ERC Synergy Grant 2025
Research Proposal [Part B2]

Christoph Thiele (cPI), Rheinische Friedrich-Wilhelms-Universität Bonn (cHI)
Floris van Doorn (PI 2), Rheinische Friedrich-Wilhelms-Universität Bonn (HI 2)

Part B2: ***The scientific proposal***

Section a. State-of-the-art and objectives

The formalization of mathematics, often situated within computer science departments, has evolved as a field that largely operates independently of traditional research mathematics. This separation has been challenged by recent developments, such as the Lean Carleson project, which showed that real-time formalization of research in harmonic analysis is now within reach. This development surprised both the harmonic analysis and Lean communities. Having witnessed firsthand the transformative potential of collaboration between harmonic analysis and formalization for both disciplines, we aim to leverage these groundbreaking synergies through HALF. Our goal is to accelerate the formalization of mathematical research and thereby open new pathways for computer-aided discovery.

We divide HALF into five areas, each with a major fundamental **challenge**. These challenges are not part of HALF's deliverables, but we aim to solve at least one of them. For each challenge, we will make substantial progress by addressing two **objectives**. As detailed in Section b, each objective will result in multiple publications and may evolve through new discoveries, mitigating the risk that our predictions are either too optimistic or too conservative.

a.1. Formalizing research mathematics in Lean

Just as chess consists of a board that represents the state of the game and players who make decisions respecting established rules, mathematics comprises a status quo that reflects past discoveries and researchers determining the next steps. AI, in the form of deep neural networks, presents itself as a new player in the field of mathematics. However, the vast status quo of mathematics is primarily documented too informally for reliable autonomous AI applications. Human effort is still needed to formalize this status quo in mathematics before we can fully harness AI for further discoveries.

Now is an opportune time to invest in formalization, as significant advancements in AI for finding and writing mathematical proofs promise rewards from this investment in the near future. In July 2024, Google DeepMind published a blog post about their program, AlphaProof [At24]. AlphaProof successfully solved multiple problems at the level of the International Mathematics Olympiad, which is a highly challenging competition for high school students. However, advancements of these technologies at the research level in mathematics are still lacking due to the much sparser training material available.

The proof assistant Lean serves as a mathematical counterpart to a chessboard. It rigorously accepts and checks coded mathematics while assisting with the encoding process. Since its launch in 2013, Lean has developed a wide ecosystem of users. Amazon Web Services utilizes Lean for software verification, notably for verifying cryptographic protocols. The company Harmonic aims to integrate large language models with Lean to advance the state of the art in AI.

In 2017, a community of mathematicians initiated a collaborative project to create a large-scale, unified open-source library of mathematical definitions and proofs in Lean, called `Mathlib` [mat20]. This library has been developed and maintained by mathematicians such as Jeremy Avigad, Kevin Buzzard, Sébastien Gouëzel, Patrick Massot, and van Doorn, and it serves as the standard mathematical library for Lean. In total, there are 27 maintainers overseeing design decisions and approving new contributions, supported by 21 additional reviewers. Among these reviewers are two postdocs at Bonn, María Inés de Frutos Fernández and Michael Rothgang, making Bonn the university with the largest group of Lean reviewers in the European Union. Van Doorn has formalized results in a variety of areas, including spectral sequences in algebraic topology, the independence of the continuum hypothesis from the Zermelo–Fraenkel axioms and the axiom of choice [HvD20], the existence and uniqueness of Haar measures in functional analysis [vD21], and Gromov's h-principle for open and ample differential relations in differential geometry [vDMN23]. Several theorems were formalized in Lean shortly after a preprint containing the proof was made public [DHL19, CT$^+$22, Met23, T$^+$23, Blo24, T$^+$24]. With a few exceptions, all of these results pertained to combinatorics, suggesting that `Mathlib` provides sufficient material to facilitate the formalization of some arguments in current research in combinatorics.

In 2023, Thiele and van Doorn initiated the Lean Carleson project to formalize a new theorem in harmonic analysis [BvDJ+24]. This theorem includes as a special case a renowned result from 1966 [Car66] by Lennart Carleson regarding the convergence of Fourier series, which had not been previously formalized. Carleson's theorem asserts that for almost every $x \in [0,1]$ one has

$$\lim_{N\to\infty} \sum_{n=-N}^{N(x)} \widehat{f}_n e^{2\pi i n x} = f(x)\,, \tag{1}$$

where $\widehat{f}_n$ is the n-th Fourier coefficients of a continuous one-periodic function f on $\mathbb{R}$. Notable alternative proofs of this theorem were provided in [Fef73] and [LT00]. More recent generalizations include, but are not limited to, results concerning maximal multiplier norms [DLTT08], variation norms [OST+12], polynomial phases [Lie20], polynomials in higher dimensions [ZK21], phase unwinding [Mna22], and Radon transforms [Bec24]. These modern approaches interpret Carleson's theorem as estimates for the Carleson operator, which is expressed using a principal value singular integral as

$$Cf(x) := \sup_N \Big| \int_{\mathbb{R}} f(x-y) e^{iNy} \frac{1}{y}\, dy \Big|\,. \tag{2}$$

The preprint [BvDJ+24] generalizes this operator to doubling metric measure spaces, which consist of a space X along with a metric ρ and a compatible measure μ. The modulation functions are selected from a suitably defined class Θ. The generalized Carleson operator is then expressed as

$$\sup_{\theta\in\Theta} \Big| \int_X f(x,y) e^{i\theta(y)} K(y)\, d\mu(y) \Big| \tag{3}$$

with K is a singular integral kernel. An axiomatic approach to both doubling metric measure spaces and the class of modulation functions makes this generalization particularly suitable for formalization.

In May 2024, the Lean Carleson project went public [vD+] to solicit volunteers from the Lean community for the coding of the formalization. This event was accompanied by a post from Terence Tao on his Mathstodon blog. Tao also helped alleviate programmers' concerns about formalizing an unpublished and unreviewed result, which could contain critical errors. Indeed, Tao assured:

> Thiele is one of the leading world experts in the subject and has not to my knowledge published anything with major errors in it.

By September 2024, a milestone was reached in which half of the lemmas in the blueprint [BvDJ+24] had a fully formalized proof in Lean. As this milestone was celebrated on the Zulip channel coordinating the project, Kevin Buzzard commented:

> This is amazing progress. I had no idea until recently that the project was going so quickly! To be honest, after having talked to Hairer about the nature of hard analysis, I was a bit scared that formalizing hard analysis might be a fair bit trickier than formalizing hard commutative algebra and that this project would be a great test to see if this were true.

The Lean Carleson project continues to make steady progress, although at a slower pace. Current estimates suggest completion in 2025. This slowdown is partly due to competition for volunteers from other formalization projects in the Lean community.

The timeline for the Lean Carleson project should be compared to the more than eight years it took from the initial posting of the related polynomial Carleson result by V. Lie until its publication [Lie20] in the Annals of Mathematics. While formalization currently requires about ten times more human resources, it is approaching a level where it can compete with the refereeing process.

Challenge I. *Shift the culture in mathematics to make it routine to accompany new results with computer verification prior to submission to a reviewing process.*

The Lean Carleson project shows that harmonic analysis is a good area to work on Challenge I. This is the first objective of HALF.

Objective 1. *Assemble a team of harmonic analysts and formalization experts who can generate new results in harmonic analysis and formalize these findings within a reasonable timeframe before submitting the research paper.*

The process of formalization also leads to new ideas in harmonic analysis by offering a structured approach to optimizing proofs and concepts. This occurred during the Lean Carleson project, where the overall theorem was improved twice due to feedback from the formalization process, resulting in a much more general statement that also encompasses a variant of the classical Carleson theorem in finite characteristic. The next objective is designed to utilize this effect.

Objective 2. *Formalize the paper [DST24] on quantitative norm convergence for ergodic averages for three commuting transformations.*

This technical paper allows for generalizations, one of which is stated as Objective 8 of HALF. We will address Objective 8 by utilizing the expertise gained from Objective 2.

a.2. Libraries and tactics for harmonic analysis

The formalization of mathematical theorems requires considerable effort. Key challenges include the extensive training needed to master proof assistants like Lean, the numerous prerequisites referenced in standard proofs that may not yet be formalized, and the many tedious but straightforward steps often omitted in conventional proofs. Our focus will be on addressing these issues in the field of harmonic analysis, with the following challenge as our ultimate goal.

Challenge II. *Reduce the effort needed to formalize a new theorem in harmonic analysis to be less than that required to find and prove it.*

As of October 2024, `Mathlib` has formalized over 170,000 lemmas and theorems with proofs, created 85,000 definitions, and consists of more than 1.5 million lines of code organized into over 5,000 files.All submissions to `Mathlib` undergo a review process to ensure they meet high-quality standards and that the lemmas are proven generally enough for reuse in various contexts. These standards also promote the unification of equivalent definitions and consistent notation and terminology throughout the library. `Mathlib` stands out as a coherent mathematical library for Lean, unlike other proof assistants such as Coq, which has multiple incompatible analysis libraries [Aff24, BLM15], and Isabelle, which features several libraries for algebra and category theory [BPL22, Section 2].

The `Mathlib` library encompasses nearly all topics found in a typical undergraduate curriculum, including abstract algebra, linear algebra, multivariable analysis, measure theory, and many advanced subjects. It includes significant material for harmonic analysis, such as Lebesgue L^p spaces, basic Fourier analysis theory like the Riemann–Lebesgue lemma and the Fourier inversion theorem, and integral inequalities like Hölder's inequality. Van Doorn has advanced the formalization of Sobolev spaces by proving the Sobolev inequality and has established the existence and uniqueness of Haar measures for locally compact groups. However, many commonly used concepts in harmonic analysis are still not included in `Mathlib`.

Objective 3. *Build a corpus of prerequisites for modern harmonic analysis and incorporate their formalization in* `Mathlib`*.*

In addition to recording mathematical data such as definitions, theorems, and proofs, the `Mathlib` library includes tactics—programs that assist in writing proofs by automating repetitive reasoning. Currently, tactics primarily utilize "good old-fashioned AI," which follows prescribed instructions to achieve specific goals.

One tactic in Lean is the `simp` tactic, which simplifies mathematical statements using over 30,000 tagged simplification rules from `Mathlib`. It employs an efficient data structure known as *discrimination trees* [SUdM20] to focus only on relevant rules for the statement being simplified. Introduced in Lean 4, discrimination trees significantly speed up tactics compared to Lean 3. An extension of `simp` is the `aesop` tactic, which performs limited reasoning on top of `simp` [LF23]. Another tactic is `ring`, used to simplify algebraic expressions based on commutative ring axioms, transforming expressions into a specified normal form. It provides a complete decision procedure for equations in a commutative ring, solving them by turning both sides to normal form. While `ring` works exhaustively with commutativity and associativity, it is not extensible like `simp`.

In formalization, identical mathematical concepts often arise in different contexts and notations. For instance, group laws may be presented in either additive or multiplicative forms. Van Doorn has developed the tactic `to_additive`, which converts lemmas about groups or monoids in multiplicative notation into their additive counterparts. This command maintains a dictionary mapping between these structures, enabling it to automatically translate the statements and proofs. It intelligently avoids translating multiplication in fixed groups like $\mathbb{N}$ and $\mathbb{R}$ into addition, preventing nonsensical statements. For example, the statement

$$x^{nm} = (x^n)^m, \tag{4}$$

where x is the element of some group and $n, m \in \mathbb{N}$, is translated into

$$nm \cdot x = m \cdot (n \cdot x)\,, \tag{5}$$

leaving the multiplication nm unchanged. Another common challenge in formalization is that natural numbers are defined as an *inductive type*, but they also exist as subsets of the integers, rationals, reals, and complex numbers, each with separate definitions. To facilitate transitions between these contexts in proofs, Van Doorn has created the tactic `lift`.

In HALF, we will identify repetitive tasks related to harmonic analysis and address them with tailored tactics, using specifically the expertise of van Doorn. These tactics will be applied to the formalization objectives of HALF.

Objective 4. *Develop tactics to automate routine steps that occur in proofs in harmonic analysis.*

Other proof assistants, especially Isabelle, also have powerful automation tools. In particular the Sledgehammer tool [BBP13] is popular, and an implementation in Lean which extends it to dependent type theory is currently under development.

a.3. Multilinear singular integrals

The Brascamp–Lieb inequalities are a fundamental class of multilinear inequalities of the form

$$\int_{\mathbb{R}^m} \Big[\prod_{j=1}^{n} f_j(\Pi_j x)\Big]\, dx \le C \prod_{j=1}^{n} \|f\|_{p_j} \tag{6}$$

with linear surjections Π_j and suitable Lebesgue-norm exponents p_j. These inequalities generalize classical cases known as the Hölder, Hausdorff–Young, and Loomis–Whitney inequalities. The conditions on the parameters Π_j and p_j required for the validity of a Brascamp–Lieb inequality are now well understood [BCCT08].

However, general criteria for the validity of *singular* Brascamp–Lieb inequalities are not yet known. Singular Brascamp–Lieb inequalities, as defined by Durcik and Thiele in [DT21], take the form

$$\int_{\mathbb{R}^m} \Big[\prod_{j=1}^{n} f_j(\Pi_j x)\Big]\, K(\Pi x) dx \le C \prod_{j=1}^{n} \|f\|_{p_j} \tag{7}$$

with a Calderón-Zygmund singular integral kernel K and a further linear surjection Π.

Even for given parameters n, m, and ranks of Π_j and Π, the relative positioning of the projections affects both the validity and, if known, the nature of the proof for the corresponding Brascamp–Lieb inequality. Algebraically, describing these relative positions of projections is a version of the $(n+1)$-subspace problem, which is tame for $n \le 3$ and wild for $n \ge 4$. Generally, the inequalities are more challenging to prove when the rank of Π is small relative to the sum of the ranks of the Π_j.

At the one extreme of largest possible rank of Π, the singular Brascamp–Lieb integrals can be identified with Coifman–Meyer paraproducts [MC97] and their estimates are well understood. At the other extreme, one finds the simplex Hilbert forms

$$\Lambda(f_1, \ldots, f_m) := \int_{\mathbb{R}^m} \Big[\prod_{j=1}^{m} f_j(x_1, \ldots, \hat{x}_j, \ldots x_m)\Big] \frac{1}{x_1 + \cdots + x_m} dx \tag{8}$$

with $\hat{x}_j$ denoting omission of the variable x_j. The conjecture is that simplex Hilbert forms satisfy Brascamp–Lieb inequalities with $p_j = \frac{1}{m}$ for all m. This conjecture is only known to hold in the trivial case $n = 1$ and in the classical case $n = 2$ of the Hilbert transform. The case $n = 3$, known as the triangular Hilbert form, remains the most fundamental open problem in the field.

Challenge III. *Prove for Λ defined in (8) with $m = 3$, some universal C, and all test tuples f_1, f_2, f_3,*

$$\Lambda(f_1, f_2, f_3) \leq \prod_{j=1}^{3} C\|f_j\|_3 \, . \tag{9}$$

A series of past results represents partial progress towards this conjecture. By specializing the three input functions to elementary tensors in suitable variables and integrating in one direction trivially, the conjectural bounds for the triangle Hilbert form imply bounds for the dual of the so-called bilinear Hilbert transform,

$$\int f_1(x + \alpha_1 t) f_1(x + \alpha_2 t,) f_1(x + \alpha_3 t) \frac{1}{t} \, dt \tag{10}$$

for a generic vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. Thus, the bounds on the bilinear Hilbert transform by Lacey and Thiele [LT97, LT99], and even more so the uniform bounds in the parameter α investigated in a series of papers [Thi02, GL04, UW22, FST24b] by Thiele and others, represent partial progress towards Challenge III. Specializing all three functions in a different way, using expressions such as

$$f(x, y) = f(x)e^{iN(x)y} \tag{11}$$

allows one to deduce estimates for Carleson's operator from the triangle Hilbert form. In [KTZK15] by Thiele and coauthors, bounds are provided for a dyadic model of the triangle Hilbert form when specializing only one function to an elementary tensor. The methods in [KTZK15] strongly suggest that any future solution of Challenge III will require estimates on singular Brascamp–Lieb forms with cubical structure as a tool.

Further progress towards Challenge III, from a different perspective, was achieved in [ZK17] and in [DKT19] by Thiele and coauthors, who estimated truncations of the singular kernel to a finite number N of scales. Since bounds for each scale are trivial, estimates of the truncated form with growth $O(N)$ follow by the triangle inequality, while estimates of the order $O(1)$ are equivalent to Challenge III. Any intermediate growth between these two orders is referred to as a cancellation estimate and represents progress toward the challenge. Building on ideas from [Tao16] and streamlining the approach, the paper [ZK17] proves cancellation estimates of the form $o(N)$, while [DKT19] establishes cancellation estimates with growth $O(N^{\frac{1}{2}})$. Both papers extend their results to the general simplex Hilbert forms, albeit with weaker growth estimates.

Roughly halfway between the above extremes of singular Brascamp–Lieb inequalities, in terms of a count of dimensions, is a remarkable class of forms known as cubical singular Brascamp–Lieb forms [DT20, DT21, DST22].

$$\int_{\mathbb{R}^{2k}} \prod_{j:\{1,\ldots,k\}\to\{0,1\}} f_j(x_{1,j(1)}, \ldots, x_{k,j(k)}) K(x_{1,0} - x_{1,1}, \ldots, x_{k,0} - x_{k,1}) dx, \tag{12}$$

where the product is taken over the set of functions from $1, \ldots, k$ to $0, 1$, which can be identified with the corners of a cube. Thanks to the symmetries of the cube and the techniques of Gowers box norms, there is a satisfactory theory of estimates for these cubical forms. The cubical forms have found applications in the norm convergence of multiple ergodic means [DKŠT19, DST24] and in Ramsey-type density problems [DK22].

A distinguishing feature of the cubical singular Brascamp–Lieb forms is that, despite the non-existence of linear projections on doubling metric measure spaces, one can formulate these forms on doubling metric measure spaces. Namely, the only algebraic operations in (12) are differences, which can be interpreted as distances. One objective is to extend the theory of cubical forms to the setting of doubling metric measure spaces.

Objective 5. *For a metric space X with doubling measure μ, prove a priori L^p estimates for cubical singular Brascamp–Lieb forms*

$$\int_{X^{2k}} \prod_{j:\{1,\ldots,k\}\to\{0,1\}} f_j(x_{1,j(1)}, \ldots, x_{k,j(k)}) K((x_{1,0}, \ldots x_{k,0}), (x_{1,1}, \ldots, k, 1)) d\mu^{2k}, \tag{13}$$

where $K : X^k \times X^k \to \mathbb{R}$ is a singular integral kernel in the sense of doubling metric measure spaces relative to the k fold Cartesian product of X with itself.

The cubic singular integrals do not require time-frequency analysis and thus do not imply bounds for objects such as the bilinear Hilbert transform or Carleson's operator. To make progress on multilinear singular integrals in the realm of time-frequency analysis, we propose:

Objective 6. *Prove bounds for the simplex Hilbert form with some additional structural properties on the input functions.*

Here, the first goal will be to prove the continuous analogue of the dyadic model for the triangle Hilbert form in [KTZK15]. The second step will be to identify suitable assumptions on the input functions for higher simplex Hilbert forms.

a.4. Convergence of ergodic averages

The Calderón transference principle allows for the use of estimates in harmonic analysis to prove quantitative results on ergodic means. Questions in ergodic theory thus provide motivation to investigate specific types of results in harmonic analysis, many of which are closely related to multilinear singular integrals. The relevant results in harmonic analysis also include estimates on generalizations of Carleson's operator. Such a generalization was used by Thiele and coauthors [DLTT08] to extend the parameter range for return times theorems.

Let X be a probability space, and let T be a measure-preserving transformation on X that we assume to be bijective for simplicity. The classical ergodic averages of a bounded measurable function f on X with respect to the transformation T are given by

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x). \tag{14}$$

A result by von Neumann shows the convergence of these averages in the sense of $L^2(X)$ as N tends to infinity, while Birkhoff's stronger theorem provides pointwise convergence almost everywhere. Both theorems come with improvements regarding variation norm bounds, as elaborated in [JOR96].

In [KMT22], a far-reaching generalization of Birkhoff's theorem has been referred to as the Furstenberg–Bergelson–Leibman conjecture, which was first promoted in person by Furstenberg and later published by Bergelson and Leibman in [BL02]. This conjecture concerns averages of the form

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{j=1}^{k} f_j(T_1^{p_{j,1}(n)}\dots T_d^{p_{j,d}(n)}x) \tag{15}$$

with k functions f_j on X, d transformations $T_1,\dots,T_d$ that span a nilpotent group of actions on X, and kd polynomials $p_{j,1},\dots,p_{j,d}$. The conjecture asks for pointwise almost everywhere convergence of the averages (15). Much progress has been made recently on these averages in the case where at least some of the polynomials are not linear. Such higher-degree polynomials allow for curvature-related tools in harmonic analysis; see, for example, [KMT22, IMMS23a, IMMS23b].

The case of all polynomials being linear is substantially different and, in fact, more difficult because it does not allow for any curvature-related methods. Even one of the most basic questions regarding linear polynomials remains a celebrated open problem.

Challenge IV. *Given two commuting transformations T_1 and T_2, prove pointwise almost everywhere convergence of the ergodic means*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T_1^n x) f_2(T_2^n x)\,. \tag{16}$$

Questions on averages (15) with only linear polynomials are closely related to multilinear Brascamp–Lieb integrals, as discussed in the previous section. Only a special case in Challenge IV, with T_1 and T_2 being powers of the same transformation, is known due to Bourgain [Bou90]. This result was improved quantitatively in [Lac00, DOP17] in work related to the bilinear Hilbert transform.

A more approachable problem arises when one averages with additional parameters. This was proposed in [DS18], where qualitative almost everywhere convergence for such extended averages for commuting transformations T_1 and T_2 was shown. We propose to generalize these results to more transformations while simultaneously establishing quantitative improvements. We state one interesting example as an objective.

Objective 7. *Let T_1, T_2, T_3 be commuting or span the Heisenberg group. Prove quantitative almost everywhere convergence results for ergodic averages of the form*

$$\frac{1}{N^4} \sum_{n_1,n_2,n_3,j=1}^{N} f_0(T_1^{n_1} T_2^{n_2} T_3^{n_3} x) f_1(T_1^{n_1+j} T_2^{n_2} T_3^{n_3} x) f_2(T_1^{n_1} T_2^{n_2+j} T_3^{n_3} x) f_3(T_1^{n_1} T_2^{n_2} T_3^{n_3+j} x) \,. \tag{17}$$

Reducing the number of averaging parameters in this last objective will result in more difficult problems and provide partial targets toward the goal of a single averaging parameter.

Returning to the ergodic averages with a single averaging parameter, as in the Furstenberg–Bergelson–Leibman conjecture, a simpler task than pointwise convergence is convergence in the Hilbert space norm. For two commuting transformations, such norm convergence results trace back to influential work in ergodic theory in the 1970s. In the case of more than two commuting transformations, this was pioneered by Host and Kra [HK05] for powers of a single transformation and by Tao [Tao08] for general commuting transformations, with more ergodic theoretic and structure theoretic proofs found in [Hos09, Aus10]. A far-reaching generalization to transformations spanning a nilpotent group was proved by Walsh [Wal12].

The question of stronger quantitative norm convergence bounds, such as variation norm bounds, was raised in [AR15]. For two and three commuting transformations, such quantitative norm convergence bounds were demonstrated by Thiele and coauthors in [DKŠT19, DST24]. We propose to generalize the latter to three transformations spanning the Heisenberg group, a specific instance of the theorem in [Wal12].

Objective 8. *Given three transformations T_1, T_2, T_3 that generate an action of the discrete Heisenberg group on a measure space X, prove that for any sequence $N_0 < N_1 < \cdots < N_J$ the following holds:*

$$\sum_{j=1}^{J} \|M_{N_j}(f_1, f_2, f_3) - M_{N_{j-1}}(f_1, f_2, f_3)\|_2^2 \le C J^{\frac{3}{4}} \prod_{i=1}^{3} \|f_i\|_\infty \,, \tag{18}$$

where

$$M_N(f_1, f_2, f_3)(x) = \frac{1}{N} \sum_{n=1}^{N} f_1(T_1^n x) f_2(T_2^n x) f_3(T_3^n x) \,. \tag{19}$$

The natural but challenging generalization to four or more commuting transformations of the result in [DST24] is project C06 by Thiele in a CRC grant proposal, albeit without formalization. Therefore, this will not be part of HALF if the CRC project is funded.

a.5 Nonlinear Fourier analysis

The partial Fourier integral of an integrable function f on the real line is given by

$$S(\xi, x) := \int_{-\infty}^{x} f(t) e^{-2\pi i t \xi} dt \,. \tag{20}$$

When x is equal to $-\infty$, this partial Fourier integral vanishes, while at $x = \infty$, it represents the actual Fourier transform. Taking the exponential of the partial Fourier integral (20),

$$G(\xi, x) := e^{S(\xi, x)} \,, \tag{21}$$

turns it into an infinitesimal product, which is described by the ordinary differential equation

$$\partial_x G(\xi, x) = G(\xi, x) f(x) e^{-2\pi i x \xi} \tag{22}$$

with the initial condition of being constant at 1 when x equals $-\infty$, and being the exponential of the Fourier transform at ∞. The multiplicative interpretation allows for matrix-valued generalizations. Since matrix multiplication is not commutative, these generalizations can no longer be expressed as the exponential of a linear Fourier integral, resulting in genuinely nonlinear Fourier transforms.

These matrix differential equations evolve on Lie groups and depend on a suitable embedding of the complex driving force in (22) into the Lie algebra. Prominent examples are the simple groups

$SU(2)$ and $SU(1,1)$, the latter of which is isomorphic to $Sl_2(\mathbb{R})$. These two notable models can be expressed as

$$\frac{d}{dx}\begin{pmatrix} a(\xi,x) & b(\xi,x) \\ \pm\bar{b}(\xi,x) & \bar{a}(\xi,x) \end{pmatrix} = \begin{pmatrix} a(\xi,x) & b(\xi,x) \\ \pm\bar{b}(\xi,x) & \bar{a}(\xi,x) \end{pmatrix}\begin{pmatrix} 0 & f(x)e^{-2\pi i x\xi} \\ \pm\overline{f}(x)e^{2\pi i x\xi} & 0 \end{pmatrix}, \tag{23}$$

with the upper sign corresponding to the $SU(1,1)$ model and the lower sign corresponding to the $SU(2)$ model. The basic analysis of the $SU(1,1)$ model is surveyed in the Park City lecture notes by Tao and Thiele [TT12], and that of the $SU(2)$ model is covered in the thesis of Thiele's former PhD student, Ya-Ju Tsai.

In first-order approximation, the function b becomes the partial Fourier integral. It has many properties analogous to those of the Fourier integral. In particular, it satisfies a nonlinear Plancherel identity, which in the $SU(1,1)$ model reads as

$$\int_{\mathbb{R}} \log(1+|b(\xi,\infty)|^2)\, d\xi = \int_{\mathbb{R}} |f(x)|^2\, dx\ . \tag{24}$$

Nonlinear analogues of estimates for functions f in L^1, such as the Riemann–Lebesgue theorem, are basic. In contrast, for f in L^p with $1<p<2$, nonlinear analogues of the Hausdorff–Young inequality and the Menshov–Paley–Zygmund inequality are much harder to prove and are established in [CK01], with an alternative proof provided by Thiele and coauthors in [OST+12]. The paper [MTT03b] by Thiele and coauthors demonstrates that the approach in [CK01] cannot be extended to the endpoint where $p=2$, and, in particular, cannot yield uniform bounds for the Hausdorff–Young inequality as p approaches 2.

The paper [MTT03a] by Thiele and coauthors proves a nonlinear analogue of Carleson's theorem in a model on fields with finite characteristic, while a series of papers culminating in [DMT17] by Thiele and coauthors proves an easier nonlinear analogue of Carleson's theorem for some nonlinear Fourier transforms on nilpotent Lie groups. The paper [DMT17] utilizes the theory of outer L^p spaces developed by Do and Thiele [DT15], which, in its fundamental nature, is a good target for formalization in Lean and incorporation into `Mathlib`. The question of a generalization of Carleson's theorem that would be strictly stronger than Carleson's original theorem remains open, despite the most recent progress by Poltoratski in [Pol24] on a weak form of a nonlinear Carleson's theorem.

Challenge V. *Generalize Carleson's theorem to the nonlinear Fourier transforms* (23)*.*

The differential equation (23) has a discrete analogue in the form of the recursion equation

$$\begin{pmatrix} a(z,n) & b(z,n) \\ \pm\bar{b}(z,n) & \bar{a}(z,n) \end{pmatrix} = \begin{pmatrix} a(z,n-1) & b(z,n-1) \\ \pm\bar{b}(z,n-1) & \bar{a}(z,n-1) \end{pmatrix}\frac{1}{\sqrt{1\mp|f(n)|^2}}\begin{pmatrix} 1 & f(n)z^n \\ \pm\overline{f(n)z^n} & 0 \end{pmatrix} \tag{25}$$

for summable sequences f, producing Lie group-valued functions on the unit circle. This is a nonlinear analogue of Fourier series. A reordering of the product, analogous to the repeated application of the distributive law to the Fourier series,

$$a_nz^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_0 = (\ldots((a_nz+a_{n-1})z+a_{n-2})z+\ldots)z+a_0\,, \tag{26}$$

essentially turns the recursion (25) into a product ordered by increasing n. In the $SU(2)$ case, it is:

$$\prod_n \begin{pmatrix} z^{\frac{1}{2}} & 0 \\ 0 & z^{-\frac{1}{2}} \end{pmatrix}\frac{1}{\sqrt{1+|f(n)|^2}}\begin{pmatrix} 1 & f(n) \\ -\overline{f(n)} & 0 \end{pmatrix}. \tag{27}$$

This product, after changing to variables typical for quantum computing, identifies the nonlinear Fourier series with an important algorithm called quantum signal processing [DLNW22]. The algorithm alternates the application of gates taken from a sequence of tuning parameters, here f_n, with the application of a fixed gate depending on the input signal, here z, in order to compute the desired output signal, here $b(z)$. This recent and surprising observation by Thiele and coauthors [AMT24] that quantum signal processing and the $SU(2)$ model of nonlinear Fourier series are essentially the same has led to a transfer of ideas between harmonic analysis and quantum computing. It resulted in the first provably stable algorithm [ALM+24] by Thiele and coauthors to compute the tuning parameters of quantum signal processing in polynomial time for general signals consistent with (24). This algorithm was accelerated by Ni and Ying [NY24] using fast Toeplitz solvers to quadratic time in the number of parameters.

The focus on the signal $b(z)$ in quantum computing allows for the artificial selection of $a(z)$ as an outer function. This choice gives the unique set of minimal tuning parameters. Additionally, this choice gives rise to new types of questions regarding nonlinear Fourier series.

Objective 9. *Understand analytic properties of the nonlinear Fourier transform under an outerness assumption on a.*

Higher-dimensional variants of quantum signal processing are of interest, both in terms of the number of input parameters [RC22] and in the dimension of the output variable [GSLW19].

Objective 10. *Develop a higher-dimensional theory of the nonlinear Fourier transform and relate it to quantum signal processing.*

The discovery of the connection between nonlinear Fourier analysis and quantum signal processing has inspired an Oberwolfach Arbeitsgemeinschaft in October 2024, specifically devoted to this connection and bringing together many young researchers from harmonic analysis and quantum computing.

Section b. Methodology

b.1. Formalizing research mathematics in Lean

Objective 1 of HALF is to build a team that combines harmonic analysts led by Thiele with formalization experts led by van Doorn. This team will collaborate closely, as we will describe in this section. Each of the Objectives 5–10 will be implemented in four stages. These stages have been tested in the pilot Lean Carleson project, showcasing the synergistic nature of the HALF collaboration. As of October 2024, Stage 3 of the pilot project is over halfway completed, and materials for Stage 4 are taking shape. Each stage will include its own dissemination materials, potentially with different lists of authors.

Stage 1, Prove: The expected theorem in harmonic analysis will be proven and elaborated upon in a standard research paper. This stage will be primarily led by experts in harmonic analysis, although the insights gained from formalizing previous results will be beneficial. For instance, the experience with doubling metric measure spaces—a key element of the Lean Carleson project—will be useful for Objectives 5, 6, 7, and 8. The research paper will be published through conventional channels for harmonic analysis, starting with a preprint on arXiv for open access, followed by publication in a peer-reviewed journal. However, the journal publication will be postponed until after the completion of the computer verification, as one of the goals of HALF is to shift the responsibility of checking for correctness from the standard referee to in-house formalization. The reviewer in harmonic analysis will only need to assess the significance and relevance of the results.

Stage 2, Blueprint: The harmonic analysis team produces a second document known as a blueprint. This blueprint provides a more detailed proof of the same result, maintaining standard mathematical language. It breaks down the formalization into individual lemmas that very explicitly state their assumptions and conclusions, allowing contributors to work on each part independently. The logical dependencies among the lemmas are illustrated in another document called the dependency graph, which is generated by a computer tool developed by Patrick Massot.

Both the blueprint and the dependency graph are not typically found in traditional mathematical research, but have proven themselves to be useful tools within the standard mathematics community. In the Lean Carleson project, the blueprint spans approximately 150 pages, compared to about 30 pages for the standard proof, while the dependency graph contains nearly 200 individual lemmas.

Close cooperation between the harmonic analysis and formalization teams is essential for adjusting the blueprint to align with the design decisions of `Mathlib`. The best practices learned from the blueprint will be an output of HALF. The blueprint will first be published on the open-access server arXiv and subsequently in appropriate peer-reviewed outlets, such as the newly established journal Annals of Formalized Mathematics.

Stage 3, Formalize: To formalize the blueprint, the first step is to write the definitions and statements for all theorems and lemmas in Lean. The choice of formulation of these definitions significantly impacts the difficulty of formalizing the proofs, so this step should be managed by an expert team leader, either van Doorn or an experienced postdoc. This leader must have advanced skills, and it is essential for HALF that van Doorn is a world-leading expert in this field.

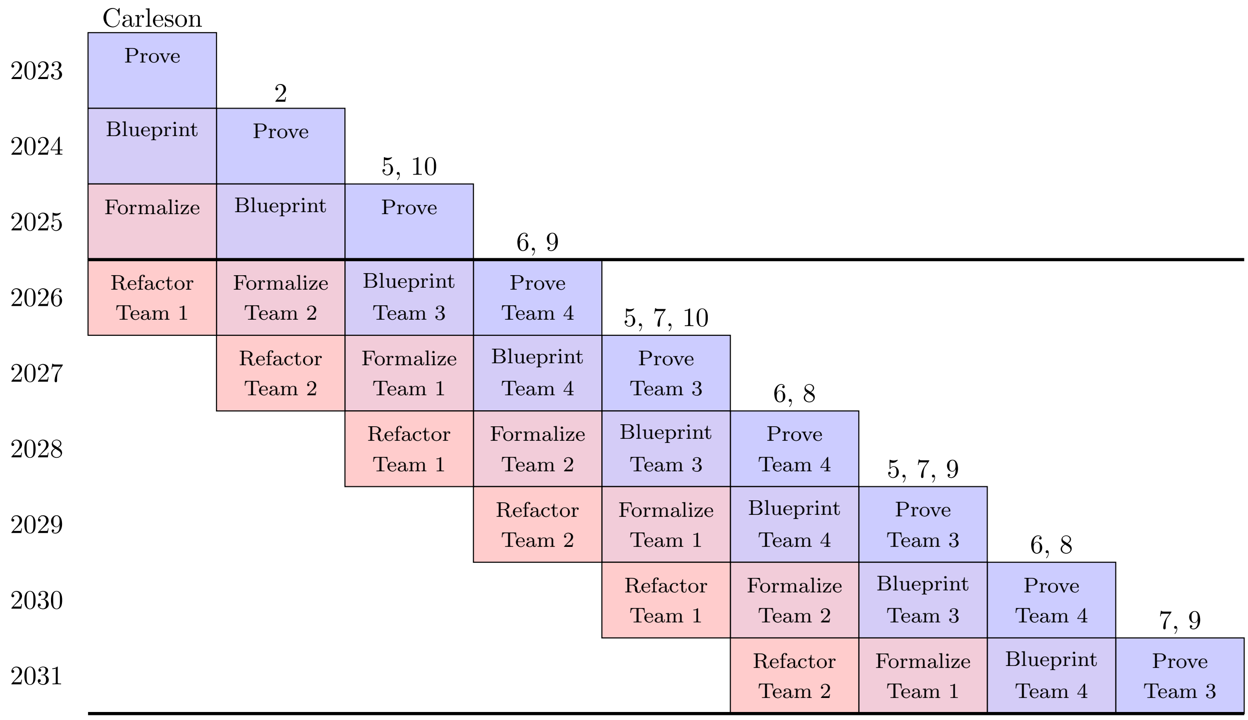
Once the definitions are established, the proof of each lemma or theorem can be addressed by the rest of the team. Task assignments and coordination among the various team members, including research student assistants, occur through GitHub projects and a Zulip chat channel, with the ongoing formalization shared via GitHub. Throughout this process, numerous questions will inevitably arise, including mathematical ones related to minor inaccuracies in the blueprint. Some of these can be answered by van Doorn, while others will need to be directed to the harmonic analysis group, requiring prompt collaboration to maintain momentum. For the Lean Carleson project, our group took great care to output a high quality blueprint, enabling local corrections without necessitating major revisions.

The outcome of this stage will be open-source code available through GitHub and the database `bonndata` established by the University of Bonn.

Stage 4, Refactor: Code developed for one of the harmonic analysis objectives in HALF does not need to adhere to the usability and generality coding standards required for inclusion in the `Mathlib` library. For the purpose of verifying research results, it is sufficient for the code to compile. This is crucial for optimizing human resource scalability within HALF.

The final stage will involve incorporating new material that can be reused in future formalization projects into `Mathlib`. During this stage, we will leverage the experience gained from formalization in Stage 3 to identify material that can be rewritten to meet higher coding standards and pinpoint tasks that could be automated through a tactic, subsequently developing those tactics. This stage will be reflected in the `Mathlib` library, and we will also write a paper discussing the new theories, tactics, and insights we have gathered throughout the formalization process.

The anticipated global timeline for HALF's work is shown in the figure below. Later phases may be adjusted based on new developments and insights.



To optimize lab processes, we will work concurrently on all four stages outlined above, with each year's work shown in a row. The columns represent one or more mathematical results progressing through these stages.

HALF's scientific staff will be organized into four teams, each typically composed of a postdoc and a PhD student. Teams 1 and 2 will focus on formalization under van Doorn's supervision, while Teams 3 and 4 will specialize in harmonic analysis under Thiele's guidance. Additionally, four part-time student research assistants will work on formalizing the objective currently in Stage 3.

The objectives listed in the figure outline those to be explored in Stage 1. When an objective appears multiple times, it denotes distinct natural steps, as explained in the methodology sections.

In Stage 2, we will select particularly promising results for formalization and draft blueprints. In the final years of HALF, as `Mathlib` expands, we may produce multiple blueprints for simultaneous formalization. This formalization will occur in Stage 3, where external Lean community volunteers may still contribute, though our reliance on them will decrease over time. In Stage 4, we will integrate reusable results into `Mathlib` and develop relevant tactics.

This workflow is tailored to the needs and availability of HALF's early-career personnel. Each PhD student and postdoc will engage in both traditional work and HALF's novel research approach. Given typical appointment durations, the staged approach outlined here provides a suitable rhythm for their involvement. Ensuring each stage has adequate staffing is crucial for success.

Objective 1 requires expertise to be housed in Bonn to conduct the formalization previously managed by a global community in the Lean Carleson project. We have started training young researchers in this area, with Van Doorn offering an annual practical course on Lean for bachelor and master students. This initiative has already led to significant contributions from a bachelor student, Leo Diedering, to the Lean Carleson project. Beginning in Fall 2025, Thiele and van Doorn will co-teach the Analysis sequence for incoming Bachelor students, which will include a small Lean component based on successful teaching projects worldwide. These efforts will ensure the availability of necessary human resources for HALF. The risk of temporary shortage can be mitigated by temporary shifts in the vertical structure such as replacing a PhD student by a postdoc or vive versa.

Objective 2 is a formalization objective of a result that has already appeared on arXiv, so Stage 1 is already completed. The writing of the blueprint, Stage 2 for this objective, is underway and should be ready for Stage 3 at the start of HALF.

b.2. Libraries and tactics for harmonic analysis

The formalization of mathematical results in Stage 3 will result in identification of material for the `Mathlib` library, Objective 3 and for tactics to be developed, Objective 4. These two objectives are the focus of Stage 4.

The Lean Carleson project already suggests the first step towards Objective 3. Most prominently, the Hardy–Littlewood maximal function and its boundedness in L^p spaces is used in almost all individual chapters of the blueprint, and will be incorporated into `Mathlib`. Furthermore, real and complex interpolation theory were developed as part of the Carleson project, and will also be ported to `Mathlib` and thus be made generally reusable. Finally, Calderón-Zygmund theory as for example in the first four chapters of Stein's book [Ste93] and done in the generality of doubling metric measure spaces presents suitable material for `Mathlib`.

The next formalization goal after Lean Carleson, Objective 2, suggests foremost the Calderón transference principle as a suitable item for `Mathlib`. In fact, being a principle rather than a specific theorem may lend itself to a combination of theorems and tactics to be used more generally.

Nonlinear Fourier analysis is a fundamental theory that is used in many areas of mathematics, such as bounded analytic functions, Riemann–Hilbert problems, orthogonal polynomials, scattering theory, operator theory and quantum computing. Including this theory into `Mathlib` will be beneficial for all of these applications. Some complex mathematical tools as presented in Garnett's book [Gar07] will be among the first targets, such as Hardy spaces and factorization of bounded analytic functions into inner and outer factors.

Turning to Objective 4, we plan to write tactics using so-called "good old-fashioned AI": programs that have a well-defined scope and will solve problems in that scope by explicit human-written algorithms. These tools can be quick and reliable, which is important for formalization.

In `Mathlib` there are tactics that can automatically prove that functions are measurable, continuous or differentiable by decomposing the function into simpler components. This is crucial in harmonic analysis, because they often show up as side conditions when applying theorems. These tactics are still limited, and often not directly applicable in actual formalizations. A first step of Objective 4 is to extend these tactics to apply in all common use cases and to have a similar tactic that can deal with the integrability of functions. Moreover, we want to extend the integrability-tactic to encompass reasoning that goes beyond merely decomposing the function into simpler components, allowing us to consider the behavior of the entire function as well. A common feature of many proofs in harmonic analysis is the occurrence of a constant in estimation whose specific value is often not interesting for the purpose of the estimation. For the proof, the exact interdependence of these constants is important, but usually dealt with very implicitly in standard mathematical papers. We will work on an improved implementation of the Landau asymptotic notation as well as a tactic that easily handles the dependencies in this asymptotic notation. The library `Mathlib` can already work with statements of the form $f = O(g)$, but in mathematics one often wants to write this asymptotic notation inside a formula, like $f = \exp(O(g)) + O(h)$, which `Mathlib` cannot conveniently express yet.

We will additionally work on a suite of tactics that help with manipulating finite sums, infinite sums and integrals. We currently do calculations involving these by invoking specific lemmas from the library. For example, we can invoke the specific lemma that we can reindex a sum or integral, which requires proving that the function is a bijection or a measure-preserving equivalence. Instead, we will write a tactic `reindex` that will automatically find the correct function with which to reindex and will prove these side conditions. Similarly, we will write tactics that push operations (like addition and multiplying by a constant) outside the sum or integral, or reorder multiple sums or integrals.

Another target is a tactic that can automatically apply common integral inequalities, such as the Cauchy–Schwarz, Hölder, Minkowski, and Jensen's inequalities.

We will identify more objectives that we want to solve automatically by these tactics by comparing the paper proofs with the proofs in Lean and looking for proofs that are significantly longer in the formal proof because of routine proof steps that are omitted on paper.

During HALF, we will not directly develop any neural AI. However, the tactics we write will synergize with the continued development of neural AI. Neural AI, like AlphaProof, learns to write mathematical proofs by combining tactics, and expanding the capabilities of such tactics will also expand the capabilities of neural AI.

While neural AI cannot solve many mathematical problems autonomously, they can already be used to solve certain simple problems. An example is Github Copilot [Fri21], which can automatically give code suggestions while you are typing. During the project we will carefully track the progress in the field of neural AI, and we will incorporate state-of-the-art neural AI tools in our workflow when they become available and useful for us.

b.3. Multilinear singular integrals

Many existing results for multilinear singular integrals come in pairs: one for a dyadic model that operates in a field of characteristic two, and one for a continuous model in Euclidean space. The Lean Carleson project [BvDJ+24] introduces a new framework within doubling metric measure spaces. Indeed, this new setting encompasses both of the previous cases as special instances. Therefore, [BvDJ+24] is the first theorem to unify Carleson's theorem in the classical continuous setting [Car66] and the dyadic setting [Bil67].

Work on Objective 5 will utilize both the dyadic work [KT13] by Kovač and Thiele and the continuous work [DST22] by Thiele and his coauthors as a basis for generalizations to doubling metric measure spaces. However, there are questions that must be addressed in the full generality of doubling metric measure spaces, and neither the dyadic nor the Euclidean setting directly provides the answers.

One pair of ingredients in the theory of cubical Brascamp–Lieb forms that is difficult to extend to doubling metric measure spaces is the use of exact martingale identities in the dyadic setting and, correspondingly, the use of heat kernels, specifically Gaussians, in the Euclidean setting. The attempt to use diffusion semigroups on doubling metric measure spaces as substitutes for Gaussian heat kernels appears to fall short. Although such diffusion semigroups have been extensively studied [AMS19] in recent years, they typically require additional assumptions on the doubling metric measure space, such as some mild form of differentiability structure. Adding such differentiability conditions is one approach toward Objective 5, but we aspire to work in full generality. Thus, we propose to work with compositions of operators of the form

$$P_t = \prod_{2^k < t} A_k \tag{28}$$

with averaging operators A_k at scale 2^k in place of diffusion semigroups.

A second difficulty in doubling metric measure spaces is the absence of a translation structure, which is heavily utilized in the dyadic and Euclidean settings. We will require assumptions on singular kernels that are independent of any translation structure. Our objective is to establish $T(1)$ and $T(b)$ type testing assumptions, which come in various forms with increasing levels of generality. Part of the work on Objective 5 will be to clarify the theory of $T(1)$ and $T(b)$ theorems for singular Brascamp–Lieb forms. Multilinear $T(1)$ theorems have been studied in the dyadic setting of cubical Brascamp–Lieb forms [KT13], but corresponding theorems in the Euclidean setting are still lacking and surprisingly challenging. We aim to achieve such generalizations. In the more complex realm of multilinear $T(b)$ theorems, there remain open questions in the Euclidean setting even for Coifman–Meyer paraproducts, such as a Euclidean analogue of the dyadic result [MT17] by Mirek and Thiele. We aspire to establish such an analogue and further generalizations to doubling metric measure spaces. The exact formulation of the results in Objective 5 may be adjusted based on the findings from the investigations into $T(1)$ and $T(b)$ theorems.

The first step towards Objective 6 also involves extending a dyadic result [KTZK15] to the Euclidean setting. This dyadic result provides a roadmap towards the objective. Unlike Objective 5, Objective 6 requires the use of time-frequency analysis, as the proposed result is strong enough to yield uniform bounds for the bilinear Hilbert transform. We will employ the approach of Thiele and coauthors [MTT02] in the refined version of [FST24b] regarding uniform bounds in time-frequency analysis, utilizing phase plane projections as described in [FST24a] by Thiele and coauthors.

Time-frequency analysis decomposes the object of interest into components called trees. It is necessary to estimate each individual tree, which is less singular than the overall object, and to obtain good bounds on the number of trees. While the trees of the bilinear Hilbert transform represent elements of classical Calderón-Zygmund theory, the trees of the simplex Hilbert form will involve cubical singular Brascamp–Lieb integrals, and even the bounds for an individual tree will be new. The counting and organization of the trees will involve generalizations of the phase plane projections as in [FST24a].

Following this initial step toward Objective 6, we will further explore generalizations of [KTZK15] from the dyadic and Euclidean settings to doubling metric measure spaces. Currently, there is no suitable general formulation of a bilinear Hilbert transform applicable to doubling metric measure spaces. The ideas presented in [CHL24], which focus on Lipschitz singularities, represent progress in this direction. Since the Euclidean version of [KTZK15] also implies Carleson's theorem, the theory developed in the Lean Carleson project is another step toward generalizations of [KTZK15] for doubling metric measure spaces, particularly regarding certain ways of specializing one function.

Objective 6 will extend to investigate higher-order simplices. As the algebraic structure of the projections becomes more complex, there is currently a lack of overview regarding the possible specializations of the functions. It is evident that at least two new phenomena will arise even for the three-dimensional simplex. Specifically, specializing all functions entirely into elementary tensors may lead to the trilinear Hilbert transform, which cannot currently be estimated with existing technology. Additionally, certain specializations will require some additive combinatorics alongside time-frequency analysis, as discussed in [DPT10] by Thiele and coauthors. Jianghao Zhang, a student of Thiele, is currently working to further clarify this phenomenon.

b.4. Convergence of ergodic averages

According to the Calderón transference principle, quantitative estimates for the ergodic means (16) can be derived from quantitative estimates for the following averages of planar functions:

$$M_T(f_1, f_2)(x, y) = \frac{1}{T}\int_0^T f_1(x + t, y)f_2(x, y + t)\, dt\ . \tag{29}$$

The quantitative convergence of these averages in the norm, measured by the 2-variation, requires the estimation of sums

$$\sum_{j=1}^{J} \|M_{T_j}(f_1, f_2) - M_{T_{j-1}}(f_1, f_2)\|_2^2 \tag{30}$$

$$= \sum_{j=1}^{J} \int_{\mathbb{R}^4} f_1(x + t, y)f_2(x, y + t)\psi_j(t)f_1(x + s, y)f_2(x, y + s)\psi_j(s)\, dxdydtds\ , \tag{31}$$

where ψ_j is the difference between two successive L^1-normalized indicator functions of intervals, as implied in (29) and (30). A critical change of variables with $u = x + y + t$ and $v = x + y + s$, along with the shearing of each function f_i into a new function g_i, transforms this expression into

$$\sum_{j=1}^{J} \int_{\mathbb{R}^4} g_1(u,y)g_2(x,u)g_1(v,y)g_2(x,v)\psi_j(u-x-y)\psi_j(v-x-y)\,dxdydtds\,. \tag{32}$$

The arguments of the functions g_1 and g_2 exhibit the structure of cubical multilinear singular integrals as described in (12). The bump functions ψ combine to form a multiscale object that is essentially a singular integral kernel, albeit one with a Marcinkiewicz-type multiparameter structure. The arguments of the bump functions display a somewhat skewed projection, as extensively investigated by Thiele and coauthors in [DT20, DST22]. Therefore, estimates for the expression (32) fall within the scope of cubical multilinear singular integrals, with particular features of the kernel relevant to specific questions in ergodic theory. This has been employed by Thiele and coauthors in [DKŠT19] and similarly for three commuting transformations in [DST24].

Objective 8 necessitates estimating analogous expressions for three transformations that span a Heisenberg group. The shearing of variables mentioned above can be adapted to the Heisenberg group. Apart from this shearing, the Heisenberg group can be viewed as a doubling metric measure, allowing the application of the theory of singular Brascamp–Lieb integrals on doubling metric measure spaces, as developed in Objective 5. Since the Heisenberg group has a differentiable structure, an additional option is to utilize heat kernels defined on the Heisenberg group.

Pointwise estimates for two commuting transformations, as in Challenge IV, would require estimates on a larger variant of (30), with the variation norm applied before the Hilbert space norm. For some $r > 2$, this would look like:

$$\| \sup_{T_0<T_1<\cdots<T_J} (\sum_{j=1}^{J} |M_{T_j}(f_1,f_2) - M_{T_{j-1}}(f_1,f_2)|^r)^{1/r} \|_2\,. \tag{33}$$

Estimating (33) for the averages given in (29) appears out of reach with current technology, even considering that analogous bounds have been obtained [Lac00, DOP17] in the collinear case:

$$M_T(f_1,f_2)(x) = \frac{1}{T}\int_0^T f_1(x+t)f_2(x+2t)\,dt \tag{34}$$

related for the bilinear Hilbert transform.

Further averaging the ergodic means over orbits of several actions, as in (17), produces smoother objects and more coherent truncations of cubical singular integrals that are amenable to techniques such as martingale stopping times and their continuous analogues, as discussed in [DMT12]. Elaborating on these ideas in the setting of cubical multilinear singular integrals will be crucial to Objective 7. This will initially be done for commuting transformations and subsequently for transformations spanning a Heisenberg group. Further investigation will focus on reducing the number of orbit averages, which will address intermediate problems between Objective 7 and Challenge IV.

b.5 Nonlinear Fourier analysis

The starting point for Objective 9 is the theory developed in [ALM+24]. By restricting the nonlinear Fourier series (25) to sequences f supported only on either the positive or negative integers, a homeomorphism is formed [TT12] between square-integrable sequences and a corresponding nonlinear space that preserves the Plancherel identity (24).

For sequences on the full line, the forward nonlinear Fourier series then becomes the product of the Fourier series of f restricted to the negative and positive half-lines. Applying homeomorphisms on each half-line, inverting the nonlinear Fourier series becomes equivalent to performing a Riemann–Hilbert factorization of this product into two suitable factors.

Although this factorization is typically non-unique—indicating that the nonlinear Fourier series is not injective—the work in [ALM+24] establishes uniqueness under the condition that a is outer. This uniqueness lends stability to the forward nonlinear Fourier series, which we will utilize to derive analytic estimates.

One goal under Objective 9 is to establish a nonlinear analogue of Carleson's theorem, assuming that a is outer. We will draw on ideas from [Pol24], which proposes a weak nonlinear version of

Carleson's theorem but appears to contain a gap in the last portion of the argument, as noted in the most recent update of [Mna24]. We will adapt the ideas from both of these papers by transferring them from the $SU(1,1)$ model to the $SU(2)$ model, which involves developing a theory of left and right orthogonal polynomials for bilinear forms that are not symmetric. An intermediate objective may be a nonlinear analogue of a lacunary version of Carleson's theorem.

A second goal under Objective 9 is to establish a uniform nonlinear Hausdorff–Young inequality under the assumption of outer a. Uniform bounds were proven in a finite characteristic model of nonlinear Fourier series in [Kov12]. A third goal is to develop a structural theorem for the fibers of the $SU(2)$ nonlinear Fourier series, building on the results of Tsai using the ideas from [ALM$^+$24].

To generalize nonlinear Fourier series to higher dimensions, as desired in Objective 10, we will first focus on extending the results of [ALM$^+$24] to Fourier series valued in $SU(2n)$. For this, we interpret the definition in (25) in block matrix form. The initial step is to construct an outer matrix-valued function a for a given b. While this can be achieved through polynomial approximations, we will also seek efficient numerical methods for computing a that parallel the Weiss algorithm in the scalar case. The next step involves a matrix-valued Riemann–Hilbert factorization, which we aim to implement using fast Toeplitz solvers. An additional generalization will be to replace z with an arbitrary unitary input matrix, following the quantum singular value decomposition [GSLW19]. We will examine the implications of these constructions for quantum computing.

Similarly, one may explore generalizations of the $SU(1,1)$ nonlinear Fourier series. More broadly, we will investigate models in $SU(n,m)$ for general values of n and m, which includes the case of $SU(n)$ for odd n, and therefore does not follow the block matrix structure of (25). We will present a comprehensive analysis of all these models, extending the scalar case as in [TT12].

A further goal in Objective 10 is to extend the integer domain of the sequence f_n to the square lattice. In the context of solving the Davey–Stewartson equation [DS74], a nonlinear Fourier transform in the plane was developed in [AH75]. This transform has remarkable properties; for instance, it shares with the linear case the characteristic that, up to a reflection symmetry, it is essentially its own inverse. Tataru and coauthors [NRT20] have developed an advanced L^2 theory for this nonlinear Fourier transform. A discrete analog was introduced earlier in [GHL06]. We will discuss both models beyond the L^2 theory, examine the impact of the discrete model on quantum computing, and analyze fast algorithms for implementing this nonlinear Fourier transform.

Conclusion

HALF arrives at a pivotal moment, poised to be the first initiative to integrate formalization with research-level mathematics, fully leveraging the synergy that arises from this essential collaboration. By establishing a groundbreaking lab focused on the formalization of advanced mathematical research, HALF is charting a course toward a future where formalization becomes a standard part of mathematical exploration. Although this ambitious endeavor initially requires significant human resources, such an investment is the most effective and research-driven approach to achieving this transformative goal. A future with routine formalization is essential for mathematics, as it ensures absolute correctness and enables us to harness the full potential of modern artificial intelligence in mathematical research.

The HALF project structures its formalization efforts around a crucial agenda in harmonic analysis, tackling fundamental questions on multilinear singular integrals. This work in harmonic analysis also holds significant applications for convergence problems in ergodic theory and will pioneer the use of nonlinear Fourier analysis in quantum computing. Through these efforts, HALF not only aims to advance the frontiers of mathematics but also to establish a lasting framework that will enable future breakthroughs in both mathematics and its interdisciplinary applications.

Section c. Resources and time commitments (including project costs)

Total amount requested: 6.443.439 €

c.1 Resources and time commitments for Thiele

Total amount requested by Thiele, including indirect costs: 3.250.653 €

Personnel costs (2.407.522 €)

In every year of HALF, the team of Thiele will consist of A

- Christoph Thiele as cPI (298.819 €).
  Thiele is an expert in harmonic analysis. He will supervise the other members of his team, conduct research for HALF, coordinate with van Doorn regarding the interaction between the teams, and administer HALF as the corresponding PI. He will primarily work on Objectives 1 and 5–10.

  To fully focus on and work exclusively on the projects outlined in this proposal, Thiele requests six years of funding for 25% of his own position (full professorship at the Universität Bonn). Without grant support, he will not be able to engage in this research to the necessary extent. Thiele will dedicate 35% of his time to HALF.

- Two postdocs (NN) at 80% time each (1.000.047 €).
  Postdocs will be chosen respecting a balance of harmonic analysis expertise between the Objectives 5–10. They will conduct research on the respective objectives. At any time, one postdoc will be involved in the stage of writing a blueprint.

- Two PhD students (NN) at 75% time each (866.219 €).
  Each PhD student will conduct research on one of the Objectives 5–10. At any time, one PhD student will be involved in the stage of writing a blueprint.

- One Secretary (NN) at 50% time (242.437 €).
  The secretary will work for the administration of HALF.

Travel and subsistence (165.000 €)

We request 3.000 € per year for Thiele and for each of the postdocs and PhD students in Thiele's team. This funding will be used for travel to conferences as well as to invite guest researchers to support the team's research on the objectives of HALF.

Thiele will organize three small workshops on topics relevant to HALF, in which Thiele's team and additional young researchers from Bonn and beyond will participate. These workshops will disseminate the results of HALF and facilitate scientific exchange in areas pertinent to HALF. They will also contribute to the recruitment of scientific personnel for HALF. We request 15.000 € in travel and subsistence funds for each of these workshops.

Additionally, Thiele will organize a larger conference in the later years of HALF to further disseminate the research findings of HALF. We request 30.000 € in travel and subsistence funds for this conference.

Publications (0 €)

In line with practice in the field of harmonic analysis, all publications will be posted on the open access preprint server arXiv and submitted to journals at no cost to the authors.

Other additional costs (28.000 €)

We request additional costs of 5.000 € for each of the three smaller workshops and 10.000 € for the larger conference.

At the end of the funding period, we are required to conduct an audit to obtain a certificate of financial statements (3.000 €).

Existing resources

The corresponding host institution will cover the remaining 20% for each postdoc. Office space and a workstation are provided for all personnel in Thiele's team.

c.2 Resources and time commitments for van Doorn

Total amount requested by van Doorn, including indirect costs: 3.192.786 €

Personnel costs (2.461.229 €)

In every year of HALF, the team of van Doorn will consist of

- Floris van Doorn as PI 2 (319.032 €).
  Van Doorn is an expert in formalization and Lean. He will supervise the other members of his team, work on research for HALF and coordinate with Thiele the interaction between the teams. He will work mainly on Objectives 1, 2, 3, and 4 and the formalization of 5–10.

  To be able to fully focus and work exclusively on the projects outlined in this proposal, Van Doorn is requesting 50% of his own position for the first three years. His current non-permanent contract is set to expire in September 2028, so he is requesting 70% of his position for the fourth to sixth years. Van Doorn will dedicate 50% of his time to HALF in years 1-3 and 70% in years 4-6, which means that, on average, he will work 60% of his time for HALF over the entire duration of the project.

- Two postdocs (NN) at 80% time each (1.000.047 €).
  Postdocs will have expertise and do research in Lean and formalizations. They will work on Objectives 3 and 4, and the formalization of Objectives 2 and 5–10. An experienced postdoc will help supervise student research assistants with the formalization tasks.

- Two PhD students (NN) at 75% time each (866.219 €).
  Each PhD student will conduct research on one of the Objectives 3 and 4 and help with the formalization of Objectives 2 and 5–10.

- Four research student assistants (NN) at 10h per week (275.931 €).
  The research student assistants will work for the stage of formalizing a blueprint.

Travel and subsistence (90.000 €)

We request 3.000 € per year for van Doorn and for each of the postdocs and PhD students in van Doorn's team. This will be used for travel to conferences and to invite guest researchers to assist the team's research on the objectives of HALF. Attending conferences is important for presenting published work and exchanging ideas with external colleagues.

Publications (0 €)

In line with practice in the field of formalization, all publications will be posted on the open access preprint server arXiv and submitted to conferences and journals at no cost to the authors.

Other additional costs (3.000 €)

At the end of the funding period, we are required to conduct an audit to obtain a certificate of financial statements (3.000 €).

Existing resources

The current contract of van Doorn with the host institution ends in September 2028. However, the host institution will continue to fund the remaining percentage up to 100% of van Doorn's time for the duration of HALF. The host institution will also cover the remaining 20% of funding for each postdoc. Office space and a workstation will be provided for each member of van Doorn's team.

Appendix

All ongoing grants and submitted grant applications of each of the PIs (Funding ID)

Funding ID for Thiele

Ongoing grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
EXC 2047/1 Hausdorff Center for Mathematics	DFG	48.444.000	1/2019–12/2025	Thiele is one of 25 PIs.	This funding period will conclude before HALF starts.
Project C08 of CRC 1060	DFG	205.200	1/2021–12/2024	Thiele participates through project C08, *Multilinear estimates in geometric Fourier analysis.*	CRC 1060 is in its final funding period and will conclude before HALF starts.

Ongoing/submitted grant proposals

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
EXC 2047/2 Hausdorff Center for Mathematics	DFG	56.685.000	1/2026–12/2032	Thiele is one of 25 PIs. Thiele has no specified time commitment. No resources are allocated directly to Thiele.	Initiating the IRU Formal Mathematics, the HCM has laid the institutional base for HALF. The research proposed in HALF will not be funded by HCM.
Project C06 of CRC 1720	DFG	239.500	1/2026–12/2029	Thiele participates through project C06, *Variational Estimates in Multi- and Nonlinear Harmonic Analysis*, dedicating 6% of his time. C06 funds one PhD students supervised by Thiele.	The research will remain separate from HALF. C06 does not include formalization or focus on doubling metric measure spaces; instead, it will explore nonlinear phase unwinding. See also page 7 of B2.

Funding ID for van Doorn

Ongoing grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
NFDI 29/1 "MaRDI – Mathematische Forschungsdateninitiative."	DFG	1.261.800	10/2024–10/2027	Van Doorn is an application partner with no specified time commitment. This arrangement funds one PhD student for Van Doorn.	There is no overlap with HALF, as MaRDI focuses on managing research data derived from formalization and making it accessible via search. It does not involve the formalization of mathematics itself.

Ongoing/submitted grant proposals

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
EXC 2047/2 Hausdorff Center for Mathematics	DFG	56.685.000	1/2026–12/2032	Van Doorn is one of 25 PIs. There is no specificed percentage of time commitment.	Initiating the IRU Formal Mathematics, the HCM has laid the institutional base for HALF. The research proposed in HALF will not be funded by HCM.

References

[Aff24] Reynald Affeldt, *An introduction to MathComp-Analysis*, Lecture notes, aug 2024, version 8, https://staff.aist.go.jp/reynald.affeldt/documents/karate-coq.pdf.

[AH75] M. J. Ablowitz and R. Haberman, *Nonlinear evolution equations–two and three dimensions*, Phys. Rev. Lett. **35** (1975), no. 18, 1185–1188, doi:10.1103/PhysRevLett.35.1185.

[ALM+24] Michel Alexis, Lin Lin, Gevorg Mnatsakanyan, Christoph Thiele, and Jiasu Wang, *Infinite quantum signal processing for arbitrary Szegö functions*, 2024, arXiv:2407.05634.

[AMS19] Luigi Ambrosio, Andrea Mondino, and Giuseppe Savaré, *Nonlinear diffusion equations and curvature conditions in metric measure spaces*, Mem. Amer. Math. Soc. **262** (2019), no. 1270, v+121, doi:10.1090/memo/1270.

[AMT24] Michel Alexis, Gevorg Mnatsakanyan, and Christoph Thiele, *Quantum signal processing and nonlinear Fourier analysis*, Rev. Mat. Complut. **37** (2024), no. 3, 655–694, doi:10.1007/s13163-024-00494-5.

[AR15] Jeremy Avigad and Jason Rute, *Oscillation and the mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory Dynam. Systems **35** (2015), no. 4, 1009–1027, doi:10.1017/etds.2013.90.

[At24] AlphaProof and AlphaGeometry teams, *AI achieves silver-medal standard solving International Mathematical Olympiad problems*, blog post, July 2024, https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/.

[Aus10] Tim Austin, *On the norm convergence of non-conventional ergodic averages*, Ergodic Theory Dynam. Systems **30** (2010), no. 2, 321–338, doi:10.1017/S014338570900011X.

[BBP13] Jasmin Christian Blanchette, Sascha Böhme, and Lawrence C Paulson, *Extending sledgehammer with SMT solvers*, Journal of automated reasoning **51** (2013), no. 1, 109–128.

[BCCT08] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao, *The Brascamp–Lieb inequalities: finiteness, structure and extremals*, Geom. Funct. Anal. **17** (2008), no. 5, 1343–1415, doi:10.1007/s00039-007-0619-6.

[Bec24] Lars Becker, *A degree one Carleson operator along the paraboloid*, 2024, arXiv:2312.01134.

[Bil67] P. Billard, *Sur la convergence presque partout des séries de Fourier–Walsh des fonctions de l'espace L^2* (0, 1), Studia Math. **28** (1966/67), 363–388, doi:10.4064/sm-28-3-363-388.

[BL02] V. Bergelson and A. Leibman, *A nilpotent Roth theorem*, Invent. Math. **147** (2002), no. 2, 429–470, doi:10.1007/s002220100179.

[BLM15] Sylvie Boldo, Catherine Lelay, and Guillaume Melquiond, *Coquelicot: A user-friendly library of real analysis for Coq*, Mathematics in Computer Science **9** (2015), 41–62.

[Blo24] Thomas F. Bloom, *Unit fractions with shifted prime denominators*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics (2024), 1–11, doi:10.1017/prm.2024.42.

[Bou90] J. Bourgain, *Double recurrence and almost sure convergence*, J. Reine Angew. Math. **404** (1990), 140–161, doi:10.1515/crll.1990.404.140.

[BPL22] Anthony Bordg, Lawrence Paulson, and Wenda Li, *Simple type theory is not too simple: Grothendieck's schemes without dependent types*, Experimental Mathematics **31** (2022), no. 2, 364–382, doi:10.1080/10586458.2022.2062073.

[BvDJ+24] Lars Becker, Floris van Doorn, Asgar Jamneshan, Rajula Srivastava, and Christoph Thiele, *Carleson operators on doubling metric measure spaces*, 2024, arXiv:2405.06423.

[Car66] Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157, doi:10.1007/BF02392815.

[CHL24] Jiao Chen, Martin Hsu, and Fred Yu-Hsiang Lin, *A sharp Hörmander condition for bilinear Fourier multipliers with Lipschitz singularities*, 2024, arXiv:2403.04721.

[CK01] Michael Christ and Alexander Kiselev, *WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials*, J. Funct. Anal. **179** (2001), no. 2, 426–447, doi:10.1006/jfan.2000.3688.

[CT+22] Johan Commelin, Adam Topaz, et al., *Liquid Tensor Experiment*, Formalization project with 29 contributors, July 2022, https://github.com/leanprover-community/lean-liquid.

[DHL19] Sander R. Dahmen, Johannes Hölzl, and Robert Y. Lewis, *Formalizing the solution to the cap set problem*, 10th International Conference on Interactive Theorem Proving (ITP 2019) (Dagstuhl, Germany), Leibniz International Proceedings in Informatics (LIPIcs), vol. 141, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019, pp. 15:1–15:19, doi:10.4230/LIPIcs.ITP.2019.15.

[DK22] Polona Durcik and Vjekoslav Kovač, *A Szemerédi-type theorem for subsets of the unit cube*, Anal. PDE **15** (2022), no. 2, 507–549, doi:10.2140/apde.2022.15.507.

[DKŠT19] Polona Durcik, Vjekoslav Kovač, Kristina Ana Škreb, and Christoph Thiele, *Norm variation of ergodic averages with respect to two commuting transformations*, Ergodic Theory Dynam. Systems **39** (2019), no. 3, 658–688, doi:10.1017/etds.2017.48.

[DKT19] Polona Durcik, Vjekoslav Kovač, and Christoph Thiele, *Power-type cancellation for the simplex Hilbert transform*, J. Anal. Math. **139** (2019), no. 1, 67–82, doi:10.1007/s11854-019-0052-4.

[DLNW22] Yulong Dong, Lin Lin, Hongkang Ni, and Jiasu Wang, *Infinite quantum signal processing*, 2022, arXiv:2209.10162.

[DLTT08] Ciprian Demeter, Michael T. Lacey, Terence Tao, and Christoph Thiele, *Breaking the duality in the return times theorem*, Duke Math. J. **143** (2008), no. 2, 281–355, doi:10.1215/00127094-2008-020.

[DMT12] Yen Do, Camil Muscalu, and Christoph Thiele, *Variational estimates for paraproducts*, Rev. Mat. Iberoam. **28** (2012), no. 3, 857–878, doi:10.4171/RMI/694.

[DMT17] ———, *Variational estimates for the bilinear iterated Fourier integral*, J. Funct. Anal. **272** (2017), no. 5, 2176–2233, doi:10.1016/j.jfa.2016.09.010.

[DOP17] Yen Do, Richard Oberlin, and Eyvindur A. Palsson, *Variation-norm and fluctuation estimates for ergodic bilinear averages*, Indiana Univ. Math. J. **66** (2017), no. 1, 55–99, doi:10.1512/iumj.2017.66.5983.

[DPT10] Ciprian Demeter, Malabika Pramanik, and Christoph Thiele, *Multilinear singular operators with fractional rank*, Pacific J. Math. **246** (2010), no. 2, 293–324, doi:10.2140/pjm.2010.246.293.

[DS74] A. Davey and K. Stewartson, *On three-dimensional packets of surface waves*, Proc. Roy. Soc. London Ser. A **338** (1974), 101–110, doi:10.1098/rspa.1974.0076.

[DS18] Sebastián Donoso and Wenbo Sun, *Pointwise multiple averages for systems with two commuting transformations*, Ergodic Theory Dynam. Systems **38** (2018), no. 6, 2132–2157, doi:10.1017/etds.2016.127.

[DST22] Polona Durcik, Lenka Slavíková, and Christoph Thiele, *Local bounds for singular Brascamp–Lieb forms with cubical structure*, Math. Z. **302** (2022), no. 4, 2375–2405, doi:10.1007/s00209-022-03148-8.

[DST24] Polona Durcik, Lenka Slavíková, and Christoph Thiele, *Norm-variation of triple ergodic averages for commuting transformations*, 2024, arXiv:2307.07372.

[DT15] Yen Do and Christoph Thiele, *L^p theory for outer measures and two themes of Lennart Carleson united*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 249–296, doi:10.1090/S0273-0979-2014-01474-0.

[DT20] Polona Durcik and Christoph Thiele, *Singular Brascamp–Lieb inequalities with cubical structure*, Bull. Lond. Math. Soc. **52** (2020), no. 2, 283–298, doi:10.1112/blms.12310.

[DT21] ———, *Singular Brascamp–Lieb: a survey*, Geometric aspects of harmonic analysis, Springer INdAM Ser., vol. 45, Springer, Cham, [2021] ©2021, pp. 321–349, doi:10.1007/978-3-030-72058-2_9.

[Fef73] Charles Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. (2) **98** (1973), 551–571, doi:10.2307/1970917.

[Fri21] Nat Friedman, *Introducing GitHub Copilot: your AI pair programmer*, Blog post, jun 2021, https://github.blog/news-insights/product-news/introducing-github-copilot-ai-pair-programmer/.

[FST24a] Marco Fraccaroli, Olli Saari, and Christoph Thiele, *Phase space localizing operators*, 2024, arXiv:2210.16164.

[FST24b] ———, *Uniform bounds for bilinear symbols with linear k-quasiconformally embedded singularity*, 2024, arXiv:2402.11661.

[Gar07] John B. Garnett, *Bounded analytic functions*, first ed., Graduate Texts in Mathematics, vol. 236, Springer, New York, 2007.

[GHL06] Gegenhasi, Xing-Biao Hu, and Decio Levi, *On a discrete Davey–Stewartson system*, Inverse Problems **22** (2006), no. 5, 1677–1688, doi:10.1088/0266-5611/22/5/009.

[GL04] Loukas Grafakos and Xiaochun Li, *Uniform bounds for the bilinear Hilbert transforms. I*, Ann. of Math. (2) **159** (2004), no. 3, 889–933, doi:10.4007/annals.2004.159.889.

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe, *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*, STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2019, pp. 193–204.

[HK05] Bernard Host and Bryna Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488, doi:10.4007/annals.2005.161.397.

[Hos09] Bernard Host, *Ergodic seminorms for commuting transformations and applications*, Studia Math. **195** (2009), no. 1, 31–49, doi:10.4064/sm195-1-3.

[HvD20] Jesse Michael Han and Floris van Doorn, *A formal proof of the independence of the continuum hypothesis*, Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, Association for Computing Machinery, 2020, pp. 353–366, doi:10.1145/3372885.3373826.

[IMMS23a] Alexandru D. Ionescu, Ákos Magyar, Mariusz Mirek, and Tomasz Z. Szarek, *Polynomial averages and pointwise ergodic theorems on nilpotent groups*, Invent. Math. **231** (2023), no. 3, 1023–1140, doi:10.1007/s00222-022-01159-0.

[IMMS23b] ———, *Polynomial sequences in discrete nilpotent groups of step 2*, Adv. Nonlinear Stud. **23** (2023), no. 1, Paper No. 20230085, 24, doi:10.1515/ans-2023-0085.

[JOR96] Roger L. Jones, Iosif V. Ostrovskii, and Joseph M. Rosenblatt, *Square functions in ergodic theory*, Ergodic Theory Dynam. Systems **16** (1996), no. 2, 267–305, doi:10.1017/S0143385700008816.

[KMT22] Ben Krause, Mariusz Mirek, and Terence Tao, *Pointwise ergodic theorems for non-conventional bilinear polynomial averages*, Ann. of Math. (2) **195** (2022), no. 3, 997–1109, doi:10.4007/annals.2022.195.3.4.

[Kov12] Vjekoslav Kovač, *Uniform constants in Hausdorff–Young inequalities for the Cantor group model of the scattering transform*, Proc. Amer. Math. Soc. **140** (2012), no. 3, 915–926, doi:10.1090/S0002-9939-2011-11078-5.

[KT13] Vjekoslav Kovač and Christoph Thiele, *A* T(1) *theorem for entangled multilinear dyadic Calderón–Zygmund operators*, Illinois J. Math. **57** (2013), no. 3, 775–799, http://projecteuclid.org/euclid.ijm/1415023510.

[KTZK15] Vjekoslav Kovač, Christoph Thiele, and Pavel Zorin-Kranich, *Dyadic triangular Hilbert transform of two general functions and one not too general function*, Forum Math. Sigma **3** (2015), Paper No. e25, 27, doi:10.1017/fms.2015.25.

[Lac00] Michael T. Lacey, *The bilinear maximal functions map into L^p for $2/3 < p \leq 1$*, Ann. of Math. (2) **151** (2000), no. 1, 35–57, doi:10.2307/121111.

[LF23] Jannis Limperg and Asta Halkjær From, *Aesop: White-box best-first proof search for Lean*, Proceedings of the 12th ACM SIGPLAN International Conference on Certified Programs and Proofs (New York, NY, USA), CPP 2023, Association for Computing Machinery, 2023, p. 253–266, doi:10.1145/3573105.3575671.

[Lie20] Victor Lie, *The polynomial Carleson operator*, Ann. of Math. (2) **192** (2020), no. 1, 47–163, doi:10.4007/annals.2020.192.1.2.

[LT97] Michael Lacey and Christoph Thiele, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$*, Ann. of Math. (2) **146** (1997), no. 3, 693–724, doi:10.2307/2952458.

[LT99] ———, *On Calderón's conjecture*, Ann. of Math. (2) **149** (1999), no. 2, 475–496, doi:10.2307/120971.

[LT00] ———, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370, doi:10.4310/MRL.2000.v7.n4.a1.

[mat20] The mathlib Community, *The Lean mathematical library*, Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (New York, NY, USA), CPP 2020, Association for Computing Machinery, 2020, pp. 367–381, doi:10.1145/3372885.3373824.

[MC97] Yves Meyer and Ronald Coifman, *Wavelets*, Cambridge Studies in Advanced Mathematics, vol. 48, Cambridge University Press, Cambridge, 1997, Calderón-Zygmund and multilinear operators, Translated from the 1990 and 1991 French originals by David Salinger.

[Met23] Bhavik Metha, *Formalising modern research mathematics in real time*, Blog post, November 2023, https://xenaproject.wordpress.com/2023/11/04/formalising-modern-research-mathematics-in-real-time/.

[Mna22] Gevorg Mnatsakanyan, *On almost-everywhere convergence of Malmquist–Takenaka series*, J. Funct. Anal. **282** (2022), no. 12, Paper No. 109461, 33, doi:10.1016/j.jfa.2022.109461.

[Mna24] Gevorg Mnatsakanyan, *Some quantitative estimates for the nonlinear Fourier transform*, 2024, arXiv:2210.00309.

[MT17] Mariusz Mirek and Christoph Thiele, *A local $T(b)$ theorem for perfect multilinear Calderón–Zygmund operators*, Proc. Lond. Math. Soc. (3) **114** (2017), no. 1, 35–59, doi:10.1112/plms.12000.

[MTT02] Camil Muscalu, Terence Tao, and Christoph Thiele, *Uniform estimates on multi-linear operators with modulation symmetry*, J. Anal. Math. **88** (2002), 255–309, Dedicated to the memory of Tom Wolff, doi:10.1007/BF02786579.

[MTT03a] ______, *A Carleson theorem for a Cantor group model of the scattering transform*, Nonlinearity **16** (2003), no. 1, 219–246, doi:10.1088/0951-7715/16/1/314.

[MTT03b] ______, *A counterexample to a multilinear endpoint question of Christ and Kiselev*, Math. Res. Lett. **10** (2003), no. 2-3, 237–246, doi:10.4310/MRL.2003.v10.n2.a10.

[NRT20] Adrian Nachman, Idan Regev, and Daniel Tataru, *A nonlinear Plancherel theorem with applications to global well-posedness for the defocusing Davey–Stewartson equation and to the inverse boundary value problem of Calderón*, Invent. Math. **220** (2020), no. 2, 395–451, doi:10.1007/s00222-019-00930-0.

[NY24] Hongkang Ni and Lexing Ying, *Fast phase factor finding for quantum signal processing*, 2024, arXiv:2410.06409.

[OST+12] Richard Oberlin, Andreas Seeger, Terence Tao, Christoph Thiele, and James Wright, *A variation norm Carleson theorem*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 421–464, doi:10.4171/JEMS/307.

[Pol24] A. Poltoratski, *Pointwise convergence of the non-linear Fourier transform*, Ann. of Math. (2) **199** (2024), no. 2, 741–793, doi:10.4007/annals.2024.199.2.4.

[RC22] Zane M. Rossi and Isaac L. Chuang, *Multivariable quantum signal processing (M-QSP): prophecies of the two-headed oracle*, Quantum **6** (2022), 811, doi:10.22331/q-2022-09-20-811, https://doi.org/10.22331/q-2022-09-20-811.

[Ste93] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[SUdM20] Daniel Selsam, Sebastian Ullrich, and Leonardo de Moura, *Tabled typeclass resolution*, 2020, arXiv:2001.04301.

[T+23] Terence Tao et al., *The polynomial Freiman–Ruzsa conjecture*, Formalization project with 28 contributors, December 2023, https://teorth.github.io/pfr/.

[T+24] ______, *Equational theories project*, Formalization project with 42 contributors, October 2024, https://github.com/teorth/equational_theories.

[Tao08] Terence Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 657–688, doi:10.1017/S0143385708000011.

[Tao16] ______, *Cancellation for the multilinear Hilbert transform*, Collect. Math. **67** (2016), no. 2, 191–206, doi:10.1007/s13348-015-0162-y.

[Thi02] Christoph Thiele, *A uniform estimate*, Ann. of Math. (2) **156** (2002), no. 2, 519–563, doi:10.2307/3597197.

[TT12] Terence Tao and Christoph Thiele, *Nonlinear Fourier analysis*, 2012, arXiv:1201.5129.

[UW22] Gennady Uraltsev and Michał Warchalski, *The full range of uniform bounds for the bilinear Hilbert transform*, 2022, arXiv:2205.09851.

[vD⁺] Floris van Doorn et al., *Carleson operators on doubling metric measure spaces*, website, retrieved on 19 Oct 2024, https://florisvandoorn.com/carleson/.

[vD21] Floris van Doorn, *Formalized Haar measure*, 12th International Conference on Interactive Theorem Proving (ITP 2021), Leibniz International Proceedings in Informatics (LIPIcs), vol. 193, 2021, pp. 18:1–18:17, doi:10.4230/LIPIcs.ITP.2021.18.

[vDMN23] Floris van Doorn, Patrick Massot, and Oliver Nash, *Formalising the h-principle and sphere eversion*, Proceedings of the 12th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2023, Association for Computing Machinery, 2023, pp. 121–134, doi:10.1145/3573105.3575688.

[Wal12] Miguel N. Walsh, *Norm convergence of nilpotent ergodic averages*, Ann. of Math. (2) **175** (2012), no. 3, 1667–1688, doi:10.4007/annals.2012.175.3.15.

[ZK17] Pavel Zorin-Kranich, *Cancellation for the simplex Hilbert transform*, Math. Res. Lett. **24** (2017), no. 2, 581–592, doi:10.4310/MRL.2017.v24.n2.a16.

[ZK21] ———, *Maximal polynomial modulations of singular integrals*, Adv. Math. **386** (2021), Paper No. 107832, 40, doi:10.1016/j.aim.2021.107832.