


Fig. 29 Chaos

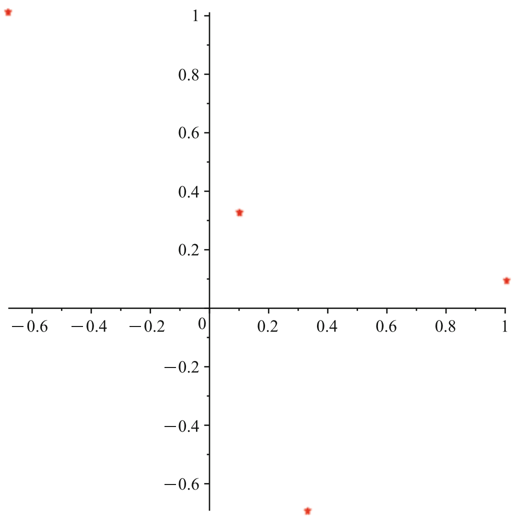


Fig. 30 4-cycle, N = 40

Holmes Cubic Map, T = 4

The Holmes cubic map is defined by the system below and is displayed in Fig. 31, the 4 cycle in Fig. 32.

$$\begin{cases} x_{n+1} = y_n \\ y_{n+1} = -0.2x + 2.77y - y^3. \end{cases}$$

Fig. 31 Chaos

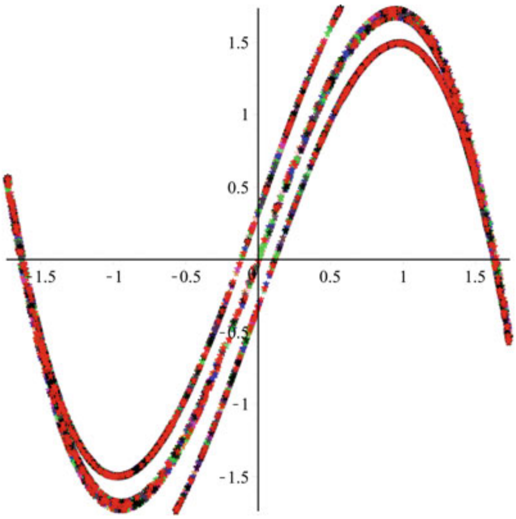
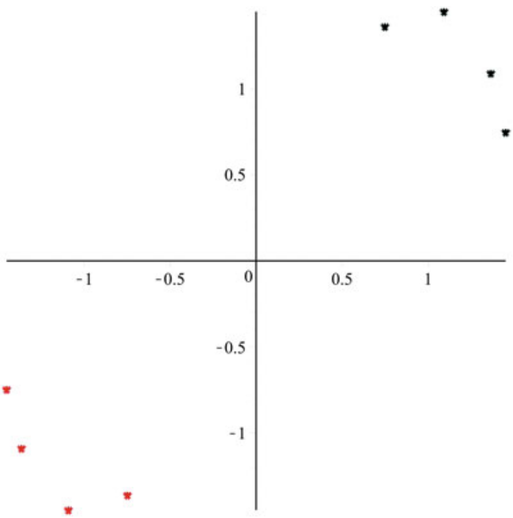


Fig. 32 Two 4-cycles, N = 18



Chirikov Map, T = 1 and T = 2

The map that describes the dynamics of the kicked rotor is given by the system below and is displayed in Fig. 33. Figures 34 and 35 display the dynamics of the systems closed by 1-cycle and 2-cycle controls correspondingly.

$$\begin{cases} x_{n+1} = 2\pi \lfloor \frac{x_n + y_n + 1.4 \sin x_n}{2\pi} \rfloor, \\ y_{n+1} = 2\pi \lfloor \frac{y_n + 1.4 \sin x_n}{2\pi} \rfloor. \end{cases}$$

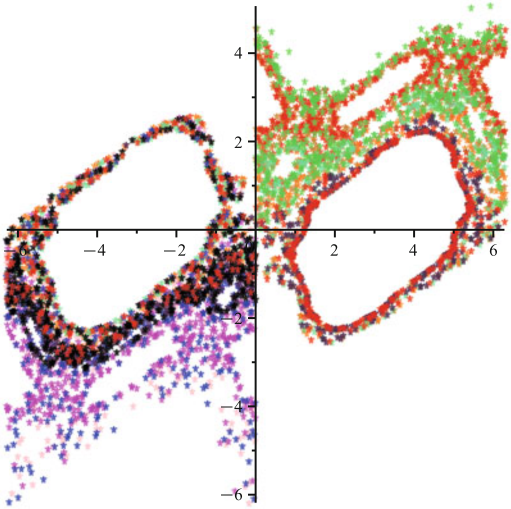


Fig. 33 Chaos

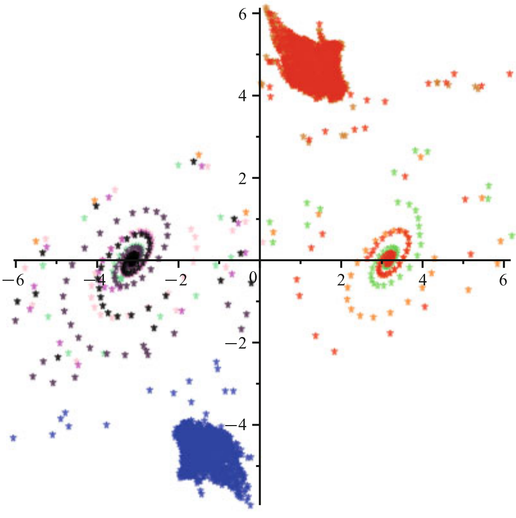


Fig. 34 $T = 1, N = 9$

Fig. 35 T = 2, N = 24

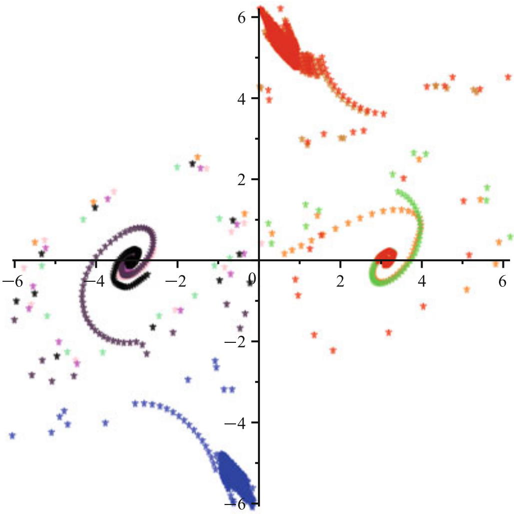
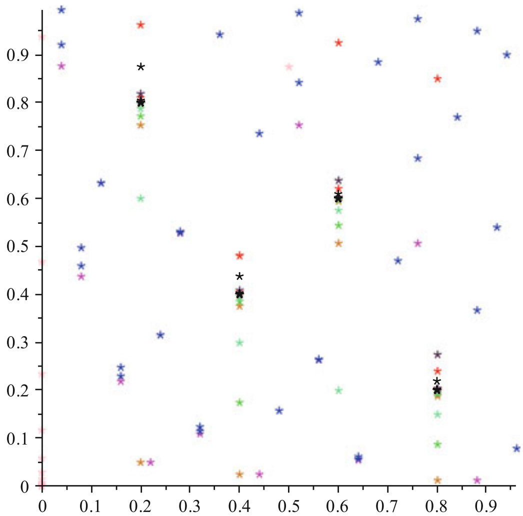


Fig. 36 Chaos

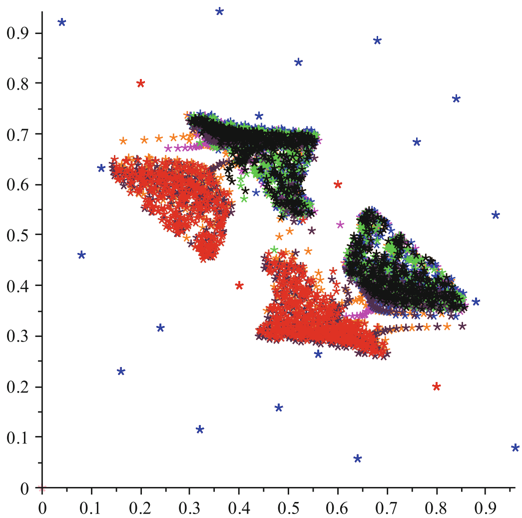


Baker's Map, T = 2

Baker's map is defined by the system below and is displayed in Fig. 36. The systems closed by 2-cycle control is displayed in Fig. 37.

$$\begin{cases} x_{n+1} = 2x_n - \lfloor 2x_n \rfloor, \\ y_{n+1} = \frac{1}{2}(y + \lfloor 2x_n \rfloor). \end{cases}$$

Fig. 37 T = 2, N = 20



It is yet another example of globally non-smooth map in addition to Arnold's Cat map and Chirikov map considered above. It is remarkable that in each case one can observe separation of initial values, thus a regularization of chaotic behavior.

Complex Multipliers, $\Re(\mu) > 0$

In this case the worst possible scenario consists of having real multipliers. Recall that even in the simplest system $x_{n+1} = \mu x_n$ the solution $x_n = C\mu^n$ blows up exponentially and our control cannot stabilize it, since there are no oscillations present. Therefore, it is natural to expect that the required N will grow very fast as the set of multipliers M gets close to the real line. The Alexander polynomials illustrate this hypothesis very well.

One piece of good news is that now we can use $q(z, T, \sigma, \tau)$ instead of Alexander polynomials. There is a choice of the parameters that allows the set (13) to cover any part of the region $\mathbb{C}\backslash[1, \infty)$ right to the imaginary axis as $N \to \infty$. The sets look like the angel wings in Fig. 39 below, or like the dragonfly wings in Fig. 41. The images of the unit discs are displayed on Figs. 38 and 40. The value of $N = 503$ is, of course, huge, but it is much better compare to $N = 20,000$ for Alexander polynomials. The values for N are selected to highlight the difference between the case of negative real part multipliers, where $N = 12$ suffices for very negative values of the real part, and the positive real part multipliers, where even for a relatively small multipliers a very large value of N is required.

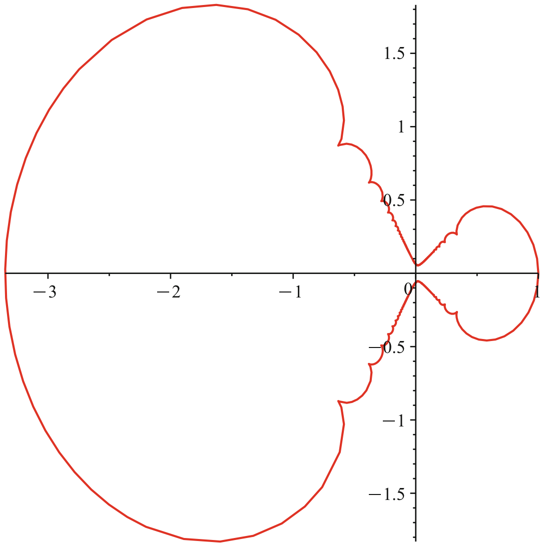


Fig. 38 $F_1(\bar{\mathbb{D}})$, $\sigma = 1.2$, $\tau = 0.5$, N $= 503$

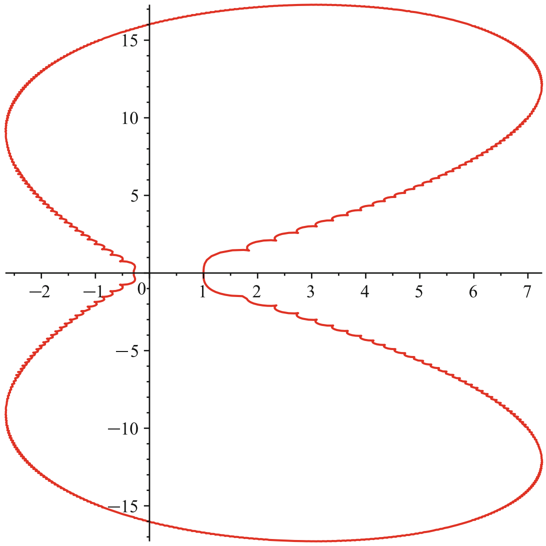


Fig. 39 $(\bar{\mathbb{C}} \backslash F_1(\bar{\mathbb{D}})^*$

Fig. 40 $F_1(\bar{\mathbb{D}})$, $\sigma = 1.9$, $\tau = 0.75$, N = 503

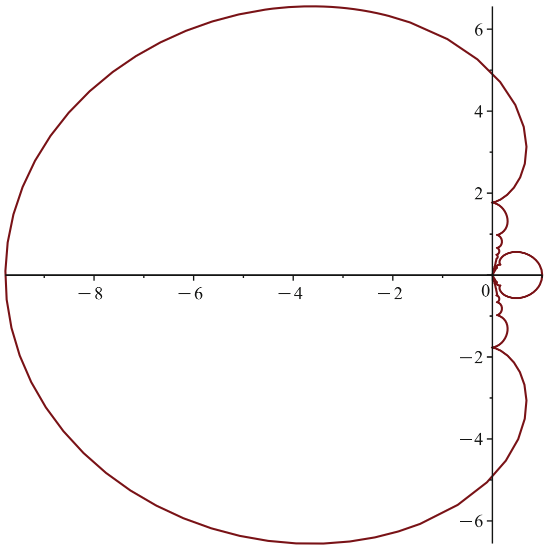
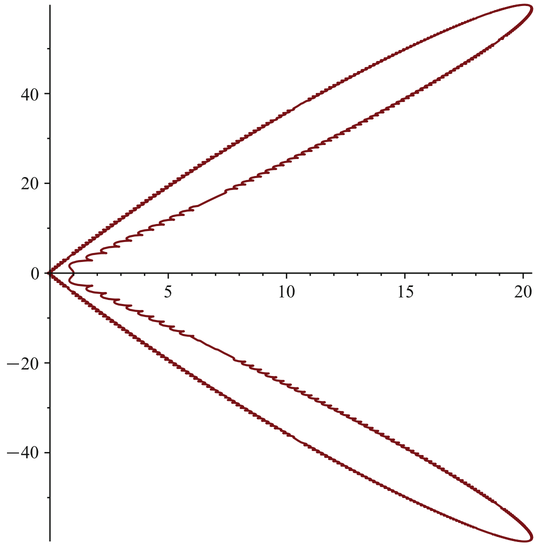


Fig. 41 $(\bar{\mathbb{C}} \backslash F_1(\bar{\mathbb{D}})^*$



Generalized Fejér Kernels

In analysis there are two types of extremal non-negative polynomials introduced by Fejér. In a closed form they can be written as

$$\Phi^{(1)}_{N-1}(t) = \left(\frac{\cos \frac{N+1}{2} t}{\cos t - \cos \frac{\pi}{N+1}} \right)^2 \quad \text{and} \quad \Phi^{(2)}_{N-1}(t) = \left(\frac{\sin \frac{N}{2} t}{\sin \frac{t}{2}} \right)^2 . \tag{19}$$

Graphically these two polynomials look similar; however, there is no explanation of that fact and no relation between these two families polynomials has been established.

Surprisingly, both polynomials turned out to be involved in the problem of optimal stability. It was showed in [8] that the polynomials $\Phi^{(1)}_{N-1}(t)$ play a central role in the problem of 1-cycle (equilibrium) stability, while $\Phi^{(2)}_{N-1}(t)$ are central to the 2-cycle stability.

The family of complex polynomials $q(z, T, 2, 2)$ that we introduced above generates a new family of trigonometric polynomials which contains both Fejér polynomials as particular cases. Denote $q(z, T, 2, 2)$ by $q^{(T)}_N(z)$ in this section, and their coefficients by $a^{(T)}_j$. Let

$$G^{(T)}_{N-1}(t) = \Im \left\{ \frac{e^{i\frac{t-\pi}{T}} q^{(T)}_N(e^{it})}{\sin \frac{t-\pi}{T}} \right\}, \ 0 < t < \pi.$$

One can check that

$$\frac{1}{G^{(T)}_{N-1}(0+)} \cdot G^{(T)}_{N-1}(t) = \frac{1}{\Phi^{(T)}_{N-1}(0)} \cdot \Phi^{(T)}_{N-1}(t), \ T = 1, 2.$$

Letting $\tau = (t - \pi)/T$ we obtain the normalized version

$$\tilde{G}^{(T)}_{N-1}(\tau) = \frac{1}{\sin \tau} \sum_{j=1}^{N} (-1)^{j-1} a^{(T)}_j \sin(1 + (j-1)T)\tau. \tag{20}$$

Some Properties of $\tilde{G}^{(T)}_{N-1}$

Note that

$$a^{(T)}_j = (1 + (j-1)T) a^{(T)}_{N-j+1},$$

and that

$$\tau_j = \frac{\pi(N - 2j)}{2 + (N-1)T}$$

are double roots of $\tilde{G}^{(T)}_{N-1}(\tau)$.

In Fig. 42 we display the plots of a few generalized Fejér polynomials (20). T = 1 coral, T = 2 blue, T = 3 green, T = 4 red.

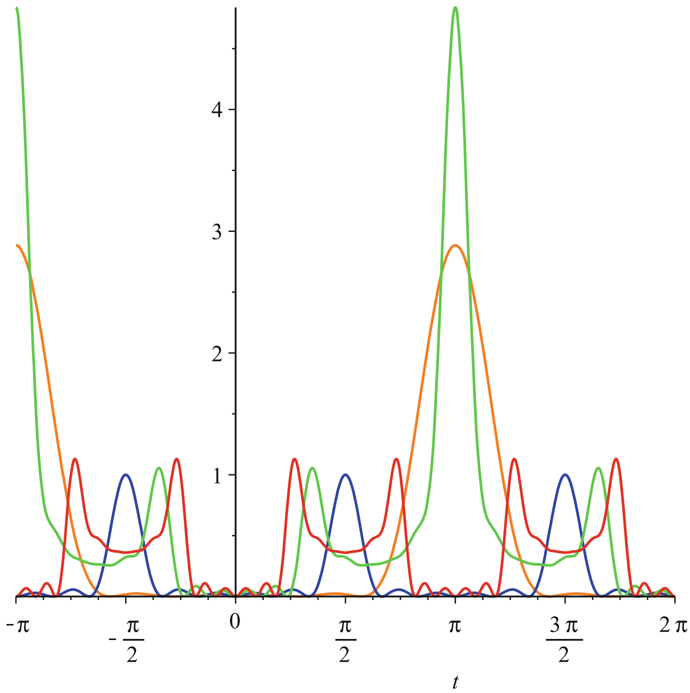


Fig. 42 Graphs of polynomials $\widetilde{G}_{N-1}^{(T)}(\tau)$, T=1,2,3,4

The classic Fejér polynomials are plotted for T $=$ 1 and T $=$ 2, however they are shifted by π and $\pi/2$ correspondently because of the τ substitution. These plots support the following

Conjecture B *The generalized Fejér polynomials*

$$\sum_{j=1}^{N}(-1)^{j-1}a_j^{(T)}\sin(1+(j-1)T)\tau$$

are non-negative on $[0,\pi]$.

Conclusion

We list below what we were able to prove, and what we conjecture to be true.

(i) We were able to find the optimal coefficients for real multipliers in the case of an equilibrium and 2-cycle $T=1,2$.

(ii) We were able to find the optimal coefficients for complex multipliers with $\Re(\mu) < 0$ in the case of an equilibrium $T = 1$.
(iii) We have found coefficients for real multipliers in the case of T-cycles, $T \geq 3$, which vastly improve the results obtained by using Alexander polynomials. We conjecture that the discovered coefficients are optimal.
(iv) We have suggested improved coefficients for the complex multipliers with $\Re(\mu) > 0$. We don't know whether they are optimal.
(v) We have found a family of trigonometric polynomials that contains Fejér polynomials as a particular case which we conjecture to be non-negative.

We believe that the developed techniques can also be useful in the attempts to solve the second part of Hilbert's 16th problem, which deals with the number and location of limit cycles of a planar polynomial vector field of degree n. The development of computational sciences made it possible to employ computers in this matter, thus discretizing the problem and reducing it to the problem of detecting of cycles of high periods in discrete settings, which is the topic of this article.

Remembrance by Alex Stokolos

I first met Cora in 1992 at the Miraflores Conference, Spain. That was just half a year after the collapse of the Soviet Union, a very difficult time in Ukraine. To participate in the meeting I got help from my friend Alexey Solyanik who was at the time a visitor at UAM and who has been a great influence on my whole career and life. At the banquet, Alexey and I sat at the same table as Cora and Carlos Segovia. I have been to many conferences and met many people since then, but I remember that night in Miraflores very well. Cora was just adorable, Carlos was great. They made fun of each other, it was an unforgettable play by two brilliant actors.

Cora told me that in her junior years, American mathematicians tried to help young talents abroad, and she gave the example of Zygmund lecturing in Argentina. She tried to follow this idea, helping many Argentinian, Venezuelan, and Ukrainian mathematicians.

In 1997, Cora took part in Krein's 90th birthday anniversary conference on operator theory and its applications that took place in Odessa, Ukraine. There she met her future collaborator Dmitry Kaliuzhnyi-Verbovetskyi. The article of Dmitry et al is in this volume. The article of another Odessiter, Nikolai Vasilevski, is included in this volume as well.

The next time I met Cora was in Williams College in 2001 while participating at a Special Session on "Harmonic Analysis since the Williamstown Conference of 1978". It was a remarkable idea of David Cruz-Uribe and Janine Wittwer to celebrate that milestone symposium. The two Williamstown's volumes conference proceedings have identified the development of harmonic analysis for many years. Some problems have been solved, many are waiting for a solution.

At that meeting Cora invited me to give a talk at Howard University. I went to Howard, gave a presentation for the seminar, met Cora's colleagues and her husband Daniel. Everything was wonderful. Then more meetings, more impressions, more memorial moments. Last time I saw Cora was at the 10th New Mexico Analysis Seminar in 2007. As always, she was surrounded by people; one had to find a path through the crowd even to say hello.

In 2010, news as a thunderbolt struck me - Cora passed away. I was shocked and depressed. Trying to pay our last tribute, my colleagues and I organized a session at the AMS meeting in Albuquerque in 2014; there was also an evening in memory of Cora during the 13th New Mexico Analysis seminar meeting. For this event I gave a lecture on stability of dynamical systems. I chose this topic because quite exciting and very subtle results in non-linear dynamics were obtained by means of classical harmonic and complex analysis - the subjects so close to Cora's heart.

Initially, the content of the talk in the preprint version was a gift to Alexey Solyanik. To mention the connection of Alexey to this volume allows me to point out that Paul Hagelstein's article in the volume is about Solyanik's estimates in harmonic analysis, the subject of Paul's talk at SEOUL ICM 2014. Since I indirectly met Cora thanks to Alexey, we decided to also dedicate this article to him.

References

1. J.W. Alexander, Functions which map the interior of the unit disc upon simple regions. Ann. Math. **17**, 12–22 (1915–1916)
2. S. Bielawski, D. Derozier, P. Glorieux, Controlling unstable periodic orbits by a delayed continuous feedback. Phys. Rev. E **49**, R971 (1994)
3. M.E. Bleich, J.E.S. Socolar, Stability of periodic orbits controlled by time-delay feedback. Phys. Lett. A **210**, 87 (1996)
4. S. Boccaletti, C. Grebogi, Y.-C. Lai, H. Mancini, D. Maza, The control of chaos: theory and applications. Phys. Rep. **329**, 103–197 (2000)
5. T. Dahms, P. Hovel, E. Scholl, Control of unstable steady states by extended time-delayed feedback. Phys. Rev. E **76**, 056201 (2007)
6. D. Dmitrishin, P. Hagelstein, A. Khamitova, A. Stokolos, On the stability of cycles by delayed feedback control. Linear Multilinear Algebra **64**(8), 1538–1549 (2014)
7. D. Dmitrishin, A. Khamitova, Methods of harmonic analysis in nonlinear dynamics. C. R. Math. **351**(9–10), 367–370 (2013)
8. D. Dmitrishin, A. Khamitova, A. Stokolos, Fejér polynomials and chaos, in *Special Functions, Partial Differential Equations, and Harmonic Analysis.* Springer Proceedings in Mathematics and Statistics, vol. 108 (Springer, Cham, 2014), pp. 49–75
9. D. Dmitrishin, A. Khamitova, A. Korenovskyi, A. Stokolos, Optimal stabilization of a cycle in nonlinear discrete systems. arXiv:1307.7369 [math.DS]
10. D. Dmitrishin, A. Khamitova, A. Stokolos, On the generalized linear and non-linear DFC in non-linear dynamics. arXiv:1407.6488 [math.DS]
11. D. Dmitrishin, A. Khamitova, A. Stokolos, M. Tohaneanu, Complex polynomials and cycles in nonlinear autonomous discrete dynamical systems (in preparation)
12. J. Hizanidis, R. Aust, E. Scholl: Delay-induced multistability near a globalbifurcation. Int. J. Bifur. Chaos **18**, 1759 (2008)

13. R.A. Horn, C. Johnson, Matrix analysis. (Cambridge University Press, Cambridge, 1985), p. 561
14. P. Hovel, J.E.S. Socolar, Stability domains for time-delay feedback control with latency. Phys. Rev. E **68**, 036–206 (2003)
15. P. Hovel, E. Scholl, Control of unstable steady states by time-delayed feedback methods. Phys. Rev. E **72**, 046–203 (2005)
16. W. Just, H. Benner, E. Scholl, Control of chaos by time–delayed feedback: a survey of theoretical and experimental aspects, in *Advances in Solid State Physics*, vol. 43, ed. by B. Kramer (Springer, Berlin, 2003), pp. 589–603
17. W. Just, T. Bernard, M. Ostheimer, E. Reibold, H. Benner, Mechanism of time-delayed feedback control. Phys. Rev. Lett. **78**, 203 (1997). Time-Delayed Feedback Control 63
18. W. Just, B. Fiedler, V. Flunkert, M. Georgi, P. Hovel, E. Scholl, Beyond odd number limitation: a bifurcation analysis of time-delayed feedback control. Phys. Rev. E **76**, 026–210 (2007)
19. Ö. Morgül, Further stability results for a generalization of delayed feedback control. Nonlinear Dyn. **70**(2), 1255–1262 (2012)
20. I.D. Murray, *Mathematical Biology. An Introduction*, 3rd edn. (Springer, New York, 2002), p. 551
21. E. Scholl, P. Hovel, V. Flunkert, M.A. Dahlem, H. Nakajima, On analytical properties of delayed feedback control of chaos. Phys. Lett. A **232**, 207 (1997)
22. E. Scholl, H. G. Schuster (eds.), *Handbook of Chaos Control* (Wiley- VCH, Weinheim, 2008). Second completely revised and enlarged edition
23. J.E.S. Socolar, D.W. Sukow, D.J. Gauthier, Stabilizing unstable periodic orbits in fast dynamical systems. Phys. Rev. E **50**, 32–45 (1994)
24. J.E.S. Socolar, D.J. Gauthier, Analysis and comparison of multiple-delay schemes for controlling unstable fixed points of discrete maps. Phys. Rev. E **57**, 65–89 (1998)
25. Y.-P. Tian, Z. Jiandong, Full characterization on limitation of generalized delayed feedback control for discrete-time systems. Phisica D **198**, 248–257 (2004)
26. d.S.M. Vieira, A.J. Lichtenberg, Controlling chaos using nonlinear feedback with delay. Phys. Rev. E **54**, 1200–1207 (1996)
27. S. Yanchuk, M. Wolfrum, P. Hovel, E. Scholl, Control of unstable steady states by long delay feedback. Phys. Rev. E **74**, 026201 (2006)

Smooth Analytic Functions and Model Subspaces

Konstantin M. Dyakonov

Dedicated to the memory of Cora Sadosky

Abstract The main themes of this survey are as follows: (a) the canonical (Riesz–Nevanlinna) factorization in various classes of analytic functions on the disk that are smooth up to its boundary, and (b) model subspaces (i.e., invariant subspaces of the backward shift) in the Hardy spaces H^p and in BMOA. It is the interrelationship and a peculiar cross-fertilization between the two topics that we wish to highlight.

2010 *Mathematics Subject Classification.* 30H10, 30H35, 30J05, 46E15, 46E35, 46J15, 47B35

Introduction

Our first topic in this survey is the multiplicative structure in spaces of *smooth analytic functions*. This phrase may sound somewhat redundant, if not downright confusing, since every analytic function is automatically smooth (in any reasonable sense) on its domain. The term becomes perfectly meaningful, though, if "smooth" is interpreted as "smooth up to the boundary". It is indeed the boundary smoothness of analytic functions that interests us here.

Supported in part by grants MTM2011-27932-C02-01, MTM2014-51834-P from El Ministerio de Economía y Competitividad (Spain) and grant 2014-SGR-289 from AGAUR (Generalitat de Catalunya).

K.M. Dyakonov (✉)
Departament de Matemàtiques i Informàtica, BGSMath and Universitat de Barcelona, Gran Via 585, E-08007 Barcelona, Spain

ICREA, Pg. Lluís Companys 23, E-08010 Barcelona, Spain
e-mail: konstantin.dyakonov@icrea.cat

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_9

Our functions will live on the disk $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. Putting the smoothness issue aside (but only for a short while), let us now recall a bit of function theory on the disk. Suppose that f is analytic on $\mathbb{D}$ and not too large near the unit circle $\mathbb{T} := \partial\mathbb{D}$. Specifically, assume that f lies in some *Hardy space* H^p with $0 < p \le \infty$. By definition, this means – in addition to analyticity – that

$$\|f\|_{H^p} := \sup_{0<r<1} \left(\int_{\mathbb{T}} |f(r\zeta)|^p dm(\zeta) \right)^{1/p} < \infty$$

if $0 < p < \infty$, or $\|f\|_{H^\infty} := \sup_{\mathbb{D}} |f| < \infty$ if $p = \infty$. Here and below, m denotes the normalized arclength measure on $\mathbb{T}$. It is well known that H^p functions have boundary values (nontangential limits) m-almost everywhere on $\mathbb{T}$. We may then identify H^p with a subspace of $L^p = L^p(\mathbb{T}, m)$ bearing in mind that the above norm, $\|\cdot\|_{H^p}$, agrees on H^p with the standard L^p-norm $\|\cdot\|_p$ over $\mathbb{T}$ (see [16, Chapter II]). When $0 < p < 1$, the two quantities should actually be called quasinorms rather than norms.

For f as above, the function $\varphi := |f|\big|_{\mathbb{T}}$ will satisfy $\varphi \in L^p$ and $\log\varphi \in L^1$. Moreover, these last two conditions characterize the moduli of H^p functions on $\mathbb{T}$. Now, letting $u := \log\varphi$ and writing $\mathcal{P}u$ for the harmonic extension (via the Poisson integral) of u from $\mathbb{T}$ into $\mathbb{D}$, we define the *outer function* $\mathcal{O}_\varphi$ as the (essentially unique) analytic function on $\mathbb{D}$ satisfying $\log|\mathcal{O}_\varphi(z)| = \mathcal{P}u(z)$. This done, we have $\mathcal{O}_\varphi \in H^p$ and $|\mathcal{O}_\varphi| = \varphi$ a.e. on $\mathbb{T}$. The ratio $f/\mathcal{O}_\varphi =: \theta$ will then be an *inner function*; that is, $\theta \in H^\infty$ and $|\theta| = 1$ a.e. on $\mathbb{T}$. Thus we arrive at the Canonical Factorization Theorem: the general form of an $f \in H^p$ is given by $f = \theta F$, where θ is inner and F outer (so that $F = \mathcal{O}_\varphi$ for some φ as above). A further factorization formula for inner functions allows us to express θ canonically in terms of its zeros $\{a_n\}$ (these are only required to satisfy $\sum_n(1 - |a_n|) < \infty$) and a certain singular measure μ on $\mathbb{T}$; see [16, Chapter II]. In summary, the original function $f \in H^p$ is fully described by the parameters φ, $\{a_n\}$ and μ that emerge; and any choice of parameters gives rise to an $f \in H^p$ via factorization.

The terms “inner function” and “outer function” were coined by Beurling. Why did he call them that? An amusing, but rather controversial, explanation I have heard is that the identity $f = \theta F$, when written in *this* specific form, has θ (the “inner factor”) inside and F (the “outer factor”) outside. Observe that in some noncommutative generalizations, which we do not touch upon, the order may become crucial; and yes, it should be θF rather than $F\theta$.

While quite a bit of modern 1-D complex analysis has evolved in an attempt to extend the H^p theory to *larger* analytic spaces, one also feels tempted to look at *smaller* (nicer) classes, in particular, at those populated by smooth analytic functions. Here, the good news is that the canonical factorization theorem applies. The bad news is, however, that the parameters cannot be chosen freely. Indeed, most inner functions – actually, all the “interesting” (i.e., nonrational) ones – are highly oscillatory, hence discontinuous, at some points of $\mathbb{T}$. Consequently, the product θF may only be smooth on $\mathbb{T}$ if the outer factor, F, is good enough and kills the singularities of the (bad) inner factor, θ. To find an explicit quantitative expression of this interplay, for a given "smooth analytic space", is therefore one problem to be dealt with.

Our second topic is the *model subspaces*, alias *star-invariant subspaces*, in H^p and in BMOA $:= \text{BMO} \cap H^1$, where BMO $= \text{BMO}(\mathbb{T})$ is the space of functions of bounded mean oscillation on $\mathbb{T}$ (see [16, Chapter VI]). In H^2, the model subspace K_θ generated by an inner function θ is, by definition, the orthogonal complement of the shift-invariant subspace θH^2. Thus,

$$K_\theta \left(= K_\theta^2\right) := H^2 \ominus \theta H^2. \tag{1}$$

It is a reproducing kernel Hilbert space, whose kernel function k_z associated with a point $z \in \mathbb{D}$ is given by

$$k_z(\zeta) = \frac{1 - \overline{\theta(z)}\theta(\zeta)}{1 - \bar{z}\zeta}.$$

This last function is therefore in K_θ for every z, and every $f \in K_\theta$ satisfies

$$f(z) = \int_{\mathbb{T}} f(\zeta)\overline{k_z(\zeta)}\, dm(\zeta), \qquad z \in \mathbb{D}.$$

It is straightforward to verify that $K_\theta = H^2 \cap \theta\, \overline{H_0^2}$, and we further define K_θ^p (the H^p-analogue of K_θ) by putting

$$K_\theta^p := H^p \cap \theta\, \overline{H_0^p}, \qquad 1 \le p \le \infty,$$

where $H_0^p := \{f \in H^p : f(0) = 0\}$ and the bar denotes complex conjugation. For smaller p's, a more reasonable definition appears to be

$$K_\theta^p := \text{clos}_{H^p} K_\theta, \qquad 0 < p < 1.$$

These subspaces play a crucial role in the Sz.-Nagy–Foiaş operator model (see [20]), which accounts for the terminology. Now, the term "star-invariant" means invariant under the backward shift operator $f \mapsto (f - f(0))/z$, and it follows from Beurling's theorem (see [16, Chapter II]) that the general form of a closed and nontrivial star-invariant subspace in H^2 is indeed given by (1), with θ inner. A similar fact is true for H^p when $1 \le p < \infty$.

Finally, we put

$$K_{*\theta} := K_\theta \cap \text{BMOA}.$$

When equipped with the BMO-norm $\|\cdot\|_*$, $K_{*\theta}$ becomes a star-invariant subspace of BMOA; in fact, it is the annihilator in BMOA of the shift-invariant subspace θH^1 in H^1. Of course, $K_{*\theta}$ contains K_θ^∞ and is contained in every K_θ^p with $0 < p < \infty$.

While each of the two topics just mentioned has received quite a bit of attention in its own right, the intimate interconnection between them does not seem to have been noticed (until recently) or explored in any detail. It is precisely the systematic exploitation of this interrelationship, perhaps a kind of duality, between the two subjects that is characteristic of our approach. In fact, the three stories told in the next three sections are intended to show that results and methods pertaining to one of our themes cast new light on the other, and vice versa.

Before moving any further, we need to recall the notions of Toeplitz and Hankel operators, since these will be crucial in what follows. We let P_+ and P_- denote the orthogonal projections from L^2 onto H^2 and onto $\overline{H_0^2}$, respectively. Thus,

$$(P_+F)(z) := \sum_{n\geq 0}\widehat{F}(n)z^n \quad \text{and} \quad (P_-F)(z) := \sum_{n<0}\widehat{F}(n)z^n,$$

where $\widehat{F}(n) := \int_{\mathbb{T}} F(\zeta)\overline{\zeta}^n dm(\zeta)$ is the nth Fourier coefficient of F. These operators are then extended to L^p with $1 < p < \infty$ (in which case they become bounded projections onto H^p and $\overline{H_0^p}$, the classical M. Riesz theorem tells us) and furthermore to L^1 (even though $P_\pm(L^1) \not\subset L^1$). Next, given a measurable function ψ on $\mathbb{T}$, we write

$$T_\psi f := P_+(\psi f) \quad \text{and} \quad H_\psi f := P_-(\psi f),$$

whenever $f \in H^1$ and $\psi f \in L^1$. The mapping T_ψ (resp., H_ψ) is called the *Toeplitz* (resp., *Hankel*) *operator with symbol* ψ.

In the special case where ψ is analytic (i.e., $\psi \in H^1$), T_ψ reduces to the multiplication map $f \mapsto f\psi$, defined at least on H^∞. The Toeplitz operators with symbols in $\overline{H^1}$ are said to be *coanalytic*. It is also worth mentioning that the model subspace K_θ^p (where $p \geq 1$) or $K_{*\theta}$, with θ an inner function, is precisely the kernel of the coanalytic Toeplitz operator $T_{\overline{\theta}}$ acting on H^p or BMOA.

Because Toeplitz and Hankel operators were among Cora Sadosky's best beloved mathematical creatures, their appearance in this survey seems to be appropriate (and is, anyway, far from incidental to the subject matter).

We conclude this introduction with a brief outline of the rest of the paper. In sections "Factorization in Lipschitz–Zygmund Spaces" and "Factorization in Dirichlet-Type Spaces", we look at certain smooth analytic spaces X and seek to characterize the pairs (f, θ), with $f \in X$ and θ inner, which satisfy

$$f\theta \in X. \tag{2}$$

Sometimes it is more natural to replace (2) by

$$f\theta^k \in X \text{ for all } k \in \mathbb{N}, \tag{3}$$

and we are led to consider some other related conditions as well. In section "Factorization in Lipschitz–Zygmund Spaces", the role of X is played by the analytic Lipschitz–Zygmund spaces A^α (see the beginning of that section for definitions), and the pairs (f, θ) with property (3) are then explicitly described by a certain *smallness condition*, to be imposed on $|f|$ near the singularities of θ. Furthermore, the same smallness condition ensures that the multiplication operator $g \mapsto fg$ acts nicely on the model space K_θ^p, or perhaps on $K_{\theta^n}^p$ with n suitably large, by improving integrability properties of the functions therein. For instance, given $1 < p < q < \infty$ and $\alpha = p^{-1} - q^{-1}$, we prove that multiplication by a function $f \in A^\alpha$ maps K_θ^p into H^q if and only if it maps θ into A^α (so that (2) holds with $X = A^\alpha$). The case of smaller p's and larger α's leads to a minor complication involving (3) in place of (2), and $K_{\theta^n}^p$ in place of K_θ^p.

In section "Factorization in Dirichlet-Type Spaces", our space X is chosen from among the so-called Dirichlet-type spaces. Each of these is formed by the functions $f \in H^2$ whose coefficient sequence, $\{\widehat{f}(n)\}$, lies in a certain weighted ℓ^2. An important special case is the classical *Dirichlet space* $\mathcal{D}$, the set of analytic functions f on $\mathbb{D}$ whose derivative, f', is square integrable over $\mathbb{D}$ with respect to the normalized area measure A; the (semi)norm $\|f\|_{\mathcal{D}}$ is then defined to be $\left(\int_{\mathbb{D}} |f'|^2 dA\right)^{1/2}$. Among other things we recover, for $f \in \mathcal{D}$ and θ inner, the identity

$$\|f\theta\|_{\mathcal{D}}^2 = \|f\|_{\mathcal{D}}^2 + \int_{\mathbb{T}} |f|^2 |\theta'| dm, \tag{4}$$

which forms part of Carleson's celebrated formula from [4]. Moreover, we obtain similar – but more sophisticated – formulas for general Dirichlet-type spaces; these yield the smallness conditions on f (in relation to θ) that are responsible for the interplay between the two factors in (2), for the current choices of X. When $X = \mathcal{D}$, the corresponding smallness condition reads $\int_{\mathbb{T}} |f|^2 |\theta'| dm < \infty$, as readily seen from (4). Our approach to (4) is based on the fact that the quantity $\|f\theta\|_{\mathcal{D}}$ coincides with the Hilbert–Schmidt norm of the Hankel operator $H_{\overline{f\theta}}$ acting from H^2 to $\overline{H_0^2}$ (and similarly for f in place of $f\theta$). Now let $\{g_n\}$ be an orthonormal basis in the model subspace K_θ. Since $H^2 = \theta H^2 \oplus K_\theta$, the family $\{\theta z^k\}_{k \geq 0} \cup \{g_n\}$ is an orthonormal basis in H^2, and we may use it to compute the Hilbert–Schmidt norm of $H_{\overline{f\theta}}$. This gives

$$\|f\theta\|_{\mathcal{D}}^2 = \sum_{k \geq 0} \left\| H_{\overline{f\theta}}(\theta z^k) \right\|_2^2 + \sum_n \left\| H_{\overline{f\theta}} g_n \right\|_2^2,$$

and a further calculation shows that the two sums above reduce to the two terms on the right-hand side of (4). A modification of the same technique allows us to handle the case of a generic Dirichlet-type space.

In section "Model Subspaces in BMOA", we consider coanalytic Toeplitz operators on the model subspace $K_{*\theta}$, and we obtain a criterion for such an operator to act boundedly from $K_{*\theta}$ to a given analytic space X, under certain assumptions on the latter. Precisely speaking, the spaces X that arise here naturally are those which enjoy the *K-property* of Havin. In other words, it will be assumed that every Toeplitz operator $T_{\overline{h}}$ with $h \in H^\infty$ maps X boundedly into itself and satisfies $\|T_{\overline{h}}\|_{X\to X} \le \text{const}\cdot\|h\|_\infty$. This property was introduced by Havin in [17], where he also verified it for a number of smooth analytic spaces. (It was further observed in [17] that every space X with the K-property admits division by inner factors: whenever $f \in X$ and I is an inner function such that $f/I \in H^1$, it follows that $f/I \in X$.) Now, the appearance of the K-property in connection with model subspaces of BMOA seems to reveal yet another link between the two topics of concern.

The content of section "Factorization in Lipschitz–Zygmund Spaces" is essentially borrowed from the author's papers [7, 8], while sections "Factorization in Dirichlet-Type Spaces" and "Model Subspaces in BMOA" are based on [10] and [12], respectively. It seems that a bit of self-plagiarism is unavoidable – and hopefully pardonable – under the circumstances.

Factorization in Lipschitz–Zygmund Spaces

This section deals with the *Lipschitz–Zygmund spaces* $\Lambda^\alpha = \Lambda^\alpha(\mathbb{T})$ and their analytic subspaces A^α. For $0 < \alpha < \infty$, the space Λ^α is defined as the set of all (complex-valued) functions $f \in C(\mathbb{T})$ that satisfy

$$\|\Delta_h^n f\|_\infty = O(|h|^\alpha), \qquad h \in \mathbb{R}, \tag{5}$$

where $\|\cdot\|_\infty$ is the sup-norm on $\mathbb{T}$, n is an integer with $n > \alpha$, and Δ_h^n denotes the nth order difference operator with step h. (As usual, the difference operators Δ_h^k are defined by induction: one puts $(\Delta_h^1 f)(\zeta) := f(e^{ih}\zeta) - f(\zeta)$ and $\Delta_h^k f := \Delta_h^1\Delta_h^{k-1} f$.) It is well known that property (5) does not depend on the choice of n, as long as $n > \alpha$, except possibly for the constant in the O-condition.

The corresponding analytic subspaces are

$$A^\alpha := \Lambda^\alpha \cap H^\infty, \qquad 0 < \alpha < \infty.$$

Equivalently, by a theorem essentially due to Hardy and Littlewood, A^α is formed by those holomorphic functions f on $\mathbb{D}$ which obey the condition

$$|f^{(n)}(z)| = O\left((1-|z|)^{\alpha-n}\right), \qquad z \in \mathbb{D},$$

for some (and then every) integer n with $n > \alpha$; here $f^{(n)}$ is the nth order derivative of f. The spaces Λ^α and A^α are then normed in a natural way.

The main result of this section is Theorem 2.1 below, which characterizes the pairs (f, θ), with $f \in A^\alpha$ and θ inner, such that f admits multiplication and/or division by every power of θ in Λ^α. The characterization involves an explicit quantitative condition saying that $|f(z)|$ must decay at a certain rate as z approaches the boundary along the sublevel set

$$\Omega(\theta, \varepsilon) := \{z \in \mathbb{D} : |\theta(z)| < \varepsilon\} \tag{6}$$

with $0 < \varepsilon < 1$. Moreover, it turns out that the same decay condition provides a criterion for the multiplication operator $T_f : g \mapsto fg$ to map the model subspace $K^p_{\theta^n}$ continuously into H^q, once the exponents are related appropriately.

Theorem 2.1 *Suppose that* $0 < p < \infty$, $\max(1, p) < q < \infty$, $\alpha = p^{-1} - q^{-1}$, *and* n *is an integer with* $np > 1$. *Assume also that* $f \in A^\alpha$ *and* θ *is an inner function. The following conditions are equivalent:*

(i) $f\overline{\theta}^k \in \Lambda^\alpha$ *for all* $k \in \mathbb{N}$.
(ii) $f\overline{\theta}^n \in \Lambda^\alpha$.
(iii) *The multiplication operator* T_f *maps* $K^p_{\theta^n}$ *boundedly into* H^q.
(iv) *For some (or every)* $\varepsilon \in (0, 1)$, *one has*

$$|f(z)| = O((1 - |z|)^\alpha) \quad \textit{for} \quad z \in \Omega(\theta, \varepsilon). \tag{7}$$

(v) $f\theta^k \in A^\alpha$ *for all* $k \in \mathbb{N}$.
(vi) $f\theta^n \in A^\alpha$.

It should be noted that the set $\Omega(\theta, \varepsilon)$ hits $\mathbb{T}$ precisely at those points which are singular for θ. Thus, (7) tells us how strongly the good factor f must vanish on the bad set of the problematic (nonsmooth) factor θ in order that the products in question be appropriately smooth.

Postponing the proof for a while, we first establish a few preliminary facts to lean upon. To begin with, we recall the Duren–Romberg–Shields theorem (see [6]) which allows us to identify A^α with the dual of the Hardy space H^r, where $r = (1 + \alpha)^{-1}$, under the pairing

$$\langle \varphi, \psi \rangle = \int_{\mathbb{T}} \varphi \overline{\psi} \, dm.$$

For a given $\psi \in A^\alpha$, the integral above is well defined at least when $\varphi \in H^\infty$, and we have

$$|\langle \varphi, \psi \rangle| \le c_\alpha \|\varphi\|_r \|\psi\|_{\Lambda^\alpha}$$

with some constant $c_\alpha > 0$. Moreover, the norm of the functional induced by ψ on H^r is actually comparable to $\|\psi\|_{\Lambda^\alpha}$.

The next three lemmas exploit this duality relation. The first of these was established by Havin in [17]; we also cite Shamoyan [24] in connection with part (b) below.

Lemma 2.2 *Let* $0 < \alpha < \infty$.

(a) *If* $h \in H^\infty$*, then the Toeplitz operator* $T_{\overline{h}}$ *maps the space* A^α *boundedly into itself, with norm at most* $\mathrm{const} \cdot \|h\|_\infty$.
(b) *If* $f \in H^1$ *and* θ *is an inner function such that* $f\theta \in A^\alpha$*, then* $f \in A^\alpha$ *and* $\|f\|_{\Lambda^\alpha} \le \mathit{const} \cdot \|f\theta\|_{\Lambda^\alpha}$.

The constants are allowed to depend only on α.

In Havin's terminology, statements (a) and (b) can be rephrased by saying that A^α has the K-property and the (weaker) f-property, respectively. To prove (a), one notes that $T_{\overline{h}}$ is the adjoint of the multiplication operator $T_h : g \mapsto gh$, which is obviously bounded on H^r with norm at most $\|h\|_\infty$. To deduce (b) from (a), observe that $f = T_{\overline{\theta}}(f\theta)$.

Lemma 2.3 *Suppose that* $0 < p < \infty$, $\max(1,p) < q < \infty$*, and* $\alpha = p^{-1} - q^{-1}$*. If* $f \in A^\alpha$*, then the Hankel operator* $H_{\overline{f}}$*, defined by*

$$H_{\overline{f}}g = P_-(\overline{f}g), \qquad g \in H^\infty,$$

can be extended to a bounded linear operator mapping H^p *into* $\overline{H^q_0}$.

Proof Put $r = (1+\alpha)^{-1}$ and $q' = q/(q-1)$. Given $g \in H^\infty$ and $h \in H^{q'}_0$, we have

$$\left|\int_{\mathbb{T}} \left(H_{\overline{f}}g\right) h\, dm\right| = \left|\int_{\mathbb{T}} P_-(\overline{f}g) \cdot h\, dm\right| = \left|\int_{\mathbb{T}} \overline{f}gh\, dm\right|$$
$$\le c_\alpha \|f\|_{\Lambda^\alpha} \|gh\|_r \le c_\alpha \|f\|_{\Lambda^\alpha} \|g\|_p \|h\|_{q'}.$$

Here, the last two inequalities rely on the Duren–Romberg–Shields duality theorem and on Hölder's inequality. Taking the supremum over the unit-norm functions h in $H^{q'}_0$, we obtain

$$\|H_{\overline{f}}g\|_q \le c_\alpha \|f\|_{\Lambda^\alpha} \|g\|_p,$$

which proves the required result. □

Lemma 2.4 *Suppose that* $0 < p < \infty$, $\max(1,p) < q \le \infty$*, and* $\alpha = p^{-1} - q^{-1}$*. Further, let* $f \in H^2$ *and let* θ *be an inner function. If* $P_-(f\overline{\theta}) \in \Lambda^\alpha$*, then the operator* $T_{\overline{f}}\big|_{K^\infty_\theta}$ *can be extended to a bounded linear operator acting from* K^p_θ *to* H^q.

Proof Given $g \in K^\infty_\theta$, put $h := \overline{z}\overline{g}\theta$ (so that $h \in H^\infty$) and $\psi := P_-(f\overline{\theta})$. The elementary identity

$$\overline{P_+F} = zP_-(\overline{z}\overline{F}), \qquad F \in L^2,$$

shows that $\overline{T_{\overline{f}}g} = zH_\psi h$. Using Lemma 2.3, we get

$$\|T_{\overline{f}}g\|_q = \|H_\psi h\|_q \le \mathrm{const} \cdot \|\psi\|_{\Lambda^\alpha} \|h\|_p = \mathrm{const} \cdot \|\psi\|_{\Lambda^\alpha} \|g\|_p,$$

which completes the proof. □

As a final preliminary result, we list some facts about the so-called Carleson curves associated with an inner function; see [16, Chapter VIII] for a proof.

Lemma 2.5 *Given an inner function θ and a number $\varepsilon \in (0,1)$, there exists a countable (possibly finite) system $\Gamma_\varepsilon = \Gamma_\varepsilon(\theta)$ of simple closed rectifiable curves in $\mathbb{D} \cup \mathbb{T}$ with the following properties.*

(a) *The interiors of the curves in Γ_ε are pairwise disjoint; the intersection of each of these curves with the circle $\mathbb{T}$ has zero length.*
(b) *One has $\eta < |\theta| < \varepsilon$ on $\Gamma_\varepsilon \cap \mathbb{D}$ for some positive $\eta = \eta(\varepsilon)$.*
(c) *The arclength $|dz|$ on $\Gamma_\varepsilon \cap \mathbb{D}$ is a Carleson measure, i.e., $H^1 \subset L^1(\Gamma_\varepsilon, |dz|)$; moreover, the norm of the corresponding embedding operator is bounded by a constant $N(\varepsilon)$ depending only on ε.*
(d) *For every $F \in H^1$, the equality*

$$\int_{\mathbb{T}} \frac{F}{\theta} dz = \int_{\Gamma_\varepsilon} \frac{F}{\theta} dz$$

holds true, provided that the curves in the family Γ_ε are oriented appropriately.

Now we are in a position to prove our main result in this section.

Proof of Theorem 2.1 The implications (i) $\Longrightarrow$ (ii) and (v) $\Longrightarrow$ (vi) being obvious, our plan is to show that (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv) $\Longrightarrow$ (i) $\Longrightarrow$ (v) and also that (vi) $\Longrightarrow$ (iii).

(ii) $\Longrightarrow$ (iii). Write $u := \theta^n$ and let $g \in K_u^\infty$. Note that

$$\bar{f}g = T_{\bar{f}}g + H_{\bar{f}}g. \tag{8}$$

Since $f \in A^\alpha$, Lemma 2.3 tells us that

$$\|H_{\bar{f}}g\|_q \le c_\alpha \|f\|_{\Lambda^\alpha} \|g\|_p.$$

Now, since $f\overline{u} \in \Lambda^\alpha$ by (ii), it follows that $P_-(f\overline{u}) \in \Lambda^\alpha$ (indeed, the operators P_+ and P_- are known to map Λ^α into itself), and Lemma 2.4 gives

$$\|T_{\bar{f}}g\|_q \le c_\alpha \|P_-(f\overline{u})\|_{\Lambda^\alpha} \|g\|_p.$$

The last two inequalities, together with (8), imply

$$\|\bar{f}g\|_q \le \text{const} \cdot \|g\|_p,$$

where the constant does not depend on g. Obviously,

$$\|T_f g\|_q = \|fg\|_q = \|\bar{f}g\|_q,$$

and since K_u^∞ is dense in K_u^p, we conclude that T_f is a bounded operator from K_u^p to H^q.

(iii) $\Longrightarrow$ (iv). Fix $z \in \mathbb{D}$ and consider the reproducing kernel k_z (for K^2_θ), given by

$$k_z(\zeta) = \frac{1 - \overline{\theta(z)}\theta(\zeta)}{1 - \bar{z}\zeta}.$$

Since $k_z \in K^\infty_\theta$, it follows that $k^n_z \in K^\infty_{\theta^n} (\subset K^p_{\theta^n})$; indeed,

$$k^n_z \overline{\theta^n} = \left(k_z \overline{\theta}\right)^n \in \overline{H^\infty_0}.$$

Therefore, by (iii),

$$\|fk^n_z\|_q \le \text{const} \cdot \|k^n_z\|_p. \tag{9}$$

In order to derive further information from this inequality, we now estimate its right-hand side from above, and the left-hand side from below. The elementary estimate

$$\int_{\mathbb{T}} \frac{dm(\zeta)}{|\zeta - z|^\gamma} \le \frac{C_\gamma}{(1 - |z|)^{\gamma - 1}} \qquad (\gamma > 1)$$

shows that

$$\begin{aligned} \|k^n_z\|_p &= \left(\int_{\mathbb{T}} \left| \frac{1 - \overline{\theta(z)}\theta(\zeta)}{1 - \bar{z}\zeta} \right|^{np} dm(\zeta) \right)^{1/p} \\ &\le 2^n \left(\int_{\mathbb{T}} \frac{dm(\zeta)}{|\zeta - z|^{np}} \right)^{1/p} \le \frac{\text{const}}{(1 - |z|)^{n - 1/p}}, \end{aligned} \tag{10}$$

since $np > 1$.

Now let F stand for the outer factor of f. Using the Cauchy integral formula, we get

$$\begin{aligned} \|fk^n_z\|_q &= \left(\int_{\mathbb{T}} |F(\zeta)|^q \left| \frac{1 - \overline{\theta(z)}\theta(\zeta)}{1 - \bar{z}\zeta} \right|^{nq} dm(\zeta) \right)^{1/q} \\ &\ge \left| \int_{\mathbb{T}} F^q(\zeta) \frac{(1 - \overline{\theta(z)}\theta(\zeta))^{nq}}{(1 - \bar{z}\zeta)^{nq-1}} \frac{dm(\zeta)}{1 - z\bar{\zeta}} \right|^{1/q} \\ &= \left(|F(z)|^q \frac{(1 - |\theta(z)|^2)^{nq}}{(1 - |z|^2)^{nq-1}} \right)^{1/q} = |F(z)| \frac{(1 - |\theta(z)|^2)^n}{(1 - |z|^2)^{n - 1/q}} \\ &\ge \text{const} \cdot |f(z)| \frac{(1 - |\theta(z)|)^n}{(1 - |z|)^{n - 1/q}}. \end{aligned} \tag{11}$$

In view of (10) and (11), inequality (9) now yields

$$|f(z)| \cdot (1 - |\theta(z)|)^n \le \text{const} \cdot (1 - |z|)^{1/p - 1/q} = \text{const} \cdot (1 - |z|)^{\alpha},$$

the constant being independent of z. Hence, for $0 < \varepsilon < 1$, we have

$$|f(z)| \le \text{const} \cdot (1 - \varepsilon)^{-n} (1 - |z|)^{\alpha}$$

whenever $z \in \Omega(\theta, \varepsilon)$, so that (iv) holds true.

(iv) $\Longrightarrow$ (i). We begin by showing that if (iv) is fulfilled with some $\varepsilon \in (0, 1)$, then $f\overline{\theta} \in \Lambda^{\alpha}$. Since

$$f\overline{\theta} = T_{\overline{\theta}} f + H_{\overline{\theta}} f$$

and $T_{\overline{\theta}} f \in A^{\alpha}$ (recall Lemma 2.2), it suffices to check that $H_{\overline{\theta}} f \in \Lambda^{\alpha}$. To this end, we take an arbitrary function $g \in H_0^{\infty}$ with $\|g\|_r = 1$, where $r = (1 + \alpha)^{-1}$, and verify that the integrals $\int_{\mathbb{T}} (H_{\overline{\theta}} f) g \, dm$ are bounded in modulus by a constant independent of g. This will mean that the function $z H_{\overline{\theta}} f$ generates a continuous linear functional on H^r, and hence lies in $\overline{A^{\alpha}}$. Writing $g_1 := g/z$ and using the Carleson curves $\Gamma_{\varepsilon} = \Gamma_{\varepsilon}(\theta)$ as described in Lemma 2.5, we obtain

$$\begin{aligned} \left| \int_{\mathbb{T}} (H_{\overline{\theta}} f) g \, dm \right| &= \left| \int_{\mathbb{T}} f \overline{\theta} g \, dm \right| = \left| \frac{1}{2\pi i} \int_{\mathbb{T}} \frac{f g_1}{\theta} \, dz \right| \\ &= \left| \frac{1}{2\pi i} \int_{\Gamma_{\varepsilon}} \frac{f g_1}{\theta} \, dz \right| \le \frac{1}{2\pi} \int_{\Gamma_{\varepsilon}} \frac{|f| |g_1|^{1-r} |g_1|^r}{|\theta|} |dz|. \end{aligned}$$

Because g_1 is a unit-norm function in H^r, it follows easily that $|g_1(z)|^r \le (1 - |z|)^{-1}$, whence

$$|g_1(z)|^{1-r} \le (1 - |z|)^{-(1-r)/r} = (1 - |z|)^{-\alpha}, \qquad z \in \mathbb{D}.$$

Plugging this into the preceding estimate and recalling that $|\theta| \ge \eta(\varepsilon)$ on $\Gamma_{\varepsilon} \cap \mathbb{D}$, we find that

$$\left| \int_{\mathbb{T}} (H_{\overline{\theta}} f) g \, dm \right| \le \frac{1}{2\pi \eta(\varepsilon)} \cdot \left(\sup_{z \in \Gamma_{\varepsilon} \cap \mathbb{D}} \frac{|f(z)|}{(1 - |z|)^{\alpha}} \right) \cdot \int_{\Gamma_{\varepsilon}} |g_1|^r \, |dz|. \tag{12}$$

Since $\Gamma_{\varepsilon} \cap \mathbb{D}$ is contained in $\Omega(\theta, \varepsilon)$, the supremum in (12) is finite by virtue of (iv). Also,

$$\int_{\Gamma_{\varepsilon}} |g_1|^r \, |dz| \le N(\varepsilon) \cdot \int_{\mathbb{T}} |g_1|^r \, dm = N(\varepsilon).$$

Taking this into account, we deduce from (12) that

$$\sup\left\{\left|\int_{\mathbb{T}}(H_{\overline{\theta}}f)g\,dm\right| : g\in H_0^\infty,\ \|g\|_r=1\right\}\le\frac{CN(\varepsilon)}{2\pi\eta(\varepsilon)},$$

where C is the constant coming from the O-condition in (iv). This means that $H_{\overline{\theta}}f\in\Lambda^\alpha$, and hence $f\overline{\theta}\in\Lambda^\alpha$.

Replacing θ by θ^k and ε by ε^k in the above argument, we similarly verify that $f\overline{\theta^k}\in\Lambda^\alpha$ for every $k\in\mathbb{N}$.

(i) $\Longrightarrow$ (v). Assuming (i), we prove first that $f\theta\in A^\alpha$, or equivalently, that

$$(f\theta)^{(n)}(z)=O((1-|z|)^{\alpha-n})\quad\text{as}\quad |z|\to 1^-. \tag{13}$$

For $z\in\mathbb{D}$ and almost all $\zeta\in\mathbb{T}$, we have the elementary identity

$$\theta^{n+1}(\zeta)=(\theta(\zeta)-\theta(z))^{n+1}+\sum_{k=0}^{n}\varphi_k(z)\theta^k(\zeta),$$

where

$$\varphi_k(z):=(-1)^{n-k}\binom{n+1}{k}\theta^{n+1-k}(z).$$

Therefore,

$$\begin{aligned}(f\theta)^{(n)}(z)&=\frac{n!}{2\pi i}\int_{\mathbb{T}}\frac{f(\zeta)\theta(\zeta)}{(\zeta-z)^{n+1}}d\zeta=\frac{n!}{2\pi i}\int_{\mathbb{T}}\frac{f(\zeta)\overline{\theta^n(\zeta)}\theta^{n+1}(\zeta)}{(\zeta-z)^{n+1}}d\zeta\\&=\frac{n!}{2\pi i}\int_{\mathbb{T}}(f\overline{\theta}^n)(\zeta)\left(\frac{\theta(\zeta)-\theta(z)}{\zeta-z}\right)^{n+1}d\zeta+\frac{n!}{2\pi i}\sum_{k=0}^{n}\varphi_k(z)\int_{\mathbb{T}}\frac{f(\zeta)\overline{\theta}^{n-k}(\zeta)}{(\zeta-z)^{n+1}}d\zeta\\&=\frac{n!}{2\pi i}\int_{\mathbb{T}}(f\overline{\theta}^n)(\zeta)\cdot\Phi_z(\zeta)d\zeta+\sum_{k=0}^{n}\varphi_k(z)\cdot\left(T_{\overline{\theta}^{n-k}}f\right)^{(n)}(z),\end{aligned}$$

where

$$\Phi_z(\zeta):=\left(\frac{\theta(\zeta)-\theta(z)}{\zeta-z}\right)^{n+1}.$$

In view of (i), $f\overline{\theta}^{n-k}\in\Lambda^\alpha$ for $k=0,\dots,n$, so that $T_{\overline{\theta}^{n-k}}f\in A^\alpha$, which implies that

$$\left(T_{\overline{\theta}^{n-k}}f\right)^{(n)}(z)=O((1-|z|)^{\alpha-n})\quad\text{as}\quad |z|\to 1^-.$$

The functions $\varphi_k(z)$ are bounded in $\mathbb{D}$, and to prove (13) it remains to verify that

$$\int_{\mathbb{T}} (f\overline{\theta}^n)(\zeta)\cdot\Phi_z(\zeta)\frac{d\zeta}{2\pi i} = O((1-|z|)^{\alpha-n}) \quad \text{as} \quad |z|\to 1^-. \tag{14}$$

Denote the integral on the left-hand side by $I_n(z)$. Since $\Phi_z \in H^\infty$, it follows that

$$|I_n(z)| = \left|\int_{\mathbb{T}} (f\overline{\theta}^n)(\zeta)\cdot\zeta\Phi_z(\zeta)\,dm(\zeta)\right| \le c_\alpha \|P_-(f\overline{\theta}^n)\|_{\Lambda^\alpha}\|\Phi_z\|_r;$$

here, as before, $r=(1+\alpha)^{-1}$. Because $n>\alpha$, we have $(n+1)r>1$ and

$$\|\Phi_z\|_r \le 2^{n+1}\left(\int_{\mathbb{T}}\frac{dm(\zeta)}{|\zeta-z|^{(n+1)r}}\right)^{1/r} \le \frac{\text{const}}{(1-|z|)^{n+1-1/r}} = \frac{\text{const}}{(1-|z|)^{n-\alpha}},$$

where the constant does not depend on z. Consequently,

$$|I_n(z)| \le \text{const}\cdot\|P_-(f\overline{\theta}^n)\|_{\Lambda^\alpha}(1-|z|)^{\alpha-n}.$$

Since

$$\|P_-(f\overline{\theta}^n)\|_{\Lambda^\alpha} \le C_\alpha\|f\overline{\theta}^n\|_{\Lambda^\alpha} < \infty$$

by virtue of (i), the estimate (14) is thereby established.

Thus, we have proved the implication

$$(f\in A^\alpha)\ \&\ \text{(i)} \implies f\theta\in A^\alpha.$$

Applying this inductively to $f\theta, f\theta^2$, etc., in place of f, we eventually deduce from (i) that $f\theta^k \in A^\alpha$ for each $k\in\mathbb{N}$.

(vi) $\implies$ (iii). Write $u := \theta^n$ and suppose that $g\in K_u^\infty$. Then $g\overline{u}\in\overline{H_0^\infty}$, and hence

$$\overline{f}\overline{u}g = P_-(\overline{f}\overline{u}g) = H_{\overline{f}\overline{u}}g.$$

Therefore,

$$\|fg\|_q = \|\overline{f}\overline{u}g\|_q = \|H_{\overline{f}\overline{u}}g\|_q \le c_\alpha\|fu\|_{\Lambda^\alpha}\|g\|_p, \tag{15}$$

where the last inequality is due to Lemma 2.3. The quantity $\|fu\|_{\Lambda^\alpha}$ is finite in view of (vi), and (15) tells us that

$$\|fg\|_q \le \text{const}\cdot\|g\|_p$$

with a constant independent of g. Thus, the multiplication operator $T_f : g\mapsto fg$ maps K_u^p boundedly into H^q, as required. □

If we wish to restrict ourselves to the issue of multiplying or dividing a function $f \in A^\alpha$ by an inner function θ (and its powers), leaving out the model subspace part, we may state the result in a more concise form as follows.

Proposition 2.6 *Suppose that* $0 < \alpha < \infty$, $n \in \mathbb{N}$, *and* $n > \alpha$. *Given* $f \in A^\alpha$ *and an inner function* θ, *the four statements below are equivalent.*

(i) $f\theta^n \in A^\alpha$.
(ii) $f\overline{\theta}^n \in \Lambda^\alpha$.
(iii) $f\theta^k \in \Lambda^\alpha$ *for all* $k \in \mathbb{Z}$.
(iv) *Condition* (7) *holds for some (or every)* $\varepsilon \in (0, 1)$.

To prove this, it suffices to choose exponents p and q (once α and n are given) so as to make the hypotheses of Theorem 2.1 true, and then invoke the theorem.

Remarks 1. An alternative route to Proposition 2.6 (but not to Theorem 2.1 in its entirety) via the pseudoanalytic extension method was found by Dyn'kin [15]. A similar technique was later used by the author in [11] to completely characterize the functions in A^α, $0 < \alpha < 1$, and in more general Lipschitz-type spaces, in terms of their moduli. (In particular, some equivalent forms of the crucial condition (7) came out as a corollary.) Subsequently, Pavlović [21] gave a more elementary proof of that result from [11].

2. Some of the conditions in Theorem 2.1 and Proposition 2.6 would become simpler if we could take $n = 1$. This can be done if $1 < p < \infty$ in Theorem 2.1, or if $0 < \alpha < 1$ in Proposition 2.6, but not in the general case. Indeed, it follows from Shirokov's work (see [28, 29]) that for each $\alpha > 1$, one can find $f \in A^\alpha$ and a Blaschke product θ such that $f/\theta \in A^\alpha$, but $f\theta \notin A^\alpha$. This means, in particular, that conditions (i) and (ii) in Proposition 2.6 are no longer equivalent when $\alpha > 1$ and $n = 1$. The equivalence does hold under certain additional assumptions, though; these are likewise discussed in [28, 29]. See also [9, 13] for an alternative study of this phenomenon.

3. Given $\alpha \in (0, \infty) \setminus \mathbb{Z}$, suppose that $f \in A^\alpha$ and θ is an inner function. Comparing our Proposition 2.6 with Shirokov's earlier results (see [27–29]), one infers that condition (7) holds if and only if

$$m(\sigma(\theta)) = 0 \qquad \& \qquad |f(\zeta)| = O\left(\frac{1}{|\theta'(\zeta)|^\alpha}\right) \text{ for } \zeta \in \mathbb{T} \setminus \sigma(\theta), \tag{16}$$

where $\sigma(\theta)$ is the set of boundary singularities for θ. The equivalence between (7) and (16) was also verified directly in [7, Section 2].

4. Theorem 2.1 and Proposition 2.6 remain valid in the case $\alpha = 0$ (with $n = 1$ and $1 < p = q < \infty$), provided that the spaces Λ^0 and A^0 are taken to be BMO and BMOA, respectively. This convention might be justified by the duality relations $A^\alpha = (H^{1/(1+\alpha)})^*$ and BMOA $= (H^1)^*$. The BMO versions of the above results are discussed in more detail in [7, Section 5].

5. In [13], we also considered the algebra $H^\infty_n := \{f : f^{(n)} \in H^\infty\}$, $n \in \mathbb{N}$, in place of A^α, and we came up with an analogue of Proposition 2.6 in that context.

Factorization in Dirichlet-Type Spaces

For a sequence $w = \{w_k\}_{k=1}^{\infty}$ of nonnegative numbers, the corresponding *Dirichlet-type space* $\mathcal{D}_w$ is formed by those functions $f \in H^2$ for which the quantity

$$\|f\|_w := \left(\sum_{k=1}^{\infty} w_k |\widehat{f}(k)|^2 \right)^{1/2} \tag{17}$$

is finite. The case $w_k = k$ corresponds to the classical *Dirichlet space* $\mathcal{D}(= \mathcal{D}_{\{k\}})$, the set of all functions $f \in H^2$ with

$$\|f\|_{\mathcal{D}} := \left(\int_{\mathbb{D}} |f'(z)|^2 dA(z) \right)^{1/2} < \infty$$

(here A is the normalized area measure on $\mathbb{D}$), and we have $\|\cdot\|_{\mathcal{D}} = \|\cdot\|_{\{k\}}$.

We begin by establishing a certain orthogonality relation involving Toeplitz operators on Dirichlet-type spaces.

Theorem 3.1 *Given numbers* $0 \le w_1 \le w_2 \le \dots$, *let* $w = \{w_k\}_{k=1}^{\infty}$ *and let* $\gamma = \{\gamma_k\}_{k=1}^{\infty}$ *be the sequence defined by*

$$\gamma_1 = w_1, \qquad \gamma_k = w_k - w_{k-1} \quad (k = 2, 3, \dots). \tag{18}$$

Suppose that $F \in H^2$, θ *is an inner function, and* $\{g_n\}$ *is an orthonormal basis in* K_θ. *If* $\Phi := zT_{\bar{z}\bar{\theta}}F$ *and* $h_n := zT_{\bar{\theta}}(Fg_n)$, *then*

$$\|F\|_w^2 = \|\Phi\|_w^2 + \sum_n \|h_n\|_\gamma^2 \tag{19}$$

(the definition of $\|\cdot\|_\gamma$ *being similar to* (17) *above).*

To keep on the safe side, we remark that sequences with unspecified index sets, which we occasionally employ, are allowed to be finite (and sometimes empty). In particular, the orthonormal basis $\{g_n\}$ in Theorem 3.1 will be finite if and only if θ is a finite Blaschke product.

The proof will make use of the notion of a *Hilbert–Schmidt operator*. Recall that, given two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, a linear operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is said to be Hilbert–Schmidt if the quantity

$$\|T\|_{\mathfrak{S}_2} := \left(\sum_n \|Te_n\|_{\mathcal{H}_2}^2 \right)^{1/2}$$

is finite for some (or each) orthonormal basis $\{e_n\}$ of $\mathcal{H}_1$. It is well known – and easily shown – that this quantity does not actually depend on the choice of $\{e_n\}$ and is therefore well defined. The set of all Hilbert–Schmidt operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ is denoted by $\mathfrak{S}_2(\mathcal{H}_1, \mathcal{H}_2)$.

Also, we need a lemma that relates Hilbert–Schmidt operators to Dirichlet-type spaces. We state and prove it now, before proceeding with the proof of Theorem 3.1.

Lemma 3.2 *Let $F \in H^2$. Suppose that $w = \{w_k\}_{k=1}^\infty$ and $\gamma = \{\gamma_k\}_{k=1}^\infty$ are two sequences of nonnegative numbers related by*

$$w_n = \sum_{k=1}^n \gamma_k \quad (n = 1, 2, \dots). \tag{20}$$

Finally, consider the multiplier map M_γ acting by the rule

$$M_\gamma\left(\sum_{k=1}^\infty a_k \overline{z}^k\right) := \sum_{k=1}^\infty \sqrt{\gamma_k} a_k \overline{z}^k, \qquad z \in \mathbb{T} \tag{21}$$

(defined initially on the set of antianalytic trigonometric polynomials $\sum_k a_k \overline{z}^k$). Then the operator $M_\gamma H_{\overline{F}}$ belongs (or has an extension belonging) to $\mathfrak{S}_2(H^2, \overline{H_0^2})$ if and only if $F \in \mathcal{D}_w$. Moreover,

$$\|M_\gamma H_{\overline{F}}\|_{\mathfrak{S}_2} = \|F\|_w. \tag{22}$$

Proof Since $\{z^n\}_{n=0}^\infty$ is an orthonormal basis in H^2, we have

$$\|M_\gamma H_{\overline{F}}\|_{\mathfrak{S}_2}^2 = \sum_{n=0}^\infty \|M_\gamma H_{\overline{F}} z^n\|_2^2, \tag{23}$$

where $\|\cdot\|_2$ is the usual L^2-norm. Letting $a_n := \widehat{F}(n)$, we find that

$$H_{\overline{F}} z^n = \sum_{k=1}^\infty \overline{a}_{n+k} \overline{z}^k,$$

whence

$$M_\gamma H_{\overline{F}} z^n = \sum_{k=1}^\infty \sqrt{\gamma_k} \overline{a}_{n+k} \overline{z}^k,$$

and, by the Parseval identity,

$$\|M_\gamma H_{\overline{F}} z^n\|_2^2 = \sum_{k=1}^\infty \gamma_k |a_{n+k}|^2.$$

Plugging this into (23) and recalling (20), we obtain

$$\|M_\gamma H_{\overline{F}}\|^2_{\mathfrak{S}_2} = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty}\gamma_k|a_{n+k}|^2 = \sum_{j=1}^{\infty}|a_j|^2\sum_{k=1}^{j}\gamma_k$$
$$= \sum_{j=1}^{\infty} w_j|a_j|^2 = \|F\|^2_w,$$

which proves (22) and the lemma. □

Proof of Theorem 3.1 Let M_γ be the multiplier map defined by (21). From Lemma 3.2 we know that

$$\|F\|^2_w = \|M_\gamma H_{\overline{F}}\|^2_{\mathfrak{S}_2}. \tag{24}$$

Consider the functions G_n defined (a.e. on $\mathbb{T}$) by $G_n := \bar{z}\bar{g}_n\theta$. Since $\{g_n\}$ is an orthonormal basis in K_θ, the same is true for $\{G_n\}$ (indeed, the map $f \mapsto \bar{z}\bar{f}\theta$ is an antilinear isometry of K_θ onto itself). Furthermore, since $H^2 = \theta H^2 \oplus K_\theta$, the family $\{\theta z^n\}_{n=0}^{\infty} \cup \{G_n\}$ forms an orthonormal basis in H^2, and we may use it to compute the Hilbert–Schmidt norm in (24). In this way we obtain

$$\|M_\gamma H_{\overline{F}}\|^2_{\mathfrak{S}_2} = \sum_{n=0}^{\infty}\|M_\gamma H_{\overline{F}}(\theta z^n)\|^2_2 + \sum_n \|M_\gamma H_{\overline{F}} G_n\|^2_2 = S_1 + S_2, \tag{25}$$

where S_1 and S_2 denote the two preceding sums, in the same order. The elementary identity

$$P_-\varphi = \bar{z}\overline{P_+(\bar{z}\bar{\varphi})}, \qquad \varphi \in L^2, \tag{26}$$

yields

$$P_-(\overline{F}\theta) = \bar{z}\overline{P_+(\bar{z}F\bar{\theta})} = \overline{\Phi},$$

whence

$$H_{\overline{F}}(\theta z^n) = P_-(\overline{F}\theta z^n) = P_-(P_-(\overline{F}\theta)\cdot z^n)$$
$$= P_-(\overline{\Phi} z^n) = H_{\overline{\Phi}} z^n.$$

Thus,

$$S_1 = \sum_{n=0}^{\infty}\|M_\gamma H_{\overline{\Phi}} z^n\|^2_2 = \|M_\gamma H_{\overline{\Phi}}\|^2_{\mathfrak{S}_2} = \|\Phi\|^2_w, \tag{27}$$

where the last equality relies on Lemma 3.2.

Another application of (26) gives

$$H_{\overline{F}}G_n = P_-(\overline{F}\bar{z}\bar{g}_n\theta) = \bar{z}\overline{P_+(Fg_n\bar{\theta})} = \bar{h}_n,$$

and so

$$\|M_\gamma H_{\overline{F}}G_n\|_2^2 = \sum_{k=1}^{\infty}\gamma_k\left|\widehat{(H_{\overline{F}}G_n)}(-k)\right|^2 = \sum_{k=1}^{\infty}\gamma_k|\widehat{h}_n(k)|^2 = \|h_n\|_\gamma^2.$$

Summing over n, we get

$$S_2 = \sum_n \|M_\gamma H_{\overline{F}}G_n\|_2^2 = \sum_n \|h_n\|_\gamma^2. \tag{28}$$

Finally, we plug the identities coming from (27) and (28) into (25). Together with (24), this yields the required formula (19). □

As a consequence of Theorem 3.1, we now deduce a result of Korenblum and Faĭvyshevskiĭ concerning the action of certain Toeplitz operators on Dirichlet-type spaces. (In all fairness, their original theorem in [19] gives a bit more than our Corollary 3.3 below. Alternative routes to that result can be found in [18] and [22].) To state it, we need a minor modification of the $\|\cdot\|_w$ norm. Namely, given a sequence $v = \{v_n\}_{n=0}^\infty$ of positive numbers and a holomorphic function $f(z) = \sum_{n=0}^\infty \widehat{f}(n)z^n$ on $\mathbb{D}$, we put

$$\|f\|_{v,0} := \left(\sum_{n=0}^\infty v_n|\widehat{f}(n)|^2\right)^{1/2}$$

(note that the value $n = 0$ is now included).

Corollary 3.3 *Let $v = \{v_n\}_{n=0}^\infty$ be a nondecreasing sequence of positive numbers, and let θ be an inner function. Then, for every $f, g \in H^2$, we have*

$$\|T_{\bar{\theta}}f\|_{v,0} \le \|f\|_{v,0} \tag{29}$$

and

$$\|g\|_{v,0} \le \|g\theta\|_{v,0}. \tag{30}$$

Proof Put $F := zf$ and define Φ as in Theorem 3.1, so that

$$\Phi = zT_{\bar{z}\bar{\theta}}F = zT_{\bar{\theta}}f.$$

For $n = 1, 2, \ldots$, let $w_n = v_{n-1}$ and $w = \{w_n\}_{n=1}^\infty$. Theorem 3.1 implies that $\|\Phi\|_w \le \|F\|_w$. Observing that $\|\Phi\|_w = \|T_{\bar{\theta}}f\|_{v,0}$ and $\|F\|_w = \|f\|_{v,0}$, we arrive at (29). To prove (30), it suffices to apply (29) with $f = g\theta$. □

The next fact is likewise a straightforward consequence of Theorem 3.1.

Theorem 3.4 *Let $w = \{w_k\}_{k=1}^{\infty}$ be a nondecreasing sequence with $w_1 \ge 0$, and let $\gamma = \{\gamma_k\}_{k=1}^{\infty}$ be defined by (18). If $f \in H^2$, θ is an inner function, and $\{g_n\}$ is an orthonormal basis in K_θ, then*

$$\|f\theta\|_w^2 = \|f\|_w^2 + \sum_n \|zfg_n\|_\gamma^2. \tag{31}$$

Proof Put $F := f\theta$, and define Φ and h_n as in Theorem 3.1. We have then

$$\Phi = zT_{\bar{z}\bar{\theta}}(f\theta) = zT_{\bar{z}}f = f - f(0),$$

whence $\|\Phi\|_w = \|f\|_w$. Also,

$$h_n = zT_{\bar{\theta}}(f\theta g_n) = zfg_n.$$

The formula (19) therefore reduces to (31), and the proof is complete. □

In some special cases, Theorem 3.4 can be used to derive a more explicit form of the (nonnegative) "discrepancy term"

$$R_w(f, \theta) := \|f\theta\|_w^2 - \|f\|_w^2. \tag{32}$$

One such case is pointed out in Theorem 3.5 below. Before stating the result, we need to recall some basic facts about angular derivatives.

Given a function $\varphi \in H^\infty$ with $\|\varphi\|_\infty = 1$, we say that φ has an *angular derivative* (in the sense of Carathéodory) at a point $\zeta \in \mathbb{T}$ if both φ and φ' have nontangential limits at ζ, the former of these being of modulus 1. (The two limits are then denoted by $\varphi(\zeta)$ and $\varphi'(\zeta)$, respectively.) The classical Julia–Carathéodory theorem (see [2, Chapter VI], [3, Chapter I] or [23, Chapter VI]) asserts that this happens if and only if

$$\liminf_{z \to \zeta} \frac{1 - |\varphi(z)|^2}{1 - |z|^2} < \infty. \tag{33}$$

And if (33) holds, the theorem tells us also that $\varphi'(\zeta)$ coincides with the limit of the difference quotient

$$\frac{\varphi(z) - \varphi(\zeta)}{z - \zeta}$$

as $z \to \zeta$ nontangentially. Moreover, $|\varphi'(\zeta)|$ will then agree with the value of the (unrestricted) lim inf in (33), and this remains true if lim inf is replaced by the corresponding nontangential limit.

Finally, if $\theta = BS$ is an inner function (with B a Blaschke product and S singular), then

$$|\theta'(\zeta)| = \sum_j \frac{1-|a_j|^2}{|\zeta - a_j|^2} + 2\int_{\mathbb{T}} \frac{d\mu(\eta)}{|\zeta-\eta|^2}, \qquad \zeta \in \mathbb{T}, \tag{34}$$

where $\{a_j\}$ is the zero sequence of B and μ is the singular measure associated with S. This formula can be found in [1]; it holds for *every* point ζ of $\mathbb{T}$, with the convention that $|\theta'(\zeta)| = \infty$ whenever θ fails to possess an angular derivative at ζ.

Theorem 3.5 *Let σ be a positive Borel measure on $[0,1]$ with $\int_{[0,1]} x^2 d\sigma(x) < \infty$. Put*

$$\gamma_k := \int_{[0,1]} x^{2k} d\sigma(x), \qquad k = 1, 2, \dots,$$

and define the sequence $w = \{w_n\}_{n=1}^{\infty}$ by (20). If $f \in H^2$ and θ is an inner function, then

$$\|f\theta\|_w^2 = \|f\|_w^2 + \int_{\mathbb{T}} dm(\zeta) \int_{[0,1]} r^2 |f(r\zeta)|^2 \frac{1-|\theta(r\zeta)|^2}{1-r^2} d\sigma(r). \tag{35}$$

Here the value of $(1-|\theta(r\zeta)|^2)/(1-r^2)$ at $r=1$ is interpreted as $|\theta'(\zeta)|$, the modulus of the angular derivative of θ at ζ.

The proof will rely on Theorem 3.4 and on the following lemma.

Lemma 3.6 *Let θ be an inner function, and let $\{g_n\}$ be an orthonormal basis in K_θ. Then*

$$\sum_n |g_n(z)|^2 = \frac{1-|\theta(z)|^2}{1-|z|^2}, \qquad z \in \mathbb{D}. \tag{36}$$

Furthermore, if $\zeta \in \mathbb{T}$ is a point at which the limits $\lim_{r\to 1^-} g_n(r\zeta) =: g_n(\zeta)$ exist for all n, then

$$\sum_n |g_n(\zeta)|^2 = |\theta'(\zeta)|. \tag{37}$$

To prove the lemma, consider the reproducing kernel

$$k_z(w) = \frac{1-\overline{\theta(z)}\theta(w)}{1-\bar{z}w}$$

of K_θ and use Parseval's identity to get

$$\sum_n |g_n(z)|^2 = \sum_n |\langle g_n, k_z\rangle|^2 = \|k_z\|_2^2 = k_z(z) = \frac{1-|\theta(z)|^2}{1-|z|^2}$$

for $z \in \mathbb{D}$. This yields (36), which in turn implies (37) upon putting $z = r\zeta$ and passing to the limit as $r \to 1^-$.

Proof of Theorem 3.5 We may assume that $f \in \mathcal{D}_w$, since otherwise both sides of (35) equal ∞. By Theorem 3.4, the "discrepancy term" (32) is given by

$$R_w(f,\theta) = \sum_n \|zfg_n\|_\gamma^2, \tag{38}$$

where $\gamma = \{\gamma_k\}_{k=1}^\infty$ and $\{g_n\}$ is some (no matter which) orthonormal basis in K_θ. This said, we proceed by considering two special cases.

Case 1: σ has no atom at 1. We may think of the disk

$$\mathbb{D} = \{r\zeta : r \in [0,1),\ \zeta \in \mathbb{T}\}$$

as of a measure space endowed with the product measure $\sigma \times m =: \nu$. The monomials z^k ($k = 1, 2, \dots$) are then mutually orthogonal in $L^2(\mathbb{D})$ and have norms $\sqrt{\gamma_k}$. Therefore, for a function $h(z) = \sum_{k=1}^\infty \widehat{h}(k) z^k$ in zH^1, we have

$$\|h\|_{L^2(\mathbb{D},\nu)}^2 = \sum_{k=1}^\infty \gamma_k |\widehat{h}(k)|^2 = \|h\|_\gamma^2.$$

Applying this to $h_n := zfg_n$ gives

$$\|h_n\|_\gamma^2 = \|h_n\|_{L^2(\mathbb{D},\nu)}^2 = \int_{\mathbb{T}} dm(\zeta) \int_{[0,1]} r^2 |f(r\zeta)|^2 |g_n(r\zeta)|^2 d\sigma(r).$$

Consequently, in view of (38),

$$R_w(f,\theta) = \sum_n \|h_n\|_\gamma^2 = \int_{\mathbb{T}} dm(\zeta) \int_{[0,1]} r^2 |f(r\zeta)|^2 \sum_n |g_n(r\zeta)|^2 d\sigma(r). \tag{39}$$

By Lemma 3.6,

$$\sum_n |g_n(r\zeta)|^2 = \frac{1 - |\theta(r\zeta)|^2}{1 - r^2},$$

and so (39) reduces to

$$R_w(f,\theta) = \int_{\mathbb{T}} dm(\zeta) \int_{[0,1]} r^2 |f(r\zeta)|^2 \frac{1 - |\theta(r\zeta)|^2}{1 - r^2} d\sigma(r),$$

which proves (35).

Case 2: σ is the unit point mass at 1. In this case, we have $\gamma_k = 1$ and $w_k = k$, so that $\|\cdot\|_\gamma = \|\cdot\|_2$ on zH^2, and $\|\cdot\|_w = \|\cdot\|_{\mathcal{D}}$. Therefore, we can rewrite (38) in the form

$$\|f\theta\|_{\mathcal{D}}^2 - \|f\|_{\mathcal{D}}^2 = \sum_n \|zfg_n\|_2^2 = \int_{\mathbb{T}} |f(\zeta)|^2 \sum_n |g_n(\zeta)|^2 dm(\zeta).$$

Combining this with (37), we finally obtain

$$\|f\theta\|_{\mathcal{D}}^2 - \|f\|_{\mathcal{D}}^2 = \int_{\mathbb{T}} |f(\zeta)|^2 |\theta'(\zeta)| dm(\zeta), \tag{40}$$

which coincides with (35) under the current hypothesis on σ.

The general case being a combination of Cases 1 and 2, the required result follows. □

Remark Recalling the identity (34) and plugging it into (40), we find that

$$\|f\theta\|_{\mathcal{D}}^2 = \|f\|_{\mathcal{D}}^2 + \int_{\mathbb{T}} |f(\zeta)|^2 \left(\sum_j \frac{1-|a_j|^2}{|\zeta - a_j|^2} + 2\int_{\mathbb{T}} \frac{d\mu(\eta)}{|\zeta-\eta|^2} \right) dm(\zeta) \tag{41}$$

(here, as before, $\{a_j\}$ is the zero sequence of θ, and μ is the associated singular measure). This was established by Carleson in [4]. In fact, the formula given there is a combination of (41) and an explicit expression for the Dirichlet integral $\|f\|_{\mathcal{D}}^2$ of an outer function f.

Model Subspaces in BMOA

It has been noticed that various smoothness properties of an inner function θ, if available, tend to be inherited (typically, in a weaker form) by functions in K_θ^p. This phenomenon becomes especially pronounced when passing from θ to

$$K_{*\theta} := K_\theta^2 \cap \mathrm{BMOA},$$

the star-invariant subspace of BMOA, in which case no loss of smoothness usually occurs. (Of course, the smoothness property in question should not be too strong – it should not even imply continuity – if we want a nontrivial inner function to have it.) A result to that effect will appear as Corollary 4.4 below; we shall deduce it from a more general theorem concerning the action of a coanalytic Toeplitz operator $T_{\bar g}$, with $g \in H^1$, on $K_{*\theta}$. However, the very meaning of the expression $T_{\bar g} f$ (with $f \in K_{*\theta}$) is not immediately clear, since the product $f\bar g$ need not be integrable. The following proposition will clarify the situation.

Proposition 4.1 *Given* $f \in K_{*\theta}$ *and* $g \in H^1$, *there exists a function* $\Phi \in \cap_{0<p<1} H^p$ *such that*

$$\|T_{\overline{g}_n} f - \Phi\|_p \to 0$$

for every $p \in (0,1)$ *and every sequence* $\{g_n\} \subset H^2$ *with* $\|g_n - g\|_1 \to 0$.

This (obviously unique) function Φ is then taken to be $T_{\overline{g}} f$, the image of f under the Toeplitz operator $T_{\overline{g}}$.

The proof relies on the following lemma due to Cohn (see Lemma 3.2 in [5, p. 731]), which in turn results from an application of the (H^1, BMOA) duality.

Lemma 4.2 *Let* θ *be inner, and let* $f \in K_{*\theta}$. *Then* $f = P_+(\overline{z}\overline{\psi}\theta)$ *for a function* $\psi \in H^\infty$. *Furthermore,* ψ *may be chosen so that* $\|f\|_* = \|\psi\|_\infty$.

Here and below, $\|\cdot\|_*$ is the dual space norm on BMOA induced by H^1.

Proof of Proposition 4.1 Let $f \in K_{*\theta}$, $g \in H^1$, and suppose $\{g_n\}$ is a sequence of H^2-functions with $\|g_n - g\|_1 \to 0$. We have then

$$T_{\overline{g}_n} f = P_+\left(\overline{g}_n P_+(\overline{z}\overline{\psi}\theta)\right) = P_+\left(\overline{g}_n \overline{z}\overline{\psi}\theta\right),$$

where ψ is related to f as in Lemma 4.2. Now put

$$\Phi := P_+\left(\overline{g}\overline{z}\overline{\psi}\theta\right).$$

This definition makes sense, since P_+ is applied to an L^1-function; besides, it does not depend on the choice of ψ. (Indeed, if ψ_1 and ψ_2 are both eligible in the sense of Lemma 4.2, then $\psi_1 - \psi_2 \in \theta H^\infty$.) And since P_+ is a continuous mapping from L^1 to every H^p with $0 < p < 1$ (cf. [16, p. 128]), we conclude that $\Phi \in H^p$ and $\|T_{\overline{g}_n} f - \Phi\|_p \to 0$ for any such p. □

Now suppose X is a Banach space of analytic functions on the disk, with $X \subset H^1$. We say that X is a *K-space* if, for each $\psi \in H^\infty$, the Toeplitz operator $T_{\overline{\psi}}$ acts boundedly from X to itself, with norm at most $\mathrm{const} \cdot \|\psi\|_\infty$. (This is essentially equivalent to saying that X enjoys the so-called *K-property* of Havin. The latter was defined in [17] by the formally weaker condition that $T_{\overline{\psi}}(X) \subset X$, for all $\psi \in H^\infty$, but the norm estimate is usually automatic.)

Following [17], we remark that X will be a K-space provided it is (isomorphic to) the dual of some Banach space Y, consisting of analytic functions on $\mathbb{D}$ and satisfying the conditions

(a) $H^\infty \cap Y$ is dense in Y, and
(b) for each $\psi \in H^\infty$, the multiplication operator $f \mapsto f\psi$ acts boundedly from Y to itself, with norm at most $\mathrm{const} \cdot \|\psi\|_\infty$.

(It is understood that the pairing between X and Y is given by $\langle f, g\rangle := \int_{\mathbb{T}} f\overline{g}\,dm$, which is meaningful at least for $f \in H^\infty \cap Y$ and $g \in X$.) The Toeplitz operator $T_{\overline{\psi}} : X \to X$ is then the adjoint of the multiplication map in (b), which justifies our claim.

As examples of K-spaces, we list the following:

- H^p with $1 < p < \infty$,
- the Hardy–Sobolev spaces $H^{p,n} := \{f \in H^p : f^{(n)} \in H^p\}$ with $1 \le p < \infty$ and $n \ge 1$,
- BMOA, and more generally, $\mathrm{BMOA}^{(n)} := \{f \in H^1 : f^{(n)} \in \mathrm{BMOA}\}$ with $n \ge 0$,
- the Dirichlet-type spaces $\mathcal{D}_w := \{f \in H^2 : \sum_{n \ge 1} w_n |\widehat{f}(n)|^2 < \infty\}$ associated with nondecreasing sequences $w = \{w_n\}$ of positive numbers,
- the analytic Besov spaces $B^s_{p,q}$ with $s > 0$, $p \ge 1$, $q \ge 1$, and in particular
- the classical Lipschitz–Zygmund spaces $A^\alpha := B^\alpha_{\infty,\infty}$ with $0 < \alpha < \infty$.

We recall that $B^s_{p,q}$ is defined as the set of those analytic f on $\mathbb{D}$ for which the function

$$r \mapsto (1-r)^{n-s} \left\| f_r^{(n)} \right\|_p \tag{42}$$

is in L^q over the interval $(0, 1)$ with respect to the measure $dr/(1-r)$; here n is some (any) fixed integer with $n > s$ and $f_r^{(n)}(\zeta) := f^{(n)}(r\zeta)$.

For most of the spaces considered, the K-property has been established by means of a duality argument, as outlined above. We refer to [17], where this is done for A^α and some special cases of Hardy–Sobolev and Besov spaces; to [25, 26] for general $H^{p,n}$ and $B^s_{p,q}$ classes, as well as for $\mathrm{BMOA}^{(n)}$; and finally to any of [18, 19, 22] in connection with $\mathcal{D}_w$ spaces.

As further examples of K-spaces, we mention K^p_θ $(1 < p < \infty)$ and $K_{*\theta}$. Indeed, for $g \in H^\infty$, one verifies the inclusion $T_{\bar{g}}(K^p_\theta) \subset K^p_\theta$ by noting that K^p_θ is the kernel of the Toeplitz operator $T_{\bar{\theta}} : H^p \to H^p$, which commutes with $T_{\bar{g}}$. Then one deduces that $T_{\bar{g}}(K_{*\theta}) \subset K_{*\theta}$, recalling that $K_{*\theta} = K^2_\theta \cap \mathrm{BMOA}$ and BMOA is a K-space. And, of course, the two inclusions are accompanied by the natural norm estimates: the norm of $T_{\bar{g}}$ is in both cases $O(\|g\|_\infty)$, just as it happens for the containing spaces H^p $(1 < p < \infty)$ and BMOA.

The main result of this section is as follows.

Theorem 4.3 *Let θ be an inner function, $g \in H^1$, and let X be a K-space. The following are equivalent.*

(i) *$T_{\bar{g}}$ acts boundedly from $K_{*\theta}$ to X.*
(ii) *$T_{\bar{g}}$ acts boundedly from K^∞_θ to X.*
(iii) *The function*

$$k(z) := \frac{\theta(z) - \theta(0)}{z}$$

satisfies $T_{\bar{g}} k \in X$.

Moreover, the operator norms $\|T_{\bar{g}}\|_{K_{*\theta} \to X}$ and $\|T_{\bar{g}}\|_{K^\infty_\theta \to X}$ are comparable to each other and to $\|T_{\bar{g}} k\|_X$.

In most – perhaps all – cases of interest, condition (iii) above can be further rephrased by saying that $T_{\bar g}\theta \in X$. In fact, since $k = T_{\bar z}\theta$ and $T_{\bar z}T_{\bar g} = T_{\bar g}T_{\bar z}$, the implication

$$T_{\bar g}\theta \in X \implies T_{\bar g}k \in X$$

holds whenever X is a K-space. The converse is true provided that $1 \in X$ and $zX \subset X$; indeed,

$$T_{\bar g}\theta = \text{const} + zT_{\bar g}k.$$

In particular, we certainly have $T_{\bar g}k \in X \iff T_{\bar g}\theta \in X$ when X is one of our smoothness classes, such as $H^{p,n}$, $B^s_{p,q}$, A^α or $\text{BMOA}^{(n)}$, let alone H^p and BMOA. The theorem then states that the inclusion $T_{\bar g}f \in X$ holds for all $f \in K_{*\theta}$ if and only if it holds for $f = \theta$.

The next fact is obtained by applying Theorem 4.3 with $g \equiv 1$, in which case $T_{\bar g}$ reduces to the identity map.

Corollary 4.4 *Given an inner function θ and a K-space X, one has*

$$K_{*\theta} \subset X \iff K^\infty_\theta \subset X \iff k \in X. \tag{43}$$

And since the latter condition, $k \in X$, is implied by (and is usually equivalent to) saying that $\theta \in X$, the nontrivial part of (43) amounts to the implication

$$\theta \in X \implies K_{*\theta} \subset X. \tag{44}$$

Proof of Theorem 4.3 The part (i) $\implies$ (ii) is trivially true, as is the inequality

$$\|T_{\bar g}\|_{K^\infty_\theta \to X} \le \|T_{\bar g}\|_{K_{*\theta}\to X}.$$

The part (ii) $\implies$ (iii), along with the estimate

$$\|T_{\bar g}\|_{K^\infty_\theta \to X} \ge \frac{1}{2}\|T_{\bar g}k\|_X,$$

is also obvious, since $k \in K^\infty_\theta$ and $\|k\|_\infty \le 2$.

What remains to be proved is the implication (iii) $\implies$ (i) and its quantitative version

$$\|T_{\bar g}\|_{K_{*\theta}\to X} \le \text{const} \cdot \|T_{\bar g}k\|_X. \tag{45}$$

To this end, we fix $f \in K_{*\theta}$ and then invoke Lemma 4.2 to find a function $\psi \in H^\infty$ such that

$$f = T_{\bar z\bar\psi}\theta, \qquad \|f\|_* = \|\psi\|_\infty.$$

Using the fact that coanalytic Toeplitz operators commute (and moreover, $T_{\bar a}T_{\bar b} = T_{\bar a \bar b}$ whenever a, b and ab are H^1-functions such that the operators involved are all well-defined), we obtain

$$T_{\bar g}f = T_{\bar g}T_{\bar z\bar\psi}\theta = T_{\bar\psi}T_{\bar g}T_{\bar z}\theta = T_{\bar\psi}T_{\bar g}k. \tag{46}$$

Finally, we recall that X is a K-space to get

$$\begin{aligned}\|T_{\bar g}f\|_X &\le \|T_{\bar\psi}\|_{X\to X}\|T_{\bar g}k\|_X\\ &\le \text{const}\cdot\|\psi\|_\infty\|T_{\bar g}k\|_X\\ &= \text{const}\cdot\|f\|_*\|T_{\bar g}k\|_X,\end{aligned}$$

which readily implies (45). □

Finally, we supplement Theorem 4.3 with the following result.

Proposition 4.5 *Let θ, g and k be as above. The operator $T_{\bar g}$ acts boundedly from $K_{*\theta}$ to itself if and only if $T_{\bar g}k \in$ BMOA. In this case we also have*

$$\|T_{\bar g}f\|_p \le C_p\|T_{\bar g}k\|_*\|f\|_p, \qquad 1 < p < \infty,$$

for all $f \in K_\theta^\infty$, so that $T_{\bar g}$ extends to a bounded operator on K_θ^p.

This might be compared to the "$T(1)$-" and/or "$T(b)$-theorem" of David, Journé and Semmes (cf. [14, Chapter 5] or [30, Chapter VII]), results that provide boundedness criteria for certain singular integral operators on L^p. Just as in those theorems, we only have to test the operator on a single function. We also remark that the assumption $T_{\bar g}k \in$ BMOA can be rewritten as $T_{\bar g}\theta \in$ BMOA, and a sufficient condition for this to happen is that

$$\sup\{|g(z)| : z \in \Omega(\theta,\varepsilon)\} < \infty$$

for some $\varepsilon \in (0,1)$, where $\Omega(\theta,\varepsilon)$ is the sublevel set defined by (6). A proof of this last assertion can be found in [8].

Proof of Proposition 4.5 The first statement, concerning the action of $T_{\bar g}$ on $K_{*\theta}$, is obtained by applying Theorem 4.3 with $X =$ BMOA (or $X = K_{*\theta}$).

Now suppose $T_{\bar g}k \in$ BMOA, and let $1 < p < \infty$. Given a function $f \in K_\theta^\infty$, put $\psi := \bar z\bar f\theta (= \widetilde{f})$ and note that $\psi \in H^\infty$. We have then

$$f = \bar z\bar\psi\theta = P_+\left(\bar z\bar\psi\theta\right) = T_{\bar z\bar\psi}\theta,$$

and so (46) remains in force. Setting $h := T_{\bar g}k$ and making use of the elementary identity

$$\overline{P_+F} = zP_-(\bar z\bar F), \qquad F \in L^1,$$

we can rewrite the resulting equality from (46) as

$$T_{\bar{g}}f = T_{\bar{\psi}}h = \bar{z}\overline{H_{\bar{z}\bar{h}}\psi}.$$

In view of Nehari's theorem (see, e.g., [20, Part B, Chapter 1]), the assumption that h, and hence zh, is in BMOA implies that the Hankel operator $H_{\bar{z}\bar{h}}$ acts boundedly from H^p to $\overline{H_0^p}$, with norm not exceeding $C_p\|h\|_*$. Consequently,

$$\|T_{\bar{g}}f\|_p = \left\|H_{\bar{z}\bar{h}}\psi\right\|_p \le C_p\|h\|_*\|\psi\|_p = C_p\|h\|_*\|f\|_p, \qquad f \in K_\theta^\infty.$$

Finally, since K_θ^∞ is dense in K_θ^p (indeed, K_θ^∞ contains the family of reproducing kernels for K_θ^2), we conclude that $T_{\bar{g}}$ extends to a bounded operator on K_θ^p, with the same norm. The proof is complete. □

References

1. P.R. Ahern, D.N. Clark, On inner functions with H^p-derivative. Michigan Math. J. **21**, 115–127 (1974)
2. R.B. Burckel, *An Introduction to Classical Complex Analysis*, vol. I (Academic Press, New York, 1979)
3. C. Carathéodory, *Theory of Functions of a Complex Variable*, vol. II (Chelsea Publishing Company, New York, 1954)
4. L. Carleson, A representation formula for the Dirichlet integral. Math. Z. **73**, 190–196 (1960)
5. W.S. Cohn, Radial limits and star invariant subspaces of bounded mean oscillation. Am. J. Math. **108**, 719–749 (1986)
6. P.L. Duren, B.W. Romberg, A.L. Shields, Linear functionals on H^p spaces with $0 < p < 1$. J. Reine Angew. Math. **238**, 32–60 (1969)
7. K.M. Dyakonov, Smooth functions and coinvariant subspaces of the shift operator. Algebra i Analiz **4**(5), 117–147 (1992); translation in St. Petersburg Math. J. **4**, 933–959 (1993)
8. K.M. Dyakonov, Division and multiplication by inner functions and embedding theorems for star-invariant subspaces. Am. J. Math. **115**, 881–902 (1993)
9. K.M. Dyakonov, Multiplication by Blaschke products and stability of ideals in Lipschitz algebras. Math. Scand. **73**, 246–258 (1993)
10. K.M. Dyakonov, Factorization of smooth analytic functions via Hilbert–Schmidt operators. Algebra i Analiz **8**(4), 1–42 (1996); translation in St. Petersburg Math. J. **8**, 543–569 (1997)
11. K.M. Dyakonov, Equivalent norms on Lipschitz-type spaces of holomorphic functions. Acta Math. **178**, 143–167 (1997)
12. K.M. Dyakonov, Two theorems on star-invariant subspaces of BMOA. Indiana Univ. Math. J. **56**, 643–658 (2007)
13. K.M. Dyakonov, Blaschke products and nonideal ideals in higher order Lipschitz algebras. Algebra i Analiz **21**(6), 182–201 (2009); translation in St. Petersburg Math. J. **21**, 979–993 (2010)
14. E.M. Dyn'kin, Methods of the theory of singular integrals: Littlewood-Paley theory and its applications, in *Commutative Harmonic Analysis, IV*, ed. by V.P. Khavin, N.K. Nikol'skii. Encyclopaedia of Mathematical Sciences, vol. 42, pp. 97–194 (Springer, Berlin, 1992)
15. E.M. Dyn'kin, The pseudoanalytic extension. J. Anal. Math. **60**, 45–70 (1993)
16. J.B. Garnett, *Bounded Analytic Functions*, Revised 1st edn. (Springer, New York, 2007)

17. V.P. Havin, The factorization of analytic functions that are smooth up to the boundary. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **22**, 202–205 (1971)
18. V.E. Katsnel'son, Remark on canonical factorization in certain analytic function spaces. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **30**, 163–164 (1972); English transl. in J. Soviet Math. **4**, 444–445 (1975)
19. B.I. Korenblum, V.M. Faĭvyshevskiĭ, A certain class of compression operators that are connected with the divisibility of analytic functions. Ukrain. Mat. Zh. **24**, 692–695 (1972); English transl. in Ukrainian Math. J. **24**, 559–561 (1973)
20. N.K. Nikolski, *Operators, Functions, and Systems: An Easy Reading, Volume 1: Hardy, Hankel, and Toeplitz*. Mathematical Surveys and Monographs, vol. 92 (American Mathematical Society, Providence, 2002)
21. M. Pavlović, On Dyakonov's paper "Equivalent norms on Lipschitz-type spaces of holomorphic functions". Acta Math. **183**, 141–143 (1999)
22. M. Rabindranathan, Toeplitz operators and division by inner functions. Indiana Univ. Math. J. **22**, 523–529 (1972/1973)
23. D. Sarason, *Sub-Hardy Hilbert Spaces in the Unit Disk* (Wiley, New York, 1994)
24. F.A. Shamoyan, Division by an inner function in certain spaces of functions that are analytic in the disc. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **22**, 206–208 (1971)
25. F.A. Shamoyan, A class of Toeplitz operators that are connected with the divisibility of analytic functions. Funktsional. Anal. i Prilozhen. **13**(1), 83 (1979); English transl. in Funct. Anal. Appl. **13**, 70–71 (1979)
26. F.A. Shamoyan, Toeplitz operators and division by an inner function in some spaces of analytic functions. Akad. Nauk Armyan. SSR Dokl. **76**(3), 109–113 (1983, Russian)
27. N.A. Shirokov, Ideals and factorization in algebras of analytic functions that are smooth up to the boundary. Trudy Mat. Inst. Steklov. **130**, 196–222 (1978)
28. N.A. Shirokov, Division and multiplication by inner functions in spaces of analytic functions smooth up to the boundary, in *Complex Analysis and Spectral Theory (Leningrad, 1979/1980)*. Lecture Notes in Mathematics, Vol. 864, pp. 413–439 (Springer, Berlin/New York, 1981)
29. N.A. Shirokov, *Analytic Functions Smooth Up to the Boundary*. Lecture Notes in Mathematics, vol. 1312 (Springer, Berlin, 1988)
30. E.M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals* (Princeton University Press, Princeton, 1993)

Rational Inner Functions on a Square-Matrix Polyball

Anatolii Grinshpan, Dmitry S. Kaliuzhnyi-Verbovetskyi, Victor Vinnikov, and Hugo J. Woerdeman

Dedicated to the blessed memory of Cora Sadosky, our dear friend and colleague.

Abstract We establish the existence of a finite-dimensional unitary realization for every matrix-valued rational inner function from the Schur–Agler class on a unit square-matrix polyball. In the scalar-valued case, we characterize the denominators of these functions. We also show that a multiple of every polynomial with no zeros in the closed domain is such a denominator. One of our tools is the Korányi–Vagi theorem generalizing Rudin's description of rational inner functions to the case of bounded symmetric domains; we provide a short elementary proof of this theorem suitable in our setting.

2010 *Mathematics Subject Classification.* 26C10, 32A10, 47A48, 47N70, 93B15

Introduction

In this paper, we study rational inner functions on the Cartesian product of square-matrix Cartan domains of type I, i.e., on a unit square-matrix polyball,

AG, DK-V, HW were partially supported by NSF grant DMS-0901628. DK-V and VV were partially supported by BSF grant 2010432.

A. Grinshpan (✉) • D.S. Kaliuzhnyi-Verbovetskyi • H.J. Woerdeman
Department of Mathematics, Drexel University, 3141 Chestnut St., Philadelphia, PA 19104, USA
e-mail: tolya@math.drexel.edu; dmitryk@math.drexel.edu; hugo@math.drexel.edu

V. Vinnikov
Department of Mathematics, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel
e-mail: vinnikov@math.bgu.ac.il

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_10

$$\mathcal{B} = \mathbb{B}^{\ell_1\times\ell_1} \times \cdots \times \mathbb{B}^{\ell_k\times\ell_k}$$
$$= \Big\{Z = (Z^{(1)}, \ldots, Z^{(k)}) \in \mathbb{C}^{\ell_1\times\ell_1} \times \cdots \times \mathbb{C}^{\ell_k\times\ell_k} \colon \|Z^{(r)}\| < 1,\ r = 1, \ldots, k\Big\}.$$

We can interpret points of $\mathcal{B}$ as block-diagonal matrices $Z = \bigoplus_{r=1}^{k} Z^{(r)}$ with $\|Z\| < 1$. Then $\mathcal{B}$ is a special case of a domain $\mathcal{D}_{\mathbf{P}}$ defined by means of a matrix polynomial $\mathbf{P}$ as the set of $z = (z_1, \ldots, z_d) \in \mathbb{C}^d$ satisfying $\|\mathbf{P}(z)\| < 1$ (see [2, 4, 11]); in this case, $\mathbf{P} = Z$ viewed as a polynomial in matrix entries $z_{ij}^{(r)}$, $i, j = 1, \ldots, \ell_r$, $r = 1, \ldots, k$, and $d = \sum_{r=1}^{k} \ell_r^2$. In particular, $\mathcal{B}$ is a special case of a Cartesian product of (not necessarily square) matrix Cartan domains of type I (see [10, 14]). The distinguished (or Shilov) boundary of $\mathcal{B}$ consists of k-tuples of unitary matrices,

$$\partial_S\mathcal{B} = \Big\{Z = (Z^{(1)}, \ldots, Z^{(k)}) \in \mathbb{C}^{\ell_1\times\ell_1} \times \cdots \times \mathbb{C}^{\ell_k\times\ell_k} \colon Z^{(r)*}Z^{(r)} = I,\ r = 1, \ldots, k\Big\},$$

which can also be interpreted as a set of block-diagonal unitary matrices. Notice that the unit polydisk $\mathbb{D}^d$ is a special case of a unit square-matrix polyball where $k = d$, and $\ell_r = 1$ for $r = 1, \ldots, k$.

We consider matrix functions in variables $z_{ij}^{(r)}$. We denote the corresponding d-tuple of variables by z and, for a function F, we identify $F(z) = F(Z)$. An $s \times s$ matrix-valued function F is *rational inner* if each matrix entry $F_{\alpha\beta}$ is a rational function in d variables $z_{ij}^{(r)}$ which is regular in $\mathcal{B}$, and F takes unitary matrix values at each of its regular points on the distinguished boundary $\partial_S\mathcal{B}$. Notice that the zero variety of the least common multiple of the denominators of the rational functions $F_{\alpha\beta}$ in their coprime fraction representation has an intersection with $\partial_S\mathcal{B}$ of relative Lebesgue measure zero, which can be proved using an argument analogous to that of [6, Lemma 6.3]; thus almost all points of $\partial_S\mathcal{B}$ are regular points of F.

Define

$$\mathcal{T}_Z = \Big\{T = \bigoplus_{r=1}^{k} T^{(r)} \colon \quad T^{(r)} = [T_{ij}^{(r)}]_{i,j=1}^{\ell_r},\ r = 1, \ldots, k,$$

$$(T_{ij}^{(r)}) \text{ is a } d\text{-tuple of commuting operators on a Hilbert space and } \|T\| < 1\Big\}.$$

For $T \in \mathcal{T}_Z$, the Taylor joint spectrum [19] of T viewed as a multioperator

$$(T_{ij}^{(r)})_{r=1,\ldots,k;\, i,j=1,\ldots,\ell_r}$$

lies in the domain $\mathcal{B}$, and for a matrix-valued function F analytic on $\mathcal{B}$ one can define $F(T)$ by means of Taylor's functional calculus [20]; see [2] and a further discussion in [4]. We say that an $s\times s$ matrix-valued function F analytic on $\mathcal{B}$ belongs to the Schur–Agler class $\mathcal{SA}_Z(\mathbb{C}^s)$ associated with $\mathcal{B}$, or rather with its defining polynomial $\mathbf{P} = Z$, if its associated Agler norm,

$$\|F\|_{\mathcal{A},Z} = \sup_{T\in\mathcal{T}_Z} \|F(T)\|$$

is at most 1. In the scalar case, $s = 1$, we simply write $\mathcal{SA}_Z$. In the case of the unit polydisk $\mathbb{D}^d$, this class coincides with the classical Schur–Agler class $\mathcal{SA}_d$ studied in the seminal paper of Agler [1]. Notice that $\mathcal{SA}_Z(\mathbb{C}^s)$ is a subclass of the Schur class $\mathcal{S}_Z(\mathbb{C}^s)$ of $s \times s$ matrix-valued contractive analytic functions on $\mathcal{B}$. It follows from [21] and [3] that these classes do not coincide, i.e., the analog of von Neumann's inequality fails when $d \geq 2$, unless $\mathcal{B} = \mathbb{D}^2$. Moreover, rational inner functions, which obviously belong to the Schur class, do not necessarily belong to the Schur–Agler class: an example of a rational inner function on $\mathbb{D}^3$ which is not Schur–Agler was given in [12, Example 5.1].

In section "Scalar-Valued Rational Inner Functions", we give a characterization of scalar-valued rational inner functions on $\mathcal{B}$ in terms of their coprime fraction representation. In section "Matrix-Valued Rational Inner Functions from the Schur–Agler Class", we describe matrix-valued rational inner functions on $\mathcal{B}$ that belong to the associated Schur–Agler class as the functions that have a finite-dimensional unitary realization. In section "Eventual Agler Denominators", we characterize eventual Agler denominators, i.e., those multivariable polynomials which can be represented as the denominators of scalar rational inner functions in the Schur–Agler class $\mathcal{SA}_Z$, in terms of certain contractive determinantal representations. We also show that a multiple of every polynomial with no zeros on the closed domain, $\overline{\mathcal{B}}$, is an eventual Agler denominator, and we end with several open questions.

Scalar-Valued Rational Inner Functions

For a polynomial p in d variables $z_{ij}^{(r)}$, $i, j = 1, \ldots, \ell_r$, $r = 1, \ldots, k$, where $d = \sum_{r=1}^{k} \ell_r^2$, we define its reverse with respect to $\mathcal{B}$ as

$$\overleftarrow{p}(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{t_r} \overline{p(Z^{*-1})},$$

where t_r is the total degree of p in the variables $z_{ij}^{(r)}$, $i, j = 1, \ldots, \ell_r$. We say that a polynomial p is $\mathcal{B}$-stable (resp., strongly $\mathcal{B}$-stable) if p does not have zeros in $\mathcal{B}$ (resp., in $\overline{\mathcal{B}}$).

The following result is a generalization of Rudin's characterization of rational inner functions on the polydisk [18] to the case of a square-matrix polyball $\mathcal{B}$. It appears in more generality in [17, Theorem 3.3], where Rudin's theorem is extended to all bounded symmetric domains. We provide a proof that applies to the specific setting of $\mathcal{B}$, and therefore requires less machinery.

Theorem 2.1 *A scalar-valued function f on $\mathcal{B}$ is rational inner if and only if there exist a $\mathcal{B}$-stable polynomial p and nonnegative integers $m_1, \dots, m_k$ such that*

$$f(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{m_r} \frac{\overleftarrow{p}(Z)}{p(Z)}. \tag{1}$$

One can choose p to be coprime with $\overleftarrow{p}$.

For the proof of Theorem 2.1, we will need the following proposition.

Proposition 2.2 *Let p be a $\mathcal{B}$-stable polynomial, and suppose that $|p(Z)| = 1$ for all $Z \in \partial_S \mathcal{B}$. Then there exist nonnegative integers $m_1, \dots, m_k$ such that*

$$p(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{m_r}.$$

Proof Notice that if $Z^{(r)}$ is unitary for each r, then

$$\overleftarrow{p}(Z)p(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{t_r}. \tag{2}$$

Since $\partial_S \mathcal{B}$ is a uniqueness set for analytic functions (see, e.g., [14]), we have that (2) holds for all $Z = (Z^{(1)}, \dots, Z^{(k)}) \in \mathbb{C}^{\ell_1 \times \ell_1} \times \dots \times \mathbb{C}^{\ell_k \times \ell_k}$. Since $\det Z^{(r)}$ is an irreducible polynomial in matrix entries $z_{ij}^{(r)}$ (see, e.g., [7, Section 61]) we obtain that $p(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{m_r}$ for some $m_r \le t_r$, $r = 1, \dots, k$. □

Proof of Theorem 2.1 The sufficiency of the representation (1) for f to be rational inner is clear. To prove the necessity, we first write $f = q/p$, with p and q coprime. Since f is analytic in $\mathcal{B}$, we have that p is $\mathcal{B}$-stable. Next, for $Z \in \partial_S \mathcal{B}$ we have that $q(Z)\overline{q(Z^{-1*})} = p(Z)\overline{p(Z^{-1*})}$. Hence the equality

$$\prod_{r=1}^{k} (\det Z^{(r)})^{\tau_r} q(Z)\overline{q(Z^{-1*})} = \prod_{r=1}^{k} (\det Z^{(r)})^{\tau_r} p(Z)\overline{p(Z^{-1*})}$$

holds for every $Z \in \partial_S \mathcal{B}$, where τ_r is the maximum of the total degrees in $z_{ij}^{(r)}$, $i, j = 1, \dots, \ell_r$, in q and p. Then it also holds for all Z. Since q and p are coprime, p is a divisor of $\prod_{r=1}^{k} (\det Z^{(r)})^{\tau_r} \overline{q(Z^{-1*})}$. Hence

$$u(Z)q(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{\tau_r} \overline{p(Z^{-1*})},$$

for some polynomial u. Observe that $|u(Z)| = 1$ on $\partial_S \mathcal{B}$. Then by Proposition 2.2 we obtain that $u(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{\mu_r}$, with some $\mu_r \leq \tau_r$, and

$$q(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{\tau_r - \mu_r} \overline{p(Z^{-1*})}.$$

The latter equality implies that $m_r := \tau_r - \mu_r - t_r \geq 0$, where t_r is the total degree of p in variables $z_{ij}^{(r)}$, $i, j = 1, \dots, \ell_r$. It follows that

$$q(Z) = \prod_{r=1}^{k} (\det Z^{(r)})^{m_r} \overleftarrow{p}(Z),$$

and (1) holds. □

Remark 2.3 Proposition 2.2 and Theorem 2.1 also hold when $\mathcal{B}$ is replaced by a Cartesian product of square-matrix Cartan domains of type II [14],

$$\mathcal{B}_{\text{sym}} = \left\{ Z = (Z^{(1)}, \dots, Z^{(k)}) \in \mathcal{B} \colon Z^{(r)} = Z^{(r)\top}, r = 1, \dots, k \right\},$$

or by a Cartesian product of mixed type involving Cartan domains of types I and II. The proofs are similar after noticing that $\det Z^{(r)}$ as a polynomial in $z_{ij}^{(r)}$, $i, j = 1, \dots, \ell_r \colon i \leq j$, is also irreducible when $Z^{(r)} = Z^{(r)\top}$; see, e.g., [7, Section 61].

Matrix-Valued Rational Inner Functions from the Schur–Agler Class

We now characterize matrix-valued rational inner functions on the unit square-matrix polyball $\mathcal{B}$, which belong to the associated Schur–Agler class, in terms of their unitary realizations. This is a generalization of the result [5, Theorem 2.1] for the unit polydisk $\mathbb{D}^d$ which, in turn, is a matrix-valued extension of the result from [16] in the scalar-valued setting (see also an earlier paper [9] for the bidisk case).

Theorem 3.1 *An $s \times s$ matrix-valued function F on $\mathcal{B}$ is rational inner and belongs to the class $\mathcal{SA}_Z(\mathbb{C}^s)$ if and only if F has a finite-dimensional unitary realization, i.e., there exist nonnegative integers $n_1, \dots, n_k$ and a unitary matrix*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{C}^{(\sum_{r=1}^{k} \ell_r n_r + s) \times (\sum_{r=1}^{k} \ell_r n_r + s)}$$

such that

$$F(Z) = D + CZ_n(I - AZ_n)^{-1}B,$$

where $Z_n = \bigoplus_{r=1}^{k}(Z^{(r)} \otimes I_{n_r})$. *If we write* $F = QP^{-1}$, *where* P *and* Q *are matrix polynomials of total degree at most* g *such that* $P^*P = Q^*Q$ *on* $\partial_S\mathcal{B}$, *then* $n_1, \dots, n_k$ *can be chosen so that* $n_r \le l_r s\binom{g+d-1}{d}$, $r = 1, \dots, k$.

For the proof of Theorem 3.1, we will need the following proposition. Recall that a linear mapping $\Phi\colon \mathbb{C}^{a\times a} \to \mathbb{C}^{b\times b}$ is said to be completely positive if, for every $m \in \mathbb{N}$, the mapping $\Phi^{(m)}\colon (\mathbb{C}^{a\times a})^{m\times m} \to (\mathbb{C}^{b\times b})^{m\times m}$ defined by $(\Phi^{(m)}(A))_{ij} = \Phi(A_{ij})$, $i,j = 1, \dots, m$, is positive, that is, it maps every positive semidefinite matrix A to a positive semidefinite matrix $\Phi^{(m)}(A)$.

Proposition 3.2 ([8, Theorem 1]) *Let* $\Phi\colon \mathbb{C}^{a\times a} \to \mathbb{C}^{b\times b}$ *be a completely positive linear mapping. Then there exists* $Y \in \mathbb{C}^{a^2b\times b}$ *so that* $\Phi(X) = Y^*(X \otimes I_{ab})Y$.

Proof of Theorem 3.1 The sufficiency part is analogous to that of [6, Theorem 6.1]. To prove the necessity, let $F = QP^{-1}$, where P and Q are matrix polynomials of total degree at most g, $P^*P = Q^*Q$ on $\partial_S\mathcal{B}$, and assume that $F \in \mathcal{SA}_Z(\mathbb{C}^s)$. Then by [4, Theorem 1.5] there exist separable Hilbert spaces $\mathcal{K}_r$ and analytic functions H_r on $\mathcal{B}$ with values linear operators from $\mathbb{C}^s$ to $\mathbb{C}^{\ell_r} \otimes \mathcal{K}_r$ such that

$$P(W)^*P(Z) - Q(W)^*Q(Z) = \sum_{r=1}^{k} H_r(W)^*\Big((I - W^{(r)*}Z^{(r)}) \otimes I_{\mathcal{K}_r}\Big)H_r(Z), \quad Z, W \in \mathcal{B}.$$

Letting $Z = W = tU$ where $|t| < 1$ and $U \in \partial_S\mathcal{B}$, we obtain

$$\frac{P(tU)^*P(tU) - Q(tU)^*Q(tU)}{1 - |t|^2} = \sum_{r=1}^{k} H_r^*(tU)H_r(tU). \tag{3}$$

Since $P(tU)^*P(tU) = Q(tU)^*Q(tU)$ for all $U = (U^{(1)}, \dots, U^{(r)}) \in \partial_S\mathcal{B}$ and $|t| = 1$, the numerator of the left-hand side of (3) is a polynomial in t and $\bar{t}$ which vanishes on the variety $1 - t\bar{t} = 0$. Therefore the left-hand side of (3) is a polynomial in t and $\bar{t}$ and a trigonometric polynomial in matrix entries $u_{ij}^{(r)}$, $i,j = 1, \dots, \ell_r$, $r = 1, \dots, k$. Let P_α and Q_α be the coefficients of z^α in the polynomials P and Q, respectively, and let $H_{r,\alpha}$ be the coefficient of z^α in the Maclaurin series for H_r, where for $z = (z_1, \dots, z_d)$ and $\alpha = (\alpha_1, \dots, \alpha_d)$ we set $z^\alpha = z_1^{\alpha_1} \cdots z_d^{\alpha_d}$. Then the zeroth Fourier coefficient of the left-hand side of (3) as a trigonometric polynomial in variables $u_{ij}^{(r)}$ is

$$\frac{1}{1 - |t|^2} \sum_{|\alpha| \le g} (P_\alpha^* P_\alpha - Q_\alpha^* Q_\alpha)|t|^{2|\alpha|},$$

where $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Note that the preceding expression is a polynomial in $|t|^2$ of degree at most $g - 1$. The zeroth Fourier coefficient of the right-hand side of (3) (for sufficiently small t),

$$\sum_{r=1}^{k} \sum_{\alpha} H_{r,\alpha}^* H_{r,\alpha} |t|^{2|\alpha|},$$

is therefore a polynomial in $|t|^2$ as well, and $H_{r,\alpha}^* H_{r,\alpha} = 0$ for $|\alpha| \ge g$.

Consider now the completely positive map $\Phi_r \colon \mathbb{C}^{\ell_r \times \ell_r} \to \mathbb{C}^{s\binom{g+d-1}{d} \times s\binom{g+d-1}{d}}$ defined via

$$\Phi(X) = \operatorname{col}_{|\alpha| \leq g-1}(H_{r,\alpha}^*)(X \otimes I_{\mathcal{K}_r})\operatorname{row}_{|\alpha| \leq g-1}(H_{r,\alpha}).$$

Then by Proposition 3.2 we can find matrices $Y_r \in \mathbb{C}^{a^2 b \times b}$ such that $\Phi_r(X) = Y_r^*(X \otimes I_{ab})Y_r$, where $a = \ell_r$ and $b = s\binom{g+d-1}{d}$. Writing $Y_r = \operatorname{row}_{|\alpha| \leq g-1}(Y_{r,\alpha})$, we can form a polynomial

$$G_r(Z) = \sum_{|\alpha| \leq g-1} Y_{r,\alpha} z^\alpha$$

with the coefficients in $\mathbb{C}^{\ell_r n_r \times s}$, where $n_r = \ell_r s\binom{g+d-1}{d}$, so that

$$H_r(W)^*\Big((I - W^{(r)*}Z^{(r)}) \otimes I_{\mathcal{K}_r}\Big)H_r(Z) = G_r(W)^*\Big((I - W^{(r)*}Z^{(r)}) \otimes I_{n_r}\Big)G_r(Z), \quad r = 1, \dots, k,$$

and

$$P(W)^*P(Z) - Q(W)^*Q(Z) = \sum_{r=1}^{k} G_r(W)^*\Big((I - W^{(r)*}Z^{(r)}) \otimes I_{n_r}\Big)G_r(Z). \tag{4}$$

Rearranging the terms in (4), we obtain

$$P(W)^*P(Z) + \sum_{r=1}^{k} G_r(W)^*\Big(W^{(r)*}Z^{(r)} \otimes I_{n_r}\Big)G_r(Z) = Q(W)^*Q(Z) + \sum_{r=1}^{k} G_r(W)^*G_r(Z).$$

Therefore

$$\begin{bmatrix} (Z^{(1)} \otimes I_{n_1})G_1(Z) \\ \vdots \\ (Z^{(k)} \otimes I_{n_k})G_k(Z) \\ P(Z) \end{bmatrix} h \mapsto \begin{bmatrix} G_1(Z) \\ \vdots \\ G_k(Z) \\ Q(Z) \end{bmatrix} h$$

is a linear and isometric map from the span of the elements on the left to the span of the elements on the right. It may be extended (if necessary) to a unitary matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ so that

$$AZ_nG(Z) + BP(Z) = G(Z)$$
$$CZ_nG(Z) + DP(Z) = Q(Z),$$

for every Z; here $Z_n = \bigoplus_{r=1}^{k}(Z^{(r)} \otimes I_{n_r})$. Solving the first equation above for $G(Z)$ and then plugging the result into the second equation yield

$$F(Z) = Q(Z)P^{-1}(Z) = D + CZ_n(I - AZ_n)^{-1}B.$$

□

Remark 3.3 An analog of Theorem 3.1, with a similar proof, is also valid for Cartesian products of Cartan domains of type II or III [14] or for Cartesian products of mixed type involving Cartan domains of types I, II, and III. We recall here that a Cartan domain of type II (resp., III) is a (lower-dimensional) subset of a square-matrix Cartan domain of type I consisting of symmetric (resp., antisymmetric) matrices.

Eventual Agler Denominators

We will say that a polynomial v in $z_{ij}^{(r)}$, $i, j = 1, \ldots, \ell_r$, $r = 1, \ldots, k$, is almost self-reversive with respect to the square-matrix polyball $\mathcal{B}$ if $\overleftarrow{v} = \gamma v$, for some scalar γ with $|\gamma| = 1$.

We have the following generalization of a result that was announced in [13] for the case of a unit polydisk.

Theorem 4.1 *Let p be a $\mathcal{B}$-stable polynomial which is coprime with $\overleftarrow{p}$. Then the following are equivalent:*

(i) *p is an eventual Agler denominator, that is, there exist nonnegative integers $s_1, \ldots, s_k$ such that the rational inner function $\prod_{r=1}^{k}(\det Z^{(r)})^{s_r}\overleftarrow{p}(Z)/p(Z)$ is in the Schur–Agler class $\mathcal{SA}_Z$.*
(ii) *There exists an almost self-reversive polynomial v of multidegree $(s_1, \ldots, s_k)$ such that $p(Z)v(Z) = \det(I - KZ_n)$ for some nonnegative integers n_r, $r = 1, \ldots, k$, and a contractive matrix K, where $Z_n = \bigoplus_{r=1}^{k}(Z^{(r)} \otimes I_{n_r})$.*

Proof (i)⇒(ii) Let $f(Z) = \prod_{r=1}^{k}(\det Z^{(r)})^{s_r}\overleftarrow{p}(Z)/p(Z)$ be in $\mathcal{SA}_Z$. By Theorem 3.1 there exists a k-tuple $n = (n_1, \ldots, n_k)$ of nonnegative integers and a unitary matrix $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ such that

$$f(Z) = D + CZ_n(I - AZ_n)^{-1}B.$$

Using the factorization

$$\begin{bmatrix} I - AZ_n & B \\ -CZ_n & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ -CZ_n(I - AZ_n)^{-1} & I \end{bmatrix} \begin{bmatrix} I - AZ_n & 0 \\ 0 & f(Z) \end{bmatrix} \begin{bmatrix} I & (I - AZ_n)^{-1}B \\ 0 & I \end{bmatrix},$$

we observe that

$$f(Z) = \frac{\det \begin{bmatrix} I - AZ_n & B \\ -CZ_n & D \end{bmatrix}}{\det(I - AZ_n)}$$

and

$$\det \begin{bmatrix} I - AZ_n & B \\ -CZ_n & D \end{bmatrix} = \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} \det \left(\begin{bmatrix} A^* & 0 \\ B^* & 0 \end{bmatrix} - \begin{bmatrix} Z_n & 0 \\ 0 & -1 \end{bmatrix} \right) = \lambda \det(A^* - Z_n),$$

where $\lambda = \det \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix} \right]$. Hence

$$\prod_{r=1}^{k} (\det Z^{(r)})^{s_r} \overleftarrow{p}(Z) \det(I - AZ_n) = \lambda p(Z) \det(A^* - Z_n). \tag{5}$$

Since p is $\mathcal{B}$-stable and coprime with $\overleftarrow{p}$, the polynomials p and $\prod_{r=1}^{k} (\det Z^{(r)})^{s_r} \overleftarrow{p}$ do not have common factors. Therefore p divides $\det(I - AZ_n)$, i.e., there exists a polynomial v so that

$$p(Z)v(Z) = \det(I - AZ_n). \tag{6}$$

Taking the polynomial reverse on both sides, we obtain

$$\det Z_n \overline{p(Z^{*-1})v(Z^{*-1})} = \det(Z_n - A^*).$$

Using (5) and (6), we obtain

$$\prod_{r=1}^{k} (\det Z^{(r)})^{s_r} \overleftarrow{p}(Z) p(Z) v(Z) = \lambda \prod_{r=1}^{k} (\det Z^{(r)})^{n_r - \deg_r p} p(Z) \overleftarrow{p}(Z) \overline{v(Z^{*-1})} (-1)^{\sum_{r=1}^{k} \ell_r n_r},$$

where $\deg_r p$ is the total degree of p in the variables $z_{ij}^{(r)}$. After dividing out we see that v is almost self-reversive and $s_r = n_r - \deg_r p - \deg_r v$. Clearly, $K = A$ is a contractive matrix.

(ii)⇒(i) Suppose there exists an almost self-reversive polynomial v such that $p(Z)v(Z) = \det(I - KZ_n)$ with a contractive matrix K and a k-tuple $n = (n_1, \dots, n_k)$ of nonnegative integers. Then by a straightforward modification of [12, Theorem 5.2] the rational inner function

$$\frac{\prod_{r=1}^{k} (\det Z^{(r)})^{n_r} \overline{p(Z^{*-1})v(Z^{*-1})}}{p(Z)v(Z)}$$

is Schur–Agler. Since v is almost self-reversive, the rational inner function

$$\frac{\prod_{r=1}^{k}(\det Z^{(r)})^{n_r-\deg_r v}\overline{p(Z^{*-1})}}{p(Z)}$$

is Schur–Agler, i.e., p is an eventual Agler denominator. □

Remark 4.2 An analog of Theorem 4.1 is valid for a Cartesian product of Cartan domains of type II, i.e., for a domain $\mathcal{B}_{\rm sym}$, or for a Cartesian product of mixed type involving Cartan domains of types I and II; see Remarks 2.3 and 3.3.

The following result is a consequence of [11, Theorem 4.1] formulated for the case of the square-matrix polyball $\mathcal{B}$.

Theorem 4.3 *Some strongly $\mathcal{B}$-stable multiple of every strongly $\mathcal{B}$-stable polynomial is an eventual Agler denominator.*

Proof By [11, Theorem 4.1] there exist a strongly $\mathcal{B}$-stable polynomial q, a k-tuple of integers $n = (n_1, \ldots, n_k)$, and a strictly contractive matrix $K \in \mathbb{C}^{(\sum_{r=1}^{k}\ell_r n_r)\times(\sum_{r=1}^{k}\ell_r n_r)}$ such that $p(Z)q(Z) = \det(I - KZ_n)$, where $Z_n = \bigoplus_{r=1}^{k}(Z^{(r)} \otimes I_{n_r})$. Then, similarly to the last paragraph in the proof of Theorem 4.1 (with $v = 1$), one shows that pq is an eventual Agler denominator.

Corollary 4.4 *Let f be a rational inner function on $\mathcal{B}$ which is regular on $\partial_S\mathcal{B}$. Then there exists a rational inner function g which is regular on $\partial_S\mathcal{B}$ such that $fg \in \mathcal{SA}_Z$.*

Proof By Theorem 2.1 there exists a stable polynomial p which is coprime with $\overleftarrow{p}$ and such that (1) holds with some nonnegative integers $m_1, \ldots, m_r$. Since f is regular on $\partial_S\mathcal{B}$, the polynomial p is strongly $\mathcal{B}$-stable. By Theorem 4.3, pq is an eventual Agler denominator for some strongly $\mathcal{B}$-stable polynomial q. Therefore $\prod_{r=1}^{k}(\det Z^{(r)})^{s_r} f \overleftarrow{q}/q \in \mathcal{SA}_Z$ for some nonnegative integers $_1, \ldots, s_k$. □

The question as to whether the assumption of regularity of f on $\partial_S\mathcal{B}$ in Corollary 4.4 can be removed is open. Another open question is whether Corollary 4.4 holds for *matrix-valued* rational inner functions. Finally, it is interesting to investigate the analogues of the results in this paper for the unbounded version of the domain $\mathcal{B}$, i.e., the Cartesian product of matrix half-planes. The Cayley transform over the matrix variables $Z^{(r)}$, $r = 1, \ldots, k$, would allow one to obtain a finite-dimensional realization formula for rational inner functions on the product of matrix half-planes; if, in addition, the Cayley transform over the values of a function is applied, then one can obtain the corresponding realization formula for rational Cayley inner functions over the product of matrix half-planes (see [5] for the case of a poly-half-plane, i.e., the product of scalar half-planes). We would also like to mention [15] where a subclass of Cayley inner functions on the product of matrix half-planes, the Bessmertnyi class, was studied.

References

1. J. Agler, On the representation of certain holomorphic functions defined on a polydisc, in *Topics in Operator Theory: Ernst D. Hellinger Memorial Volume*. Operator Theory, Advances and Applications, vol. 48 (Birkhäuser, Basel, 1990), pp. 47–66
2. C.-G. Ambrozie, D.A. Timotin, Von Neumann type inequality for certain domains in $\mathbb{C}^n$. Proc. Am. Math. Soc. **131**(3), 859–869 (electronic) (2003)
3. W. Arveson, Subalgebras of C*-algebras. III. Multivariable operator theory. Acta Math. **181**(2), 159–228 (1998)
4. J.A. Ball, V. Bolotnikov. Realization and interpolation for Schur–Agler-class functions on domains with matrix polynomial defining function in $\mathbb{C}^n$. J. Funct. Anal. **213**(1), 45–87 (2004)
5. J.A. Ball, D.S. Kaliuzhnyi-Verbovetskyi, Rational Cayley inner Herglotz–Agler functions: positive-kernel decompositions and transfer-function realizations. Linear Algebra Appl. **456**, 138–156 (2014)
6. J.A. Ball, C. Sadosky, V. Vinnikov, Scattering systems with several evolutions and multidimensional input/state/output systems. Integral Equ. Oper. Theory **52**, 323–393 (2005)
7. M. Bocher, *Introduction to Higher Algebra* (Dover Publications, Inc., New York, 1964)
8. M.D. Choi, Completely positive linear maps on complex matrices. Linear Algebra Appl. **10**, 285–290 (1975)
9. B.J. Cole, J. Wermer. Ando's theorem and sums of squares. Indiana Univ. Math. J. **48**(3), 767–791 (1999)
10. A. Grinshpan, D.S. Kaliuzhnyi-Verbovetskyi, V. Vinnikov, H.J. Woerdeman, Contractive determinantal representations of stable polynomials on a matrix polyball. Math. Z. **283**, 25–37 (2016)
11. A. Grinshpan, D.S. Kaliuzhnyi-Verbovetskyi, V. Vinnikov, H.J. Woerdeman, Matrix-valued Hermitian Positivstellensatz, lurking contractions, and contractive determinantal representations of stable polynomials. Oper. Theory Adv. Appl. **255**, 123–136 (2016)
12. A. Grinshpan, D.S. Kaliuzhnyi-Verbovetskyi, H.J. Woerdeman, Norm-constrained determinantal representations of multivariable polynomials. Complex Anal. Oper. Theory **7**, 635–654 (2013)
13. A. Grinshpan, D.S. Kaliuzhnyi-Verbovetskyi, H.J. Woerdeman, The Schwarz lemma and the Schur-Agler class, in *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems*, Groningen (2014)
14. L.K. Hua. *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains.* Translated from the Russian by Leo Ebner and Adam Korányi (American Mathematical Society, Providence, 1963), pp. iv+164
15. D.S. Kalyuzhnyi-Verbovetzkii, On the Bessmertnyi class of homogeneous positive holomorphic functions on a product of matrix halfplanes, in *Operator Theory, Systems Theory and Scattering Theory: Multidimensional Generalizations*. Operator Theory, Advances and Applications, vol. 157 (Birkhäuser, Basel, 2005), pp. 139–164
16. G. Knese, Rational inner functions in the Schur-Agler class of the polydisk. Publ. Mat. **55**, 343–357 (2011)
17. A. Korányi, S. Vagi, Rational inner functions on bounded symmetric domains. Trans. Am. Math. Soc. **254**, 179–193 (1979)
18. W. Rudin, *Function Theory in Polydiscs* (W. A. Benjamin, Inc., New York/Amsterdam, 1969)
19. J.L. Taylor, A joint spectrum for several commuting operators. J. Funct. Anal. **6**, 172–191 (1970)
20. J.L. Taylor, The analytic-functional calculus for several commuting operators. Acta Math. **125**, 1–38 (1970)
21. N.Th. Varopoulos, On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory. J. Funct. Anal. **16**, 83–100 (1974)

A Note on Local Hölder Continuity of Weighted Tauberian Functions

Paul Hagelstein and Ioannis Parissis

Abstract Let M and $\mathsf{M_S}$ respectively denote the Hardy-Littlewood maximal operator with respect to cubes and the strong maximal operator on $\mathbb{R}^n$, and let w be a nonnegative locally integrable function on $\mathbb{R}^n$. We define the associated Tauberian functions $\mathsf{C}_{\mathsf{HL},w}(\alpha)$ and $\mathsf{C}_{\mathsf{S},w}(\alpha)$ on $(0, 1)$ by

$$\mathsf{C}_{\mathsf{HL},w}(\alpha) := \sup_{\substack{E\subset\mathbb{R}^n \\ 0<w(E)<\infty}} \frac{1}{w(E)} w(\{x \in \mathbb{R}^n : \mathsf{M}\chi_E(x) > \alpha\})$$

and

$$\mathsf{C}_{\mathsf{S},w}(\alpha) := \sup_{\substack{E\subset\mathbb{R}^n \\ 0<w(E)<\infty}} \frac{1}{w(E)} w(\{x \in \mathbb{R}^n : \mathsf{M_S}\chi_E(x) > \alpha\}).$$

Utilizing weighted Solyanik estimates for M and $\mathsf{M_S}$, we show that the function $\mathsf{C}_{\mathsf{HL},w}$ lies in the local Hölder class $C^{(c_n[w]_{A_\infty})^{-1}}(0, 1)$ and $\mathsf{C}_{\mathsf{S},w}$ lies in the local Hölder class $C^{(c_n[w]_{A^*_\infty})^{-1}}(0, 1)$, where the constant $c_n > 1$ depends only on the dimension n.

2010 *Mathematics Subject Classification.* Primary 42B25, Secondary: 42B35

P. H. is partially supported by a grant from the Simons Foundation (#208831 to Paul Hagelstein).
I. P. is supported by IKERBASQUE.

P. Hagelstein (✉)
Department of Mathematics, Baylor University, Waco, TX 76798, USA
e-mail: paul_hagelstein@baylor.edu

I. Parissis
Departamento de Matemáticas, Universidad del Pais Vasco, Aptdo. 644, 48080 Bilbao, Spain

Ikerbasque, Basque Foundation for Science, Bilbao, Spain
e-mail: ioannis.parissis@ehu.es

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_11

Introduction

This note concerns how Solyanik estimates may be used to establish local Hölder continuity estimates for the Tauberian functions associated to the Hardy-Littlewood and strong maximal operators in the context of Muckenhoupt weights. In [4], Hagelstein and Parissis used Solyanik estimates to prove that the Tauberian functions $\mathsf{C}_{\mathsf{HL}}(\alpha)$ and $\mathsf{C}_{\mathsf{S}}(\alpha)$ associated to the Hardy-Littlewood and strong maximal operators in $\mathbb{R}^n$ both lie in the local Hölder class $C^{1/n}(1,\infty)$. The techniques of that paper are surprisingly robust, and we here will show how the weighted Solyanik estimates for the Hardy-Littlewood and strong maximal operators obtained in [5, 6] may be used to establish local Hölder smoothness estimates for the Tauberian functions of the Hardy-Littlewood and strong maximal operators in the weighted scenario.

We now briefly review what Solyanik estimates are and how they may be used to establish local smoothness estimates for Tauberian functions associated to geometric maximal operators in the setting of Lebesgue measure. Let $\mathcal{B}$ be a collection of sets of positive measure in $\mathbb{R}^n$, and define the associated geometric maximal operator $\mathsf{M}_{\mathcal{B}}$ by

$$\mathsf{M}_{\mathcal{B}}f(x) := \sup_{x\in R\in\mathcal{B}} \frac{1}{|R|}\int_R |f|.$$

For $0<\alpha<1$, the associated Tauberian function $\mathsf{C}_{\mathcal{B}}(\alpha)$ is given by

$$\mathsf{C}_{\mathcal{B}}(\alpha) := \sup_{\substack{E\subset\mathbb{R}^n\\ 0<|E|<\infty}} \frac{1}{|E|}|\{x\in\mathbb{R}^n : \mathsf{M}_{\mathcal{B}}\chi_E(x)>\alpha\}|.$$

Our ordinary expectation is that, provided $\mathcal{B}$ is a basis with reasonable differentiation properties, for $0<\alpha<1$ and α very close to 1, we should have $|\{x\in\mathbb{R}^n : \mathsf{M}_{\mathcal{B}}\chi_E(x)>\alpha\}|$ is very close to $|E|$ itself, and accordingly that $\mathsf{C}_{\mathcal{B}}(\alpha)$ is very close to 1. Solyanik estimates provide a quantitative validation of this expectation. In particular, we have the following theorem due to Solyanik [9]; see also [3].

Theorem 1.1 (Solyanik [9]) *We have the following* Solyanik estimates *for the Hardy-Littlewood and the strong maximal operator:*

(a) Let M *denote the uncentered Hardy-Littlewood maximal operator on* $\mathbb{R}^n$ *with respect to cubes, and define the associated Tauberian function* $\mathsf{C}_{\mathsf{HL}}(\alpha)$ *by*

$$\mathsf{C}_{\mathsf{HL}}(\alpha) = \sup_{\substack{E\subset\mathbb{R}^n\\ 0<|E|<\infty}} \frac{1}{|E|}|\{x\in\mathbb{R}^n : \mathsf{M}\chi_E(x)>\alpha\}|.$$

Then for $\alpha \in (0, 1)$ *sufficiently close to* 1 *we have*

$$\mathsf{C}_{\mathsf{HL}}(\alpha) - 1 \lesssim_n (1 - \alpha)^{1/n}.$$

(b) Let M_{S} *denote the strong maximal operator on* $\mathbb{R}^n$*, and define the associated Tauberian function* $\mathsf{C}_{\mathsf{S}}(\alpha)$ *by*

$$\mathsf{C}_{\mathsf{S}}(\alpha) := \sup_{\substack{E \subset \mathbb{R}^n \\ 0<|E|<\infty}} \frac{1}{|E|} |\{x \in \mathbb{R}^n : \mathsf{M}_{\mathsf{S}}\chi_E(x) > \alpha\}|.$$

Then for $\alpha \in (0, 1)$ *sufficiently close to* 1 *we have*

$$\mathsf{C}_{\mathsf{S}}(\alpha) - 1 \lesssim_n (1 - \alpha)^{1/n}.$$

The following theorem associated to the embedding of so-called halo sets enables us to relate Solyanik estimates to Hölder smoothness estimates.

Theorem 1.2 (Hagelstein and Parissis [4]) *Let* $\mathcal{B}$ *be a homothecy invariant collection of rectangular parallelepipeds in* $\mathbb{R}^n$*. Given a set* $E \subset \mathbb{R}^n$ *of finite measure and* $0 < \alpha < 1$*, define the associated halo set* $\mathcal{H}_\alpha(E)$ *by*

$$\mathcal{H}_{\mathcal{B},\alpha}(E) := \{x \in \mathbb{R}^n : \mathsf{M}_{\mathcal{B}}\chi_E(x) > \alpha\} .$$

Then for all $\alpha, \delta \in (0, 1)$ *with* $\alpha < 1 - \delta$*, we have*

$$\mathcal{H}_{\mathcal{B},\alpha}(E) \subset \mathcal{H}_{\mathcal{B},\alpha(1+2^{-(n+1)}\delta)}(\mathcal{H}_{\mathcal{B},1-\delta}(E)).$$

An immediate corollary of this theorem is the following.

Corollary 1.3 (Hagelstein and Parissis [4]) *Let* $\mathcal{B}$ *be a homothecy invariant collection of rectangular parallelepipeds in* $\mathbb{R}^n$ *and let* $\alpha, \delta \in (0, 1)$*. Then for* $\alpha < 1 - \delta$ *we have*

$$\mathsf{C}_{\mathcal{B}}(\alpha) \leq C_{\mathcal{B}}\big(\alpha(1 + 2^{-(n+1)}\delta)\big)C_{\mathcal{B}}(1 - \delta).$$

Now, we of course have that $\mathsf{C}_{\mathcal{B}}(\alpha)$ is nonincreasing on $(0, 1)$. If $\mathcal{B}$ is the collection of rectangular parallelepipeds in $\mathbb{R}^n$ whose sides are parallel to the axes (so that $\mathsf{M}_{\mathcal{B}} = \mathsf{M}_{\mathsf{S}}$), we can accordingly combine the above corollary with the Solyanik estimates for M_{S} provided by Theorem 1.1 to relatively easily obtain the following.

Corollary 1.4 (Hagelstein and Parissis [4]) *Let* $\mathsf{C}_{\mathsf{HL}}(\alpha)$ *and* $\mathsf{C}_{\mathsf{S}}(\alpha)$ *respectively denote the Tauberian functions associated to the Hardy-Littlewood maximal operator with respect to cubes and the strong maximal operator in* $\mathbb{R}^n$ *with respect to* α*. Then*

$$\mathsf{C}_{\mathsf{HL}} \in C^{1/n}(0, 1) \quad \textit{and} \quad \mathsf{C}_{\mathsf{S}} \in C^{1/n}(0, 1).$$

The purpose of this note is to establish weighted analogues of Corollary 1.4. To make this precise let us consider a non-negative, locally integrable function w on $\mathbb{R}^n$. The relevant Tauberian functions $\mathsf{C}_{\mathsf{HL},w}(\alpha)$ and $\mathsf{C}_{\mathsf{S},w}(\alpha)$ are defined on $(0,1)$ by

$$\mathsf{C}_{\mathsf{HL},w}(\alpha) := \sup_{\substack{E\subset\mathbb{R}^n \\ 0<w(E)<\infty}} \frac{1}{w(E)} w(\{x\in\mathbb{R}^n : \mathsf{M}\chi_E(x) > \alpha\})$$

and

$$\mathsf{C}_{\mathsf{S},w}(\alpha) := \sup_{\substack{E\subset\mathbb{R}^n \\ 0<w(E)<\infty}} \frac{1}{w(E)} w(\{x\in\mathbb{R}^n : \mathsf{M}_{\mathsf{S}}\chi_E(x) > \alpha\}).$$

It was shown in [7] that the condition $\mathsf{C}_{\mathsf{HL},w}(\alpha) < +\infty$ for *some* $\alpha \in (0,1)$ already implies that $\mathsf{M} : L^p(w) \to L^p(w)$ for *some* $1 < p < \infty$ and, similarly if $\mathsf{C}_{\mathsf{S},w}(\alpha) < +\infty$ for *some* $\alpha \in (0,1)$ then $\mathsf{M}_{\mathsf{S}} : L^p(w) \to L^p(w)$ for *some* $1 < p < \infty$. These results pose an important restriction on the kind of functions w we can consider in proving Hölder regularity estimates for $\mathsf{C}_{\mathsf{HL},w}$ and $\mathsf{C}_{\mathsf{S},w}$. In particular, it is well known that the class of functions w such that $\mathsf{M} : L^p(w) \to L^p(w)$ for some $p \in (1,\infty)$ is the Muchkenhoupt class of weights A_∞; see for example [2]. Here we use the Fujii-Wilson definition of the Muckenhoupt class A_∞. Namely, the weight w belongs to the class A_∞ if and only if

$$[w]_{A_\infty} := \sup_Q \frac{1}{w(Q)} \int_Q \mathsf{M}(w\chi_Q) < +\infty,$$

where the supremum is taken with respect to all cubes in $\mathbb{R}^n$ whose sides are parallel to the axes. This description of the class A_∞ goes back to Fujii [8], and Wilson, [10, 11]; see also [1]. Thus $w \in A_\infty$ is a necessary condition for the continuity of $\mathsf{C}_{\mathsf{HL},w}$ on $(0,1)$. It turns out that $w \in A_\infty$ is also a sufficient condition for the Hölder regularity of $\mathsf{C}_{\mathsf{HL},w}$.

Theorem 1.5 *Let $w \in A_\infty$ be a Muckenhoupt weight on $\mathbb{R}^n$. Then*

$$\mathsf{C}_{\mathsf{HL},w} \in C^{(c_n[w]_{A_\infty})^{-1}}(0,1),$$

where the constant c_n depends only on the dimension n.

Moving to the multiparameter case, the condition that $\mathsf{M}_{\mathsf{S}} : L^p(w) \to L^p(w)$ for some $p \in (1,\infty)$ is equivalent to the condition $w \in A_\infty^*$, where A_∞^* denotes the class of *multiparameter* or *strong* Muckenhoupt weights. A few words about how the multiparameter Muckenhoupt class A_∞^* is defined are in order here. For

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $1 \le j \le n$ we may associate the point $\bar{x}^j := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$. Associated to a non-negative locally integrable function w on $\mathbb{R}^n$ and $\bar{x}^j$ is the one-dimensional weight

$$w_{\bar{x}^j}(t) := w(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_n), \qquad t \in \mathbb{R}.$$

Then $[w]_{A_\infty^*}$ is defined by

$$[w]_{A_\infty^*} := \sup_{1 \le j \le n} \operatorname*{ess\,sup}_{\bar{x}^j \in \mathbb{R}^{n-1}} [w_{\bar{x}^j}]_{A_\infty}.$$

Here $[\nu]_{A_\infty}$ denotes the standard Fujii-Wilson A_∞ constant of a weight ν on $\mathbb{R}^1$, given by

$$[\nu]_{A_\infty} := \sup_I \frac{1}{w(I)} \int_I \mathsf{M}_1(\nu \chi_I),$$

where the supremum is taken over all intervals $I \subset \mathbb{R}$ and M_1 denotes the Hardy-Littlewood maximal operator on $\mathbb{R}^1$. Thus a weight w is a multiparameter Muckenhoupt weight if and only if $[w]_{A_\infty^*} < +\infty$. We refer the reader to [6] and the references therein for more details on the definition and properties of multiparameter Muckenhoupt weights.

With the definition of multiparameter Muckenhoupt weights in hand, the previous discussion shows that a necessary condition for the continuity of $\mathsf{C}_{\mathsf{S},w}$ on $(0, 1)$ is that $w \in A_\infty^*$. As in the one parameter case, we show that $w \in A_\infty^*$ is also sufficient for the Hölder continuity of $\mathsf{C}_{\mathsf{S},w}$ on $(0, 1)$.

Theorem 1.6 *Let $w \in A_\infty^*$ be a multiparameter Muckenhoupt weight on $\mathbb{R}^n$. Then*

$$\mathsf{C}_{\mathsf{S},w} \in C^{(c_n [w]_{A_\infty^*})^{-1}}(0, 1),$$

where the constant c_n depends only on the dimension n.

Notation

We use the letters C, c to denote positive numerical constants whose value might change even in the same line of text. We express the dependence of a constant C on some parameter n by writing C_n. We write $A \lesssim B$ if $A \le CB$ for some numerical constant $C > 0$. If $A \le C_n B$ we then write $A \lesssim_n B$. In this note, w will always denote a non-negative, locally integrable function on $\mathbb{R}^n$. Finally, we say that a function f lies in the Hölder class $C^p(I)$ for some interval $I \subset \mathbb{R}$ if for every compact set $K \subset I$ we have $|f(x) - f(y)| \lesssim_K |x - y|^p$ for all $x, y \in K$. In this case we will say that f is *locally Hölder continuous* with exponent p in I.

Weighed Solyanik Estimates and Hölder Regularity

In this section we show that the strategy for establishing Hölder smoothness estimates for $\mathsf{C}_{\mathsf{HL}}(\alpha)$ and $\mathsf{C}_{\mathsf{S}}(\alpha)$ may be adapted to the weighted context. To implement the above strategy, we need Solyanik estimates that provide us quantitative information as to how close $\mathsf{C}_{\mathsf{HL},w}(\alpha)$ and $\mathsf{C}_{\mathsf{S}}(\alpha)$ are to 1 for α near 1. Of course, the related estimates are expected to depend on w. Suitable Solyanik estimates in this regard were found in [5, 6] when w is a Muckenhoupt weight. In particular, we have the following:

Theorem 3.1 (Hagelstein and Parissis [5, 6]) *Let $w \in A_\infty$. We have the Solyanik estimate*

$$\mathsf{C}_{\mathsf{HL},w}(\alpha) - 1 \lesssim_n \Delta_w^2 (1-\alpha)^{(c_n[w]_{A_\infty})^{-1}} \quad \textit{whenever} \quad 1 > \alpha > 1 - e^{-c_n[w]_{A_\infty}}.$$

Here Δ_w is the doubling constant of w, and c_n and the implied constant depend only upon the dimension n.

A multiparameter analogue of Theorem 3.1 the following.

Theorem 3.2 (Hagelstein and Parissis [6]) *Let w be a non-negative, locally integrable function in $\mathbb{R}^n$. If $w \in A_\infty^*$ we have*

$$\mathsf{C}_{\mathsf{S},w}(\alpha) - 1 \lesssim_n (1-\alpha)^{(cn[w]_{A_\infty^*})^{-1}} \quad \textit{for all} \quad 1 > \alpha > 1 - e^{-cn[w]_{A_\infty^*}},$$

where $c > 0$ is a numerical constant.

With these *weighted Solyanik estimates* at our disposal we can now give the proof of the Hölder continuity estimates for $\mathsf{C}_{\mathsf{HL},w}$ and $\mathsf{C}_{\mathsf{S},w}$.

Proof of Theorem 1.5 Let K be a compact subset in $(0,1)$ and let $m_K, M_K \in (0,1)$ be such that $m_K \leq x \leq M_k$ for all $x \in K$. Since $w \in A_\infty$ there exists some $q \in (0,1)$ such that $\mathsf{M} : L^q(w) \to L^{q,\infty}(w)$ and thus $\sup_{\alpha\in K} \mathsf{C}_{\mathsf{HL},w}(\alpha) \lesssim_{w,n,K} 1$. Furthermore, by Theorem 3.1 we have that

$$\mathsf{C}_{\mathsf{HL},w}(\alpha) - 1 \lesssim_{w,n} (1-\alpha)^{(c_n[w]_{A_\infty})^{-1}} \quad \text{for all} \quad 1 > \alpha > 1 - e^{-c_n[w]_{A_\infty}} =: \alpha_o. \quad (1)$$

We first consider $x, y \in K$ with $0 < y - x < \min(\frac{1-M_K}{2^{n+1}} m_K, \frac{1-\alpha_o}{2^{n+1}} m_K) =: \eta$. We can then write

$$\mathsf{C}_{\mathsf{HL},w}(x) - \mathsf{C}_{\mathsf{HL},w}(y) = \mathsf{C}_{\mathsf{HL},w}(x) - \mathsf{C}_{\mathsf{HL},w}\Big(x\big(1 + 2^{n+1}\frac{y-x}{2^{n+1}x}\big)\Big).$$

Now observe that by our choice of x, y we have

$$2^{n+1}\frac{y-x}{x} < 2^{n+1}\frac{1-M_K}{2^{n+1}} m_K \frac{1}{m_K} \leq 1 - M_K \leq 1 - x.$$

We can thus apply Theorem 1.2 with x in the role of $\alpha := x$ and $\delta := 2^{n+1}\frac{y-x}{x}$ to get

$$\mathcal{H}_{\mathcal{B},x}(E) \subset \mathcal{H}_{\mathcal{B},y}(\mathcal{H}_{\mathcal{B},(1-\delta)}(E))$$

for all measurable E where here $\mathcal{B}$ denotes the collection of all cubes in $\mathbb{R}^n$ whose sides are parallel to the axes. This immediately implies

$$\mathsf{C}_{\mathsf{HL},w}(x) \leq \mathsf{C}_{\mathsf{HL},w}(y)\mathsf{C}_{\mathsf{HL},w}\big(1 - 2^{n+1}\frac{y-x}{x}\big).$$

Thus we can estimate

$$\begin{aligned}\mathsf{C}_{\mathsf{HL},w}(x) - \mathsf{C}_{\mathsf{HL},w}(y) &\leq \mathsf{C}_{\mathsf{HL},w}(y)\Big[\mathsf{C}_{\mathsf{HL},w}\big(1 - 2^{n+1}\frac{y-x}{x}\big) - 1\Big] \\ &\lesssim_{w,n,K} \mathsf{C}_{\mathsf{HL},w}\big(1 - 2^{n+1}\frac{y-x}{x}\big) - 1\end{aligned}$$

since $\sup_{\alpha\in K} \mathsf{C}_{\mathsf{HL},w}(\alpha) \lesssim_{w,n,K} 1$. Noting that

$$1 > 1 - 2^{n+1}\frac{y-x}{x} > 1 - 2^{n+1}\frac{1-\alpha_o}{2^{n+1}x}m_K \geq \alpha_o,$$

an appeal to (1) gives

$$\mathsf{C}_{\mathsf{HL},w}(x) - \mathsf{C}_{\mathsf{HL},w}(y) \lesssim_{w,n,K} \big(\frac{y-x}{x}\big)^{(c_n[w]_{A_\infty})^{-1}} \lesssim_K (y-x)^{(c_n[w]_{A_\infty})^{-1}}.$$

We have shown that

$$\sup_{\substack{x,y\in K\\ |y-x|<\eta}} \frac{|\mathsf{C}_{\mathsf{HL},w}(y) - \mathsf{C}_{\mathsf{HL},w}(x)|}{|y-x|^{(c_n[w]_{A_\infty})^{-1}}} \lesssim_{w,n,K} 1.$$

On the other hand, if $x, y \in K$ with $y - x \geq \eta$ then the Hölder estimate follows trivially since $\sup_{x,y\in K} |\mathsf{C}_{\mathsf{HL},w}(x) - \mathsf{C}_{\mathsf{HL},w}(y)| \lesssim_{w,n,K} 1$ so we are done. □

The proof of Theorem 1.6 is virtually identical to that of Theorem 1.5.

One may naturally wonder how sharp the above smoothness estimates are for $\mathsf{C}_{\mathsf{HL},w}(\alpha)$ and $\mathsf{C}_{\mathsf{S},w}(\alpha)$. In particular we may ask the questions: Are $\mathsf{C}_{\mathsf{HL},w}(\alpha)$ and $\mathsf{C}_{\mathsf{S},w}(\alpha)$ differentiable on $(0, 1)$? Are they in fact smooth on $(0, 1)$? To the best of our knowledge, even the question of whether or not the sharp Tauberian constant $\mathsf{C}_{\mathsf{HL}}(\alpha)$ of the Hardy-Littlewood maximal operator on $\mathbb{R}$ in the Lebesgue setting is differentiable constitutes an unsolved problem. All of these topics remain a subject of continuing research.

References

1. N. Fujii, Weighted bounded mean oscillation and singular integrals. Math. Japon. **22**(5), 529–534 (1977/1978). MR0481968 (58 #2058)
2. J. García-Cuerva, J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*. North-Holland Mathematics Studies, vol. 116 (North-Holland Publishing Co., Amsterdam, 1985). Notas de Matemática [Mathematical Notes], 104. MR807149 (87d:42023)
3. P. Hagelstein, I. Parissis, Solyanik estimates in harmonic analysis, in *Special Functions, Partial Differential Equations, and Harmonic Analysis*. Springer Proceedings in Mathematics and Statistics, vol. 108 (Springer, Heidelberg, 2014), pp. 87–103
4. P.A. Hagelstein, I. Parissis, Solyanik estimates and local Hölder continuity of halo functions of geometric maximal operators. Adv. Math. **285**, 434–453 (2015)
5. P.A. Hagelstein, I. Parissis, Weighted Solyanik estimates for the Hardy-Littlewood maximal operator and embedding of A_∞ *into* A_p. J. Geom. Anal. **26**, 924–946 (2016)
6. P.A. Hagelstein, I. Parissis, Weighted Solyanik estimates for the strong maximal function Publ. Mat. (2014). (to appear). Available at 1410.3402
7. P.A. Hagelstein, T. Luque, I. Parissis, Tauberian conditions, Muckenhoupt weights, and differentiation properties of weighted bases. Trans. Am. Math. Soc. **367**, 7999–8032 (2015)
8. T. Hytönen, C. Pérez, Sharp weighted bounds involving A_∞. Anal. PDE **6**(4), 777–818 (2013). MR3092729
9. A.A. Solyanik, On halo functions for differentiation bases. Mat. Zametki **54**(6), 82–89, 160 (1993) (Russian, with Russian summary); English trans., Math. Notes **54** (1993), no. 5–6, 1241–1245 (1994). MR1268374 (95g:42033)
10. J.M. Wilson, Weighted inequalities for the dyadic square function without dyadic A_∞. Duke Math. J. **55**(1), 19–50 (1987). MR883661 (88d:42034)
11. M. Wilson, *Weighted Littlewood-Paley Theory and Exponential-Square Integrability*. Lecture Notes in Mathematics, vol. 1924 (Springer, Berlin, 2008). MR2359017 (2008m:42034)

Three Observations on Commutators of Singular Integral Operators with BMO Functions

Carlos Pérez and Israel P. Rivera-Ríos

Abstract Three observations on commutators of Singular Integral Operators with BMO functions are exposed, namely

1. The already known subgaussian local decay for the commutator, namely

$$\frac{1}{|Q|}\left|\left\{x\in Q\,:\,|[b,T](f\chi_Q)(x)|>M^2f(x)t\right\}\right|\le ce^{-\sqrt{ct\|b\|_{BMO}}}$$

is sharp, since it cannot be better than subgaussian.
2. It is not possible to obtain a pointwise control of the commutator by a finite sum of sparse operators defined by $L\log L$ averages.
3. Motivated by the conjugation method for commutators, it is shown the failure of the following endpoint estimate, if $w\in A_p\setminus A_1$ then

$$\left\|wM\left(\frac{f}{w}\right)\right\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)}=\infty.$$

2010 *Mathematics Subject Classification.* 42B35,46E30

The first author was supported by Severo Ochoa Excellence Programme and the Spanish Government grant MTM2014-53850-P and the second author was supported by Grant MTM2012-30748, Spanish Government.

C. Pérez (✉)
Department of Mathematics, University of the Basque Country UPV/EHU, IKERBASQUE, Basque Foundation for Science, Bilbao and BCAM, Basque Center for Applied Mathematics, Bilbao, Spain
e-mail: carlos.perezmo@ehu.es

I.P. Rivera-Ríos
IMUS & Departamento de Análisis Matemático, Universidad de Sevilla, Sevilla, Spain
e-mail: petnapet@gmail.com

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_12

Introduction

The purpose of this paper is to present some observations concerning commutators of singular integral operators with BMO functions. These operators were introduced by Coifman, Rochberg and Weiss in [6] as a tool to extend the classical factorization theorem for Hardy spaces in the unit circle to $\mathbb{R}^n$. These operators are defined by the expression

$$T_bf(x) = \int_{\mathbb{R}^n} (b(x) - b(y))K(x, y)f(y)\, dy, \tag{1}$$

where K is a kernel satisfying the standard Calderón-Zygmund estimates and where b, the "symbol" of the operator, is a locally integrable function. Of course, these are special cases of the more general commutators given by the expression

$$T_b = [b, T] = M_b \circ T - T \circ M_b$$

where T is any operator and M_b is the multiplication operator $M_bf = b \cdot f$.

The classical well known result from [6] establishes that $[b, T]$ is a bounded operator on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, when the symbol b is a BMO function. We state this result.

Theorem 1.1 *Let T be a singular integral operator and b a BMO function. The commutator T_b is bounded on $L^p(\mathbb{R}^n)$ for every $1 < p < \infty$.*

In the same paper it is shown that $b \in BMO$ is also a necessary condition namely, if the commutators $[b, R_j]$, $j = 1, \cdots, n$ of b with the Riesz transforms R_j are bounded on $L^p(\mathbb{R}^n)$ for some $p \in (1, \infty)$ and every $j \in \{1, 2, \ldots, n\}$ then $b \in BMO$.

None of the different proofs of this result follows the usual scheme of the classical Calderón-Zygmund theory for proving the $L^p(\mathbb{R}^n)$ boundedness of singular integral operators T. Two proofs of Theorem 1.1 can be found in [6]. The first and main one in that paper is based on methods involving techniques similar to those used in [5] to understand the Calderón commutator. As far as we know this approach has not been so influential. However, the second proof, based on the so called conjugation method from operator theory, has been widely used. In fact, it is quite surprising that this proof was postponed to the end of the paper since it turns out to be highly interesting. Indeed, the method shows the intimate connection between these commutators and the A_p theory of weights. Furthermore, this proof can be applied to general linear operators, not only for Singular Integral Operators. As a sample we will point out the following particular L^2 case:

Theorem 1.2 *Suppose that T is a linear operator such that*

$$T : L^2(w) \longrightarrow L^2(w)$$

for every $w \in A_2$. *Then for every* $b \in BMO$,

$$[b, T] : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n).$$

The method of proof can be carried out in more generality as shown in [1]. The key initial argument of the proof is that we can write $[b, T]$ as a complex integral operator using the Cauchy integral theorem as follows

$$[b, T]f = \frac{d}{dz} e^{zb} T(fe^{-zb})\Big|_{z=0} = \frac{1}{2\pi i} \int_{|z|=\varepsilon} \frac{T_z(f)}{z^2}\, dz\,, \qquad \varepsilon > 0$$

where

$$z \to T_z(f) := e^{zb} T\left(\frac{f}{e^{zb}}\right) \qquad z \in \mathbb{C}.$$

This is called the "conjugation" of T by e^{zb} and the terminology comes most probably from group theory. Now, if $\|\cdot\|$ is a norm we can apply Minkowski inequality:

$$\|[b, T]f\| \leq \frac{1}{2\pi\, \varepsilon} \sup_{|z|=\varepsilon} \|T_z(f)\| \qquad \varepsilon > 0.$$

The effectiveness of the method can be checked in the modern context of weighted L^p estimates. Indeed, the method produces very optimal bounds of the operator norm as shown in [4] (see also [15]).

This method reveals the role played by the following operation:

$$f \to T_w(f) := w\, T\left(\frac{f}{w}\right)$$

where w is a weight which, in this context, is an A_p weight. Indeed, this is the case by the well known key property of the BMO class, if $p > 1$ and $b \in BMO$ then there is a small ε_0, such that $e^{tb} \in A_p$, for any real number t such that $|t| < \varepsilon_0$. These operators were already studied by B. Muckenhoupt and R. Wheeden in the 1970s and by E. Sawyer in the 1980s. Some of the problems they left open were solved in [8]. A consequence of the main result of [8] is that if $w \in A_1$ then T_w is of weak type $(1, 1)$, namely

$$\|T_w\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)} < \infty$$

with bound depending upon the A_1 constant of w. However, we will exhibit examples of weights $w \in A_p \setminus A_1$ in section "Third Observation: The Failure of an Endpoint Estimate Motivated by the Conjugation Method" for which T_w is **not** of weak type $(1, 1)$, namely

$$\|T_w\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)} = \infty.$$

This shows that the case $w \in A_1$ is specially relevant. Perhaps, this phenomenon can be explained by the fact that the conjugation method is closely attached to commutators with BMO functions which are **not** of weak type $(1,1)$ as observed in [25]. Indeed, the conjugation method works due to the property, already mentioned, that if $p > 1$ and $b \in BMO$ then $e^{tb} \in A_p$ for small values of t. However, this property turns out to be false in the case $p = 1$. The lack of the weak type $(1,1)$ property for commutators is replaced by a $L \log L$ inequality like (5) below and not better.

There is another proof of Theorem 1.1 based on the use of the sharp maximal function of C. Fefferman and E. Stein which has also been very influential. It seems that it was first discovered by J. O. Strömberg as mentioned by S. Janson in [17] (see also [28] pp. 417–419) The proof relies on combining the following key pointwise estimate

$$M^\sharp([b,T]f) \leq c\|b\|_{BMO}\left(M_r\,(Tf) + M_s\,(f)\right) \tag{2}$$

where $1 < r, s < \infty$ and $M_r(f) = M(|f|^r)^{1/r}$ together with the classical Fefferman-Stein inequality:

$$\|M(f)\|_{L^p} \leq c\|M^\sharp(f)\|_{L^p}.$$

Here we use standard notation, M is the Hardy-Littlewood maximal function and $M^\sharp$ is the sharp maximal function. The L^p boundedness of M and T yields the alternative proof of Theorem 1.1. Proceeding in the same way we obtain the corresponding estimates for A_p weights.

This approach was considered by S. Bloom in [2] extending in an interesting way Theorem 1.1 but only on the real line.

Theorem 1.3 *Let $\mu, \lambda \in A_p$ and let H be the Hilbert transform:*

$$[b,H]\,:\,L^p(\mu) \longrightarrow L^p(\lambda)$$

where $\nu = \mu^{\frac{1}{p}}\lambda^{-\frac{1}{p}}$ if and only if

$$\|b\|_{BMO(\nu)} = \sup_Q \frac{1}{\nu(Q)}\int_Q |b - b_Q| < \infty. \tag{3}$$

The power of the pointwise estimate (2) is reflected in many situations, for instance in [12], where similar results were derived for commutators of strongly singular integral with symbol in the new BMO class (3) (see also [13, 14] for an alternative approach based on dyadic shifts).

However, estimate (2) is not sharp enough for many purposes and much better results can be obtained with the following variation:

$$M_\delta^\sharp([b,T]f) \leq c\|b\|_{BMO}\left(M_\varepsilon\,(Tf) + M^2\,(f)\right) \quad 0 < \delta < \varepsilon < 1 \tag{4}$$

where M^2 stands for $M \circ M$ (see [25]). Here, the key difference is that we are considering small parameters δ and ε. The estimate is sharp since M^2 cannot be replaced by the (pointwise) smaller operator M. Indeed, otherwise these commutators would be of weak type (1, 1) but, as we mentioned above, this is not the case [25] where it is shown that commutators satisfy the following "$L \log L$" type estimate,

$$w\left(\{x \in \mathbb{R}^n : |[b,T]f(x)| > \lambda\}\right) \le c \int_{\mathbb{R}^n} \Phi\left(\frac{|f|}{\lambda}\|b\|_{BMO}\right) w dx \quad \lambda > 0, \tag{5}$$

where $w \in A_1$, $\Phi(t) = t\log(e+t)$ and where $c > 0$ depends upon the A_1 constant. This shows that these commutators are "more singular" than Calderón-Zygmund operators. The original proof of (5) follows from the key pointwise (4) combined with a good-λ type argument, but an alternative proof was obtained by the first author and G. Pradolini in [26] with the bonus that non A_∞ weights can be considered. This argument is based on a variation of the classical scheme used to prove the weak type (1, 1) for Calderón-Zygmund operators. The statement of the result is the following.

Theorem 1.4 *Let T be a Calderón-Zygmund operator and $b \in BMO$. If w is an arbitrary weight the following inequality holds*

$$w\left(\{x \in \mathbb{R}^n : |[b,T]f(x)| > \lambda\}\right) \le C_{\varepsilon,T} \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO}\frac{|f(x)|}{\lambda}\right) M_{L(\log L)^{1+\varepsilon}} w(x) dx$$

for every $\varepsilon > 0$.

Very recently (cf. [27]) the authors have obtained a quantitative version of the endpoint estimate for arbitrary weights, namely Theorem 1.4. This result is analogous to the one obtained by the first author and T. Hytönen for singular integrals in [15].

Theorem 1.5 *Let T be a Calderón-Zygmund operator and $b \in BMO$. If $w \ge 0$ is a weight then, for every $\varepsilon > 0$*

$$w\left(\{x \in \mathbb{R}^n : |[b,T]f(x)| > \lambda\}\right) \le \frac{c}{\varepsilon^2} \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO}\frac{|f|}{\lambda}\right) M_{L(\log L)^{1+\varepsilon}} w dx.$$

The main novelty here is the appearance of the sharp factor $\frac{1}{\varepsilon^2}$ reflecting again the higher singularity of the operator. As a corollary of this result we can derive the following result obtained previously by C. Ortiz-Caraballo in [23],

$$w\left(\{x \in \mathbb{R}^n : |[b,T]f(x)| > \lambda\}\right) \le C\Phi\left([w]_{A_1}\right)^2 \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO}\frac{|f|}{\lambda}\right) w dx.$$

We remark that it seems that the conjugation method cannot be applied to prove this estimate. Therefore, estimate (5) or Theorem 1.4 works, so far, for Calderón-Zygmund operators not for general linear operators assuming a minimal appropriate weighted weak type estimate.

Another interesting difference between Calderón-Zygmund operators and commutators concerns their local behavior. A very nice way of expressing this is by means of the following estimate due to Karagulyan [18]: there exists a constant $c > 0$ such that for each cube Q and for each function f supported on the cube Q

$$\frac{1}{|Q|}|\{x \in Q : |Tf(x)| > tMf(x)\}| \leq c\, e^{-ct} \qquad t > 0. \tag{6}$$

This result can be seen as an improvement of Buckley's exponential decay theorem [3] which is a very useful result. For instance, it allows to improve in a quantitative way the classical good-λ inequality between T and M: if $p \in (0, \infty)$ and $w \in A_\infty$

$$\|Tf\|_{L^p(w)} \leq c_T\, p\, [w]_{A_\infty} \|M(f)\|_{L^p(w)}.$$

Motivated by this result of Karagulyan, Ortiz-Caraballo, Rela and the first author developed a new method for proving (6) in [24]. This method is flexible enough to deal with other operators including the commutators. In particular, we have the following sub-gaussian estimate.

Theorem 1.6 *Let T be a Calderón-Zygmund operator and $b \in BMO$, then there exists a constant $c > 0$ such that for each f*

$$\sup_Q \frac{1}{|Q|}|\{x \in Q : |[b, T](f\chi_Q)(x)| > tM^2 f(x)\}| \leq c\, e^{-\sqrt{c\, t \|b\|_{BMO}}} \qquad t > 0. \tag{7}$$

We will show in section "First Observation: Sharpness of the Subexponential Local Decay" that this subexponential decay is fully sharp. In section "Second Observation: A "natural" but False Sparse Domination Result for Commutators", we will provide a new proof of (6) based on the pointwise domination: if T is a Calderón-Zygmund operator, then it is possible to find a finite set of η-sparse families $\{\mathcal{S}_j\}_{j=1}^{3^n}$ (see section "Second Observation: A "natural" but False Sparse Domination Result for Commutators" for the definitions) contained in the same or in different dyadic lattices $\mathcal{D}_j$ and depending on f such that

$$|Tf(x)| \leq c_T \sum_{j=1}^{3^n} A_{\mathcal{S}_j} f(x) \tag{8}$$

where

$$A_{\mathcal{S}_j} f(x) = \sum_{Q \in \mathcal{S}_j} \frac{1}{|Q|} \int_Q |f|\, \chi_Q(x).$$

See section "Second Observation: A "natural" but False Sparse Domination Result for Commutators" for details, in particular Theorem 3.6.

In view of the interest of an estimate like (8) it would be relevant to produce a counterpart for commutators. The “natural” sparse operator for these commutators would be

$$B_{\mathcal{S}}f(x) = \sum_{Q\in\mathcal{S}} \|f\|_{L\log L,Q}\,\chi_Q(x).$$

The reason that leads to consider this sparse operator in terms of the average $\|\cdot\|_{L\log L,Q}$ is due to the intimate relationship of commutators and M^2 which is an operator pointwise equivalent to $M_{L\log L}$. In section “Second Observation: A “natural” but False Sparse Domination Result for Commutators” we prove the impossibility of having a domination theorem for commutators by these “sparse” operators.

First Observation: Sharpness of the Subexponential Local Decay

We prove in this section that Theorem 1.6 is sharp, i.e., we can find a Calderón-Zygmund operator T, a symbol $b\in BMO$ a function f and a cube Q such that

$$\frac{1}{|Q|}|\{x\in Q : |[b,T]f(x)| > tM^2f(x)\}| \geq c\,e^{-\sqrt{c\,t\|b\|_{BMO}}}$$

for some constant $c>0$. More precisely we have the following.

Observation 1 *Let $b(x)=\log|x|$, then we can find a constant $c>0$ such that*

$$|\{x\in(0,1) : |[b,H](\chi_{(0,1)})(x)| > t\}| \geq e^{-\sqrt{ct}}$$

where H stands for the Hilbert transform.

Proof Let $f(x)=\chi_{(0,1)}(x)$. We are going to show that

$$|\{x\in(0,1) : |[b,H]f(x)| > tM^2f(x)\}| = |\{x\in(0,1) : |[b,H]f(x)| > t\}| \geq c\,e^{-\sqrt{\alpha\,t}} \qquad t>0.$$

For $x\in(0,1)$ we have that

$$[b,H]f(x) = \int_0^1 \frac{\log(x)-\log(y)}{x-y}\,dy = \int_0^1 \frac{\log(\frac{x}{y})}{x-y}\,dy = \int_0^{1/x} \frac{\log(\frac{1}{t})}{1-t}\,dt.$$

Now we observe that

$$\int_0^{1/x} \frac{\log(\frac{1}{t})}{1-t}\,dt = \int_0^1 \frac{\log(\frac{1}{t})}{1-t}\,dt + \int_1^{1/x} \frac{\log(\frac{1}{t})}{1-t}\,dt$$

and since $\frac{\log(\frac{1}{t})}{1-t}$ is positive for $(0,1) \cup (1,\infty)$ we have for $0 < x < 1$ that

$$|[b,H]f(x)| > \int_1^{1/x} \frac{\log(\frac{1}{t})}{1-t}\,dt.$$

Finally, a computation shows that

$$\int_1^{1/x} \frac{\log(\frac{1}{t})}{1-t}\,dt \approx \left(\log\frac{1}{x}\right)^2 \qquad x \to 0.$$

Consequently, we have that for some $x_0 < 1$

$$|[b,H]f(x)| > c\left(\log\frac{1}{x}\right)^2 \qquad 0 < x < x_0.$$

and then for some $t_0 > 0$,

$$|\{x \in (0,1) : |[b,H]f(x)| > t\}| \geq \left|\left\{x \in (0,x_0) : c\left(\log\frac{1}{x}\right)^2 > t\right\}\right| = e^{-\sqrt{t/c}} \qquad t > t_0 \quad (9)$$

as we wanted to prove. □

Second Observation: A "natural" but False Sparse Domination Result for Commutators

Before stating the result we are going to prove in this section we need some notation. We borrow it from [21].

Definition 3.1 (Dyadic child) Let Q be a cube (with sides parallel to the axis). We call dyadic child any of the 2^n cubes obtained by partitioning Q by n "median hyperplanes" (planes parallel to the faces of Q and dividing each edge into 2 equal parts).

If we iterate the partition process of the preceding definition we obtain a standard dyadic grid $\mathcal{D}(Q)$ of subcubes of Q which has the usual properties:

1. For each $k = 0, 1, 2, \ldots$ cubes in the k-th generation have sidelength 2^{-k} and tile Q in a regular way.
2. Each Q' in the k-th generation has 2^n children in the in the $(k+1)$-th generation contained in it and one and only one parent in the $(k-1)$-th generation containing it (unless it is Q itself).
3. If $Q', Q'' \in \mathcal{D}(Q)$, then $Q' \cap Q'' = \emptyset$ or $Q' \subseteq Q''$ or $Q'' \subseteq Q'$.
4. If $Q' \in \mathcal{D}(Q)$, then $\mathcal{D}(Q') \subseteq \mathcal{D}(Q)$.

Definition 3.2 (Dyadic lattice) A dyadic lattice $\mathcal{D}$ in $\mathbb{R}^n$ is any collection of cubes such that

(DL-1) If $Q \in \mathcal{D}$ then each dyadic child of Q is in $\mathcal{D}$ as well.
(DL-2) If $Q', Q'' \in \mathcal{D}$ there exists $Q \in \mathcal{D}$ such that $Q', Q'' \in \mathcal{D}(Q)$.
(DL-3) If K is a compact set of $\mathbb{R}^n$ there exists $Q \in \mathcal{D}$ such that $K \subseteq Q$.

There is an easy way to build a dyadic lattice by considering a increasing sequence of dyadic cubes Q_j such that $\cup_{j=1}^{\infty} Q_j = \mathbb{R}^n$. Then

$$\mathcal{D} = \bigcup_{j=1}^{\infty} \mathcal{D}(Q_j)$$

is a dyadic lattice.

Definition 3.3 Let $\eta \in (0, 1)$. We say that a family of cubes $\mathcal{S} \subseteq \mathcal{D}$ is η-sparse if for each $Q \in \mathcal{S}$ we can find a measurable subset $E(Q) \subset Q$ such that:

1. $E(Q)$'s are pairwise disjoint.
2. $\eta|Q| \leq |E(Q)|$

Definition 3.4 Let $\Lambda > 1$. We say a family of cubes $\mathcal{S}$ is Λ-Carleson if for every cube $Q \in \mathcal{D}$ we have

$$\sum_{P \in \mathcal{S}, P \subset Q} |P| \leq \Lambda |Q|.$$

There is an interesting relation between Carleson and sparse families that we summarize in the following lemma that we borrow from [21]

Lemma 3.5 *If $\mathcal{S}$ is a Λ-Carleson family of cubes then it is $\frac{1}{\Lambda}$-sparse. Conversely if $\mathcal{S}$ is a η-sparse family of cubes then it is a $\frac{1}{\eta}$-Carleson family of cubes.*

Armed with all these definitions we can state the following pointwise domination theorem.

Theorem 3.6 *Let T be a Calderón-Zygmund operator. There is a finite set of η-sparse families $\{\mathcal{S}_j\}_{j=1}^{3^n}$ contained in the same or in different dyadic lattices $\mathcal{D}_j$ and depending on f such that*

$$T^* f(x) \leq c_{T,n} \sum_{j=1}^{3^n} A_{\mathcal{S}_j} f(x) \tag{10}$$

where $A_{\mathcal{S}_j} f(x) = \sum_{Q \in \mathcal{S}_j} \frac{1}{|Q|} \int_Q |f| \, \chi_Q(x)$.

The proof of this result can be found in [21] and [7]. In [19] M. Lacey obtains the same estimate for Calderón-Zygmund operators that satisfy a Dini condition.

Recently a fully quantitative version of Lacey's result was obtained in [16] and even more recently this quantitative version has been simplified in [20].

As a sample of the interest of this result we give a different proof of the exponential estimate (6): there exists a constant $c > 0$ such that for each cube Q and for each f supported on the cube Q

$$\frac{1}{|Q|}|\{x \in Q : T^*f(x) > tMf(x)\}| \leq c\, e^{-c\,t} \qquad t > 0, \tag{11}$$

To prove this result we will use the classical *vector-valued extension of the maximal function* introduced by Fefferman and Stein in [11] that can be written as follows:

$$\overline{M}_q f(x) = \Big(\sum_{j=1}^{\infty}(Mf_j(x))^q\Big)^{1/q} = |Mf(x)|_q,$$

where $f = \{f_j\}_{j=1}^{\infty}$ is a vector–valued function.

Taking into account that T^* is controlled by a finite sum of sparse operators it suffices to establish (11) for those operators. Assume that supp$f \subseteq Q$ for a certain cube. It is clear that we can find c_n pairwise disjoint cubes $Q_j \in \mathcal{D}$ which union covers Q and such that $|Q_j| \simeq |Q|$ We can assume those cubes to belong to any sparse family $\mathcal{S} \subset \mathcal{D}$, since it's easy to check, taking into account Lemma 3.5, that adding a finite number of pairwise disjoints cubes to a sparse family the resulting family is again a sparse family. First we are going to prove that if $\mathcal{S}$ is and $Q_j \in \mathcal{S}$ with $|Q_j| \simeq |Q|$, as we have just showed that we can assume, then

$$\frac{1}{|Q|}\left|\left\{x \in Q : \sum_{P\in\mathcal{S},\, P\subseteq Q_j} \chi_P(x) > t\right\}\right| \leq ce^{-ct}. \tag{12}$$

We begin observing that

$$\begin{aligned}
&\frac{1}{|Q|}\left|\left\{x \in Q : \sum_{P\in\mathcal{S},\, P\subseteq Q_j} \chi_P(x) > t\right\}\right| \\
&= \frac{1}{|Q|}\left|\left\{x \in Q \cap Q_j : \sum_{P\in\mathcal{S},\, P\subseteq Q_j} \chi_P(x) > t\right\}\right| \\
&\leq c\frac{1}{|Q_j|}\left|\left\{x \in Q_j : \sum_{P\in\mathcal{S},\, P\subseteq Q_j} \chi_P(x) > t\right\}\right| = C_Q.
\end{aligned}$$

We use now one of the key estimates from [24]. Indeed, let $\{E(P)\}_{P\in\mathcal{S},\,P\subseteq Q_j}$ be the family of sets from Definition 3.3. We have then for some $c>0$ that

$$\begin{aligned}\sum_{P\in\mathcal{S}}\chi_P(x) &= \sum_{Q\in\mathcal{S}}\left(\frac{1}{|P|}\,|P|\right)^q\chi_P(x)\\ &\le c\sum_{P\in\mathcal{S},\,P\subseteq Q_j}\left(\frac{1}{|P|}\,|E(P)|\right)^q\chi_P(x)\\ &\le c\sum_{P\in\mathcal{S},\,P\subseteq Q_j}\left(\frac{1}{|P|}\int_P\chi_{E(P)}(y)\,dy\right)^q\chi_P(x)\\ &\le c\left(\overline{M}_q\left(\left\{\chi_{E(P)}\right\}_{P\in\mathcal{S},\,P\subseteq Q_j}\right)(x)\right)^q\\ &\le c\left(\overline{M}_q g_j(x)\right)^q,\end{aligned}$$

where $g_j=\left\{\chi_{E(P)}\right\}_{P\in\mathcal{S},\,P\subseteq Q_j}$ is supported in Q_j. Now, since $\{E(Q)\}_{P\in\mathcal{S},\,P\subseteq Q_j}$ is a pairwise disjoint family of subsets, we have that for any j

$$\|g_j(x)\|_{\ell^q}=\left(\sum_{P\in\mathcal{S},\,P\subseteq Q_j}\left(\chi_{E(Q)}(x)\right)^q\right)^{1/q}\le 1. \tag{13}$$

We finish the proof of (12) recalling that if $|g_j|_{\ell^q}\in L^\infty$, then $\left(\overline{M}_q g_j(x)\right)^q\in ExpL$ (see [11]) from which we conclude that:

$$C_Q\le ce^{-ct},\qquad t>0.$$

Now we go back to the proof of the estimate. We first observe that

$$\begin{aligned}&\frac{1}{|Q|}\,|\{x\in Q: A_{\mathcal{S}}f(x)>tMf(x)\}|\\ &\le\sum_{j=1}^{c_n}\frac{1}{|Q|}\left|\left\{x\in Q: A_{\mathcal{S}}\left(f\chi_{Q_j}\right)(x)>\frac{t}{c_n}M(f\chi_{Q_j})(x)\right\}\right|.\end{aligned}$$

Hence it suffices to obtain an estimate for each term of the sum. First we may assume that $|Q_j\cap Q|\ne 0$ since otherwise, $\int_E f\chi_{Q_j}=0$ for every measurable set and the corresponding term in the sum equals zero. Now we split $A_{\mathcal{S}}(f\chi_{Q_j})$ as follows

$$A_{\mathcal{S}}(f\chi_{Q_j})(x)=\sum_{P\in\mathcal{S},\,P\subsetneq Q_j}\frac{1}{|P|}\int_P f\chi_{Q_j}\chi_P(x)+\sum_{P\in\mathcal{S},\,P\supseteq Q_j}\frac{1}{|P|}\int_P f\chi_{Q_j}\chi_P(x). \tag{14}$$

We observe that for the first term

$$\frac{\sum_{P\in\mathcal{S},\,P\subsetneq Q_j}\frac{1}{|P|}\int_P f\chi_{Q_j}\chi_P(x)}{M\left(f\chi_{Q_j}\right)(x)}\leq\sum_{P\in\mathcal{S},\,P\subseteq Q_j}\chi_P(x).$$

For the second term, we have that that since $Q_j\cap Q\neq\emptyset$ and $|Q|\simeq|Q_j|$ then $Q\subset 5Q_j$. Consequently,

$$\begin{aligned}\frac{\sum_{P\in\mathcal{S},\,P\supseteq Q_j}\frac{1}{|P|}\int_P f\chi_{Q_j}\chi_P(x)}{M(f\chi_{Q_j})(x)}&\leq\frac{\sum_{P\in\mathcal{S},\,P\supseteq Q_j}\frac{1}{|P|}\int_{Q_j}f\chi_P(x)}{\frac{1}{|5Q_j|}\int_{Q_j}f}\\&\leq c_n\sum_{P\in\mathcal{S},\,P\supseteq Q_j}\frac{|Q_j|}{|P|}\chi_P(x)\\&\leq c_n\sum_{k=0}^{\infty}\frac{1}{2^{nk}}.\end{aligned}$$

Then, combining the estimates obtained for each of the terms of (14), we have that

$$\begin{aligned}&\frac{1}{|Q|}\left|\left\{x\in Q:A_{\mathcal{S}}\left(f\chi_{Q_j}\right)(x)>\frac{t}{c_n}M(f\chi_{Q_j})(x)\right\}\right|\\\leq&\frac{1}{|Q|}\left|\left\{x\in Q:\sum_{P\in\mathcal{S},\,P\subseteq Q_j}\chi_P(x)>\frac{t}{c_n}-c_n\sum_{k=0}^{\infty}\frac{1}{2^{nk}}\right\}\right|\end{aligned}$$

and the desired conclusion, namely (11), follows from (12).

Observation 2 *Let T be a Calderón-Zygmund operator and $b\in BMO$. It is not possible to find a finite set of η-sparse families $\left\{\mathcal{S}_j\right\}_{j=1}^N$, with N dimensional, contained in the same or in different dyadic lattices $\mathcal{D}_j$ and depending on f such that*

$$|[b,T]f(x)|\leq c_{b,T}\sum_{j=1}^{N}B_{\mathcal{S}_j}f(x)\qquad a.e.\ x\in\mathbb{R}^n\tag{15}$$

where $B_{\mathcal{S}_j}f(x)=\sum_{Q\in\mathcal{S}_j}\|f\|_{L\log L,Q}\,\chi_Q(x)$.

We are going to give two proofs of this result. The first one is based on the Rubio de Francia algorithm.

Proof 1 Suppose that (15) holds, then we can prove the following L^1 inequality

$$\|[b,T]f\|_{L^1(w)}\leq c[w]_{A_1}\|M^2f\|_{L^1(w)}.\tag{16}$$

Indeed,

$$\begin{aligned}
\|[b,T]f\|_{L^1(w)} &\leq c_{b,T}\sum_{j=1}^{N}\|B_{\mathcal{S}_j}f\|_{L^1(w)}\\
&\leq c_{b,T}\sum_{j=1}^{N}\sum_{Q\in\mathcal{S}_j}\|f\|_{L\log L,Q}\frac{w(Q)}{|Q|}|Q|\\
&\leq \frac{c_{b,T}}{\eta}\sum_{j=1}^{N}\sum_{Q\in\mathcal{S}_j}\|f\|_{L\log L,Q}\frac{w(Q)}{|Q|}|E(Q)|\\
&\leq \frac{c_{b,T}}{\eta}\sum_{j=1}^{N}\sum_{Q\in\mathcal{S}_j}\int_{E(Q)}M_{L\log L}f(x)Mw(x)dx\\
&\leq N\frac{c_{b,T}}{\eta}[w]_{A_1}\|M^2f\|_{L^1(w)},
\end{aligned}$$

since $M^2 \approx M_{L\log L}$. We claim now the L^p version,

$$\|[b,T]f\|_{L^p(\mathbb{R}^n)} \leq c_n p\|M^2f\|_{L^p(\mathbb{R}^n)} \qquad p>1. \tag{17}$$

Indeed, by duality we can find $g\geq 0$ in $L^{p'}(\mathbb{R}^n)$ with unit norm such that

$$\|[b,T]f\|_{L^p(\mathbb{R}^n)} = \int_{\mathbb{R}^n}|[b,T]f(x)|g(x)dx.$$

We consider the Rubio de Francia algorithm

$$Rg = \sum_{k=0}^{\infty}\frac{M^k(g)}{\|M\|^k_{L^{p'}(\mathbb{R}^n)}}.$$

It's a straightforward computation that $R(g)$ is an A_1 weight with constant

$$[Rg]_{A_1} \leq 2\|M\|_{L^{p'}} \leq c_n p$$

and also that $g\leq Rg$ and $\|Rg\|_{L^{p'}} \leq 2\|g\|_{L^{p'}(\mathbb{R}^n)} = 2$. Then have that

$$\int_{\mathbb{R}^n}|[b,T]f(x)|g(x)dx \leq \int_{\mathbb{R}^n}|[b,T]f(x)|Rg(x)dx$$

and using (16) and Hölder inequality

$$\begin{aligned}
&\int_{\mathbb{R}^n}|[b,T]f(x)|Rg(x)dx \leq c\,[Rg]_{A_1}\int_{\mathbb{R}^n}M^2f(x)Rg(x)dx\\
&\leq cp\int_{\mathbb{R}^n}M^2f(x)Rg(x)dx \leq cp\|M^2f\|_{L^p(\mathbb{R}^n)}\|Rg\|_{L^{p'}(\mathbb{R}^n)}\\
&\leq cp\|M^2f\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

Hence (17) is established. Now since

$$\|M^2\|_{L^p(\mathbb{R}^n)} \leq c_n \left(p'\right)^2 \qquad p > 1$$

we have that

$$\|[b,T]\|_{L^p(\mathbb{R}^n)} \leq cp\left(p'\right)^2 \qquad p > 1 \tag{18}$$

Now let us observe that if we take $[b,H]f$ with $b(x) = \log|x|$ and $f(x) = \chi_{(0,1)}(x)$ then

$$\|[b,H]f\|_{L^p(\mathbb{R})} \geq cp^2 \qquad p > 1,$$

and this leads to a contradiction when $p \to \infty$. To prove this lower estimate we use estimate (9) from Theorem 1. Indeed, for some $t_0 > 0$

$$\begin{aligned}
\|[b,H]f\|_{L^p(\mathbb{R})} &\geq \|\,[b,H]f\|_{L^{p,\infty}(\mathbb{R})} = \sup_{t>0} t|\{x \in \mathbb{R} : |[b,H]f(x)| > t\}|^{\frac{1}{p}} \\
&\geq \sup_{t>t_0} t\left|\left\{x \in (0,x_0) : c\left(\log\frac{1}{x}\right)^2 > t\right\}\right|^{\frac{1}{p}} \\
&\geq \sup_{t>t_0} tce^{\frac{-\sqrt{t}}{p}} \geq c\,p^2\,t_0 e^{-\sqrt{t_0}}
\end{aligned}$$

and this concludes the first proof. □

For the second proof we will rely on the sharpness result that was settled in the previous section.

Proof 2 Assume again that (15) holds. Then, for some $c > 1$

$$\left|\left\{x \in Q : |[b,T]f(x)| > tM^2f(x)\right\}\right| \leq \left|\left\{x \in Q : \sum_{j=1}^{N}\sum_{P\in\mathcal{S}_j}\|f\|_{L\log L,P}\chi_P(x) > \frac{t}{c}M^2f(x)\right\}.\right|$$

It will be enough for our purposes to work on each term of the inner sum, namely to control

$$\left|\left\{x \in Q : \sum_{P\in\mathcal{S}_j}\|f\|_{L\log L,P}\chi_P(x) > tM^2f(x)\right\}.\right|$$

Now, recalling that $M^2f \simeq M_{L\log L}f$, is not hard to see that essentially the same argument we used to prove (11) yields that

$$\frac{1}{|Q|}\left|\left\{x \in Q : \sum_{P\in\mathcal{S}_j}\|f\|_{L\log L,P}\chi_P(x) > tM^2f(x)\right\}\right| \leq ce^{-\alpha t}.$$

Combining the preceding estimates we arrive to

$$\frac{1}{|Q|}\left|\left\{x \in Q : |[b,T]f(x)| > tM^2 f(x)\right\}\right| \le ce^{-\alpha t} \qquad t > 0$$

which is a contradiction by Observation 1. □

The correct pointwise control for the commutator seems to be the following one

Conjecture 1 *Let T be a Calderón-Zygmund operator and $b \in BMO$. Then*

$$|[b,T]f(x)| \le C(n,T)\|b\|_{BMO} \sum_{i,j=1}^{N} A_{\mathcal{S}_i}\left(A_{\mathcal{S}_j} f\right)(x)$$

where $A_{\mathcal{S}_j} f(x) = \sum_{Q \in \mathcal{S}_j} \frac{1}{|Q|}\int_Q |f(y)|\,dy \chi_Q(x)$ and the sparse families $\mathcal{S}_j$ are not necessarily subfamilies of the same dyadic lattice.

If this conjecture holds it would be very easy to recover the main theorem from [4] since it suffices to iterate the following estimate:

$$\left\|A_{\mathcal{S}_j} f\right\|_{L^p(w)} \le C_{n,p}[w]_{A_p}^{\max\left\{1,\frac{1}{p-1}\right\}}\|f\|_{L^p}.$$

which was studied in [9, 10] (see also [21]).

Third Observation: The Failure of an Endpoint Estimate Motivated by the Conjugation Method

In this section we consider the following family of operators:

$$f \to T_w(f) := w\,T\left(\frac{f}{w}\right) \tag{19}$$

where w is a weight and T is a Calderón-Zygmund operator. We already mentioned in the introduction that these operators are of interest since they are very much related to commutators due to the conjugation method. We emphasized that the case $w \in A_1$ is special since T_w is of weak type $(1,1)$ as a consequence of the main results from [8]. Understanding the case $w \in A_p$ would be more interesting due to its connection with the conjugation method. However, T_w is **not** of weak type $(1,1)$ in general since there are weights $w \in A_p \setminus A_1$ for which

$$\|T_w\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)} = \infty,$$

being the purpose of this section to show the existence of such weights. In fact we are going to show something worst replacing T by the less singular operator M.

Observation 3 *Let* $1 < p < \infty$*, then there is* $w \in A_p \setminus A_1$ *such that*

$$\|M_w\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)} = \infty.$$

Proof In dimension 1 we choose the A_p weight $w(x) = |x|^{-\delta(1-p)}$ with $\delta \in \left(0, \min\left\{1, \frac{1}{p-1}\right\}\right)$ and $f = \chi_{[0,1]}$ so that $f \in L^1(w)$. We prove that

$$\left\| wM\left(\frac{f}{w}\right)\right\|_{L^{1,\infty}(\mathbb{R})} = \infty.$$

Indeed, a computation shows that for $x > 1$

$$M\left(\frac{\chi_{(0,1)}}{w}\right)(x) \geq \frac{1}{x}\frac{1}{\beta}$$

with $\beta = 1 + \delta(1-p)$ and then

$$\left\| wM\left(\frac{f}{w}\right)\right\|_{L^{1,\infty}(\mathbb{R})} \geq \frac{1}{\beta}\sup_{t>0} t\left|\left\{x > 1 : x^{-\delta(1-p)-1} > t\right\}\right| = \frac{1}{\beta}\sup_{1>t>0} t\left(\left(\frac{1}{t}\right)^{\frac{1}{\beta}} - 1\right) = \infty$$

since $\beta \in (0,1)$. □

An interesting question is to find a necessary and sufficient condition for the boundedness of this operator, namely, characterize the weights w for which

$$\|M_w\|_{L^1(\mathbb{R}^n)\to L^{1,\infty}(\mathbb{R}^n)} < \infty.$$

In [22] Muckenhoupt and Wheeden proved that this inequality holds for $w \in A_1$ in the real line and also obtained a necessary condition on the weights, namely

$$\left\| \frac{w\chi_Q}{|\cdot - x|^n}\right\|_{L^{1,\infty}(\mathbb{R}^n)} \leq cw(x) \qquad \text{a.e. } x \in \mathbb{R}^n$$

but we don't know whether is sufficient or not.

To end this section we show that can go further and prove a negative result for possible $L\log L$ type estimates.

Observation 4 *Let* $1 < p < \infty$*, and let* $\Phi(t) = t\log(e+t)^\alpha$*,* $\alpha > 0$*. Then we can find* $w \in A_p \setminus A_1$ *and* f *such that there's no* $c > 0$ *for which*

$$\left|\left\{x \in \mathbb{R}^n : wM\left(\frac{f}{w}\right) > t\right\}\right| \leq c\int_{\mathbb{R}^n}\Phi\left(\frac{|f(x)|}{t}\right)dx. \tag{20}$$

Proof As above we do it for the case $n = 1$. We assume the contrary, namely there is a finite constant $c > 0$ such that (20) holds for any nonnegative f. Let $f = \chi_{(0,1)}$. For this choice of f the right hand side of (20) equals $\Phi\left(\frac{1}{t}\right)$ and we have that

$$\sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} \left|\left\{x \in \mathbb{R} \,:\, wM\left(\frac{\chi_{(0,1)}}{w}\right) > t\right\}\right| < \infty.$$

Choose again the A_p weight $w(x) = |x|^{-\delta(1-p)}$ with $\delta \in \left(0, \min\left\{1, \frac{1}{p-1}\right\}\right)$. Proceeding and using the same notation as in the proof of Observation 3 we have that

$$\sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} \left|\left\{x \in \mathbb{R} \,:\, wM\left(\frac{\chi_{(0,1)}}{w}\right) > t\right\}\right| \geq c \sup_{0<t<1} \frac{1}{\Phi\left(\frac{1}{t}\right)} \left[\left(\frac{1}{t}\right)^{\frac{1}{\beta}} - 1\right]$$

$$= c \sup_{0<t<1} \frac{t}{\log\left(e + \frac{1}{t}\right)^{\alpha}} \left[\left(\frac{1}{t}\right)^{\frac{1}{\beta}} - 1\right] = \infty.$$

since $\beta \in (0, 1)$. □

References

1. J. Álvarez, R.J. Bagby, D.S. Kurtz, C. Pérez, Weighted estimates for commutators of linear operators. Studia Math. **104**(2), 195–209 (1993)
2. S. Bloom, A commutator theorem and weighted BMO. Trans. Am. Math. Soc. **292**(1), 103–122 (1985)
3. S.M. Buckley, Estimates for operator norms on weighted spaces and reverse Jensen inequalities. Trans. Am. Math. Soc. **340**(1), 253–272 (1993)
4. D. Chung, M.C. Pereyra, C. Perez, Sharp bounds for general commutators on weighted Lebesgue spaces. Trans. Am. Math. Soc. **364**(3), 1163–1177 (2012)
5. R.R. Coifman, Y. Meyer, On commutators of singular integrals and bilinear singular integrals. Trans. Am. Math. Soc. **212**, 315–331 (1975)
6. R.R. Coifman, R. Rochberg, G. Weiss, Factorization theorems for Hardy spaces in several variables. Ann. of Math. (2) **103**(3), 611–635 (1976)
7. J.M. Conde-Alonso, G. Rey, A pointwise estimate for positive dyadic shifts and some applications. Math. Ann. **365**(3-4), 1111–1135 (2016)
8. D. Cruz-Uribe, J.M. Martell, C. Pérez, Weighted weak-type inequalities and a conjecture of Sawyer. Int. Math. Res. Not. **30**, 1849–1871 (2005)
9. D. Cruz-Uribe, J.M. Martell, C. Pérez, Sharp weighted estimates for approximating dyadic operators. Electron. Res. Announc. Math. Sci. **17**, 12–19 (2010)
10. D. Cruz-Uribe, J.M. Martell, C. Pérez, Sharp weighted estimates for classical operators. Adv. Math. **229**(1), 408–441 (2012)
11. C. Fefferman, E.M. Stein, Some maximal inequalities. Am. J. Math. **93**, 107–115 (1971)
12. J. García-Cuerva, E. Harboure, C. Segovia, J.L. Torrea, Weighted norm inequalities for commutators of strongly singular integrals. Indiana Univ. Math. J. **40**(4), 1397–1420 (1991)
13. I. Holmes, M.T. Lacey, B.D. Wick, Commutators in the two-weight setting. Math. Ann. **367**(1–2), 51–80 (2017)

14. I. Holmes, B.D. Wick, Two weight inequalities for iterated commutators with Calderón-Zygmund operators (Sept 2015). ArXiv e-prints
15. T. Hytönen, C. Pérez, Sharp weighted bounds involving A_∞. Anal. PDE **6**(4), 777–818 (2013)
16. T.P. Hytönen, L. Roncal, O. Tapiola, Quantitative weighted estimates for rough homogeneous singular integrals. Isr. J. Math. **218**, 133 (2017). doi:10.1007/s11856-017-1462-6
17. S. Janson, Mean oscillation and commutators of singular integral operators. Ark. Mat. **16**(2), 263–270 (1978)
18. G.A. Karagulyan, Exponential estimates for the Calderón-Zygmund operator and related problems of Fourier series. Mat. Zametki **71**(3), 398–411 (2002)
19. M.T. Lacey, An elementary proof of the A_2 bound. Isr. J. Math. **217**, 181 (2017). doi:10.1007/s11856-017-1442-x
20. A.K. Lerner, On pointwise estimates involving sparse operators. New York J. Math. **22**, 341–349 (2016)
21. A.K. Lerner, F. Nazarov, Intuitive dyadic calculus: the basics (Aug 2015). ArXiv e-prints
22. B. Muckenhoupt, R.L. Wheeden, Some weighted weak-type inequalities for the Hardy-Littlewood maximal function and the Hilbert transform. Indiana Univ. Math. J. **26**(5), 801–816 (1977)
23. C. Ortiz-Caraballo, Quadratic A_1 bounds for commutators of singular integrals with BMO functions. Indiana Univ. Math. J. **60**(6), 2107–2129 (2011)
24. C. Ortiz-Caraballo, C. Pérez, E. Rela, Exponential decay estimates for singular integral operators. Math. Ann. **357**(4), 1217–1243 (2013)
25. C. Pérez, Endpoint estimates for commutators of singular integral operators. J. Funct. Anal. **128**(1), 163–185 (1995)
26. C. Pérez, G. Pradolini, Sharp weighted endpoint estimates for commutators of singular integrals. Mich. Math. J. **49**(1), 23–37 (2001)
27. C. Pérez, I.P. Rivera-Ríos, Borderline weighted estimates for commutators of singular integrals. Isr. J. Math. **217**, 435 (2017). doi:10.1007/s11856-017-1454-6
28. A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*. Volume 123 of Pure and Applied Mathematics (Academic, Orlando, 1986)

A Two Weight Fractional Singular Integral Theorem with Side Conditions, Energy and k-Energy Dispersed

Eric T. Sawyer, Chun-Yen Shen, and Ignacio Uriarte-Tuero

This paper is dedicated to the memory of Cora Sadosky and her contributions to the theory of weighted inequalities and her promotion of women in mathematics.

Abstract This paper is a sequel to our paper Sawyer et al. (Revista Mat Iberoam 32(1):79–174, 2016). Let σ and ω be locally finite positive Borel measures on $\mathbb{R}^n$ (possibly having common point masses), and let T^α be a standard α-fractional Calderón-Zygmund operator on $\mathbb{R}^n$ with $0 \leq \alpha < n$. Suppose that $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ is a globally biLipschitz map, and refer to the images ΩQ of cubes Q as *quasicubes*. Furthermore, assume as side conditions the $\mathcal{A}_2^\alpha$ conditions, punctured A_2^α conditions, and certain α*-energy conditions* taken over quasicubes. Then we show that T^α is bounded from $L^2(\sigma)$ to $L^2(\omega)$ if the quasicube testing conditions hold for T^α and its dual, and if the quasiweak boundedness property holds for T^α.

Conversely, if T^α is bounded from $L^2(\sigma)$ to $L^2(\omega)$, then the quasitesting conditions hold, and the quasiweak boundedness condition holds. If the vector of α-fractional Riesz transforms $\mathbf{R}_\sigma^\alpha$ (or more generally a strongly elliptic vector of transforms) is bounded from $L^2(\sigma)$ to $L^2(\omega)$, then both the $\mathcal{A}_2^\alpha$ conditions and the punctured A_2^α conditions hold.

Research supported in part by NSERC.

C.-Y. Shen supported in part by the NSC, through grant NSC 104-2628-M-008 -003 -MY4.

I. Uriarte-Tuero has been partially supported by grants DMS-1056965 (US NSF), MTM2010-16232, MTM2015-65792-P (MINECO, Spain), and a Sloan Foundation Fellowship.

E.T. Sawyer (✉)
Department of Mathematics & Statistics, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
e-mail: sawyer@mcmaster.ca

C.-Y. Shen
Department of Mathematics, National Central University, Chungli 32054, Taiwan
e-mail: cyshen@math.ncu.edu.tw

I. Uriarte-Tuero
Department of Mathematics, Michigan State University, East Lansing, MI, USA
e-mail: ignacio@math.msu.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_13

Our quasienergy conditions are not in general necessary for elliptic operators, but are known to hold for certain situations in which one of the measures is one-dimensional (Lacey et al., Two weight inequalities for the Cauchy transform from $\mathbb{R}$ to $\mathbb{C}_+$, arXiv:1310.4820v4; Sawyer et al., The two weight $T1$ theorem for fractional Riesz transforms when one measure is supported on a curve, arXiv:1505.07822v4), and for certain side conditions placed on the measures such as doubling and k-energy dispersed, which when $k = n-1$ is similar to the condition of uniformly full dimension in Lacey and Wick (Two weight inequalities for the Cauchy transform from $\mathbb{R}$ to $\mathbb{C}_+$, arXiv:1310.4820v1, versions 2 and 3).

Introduction

The boundedness of the Hilbert transform $Hf(x) = \int_{\mathbb{R}} \frac{f(y)}{y-x} dy$ on the real line $\mathbb{R}$ in the Hilbert space $L^2(\mathbb{R})$ has been known for at least a century (perhaps dating back to A & E[1]):

$$\|Hf\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}, \quad f \in L^2(\mathbb{R}). \tag{1}$$

This inequality has been the subject of much generalization, to which we now turn.

A Brief History of the $T1$ Theorem

The celebrated $T1$ theorem of David and Journé [3] extends (1) to more general kernels by characterizing those singular integral operators T on $\mathbb{R}^n$ that are bounded on $L^2(\mathbb{R}^n)$, and does so in terms of a weak boundedness property, and the membership of the two functions $T\mathbf{1}$ and $T^*\mathbf{1}$ in the space of bounded mean oscillation,

$$\|T\mathbf{1}\|_{BMO(\mathbb{R}^n)} \lesssim \|\mathbf{1}\|_{L^\infty(\mathbb{R}^n)} = 1,$$
$$\|T^*\mathbf{1}\|_{BMO(\mathbb{R}^n)} \lesssim \|\mathbf{1}\|_{L^\infty(\mathbb{R}^n)} = 1.$$

These latter conditions are actually the following *testing conditions* in disguise,

$$\|T\mathbf{1}_Q\|_{L^2(\mathbb{R}^n)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n)} = \sqrt{|Q|},$$
$$\|T^*\mathbf{1}_Q\|_{L^2(\mathbb{R}^n)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n)} = \sqrt{|Q|},$$

[1] Peter Jones used A&E to stand for Adam and Eve.

tested uniformly over all indicators of cubes Q in $\mathbb{R}^n$ for both T and its dual operator T^*. This theorem was the culmination of decades of investigation into the nature of cancellation conditions required for boundedness of singular integrals.[2]

A parallel thread of investigation had begun even earlier with the equally celebrated theorem of Hunt, Muckenhoupt and Wheeden [4] that extended (1) to measures more general than Lebesgue's by characterizing boundedness of the Hilbert transform on weighted spaces $L^2(\mathbb{R}; w)$. This thread culminated in the theorem of Coifman and Fefferman[3] [2] that characterizes those nonnegative weights w on $\mathbb{R}^n$ for which all of the 'nicest' of the $L^2(\mathbb{R}^n)$ bounded singular integrals T above are bounded on weighted spaces $L^2(\mathbb{R}^n; w)$, and does so in terms of the A_2 condition of Muckenhoupt,

$$\left(\frac{1}{|Q|}\int_Q w(x)\,dx\right)\left(\frac{1}{|Q|}\int_Q \frac{1}{w(x)}dx\right) \lesssim 1,$$

taken uniformly over all cubes Q in $\mathbb{R}^n$. This condition is also a testing condition in disguise, in particular it follows from

$$\left\| T\left(\mathbf{s}_Q \frac{1}{w}\right)\right\|_{L^2(\mathbb{R}^n;w)} \lesssim \left\| \mathbf{s}_Q \frac{1}{w}\right\|_{L^2(\mathbb{R}^n;w)},$$

tested over all 'indicators with tails' $\mathbf{s}_Q(x) = \frac{\ell(Q)}{\ell(Q)+|x-c_Q|}$ of cubes Q in $\mathbb{R}^n$.

A natural synthesis of these two threads leads to the 'two weight' question of characterizing those pairs of weights (σ, ω) having the property that nice singular integrals are bounded from $L^2(\mathbb{R}^n; \sigma)$ to $L^2(\mathbb{R}^n; \omega)$. Returning to the simplest (nontrivial) singular integral of all, namely the Hilbert transform $Hf(x) = \int_{\mathbb{R}} \frac{f(y)}{y-x}dy$ on the real line, Cotlar and Sadosky gave a beautiful function theoretic characterization of the weight pairs (σ, ω) for which H is bounded from $L^2(\mathbb{R}; \sigma)$ to $L^2(\mathbb{R}; \omega)$, namely a two-weight extension of the Helson-Szegö theorem. This characterization illuminated a deep connection between two quite different function theoretic conditions, but failed to shed much light on when either of them held. On the other hand, the two weight inequality for positive fractional integrals, Poisson integrals and maximal functions were characterized using testing conditions by one of us in [24] (see also [6]) and [23], but relying in a very strong way on the positivity of the kernel, something the Hilbert kernel lacks. In light of these considerations, Nazarov, Treil and Volberg formulated the two weight question for the Hilbert transform [35], that in turn led to the following NTV conjecture:

[2]See e.g. chapter VII of Stein [34] and the references given there for a historical background.

[3]See e.g. chapter V of [34] and the references given there for the long history of this investigation.

Conjecture 1 [35] *The Hilbert transform is bounded from* $L^2(\mathbb{R}^n;\sigma)$ *to* $L^2(\mathbb{R}^n;\omega)$*, i.e.*

$$\|H(f\sigma)\|_{L^2(\mathbb{R}^n;\omega)} \lesssim \|f\|_{L^2(\mathbb{R}^n;\sigma)}, \quad f \in L^2(\mathbb{R}^n;\sigma), \tag{2}$$

if and only if the two weight A_2 *condition with two tails holds,*

$$\left(\frac{1}{|Q|}\int_Q \mathbf{s}_Q^2 d\omega(x)\right)\left(\frac{1}{|Q|}\int_Q \mathbf{s}_Q^2 d\sigma(x)\right) \lesssim 1,$$

uniformly over all cubes Q*, and the two testing conditions hold,*

$$\|H\mathbf{1}_Q\sigma\|_{L^2(\mathbb{R}^n;\omega)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n;\sigma)} = \sqrt{|Q|_\sigma},$$
$$\|H^*\mathbf{1}_Q\omega\|_{L^2(\mathbb{R}^n;\sigma)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n;\omega)} = \sqrt{|Q|_\omega},$$

uniformly over all cubes Q*.*

In a groundbreaking series of papers including [19, 20] and [21], Nazarov, Treil and Volberg used weighted Haar decompositions with random grids, introduced their 'pivotal' condition, and proved the above conjecture under the side assumption that the pivotal condition held. Subsequently, in joint work of two of us, Sawyer and Uriarte-Tuero, with Lacey [16], it was shown that the pivotal condition was not necessary in general, a necessary 'energy' condition was introduced as a substitute, and a hybrid merging of these two conditions was shown to be sufficient for use as a side condition. Eventually, these three authors with Shen established the NTV conjecture in a two part paper; Lacey, Sawyer, Shen and Uriarte-Tuero [13] and Lacey [8]. A key ingredient in the proof was an 'energy reversal' phenomenon enabled by the Hilbert transform kernel equality

$$\frac{1}{y-x} - \frac{1}{y-x'} = \frac{x-x'}{(y-x)(y-x')},$$

having the remarkable property that the denominator on the right hand side remains *positive* for all y outside the smallest interval containing both x and x'. This proof of the NTV conjecture was given in the special case that the weights σ and ω had no point masses in common, largely to avoid what were then thought to be technical issues. However, these issues turned out to be considerably more interesting, and this final assumption of no common point masses was removed shortly after by Hytönen [6], who also simplified some aspects of the proof.

At this juncture, attention naturally turned to the analogous two weight inequalities for higher dimensional singular integrals, as well as α-fractional singular integrals such as the Cauchy transform in the plane. In a long paper begun in [28] on the *arXiv* in 2013, and subsequently appearing in [30], the authors introduced the appropriate notions of Poisson kernel to deal with the A_2^α condition on the one hand,

and the α-energy condition on the other hand (unlike for the Hilbert transform, these two Poisson kernels differ in general). The main result of that paper established the $T1$ theorem for 'elliptic' vectors of singular integrals under the side assumption that an energy condition and its dual held, thus identifying the *culprit* in higher dimensions as the energy conditions. A general $T1$ conjecture is this (see below for definitions).

Conjecture 2 *Let $\mathbf{T}^{\alpha,n}$ denote an elliptic vector of standard α-fractional singular integrals in $\mathbb{R}^n$. Then $\mathbf{T}^{\alpha,n}$ is bounded from $L^2(\mathbb{R}^n;\sigma)$ to $L^2(\mathbb{R}^n;\omega)$, i.e.*

$$\left\|\mathbf{T}^{\alpha,n}(f\sigma)\right\|_{L^2(\mathbb{R}^n;\omega)} \lesssim \|f\|_{L^2(\mathbb{R}^n;\sigma)}, \quad f\in L^2(\mathbb{R}^n;\sigma), \tag{3}$$

if and only if the two one-tailed $\mathcal{A}_2^\alpha$ conditions with holes hold, the punctured $A_2^{\alpha,\mathrm{punct}}$ conditions hold, and the two testing conditions hold,

$$\left\|\mathbf{T}^{\alpha,n}\mathbf{1}_Q\sigma\right\|_{L^2(\mathbb{R}^n;\omega)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n;\sigma)} = \sqrt{|Q|_\sigma},$$
$$\left\|\mathbf{T}^{\alpha,n,\mathrm{dual}}\mathbf{1}_Q\omega\right\|_{L^2(\mathbb{R}^n;\sigma)} \lesssim \|\mathbf{1}_Q\|_{L^2(\mathbb{R}^n;\omega)} = \sqrt{|Q|_\omega},$$

for all cubes Q in $\mathbb{R}^n$ (whose sides need not be parallel to the coordinate axes).

In [32], the authors have recently shown that the energy conditions are *not* necessary for boundedness of elliptic vectors of singular integrals in general, but have left open the following conjecture, which in view of the aforementioned main result in [30], would yield the $T1$ theorem for gradient elliptic operators. An elliptic α-fractional singular integral vector $\mathbf{T}^{\alpha,n}$ in $\mathbb{R}^n$ is said to be *gradient elliptic* if both $\left|\nabla_x\mathbf{K}^\alpha(x,y)\right| \gtrsim |x-y|^{\alpha-n-1}$ and $\left|\nabla_y\mathbf{K}^\alpha(x,y)\right| \gtrsim |x-y|^{\alpha-n-1}$.

Conjecture 3 *Let $\mathbf{T}^{\alpha,n}$ denote a gradient elliptic vector of standard α-fractional singular integrals in $\mathbb{R}^n$. If $\mathbf{T}^{\alpha,n}$ is bounded from $L^2(\mathbb{R}^n;\sigma)$ to $L^2(\mathbb{R}^n;\omega)$, then the energy conditions hold as defined in Definition 2.6 below.*

While the energy conditions are not necessary for elliptic operators in general [32], there are some cases in which they have been proved to hold. Of course, they hold for the Hilbert transform on the line [16], and in recent joint work with M. Lacey and B. Wick, the five of us have established that the energy conditions hold for the Cauchy transform in the plane in the special case where one of the measures is supported on either a straight line or a circle, thus proving the $T1$ theorem in this case. The key to this result was an extension of the energy reversal phenomenon for the Hilbert transform to the setting of the Cauchy transform, and here the one-dimensional nature of the line and circle played a critical role. In particular, a special decomposition of a 2-dimensional measure into 'end' and 'side' pieces played a crucial role, and was in fact discovered independently in both [26] and [17]. A further instance of energy reversal occurs in our $T1$ theorem [31] when one measure is compactly supported on a $C^{1,\delta}$ curve in $\mathbb{R}^n$.

The paper [18, v3] by Lacey and Wick overlaps both our paper [30] and this paper to some extent, and we refer the reader to [30] for a more detailed discussion.

Finally, we mention an entirely different approach to investigating the two weight problem that has attracted even more attention than the $T1$ approach we just described. Nazarov has shown that the two-tailed $\mathcal{A}_2^\alpha$ condition of Muckenhoupt (see below) is insufficient for (3), and this begs the question of strengthening the Muckenhoupt condition enough to make it sufficient for (3). The great advantage of this approach is that strengthened Muckenhoupt conditions are generally 'easy' to check as compared to the highly unstable testing conditions. The disadvantage of course is that such conditions have never been shown to characterize (3). The literature devoted to these issues, beginning with that of Pérez [22], and continuing more recently with work of many groups involving, among others, D. Cruz-Uribe, M. Lacey, A. K. Lerner, J. M. Martell, F. Nazarov, C. Pérez, A. Reznikov and A. Volberg, is both too vast and too tangential to this paper to record here, and we encourage the reader to search the web for more on 'bumped-up' Muckenhoupt conditions.[4]

This paper is concerned with the $T1$ approach and is a sequel to our first paper [30]. We prove here a two weight inequality for standard α-fractional Calderón-Zygmund operators T^α in Euclidean space $\mathbb{R}^n$, where we assume n-dimensional $\mathcal{A}_2^\alpha$ conditions (with holes), punctured $A_2^{\alpha.\text{punct}}$ conditions, and certain *α-energy conditions* as side conditions on the weights (in higher dimensions the Poisson kernels used in these two conditions differ). The two main differences in this theorem here are that we state and prove[5] our theorem in the more general setting of *quasicubes* (as in [28]), and more notably, we now permit the weights, or measures, to have common point masses, something not permitted in [30] (and only obtained for a partial range of α in [18, version 3]). As a consequence, we use $\mathcal{A}_2^\alpha$ conditions with holes as in the one-dimensional setting of Hytönen [6], together with punctured $A_2^{\alpha,\text{punct}}$ conditions, as the usual A_2^α '*without punctures*' fails whenever the measures have a common point mass. The extension to permitting common point masses uses the two weight Poisson inequality in [24] to derive functional energy, together with a delicate adaptation of arguments in [28]. The key point here is the use of the (typically necessary) 'punctured' Muckenhoupt $A_2^{\alpha,\text{punct}}$ conditions below. They turn out to be crucial in estimating the backward Poisson testing condition later in the paper. We remark that Hytönen's bilinear dyadic Poisson operator and shifted dyadic grids [6] in dimension $n = 1$ can be extended to derive functional energy in higher dimensions, but at a significant cost of increased complexity. See the previous version of this paper on the *arXiv* for this approach,[6] and also [18] where Lacey and Wick use this approach. Finally, we point out that our use of punctured Muckenhoupt conditions provides a simpler alternative to Hytönen's method of extending to common point masses the NTV

[4] Starting e.g. with the recent articles [1] and [10].

[5] Very detailed proofs of all of the results here can be found on the arXiv [29].

[6] Additional small arguments are needed to complete the shifted dyadic proof given there, but we omit them in favour of the simpler approach here resting on punctured Muckenhoupt conditions instead of holes. The authors can be contacted regarding completion of the shifted dyadic proof.

conjecture for the Hilbert transform [6]. The Muckenhoupt $\mathcal{A}_2^\alpha$ conditions (with holes) are also typically necessary for the norm inequality, but the proofs require extensive modification when quasicubes and common point masses are included.

On the other hand, the extension to quasicubes in the setting of *no* common point masses turns out to be, after checking all the details, mostly a cosmetic modification of the proof in [30], as demonstrated in [28]. The use of quasicubes is however crucial in our $T1$ theorem when one of the measures is compactly supported on a $C^{1,\delta}$ curve [31], and this accounts for their inclusion here.

We also introduce a new side condition on a measure, that we call *k-energy dispersed*, which captures the notion that a measure is *not* supported too near a k-dimensional plane at any scale. When $0 \leq \alpha < n$ is appropriately related to k, we are able to obtain the necessity of the energy conditions for k-energy dispersed measures, and hence a $T1$ theorem for strongly elliptic operators $\mathbf{T}^\alpha$. The case $k = n - 1$ is similar to the condition of uniformly full dimension introduced in [18, versions 2 and 3].

We begin by recalling the notion of quasicube used in [28] - a special case of the classical notion used in quasiconformal theory.

Definition 1.1 We say that a homeomorphism $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ is a globally biLipschitz map if

$$\|\Omega\|_{Lip} \equiv \sup_{x,y\in\mathbb{R}^n} \frac{\|\Omega(x) - \Omega(y)\|}{\|x-y\|} < \infty, \tag{4}$$

and $\left\|\Omega^{-1}\right\|_{Lip} < \infty$.

Note that a globally biLipschitz map Ω is differentiable almost everywhere, and that there are constants $c, C > 0$ such that

$$c \leq J_\Omega(x) \equiv |\det D\Omega(x)| \leq C, \qquad x \in \mathbb{R}^n.$$

Example 1.2 Quasicubes can be wildly shaped, as illustrated by the standard example of a logarithmic spiral in the plane $f_\varepsilon(z) = z|z|^{2\varepsilon i} = ze^{i\varepsilon \ln(z\bar{z})}$. Indeed, $f_\varepsilon : \mathbb{C} \to \mathbb{C}$ is a globally biLipschitz map with Lipschitz constant $1 + C\varepsilon$ since $f_\varepsilon^{-1}(w) = w|w|^{-2\varepsilon i}$ and

$$\nabla f_\varepsilon = \left(\frac{\partial f_\varepsilon}{\partial z}, \frac{\partial f_\varepsilon}{\partial \bar{z}}\right) = \left(|z|^{2\varepsilon i} + i\varepsilon |z|^{2\varepsilon i}, i\varepsilon \frac{z}{\bar{z}} |z|^{2\varepsilon i}\right).$$

On the other hand, f_ε behaves wildly at the origin since the image of the closed unit interval on the real line under f_ε is an infinite logarithmic spiral.

Notation 1 *We define $\mathcal{P}^n$ to be the collection of half open, half closed cubes in $\mathbb{R}^n$ with sides parallel to the coordinate axes. A half open, half closed cube Q*

in $\mathbb{R}^n$ *has the form* $Q = Q(c,\ell) \equiv \prod_{k=1}^{n}\left[c_k - \frac{\ell}{2}, c_k + \frac{\ell}{2}\right)$ *for some* $\ell > 0$ *and* $c = (c_1,\ldots,c_n) \in \mathbb{R}^n$. *The cube* $Q(c,\ell)$ *is described as having center* c *and sidelength* ℓ.

We repeat the natural *quasi* definitions from [28].

Definition 1.3 Suppose that $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ is a globally biLipschitz map.

1. If E is a measurable subset of $\mathbb{R}^n$, we define $\Omega E \equiv \{\Omega(x) : x \in E\}$ to be the image of E under the homeomorphism Ω.
   (a) In the special case that $E = Q$ is a cube in $\mathbb{R}^n$, we will refer to ΩQ as a quasicube (or Ω-quasicube if Ω is not clear from the context).
   (b) We define the center $c_{\Omega Q} = c(\Omega Q)$ of the quasicube ΩQ to be the point Ωc_Q where $c_Q = c(Q)$ is the center of Q.
   (c) We define the side length $\ell(\Omega Q)$ of the quasicube ΩQ to be the sidelength $\ell(Q)$ of the cube Q.
   (d) For $r > 0$ we define the 'dilation' $r\Omega Q$ of a quasicube ΩQ to be ΩrQ where rQ is the usual 'dilation' of a cube in $\mathbb{R}^n$ that is concentric with Q and having side length $r\ell(Q)$.
2. If $\mathcal{K}$ is a collection of cubes in $\mathbb{R}^n$, we define $\Omega\mathcal{K} \equiv \{\Omega Q : Q \in \mathcal{K}\}$ to be the collection of quasicubes ΩQ as Q ranges over $\mathcal{K}$.
3. If $\mathcal{F}$ is a grid of cubes in $\mathbb{R}^n$, we define the inherited quasigrid structure on $\Omega\mathcal{F}$ by declaring that ΩQ is a child of $\Omega Q'$ in $\Omega\mathcal{F}$ if Q is a child of Q' in the grid $\mathcal{F}$.

Note that if ΩQ is a quasicube, then $|\Omega Q|^{\frac{1}{n}} \approx |Q|^{\frac{1}{n}} = \ell(Q) = \ell(\Omega Q)$. For a quasicube $J = \Omega Q$, we will generally use the expression $|J|^{\frac{1}{n}}$ in the various estimates arising in the proofs below, but will often use $\ell(J)$ when defining collections of quasicubes. Moreover, there are constants R_{big} and R_{small} such that we have the comparability containments

$$Q + \Omega x_Q \subset R_{big}\Omega Q \text{ and } R_{small}\Omega Q \subset Q + \Omega x_Q.$$

Given a fixed globally biLipschitz map Ω on $\mathbb{R}^n$, we will define below the n-dimensional $\mathcal{A}_2^\alpha$ conditions (with holes), punctured Muckenhoupt conditions $A_2^{\alpha,\text{punct}}$, testing conditions, and energy conditions using Ω-quasicubes in place of cubes, and we will refer to these new conditions as quasi$\mathcal{A}_2^\alpha$, quasitesting and quasienergy conditions. We will then prove a $T1$ theorem with quasitesting and with quasi$\mathcal{A}_2^\alpha$ and quasienergy side conditions on the weights. Since quasi$\mathcal{A}_2^\alpha \cap$ quasi$A_2^{\alpha,\text{punct}} = \mathcal{A}_2^\alpha \cap A_2^{\alpha,\text{punct}}$ (see [31]), we usually drop the prefix quasi from the various Muckenhoupt conditions (warning: quasi$\mathcal{A}_2^\alpha \neq \mathcal{A}_2^\alpha$).

Since the $\mathcal{A}_2^\alpha$ and punctured Muckenhoupt conditions typically hold, this identifies the culprit in higher dimensions as the pair of quasienergy conditions. We point out that these quasienergy conditions are implied by higher dimensional analogues of essentially all the other side conditions used previously in two weight theory, in particular doubling conditions and the Energy Hypothesis (1.16) in [16], as well as the condition of k-energy dispersed measures that is introduced below. This leads to our second theorem, which establishes the $T1$ theorem for strongly elliptic operators $\mathbf{T}^{\alpha}$ when both measures are k-energy dispered with k and α appropriately related.

It turns out that in higher dimensions, there are two natural 'Poisson integrals' P^{α} and $\mathcal{P}^{\alpha}$ that arise, the usual Poisson integral P^{α} that emerges in connection with energy considerations, and a different Poisson integral $\mathcal{P}^{\alpha}$ that emerges in connection with size considerations. The standard Poisson integral P^{α} appears in the energy conditions, and the reproducing Poisson integral $\mathcal{P}^{\alpha}$ appears in the $\mathcal{A}_2^{\alpha}$ condition. These two kernels coincide in dimension $n = 1$ for the case $\alpha = 0$ corresponding to the Hilbert transform.

Statements of Results

Now we turn to a precise description of our main two weight theorem.

Assumption: We fix once and for all a globally biLipschitz map $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ for use in all of our quasi-notions.

We will prove a two weight inequality for standard α-fractional Calderón-Zygmund operators T^{α} in Euclidean space $\mathbb{R}^n$, where we assume the n-dimensional $\mathcal{A}_2^{\alpha}$ conditions, new punctured A_2^{α} conditions, and certain *α-quasienergy conditions* as side conditions on the weights. In particular, we show that for positive locally finite Borel measures σ and ω in $\mathbb{R}^n$, and assuming that both the *quasienergy condition* and its dual hold, a strongly elliptic vector of standard α-fractional Calderón-Zygmund operators $\mathbf{T}^{\alpha}$ is bounded from $L^2(\sigma)$ to $L^2(\omega)$ *if and only if* the $\mathcal{A}_2^{\alpha}$ condition and its dual hold (we assume a mild additional condition on the quasicubes for this), the punctured Muckenhoupt condition $A_2^{\alpha,\mathrm{punct}}$ and its dual hold, the quasicube testing condition for $\mathbf{T}^{\alpha}$ and its dual hold, and the quasiweak boundedness property holds. In order to state our theorem precisely, we define these terms in the following subsections.

Remark 2.1 It is possible to collect our various Muckenhoupt and quasienergy assumptions on the weight pair (σ, ω) into just *two* compact side conditions of Muckenhoupt and quasienergy type. We prefer however, to keep the individual conditions separate so that the interested reader can track their use below.

Standard Fractional Singular Integrals and the Norm Inequality

Let $0 \leq \alpha < n$. We define a standard α-fractional CZ kernel $K^{\alpha}(x, y)$ to be a function defined on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying the following fractional size and smoothness conditions of order $1 + \delta$ for some $\delta > 0$,

$$|K^{\alpha}(x,y)| \leq C_{CZ}|x-y|^{\alpha-n} \text{ and } |\nabla K^{\alpha}(x,y)| \leq C_{CZ}|x-y|^{\alpha-n-1}, \quad (5)$$

$$\left|\nabla K^{\alpha}(x,y)-\nabla K^{\alpha}(x',y)\right| \leq C_{CZ}\left(\frac{|x-x'|}{|x-y|}\right)^{\delta}|x-y|^{\alpha-n-1}, \quad \frac{|x-x'|}{|x-y|} \leq \frac{1}{2},$$

and the last inequality also holds for the adjoint kernel in which x and y are interchanged. We note that a more general definition of kernel has only order of smoothness $\delta > 0$, rather than $1+\delta$, but the use of the Monotonicity and Energy Lemmas below, which involve first order Taylor approximations to the kernel functions $K^{\alpha}(\cdot,y)$, requires order of smoothness more than 1.

Defining the Norm Inequality

We now turn to a precise definition of the weighted norm inequality

$$\|T^{\alpha}_{\sigma}f\|_{L^2(\omega)} \leq \mathfrak{N}_{T^{\alpha}_{\sigma}}\|f\|_{L^2(\sigma)}, \quad f \in L^2(\sigma). \quad (6)$$

For this we introduce a family $\left\{\eta^{\alpha}_{\delta,R}\right\}_{0<\delta<R<\infty}$ of nonnegative functions on $[0,\infty)$ so that the truncated kernels $K^{\alpha}_{\delta,R}(x,y) = \eta^{\alpha}_{\delta,R}(|x-y|)K^{\alpha}(x,y)$ are bounded with compact support for fixed x or y. Then the truncated operators

$$T^{\alpha}_{\sigma,\delta,R}f(x) \equiv \int_{\mathbb{R}^n} K^{\alpha}_{\delta,R}(x,y)f(y)\,d\sigma(y), \quad x \in \mathbb{R}^n,$$

are pointwise well-defined, and we will refer to the pair $\left(K^{\alpha},\left\{\eta^{\alpha}_{\delta,R}\right\}_{0<\delta<R<\infty}\right)$ as an α-fractional singular integral operator, which we typically denote by T^{α}, suppressing the dependence on the truncations.

Definition 2.2 We say that an α-fractional singular integral operator $T^{\alpha} = \left(K^{\alpha},\left\{\eta^{\alpha}_{\delta,R}\right\}_{0<\delta<R<\infty}\right)$ satisfies the norm inequality (6) provided

$$\left\|T^{\alpha}_{\sigma,\delta,R}f\right\|_{L^2(\omega)} \leq \mathfrak{N}_{T^{\alpha}_{\sigma}}\|f\|_{L^2(\sigma)}, \quad f \in L^2(\sigma), 0<\delta<R<\infty.$$

It turns out that, in the presence of Muckenhoupt conditions, the norm inequality (6) is essentially independent of the choice of truncations used, and we now explain this in some detail. A *smooth truncation* of T^{α} has kernel $\eta_{\delta,R}(|x-y|)K^{\alpha}(x,y)$ for a smooth function $\eta_{\delta,R}$ compactly supported in (δ,R), $0<\delta<R<\infty$, and satisfying standard CZ estimates. A typical example of an α-fractional transform is the α-fractional *Riesz* vector of operators

$$\mathbf{R}^{\alpha,n} = \left\{R^{\alpha,n}_{\ell} : 1 \leq \ell \leq n\right\}, \quad 0 \leq \alpha < n.$$

The Riesz transforms $R_\ell^{n,\alpha}$ are convolution fractional singular integrals $R_\ell^{n,\alpha} f \equiv K_\ell^{n,\alpha} * f$ with odd kernel defined by

$$K_\ell^{\alpha,n}(w) \equiv \frac{w^\ell}{|w|^{n+1-\alpha}} \equiv \frac{\Omega_\ell(w)}{|w|^{n-\alpha}}, \qquad w = \left(w^1, \ldots, w^n\right).$$

However, in dealing with energy considerations, and in particular in the Monotonicity Lemma below where first order Taylor approximations are made on the truncated kernels, it is necessary to use the *tangent line truncation* of the Riesz transform $R_\ell^{\alpha,n}$ whose kernel is defined to be $\Omega_\ell(w)\,\psi_{\delta,R}^\alpha(|w|)$ where $\psi_{\delta,R}^\alpha$ is continuously differentiable on an interval $(0,S)$ with $0 < \delta < R < S$, and where $\psi_{\delta,R}^\alpha(r) = r^{\alpha-n}$ if $\delta \le r \le R$, and has constant derivative on both $(0,\delta)$ and (R,S) where $\psi_{\delta,R}^\alpha(S) = 0$. Here S is uniquely determined by R and α. Finally we set $\psi_{\delta,R}^\alpha(0) = 0$ as well, so that the kernel vanishes on the diagonal and common point masses do not 'see' each other. Note also that the tangent line extension of a $C^{1,\delta}$ function on the line is again $C^{1,\delta}$ with no increase in the $C^{1,\delta}$ norm.

It was shown in the one dimensional case with no common point masses in [13], that boundedness of the Hilbert transform H with one set of appropriate truncations together with the A_2^α condition without holes, is equivalent to boundedness of H with any other set of appropriate truncations. We need to extend this to $\mathbf{R}^{\alpha,n}$ and more general operators in higher dimensions and to permit common point masses, so that we are free to use the tangent line truncations throughout the proof of our theorem. For this purpose, we note that the difference between the tangent line truncated kernel $\Omega_\ell(w)\,\psi_{\delta,R}^\alpha(|w|)$ and the corresponding cutoff kernel $\Omega_\ell(w)\,\mathbf{1}_{[\delta,R]}(|w|)\,|w|^{\alpha-n}$ satisfies (since both kernels vanish at the origin)

$$\begin{aligned}
&\left|\Omega_\ell(w)\,\psi_{\delta,R}^\alpha(|w|) - \Omega_\ell(w)\,\mathbf{1}_{[\delta,R]}(|w|)\,|w|^{\alpha-n}\right| \\
&\lesssim \sum_{k=0}^\infty 2^{-k(n-\alpha)}\left\{\left(2^{-k}\delta\right)^{\alpha-n}\mathbf{1}_{[2^{-k-1}\delta,2^{-k}\delta]}(|w|)\right\} + \sum_{k=1}^\infty 2^{-k(n-\alpha)}\left\{\left(2^k R\right)^{\alpha-n}\mathbf{1}_{[2^{k-1}R,2^k R]}(|w|)\right\} \\
&\equiv \sum_{k=0}^\infty 2^{-k(n-\alpha)} K_{2^{-k}\delta}(w) + \sum_{k=1}^\infty 2^{-k(n-\alpha)} K_{2^k R}(w),
\end{aligned}$$

where the kernels $K_\rho(w) \equiv \frac{1}{\rho^{n-\alpha}}\mathbf{1}_{[\rho,2\rho]}(|w|)$ are easily seen to satisfy, uniformly in ρ, the norm inequality (12) with constant controlled by the offset A_2^α condition (7) below. The equivalence of the norm inequality for these two families of truncations now follows from the summability of the series $\sum_{k=0}^\infty 2^{-k(n-\alpha)}$ for $0 \le \alpha < n$. The case of more general families of truncations and operators is similar.

Quasicube Testing Conditions

The following 'dual' quasicube testing conditions are necessary for the boundedness of T^α from $L^2(\sigma)$ to $L^2(\omega)$,

$$\mathfrak{T}_{T^\alpha}^2 \equiv \sup_{Q\in\Omega\mathcal{P}^n} \frac{1}{|Q|_\sigma} \int_Q \left|T^\alpha\left(\mathbf{1}_Q\sigma\right)\right|^2 \omega < \infty,$$

$$\left(\mathfrak{T}_{T^\alpha}^*\right)^2 \equiv \sup_{Q\in\Omega\mathcal{P}^n} \frac{1}{|Q|_\omega} \int_Q \left|(T^\alpha)^*\left(\mathbf{1}_Q\omega\right)\right|^2 \sigma < \infty,$$

and where we interpret the right sides as holding uniformly over all tangent line truncations of T^α.

Remark 2.3 We alert the reader that the symbols Q, I, J, K will all be used to denote either cubes or quasicubes, and the context will make clear which is the case. Throughout most of the proof of the main theorem only quasicubes are considered.

Quasiweak Boundedness Property

The quasiweak boundedness property for T^α with constant C is given by

$$\left|\int_Q T^\alpha\left(1_{Q'}\sigma\right) d\omega\right| \leq \mathcal{WBP}_{T^\alpha}\sqrt{|Q|_\omega\,|Q'|_\sigma},$$

$$\text{for all quasicubes } Q, Q' \text{ with } \frac{1}{C} \leq \frac{|Q|^{\frac{1}{n}}}{|Q'|^{\frac{1}{n}}} \leq C,$$

$$\text{and either } Q \subset 3Q' \setminus Q' \text{ or } Q' \subset 3Q \setminus Q,$$

and where we interpret the left side above as holding uniformly over all tangent line trucations of T^α. Note that the quasiweak boundedness property is implied by either the *tripled* quasicube testing condition,

$$\left\|\mathbf{1}_{3Q}\mathbf{T}^\alpha\left(\mathbf{1}_Q\sigma\right)\right\|_{L^2(\omega)} \leq \mathfrak{T}_{\mathbf{T}^\alpha}^{\text{triple}}\left\|\mathbf{1}_Q\right\|_{L^2(\sigma)}, \quad \text{for all quasicubes } Q \text{ in } \mathbb{R}^n,$$

or its dual defined with σ and ω interchanged and the dual operator $\mathbf{T}^{\alpha,*}$ in place of $\mathbf{T}^\alpha$. In turn, the tripled quasicube testing condition can be obtained from the quasicube testing condition for the truncated weight pairs $\left(\omega, \mathbf{1}_Q\sigma\right)$.

Poisson Integrals and $\mathcal{A}_2^\alpha$

Let μ be a locally finite positive Borel measure on $\mathbb{R}^n$, and suppose Q is an Ω-quasicube in $\mathbb{R}^n$. Recall that $|Q|^{\frac{1}{n}} \approx \ell(Q)$ for a quasicube Q. The two α-fractional Poisson integrals of μ on a quasicube Q are given by:

$$\mathrm{P}^{\alpha}(Q,\mu) \equiv \int_{\mathbb{R}^n} \frac{|Q|^{\frac{1}{n}}}{\left(|Q|^{\frac{1}{n}} + |x - x_Q|\right)^{n+1-\alpha}} d\mu(x),$$

$$\mathcal{P}^{\alpha}(Q,\mu) \equiv \int_{\mathbb{R}^n} \left(\frac{|Q|^{\frac{1}{n}}}{\left(|Q|^{\frac{1}{n}} + |x - x_Q|\right)^{2}} \right)^{n-\alpha} d\mu(x),$$

where we emphasize that $|x - x_Q|$ denotes Euclidean distance between x and x_Q and $|Q|$ denotes the Lebesgue measure of the quasicube Q. We refer to P^{α} as the *standard* Poisson integral and to $\mathcal{P}^{\alpha}$ as the *reproducing* Poisson integral.

We say that the pair K, K' in $\mathcal{P}^n$ are *neighbours* if K and K' live in a common dyadic grid and both $K \subset 3K' \setminus K'$ and $K' \subset 3K \setminus K$, and we denote by $\mathcal{N}^n$ the set of pairs (K, K') in $\mathcal{P}^n \times \mathcal{P}^n$ that are neighbours. Let

$$\Omega\mathcal{N}^n = \left\{ \left(\Omega K, \Omega K'\right) : \left(K, K'\right) \in \mathcal{N}^n \right\}$$

be the corresponding collection of quasineighbour pairs of quasicubes. Let σ and ω be locally finite positive Borel measures on $\mathbb{R}^n$, possibly having common point masses, and suppose $0 \leq \alpha < n$. Then we define the classical *offset* A_2^{α} *constants* by

$$A_2^{\alpha}(\sigma,\omega) \equiv \sup_{(Q,Q') \in \Omega\mathcal{N}^n} \frac{|Q|_{\sigma}}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q'|_{\omega}}{|Q'|^{1-\frac{\alpha}{n}}}. \tag{7}$$

Since the cubes in $\mathcal{P}^n$ are products of half open, half closed intervals $[a, b)$, the neighbouring quasicubes $(Q, Q') \in \Omega\mathcal{N}^n$ are disjoint, and the common point masses of σ and ω do not simultaneously appear in each factor.

We now define the *one-tailed* $\mathcal{A}_2^{\alpha}$ constant using $\mathcal{P}^{\alpha}$. The energy constants $\mathcal{E}_{\alpha}^{\mathrm{strong}}$ introduced in the next subsection will use the standard Poisson integral P^{α}.

Definition 2.4 The one-tailed constants $\mathcal{A}_2^{\alpha}$ and $\mathcal{A}_2^{\alpha,*}$ for the weight pair (σ, ω) are given by

$$\mathcal{A}_2^{\alpha} \equiv \sup_{Q \in \Omega\mathcal{P}^n} \mathcal{P}^{\alpha}\left(Q, \mathbf{1}_{Q^c}\sigma\right) \frac{|Q|_{\omega}}{|Q|^{1-\frac{\alpha}{n}}} < \infty,$$

$$\mathcal{A}_2^{\alpha,*} \equiv \sup_{Q \in \Omega\mathcal{P}^n} \mathcal{P}^{\alpha}\left(Q, \mathbf{1}_{Q^c}\omega\right) \frac{|Q|_{\sigma}}{|Q|^{1-\frac{\alpha}{n}}} < \infty.$$

Note that these definitions are the analogues of the corresponding conditions with ‘holes’ introduced by Hytönen [5] in dimension $n = 1$ – the supports of the measures $\mathbf{1}_{Q^c}\sigma$ and $\mathbf{1}_{Q}\omega$ in the definition of $\mathcal{A}_2^{\alpha}$ are disjoint, and so the common point masses of σ and ω do not appear simultaneously in each factor. Note also that, unlike in [28], where common point masses were not permitted, we can no longer assert the equivalence of $\mathcal{A}_2^{\alpha}$ with holes taken over *quasicubes* with $\mathcal{A}_2^{\alpha}$ with holes taken over *cubes*.

Punctured A_2^α Conditions

As mentioned earlier, the *classical* A_2^α characteristic $\sup_{Q\in\Omega\mathcal{Q}^n}\frac{|Q|_\omega}{|Q|^{1-\frac{\alpha}{n}}}\frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}}$ fails to be finite when the measures σ and ω have a common point mass - simply let Q in the sup above shrink to a common mass point. But there is a substitute that is quite similar in character that is motivated by the fact that for large quasicubes Q, the sup above is problematic only if just *one* of the measures is *mostly* a point mass when restricted to Q. The one-dimensional version of the condition we are about to describe arose in Conjecture 1.12 of Lacey [9], and it was pointed out in [6] that its necessity on the line follows from the proof of Proposition 2.1 in [16]. We now extend this condition to higher dimensions, where its necessity is more subtle.

Given an at most countable set $\mathfrak{P}=\{p_k\}_{k=1}^\infty$ in $\mathbb{R}^n$, a quasicube $Q\in\Omega\mathcal{P}^n$, and a positive locally finite Borel measure μ, define

$$\mu(Q,\mathfrak{P})\equiv|Q|_\mu-\sup\{\mu(p_k):p_k\in Q\cap\mathfrak{P}\},$$

where the supremum is actually achieved since $\sum_{p_k\in Q\cap\mathfrak{P}}\mu(p_k)<\infty$ as μ is locally finite. The quantity $\mu(Q,\mathfrak{P})$ is simply the $\widetilde{\mu}$ measure of Q where $\widetilde{\mu}$ is the measure μ with its largest point mass from $\mathfrak{P}$ in Q removed. Given a locally finite measure pair (σ,ω), let $\mathfrak{P}_{(\sigma,\omega)}=\{p_k\}_{k=1}^\infty$ be the at most countable set of common point masses of σ and ω. Then the weighted norm inequality (6) typically implies finiteness of the following *punctured* Muckenhoupt conditions:

$$A_2^{\alpha,\text{punct}}(\sigma,\omega)\equiv\sup_{Q\in\Omega\mathcal{P}^n}\frac{\omega\left(Q,\mathfrak{P}_{(\sigma,\omega)}\right)}{|Q|^{1-\frac{\alpha}{n}}}\frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}},$$

$$A_2^{\alpha,*,\text{punct}}(\sigma,\omega)\equiv\sup_{Q\in\Omega\mathcal{P}^n}\frac{|Q|_\omega}{|Q|^{1-\frac{\alpha}{n}}}\frac{\sigma\left(Q,\mathfrak{P}_{(\sigma,\omega)}\right)}{|Q|^{1-\frac{\alpha}{n}}}.$$

Lemma 2.5 *Let* $\mathbf{T}^\alpha$ *be an* α*-fractional singular integral operator as above, and suppose that there is a positive constant* C_0 *such that*

$$A_2^\alpha(\sigma,\omega)\leq C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\sigma,\omega),$$

for all pairs (σ,ω) *of positive locally finite measures* ***having no common point masses****. Now let* σ *and* ω *be positive locally finite Borel measures on* $\mathbb{R}^n$ *and let* $\mathfrak{P}_{(\sigma,\omega)}$ *be the possibly nonempty set of common point masses. Then we have*

$$A_2^{\alpha,\text{punct}}(\sigma,\omega)+A_2^{\alpha,*,\text{punct}}(\sigma,\omega)\leq 4C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\sigma,\omega).$$

Proof Fix a quasicube $Q\in\Omega\mathcal{P}^n$. Suppose first that $\mathfrak{P}_{(\sigma,\omega)}\cap Q=\{p_k\}_{k=1}^{2N}$ is finite with an even number of points. Choose $k_1\in\mathbb{N}_{2N}=\{1,2,\ldots,2N\}$ so that

$$\sigma(p_{k_1})=\max_{k\in\mathbb{N}_{2N}}\sigma(p_k).$$

Then choose $k_2 \in \mathbb{N}_{2N} \setminus \{k_1\}$ such that

$$\omega\left(p_{k_2}\right) = \max_{k\in\mathbb{N}_{2N}\setminus\{k_1\}} \omega\left(p_k\right).$$

Repeat this procedure so that

$$\sigma\left(p_{k_{2m+1}}\right) = \max_{k\in\mathbb{N}_{2N}\setminus\{k_1,\dots k_{2m}\}} \sigma\left(p_k\right), \qquad k_{2m+1} \in \mathbb{N}_{2N} \setminus \{k_1, \dots k_{2m}\},$$
$$\omega\left(p_{k_{2m+2}}\right) = \max_{k\in\mathbb{N}_{2N}\setminus\left\{k_1,\dots k_{2m+1}\right\}} \omega\left(p_k\right), \qquad k_{2m+2} \in \mathbb{N}_{2N} \setminus \{k_1, \dots k_{2m+1}\},$$

for each $m \leq N-1$. It is now clear that both

$$\sum_{i=0}^{N-1} \sigma\left(p_{k_{2i+1}}\right) \geq \frac{1}{2}\sigma\left(Q \cap \mathfrak{P}_{(\sigma,\omega)}\right) \text{ and } \sum_{i=0}^{N-1} \omega\left(p_{k_{2i+2}}\right) \geq \frac{1}{2}\left[\omega\left(Q \cap \mathfrak{P}_{(\sigma,\omega)}\right) - \omega\left(p_1\right)\right].$$

In the case of an odd number $2N-1$ of common point masses, the second inequality will have $N-1$ replaced with $N-2$.

Now, returning to the case of $2N$ common point masses, define new measures $\widetilde{\sigma}$ and $\widetilde{\omega}$ by

$$\widetilde{\sigma} \equiv \mathbf{1}_Q\sigma - \sum_{i=0}^{N-1} \sigma\left(p_{k_{2i+2}}\right)\delta_{p_{k_{2i+2}}} \text{ and } \widetilde{\omega} = \mathbf{1}_Q\omega - \sum_{i=0}^{N-1} \omega\left(p_{k_{2i+1}}\right)\delta_{p_{k_{2i+1}}}$$

so that

$$|Q|_{\widetilde{\sigma}} \geq \frac{1}{2}|Q|_\sigma \text{ and } |Q|_{\widetilde{\omega}} \geq \frac{1}{2}\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right)$$

Now $\widetilde{\sigma}$ and $\widetilde{\omega}$ have no common point masses and $\mathfrak{N}_{\mathbf{T}^\alpha}(\sigma,\omega)$ is monotone in each measure separately, so we have

$$\frac{\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right)}{|Q|^{1-\frac{\alpha}{n}}}\frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}} \leq 4A_2^\alpha(\widetilde{\sigma},\widetilde{\omega}) \leq 4C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\widetilde{\sigma},\widetilde{\omega}) \leq 4C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\sigma,\omega).$$

Thus $A_2^{\alpha,\text{punct}}(\sigma,\omega) \leq 4C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\sigma,\omega)$ if the number of common point masses in Q is finite. A limiting argument proves the general case. The dual inequality $A_2^{\alpha,*,\text{punct}}(\sigma,\omega) \leq 4C_0\mathfrak{N}_{\mathbf{T}^\alpha}^2(\sigma,\omega)$ now follows upon interchanging the measures σ and ω. □

Now we turn to the definition of a quasiHaar basis of $L^2(\mu)$.

A Weighted QuasiHaar Basis

We will use a construction of a quasiHaar basis in $\mathbb{R}^n$ that is adapted to a measure μ (cf. [20] for the nonquasi case and [7] for the geometrically doubling quasi-metric space case). Given a dyadic quasicube $Q \in \Omega\mathcal{D}$, where $\mathcal{D}$ is a dyadic grid of cubes from $\mathcal{P}^n$, let $\triangle_Q^\mu$ denote orthogonal projection onto the finite dimensional subspace $L_Q^2(\mu)$ of $L^2(\mu)$ that consists of linear combinations of the indicators of the children $\mathfrak{C}(Q)$ of Q that have μ-mean zero over Q:

$$L_Q^2(\mu) \equiv \left\{ f = \sum_{Q' \in \mathfrak{C}(Q)} a_{Q'} \mathbf{1}_{Q'} : a_{Q'} \in \mathbb{R}, \int_Q f d\mu = 0 \right\}.$$

Then we have the important telescoping property for dyadic quasicubes $Q_1 \subset Q_2$ (where $[Q_1, Q_2) \equiv \{Q \text{ dyadic} : Q_1 \subset Q \subsetneqq Q_2\}$):

$$\mathbf{1}_{Q_0}(x) \left(\sum_{Q \in [Q_1, Q_2)} \triangle_Q^\mu f(x) \right) = \mathbf{1}_{Q_0}(x) \left(\mathbb{E}_{Q_0}^\mu f - \mathbb{E}_{Q_2}^\mu f \right), \quad Q_0 \in \mathfrak{C}(Q_1), f \in L^2(\mu). \tag{8}$$

We will at times find it convenient to use a fixed orthonormal basis $\{h_Q^{\mu,a}\}_{a \in \Gamma_n}$ of $L_Q^2(\mu)$ where $\Gamma_n \equiv \{0,1\}^n \setminus \{\mathbf{1}\}$ is a convenient index set with $\mathbf{1} = (1, 1, \ldots, 1)$. Then $\{h_Q^{\mu,a}\}_{a \in \Gamma_n \text{ and } Q \in \Omega\mathcal{D}}$ is an orthonormal basis for $L^2(\mu)$, with the understanding that we add the constant function $\mathbf{1}$ if μ is a finite measure. In particular we have

$$\|f\|_{L^2(\mu)}^2 = \sum_{Q \in \Omega\mathcal{D}} \left\| \triangle_Q^\mu f \right\|_{L^2(\mu)}^2 = \sum_{Q \in \Omega\mathcal{D}} \left| \widehat{f}(Q) \right|^2, \quad \left| \widehat{f}(Q) \right|^2 \equiv \sum_{a \in \Gamma_n} \left| \left\langle f, h_Q^{\mu,a} \right\rangle_\mu \right|^2,$$

where the measure is suppressed in the notation $\widehat{f}$. Indeed, this follows from (8) and Lebesgue's differentiation theorem for quasicubes. We also record the following useful estimate. If I' is any of the 2^n $\Omega\mathcal{D}$-children of I, and $a \in \Gamma_n$, then

$$\left| \mathbb{E}_{I'}^\mu h_I^{\mu,a} \right| \leq \sqrt{\mathbb{E}_{I'}^\mu \left(h_I^{\mu,a} \right)^2} \leq \frac{1}{\sqrt{|I'|_\mu}}. \tag{9}$$

The Strong Quasienergy Conditions

Given a dyadic quasicube $K \in \Omega\mathcal{D}$ and a positive measure μ we define the quasiHaar projection $\mathsf{P}_K^\mu \equiv \sum_{J \in \Omega\mathcal{D}:\, J \subset K} \triangle_J^\mu$ on K by

$$\mathsf{P}_K^{\mu} f = \sum_{J \in \Omega\mathcal{D}:\ J \subset K} \sum_{a \in \Gamma_n} \left\langle f, h_J^{\mu,a} \right\rangle_{\mu} h_J^{\mu,a} \text{ and } \left\| \mathsf{P}_K^{\mu} f \right\|_{L^2(\mu)}^2 = \sum_{J \in \Omega\mathcal{D}:\ J \subset K} \sum_{a \in \Gamma_n} \left| \left\langle f, h_J^{\mu,a} \right\rangle_{\mu} \right|^2,$$

and where a quasiHaar basis $\left\{ h_J^{\mu,a} \right\}_{a \in \Gamma_n \text{ and } J \in \Omega\mathcal{D}\Omega}$ adapted to the measure μ was defined in the subsubsection on a weighted quasiHaar basis above.

Now we define various notions for quasicubes which are inherited from the same notions for cubes. The main objective here is to use the familiar notation that one uses for cubes, but now extended to Ω-quasicubes. We have already introduced the notions of quasigrids $\Omega\mathcal{D}$, and center, sidelength and dyadic associated to quasicubes $Q \in \Omega\mathcal{D}$, as well as quasiHaar functions, and we will continue to extend to quasicubes the additional familiar notions related to cubes as we come across them. We begin with the notion of *deeply embedded*. Fix a quasigrid $\Omega\mathcal{D}$. We say that a dyadic quasicube J is $(\mathbf{r}, \varepsilon)$*-deeply embedded* in a (not necessarily dyadic) quasicube K, which we write as $J \Subset_{\mathbf{r},\varepsilon} K$, when $J \subset K$ and both

$$\begin{aligned} \ell(J) &\leq 2^{-\mathbf{r}} \ell(K), \qquad (10) \\ \operatorname{qdist}(J, \partial K) &\geq \frac{1}{2} \ell(J)^{\varepsilon} \ell(K)^{1-\varepsilon}, \end{aligned}$$

where we define the quasidistance $\operatorname{qdist}(E, F)$ between two sets E and F to be the Euclidean distance $\operatorname{dist}\left(\Omega^{-1}E, \Omega^{-1}F\right)$ between the preimages $\Omega^{-1}E$ and $\Omega^{-1}F$ of E and F under the map Ω, and where we recall that $\ell(J) \approx |J|^{\frac{1}{n}}$. For the most part we will consider $J \Subset_{\mathbf{r},\varepsilon} K$ when J and K belong to a common quasigrid $\Omega\mathcal{D}$, but an exception is made when defining the strong energy constants below.

Recall that in dimension $n = 1$, and for $\alpha = 0$, the energy condition constant was defined by

$$\mathcal{E}^2 \equiv \sup_{I = \dot{\cup} I_r} \frac{1}{|I|_{\sigma}} \sum_{r=1}^{\infty} \left(\frac{\mathrm{P}^{\alpha}(I_r, \mathbf{1}_I \sigma)}{|I_r|} \right)^2 \left\| \mathsf{P}_{I_r}^{\omega} \mathbf{x} \right\|_{L^2(\omega)}^2,$$

where I and I_r are intervals in the real line, and $\dot{\cup}$ denotes a pairwise disjoint union. The extension to higher dimensions we use here is that of 'strong quasienergy condition' below. Later on, in the proof of the theorem, we will break down this strong quasienergy condition into various smaller quasienergy conditions, which are then used in different ways in the proof.

We define a quasicube K (not necessarily in $\Omega\mathcal{D}$) to be an *alternate* $\Omega\mathcal{D}$-quasicube if it is a union of 2^n $\Omega\mathcal{D}$-quasicubes K' with side length $\ell(K') = \frac{1}{2}\ell(K)$ (such quasicubes were called shifted in [28], but that terminology conflicts with the more familiar notion of shifted quasigrid). Thus for any $\Omega\mathcal{D}$-quasicube L there are exactly 2^n alternate $\Omega\mathcal{D}$-quasicubes of twice the side length that contain L, and one of them is of course the $\Omega\mathcal{D}$-parent of L. We denote the collection of alternate $\Omega\mathcal{D}$-quasicubes by $\mathcal{A}\Omega\mathcal{D}$.

The extension of the energy conditions to higher dimensions in [28] used the collection

$$\mathcal{M}_{\mathbf{r},\varepsilon-\text{deep}}(K) \equiv \{\text{maximal } J \Subset_{\mathbf{r},\varepsilon} K\}$$

of *maximal* $(\mathbf{r},\varepsilon)$-deeply embedded dyadic subquasicubes of a quasicube K (a subquasicube J of K is a *dyadic* subquasicube of K if $J \in \Omega\mathcal{D}$ when $\Omega\mathcal{D}$ is a dyadic quasigrid containing K). This collection of dyadic subquasicubes of K is of course a pairwise disjoint decomposition of K.We also defined there a refinement and extension of the collection $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(K)$ for certain K and each $\ell \geq 1$. For an alternate quasicube $K \in \mathcal{A}\Omega\mathcal{D}$, define $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K)$ to consist of the *maximal* $\mathbf{r}$-deeply embedded $\Omega\mathcal{D}$-dyadic subquasicubes J of K. (In the special case that K itself belongs to $\Omega\mathcal{D}$, then $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K) = \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(K)$.) Then in [28] for $\ell \geq 1$ we defined the refinement (where $\pi^\ell K'$ denotes the ℓ^{th} ancestor of K' in the grid):

$$\begin{aligned}\mathcal{M}^{\ell}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K) \equiv \Big\{J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}\left(\pi^\ell K'\right) &\text{ for some } K' \in \mathfrak{C}_{\Omega\mathcal{D}}(K) : \\ &J \subset L \text{ for some } L \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(K)\Big\},\end{aligned}$$

where $\mathfrak{C}_{\Omega\mathcal{D}}(K)$ is the obvious extension to alternate quasicubes of the set of $\Omega\mathcal{D}$-dyadic children. Thus $\mathcal{M}^{\ell}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K)$ is the union, over all quasichildren K' of K, of those quasicubes in $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}\left(\pi^\ell K'\right)$ that happen to be contained in some $L \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K)$. We then define the *strong* quasienergy condition as follows.

Definition 2.6 Let $0 \leq \alpha < n$ and fix parameters $(\mathbf{r},\varepsilon)$. Suppose σ and ω are positive Borel measures on $\mathbb{R}^n$ possibly with common point masses. Then the *strong* quasienergy constant $\mathcal{E}^{\text{strong}}_\alpha$ is defined by[7]

$$\begin{aligned}\left(\mathcal{E}^{\text{strong}}_\alpha\right)^2 \equiv \sup_{\Omega\mathcal{D}} \sup_{\substack{I=\dot\cup I_r \\ I,I_r\in\Omega\mathcal{D}}} \frac{1}{|I|_\sigma} \sum_{r=1}^{\infty} \sum_{J\in\mathcal{M}_{\mathbf{r},\varepsilon-\text{deep}}(I_r)} \left(\frac{\mathrm{P}^\alpha(J,\mathbf{1}_I\sigma)}{|J|^{\frac{1}{n}}}\right)^2 \left\|\mathsf{P}^\omega_J \mathbf{x}\right\|^2_{L^2(\omega)} \\ + \sup_{\Omega\mathcal{D}} \sup_{I\in\mathcal{A}\Omega\mathcal{D}} \sup_{\ell\geq 0} \frac{1}{|I|_\sigma} \sum_{J\in\mathcal{M}^{\ell}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(I)} \left(\frac{\mathrm{P}^\alpha(J,\mathbf{1}_I\sigma)}{|J|^{\frac{1}{n}}}\right)^2 \left\|\mathsf{P}^\omega_J \mathbf{x}\right\|^2_{L^2(\omega)}.\end{aligned}$$

Similarly we have a dual version of $\mathcal{E}^{\text{strong}}_\alpha$ denoted $\mathcal{E}^{\text{strong},*}_\alpha$, and both depend on $\mathbf{r}$ and ε as well as on n and α. An important point in this definition is that the quasicube I in the second line is permitted to lie *outside* the quasigrid $\Omega\mathcal{D}$, but only as an alternate dyadic quasicube $I \in \mathcal{A}\Omega\mathcal{D}$. In the setting of quasicubes we continue to use the linear function $\mathbf{x}$ in the final factor $\left\|\mathsf{P}^\omega_J \mathbf{x}\right\|^2_{L^2(\omega)}$ of each line, and not the

[7]The first line in the display in Definition 5 in [29] is missing notation that is corrected here.

pushforward of $\mathbf{x}$ by Ω. The reason of course is that this condition is used to capture the first order information in the Taylor expansion of a singular kernel. There is a logically weaker form of the quasienergy conditions that we discuss after stating our main theorem, but these refined quasienergy conditions are more complicated to state, and have as yet found no application - the strong energy conditions above suffice for use when one measure is compactly supported on a $C^{1,\delta}$ curve as in [31].

Statement of the Theorems

We can now state our main quasicube two weight theorem for general measures allowing common point masses, as well as our application to energy dispersed measures. Recall that $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ is a globally biLipschitz map, and that $\Omega\mathcal{P}^n$ denotes the collection of all quasicubes in $\mathbb{R}^n$ whose preimages under Ω are usual cubes with sides parallel to the coordinate axes. Denote by $\Omega\mathcal{D} \subset \Omega\mathcal{P}^n$ a dyadic quasigrid in $\mathbb{R}^n$. For the purpose of obtaining necessity of $\mathcal{A}_2^\alpha$ for $\frac{n}{2} \leq \alpha < n$, we adapt the notion of strong ellipticity from [30].

Definition 2.7 Fix a globally biLipschitz map Ω. Let $\mathbf{T}^\alpha = \left\{T_j^\alpha\right\}_{j=1}^J$ be a vector of singular integral operators with standard kernels $\left\{K_j^\alpha\right\}_{j=1}^J$. We say that $\mathbf{T}^\alpha$ is *strongly elliptic* with respect to Ω if for each $m \in \{1, -1\}^n$, there is a sequence of coefficients $\left\{\lambda_j^m\right\}_{j=1}^J$ such that

$$\left|\sum_{j=1}^J \lambda_j^m K_j^\alpha (x, x + t\mathbf{u})\right| \geq ct^{\alpha-n}, \qquad t \in \mathbb{R}, \tag{11}$$

holds for *all* unit vectors $\mathbf{u}$ in the quasi-n-ant ΩV_m (i.e. an n-dimensional quasi-quadrant) where

$$V_m = \{x \in \mathbb{R}^n : m_i x_i > 0 \text{ for } 1 \leq i \leq n\}, \qquad m \in \{1, -1\}^n.$$

Theorem 2.8 *Suppose that T^α is a standard α-fractional singular integral operator on $\mathbb{R}^n$, and that ω and σ are positive Borel measures on $\mathbb{R}^n$ (possibly having common point masses). Set $T_\sigma^\alpha f = T^\alpha(f\sigma)$ for any smooth truncation of T_σ^α. Let $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ be a globally biLipschitz map.*

1. *Suppose $0 \leq \alpha < n$. Then the operator T_σ^α is bounded from $L^2(\sigma)$ to $L^2(\omega)$, i.e.*

$$\|T_\sigma^\alpha f\|_{L^2(\omega)} \leq \mathfrak{N}_{T_\sigma^\alpha} \|f\|_{L^2(\sigma)}, \tag{12}$$

uniformly in smooth truncations of T^α, and moreover

$$\mathfrak{N}_{T^\alpha_\sigma} \leq C_\alpha \left(\sqrt{\mathcal{A}_2^\alpha + \mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\text{punct}} + A_2^{\alpha,*,\text{punct}}} + \mathfrak{T}_{T^\alpha} + \mathfrak{T}^*_{T^\alpha} + \mathcal{E}_\alpha^{\text{strong}} + \mathcal{E}_\alpha^{\text{strong},*} + \mathcal{WBP}_{T^\alpha} \right),$$

provided that the two dual $\mathcal{A}_2^\alpha$ conditions and the two dual punctured Muckenhoupt conditions all hold, and the two dual quasitesting conditions for T^α hold, the quasiweak boundedness property for T^α holds for a sufficiently large constant C depending on the goodness parameter $\mathbf{r}$*, and provided that the two dual strong quasienergy conditions hold uniformly over all dyadic quasigrids $\Omega\mathcal{D} \subset \Omega\mathcal{P}^n$, i.e. $\mathcal{E}_\alpha^{\text{strong}} + \mathcal{E}_\alpha^{\text{strong},*} < \infty$, and where the goodness parameters* $\mathbf{r}$ *and ε in the definition of the collections $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(K)$ and $\mathcal{M}^{\ell}_{(\mathbf{r},\varepsilon)-\text{deep},\Omega\mathcal{D}}(K)$ appearing in the strong energy conditions, are fixed sufficiently large and small respectively depending only on n and α.*

2. *Conversely, suppose $0 \leq \alpha < n$ and that $\mathbf{T}^\alpha = \left\{T_j^\alpha\right\}_{j=1}^J$ is a vector of Calderón-Zygmund operators with standard kernels $\left\{K_j^\alpha\right\}_{j=1}^J$. In the range $0 \leq \alpha < \frac{n}{2}$, we assume the* ellipticity *condition from (*[30]*): there is $c > 0$ such that for* each *unit vector* $\mathbf{u}$ *there is j satisfying*

$$\left|K_j^\alpha(x, x + t\mathbf{u})\right| \geq ct^{\alpha-n}, \qquad t \in \mathbb{R}. \tag{13}$$

For the range $\frac{n}{2} \leq \alpha < n$, we assume the strong ellipticity condition in Definition 2.7 *above. Furthermore, assume that each operator T_j^α is bounded from $L^2(\sigma)$ to $L^2(\omega)$,*

$$\left\|\left(T_j^\alpha\right)_\sigma f\right\|_{L^2(\omega)} \leq \mathfrak{N}_{T_j^\alpha} \|f\|_{L^2(\sigma)}.$$

Then the fractional $\mathcal{A}_2^\alpha$ conditions (with 'holes') hold as well as the punctured Muckenhoupt conditions, and moreover,

$$\sqrt{\mathcal{A}_2^\alpha + \mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\text{punct}} + A_2^{\alpha,*,\text{punct}}} \leq C\mathfrak{N}_{\mathbf{T}^\alpha}.$$

Problem 1 *Given any strongly elliptic vector $\mathbf{T}^\alpha$ of classical α-fractional Calderón-Zygmund operators, it is an open question whether or not the usual (quasi or not) energy conditions are necessary for boundedness of $\mathbf{T}^\alpha$. See* [27] *for a failure of* energy reversal *in higher dimensions – such an energy reversal was used in dimension $n = 1$ to prove the necessity of the energy condition for the Hilbert transform, and also in* [26] *and* [14] *for the Riesz transforms and Cauchy transforms respectively when one of the measures is supported on a line, and in* [31] *when one of the measures is supported on a $C^{1,\delta}$ curve.*

Remark 2.9 If Definition 2.7 holds for some $\mathbf{T}^\alpha$ and Ω, then Ω must be fairly tame, in particular the logarithmic spirals in Example 1.2 are ruled out! On the other hand, the vector of Riesz transforms $\mathbf{R}^{\alpha,n}$ is easily seen to be strongly elliptic with respect to Ω if Ω satisfies the following *sector separation property*. Given a hyperplane H

and a perpendicular line L intersecting at point P, there exist spherical cones S_H and S_L intersecting only at the point $P' = \Omega(P)$, such that $H' \equiv \Omega H \subset S_H$ and $L' \equiv \Omega L \subset S_L$ and

$$\operatorname{dist}(x, \partial S_H) \approx |x|, \ x \in H \text{ and } \operatorname{dist}(x, \partial S_L) \approx |x|, \ x \in L.$$

Examples of globally biLipshcitz maps Ω that satisfy the sector separation property include finite compositions of maps of the form

$$\Omega\left(x_1, x'\right) = \left(x_1, x' + \psi(x_1)\right), \qquad \left(x_1, x'\right) \in \mathbb{R}^n,$$

where $\psi : \mathbb{R} \to \mathbb{R}^{n-1}$ is a Lipschitz map with sufficiently small Lipschitz constant.

In order to state our application to energy dispersed measures, we introduce some notation and a definition. Fix a globally biLipschitz map $\Omega : \mathbb{R}^n \to \mathbb{R}^n$. For $0 \le k \le n-1$, denote by $\mathcal{L}_k^n$ the collection of all k-dimensional planes in $\mathbb{R}^n$. If in addition J is an Ω-quasicube in $\mathbb{R}^n$, denote by $M_k^n(J,\mu)$ the 'moments'

$$\mathsf{M}_k^n(J,\mu)^2 \equiv \inf_{L \in \mathcal{L}_k^n} \int_J \operatorname{dist}(x, L)^2 \, d\mu(x),$$

and note that $\mathsf{M}_0^n(J,\mu)$ is related to the energy $\mathsf{E}(J,\mu) \equiv \sqrt{\mathbb{E}_J^{\mu} \left| \frac{x - \mathbb{E}_J^{\mu} x}{|J|^{\frac{1}{n}}} \right|^2}$, $\mathbb{E}_J^{\mu} x = \frac{1}{|J|_{\mu}} \int_J x d\mu(x)$:

$$\mathsf{M}_0^n(J,\mu)^2 = \int_J \left|x - \mathbb{E}_J^{\mu} x\right|^2 d\mu(x) = |J|_{\mu} |J|^{\frac{2}{n}} \mathsf{E}(J,\mu)^2.$$

Clearly the moments decrease in k and we now give a name to various reversals of this decrease.

Definition 2.10 Suppose μ is a locally finite Borel measure on $\mathbb{R}^n$, and let k be an integer satisfying $0 \le k \le n-1$. We say that μ is *k-energy dispersed* if there is a positive constant $C = C_{k,n}$ such that for all Ω-quasicubes J,

$$\mathsf{M}_0^n(J,\mu) \le C \mathsf{M}_k^n(J,\mu).$$

If both σ and ω are appropriately energy dispersed relative to the order $0 \le \alpha < n$, then the $T1$ theorem holds for the α-fractional Riesz vector transform $\mathbf{R}^{\alpha,n}$.

Theorem 2.11 *Let* $0 \le \alpha < n$ *and* $0 \le k \le n-1$ *sastisfy*

$$\begin{cases} n-k < \alpha < n, \ \alpha \ne n-1 & \text{if } \ 1 \le k \le n-2 \\ 0 \le \alpha < n, \ \alpha \ne 1, n-1 & \text{if } \quad k = n-1 \end{cases}.$$

Suppose that $\mathbf{R}^{\alpha,n}$ *is the* α*-fractional Riesz vector transform on* $\mathbb{R}^n$*, and that* ω *and* σ *are* k-energy dispersed *locally finite positive Borel measures on* $\mathbb{R}^n$ *(possibly having common point masses). Set* $\mathbf{R}_\sigma^{\alpha,n}f = \mathbf{R}^{\alpha,n}(f\sigma)$ *for any smooth truncation of* $\mathbf{R}^{\alpha,n}$*. Let* $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ *be a globally biLipschitz map. Then the operator* $\mathbf{R}_\sigma^{\alpha,n}$ *is bounded from* $L^2(\sigma)$ *to* $L^2(\omega)$*, i.e.*

$$\|\mathbf{R}_\sigma^\alpha f\|_{L^2(\omega)} \le \mathfrak{N}_{\mathbf{R}_\sigma^\alpha} \|f\|_{L^2(\sigma)} ,$$

uniformly in smooth truncations of $\mathbf{R}^{\alpha,n}$*, if and only if the Muckenhoupt conditions hold, the testing conditions hold and the weak boundedness property holds. Moreover, we have the equivalence*

$$\mathfrak{N}_{\mathbf{R}_\sigma^{\alpha,n}} \approx \sqrt{\mathcal{A}_2^\alpha + \mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\text{punct}} + A_2^{\alpha,*,\text{punct}}} + \mathfrak{T}_{\mathbf{R}^{\alpha,n}} + \mathfrak{T}_{\mathbf{R}^{\alpha,n}}^* + \mathcal{WBP}_{\mathbf{R}^{\alpha,n}}.$$

The case $k = n - 1$ of k-energy dispersed is similar to the notion of uniformly full dimension introduced by Lacey and Wick in [18, versions 2 and 3]. The proof of Theorem 2.11 shows that we can also take ω and σ to be k_1 and k_2 energy dispersed respectively, provided α satisfies the hypotheses with respect to *both* k_1 and k_2.

Proof of Theorem 2.8

We now give the proof of Theorem 2.8 in the following sections. Sections "Monotonicity Lemma and Energy Lemma", "Corona Decompositions and Splittings" and "Energy Dispersed Measures" are largely taken verbatim from the corresponding sections of [28], but are included here since their omission here would hinder the readability of an already complicated argument.

Good Quasicubes and Energy Muckenhoupt Conditions

First we extend the notion of goodness to quasicubes.

Definition 3.1 Let $\mathbf{r} \in \mathbb{N}$ and $0 < \varepsilon < 1$. Fix a quasigrid $\Omega\mathcal{D}$. A dyadic quasicube J is $(\mathbf{r}, \varepsilon)$-*good*, or simply *good*, if for *every* dyadic superquasicube I, it is the case that **either** J has side length greater than $2^{-\mathbf{r}}$ times that of I, **or** $J \Subset_{\mathbf{r},\varepsilon} I$ is $(\mathbf{r}, \varepsilon)$-deeply embedded in I.

Note that this definition simply asserts that a dyadic quasicube $J = \Omega J'$ is $(\mathbf{r}, \varepsilon)$-good if and only if the cube J' is $(\mathbf{r}, \varepsilon)$-good in the usual sense. Finally, we say that J is $\mathbf{r}$-*nearby* in K when $J \subset K$ and

$$\ell(J) > 2^{-\mathbf{r}}\ell(K).$$

The parameters $\mathbf{r}, \varepsilon$ will be fixed sufficiently large and small respectively later in the proof, and we denote the set of such good dyadic quasicubes by $\Omega\mathcal{D}_{(\mathbf{r},\varepsilon)-\text{good}}$, or simply $\Omega\mathcal{D}_{\text{good}}$ when the goodness parameters $(\mathbf{r}, \varepsilon)$ are understood. Note that if $J' \in \Omega\mathcal{D}_{(\mathbf{r},\varepsilon)-\text{good}}$ and if $J' \subset K \in \Omega\mathcal{D}$, then **either** J' is $\mathbf{r}$-*nearby* in K **or** $J' \subset J \Subset_{\mathbf{r},\varepsilon} K$.

Throughout the proof, it will be convenient to also consider pairs of quasicubes J, K where J is $(\boldsymbol{\rho}, \varepsilon)$-*deeply embedded* in K, written $J \Subset_{\boldsymbol{\rho},\varepsilon} K$ and meaning (10) holds with the same $\varepsilon > 0$ but with $\boldsymbol{\rho}$ in place of $\mathbf{r}$; as well as pairs of quasicubes J, K where J is $\boldsymbol{\rho}$-*nearby* in K, $\ell(J) > 2^{-\boldsymbol{\rho}}\ell(K)$, for a parameter $\boldsymbol{\rho} \gg \mathbf{r}$ that will be fixed later.

Notation 2 *We will typically use the side length $\ell(J)$ of a Ω-quasicube when we are describing collections of quasicubes, and when we want $\ell(J)$ to be a dyadic or related number; while we will typically use $|J|^{\frac{1}{n}}$ in estimates, and when we want to compare powers of volumes of quasicubes. We will continue to use the prefix 'quasi' when discussing quasicubes, quasiHaar, quasienergy and quasidistance in the text, but will not use the prefix 'quasi' when discussing other notions. In particular, since* $\text{quasi}\mathcal{A}_2^\alpha + \text{quasi}A_2^{\alpha,\text{punct}} \approx \mathcal{A}_2^\alpha + A_2^{\alpha,\text{punct}}$ *(see e.g. [31] for a proof) we do not use* quasi *as a prefix for the Muckenhoupt conditions, even though* $\text{quasi}\mathcal{A}_2^\alpha$ *alone is not comparable to $\mathcal{A}_2^\alpha$. Finally, we will not modify any mathematical symbols to reflect quasinotions, except for using $\Omega\mathcal{D}$ to denote a quasigrid, and* $\text{qdist}(E, F) \equiv \text{dist}\left(\Omega^{-1}E, \Omega^{-1}F\right)$ *to denote quasidistance between sets E and F, and using* $|x-y|_{\text{qdist}} \equiv \left|\Omega^{-1}x - \Omega^{-1}y\right|$ *to denote quasidistance between points x and y. This limited use of quasi in the text serves mainly to remind the reader we are working entirely in the 'quasiworld'.*

Energy Muckenhoupt Conditions

We now show that the punctured Muckenhoupt conditions $A_2^{\alpha,\text{punct}}$ and $A_2^{\alpha,*,\text{punct}}$ control respectively the 'energy A_2^α conditions', denoted $A_2^{\alpha,\text{energy}}$ and $A_2^{\alpha,*,\text{energy}}$ where

$$A_2^{\alpha,\text{energy}}(\sigma, \omega) \equiv \sup_{Q \in \Omega\mathcal{P}^n} \frac{\left\| \mathsf{P}_Q^\omega \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\omega)}^2}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}}, \tag{14}$$

$$A_2^{\alpha,*,\text{energy}}(\sigma, \omega) \equiv \sup_{Q \in \Omega\mathcal{P}^n} \frac{|Q|_\omega}{|Q|^{1-\frac{\alpha}{n}}} \frac{\left\| \mathsf{P}_Q^\sigma \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\sigma)}^2}{|Q|^{1-\frac{\alpha}{n}}}.$$

These energy A_2^α conditions play a critical role in controlling local parts of functional energy later in the paper, and it is a crucial requirement that they are necessary conditions, as shown by the next lemma.

Lemma 3.2 *For any positive locally finite Borel measures σ, ω we have*

$$A_2^{\alpha,\text{energy}}(\sigma,\omega) \leq \max\{n,3\} A_2^{\alpha,\text{punct}}(\sigma,\omega),$$
$$A_2^{\alpha,*,\text{energy}}(\sigma,\omega) \leq \max\{n,3\} A_2^{\alpha,*,\text{punct}}(\sigma,\omega).$$

Proof Fix a quasicube $Q \in \Omega\mathcal{D}$. If $\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right) \geq \frac{1}{2}|Q|_\omega$, then we trivially have

$$\begin{aligned}\frac{\left\|\mathsf{P}_Q^\omega \frac{\mathbf{x}}{\ell(Q)}\right\|_{L^2(\omega)}^2}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}} &\leq n \frac{|Q|_\omega}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}} \\ &\leq 2n \frac{\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right)}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_\sigma}{|Q|^{1-\frac{\alpha}{n}}} \leq 2n A_2^{\alpha,\text{punct}}(\sigma,\omega).\end{aligned}$$

On the other hand, if $\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right) < \frac{1}{2}|Q|_\omega$ then there is a point $p \in Q \cap \mathfrak{P}_{(\sigma,\omega)}$ such that

$$\omega(\{p\}) > \frac{1}{2}|Q|_\omega,$$

and consequently, p is the largest ω-point mass in Q. Thus if we define $\widetilde{\omega} = \omega - \omega(\{p\})\,\delta_p$, then we have

$$\omega\left(Q, \mathfrak{P}_{(\sigma,\omega)}\right) = |Q|_{\widetilde{\omega}}.$$

Now we observe from the construction of Haar projections that

$$\triangle_J^{\widetilde{\omega}} = \triangle_J^\omega, \qquad \text{for all } J \in \Omega\mathcal{D} \text{ with } p \notin J.$$

So for each $s \geq 0$ there is a unique quasicube $J_s \in \Omega\mathcal{D}$ with $\ell(J_s) = 2^{-s}\ell(Q)$ that contains the point p. For this quasicube we have, if $\left\{h_J^{\omega,a}\right\}_{J\in\Omega\mathcal{D},\ a\in\Gamma_n}$ is a basis for $L^2(\omega)$,

$$\begin{aligned}\left\|\triangle_{J_s}^\omega \mathbf{x}\right\|_{L^2(\omega)}^2 &= \sum_{a\in\Gamma_n}\left|\left\langle h_{J_s}^{\omega,a}, x\right\rangle_\omega\right|^2 = \sum_{a\in\Gamma_n}\left|\left\langle h_{J_s}^{\omega,a}, x-p\right\rangle_\omega\right|^2 \\ &= \sum_{a\in\Gamma_n}\left|\int_{J_s} h_{J_s}^{\omega,a}(x)\,(x-p)\,d\omega(x)\right|^2 = \sum_{a\in\Gamma_n}\left|\int_{J_s} h_{J_s}^{\omega,a}(x)\,(x-p)\,d\widetilde{\omega}(x)\right|^2 \\ &\leq \sum_{a\in\Gamma_n}\left\|h_{J_s}^{\omega,a}\right\|_{L^2(\widetilde{\omega})}^2 \left\|\mathbf{1}_{J_s}(x)\,(x-p)\right\|_{L^2(\widetilde{\omega})}^2 \leq \sum_{a\in\Gamma_n}\left\|h_{J_s}^{\omega,a}\right\|_{L^2(\omega)}^2 \left\|\mathbf{1}_{J_s}(x)\,(x-p)\right\|_{L^2(\widetilde{\omega})}^2 \\ &\leq n2^n\ell(J_s)^2\,|J_s|_{\widetilde{\omega}} \leq 2^{-2s}\ell(Q)^2\,|Q|_{\widetilde{\omega}}.\end{aligned}$$

Thus we can estimate

$$\begin{aligned}
\left\| \mathsf{P}_Q^{\omega} \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\omega)}^2 &= \frac{1}{\ell(Q)^2} \sum_{J \in \Omega\mathcal{D}:\ J \subset Q} \left\| \triangle_J^{\omega} \mathbf{x} \right\|_{L^2(\omega)}^2 \\
&= \frac{1}{\ell(Q)^2} \left(\sum_{J \in \Omega\mathcal{D}:\ p \notin J \subset Q} \left\| \triangle_J^{\widetilde{\omega}} \mathbf{x} \right\|_{L^2(\widetilde{\omega})}^2 + \sum_{s=0}^{\infty} \left\| \triangle_{J_s}^{\omega} \mathbf{x} \right\|_{L^2(\omega)}^2 \right) \\
&\le \frac{1}{\ell(Q)^2} \left(\left\| \mathsf{P}_Q^{\widetilde{\omega}} \mathbf{x} \right\|_{L^2(\widetilde{\omega})}^2 + \sum_{s=0}^{\infty} 2^{-2s} \ell(Q)^2 \left| Q \right|_{\widetilde{\omega}} \right) \\
&\le \frac{1}{\ell(Q)^2} \left(\ell(Q)^2 \left| Q \right|_{\widetilde{\omega}} + \sum_{s=0}^{\infty} 2^{-2s} \ell(Q)^2 \left| Q \right|_{\widetilde{\omega}} \right) \\
&\le 3 \left| Q \right|_{\widetilde{\omega}} \le 3\omega \left(Q, \mathfrak{P}_{(\sigma,\omega)} \right),
\end{aligned}$$

and so

$$\frac{\left\| \mathsf{P}_Q^{\omega} \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\omega)}^2}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_{\sigma}}{|Q|^{1-\frac{\alpha}{n}}} \le \frac{3\omega \left(Q, \mathfrak{P}_{(\sigma,\omega)} \right)}{|Q|^{1-\frac{\alpha}{n}}} \frac{|Q|_{\sigma}}{|Q|^{1-\frac{\alpha}{n}}} \le 3A_2^{\alpha,\mathrm{punct}}(\sigma,\omega).$$

Now take the supremum over $Q \in \Omega\mathcal{D}$ to obtain $A_2^{\alpha,\mathrm{energy}}(\sigma,\omega) \le \max\{n,3\} A_2^{\alpha,\mathrm{punct}}(\sigma,\omega)$. The dual inequality follows upon interchanging the measures σ and ω. □

Plugged $\mathcal{A}_2^{\alpha,\mathrm{energyplug}}$ Conditions

Using Lemma 3.2 we can control the 'plugged' energy $\mathcal{A}_2^{\alpha}$ conditions:

$$\mathcal{A}_2^{\alpha,\mathrm{energyplug}}(\sigma,\omega) \equiv \sup_{Q \in \Omega\mathcal{P}^n} \frac{\left\| \mathsf{P}_Q^{\omega} \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\omega)}^2}{|Q|^{1-\frac{\alpha}{n}}} \mathcal{P}^{\alpha}(Q,\sigma),$$

$$\mathcal{A}_2^{\alpha,*,\mathrm{energyplug}}(\sigma,\omega) \equiv \sup_{Q \in \Omega\mathcal{P}^n} \mathcal{P}^{\alpha}(Q,\omega) \frac{\left\| \mathsf{P}_Q^{\sigma} \frac{\mathbf{x}}{\ell(Q)} \right\|_{L^2(\sigma)}^2}{|Q|^{1-\frac{\alpha}{n}}}.$$

Lemma 3.3 *We have*

$$\mathcal{A}_2^{\alpha,\mathrm{energyplug}}(\sigma,\omega) \lesssim \mathcal{A}_2^{\alpha}(\sigma,\omega) + A_2^{\alpha,\mathrm{energy}}(\sigma,\omega),$$

$$\mathcal{A}_2^{\alpha,*,\mathrm{energyplug}}(\sigma,\omega) \lesssim \mathcal{A}_2^{\alpha,*}(\sigma,\omega) + A_2^{\alpha,*,\mathrm{energy}}(\sigma,\omega).$$

Proof We have

$$\begin{aligned}
\frac{\left\|\mathsf{P}^{\omega}_{Q}\frac{\mathbf{x}}{\ell(Q)}\right\|^{2}_{L^{2}(\omega)}}{|Q|^{1-\frac{\alpha}{n}}}\mathcal{P}^{\alpha}(Q,\sigma) &= \frac{\left\|\mathsf{P}^{\omega}_{Q}\frac{\mathbf{x}}{\ell(Q)}\right\|^{2}_{L^{2}(\omega)}}{|Q|^{1-\frac{\alpha}{n}}}\mathcal{P}^{\alpha}\left(Q,\mathbf{1}_{Q^{c}}\sigma\right)+\frac{\left\|\mathsf{P}^{\omega}_{Q}\frac{\mathbf{x}}{\ell(Q)}\right\|^{2}_{L^{2}(\omega)}}{|Q|^{1-\frac{\alpha}{n}}}\mathcal{P}^{\alpha}\left(Q,\mathbf{1}_{Q}\sigma\right)\\
&\lesssim \frac{|Q|_{\omega}}{|Q|^{1-\frac{\alpha}{n}}}\mathcal{P}^{\alpha}\left(Q,\mathbf{1}_{Q^{c}}\sigma\right)+\frac{\left\|\mathsf{P}^{\omega}_{Q}\frac{\mathbf{x}}{\ell(Q)}\right\|^{2}_{L^{2}(\omega)}}{|Q|^{1-\frac{\alpha}{n}}}\frac{|Q|_{\sigma}}{|Q|^{1-\frac{\alpha}{n}}}\\
&\lesssim \mathcal{A}^{\alpha}_{2}(\sigma,\omega)+A^{\alpha,\text{energy}}_{2}(\sigma,\omega)\,.
\end{aligned}$$

□

Random Grids and Shifted Grids

Using the analogue for dyadic quasigrids of the good random grids of Nazarov, Treil and Volberg, a standard argument of NTV, see e.g. [35], reduces the two weight inequality (12) for T^{α} to proving boundedness of a bilinear form $\mathcal{T}^{\alpha}(f,g)$ with uniform constants over dyadic quasigrids, and where the quasiHaar supports $\operatorname{supp}\widehat{f}$ and $\operatorname{supp}\widehat{g}$ of the functions f and g are contained in the collection $\Omega\mathcal{D}^{\text{good}}$ of good quasicubes, whose children are all good as well, with goodness parameters $\mathbf{r}<\infty$ and $\varepsilon>0$ chosen sufficiently large and small respectively depending only on n and α. Here the quasiHaar support of f is $\operatorname{supp}\widehat{f}\equiv\left\{I\in\Omega\mathcal{D}:\triangle^{\sigma}_{I}f\neq 0\right\}$, and similarly for g. In fact we can assume even more, namely that the quasiHaar supports $\operatorname{supp}\widehat{f}$ and $\operatorname{supp}\widehat{g}$ of f and g are contained in the collection of $\boldsymbol{\tau}$-*good* quasicubes

$$\Omega\mathcal{D}^{\boldsymbol{\tau}}_{(\mathbf{r},\varepsilon)-\text{good}}\equiv\left\{K\in\Omega\mathcal{D}:\mathfrak{C}_{K}\subset\Omega\mathcal{D}_{(\mathbf{r},\varepsilon)-\text{good}}\text{ and }\pi^{\ell}_{\Omega\mathcal{D}}K\in\Omega\mathcal{D}_{(\mathbf{r},\varepsilon)-\text{good}}\text{ for all }0\leq\ell\leq\boldsymbol{\tau}\right\}, \tag{15}$$

that are $(\mathbf{r},\varepsilon)$-good and whose children are also $(\mathbf{r},\varepsilon)$-good, and whose ℓ-parents up to level $\boldsymbol{\tau}$ are also $(\mathbf{r},\varepsilon)$-good. Here $\boldsymbol{\tau}>\mathbf{r}$ is a parameter to be fixed later. We may assume this restriction on the quasiHaar supports of f and g by the following lemma. See [29] for a proof.[8]

Lemma 3.4 *Given* $\mathbf{r}\geq 3$, $\boldsymbol{\tau}\geq 1$ *and* $\frac{1}{\mathbf{r}}<\varepsilon<1-\frac{1}{\mathbf{r}}$, *we have*

$$\Omega\mathcal{D}_{(\mathbf{r}-1,\delta)-\text{good}}\subset\Omega\mathcal{D}^{\boldsymbol{\tau}}_{(\mathbf{r},\varepsilon)-\text{good}},$$

provided

$$0<\delta\leq\frac{\mathbf{r}\varepsilon-1}{\mathbf{r}+\boldsymbol{\tau}}. \tag{16}$$

[8]This lemma is misstated in [30].

For convenience in notation we will sometimes suppress the dependence on α in our nonlinear forms, but will retain it in the operators, Poisson integrals and constants. More precisely, let $\Omega\mathcal{D}^{\sigma} = \Omega\mathcal{D}^{\omega}$ be an $(\mathbf{r},\varepsilon)$-good quasigrid on $\mathbb{R}^n$, and let $\left\{h_I^{\sigma,a}\right\}_{I\in\Omega\mathcal{D}^{\sigma},\, a\in\Gamma_n}$ and $\left\{h_J^{\omega,b}\right\}_{J\in\Omega\mathcal{D}^{\omega},\, b\in\Gamma_n}$ be corresponding quasiHaar bases as described above, so that

$$f=\sum_{I\in\Omega\mathcal{D}^{\sigma}}\triangle_I^{\sigma}f=\sum_{I\in\Omega\mathcal{D}^{\sigma},\, a\in\Gamma_n}\left\langle f,h_I^{\sigma,a}\right\rangle h_I^{\sigma,a}=\sum_{I\in\Omega\mathcal{D}^{\sigma},\, a\in\Gamma_n}\widehat{f}(I;a)\ h_I^{\sigma,a},$$

$$g=\sum_{J\in\Omega\mathcal{D}^{\omega}}\triangle_J^{\omega}g=\sum_{J\in\Omega\mathcal{D}^{\omega},\, b\in\Gamma_n}\left\langle g,h_J^{\omega,b}\right\rangle h_J^{\omega,b}=\sum_{J\in\Omega\mathcal{D}^{\omega},\, b\in\Gamma_n}\widehat{g}(J;b)\ h_J^{\omega,b},$$

where the appropriate measure is understood in the notation $\widehat{f}(I;a)$ and $\widehat{g}(J;b)$, and where these quasiHaar coefficients $\widehat{f}(I;a)$ and $\widehat{g}(J;b)$ vanish if the quasicubes I and J are not good. Inequality (12) is equivalent to boundedness of the bilinear form

$$\mathcal{T}^{\alpha}(f,g)\equiv\left\langle T_{\sigma}^{\alpha}(f),g\right\rangle_{\omega}=\sum_{I\in\Omega\mathcal{D}^{\sigma}\text{ and }J\in\Omega\mathcal{D}^{\omega}}\left\langle T_{\sigma}^{\alpha}\left(\triangle_I^{\sigma}f\right),\triangle_J^{\omega}g\right\rangle_{\omega}$$

on $L^2(\sigma)\times L^2(\omega)$, i.e.

$$\left|\mathcal{T}^{\alpha}(f,g)\right|\leq\mathfrak{N}_{T^{\alpha}}\left\|f\right\|_{L^2(\sigma)}\left\|g\right\|_{L^2(\omega)},\tag{17}$$

uniformly over all quasigrids and appropriate truncations. We may assume the two quasigrids $\Omega\mathcal{D}^{\sigma}$ and $\Omega\mathcal{D}^{\omega}$ are equal here, and this we will do throughout the paper, although we sometimes continue to use the measure as a superscript on $\Omega\mathcal{D}$ for clarity of exposition. Roughly speaking, we analyze the form $\mathcal{T}^{\alpha}(f,g)$ by splitting it in a nonlinear way into three main pieces, following in part the approach in [12] and [13]. The first piece consists of quasicubes I and J that are either disjoint or of comparable side length, and this piece is handled using the section on preliminaries of NTV type. The second piece consists of quasicubes I and J that overlap, but are 'far apart' in a nonlinear way, and this piece is handled using the sections on the Intertwining Proposition and the control of the functional quasienergy condition by the quasienergy condition. Finally, the remaining local piece where the overlapping quasicubes are 'close' is handled by generalizing methods of NTV as in [11], and then splitting the stopping form into two sublinear stopping forms, one of which is handled using techniques of [16], and the other using the stopping time and recursion of M. Lacey [8]. See the schematic diagram in section "Doubly Iterated Coronas and the NTV Quasicube Size Splitting" below.

We summarize our assumptions on the Haar supports of f and g, and on the dyadic quasigrids $\Omega\mathcal{D}$.

Condition 1 (on Haar supports and quasigrids) *We suppose the quasiHaar supports of the functions f and g satisfy* $\mathrm{supp}\widehat{f}, \mathrm{supp}\widehat{g} \subset \Omega\mathcal{D}^{\tau}_{(\mathbf{r},\varepsilon)-\mathrm{good}}$. *We also assume that $|\partial Q|_{\sigma+\omega} = 0$ for all dyadic quasicubes Q in the grids $\Omega\mathcal{D}$ (since this property holds with probability* 1 *for random grids $\Omega\mathcal{D}$).*

Necessity of the $\mathcal{A}_2^{\alpha}$ Conditions

Here we consider in particular the necessity of the fractional $\mathcal{A}_2^{\alpha}$ condition (with holes) when $0 \leq \alpha < n$, for the boundedness from $L^2(\sigma)$ to $L^2(\omega)$ (where σ and ω may have common point masses) of the α-fractional Riesz vector transform $\mathbf{R}^{\alpha}$ defined by

$$\mathbf{R}^{\alpha}(f\sigma)(x) = \int_{\mathbb{R}^n} K_j^{\alpha}(x,y) f(y)\, d\sigma(y), \qquad K_j^{\alpha}(x,y) = \frac{x^j - y^j}{|x-y|^{n+1-\alpha}},$$

whose kernel $K_j^{\alpha}(x,y)$ satisfies (5) for $0 \leq \alpha < n$. More generally, necessity holds for elliptic operators as in the next lemma. See [30] for the easier proof in the case without holes.

Lemma 4.1 *Suppose $0 \leq \alpha < n$. Let T^{α} be any collection of operators with α-standard fractional kernel satisfying the ellipticity condition* (13)*, and in the case $\frac{n}{2} \leq \alpha < n$, we also assume the more restrictive condition* (11)*. Then for $0 \leq \alpha < n$ we have*

$$\sqrt{\mathcal{A}_2^{\alpha}} \lesssim \mathfrak{N}_{\alpha}(T^{\alpha}).$$

Proof First we give the proof for the case when T^{α} is the α-fractional Riesz transform $\mathbf{R}^{\alpha}$, whose kernel is $\mathbf{K}^{\alpha}(x,y) = \frac{x-y}{|x-y|^{n+1-\alpha}}$. Define the 2^n generalized n-ants $\mathcal{Q}_m$ for $m \in \{-1,1\}^n$, and their translates $\mathcal{Q}_m(w)$ for $w \in \mathbb{R}^n$ by

$$\mathcal{Q}_m = \{(x_1,\ldots,x_n) : m_k x_k > 0\}, \quad \mathcal{Q}_m(w) = \{z : z - w \in \mathcal{Q}_m\}, \qquad w \in \mathbb{R}^n.$$

Fix $m \in \{-1,1\}^n$ and a quasicube I. For $a \in \mathbb{R}^n$ and $r > 0$ let

$$s_I(x) = \frac{\ell(I)}{\ell(I) + |x - \zeta_I|}, \qquad f_{a,r}(y) = \mathbf{1}_{\mathcal{Q}_{-m}(a)\cap B(0,r)}(y)\, s_I(y)^{n-\alpha},$$

where ζ_I is the center of the cube I. Now

$$\ell(I)\,|x-y| \leq \ell(I)\,|x-\zeta_I| + \ell(I)\,|\zeta_I - y| \leq [\ell(I) + |x-\zeta_I|]\,[\ell(I) + |\zeta_I - y|]$$

implies

$$\frac{1}{|x-y|} \geq \frac{1}{\ell(I)} s_I(x) s_I(y), \qquad x, y \in \mathbb{R}^n.$$

Now the key observation is that with $L\zeta \equiv m \cdot \zeta$, we have

$$L(x-y) = m \cdot (x-y) \geq |x-y|, \qquad x \in \mathcal{Q}_m(y),$$

which yields

$$L(\mathbf{K}^\alpha(x,y)) = \frac{L(x-y)}{|x-y|^{n+1-\alpha}} \geq \frac{1}{|x-y|^{n-\alpha}} \geq \ell(I)^{\alpha-n} s_I(x)^{n-\alpha} s_I(y)^{n-\alpha}, \quad (18)$$

provided $x \in \mathcal{Q}_m(y)$. Now we note that $x \in \mathcal{Q}_m(y)$ when $x \in \mathcal{Q}_m(a)$ and $y \in \mathcal{Q}_{-m}(a)$ to obtain that for $x \in \mathcal{Q}_m(a)$,

$$\begin{aligned} L(T^\alpha(f_{a,r}\sigma)(x)) &= \int_{\mathcal{Q}_{-m}(a)\cap B(0,r)} \frac{L(x-y)}{|x-y|^{n+1-\alpha}} s_I(y)\, d\sigma(y) \\ &\geq \ell(I)^{\alpha-n} s_I(x)^{n-\alpha} \int_{\mathcal{Q}_{-m}(a)\cap B(0,r)} s_I(y)^{2n-2\alpha}\, d\sigma(y). \end{aligned}$$

Applying $|L\zeta| \leq \sqrt{n}\,|\zeta|$ and our assumed two weight inequality for the fractional Riesz transform, we see that for $r > 0$ large,

$$\begin{aligned} &\ell(I)^{2\alpha-2n} \int_{\mathcal{Q}_m(a)} s_I(x)^{2n-2\alpha} \left(\int_{\mathcal{Q}_{-m}(a)\cap B(0,r)} s_I(y)^{2n-2\alpha}\, d\sigma(y) \right)^2 d\omega(x) \\ &\leq \|LT(\sigma f_{a,r})\|^2_{L^2(\omega)} \lesssim \mathfrak{N}_\alpha(\mathbf{R}^\alpha)^2 \|f_{a,r}\|^2_{L^2(\sigma)} = \mathfrak{N}_\alpha(\mathbf{R}^\alpha)^2 \int_{\mathcal{Q}_{-m}(a)\cap B(0,r)} s_I(y)^{2n-2\alpha}\, d\sigma(y). \end{aligned}$$

Rearranging the last inequality, and upon letting $r \to \infty$, we obtain

$$\int_{\mathcal{Q}_m(a)} \frac{\ell(I)^{n-\alpha}}{(\ell(I) + |x-\zeta_I|)^{2n-2\alpha}} d\omega(x) \int_{\mathcal{Q}_{-m}(a)} \frac{\ell(I)^{n-\alpha}}{(\ell(I) + |y-\zeta_I|)^{2n-2\alpha}} d\sigma(y) \lesssim \mathfrak{N}_\alpha(\mathbf{R}^\alpha)^2.$$

Note that the ranges of integration above are pairs of opposing n-ants.

Fix a quasicube Q, which without loss of generality can be taken to be centered at the origin, $\zeta_Q = 0$. Then choose $a = (2\ell(Q), 2\ell(Q))$ and $I = Q$ so that we have

$$\begin{aligned} &\left(\int_{\mathcal{Q}_m(a)} \frac{\ell(Q)^{n-\alpha}}{(\ell(Q) + |x|)^{2n-2\alpha}} d\omega(x) \right) \left(\ell(Q)^{\alpha-n} \int_Q d\sigma \right) \\ &\leq C_\alpha \int_{\mathcal{Q}_m(a)} \frac{\ell(Q)^{n-\alpha}}{(\ell(Q) + |x|)^{2n-2\alpha}} d\omega(x) \int_{\mathcal{Q}_{-m}(a)} \frac{\ell(Q)^{n-\alpha}}{(\ell(Q) + |y|)^{2n-2\alpha}} d\sigma(y) \lesssim \mathfrak{N}_\alpha(\mathbf{R}^\alpha)^2. \end{aligned}$$

Now fix $m = (1, 1, \ldots, 1)$ and note that there is a fixed N (independent of $\ell(Q)$) and a fixed collection of rotations $\{\rho_k\}_{k=1}^N$, such that the rotates $\rho_k \mathcal{Q}_m(a)$, $1 \le k \le N$, of the n-ant $\mathcal{Q}_m(a)$ cover the complement of the ball $B\left(0, 4\sqrt{n}\ell(Q)\right)$:

$$B\left(0, 4\sqrt{n}\ell(Q)\right)^c \subset \bigcup_{k=1}^{N} \rho_k \mathcal{Q}_m(a).$$

Then we obtain, upon applying the same argument to these rotated pairs of n-ants,

$$\left(\int_{B\left(0,4\sqrt{n}\ell(Q)\right)^c} \frac{\ell(Q)^{n-\alpha}}{\left(\ell(Q)+|x|\right)^{2n-2\alpha}} d\omega(x)\right)\left(\ell(Q)^{\alpha-n}\int_Q d\sigma\right) \lesssim \mathfrak{N}_\alpha\left(\mathbf{R}^\alpha\right)^2. \quad (19)$$

Now we assume for the moment the offset A_2^α condition

$$\ell(Q)^{2(\alpha-n)}\left(\int_{Q'} d\omega\right)\left(\int_Q d\sigma\right) \le A_2^\alpha,$$

where Q' and Q are neighbouring quasicubes, i.e. $(Q', Q) \in \Omega\mathcal{N}^n$. If we use this offset inequality with Q' ranging over $3Q \setminus Q$, and then use the separation of $B\left(0, 4\sqrt{n}\ell(Q)\right) \setminus 3Q$ and Q to obtain the inequality

$$\ell(Q)^{2(\alpha-n)}\left(\int_{B\left(0,4\sqrt{n}\ell(Q)\right)\setminus 3Q} d\omega\right)\left(\int_Q d\sigma\right) \lesssim A_2^\alpha,$$

together with (19), we obtain

$$\left(\int_{\mathbb{R}^n\setminus Q} \frac{\ell(Q)^{n-\alpha}}{\left(\ell(Q)+|x|\right)^{2n-2\alpha}} d\omega(x)\right)^{\frac{1}{2}}\left(\ell(Q)^{\alpha-n}\int_Q d\sigma\right)^{\frac{1}{2}} \lesssim \mathfrak{N}_\alpha\left(\mathbf{R}^\alpha\right) + \sqrt{A_2^\alpha}.$$

Clearly we can reverse the roles of the measures ω and σ and obtain

$$\sqrt{\mathcal{A}_2^{\alpha,*}} \lesssim \mathfrak{N}_\alpha\left(\mathbf{R}^\alpha\right) + \sqrt{A_2^\alpha}$$

for the kernels $\mathbf{K}^\alpha$, $0 \le \alpha < n$.

More generally, to obtain the case when T^α is elliptic and the offset A_2^α condition holds, we note that the key estimate (18) above extends to the kernel $\sum_{j=1}^J \lambda_j^m K_j^\alpha$ of $\sum_{j=1}^J \lambda_j^m T_j^\alpha$ in (11) if the n-ants above are replaced by thin cones of sufficently small aperture, and there is in addition sufficient separation between opposing cones, which in turn may require a larger constant than $4\sqrt{n}$ in the choice of Q' above.

Finally, we turn to showing that the offset A_2^α condition is implied by the norm inequality, i.e.

$$\sqrt{A_2^\alpha} \equiv \sup_{(Q',Q)\in\Omega\mathcal{N}^n} \ell(Q)^\alpha \left(\frac{1}{|Q'|}\int_{Q'} d\omega\right)^{\frac{1}{2}} \left(\frac{1}{|Q|}\int_{Q} d\sigma\right)^{\frac{1}{2}} \lesssim \mathfrak{N}_\alpha(\mathbf{R}^\alpha);$$

$$\text{i.e. } \left(\int_{Q'} d\omega\right)\left(\int_{Q} d\sigma\right) \lesssim \mathfrak{N}_\alpha(\mathbf{R}^\alpha)^2 |Q|^{2-\frac{2\alpha}{n}}, \quad (Q',Q)\in\Omega\mathcal{N}^n.$$

In the range $0 \le \alpha < \frac{n}{2}$ where we only assume (13), we adapt a corresponding argument from [15].

The 'one weight' argument on page 211 of Stein [34] yields the *asymmetric* two weight A_2^α condition

$$\left|Q'\right|_\omega |Q|_\sigma \le C\mathfrak{N}_\alpha(\mathbf{R}^\alpha)|Q|^{2\left(1-\frac{\alpha}{n}\right)}, \tag{20}$$

where Q and Q' are quasicubes of equal side length r and distance $C_0 r$ apart for some (fixed large) positive constant C_0 (for this argument we choose the unit vector $\mathbf{u}$ in (13) to point in the direction from Q to Q'). In the one weight case treated in [34] it is easy to obtain from this (even for a *single* direction $\mathbf{u}$) the usual (symmetric) A_2 condition. Here we will have to employ a different approach.

Now recall (see Sec 2 of [24] for the case of usual cubes, and the case of half open, half closed quasicubes here is no different) that given an open subset Φ of $\mathbb{R}^n$, we can choose $R \ge 3$ sufficiently large, depending only on the dimension, such that if $\left\{Q_j^k\right\}_j$ are the dyadic quasicubes maximal among those dyadic quasicubes Q satisfying $RQ \subset \Phi$, then the following properties hold:

$$\begin{cases} \text{(disjoint cover)} & \Phi = \bigcup_j Q_j \text{ and } Q_j \cap Q_i = \emptyset \text{ if } i \ne j \\ \text{(Whitney condition)} & RQ_j \subset \Phi \text{ and } 3RQ_j \cap \Phi^c \ne \emptyset \text{ for all } j \\ \text{(finite overlap)} & \sum_j \chi_{3Q_j} \le C\chi_\Phi \end{cases}. \tag{21}$$

So fix a pair of neighbouring quasicubes $\left(Q_0', Q_0\right) \in \Omega\mathcal{N}^n$, and let $\{\mathsf{Q}_i\}_i$ be a Whitney decomposition into quasicubes of the set $\Phi \equiv \left(Q_0' \times Q_0\right) \setminus \mathfrak{D}$ relative to the diagonal $\mathfrak{D}$ in $\mathbb{R}^n \times \mathbb{R}^n$. Of course, there are no common point masses of ω in Q_0' and σ in Q_0 since the quasicubes Q_0' and Q_0 are disjoint. Note that if $\mathsf{Q}_i = Q_i' \times Q_i$, then (20) can be written

$$|\mathsf{Q}_i|_{\omega\times\sigma} \le C\mathfrak{N}_\alpha(\mathbf{R}^\alpha)|\mathsf{Q}_i|^{1-\frac{\alpha}{n}}, \tag{22}$$

where $\omega \times \sigma$ denotes product measure on $\mathbb{R}^n \times \mathbb{R}^n$. We choose R sufficiently large in the Whitney decomposition (21), depending on C_0, such that (22) holds for all the Whitney quasicubes Q_i. We have $\sum_i |\mathsf{Q}_i| = |Q' \times Q| = |Q|^2$.

Moreover, if $\mathsf{R} = Q' \times Q$ is a rectangle in $\mathbb{R}^n \times \mathbb{R}^n$ (i.e. Q', Q are quasicubes in $\mathbb{R}^n$), and if $\mathsf{R} = \dot{\cup}_i \mathsf{R}_i$ is a finite disjoint union of rectangles R_α, then by additivity of the product measure $\omega \times \sigma$,

$$|\mathsf{R}|_{\omega\times\sigma} = \sum_i |\mathsf{R}_i|_{\omega\times\sigma} \,.$$

Let $\mathsf{Q}_0 = Q_0' \times Q_0$ and set

$$\Lambda \equiv \left\{\mathsf{Q} = Q' \times Q : \mathsf{Q} \subset \mathsf{Q}_0, \ell(Q) = \ell\left(Q'\right) \approx C_0^{-1} \text{qdist}\left(Q, Q'\right) \text{ and (20) holds}\right\}.$$

Divide Q_0 into $2n \times 2n = 4n^2$ congruent subquasicubes $\mathsf{Q}_0^1, \ldots, \mathsf{Q}_0^{4^n}$ of side length $\frac{1}{2}$, and set aside those $\mathsf{Q}_0^j \in \Lambda$ (those for which (20) holds) into a collection of stopping cubes Γ. Continue to divide the remaining $\mathsf{Q}_0^j \in \Lambda$ of side length $\frac{1}{4}$, and again, set aside those $\mathsf{Q}_0^{j,i} \in \Phi$ into Γ, and continue subdividing those that remain. We continue with such subdivisions for N generations so that all the cubes *not* set aside into Γ have side length 2^{-N}. The important property these latter cubes have is that they all lie within distance $r2^{-N}$ of the diagonal $\mathfrak{D} = \left\{(x,x) : (x,x) \in Q_0' \times Q_0\right\}$ in $\mathsf{Q}_0 = Q_0' \times Q_0$ since (20) holds for all pairs of cubes Q' and Q of equal side length r having distance at least $C_0 r$ apart. Enumerate the cubes in Γ as $\{\mathsf{Q}_i\}_i$ and those remaining that are not in Γ as $\left\{\mathsf{P}_j\right\}_j$. Thus we have the pairwise disjoint decomposition

$$\mathsf{Q}_0 = \left(\bigcup_i \mathsf{Q}_i\right) \bigcup \left(\bigcup_j \mathsf{P}_j\right).$$

The countable additivity of the product measure $\omega \times \sigma$ shows that

$$|\mathsf{Q}_0|_{\omega\times\sigma} = \sum_i |\mathsf{Q}_i|_{\omega\times\sigma} + \sum_j \left|\mathsf{P}_j\right|_{\omega\times\sigma}.$$

Now we have

$$\sum_i |\mathsf{Q}_i|_{\omega\times\sigma} \lesssim \sum_i \mathfrak{N}_\alpha \left(\mathbf{R}^\alpha\right)^2 |\mathsf{Q}_i|^{1-\frac{\alpha}{n}} ,$$

and

$$\begin{aligned}\sum_i |\mathsf{Q}_i|^{1-\frac{\alpha}{n}} &= \sum_{k\in\mathbb{Z}:\ 2^k \le \ell(Q_0)} \ \sum_{i:\ \ell(Q_i)=2^k} \left(2^{2nk}\right)^{1-\frac{\alpha}{n}} \approx \sum_{k\in\mathbb{Z}:\ 2^k \le \ell(Q_0)} \left(\frac{2^k}{\ell(Q_0)}\right)^{-n} \left(2^{2nk}\right)^{1-\frac{\alpha}{n}} \quad \text{(Whitney)} \\ &= \ell(Q_0)^n \sum_{k\in\mathbb{Z}:\ 2^k \le \ell(Q_0)} 2^{nk\left(-1+2-\frac{2\alpha}{n}\right)} \le C_\alpha \ell(Q_0)^n \ell(Q_0)^{n\left(1-\frac{2\alpha}{n}\right)} = C_\alpha |Q_0 \times Q_0|^{2-\frac{2\alpha}{n}} = C_\alpha |\mathsf{Q}_0|^{1-\frac{\alpha}{n}} ,\end{aligned}$$

provided $0 \leq \alpha < \frac{n}{2}$. Using that the side length of $\mathsf{P}_j = P_j \times P'_j$ is 2^{-N} and $dist\left(\mathsf{P}_j, \mathfrak{D}\right) \leq C_r 2^{-N}$, we have the following limit,

$$\sum_j \left|\mathsf{P}_j\right|_{\omega\times\sigma} = \left|\bigcup_j \mathsf{P}_j\right|_{\omega\times\sigma} \to 0 \text{ as } N \to \infty,$$

since $\bigcup_j \mathsf{P}_j$ shrinks to the empty set as $N \to \infty$, and since locally finite measures such as $\omega \times \sigma$ are regular in Euclidean space. This completes the proof that $\sqrt{A_2^\alpha} \lesssim \mathfrak{N}_\alpha\left(\mathbf{R}^\alpha\right)$ for the range $0 \leq \alpha < \frac{n}{2}$.

Now we turn to proving $\sqrt{A_2^\alpha} \lesssim \mathfrak{N}_\alpha\left(\mathbf{R}^\alpha\right)$ for the range $\frac{n}{2} \leq \alpha < n$, where we assume the stronger ellipticity condition (11). So fix a pair of neighbouring quasicubes $(K', K) \in \Omega\mathcal{N}^n$, and assume that $\sigma + \omega$ doesn't charge the intersection $\overline{K'} \cap \overline{K}$ of the closures of K' and K. It will be convenient to replace n by $n+1$, i.e to introduce an additional dimension, and work with the preimages $Q' = \Omega^{-1}K'$ and $Q = \Omega^{-1}K$ that are usual cubes, and with the corresponding pullbacks $\widetilde{\omega} = m_1 \times \Omega^*\omega$ and $\widetilde{\sigma} = m_1 \times \Omega^*\sigma$ of the measures ω and σ where m_1 is Lebesgue measure on the line. We may also assume that

$$Q' = [-1, 0) \times \prod_{i=1}^n Q_i, \quad Q = [0, 1) \times \prod_{i=1}^n Q_i.$$

where $Q_i = [a_i, b_i]$ for $1 \leq i \leq n$ (since the other cases are handled in similar fashion). It is important to note that we are considering the intervals Q_i here to be closed, and we will track this difference as we proceed.

Choose $\theta_1 \in [a_1, b_1]$ so that both

$$\left|[-1, 0) \times [a_1, \theta_1] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\omega}}, \quad \left|[-1, 0) \times [\theta_1, b_1] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\omega}} \geq \frac{1}{2}\left|Q'\right|_{\widetilde{\omega}}.$$

Now denote the two intervals $[a_1, \theta_1]$ and $[\theta_1, b_1]$ by $\left[a_1^*, b_1^*\right]$ and $\left[a_1^{**}, b_1^{**}\right]$ where the order is chosen so that

$$\left|[0, 1) \times \left[a_1^*, b_1^*\right] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\sigma}} \leq \left|[0, 1) \times \left[a_1^{**}, b_1^{**}\right] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\sigma}}.$$

Then we have both

$$\left|[-1, 0) \times \left[a_1^*, b_1^*\right] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\omega}} \geq \frac{1}{2}|Q|_{\widetilde{\omega}} \text{ and } \left|[0, 1) \times \left[a_1^{**}, b_1^{**}\right] \times \prod_{i=2}^n Q_i\right|_{\widetilde{\sigma}} \geq \frac{1}{2}|Q|_{\widetilde{\sigma}}.$$

Now choose $\theta_2 \in [a_2, b_2]$ so that both

$$\left|[-1,0) \times \left[a_1^*, b_1^*\right] \times [a_2, \theta_2] \times \prod_{i=3}^{n} Q_i\right|_{\widetilde{\omega}}, \quad \left|[-1,0) \times \left[a_1^*, b_1^*\right] \times [\theta_2, b_2] \times \prod_{i=3}^{n} Q_i\right|_{\widetilde{\omega}} \geq \frac{1}{4} |Q|_{\widetilde{\omega}},$$

and denote the two intervals $[a_2, \theta_2]$ and $[\theta_2, b_2]$ by $\left[a_2^*, b_2^*\right]$ and $\left[a_2^{**}, b_2^{**}\right]$ where the order is chosen so that

$$[0,1) \times \left|\left[a_1^{**}, b_1^{**}\right] \times \left[a_2^*, b_2^*\right] \times \prod_{i=2}^{n} Q_i\right|_{\widetilde{\sigma}} \leq \left|[0,1) \times \left[a_1^{**}, b_1^{**}\right] \times \left[a_2^{**}, b_2^{**}\right] \times \prod_{i=2}^{n} Q_i\right|_{\widetilde{\sigma}}.$$

Then we have both

$$\left|[-1,0) \times \left[a_1^*, b_1^*\right] \times \left[a_2^*, b_2^*\right] \times \prod_{i=3}^{n} Q_i\right|_{\widetilde{\omega}} \geq \frac{1}{4} |Q|_{\widetilde{\omega}},$$
$$\left|[0,1) \times \left[a_1^{**}, b_1^{**}\right] \times \left[a_2^{**}, b_2^{**}\right] \times \prod_{i=3}^{n} Q_i\right|_{\widetilde{\sigma}} \geq \frac{1}{4} |Q|_{\widetilde{\sigma}},$$

and continuing in this way we end up with two rectangles,

$$G \equiv [-1,0) \times \left[a_1^*, b_1^*\right] \times \left[a_2^*, b_2^*\right] \times \ldots \times \left[a_n^*, b_n^*\right],$$
$$H \equiv [0,1) \times \left[a_1^{**}, b_1^{**}\right] \times \left[a_2^{**}, b_2^{**}\right] \times \ldots \times \left[a_n^{**}, b_n^{**}\right],$$

that satisfy

$$|G|_{\widetilde{\omega}} = \left|[-1,0) \times \left[a_1^*, b_1^*\right] \times \left[a_2^*, b_2^*\right] \times \ldots \times \left[a_n^*, b_n^*\right]\right|_{\widetilde{\omega}} \geq \frac{1}{2^n} |Q|_{\widetilde{\omega}},$$
$$|H|_{\widetilde{\sigma}} = \left|[0,1) \times \left[a_1^{**}, b_1^{**}\right] \times \left[a_2^{**}, b_2^{**}\right] \times \ldots \times \left[a_n^{**}, b_n^{**}\right]\right|_{\widetilde{\sigma}} \geq \frac{1}{2^n} |Q|_{\widetilde{\sigma}}.$$

However, the quasirectangles ΩG and ΩH lie in opposing quasi-n-ants at the vertex $\Omega\theta = \Omega(\theta_1, \theta_2, \ldots, \theta_n)$, and so we can apply (11) to obtain that for $x \in \Omega G$,

$$\left|\sum_{j=1}^{J} \lambda_j^m T_j^\alpha (\mathbf{1}_{\Omega H}\sigma)(x)\right| = \left|\int_{\Omega H} \sum_{j=1}^{J} \lambda_j^m K_j^\alpha (x,y)\, d\sigma(y)\right| \gtrsim \int_{\Omega H} |x-y|^{\alpha-n}\, d\sigma(y) \gtrsim |\Omega Q|^{\frac{\alpha}{n}-1} |\Omega H|_\sigma.$$

For the inequality above, we need to know that the distinguished point $\Omega\theta$ is not a common point mass of σ and ω, but this follows from our assumption that $\sigma + \omega$ doesn't charge the intersection $\overline{K'} \cap \overline{K}$ of the closures of K' and K. Then from the norm inequality we get

$$
\begin{aligned}
|\Omega G|_{\omega}\left(|\Omega Q|^{\frac{\alpha}{n}-1}|\Omega H|_{\sigma}\right)^{2} &\lesssim \int_{G}\left|\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}\left(\mathbf{1}_{\Omega H}\sigma\right)\right|^{2}d\omega \\
&\lesssim \mathfrak{N}_{\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}}^{2}\int\mathbf{1}_{\Omega H}^{2}d\sigma=\mathfrak{N}_{\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}}^{2}|\Omega H|_{\sigma}\,,
\end{aligned}
$$

from which we deduce that

$$
\begin{aligned}
|\Omega Q|^{2\left(\frac{\alpha}{n}-1\right)}\left|\Omega Q'\right|_{\omega}|\Omega Q|_{\sigma} &\lesssim 2^{2n}|\Omega Q|^{2\left(\frac{\alpha}{n}-1\right)}|\Omega G|_{\omega}|\Omega H|_{\sigma}\lesssim 2^{2n}\mathfrak{N}_{\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}}^{2}; \\
|K|^{2\left(\frac{\alpha}{n}-1\right)}\left|K'\right|_{\omega}|K|_{\sigma} &\lesssim 2^{2n}\mathfrak{N}_{\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}}^{2},
\end{aligned}
$$

and hence

$$
A_{2}^{\alpha}\lesssim 2^{2n}\mathfrak{N}_{\sum_{j=1}^{J}\lambda_{j}^{m}T_{j}^{\alpha}}^{2}.
$$

Thus we have obtained the offset A_2^α condition for pairs $(K', K) \in \Omega\mathcal{N}^n$ such that $\sigma + \omega$ doesn't charge the intersection $\overline{K'} \cap \overline{K}$ of the closures of K' and K. From this and the argument at the beginning of this proof, we obtain the one-tailed $\mathcal{A}_2^\alpha$ conditions. Indeed, we note that $|\partial(rQ)|_{\sigma+\omega} > 0$ for only a countable number of dilates $r > 1$, and so a limiting argument applies. This completes the proof of Lemma 4.1. □

Monotonicity Lemma and Energy Lemma

The Monotonicity Lemma below will be used to prove the Energy Lemma, which is then used in several places in the proof of Theorem 2.8. The formulation of the Monotonicity Lemma with $m = 2$ for cubes is due to M. Lacey and B. Wick [18], and corrects that used in early versions of our paper [28].

The Monotonicity Lemma

For $0 \le \alpha < n$ and $m \in \mathbb{R}_+$, we recall the m-weighted fractional Poisson integral

$$
\mathrm{P}_{m}^{\alpha}(J,\mu)\equiv\int_{\mathbb{R}^{n}}\frac{|J|^{\frac{m}{n}}}{\left(|J|^{\frac{1}{n}}+|y-c_{J}|\right)^{n+m-\alpha}}d\mu(y)\,,
$$

where $\mathrm{P}_1^\alpha(J,\mu) = \mathrm{P}^\alpha(J,\mu)$ is the standard Poisson integral. The next lemma holds for quasicubes and common point masses with the same proof as in [30].

Lemma 5.1 (Monotonicity) *Suppose that I and J are quasicubes in $\mathbb{R}^n$ such that $J \subset 2J \subset I$, and that μ is a signed measure on $\mathbb{R}^n$ supported outside I. Finally suppose that T^α is a standard α-fractional singular integral on $\mathbb{R}^n$ with $0 < \alpha < n$. Then we have the estimate*

$$\left\| \triangle_J^\omega T^\alpha \mu \right\|_{L^2(\omega)} \lesssim \Phi^\alpha \left(J, |\mu| \right), \tag{23}$$

where for a positive measure ν,

$$\Phi^\alpha \left(J, \nu \right)^2 \equiv \left(\frac{\mathrm{P}^\alpha \left(J, \nu \right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \triangle_J^\omega \mathbf{x} \right\|_{L^2(\omega)}^2 + \left(\frac{\mathrm{P}_{1+\delta}^\alpha \left(J, \nu \right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathbf{x} - \mathbf{m}_J \right\|_{L^2(\mathbf{1}_J \omega)}^2 ,$$

$$\mathbf{m}_J \equiv \mathbb{E}_J^\omega \mathbf{x} = \frac{1}{|J|_\omega} \int_J \mathbf{x} d\omega .$$

The Energy Lemma

Suppose now we are given a subset $\mathcal{H}$ of the dyadic quasigrid $\Omega\mathcal{D}^\omega$. Let $\mathsf{P}_{\mathcal{H}}^\omega = \sum_{J \in \mathcal{H}} \triangle_J^\omega$ be the corresponding ω-quasiHaar projection. We define $\mathcal{H}^* \equiv \bigcup_{J \in \mathcal{H}} \left\{ J' \in \Omega\mathcal{D}^\omega : J' \subset J \right\}$. The next lemma also holds for quasicubes and common point masses with the same proof as in [30].

Lemma 5.2 (Energy Lemma) *Let J be a quasicube in $\Omega\mathcal{D}^\omega$. Let Ψ_J be an $L^2(\omega)$ function supported in J and with ω-integral zero, and denote its quasiHaar support by $\mathcal{H} = \operatorname{supp} \widehat{\Psi_J} \equiv \left\{ K \in \Omega\mathcal{D}^\omega : \widehat{\Psi_J}(K) \neq 0 \right\}$. Let ν be a positive measure supported in $\mathbb{R}^n \setminus \gamma J$ with $\gamma \geq 2$, and for each $J' \in \mathcal{H}$, let $\nu_{J'} = \varphi_{J'} \nu$ with $|\varphi_{J'}| \leq 1$. Let T^α be a standard α-fractional singular integral operator with $0 \leq \alpha < n$. Then with $\delta' = \frac{\delta}{2}$ we have*

$$\begin{aligned} \left| \sum_{J' \in \mathcal{H}} \left\langle T^\alpha \left(\nu_{J'} \right), \triangle_{J'}^\omega \Psi_J \right\rangle_\omega \right| &\lesssim \left\| \Psi_J \right\|_{L^2(\omega)} \left(\frac{\mathrm{P}^\alpha \left(J, \nu \right)}{|J|^{\frac{1}{n}}} \right) \left\| \mathsf{P}_{\mathcal{H}}^\omega \mathbf{x} \right\|_{L^2(\omega)} \\ &\quad + \left\| \Psi_J \right\|_{L^2(\omega)} \frac{1}{\gamma^{\delta'}} \left(\frac{\mathrm{P}_{1+\delta'}^\alpha \left(J, \nu \right)}{|J|^{\frac{1}{n}}} \right) \left\| \mathsf{P}_{\mathcal{H}^*}^\omega \mathbf{x} \right\|_{L^2(\omega)} \\ &\lesssim \left\| \Psi_J \right\|_{L^2(\omega)} \left(\frac{\mathrm{P}^\alpha \left(J, \nu \right)}{|J|^{\frac{1}{n}}} \right) \left\| \mathsf{P}_{\mathcal{H}^*}^\omega \mathbf{x} \right\|_{L^2(\omega)} , \end{aligned}$$

and in particular the 'pivotal' bound

$$\left|\left\langle T^{\alpha}\left(\nu\right),\Psi_{J}\right\rangle_{\omega}\right| \leq C\left\|\Psi_{J}\right\|_{L^{2}(\omega)} \mathrm{P}^{\alpha}\left(J,|\nu|\right) \sqrt{|J|_{\omega}}.$$

Remark 5.3 The first term on the right side of the energy inequality above is the 'big' Poisson integral P^{α} times the 'small' energy term $\left\|\mathsf{P}_{\mathcal{H}}^{\omega}\mathbf{x}\right\|_{L^{2}(\omega)}^{2}$ that is additive in $\mathcal{H}$, while the second term on the right is the 'small' Poisson integral $\mathrm{P}_{1+\delta'}^{\alpha}$ times the 'big' energy term $\left\|\mathsf{P}_{\mathcal{H}^{*}}^{\omega}\mathbf{x}\right\|_{L^{2}(\omega)}$ that is no longer additive in $\mathcal{H}$. The first term presents no problems in subsequent analysis due solely to the additivity of the 'small' energy term. It is the second term that must be handled by special methods. For example, in the Intertwining Proposition below, the interaction of the singular integral occurs with a pair of quasicubes $J \subset I$ at *highly separated* levels, where the goodness of J can exploit the decay δ' in the kernel of the 'small' Poisson integral $\mathrm{P}_{1+\delta'}^{\alpha}$ relative to the 'big' Poisson integral P^{α}, and results in a bound directly by the quasienergy condition. On the other hand, in the local recursion of M. Lacey at the end of the paper, the separation of levels in the pairs $J \subset I$ can be as *little* as a fixed parameter $\boldsymbol{\rho}$, and here we must first separate the stopping form into two sublinear forms that involve the two estimates respectively. The form corresponding to the smaller Poisson integral $\mathrm{P}_{1+\delta'}^{\alpha}$ is again handled using goodness and the decay δ' in the kernel, while the form corresponding to the larger Poisson integral P^{α} requires the stopping time and recursion argument of M. Lacey.

Preliminaries of NTV Type

An important reduction of our theorem is delivered by the following two lemmas, the first of which is due to Nazarov, Treil and Volberg in the case of one dimension (see [21] and [35]), and the second of which is a bilinear Carleson embedding. The proofs given there do not extend in standard ways to higher dimensions with common point masses, and we use the quasiweak boundedness property to handle the case of touching quasicubes, and an application of Schur's Lemma to handle the case of separated quasicubes. The first lemma below is Lemmas 8.1 and 8.7 in [18] but with the larger constant $\mathcal{A}_{2}^{\alpha}$ there in place of the smaller constant A_{2}^{α} here. We emphasize that only the offset A_{2}^{α} condition is needed with testing and weak boundedness in these preliminary estimates.

Lemma 6.1 *Suppose T^{α} is a standard fractional singular integral with $0 \leq \alpha < n$, and that all of the quasicubes $I \in \Omega\mathcal{D}^{\sigma}, J \in \Omega\mathcal{D}^{\omega}$ below are good with goodness parameters ε and $\mathbf{r}$. Fix a positive integer $\boldsymbol{\rho} > \mathbf{r}$. For $f \in L^{2}(\sigma)$ and $g \in L^{2}(\omega)$ we have*

$$\sum_{\substack{(I,J)\in\Omega\mathcal{D}^{\sigma}\times\Omega\mathcal{D}^{\omega}\\ 2^{-\boldsymbol{\rho}}\ell(I)\leq\ell(J)\leq 2^{\boldsymbol{\rho}}\ell(I)}} \left|\left\langle T_{\sigma}^{\alpha}\left(\triangle_{I}^{\sigma}f\right),\triangle_{J}^{\omega}g\right\rangle_{\omega}\right| \lesssim \left(\mathfrak{T}_{\alpha}+\mathfrak{T}_{\alpha}^{*}+\mathcal{WBP}_{T^{\alpha}}+\sqrt{A_{2}^{\alpha}}\right)\left\|f\right\|_{L^{2}(\sigma)}\left\|g\right\|_{L^{2}(\omega)} \tag{24}$$

and

$$\sum_{\substack{(I,J)\in\Omega\mathcal{D}^{\sigma}\times\Omega\mathcal{D}^{\omega}\\ I\cap J=\emptyset \text{ and } \frac{\ell(J)}{\ell(I)}\notin[2^{-\rho},2^{\rho}]}} \left|\left\langle T_{\sigma}^{\alpha}\left(\triangle_I^{\sigma}f\right),\triangle_J^{\omega}g\right\rangle_{\omega}\right| \lesssim \sqrt{A_2^{\alpha}}\,\|f\|_{L^2(\sigma)}\,\|g\|_{L^2(\omega)}\,, \tag{25}$$

where the implied constants depend only on n, α and T^{α}.

Lemma 6.2 *Suppose T^{α} is a standard fractional singular integral with $0\le\alpha<n$, that all of the quasicubes $I\in\Omega\mathcal{D}^{\sigma}, J\in\Omega\mathcal{D}^{\omega}$ below are good, that $\boldsymbol{\rho}>\mathbf{r}$, that $f\in L^2(\sigma)$ and $g\in L^2(\omega)$, that $\mathcal{F}\subset\Omega\mathcal{D}^{\sigma}$ and $\mathcal{G}\subset\Omega\mathcal{D}^{\omega}$ are σ-Carleson and ω-Carleson collections respectively, i.e.,*

$$\sum_{F'\in\mathcal{F}:\,F'\subset F}\left|F'\right|_{\sigma}\lesssim|F|_{\sigma}\,,\qquad F\in\mathcal{F},\text{ and }\sum_{G'\in\mathcal{G}:\,G'\subset G}\left|G'\right|_{\omega}\lesssim|G|_{\omega}\,,\qquad G\in\mathcal{G},$$

that there are numerical sequences $\{\alpha_{\mathcal{F}}(F)\}_{F\in\mathcal{F}}$ and $\{\beta_{\mathcal{G}}(G)\}_{G\in\mathcal{G}}$ such that

$$\sum_{F\in\mathcal{F}}\alpha_{\mathcal{F}}(F)^2\,|F|_{\sigma}\le\|f\|_{L^2(\sigma)}^2 \text{ and } \sum_{G\in\mathcal{G}}\beta_{\mathcal{G}}(G)^2\,|G|_{\sigma}\le\|g\|_{L^2(\sigma)}^2\,, \tag{26}$$

and finally that for each pair of quasicubes $(I,J)\in\Omega\mathcal{D}^{\sigma}\times\Omega\mathcal{D}^{\omega}$, there are bounded functions $\beta_{I,J}$ and $\gamma_{I,J}$ supported in $I\setminus 2J$ and $J\setminus 2I$ respectively, satisfying

$$\|\beta_{I,J}\|_{\infty}\,,\|\gamma_{I,J}\|_{\infty}\le 1.$$

Then

$$\sum_{\substack{(F,J)\in\mathcal{F}\times\Omega\mathcal{D}^{\omega}\\ F\cap J=\emptyset \text{ and } \ell(J)\le 2^{-\rho}\ell(F)}}\left|\left\langle T_{\sigma}^{\alpha}\left(\beta_{F,J}\mathbf{1}_F\alpha_{\mathcal{F}}(F)\right),\triangle_J^{\omega}g\right\rangle_{\omega}\right| + \sum_{\substack{(I,G)\in\Omega\mathcal{D}^{\sigma}\times\mathcal{G}\\ I\cap G=\emptyset \text{ and } \ell(I)\le 2^{-\rho}\ell(G)}}\left|\left\langle T_{\sigma}^{\alpha}\left(\triangle_I^{\sigma}f\right),\gamma_{I,G}\mathbf{1}_G\beta_{\mathcal{G}}(G)\right\rangle_{\omega}\right|$$

$$\lesssim\sqrt{A_2^{\alpha}}\,\|f\|_{L^2(\sigma)}\,\|g\|_{L^2(\omega)}\,. \tag{27}$$

See [29] for complete details of the proofs when common point masses are permitted.

Remark 6.3 If $\mathcal{F}$ and $\mathcal{G}$ are σ-Carleson and ω-Carleson collections respectively, and if $\alpha_{\mathcal{F}}(F)=\mathbb{E}_F^{\sigma}|f|$ and $\beta_{\mathcal{G}}(G)=\mathbb{E}_G^{\omega}|g|$, then the 'quasi' orthogonality condition (26) holds (here 'quasi' has a different meaning than quasi), and this special case of Lemma 6.2 serves as a basic example.

Remark 6.4 Lemmas 6.1 and 6.2 differ mainly in that an orthogonal collection of quasiHaar projections is replaced by a 'quasi' orthogonal collection of indicators $\{\mathbf{1}_F\alpha_{\mathcal{F}}(F)\}_{F\in\mathcal{F}}$. More precisely, the main difference between (25) and (27) is that a quasiHaar projection $\triangle_I^{\sigma}f$ or $\triangle_J^{\omega}g$ has been replaced with a constant multiple of an indicator $\mathbf{1}_F\alpha_{\mathcal{F}}(F)$ or $\mathbf{1}_G\beta_{\mathcal{G}}(G)$, and in addition, a bounded function is permitted to multiply the indicator of the quasicube having larger sidelength.

Corona Decompositions and Splittings

We will use two different corona constructions, namely a Calderón-Zygmund decomposition and an energy decomposition of NTV type, to reduce matters to the stopping form, the main part of which is handled by Lacey's recursion argument. We will then iterate these coronas into a double corona. We first recall our basic setup. For convenience in notation we will sometimes suppress the dependence on α in our nonlinear forms, but will retain it in the operators, Poisson integrals and constants. We will assume that the good/bad quasicube machinery of Nazarov, Treil and Volberg [35] is in force here as in [30]. Let $\Omega\mathcal{D}^{\sigma} = \Omega\mathcal{D}^{\omega}$ be an $(\mathbf{r}, \varepsilon)$-good quasigrid on $\mathbb{R}^n$, and let $\left\{h_I^{\sigma,a}\right\}_{I\in\Omega\mathcal{D}^{\sigma},\, a\in\Gamma_n}$ and $\left\{h_J^{\omega,b}\right\}_{J\in\Omega\mathcal{D}^{\omega},\, b\in\Gamma_n}$ be corresponding quasiHaar bases as described above, so that

$$f = \sum_{I\in\Omega\mathcal{D}^{\sigma}} \triangle_I^{\sigma} f \text{ and } g = \sum_{J\in\Omega\mathcal{D}^{\omega}} \triangle_J^{\omega} g,$$

where the quasiHaar projections $\triangle_I^{\sigma} f$ and $\triangle_J^{\omega} g$ vanish if the quasicubes I and J are not good. Recall that we must show the bilinear inequality (17), i.e. $|\mathcal{T}^{\alpha}(f, g)| \leq \mathfrak{N}_{T^{\alpha}} \|f\|_{L^2(\sigma)} \|g\|_{L^2(\omega)}$.

We now proceed for the remainder of this section to follow the development in [30], pointing out just the highlights, and referring to [30] for proofs, when no changes are required by the inclusion of quasicubes and common point masses.

The Calderón-Zygmund Corona

We now introduce a stopping tree $\mathcal{F}$ for the function $f \in L^2(\sigma)$. Let $\mathcal{F}$ be a collection of Calderón-Zygmund stopping quasicubes for f, and let $\Omega\mathcal{D}^{\sigma} = \bigcup_{F\in\mathcal{F}} \mathcal{C}_F$ be the associated corona decomposition of the dyadic quasigrid $\Omega\mathcal{D}^{\sigma}$. See below and also [30] for the standard definitions of corona, etc.

For a quasicube $I \in \Omega\mathcal{D}^{\sigma}$ let $\pi_{\Omega\mathcal{D}^{\sigma}} I$ be the $\Omega\mathcal{D}^{\sigma}$-parent of I in the quasigrid $\Omega\mathcal{D}^{\sigma}$, and let $\pi_{\mathcal{F}} I$ be the smallest member of $\mathcal{F}$ that contains I. For $F, F' \in \mathcal{F}$, we say that F' is an $\mathcal{F}$-child of F if $\pi_{\mathcal{F}}(\pi_{\Omega\mathcal{D}^{\sigma}} F') = F$ (it could be that $F = \pi_{\Omega\mathcal{D}^{\sigma}} F'$), and we denote by $\mathfrak{C}_{\mathcal{F}}(F)$ the set of $\mathcal{F}$-children of F. For $F \in \mathcal{F}$, define the projection $\mathsf{P}_{\mathcal{C}_F}^{\sigma}$ onto the linear span of the quasiHaar functions $\left\{h_I^{\sigma,a}\right\}_{I\in\mathcal{C}_F,\, a\in\Gamma_n}$ by

$$\mathsf{P}_{\mathcal{C}_F}^{\sigma} f = \sum_{I\in\mathcal{C}_F} \triangle_I^{\sigma} f = \sum_{I\in\mathcal{C}_F,\, a\in\Gamma_n} \left\langle f, h_I^{\sigma,a}\right\rangle_{\sigma} h_I^{\sigma,a}.$$

The standard properties of these projections are

$$f = \sum_{F \in \mathcal{F}} \mathsf{P}^{\sigma}_{\mathcal{C}_F} f, \qquad \int \left(\mathsf{P}^{\sigma}_{\mathcal{C}_F} f \right) \sigma = 0, \qquad \|f\|^2_{L^2(\sigma)} = \sum_{F \in \mathcal{F}} \left\| \mathsf{P}^{\sigma}_{\mathcal{C}_F} f \right\|^2_{L^2(\sigma)} .$$

The Energy Corona

We also impose a quasienergy corona decomposition as in [21] and [16].

Definition 7.1 Given a quasicube S_0, define $\mathcal{S}(S_0)$ to be the maximal subquasicubes $I \subset S_0$ such that

$$\sum_{J \in \mathcal{M}_{\tau-\mathrm{deep}}(I)} \left(\frac{\mathrm{P}^{\alpha}\left(J, \mathbf{1}_{S_0 \setminus \gamma J} \sigma\right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}^{\mathrm{subgood},\omega}_J \mathbf{x} \right\|^2_{L^2(\omega)} \geq C_{\mathrm{energy}} \left[\left(\mathcal{E}^{\mathrm{strong}}_{\alpha} \right)^2 + A^{\alpha}_2 + A^{\alpha,\mathrm{punct}}_2 \right] |I|_{\sigma} , \tag{28}$$

where $\mathcal{E}^{\mathrm{strong}}_{\alpha}$ is the constant in the strong quasienergy condition defined in Definition 2.6, and C_{energy} is a sufficiently large positive constant depending only on $\boldsymbol{\tau} \geq \mathbf{r}, n$ and α. Then define the σ-energy stopping quasicubes of S_0 to be the collection

$$\mathcal{S} = \{S_0\} \cup \bigcup_{n=0}^{\infty} \mathcal{S}_n$$

where $\mathcal{S}_0 = \mathcal{S}(S_0)$ and $\mathcal{S}_{n+1} = \bigcup_{S \in \mathcal{S}_n} \mathcal{S}(S)$ for $n \geq 0$.

From the quasienergy condition in Definition 2.6 we obtain the σ-Carleson estimate

$$\sum_{S \in \mathcal{S}:\ S \subset I} |S|_{\sigma} \leq 2\, |I|_{\sigma} , \qquad I \in \Omega \mathcal{D}^{\sigma} . \tag{29}$$

Finally, we record the reason for introducing quasienergy stopping times. If

$$X_{\alpha}\left(\mathcal{C}_S\right)^2 \equiv \sup_{I \in \mathcal{C}_S} \frac{1}{|I|_{\sigma}} \sum_{J \in \mathcal{M}_{\tau-\mathrm{deep}}(I)} \left(\frac{\mathrm{P}^{\alpha}\left(J, \mathbf{1}_{S \setminus \gamma J} \sigma\right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}^{\mathrm{subgood},\omega}_J \mathbf{x} \right\|^2_{L^2(\omega)} \tag{30}$$

is (the square of) the *α-stopping quasienergy* of the weight pair (σ, ω) with respect to the corona $\mathcal{C}_S$, then we have the *stopping quasienergy bounds*

$$X_{\alpha}\left(\mathcal{C}_S\right) \leq \sqrt{C_{\mathrm{energy}}} \sqrt{\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2 + A^{\alpha}_2 + A^{\alpha,\mathrm{punct}}_2}, \qquad S \in \mathcal{S}, \tag{31}$$

where $A^{\alpha}_2 + A^{\alpha,\mathrm{punct}}_2$ and the strong quasienergy constant $\mathcal{E}^{\mathrm{strong}}_{\alpha}$ are controlled by assumption.

General Stopping Data

It is useful to extend our notion of corona decomposition to more general stopping data. Our general definition of stopping data will use a positive constant $C_0 \geq 4$.

Definition 7.2 Suppose we are given a positive constant $C_0 \geq 4$, a subset $\mathcal{F}$ of the dyadic quasigrid $\Omega\mathcal{D}^\sigma$ (called the stopping times), and a corresponding sequence $\alpha_{\mathcal{F}} \equiv \{\alpha_{\mathcal{F}}(F)\}_{F\in\mathcal{F}}$ of nonnegative numbers $\alpha_{\mathcal{F}}(F) \geq 0$ (called the stopping data). Let $(\mathcal{F}, \prec, \pi_{\mathcal{F}})$ be the tree structure on $\mathcal{F}$ inherited from $\Omega\mathcal{D}^\sigma$, and for each $F \in \mathcal{F}$ denote by $\mathcal{C}_F = \{I \in \Omega\mathcal{D}^\sigma : \pi_{\mathcal{F}} I = F\}$ the corona associated with F:

$$\mathcal{C}_F = \left\{I \in \Omega\mathcal{D}^\sigma : I \subset F \text{ and } I \not\subset F' \text{ for any } F' \prec F\right\}.$$

We say the triple $(C_0, \mathcal{F}, \alpha_{\mathcal{F}})$ constitutes *stopping data* for a function $f \in L^1_{loc}(\sigma)$ if

1. $\mathbb{E}^\sigma_I |f| \leq \alpha_{\mathcal{F}}(F)$ for all $I \in \mathcal{C}_F$ and $F \in \mathcal{F}$,
2. $\sum_{F' \preceq F} |F'|_\sigma \leq C_0 |F|_\sigma$ for all $F \in \mathcal{F}$,
3. $\sum_{F\in\mathcal{F}} \alpha_{\mathcal{F}}(F)^2 |F|_\sigma \leq C_0^2 \|f\|^2_{L^2(\sigma)}$,
4. $\alpha_{\mathcal{F}}(F) \leq \alpha_{\mathcal{F}}(F')$ whenever $F', F \in \mathcal{F}$ with $F' \subset F$.

Definition 7.3 If $(C_0, \mathcal{F}, \alpha_{\mathcal{F}})$ constitutes (general) *stopping data* for a function $f \in L^1_{loc}(\sigma)$, we refer to the orthogonal decomposition

$$f = \sum_{F\in\mathcal{F}} \mathsf{P}^\sigma_{\mathcal{C}_F} f; \qquad \mathsf{P}^\sigma_{\mathcal{C}_F} f \equiv \sum_{I\in\mathcal{C}_F} \triangle^\sigma_I f,$$

as the (general) *corona decomposition* of f associated with the stopping times $\mathcal{F}$.

Property (1) says that $\alpha_{\mathcal{F}}(F)$ bounds the quasiaverages of f in the corona $\mathcal{C}_F$, and property (2) says that the quasicubes at the tops of the coronas satisfy a Carleson condition relative to the weight σ. Note that a standard 'maximal quasicube' argument extends the Carleson condition in property (2) to the inequality

$$\sum_{F'\in\mathcal{F}:\ F'\subset A} \left|F'\right|_\sigma \leq C_0 |A|_\sigma \text{ for all open sets } A \subset \mathbb{R}^n. \tag{32}$$

Property (3) is the 'quasi' orthogonality condition that says the sequence of functions $\{\alpha_{\mathcal{F}}(F) \mathbf{1}_F\}_{F\in\mathcal{F}}$ is in the vector-valued space $L^2\left(\ell^2;\sigma\right)$, and property (4) says that the control on stopping data is nondecreasing on the stopping tree $\mathcal{F}$. We emphasize that we are *not* assuming in this definition the stronger property that there is $C > 1$ such that $\alpha_{\mathcal{F}}(F') > C\alpha_{\mathcal{F}}(F)$ whenever $F', F \in \mathcal{F}$ with $F' \subsetneqq F$. Instead, the properties (2) and (3) substitute for this lack. Of course the stronger property *does* hold for the familiar *Calderón-Zygmund* stopping data determined by the following requirements for $C > 1$,

$$\mathbb{E}^{\sigma}_{F'}|f| > C\mathbb{E}^{\sigma}_{F}|f| \text{ whenever } F', F \in \mathcal{F} \text{ with } F' \subsetneqq F, \quad \mathbb{E}^{\sigma}_{I}|f| \leq C\mathbb{E}^{\sigma}_{F}|f| \text{ for } I \in \mathcal{C}_F,$$

which are themselves sufficiently strong to automatically force properties (2) and (3) with $\alpha_{\mathcal{F}}(F) = \mathbb{E}^{\sigma}_{F}|f|$.

We have the following useful consequence of (2) and (3) that says the sequence $\{\alpha_{\mathcal{F}}(F)\mathbf{1}_F\}_{F\in\mathcal{F}}$ has a *'quasi' orthogonal* property relative to f with a constant C_0' depending only on C_0:

$$\left\|\sum_{F\in\mathcal{F}} \alpha_{\mathcal{F}}(F)\mathbf{1}_F\right\|^2_{L^2(\sigma)} \leq C_0' \|f\|^2_{L^2(\sigma)}. \tag{33}$$

We will use a construction that permits *iteration* of general corona decompositions.

Lemma 7.4 *Suppose that* $(C_0, \mathcal{F}, \alpha_{\mathcal{F}})$ *constitutes* stopping data *for a function* $f \in L^1_{loc}(\sigma)$*, and that for each* $F \in \mathcal{F}$*,* $(C_0, \mathcal{K}(F), \alpha_{\mathcal{K}(F)})$ *constitutes* stopping data *for the corona projection* $\mathsf{P}^{\sigma}_{\mathcal{C}_F} f$*, where in addition* $F \in \mathcal{K}(F)$*. There is a positive constant* C_1*, depending only on* C_0*, such that if*

$$\begin{aligned}
\mathcal{K}^*(F) &\equiv \left\{K \in \mathcal{K}(F) \cap \mathcal{C}_F : \alpha_{\mathcal{K}(F)}(K) \geq \alpha_{\mathcal{F}}(F)\right\}\\
\mathcal{K} &\equiv \bigcup_{F\in\mathcal{F}} \mathcal{K}^*(F) \cup \{F\},\\
\alpha_{\mathcal{K}}(K) &\equiv \begin{array}{cc} \alpha_{\mathcal{K}(F)}(K) & \text{for } K \in \mathcal{K}^*(F)\setminus\{F\} \\ \max\left\{\alpha_{\mathcal{F}}(F), \alpha_{\mathcal{K}(F)}(F)\right\} & \text{for } \quad K = F \end{array}, \quad \text{for } F \in \mathcal{F},
\end{aligned}$$

the triple $(C_1, \mathcal{K}, \alpha_{\mathcal{K}})$ *constitutes* stopping data *for* f*. We refer to the collection of quasicubes* $\mathcal{K}$ *as the* iterated *stopping times, and to the orthogonal decomposition* $f = \sum_{K\in\mathcal{K}} P_{\mathcal{C}^{\mathcal{K}}_K} f$ *as the* iterated *corona decomposition of* f*, where*

$$\mathcal{C}^{\mathcal{K}}_K \equiv \left\{I \in \Omega\mathcal{D} : I \subset K \text{ and } I \not\subset K' \text{ for } K' \prec_{\mathcal{K}} K\right\}.$$

Note that in our definition of $(C_1, \mathcal{K}, \alpha_{\mathcal{K}})$ we have 'discarded' from $\mathcal{K}(F)$ all of those $K \in \mathcal{K}(F)$ that are not in the corona $\mathcal{C}_F$, and also all of those $K \in \mathcal{K}(F)$ for which $\alpha_{\mathcal{K}(F)}(K)$ is strictly less than $\alpha_{\mathcal{F}}(F)$. Then the union over F of what remains is our new collection of stopping times. We then define stopping data $\alpha_{\mathcal{K}}(K)$ according to whether or not $K \in \mathcal{F}$: if $K \notin \mathcal{F}$ but $K \in \mathcal{C}_F$ then $\alpha_{\mathcal{K}}(K)$ equals $\alpha_{\mathcal{K}(F)}(K)$, while if $K \in \mathcal{F}$, then $\alpha_{\mathcal{K}}(K)$ is the larger of $\alpha_{\mathcal{K}(F)}(F)$ and $\alpha_{\mathcal{F}}(K)$. See [30] for a proof.

Doubly Iterated Coronas and the NTV Quasicube Size Splitting

Let

$$\mathcal{NTV}_\alpha \equiv \sqrt{\mathcal{A}_2^\alpha + \mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\text{punct}} + A_2^{\alpha,*,\text{punct}}} + \mathfrak{T}_{T^\alpha} + \mathfrak{T}_{T^\alpha}^*.$$

Here is a brief schematic diagram of the decompositions, with bounds in $\boxed{}$:

$\langle T_\sigma^\alpha f, g\rangle_\omega$			
$\downarrow$			
$\mathsf{B}_{\Subset_\rho}(f,g)$ +	$\mathsf{B}_{_\rho\Supset}(f,g)$ +	$\mathsf{B}_\cap(f,g)$ +	$\mathsf{B}_\diagup(f,g)$
$\downarrow$	$\boxed{\text{dual}}$	$\boxed{\mathcal{NTV}_\alpha}$	$\boxed{\mathcal{NTV}_\alpha}$
$\downarrow$			
$\mathsf{T}_{\text{diagonal}}(f,g)$ +	$\mathsf{T}_{\text{farbelow}}(f,g)$ +	$\mathsf{T}_{\text{farabove}}(f,g)$ +	$\mathsf{T}_{\text{disjoint}}(f,g)$
$\downarrow$	$\downarrow$	$\boxed{\emptyset}$	$\boxed{\emptyset}$
$\downarrow$	$\downarrow$		
$\mathsf{B}_{\Subset_\rho}^A(f,g)$	$\mathsf{T}_{\text{farbelow}}^1(f,g)$ +	$\mathsf{T}_{\text{farbelow}}^2(f,g)$	
$\downarrow$	$\boxed{\mathcal{NTV}_\alpha + \mathcal{E}_\alpha^{\text{strong}}}$	$\boxed{\mathcal{NTV}_\alpha}$	
$\downarrow$			
$\mathsf{B}_{stop}^A(f,g)$ +	$\mathsf{B}_{paraproduct}^A(f,g)$ +	$\mathsf{B}_{neighbour}^A(f,g)$	
$\boxed{\mathcal{E}_\alpha^{\text{strong}} + \sqrt{A_2^\alpha}}$	$\boxed{\mathfrak{T}_{T^\alpha}}$	$\boxed{\sqrt{A_2^\alpha}}$	

We begin with the NTV *quasicube size splitting* of the inner product $\langle T_\sigma^\alpha f, g\rangle_\omega$ – and later apply the iterated corona construction to the Calderón–Zygmund corona and the energy corona in order to bound the below form $\mathsf{B}_{\Subset_\rho}(f,g)$ – that splits the pairs of quasicubes (I,J) in a simultaneous quasiHaar decomposition of f and g into four groups, namely those pairs that:

1. are below the size diagonal and $\boldsymbol{\rho}$-deeply embedded,
2. are above the size diagonal and $\boldsymbol{\rho}$-deeply embedded,
3. are disjoint, and
4. are of $\boldsymbol{\rho}$-comparable size.

More precisely we have

$$\begin{aligned}
\langle T_\sigma^\alpha f, g\rangle_\omega &= \sum_{I\in\Omega\mathcal{D}^\sigma,\, J\in\Omega\mathcal{D}^\omega} \left\langle T_\sigma^\alpha\left(\triangle_I^\sigma f\right), \left(\triangle_J^\omega g\right)\right\rangle_\omega \\
&= \sum_{\substack{I\in\Omega\mathcal{D}^\sigma,\, J\in\Omega\mathcal{D}^\omega\\ J\Subset_\rho I}} \left\langle T_\sigma^\alpha\left(\triangle_I^\sigma f\right), \left(\triangle_J^\omega g\right)\right\rangle_\omega + \sum_{\substack{I\in\Omega\mathcal{D}^\sigma,\, J\in\Omega\mathcal{D}^\omega\\ J_\rho\Supset I}} \left\langle T_\sigma^\alpha\left(\triangle_I^\sigma f\right), \left(\triangle_J^\omega g\right)\right\rangle_\omega \\
&\quad + \sum_{\substack{I\in\Omega\mathcal{D}^\sigma,\, J\in\Omega\mathcal{D}^\omega\\ J\cap I=\emptyset}} \left\langle T_\sigma^\alpha\left(\triangle_I^\sigma f\right), \left(\triangle_J^\omega g\right)\right\rangle_\omega + \sum_{\substack{I\in\Omega\mathcal{D}^\sigma,\, J\in\Omega\mathcal{D}^\omega\\ 2^{-\rho}\le\ell(J)\diagup\ell(I)\le 2^\rho}} \left\langle T_\sigma^\alpha\left(\triangle_I^\sigma f\right), \left(\triangle_J^\omega g\right)\right\rangle_\omega \\
&= \mathsf{B}_{\Subset_\rho}(f,g) + \mathsf{B}_{_\rho\Supset}(f,g) + \mathsf{B}_\cap(f,g) + \mathsf{B}_\diagup(f,g).
\end{aligned}$$

Lemma 6.1 in the section on NTV preliminaries show that the *disjoint* and *comparable* forms $\mathsf{B}_{\cap}(f,g)$ and $\mathsf{B}_{\diagup}(f,g)$ are both bounded by the $\mathcal{A}_2^{\alpha}+A_2^{\alpha,\text{punct}}$, quasitesting and quasiweak boundedness property constants. The *below* and *above* forms are clearly symmetric, so we need only consider the form $\mathsf{B}_{\Subset_{\rho}}(f,g)$, to which we turn for the remainder of the proof. For this we need functional energy.

Definition 7.5 Let $\mathfrak{F}_{\alpha}$ be the smallest constant in the '**f**unctional quasienergy' inequality below, holding for all $h\in L^2(\sigma)$ and all σ-Carleson collections $\mathcal{F}$ with Carleson norm $C_{\mathcal{F}}$ bounded by a fixed constant C:

$$\sum_{F\in\mathcal{F}}\sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)}\left(\frac{\mathrm{P}^{\alpha}(J,h\sigma)}{|J|^{\frac{1}{n}}}\right)^2\left\|\mathsf{P}^{\omega}_{\mathcal{C}_F^{\text{good},\tau-\text{shift}};J}\mathbf{x}\right\|^2_{L^2(\omega)}\leq\mathfrak{F}_{\alpha}\|h\|_{L^2(\sigma)}. \tag{34}$$

Several ingredients now come into play in order to reduce control of the below form $\mathsf{B}_{\Subset_{\rho}}(f,g)$ to the functional energy constant $\mathcal{F}_{\alpha}$ and the stopping form $\mathsf{B}^{A}_{stop}(f,g)$;

1. starting with the doubly iterated corona of Calderón-Zygmund and energy in Lemma 7.4 in order to obtain the decomposition into $\mathsf{T}_{\text{diagonal}}$, $\mathsf{T}_{\text{farbelow}}$, $\mathsf{T}_{\text{farabove}}$ and $\mathsf{T}_{\text{disjoint}}$,
2. continuing with an adaptation of the Intertwining Proposition from [30] to include quasicubes and common point masses so as to bound the forms $\mathsf{T}^1_{\text{farbelow}}$ and $\mathsf{T}^2_{\text{farbelow}}(f,g)$ using the functional energy constant $\mathcal{F}_{\alpha}$,
3. and followed by the NTV decomposition into paraproduct, neighbour and stopping forms.

The adaptation of the Intertwining Proposition to include quasicubes and common point masses is easy because the measures ω and σ only 'see each other' in the proof through the energy Muckenhoupt conditions $A_2^{\alpha,\text{energy}}$ and $A_2^{\alpha,*,\text{energy}}$, and the straightforward details can be found in [29]. Thus we now turn to the difficult task of controlling the functional energy constant $\mathcal{F}_{\alpha}$ by the Muckenhoupt and energy side conditions.

Control of Functional Energy by Energy Modulo $\mathcal{A}_2^{\alpha}$ and $A_2^{\alpha,\text{punct}}$

Now we arrive at one of our main propositions in the proof of our theorem. We show that the functional quasienergy constants $\mathfrak{F}_{\alpha}$ as in (34) are controlled by $\mathcal{A}_2^{\alpha}$, $A_2^{\alpha,\text{punct}}$ and both the *strong* quasienergy constant $\mathcal{E}_{\alpha}^{\text{strong}}$ defined in Definition 2.6. The proof of this fact is further complicated when common point masses are permitted, accounting for the inclusion of the punctured Muckenhoupt condition $A_2^{\alpha,\text{punct}}$. But apart from this difference, the proof here is essentially the same as that in [30], where common point masses were prohibited. As a consequence we will refer to

[30] in many of the places where the arguments are unchanged. A complete and detailed proof can of course be found in [29].

Proposition 8.1 *We have*

$$\mathfrak{F}_\alpha \lesssim \mathcal{E}_\alpha^{\text{strong}} + \sqrt{\mathcal{A}_2^\alpha} + \sqrt{\mathcal{A}_2^{\alpha,*}} + \sqrt{A_2^{\alpha,\text{punct}}} \text{ and } \mathfrak{F}_\alpha^* \lesssim \mathcal{E}_\alpha^{\text{strong},*} + \sqrt{\mathcal{A}_2^\alpha} + \sqrt{\mathcal{A}_2^{\alpha,*}} + \sqrt{A_2^{\alpha,*,\text{punct}}}.$$

To prove this proposition, we fix $\mathcal{F}$ as in (34), and set

$$\mu \equiv \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)} \left\| \mathsf{P}_{F,J}^{\omega}\mathbf{x} \right\|_{L^2(\omega)}^2 \cdot \delta_{(c_J,\ell(J))} \text{ and } d\overline{\mu}(x,t) \equiv \frac{1}{t^2} d\mu(x,t), \quad (35)$$

where $\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)$ consists of the maximal $\mathbf{r}$-deeply embedded subquasicubes of F, and where $\delta_{(c_J,\ell(J))}$ denotes the Dirac unit mass at the point $(c_J,\ell(J))$ in the upper half-space $\mathbb{R}_+^{n+1}$. Here J is a dyadic quasicube with center c_J and side length $\ell(J)$. For convenience in notation, we denote for any dyadic quasicube J the localized projection $\mathsf{P}^{\omega}_{\mathcal{C}_F^{\text{good},\boldsymbol{\tau}-\text{shift}};J}$ given by

$$\mathsf{P}_{F,J}^{\omega} \equiv \mathsf{P}^{\omega}_{\mathcal{C}_F^{\text{good},\boldsymbol{\tau}-\text{shift}};J} = \sum_{J'\subset J:\ J'\in\mathcal{C}_F^{\text{good},\boldsymbol{\tau}-\text{shift}}} \triangle_{J'}^{\omega}.$$

We emphasize that the quasicubes $J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)$ are not necessarily good, but that the subquasicubes $J' \subset J$ arising in the projection $\mathsf{P}_{F,J}^{\omega}$ are good. We can replace $\mathbf{x}$ by $\mathbf{x}-\mathbf{c}$ inside the projection for any choice of $\mathbf{c}$ we wish; the projection is unchanged. More generally, δ_q denotes a Dirac unit mass at a point q in the upper half-space $\mathbb{R}_+^{n+1}$.

We prove the two-weight inequality

$$\left\|\mathbb{P}^\alpha(f\sigma)\right\|_{L^2(\mathbb{R}_+^{n+1},\overline{\mu})} \lesssim \left(\mathcal{E}_\alpha^{\text{strong}} + \sqrt{\mathcal{A}_2^\alpha} + \sqrt{\mathcal{A}_2^{\alpha,*}} + \sqrt{A_2^{\alpha,\text{punct}}}\right) \|f\|_{L^2(\sigma)}, \quad (36)$$

for all nonnegative f in $L^2(\sigma)$, noting that $\mathcal{F}$ and f are *not* related here. Above, $\mathbb{P}^\alpha(\cdot)$ denotes the α-fractional Poisson extension to the upper half-space $\mathbb{R}_+^{n+1}$,

$$\mathbb{P}^\alpha\nu(x,t) \equiv \int_{\mathbb{R}^n} \frac{t}{\left(t^2+|x-y|^2\right)^{\frac{n+1-\alpha}{2}}} d\nu(y),$$

so that in particular

$$\left\|\mathbb{P}^\alpha(f\sigma)\right\|_{L^2(\mathbb{R}_+^{n+1},\overline{\mu})}^2 = \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F)} \mathbb{P}^\alpha(f\sigma)(c(J),\ell(J))^2 \left\| \mathsf{P}_{F,J}^{\omega} \frac{x}{|J|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2,$$

and so (36) proves the first line in Proposition 8.1 upon inspecting (34). Note also that we can equivalently write $\|\mathbb{P}^{\alpha}(f\sigma)\|_{L^2(\mathbb{R}^{n+1}_+,\overline{\mu})} = \left\|\widetilde{\mathbb{P}^{\alpha}}(f\sigma)\right\|_{L^2(\mathbb{R}^{n+1}_+,\mu)}$ where $\widetilde{\mathbb{P}^{\alpha}}\nu(x,t) \equiv \frac{1}{t}\mathbb{P}^{\alpha}\nu(x,t)$ is the renormalized Poisson operator. Here we have simply shifted the factor $\frac{1}{t^2}$ in $\overline{\mu}$ to $\left|\widetilde{\mathbb{P}^{\alpha}}(f\sigma)\right|^2$ instead, and we will do this shifting often throughout the proof when it is convenient to do so.

The characterization of the two-weight inequality for fractional and Poisson integrals in [24] was stated in terms of the collection $\mathcal{P}^n$ of cubes in $\mathbb{R}^n$ with sides parallel to the coordinate axes. It is a routine matter to pullback the Poisson inequality under a globally biLipschitz map $\Omega : \mathbb{R}^n \to \mathbb{R}^n$, then apply the theorem in [24] (as a black box), and then to pushforward the conclusions of the theorems so as to extend these characterizations of fractional and Poisson integral inequalities to the setting of quasicubes $Q \in \Omega\mathcal{P}^n$ and quasitents $Q \times [0, \ell(Q)] \subset \mathbb{R}^{n+1}_+$ with $Q \in \Omega\mathcal{P}^n$. Using this extended theorem for the two-weight Poisson inequality, we see that inequality (36) requires checking these two inequalities for dyadic quasicubes $I \in \Omega\mathcal{D}$ and quasiboxes $\widehat{I} = I \times [0, \ell(I))$ in the upper half-space $\mathbb{R}^{n+1}_+$:

$$\int_{\mathbb{R}^{n+1}_+} \mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)(x,t)^2\, d\overline{\mu}(x,t) \equiv \|\mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)\|^2_{L^2(\widehat{I},\overline{\mu})} \lesssim \left(\left(\mathcal{E}^{\text{strong}}_{\alpha}\right)^2 + \mathcal{A}^{\alpha}_2 + \mathcal{A}^{\alpha,*}_2 + A^{\alpha,\text{punct}}_2\right)\sigma(I), \tag{37}$$

$$\int_{\mathbb{R}^n} [\mathbb{Q}^{\alpha}(t\mathbf{1}_{\widehat{I}}\overline{\mu})]^2 d\sigma(x) \lesssim \left(\left(\mathcal{E}^{\text{strong}}_{\alpha}\right)^2 + \mathcal{A}^{\alpha}_2 + \mathcal{A}^{\alpha,*}_2 + A^{\alpha,\text{punct}}_2\right)\int_{\widehat{I}} t^2 d\overline{\mu}(x,t), \tag{38}$$

for all *dyadic* quasicubes $I \in \Omega\mathcal{D}$, and where the dual Poisson operator $\mathbb{Q}^{\alpha}$ is given by

$$\mathbb{Q}^{\alpha}(t\mathbf{1}_{\widehat{I}}\overline{\mu})(x) = \int_{\widehat{I}} \frac{t^2}{(t^2 + |x-y|^2)^{\frac{n+1-\alpha}{2}}} d\overline{\mu}(y,t)\ .$$

It is important to note that we can choose for $\Omega\mathcal{D}$ any fixed dyadic quasigrid, the compensating point being that the integrations on the left sides of (37) and (38) are taken over the entire spaces $\mathbb{R}^{n+1}_+$ and $\mathbb{R}^n$ respectively.

Remark 8.2 There is a gap in the proof of the Poisson inequality at the top of page 542 in [24]. However, this gap can be fixed as in [33, p. 861].

Poisson Testing

We now turn to proving the Poisson testing conditions (37) and (38). The same testing conditions have been considered in [28] but in the setting of no common point masses, and the proofs there carry over to the situation here, but careful attention must now be paid to the possibility of common point masses. In [6] Hytönen circumvented this difficulty by introducing a Poisson operator ‘with holes’, which was then analyzed using shifted dyadic grids, but part of his argument was heavily dependent on the dimension being $n = 1$, and the extension of this argument to higher dimensions is feasible (see earlier versions of this paper on the *arXiv*), but technically very involved. We circumvent the difficulty of permitting common point masses here instead by using the energy Muckenhoupt constants $A_2^{\alpha,\text{energy}}$ and $A_2^{\alpha,*,\text{energy}}$, which require control by the punctured Muckenhoupt constants $A_2^{\alpha,\text{punct}}$ and $A_2^{\alpha,*,\text{punct}}$. The following elementary Poisson inequalities (see e.g. [35]) will be used extensively.

Lemma 8.3 *Suppose that J, K, I are quasicubes in $\mathbb{R}^n$, and that μ is a positive measure supported in $\mathbb{R}^n \setminus I$. If $J \subset K \subset 2K \subset I$, then*

$$\frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}} \lesssim \frac{\mathrm{P}^\alpha(K,\mu)}{|K|^{\frac{1}{n}}} \lesssim \frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}},$$

while if $2J \subset K \subset I$, then

$$\frac{\mathrm{P}^\alpha(K,\mu)}{|K|^{\frac{1}{n}}} \lesssim \frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}}.$$

Now we record the bounded overlap of the projections $\mathsf{P}^\omega_{F,J}$.

Lemma 8.4 *Suppose $\mathsf{P}^\omega_{F,J}$ is as above and fix any $I_0 \in \Omega\mathcal{D}$, so that I_0, F and J all lie in a common quasigrid. If $J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)$ for some $F \in \mathcal{F}$ with $F \supsetneqq I_0 \supset J$ and $\mathsf{P}^\omega_{F,J} \neq 0$, then*

$$F = \pi^{(\ell)}_{\mathcal{F}} I_0 \textit{ for some } 0 \leq \ell \leq \boldsymbol{\tau}.$$

As a consequence we have the bounded overlap,

$$\#\left\{F \in \mathcal{F} : J \subset I_0 \subsetneqq F \textit{ for some } J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F) \ \textit{ with } \mathsf{P}^\omega_{F,J} \neq 0\right\} \leq \boldsymbol{\tau}.$$

Finally we record the only places in the proof where the *refined* quasienergy conditions are used. This lemma will be used in bounding both of the local Poisson testing conditions. Recall that $\mathcal{A}\Omega\mathcal{D}$ consists of all alternate $\Omega\mathcal{D}$-dyadic quasicubes where K is alternate dyadic if it is a union of 2^n $\Omega\mathcal{D}$-dyadic quasicubes K' with $\ell(K') = \frac{1}{2}\ell(K)$. See [30] for a proof when common point masses are prohibited, and the presence of common point masses here requires no change.

Remark 8.5 The following lemma is another of the key results on the way to the proof of our theorem, and is an analogue of the corresponding lemma from [28], but with the right hand side involving only the plugged energy constants and the energy Muckenhoupt constants.

Lemma 8.6 *Let $\Omega\mathcal{D}, \mathcal{F} \subset \Omega\mathcal{D}$ be quasigrids and $\left\{\mathsf{P}^{\omega}_{F,J}\right\}_{\substack{F\in\mathcal{F}\\ J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}}$ be as above with J, F in the dyadic quasigrid $\Omega\mathcal{D}$. For any alternate quasicube $I \in \mathcal{A}\Omega\mathcal{D}$ define*

$$B(I) \equiv \sum_{F\in\mathcal{F}:\ F\supsetneqq I' \text{ for some } I'\in\mathfrak{C}(I)} \ \sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F):\ J\subset I} \left(\frac{\mathrm{P}^{\alpha}\left(J, \mathbf{1}_I\sigma\right)}{|J|^{\frac{1}{n}}}\right)^2 \left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)} . \tag{39}$$

Then

$$B(I) \lesssim \boldsymbol{\tau}\left(\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2 + A_2^{\alpha,\mathrm{energy}}\right)|I|_{\sigma} . \tag{40}$$

The Forward Poisson Testing Inequality

Fix $I \in \Omega\mathcal{D}$. We split the integration on the left side of (37) into a local and global piece:

$$\int_{\mathbb{R}^{n+1}_{+}} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)^2 d\overline{\mu} = \int_{\widehat{I}} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)^2 d\overline{\mu} + \int_{\mathbb{R}^{n+1}_{+}\setminus\widehat{I}} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)^2 d\overline{\mu} \equiv \mathbf{Local}(I) + \mathbf{Global}(I),$$

where more explicitly,

$$\mathbf{Local}(I) \equiv \int_{\widehat{I}} \left[\mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)(x,t)\right]^2 d\overline{\mu}(x,t); \qquad \overline{\mu} \equiv \frac{1}{t^2}\mu, \tag{41}$$

$$\text{i.e. } \overline{\mu} \equiv \sum_{J\in\Omega\mathcal{D}} \frac{1}{\ell(J)^2} \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)} \left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)} \cdot \delta_{(c_J,\ell(J))}.$$

Here is a brief schematic diagram of the decompositions, with bounds in $\square$, used in this subsection:

$$\boxed{\begin{array}{ccc} \mathbf{Local}(I) & & \\ \downarrow & & \\ \mathbf{Local}^{\mathrm{plug}}(I) & + & \mathbf{Local}^{\mathrm{hole}}(I) \\ & & \boxed{\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2} \\ \downarrow & & \\ \downarrow & & \\ A & + & B \\ \boxed{\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2} & & \boxed{\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2 + A_2^{\alpha,\mathrm{energy}}} \end{array}}$$

and

$$\boxed{\begin{array}{ccccccc} \textbf{Global}\,(I) & & & & & & \\ \downarrow & & & & & & \\ A & + & B & + & C & + & D \\ \boxed{A_2^\alpha} & & \boxed{A_2^\alpha + A_2^{\alpha,\text{energy}}} & & \boxed{\mathcal{A}_2^{\alpha,*}} & & \boxed{\mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\text{energy}} + A_2^{\alpha,\text{punct}}} \end{array}}.$$

An important consequence of the fact that I and J lie in the same quasigrid $\Omega\mathcal{D} = \Omega\mathcal{D}^\omega$, is that

$$(c(J), \ell(J)) \in \widehat{I} \textbf{ if and only if } J \subset I. \tag{42}$$

We thus have

$$\begin{aligned}
&\textbf{Local}\,(I) = \int_{\widehat{I}} \mathbb{P}^\alpha(\mathbf{1}_I\sigma)(x,t)^2\, d\overline{\mu}(x,t) \\
&= \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F):\ J\subset I} \mathbb{P}^\alpha(\mathbf{1}_I\sigma)\left(c_J, |J|^{\frac{1}{n}}\right)^2 \left\| \mathsf{P}_{F,J}^\omega \frac{\mathbf{x}}{|J|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2 \\
&\approx \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F):\ J\subset I} \mathrm{P}^\alpha(J, \mathbf{1}_I\sigma)^2 \left\| \mathsf{P}_{F,J}^\omega \frac{\mathbf{x}}{|J|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2 \\
&\lesssim \textbf{Local}^{\text{plug}}(I) + \textbf{Local}^{\text{hole}}(I),
\end{aligned}$$

where the 'plugged' local sum $\textbf{Local}^{\text{plug}}(I)$ is given by

$$\begin{aligned}
\textbf{Local}^{\text{plug}}(I) &\equiv \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F):\ J\subset I} \left(\frac{\mathrm{P}^\alpha(J, \mathbf{1}_{F\cap I}\sigma)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}_{F,J}^\omega \mathbf{x} \right\|_{L^2(\omega)}^2 \\
&= \left\{ \sum_{F\in\mathcal{F}:\ F\subset I} + \sum_{F\in\mathcal{F}:\ F\supsetneqq I} \right\} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F):\ J\subset I} \left(\frac{\mathrm{P}^\alpha(J, \mathbf{1}_{F\cap I}\sigma)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}_{F,J}^\omega \mathbf{x} \right\|_{L^2(\omega)}^2 \\
&= A + B.
\end{aligned}$$

Then a *trivial* application of the deep quasienergy condition (where 'trivial' means that the outer decomposition is just a single quasicube) gives

$$\begin{aligned}
A &\le \sum_{F\in\mathcal{F}:\ F\subset I} \sum_{J\in\mathcal{M}_{\mathbf{r}-\text{deep}}(F)} \left(\frac{\mathrm{P}^\alpha(J, \mathbf{1}_F\sigma)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}_{F,J}^\omega \mathbf{x} \right\|_{L^2(\omega)}^2 \\
&\le \sum_{F\in\mathcal{F}:\ F\subset I} \left(\mathcal{E}_\alpha^{\text{strong}}\right)^2 |F|_\sigma \lesssim \left(\mathcal{E}_\alpha^{\text{strong}}\right)^2 |I|_\sigma ,
\end{aligned}$$

since $\left\| \mathsf{P}^{\omega}_{F,J} x \right\|^2_{L^2(\omega)} \leq \left\| \mathsf{P}^{\text{good},\omega}_{J} \mathbf{x} \right\|^2_{L^2(\omega)}$, where we recall that the quasienergy constant $\mathcal{E}^{\text{strong}}_{\alpha}$ is defined in Definition 2.6. We also used that the stopping quasicubes $\mathcal{F}$ satisfy a σ-Carleson measure estimate,

$$\sum_{F \in \mathcal{F}:\ F \subset F_0} |F|_{\sigma} \lesssim |F_0|_{\sigma} .$$

Lemma 8.6 applies to the remaining term B to obtain the bound

$$B \lesssim \boldsymbol{\tau} \left(\left(\mathcal{E}^{\text{strong}}_{\alpha} \right)^2 + A^{\alpha,\text{energy}}_{2} \right) |I|_{\sigma} .$$

It remains then to show the inequality with ‘holes’, where the support of σ is restricted to the complement of the quasicube F. Thus for $J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)$ we may use $I \setminus F$ in the argument of the Poisson integral. We consider

$$\mathbf{Local}^{\text{hole}}(I) = \sum_{F \in \mathcal{F}} \sum_{J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F):\ J \subset I} \left(\frac{\mathrm{P}^{\alpha}\left(J, \mathbf{1}_{I \setminus F} \sigma\right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}^{\omega}_{F,J} \mathbf{x} \right\|^2_{L^2(\omega)} .$$

Lemma 8.7 *We have*

$$\mathbf{Local}^{\text{hole}}(I) \lesssim \left(\mathcal{E}^{\text{strong}}_{\alpha} \right)^2 |I|_{\sigma} . \tag{43}$$

Details are left to the reader,or see [30] or [29] for a proof. This completes the proof of

$$\begin{aligned} \mathbf{Local}(L) &\approx \sum_{F \in \mathcal{F}} \sum_{J \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F):\ J \subset L} \left(\frac{\mathrm{P}^{\alpha}\left(J, \mathbf{1}_{L} \sigma\right)}{|J|^{\frac{1}{n}}} \right)^2 \left\| \mathsf{P}^{\omega}_{F,J} \mathbf{x} \right\|^2_{L^2(\omega)} \\ &\lesssim \left(\left(\mathcal{E}^{\text{strong}}_{\alpha} \right)^2 + A^{\alpha,\text{energy}}_{2} \right) |L|_{\sigma} , \quad L \in \Omega\mathcal{D}. \end{aligned} \tag{44}$$

The Alternate Local Estimate

For future use, we prove a strengthening of the local estimate $\mathbf{Local}(L)$ to *alternate* quasicubes $M \in \mathcal{A}\Omega\mathcal{D}$.

Lemma 8.8 *With notation as above and $M \in \mathcal{A}\Omega\mathcal{D}$ an alternate quasicube, we have*

$$\begin{aligned} \mathbf{Local}(M) &\equiv \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F):\, J\subset M} \left(\frac{\mathrm{P}^{\alpha}(J,\mathbf{1}_M\sigma)}{|J|^{\frac{1}{n}}}\right)^2 \left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)} \\ &\lesssim \left(\left(\mathcal{E}^{\mathrm{strong}}_{\alpha}\right)^2 + A_2^{\alpha,\mathrm{energy}}\right)|M|_{\sigma}, \quad M\in\mathcal{A}\Omega\mathcal{D}. \end{aligned} \tag{45}$$

Again details are left to the reader, or see [30] or [29] for a proof.

The Global Estimate

Now we turn to proving the following estimate for the global part of the first testing condition (37):

$$\mathbf{Global}(I) = \int_{\mathbb{R}^{n+1}_+\setminus\widehat{I}} \mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)^2\, d\overline{\mu} \lesssim \mathcal{A}_2^{\alpha,*}|I|_{\sigma}.$$

We begin by decomposing the integral on the right into four pieces. As a particular consequence of Lemma 8.4, we note that given J, there are at most a fixed number $\boldsymbol{\tau}$ of $F\in\mathcal{F}$ such that $J\in\mathcal{M}_{\mathbf{r}-\mathrm{deep}}(F)$. We have:

$$\begin{aligned} &\int_{\mathbb{R}^{n+1}_+\setminus\widehat{I}} \mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)^2\, d\mu \le \sum_{J:\,(c_J,\ell(J))\in\mathbb{R}^{n+1}_+\setminus\widehat{I}} \mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)(c_J,\ell(J))^2 \sum_{\substack{F\in\mathcal{F}\\ J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \left\|\mathsf{P}^{\omega}_{F,J}\frac{\mathbf{x}}{|J|^{\frac{1}{n}}}\right\|^2_{L^2(\omega)} \\ &= \left\{\sum_{\substack{J\cap 3I=\emptyset\\ \ell(J)\le\ell(I)}} + \sum_{J\subset 3I\setminus I} + \sum_{\substack{J\cap I=\emptyset\\ \ell(J)>\ell(I)}} + \sum_{J\supsetneqq I}\right\} \mathbb{P}^{\alpha}(\mathbf{1}_I\sigma)(c_J,\ell(J))^2 \sum_{\substack{F\in\mathcal{F}:\\ J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \left\|\mathsf{P}^{\omega}_{F,J}\frac{\mathbf{x}}{|J|^{\frac{1}{n}}}\right\|^2_{L^2(\omega)} \\ &= A+B+C+D. \end{aligned}$$

Terms A, B and C are handled almost the same as in [30], and we leave them for the reader. As always complete details are in [29].

Finally, we turn to term D which is significantly different due to the presence of common point masses, more precisely a new *'preparation to puncture'* argument arises which is explained in detail below. The quasicubes J occurring here are included in the set of ancestors $A_k \equiv \pi^{(k)}_{\Omega\mathcal{D}}I$ of I, $1\le k<\infty$.

$$\begin{aligned}
D &= \sum_{k=1}^{\infty} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)\left(c\left(A_k\right), \left|A_k\right|^{\frac{1}{n}}\right)^2 \sum_{\substack{F\in\mathcal{F}:\\ A_k\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \left\| \mathsf{P}^{\omega}_{F,A_k} \frac{\mathbf{x}}{\left|A_k\right|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&= \sum_{k=1}^{\infty} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)\left(c\left(A_k\right), \left|A_k\right|^{\frac{1}{n}}\right)^2 \sum_{\substack{F\in\mathcal{F}:\\ A_k\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \sum_{J'\in\mathcal{C}_F^{\mathrm{good},\boldsymbol{\tau}-\mathrm{shift}}:\ J'\subset A_k\setminus I} \left\| \triangle^{\omega}_{J'} \frac{\mathbf{x}}{\left|A_k\right|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&\quad + \sum_{k=1}^{\infty} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)\left(c\left(A_k\right), \left|A_k\right|^{\frac{1}{n}}\right)^2 \sum_{\substack{F\in\mathcal{F}:\\ A_k\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \sum_{J'\in\mathcal{C}_F^{\mathrm{good},\boldsymbol{\tau}-\mathrm{shift}}:\ J'\subset I} \left\| \triangle^{\omega}_{J'} \frac{\mathbf{x}}{\left|A_k\right|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&\quad + \sum_{k=1}^{\infty} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)\left(c\left(A_k\right), \left|A_k\right|^{\frac{1}{n}}\right)^2 \sum_{\substack{F\in\mathcal{F}:\\ A_k\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \sum_{J'\in\mathcal{C}_F^{\mathrm{good},\boldsymbol{\tau}-\mathrm{shift}}:\ I\subsetneqq J'\subset A_k} \left\| \triangle^{\omega}_{J'} \frac{\mathbf{x}}{\left|A_k\right|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&\equiv D_{\mathrm{disjoint}} + D_{\mathrm{descendent}} + D_{\mathrm{ancestor}}.
\end{aligned}$$

We thus have from Lemma 8.4 again,

$$\begin{aligned}
D_{\mathrm{disjoint}} &= \sum_{k=1}^{\infty} \mathbb{P}^{\alpha}\left(\mathbf{1}_I\sigma\right)\left(c\left(A_k\right), \left|A_k\right|^{\frac{1}{n}}\right)^2 \\
&\qquad\qquad \times \sum_{\substack{F\in\mathcal{F}:\\ A_k\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\mathrm{deep}}(F)}} \sum_{J'\in\mathcal{C}_F^{\mathrm{good},\boldsymbol{\tau}-\mathrm{shift}}:\ J'\subset A_k\setminus I} \left\| \triangle^{\omega}_{J'} \frac{\mathbf{x}}{\left|A_k\right|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&\lesssim \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma \left|A_k\right|^{\frac{1}{n}}}{\left|A_k\right|^{1+\frac{1-\alpha}{n}}} \right)^2 \boldsymbol{\tau}\ \left|A_k\setminus I\right|_\omega = \boldsymbol{\tau} \left\{ \frac{|I|_\sigma}{|I|^{1-\frac{\alpha}{n}}} \sum_{k=1}^{\infty} \frac{|I|^{1-\frac{\alpha}{n}}}{\left|A_k\right|^{2\left(1-\frac{\alpha}{n}\right)}} \left|A_k\setminus I\right|_\omega \right\} |I|_\sigma \\
&\lesssim \boldsymbol{\tau} \left\{ \frac{|I|_\sigma}{|I|^{1-\frac{\alpha}{n}}} \mathcal{P}^{\alpha}\left(I, \mathbf{1}_{I^c}\omega\right) \right\} |I|_\sigma \lesssim \boldsymbol{\tau} \mathcal{A}_2^{\alpha,*} |I|_\sigma ,
\end{aligned}$$

since

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{|I|^{1-\frac{\alpha}{n}}}{\left|A_k\right|^{2\left(1-\frac{\alpha}{n}\right)}} \left|A_k\setminus I\right|_\omega &= \int \sum_{k=1}^{\infty} \frac{|I|^{1-\frac{\alpha}{n}}}{\left|A_k\right|^{2\left(1-\frac{\alpha}{n}\right)}} \mathbf{1}_{A_k\setminus I}(x)\, d\omega(x) \\
&= \int \sum_{k=1}^{\infty} \frac{1}{2^{2\left(1-\frac{\alpha}{n}\right)k}} \frac{|I|^{1-\frac{\alpha}{n}}}{|I|^{2\left(1-\frac{\alpha}{n}\right)}} \mathbf{1}_{A_k\setminus I}(x)\, d\omega(x) \\
&\lesssim \int_{I^c} \left(\frac{|I|^{\frac{1}{n}}}{\left[|I|^{\frac{1}{n}} + \mathrm{quasidist}(x, I)\right]^2} \right)^{n-\alpha} d\omega(x) = \mathcal{P}^{\alpha}\left(I, \mathbf{1}_{I^c}\omega\right).
\end{aligned}$$

The next term $D_{\mathrm{descendent}}$ satisfies

$$D_{\text{descendent}} \lesssim \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma |A_k|^{\frac{1}{n}}}{|A_k|^{1+\frac{1-\alpha}{n}}} \right)^2 \boldsymbol{\tau} \left\| \mathsf{P}_I^{\text{good},\omega} \frac{\mathbf{x}}{2^k |I|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2$$
$$= \boldsymbol{\tau} \sum_{k=1}^{\infty} 2^{-2k(n-\alpha+1)} \left(\frac{|I|_\sigma}{|I|^{1-\frac{\alpha}{n}}} \right)^2 \left\| \mathsf{P}_I^{\text{good},\omega} \frac{\mathbf{x}}{|I|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2$$
$$\lesssim \boldsymbol{\tau} \left\{ \frac{|I|_\sigma \left\| \mathsf{P}_I^{\text{good},\omega} \frac{\mathbf{x}}{|I|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2}{|I|^{2\left(1-\frac{\alpha}{n}\right)}} \right\} |I|_\sigma \lesssim \boldsymbol{\tau} A_2^{\alpha,\text{energy}} |I|_\sigma .$$

Finally for D_{ancestor} we note that each J' is of the form $J' = A_\ell \equiv \pi_{\Omega\mathcal{D}}^{(\ell)} I$ for some $\ell \geq 1$, and that there are at most $C\boldsymbol{\tau}$ pairs (F, A_k) with $k \geq \ell$ such that $A_k \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)$ and $J' = A_\ell \in \mathcal{C}_F^{\text{good},\boldsymbol{\tau}-\text{shift}}$. Now we write

$$D_{\text{ancestor}} = \sum_{k=1}^{\infty} \mathbb{P}^\alpha (\mathbf{1}_I \sigma) \left(c(A_k), |A_k|^{\frac{1}{n}} \right)^2 \sum_{\substack{F \in \mathcal{F}: \\ A_k \in \mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)}} \sum_{J' \in \mathcal{C}_F^{\text{good},\boldsymbol{\tau}-\text{shift}} :\, I \subsetneqq J' \subset A_k} \left\| \triangle_{J'}^{\omega} \frac{\mathbf{x}}{|A_k|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2$$
$$\lesssim \boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma |A_k|^{\frac{1}{n}}}{|A_k|^{1+\frac{1-\alpha}{n}}} \right)^2 \sum_{\ell=1}^{k} \left\| \triangle_{A_\ell}^{\omega} \frac{\mathbf{x}}{|A_k|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2$$
$$\leq \boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma |A_k|^{\frac{1}{n}}}{|A_k|^{1+\frac{1-\alpha}{n}}} \right)^2 \left\| \mathsf{P}_{A_k}^{\text{good},\omega} \frac{\mathbf{x}}{|A_k|^{\frac{1}{n}}} \right\|_{L^2(\omega)}^2 .$$

It is at this point that we must invoke a new *'prepare to puncture'* argument. Now define $\widetilde{\omega} = \omega - \omega(\{p\}) \delta_p$ where p is an atomic point in I for which

$$\omega(\{p\}) = \sup_{q \in \mathfrak{P}_{(\sigma,\omega)} :\, q \in I} \omega(\{q\}) .$$

(If ω has no atomic point in common with σ in I set $\widetilde{\omega} = \omega$.) Then we have $|I|_{\widetilde{\omega}} = \omega\left(I, \mathfrak{P}_{(\sigma,\omega)}\right)$ and

$$\frac{|I|_{\widetilde{\omega}}}{|I|^{\left(1-\frac{\alpha}{n}\right)}} \frac{|I|_\sigma}{|I|^{\left(1-\frac{\alpha}{n}\right)}} = \frac{\omega\left(I, \mathfrak{P}_{(\sigma,\omega)}\right)}{|I|^{\left(1-\frac{\alpha}{n}\right)}} \frac{|I|_\sigma}{|I|^{\left(1-\frac{\alpha}{n}\right)}} \leq A_2^{\alpha,\text{punct}}.$$

A key observation, already noted in the proof of Lemma 3.2 above, is that

$$\left\| \triangle_K^\omega \mathbf{x} \right\|_{L^2(\omega)}^2 = \begin{cases} \left\| \triangle_K^\omega (\mathbf{x} - \mathbf{p}) \right\|_{L^2(\omega)}^2 & \text{if } p \in K \\ \left\| \triangle_K^\omega \mathbf{x} \right\|_{L^2(\widetilde{\omega})}^2 & \text{if } p \notin K \end{cases} \leq \ell(K)^2 |K|_{\widetilde{\omega}} , \qquad \text{for all } K \in \Omega\mathcal{D}, \tag{46}$$

and so, as in the proof of Lemma 3.2,

$$\left\| \mathsf{P}^{\mathrm{good},\omega}_{A_k} \frac{\mathbf{x}}{|A_k|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \leq 3 \, |A_k|_{\widetilde{\omega}} \, .$$

Then we continue with

$$\begin{aligned}
&\boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma \, |A_k|^{\frac{1}{n}}}{|A_k|^{1+\frac{1-\alpha}{n}}} \right)^2 \left\| \mathsf{P}^{\mathrm{good},\omega}_{A_k} \frac{\mathbf{x}}{|A_k|^{\frac{1}{n}}} \right\|^2_{L^2(\omega)} \\
&\lesssim \boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma \, |A_k|^{\frac{1}{n}}}{|A_k|^{1+\frac{1-\alpha}{n}}} \right)^2 |A_k|_{\widetilde{\omega}} \\
&= \boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma}{|A_k|^{1-\frac{\alpha}{n}}} \right)^2 |A_k \setminus I|_\omega + \boldsymbol{\tau} \sum_{k=1}^{\infty} \left(\frac{|I|_\sigma}{2^{k(n-\alpha)} \, |I|^{1-\frac{\alpha}{n}}} \right)^2 |I|_{\widetilde{\omega}} \\
&\lesssim \boldsymbol{\tau} \left(\mathcal{A}_2^{\alpha,*} + A_2^{\alpha,\mathrm{punct}} \right) |I|_\sigma \, ,
\end{aligned}$$

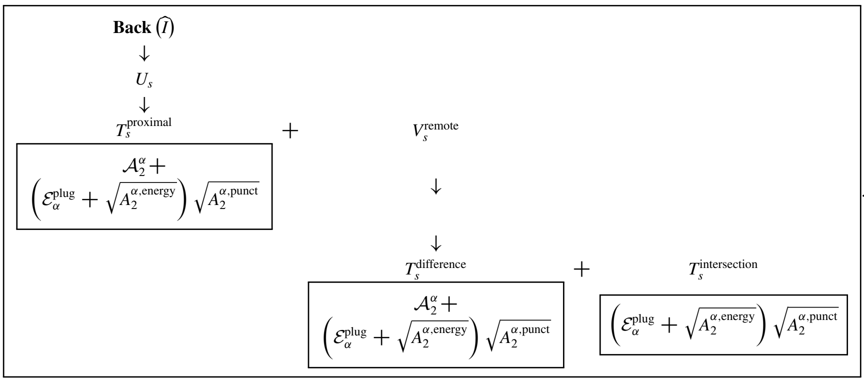
where the inequality $\sum_{k=1}^{\infty} \left(\frac{|I|_\sigma}{|A_k|^{1-\frac{\alpha}{n}}} \right)^2 |A_k \setminus I|_\omega \lesssim \mathcal{A}_2^{\alpha,*} \, |I|_\sigma$ is already proved above in the estimate for D_{disjoint}.

The Backward Poisson Testing Inequality

Fix $I \in \Omega\mathcal{D}$. It suffices to prove

$$\mathbf{Back}\left(\widehat{I}\right) \equiv \int_{\mathbb{R}^n} \left[\mathbb{Q}^\alpha \left(t \mathbf{1}_{\widehat{I}} \overline{\mu} \right)(y) \right]^2 d\sigma(y) \lesssim \left\{ \mathcal{A}_2^\alpha + \left(\mathcal{E}_\alpha^{\mathrm{plug}} + \sqrt{A_2^{\alpha,\mathrm{energy}}} \right) \sqrt{A_2^{\alpha,\mathrm{punct}}} \right\} \int_{\widehat{I}} t^2 d\overline{\mu}(x,t). \tag{47}$$

Note that in dimension $n = 1$, Hytönen obtained in [6] the simpler bound A_2^α for the term analogous to (47). Here is a brief schematic diagram of the decompositions, with bounds in $\square$, used in this subsection:



Using (42) we see that the integral on the right hand side of (47) is

$$\int_{\widehat{I}} t^2 d\overline{\mu} = \sum_{F\in\mathcal{F}} \sum_{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F):\ J\subset I} \left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)}. \tag{48}$$

where $\mathsf{P}^{\omega}_{F,J}$ was defined earlier.

We now compute using (42) again that

$$\begin{aligned}\mathbb{Q}^{\alpha}\left(t\mathbf{1}_{\widehat{I}}\overline{\mu}\right)(y) &= \int_{\widehat{I}} \frac{t^2}{\left(t^2+|x-y|^2\right)^{\frac{n+1-\alpha}{2}}} d\overline{\mu}(x,t) \\ &\approx \sum_{F\in\mathcal{F}} \sum_{\substack{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)\\ J\subset I}} \frac{\left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)}}{\left(|J|^{\frac{1}{n}}+|y-c_J|\right)^{n+1-\alpha}},\end{aligned} \tag{49}$$

and then expand the square and integrate to obtain that the term $\mathbf{Back}\left(\widehat{I}\right)$ is

$$\sum_{\substack{F\in\mathcal{F}\\ J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F)\\ J\subset I}} \sum_{\substack{F'\in\mathcal{F}:\\ J'\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F')\\ J'\subset I}} \int_{\mathbb{R}^n} \frac{\left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)}}{\left(|J|^{\frac{1}{n}}+|y-c_J|\right)^{n+1-\alpha}} \frac{\left\|\mathsf{P}^{\omega}_{F',J'}\mathbf{x}\right\|^2_{L^2(\omega)}}{\left(|J'|^{\frac{1}{n}}+|y-c_{J'}|\right)^{n+1-\alpha}} d\sigma(y).$$

By symmetry we may assume that $\ell(J') \leq \ell(J)$. We fix an integer s, and consider those quasicubes J and J' with $\ell(J') = 2^{-s}\ell(J)$. For fixed s we will control the expression

$$\begin{aligned}U_s \equiv &\sum_{F,F'\in\mathcal{F}} \sum_{\substack{J\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F),\ J'\in\mathcal{M}_{(\mathbf{r},\varepsilon)-\text{deep}}(F')\\ J,J'\subset I,\ \ell(J')=2^{-s}\ell(J)}} \\ &\times \int_{\mathbb{R}^n} \frac{\left\|\mathsf{P}^{\omega}_{F,J}\mathbf{x}\right\|^2_{L^2(\omega)}}{\left(|J|^{\frac{1}{n}}+|y-c_J|\right)^{n+1-\alpha}} \frac{\left\|\mathsf{P}^{\omega}_{F',J'}\mathbf{x}\right\|^2_{L^2(\omega)}}{\left(|J'|^{\frac{1}{n}}+|y-c_{J'}|\right)^{n+1-\alpha}} d\sigma(y),\end{aligned}$$

by proving that

$$U_s \lesssim 2^{-\delta s}\left\{\mathcal{A}_2^{\alpha} + \left(\mathcal{E}_{\alpha}^{\text{strong}} + \sqrt{A_2^{\alpha,\text{energy}}}\right)\sqrt{A_2^{\alpha,\text{punct}}}\right\} \int_{\widehat{I}} t^2 d\overline{\mu}, \quad \text{where } \delta = \frac{1}{2n}. \tag{50}$$

With this accomplished, we can sum in $s \geq 0$ to control the term $\mathbf{Back}\left(\widehat{I}\right)$. The remaining details of the proof are very similar to the corresponding arguments in

[30], with the only exception being the repeated use of the *'prepare to puncture'* argument above whenever the measures σ and ω can 'see each other' in an estimate. We refer the reader to [29] for complete details.[9]

The Stopping Form

This section is virtually unchanged from the corresponding section in [30], so we content ourselves with a brief recollection. In the one-dimensional setting of the Hilbert transform, Hytönen [6] observed that "…the innovative verification of the local estimate by Lacey [8] is already set up in such a way that it is ready for us to borrow as a black box." The same observation carried over in spirit regarding the adaptation of Lacey's recursion and stopping time to proving the local estimate in [30]. However, that adaptation involved the splitting of the stopping form into two sublinear forms, the first handled by methods in [16], and the second by the methods in [8]. The arguments are little changed when including common point masses, and we leave them for the reader (or see [29] for the proofs written out in detail).

Energy Dispersed Measures

In this final section we prove that the energy side conditions in our main theorem hold if both measures are appropriately energy dispersed. We begin with the definitions of energy dispersed and reversal of energy.

Energy Dispersed Measures and Reversal of Energy

Let μ be a locally finite positive Borel measure on $\mathbb{R}^n$. Recall that for $0 \leq k \leq n$, we denote by $\mathcal{L}_k^n$ the collection of all k-dimensional planes in $\mathbb{R}^n$, and for a quasicube J, we define the *k-dimensional second moment* $\mathsf{M}_k^n(J,\mu)$ of μ on J by

$$\mathsf{M}_k^n(J,\mu)^2 \equiv \inf_{L\in\mathcal{L}_k^n} \int_J \operatorname{dist}(x,L)^2\, d\mu(x)\,.$$

Finally we defined μ to be *k-energy dispersed* if there is $c>0$ such that

$$\mathsf{M}_k^n(J,\mu) \geq c\mathsf{M}_0^n(J,\mu)\,, \qquad \text{for all quasicubes } J \text{ in } \mathbb{R}^n.$$

[9]In [28] and [30] the bound for term B in the global estimate was mistakenly claimed without proof to be simply $\mathcal{A}_2^\alpha$ instead of the correct bound $\mathcal{A}_2^\alpha + \left(\mathcal{E}_\alpha^{\text{plug}} + \sqrt{A_2^{\alpha,\text{energy}}}\right)\sqrt{A_2^{\alpha,\text{punct}}}$ given in [29].

In order to introduce a useful reformulation of the k-dimensional second moment, we will use the observation that minimizing k-planes L pass through the center of mass. More precisely, for any k-plane $L \in \mathcal{L}_k^n$ such that $\int_A \operatorname{dist}(x,L)^2\, d\mu(x)$ is minimized, where A is a set of positive μ-measure, we claim that

$$\mathbb{E}_A^{\mu} x \in L.$$

Indeed, if we rotate coordinates so that $L = \left\{\left(x^1,\dots,x^k,a^{k+1},\dots,a^n\right) : \left(x^1,\dots,x^k\right) \in \mathbb{R}^k\right\}$, then

$$\begin{aligned}
\int_A \operatorname{dist}(x,L)^2\, d\mu(x) &= \int_A \sum_{j=k+1}^{n} \left(x^j - a^j\right)^2 d\mu(x) \\
&= \sum_{j=k+1}^{n} \left[\int_A \left(x^j\right)^2 d\mu(x) - 2a^j \int_A x^j d\mu(x) + \left(a^j\right)^2 \int_A d\mu(x)\right] \\
&= \sum_{j=k+1}^{n} \left[\int_A \left(x^j\right)^2 d\mu(x) + \left(\int_A d\mu(x)\right)\left\{\left(a^j\right)^2 - 2\frac{\int_A x^j d\mu(x)}{\int_A d\mu(x)} a^j\right\}\right]
\end{aligned}$$

is minimized over $a^{k+1},\dots,a^n$ when

$$a^j = \frac{\int_A x^j d\mu(x)}{\int_A d\mu(x)} = \left(\mathbb{E}_A^{\mu} x\right)^j, \qquad k+1 \le j \le n.$$

This shows that the point $\mathbb{E}_A^{\mu} x$ belongs to the k-plane L.

Now we can obtain our reformulation of the k-dimensional second moment. Let $\mathcal{S}_k^n$ denote the collection of k-dimenional subspaces in $\mathbb{R}^n$. If $\mathcal{P}_S$ denotes orthogonal projection onto the subspace $S \in \mathcal{S}_{n-k}^n$ where $S = L_0^{\perp}$ and $L_0 \in \mathcal{S}_k^n$ is the subspace parallel to L, then we have the variance identity,

$$\begin{aligned}
\mathsf{M}_k^n(J,\mu)^2 &= \inf_{L \in \mathcal{L}_k^n} \int_J \operatorname{dist}(x,L)^2\, d\mu(x) = \inf_{S \in \mathcal{S}_{n-k}^n} \int_J \left|\mathcal{P}_S x - \mathcal{P}_S\left(\mathbb{E}_J^{\mu} x\right)\right|^2 d\mu(x) \quad (51) \\
&= \frac{1}{2} \inf_{S \in \mathcal{S}_{n-k}^n} \frac{1}{|J|_{\mu}} \int_J \int_J \left|\mathcal{P}_S x - \mathcal{P}_S y\right|^2 d\mu(x)\, d\mu(y) \\
&= \frac{1}{2} \inf_{L_0 \in \mathcal{S}_k^n} \frac{1}{|J|_{\mu}} \int_J \int_J \operatorname{dist}(x, L_0 + y)^2\, d\mu(x)\, d\mu(y),
\end{aligned}$$

since $\mathcal{P}_S\left(\mathbb{E}_J^{\mu} x\right) = \mathbb{E}_J^{\mu}(\mathcal{P}_S x)$. Here we have used in the first line the fact that the minimizing k-planes L pass through the center of mass $\mathbb{E}_J^{\mu} x$ of x in J.

Note that if μ is supported on a k-dimensional plane L in $\mathbb{R}^n$, then $\mathsf{M}_k^n(J,\mu)$ vanishes for all quasicubes J. On the other hand, $\mathsf{M}_0^n(J,\mu)$ is positive for any quasicube J on which the restriction of μ is *not* a point mass, and we conclude that measures μ supported on a k-plane. and whose restriction to J is not a point mass, are *not* k-energy dispersed. Thus $\mathsf{M}_k^n(J,\mu)$ measures the extent to which a certain 'energy' of μ is not localized to a k-plane. In this final section we will prove the necessity of the energy conditions for boundedness of the vector Riesz transform $\mathbf{R}^{\alpha,n}$ when the locally finite Borel measures σ and ω on $\mathbb{R}^n$ are k-energy dispersed with

$$\begin{cases} n-k<\alpha<n,\ \alpha\neq n-1 & \text{if } 1\leq k\leq n-2 \\ 0\leq\alpha<n,\ \alpha\neq 1,n-1 & \text{if } \quad k=n-1 \end{cases}. \tag{52}$$

Now we recall the definition of strong energy reversal from [25]. We say that a vector $\mathbf{T}^\alpha=\left\{T_\ell^\alpha\right\}_{\ell=1}^2$ of α-fractional transforms in the plane has *strong* reversal of ω-energy on a cube J if there is a positive constant C_0 such that for all $2\leq\gamma\leq 2^{\mathbf{r}(1-\varepsilon)}$ and for all positive measures μ supported outside γJ, we have the inequality

$$\mathbb{E}_J^\omega\left[\left(\mathbf{x}-\mathbb{E}_J^\omega\mathbf{x}\right)^2\right]\left(\frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}}\right)^2=\mathsf{E}(J,\omega)^2\,\mathrm{P}^\alpha(J,\mu)^2\leq C_0\,\mathbb{E}_J^\omega\left|\mathbf{T}^\alpha\mu-\mathbb{E}_J^{d\omega}\mathbf{T}^\alpha\mu\right|^2, \tag{53}$$

Now note that if ω is k-energy dispersed, then we have

$$\mathsf{E}(J,\omega)^2=\frac{1}{|J|_\omega|J|^{\frac{2}{n}}}\mathsf{M}_0^n(J,\omega)^2\lesssim\frac{1}{|J|_\omega|J|^{\frac{2}{n}}}\mathsf{M}_k^n(J,\omega)^2\equiv\mathsf{E}_k(J,\omega)^2,$$

and where we have defined on the right hand side the analogous notion of energy $\mathsf{E}_k(J,\omega)$ in terms of $\mathsf{M}_k(J,\omega)$, and which is smaller than $\mathsf{E}(J,\omega)$. We now state the main result of this first subsection.

Lemma 10.1 *Let* $0\leq\alpha<n$. *Suppose that* ω *is* k*-energy dispersed and that* k *and* α *satisfy* (52). *Then the* α*-fractional Riesz transform* $\mathbf{R}^{\alpha,n}=\left\{R_\ell^{n,\alpha}\right\}_{\ell=1}^n$ *has strong reversal* (53) *of* ω*-energy on* all *cubes* J *provided* γ *is chosen large enough depending only on* n *and* α.

In [27] we showed that energy reversal can fail spectacularly for measures in general, but left open the possibility of reversing at least one direction in the energy for $\mathbf{R}^{\alpha,n}$ when $\alpha\neq 1$ in the plane $n=2$, and we will show in the next subsection that this is indeed possible, with even more directions included in higher dimensions.

Fractional Riesz Transforms and Semi-harmonicity

Now we fix $1\leq\ell\leq n$ and write $x=(x',x'')$ with $x'=(x_1,\ldots,x_\ell)\in\mathbb{R}^\ell$ and $x''=(x_{\ell+1},\ldots,x_n)\in\mathbb{R}^{n-\ell}$ (when $\ell=n$ we have $x=x'$). Then we compute for β real that

$$\begin{aligned}
\triangle_{x'}|x|^{\beta} &= \triangle_{x'}\left(|x'|^2+|x''|^2\right)^{\frac{\beta}{2}} = \nabla_{x'}\cdot\nabla_{x'}\left(|x'|^2+|x''|^2\right)^{\frac{\beta}{2}} \\
&= \nabla_{x'}\cdot\left\{\frac{\beta}{2}\left(|x'|^2+|x''|^2\right)^{\frac{\beta}{2}-1}2x'\right\} = \beta\nabla_{x'}\cdot\left\{x'\left(|x'|^2+|x''|^2\right)^{\frac{\beta}{2}-1}\right\} \\
&= \beta\left\{\left(\nabla_{x'}\cdot x'\right)\left(|x'|^2+|x''|^2\right)^{\frac{\beta-2}{2}} + x'\cdot\nabla_{x'}\left(|x'|^2+|x''|^2\right)^{\frac{\beta-2}{2}}\right\} \\
&= \beta\left\{\ell\left(|x'|^2+|x''|^2\right)^{\frac{\beta-2}{2}} + x'\cdot\frac{\beta-2}{2}\left(|x'|^2+|x''|^2\right)^{\frac{\beta-2}{2}-1}2x'\right\} \\
&= \beta\left\{\ell\left(|x'|^2+|x''|^2\right)^{\frac{\beta-2}{2}} + (\beta-2)|x'|^2\left(|x'|^2+|x''|^2\right)^{\frac{\beta-4}{2}}\right\} \\
&= \beta\left\{\ell\left(|x'|^2+|x''|^2\right)\left(|x'|^2+|x''|^2\right)^{\frac{\beta-4}{2}} + (\beta-2)|x'|^2\left(|x'|^2+|x''|^2\right)^{\frac{\beta-4}{2}}\right\} \\
&= \beta\left\{(\ell+\beta-2)|x'|^2+\ell|x''|^2\right\}\left(|x'|^2+|x''|^2\right)^{\frac{\beta-4}{2}}.
\end{aligned}$$

The case of interest for us is when $\beta=\alpha-n+1$, since then

$$\triangle_{x'}|x|^{\beta} = \nabla_{x'}\cdot\nabla_{x'}|x|^{\alpha-n+1} = \nabla_{x'}\cdot\nabla|x|^{\alpha-n+1} = c_{\alpha,n}\nabla_{x'}\cdot\mathbf{K}^{\alpha,n}(x), \tag{54}$$

where $\mathbf{K}^{\alpha,n}$ is the vector convolution kernel of the α-fractional Riesz transform $\mathbf{R}^{\alpha,n}$. Now if $\ell=1$ in this case, then the factor

$$F_{\ell,\beta}(x) \equiv (\ell+\beta-2)\left|x'\right|^2+\ell\left|x''\right|^2$$

is $(\beta-1)|x'|^2+|x''|^2$, and thus in dimension $n\geq 2$, the factor $F_{1,\beta}(x)$ will be of one sign for all x if and only if $\alpha-n+1=\beta>1$, i.e. $\alpha>n$, which is of no use since the Riesz transform $\mathbf{R}^{\alpha,n}$ is defined only for $0\leq\alpha<n$.

Thus we must assume $\ell\geq 2$ and $\beta=\alpha-n+1$ when $n\geq 2$. Under these assumptions, we then note that $F_{\ell,\beta}(x)$ will be of one sign for all x if $\ell+\beta-2>0$, i.e. $\alpha>n+1-\ell$, in which case we conclude that

$$\begin{aligned}
\left|\triangle_{x'}|x|^{\alpha-n+1}\right| &= |\alpha-n+1|\left\{(\ell+\alpha-n-1)|x'|^2+\ell|x''|^2\right\}\left(|x'|^2+|x''|^2\right)^{\frac{\alpha-n-3}{2}} \quad (55)\\
&\approx \left(|x'|^2+|x''|^2\right)^{\frac{\alpha-n-1}{2}} = |x|^{\alpha-n-1}, \quad \text{for } \alpha\neq n-1.
\end{aligned}$$

When $\ell=n$, this shows that $\left|\triangle_x|x|^{\alpha-n+1}\right|\approx|x|^{\alpha-n-1}$ for $\alpha>1$ with $\alpha\neq n-1$. But in the case $\ell=n$ we can obtain more. Indeed, since x'' is no longer present, we have for $0\leq\alpha<1$ that

$$\triangle_x |x|^{\alpha-n+1} \approx |x|^{\alpha-n-1} .$$

(This includes dimension $n = 1$ but only for $0 < \alpha < 1$).

We summarize these results as follows. For dimension $n \geq 2$ and $x = (x', x'')$ with $x' \in \mathbb{R}^{\ell}$ and $x'' \in \mathbb{R}^{n-\ell}$, we have

$$\left| \triangle_{x'} |x|^{\alpha-n+1} \right| \approx |x|^{\alpha-n-1} ,$$

provided

$$\begin{aligned} &\textbf{either } 2 \leq \ell \leq n-1 \text{ and } n+1-\ell < \alpha < n \text{ with } \alpha \neq n-1, \\ &\textbf{or } \ell = n \text{ and } 0 \leq \alpha < n \text{ with } \alpha \neq 1, n-1. \end{aligned} \tag{56}$$

Thus the two cases not included are $\alpha = 1$ and $\alpha = n - 1$. The case $\alpha = 1$ is not included since $|x|^{\alpha-n+1} = |x|^{2-n}$ is the fundamental solution of the Laplacian for $n > 2$ and constant for $n = 2$. The case $\alpha = n-1$ is not included since $|x|^{\alpha-n+1} = 1$ is constant.

So we now suppose that α and ℓ are as in (56), and we consider ℓ-planes L intersecting the cube J. Recall that the trace of a matrix is invariant under rotations. Thus for each such ℓ-plane L, and for $z \in J \cap L$, we have from (54) and (55), and with $\mathbf{I}^{\alpha+1,n}\mu(z) \equiv \int_{\mathbb{R}^n} |z-y|^{\alpha+1-n} d\mu(y)$ denoting the convolution of $|x|^{\alpha+1-n}$ with μ, that

$$|\nabla_L \mathbf{R}^{\alpha,n}\mu(z)| \gtrsim |\text{trace}\nabla_L \mathbf{R}^{\alpha,n}\mu(z)| = \left|\triangle_L \mathbf{I}^{\alpha+1,n}\mu(z)\right| \approx \int |y-z|^{\alpha-n-1} d\mu(y) \approx \frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}, \tag{57}$$

where ∇_L denotes the gradient in the ℓ-plane L, i.e. $\nabla_L = \mathcal{P}_S \nabla$ where S is the subspace parallel to L and $\mathcal{P}_S$ is orthogonal projection onto S, and where we assume that the positive measure μ is supported outside the expanded cube γJ.

We now claim that for every $z \in J \cap L$, the full matrix gradient $\nabla \mathbf{R}^{\alpha,n}\mu(z)$ is 'missing' at most $\ell - 1$ 'large' directions, i.e. has at least $n - \ell + 1$ eigenvalues each of size at least $c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}$. Indeed, to see this, suppose instead that the matrix $\nabla \mathbf{R}^{\alpha,n}\mu(z)$ has at most $n - \ell$ eigenvalues of size at least $c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}$. Then there is an ℓ-dimensional subspace S such that

$$|\nabla_S \mathbf{R}^{\alpha,n}\mu(z)| = |(\mathcal{P}_S\nabla)\,\mathbf{R}^{\alpha,n}\mu(z)| = |\mathcal{P}_S(\nabla \mathbf{R}^{\alpha,n}\mu(z))| \leq c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}},$$

which contradicts (57) if c is chosen small enough. This proves our claim, and moreover, it satisfies the quantitative quadratic estimate

$$|\xi \cdot \nabla \mathbf{R}^{\alpha,n}\mu(z)\,\xi| \geq c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}} |\xi|^2 ,$$

for all vectors ξ in some $(n-\ell+1)$-dimensional subspace

$$\mathsf{S}_z^{n-\ell+1} \equiv \text{Span}\left\{\mathbf{v}_z^1, \dots, \mathbf{v}_z^{n-\ell+1}\right\} \in \mathcal{S}_{n-\ell+1}^n,$$

with $\mathbf{v}_z^j \in \mathbb{S}^{n-1}$ for $1 \leq j \leq n-\ell+1$.

It is convenient at this point to let

$$k = \ell - 1,$$

so that $1 \leq k \leq n-1$ and the assumptions (56) become

$$\begin{aligned} &\textbf{either } 1 \leq k \leq n-2 \text{ and } n-k < \alpha < n \text{ with } \alpha \neq n-1, \\ &\textbf{or } k = n-1 \text{ and } 0 \leq \alpha < n \text{ with } \alpha \neq 1, n-1, \end{aligned} \tag{58}$$

and our conclusion becomes

$$\left|\xi \cdot \nabla \mathbf{R}^{\alpha,n}\mu\left(z\right)\xi\right| \geq c \frac{\mathrm{P}^\alpha\left(J,\mu\right)}{\left|J\right|^{\frac{1}{n}}} \left|\xi\right|^2, \qquad \xi \in \mathsf{S}_z^{n-k},\ z \in J. \tag{59}$$

Proof of Strong Reversal of Energy

We are now in a position to prove the strong reversal of energy for Riesz transforms in Lemma 10.1.

Proof (of Lemma 10.1) Recall that $\mathsf{E}_k\left(J,\omega\right)^2 = \inf_{L \in \mathcal{L}_k^n} \frac{1}{|J|_\omega} \int_J \left(\frac{\text{dist}(x,L)}{|J|^{\frac{1}{n}}}\right)^2 d\omega\left(x\right)$ and

$$\frac{1}{|J|_\omega} \int_J \left(\frac{\text{dist}\left(x,L\right)}{|J|^{\frac{1}{n}}}\right)^2 d\omega\left(x\right) = \frac{1}{2} \frac{1}{|J|_\omega} \int_J \frac{1}{|J|_\omega} \int_J \left(\frac{\text{dist}\left(x,z+L_0\right)}{|J|^{\frac{1}{n}}}\right)^2 d\omega\left(x\right) d\omega\left(z\right), \tag{60}$$

where we recall that $L_0 \in \mathcal{S}_k^n$ is parallel to L. The real matrix

$$M\left(x\right) \equiv \nabla \mathbf{R}^{\alpha,n}\mu\left(x\right), \qquad x \in J, \tag{61}$$

is a scalar multiple of the Hessian of $|x|^{\alpha+1}$, hence is symmetric, and so we can rotate coordinates to diagonalize the matrix,

$$M\left(x\right) = \begin{bmatrix} \lambda_1\left(x\right) & 0 & \cdots & 0 \\ 0 & \lambda_2\left(x\right) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n\left(x\right) \end{bmatrix},$$

where $|\lambda_1(x)| \leq |\lambda_2(x)| \leq \ldots \leq |\lambda_n(x)|$. We now fix $x = c_J$ to be the center of J in the matrix $M(c_J)$ and fix the eigenvalues corresponding to $M(c_J)$:

$$|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_n|, \qquad \lambda_j \equiv \lambda_j(c_J),$$

and define also the subspaces S^{n-i} to be $\mathsf{S}^{n-i}_{c_J}$ for $1 \leq i \leq k$. Note that we then have $\mathsf{S}^{n-i} = \mathrm{Span}\{\mathbf{e}_{i+1}, \ldots, \mathbf{e}_n\}$. Let L^i_z be the i-plane

$$L^i_z \equiv z + \left(\mathsf{S}^{n-i}\right)^{\perp} = \left\{\left(u^1, \ldots, u^i, z^{i+1}, \ldots, z^n\right) : \left(u^1, \ldots, u^i\right) \in \mathbb{R}^i\right\}. \qquad (62)$$

By (59) we have

$$|\lambda_{k+1}| \geq c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}.$$

For convenience define $|\lambda_0| \equiv 0$ and then define $0 \leq m \leq k$ be the unique integer such that

$$|\lambda_m| < c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}} \leq |\lambda_{m+1}|. \qquad (63)$$

Now consider the largest $0 \leq \ell \leq m$ that satisfies

$$|\lambda_{\ell}| \leq \gamma^{-\frac{1}{2n}} |\lambda_{\ell+1}|. \qquad (64)$$

Note that this use of ℓ is quite different than that used in (56).

So suppose first that ℓ satisfies $1 \leq \ell \leq m$ and is the largest index satisfying (64). Then if $\ell < m$ we have $|\lambda_i| > \gamma^{-\frac{1}{2n}} |\lambda_{i+1}|$ for $\ell + 1 \leq i \leq m$, and so both

$$|\lambda_{\ell+1}| > \gamma^{-\frac{1}{2n}} |\lambda_{\ell+2}| > \ldots > \gamma^{-\frac{m-\ell}{2n}} |\lambda_{m+1}| \geq \gamma^{-\frac{m-\ell}{2n}} c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}, \qquad (65)$$

$$|\lambda_1| \leq \ldots \leq |\lambda_{\ell}| \leq \gamma^{-\frac{1}{2n}} |\lambda_{\ell+1}|.$$

Both inequalities in the display above also hold for $\ell = m$ by (63) and (64). Roughly speaking, in this case where $1 \leq \ell \leq m$, the gradient of $\mathbf{R}^{\alpha,n}\mu$ has modulus at least $|\lambda_{\ell+1}|$ in the directions of $\mathbf{e}_{\ell+1}, \ldots, \mathbf{e}_n$, while the gradient of $\mathbf{R}^{\alpha,n}\mu$ has modulus at most $\gamma^{-\frac{1}{2n}} |\lambda_{\ell+1}|$ in the directions of $\mathbf{e}_1, \ldots, \mathbf{e}_{\ell}$.

Recall that $\mathsf{S}^{n-\ell} = \mathsf{S}^{n-\ell}_{c_J}$ is the subspace on which the symmetric matrix $M(c_J) = \nabla(\mathbf{R}^{\alpha,n}\mu)(c_J)$ has energy $\xi^{\mathrm{tr}} M(c_J)\xi$ bounded below by $|\lambda_{\ell+1}|$. Now we proceed to show that

$$|\lambda_{\ell+1}|^2 |J|^{\frac{2}{n}} \mathsf{E}(J,\omega)^2 \lesssim \frac{1}{|J|^2_{\omega}} \int_J \int_J |\mathbf{R}^{\alpha,n}\mu(x) - \mathbf{R}^{\alpha,n}\mu(z)|^2 \, d\omega(x) \, d\omega(z). \quad (66)$$

We will use our hypothesis that ω is k-energy dispersed to obtain

$$\mathsf{E}(J,\omega) \le \mathsf{E}_k(J,\omega) \le \mathsf{E}_m(J,\omega) \le \mathsf{E}_\ell(J,\omega)$$

since $\ell \le m \le k$. To prove (66), we take $L_z \equiv L_z^\ell$ as in (62) and begin with

$$\begin{aligned} \operatorname{dist}(x, L_z)^2 &= \operatorname{dist}\left(x, z + \left(\mathsf{S}^{n-\ell}\right)^\perp\right)^2 \qquad (67) \\ &= (x_{\ell+1} - z_{\ell+1})^2 + \ldots + (x_n - z_n)^2 = \left|x'' - z''\right|^2, \end{aligned}$$

where $x = (x', x'')$ with $x' \in \mathbb{R}^\ell$ and $x'' \in \mathbb{R}^{n-\ell}$, and $L_z = \left\{(u', z'') : u' \in \mathbb{R}^\ell\right\}$. Now for $x, z \in J$ we take $\xi \equiv \left(0, \frac{x''-z''}{|x''-z''|}\right) \in \mathsf{S}^{n-\ell}$ (where $\frac{0}{0} = 0$). We use the estimate

$$|J|^{\frac{1}{n}} \left\|\nabla^2 \mathbf{R}^{\alpha,n}\mu\right\|_{L^\infty(J)} \lesssim |J|^{\frac{1}{n}} \int_{\mathbb{R}^n \setminus \gamma J} \frac{d\mu(y)}{|y - c_J|^{n-\alpha+2}} \lesssim \frac{1}{\gamma} \int_{\mathbb{R}^n \setminus \gamma J} \frac{d\mu(y)}{|y - c_J|^{n-\alpha+1}} \approx \frac{1}{\gamma} \frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}}, \qquad (68)$$

to obtain

$$\begin{aligned} &\frac{1}{|J|_\omega^2} \int_J \int_J \left(\left\|\nabla^2 \mathbf{R}^{\alpha,n}\mu\right\|_{L^\infty(J)} |x - z|\, |J|^{\frac{1}{n}}\right)^2 d\omega(x)\, d\omega(z) \qquad (69) \\ &\lesssim \frac{1}{\gamma^2} \mathrm{P}^\alpha(J,\mu)^2 \frac{1}{|J|_\omega^2} \int_J \int_J \left(\frac{|x - z|}{|J|^{\frac{1}{n}}}\right)^2 d\omega(x)\, d\omega(z) = \frac{1}{\gamma^2} \mathrm{P}^\alpha(J,\mu)^2\, \mathsf{E}(J,\omega)^2. \end{aligned}$$

We then start with a decomposition into big B and small S pieces,

$$\begin{aligned} &\frac{1}{|J|_\omega^2} \int_J \int_J \left|\mathbf{R}^{\alpha,n}\mu(x) - \mathbf{R}^{\alpha,n}\mu(z)\right|^2 d\omega(x)\, d\omega(z) \\ &\gtrsim \frac{1}{|J|_\omega^2} \int_J \int_J \left|\mathbf{R}^{\alpha,n}\mu\left(z', x''\right) - \mathbf{R}^{\alpha,n}\mu\left(z', z''\right)\right|^2 d\omega(x)\, d\omega(z) \\ &\quad - \frac{1}{|J|_\omega^2} \int_J \int_J \left|\mathbf{R}^{\alpha,n}\mu\left(x', x''\right) - \mathbf{R}^{\alpha,n}\mu\left(z', x''\right)\right|^2 d\omega(x)\, d\omega(z) \\ &\equiv B - S. \end{aligned}$$

For $w \in J$ we have

$$\begin{aligned} \left|\nabla \mathbf{R}^{\alpha,n}\mu(w) - M(c_J)\right| &= \left|\nabla \mathbf{R}^{\alpha,n}\mu(w) - \nabla \mathbf{R}^{\alpha,n}\mu(c_J)\right| \qquad (70) \\ &\lesssim |w - c_J| \left\|\nabla^2 \mathbf{R}^{\alpha,n}\mu\right\|_{L^\infty(J)} \lesssim \frac{1}{\gamma} \frac{\mathrm{P}^\alpha(J,\mu)}{|J|^{\frac{1}{n}}}, \end{aligned}$$

from (68), and this inequality will allow us to replace x or z by c_J at appropriate places in the estimates below, introducing a harmless error. We now use the second inequality in (65) with the diagonal form of $M(c_J) = \nabla \mathbf{R}^{\alpha,n}\mu(c_J)$, along with the error estimates (69) and (70), to control S by

$$\begin{aligned} S &\leq \frac{1}{|J|_\omega^2} \int_J \int_J \left| \left(x' - z'\right) \cdot \nabla' \mathbf{R}^{\alpha,n}\mu(x) \right|^2 d\omega(x)\, d\omega(z) \\ &\quad + \frac{1}{|J|_\omega^2} \int_J \int_J \left\{ \left\| \nabla^2 \mathbf{R}^{\alpha,n}\mu \right\|_{L^\infty(J)} \left| x' - z' \right|^2 \right\}^2 d\omega(x)\, d\omega(z) \\ &\lesssim \frac{1}{|J|_\omega^2} \int_J \int_J \left| \left(x' - z'\right) \cdot \nabla' \mathbf{R}^{\alpha,n}\mu(c_J) \right|^2 d\omega(x)\, d\omega(z) \\ &\quad + \frac{1}{|J|_\omega^2} \int_J \int_J \left\{ \left\| \nabla^2 \mathbf{R}^{\alpha,n}\mu \right\|_{L^\infty(J)} \left| x' - z' \right| |J|^{\frac{1}{n}} \right\}^2 d\omega(x)\, d\omega(z), \end{aligned}$$

and then continuing with

$$\begin{aligned} S &\lesssim \frac{1}{|J|_\omega^2} \int_J \int_J \left\{ \left| x' - z' \right| \ |\lambda_\ell| \right\}^2 d\omega(x)\, d\omega(z) + \frac{1}{\gamma^2} \mathrm{P}^\alpha(J,\mu)^2 \mathsf{E}(J,\omega)^2 \\ &\lesssim \frac{1}{\gamma} |\lambda_{\ell+1}|^2 \frac{1}{|J|_\omega^2} \int_J \int_J |x - z|^2 d\omega(x)\, d\omega(z) + \frac{1}{\gamma^2} \mathrm{P}^\alpha(J,\mu)^2 \mathsf{E}(J,\omega)^2 \\ &= \frac{1}{\gamma} |J|^{\frac{2}{n}} |\lambda_{\ell+1}|^2 \mathsf{E}(J,\omega)^2 + \frac{1}{\gamma^2} \mathrm{P}^\alpha(J,\mu)^2 \mathsf{E}(J,\omega)^2, \end{aligned}$$

which is small enough to be absorbed later on in the proof. To bound term B from below we use (70) in

$$\begin{aligned} \mathbf{R}^{\alpha,n}\mu(z', x'') - \mathbf{R}^{\alpha,n}\mu(z', z'') &= \left(x'' - z''\right) \cdot \nabla'' \mathbf{R}^{\alpha,n}\mu(z) + O\left(\left\| \nabla^2 \mathbf{R}^{\alpha,n}\mu \right\|_{L^\infty(J)} |x - z|^2 \right) \\ &= \left(x'' - z''\right) \cdot \nabla'' \mathbf{R}^{\alpha,n}\mu(c_J) + O\left(\left\| \nabla^2 \mathbf{R}^{\alpha,n}\mu \right\|_{L^\infty(J)} |x - z| \, |J|^{\frac{1}{n}} \right), \end{aligned}$$

and then (59) with the choice $\xi \equiv \left(0, \frac{x''-z''}{|x''-z''|}\right) \in \mathsf{S}^{n-\ell}$, to obtain

$$\begin{aligned} \left| x'' - z'' \right| |\lambda_{\ell+1}| &\leq \left| x'' - z'' \right| \left| \left(\xi \cdot \nabla''\right) \mathbf{R}^{\alpha,n}\mu(c_J) \cdot \xi \right| \\ &= \left| \left(x'' - z''\right) \cdot \nabla'' \mathbf{R}^{\alpha,n}\mu(c_J) \cdot \xi \right| \\ &\leq \left| \left(x'' - z''\right) \cdot \nabla'' \mathbf{R}^{\alpha,n}\mu(c_J) \right| \\ &\leq \left| \mathbf{R}^{\alpha,n}\mu(z', x'') - \mathbf{R}^{\alpha,n}\mu(z', z'') \right| + O\left(\left\| \nabla^2 \mathbf{R}^{\alpha,n}\mu \right\|_{L^\infty(J)} |x - z| \, |J|^{\frac{1}{n}} \right). \end{aligned}$$

Then using (69) and (70) we continue with

$$\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}\left|\mathbf{R}^{\alpha,n}\mu\left(z',x''\right)-\mathbf{R}^{\alpha,n}\mu\left(z',z''\right)\right|^{2}d\omega\left(x\right)d\omega\left(z\right)$$
$$\gtrsim|\lambda_{\ell+1}|^{2}\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}\left|x''-z''\right|^{2}d\omega\left(x\right)d\omega\left(z\right)-\frac{1}{\gamma^{2}}\mathrm{P}^{\alpha}\left(J,\mu\right)^{2}\mathsf{E}\left(J,\omega\right)^{2},$$

and then

$$|\lambda_{\ell+1}|^{2}|J|^{\frac{2}{n}}\mathsf{E}\left(J,\omega\right)^{2}\leq C|\lambda_{\ell+1}|^{2}|J|^{\frac{2}{n}}\mathsf{E}_{\ell}\left(J,\omega\right)^{2} \tag{71}$$
$$=|\lambda_{\ell+1}|^{2}\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}\mathrm{dist}\left(x,L_{z}\right)^{2}d\omega\left(x\right)d\omega\left(z\right)=|\lambda_{\ell+1}|^{2}\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|x''-z''|^{2}d\omega\left(x\right)d\omega\left(z\right)$$
$$\lesssim\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(z',x''\right)-\mathbf{R}^{\alpha,n}\mu\left(z',z''\right)|^{2}d\omega\left(x\right)d\omega\left(z\right)+\frac{1}{\gamma^{2}}\mathrm{P}^{\alpha}\left(J,\mu\right)^{2}\mathsf{E}\left(J,\omega\right)^{2}$$
$$\lesssim\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(x\right)-\mathbf{R}^{\alpha,n}\mu\left(z\right)|^{2}d\omega\left(x\right)d\omega\left(z\right)+S+\frac{1}{\gamma^{2}}\mathrm{P}^{\alpha}\left(J,\mu\right)^{2}\mathsf{E}\left(J,\omega\right)^{2}$$
$$\lesssim\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(x\right)-\mathbf{R}^{\alpha,n}\mu\left(z\right)|^{2}d\omega\left(x\right)d\omega\left(z\right)+\frac{1}{\gamma}|\lambda_{\ell+1}|^{2}|J|^{\frac{2}{n}}\mathsf{E}\left(J,\omega\right)^{2},$$

since $\frac{1}{\gamma^{2}}\mathrm{P}^{\alpha}\left(J,\mu\right)^{2}\mathsf{E}\left(J,\omega\right)^{2}\leq\frac{1}{\gamma}|J|^{\frac{2}{n}}|\lambda_{\ell+1}|^{2}\mathsf{E}\left(J,\omega\right)^{2}$ for γ large enough depending only on n and α. Finally then, for γ large enough depending only on n and α we can absorb the last term on the right hand side of (71) into the left hand side to obtain (66):

$$|\lambda_{\ell+1}|^{2}|J|^{\frac{2}{n}}\mathsf{E}\left(J,\omega\right)^{2}\lesssim\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(x\right)-\mathbf{R}^{\alpha,n}\mu\left(z\right)|^{2}d\omega\left(x\right)d\omega\left(z\right).$$

But since $\gamma^{-\frac{m-\ell}{2n}}c\frac{\mathrm{P}^{\alpha}(J,\mu)}{|J|^{\frac{1}{n}}}\leq|\lambda_{\ell+1}|$ by (65), we have obtained

$$\mathrm{P}^{\alpha}\left(J,\mu\right)^{2}\mathsf{E}\left(J,\omega\right)^{2}\leq\frac{1}{c^{2}}\gamma|\lambda_{\ell+1}|^{2}|J|^{\frac{2}{n}}\mathsf{E}\left(J,\omega\right)^{2}$$
$$\lesssim\frac{1}{|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(x\right)-\mathbf{R}^{\alpha,n}\mu\left(z\right)|^{2}d\omega\left(x\right)d\omega\left(z\right),$$

which is the strong reverse energy inequality for J since

$$\frac{1}{2|J|_{\omega}^{2}}\int_{J}\int_{J}|\mathbf{R}^{\alpha,n}\mu\left(x\right)-\mathbf{R}^{\alpha,n}\mu\left(z\right)|^{2}d\omega\left(x\right)d\omega\left(z\right)=\mathbb{E}_{J}^{\omega}\left|\mathbf{R}^{\alpha,n}\mu-\mathbb{E}_{J}^{d\omega}\mathbf{R}^{\alpha,n}\mu\right|^{2}.$$

This completes the proof of strong reversal of energy under the assumption that $1\leq\ell\leq m$.

If instead $\ell = 0$, then $|\lambda_i| > \gamma^{-\frac{1}{2n}} |\lambda_{i+1}|$ for all $1 \leq i \leq m$, and so the smallest eigenvalue satisfies

$$|\lambda_1| > \gamma^{-\frac{1}{2n}} |\lambda_2| > \gamma^{-\frac{2}{2n}} |\lambda_3| > \ldots > \gamma^{-\frac{k}{2n}} |\lambda_{m+1}| > \gamma^{-\frac{1}{2}} c \frac{\mathrm{P}^\alpha (J, \mu)}{|J|^{\frac{1}{n}}}.$$

In this case the arguments above show that

$$\left(\gamma^{-\frac{1}{2}} c \frac{\mathrm{P}^\alpha (J, \mu)}{|J|^{\frac{1}{n}}}\right)^2 \mathsf{E}(J, \omega)^2 \lesssim \frac{1}{|J|_\omega^2} \int_J \int_J |\mathbf{R}^{\alpha,n}\mu(x) - \mathbf{R}^{\alpha,n}\mu(z)|^2 \, d\omega(x) \, d\omega(z) + \frac{1}{\gamma^2} \mathrm{P}^\alpha (J, \mu)^2 \mathsf{E}(J, \omega)^2,$$

which again yields the strong reverse energy inequality for J since the second term on the right hand side can then be absorbed into the left hand side for γ sufficiently large depending only on n and α. □

Necessity of the Energy Conditions

Now we demonstrate in a standard way the necessity of the energy conditions for the vector Riesz transform $\mathbf{R}^{\alpha,n}$ when the measures σ and ω are appropriately energy dispersed. Indeed, we can then establish the inequality

$$\mathcal{E}_\alpha^{\text{strong}} \lesssim \sqrt{\mathcal{A}_2^\alpha} + \mathfrak{T}_{\mathbf{R}^{\alpha,n}}.$$

So assume that (58) holds. We use Lemma 10.1 to obtain that the α-fractional Riesz transform $\mathbf{R}^{\alpha,n}$ has strong reversal of ω-energy on *all* quasicubes J. Then we use the next lemma to obtain the energy condition $\mathcal{E}_\alpha^{\text{strong}} \lesssim \mathfrak{T}_{\mathbf{R}^{n,\alpha}} + \sqrt{A_2^\alpha}$.

Lemma 10.2 *Let* $0 \leq \alpha < n$ *and suppose that* $\mathbf{R}^{\alpha,n}$ *has strong reversal of* ω*-energy on* all *quasicubes* J*. Then we have the energy condition inequality,*

$$\mathcal{E}_\alpha^{\text{strong}} \lesssim \mathfrak{T}_{\mathbf{T}^{n,\alpha}} + \sqrt{A_2^{\alpha,\text{punct}}}.$$

Proof Fix $\gamma \geq 2$ large enough depending only on n and α, and fix goodness parameters $\mathbf{r}$ and ε so that $\gamma \leq 2^{\mathbf{r}(1-\varepsilon)}$. Then Lemma 10.1 holds. From the strong reversal of ω-energy with $d\mu \equiv \mathbf{1}_{I_r \setminus \gamma J} d\sigma$, we have

$$\begin{aligned} &\mathsf{E}(J, \omega)^2 \, \mathrm{P}^\alpha \left(J, \mathbf{1}_{I_r \setminus \gamma J} d\sigma\right)^2 \\ &\leq C \, \mathbb{E}_J^\omega \left|\mathbf{T}^\alpha \left(\mathbf{1}_{I_r \setminus \gamma J} d\sigma\right) - \mathbb{E}_J^{d\omega} \mathbf{T}^\alpha \left(\mathbf{1}_{I_r \setminus \gamma J} d\sigma\right)\right|^2 \\ &\lesssim \mathbb{E}_J^\omega \left|\mathbf{T}^\alpha \left(\mathbf{1}_{I_r \setminus \gamma J} d\sigma\right)\right|^2 \lesssim \mathbb{E}_J^\omega \left|\mathbf{T}^\alpha \left(\mathbf{1}_{I_r} d\sigma\right)\right|^2 + \mathbb{E}_J^\omega \left|\mathbf{T}^\alpha \left(\mathbf{1}_{\gamma J} d\sigma\right)\right|^2, \end{aligned}$$

and so

$$\sum_{J\in M_{(\mathbf{r},\varepsilon)-\text{deep}}(I_r)} |J|_\omega \, \mathsf{E}(J,\omega)^2 \, \mathrm{P}^\alpha(J,\mu)^2 \lesssim \sum_J \int_J \left|\mathbf{T}^\alpha\left(\mathbf{1}_{I_r}d\sigma\right)(x)\right|^2 d\omega(x) + \sum_J \int_J \left|\mathbf{T}^\alpha\left(\mathbf{1}_{\gamma J}d\sigma\right)(x)\right|^2 d\omega(x)$$

$$\lesssim \int_{I_r} \left|\mathbf{T}^\alpha\left(\mathbf{1}_{I_r}d\sigma\right)(x)\right|^2 d\omega(x) + \sum_J \int_{\gamma J} \left|\mathbf{T}^\alpha\left(\mathbf{1}_{\gamma J}d\sigma\right)(x)\right|^2 d\omega(x)$$

$$\lesssim \mathfrak{T}_{\mathbf{T}^{n,\alpha}} |I_r|_\sigma + \sum_J \mathfrak{T}_{\mathbf{T}^{n,\alpha}} |\gamma J|_\sigma \lesssim \mathfrak{T}_{\mathbf{T}^{n,\alpha}} |I_r|_\sigma$$

since $\gamma J \subset I_r$ for $\gamma \leq 2^{\mathbf{r}(1-\varepsilon)}$, and since the quasicubes γJ have bounded overlap (see [29, Lemma 2 in v3]). We also have

$$\sum_{J\in M_{(\mathbf{r},\varepsilon)-\text{deep}}(I_r)} |J|_\omega \, \mathsf{E}(J,\omega)^2 \, \mathrm{P}^\alpha\left(J,\mathbf{1}_{\gamma J}d\sigma\right)^2 \lesssim \sum_{J\in M_{(\mathbf{r},\varepsilon)-\text{deep}}(I_r)} A_2^{\alpha,\text{energy}} |\gamma J|_\sigma \lesssim A_2^{\alpha,\text{energy}} |I_r|_\sigma$$

by the bounded overlap of the quasicubes γJ in I_r once more. We can now easily complete the proof of $\mathcal{E}_\alpha^{\text{strong}} \lesssim \mathfrak{T}_{\mathbf{T}^{n,\alpha}} + \sqrt{A_2^{\alpha,\text{punct}}}$. □

Acknowledgements We thank the referee for a careful reading of the paper and numerous suggestions to improve the exposition.

References

1. T.C. Anderson, A new sufficent two-weighted bump assumption for L^p boundedness of Calderó n-Zygmund operators. Proc. Am. Math. Soc. **143**(8), 3573–3586 (2015)
2. R.R. Coifman, C.L. Fefferman, Weighted norm inequalities for maximal functions and singular integrals. Stud. Math. **51**, 241–250 (1974)
3. G. David, J.-L. Journé, A boundedness criterion for generalized Calderón-Zygmund operators. Ann. Math. **120**(2), 371–397 (1984). MR763911 (85k:42041)
4. R. Hunt, B. Muckenhoupt, R.L. Wheeden, Weighted norm inequalities for the conjugate function and the Hilbert transform. Trans. Am. Math. Soc. **176**, 227–251 (1973)
5. T. Hytönen, On Petermichl's dyadic shift and the Hilbert transform. C. R. Math. Acad. Sci. Paris **346**, 1133–1136 (2008). MR2464252
6. T. Hytönen, The two weight inequality for the Hilbert transform with general measures. arXiv:1312.0843v2
7. A. Kairema, J. Li, M.C. Pereyra, L. Ward, Haar bases on quasi-metric spaces, and dyadic structure theorems for function spaces on product spaces of homogeneous type. J. Funct. Anal. **271**(7), 1793–1843 (2016)
8. M.T. Lacey, Two weight inequality for the Hilbert transform: a real variable characterization, II. Duke Math. J. **163**(15), 2821–2840 (2014)
9. M.T. Lacey, The two weight inequality for the Hilbert transform: a primer, to appear in this volume in honor of Cora Sadosky. arXiv:1304.5004v1
10. M.T. Lacey, On the separated bumps conjecture for Calderón-Zygmund operators. arXiv:1310.3507v4
11. M.T. Lacey, E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, The two weight inequality for Hilbert transform, coronas, and energy conditions (2011). arXiv

12. M.T. Lacey, E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, Two weight inequality for the Hilbert transform: a real variable characterization (2012). arXiv:1201.4319
13. 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**(15), 2795–2820 (2014)
14. M.T. Lacey, E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, B.D. Wick, Two weight inequalities for the Cauchy transform from $\mathbb{R}$ to $\mathbb{C}_+$. arXiv:1310.4820v4
15. M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, A characterization of two weight norm inequalities for maximal singular integrals with one doubling measure. Anal. PDE **5**(1), 1–60 (2012)
16. M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, A two weight inequality for the Hilbert transform assuming an energy hypothesis. J. Funct. Anal. **263**(2), 305–363 (2012)
17. M.T. Lacey, B.D. Wick, Two weight inequalities for the Cauchy transform from $\mathbb{R}$ to $\mathbb{C}_+$. arXiv:1310.4820v1
18. M.T. Lacey, An elementary proof of the A2 Bound, 2015, Available at http://arxiv.org/abs/1501.05818
19. F. Nazarov, S. Treil, A. Volberg, The Bellman function and two weight inequalities for Haar multipliers. J. Am. Math. Soc. **12**, 909–928 (1999). MR{1685781 (2000k:42009)}
20. F. Nazarov, S. Treil, A. Volberg, The *Tb*-theorem on non-homogeneous spaces. Acta Math. **190**(2), 151 (2003). MR 1998349 (2005d:30053)
21. F. Nazarov, S. Treil, A. Volberg, Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures (2004, preprint). arxiv:1003.1596
22. C. Pérez, Two weighted inequalities for potential and fractional type maximal operators. Indiana Univ. Math. J. **43**(2), 663–683 (1994)
23. E. Sawyer, A characterization of a two weight norm inequality for maximal operators. Studia Math. **75**, 1–11 (1982). MR{676801 (84i:42032)}
24. E. Sawyer, A characterization of two weight norm inequalities for fractional and Poisson integrals. Trans. A.M.S. **308**, 533–545 (1988). MR{930072 (89d:26009)}
25. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A two weight theorem for α-fractional singular integrals with an energy side condition. arXiv:1302.5093v8
26. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A geometric condition, necessity of energy, and two weight boundedness of fractional Riesz transforms. arXiv:1310.4484v1
27. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A note on failure of energy reversal for classical fractional singular integrals. IMRN **2015**(19), 9888–9920 (2015)
28. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A two weight theorem for α-fractional singular integrals with an energy side condition and quasicube testing. arXiv:1302.5093v10
29. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A two weight theorem for α-fractional singular integrals with an energy side condition, quasicube testing and common point masses. arXiv:1505.07816v2,v3
30. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, A two weight theorem for α -fractional singular integrals with an energy side condition. Revista Mat. Iberoam. **32**(1), 79–174 (2016)
31. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, The two weight $T1$ theorem for fractional Riesz transforms when one measure is supported on a curve. arXiv:1505.07822v4
32. E.T. Sawyer, C.-Y. Shen, I. Uriarte-Tuero, Failure of necessity of the energy condition. arXiv:16072.06071v2
33. E. Sawyer, R.L. Wheeden, Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces. Am. J. Math. **114**, 813–874 (1992)
34. E.M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals* (Princeton University Press, Princeton, 1993)
35. A. Volberg, *Calderón-Zygmund Capacities and Operators on Nonhomogeneous Spaces*. CBMS Regional Conference Series in Mathematics (American Mathematical Society, Providence, 2003). MR{2019058 (2005c:42015)}

A Partition Function Connected with the Göllnitz–Gordon Identities

Nicolas Allen Smoot

Abstract We use the celebrated circle method of Hardy and Ramanujan to develop convergent formulæ for counting a restricted class of partitions that arise from the Göllnitz–Gordon identities.

Introduction

The purpose of this article is to illustrate a beautiful application of the tools of complex analysis to a discrete subject: the theory of addition over the integers, also known as partition theory.

A partition of a positive integer n is simply an expression of n as a sum of other positive integers. For example, taking the number 5, we find 7 different partitions: 5, $4 + 1$, $3 + 2$, $3 + 1 + 1$, $2 + 2 + 1$, $2 + 1 + 1 + 1$, and $1 + 1 + 1 + 1 + 1$.

The number of partitions of n is denoted by $p(n)$, and is often called the partition function. In our example, we have $p(5) = 7$.

While partitions have been studied since the time of Euler [6], very little was known about the partition function itself before the twentieth century. Indeed, at the end of the nineteenth century, attempts to study the behavior of the prime counting function [14] had led to a general sense of pessimism in number theory [5]; it was expected that any careful analysis of $p(n)$ would produce an asymptotic formula that was approximate at best, and certainly not useful for direct computation.

It was not until 1918 that Hardy and Ramanujan developed the techniques to conduct a detailed study of $p(n)$ [5]. The results of their work were astonishing: not only were they capable of achieving a formula that could give the exact value of $p(n)$ with relative efficiency, but the formula itself is an utterly bizarre object, as an infinite series containing Bessel functions, coprime sums over roots of unity, and π—analytic entities that seem wholly irrelevant to the question of simple addition over the natural numbers.

N. Allen Smoot (✉)
Georgia Southern University, Statesboro, GA, USA
e-mail: nsmoot@georgiasouthern.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_14

The techniques that Hardy and Ramanujan had developed are embodied in what is now known as the circle method. This method has since become one of the most basic tools in analytic number theory [11, 15].

Notably, the circle method has continued to contribute to the theory of partitions and q-series. Hardy and Ramanujan's formula was carefully refined by Rademacher, first in 1936 [9] to make their formula for $p(n)$ convergent, and again in 1943 [10] as an adjustment of the method itself. Soon thereafter, it was realized that the techniques embodying the circle method could be used to develop formulæ for a variety of more restricted partition functions (two notable examples are [8] and [7]).

We are interested here in one such partition function, associated with the Göllnitz–Gordon identities [3, 4], which we provide here for reference:

Theorem 1.1 (Göllnitz–Gordon Identities) *Fix a to be either* 1 *or* 3*. Given an integer n, the number of partitions of n in which parts are congruent to* $4, \pm a \pmod 8$*, is equal to the number of partitions of n in which parts are non-repeating and non-consecutive, with any two even parts differing by at least* 4*, and with all parts* $\geq a$*.*

Each identity—one for either value of a—equates the sizes of two different classes of partitions of n, while not actually indicating the class size itself. We will use Hardy and Ramanujan's method, together with Rademacher's refinements, to formulate a convergent expression for the number of partitions associated with these identities.

Definition 1.2 Fix a at either 1 or 3. A Göllnitz–Gordon partition of type a is composed of parts of the form $4, \pm a \pmod 8$. The generating function for such partitions is expressed as $F_a(q)$, and the actual number of such partitions of n is given as $g_a(n)$.

We seek a formula for $g_a(n)$. The author wishes to note his deep appreciation for the guidance and encouragement of Professor Andrew Sills, who first suggested this problem.

In keeping with the theory of q-series [1, Chapter 2], we have

$$F_a(q) = \sum_{k=0}^{\infty} g_a(k) q^k \tag{1}$$

$$= \prod_{m=0}^{\infty} (1 - q^{8m+a})^{-1} (1 - q^{8m+4})^{-1} (1 - q^{8m+8-a})^{-1} \tag{2}$$

$$= \frac{1}{(q^a; q^8)_\infty (q^4; q^8)_\infty (q^{8-a}; q^8)_\infty}, \tag{3}$$

with

$$(a; q)_\infty = \prod_{j=0}^{\infty} \left(1 - aq^j\right). \tag{4}$$

Cauchy's residue theorem [14, Chapter 3] gives us a means of calculating—at least in principle—the value of $g_a(n)$. Dividing $F_a(q)$ by q^{n+1}, we find that $g_a(n)$ is the coefficient of q^{-1}, and is therefore the residue of $F_a(q)/q^{n+1}$.

Theorem 1.3

$$g_a(n) = \frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{F_a(q)}{q^{n+1}} dq, \tag{5}$$

for $\mathcal{C}$ some curve inside the unit circle of the q-plane, encompassing $q = 0$.

We must choose an appropriate contour for $\mathcal{C}$. We then study $F_a(q)$ itself, including some of its useful transformation properties. Next, we will employ the circle method in reducing our integral (5) to something far more accessible to integration. We finish our integration using the theory of Bessel functions.

Rademacher's Contour

Casual inspection of (3) suggests that $F_a(q)$ has important structure near the roots of unity of the unit circle. We will construct a contour that remains inside the unit circle, but approaches the roots of unity $e^{2\pi ih/k}$ in a controlled way. This contour was first used by Rademacher [10]. We use the contour in a form slightly modified by Sills [13].

Definition 2.1 For a given $h/k \in \mathcal{F}_N$, define the Ford circle $C(h, k)$ as the curve given by

$$\left| \tau - \left(\frac{h}{k} + \frac{i}{2k^2} \right) \right| = \frac{1}{2k^2}. \tag{6}$$

Given the set of Ford circles corresponding to the Farey sequence of degree N, let $\gamma(h, k)$ be defined as the upper arc of $C(h, k)$ from

$$\tau_I(h, k) = \frac{h}{k} - \frac{k_p}{k(k^2 + k_p^2)} + \frac{1}{k^2 + k_p^2} i$$

to

$$\tau_T(h, k) = \frac{h}{k} + \frac{k_s}{k(k^2 + k_s^2)} + \frac{1}{k^2 + k_s^2} i,$$

with h_p/k_p and h_s/k_s the immediate predecessor and successor (respectively) of $h/k \in \mathcal{F}_N$ (let $0_p/1_p = (N-1)/N$; similarly, let $(N-1)_s/N_s = 0/1$).

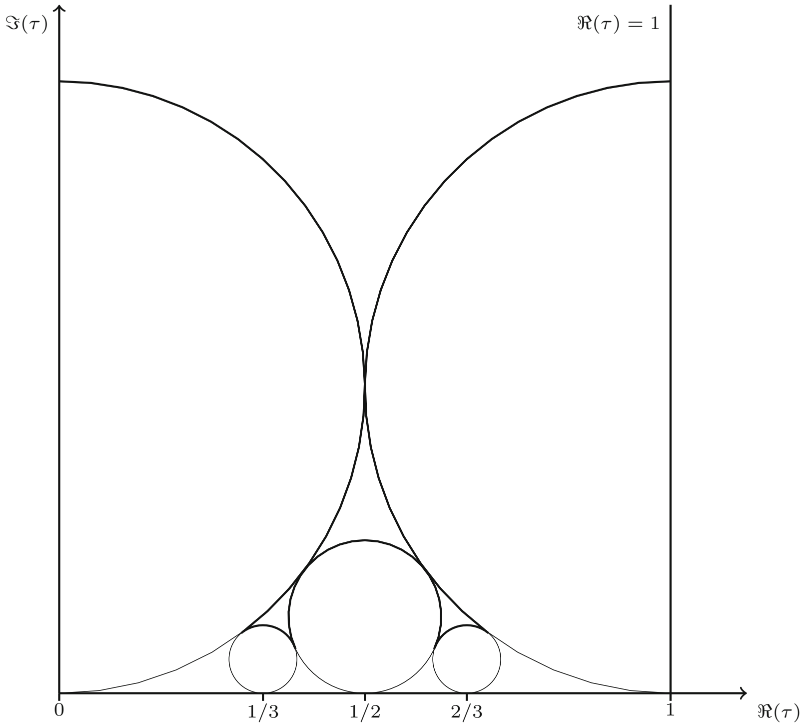


Fig. 1 Ford circles $C(h, k)$ for $h/k \in \mathcal{F}_3$, with $P(3)$ highlighted

Definition 2.2 The Rademacher path of order N, $P(N)$, is the union of all upper arcs $\gamma(h, k)$ from $\tau = i$ to $\tau = i + 1$:

$$P(N) = \bigcup_{h/k \in \mathcal{F}_N} \gamma(h, k). \tag{7}$$

We give an illustration of $P(3)$ in Fig. 1.

It may be easily demonstrated that consecutive Ford circles corresponding to $\mathcal{F}_N$ are tangent to one another, so that $P(N)$ is a connected curve. Moreover, for τ in the upper arc $\gamma(h, k)$, $\Im(\tau) > 0$; therefore, $\gamma(h, k)$ lies entirely in $\mathbb{H}$ for every $h/k \in \mathcal{F}_N$. Therefore, $P(N)$ is a connected curve that lies entirely in $\mathbb{H}$.

So if we define $q = e^{2\pi i \tau}$, then we may define our curve $\mathcal{C}$ from (5) as the preimage of $P(N)$. We will make one more helpful change of variables:

$$\tau = \frac{h}{k} + \frac{iz}{k}, \tag{8}$$

with $\Re(z) > 0$. This change maps $C(h,k)$ (with $\gamma(h,k)$) to the circle

$$K_k^{(-)} : \left| z - \frac{1}{2k} \right| = \frac{1}{2k}. \tag{9}$$

Notice that the initial and terminal points of $\gamma(h,k)$ are mapped to $z_I(h,k)$ and $z_T(h,k)$ by the following:

$$\tau_I(h,k) \mapsto z_I(h,k) = \frac{k}{k^2 + k_p^2} + \frac{k_p}{k^2 + k_p^2} i, \tag{10}$$

$$\tau_T(h,k) \mapsto z_T(h,k) = \frac{k}{k^2 + k_s^2} - \frac{k_s}{k^2 + k_s^2} i. \tag{11}$$

We finish this section by referencing an important lemma, which can be proved quickly from the properties of the Farey fractions [6].

Lemma 2.3 *Let $N \in \mathbb{N}$ be given, with $h/k \in \mathcal{F}_N$. Let $z_I(h,k)$, $z_T(h,k)$ be the images of $\tau_I(h,k)$, $\tau_T(h,k)$, respectively, from $C(h,k)$ to $K_k^{(-)}$. Then for any z on the chord connecting $z_I(h,k)$ to $z_T(h,k)$, we have*

$$|z| = O\left(N^{-1}\right). \tag{12}$$

Transformation Equations

We begin by expressing $F_a(q)$ in terms of automorphic forms—in particular, as a quotient of eta functions by a theta function. Let $q = e^{2\pi i \tau}$, with τ a complex variable, $\Im(\tau) > 0$.

Recall that Ramanujan's theta function [2, Chapter 1] has the following product expansion:

$$f(-q^\alpha, -q^\beta) = (q^\alpha; q^{\alpha+\beta})_\infty (q^\beta; q^{\alpha+\beta})_\infty (q^{\alpha+\beta}; q^{\alpha+\beta})_\infty. \tag{13}$$

Moreover, Ramanujan's theta function is related to the standard theta function ϑ_1 by the following, which can be verified by the series representations of both functions [11, Chapter 10]:

$$f(-q^\alpha, -q^\beta) = -ie^{\pi i \tau (3\alpha - \beta)/4} \vartheta_1(\alpha\tau | (\alpha + \beta)\tau). \tag{14}$$

We then have

$$F_a(q) = \frac{(q^8; q^8)_\infty^2}{(q^4; q^4)_\infty f(-q^a, -q^{8-a})} \tag{15}$$

$$= i \exp(\pi i \tau (2 - a)) \frac{(q^8; q^8)_\infty^2}{(q^4; q^4)_\infty \vartheta_1(a\tau | 8\tau)}. \tag{16}$$

Since $(q^\alpha; q^\alpha)_\infty = e^{-\alpha\pi i\tau/12}\eta(\alpha\tau)$, we can rewrite the remaining q-Pochhammer symbols in terms of eta functions in the following way:

$$F_a(q) = i\exp(\pi i\tau(1-a))\frac{\eta(8\tau)^2}{\eta(4\tau)\vartheta_1(a\tau|8\tau)}. \tag{17}$$

We will now study the behavior of $F_a(q)$ near the arbitrary singularity $e^{2\pi ih/k}$, with $0 \leq h < k$, and $(h,k) = 1$. To do this, we will divide our work into four cases, depending on the divisibility properties of k with respect to 8, and then take advantage of the modular symmetries of the η and ϑ_1 functions.

$GCD(k, 8) = 8$

The simplest transformation formula relevant to our problem occurs for $(k, 8) = 8$. Let H_8 be defined as the negative inverse of h modulo $16k$:

$$hH_8 \equiv -1 \pmod{16k}. \tag{18}$$

Notice that since $8|k$ by hypothesis, and $(h,k) = 1$, therefore $(h, 16k) = (h,k) = 1$, so that H_8 exists. Then the following are elements of $SL(2,\mathbb{Z})$:

$$\begin{pmatrix} h & -\frac{8}{k}(hH_8+1) \\ \frac{k}{8} & -H_8 \end{pmatrix}, \tag{19}$$

$$\begin{pmatrix} h & -\frac{4}{k}(hH_8+1) \\ \frac{k}{4} & -H_8 \end{pmatrix}, \tag{20}$$

We will allow

$$\tau' = \frac{H_8}{k} + \frac{iz^{-1}}{k}. \tag{21}$$

Applying (19) as a modular transformation to $8\tau'$, we have

$$\frac{8h\tau' - \frac{8}{k}(hH_8+1)}{8\frac{k}{8}\tau' - H_8} = 8\tau.$$

Similarly, applying (20) to $4\tau'$, we get 4τ.

Therefore, we will transform $\eta(8\tau)$ to $\eta(8\tau')$, using (19). Similarly, we transform $\eta(4\tau)$ to $\eta(4\tau')$ using (20).

Invoking these transformations, we must contend with the roots of unity associated with the η and ϑ_1 functions. As a shorthand, we will refer to the roots of unity as the following:

$$\epsilon(8,8) = \epsilon\left(h, -\frac{8}{k}(hH_8+1), \frac{k}{8}, -H_8\right), \tag{22}$$

$$\epsilon(8,4) = \epsilon\left(h, -\frac{4}{k}(hH_8+1), \frac{k}{4}, -H_8\right), \tag{23}$$

where $\epsilon(a,b,c,d)$ is the root of unity given by

$$\epsilon(a,b,c,d) = \begin{cases} \left(\frac{d}{c}\right) i^{(1-c)/2} \exp\left(\frac{\pi i}{12}(bd(1-c^2)+c(a+d))\right), & 2 \nmid c \\ \left(\frac{c}{d}\right) \exp\left(\frac{\pi i d}{4} + \frac{\pi i}{12}(ac(1-d^2)+d(b-c))\right), & 2 \nmid d \end{cases}, \tag{24}$$

and $\left(\frac{m}{n}\right)$ is the Legendre–Jacobi character. See [11, Chapter 9].

Invoking the functional equation for η [11, Chapter 9], it follows that

$$\frac{\eta(8\tau)^2}{\eta(4\tau)} = \frac{1}{z^{1/2}} \frac{\epsilon(8,8)^2}{\epsilon(8,4)} \frac{\eta(8\tau')^2}{\eta(4\tau')}. \tag{25}$$

Handling ϑ_1 turns out to be more difficult, due to the presence of a second complex variable. We will mimic our work with $\eta(8\tau)$, using (19), and setting

$$v = a\tau i z^{-1} = \frac{a(hiz^{-1}-1)}{k}. \tag{26}$$

The functional equation for ϑ_1 [11, Chapter 10] gives us

$$\vartheta_1(a\tau|8\tau) = \vartheta_1\left(\frac{v}{iz^{-1}}\middle|8\tau\right) = -i\epsilon(8,8)^3 \frac{1}{z^{1/2}} e^{z\pi a^2(hiz^{-1}-1)^2/8k} \vartheta_1\left(v|8\tau'\right). \tag{27}$$

Recall that $hH_8 \equiv -1 \pmod{16k}$. We can therefore write

$$-1 = hH_8 + 16kM, \tag{28}$$

with $M \in \mathbb{Z}$. We then have

$$v = ah\tau' + 16aM. \tag{29}$$

If we also take advantage of the fact that $\vartheta_1(v+1|\tau) = -\vartheta_1(v|\tau)$ [11, Chapter 10], then we have

$$\vartheta_1\left(v|8\tau'\right) = \vartheta_1\left(ah\tau' + 16aM|8\tau'\right) = \vartheta_1\left(ah\tau'|8\tau'\right). \tag{30}$$

Again considering that $(k, 8) = 8$ and $(h, k) = 1$, and $a = 1, 3$, we also have $ah \equiv 1, 3, 5, 7 \pmod 8$. We therefore write

$$\vartheta_1(ah\tau'|8\tau') = \vartheta_1(b\tau' + 8N\tau'|8\tau'), \tag{31}$$

with b the least positive residue of ah modulo 8.

We now make use of the fact that for $N \in \mathbb{N}$,

$$\vartheta_1(v + N\tau|\tau) = (-1)^N \exp\left(-\pi i N(2v + N\tau)\right) \vartheta_1(v|\tau) \tag{32}$$

[11, Chapter 10], so that

$$\vartheta_1(ah\tau'|8\tau') = (-1)^N \exp(-\pi i N(2b\tau' + 8N\tau'))\vartheta_1(b\tau'|8\tau'). \tag{33}$$

Combining (27), (30), (33), and inverting, we have the following:

$$\frac{1}{\vartheta_1(a\tau|8\tau)} = i\frac{(-1)^N}{\epsilon(8,8)^3} z^{1/2} e^{-z\pi a^2(hiz^{-1}-1)^2/8k} \frac{\exp(\pi i N(2b\tau' + 8N\tau'))}{\vartheta_1(b\tau'|8\tau')}. \tag{34}$$

We now have sufficient information, in (25) and (34), to reassemble the transformed generating function.

$$F_a(q) = i\exp(\pi i\tau(1-a))\frac{1}{z^{1/2}}\frac{\epsilon(8,8)^2}{\epsilon(8,4)}\frac{\eta(8\tau')^2}{\eta(4\tau')}\frac{1}{\vartheta_1(a\tau|8\tau)} \tag{35}$$

$$\times \exp(\pi i N(2b\tau' + 8N\tau'))\frac{\eta(8\tau')^2}{\eta(4\tau')\vartheta_1(b\tau'|8\tau')}. \tag{36}$$

Here $b = 1, 3, 5, 7$. However, noting from (13) that

$$f(-q^\alpha, -q^\beta) = f(-q^\beta, -q^\alpha), \tag{37}$$

we may define $F_5(q) = F_3(q),\ F_7(q) = F_1(q)$. We therefore have

$$F_a(q) = \frac{i(-1)^N}{\epsilon(8,8)\epsilon(8,4)} \times \exp\Big(\pi i\tau(1-a) - z\pi a^2(hiz^{-1}-1)^2/8k$$
$$+ \pi i N(2b\tau' + 8N\tau') + \pi i\tau'(a-1)\Big)F_b(y), \tag{38}$$

with $y = \exp(2\pi i\tau')$.

Remembering that

$$N = \left\lfloor \frac{ah}{8} \right\rfloor = \frac{ah-b}{8}, \tag{39}$$

and that

$$a^2 - 4a + 3 = (a-1)(a-3) = 0, \tag{40}$$

we may collect and reorganize the coefficients of 1, z, and $1/z$ in the exponential of (38). Doing so gives the following transformation formula:

$$F_a(q) = \omega_{a,8}(h,k) \exp\left(\frac{\pi}{8k}\left(\frac{(b-4)^2-8}{z} + z(4a-5)\right)\right) F_b(y), \tag{41}$$

where

$$\omega_{a,8}(h,k) = \frac{i(-1)^{\lfloor \frac{ah}{8} \rfloor}}{\epsilon(8,8)\epsilon(8,4)} \exp\left(\frac{\pi i}{8k}(h(5-4a) - H_8((b-4)^2-8))\right). \tag{42}$$

We note that we can extend this result to prove the modularity of $F_a(q)$ relative to a certain subgroup of $SL(2,\mathbb{Z})$. We do not give the proof here.

$GCD(k, 8) < 8$

The result of the section "$GCD(k,8) = 8$" suggests that $F_a(q)$ is modular, at least with respect to a subgroup of the modular group. While such a property does not carry over exactly to the remaining three cases, it is only necessary to show that $F_a(q) = f(z)\Psi(y)$, with $\Psi(q)$ a suitable quotient of q-series.

For each case $(k,8) = d$, we will define

$$\tau' = \frac{H_d}{k} + \frac{diz^{-1}}{8k}, \tag{43}$$

where

$$\frac{8hH_d}{d} \equiv -1 \pmod{k/d}, \tag{44}$$

and

$$y = e^{2\pi i \tau'}. \tag{45}$$

We consider the following matrices, which can easily be shown to be in $SL(2,\mathbb{Z})$:

$$\begin{pmatrix} 8h/d & -\frac{d}{k}(8hH_d/d+1) \\ k/d & -H_d \end{pmatrix}, \tag{46}$$

$$\begin{pmatrix} 4h/d & -\frac{d}{k}(8hH_d/d+1) \\ k/d & -2H_d \end{pmatrix}. \tag{47}$$

We also define

$$\epsilon(d,8) = \epsilon\left(8h/d, -\frac{d}{k}(8hH_d/d+1), \frac{k}{d}, -H_d\right), \tag{48}$$

$$\epsilon(d,4) = \epsilon\left(4h/d, -\frac{d}{k}(8hH_d/d+1), \frac{k}{d}, -2H_d\right), \tag{49}$$

with $\epsilon(a,b,c,d)$ defined by (24). Remembering (44), we also let

$$v = \frac{da(hiz^{-1}-1)}{8k} = ah\tau' + \frac{aM}{8}, \tag{50}$$

with $M \in \mathbb{Z}$. Finally, we write

$$\rho_{a,d} = \exp\left(\frac{\pi iad}{4k}\left(\frac{8hH_d}{d}+1\right)\right). \tag{51}$$

GCD(k, 8) = 4

With $d = 4$, we apply (46) to $4\tau'$, and (47) to $8\tau'$, so that we have

$$\frac{\eta(8\tau)^2}{\eta(4\tau)} = \frac{1}{2z^{1/2}}\frac{\epsilon(4,8)^2}{\epsilon(4,4)}\frac{\eta(4\tau')^2}{\eta(8\tau')}. \tag{52}$$

As with the case of $d = 8$, ϑ_1 requires the most work by far. The initial transformation through (46) gives us

$$\vartheta_1(a\tau|8\tau) = -i\epsilon(4,8)^3\frac{1}{(2z)^{1/2}}e^{z\pi a^2(hiz^{-1}-1)^2/8k}\vartheta_1\left(v|4\tau'\right). \tag{53}$$

And

$$\vartheta_1\left(v|4\tau'\right) = \vartheta_1\left(ah\tau' + \frac{aM}{8}\middle|4\tau'\right). \tag{54}$$

We may now allow $b \equiv ah \pmod 4$, letting $ah = 4N + b$, so that (54), together with (32), gives

$$\vartheta_1\left(v|4\tau'\right) = (-1)^N\exp(-\pi iN(2\tau'(2N+b) + aM/4)) \times \vartheta_1\left(b\tau' + \frac{aM}{8}\middle|4\tau'\right). \tag{55}$$

We now shift from ϑ_1 to ϑ_4 [11, Chapter 10]:

$$\vartheta_1(v|\tau) = i\exp(-\pi i\tau/4 - \pi iv)\vartheta_4(v|\tau).$$

Write

$$\vartheta_1\left(b\tau' + \frac{aM}{8}\middle|4\tau'\right) = \vartheta_1\left((b-2)\tau' + \frac{aM}{8} + 2\tau'\middle|4\tau'\right) \tag{56}$$

$$= i\exp(-\pi i((b-1)\tau' + aM/8)) \tag{57}$$

$$\times\, \vartheta_4\left((b-2)\tau' + \frac{aM}{8}\middle|4\tau'\right). \tag{58}$$

We now express ϑ_4 as an infinite product [11, Chapter 10]:

$$\vartheta_4\left((b-2)\tau' + \frac{aM}{8}\middle|4\tau'\right) = \prod_{m=1}^{\infty}(1-y^{4m})(1-\rho_{a,4}y^{4m-4+b})(1-\rho_{a,4}^{-1}y^{4m-b})$$

$$= (y^4;y^4)_\infty(\rho_{a,4}y^b;y^4)_\infty(\rho_{a,4}^{-1}y^{4-b};y^4)_\infty. \tag{59}$$

Combining (52), (53), (55), (58), (59), and simplifying, we have

$$F_a(q) = \frac{1}{\sqrt{2}}\omega_{a,4}(h,k)\exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(4a-5)\right)\right)\Psi_{a,4}(y), \tag{60}$$

with

$$\Psi_{a,4}(q) = \frac{(q^4;q^4)_\infty^2}{(q^8;q^8)_\infty f(-\rho_{a,4}q^b; -\rho_{a,4}^{-1}q^{4-b})}, \tag{61}$$

and

$$\omega_{a,4}(h,k) = \frac{i(-1)^{\lfloor \frac{ah}{4} \rfloor}}{\epsilon(4,8)\epsilon(4,4)}$$

$$\times \exp\left(\frac{\pi i}{4k}(h - H_4 - h(4a-3)(hH_4+1) + a(2hH_4+1)(b-2)\right). \tag{62}$$

GCD(k, 8) = 2

With $d = 4$, we apply (46) to $2\tau'$, and (47) to $4\tau'$, so that we have

$$\frac{\eta(8\tau)^2}{\eta(4\tau)} = \frac{1}{2\sqrt{2}z^{1/2}}\frac{\epsilon(2,8)^2}{\epsilon(2,4)}\frac{\eta(2\tau')^2}{\eta(4\tau')}. \tag{63}$$

Once again, ϑ_1 requires the most work by far. The initial transformation through (46) gives us

$$\vartheta_1(a\tau|8\tau) = -i\epsilon(2,8)^3 \frac{1}{2z^{1/2}} e^{z\pi a^2(hiz^{-1}-1)^2/8k} \vartheta_1\left(v|2\tau'\right). \tag{64}$$

And

$$\vartheta_1\left(v|2\tau'\right) = \vartheta_1\left(ah\tau' + \frac{aM}{8}\middle|2\tau'\right). \tag{65}$$

Notice that both a and h are odd. We may therefore write $ah = 2N + 1$, so that

$$\vartheta_1\left(v|2\tau'\right) = (-1)^N \exp(-\pi i N(2\tau' + aM/4 + 2\tau' N))\vartheta_1\left(\tau' + \frac{aM}{4}\middle|2\tau'\right). \tag{66}$$

We now shift from ϑ_1 to ϑ_4. Write

$$\vartheta_1\left(\tau' + \frac{aM}{8}\middle|2\tau'\right) = i\exp\left(\frac{-\pi i}{8}(12\tau' + aM/8)\right)\vartheta_4\left(\frac{aM}{8}\middle|2\tau'\right). \tag{67}$$

We express ϑ_4 as an infinite product:

$$\vartheta_4\left(\frac{aM}{8}\middle|2\tau'\right) = (y^2;y^2)_\infty(\rho_{a,2}y;y^2)_\infty(\rho_{a,2}^{-1}y;y^2)_\infty. \tag{68}$$

Combining (63), (64), (66), (67), (68), and simplifying, we have

$$F_a(q) = \frac{1}{\sqrt{2}}\omega_{a,2}(h,k)\exp\left(\frac{\pi}{8k}\left(z(4a-5)\right)\right)\Psi_{a,2}(y), \tag{69}$$

where

$$\Psi_{a,2}(q) = \frac{(q^2;q^2)_\infty^2}{(q^4;q^4)_\infty f(-\rho_{a,2}q;-\rho_{a,2}^{-1}q)}, \tag{70}$$

and

$$\omega_{a,2}(h,k) = \frac{i(-1)^{\lfloor\frac{ah}{2}\rfloor}}{\epsilon(2,8)\epsilon(2,4)}\exp\left(\frac{\pi i}{4k}\left(1-(4a-3)(2hH_2+1)\right)\right). \tag{71}$$

$GCD(k, 8) = 1$

With $d = 4$, we apply (46) to $2\tau'$, and (47) to $4\tau'$, so that we have

$$\frac{\eta(8\tau)^2}{\eta(4\tau)} = \frac{1}{4z^{1/2}} \frac{\epsilon(1,8)^2}{\epsilon(1,4)} \frac{\eta(\tau')^2}{\eta(2\tau')}. \tag{72}$$

Returning to ϑ_1,

$$\vartheta_1(a\tau|8\tau) = -i\epsilon(1,8)^3 \frac{1}{2\sqrt{2}z^{1/2}} e^{8\pi kzv^2} \vartheta_1\left(v|\tau'\right). \tag{73}$$

And

$$\vartheta_1\left(v|\tau'\right) = \vartheta_1\left(ah\tau' + \frac{aM}{8}\middle|\tau'\right). \tag{74}$$

Recognizing that we may extract $ah\tau'$ altogether from our first variable, and recognizing that $(-1)^{ah} = (-1)^h$, we have

$$\vartheta_1\left(v|\tau'\right) = (-1)^h \exp(-\pi iah(ah\tau' + aM/4))\vartheta_1\left(\frac{aM}{8}\middle|\tau'\right). \tag{75}$$

We may now write $\vartheta_1\left(\frac{aM}{8}\middle|\tau'\right)$ in its classic product form [11, Chapter 10]:

$$\vartheta_1\left(\frac{aM}{8}\middle|\tau'\right) = 2e^{\pi i\tau'/4}\sin(\pi aM/8)$$
$$\times \prod_{m=1}^{\infty}(1-e^{2\pi im\tau'})(1-e^{2\pi im\tau'+2\pi iaM/8})(1-e^{2\pi im\tau'-2\pi iaM/8}) \tag{76}$$
$$= 2e^{\pi i\tau'/4}\sin(\pi aM/8)(y;y)_\infty(\rho_{a,1}y;y)_\infty(\rho_{a,1}^{-1}y;y)_\infty. \tag{77}$$

Examining the sine function, let $aM = 8N + c$, with c the least positive residue of $aM \pmod 8$. Then

$$\sin\left(\frac{\pi aM}{8}\right) = (-1)^N \sin\left(\frac{\pi c}{8}\right). \tag{78}$$

Notice that $\sin\left(\frac{\pi c}{8}\right) > 0$. We know that since

$$M = -\frac{1}{k}(8hH_1 + 1), \tag{79}$$

and since $(k, 8) = 1$, therefore

$$c \equiv -ak^{-1} \pmod 8. \tag{80}$$

Moreover, k is odd, so $k^{-1} \equiv k \pmod 8$. So

$$\sin\left(\frac{\pi c}{8}\right) = \left|\sin\left(\frac{\pi ak}{8}\right)\right|. \tag{81}$$

Combining (72), (73), (74), (75), (76), (77), (78), and simplifying, we have:

$$F_a(q) = \frac{1}{2\sqrt{2}}\omega_{a,1}(h,k)\left|\csc\left(\frac{\pi ak}{8}\right)\right|\exp\left(\frac{\pi}{8k}\left(\frac{1}{4z}+z(4a-5)\right)\right)\Psi_{a,1}(y), \tag{82}$$

where

$$\Psi_{a,1}(y) = \frac{(y;y)_\infty}{(y^2;y^2)_\infty(\rho_{a,1}y;y)_\infty(\rho_{a,1}^{-1}y;y)_\infty}, \tag{83}$$

and

$$\omega_{a,1}(h,k) = \frac{(-1)^{\lfloor\frac{-a(8hH_1+1)}{8k}\rfloor+h-1}}{\epsilon(1,8)\epsilon(1,4)} \times \exp\left(\frac{\pi i}{4k}(4h(1-a+hH_1(3-4a))-H_1)\right). \tag{84}$$

Integration

Recall from section "Introduction" that

$$g_a(n) = \frac{1}{2\pi i}\oint_{\mathcal{C}}\frac{F_a(q)}{q^{n+1}}dq,$$

while in section "Rademacher's Contour" we described a contour for $\mathcal{C}$ that will prove useful for integration. We will now begin the integration proper.

Let N be some large positive integer, and let the corresponding Rademacher curve $P(N)$ be given. Then we have the following:

$$g_a(n) = \frac{1}{2\pi i}\oint_{\mathcal{C}}\frac{F_a(q)}{q^{n+1}}dq = \sum_{k=1}^{N}\sum_{\substack{0\le h<k,\\(h,k)=1}}\frac{1}{2\pi i}\int_{\gamma(h,k)}\frac{F_a(q)}{q^{n+1}}dq. \tag{85}$$

In section "Transformation Equations", we gave transformation equations for $F_a(q)$ depending on the divisibility properties of k. We now separate our integral into the corresponding cases:

$$g_a(n) = g_a^{(8)}(n) + g_a^{(4)}(n) + g_a^{(2)}(n) + g_a^{(1)}(n), \tag{86}$$

with

$$g_a^{(d)}(n) = \sum_{\substack{(k,8)=d,\\ k\leq N}} \sum_{\substack{0\leq h<k,\\ (h,k)=1}} \frac{1}{2\pi i} \int_{\gamma(h,k)} \frac{F_a(q)}{q^{n+1}} dq. \tag{87}$$

In each case, we will tranform $F_a(q)$ by the following:

$$g_a^{(d)}(n) = \sum_{\substack{(k,8)=d\\ k\leq N}} \frac{i}{k} \sum_{\substack{0\leq h<k,\\ (h,k)=1}} e^{-2\pi inh/k}$$

$$\times \int_{z_I(h,k)}^{z_T(h,k)} F_a\left(\exp\left(2\pi i\left(\frac{h}{k}+\frac{iz}{k}\right)\right)\right) e^{2\pi nz/k} dz \tag{88}$$

$$=2^{(\alpha-3)/2} \sum_{\substack{(k,8)=d\\ k\leq N}} \frac{i}{k} T_{a,d}(k) \sum_{\substack{0\leq h<k,\\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi inh/k}$$

$$\times \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z} + z(16n+4a-5)\right)\right) \Psi_{a,d}(y) dz, \tag{89}$$

with $\omega_{a,d}(h,k)$ defined by (42), (62), (71), (84), $\Psi_{a,d}(y) = \sum_{j=0}^{\infty} \psi_{a,d}(j) y^j$ defined as $F_b(y)$ for $d = 8$, and (61), (70), (83), otherwise (note that $\psi_{a,d}(0) = 1$);

$$\Lambda(a,d) = \begin{cases} (b-4)^2 - 8 & \text{if } d = 8 \\ 1 & \text{if } d = 4 \\ 0 & \text{if } d = 2 \\ 1/4 & \text{if } d = 1. \end{cases}$$

$$\alpha = \log_2(d), \tag{90}$$

and

$$T_{a,d}(k) = \begin{cases} |\csc(\pi ak/8)| & \text{if } d = 1 \\ 1 & \text{otherwise.} \end{cases}$$

Thereafter,

$$g_a^{(d)}(n) = 2^{(\alpha-3)/2} \sum_{\substack{(k,8)=d \\ k \leq N}} \frac{i}{k} T_{a,d}(k) \sum_{\substack{0 \leq h < k, \\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi i n h/k} \times \left(I_{a,d}^{(1)}(h,k) + I_{a,d}^{(0)}(h,k) \right), \quad (91)$$

where

$$I_{a,d}^{(1)}(h,k) = \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k} \left(\frac{\Lambda(a,d)}{z} + z(16n + 4a - 5) \right) \right) dz, \quad (92)$$

and

$$I_{a,d}^{(0)}(h,k) = \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k} \left(\frac{\Lambda(a,d)}{z} + z(16n + 4a - 5) \right) \right) \sum_{j=1}^{\infty} \psi_{a,d}(j) y^j dz. \quad (93)$$

In each of our cases, we will show that $I_{a,d}^{(0)}(h,k)$ will contribute nothing to our final formula.

Lemma 4.1 *For* $d = 8, 4, 2, 1$*,*

$$\left| \sum_{\substack{0 \leq h < k, \\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi n h/k} \right| = O\left(k^{2/3+\epsilon} n^{1/3} \right). \quad (94)$$

This result can be shown through Kloosterman sum estimation, using the techniques of Salié [12].

Lemma 4.2

$$\left| I_{a,d}^{(0)}(h,k) \right| = O\left(\exp(3n\pi) N^{-1} \right). \quad (95)$$

Proof We may interchange the summation with the integration. Also, remembering that $y = \exp\left(2\pi i \left(\frac{H_d}{k} + \frac{diz^{-1}}{8k} \right)\right)$,

$$I_{a,d}^{(0)}(h,k)$$

$$= \sum_{j=1}^{\infty} \psi_{a,d}(j) e^{2\pi i H_d j/k}$$

$$\times \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z} + z(16n+4a-5)\right)\right) e^{-2dj\pi z^{-1}/8k} dz \tag{96}$$

$$= \sum_{j=1}^{\infty} \psi_{a,d}(j) e^{2\pi i H_d j/k}$$

$$\times \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)-2dj}{z} + z(16n+4a-5)\right)\right) dz. \tag{97}$$

Notice that no matter the permitted value of d, the coefficient of $1/z$ in the exponent of the integrand is now always negative.

Taking advantage of the fact that on and within $K_k^{(-)}$, $\Re(1/z) \geq k$ and $\Re(z) \leq 1/k$, we now examine the magnitude of the integrand:

$$\left|\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)-2dj}{z} + z(16n+4a-5)\right)\right)\right| \tag{98}$$

$$= \exp\left(\frac{\pi}{8k}(\Lambda(a,d)-2dj)\Re(1/z) + \frac{\pi}{8k}(16n+4a-5)\Re(z)\right) \tag{99}$$

$$\leq \exp\left(\frac{\pi(1-2dj)}{8} + \frac{\pi(16n+4a-5)}{8k^2}\right) \tag{100}$$

$$\leq \exp(-\pi j/8 + 3n\pi). \tag{101}$$

We therefore have

$$|I_{a,d}^{(0)}(h,k)|$$

$$\leq \sum_{j=1}^{\infty} |\psi_{a,d}(j)||e^{2\pi i H_d j/k}|$$

$$\times \int_{z_I(h,k)}^{z_T(h,k)} \left|\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)-2dj}{z} + z(16n+4a-5)\right)\right)\right| dz \tag{102}$$

$$\leq \sum_{j=1}^{\infty} |\psi_{a,d}(j)| \exp(-\pi j/8 + 3n\pi) \int_{z_I(h,k)}^{z_T(h,k)} dz. \tag{103}$$

Recall that we are integrating along the circle $K_k^{(-)}$ in the z-plane. We will now deform our contour so that it is a chord connecting z_I and z_T along $K_k^{(-)}$ (Fig. 2).

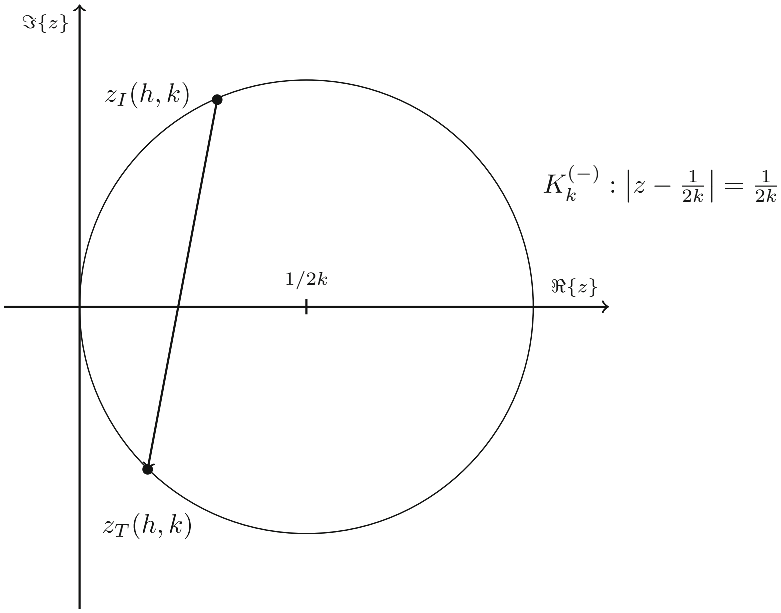


Fig. 2 $K_k^{(-)}$ with the chord connecting $z_I(h,k)$ to $z_T(h,k)$

Recognizing from Lemma 2.3 that the length of such a chord is bounded above by a constant multiple of N^{-1}, we have

$$|I_{a,d}^{(0)}(h,k)| = O\left(\sum_{j=1}^{\infty} |\psi_{a,d}(j)| \exp(-\pi j/8 + 3n\pi)N^{-1}\right) \tag{104}$$

$$= O\left(\exp(3n\pi)N^{-1}\sum_{j=1}^{\infty} |\psi_{a,d}(j)| \exp(-\pi j)\right) \tag{105}$$

$$= O(\exp(3n\pi)N^{-1}). \tag{106}$$

□

Lemma 4.3 *Let* $\epsilon > 0$. *Then*

$$\left|2^{(\alpha-3)/2} \sum_{\substack{(k,8)=d \\ k\leq N}} \frac{i}{k} T_{a,d}(k) \sum_{\substack{0\leq h<k, \\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi i n h/k} I_{a,d}^{(0)}(h,k)\right|$$
$$= O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \tag{107}$$

Proof We note that since $2^{(\alpha-3)/2}$ and $T_{a,d}(k)$ are bounded, we may disregard both in our estimation. We now take the previous result into account:

$$\left|\sum_{\substack{(k,8)=d\\k\le N}}\frac{i}{k}\sum_{\substack{0\le h<k,\\(h,k)=1}}\omega_{a,d}(h,k)e^{-2\pi inh/k}I_{a,d}^{(0)}(h,k)\right|$$

$$\le\sum_{\substack{(k,8)=d\\k\le N}}\frac{1}{kN}e^{3n\pi}\left|\sum_{\substack{0\le h<k,\\(h,k)=1}}\omega_{a,d}(h,k)e^{-2\pi inh/k}\right|. \tag{108}$$

With Lemma 4.2, we know that

$$\left|\sum_{\substack{0\le h<k,\\(h,k)=1}}\omega_{a,d}(h,k)e^{-2\pi inh/k}\right| = O\left(k^{2/3+\epsilon}n^{1/3}\right). \tag{109}$$

This gives us

$$\left|\sum_{\substack{(k,8)=d\\k\le N}}\frac{i}{k}\sum_{\substack{0\le h<k,\\(h,k)=1}}\omega_{a,d}(h,k)e^{-2\pi inh/k}I_{a,d}^{(0)}(h,k)\right|$$

$$=O\left(\left|\sum_{\substack{(k,8)=d\\k\le N}}\frac{1}{kN}e^{3n\pi}k^{2/3+\epsilon}n^{1/3}\right|\right)=O\left(\left|e^{3n\pi}n^{1/3}N^{-1}\sum_{k=1}^{N}\frac{k^{2/3+\epsilon}}{k}\right|\right). \tag{110}$$

Recognizing that

$$\sum_{k=1}^{N}\frac{k^{2/3+\epsilon}}{k}=\sum_{k=1}^{N}\frac{k^{2/3+2\epsilon}}{k^{1+\epsilon}}\le\sum_{k=1}^{N}\frac{N^{2/3+2\epsilon}}{k^{1+\epsilon}}=N^{2/3+2\epsilon}\sum_{k=1}^{N}\frac{1}{k^{1+\epsilon}}, \tag{111}$$

that $\sum_{k=1}^{N}\frac{1}{k^{1+\epsilon}}$ is bounded above as N gets large, and finally noting that we may replace 2ϵ with ϵ, we now have

$$O\left(\left|e^{3n\pi}n^{1/3}N^{-1}\sum_{k=1}^{N}\frac{k^{2/3+\epsilon}}{k}\right|\right)=O(e^{3n\pi}n^{1/3}N^{-1/3+\epsilon}), \tag{112}$$

and the proof is completed. □

We now have

$$g_a^{(d)}(n) = 2^{(\alpha-3)/2} \sum_{\substack{(k,8)=d \\ k\leq N}} \frac{i}{k} T_{a,d}(k) \sum_{\substack{0\leq h<k, \\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi i n h/k} I_{a,d}^{(1)}(h,k)$$
$$+ O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (113)$$

Our object now will be to put $I_{a,d}^{(1)}(h,k)$ into a form approachable from the theory of Bessel functions.

We now return to the original Rademacher contour of $I_{a,d}^{(1)}(h,k)$, along a portion of $K_k^{(-)}$. The brilliance of the contour becomes clear once it is realized that $\Re(1/z) = k$, i.e. is a constant, provided we remain along $K_k^{(-)}$ (and avoid $z = 0$, of course). We wish to make use of the whole of $K_k^{(-)}$, so we will make adjustments to the contour as follows:

$$I_{a,d}^{(1)}(h,k) = \left(\oint_{K_k^{(-)}} - \int_0^{z_I(h,k)} - \int_{z_T(h,k)}^{0} \right)$$
$$\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z} + z(16n+4a-5)\right)\right) dz. \quad (114)$$

Notice that $\int_0^{z_I(h,k)}$ and $\int_{z_T(h,k)}^{0}$ are improper: the integrand is not defined at $z = 0$. We interpret these integrals as limits in which a variable approaches 0. We will now show that $\int_0^{z_I(h,k)}$ and $\int_{z_T(h,k)}^{0}$ will not contribute anything of importance:

Lemma 4.4

$$\left| \int_{z_T(h,k)}^{0} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz \right|,$$
$$\left| \int_0^{z_I(h,k)} \exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z} + z(16n+4a-5)\right)\right) dz \right|$$
$$= O\big(\exp(3n\pi) N^{-1}\big). \quad (115)$$

Proof We will keep on $K_k^{(-)}$ for these estimations. Since the estimation is almost identical in either case, we will work with the integral $\int\limits_{z_T(h,k)}^{0}$. We begin by estimating the integrand of the integral:

$$\left|\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z}+z(16n+4a-5)\right)\right)\right|$$

$$= \exp\left(\frac{\pi}{8k}\left(\Re(1/z)+\Re(z)(16n+4a-5)\right)\right) \tag{116}$$

$$\leq \exp\left(\frac{\pi}{8k}\left(k+\frac{16n+4a-5}{k}\right)\right) \tag{117}$$

$$\leq \exp\left(\frac{\pi}{8}+\frac{\pi(16n+4a-5)}{8k^2}\right) \tag{118}$$

$$\leq \exp(3n\pi). \tag{119}$$

We now estimate the path of integration:

The chord connecting 0 with $z_T(h,k)$ can be no longer than the diameter of $K^{(-)}$, so the length along the arc from 0 to $z_T(h,k)$ can be no longer than $|z_T(h,k)|\frac{\pi}{2}$. Since $|z_T(h,k)| < \sqrt{2}/N$, we have a path length that is $O(N^{-1})$. This gives us

$$\left|\int\limits_{z_T(h,k)}^{0}\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z}+z(16n+4a-5)\right)\right)dz\right|$$

$$\leq \int\limits_{z_T(h,k)}^{0}\left|\exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z}+z(16n+4a-5)\right)\right)\right|dz \tag{120}$$

$$\leq \exp(3n\pi)\int\limits_{z_T(h,k)}^{0} dz \tag{121}$$

$$= O\left(\exp(3n\pi)N^{-1}\right). \tag{122}$$

The case for $\int\limits_{0}^{z_I(h,k)}$ is virtually identical. □

As a consequence of the previous Lemmas 4.1, 4.2, 4.3, and 4.4, we have

Theorem 4.5

$$g_a^{(d)}(n) = 2^{(\alpha-3)/2} \sum_{\substack{(k,8)=d,\\ k\le N}} \frac{i}{k} T_{a,d}(k) \sum_{\substack{0\le h<k,\\ (h,k)=1}} \omega_{a,d}(h,k) e^{-2\pi i n h/k}$$
$$\times \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{\Lambda(a,d)}{z} + z(16n+4a-5)\right)\right) dz$$
$$+ O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (123)$$

Estimating $g_a^{(8)}(n)$

In the case of $d = 8$ we can discard a large portion of what remains. Notice that

$$I_{a,8}^{(1)}(h,k) = \int_{z_I(h,k)}^{z_T(h,k)} \exp\left(\frac{\pi}{8k}\left(\frac{(b-4)^2-8}{z} + z(16n+4a-5)\right)\right) dz, \quad (124)$$

with $b \equiv ah \pmod 8$. If $b = 1, 7$, then the coefficient of $1/z$ in the exponent is 1. However, if $b = 3, 5$, then the coefficient is -7, and by almost identical reasoning of Lemmas 4.1, 4.2, and 4.3, applied to $I_{a,8}^{(1)}(h,k)$, we have

$$\left|I_{a,8}^{(1)}(h,k)\right| = O\left(\exp(3n\pi)N^{-1}\right). \quad (125)$$

Now $\alpha = 3$ and $T_{a,d}(k) = 1$. Since $b = 1, 7$ implies $h \equiv \pm a \pmod 8$, we have

$$g_a^{(8)}(n) = \sum_{\substack{(k,8)=8\\ k\le N}} \frac{i}{k} \sum_{\substack{0\le h<k,\\ (h,k)=1\\ h\equiv\pm a \pmod 8}} \omega_{a,8}(h,k) e^{-2\pi i n h/k}$$
$$\times \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz + O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (126)$$

Estimating $g_a^{(4)}(n)$

Lemmas 4.1, 4.2, 4.3, and 4.4 are sufficient to complete the estimation of $g_a^{(4)}(n)$. With $\alpha = 2$ and $T_{a,d}(k) = 1$, we have

Theorem 4.6

$$g_a^{(4)}(n) = \frac{1}{\sqrt{2}} \sum_{\substack{(k,8)=4,\\ k\le N}} \frac{i}{k} \sum_{\substack{0\le h<k,\\ (h,k)=1}} \omega_{a,4}(h,k) e^{-2\pi i nh/k} \times \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz + O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (127)$$

Estimating $g_a^{(2)}(n)$

In the case of $d = 2$, the coefficient of $1/z$ in the exponential of the integrand is never positive. Therefore, we may immediately apply the reasoning of Lemmas 4.1, 4.2, and 4.3 to both $I_{a,2}^{(0)}(h,k)$ and $I_{a,2}^{(1)}(h,k)$:

$$|I_{a,2}^{(0)}(h,k)| = |I_{a,2}^{(1)}(h,k)| = O\left(\exp(3n\pi)N^{-1}\right). \quad (128)$$

Therefore,

Theorem 4.7

$$g_a^{(2)}(n) = O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (129)$$

Estimating $g_a^{(1)}(n)$

Lemmas 4.1, 4.2, 4.3, and 4.4 are sufficient to complete the estimation of $g_a^{(1)}(n)$. Noting that $\alpha = 0$ and $T_{a,d}(k) = |\csc(\pi ak/8)|$, we have

Theorem 4.8

$$g_a^{(1)}(n) = \frac{1}{2\sqrt{2}} \sum_{\substack{(k,8)=1\\ k\le N}} \frac{i}{k} \left|\csc\left(\frac{\pi ak}{8}\right)\right| \sum_{\substack{0\le h<k,\\ (h,k)=1}} \omega_{a,1}(h,k) e^{-2\pi i nh/k} \times \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{4z} + z(16n+4a-5)\right)\right) dz + O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right). \quad (130)$$

Complete Formula

Combining (126), (127), (129), (130), and collecting the error terms, we have:

$$g_a(n) = \frac{1}{2\sqrt{2}} \sum_{\substack{(k,8)=1\\ k\le N}} \frac{i}{k} \left|\csc\left(\frac{\pi a k}{8}\right)\right| A_{a,1}(n,k) \times \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{4z} + z(16n+4a-5)\right)\right) dz$$
$$+ \frac{1}{\sqrt{2}} \sum_{\substack{(k,8)=4\\ k\le N}} \frac{i}{k} A_{a,4}(n,k) \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz$$
$$+ \sum_{\substack{(k,8)=8\\ k\le N}} \frac{i}{k} A_{a,8}(n,k) \oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz$$
$$+ O\left(e^{3n\pi} n^{1/3} N^{-1/3+\epsilon}\right), \quad (131)$$

with

$$A_{a,d}(n,k) = \sum_{\substack{0\le h<k,\\ (h,k)=1,\\ h\equiv \pm a\ (mod\ d)}} \omega_{a,d}(h,k) e^{-2\pi i n h/k}. \quad (132)$$

We represent the remaining integrals with modified Bessel functions [16]:

Lemma 5.1

$$\oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{z} + z(16n+4a-5)\right)\right) dz = \frac{-2\pi i}{\sqrt{16n+4a-5}} I_1\left(\frac{\pi\sqrt{16n+4a-5}}{4k}\right), \quad (133)$$

$$\oint_{K_k^{(-)}} \exp\left(\frac{\pi}{8k}\left(\frac{1}{4z} + z(16n+4a-5)\right)\right) dz = \frac{-\pi i}{\sqrt{16n+4a-5}} I_1\left(\frac{\pi\sqrt{16n+4a-5}}{8k}\right). \quad (134)$$

The first equality may be proved by changing variables, first by $z = 1/w$, and then by $w = 8kt/\pi$. We may then represent the integral with the modified Bessel function I_1 [16]: The second equality may be similarly proved by changing variables by $z = 1/w$, and then by $w = 32kt/\pi$.

Finishing the Limit Process

We now take (131), with $A_{a,d}(n,k)$ defined by (132), substitute and simplify through Lemma 5.1, and let $N \to \infty$. We now have our final formula.

Theorem 5.2 *Let $g_a(n)$ be the number of type-a Göllnitz–Gordon partitions of n, with $a = 1$ or 3. Then*

$$g_a(n) = \frac{\pi\sqrt{2}}{4\sqrt{16n+4a-5}} \sum_{(k,8)=1} \left|\csc\left(\frac{\pi ak}{8}\right)\right| \frac{A_{a,1}(n,k)}{k} I_1\left(\frac{\pi\sqrt{16n+4a-5}}{8k}\right)$$
$$+ \frac{\pi\sqrt{2}}{\sqrt{16n+4a-5}} \sum_{(k,8)=4} \frac{A_{a,4}(n,k)}{k} I_1\left(\frac{\pi\sqrt{16n+4a-5}}{4k}\right)$$
$$+ \frac{2\pi}{\sqrt{16n+4a-5}} \sum_{(k,8)=8} \frac{A_{a,8}(n,k)}{k} I_1\left(\frac{\pi\sqrt{16n+4a-5}}{4k}\right). \quad (135)$$

Numerical Tests

Mathematica was used to test (135), for k truncated. See the tables below (Tables 1 and 2). For the first 200 positive integers, our formula with $k \le 3\sqrt{n}$ gives the correct value with an absolute error less than 0.33. Since $g_a(n) \in \mathbb{Z}$, we need only round our formulaic value to the nearest integer to achieve the correct answer.

Table 1 $g_1(n)$ compared to (135) truncated for k, with $a = 1$

n	$g_1(n)$	Eq. (135), $a = 1$, $1 \le k \le 3\sqrt{n}$	Absolute error of (135)
1	1	0.7784305652	0.2215694348
2	1	0.7196351376	0.2803648624
3	1	1.114485490	0.114485490
4	2	1.890769460	0.109230540
5	2	2.146945231	0.146945231
6	2	2.174897898	0.174897898
7	3	2.917027886	0.082972114
8	4	3.994864237	0.005135763
9	5	4.903833678	0.096166322
10	5	5.108441112	0.108441112
20	26	26.07125673	0.07125673
40	288	287.9388309	0.0611691
60	1989	1988.942843	0.057157
80	10,570	10,569.99993	0.00007
100	47,091	47,090.99132	0.00868
150	1,191,854	1,191,853.996	0.004
200	18,900,623	18,900,622.99	0.001

Table 2 $g_3(n)$ compared to (135) truncated for k, with $a = 3$

n	$g_3(n)$	Eq. (135), $a = 3$, $1 \le k \le 3\sqrt{n}$	Absolute error of (135)
1	0	0.2908871603	0.2908871603
2	0	0.1385488254	0.1385488254
3	1	0.8129880460	0.1870119540
4	1	0.9584818018	0.0415181982
5	1	0.8666320258	0.1333679742
6	1	0.9177374697	0.0822625303
7	1	1.323340028	0.323340028
8	2	2.095679009	0.095679009
9	2	2.042654099	0.042654099
10	2	1.953812941	0.046187059
20	12	12.01649403	0.01649403
40	127	126.9760443	0.0239557
60	865	865.0090307	0.0090307
80	4560	4560.002784	0.002784
100	20,223	20,223.00416	0.00416
150	508,454	508,454.0481	0.0481
200	8,034,534	8,034,534.006	0.006

References

1. G.E. Andrews, in *The Theory of Partitions*. Encyclopedia of Mathematics and Its Applications, vol. 2 (Addison-Wesley, Reading, 1976). Reissued, Cambridge, 1998
2. B.C. Berndt, *Number Theory in the Spirit of Ramanujan* (American Mathematical Society, Providence, 2006)
3. H. Göllnitz, Partitionen mit Differenzenbedingungen. Journal für die reine und angewandte Mathematik **225**, 154–190 (1967)
4. B. Gordon, Some continued fractions of the Rogers–Ramanujan type. Duke Math. J. **32**, 741–748 (1965)
5. G.H. Hardy, S. Ramanujan, Asymptotic formulæ in combinatory analysis. Proc. Lond. Math. Soc. **2**(17), 75–115 (1918)
6. G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers*, 6th edn. (Oxford University Press, Oxford/New York, 2008)
7. J. Lehner, A partition function connected with the modulus five. Duke Math. J. **8**, 631–655 (1941)
8. I. Niven, On a certain partition function. Am. J. Math. **62**, 353–364 (1940)
9. H. Rademacher, On the partition function $p(n)$. Proc. Lond. Math. Soc. **43**(2), 241–254 (1937)
10. H. Rademacher, On the expansion of the partition function in a series. Ann. Math. **44**(2), 416–422 (1943)
11. H. Rademacher, in *Topics in Analytic Number Theory*. Die Grundelhren der Mathematischen Wissenschaften, Bd. 169 (Springer, Berlin/New York, 1973)
12. H. Salié, Zur Abschätzung der Fourierkoeffizienten ganzer Modulformen. Mathematische Zeitschrift **36**, 263–278 (1933)
13. A.V. Sills, Towards an automation of the circle method, in *Gems in Experimental Mathematics*. Contemporary Mathematics, vol. 517 (American Mathematical Society, Providence, 2010), pp. 321–338
14. E.M. Stein, R. Shakarchi, in *Complex Analysis*. Princeton Lectures in Analysis, vol. 2 (Princeton University Press, Princeton, 2003)
15. R.C. Vaughan, *The Hardy–Littlewood Method*, 1st edn. (Cambridge University Press, Cambridge, 1981)
16. G.N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd edn. (Cambridge University Press, Cambridge, 1944)

On Toeplitz Operators with Quasi-radial and Pseudo-homogeneous Symbols

Nikolai Vasilevski

In memory of the brilliant mathematician and personality
Cora Sadosky

Abstract We explore a new wide class of symbols that generate commutative Banach algebras on each weighted Bergman space on the unit ball in $\mathbb{C}^n$. These symbols are a natural extension of the previously studied quasi-radial quasi-homogeneous symbols, and contain them as a very special particular case. Roughly speaking, instead of the fixed specific bounded continuous functions we admit now any L_∞-functions.

Mathematics Subject Classification (2010): Primary 47B35; Secondary 47L80, 32A36.

Introduction

After a successful description of commutative C^*-algebras generated by Toeplitz operators on the unit ball $\mathbb{B}^n$ [6, 7], it was unexpectedly observed that for $n \geq 2$ there exist many others algebras generated by Toeplitz operators that are commutative on each weighted Bergman space. All of them were Banach (not C^*), and their description was done in terms of generating Toeplitz operators. The first result in this direction [9] deals with the so-called quasi-radial quasi-homogeneous symbols. The subsequent results [1, 2, 8] were based on similar ideas exploiting a quasi-homogeneity of corresponding symbols. Note that the notion of quasi-homogeneous symbols is based on the use of the spherical coordinates.

This work was partially supported by CONACYT Project 238630, México.

N. Vasilevski (✉)
Departamento de Matemáticas, CINVESTAV, Apartado Postal 14-740,
07000 México, D.F., México
e-mail: nvasilev@math.cinvestav.mx

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_15

Quite recently [5], for the case of two-dimensional ball $\mathbb{B}^2$, the much more wide class of quasi-homogeneous like symbols was described. These symbols also generate via corresponding Toeplitz operators commutative Banach algebras on each weighted Bergman space, and their existence was hidden just by the use of the spherical coordinates in the approach of [9].

This paper extends the results of [5] for the case of the unit ball $\mathbb{B}^n$ with $n > 2$, and explores a new wide class of symbols that generate commutative Banach algebras on each weighted Bergman space on these balls. We call these new symbols *pseudo-homogeneous*, they include the previous quasi-homogeneous symbols as a very special particular case. Roughly speaking, instead of a fixed specific bounded continuous functions we admit now any L_∞-functions.

Section "Preliminaries" collects notation used throughout the paper. In section "Quasi-radial and Pseudo-homogeneous Symbols" we introduce pseudo-homogeneous symbols and describe the action of Toeplitz operators with quasi-radial and pseudo-homogeneous symbols. Section "Commutative Algebras" devoted to the description of commutative Banach algebras generated by Toeplitz operators with above symbols and of some their common properties. Yet another option to build commutative Banach Toeplitz operator algebras is presented in the last section "Yet Another Option".

Preliminaries

Let $\mathbb{B}^n$ be the unit ball in $\mathbb{C}^n$,

$$\mathbb{B}^n = \{z = (z_1, \ldots, z_n) \in \mathbb{C}^n : \ |z|^2 = |z_1|^2 + \ldots + |z_n|^2 < 1\},$$

and let S^{2n-1} be the corresponding (real) unit sphere, the boundary of the unit ball $\mathbb{B}^n$.

In what follows we will use the notation $\tau(\mathbb{B}^m)$ for the base of the unit ball $\mathbb{B}^m$, considered as a Reinhard domain, i.e.,

$$\tau(\mathbb{B}^m) = \{(r_1, \ldots, r_m) = (|z_1|, \ldots, |z_m|) \ : \ r^2 = r_1^2 + \ldots + r_m^2 \in [0, 1)\}.$$

We denote as well by $\mathbf{B}^m$ the *real* m-dimensional unit ball,

$$\mathbf{B}^m = \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m : \ x^2 = x_1^2 + \ldots + x_m^2 < 1\}.$$

Then, of course, $\tau(\mathbb{B}^m) = \mathbf{B}^m \cap \mathbb{R}^m_+ =: \mathbf{B}^m_+$.

Given a multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}_+^n$ we will use the standard notation,

$$\alpha! = \alpha_1!\,\alpha_2! \cdots \alpha_n!,$$
$$z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n},$$

for $p = (p_1, p_2, \ldots, p_n) \in \mathbb{Z}^n$ we set

$$|p| = p_1 + p_2 + \ldots + p_n \qquad \text{and} \qquad \|p\| = |p_1| + |p_2| + \ldots + |p_n|.$$

Denote by $dV = dx_1dy_1 \ldots dx_ndy_n$, where $z_l = x_l + iy_l$, $l = 1, 2, \ldots, n$, the standard Lebesgue measure in $\mathbb{C}^n$; and let dS be the corresponding surface measure on S^{2n-1}. We introduce the standard one-parameter family of weighted measures,

$$dv_\lambda(z) = \frac{\Gamma(n+\lambda+1)}{\pi^n\,\Gamma(\lambda+1)}\,(1-|z|^2)^\lambda\,dV(z), \quad \lambda > -1,$$

which are probability ones in $\mathbb{B}^n$; and recall (see, for example, [10, Section 1.3]) that

$$\int_{\mathbb{B}^n} z^\alpha \overline{z}^\beta dv_\lambda(z) = \delta_{\alpha,\beta}\,\frac{\alpha!\,\Gamma(n+\lambda+1)}{\Gamma(n+|\alpha|+\lambda+1)}. \tag{1}$$

We introduce the weighted space $L_2(\mathbb{B}^n, dv_\lambda)$ and its subspace, the weighted Bergman space $\mathcal{A}_\lambda^2 = \mathcal{A}_\lambda^2(\mathbb{B}^n)$, which consists of all functions analytic in $\mathbb{B}^n$. The (orthogonal) Bergman projection B_λ of $L_2(\mathbb{B}^n, dv_\lambda)$ onto $\mathcal{A}_\lambda^2(\mathbb{B}^n)$ is given by

$$(B_\lambda\varphi)(z) = \int_{\mathbb{B}^n} \frac{\varphi(\zeta)\,dv_\lambda(\zeta)}{(1-z\cdot\overline{\zeta})^{n+\lambda+1}}.$$

Finally, given a function $a(z) \in L_\infty(\mathbb{B}^n)$, the Toeplitz operator T_a with symbol a acts on $\mathcal{A}_\lambda^2(\mathbb{B}^n)$ as follows

$$T_a : \varphi \in \mathcal{A}_\lambda^2(\mathbb{B}^n) \longmapsto B_\lambda(a\varphi) \in \mathcal{A}_\lambda^2(\mathbb{B}^n).$$

Quasi-radial and Pseudo-homogeneous Symbols

Let $k = (k_1, \ldots, k_m)$ be a tuple of positive integers whose sum is equal to n: $k_1 + \ldots + k_m = n$. The length of such a tuple may obviously vary from 1, for $k = (n)$, to n, for $k = (1, \ldots, 1)$.

Throughout the paper we fix a tuple $k = (k_1, \ldots, k_m)$, with $k_1 \leq k_2 \leq \ldots \leq k_m$, and rearrange the n coordinates of $z \in \mathbb{B}^n$ in m groups, each one of which has k_j, $j = 1, \ldots, m$, entries. We will use the notation

$$z_{(1)} = (z_{1,1}, \ldots, z_{1,k_1}),\ z_{(2)} = (z_{2,1}, \ldots, z_{2,k_2}),\ \ldots,\ z_{(m)} = (z_{m,1}, \ldots, z_{m,k_m})$$

with

$$\begin{aligned} &z_{1,1} = z_1,\ \ z_{1,2} = z_2,\ \ \ldots,\ \ z_{1,k_1} = z_{k_1}, \\ &z_{2,1} = z_{k_1+1},\ \ \ldots,\ \ z_{2,k_2} = z_{k_1+k_2} \ldots \end{aligned} \tag{2}$$

Note that the above ordering condition on the tuple k can be easily fulfilled for an arbitrary (non ordered tuple) by making the biholomorphism of the unit ball that interchanges the coordinates of z.

We will also use an alternative representation of a point $z = (z_1, \ldots, z_n) \in \mathbb{B}^n$:

$$z = (z_{(1)}, \ldots, z_{(m)}), \qquad \text{where} \quad z_{(j)} \in \mathbb{B}^{k_j},\ \ j = 1, \ldots, m.$$

In general, given any n-tuple u, we will also use two its alternative representations

$$u = (u_1, \ldots, u_n) = (u_{(1)}, \ldots, u_{(m)}),$$

where

$$\begin{aligned} &u_{1,1} = u_1,\ u_{1,2} = u_2,\ \ \ldots,\ \ u_{1,k_1} = u_{k_1}, \\ &u_{2,1} = u_{k_1+1},\ \ \ldots,\ \ u_{2,k_2} = u_{k_1+k_2},\ \ldots,\ u_{m,k_m} = u_n. \end{aligned} \tag{3}$$

We represent then each coordinate of $z \in \mathbb{B}^n$ (which is the same as each coordinate of $z_{(j)}, j = 1, \ldots, m$) in the form

$$z_i = |z_i| t_i \qquad \text{or} \qquad z_{j,\ell} = |z_{j,\ell}| t_{j,\ell},$$

where t_i and $t_{j,\ell}$ belong to $\mathbb{T} = S^1$. For each portion $z_{(j)}, j = 1, \ldots, m$, of a point z we introduce its "common" radius

$$r_j = \sqrt{|z_{j,1}|^2 + \ldots |z_{j,k_j}|^2},$$

and represent the coordinates of $z_{(j)}$ in the form

$$z_{j,\ell} = r_j\, s_{j,\ell}\, t_{j,\ell},\ \ \text{where}\ \ \ell = 1, \ldots, k_j,\ \ s_{(j)} = (s_{j,1}, \ldots, s_{j,k_j}) \in S_+^{k_j-1} := S^{k_j-1} \cap \mathbb{R}_+^{k_j}.$$

Recall [9] that a bounded measurable function $a = a(z)$, $z \in \mathbb{B}^n$, is called *k-quasi-radial* if it depends only on $r_1, \ldots, r_m$.

The following result is Lemma 3.1 of [9].

Lemma 3.1 *Given a bounded measurable k-quasi-radial function* $a = a(r_1, \ldots, r_m)$, *we have*

$$T_a z^\alpha = \gamma_{a,k,\lambda}(\alpha) z^\alpha, \quad \alpha \in \mathbb{Z}_+^n,$$

with

$$\begin{aligned}
&\gamma_{a,k,\lambda}(\alpha) = \gamma_{a,k,\lambda}(|\alpha_{(1)}|, \ldots, |\alpha_{(m)}|) \\
&= \frac{2^m \Gamma(n + |\alpha| + \lambda + 1)}{\Gamma(\lambda + 1) \prod_{j=1}^m (k_j - 1 + |\alpha_{(j)}|)!} \int_{\tau(\mathbb{B}^m)} a(r_1, \ldots, r_m)(1 - |r|^2)^\lambda \\
&\quad \times \prod_{j=1}^m r_j^{2|\alpha_{(j)}|+2k_j-1} dr_j \\
&= \frac{\Gamma(n + |\alpha| + \lambda + 1)}{\Gamma(\lambda + 1) \prod_{j=1}^m \Gamma(k_j + |\alpha_{(j)}|)} \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m}) \\
&\quad \times \prod_{j=1}^m r_j^{|\alpha_{(j)}|+k_j-1} (1 - (r_1 + \ldots + r_m))^\lambda dr_1 \cdots r_m,
\end{aligned}$$

where $\Delta_m = \{(r_1, \ldots, r_m) \in \mathbb{R}_+^m \, : \, r_1 + \ldots + r_m \in [0, 1)\}$.

We introduce now an extension of the class of quasi-homogeneous functions defined in [9]. Note that here we use a slightly different notation for multi-indices, instead of a pair of the orthogonal multi-indices p and q of [9], we will consider just one multi-index $p = (p_1, \ldots, p_n)$ that satisfies the condition $|p| = 0$.

A function ψ is called *pseudo-homogeneous* (or *k-pseudo-homogeneous*) if it has the form

$$\psi(z) = b(s_{(1)}, \ldots, s_{(m)}) \, t^p = b(s_{(1)}, \ldots, s_{(m)}) \prod_{j=1}^m t_{(j)}^{p_{(j)}},$$

where $b(s_{(1)}, \ldots, s_{(m)}) \in L_\infty(S_+^{k_1-1} \times \ldots \times S_+^{k_m-1})$ and $p = (p_1, \ldots, p_n) = (p_{(1)}, \ldots, p_{(m)}) \in \mathbb{Z}^n$.

Most frequently we will consider the case when $b(s_{(1)}, \ldots, s_{(m)}) = \prod_{j=1}^m b_j(s_{(j)})$ with $b_j \in L_\infty(S_+^{k_j-1})$, for all $j = 1, \ldots, m$.

Note that quasi-homogeneous functions, introduced in [9], correspond to the case when each $b_j = b_j(s_{(j)})$ has the form

$$b_j(s_{(j)}) = s_{j,1}^{|p_{j,1}|} \cdots s_{j,k_j}^{|p_{j,k_j}|},$$

so that

$$\psi(z) = \prod_{j=1}^m s_{j,1}^{|p_{j,1}|} t_{j,1}^{p_{j,1}} \cdots s_{j,k_j}^{|p_{j,k_j}|} t_{j,k_j}^{p_{j,k_j}}. \tag{4}$$

Consider now the following k-quasi-radial pseudo-homogeneous symbol

$$\varphi(z) = a(r_1, \ldots, r_m) \prod_{j=1}^{m} b_j(s_{(j)})\, t^p = a(r_1, \ldots, r_m) \prod_{j=1}^{m} b_j(s_{(j)})\, t_{(j)}^{p_{(j)}},$$

where $a = a(r_1, \ldots, r_m) \in L_\infty(\tau(\mathbb{B}^m))$, $b_j = b_j(s_{(j)}) \in L_\infty(S_+^{k_j-1})$, $j = 1, \ldots, m$, and $t = (t_1, \ldots, t_n) = (t_{(1)}, \ldots, t_{(m)}) \in \mathbb{T}^n$, $p = (p_1, \ldots, p_n) = (p_{(1)}, \ldots, p_{(m)}) \in \mathbb{Z}^n$.

In what follows we will use the parametrization of each $S_+^{k_j-1}$ by its first $k_j - 1$ coordinates $s_{j_1}, \ldots, s_{j,k_j-1}$, so that $s_{j,k_j} = \sqrt{1 - (s_{j_1}^2 + \ldots + s_{j,k_j-1}^2)}$, and $b_j = b_j(s_{(j)}) = b_j(s_{j_1}, \ldots, s_{j,k_j-1}) \in L_\infty(\mathbf{B}_+^{k_j-1})$. In this parametrization the standard Euclidean volume element dS_j on $S_+^{k_j-1}$, used in the next lemma, is given by

$$dS_j = \frac{ds_{j_1} \cdots ds_{j,k_j-1}}{\sqrt{1 - (s_{j_1}^2 + \ldots + s_{j,k_j-1}^2)}}.$$

Lemma 3.2 *The Toeplitz operator T_φ with the above symbol $\varphi(z)$ acts on monomials z^α, $\alpha \in \mathbb{Z}_+^n$, as follows*

$$T_\varphi\, z^\alpha = \begin{cases} 0, & \text{if } \exists\ i \text{ such that } \ \alpha_i + p_i < 0 \\ \widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha)\, z^{\alpha+p}, & \text{if } \forall\ i \ \ \alpha_i + p_i \geq 0 \end{cases},$$

with

$$\begin{aligned}
\widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha) &= \frac{\Gamma(n + |\alpha + p| + \lambda + 1)}{\Gamma(\lambda + 1) \prod_{j=1}^{m} \Gamma(|\alpha_{(j)}| + \frac{1}{2}|p_{(j)}| + k_j)} \\
&\times \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m})\, r_j^{|\alpha_{(j)}| + \frac{1}{2}|p_{(j)}| + k_j - 1} (1 - (r_1 + \ldots + r_m))^\lambda\, dr_1 \cdots dr_m \\
&\times \prod_{j=1}^{m} \frac{\Gamma(|\alpha_{(j)}| + \frac{1}{2}|p_{(j)}| + k_j)}{\prod_{\ell=1}^{k_j} \Gamma(\alpha_{j,\ell} + p_{j,\ell} + 1)} \int_{\Delta_{k_j-1}} b(s_{(j)}^{1/2}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell} + \frac{1}{2} p_{j,\ell}} \\
&\times (1 - (s_{j,1} + \ldots + s_{j,k_j-1}))^{\alpha_{j,k_j} + \frac{1}{2} p_{j,k_j}}\, ds_{j,1} \cdots ds_{j,k_j-1},
\end{aligned}$$

where $b(s_{(j)}^{1/2}) = b(\sqrt{s_{j,1}}, \ldots, \sqrt{s_{j,k_j-1}})$.

Proof Given two multi-indices $\alpha,\ \beta \in \mathbb{Z}_+^n$, we calculate

$$\begin{aligned}
\langle T_\varphi z^\alpha, z^\beta \rangle &= \langle a \prod_{j=1}^{m} b_j\, t^p\, z^\alpha, z^\beta \rangle \\
&= \frac{\Gamma(n + \lambda + 1)}{\pi^n \Gamma(\lambda + 1)} \int_{\mathbb{B}^n} a \prod_{j=1}^{m} b_j \prod_{l=1}^{n} |z_l|^{\alpha_l + \beta_l} (1 - |r|^2)^\lambda\, t^{\alpha - \beta + p}\, dV(z)
\end{aligned}$$

$$= \frac{\Gamma(n+\lambda+1)}{\pi^n\Gamma(\lambda+1)} \int_{\tau(\mathbb{B}^n)} a \prod_{j=1}^{m} b_j \prod_{l=1}^{n} |z_l|^{\alpha_l+\beta_l+1}(1-|r|^2)^{\lambda}\, d|z|_1 \cdots d|z|_n$$

$$\times \prod_{l=1}^{n} \int_{\mathbb{T}} t_l^{\alpha_l-\beta_l+p_l} \tfrac{dt_l}{it_l}.$$

The last n integrals are different from zero if and only if $\beta = \alpha + p$, in the last case each of them is equal to 2π. Thus, setting $\beta = \alpha + p$, we have

$$\langle T_{\varphi} z^{\alpha}, z^{\alpha+p}\rangle = \langle a \prod_{j=1}^{m} b_j\, t^p\, z^{\alpha}, z^{\alpha+p}\rangle$$

$$= 2^n \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\tau(\mathbb{B}^n)} a \prod_{j=1}^{m} b_j \prod_{l=1}^{n} |z_l|^{2\alpha_l+p_l+1}(1-|r|^2)^{\lambda}\, d|z|_1 \cdots d|z|_n.$$

Changing the variables $r_{j,\ell} = r_j s_{j,\ell}$ (with $d|z|_{j,1} \cdots d|z|_{j,k_j} = r_j^{k_j-1} dr_j dS_j$), we obtain

$$\langle T_{\varphi} z^{\alpha}, z^{\beta}\rangle = 2^n \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\tau(\mathbb{B}^m)\times\prod_{j=1}^{m} S_+^{k_j-1}} a(r_1, \ldots, r_m)$$

$$\times \prod_{j=1}^{m} r_j^{2|\alpha_{(j)}|+|p_{(j)}|+2k_j-1} (1-|r|^2)^{\lambda} \prod_{j=1}^{m} b_j(s_{(j)}) \prod_{\ell=1}^{k_j} s_{j,\ell}^{2\alpha_{j,\ell}+p_{j,\ell}+1}\, dr_j\, dS_j$$

$$= 2^{n-m} \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m}) \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j-1}$$

$$\times\ (1-(r_1+\ldots+r_m))^{\lambda}\, dr_1 \cdots dr_m \prod_{j=1}^{m} \int_{S_+^{k_j-1}} b_j(s_{(j)}) \prod_{\ell=1}^{k_j} s_{j,\ell}^{2\alpha_{j,\ell}+p_{j,\ell}+1}\, dS_j$$

$$= \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m}) \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j-1}$$

$$\times\ (1-(r_1+\ldots+r_m))^{\lambda}\, dr_1 \cdots dr_m \prod_{j=1}^{m} 2^{k_j-1} \int_{S_+^{k_j-1}} b_j(s_{(j)}) \prod_{\ell=1}^{k_j} s_{j,\ell}^{2\alpha_{j,\ell}+p_{j,\ell}+1}\, dS_j$$

$$= \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m}) \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j-1}$$

$$\times\ (1-(r_1+\ldots+r_m))^{\lambda}\, dr_1 \cdots dr_m$$

$$\times \prod_{j=1}^{m} 2^{k_j-1} \int_{\mathbf{B}_+^{k_j-1}} b(s_{(j)}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{2\alpha_{j,\ell}+p_{j,\ell}+1}$$

$$\times \; (1-(s_{j,1}^2+\ldots+s_{j,k_j-1}^2))^{\alpha_{j,k_j}+\frac{1}{2}p_{j,k_j}} \, ds_{j,1}\cdots ds_{j,k_j-1}$$

$$= \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\Delta_m} a(\sqrt{r_1},\ldots,\sqrt{r_m}) \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j-1}$$

$$\times \; (1-(r_1+\ldots+r_m))^{\lambda} \, dr_1\cdots dr_m$$

$$\times \prod_{j=1}^{m} \int_{\Delta_{k_j-1}} b(s_{(j)}^{1/2}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell}+\frac{1}{2}p_{j,\ell}}$$

$$\times \; (1-(s_{j,1}+\ldots+s_{j,k_j-1}))^{\alpha_{j,k_j}+\frac{1}{2}p_{j,k_j}} \, ds_{j,1}\cdots ds_{j,k_j-1},$$

where $b(s_{(j)}^{1/2}) = b(\sqrt{s_{j,1}},\ldots,\sqrt{s_{j,k_j-1}})$.

For $\beta = \alpha + p$, by (1), we have

$$\begin{aligned}\langle z^{\alpha+p}, z^{\alpha+p}\rangle &= \frac{(\alpha+p)!\Gamma(n+\lambda+1)}{\Gamma(n+|\alpha+p|+\lambda+1)} \\ &= \prod_{l=1}^{n} \Gamma(\alpha_l+p_l+1) \frac{\Gamma(n+\lambda+1)}{\Gamma(n+|\alpha+p|+\lambda+1)}.\end{aligned}$$

That is

$$T_\varphi z^\alpha = \begin{cases} 0, & \text{if } \exists\, i \text{ such that } \alpha_i+p_i<0 \\ \widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha)\, z^{\alpha+p}, & \text{if } \forall\, i \;\; \alpha_i+p_i \geq 0 \end{cases},$$

where

$$\widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha) = \frac{\Gamma(n+|\alpha+p|+\lambda+1}{\Gamma(\lambda+1)\prod_{j=1}^{m}\Gamma(|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j)} \int_{\Delta_m} a(\sqrt{r_1},\ldots,\sqrt{r_m})$$

$$\times \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j-1} \, (1-(r_1+\ldots+r_m))^{\lambda} \, dr_1\cdots dr_m$$

$$\times \prod_{j=1}^{m} \frac{\Gamma(|\alpha_{(j)}|+\frac{1}{2}|p_{(j)}|+k_j)}{\prod_{\ell=1}^{k_j}\Gamma(\alpha_{j,\ell}+p_{j,\ell}+1)} \int_{\Delta_{k_j-1}} b(s_{(j)}^{1/2}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell}+\frac{1}{2}p_{j,\ell}}$$

$$\times \; (1-(s_{j,1}+\ldots+s_{j,k_j-1}))^{\alpha_{j,k_j}+\frac{1}{2}p_{j,k_j}} \, ds_{j,1}\cdots ds_{j,k_j-1}.$$

□

For the case when $|p_{(j)}| = 0$, for each $j = 1,\ldots,m$, Lemma 3.2 yields

Corollary 3.3 *The Toeplitz operator T_φ with symbol*

$$\varphi(z) = a(r_1, \ldots, r_m) \prod_{j=1}^{m} b_j(s_{(j)})\, t_{(j)}^{p_{(j)}},$$

where $|p_{(j)}| = 0$ for each $j = 1, \ldots, m$, acts on monomials z^α, $\alpha \in \mathbb{Z}_+^n$, as follows

$$T_\varphi\, z^\alpha = \begin{cases} 0, & \text{if } \exists\ \ell \text{ such that } \alpha_\ell + p_\ell < 0 \\ \widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha)\, z^{\alpha+p}, & \text{if } \forall\ \ell\ \ \alpha_\ell + p_\ell \geq 0 \end{cases},$$

with

$$\widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha) = \frac{\Gamma(n + |\alpha| + \lambda + 1)}{\Gamma(\lambda + 1) \prod_{j=1}^{m} \Gamma(|\alpha_{(j)}| + k_j)}$$
$$\times \int_{\Delta_m} a(\sqrt{r_1}, \ldots, \sqrt{r_m}) \prod_{j=1}^{m} r_j^{|\alpha_{(j)}|+k_j-1} (1 - (r_1 + \ldots + r_m))^\lambda\, dr_1 \cdots dr_m$$
$$\times \prod_{j=1}^{m} \frac{\Gamma(|\alpha_{(j)}| + k_j)}{\prod_{\ell=1}^{k_j} \Gamma(\alpha_{j,\ell} + p_{j,\ell} + 1)} \int_{\Delta_{k_j-1}} b_j(s_{(j)}^{1/2}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell}+\frac{1}{2}p_{j,\ell}}$$
$$\times (1 - (s_{j,1} + \ldots + s_{j,k_j-1}))^{\alpha_{j,k_j}+\frac{1}{2}p_{j,k_j}}\, ds_{j,1} \cdots ds_{j,k_j-1},$$

where $b(s_{(j)}^{1/2}) = b(\sqrt{s_{j,1}}, \ldots, \sqrt{s_{j,k_j-1}})$.

Observe now that for $\varphi(z) = a(r_1, \ldots, r_m)$ we have that $\widetilde{\gamma}_{a,k,p,\lambda}(\alpha) = \gamma_{a,k,p,\lambda}(\alpha)$ (see Lemma 3.1); and for $\varphi(z) = b_j(s_{(j)})\, t_{(j)}^{p_{(j)}}$, $|p_{(j)}| = 0$ and $p_{(l)} = 0$ for all $l \neq j$, we have that

$$T_{b_j(s_{(j)})\, t_{(j)}^{p_{(j)}}} z^\alpha = \widetilde{\gamma}_{b_j,k,p_j}(\alpha) z^{\alpha+p},$$

where

$$\widetilde{\gamma}_{b_j,k,p_{(j)}}(\alpha) = \frac{\Gamma(|\alpha_{(j)}| + k_j)}{\prod_{\ell=1}^{k_j} \Gamma(\alpha_{j,\ell} + p_{j,\ell} + 1)} \int_{\Delta_{k_j-1}} b_j(s_{(j)}^{1/2}) \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell}+\frac{1}{2}p_{j,\ell}}$$
$$\times (1 - (s_{j,1} + \ldots + s_{j,k_j-1}))^{\alpha_{j,k_j}+\frac{1}{2}p_{j,k_j}}\, ds_{j,1} \cdots ds_{j,k_j-1}.$$

The last formula implies, in particular, that the action of the operator $T_{b_j(s_{(j)})\, t_{(j)}^{p_{(j)}}}$ does not depend on the weight parameter λ.

We mention as well that if $k_j = 1$, then $|p_{(j)}| = 0$ implies that $p_{(j)} = (0)$ and that $b_j(s_{(j)}) = \text{const}$. That is, the corresponding part in a pseudo-homogeneous symbol can be just omitted.

Corollary 3.4 *The Toeplitz operator T_φ with symbol*

$$\varphi(z) = a(r_1, \ldots, r_m) \prod_{j=1}^{m} b_j(s_{(j)})\, t_{(j)}^{p_{(j)}},$$

where $|p_{(j)}| = 0$ for each $j = 1, \ldots, m$, acts on monomials z^α, $\alpha \in \mathbb{Z}_+^n$, as follows

$$T_\varphi\, z^\alpha = \begin{cases} 0, & \text{if } \exists\ \ell \text{ such that } \alpha_\ell + p_\ell < 0 \\ \widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha)\, z^{\alpha+p}, & \text{if } \forall\ \ell\ \ \alpha_\ell + p_\ell \geq 0 \end{cases},$$

with

$$\widetilde{\gamma}_{\varphi,k,p,\lambda}(\alpha) = \gamma_{a,k,p,\lambda}(\alpha) \prod_{j=1}^{m} \widetilde{\gamma}_{b_j,k,p_{(j)}}(\alpha).$$

Thus the Toeplitz operators T_a, $T_{b_j(s_{(j)})\, t_{(j)}^{p_{(j)}}}$, $j = 1, \ldots, m$, pairwise commute and

$$T_{a \prod_{j=1}^{m} b_j\, t_{(j)}^{p_{(j)}}} = T_a \prod_{j=1}^{m} T_{b_j\, t_{(j)}^{p_{(j)}}}.$$

Note that for the quasi-homogeneous symbol (4), considered in [9], with $|p_{(j)}| = 0$, for each $j = 1, \ldots, m$, we have

$$\widetilde{\gamma}_{\psi,k,p,\lambda}(\alpha) = \prod_{j=1}^{m} \frac{\Gamma(|\alpha_{(j)}| + k_j)}{\prod_{\ell=1}^{k_j} \Gamma(\alpha_{j,\ell} + p_{j,\ell} + 1)} \int_{\Delta_{k_j-1}} \prod_{\ell=1}^{k_j-1} s_{j,\ell}^{\alpha_{j,\ell} + \frac{1}{2}(p_{j,\ell} + |p_{j,\ell}|)}$$

$$\times\ (1 - (s_{j,1} + \ldots + s_{j,k_j-1}))^{\alpha_{j,k_j} + \frac{1}{2}(p_{j,k_j} + |p_{j,k_j}|)}\, ds_{j,1} \cdots ds_{j,k_j-1}$$

$$= \prod_{j=1}^{m} \frac{\Gamma(|\alpha_{(j)}| + k_j)}{\Gamma(|\alpha_{(j)}| + \frac{1}{2}\|p_{(j)}\| + k_j)} \prod_{\ell=1}^{k_j-1} \frac{\Gamma(\alpha_{j,\ell} + \frac{1}{2}(p_{j,\ell} + |p_{j,\ell}|) + 1)}{\Gamma(\alpha_{j,\ell} + p_{j,\ell} + 1)},$$

which, after returning to the notation of [9], recovers formula (4.1) of [9] with $a \equiv 1$.

Commutative Algebras

The results of the previous section, and especially Corollary 3.4, permit us to describe many new commutative Banach algebras generated by Toeplitz operators. To characterize them we proceed as follows.

We start again with a tuple $k = (k_1, k_2, \ldots, k_m)$ of positive integers with $k_1 + k_2 + \ldots + k_m = n$ and $k_1 \leq k_2 \leq \ldots \leq k_m$. Consider then the set $L_\infty(\tau(\mathbb{B}^m))$ of all k-quasi-radial symbols $a = a(r_1, \ldots, r_m)$, and denote by $\mathcal{T}_\lambda(k\text{-}qr)$ the C^*-algebra generated by all Toeplitz operators T_a with k-quasi-radial symbols $a \in L_\infty(\tau(\mathbb{B}^m))$.

We fix then a tuple $b = (b_1, b_2, .., b_m)$ of functions $b_j = b_j(s_{(j)}) \in L_\infty(S_+^{k_j-1})$ (or $b_j = b_j(s_{j_1}, \ldots, s_{j,k_j-1}) \in L_\infty(\mathbf{B}_+^{k_j-1})$, after the parametrization of $S_+^{k_j-1}$) and a tuple $p = (p_1, p_2, \ldots, p_n) \in \mathbb{Z}^n$ with the property $|p_{(j)}| = 0$, for all $j = 1, \ldots, m$, and denote by $\mathcal{T}(k, b, p)$ the unital Banach algebra generated by Toeplitz operators $T_{b_j t_{(j)}^{p_{(j)}}}$, for $j = 1, \ldots, m$.

Note that, contrary to the case of the algebra $\mathcal{T}_\lambda(k\text{-}qr)$, the action of generators $T_{b_j t_{(j)}^{p_{(j)}}}$ and thus the properties of the algebra $\mathcal{T}(k, b, p)$ do not depend on the weight parameter λ.

Then, by Corollary 3.4, the Banach algebra generated by elements of $\mathcal{T}_\lambda(k\text{-}qr)$ and elements of $\mathcal{T}(k, b, p)$ is commutative. We group all the ingredients that define this commutative algebra into one set $\mathbf{d} = \{k, L_\infty(\tau(\mathbb{B}^m)), b, p\}$ and denote this algebra by $\mathcal{T}_\lambda(\mathbf{d})$. That is the Banach algebra $\mathcal{T}_\lambda(\mathbf{d})$ is generated by the elements of the algebras $\mathcal{T}_\lambda(k\text{-}qr)$ and $\mathcal{T}(k, b, p)$.

In the rest of the section we will list some common properties of the algebras $\mathcal{T}_\lambda(\mathbf{d})$; for the specific case of [9], see [3].

For each $\kappa = (\kappa_1, \ldots, \kappa_m) \in \mathbb{Z}_+^m$ we introduce the finite dimensional subspace H_κ of the Bergman space $\mathcal{A}_\lambda^2(\mathbb{B}^n)$:

$$H_\kappa := \operatorname{span}\left\{e_\alpha \ : \ |\alpha_{(j)}| = \kappa_j, \quad j = 1, \ldots, m\right\}.$$

We have that

$$\mathcal{A}_\lambda^2(\mathbb{B}^n) = \bigoplus_{|\kappa|=0}^{\infty} H_\kappa,$$

and that the orthogonal projections P_κ of $\mathcal{A}_\lambda^2(\mathbb{B}^n)$ onto H_κ belong [3, Corollary 3.3] to the algebra $\mathcal{T}_\lambda(k\text{-}qr)$.

Observe that each space H_κ is invariant for all operators from $\mathcal{T}_\lambda(\mathbf{d})$, but the action of the generators of $\mathcal{T}_\lambda(\mathbf{d})$ on the space H_κ is quite different. The Toeplitz operator T_a with k-quasi-radial symbol a acts on H_κ as the multiplication operator $\gamma_{a,k,\lambda}(\kappa_1, \ldots, \kappa_m)I$, while each Toeplitz operator $T_{b_j t_{(j)}^{p_{(j)}}}$, $j = 1, \ldots, m$, acts on H_κ as a certain weighted shift operator. Moreover each operator $T_{b_j t_{(j)}^{p_{(j)}}}$, restricted to H_κ, is nilpotent

$$\left(T_{b_j t_{(j)}^{p_{(j)}}}\Big|_{H_\kappa}\right)^{\left[\frac{2\kappa_j}{\|p_{(j)}\|}\right]+1} = 0, \tag{5}$$

where $[x]$ denotes the integer part of $x > 0$.

We remark next that the algebra $\mathcal{T}_\lambda(\mathbf{d})$ is not semi-simple, it has a sufficiently large radical. The next lemma describes some of its elements.

Lemma 4.1 *The radical* $\mathrm{Rad}\,\mathcal{T}_\lambda(\mathbf{d})$ *of the algebra* $\mathcal{T}_\lambda(\mathbf{d})$ *contains the following operators*

- $D_\gamma T_{b_j t_{(j)}^{p(j)}}$, *where the eigenvalue sequence* $\gamma = \{\gamma(\kappa)\}_{\kappa\in\mathbb{Z}_+^m}$ *of the diagonal operator* $D_\gamma \in \mathcal{T}_\lambda(k\text{-}qr)$ *depends only on the component* κ_j *of* κ, *i.e.,* $\gamma(\kappa) = \gamma(\kappa_j)$ *and* $\{\gamma(\kappa_j)\}_{\kappa\in\mathbb{Z}_+} \in c_0$*;*
- $D_\gamma \prod_{j=1}^m \left(T_{b_j t_{(j)}^{p(j)}}\right)^{h_j}$, *where* $h_j \in \mathbb{Z}_+$ *and the eigenvalue sequence* $\gamma = \{\gamma(\kappa)\}_{\kappa\in\mathbb{Z}_+^m}$ *of the diagonal operator* $D_\gamma \in \mathcal{T}_\lambda(k\text{-}qr)$ *is such that* $\gamma(\kappa) \to 0$ *when* $|\kappa| \to \infty$ *under the condition that* $\kappa_{j_0} \to \infty$ *for at list one* $j_0 \in \{1, 2, \dots, m\}$ *with* $h_{j_0} \geq 1$*;*
- $D_\gamma \prod_{j=1}^m \left(T_{b_j t_{(j)}^{p(j)}}\right)^{h_j}$, $h_j \in \mathbb{Z}_+$ *with at list one* $j_0 \in \{1, 2, \dots, m\}$ *with* $h_{j_0} > 0$ *and* $\gamma \in c_0$.

Proof For a special case of quasi-homogeneous symbols of [9] the result has been proved in [4, Lemma 6.7], and that proof almost literally extends to the above general case of pseudo-homogeneous symbols. □

We note that the linear span of the operators considered in the last two items of Lemma 4.1, i.e., the set of all operators of the form

$$\sum_{\ell=1}^{g} D_{\gamma_\ell} \prod_{j=1}^{m} \left(T_{b_{j,\ell} t_{(j)}^{p(j)}}\right)^{h_{j,\ell}} \tag{6}$$

where $D_{\gamma_\ell} \in \mathcal{T}_\lambda(k\text{-}qr)$ and $h_{j,\ell} \in \mathbb{Z}_+$, for $\ell = 1, \dots, g$ and $j = 1, \dots, m$, forms a dense subalgebra $\mathcal{D}_\lambda(\mathbf{d})$ in the algebra $\mathcal{T}_\lambda(\mathbf{d})$. At the same time, as in [4], the representation of operators from $\mathcal{D}_\lambda(\mathbf{d})$ in the form (6) is not unique. The next lemma explains the source of such an ambiguity.

Lemma 4.2 *Let*

$$A = \sum_{\ell=1}^{g} D_{\gamma_\ell} \prod_{j=1}^{m} \left(T_{b_{j,\ell} t_{(j)}^{p(j)}}\right)^{h_{j,\ell}},$$

where all tuples $h_\ell = (h_{1,\ell}, \dots, h_{m,\ell})$ *are different, then the following statements are equivalent.*

(i) $A = 0$.

(ii) *For each* $\ell = 1, \dots, g$ *we have that* $D_{\gamma_\ell} \prod_{j=1}^m \left(T_{b_{j,\ell} t_{(j)}^{p(j)}}\right)^{h_{j,\ell}} = 0$.

(iii) *For each* $\ell = 1, \dots, g$ *we have that* $\gamma_\ell(|\alpha_{(1)}|, \dots, |\alpha_{(m)}|) = 0$ *if* $|\alpha_{(j)}| \geq \frac{h_{j,\ell}\|p_{(j)}\|}{2}$ *for all* $j = 1, \dots, m$.

Proof Follows the same arguments as the proof of Lemma 2.8 in [4].
The part (ii) ⇒ (i) is trivial.
The part (iii) ⇒ (ii) follows from (5) and the conditions posed in (iii).
To prove (i) ⇒ (iii), for a fixed $\alpha \in \mathbb{Z}_+^n$, we introduce the set of polynomials

$$F_\alpha = \left\{ f_{\ell,\alpha} = \prod_{j=1}^{m} \left(T_{b_{j,\ell}\, t_{(j)}^{p_{(j)}}} \right)^{h_{j,\ell}} e_\alpha \; : \; \ell = 1, .., g, \quad f_{\ell,\alpha} \neq 0 \right\}.$$

This set is either empty or consists of orthogonal elements of the space $\mathcal{A}_\lambda^2(\mathbb{B}^n)$. Observe now that among multi-indices $\alpha \in \mathbb{Z}_+^n$, with $|\alpha_{(j)}| \geq \frac{h_{j,\ell}\|p_{(j)}\|}{2}$ for all $j = 1, \ldots, m$, there exists such α that the set F_α is not empty. Then the orthogonality of elements of F_α for that α together with $Ae_\alpha = 0$ imply

$$D_{\gamma_\ell} f_{\ell,\alpha} = \gamma(|\alpha_{(1)}|, \ldots, |\alpha_{(m)}|) f_{\ell,\alpha} = 0, \qquad \text{for all} \quad \ell = 1, \ldots, g,$$

or $\gamma(|\alpha_{(1)}|, \ldots, |\alpha_{(m)}|) = 0$ for all $\ell = 1, \ldots, g$. □

Yet Another Option

We start again with a fixed tuple on natural numbers $k = (k_1, \ldots, k_m)$ with $k_1 + \ldots + k_m = n$, and again rearrange the n coordinates of $z \in \mathbb{B}^n$ in m groups, each one of which has $k_j, j = 1, \ldots, m$, entries. We keep using the notation

$$z_{(1)} = (z_{1,1}, \ldots, z_{1,k_1}),\; z_{(2)} = (z_{2,1}, \ldots, z_{2,k_2}),\; \ldots,\; z_{(m)} = (z_{m,1}, \ldots, z_{m,k_m}),$$

so that we have two different representations of $z \in \mathbb{B}^n$:

$$z = (z_1, \ldots, z_n) \qquad \text{and} \qquad z = (z_{(1)}, \ldots, z_{(m)}).$$

In general, having any n-tuple $x = (x_1, \ldots, x_n)$ and the above fixed tuple k, we write $x = (x_{(1)}, \ldots, x_{(m)})$.

Introduce now the *indicator* sets

$$\chi_j = \{\ell \in \{1, 2, \ldots, n\} \; : \; z_\ell \in z_{(j)}\}, \qquad j = 1, \ldots, m,$$

which specify the places of elements of $z_{(j)}$ in the n-tuple $z = (z_1, \ldots, z_n)$.

As in the previous section we consider a tuple $p = (p_1, \ldots, p_n) \in \mathbb{Z}^n$ such that each its portion $p_{(j)}, j = 1, \ldots, m$ satisfies the condition $|p_{(j)}| = 0$. We introduce as well n-tuples $\widetilde{p}_{(j)}, j = 1, \ldots, m$, by

$$\widetilde{p}_{(j),\ell} = \begin{cases} p_\ell, & \text{if} \quad \ell \in \chi_j \\ 0, & \text{otherwise} \end{cases}.$$

We represent now each coordinate of $z \in \mathbb{B}^n$ in the form

$$z_\ell = r_\ell t_\ell, \qquad \text{where} \quad r_\ell = |z_\ell| \quad \text{and} \quad t_\ell \in \mathbb{T};$$

then we represent each r_ℓ as $r_\ell = rs_\ell$, where $r = \sqrt{r_1^2 + \ldots + r_n^2}$ and $s = (s_1, \ldots, s_n) \in S_+^{n-1}$.

Recall that n-tuples $t = (t_1, \ldots, t_n) \in \mathbb{T}^n$ and $s = (s_1, \ldots, s_n) \in S_+^{n-1}$ admit the alternative representations

$$t = (t_{(1)}, \ldots, t_{(m)}) \qquad \text{and} \qquad s = (s_{(1)}, \ldots, s_{(m)}).$$

Remark 5.1 The principal difference of this section compared with the previous one is that, contrary to the previous case of $s = (s_{(1)}, \ldots, s_{(m)}) \in S_+^{k_1-1} \times \ldots \times S_+^{k_m-1}$, we have now "just one big sphere" $s = (s_1, \ldots, s_n) \in S_+^{n-1}$.

Introduce the symbol $\phi_j = b_j(s_{(j)})\, t_{(j)}^{p_{(j)}}$, where $b_j = b_j(s_{(j)})$ is bounded and measurable, and the portion $p_{(j)}$ of a tuple $p = (p_1, \ldots, p_n) \in \mathbb{Z}_+^n$ satisfies the condition $|p_{(j)}| = 0$. Given two multi-indices $\alpha,\ \beta \in \mathbb{Z}_+^n$, we calculate

$$\begin{aligned}
&\langle T_{\psi_j} z^\alpha, z^\beta \rangle = \langle \psi_j\, z^\alpha, z^\beta \rangle \\
&= \frac{\Gamma(n+\lambda+1)}{\pi^n \Gamma(\lambda+1)} \int_{\mathbb{B}^n} b_j \prod_{\ell=1}^n |z_\ell|^{\alpha_\ell+\beta_\ell} (1-|r|^2)^\lambda\, t^{\alpha-\beta+\widetilde{p}_{(j)}}\, dV(z) \\
&= \frac{\Gamma(n+\lambda+1)}{\pi^n \Gamma(\lambda+1)} \int_{\tau(\mathbb{B}^n)} b_j \prod_{\ell=1}^n r_\ell^{\alpha_\ell+\beta_\ell+1} (1-|r|^2)^\lambda\, dr_1 \cdots dr_n \prod_{\ell=1}^n \int_{\mathbb{T}} t_\ell^{\alpha_l-\beta_\ell+\widetilde{p}_{(j),\ell}} \frac{dt_\ell}{it_\ell}.
\end{aligned}$$

The last n integrals are different from zero if and only if $\beta = \alpha + \widetilde{p}_{(j)}$, in the last case each of them is equal to 2π. Thus, setting $\beta = \alpha + \widetilde{p}_{(j)}$, we have

$$\begin{aligned}
&\langle T_{\psi_j} z^\alpha, z^{\alpha+\widetilde{p}_{(j)}} \rangle = \langle \psi_j\, z^\alpha, z^{\alpha+\widetilde{p}_{(j)}} \rangle \\
&= \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_{\tau\mathbb{B}^n} b_j \prod_{\ell=1}^n r_\ell^{2\alpha_\ell+\widetilde{p}_{(j),\ell}+1} (1-|r|^2)^\lambda\, dr_1 \cdots dr_n \\
&= 2^n \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_0^1 r^{2|\alpha|+2n-1} (1-|r|^2)^\lambda\, dr \int_{S_+^{n-1}} b_j \prod_{\ell=1}^n s_\ell^{2\alpha_\ell+\widetilde{p}_{(j),\ell}+1} dS \\
&= 2^{n-1} \frac{\Gamma(n+\lambda+1)}{\Gamma(\lambda+1)} \int_0^1 r^{|\alpha|+n-1} (1-r)^\lambda\, dr \int_{S_+^{n-1}} b_j \prod_{\ell=1}^n s_\ell^{2\alpha_\ell+\widetilde{p}_{(j),\ell}+1} dS \\
&= 2^{n-1} \frac{\Gamma(n+\lambda+1)\Gamma(|\alpha|+n)}{\Gamma(n+|\alpha|+\lambda+1)} \int_{S_+^{n-1}} b_j \prod_{\ell=1}^n s_\ell^{2\alpha_\ell+\widetilde{p}_{(j),\ell}+1} dS.
\end{aligned}$$

We parametrize then S_+^{n-1} by $n-1$ elements of $s = (s_1, \ldots, s_n)$ selecting them so that the remaining one *does not belong* to $s_{(j)}$. To simplify the notation in further calculations we assume that $s_n \notin s_{(j)}$ ($\widetilde{p}_{(j),n} = 0$ in this case), and we parametrize thus S_+^{n-1} by $s_1, \ldots, s_{n-1}$. That is,

$$\begin{aligned}\langle T_{\psi_j} z^\alpha, z^{\alpha+\widetilde{p}_{(j)}}\rangle &= 2^{n-1}\frac{\Gamma(n+\lambda+1)\Gamma(|\alpha|+n)}{\Gamma(n+|\alpha|+\lambda+1)}\int_{\mathbf{B}_+^{n-1}} b_j(s_{(j)})\prod_{\ell=1}^{n-1} s_\ell^{2\alpha_\ell+\widetilde{p}_{(j),\ell}+1}\\ &\quad\times \left(1-(s_1^2+\ldots+s_{n-1}^2)\right)^{2\alpha_n} ds_1\cdots ds_{n-1}\\ &= \frac{\Gamma(n+\lambda+1)\Gamma(|\alpha|+n)}{\Gamma(n+|\alpha|+\lambda+1)}\int_{\Delta_{n-1}} b_j(s_{(j)}^{1/2})\prod_{\ell=1}^{n-1} s_\ell^{\alpha_\ell+\frac{1}{2}\widetilde{p}_{(j),\ell}}\\ &\quad\times (1-(s_1+\ldots+s_{n-1}))^{\alpha_n} ds_1\cdots ds_{n-1}\end{aligned}$$

Changing the variables: $s_\ell = u_\ell$, for $\ell \in \chi_j$, and $s_q = \left(1-(u_{(j),1}+\ldots+u_{(j),k_j})\right) u_q$, for $q \notin \chi_j \cup \{n\}$, we have

$$\begin{aligned}\langle T_{\psi_j} z^\alpha, z^{\alpha+\widetilde{p}_{(j)}}\rangle &= \frac{\Gamma(n+\lambda+1)\Gamma(|\alpha|+n)}{\Gamma(n+|\alpha|+\lambda+1)}\int_{\Delta_{k_j}} b_j(u_{(j)}^{1/2})\prod_{\ell\in\chi_j} u_\ell^{\alpha_\ell+\frac{1}{2}p_\ell}\\ &\times \left(1-(u_{(j),1}+\ldots+u_{(j),k_j})\right)^{|\alpha|-|\alpha_{(j)}|+n-k_j-1}\prod_{\ell\in\chi_j} du_\ell\\ &\times \int_{\Delta_{n-k_j-1}}\prod_{q\notin\chi_j\cup\{n\}} u_q^{\alpha_q}\left(1-\sum_{q\notin\chi_j\cup\{n\}} u_q\right)^{\alpha_n}\prod_{q\notin\chi_j\cup\{n\}} du_q\\ &= \frac{\Gamma(n+\lambda+1)\Gamma(|\alpha|+n)}{\Gamma(n+|\alpha|+\lambda+1)}\,\frac{\prod_{q\notin\chi_j}\Gamma(\alpha_q+1)}{\Gamma(|\alpha|-|\alpha_{(j)}|+n-k_j)}\\ &\times\int_{\Delta_{k_j}} b_j(u_{(j)}^{1/2})\prod_{\ell\in\chi_j} u_\ell^{\alpha_\ell+\frac{1}{2}p_\ell}\left(1-(u_{(j),1}+\ldots+u_{(j),k_j})\right)^{|\alpha|-|\alpha_{(j)}|+n-k_j-1}\prod_{\ell\in\chi_j} du_\ell.\end{aligned}$$

Taking into account that

$$\langle z^{\alpha+\widetilde{p}_{(j)}}, z^{\alpha+\widetilde{p}_{(j)}}\rangle = \frac{\Gamma(n+\lambda+1)\prod_{\ell\in\chi_j}\Gamma(\alpha_\ell+p_\ell+1)\prod_{q\notin\chi_j}\Gamma(\alpha_q+1)}{\Gamma(n+|\alpha|+\lambda+1)}$$

we come to he following lemma (where we change u_ℓ for s_ℓ).

Lemma 5.2 *Let $b_j = b_j(s_{(j)}) \in L_\infty(\mathbf{B}_+^{k_j})$ and let the portion $p_{(j)}$ of a tuple $p = (p_1, \ldots, p_n) \in \mathbb{Z}_+^n$ satisfies the condition $|p_{(j)}| = 0$. Then the Toeplitz operator $T_{b_j t_{(j)}^{p_{(j)}}}$ acts on monomials z^α as follows*

$$T_{b_j t_{(j)}^{p(j)}} z^\alpha = \begin{cases} 0, & \text{if } \exists\, \ell \in \chi_j \text{ such that } \alpha_\ell + p_\ell < 0 \\ \widetilde{\gamma}_{b_j t_{(j)}^{p(j)},k,p,\lambda}(\alpha)\, z^{\alpha + \widetilde{p}_{(j)}}, & \text{if } \forall\, \ell \in \chi_j \;\; \alpha_\ell + p_\ell \geq 0 \end{cases},$$

where

$$\widetilde{\gamma}_{b_j t_{(j)}^{p(j)},k,p,\lambda}(\alpha) = \frac{\Gamma(|\alpha| + n)}{\prod_{\ell \in \chi_j} \Gamma(\alpha_\ell + p_\ell + 1)\Gamma(|\alpha| - |\alpha_{(j)}| + n - k_j)}$$

$$\times \int_{\Delta_{k_j}} b_j(s_{(j)}^{1/2}) \prod_{\ell \in \chi_j} s_\ell^{\alpha_\ell + \frac{1}{2}p_\ell} \left(1 - (s_{(j),1} + \ldots + s_{(j),k_j})\right)^{|\alpha| - |\alpha_{(j)}| + n - k_j - 1} \prod_{\ell \in \chi_j} ds_\ell$$

and $b(s_{(j)}^{1/2}) = b(\sqrt{s_{j,1}}, \ldots, \sqrt{s_{j,k_j}})$.

Corollary 5.3 *The action of the operator* $T_{b_j t_{(j)}^{p(j)}}$ *does not depend on the weight parameter* λ.

Corollary 5.4 *Let* a *be a bounded measurable* k*-quasi-radial function, let* $b_j = b_j(s_{(j)}) \in L_\infty(\mathbf{B}_+^{k_j})$ *for each* $j = 1, \ldots, m$*, and let the tuple* $p = (p_1, \ldots, p_n) \in \mathbb{Z}_+^n$ *satisfy the condition* $|p_{(j)}| = 0$ *for each* $j = 1, \ldots, m$*. Then the Toeplitz operators* T_a*,* $T_{b_j t_{(j)}^{p(j)}}$*, for* $j = 1, \ldots, m$*, pairwise commute.*

Remark 5.5 It is straightforward to check that, contrary to the case of Corollary 3.4, we have that neither $T_a T_{b_j t_{(j)}^{p(j)}} = T_{a b_j t_{(j)}^{p(j)}}$, nor $T_{b_j t_{(j)}^{p(j)}} T_{b_\ell T_{(\ell)}^{p(\ell)}} = T_{b_j t_{(j)}^{p(j)} b_\ell T_{(\ell)}^{p(\ell)}}$, for all $j \neq \ell$ both from $1, \ldots, m$.

Now we can introduce new commutative Banach algebras which, in the setting of this section, are defined by the following data. We start with a tuple $k = (k_1, \ldots, k_m)$ of positive integers with $k_1 + \ldots + k_m = n$. Consider the set $L_\infty(\tau(\mathbb{B}^m))$ of all k-quasi-radial symbols a, then fix a tuple $b = (b_1, \ldots, b_m)$ of functions $b_j = b_j(s_{(j)}) \in L_\infty(\mathbf{B}_+^{k_j})$, $j = 1, \ldots, m$, and a tuple $p = (p_1, \ldots, p_n) \in \mathbb{Z}^n$ with $|p(j)| = 0$ for all $j = 1, \ldots, m$. We group all the above ingredients into a single set $\mathfrak{d} = \{k, L_\infty(\tau(\mathbb{B}^m)), b, p\}$ and denote by $\mathcal{T}_\lambda(\mathfrak{d})$ the unital Banach algebra generated by all Toeplitz operators

$$T_a, \quad \text{with} \quad a \in L_\infty(\tau(\mathbb{B}^m)), \quad \text{and} \quad T_{b_j t_{(j)}^{p(j)}}, \quad \text{for all} \quad j = 1, \ldots, m,$$

acting on the weighted Bergman space $\mathcal{A}_\lambda^2(\mathbb{B}^n)$. By Corollary 5.4 each algebra of this type is commutative.

Not entering into details, we mention just that the algebras $\mathcal{T}_\lambda(\mathfrak{d})$ possess the same properties as the algebras $\mathcal{T}_\lambda(\mathbf{d})$ of section “Commutative Algebras”. All of them have the same invariant subspaces H_κ, with $\kappa \in \mathbb{Z}_+$, they are not semi-simple, and have the same type on an ambiguity of representations of elements from corresponding dense subalgebras.

Remark 5.6 Two different procedures of the construction of commutative Banach algebras $\mathcal{T}_\lambda(\mathbf{d})$ and $\mathcal{T}_\lambda(\mathfrak{d})$ can be even combined into a mixed single one. That is, given a tuple $k = (k_1, \ldots, k_m)$ with $k_1 + \ldots + k_m = n$, we start as in section "Quasi-radial and Pseudo-homogeneous Symbols" by representing $z \in \mathbb{B}^n$ in the form $z = (z_{(1)}, \ldots, z_{(m)})$, and then we represent each coordinate $z_{j,\ell}$ of $z_{(j)}$ in the form $z_{j,\ell} = r_j s_{j,\ell} t_{j,\ell}$, where

$$r_j = \sqrt{|z_{j,1}|^2 + \ldots + |z_{j,k_j}|^2}, \qquad t_{j,\ell} \in \mathbb{T}, \qquad s_{(j)} = (s_{j,1}, \ldots, s_{j,k_j}) \in S_+^{k_j-1}.$$

After that we proceed with the recipe of section "Yet Another Option" on each "small sphere" $S_+^{k_j-1}$, $j = 1, \ldots, m$ (see Remark 5.1), i.e., we separate each portion $p_{(j)}$ of a $p = (p_1, \ldots, p_n) \in \mathbb{Z}^n$ onto smaller sub-portions $p_{(j,h)}$ with $|p_{(j,h)}| = 0$ and use, as generators, the Toeplitz operators with symbols $\phi_{j,h} = b_{j,h}(s_{(j,h)})\, t_{(j,h)}^{p_{(j,h)}}$ (and the Toeplitz operators with corresponding quasi-radial symbols).

References

1. W. Bauer, N. Vasilevki, Commutative Toeplitz Banach algebras on the ball and quasi-nilpotent group action. Integr. Equ. Oper. Theory **72**(2), 223–240 (2012)
2. W. Bauer, N. Vasilevki, Banach algebras of commuting Toeplitz operators on the unit ball via the quasi-hyperbolic group. Oper. Theory Adv. Appl. **218**, 155–175 (2012)
3. W. Bauer, N. Vasilevki, On the structure of commutative Banach algebras generated by Toeplitz operators on the unit ball. Quasi-elliptic case. I: generating subalgebras. J. Funct. Anal. **265**, 2956–2990 (2013)
4. W. Bauer, N. Vasilevki, On the structure of commutative Banach algebras generated by Toeplitz operators on the unit ball. Quasi-elliptic case. II: Gelfand theory. Compl. Anal. Oper. Theory **9**, 593–630 (2015)
5. A. Garcia, N. Vasilevki, Toeplitz operators on the weighted Bergman space over the two-dimensional unit ball. J. Funct. Spaces **2015**, 1–10 (2015). doi: 10.1155/2015/306168. Article ID 306168
6. R. Quiroga-Barranco, N. Vasilevski, Commutative C^*-algebras of Toeplitz operators on the unit ball, I. Bargmann-type transforms and spectral representations of Toeplitz operators. Integr. Equ. Oper. Theory **59**(3), 379–419 (2007)
7. R. Quiroga-Barranco, N. Vasilevski, Commutative C^*-algebras of Toeplitz operators on the unit ball, II. Geometry of the level sets of symbols. Integr. Equ. Oper. Theory **59**(1), 89–132 (2008)
8. N. Vasilevski, Parabolic quasi-radial quasi-homogeneous symbols and commutative algebras of toeplitz operators. Oper. Theory Adv. Appl. **202**, 553–568 (2010)
9. N. Vasilevski, Quasi-radial quasi-homogeneous symbols and commutative Banach algebras of Toeplitz operators. Integr. Equ. Oper. Theory **66**(1), 141–152 (2010)
10. K. Zhu, *Spaces of Holomorphic Functions in the Unit Ball* (Springer, New York, 2005)

A Bump Theorem for Weighted Embeddings and Maximal Operator: The Bellman Function Approach

Alexander Volberg

To Cora Sadosky with Gratitude for Her Friendship and Encouragements

Abstract We give here an "automatic" proof of a weighted embedding theorem with a bumping of the weight. It implies a well-known weighted theorem of C. Pérez.

2010 *Mathematics Subject Classification.* 42B20, 42B35, 47A30.

Introduction

Preliminaries

In this note we give a simple Bellman function proof of Carlos Pérez's "bump theorem" for the two weight estimates of maximal operators.

The original question about two weight estimates for the singular integral operators is to find a necessary and sufficient condition on the weights v and u such that a Calderón–Zygmund operator $T : L^p(v) \to L^p(u)$, $1 < p < \infty$, is bounded, i.e. the inequality

$$\int |Tf|^p u dx \leq C \int |f|^p v dx \qquad \forall f \in L^p(u) \tag{1}$$

holds. By weight we understand here non-negative locally integrable function.

AV is partially supported by the NSF grant DMS-1265549 and by the Hausdorff Institute for Mathematics, Bonn, Germany

A. Volberg (✉)
Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA
e-mail: volberg@math.msu.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_16

There is a famous solution of this problem when T is the Hilbert transform that belongs to Misha Cotlar and Cora Sadosky [1]. This characterization is in the spirit of Helson–Szegö characterization of the *one weight* boundedness of the Hilbert transform. Helson–Szegö characterization can be expressed e.g. as the existence of an H^1-function (here H^1 is a Hardy class of analytic functions in the disc) that "closely follows" the weight $u = v$. When two weights are present Cotlar–Sadosky [1] found the condition equivalent to (1) (with T being the Hilbert transform) in terms of the existence of an H^1-function that "closely follows" the right combination of two weights. Parallel to Helson–Szegö's characterization, the equivalent condition of Hunt–Muckenhoupt–Wheeden was quite important for the one weight theory. And then Sawyer's characterization for the two weight problem for maximal operator appeared. All this asked for the parallel Muckenhoupt–Hunt–Wheeden or Sawyer's language for two weight problems for general Calderón–Zygmund operators. However, it is worthwhile to mention that even in the *one weight situation* one has two different characterization of the boundedness of the Hilbert transform: (1) the Helson–Szegö's characterization and (2) the Hunt–Muckenhoupt–Wheeden characterization. They are equivalent of course, but the direct analytic proof of their equivalence is still unknown. In the two weight situation we consider here this is even more so that one can express the answer in several different languages. Cotlar–Sadosky characterization is the extension of Helson–Szegö's one.

The problem of finding the two weight characterization for the boundedness of the Calderón–Zygmund operators T in Sawyer's language has been recently solved for $p = 2$ when T is the Hilbert transform by M. Lacey, thus culminating a long search for the two weight $T1$ theorem by the group that included I. Uriarto-Tuero, E. Sawyer, C.-Y. Shen and earlier efforts of F. Nazarov, S. Treil, A. Volberg. For short range dyadic singular operators it has been solved by F. Nazarov, S. Treil, A. Volberg in 2008, [12], see also [14]. Of course the story of two weight estimates for *positive* operators can be traced to works of E. Sawyer [19, 20], the latter paper is devoted to the characterization of the two weight boundedness of the maximal operator.

We are working here with the maximal operator again. But we are interested in simple and sharp sufficient conditions rather than in necessary and sufficient ones. As Pérez has shown this can be done by the use of Hunt–Muckenhoupt–Wheeden language by modifying it with the introduction of Orlicz type norms. We wish to mention here that the Pérez "bumping" by Orlicz norms [15–17] has been replaced recently by the entropy norms "bumping", see [21]. We describe below the Orlicz bumping that we will be using in this paper.

For the interesting operators acting on functions defined on $\mathbb{R}^n$ the following two weight analogue of the A_p condition is necessary for the boundedness of the operator $M_{u^{1/p}} T M_{\sigma^{1/p'}}$. Below Q denotes any cube in $\mathbb{R}^n$.

$$\sup_Q \left(|Q|^{-1} \int_Q u dx \right) \left(|Q|^{-1} \int_Q \sigma \, dx \right)^{p/p'} < \infty \tag{2}$$

or in the symmetric form

$$\sup_Q \left(|Q|^{-1}\int_Q u dx\right)^{1/p}\left(|Q|^{-1}\int_Q \sigma dx\right)^{1/p'} < \infty \tag{3}$$

Simple counterexamples show that this condition is not sufficient for the boundedness. This holds even for the simplest "singular" operator: the Hardy–Littlewood maximal operator, and even for its dyadic counterpart. So a natural way to get a sufficient condition is to replace the L^1 norms of u and σ in (3) (or the L^p and $L^{p'}$ norms of $u^{1/p}$ and $\sigma^{1/p'}$) by some stronger Orlicz norms (this is called "bumping" the L^p norms).

This ideology of bumping can be traced to the work of C. Fefferman [9]. It has been widely used by D. Cruz-Uribe, C. Pérez, J. M. Martell, [2–8, 15–18].

Namely, given a Young function Φ (convex increasing function) and a cube Q one can consider the normalized on Q Orlicz space $L^\Phi(Q)$ with the norm given by

$$\|f\|_{L^\Phi(Q)} := \inf\left\{\lambda > 0 : \int_Q \Phi\left(\frac{f(x)}{\lambda}\right)\frac{dx}{|Q|} \le 1\right\}.$$

And it was conjectured (for $p = 2$) that if the Young functions Φ_1 and Φ_2 satisfy the condition

$$\int^\infty \frac{dx}{\Phi(x)} < \infty, \tag{4}$$

then the condition

$$\sup_Q \|u\|_{L^{\Phi_1}(Q)}\|\sigma\|_{L^{\Phi_2}(Q)} < \infty \tag{5}$$

implies that for any bounded Calderón–Zygmund operator T the operator $M_{u^{1/2}}TM_{\sigma^{1/2}}$ is bounded in L^2.

This conjecture (belonging to C. Pérez and D. Cruz-Uribe) was, in fact, proved by different methods in the paper of Lerner [10, 11] and in the paper [13]. The first paper has the advantage of giving the result for all p's. The second paper deals only with $p = 2$ (which is still quite an interesting case) and it demonstrates an "automatic" proof of the bump conjecture. Namely, a function is constructed (on a certain infinite dimensional space) such, that after feeding into this function the distribution functions of the weights and several other data, we obtain the desired estimate just by applying Green's formula to this function (we call it a Bellman function of the problem).

But the paper [13] did not deal with the simplest "singular" operator: the maximal operator. This was just because for maximal operator the right bump conjecture has been already proved by Carlos Pérez, see [15–17]. But it seems to us that to give an "automatic" proof of Pérez's result can be interesting too. Moreover, the main Theorem 2.1 formally proves a seemingly slightly more general result, which implies the maximal estimate with a bumping proved by Pérez.

So in the present note we are concerned with the simplest case, namely we deal with a new "automatic" proof of the sharp bump result proved by Pérez [15–17] for maximal operators (and a slight generalization of this result). The bump condition here is slightly different, it is a *one-sided bump condition*.

Our result differs in a way from Pérez' theorem in that it has stronger regularity assumptions on the bump. We assume that Φ is convex, and Pérez assumed that $\Phi(t^2)$ is convex. The comparison between (4) and Pérez assumption $\int \frac{t}{B(t)} dt < \infty$ plus a simple change of variables show that $\Phi(s) = B(\sqrt{s})$. This relationship between our "bump" function Φ (applied directly to our weight) and Pérez' "bump" function B (applied to the square of the weight) shows that our assumption is stronger, but one should remember that the main assumption is (4), and thus the interesting case is $\Phi(t) = td(t)$, where d is a sort of logarithmic correction. If d is regular enough there is no difference between these two convexities. This is the case, for example, if $s\frac{d'(s)}{d(s)}$ tends monotonically to zero when s tends to infinity. There is one more regularity assumption that we choose to impose on Φ. It is formulated in (28) in terms of Ψ that will be defined by Φ. It is again satisfied for all sufficiently regular d.

The proof has some value because of the above mentioned slight generalization (Theorem 2.1) and because it illustrates how one can give an "automatic" proof by presenting a formula for a certain function, which, in its turn, "automatically stops the time" (so no stopping time argument should be invented). On the other hand, the accent in difficulty is moved now to building such a function. And the main thrust is to ensure its specific concavity properties.

Sharp Bump Conditions for the Maximal Operators and Weighted Embedding

Let us consider the maximal operator acting on function on $\mathbb{R}^n, n = 1$. This is just for the sake of simplicity (the argument can be carried on for all n). In what follows J denotes an interval of the real line $\mathbb{R}$. We prove here Pérez' theorem [15–17]:

$$\int \mathcal{M}(\varphi\sigma)^2 u\, dt \le C \int \varphi^2 \sigma\, dt, \tag{6}$$

under the condition

$$\sup_J \langle u \rangle_J \cdot \|\sigma\|_{L^\Phi(J)} < C\,, \tag{7}$$

which ports the name *one-sided bump condition*.

Notice that we made here a change of variable in comparison with (1) and (5). In fact, for $p = 2$ inequality (1) can be rewritten as the boundedness of the operator $u^{1/2}Tv^{-1/2}$ in unweighted L^2. So when one writes that one is interested in two weight boundedness of $u^{1/2}T\sigma^{1/2}$, one should identify v^{-1} in (1) with σ: this identification has been done while writing down (5).

In this section again we use a convenient change of variable, transforming the two weighted inequality in the form (1)

$$\int \mathcal{M}(f)^2 u\,dt \le C\int f^2 v\,dt,$$

into

$$\int [\mathcal{M}(fv\frac{1}{v}])^2 u\,dt \le C\int (fv)^2\frac{1}{v}\,dt\,.$$

Now we change the variable $\sigma = v^{-1}, \varphi = fv$ and we obtain the form (6) of two weight inequality with which we will be working now.

The reader should keep this change of variables in mind because the two weight problems used to be formulated in different forms: in the form (1), or in the form of the boundedness of $u^{1/2}T\sigma^{1/2}$ in unweighted L^2, or, at last, in the form of (6), where the integration is with respect to *the same measure* σdx in the right hand side and *inside* the operator. Notice that all forms are basically equivalent, but the latter has the advantage that measure σdx can be easily made a general measure (and not just Lebesgue measure density).

Condition (7) above is a strengthening of the classical A_2 condition. Some strengthening is unavoidable as we are dealing here with a two-weight problem. The sharp condition on Φ (called B_2 condition) will be quoted below, in essence it means $\int^\infty \frac{1}{\Phi(t)}dt < \infty$.

We prove this result of Carlos Pérez by a new method, actually by a formula. What follows is a sort of *automatic proof* of the result.

It is well known that the problem for the classical maximal operator can be reduced to dyadic maximal operator (in this setting it is easy and follows by a simple trick with averaging over random dyadic lattices).

Let us denote by $\mathcal{D}$ a standard dyadic lattice. Then it is also quite known that we can reduce our problem to the following: If $\{a_I|I|\}_{I\in\mathcal{D}}$ is a u-Carleson sequence, then

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}, I\subset J} a_I\langle\sigma\rangle_I^2|I| \le C\langle\sigma\rangle_J. \tag{8}$$

The requirement that $a_I|I|$ is a u-Carleson sequence means that

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}, I\subset J} a_I|I| \le C\langle u\rangle_J. \tag{9}$$

Since we bump v, we should care about

$$N_J(t) = \frac{1}{|J|}|\{x \in J : \sigma(x) \geq t\}|.$$

Also we denote

$$A_J = \frac{1}{|J|}\sum_{I\in\mathcal{D}, I\subset J} a_I|I|.$$

The u-Carleson property of $\{a_I \cdot |I|\}_{I\in\mathcal{D}}$ means $A_J \leq C\langle u\rangle_J$ for all dyadic J.

The combination of (7) and the u-Carleson property of $\{a_I \cdot |I|\}_{I\in\mathcal{D}}$ means that

$$\forall J \in \mathcal{D} \text{ we have } A_J \cdot \|\sigma\|_{L^\Phi(J)} \leq C < \infty\,. \tag{10}$$

We want to deduce (8) from (10) under the sharp assumptions on the gauge function Φ. These assumptions are as follows: Φ is increasing convex function satisfying (4).

So here is our theorem.

Theorem 2.1 *Let the sequence of nonnegative numbers* $\{a_I\}_{I\in\mathcal{D}}$ *satisfy* (10) *for every dyadic interval J, where* $A_J = \frac{1}{|J|}\sum_{I\in\mathcal{D},I\subset J} a_I|I|$. *Then if* Φ *is increasing convex function satisfying*

$$\int^{\infty} \frac{1}{\Phi(t)}dt < \infty$$

and satisfying a mild regularity condition then for every dyadic interval J the following inequality holds

$$\frac{1}{|J|}\sum_{I\in\mathcal{D},I\subset J} a_I\langle\sigma\rangle_I^2|I| \leq C\langle\sigma\rangle_J.$$

Notice that in the theorem there is no second weight u, there is no assumption of the type (9) whatsoever. In the standard deduction of two weight maximal theorem from Theorem 2.1 of course the second weight u will be present and the standard linearization of maximal operator will bring the sequence of nonnegative numbers $\{a_I\}_{I\in\mathcal{D}}$ that will satisfy (9). Then the application of Theorem 2.1 will prove the maximal estimate (in its form (8), but we know by the Sawyer's result [19] that this is enough).

However, it is not clear how to get Theorem 2.1 from the corresponding maximal result. It seems like for that one would need to build some special weight u, which is not present in Theorem 2.1. This is why this theorem seems to be a slight generalization of the result of Pérez.

The mild regularity condition on Φ is formulated in terms of its transform Ψ that we are introducing in the next section "Orlicz Norms and Distribution Functions". In terms of Ψ this mild regularity condition is formulated at the very end of the paper in (28). In fact we believe that for every increasing convex Φ satisfying the integrability condition in the theorem, there is a smaller function also satisfying the integrability condition and such that the new Ψ will satisfy this mild regularity (28).

Notice that the mild regularity condition (28) is satisfied for all examples of Φ presented at the end of the next section "Orlicz Norms and Distribution Functions".

We explained that this result gives the bump result of Carlos Pérez. The latter is the sharp result and the integrability of $1/\Phi$ cannot be weakened. On the other hand, notice that the statement of the theorem requires only (10) and does not require (9). So it is more the statement of "Carleson embedding theorem", than the statement about maximal operator. After all, it might be important to have simple conditions on the sequence of nonnegative numbers $\{a_I\}_{I\in\mathcal{D}}$ and a weights σ that would imply inequality (8). Notice that by Sawyer's arguments we know that (8) then implies a general embedding theorem localized to interval J:

$$\frac{1}{|J|}\sum_{I\in\mathcal{D}, I\subset J} a_I \langle\varphi\sigma\rangle_I^2 |I| \le C\langle\varphi^2\sigma\rangle_J. \tag{11}$$

Orlicz Norms and Distribution Functions

Here we repeat verbatim some results from [13]. Orlicz norm is not very convenient to work with, so we would like to replace it, and to work with something more tractable.

A Lower Bound for the Orlicz Norm

Let Φ be a continuous non-negative increasing convex function such that $\Phi(0) = 0$ and $\int^{+\infty} \frac{dt}{\Phi(t)} < +\infty$. Define $\Psi(s)$ parametrically by $\Psi(s) = \Phi'(t)$ when $s = \frac{1}{\Phi(t)\Phi'(t)}$ $(t > 0)$. Then $\Psi(s)$ is positive and decreasing for $s > 0$ and $s\Psi(s)$ is increasing. Moreover $\int_0 \frac{ds}{s\Psi(s)} < +\infty$. Indeed, using our parameterization we can rewrite the last integral as

$$\int^{+\infty} \left(\frac{1}{\Phi(t)} + \frac{\Phi''(t)}{\Phi'(t)^2}\right) dt\,.$$

The first integral converges by our assumption and the second integrand has a bounded near $+\infty$ antiderivative $\frac{-1}{\Phi'(t)}$.

Let $w \ge 0$ on $J \subset \mathbb{R}^n$. Define the normalized distribution function N of w by

$$N(t) = N_J^w(t) = \frac{1}{|J|} \left|\{x \in I : w(x) > t\}\right| \tag{12}$$

Lemma 2.2 *Let* $\Psi : (0, 1] \to \mathbb{R}_+$ *be a decreasing function such that the function* $s \mapsto s\Psi(s)$ *is increasing. Let* Φ *be a Young function and let*

$$\Psi(s) \le C\Phi'(t) \qquad \textit{where} \quad s = \frac{1}{\Phi(t)\Phi'(t)}$$

for all sufficiently large t. Then for $N = N_J^w$

$$\mathbf{n}_\Psi(N) := \int_0^\infty N(t)\Psi(N(t))\, dt \le C\|w\|_{L^\Phi(J)}\,. \tag{13}$$

Proof The left hand side scales like a norm under multiplication by constants, so it is enough to show that if $\|w\|_{L_\Phi(J)} \le 1$, i.e.,

$$\frac{1}{|J|}\int_J \Phi(w) = \int_0^\infty N(t)\Phi'(t)\, dt \le 1$$

then $\mathbf{n}_\Psi(N)$ is bounded by a constant. Since $s\Psi(s)$ increases, we may have trouble only at $+\infty$ It is cleat that it suffices to estimate the integral over the set where $\Psi(N(t)) > \Phi'(t)$ but since Ψ is decreasing this means that $N(t) \le C/(\Phi(t)\Phi'(t))$, so we get at most $\int^{+\infty} \Phi(t)^{-1} dt$ and we are done. □

Remark In the above Lemma 2.2 we do not need the assumption that

$$\int_0 \frac{1}{s\Psi(s)} ds < \infty. \tag{14}$$

But in what follows this assumption will be needed, and the reasoning in the beginning of this section shows that for any Young function Φ satisfying $\int^\infty (\Phi(t))^{-1} dt < \infty$ we can find Ψ from Lemma 2.2 satisfying (14).

Examples

In the above section only the behavior of Φ at $+\infty$ and the behavior of Ψ near 0 were important, so we will concentrate our attention there.

Let $\Phi(t) = t(\ln t)^\alpha$, $\alpha > 1$ near ∞. Then

$$\Phi'(t) \sim (\ln t)^\alpha, \qquad \Phi(t)\Phi'(t) \sim t(\ln t)^{2\alpha},$$

so $\Psi(s) := (\ln(1/s))^\alpha$ satisfies the assumptions of Lemma 2.2: to see that we notice

$$\ln(\Phi(t)\Phi'(t)) \sim \ln t.$$

If $\Phi(t) = t \ln t (\ln\ln t)^\alpha$, $\alpha > 1$, then

$$\Phi'(t) \sim \ln t (\ln\ln t)^\alpha, \qquad \Phi(t)\Phi'(t) \sim t(\ln t)^2 (\ln\ln t)^{2\alpha}$$

and $\Psi(s) = \ln(1/s)(\ln\ln(1/s))^\alpha$ works because again $\ln(\Phi(t)\Phi'(t)) \sim \ln t$.

Note that in both examples $\int_0 (s\Psi(s))^{-1} ds < \infty$.

The examples of Young functions with higher order logarithms are treated similarly.

Bellman Function of a Problem

Let us consider a function $D(A, t, N)$ of three variables, where the last variable N is in fact any decreasing function on $[0, \infty)$ taking values in $[0, 1]$) and such that

$$D(A, t, N) \le C \cdot N \tag{15}$$

$$d^2_{A,N} D \le 0 \tag{16}$$

$$\int_0^\infty \frac{\partial D}{\partial A}(A, t, N(t))dt \ge K\left(\int N(t)dt\right)^2, \tag{17}$$

where what is K will be determined later. The middle inequality means that for every t the function of A, N is concave.

We will be looking for D of the following form $D(A, t, N) = B(At, N)$ (scaling in (8) hints us to do that) , and we put for an arbitrary dyadic interval $I \in \mathcal{D}$

$$\mathcal{B}(I) = \int B(A_I \cdot t, N_I(t))dt.$$

What is a Priori Bounded?

In this section we are going to discuss the following question. If (10) is satisfied, that is (after normalization) $\|\sigma\|_{L^\Phi(I)} A_I \le 1$, then what quantities, involving Φ and Ψ from section "Orlicz Norms and Distribution Functions", are bounded?

First of all, we know that

$$\int N_I(t)\Psi(N_I(t))dt \le c\|\sigma\|_{L^\Phi(I)}$$

and so our first inequality is

$$A_I \int N_I(t)\Psi(N_I(t))dt \le c. \tag{18}$$

What else? Well, to calculate the norm of σ, we should integrate something like $N_I(t)\Phi'(t)$, but not exactly! Let us be careful, and it will be the thing that we (at least I) was missing all this time.

We know that $\|\sigma\|_{L^\Phi(I)} \le \frac{1}{A_I}$. We also know that

$$\|\sigma\|_{L^\Phi(I)} = \inf\left\{\lambda : \langle\Phi\left(\frac{\sigma}{\lambda}\right)\rangle_I \le 1\right\}.$$

Therefore (the average in the inf decreases when λ increases), we get

$$\langle\Phi\left(A_I\sigma\right)\rangle_I \le 1.$$

We now write the last inequality as follows

$$\int_0^\infty \Phi'(t)N_I\left(\frac{t}{A_I}\right)dt \le 1,$$

or

$$A_I \int_0^\infty N_I(t)\Phi'(A_I t)dt \le 1. \tag{19}$$

This is the second inequality. Notice that we have $\Phi'(A_I t)$ instead of $\Phi'(t)$, and this will be crucial.

Next Step is Green's Formula on the Tree and the Use of Concavity

So,

$$\mathcal{B}(I) = \int B(A_I \cdot t, N_I(t))dt,$$

and using that $N(t) = \frac{N_{I+}(t)+N_{I-}(t)}{2}$ we can write

$$\mathcal{B}(I) - \frac{\mathcal{B}(I_+) + \mathcal{B}(I_-)}{2} = \int \left[B(A_I \cdot t, N_I(t)) - B\left(\frac{A_{I+} + A_{I-}}{2} \cdot t, N_I(t)\right) \right] dt +$$

$$\int \left[B\left(\frac{A_{I+} + A_{I-}}{2} \cdot t, \frac{N_{I+}(t) + N_{I-}(t)}{2}\right) - \frac{1}{2}\Big(B(A_{I+} \cdot t, N_{I+}(t)) + B(A_{I-} \cdot t, N_{I-}(t))\Big)\right].$$

In the second line we use concavity of B, thus, this term is nonnegative. In the first line we use the mean value theorem and the fact that as $\frac{\partial^2 B}{\partial A^2} \leq 0$ (concavity in variable A), we are ensured that $\frac{\partial B}{\partial A}$ at the intermediate point is at least $\frac{\partial B}{\partial A}(A_I t, N_I(t))$.

Then we can continue (using that $A_I - \frac{A_{I+} + A_{I-}}{2} = a_I$)

$$\mathcal{B}(I) - \frac{\mathcal{B}(I_+) + \mathcal{B}(I_-)}{2} \geq a_I \int t \frac{\partial B}{\partial A}(A_I t, N_I(t)) dt \geq a_I \langle \sigma \rangle_I^2 \left(\int \frac{N_I(t)^2}{t \frac{\partial B}{\partial A}(A_I t, N_I(t))} dt \right)^{-1}.$$

The last inequality is just Hölder inequality applied to

$$\langle \sigma \rangle_I^2 = \left(\int_0^\infty N_I(t) dt \right)^2 = \left(\int_0^\infty \frac{N(t)}{\sqrt{t \frac{\partial B}{\partial A}(A_I t, N_I(t))}} \cdot \sqrt{t \frac{\partial B}{\partial A}(A_I t, N_I(t)) dt} \right)^2.$$

We want

$$\int \frac{N_I(t)^2}{t \frac{\partial B}{\partial A}(A_I t, N_I(t))} dt \leq C, \tag{20}$$

which will be satisfied if we take (18), (19) in combination with

$$t \frac{\partial B}{\partial A}(A_I t, N_I(t)) \geq \frac{N_I(t)}{\Psi(N_I(t)) + \Phi'(A_I \cdot t)} \cdot \frac{1}{A_I}, \tag{21}$$

After dividing by t becomes

$$\frac{\partial B}{\partial A}(A_I t, N_I(t)) \geq \frac{N_I(t)}{\Psi(N_I(t)) + \Phi'(A_I \cdot t)} \cdot \frac{1}{A_I \cdot t}.$$

So it is tempting to put

$$\frac{\partial B}{\partial A}(A, N) = \frac{N}{\Psi(N) + \Phi'(A)} \cdot \frac{1}{A}.$$

In fact this formula cannot give us the desired B. But in section "Here is B" we will modify the formula.

Here is B

Recall the notations: we start with convex Φ, whose reciprocal is integrable at ∞, and we build $\Psi(s), \phi(s) = s\Psi(s), s \approx 0$, in such a way that

$$s\Psi(s) = \frac{1}{\Phi(t)}\,, \text{ if } s = \frac{1}{\Phi'(t)\Phi(t)}\,, \tag{22}$$

which implies that

$$\Psi(s) = \Phi'(t)\,, \text{ if } s = \frac{1}{\Phi'(t)\Phi(t)}\,. \tag{23}$$

We want to combine that with

$$\Phi'(t) \le \Phi'\left(\frac{1}{s}\right), \text{ if } s = \frac{1}{\Phi'(t)\Phi(t)}\,, \tag{24}$$

Together (23) and (24) give that

$$\Psi(s) \le \Phi'\left(\frac{1}{s}\right), \text{if } s \approx 0\,. \tag{25}$$

which we will use in what follows.

But we need to check first (24), which is the same (as Φ' is increasing) as to verify that

$$t \le \frac{1}{s}\,, \text{ if } s = \frac{1}{\Phi'(t)\Phi(t)}\,. \tag{26}$$

This is immediate: $\frac{1}{s} = \Phi'(t)\Phi(t) \ge t$ just because $\Phi(t) \ge t, \Phi'(t) \ge 1$ if t is sufficiently large (this is a simple consequence of convexity of Φ and condition (4)).

Let us introduce now

$$\Psi_{new}(s) := \begin{cases} \Psi(s), s \in (0,1], \\ \Psi(\frac{1}{s}), s \ge 1\,. \end{cases}$$

Recall that we are looking for $D(A,t,N) = B(At,N)$ with properties listed at the beginning of the Section. Here is B that will give us the function D we need:

$$B(\tau, N) := N\int_0^{\frac{\tau}{N}} \frac{1}{\Psi_{new}(s)} \cdot \frac{ds}{s}\,.$$

Obviously

$$0 \le B \le CN$$

because $\frac{1}{\Psi(s)}$ is $\frac{ds}{s}$ integrable on (0, 1], and $\frac{1}{\Psi_{new}(s)} = \frac{1}{\Psi(\frac{1}{s})}$ is $\frac{ds}{s}$ integrable on $[1,\infty)$.

The determinant of Hessian matrix is zero, so concavity is equivalent to $B_{\tau\tau} \le 0$. But

$$B_{\tau\tau} = -\frac{N}{\tau^2\Psi(\frac{\tau}{N})}\left(1 + \frac{\tau}{N}\frac{\Psi'(\frac{\tau}{N})}{\Psi(\frac{\tau}{N})}\right) \le 0$$

when $\tau/N \le 1$ because $s\Psi(s)$ is increasing on [0, 1]. In the other case $N/\tau < 1$ we have $B(\tau,N) := N\int_{\frac{N}{\tau}}^{\infty} \frac{1}{\Psi(s)} \cdot \frac{ds}{s}$

$$B_{\tau\tau} = -\frac{N}{\tau^2\Psi(\frac{N}{\tau})}\left(1 - \frac{N}{\tau}\frac{\Psi'(\frac{N}{\tau})}{\Psi(\frac{N}{\tau})}\right) \le 0$$

because $\Psi(s)$ is decreasing on (0, 1].

We are left to prove that (20) holds, namely, that

$$\int_0^\infty \frac{N^2(t)}{tB'_\tau(At,N(t))}\,dt \le C_0\,. \tag{27}$$

This is the same as

$$S := A\int_0^\infty N(t)\Psi_{new}\left(\frac{At}{N(t)}\right)\,dt \le C_0\,.$$

We split integral S to three parts. Integral S_1, where $\frac{At}{N(t)} \le 1$, so $At \le N(t) \le 1$ (we use that Ψ is decreasing on (0, 1]):

$$S_1 \le A\int_0^{\frac{1}{A}} N(t)\Psi\left(\frac{At}{N(t)}\right)\,dt \le A\int_0^{\frac{1}{A}} \Psi(At)\,dt = \int_0^1 \Psi(s)ds = C_1 < \infty\,.$$

In fact, using, for example, (24) we see that $\int_0 \Psi(s)\,ds \le C\int^\infty \frac{\Phi'(t)}{t^2}dt \approx \int^\infty \frac{\Phi(t)}{t^3}\,dt$. The latter integral is finite because in all interesting cases we can assume $\Phi(t) \le t^{3/2}$. In fact, for power bump functions Φ the result we are proving can be proved in a much easier way, see for example [9], (and, in fact, it follows of course from the result for more complicated bumps).

Integral S_2 is where $\frac{At}{N(t)} > 1, At \le 1$. We use that Ψ is decreasing on (0, 1].

$$S_2 = A\int_0^\infty N(t)\Psi\left(\frac{N(t)}{At}\right)\,dt \le A\int_0^\infty N(t)\Psi(N(t))\,dt \le C_2\,,$$

which we know from section "What is a Priori Bounded?".

Finally we are left with integral S_3, where $At > 1$.

$$S_3 \le A \int_0^\infty N(t)\Psi\left(\frac{N(t)}{At}\right)\, dt\,.$$

Here we use the following property of Ψ, which is satisfied for all reasonable Ψ obtained from Φ (it is a small restriction on regularity of Φ, but notice that all interesting Ψ's are (sub)-logarithmic and so have this property):

$$\forall s_2, s_2 \in (0,1],\;\; \Psi(s_1 s_2) \le C(\Psi(s_1) + \Psi(s_2))\,. \tag{28}$$

Using (28) we continue

$$I_3 \le CA \int_0^\infty N(t)\Psi(N(t))\, dt + CA \int_0^\infty N(t)\Psi\left(\frac{1}{At}\right)\, dt\,.$$

The first integral is $\le CC_3$ by (18) of section "What is a Priori Bounded?". For the second integral we use (25) and (19):

$$A\int_0^\infty N(t)\Psi\left(\frac{1}{At}\right)\, dt \le A\int_0^\infty N(t)\Phi'(At)\, dt\,,$$

which is again bounded by C_4 by (19) of section "What is a Priori Bounded?".

We are done.

Acknowledgements The author is grateful to F. Nazarov and A. Reznikov for the important discussions and valuable advices. The author also thanks the referee for many suggestions that improved the exposition.

References

1. M. Cotlar, C. Sadosky, On the Helson–Szegö theorem and a related class of modified Toeplitz kernels. Harmonic analysis in Euclidean spaces, in *Proceedings of Symposia in Pure Mathematics*, Part 1 (Williams Coll., Williamstown, 1978), pp. 383–407; *Proceedings of Symposia in Pure Mathematics*, XXXV Part (American Mathematical Society, Providence, 1979)
2. D. Cruz-Uribe, J.M. Martell, C. Pérez, Sharp weighted estimates for classical operators. Adv. Math. **229**, 408–441 (2012)
3. D. Cruz-Uribe, J.M. Martell, C. Pérez, *Weights, Extrapolation and the Theory of Rubio de Francia*. Operator Theory: Advances and Applications, vol. 215 (Birkhauser, Basel, 2011)
4. D. Cruz-Uribe, J.M. Martell, C. Pérez, Sharp two-weight inequalities for singular integrals, with applications to the Hilbert transform and the Sarason conjecture. Adv. Math. **216**(2), 647–676 (2007)
5. D. Cruz-Uribe, C. Pérez, Sharp two-weight, weak-type norm inequalities for singular integral operators. Math. Res. Lett. **6**(3–4), 417–427 (1999)

6. D. Cruz-Uribe, C. Pérez, Two-weight, weak-type norm inequalities for fractional integrals, Calderón–Zygmund operators and commutators. Indiana Univ. Math. J. **49**(2), 697–721 (2000)
7. D. Cruz-Uribe, C. Pérez, On the two-weight problem for singular integral operators. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **1**(4), 821–849 (2002)
8. D. Cruz-Uribe, A. Reznikov, A. Volberg, Logarithmic bump conditions and the two-weight boundedness of Calderón–Zygmund operators. Adv. Math. **255**, 706–729 (2014)
9. C. Fefferman, The uncertainty principle. Bull. Am. Math. Soc. **9**, 129–206 (1983)
10. A. Lerner, A pointwise estimate for local sharp maximal function with applications to singular integrals. Bull. Lond. Math. Soc. **42**, 843–856 (2010)
11. A.K. Lerner, On an estimate of Calderón-Zygmund operators by dyadic positive operators. J. Anal. Math. **121**, 141–161 (2013)
12. F. Nazarov, S. Treil, A. Volberg, Two weight inequalities for individual Haar multipliers and other well localized operators. Math. Res. Lett. **15**(3), 583–597 (2008)
13. F. Nazarov, A. Reznikov, S Treil, A. Volberg, A Bellman function proof of the L^2 bump conjecture. J. Anal. Math. **121**, 255–277 (2013)
14. F. Nazarov, S. Treil, A. Volberg, Bellman function and two-weight inequality for martingale transform. J. Am. Math. Soc. **12**(4), 909–928 (1999)
15. C. Pérez, Two weighted inequalities for potential and fractional type maximal operators. Indiana Univ. Math. J. **43**(2), 663–684 (1994)
16. C. Pérez, Weighted norm inequalities for singular integral operators. J. Lond. Math. Soc. **49**(2), 296–308 (1994)
17. C. Pérez, On sufficient conditions for the boundedness of the Hardy-Littlewood maximal operator between weighted L^p-spaces with different weights. Proc. Lond. Math. Soc. **71**(3), 135–157 (1995)
18. C. Pérez, R. Wheeden, Uncertainty principle estimates for vector fields. J. Funct. Anal. **181**, 146–188 (2001)
19. E. Sawyer, A characterization of a two-weight norm inequality for maximal operators. Stud. Math. **75**(1), 1–11 (1982)
20. E. Sawyer, A characterization of two weight norm inequalities for fractional and Poisson integrals. Trans. Am. Math. Soc. **308**, 533–545 (1988)
21. S. Treil, A. Volberg, Entropy conditions in two weight inequalities for singular integral operators, pp. 1–37. arXiv:1408.0385 (2014)

The Necessity of A_∞ for Translation and Scale Invariant Almost-Orthogonality

Michael Wilson

Abstract If ν is a measure, we say a set $\{\psi_k\}_k \subset L^2(\nu)$ is almost-orthogonal in $L^2(\nu)$ if there is an $R < \infty$ such that, for all finite linear sums $\sum \lambda_k \psi_k$,

$$\int \left|\sum \lambda_k \psi_k\right|^2 d\nu \le R \sum |\lambda_k|^2.$$

If $z = (t, y) \in \mathbf{R}_+^{d+1} \equiv \mathbf{R}^d \times (0, \infty)$ and $f : \mathbf{R}^d \to \mathbf{C}$, define $f_z(x) \equiv f((x-t)/y)$. If $Q \subset \mathbf{R}^d$ is a cube with sidelength $\ell(Q)$, define $T(Q) \equiv Q \times [\ell(Q)/2, \ell(Q))$. We say that $\{\phi_k\}_1^n$, a finite set of bounded, complex-valued functions with supports contained in $B(0; 1)$, satisfies the *collective non-degeneracy condition* (CNDC) if there is no ray emanating from the origin on which the Fourier transform of every ϕ_k vanishes identically. We prove: If μ is a doubling measure on $\mathbf{R}^d$ with the property that, for some family $\{\phi_k\}_1^n$ satisfying CNDC, it is the case that, for every $1 \le k \le n$ and every choice of points $\zeta(Q) \in \overline{T(Q)}$, $Q \in \mathcal{D}$ (where $\mathcal{D}$ is the family of dyadic cubes), the set

$$\left\{ \frac{(\phi_k)_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right\}_{Q \in \mathcal{D}}$$

is almost-orthogonal in $L^2(\mu)$, then μ is a Muckenhoupt A_∞ measure.

AMS Subject Classification (2000): 42B25 (primary); 42C15, 42C40 (secondary).

M. Wilson (✉)
Department of Mathematics, University of Vermont, Burlington, VT 05405, USA
e-mail: jmwilson@uvm.edu

M.C. Pereyra et al. (eds.), *Harmonic Analysis, Partial Differential Equations, Banach Spaces, and Operator Theory (Volume 2)*, Association for Women in Mathematics Series 5, DOI 10.1007/978-3-319-51593-9_17

Introduction

We recall that a non-trivial Radon measure ν on $\mathbf{R}^d$ is said to be A_∞ (in symbols: $\nu \in A_\infty$) if, for every $\epsilon > 0$, there is a $\delta > 0$ such that, for every cube $Q \subset \mathbf{R}^d$ and every measurable $E \subset Q$, having $|E|/|Q| < \delta$ implies $\nu(E) \leq \epsilon\nu(Q)$; where, here and in the future, we use $|\cdot|$ to denote a set's Lebesgue measure. A non-trivial Radon measure ν on $\mathbf{R}^d$ is said to be doubling if there is a finite C so that, for all cubes $Q \subset \mathbf{R}^d$, $\nu(2Q) \leq C\nu(Q)$, where $2Q$ denotes Q's concentric double. It is easy to see that $\nu \in A_\infty$ implies that ν is doubling; it is not so easy (but classical) that the converse fails. If $\nu \in A_\infty$ then $d\nu = v\,dx$ for some non-negative $v \in L^1_{loc}(\mathbf{R}^d)$. In such a case we say that $v \in A_\infty$. It is well known that $v \in A_\infty$ if and only if there is a $p > 1$ and a finite K_p such that, for all cubes Q,

$$\left(\frac{1}{|Q|}\int_Q v^p\,dx\right)^{1/p} \leq \frac{K_p}{|Q|}\int_Q v\,dx, \tag{1}$$

which is the so-called "reverse-Hölder inequality".

In a recent paper [9] the author proved that, if $\mu \in A_\infty$, then, in a precise sense to be explained shortly, $L^2(\mu)$ and ordinary, Lebesgue-measure L^2 have the same almost-orthogonal systems; where we say that a collection of functions $\{\psi_k\}_k$ is almost-orthogonal in $L^2(\nu)$ if there is a finite R so that, for all finite linear sums $\sum \lambda_k\psi_k$,

$$\int \left|\sum \lambda_k\psi_k\right|^2 d\nu \leq R\sum |\lambda_k|^2. \tag{2}$$

He also proved that if μ is a doubling measure and L^2 and $L^2(\mu)$ have (in a precise sense) the same almost-orthogonal systems, then μ must be A_∞.

Let us explain what this "precise sense" is.

If $z = (t, y) \in \mathbf{R}^{d+1}_+ \equiv \mathbf{R}^d \times (0, \infty)$ and $f : \mathbf{R}^d \to \mathbf{C}$, we define $f_z(x)$ to be $f((x-t)/y)$. This is the function f dilated and translated relative to the ball $B(t; y)$, but without any measure-based normalization. If $0 < \alpha \leq 1$ we say that $\phi \in \mathcal{C}_\alpha$ if $\phi : \mathbf{R}^d \to \mathbf{C}$ has support contained in $B(0; 1)$ and, for all x and x' in $\mathbf{R}^d$, $|\phi(x)-\phi(x')| \leq |x-x'|^\alpha$. We write $\mathcal{C}_{\alpha,0}$ to mean the subspace of ϕ's in $\mathcal{C}_\alpha$ satisfying $\int \phi\,dx = 0$. We call a cube Q dyadic if $Q = [j_1 2^k, (j_1+1)2^k)\times\cdots\times[j_d 2^k, (j_d+1)2^k)$ for some integers $j_1, \ldots, j_d$, and k, and we write $\ell(Q)$ for Q's sidelength (which is 2^k). We call the set of all dyadic cubes $\mathcal{D}$. If $Q \in \mathcal{D}$ we put $z_Q \equiv (x_Q, \ell(Q)) \in \mathbf{R}^{d+1}_+$, where x_Q is Q's center. If $\{\phi^{(Q)}\}_{\mathcal{D}} \subset \mathcal{C}_\alpha$, then

$$\left\{\frac{\phi^{(Q)}_{z_Q}}{\sqrt{|Q|}}\right\}_{Q\in\mathcal{D}} \tag{3}$$

is a family of Hölder-smooth functions, indexed over $\mathcal{D}$, with each one dilated, translated, and (Lebesgue) measure-normalized to "fit" a dyadic cube Q. If each $\phi^{(Q)} \in \mathcal{C}_{\alpha,0}$ then it is easy to see that (3) is almost-orthogonal in L^2, with an R (as in (2)) that only depends on α and d. If each $\phi^{(Q)}$ equals a fixed $\phi \in \mathcal{C}_{\alpha,0}$ (a "mother wavelet") then (3) is sometimes called a wavelet system [2].

We could also consider the collection

$$\left\{ \frac{\phi^{(Q)}_{z_Q}}{\sqrt{\mu(Q)}} \right\}_{Q\in\mathcal{D}}. \tag{4}$$

In [9] the author showed that, if $\mu \in A_\infty$ then, for every family $\{\phi^{(Q)}\}_{\mathcal{D}} \subset \mathcal{C}_\alpha$, the set (3) is almost-orthogonal in L^2 if and only if (4) is almost-orthogonal in $L^2(\mu)$. He showed that this result has a partial converse: if μ is a doubling measure and it is the case that, for every $\{\phi^{(Q)}\}_{\mathcal{D}} \subset \mathcal{C}_\alpha$, the L^2 almost-orthogonality of (3) implies the $L^2(\mu)$ almost-orthogonality of (4), then $\mu \in A_\infty$.

In a later paper [10] the author strengthened the converse. We define a *T-sequence* to be a function ζ mapping from $\mathcal{D}$ into $\mathbf{R}^{d+1}_+$ such that $\zeta(Q) \in \overline{T(Q)}$ for all $Q \in \mathcal{D}$. In [10] the author proved that if μ is doubling, and ϕ is any non-trivial, real, radial function in $\mathcal{C}_{\alpha,0}$ such that, for all T-sequences ζ, the family

$$\left\{ \frac{\phi_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right\}_{Q\in\mathcal{D}} \tag{5}$$

is almost-orthogonal in $L^2(\mu)$, then $\mu \in A_\infty$.

The hypotheses that ϕ be real and radial are unnecessary. The "real" assumption is a computational convenience. The "radial" hypothesis (combined with non-triviality) simply ensures that $\widehat{\phi}$ (the Fourier transform of ϕ) does not vanish identically on any ray emanating from the origin. It turns out that smoothness and cancelation are also red herrings, at least for showing *necessity* of $\mu \in A_\infty$. In the current work we replace these hypotheses with a non-degeneracy condition that can be applied to subsets of $L^\infty(B(0;1))$ (bounded functions with supports contained in $B(0;1)$). This condition allows individual functions in the set to have Fourier transforms with bad directions. It only requires that no direction be bad for *all* of them. Precisely, we say that $\{\phi_k\}_1^n \subset L^\infty(B(0;1))$ satisfies the *collective non-degeneracy condition* (CNDC) if there is no ray from the origin on which every $\widehat{\phi_k}$ is identically 0.

Our main result is:

Theorem 1.1 *Let μ be a doubling measure on $\mathbf{R}^d$ and let $\{\phi_k\}_1^n \subset L^\infty(B(0;1))$ satisfy CNDC. If, for every $1 \le k \le n$ and every T-sequence ζ, the set*

$$\left\{ \frac{(\phi_k)_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right\}_{Q\in\mathcal{D}} \tag{6}$$

is almost-orthogonal in $L^2(\mu)$, then $\mu \in A_\infty$.

The meaning of the theorem seems to be: If μ is doubling and $L^2(\mu)$ has a reasonable wavelet basis (one given by normalized translates/dilates of a finite set of mother wavelets), then μ must be A_∞.

The proof uses a slightly non-standard characterization of A_∞; or, to be more precise, dyadic A_∞. We recall that a measure ν belongs to dyadic A_∞ (in symbols: $\nu \in A_\infty^d$) if, for every $\epsilon > 0$, there is a $\delta > 0$ so that, for all *dyadic* cubes Q and all measurable $E \subset Q$, $|E|/|Q| < \delta$ implies $\nu(E) \le \epsilon\nu(Q)$. Obviously $A_\infty \subset A_\infty^d$. It is not hard to show that if $\nu \in A_\infty^d$ and ν is doubling then $\nu \in A_\infty$. To prove Theorem 1.1, it suffices to show that its hypotheses imply $\mu \in A_\infty^d$.

We will call $\{c_Q\}_{\mathcal{D}}$, a sequence of non-negative numbers indexed over $\mathcal{D}$, a *Carleson sequence* if, for all $Q' \in \mathcal{D}$,

$$\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q'}} c_Q|Q| \le |Q'|. \tag{7}$$

This is the same as saying that, for every $Q' \in \mathcal{D}$,

$$\int_{Q'} \left(\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q'}} c_Q\chi_Q \right) dx \le |Q'|.$$

In section "The One-Dimensional, Dyadic Case" we show that $\nu \in A_\infty^d$ if and only if there is a finite R so that, for all Carleson sequences $\{c_Q\}_{\mathcal{D}}$ and all $Q' \in \mathcal{D}$,

$$\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q'}} c_Q\nu(Q) \le R\nu(Q'); \tag{8}$$

which, the reader will note, is the same as

$$\int_{Q'} \left(\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q'}} c_Q\chi_Q \right) d\nu \le R\nu(Q').$$

We prove Theorem 1.1 by showing that, given its hypotheses, μ must satisfy (8), for some fixed R, for all $Q' \in \mathcal{D}$ and all Carleson sequences.

Aside from some technical lemmas, the proof turns on a simple observation. Suppose that $(\Omega, \mathcal{M}, \nu)$ is a measure space, and $f : \Omega \to \mathbf{C}$ satisfies

$$\int_\Omega |f|^2\, d\nu \le R \int_\Omega |f|\, d\nu < \infty \tag{9}$$

for some finite R. Then the Cauchy-Schwarz inequality implies

$$\int_\Omega |f|\, d\nu \le R\nu(\Omega). \tag{10}$$

(We need the '$< \infty$' in (9): consider $f(x) = 1/x$ on $(0, 1)$ with Lebesgue measure.) In the proof of Theorem 1.1, Ω will be a certain "nearly optimal" $Q' \in \mathcal{D}$ and f will *essentially* be a function of the form

$$\sum_{\substack{Q \in \mathcal{D} \\ Q \subset Q'}} c_Q \chi_Q,$$

with $\{c_Q\}_{\mathcal{D}}$ a "nearly optimal" Carleson sequence, carefully defined to have the second inequality in (9). After some work, Theorem 1.1's almost-orthogonality hypothesis will yield the first inequality in (9), giving us (10) (and (8)).

What seems to be going on here is a sneaky version of the self-improving ("John-Nirenberg") property of *BMO*. Recall that $f \in L^1_{loc}(\mathbf{R}^d)$ is said to belong to *BMO* if

$$\sup_{\substack{Q \subset \mathbf{R}^d \\ Q \text{ a cube}}} \frac{1}{|Q|} \int_Q |f - f_Q| \, dx \equiv \|f\|_* < \infty, \tag{11}$$

where f_Q denotes $\frac{1}{|Q|} \int_Q f \, dx$, f's average over Q. The John-Nirenberg theorem ([4], p. 144) states that there are postive constants $c_1(d)$ and $c_2(d)$ such that, if $f \in BMO$, then for all cubes Q and all numbers $\lambda > 0$,

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \leq c_1(d) \exp(-c_2(d)\lambda/\|f\|_*)|Q|.$$

This implies that if (11) holds then

$$\sup_{\substack{Q \subset \mathbf{R}^d \\ Q \text{ a cube}}} \frac{1}{|Q|} \int_Q |f - f_Q|^2 \, dx \leq C\|f\|_*^2$$

for some C depending only on d. In other words,

$$\sup_{\substack{Q \subset \mathbf{R}^d \\ Q \text{ a cube}}} \frac{1}{|Q|} \int_Q |f - f_Q|^2 \, dx \leq C \left(\sup_{\substack{Q \subset \mathbf{R}^d \\ Q \text{ a cube}}} \frac{1}{|Q|} \int_Q |f - f_Q| \, dx \right)^2 :$$

"the L^1 norm controls the L^2 norm."

Because we will need it later, we recall that $f \in L^1_{loc}(\mathbf{R}^d)$ is said to belong to *dyadic BMO* ("$f \in BMO_d$") if the inequalty (11) holds when the supremum is taken over all dyadic cubes. We write the resulting (finite) supremum as $\|f\|_{*,d}$. The analogous John-Nirenberg properties also hold for $f \in BMO_d$, with the cubes now required to belong to $\mathcal{D}$.

In section "The One-Dimensional, Dyadic Case" we state and prove a dyadic version of our main result, hoping it will illuminate the main ideas in the proof of Theorem 1.1.

In section "Technical Lemmas" we prove some technical lemmas.

In section "Proof of Theorem 1.1" we prove Theorem 1.1 and give, as a corollary, an application to wavelet representations of linear operators.

Notations. If A and B are positive quantities depending on some parameters, we write '$A \sim B$' ("A and B are comparable") to mean that there are positive numbers c_1 and c_2 ("comparability constants") so that

$$c_1 A \leq B \leq c_2 A; \tag{12}$$

and, if c_1 and c_2 depend on parameters, they do not do so in a way that makes (12) trivial. We often use 'C' to denote a constant that might change from occurrence to occurrence; we will not always say how C changes or what it depends on. If E and F are sets, we write $E \subset F$ to express $E \subseteq F$.

We will refer to "finite linear sums" of the form $\sum_{\gamma \in \Gamma} \lambda_\gamma g_\gamma(x)$, where $\{\lambda_\gamma\}_\Gamma$ is a set of numbers and $\{g_\gamma\}_\Gamma$ is a set of functions, both indexed over an infinite set Γ (typically $\mathcal{D}$). "Finite linear sum" will mean a sum in which all but finitely many of the λ_γ's are 0. Similarly, a "finite sequence" $\{\lambda_\gamma\}_\Gamma$ indexed over Γ will be one in which all but finitely many λ_γ's are 0.

We indicate the end of a proof with the symbol ♣.

The One-Dimensional, Dyadic Case

First we prove our characterization of A_∞^d (8) (see [7] and [11] for its original form).

Lemma 2.1 *A Radon measure μ belongs to A_∞^d if and only if there is a finite R so that* (8) *holds for all Carleson sequences $\{c_Q\}_\mathcal{D}$ and all $Q' \in \mathcal{D}$.*

Proof of Lemma 2.1 Suppose that $\mu \in A_\infty^d$. Then μ is absolutely continuous, and we can write $d\mu = v\,dx$, with $v \in A_\infty^d$. Classical arguments (see [1]) show that v satisfies (1) with respect to dyadic cubes, for some $p > 1$. Let $M_d(\cdot)$ denote the dyadic Hardy-Littlewood maximal operator:

$$M_d(g)(x) \equiv \sup_{x \in Q \in \mathcal{D}} \frac{1}{|Q|} \int_Q |g(t)|\,dt.$$

The L^p-boundedness of $M_d(\cdot)$ and Hölder's inequality imply, for any $Q' \in \mathcal{D}$,

$$\begin{aligned}
\frac{1}{|Q'|} \int_{Q'} M_d(\chi_{Q'} v)\,dx &\leq \left(\frac{1}{|Q'|} \int_{Q'} (M_d(\chi_{Q'} v))^p\,dx \right)^{1/p} \\
&\leq C_p \left(\frac{1}{|Q'|} \int_{Q'} (v(x))^p\,dx \right)^{1/p} \\
&\leq \frac{C_p K_p}{|Q'|} \int_{Q'} v(x)\,dx;
\end{aligned}$$

i.e.,

$$\int_{Q'} M_d(\chi_{Q'} v)\, dx \leq C v(Q')$$

for all $Q' \in \mathcal{D}$. Now let $\{c_Q\}_{\mathcal{D}}$ be a Carleson sequence. If $Q' \in \mathcal{D}$ then

$$\sum_{Q \subset Q'} c_Q v(Q) = \sum_{Q \subset Q'} c_Q |Q| \left(\frac{1}{|Q|} v(Q) \right) \leq \int_{Q'} M_d(\chi_{Q'} v)\, dx,$$

by standard tent-space arguments (see, e.g., Theorem 2 on page 59 of [4]). Therefore $\mu \in A_\infty^d$ implies (8).

Suppose (8) holds. First we will show that μ is absolutely continuous with respect to Lebesgue measure. Then we will finish the lemma's proof.

Suppose E is measurable, $|E| = 0$ and, without loss of generality, $E \subset Q_0 \in \mathcal{D}$. Cover E with countably many disjoint cubes $Q_1^j \subset Q_0$ such that

$$\sum_j |Q_1^j| \leq (1/2)|Q_0|.$$

Now, having chosen the cubes $\{Q_k^j\}_j$, let $\{Q_{k+1}^{j'}\}_{j'}$ be a family of disjoint dyadic cubes such that: a) $E \subset \cup_{j'} Q_{k+1}^{j'}$; b) each $Q_{k+1}^{j'}$ is a subset of some Q_k^j; c) for all Q_k^j,

$$\sum_{Q_{k+1}^{j'} \subset Q_k^j} |Q_{k+1}^{j'}| \leq (1/2)|Q_k^j|. \tag{13}$$

We can do this for all k because $|E| = 0$. Inequality (13) implies that, for any $Q \in \mathcal{D}$,

$$\sum_{Q_k^j \subset Q} |Q_k^j| \leq 2|Q|. \tag{14}$$

We give the quick (and well known) proof of (14). By induction, for any Q_k^j and any $n \geq 0$,

$$\sum_{Q_{k+n}^{j'} \subset Q_k^j} |Q_{k+n}^{j'}| \leq 2^{-n} |Q_k^j|,$$

which implies that

$$\sum_{Q_{k'}^{j'} \subset Q_k^j} |Q_{k'}^{j'}| \leq 2|Q_k^j|$$

for every Q_k^j. If Q is arbitrary let $\{Q_{k^*}^{j^*}\}_{j^*,k^*}$ the *maximal* Q_k^j's contained in Q. The cubes $Q_{k^*}^{j^*}$ are disjoint. Therefore

$$\sum_{Q_k^j \subset Q} |Q_k^j| = \sum_{j^*,k^*} \sum_{Q_j^k \subset Q_{k^*}^{j^*}} |Q_j^k| \leq 2 \sum_{j^*,k^*} |Q_{k^*}^{j^*}| \leq 2|Q|,$$

proving (14).

Define:

$$c_Q = \begin{cases} 1/2 & \text{if } Q \in \{Q_k^j\}_{j,k}; \\ 0 & \text{otherwise.} \end{cases}$$

Inequalities (13) and (14) imply that $\{c_Q\}_{\mathcal{D}}$ is Carleson. Therefore there is a finite R such that

$$\sum_{j,k} (1/2)\mu(Q_k^j) \leq R\mu(Q_0) < \infty.$$

But, because of a), for all N,

$$N\mu(E) \leq \sum_{k=1}^{N} \sum_j \mu(Q_k^j) \leq 2R\mu(Q_0),$$

forcing $\mu(E) = 0$.

The rest of the proof that $\mu \in A_\infty^d$ is like what we just saw, only more careful. Let $Q_0 \in \mathcal{D}$, $E \subset Q_0$, and $|E|/|Q_0| < \eta << 1$. For $k \geq 1$, let $\{Q_k^j\}_j$ be the maximal dyadic subcubes of Q_0 such that

$$\frac{|E \cap Q_k^j|}{|Q_k^j|} > 2^{(d+1)k}\eta.$$

These are the Calderón-Zygmund cubes, taken at "height" $2^{(d+1)k}\eta$, of χ_E relative to Q_0. Because of their maximality, for each Q_k^j,

$$\frac{|E \cap Q_k^j|}{|Q_k^j|} \leq 2^d 2^{(d+1)k}\eta = (1/2)2^{(d+1)(k+1)}\eta,$$

which implies that every cube $Q_{k+1}^{j'}$ is contained in some Q_k^j, and that, for every Q_k^j,

$$\sum_{Q_{k+1}^{j'} \subset Q_k^j} |Q_{k+1}^{j'}| \leq (1/2)|Q_k^j|,$$

which is the condition (13) we saw earlier. The same reasoning as before implies that, for all $Q \in \mathcal{D}$,

$$\sum_{Q_j^k \subset Q} |Q_j^k| \leq 2|Q|.$$

Almost every point of E is a point of density. Therefore we will keep getting cubes Q_k^j as long as $2^{(d+1)k}\eta$ is less than 1: there is a $K_0 \sim \log(1/\eta)$ such that, for all $1 \leq k \leq K_0$, $|E \setminus \cup_j Q_k^j| = 0$, and hence $\mu(E \setminus \cup_j Q_k^j) = 0$. (The union $\cup_j Q_k^j$ "almost contains" E.) Define:

$$c_Q = \begin{cases} 1/2 & \text{if } Q \in \{Q_j^k\}_{j,k}; \\ 0 & \text{otherwise.} \end{cases}$$

The sequence $\{c_Q\}_{\mathcal{D}}$ is Carleson; therefore

$$\sum_{Q \subset Q_0} c_Q \mu(Q) \leq R\mu(Q_0).$$

But

$$\sum_{Q \subset Q_0} c_Q \mu(Q) = (1/2) \sum_{j,k} \mu(Q_j^k) \geq (1/2) \sum_{k=1}^{K_0} \sum_j \mu(Q_k^j) \geq (1/2) K_0 \mu(E),$$

because, for each $k \leq K_0$, the part of E outside $\cup_j Q_k^j$ has μ-measure 0. Thus,

$$\mu(E) \leq \frac{2R}{K_0} \mu(Q_0),$$

and $2R/K_0 \to 0$ as $\eta \to 0^+$: $\mu \in A_\infty^d$. ♣

If $I = [j2^k, (j+1)2^k) \subset \mathbf{R}$ is a dyadic interval, define $I^+ \equiv [2j2^{k-1}, (2j+1)2^{k-1})$ (I's left half) and $I^- \equiv [(2j+1)2^{k-1}, (2j+2)2^{k-1})$ (I's right half), and set

$$h_{(I)} \equiv \chi_{I^+} - \chi_{I^-}.$$

The functions $\{h_{(I)}/|I|^{1/2}\}_{I \in \mathcal{D}}$ are known as the Haar functions, which comprise an orthonormal basis for $L^2(\mathbf{R})$.

The dyadic analogue of Theorem 1.1 is

Theorem 2.2 *Let μ be a non-trivial Radon measure on* **R**. *If*

$$\left\{ \frac{h_{(I)}}{\sqrt{\mu(I)}} \right\}_{I \in \mathcal{D}} \tag{15}$$

is almost-orthogonal in $L^2(\mu)$ then $\mu \in A_\infty^d$.

Proof of Theorem 2.2. The reader might want to look back at (9) and (10).

Fix $I_0 \in \mathcal{D}$ and $0 < \epsilon << \ell(I_0)$. Let $\mathcal{F}(I_0, \epsilon)$ be the familiy of Carleson sequences $\{c_I\}_{\mathcal{D}}$ such that $c_I = 0$ if $I \not\subset I_0$ or $\ell(I) < \epsilon$. By compactness, there is a Carleson sequence $\{\tilde{c}_I\}_{\mathcal{D}} \in \mathcal{F}(I_0, \epsilon)$ such that

$$\sum_{\mathcal{D}} \tilde{c}_I \mu(I) = \sup \left\{ \sum_{\mathcal{D}} c_I \mu(I) : \{c_I\}_{\mathcal{D}} \in \mathcal{F}(I_0, \epsilon) \right\} < \infty.$$

Call the supremum L. Define

$$f(x) \equiv \sum_{\mathcal{D}} \tilde{c}_I \chi_I(x) - \left(\sum_{\mathcal{D}} \tilde{c}_I |I| \right) \frac{\chi_{I_0}(x)}{|I_0|}.$$

Notice that, because $\{\tilde{c}_I\}_{\mathcal{D}}$ is Carleson,

$$\frac{1}{|I_0|} \left(\sum_{\mathcal{D}} \tilde{c}_I |I| \right) \leq 1.$$

The function f is supported on I_0 and satisfies $\int f\, dx = 0$. Also, f belongs to BMO_d, with $\|f\|_{*,d} \leq 2$. Let us prove this fact. Take $J \in \mathcal{D}$. If $J \cap I_0 = \emptyset$ we have nothing to prove. If $I_0 \subset J$ then $f_J = 0$ and

$$\int_J |f - f_J|\, dx \leq 2 \sum_{\mathcal{D}} \tilde{c}_I |I| \leq 2|I_0| \leq 2|J|.$$

If $J \subset I_0$ then

$$\int_J |f - f_J|\, dx \leq 2 \sum_{I \in \mathcal{D}: I \subset J} \tilde{c}_I |I| \leq 2|J|.$$

By the John-Nirenberg theorem, there exists an absolute constant—which we call C—so that, for all $J \in \mathcal{D}$,

$$\int_J |f - f_J|^2\, dx = \sum_{I \in \mathcal{D}:\ I \subset J} \frac{|\langle f, h_{(I)}\rangle|^2}{|I|} \leq C|J|, \tag{16}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual (Lebesgue) L^2 inner product. Because of how we defined f, the inner products $\langle f, h_{(I)}\rangle = 0$ if $I \not\subset I_0$ or $\ell(I) < \epsilon$. Therefore the sequence defined by

$$\alpha_I \equiv \frac{|\langle f, h_{(I)}\rangle|^2}{|I|^2}$$

is a bounded multiple of a sequence from $\mathcal{F}(I_0, \epsilon)$, implying

$$\sum_{\mathcal{D}} \frac{|\langle f, h_{(I)}\rangle|^2}{|I|^2} \mu(I) \leq CL,$$

with C an absolute constant.

We can write

$$f = \sum_{\mathcal{D}} \frac{\langle f, h_{(I)}\rangle}{|I|} h_{(I)},$$

and this is an exact, finite sum, because of f's special form. We rewrite it as

$$f = \sum_{\mathcal{D}} \gamma_I \frac{h_{(I)}}{\sqrt{\mu(I)}},$$

where

$$\gamma_I = \langle f, h_{(I)}\rangle \frac{\sqrt{\mu(I)}}{|I|}.$$

The $L^2(\mu)$ almost-orthogonality of (15) implies that

$$\begin{aligned}\int |f|^2\, d\mu \leq R \sum_{\mathcal{D}} |\gamma_I|^2 = R \sum |\langle f, h_{(I)}\rangle|^2 \frac{\mu(I)}{|I|^2} \\ = R \sum_{\mathcal{D}} \frac{|\langle f, h_{(I)}\rangle|^2}{|I|^2} \mu(I) \\ \leq RCL.\end{aligned}$$

But

$$L = \sum_{\mathcal{D}} \tilde{c}_I \mu(I) = \int_{I_0} (f + c_0)\, d\mu,$$

where

$$c_0 = \frac{1}{|I_0|} \sum_{\mathcal{D}} \tilde{c}_I |I| \leq 1.$$

Therefore

$$\int |f|^2\, d\mu \leq RC \left(\int |f|\, d\mu + \mu(I_0) \right),$$

which implies

$$\int |f|\, d\mu \leq C' \mu(I_0),$$

and

$$\sum_{\mathcal{D}} \tilde{c}_I \mu(I) \leq C'' \left(\int |f|\, d\mu + \mu(I_0) \right) \leq \tilde{C} \mu(I_0). \tag{17}$$

The sequence $\{\tilde{c}_I\}_{\mathcal{D}}$ is optimal for sequences from $\mathcal{F}(I_0, \epsilon)$. Therefore (17) holds for every sequence in $\mathcal{F}(I_0, \epsilon)$. But the bound holds independent of I_0 and ϵ; therefore, by an obvious limiting argument, it holds for all Carleson sequences $\{c_I\}_{\mathcal{D}}$. By Lemma 2.1, the measure μ belongs to A_∞^d. ♣

Remark We ask the reader to note how, in the interaction between (16) and (17), the John-Nirenberg theorem lets us bound an L^2 norm by an L^1 norm—which is the heart of the proof.

Technical Lemmas

The first lemma in this section says that, if every family of the form (6) is almost-orthogonal in $L^2(\mu)$, then these families must be, in an obvious sense, *uniformly* almost-orthogonal.

Lemma 3.1 *Let $\psi \in L^\infty(B(0; 1))$. Suppose that, for every T-sequence ζ, there is a finite $R = R(\zeta, \mu, \psi)$ such that, for all finite linear sums*

$$\sum_{\mathcal{D}} \lambda_Q \frac{\psi_{\zeta(Q)}}{\sqrt{\mu(Q)}},$$

we have

$$\int \left| \sum_{\mathcal{D}} \lambda_Q \frac{\psi_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu \leq R \sum_{\mathcal{D}} |\lambda_Q|^2. \tag{18}$$

Then there is a finite $\tilde{R} = \tilde{R}(\mu, \psi)$ *such that* (18) *holds for all* T*-sequences* ζ.

Proof of Lemma 3.1 For every T-sequence ζ, we can define a linear map L_ζ : $\ell^2(\mathcal{D}) \to L^2(\mu)$ by

$$L_\zeta(\{\lambda_Q\}_{\mathcal{D}}) \equiv \sum_{\mathcal{D}} \lambda_Q \frac{\psi_{\zeta(Q)}}{\sqrt{\mu(Q)}}. \tag{19}$$

Inequality (18) shows that the series in (19) converges unconditionally to an $f \in L^2(\mu)$, and that $\int |f|^2 \, d\mu \leq R \sum_{\mathcal{D}} |\lambda_Q|^2$. By the Uniform Boundedness Principle, if no universal $\tilde{R}$ exists, then there is a sequence $\{\lambda_Q\}_{\mathcal{D}} \in \ell^2(\mathcal{D})$ such that $\sum_{\mathcal{D}} |\lambda_Q|^2 \leq 1$, and there is a sequence of T-sequences ζ_k, such that

$$\int \left| \sum_{\mathcal{D}} \lambda_Q \frac{\psi_{\zeta_k(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu \to \infty. \tag{20}$$

We will patch together a T-sequence $\tilde{\zeta}$ such that $\{\frac{\psi_{\tilde{\zeta}(Q)}}{\sqrt{\mu(Q)}}\}_{\mathcal{D}}$ is not almost-orthogonal. Fix the sequence $\{\lambda_Q\}_{\mathcal{D}}$. If $\mathcal{F} \subset \mathcal{D}$ is finite, there is an $N = N(\mathcal{F})$ such that

$$\int \left| \sum_{Q \in \mathcal{F}} \lambda_Q \frac{\psi_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu \leq N$$

for all T-sequences ζ. Thus, because of (20), we know that, if $\mathcal{F}_0 \subset \mathcal{D}$ is finite and R is any large number, there is a finite subset $\mathcal{F}_1 \subset \mathcal{D}$, disjoint from $\mathcal{F}_0$, and there is a T-sequence ζ_1, such that

$$\int \left| \sum_{Q \in \mathcal{F}_1} \lambda_Q \frac{\psi_{\zeta_1(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu > R.$$

Let $R_k \to \infty$. Let $\mathcal{F}_1 \subset \mathcal{D}$ be a finite subset and ζ_1 a T-sequence such that

$$\int \left| \sum_{Q \in \mathcal{F}_1} \lambda_Q \frac{\psi_{\zeta_1(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu > R_1.$$

Having defined $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$, let $\mathcal{F}_{n+1} \subset \mathcal{D}$ be a finite subset disjoint from $\cup_1^n \mathcal{F}_k$, and ζ_{n+1} a T-sequence such that

$$\int \left| \sum_{Q \in \mathcal{F}_{n+1}} \lambda_Q \frac{\psi_{\zeta_{n+1}(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu > R_{n+1}.$$

Define $\tilde{\zeta} : \mathcal{D} \to \mathbf{R}_+^{d+1}$ by

$$\tilde{\zeta}(Q) = \begin{cases} \zeta_k(Q) & \text{if } Q \in \mathcal{F}_k; \\ z_Q & \text{if } Q \notin \cup_k \mathcal{F}_k. \end{cases}$$

Then $\tilde{\zeta}$ is a T-sequence for which (18) fails. ♣

The proof of Theorem 1.1 uses a general form of the Calderón reproducing formula. Our approach is based on ideas and methods of Frazier, Jawerth, and Weiss [3]. We gratefully acknowledge their influence and inspiration.

Recall that if $\psi \in \mathcal{C}_{\alpha,0}$ is real, radial, non-trivial, and normalized so that

$$\int_0^\infty |\widehat{\psi}(y\xi)|^2 \, \frac{dy}{y} = 1$$

for all $\xi \neq 0$, then, if $f \in \cup_{1<p<\infty} L^p(\mathbf{R}^d)$, we have

$$f(x) = \int_{\mathbf{R}_+^{d+1}} (f * y^{-d}\psi_{(0,y)}(t))\, y^{-d}\psi_{(0,y)}(x-t)\, \frac{dt\, dy}{y}$$

in various senses [8, 11]. To be consistent with the notation in the introduction, we have written "$y^{-d}\psi_{(0,y)}$" in place of the more traditional "ψ_y". We will continue to follow this convention.

We define $\Phi(x)$ to be the inverse Fourier transform of $\exp(-|\xi|^2 - |\xi|^{-2})$. We notice that Φ belongs to the Schwartz class $\mathcal{S}$, and that $\widehat{\Phi}(\xi)$ and all of $\widehat{\Phi}$'s derivatives vanish to infinite order at the origin.

It is important that $\widehat{\Phi}(\xi) > 0$ on all of $\mathbf{R}^d \setminus \{0\}$.

Lemma 3.2 *Suppose that* $\{\phi_k\}_1^n \subset L^\infty(B(0;1))$ *satisfies CNDC. For* $\xi \in \mathbf{R}^d \setminus \{0\}$ *define*

$$G(\xi) \equiv \int_0^\infty \widehat{\Phi}(y\xi) \left(\sum_1^n |\widehat{\phi_k}(y\xi)|^2 \right) \frac{dy}{y}. \tag{21}$$

The function $G(\xi)$ *is infinitely differentiable on* $\mathbf{R}^d \setminus \{0\}$ *and homogeneous of degree 0:* $G(t\xi) = G(\xi)$ *for all* $t > 0$. *There are positive numbers* c_1 *and* c_2 *such that* $c_1 \leq G(\xi) \leq c_2$ *for all* $\xi \neq 0$.

Proof of lemma. The homogeneity is obvious. Every $\widehat{\phi_k}$ is infinitely differentiable, and $D^\alpha \widehat{\phi_k} \in L^\infty$ for every k and multi-index α. The function $\widehat{\Phi}$ is also infinitely differentiable, and, for all α, $D^\alpha \widehat{\Phi}$ vanishes rapidly at 0 and infinity. These imply that G is infinitely differentiable. The CNDC implies that $G(\xi)$ never vanishes on $S^{d-1} \equiv \{\xi : |\xi| = 1\}$. The smoothness of G and the compactness of S^{d-1} imply that G lies between two positive constants there, hence on all of $\mathbf{R}^d \setminus \{0\}$. ♣

Now, given $\{\phi_k\}_1^n \subset L^\infty(B(0;1))$ satisfying CNDC, and G as defined by (21), we set

$$m(\xi) \equiv \frac{1}{G(\xi)} \tag{22}$$

for $\xi \neq 0$, and undefined at the origin. By standard arguments ([4], p. 26), the Fourier multiplier operators given by

$$\widehat{T_G f}(\xi) \equiv G(\xi)\widehat{f}(\xi)$$

and

$$\widehat{T_m f}(\xi) \equiv m(\xi)\widehat{f}(\xi),$$

initially defined for $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$, extend to bounded operators on $L^p(\mathbf{R}^d)$ for every $1 < p < \infty$. On these domains they are inverses of each other: $T_G T_m = T_m T_G = I$, the identity.

For each ϕ_k, define $\tilde{\phi}_k(x) \equiv \overline{\phi_k(-x)}$, and recall that $\widehat{\tilde{\phi}_k}(\xi) = \overline{\widehat{\phi_k}}(\xi)$. If $f \in L^2(\mathbf{R}^d)$ then

$$T_G f = \sum_1^n \int_{\mathbf{R}_+^{d+1}} (f * y^{-d}\Phi_{(0,y)} * (y^{-d}\tilde{\phi}_k)_{(0,y)}(t))\,(y^{-d}\phi_k)_{(0,y)}(x-t)\,\frac{dt\,dy}{y},$$

where we interpret each integral as

$$\lim_{\substack{\epsilon \searrow 0 \\ R \nearrow \infty}} \int_\epsilon^R \left(\int_{\mathbf{R}^d} (f * y^{-d}\Phi_{(0,y)} * (y^{-d}\tilde{\phi}_k)_{(0,y)}(t))\,(y^{-d}\phi_k)_{(0,y)}(x-t)\,dt\right)\frac{dy}{y},$$

with the limit existing in L^2. As we shall see, if $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$, the limit also exists pointwise in x, with the integral being, in a natural sense, absolutely convergent.

Because T_m and T_G are inverses of each other, if $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$,

$$f = \sum_1^n \int_{\mathbf{R}_+^{d+1}} (f * T_m(y^{-d}\Phi_{(0,y)}) * (y^{-d}\tilde{\phi}_k)_{(0,y)}(t))\,(y^{-d}\phi_k)_{(0,y)}(x-t)\,\frac{dt\,dy}{y},$$

where the integrals converge (in the above sense) in L^2. Let us define

$$\Psi(x) \equiv T_m(\Phi)(x).$$

With this notation, we can rewrite the preceding integral formula as

$$f = \sum_1^n \int_{\mathbf{R}_+^{d+1}} (f * y^{-d}\Psi_{(0,y)} * (y^{-d}\tilde{\phi}_k)_{(0,y)}(t))\,(y^{-d}\phi_k)_{(0,y)}(x-t)\,\frac{dt\,dy}{y}.$$

(We have used the dilation-invariance of T_m.)

A look at Ψ's Fourier transform shows that $\Psi \in \mathcal{S}$ and $\int \Psi\,dx = 0$. The same are true of Ψ_k, which we define as

$$\Psi_k(x) \equiv \Psi * \tilde{\phi}_k(x).$$

With this convention we can compress our integral formula to

$$f = \sum_1^n \int_{\mathbf{R}_+^{d+1}} (f * y^{-d}(\Psi_k)_{(0,y)}(t))\,(y^{-d}\phi_k)_{(0,y)}(x-t)\,\frac{dt\,dy}{y}. \tag{23}$$

We now prove two lemmas relating to (23).

Lemma 3.3 *Suppose that $\Gamma \in \mathcal{S}$, $\int \Gamma\,dx = 0$, and $\gamma \in L^\infty(B(0;1))$. There is a $C = C(\Gamma, \gamma)$ such that, if $f \in C_0^\infty(\mathbf{R}^d)$ satisfies $|\nabla f| \le A$ pointwise and B is any positive number, then*

$$\int_0^B \left(\int_{\mathbf{R}^d} \left| (f * y^{-d}\Gamma_{(0,y)}(t))\,(y^{-d}\gamma)_{(0,y)}(x-t) \right| dt \right) \frac{dy}{y} \le CAB.$$

Remark In our applications of Lemma 3.3, $\Gamma = \Psi_k$, $\gamma = \phi_k$, and $AB \sim 1$.

Proof of Lemma 3.3 The function Γ satisfies

$$|\Gamma(x)| \le C(1+|x|)^{-d-2}$$
$$|\nabla\Gamma(x)| \le C(1+|x|)^{-d-3}$$
$$\int \Gamma(x)\,dx = 0,$$

for a fixed constant C. A lemma of Uchiyama [6] says that we can decompose Γ into a rapidly converging sum of dilates of smooth, compactly supported functions, with integrals equal to 0. Precisely:

$$\Gamma(x) = C \sum_{j=0}^\infty 2^{-j(d+2)} (F_j)_{(0,2^j)}(x),$$

for an appropriate C, where each F_j has support contained in $B(0;1)$ and satisfies

$$\|F_j\|_\infty \leq C$$
$$\int F_j\,dx = 0.$$

(Uchiyama's lemma actually yields $\|\nabla F_j\|_\infty \leq C$, but we don't need that.) The function $(F_j)_{(0,2^j)}$ has support contained in $B(0;2^j)$ and the function $((F_j)_{(0,2^j)})_{(0,y)}$ has support contained in $B(0;2^j y)$. The smoothness of f and the cancelation in F_j imply that

$$\begin{aligned}|f * y^{-d}((F_j)_{(0,2^j)})_{(0,y)}(t)| &\leq CA2^j y\|y^{-d}((F_j)_{(0,2^j)})_{(0,y)}\|_1 \\ &\leq CA2^j y 2^{jd} = CA2^{j(d+1)}y\end{aligned}$$

for any t, and therefore

$$\begin{aligned}|f * y^{-d}\Gamma_{(0,y)}(t)| &\leq CA\sum_{j=0}^{\infty} 2^{-j(d+2)}2^{j(d+1)}y \\ &= CAy.\end{aligned}$$

Since $\|\gamma\|_1 \leq C(\gamma)$,

$$\int_{\mathbf{R}^d} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right|\,dt \leq CAy,$$

implying

$$\begin{aligned}\int_0^B \left(\int_{\mathbf{R}^d} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right|\,dt\right)\frac{dy}{y} &\leq \int_0^B (CAy)\,\frac{dy}{y} \\ &= CAB,\end{aligned}$$

proving the lemma. ♣

The next lemma uses a standard definition and one derived from it.

Definition 3.4 If $Q \subset \mathbf{R}^d$ is a cube then we set $\widehat{Q} \equiv Q \times (0, \ell(Q)) \subset \mathbf{R}^{d+1}_+$ (sometimes called the "Carleson box" above Q) and $R(Q) \equiv \{(t,y) \in \mathbf{R}^{d+1}_+ : d((t,y),\widehat{Q}) \geq \ell(Q)\}$, where $d(\cdot,\cdot)$ denotes the usual Euclidean distance to a set in $\mathbf{R}^{d+1}_+$.

Lemma 3.5 *Let $\Gamma \in \mathcal{S}$ and $\gamma \in L^\infty(B(0;1))$. There is constant $C = C(\Gamma,\gamma)$ such that if $f \in L^1(\mathbf{R}^d)$ and the support of f is contained in a cube Q then*

$$\int_{R(Q)} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right| \,\frac{dt\,dy}{y} \le \frac{C}{|Q|}\int |f|\,dt$$

for all $x \in Q$.

Proof of Lemma 3.5. For $j = 0, 1, 2\ldots$, define $R_j(Q) \equiv \{(t,y) \in R(Q) : 2^j\ell(Q) \le d((t,y),\widehat{Q}) < 2^{j+1}\ell(Q)\}$, and observe that $R(Q) = \cup_0^\infty R_j(Q)$. Since γ has its support contained in $B(0;1)$, $\gamma_{(0,y)}(x-t) = \gamma(\frac{x-t}{y})$ can be non-zero only if $|x-t| < y$. Therefore there is a positive $c = c(d)$ such that, if $x \in Q$ and $(t,y) \in R_j(Q)$, $\gamma_{(0,y)}(x-t)$ will be zero unless $y > c2^j\ell(Q)$. If $y > c2^j\ell(Q)$, Hölder's inequality implies

$$|f * y^{-d}\Gamma_{(0,y)}(t)| \le C(2^j\ell(Q))^{-d}\|f\|_1$$

and

$$\int_{\mathbf{R}^d} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right| \,dt \le C(2^j\ell(Q))^{-d}\|f\|_1.$$

If $(t,y) \in R_j(Q)$ then $y < 2^{j+2}\ell(Q)$. Therefore:

$$\begin{aligned}&\textstyle\int_{R_j(Q)} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right| \,\frac{dt\,dy}{y}\\ &\le C\textstyle\int_{c2^j\ell(Q)}^{2^{j+2}\ell(Q)} \left(\int_{\mathbf{R}^d} \left|(f * y^{-d}\Gamma_{(0,y)}(t))\, y^{-d}\gamma_{(0,y)}(x-t)\right| \,dt\right) \frac{dy}{y}\\ &\le C(2^j\ell(Q))^{-d}\|f\|_1.\end{aligned}$$

Summing over j finishes the proof. ♣

Proof of Theorem 1.1.

For the rest of this section, μ will be a fixed doubling measure.

The proof of Theorem 1.1 works by rewriting each of the n summands in (23) as an average of sums of the form

$$\sum_{\mathcal{D}} \lambda_Q \frac{(\phi_k)_{\zeta(Q)}}{\sqrt{\mu(Q)}}$$

where ζ is a T-sequence. We now describe how this rewriting will go. If $Q = [j_1 2^k, (j_1+1)2^k) \times \cdots \times [j_d 2^k, (j_d+1)2^k) \in \mathcal{D}$ we set $t_Q \equiv (j_1 2^k, j_2 2^k, \ldots, j_d 2^k)$, the "left-most corner" of Q. Define $V_0 \equiv [0,1)^d$, the "unit" dyadic cube. If $Q \in \mathcal{D}$, we define a bijective mapping $\sigma(Q,\cdot,\cdot) : T(V_0) \to T(Q)$ by

$$\sigma(Q, \tau, \eta) \equiv (t_Q + \ell(Q)\tau, \ell(Q)\eta).$$

We point out some properties of this mapping. If $g : T(Q) \to \mathbf{C}$ is measurable we can define $h : T(V_0) \to \mathbf{C}$ by $h(\tau, \eta) \equiv g(\sigma(Q, \tau, \eta))$. By the change-of-variables formula,

$$\int_{T(Q)} g(t, y) \frac{dt\,dy}{y} = |Q| \int_{T(V_0)} h(\tau, \eta) \frac{d\tau\,d\eta}{\eta}. \tag{24}$$

We can write

$$\int_{T(Q)} (f * y^{-d}(\Psi_k)_{(0,y)}(t))\, y^{-d}(\phi_k)_{(0,y)}(x - t) \frac{dt\,dy}{y}$$

as

$$\int_{T(Q)} y^{-2d} \langle f, \overline{(\Psi_k)_{(t,y)}}\rangle\, (\phi_k)_{(t,y)}(x) \frac{dt\,dy}{y},$$

where $\langle\cdot,\cdot\rangle$ is the ordinary L^2 inner product. Because of (24), this is equal to

$$\begin{aligned} &|Q| \textstyle\int_{T(V_0)} (\ell(Q)\eta)^{-2d} \langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle\, (\phi_k)_{\sigma(Q,\tau,\eta)}(x) \frac{d\tau\,d\eta}{\eta} \\ &= |Q|^{-1} \textstyle\int_{T(V_0)} \eta^{-2d} \langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle\, (\phi_k)_{\sigma(Q,\tau,\eta)}(x) \frac{d\tau\,d\eta}{\eta}. \end{aligned}$$

Therefore, we can *formally* rewrite the integral in (23) as:

$$\begin{aligned} &\textstyle\sum_{\mathcal{D}} \int_{T(Q)} y^{-2d} \langle f, \overline{(\Psi_k)_{(t,y)}}\rangle\, (\phi_k)_{(t,y)}(x) \frac{dt\,dy}{y} \\ &= \textstyle\int_{T(V_0)} \left(\sum_{\mathcal{D}} |Q|^{-1} \langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle\, (\phi_k)_{\sigma(Q,\tau,\eta)}(x)\right) \eta^{-2d} \frac{d\tau\,d\eta}{\eta}. \end{aligned} \tag{25}$$

Of course, if the summation only runs over a finite set of Q's (as it will for us), the equality is literal.

In proving Theorem 1.1, it will be more convenient to write (25) as

$$\int_{T(V_0)} \sum_{\mathcal{D}} \left[\left(|Q|^{-1} \langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle \sqrt{\mu(Q)}\right) \frac{(\phi_k)_{\sigma(Q,\tau,\eta)}(x)}{\sqrt{\mu(Q)}} \right] \eta^{-2d} \frac{d\tau\,d\eta}{\eta}.$$

Proof of Theorem 1.1 We fix, once and for all, a function $b \in \mathcal{C}_0^\infty(\mathbf{R}^d)$ that is non-negative, has support contained in $B(0; 1/2)$, and satisfies $\int b\,dx = 1$. Recall our definition of $z_Q \equiv (x_Q, \ell(Q))$, where x_Q is Q's center and $\ell(Q)$ is Q's sidelength. If $Q \subset \mathbf{R}^d$ is any cube then b_{z_Q} is supported in Q and satisfies $\int b_{z_Q}\,dx = |Q|$. If ν is any doubling measure then

$$\int b_{z_Q}\,d\nu \sim \nu(Q), \tag{26}$$

with comparability constants depending on b and ν. If $Q_0 \in \mathcal{D}$ and $2^j << 1$, we define $\mathcal{F}(Q_0, 2^j)$ to be the family of Carleson sequences $\{c_Q\}_{\mathcal{D}}$ such that $c_Q = 0$ if $Q \not\subset Q_0$ or $\ell(Q) < 2^j \ell(Q_0)$. It is clear that the set of numbers

$$\left\{\mu(Q_0)^{-1} \sum_{\mathcal{D}} c_Q \mu(Q) : \; Q_0 \in \mathcal{D}, \; \{c_Q\}_{\mathcal{D}} \in \mathcal{F}(Q_0, 2^j)\right\} \tag{27}$$

is bounded above by $1 + |j|$. Call the actual supremum $L(j)$. Theorem 1.1 will follow once we show that $\sup_j L(j) < \infty$.

Fix j. There exist a $Q_0 \in \mathcal{D}$ and a Carleson sequence $\{\tilde{c}_Q\}_{\mathcal{D}} \in \mathcal{F}(Q_0, 2^j)$ such that

$$\mu(Q_0)^{-1} \sum_{\mathcal{D}} \tilde{c}_Q \mu(Q) \geq (1/2) L(j).$$

Fix Q_0 and $\{\tilde{c}_Q\}$. Theorem 1.1 will follow if we show that $\mu(Q_0)^{-1} \sum_{\mathcal{D}} \tilde{c}_Q \mu(Q)$ is bounded by a number independent of Q_0 and j.

Define

$$f(x) \equiv \sum_{\mathcal{D}} \tilde{c}_Q b_{z_Q}(x).$$

Because of (26),

$$\int f \, d\mu \sim \sum_{\mathcal{D}} \tilde{c}_Q \mu(Q) \sim L(j) \mu(Q_0). \tag{28}$$

As with Theorem 2.2, the "game" now is to show that

$$\int |f|^2 \, d\mu \leq C \int |f| \, d\mu, \tag{29}$$

for some $C < \infty$ independent of Q_0 and j; because, as we have seen, the Cauchy-Schwarz inequality will imply

$$\int |f| \, d\mu \leq C \mu(Q_0);$$

which, with (28), will yield

$$L(j) \leq C,$$

for some absolute C independent of Q_0 and j.

Because of (28), (29) will follow from

$$\int |f|^2\, d\mu \leq CL(j)\mu(Q_0).$$

It is obvious that f is supported inside Q_0 and satisfies $\int |f|\, dx \leq |Q_0|$. It will be important to us that $f \in BMO$, with a BMO norm bounded by a constant depending only on b and d; so let us prove this. Write $f = \sum_k f_k$, where

$$f_k(x) = \sum_{Q:\ \ell(Q)=2^k} \tilde{c}_Q b_{z_Q}(x).$$

Each f_k is infinitely differentiable and satisfies: (i) $\|f_k\|_\infty \leq 1$; and (ii) $\|\nabla f_k\|_\infty \leq C2^{-k}$. We note that inequality (ii) implies $|\nabla f| \leq C(2^j\ell(Q_0))^{-1}$ pointwise.

Let Q' be a cube and write

$$f = \sum_{k:\,2^k \geq \ell(Q')} f_k + \sum_{k:\,2^k < \ell(Q')} f_k \equiv F_1 + F_2.$$

We can cover Q' with $C(d)$ congruent dyadic cubes $\{Q_j^*\}_1^{C(d)}$ such that $(1/2)\ell(Q') \leq \ell(Q_j^*) < \ell(Q')$, which implies that, if $Q \in \mathcal{D}$ and $\ell(Q) < \ell(Q')$, then $\ell(Q) \leq \ell(Q_j^*)$ for every j; hence, if $Q \cap Q' \neq \emptyset$ then $Q \subset Q_j^*$ for some j. Then:

$$\begin{aligned}
\int_{Q'} |F_2(x)|\, dx &= \int_{Q'} \left(\sum_{Q:\ell(Q)<\ell(Q')} \tilde{c}_Q b_{z_Q}(x) \right) dx \\
&\leq \sum_{j=1}^{C(d)} \int_{Q_j^*} \left(\sum_{Q:Q\subset Q_j^*} \tilde{c}_Q b_{z_Q}(x) \right) dx \\
&\leq \sum_{j=1}^{C(d)} \sum_{Q:Q\subset Q_j^*} \tilde{c}_Q |Q| \\
&\leq \sum_{1}^{C(d)} |Q_j^*| \\
&\leq C|Q'|.
\end{aligned}$$

On the other hand, $|\nabla F_1(x)| \leq C/\ell(Q')$, implying that

$$\int_{Q'} |F_1(x) - (F_1)_{Q'}|\, dx \leq C|Q'|.$$

Therefore f belongs to BMO, with a norm $\leq C$.

We invoke a standard fact about *BMO* ([4], p. 159): If $h \in BMO$, $\Gamma \in \mathcal{S}$, and $\int \Gamma\, dx = 0$, then, for all cubes $Q \subset \mathbf{R}^d$,

$$\frac{1}{|Q|} \int_{\widehat{Q}} |h * y^{-d} \Gamma_{(0,y)}(t)|^2 \frac{dt\, dy}{y} \leq C \|h\|_*^2,$$

where the constant C only depends on Γ. This implies that, for $h \in BMO$, the sequence of numbers $\{c_Q\}_{\mathcal{D}}$ defined by

$$c_Q \equiv \frac{1}{|Q|} \int_{T(Q)} |h * y^{-d} \Gamma_{(0,y)}(t)|^2 \frac{dt\, dy}{y}$$

is a bounded multiple of a Carleson sequence.

We can write $f = g_1 + g_2 + g_3 + g_4$, where

$$g_1(x) \equiv \sum_1^n \int_{\{(t,y):\ y<2^{j-1}\ell(Q_0)\}} (f * y^{-d} (\Psi_k)_{(0,y)}(t))\, y^{-d} (\phi_k)_{(0,y)}(x-t) \frac{dt\, dy}{y}$$

$$g_2(x) \equiv \sum_1^n \int_{\{(t,y):\ 2^{j-1}\ell(Q_0)\leq y<\ell(Q_0)\}} (f * y^{-d} (\Psi_k)_{(0,y)}(t))\, y^{-d} (\phi_k)_{(0,y)}(x-t) \frac{dt\, dy}{y}$$

$$g_3(x) \equiv \sum_1^n \int_{\{(t,y):\ \ell(Q_0)\leq y\leq 3\ell(Q_0)\}} (f * y^{-d} (\Psi_k)_{(0,y)}(t))\, y^{-d} (\phi_k)_{(0,y)}(x-t) \frac{dt\, dy}{y}$$

$$g_4(x) \equiv \sum_1^n \int_{\{(t,y):\ y>3\ell(Q_0)\}} (f * y^{-d} (\Psi_k)_{(0,y)}(t))\, y^{-d} (\phi_k)_{(0,y)}(x-t) \frac{dt\, dy}{y}.$$

Lemmas 3.3 and 3.5 imply that the integrals on the right-hand sides all converge absolutely. By Lemma 3.5, g_4 is pointwise bounded by $C|Q_0|^{-1} \int |f|\, dx \leq C$ for $x \in Q_0$, and it is easy to see that the same bound holds for g_3. Since $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$ and $|\nabla f| \leq C(2^j \ell(Q_0))^{-1}$ pointwise, Lemma 3.3 implies that $|g_1|$ is bounded by an absolute constant in Q_0. Thus, for $x \in Q_0$, we may write $f = g_2 + G$, where $|G| \leq C$, and C does not depend on Q_0 or j.

By Lemma 3.1, there is an R such that, for every $1 \leq k \leq n$, every T-sequence ζ, and every finite sequence $\{\lambda_Q\}_{\mathcal{D}} \subset \mathbf{C}$,

$$\int \left| \sum_{\mathcal{D}} \lambda_Q \frac{(\phi_k)_{\zeta(Q)}}{\sqrt{\mu(Q)}} \right|^2 d\mu \leq R \sum_{\mathcal{D}} |\lambda_Q|^2.$$

We claim that

$$\int_{Q_0} |g_2|^2\, d\mu \leq CRL(j) \mu(Q_0) \tag{30}$$

for a constant C depending on μ and d, but not on Q_0 or j. Since $\int_{Q_0} |G|^2\, d\mu \leq C\mu(Q_0)$, proving (30) will finish the proof.

There exist $N = N(d)$ dyadic cubes $\{Q_i\}_1^N$, congruent to Q_0, such that $\overline{Q_i} \cap \overline{Q_0} \neq \emptyset$. If $x \in Q_0$ then the support restriction on the ϕ_k's implies that

$$g_2(x) = \sum_{k=1}^{n} \int_{\{(t,y):\ t \in \cup_0^N Q_i,\ 2^{j-1}\ell(Q_0) \le y < \ell(Q_0)\}} (f * y^{-d}(\Psi_k)_{(0,y)}(t))\, y^{-d}(\phi_k)_{(0,y)}(x-t)\, \frac{dt\,dy}{y}.$$

For each $0 \le i \le N$ and $1 \le k \le n$, define

$$\gamma_{i,k}(x) \equiv \int_{\{(t,y):\ t \in Q_i,\ 2^{j-1}\ell(Q_0) \le y < \ell(Q_0)\}} (f * y^{-d}(\Psi_k)_{(0,y)}(t))\, y^{-d}(\phi_k)_{(0,y)}(x-t)\, \frac{dt\,dy}{y}.$$

Inequality (30) will follow once we show

$$\int |\gamma_{i,k}|^2\, d\mu \le CRL(j)\mu(Q_i), \tag{31}$$

because μ's doubling property implies $\mu(Q_i) \le C\mu(Q_0)$.

For $0 \le i \le N$, we define $\mathcal{F}_i$ to be the (finite!) family of dyadic subcubes Q of Q_i such that $2^j\ell(Q_i) \le \ell(Q) \le \ell(Q_i)$. We can then write:

$$\gamma_{i,k}(x) = \sum_{Q \in \mathcal{F}_i} \int_{T(Q)} (f * y^{-d}(\Psi_k)_{(0,y)}(t))\, y^{-d}(\phi_k)_{(0,y)}(x-t)\, \frac{dt\,dy}{y}.$$

We rewrite the last equation as

$$\gamma_{i,k}(x) = \int_{T(V_0)} \sum_{Q \in \mathcal{F}_i} \left[\left(|Q|^{-1}\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}} \rangle \sqrt{\mu(Q)} \right) \frac{(\phi_k)_{\sigma(Q,\tau,\eta)}(x)}{\sqrt{\mu(Q)}} \right] \eta^{-2d}\, \frac{d\tau\, d\eta}{\eta}.$$

For each $(\tau, \eta) \in T(V_0)$,

$$\int \left| \sum_{Q \in \mathcal{F}_i} \left[|Q|^{-1}\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}} \rangle \sqrt{\mu(Q)} \right] \frac{(\phi_k)_{\sigma(Q,\tau,\eta)}(x)}{\sqrt{\mu(Q)}} \right|^2 d\mu$$

is less than or equal to R times

$$\sum_{Q \in \mathcal{F}_i} \left| |Q|^{-1}\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}} \rangle \sqrt{\mu(Q)} \right|^2 = \sum_{Q \in \mathcal{F}_i} \left(\frac{|\langle f, (\Psi_k)_{\sigma(Q,\tau,\eta)} \rangle|^2}{|Q|^2} \right) \mu(Q).$$

Thus, by the generalized Minkowski inequality,

$$\left(\int |\gamma_{i,k}|^2\, d\mu \right)^{1/2} \le R^{1/2} \int_{T(V_0)} \left(\sum_{Q \in \mathcal{F}_i} \left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}} \rangle|^2}{|Q|^2} \right) \mu(Q) \right)^{1/2} \eta^{-2d}\, \frac{d\tau\, d\eta}{\eta}.$$

But $(T(V_0), \eta^{-2d}\,\frac{d\tau\,d\eta}{\eta})$ is a finite measure space (with a total measure only depending on d); therefore,

$$\int_{T(V_0)}\left(\sum_{Q\in\mathcal{F}_i}\left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle|^2}{|Q|^2}\right)\mu(Q)\right)^{1/2}\eta^{-2d}\,\frac{d\tau\,d\eta}{\eta}$$

is less than or equal to a dimensional constant times

$$\left(\sum_{Q\in\mathcal{F}_i}\int_{T(V_0)}\left[\left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle|^2}{|Q|^2}\right)\mu(Q)\right]\eta^{-2d}\,\frac{d\tau\,d\eta}{\eta}\right)^{1/2};$$

which implies that

$$\begin{aligned}\int|\gamma_{i,k}|^2\,d\mu &\le CR\sum_{Q\in\mathcal{F}_i}\int_{T(V_0)}\left[\left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle|^2}{|Q|^2}\right)\mu(Q)\right]\eta^{-2d}\,\frac{d\tau\,d\eta}{\eta}\\ &= CR\sum_{Q\in\mathcal{F}_i}\left(\int_{T(V_0)}\left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle|^2}{|Q|^2}\right)\eta^{-2d}\,\frac{d\tau\,d\eta}{\eta}\right)\mu(Q).\end{aligned}$$

But, for each $Q\in\mathcal{F}_i$, by the change of variables formula (24),

$$\int_{T(V_0)}\left(\frac{|\langle f, \overline{(\Psi_k)_{\sigma(Q,\tau,\eta)}}\rangle|^2}{|Q|^2}\right)\eta^{-2d}\,\frac{d\tau\,d\eta}{\eta} = |Q|^{-1}\int_{T(Q)}|f*y^{-d}(\Psi_k)_{(0,y)}(t)|^2\,\frac{dt\,dy}{y};$$

and, because $f\in BMO$, with $\|f\|_*\le C$, the sequence defined by

$$c_{Q,i}\equiv|Q|^{-1}\int_{T(Q)}|f*y^{-d}(\Psi_k)_{(0,y)}(t)|^2\,\frac{dt\,dy}{y}$$

is a bounded multiple of a Carleson sequence. By our definition of $L(j)$,

$$\sum_{Q\in\mathcal{F}_i}c_{Q,i}\mu(Q)\le CL(j)\mu(Q_i)$$

(because all of the Q's occurring in the sum satisfy $\ell(Q)\ge 2^j\ell(Q_i)$ and are contained in Q_i). Therefore

$$\begin{aligned}\int |\gamma_{i,k}|^2\, d\mu &\leq CR \sum_{Q\in\mathcal{F}_i} \left(|Q|^{-1} \int_{T(Q)} |f * y^{-d}(\Psi_k)_{(0,y)}(t)|^2 \,\frac{dt\,dy}{y}\right) \mu(Q)\\ &= CR \sum_{Q\in\mathcal{F}_i} c_{Q,i}\mu(Q)\\ &\leq CRL(j)\mu(Q_i),\end{aligned}$$

finishing the proof of Theorem 1.1. ♣

We present an easy corollary of Theorem 1.1. We first note that, by duality, if $\{\psi_k\}_k \subset L^2(\nu)$ satisfies (2), then, for all $f \in L^2(\nu)$,

$$\sum_k |\langle f, \psi_k\rangle_\nu|^2 \leq R \int |f|^2\, d\nu \tag{32}$$

(where we use $\langle\cdot,\cdot\rangle_\nu$ to denote the inner product in $L^2(\nu)$); and, conversely, if $\{\psi_k\}_k \subset L^2(\nu)$ satisfies (32), it satisfies (2).

In [9] the author looked at linear operators of the form

$$\sum_{\mathcal{D}} \frac{\langle f, \psi^{(Q)}_{\zeta(Q)}\rangle_\nu}{\nu(Q)} \phi^{(Q)}_{\zeta'(Q)}(x), \tag{33}$$

for a doubling measure ν, sequences of functions $\{\psi^{(Q)}\}_{\mathcal{D}}$ and $\{\phi^{(Q)}\}_{\mathcal{D}}$ in $\mathcal{C}_\alpha$, and T-sequences ζ and ζ'. One can think of (33) as a simple model for a wavelet representation of a Calderón-Zygmund singular integral operator (see [5] and references cited there). By Littlewood-Paley theory, if the $\psi^{(Q)}$'s and $\phi^{(Q)}$'s lie in $\mathcal{C}_{\alpha,0}$ and $\nu \in A_\infty$ then (33) defines a bounded linear operator on $L^2(\nu)$ in the following sense: If $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots$ is any increasing sequence of finite subsets of $\mathcal{D}$ such that $\mathcal{D} = \cup_i \mathcal{F}_i$ then, for all $f \in L^2(\nu)$,

$$T(f)(x) \equiv \lim_{i\to\infty} \sum_{Q\in\mathcal{F}_i} \frac{\langle f, \psi^{(Q)}_{\zeta(Q)}\rangle_\nu}{\nu(Q)} \phi^{(Q)}_{\zeta'(Q)}(x) \tag{34}$$

exists in $L^2(\nu)$ and $\|T(f)\|_{L^2(\nu)} \leq C(\nu,\alpha)\|f\|_{L^2(\nu)}$.[1] We present a partial converse:

Corollary 4.1 *Suppose that μ is doubling. Let $\{\phi_k\}_1^n \subset L^\infty(B(0;1))$ satisfy CNDC and suppose that, for each $1 \leq k \leq n$ and each T-sequence ζ, the series*

$$\sum_{\mathcal{D}} \frac{\langle f, (\phi_k)_{\zeta(Q)}\rangle_\mu}{\mu(Q)} (\phi_k)_{\zeta(Q)}(x), \tag{35}$$

defined as in (34), yields an $L^2(\mu)$ bounded linear operator. Then $\mu \in A_\infty$.

[1] This also holds in $L^p(\nu)$, $1 < p < \infty$, and the cancelation hypotheses can be weakened [9].

Proof Call the operator defined by (35) T. If T is $L^2(\mu)$ bounded then $|\int T(f)\bar{f}\,d\mu| \le C\int |f|^2\,d\mu$ for all $f \in L^2(\mu)$. But

$$\int T(f)\bar{f}\,d\mu = \sum_{\mathcal{D}} \frac{|\langle f, (\phi_k)_{\zeta(Q)}\rangle_\mu|^2}{\mu(Q)}.$$

Therefore, by the converse to (32), (6) is almost-orthogonal in $L^2(\mu)$. QED. ♣

Remark We believe the most natural application of Corollary 4.1 is this. Let $\psi \in \mathcal{C}_{\alpha,0}$ be real, radial, and non-trivial. If μ is doubling and the series

$$\sum_{\mathcal{D}} \frac{\langle f, \psi_{\zeta(Q)}\rangle_\mu}{\mu(Q)} \psi_{\zeta(Q)}(x)$$

(with the sum defined as above) gives an $L^2(\mu)$ bounded operator for every T-sequence ζ, then $\mu \in A_\infty$.

Acknowledgements We are grateful to the referee for spotting a gap in the proof of Lemma 2.1 and for valuable suggestions on improving the paper's exposition.

References

1. R.R. Coifman, C.L. Fefferman, Weighted norm inequalities for maximal functions and singular integrals. Stud. Math. **51**, 241–250 (1974)
2. I. Daubechies, *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61 (Society for Industrial and Applied Mathematics, Philadelphia, 1992)
3. M. Frazier, B. Jawerth, G. Weiss, *Littlewood-Paley Theory and the Study of Function Spaces*. CBMS-NSF Regional Conference Series in Mathematics, vol. 79 (American Mathematical Society, Providence, 1991)
4. E.M. Stein, *Harmonic Analysis* (Princeton University Press, Princeton, 1993)
5. P. Tchamitchian, Wavelets and differential operators, in *Different Perspectives on Wavelets*, ed. by I. Daubechies (American Mathematical Society, Providence 1993), pp. 77–88
6. A. Uchiyama, A constructive proof of the Fefferman-Stein decomposition of $BMO(\mathbf{R}^n)$. Acta Math. **148**, 215–241 (1982)
7. M. Wilson, Weighted norm inequalities for the continuous square function. Trans. Am. Math. Soc. **314**, 661–692 (1989)
8. M. Wilson, How fast and in what sense(s) does the Calderón reproducing formula converge? J. Fourier Anal. Appl. **16**, 768–785 (2010)
9. M. Wilson, Invariance and stability of almost-orthogonal systems. Trans. Am. Math. Soc. **368**, 2515–2546 (2016)
10. M. Wilson, Almost-orthogonal systems and A_∞ measures (unpublished preprint)
11. M. Wilson, *Weighted Littlewood-Paley Theory and Exponential-Square Integrability*. Lecture Notes in Mathematics, vol. 1924 (Springer, New York, 2007)