Waves, Spheres, and Tubes
A Selection of Fourier Restriction Problems, Methods, and Applications

Betsy Stovall

Introduction

The Fourier transform on $\mathbb{R}^d$

$$\hat{f}(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx, \qquad \xi \in \mathbb{R}^d, \tag{1}$$

is a fundamental operator that, via the Fourier inversion formula, provides coefficients to represent nice functions as superpositions of plane waves,

$$f(x) = \tfrac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi)\, e^{ix\cdot\xi} d\xi. \tag{2}$$

Aside from the basic utility of decomposing functions into simpler parts, this operator is powerful in the study of differential equations because it intertwines the operations of differentiation and polynomial multiplication via the formulae

$$\widehat{(-ix_j f)} = \partial_{\xi_j} \hat{f}, \qquad \widehat{(\partial_{x_j} f)} = i\xi_j \hat{f}. \tag{3}$$

Surprisingly, given its importance, much remains unknown

Betsy Stovall is an associate professor at the University of Wisconsin–Madison. Her email address is stovall@math.wisc.edu.

This work was supported in part by NSF DMS-1653264. The author is indebted to a number of colleagues, including the referees, who provided invaluable feedback on earlier versions.

Communicated by Notices *Associate Editor Daniela De Silva.*

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: https://doi.org/10.1090/noti1906

about how the Fourier transform interacts with simple geometric objects.

The Fourier restriction problem for the sphere asks for which exponents $(p, q) \in [1, \infty]^2$ does there hold an inequality

$$\|\widehat{f}|_{S^{d-1}}\|_{L^q(S^{d-1})} \leq C\|f\|_{L^p(\mathbb{R}^d)}, \tag{4}$$

with a constant C uniform over all smooth, compactly supported functions f. Such an *a priori* inequality would imply that the restriction operator, $f \mapsto \widehat{f}|_{S^{d-1}}$, uniquely extends to all of L^p, providing a reasonable interpretation of the restriction of the Fourier transform of L^p functions to the sphere. The mismatch between the linearity of planar waves and the curvature of the sphere makes this problem highly nontrivial and leads to deep connections among seemingly disparate areas of mathematics, such as geometric measure theory, dispersive PDE, and analytic number theory.

The restriction problem originated from work of Elias M. Stein, who in the 1960s established the first positive results in unpublished work. Despite its quick resolution in two dimensions, in work by Fefferman–Stein and Zygmund, the restriction conjecture for the sphere, that (4) holds if and only if

$$p < \tfrac{2d}{d+1} \quad \text{and} \quad \tfrac{d-1}{p'} \geq \tfrac{d+1}{q}, \tag{5}$$

remains open in all dimensions $d \geq 3$. Recent developments have substantially narrowed the gap between restriction theorems and the conjecture, and these new techniques have facilitated or promise to enable progress on a vast circle of related questions. The purpose of this article is to give a brief introduction for non-experts to some of the ideas and applications surrounding Stein's restriction problem.

The article is structured around a selection of methods, which we describe, for clarity, in the context of the restriction problem for the paraboloid, a simpler model than the sphere. In the elliptic setting (which contains both the sphere and paraboloid), most of these methods have by now been surpassed, but the field has become sufficiently broad that each method is at, or even beyond, the state-of-the-art for a large number of interesting restriction problems. Giving a sense of this breadth is a particular goal of this article, so we will describe and motivate some of these related questions as we reach their record-setting methods.

The competing demands of limited space, topical breadth, and (especially) clarity for non-experts, in a time of rapid development, necessitate the neglect of a number of important results. The author apologizes for these omissions and encourages the reader to view the bibliography as merely a starting point for further reading.

Background: The Restriction and Extension Operators

Our integral definition (1) of the Fourier transform only makes sense for a small class of functions, those in $L^1(\mathbb{R}^d)$, but for applications one often needs to extend the definition to a larger class. For $f \in L^1 \cap L^2$, the Plancherel theorem states that $\|f\|_2 = \frac{1}{(2\pi)^{d/2}}\|\widehat{f}\|_2$, and since $L^1 \cap L^2$ is dense in L^2, the Fourier transform extends uniquely to a bounded linear operator of L^2 onto itself. By interpolation, we obtain the Hausdorff–Young inequality, which states that for $1 \leq p \leq 2$, this extension maps L^p boundedly into $L^{p'}$ and obeys $\|\widehat{f}\|_{p'} \leq C_{p,d}\|f\|_p$; here p' is the dual exponent $p' = \frac{p}{p-1}$. This range of $L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ estimates is the best possible.

For $f \in L^1$, $\widehat{f}$ makes sense pointwise, but for $f \in L^p$, $1 < p \leq 2$, we usually interpret $\widehat{f}$ as an $L^{p'}$ limit, $\widehat{f} = \lim_{n\to\infty} \widehat{f}_n$, where f_n is a sequence of integrable functions converging to f in L^p. This interpretation leads to an obvious obstruction to restricting a Fourier transform to sets of Lebesgue measure zero: $L^{p'}$ convergence does not "see" sets of measure zero. Indeed, $L^{p'}$ consists of equivalence classes within which members are allowed to differ off of sets of measure zero. Thus it makes no sense to define Fourier restriction to a set of measure zero as a simple composition. This nontriviality leads to a generalization of Stein's original restriction problem. Given a set Σ of Lebesgue measure zero, for which p, q does there hold an inequality

$$\|\widehat{f}|_{\Sigma}\|_{L^q(\Sigma)} \leq C\|f\|_p, \tag{6}$$

for smooth, compactly supported functions f? Such an *a priori* estimate would imply that one can uniquely define an extension, $\mathcal{R}_\Sigma : L^p \to L^q$, which agrees with $f \mapsto \widehat{f}|_\Sigma$ on $L^1 \cap L^p$. (One needs a measure σ on Σ to define the space $L^q(\Sigma)$. Which measure to use is an interesting question, but here we will focus on the case that σ is comparable to the natural Hausdorff measure on Σ.)

Finding the optimal (smallest) constant C in (6) seems to be a very difficult problem, which has only been solved in a few cases. As such questions of sharp constants are beyond the scope of this article, we will sweep operator norms and other constants under the rug by writing $A \lesssim B$ to mean that $A \leq CB$, for a constant C depending only on admissible parameters, such as the exponents p, q and the set Σ.

It was observed early on that even when Σ has measure zero, (6) is often uninteresting. The restriction problem for $p = 1$ is trivial because the Fourier transform of an L^1 function is bounded and continuous on all of $\mathbb{R}^d$, and so restricts to any set $\Sigma \subseteq \mathbb{R}^d$ in the natural way. The Fourier transform is an L^2 isometry by Plancherel, so a function can have small L^2 norm, while being large near a given set of measure zero: the restriction problem for $p = 2$ (and hence $p \geq 2$) is impossible. Finally, for $p > 1$ and $0 \not\equiv \phi$ smooth, nonnegative, and compactly supported, the sequence $\phi_n(x) := n^{-\frac{1}{p}}\phi(x', x_d/n)$, where $x = (x', x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$, remains bounded in L^p, but its Fourier transform, $\hat{\phi}_n(\xi) = n^{\frac{1}{p'}}\hat{\phi}(\xi', n\xi_d)$, blows up (in every L^q space) on the hyperplane $\{\xi_d = 0\}$. Similar examples show that nontrivial restriction theorems are impossible for any subset of any affine hyperplane and, moreover, that interesting sets for Fourier restriction must escape from hyperplanes in the sense of having small measure trapped inside thin planar slabs. We think of such sets as being curved (even when they have no manifold structure); for instance, elliptic hypersurfaces, such as the sphere, escape from every hyperplane at a quadratic or faster rate.

Taken together, these examples allow us to recast the restriction problem: Given a set Σ of Lebesgue measure zero that escapes from hyperplanes, what is the largest p for which (6) holds for some exponent $q \geq 1$, and, given such a p, what is the largest possible q? Raising the exponents p and q has a heuristic interpretation, to which we now turn.

The identities in (3) imply a relationship between decay of a function and smoothness of its Fourier transform, while the Lebesgue inequalities in Hausdorff–Young do as well, but in a more subtle way. Chebyshev's inequality states that $\{|f| > \lambda\}$ has measure at most $\lambda^{-p}\|f\|_{L^p}^p$. For large λ, this inequality controls the rate of blowup of f and is strongest for large p; $f \in L^p$ with p large is thus a kind of averaged smoothness condition. For small λ, Chebyshev's inequality is a kind of decay condition, as it is strongest for small p. Since p' increases as p decreases, we may reinterpret Hausdorff–Young as a version of the principle that faster decay leads to a smoother Fourier transform.

Restriction inequalities are a finer manifestation of the link between decay and smoothness. *A priori*, the restriction problem for a set Σ is to determine precisely how little L^p decay is needed to restrict the Fourier transform and then to obtain the strongest L^q averaged smoothness along Σ for Fourier transforms of functions with that decay. Heuristically, bounds on the average smoothness along a measure zero set are quite strong because the denominator used in the average is small. (This interpretation also points to an understanding of sets with nontrivial restriction theorems, such as the sphere or even the helix, as larger than hyperplanes, but such a discussion is beyond the scope of this article.)

The preceding paragraph casts restriction inequalities as tangential smoothness conditions on Fourier transforms of functions with a given decay; recently, maximal Fourier restriction inequalities, due to Müller–Ricci–Wright [17] and others, also demonstrate a kind of transverse smoothness. An application of these results is a more intrinsic interpretation of the restriction operator than one obtains by using *a priori* inequalities to approximate and take L^q limits: for certain Σ and exponents p, the restriction operator on L^p can be understood by using averages over shrinking Euclidean balls,

$$\mathcal{R}_\Sigma f(\xi) = \lim_{r \searrow 0} \frac{1}{|B_r(\xi)|} \int_{B_r(\xi)} \hat{f}(\eta)\, d\eta, \quad f \in L^p(\mathbb{R}^d), \quad \text{a.e. } \xi \in S^{d-1}.$$

The term inside the limit always makes sense when $1 \leq p \leq 2$, but a.e. existence of the limit fails when $p > 1$ for subsets of hyperplanes, because (as we have seen) the Fourier transform of an L^p function may lack smoothness in the (fixed) direction perpendicular to the hyperplane.

Whereas the restriction operator discards a great deal of information, its dual, which is called the extension operator and is given by

$$(\mathcal{R}_\Sigma)^* g(x) =: \mathcal{E}_\Sigma g(x) = \int_\Sigma e^{ix\cdot\xi} g(\xi)\, d\sigma(\xi),$$

does not and is thus easier to work with. Every restriction theorem, $\mathcal{R}_\Sigma : L^p(\mathbb{R}^d) \to L^q(\Sigma)$, has an equivalent extension counterpart, $\mathcal{E}_\Sigma : L^{q'}(\Sigma) \to L^{p'}(\mathbb{R}^d)$, and, as of this writing, almost all restriction theorems are really proved as extension theorems. For the remainder of the article, we will consider only the extension problem, which is to prove *a priori* inequalities,

$$\|\mathcal{E}_\Sigma f\|_{L^q(dx)} \lesssim \|f\|_{L^p(\sigma)}, \tag{7}$$

with f in some appropriate dense class. Dualizing our earlier remarks, the case $q = \infty$ is trivial, the case $q \leq 2$ is impossible, and the goal is to decrease q as much as possible (show more decay), which may require increasing p (assuming more smoothness), though, optimally, we would like to increase p as little as possible.

We will focus on a specific set Σ that is easy to parametrize and visualize and for which current techniques are most advanced, the extension operator associated to the paraboloid $\mathbb{P} = \{(\xi, |\xi|^2) : \xi \in \mathbb{R}^{d-1}\}$:

$$\mathcal{E}_\mathbb{P} f(x) = \int_{\mathbb{R}^{d-1}} e^{ix\cdot(\xi,|\xi|^2)} f(\xi)\, d\xi,$$

initially defined for smooth functions f with support contained in some fixed ball centered at 0 in $\mathbb{R}^{d-1}$. (We will sweep under the rug the effects of the size of this ball.)

Most of the methods for bounding $\mathcal{E}_\mathbb{P}$ readily generalize to compact *elliptic* hypersurfaces, such as the sphere; such surfaces are locally, after an affine transformation, small perturbations of the paraboloid. However, none of these methods have been fully generalized to all "interesting" measure zero sets, nor even all hypersurfaces, as we will see.

Classical Methods and the L^2 Theory

The paraboloid has a large affine symmetry group. Since the dot product interacts well with affine transformations, so too does the extension operator $\mathcal{E}_{\mathbb{P}}$. Three particularly important symmetries are:

$$\textit{Parabolic dilations: } \mathcal{E}_{\mathbb{P}}[\varepsilon^{-(d-1)} f(\varepsilon^{-1}\xi)](x) = \mathcal{E}_{\mathbb{P}} f(\varepsilon x', \varepsilon^2 x_d); \tag{8}$$

$$\textit{Frequency translations: } \mathcal{E}_{\mathbb{P}}[f(\xi - \xi_0)](x) = e^{ix\cdot(\xi_0,|\xi_0|^2)} \mathcal{E}_{\mathbb{P}} f(x' + 2x_d\xi_0, x_d); \tag{9}$$

$$\textit{Spatial translations: } \mathcal{E}_{\mathbb{P}}[e^{-ix_0(\xi,|\xi|^2)} f(\xi)](x) = \mathcal{E}_{\mathbb{P}} f(x - x_0). \tag{10}$$

The parabolic and frequency translations are associated, respectively, with the dilation $(\xi, |\xi|^2) \mapsto (\varepsilon\xi, \varepsilon^2|\xi|^2)$ and Galilean $(\xi, |\xi|^2) \mapsto (\xi + \xi_0, |\xi + \xi_0|^2)$ symmetries of $\mathbb{P}$.

We can use the symmetries to illustrate some of the geometry associated to $\mathcal{E}_{\mathbb{P}}$. Let φ be a smooth, nonnegative function with integral 1 and support in $\{|\xi| < 0.1\}$. We begin by estimating $\mathcal{E}_{\mathbb{P}}\varphi$. For $|x'|, |x_d| < 1$ the integral defining $\operatorname{Re} \mathcal{E}_{\mathbb{P}}\varphi$ has no cancellation, and hence $|\mathcal{E}_{\mathbb{P}}\varphi(x)|$ is nearly as large as possible:

$$|\mathcal{E}_{\mathbb{P}}\varphi(x)| \geq \int \cos(x \cdot (\xi, |\xi|^2))\varphi(\xi)\, d\xi \geq \int \cos(0.11)\varphi(\xi)\, d\xi \approx 1.$$

For large $|x|$ the integrand oscillates rapidly in ξ, leading to cancellation in the integral, and hence a small extension: $|\mathcal{E}_{\mathbb{P}}\varphi(x)| \ll 1$ for $|x| \gg 1$.

Now we apply our symmetries, first a dilation, then a frequency translation, and finally a spatial translation,

$$\varphi^{\varepsilon}_{\xi_0,x_0}(\xi) = \varepsilon^{-(d-1)} e^{ix_0\cdot(\xi,|\xi|^2)} \varphi(\varepsilon^{-1}(\xi - \xi_0)).$$

Here, due to our assumption that functions are supported in some fixed ball, we think of ε as small. The function $\varphi^{\varepsilon}_{\xi_0,x_0}$ is supported on the set $\{|\xi - \xi_0| < 0.1\varepsilon\}$, which is naturally associated to the parabolic subset

$$\kappa^{\varepsilon}_{\xi_0} := \{\zeta \in \mathbb{P} : 0 \leq (\zeta - \zeta_0) \cdot \nu_{\xi_0} < 0.01\varepsilon^2\},$$

where $\zeta_0 := (\xi_0, |\xi_0|^2)$ and ν_{ξ_0} is the upward normal $\langle -2\xi_0, 1\rangle$ to $\mathbb{P}$ at ζ_0. The set $\kappa^{\varepsilon}_{\xi_0}$ thus resembles a tilted, shallow bowl of width about ε in directions tangent to $\mathbb{P}$ and depth about ε^2 in the direction ν_{ξ_0}, the relationship between width and thickness reflecting the curvature of $\mathbb{P}$. We call both the set $\kappa^{\varepsilon}_{\xi_0}$ and the function $\varphi^{\varepsilon}_{\xi_0,0}$ ε-caps on $\mathbb{P}$ centered at ζ_0, while we call the more general $\varphi^{\varepsilon}_{\xi_0,x_0}$ a modulated ε-cap. (The comparison to headgear is more convincing for caps on $-\mathbb{P}$ or on the northern hemisphere of S^{d-1}.)

The extension of a modulated ε-cap is, per (8-10), given by

$$\mathcal{E}_{\mathbb{P}}\varphi^{\varepsilon}_{\xi_0,x_0}(x) = e^{i(x-x_0)\cdot\zeta_0} \mathcal{E}_{\mathbb{P}}\varphi(\varepsilon[(x - x_0)' + 2(x - x_0)_d\, \xi_0], \varepsilon^2(x - x_0)_d).$$

Using our estimates for $\mathcal{E}_{\mathbb{P}}\varphi$, $|\mathcal{E}_{\mathbb{P}}\varphi^{\varepsilon}_{\xi_0,x_0}|$ is comparable to 1 on the long tube

$$T^{\varepsilon}_{\xi_0,x_0} = \{x \in \mathbb{R}^d : |(x - x_0)' + 2(x - x_0)_d\, \xi_0| < \varepsilon^{-1}, |x_d - (x_0)_d| < \varepsilon^{-2}\},$$

centered at x_0 and having width $\varepsilon^{-1} > 1$ and length $\varepsilon^{-2} > \varepsilon^{-1}$ in the direction ν_{ξ_0} (this geometry again reflects the curvature of $\mathbb{P}$), and decays rapidly off of this tube. For $T = T^{\varepsilon}_{\xi_0,x_0}$ and $\varphi_T = \varphi^{\varepsilon}_{\xi_0,x_0}$, the extension $\mathcal{E}_{\mathbb{P}}\varphi_T$ is called a *wave packet* associated to T.

For any $\varepsilon \ll 1$, a partition of unity directly decomposes our original function φ as a sum of (unmodulated) ε-caps, indexed by a collection of $O(\varepsilon^{-(d-1)})$ tubes,

$$\varphi = \sum_T c_T \varphi_T, \qquad T = T^{\varepsilon}_{\xi_T,0},$$

most of whose coefficients c_T are about ε^{d-1} (the rest are smaller). The curvature of $\mathbb{P}$ means that distinct tubes T, T' have directions $\nu_T, \nu_{T'}$ separated by at least ε. Because two tubes begin to separate at the length scale $\frac{width}{angle} \gtrsim \frac{\varepsilon^{-1}}{\varepsilon}$, tubes centered at 0 are nearly pairwise disjoint on $\{|x| \sim \varepsilon^{-2}\}$. Taking $\varepsilon \sim (1 + |x|)^{-\frac{1}{2}}$, we can morally approximate $\mathcal{E}_{\mathbb{P}}\varphi(x)$ using the contribution from a single tube:

$$|\mathcal{E}_{\mathbb{P}}\varphi(x)| \approx |\mathcal{E}_{\mathbb{P}}(c_T\phi_T)(x)| \approx \varepsilon^{d-1} \approx (1 + |x|)^{-(d-1)/2}. \tag{11}$$

A rigorous version of (11) can be proved using the method of stationary phase.

The estimates $\|\varphi_T\|_p \approx \varepsilon^{-(d-1)/p'}$ and $\|\mathcal{E}_{\mathbb{P}}\varphi_T\|_q \gtrsim |T|^{1/q}$ (the Knapp example), and the lower bound (11), lead to the necessary conditions

$$\tfrac{d-1}{p'} \geq \tfrac{d+1}{q}, \qquad q > \tfrac{2d}{d-1} \tag{12}$$

for validity of (7). By dualizing, (12) is equivalent to the necessary condition (5) for boundedness of the restriction operator, and these are currently the only known necessary conditions. (The introduction concerned the sphere; the conjecture for the paraboloid is the same.)

Self-duality of L^p when $p = 2$ allows one to recast the extension problem for $p = 2$ as a question about the convolution operator

$$\mathscr{E}_{\mathbb{P}} \circ \mathcal{R}_{\mathbb{P}} f(x) = \int_{\mathbb{R}^{d-1}} (\mathscr{E}_{\mathbb{P}} \psi)(x-y) f(y)\, dy,$$

where ψ is smooth, compactly supported, and is identically 1 on some large ball. This formulation has the advantage that the difficult-to-understand extension operator is applied to a fixed function ψ, rather than the arbitrary function f. An argument of Stein and Tomas from the 1970s uses the decay estimate (11) together with the simple observation that an ε-cap on the paraboloid has area ε^{d-1} to show that $\mathscr{E}_{\mathbb{P}}$ maps L^2 boundedly into L^q for $q \leq \frac{2(d+1)}{d-1}$, which by (12) is the optimal L^2-based result.

The Stein–Tomas Theorem generalizes to other sets, including sets of fractional dimension, so long as decay and volume estimates are known ([1], and the references therein). However, sharp decay estimates can be quite intricate and caps much more complicated for general sets, so sharp L^2-based estimates are known in a relatively narrow collection of examples. For example, what the optimal L^2-based extension estimates for polynomial hypersurfaces are in dimension $d \geq 4$ is largely an open question, with the $d = 3$ case having been only recently resolved by Ikromov–Müller [14].

One reason the L^2 theory for the extension problem for hypersurfaces has been so well studied is that the equations (3) lead to a connection with certain PDE. For instance, for a sufficiently nice function u_0 on $\mathbb{R}^{d-1}$, the extension $\mathscr{E}_{\mathbb{P}} \hat{u}_0$ solves the linear Schrödinger equation $iu_t = -\Delta_x u$ on $\mathbb{R}^{d-1}_x \times \mathbb{R}_t$ with initial data $u(x,0) = u_0(x)$, and extension operators for higher order polynomial surfaces are associated to higher order equations. Via Plancherel, an $L^2 \to L^q$ extension estimate leads to an $L^2 \to L^q$ inequality for the data-to-solution map (known as a Strichartz estimate), and such estimates have been extremely important in the study of a variety of dispersive and wave equations.

The Wave Packet Decomposition and Kakeya

For $p > 2$, a larger range of $L^p \to L^q$ extension estimates for $\mathscr{E}_{\mathbb{P}}$ should be possible, but the problem cannot be readily reduced to a question about convolutions. Instead, one wants to directly study the extension $\mathscr{E}_{\mathbb{P}} f$, where the function f may be very rough, even on small scales. After using a partition of unity to localize $f = \sum f^{\varepsilon}_{\xi_j}$, Fourier series methods resolve each $f^{\varepsilon}_{\xi_j}$ as a superposition of modulated caps; thus

$$f = \sum_T c_T \varphi_T, \qquad T = T^{\varepsilon}_{\xi_T, x_T}. \tag{13}$$

This *wave packet decomposition* lies at the heart of modern approaches to the subject, to which we turn in the coming sections. It first appeared in work of Bourgain in 1991, though it is closely connected with the earlier combinatorial approach of Fefferman and Córdoba. (A detailed implementation and more precise references may be found in e.g. [12].)

When $p = 2$, different modulations of the same cap are mutually orthogonal, and we can approximate

$$\|f\|_2 \approx \Big(\sum_T \|c_T \varphi_T\|_2^2\Big)^{\frac{1}{2}} \approx \varepsilon^{-(d-1)/2}\Big(\sum_T |c_T|^2\Big)^{\frac{1}{2}}; \tag{14}$$

this is analogous to the Plancherel theorem for the Fourier transform. For $p \neq 2$, however, the analogue to (14) is valid if and only if each cap is associated to at most one tube (instead of many parallel translates).

Let $\mathcal{T}^{\varepsilon}$ be a collection of length ε^{-2}, width ε^{-1} tubes in one-to-one correspondence with the set of ε caps. Understanding how these tubes can pile up and interact with one another is a very difficult geometric problem, with implications for bounds on $\mathscr{E} \sum_{T \in \mathcal{T}^{\varepsilon}} \varphi_T$. To illustrate this, we adapt an argument from [2].

In the previous section, we used the approximation $|\mathscr{E}_{\mathbb{P}} \varphi_T| \approx \chi_T$, but this ignores oscillatory factors, which might lead to cancellation (and hence a smaller sum) if many tubes overlap. Nevertheless, randomization (Khintchine's inequality) allows us to obtain a lower bound

$$\Big\| \sum_{T \in \mathcal{T}^{\varepsilon}} \chi_T \Big\|_{q/2}^{1/2} = \Big\| \Big(\sum_{T \in \mathcal{T}^{\varepsilon}} |\chi_T|^2 \Big)^{1/2} \Big\|_q \lesssim \|\mathscr{E}_{\mathbb{P}} f\|_q, \tag{15}$$

for some $f := \sum_{T \in \mathcal{T}^{\varepsilon}} \pm \phi_T$. The $L^{q/2}$ norm on the left of (15) is closely connected, via Hölder's inequality, with the volume of the union of tubes:

$$\sum_{T \in \mathcal{T}^{\varepsilon}} |T| = \Big\| \sum_{T \in \mathcal{T}^{\varepsilon}} \chi_T \Big\|_{L^1} \leq \Big| \bigcup_{T \in \mathcal{T}^{\varepsilon}} T \Big|^{1-\frac{2}{q}} \Big\| \sum_{T \in \mathcal{T}^{\varepsilon}} \chi_T \Big\|_{q/2}. \tag{16}$$

Meanwhile, validity of (7) would allow us to bound the right-hand side of (15) by $\|f\|_p \approx \varepsilon^{-(d-1)}$. After a bit of arithmetic, (7), (15), and (16) thus imply

$$\varepsilon^{\alpha} \sum_{T \in \mathcal{T}^{\varepsilon}} |T| \lesssim \Big| \bigcup_{T \in \mathcal{T}^{\varepsilon}} T \Big|, \quad \alpha := \frac{2[q(d-1)-2d]}{q-2}, \tag{17}$$

with implicit constants uniform in $\varepsilon \ll 1$.

The case $\alpha = 0$ of (17) corresponds to the case $q = \frac{2d}{d-1}$, the conjectured endpoint for the extension problem. The brief remainder of this section will be devoted to some geometric heuristics around (17), which is easier to visualize in the (equivalent) rescaled version

$$\varepsilon^{\alpha} \lesssim \Big| \bigcup_{T \in \mathcal{T}^{\varepsilon}} \tilde{T} \Big|, \tag{18}$$

in which tubes $\tilde{T} := \varepsilon^2 T$ have length 1 and width ε, and the sum of their volumes is about 1.

Inequality (18) is clearly false when $\alpha < 0$: a union cannot be larger than the sum of the volumes. Overly simplistic reasoning suggests that (18) might be true for $\alpha = 0$. Indeed, since two distinct tubes separate (as we have seen), it seems unlikely that a collection $\mathcal{T}^\varepsilon$ could possess sufficient overlap for the union to have much smaller measure than the sum of the volumes. However, a surprising 1920s example of Besicovitch showed that for all $d \geq 2$, there are configurations of tubes with such strong overlap that

$$|\bigcup_{T\in\mathcal{T}^\varepsilon} \widetilde{T}| = o_\varepsilon(1), \quad \text{as} \quad \varepsilon \searrow 0.$$

As the tubes $\widetilde{T}$, with $T \in \mathcal{T}^\varepsilon$, together contain length-one line segments pointing in (a large portion of) all possible directions, Besicovitch's example leads to the conclusion that there exist compact sets of measure zero containing a unit line segment in each direction. Such sets are called Besicovitch sets (and sets containing a unit line segment in every direction are called Kakeya sets).

One version of the Kakeya Conjecture is that for all $d \geq 2$, Besicovitch sets, though measure zero, are not too small, in the sense that (18) holds for all configurations $\mathcal{T}^\varepsilon$ and all $\alpha > 0$. This conjecture is only settled in dimension $d = 2$, and there is good evidence that the remaining cases are extremely difficult, with the current record for $d = 3$ only slightly better than $\alpha = \frac{1}{2}$. (The current record for α, which is due either to Katz–Łaba–Tao or to Katz–Zahl, has not been precisely computed.)

Capitalizing on Transversality: Multilinear Methods

Let $\tau_1, \dots, \tau_k$, $2 \leq k \leq d$ be balls in $\{|\xi| < 1\}$, of diameter comparable to 1, whose corresponding parabolic caps are k-transversal in the sense that $\|\nu_1 \wedge \dots \wedge \nu_k\| \sim 1$ whenever each ν_j is normal to the portion of $\mathbb{P}$ above τ_j. We can define a multilinear extension operator (somewhat confusingly called a multilinear restriction operator in the literature)

$$\mathcal{M}_{\mathbb{P}}(f_1, \dots, f_k)(x) := \prod_{j=1}^{k} \mathcal{E}_{\mathbb{P}}(f_j \chi_{\tau_j})(x), \tag{19}$$

and ask for inequalities of the form

$$\|\mathcal{M}_{\mathbb{P}}(f_1, \dots, f_k)\|_{\frac{q}{k}} \lesssim \prod_{j=1}^{k} \|f_j\|_p. \tag{20}$$

If (7) holds for p and q, then (20) is an immediate consequence of Hölder's inequality. Thus the interesting case is when the left side exhibits better-than-expected decay, $q < \frac{d+1}{d-1}p'$. Earlier, we sketched a heuristic argument that slight differences in the directions of tubes lead to separation and consequently the decay estimate (11), which in turn leads to the Stein–Tomas theorem. Transverse tubes separate more rapidly, and the relatively small size of their intersection makes improved decay plausible in (20). Improved decay is extremely useful because it makes linear inequalities that improve on Stein–Tomas accessible to L^2-based wave packet methods.

Variations on the multilinear idea have been around since early in the development of the subject, having played a role in the L^4 theory of Fefferman, Sjölin, and Carleson in the 1970s, and the work of Prestini, Drury, and Christ on restriction to curves in the 1980s. Bourgain, in the 1990s, was the first to improve on Stein–Tomas in higher dimensions by using bilinear estimates. These ideas were further developed by Moyua–Vargas–Vega and Tao–Vargas–Vega, culminating in near-optimal bilinear restriction inequalities for the cone (Wolff) and elliptic hypersurfaces (Tao) in the early 2000s. Since that time, there has been significant progress toward a bilinear theory for more general hypersurfaces, though major open questions remain. Somewhat better known are the developments on multilinear restriction estimates for elliptic surfaces, to which we devote the majority of this section. We will describe some challenges for more general hypersurfaces at the end of the section.

In the special case $p = 2$, it is conjectured [3] that (20) should hold for $q > \frac{2(d+k)}{d+k-2}$. Below that endpoint, the inequality fails, as can be seen by considering functions f_j formed by summing ε-caps with normal directions lying within ε of a fixed k-plane. There are $\varepsilon^{-(k-1)}$ such caps, each with norm roughly $\varepsilon^{-\frac{d-1}{2}}$, so

$$\prod_j \|f_j\|_2 \sim (\varepsilon^{-\frac{d+k-2}{2}})^k.$$

The extension of each f_j is comparable to 1 on the ε^{-1} neighborhood of an ε^{-2}-ball in the k-plane. As this (shared) cylinder has volume $(\varepsilon^{-2})^k(\varepsilon^{-1})^{d-k}$,

$$\|\mathcal{M}(f_1, \dots, f_k)\|_{\frac{q}{k}} \sim \varepsilon^{-(d+k)\frac{k}{q}},$$

and sending $\varepsilon \searrow 0$ in (20) leads to the claimed necessary condition.

The above-described multilinear restriction conjecture for $\mathbb{P}$ is settled when $k = 1$ (Stein–Tomas), when $k = 2$ (Tao), and when $k = d$ (Bennett–Carbery–Tao, [4]). The latter result also gives a nontrivial improvement for all $k \geq 3$, which we now state.

Theorem 1 ([4]). *Inequality* (20) *holds when* $p = 2$ *and* $q > \frac{2k}{k-1}$.

Of particular note is the fact that for $k = d$, the endpoint $\frac{2k}{k-1}$ matches the endpoint in the extension conjecture (12). Theorem 1 is based only on the transversality of the surface patches, not their curvature, and hence in its full generality (for which (19) is a special case) it is optimal.

We turn now to a heuristic description of the process of extracting linear inequalities from the bi- and multi-linear ones. The "bilinear-to-linear" argument of [19] expresses $|\mathcal{E}f|^q = |(\mathcal{E}f)^2|^{q/2}$ as the $\frac{q}{2}$ power of a double integral over $\mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$. The diagonal in this space has measure zero, so the double integral may be taken over (ξ, η) with $\xi \neq \eta$. Each such pair is associated with a pair of balls of radii about $|\xi - \eta|$, separated from one another by a distance about $|\xi - \eta|$, one containing ξ and one containing η. Separated balls have transverse normals because the gradient map is a diffeomorphism (thanks to the curvature). The bilinear theory applies to such pairs of balls after rescaling. Orthogonality and convexity arguments then make it possible to sum.

The geometry described above does not readily generalize to (for instance) the d-linear case, because d distinct points on $\mathbb{P}$ could have normals all transverse to one another, all arranged along some k-plane, or lying in some intermediate configuration. Because of the variety in the possible configurations, it took some years before multilinear restriction was used to obtain new linear estimates by Bourgain–Guth [7].

The Bourgain–Guth argument is via a transverse vs. linearly dependent dichotomy and induction on scales, for which we give a brief cartoon (which owes much to the exposition in [16]). The extension $\mathcal{E}f(x)$ may be thought of as a superposition of many tubes though x. There is either a significant contribution from d transverse tubes, or the bulk of the tubes lie near some hyperplane passing through x. The d-transversal case is favorable because d-linear extension estimates come with an exponent $q = \frac{2d}{d-1}$, which matches the extension conjecture. The hyperplane case is refined further to determine the right dimension of transversality for the tubes contributing to $\mathcal{E}f(x)$. We have already seen that one extreme, the d-transversal case, is favorable. At the other extreme, $\mathcal{E}f(x)$ is 1-transversal, i.e., parallel: it primarily comes from tubes whose corresponding caps lie within some small ball. This case is favorable as well, because of an argument known as induction on scales: parabolic rescaling expands the small ball to have radius 1 (so we nearly have the same problem again), but introduces an additional favorable term. This term arises because the method aims for $L^p \to L^q$ inequalities that do not scale (meaning that $\|\mathcal{E}_{\mathbb{P}}f\|_q$ and $\|f\|_p$ behave differently under rescaling).

The k-transversal case is intermediate between the two extremes. On the one hand, k-linear restriction comes with a minimal exponent of $q = \frac{2k}{k-1}$, which is smaller, hence better (closer to the conjectured $q > \frac{2d}{d-1}$) for large k; for small k, the exponent must be reduced further, leading to a loss. On the other hand, the tubes correspond to caps lying near a k-plane, and the k-plane can be covered by small balls, which can be parabolically rescaled. For this estimate, smaller k corresponds to fewer small balls in the cover, and hence a more favorable term from parabolic rescaling. For a certain range of exponents, the gains compensate for the losses, leading to new (at the time, but by now surpassed) estimates.

Among hypersurfaces, the elliptic case is generally the best studied, and there are currently few linear results beyond Stein–Tomas for hyperbolic surfaces and surfaces whose curvature varies. (As already noted, even the L^2 theory is not fully developed for such surfaces.) We give two examples to illustrate some of the challenges.

Hyperbolic surfaces may be ruled (expressible as unions of lines, such as the one-sheeted hyperboloid or the hyperbolic paraboloid), or have high order contact with certain lines. It is easy to produce lots of overlap among tubes with directions normal to a surface along some line within the surface, necessitating additional separation hypotheses for nontrivial bilinear restriction. In particular, in the case $k = 2$ of (20), to obtain a result matching that for the paraboloid, the images of the balls τ_1, τ_2 on the surface cannot be nearly collinear. Stronger hypotheses make for weaker theorems, leading to a more difficult bilinear-to-linear deduction; indeed, the natural analogue of the Tao–Vargas–Vega deduction has only been fully carried out for the hyperbolic paraboloid in $\mathbb{R}^3$. Moreover, the hyperbolic case does not perturb as well as the elliptic case, because the natural caps for the hyperbolic paraboloid are rectangles, which may be long and thin. Whereas perturbed balls (such as arise in the elliptic case) still look roughly like balls, perturbed rectangles may curve, leading to new difficulties. Some of these issues are described in [8] and the references therein.

An advantage in the hyperbolic case is that the multilinear extension result of Bennett–Carbery–Tao, Theorem 1, depends only on transversality, and hence readily generalizes to any surface (or subset thereof) on which the unit normal map is a local diffeomorphism. This property does not hold for surfaces on which the Gaussian curvature may vanish, and, moreover, variable curvature can lead to very different behavior on different parts of the surface. Even for convex polynomial hypersurfaces in dimension $d \geq 3$, how to deal with these issues to achieve results matching the elliptic case is a largely open question.

Discretization Via Decoupling

An important application of the multilinear perspective is in proving decoupling inequalities. These inequalities were introduced by Wolff in his work on the L^p regularity properties of the wave equation (the local smoothing problem), and a recent result of Bourgain–Demeter [5] established a stronger version of Wolff's inequalities in the optimal range. In this section, we will state and provide some background on the Bourgain–Demeter result and mention two recent applications of the method, one in analytic number theory, and one for restriction to fractal sets; an application to PDE is discussed in the final section.

Let f be a function on $\mathbb{R}^d$ with $\widehat{f}$ having compact support contained in the δ-neighborhood of $\mathbb{P}$, $N_\delta(\mathbb{P})$. We can discretize $N_\delta(\mathbb{P}) = \bigcup_{\theta \in \mathcal{P}_\delta} \theta$, where the θ's are finitely overlapping caps of thickness δ and width $\delta^{\frac{1}{2}}$. A partition of unity decomposes $f = \sum_\theta f_\theta$, with each $\widehat{f}_\theta$ supported on θ. By the triangle inequality, the simple estimate $\#\mathcal{P}_\delta \approx \delta^{-(d-1)/2}$, and Cauchy–Schwarz we have

$$\|f\|_\infty \leq \delta^{-(d-1)/4} (\sum_\theta \|f_\theta\|_\infty^2)^{\frac{1}{2}}, \tag{21}$$

while Plancherel and finite overlap of the caps immediately implies

$$\|f\|_2 \lesssim (\sum_{\theta \in \mathcal{P}_\delta} \|f_\theta\|_2^2)^{\frac{1}{2}}. \tag{22}$$

Both of these results are optimal, and an interpolation argument yields

$$\|f\|_p \lesssim \delta^{-\frac{d-1}{2}(\frac{1}{2}-\frac{1}{p})} (\sum_\theta \|f_\theta\|_p^2)^{\frac{1}{2}}, \quad 2 \leq p \leq \infty. \tag{23}$$

Inequality (23), it turns out, is not optimal, due to the curvature of $\mathbb{P}$.

The ℓ^2 decoupling theorem of Bourgain–Demeter [5] states that for f as above, the inequality

$$\|f\|_p \lesssim \delta^{-\frac{d-1}{4}+\frac{d+1}{2p}-\varepsilon} (\sum_{\theta \in \mathcal{P}_\delta} \|f_\theta\|_p^2)^{\frac{1}{2}} \tag{24}$$

holds for all $p \geq \frac{2(d+1)}{d-1}$ and $\varepsilon > 0$, which nearly eliminates the negative power of δ in (23) when $p = \frac{2(d+1)}{d-1}$.

It is no coincidence that the critical exponent $p = \frac{2(d+1)}{d-1}$ for (24) matches that for the Stein–Tomas theorem (i.e. the optimal L^2-based extension theorem), and an important application of (24) was to prove the discrete analogue of the Stein–Tomas theorem (which has applications to the periodic Schrödinger equation). Discrete versions of inequalities in harmonic analysis tend to be harder than the continuous versions, because delicate number theoretic issues come into play. (Compare, for instance, the difficulty in finding the measure of the real solution set of $x^a + y^a = z^a$ with finding the cardinality of its integer solution set.) These issues are somewhat easier to see in cases wherein the exponents involved are integers, so we now turn to a related example.

The Main Conjecture in Vinogradov's Mean Value Theorem is (was) the upper bound $\#J_{s,d}(N) \leq C_{\varepsilon,s,d} N^{s+\varepsilon} + N^{2s-\frac{d(d+1)}{2}+\varepsilon}$ on the number, $J_{s,d}(N)$, of solutions to the system of equations

$$j_1^i + \cdots + j_s^i = j_{s+1}^i + \cdots + j_{2s}^i, \quad 1 \leq i \leq d, \quad j \in [1,N]^{2s}.$$

This was a classical result for $d = 1, 2$, and was proved by Wooley in the case $d = 3$, with partial progress in higher dimensions, via efficient congruencing. However, a different approach, via decoupling associated to the curve $(t, t^2, \ldots, t^d)$, was used by Bourgain–Demeter–Guth in [6] to affirmatively resolve the conjecture in all dimensions.

The connection between the number theoretic problem and restriction can be seen by direct computation. Expanding

$$J_{s,d}(N) = \int_{[0,1]^d} |\sum_{j=1}^N \exp[2\pi i(x_1 j + x_2 j^2 + \cdots + x_d j^d)]|^{2s}\, dx,$$

and realizing the right as an L^q norm of the discrete extension operator

$$\mathscr{E}_\gamma a(x) := \sum_{j=1}^N \exp[2\pi i(x \cdot \gamma(j))] a_j,$$
$$\gamma(j) := (j, j^2, \ldots, j^d),\ a \in \mathbb{C}^N,$$

the conjectured bound on $J_{s,d}(N)$ is equivalent to the estimate

$$\|\mathscr{E}_\gamma a\|_{L^p([0,1]^d)} \lesssim (N^\varepsilon + N^{\frac{1}{2}-\frac{d(d+1)}{2p}+\varepsilon}) \|a\|_{\ell^2([1,N])},$$
$$p = 2s \geq 2, \tag{25}$$

where $a := (1, \ldots, 1)$. The estimate (25) in turn follows from decoupling.

Ideas from decoupling and work of Bourgain on $\Lambda(p)$ sets have recently been used by Łaba–Wang [15] to construct examples of fractal sets in $\mathbb{R}$, having arbitrary Hausdorff dimension in the interval $(0, 1)$, that obey extension inequalities in a range inaccessible to the generalized Stein–Tomas theorem from [1]. These examples are Cantor-like sets, wherein the deletion process in the standard Cantor construction is randomized. Randomization has been used in the past to construct interesting sets for Fourier restriction (further references are in [15]); the advantage of decoupling is that it helps to separate out the different pieces of the set at each scale.

Lines, Curves, and Varieties: Polynomial and Kakeya Methods

In this final section, we discuss some techniques that comprise the current state of the art for linear restriction to the paraboloid. These techniques use, in various ways, the fact that a tube is a thickened version of an algebraic object (a line) and hence should not intersect a given low-degree algebraic variety too many times. One reason algebraic varieties are relevant is that, as noted earlier, tubes lying close to k-planes provide the conjectured sharp examples for multilinear extension inequalities in the presence of curvature; controlling the extension of a function on lower dimensional sets is also relevant in PDE, as described at the end of this section.

The current best (smallest q) results for $\mathscr{E}_\mathbb{P}$ involve bounds for quantities called k-broad norms, which were introduced by Guth [11, 12]. These linearize the transversality utilized in multilinear estimates. We will not give the precise definition, but roughly, the k-broad norm counts only the part of $\mathscr{E}_{\mathbb{P}}f$ that genuinely comes from k-transversal parts of $\mathbb{P}$, and thus is closely related to the k-linear extension operator (19). Bounding these norms involves counting intersections of transversal tubes (a combinatorial perspective that traces back to the work of Bourgain and Wolff), and optimal estimates were proved for these norms by Guth in [11, 12] using a technique called polynomial partitioning.

Polynomial partitioning is a divide-and-conquer strategy for bounding the number of certain types of incidences (such as k-transversal intersections of tubes). Very roughly, the "Polynomial Ham Sandwich Theorem" allows one to find a low-ish degree algebraic variety whose complement is a union of cells that each contain roughly the same number of incidences. The cells are good objects for induction on the number of tubes: each tube is like a line and so by the Fundamental Theorem of Algebra only enters a few cells; therefore some cell meets only a small number of tubes, yielding a simpler problem. The variety contains two types of incidences: those occurring exclusively among tubes that are tangent to the variety and those involving at least one tube that intersects the variety transversally. The former situation sets up an induction on dimension, while transversal intersections are not too common because any given tube can only cross the variety a few times. The polynomial partitioning method originated in work of Guth–Katz, with important precursors in the works of Dvir and Clarkson–Edelsbrunner–Guibas–Sharir–Welzl; a detailed history of the method may be found in the introduction of [11].

The estimates for the k-broad norms in [11, 12] correspond to the conjectured optimal range for multilinear restriction to $\mathbb{P}$, and yield the same bounds for $\mathscr{E}_{\mathbb{P}}$ as would have been achieved via the methods of [7] using the optimal multilinear restriction conjecture. These bounds on $\mathscr{E}_{\mathbb{P}}$ have been improved in all except a handful of dimensions ([13, 20]), and are thought not to be optimal in any dimension, but their analogues are optimal for a variable coefficient generalization recently obtained by Guth–Hickman–Iliopoulou. In the variable coefficient generalization, tubes are replaced by neighborhoods of curves that can compress into low-dimensional sets in ways that lines cannot, as observed in [7] and earlier work of Wisewell.

Improvements over [12] have used polynomial partitioning in conjunction with additional geometric arguments that further limit the ability of straight-line tubes to compress into small sets. These include an argument of Wang in [20] that organizes the tubes lying along a hypersurface in $\mathbb{R}^3$ into "brooms," and one of Hickman–Rogers in higher dimensions [13] that utilizes a theorem of Katz–Rogers verifying what are called "polynomial Wolff axioms," which bound from below the size of a sub-algebraic set containing a large number of tubes.

The extension problem for the cone $\{(\xi, |\xi|) : 1 \leq |\xi| \leq 2\}$ is almost as well studied as that for the paraboloid, and the former problem is closely connected with the wave equation. The cone and the cylinder $[1,2] \times S^{d-2} \subseteq \mathbb{R}^d$ both have $d-2$ nonvanishing principal curvatures, and a variation on the Knapp example suggests the same restriction conjecture for these surfaces as for the sphere of one lower dimension. In fact, the product structure makes the extension problem for the cylinder exactly the same as the sphere of one lower dimension. However, the changing slope of the flat direction on the cone has facilitated much better progress on its restriction theory than that of the lower dimensional sphere, and the cone restriction problem has also been completely solved in dimensions $d = 3$ (Barceló, classical), $d = 4$ (Wolff, using the bilinear theory), and $d = 5$ (Ou–Wang, [18]). The latter result used polynomial partitioning to reduce q to the optimal range and the bilinear theory to obtain the optimal p range.

We close by describing the recent resolution of a question in PDE in which the size of a Fourier extension near some lower-dimensional set plays an important role. The solution to the linear Schrödinger equation $iu_t + \Delta u = 0$, with initial data $u(0) = u_0 \in L^2(\mathbb{R}^{d-1})$, is simply the extension $u(x,t) = (\mathscr{E}_{\mathbb{P}}\hat{u}_0)(x,t)$. Here, u_0 is the initial data in the sense that $\lim_{t\to 0} \|u(t) - u_0\|_{L^2} = 0$. Carleson's problem, posed in the late 1970s, asks for the precise number of derivatives needed to ensure that $\lim_{t\to 0} u(t,x) = u_0(x)$, for almost every x. In dimension $d = 2$, the problem was resolved relatively quickly by Carleson himself ($\frac{1}{4}$ derivative is sufficient) and Kenig–Dahlberg ($\frac{1}{4}$ derivative is necessary), while its (almost) complete resolution took nearly 40 years longer.

A.e. convergence follows from bounding the maximal operator

$$f \mapsto \sup_{0<t\leq 1} |\mathscr{E}_{\mathbb{P}}\hat{f}(x,t)|.$$

This maximal operator can be linearized as $\mathscr{E}_{\mathbb{P}}\hat{f}(x, t(x))$, which looks like the (functional) restriction of $\mathscr{E}_{\mathbb{P}}\hat{f}$ to a $(d-1)$-dimensional set. Du–Guth–Li [9] proved in dimension $d = 2$ (using polynomial partitioning), and Du–Zhang [10] proved in dimension $d \geq 3$ (using decoupling and multilinear restriction) that having more than $\frac{d-1}{2d}$ L^2 derivatives (in other words for $u_0 \in H^s$, $s > \frac{d-1}{2d}$), suffices for a.e. convergence of u to u_0. An earlier result of Bourgain showed that $s \geq \frac{d-1}{2d}$ derivatives is necessary, so Carleson's problem is resolved except at the critical value.

ACKNOWLEDGMENT. Jean Bourgain and Elias Stein passed away shortly before this article was submitted, and the author would like to express her deep gratitude and admiration for their beautiful mathematics and the subsequent work that it inspired.

References

[1] Bak J-G, Seeger A. Extensions of the Stein–Tomas theorem, *Math. Res. Lett.*, (18): 767–781, 2011. MR2831841

[2] Beckner W, Carbery A, Semmes S, Soria F. A note on restriction of the Fourier transform to spheres, *Bull. London Math. Soc.*, (21): 394–398, 1989. MR998638

[3] Bennett J. Aspects of multilinear harmonic analysis related to transversality, Harmonic analysis and partial differential equations, 1–28, *Contemp. Math.* 612, Amer. Math. Soc., Providence, RI, 2014. MR3204854

[4] Bennett J, Carbery A, Tao T. On the multilinear restriction and Kakeya conjectures, *Acta Math.*, (196): 261–302, 2006. MR2275834

[5] Bourgain J, Demeter C. The proof of the ℓ^2 decoupling conjecture, *Ann. of Math.*, (182): 351–389, 2015. MR3374964

[6] Bourgain J, Demeter C, Guth L. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, *Ann. of Math.*, (184): 633–682, 2016. MR3548534

[7] Bourgain J, Guth L. Bounds on oscillatory integral operators based on multilinear estimates, *Geom. Funct. Anal.*, (21): 1239–1295, 2011. MR2860188

[8] Buschenhenke S, Müller D, Vargas A. A Fourier restriction theorem for a perturbed hyperboloid, preprint, arXiv:1803.02711.

[9] Du X, Guth L, Li X. A sharp Schrödinger maximal estimate in $\mathbb{R}^2$, *Ann. of Math.*, (186): 607–640, 2017. MR3702674

[10] Du X, Zhang R. Sharp L^2 estimate of Schrödinger maximal function in higher dimensions, preprint, arXiv:1805.02775, to appear in the *Annals of Mathematics*.

[11] Guth L. A restriction estimate using polynomial partitioning, *J. Amer. Math. Soc.*, (29): 371–413, 2016. MR3454378

[12] Guth L. Restriction estimates using polynomial partitioning II, *Acta Math.*, (221): 81–142, 2018. MR3877019

[13] Hickman J, Rogers KM. Improved Fourier restriction estimates in higher dimensions, preprint, arXiv:1807.10940.

[14] Ikromov IA, Müller D. Fourier restriction for hypersurfaces in three dimensions and Newton polyhedra, *Annals of Mathematics Studies*, (194), Princeton University Press, Princeton, NJ, 2016. MR3524103

[15] Łaba I, Wang H. Decoupling and near-optimal restriction estimates for Cantor sets, *Int. Math. Res. Not.*, 2944–2966, 2018. MR3801501

[16] Lee S. Square function estimates for the Bochner-Riesz means, *Anal. PDE*, (11): 1535–1586, 2018. MR3803718

[17] Müller D, Ricci F, Wright JR. A maximal restriction theorem and Lebesgue points of functions in $\mathcal{F}(L^p)$, preprint, arXiv:1612.04880, to appear in *Revista Math. Iberoamericana*.

[18] Ou Y, Wang H. A cone restriction estimate using polynomial partitioning, preprint, arXiv:1704.05485.

[19] Tao T, Vargas A, Vega L. A bilinear approach to the restriction and Kakeya conjectures, *J. Amer. Math. Soc.*, (11): 967–1000, 1998. MR1625056

[20] Wang H. A restriction estimate in $\mathbb{R}^3$ using brooms, preprint, arXiv:1802.04312.

Betsy Stovall

Credits

The opening image is courtesy of Getty Images.
Author photo is courtesy of Betsy Stovall.