Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 2 (1535–1556)

DECOUPLINGS AND APPLICATIONS

Ciprian Demeter

Abstract

We describe a Fourier analytic tool that has found a large number of applications in Number Theory, Harmonic Analysis and PDEs.

1 Introduction

The circle of ideas described in this note have grown inside the framework of *restriction theory*. This area of harmonic analysis was born in the late 60s, when Elias Stein has considered the problem of restricting the Fourier transform of an L^p function $F : \mathbb{R}^n \to \mathbb{C}$ to the sphere $\mathbb{S}^{n-1}$. When $p = 1$, the Fourier transform is always a continuous function, its value is well defined at each point. At the other extreme, when $p = 2$, $\widehat{F}$ is merely measurable, so restricting it to a set of Lebesgue measure zero such as $\mathbb{S}^{n-1}$ is meaningless. It turns out that the range $1 < p < 2$ hosts a completely new phenomenon. A plethora of restriction-type estimates exist in this range, for a wide variety of curved manifolds other than the sphere. These are quantified by various operator bounds on the so-called *extension operator*, to be introduced momentarily.

A major breakthrough in the analysis of the extension operator came with the discovery of its relation to the quantitative forms of the *Kakeya set conjecture*. In one of its simplest forms, this conjecture asserts that each subset of $\mathbb{R}^n$ containing a unit line segment in every direction must have full Hausdorff dimension n. This is trivial when $n = 1$, relatively easy to prove when $n = 2$, and wide open for $n \geq 3$. The quantitative formulations of the conjecture involve estimating the overlap of collections of congruent tubes of arbitrary orientations. The afore-mentioned extension operator has an oscillatory nature, but it can be decomposed into pieces which have roughly constant magnitude on appropriate tubes. Then one can gain valuable information by understanding the worst conspiracies that tubes can use to maximize their overlap. Using the intuition from the case when tubes

The author is partially supported by the NSF Grant DMS-1161752.
MSC2010: primary 42B20; secondary 11L05, 42B37.

are replaced with lines can sometimes be helpful, though it has been well documented that the thickness of the tubes creates significant additional complications.

It turns out that the overlap question is easier to understand if one intersects families of tubes with very separated directions. This property will be called *transversality*. A critical role in the arguments presented in the following is played by the multilinear Kakeya estimate of Bennett, Carbery, and Tao [2006], which proves a sharp bound on the intersection of n transverse families of tubes. The way to harness the power of multilinear estimates in order to prove linear ones was explained by Bourgain and Guth in the fundamental paper Bourgain and Guth [2011].

A built-in feature of any restriction estimate is that of *scale*. Scales arise by localizing the operator to spatial balls of finite radius. The operator norms at various scales are typically compared using a process called *induction on scales*. A bootstrapping argument forces these operator norms to only grow mildly with the scale. Sometimes an ϵ removal argument is available to completely eliminate this dependence. In other cases, such as with decouplings, finding an ϵ removal mechanism continues to remain a challenge. *Parabolic rescaling* and its variants is a key tool that allows moving back and forth between different scales. This exploits the invariance of the manifold under certain affine transformations which interact well with the Fourier transform.

We will start by analyzing a few classical exponential sum estimates and will continue by showing how decouplings lead to new ones. We close by presenting the proof of the simplest decoupling at critical exponent, an L^6 result for the parabola.

2 Stein–Tomas–Strichartz and exponential sum estimates on small balls

We will denote by $e(z)$ the quantity $e^{2\pi i z}$, $z \in \mathbb{R}$. For $F \in L^1(\mathbb{R}^n)$ we recall its Fourier transform

$$\widehat{F}(\xi) = \int_{\mathbb{R}^n} F(x) e(-x \cdot \xi) dx, \quad \xi \in \mathbb{R}^n.$$

Let D be an open cube, ball or annulus in $\mathbb{R}^m$, $1 \le m \le n-1$. Given a smooth function $\psi : D \to \mathbb{R}^{n-m}$ we define the manifold

$$\mathfrak{M} = \mathfrak{M}^{\psi} = \{(\xi, \psi(\xi)) : \xi \in D\} \tag{1}$$

and its associated extension operator for $f : D \to \mathbb{C}$

$$Ef(x) = E^{\mathfrak{M}} f(x) = \int_D f(\xi) e(\bar{x} \cdot \xi + x^* \cdot \psi(\xi)) d\xi, \quad x = (\bar{x}, x^*) \in \mathbb{R}^m \times \mathbb{R}^{n-m}.$$

For a subset $S \subset D$ we will denote $E(f 1_S)$ by $E_S f$. The defining formula shows that $E_S f$ is the Fourier transform of the pullback of the measure $f d\xi$ from $\mathbb{R}^m$ to the manifold.

Examples of interesting manifolds arising this way include the truncated paraboloid

$$\mathbb{P}^{n-1} = \{(\xi_1, \ldots, \xi_{n-1}, \xi_1^2 + \ldots + \xi_{n-1}^2) : |\xi_i| < 1\},$$

the hemispheres

$$\mathbb{S}_{\pm}^{n-1} = \{(\xi, \pm\sqrt{1-|\xi|^2}),\ |\xi| < 1\},$$

the truncated cone

$$\mathbb{C}o^{n-1} = \{(\xi, |\xi|) :\ 1 < |\xi| < 2\}$$

and the moment curve

$$\Gamma_n = \{(\xi, \xi^2, \ldots, \xi^n) :\ \xi \in (0,1)\}.$$

To provide the reader with some motivation for considering the extension operator, let $\Psi(\bar{x}, t)$ be the solution of the free Schrödinger equation with initial data g

$$\begin{cases} 2\pi i \Psi_t = \Delta_{\bar{x}} \Psi, & (\bar{x}, t) \in \mathbb{R}^{n-1} \times \mathbb{R} \\ \Psi(\bar{x}, 0) = g(\bar{x}) \end{cases}.$$

A simple computation reveals that $\Psi(\bar{x}, t) = E^{\mathbb{P}^{n-1}} f(\bar{x}, t)$, where $f = \widehat{g}$. A similar relation exists between the cone and the wave equation and also between the sphere and the Helmholtz equation.

The following theorem provides the first wave of restriction estimates that were ever obtained. They are due to Stein and Tomas in the case of the (hemi)sphere, and to Strichartz in the case of the paraboloid. What makes them special is the fact that the function f is estimated in L^2. The core of the argument relies on the TT^* method.

Theorem 2.1 (Stein [1993], Strichartz [1977], Tomas [1975]). *Let E be the extension operator associated with either $\mathbb{S}_{\pm}^{n-1}$ or $\mathbb{P}^{n-1}$. Then for each $p \geq \frac{2(n+1)}{n-1}$ and $f \in L^2(D)$ we have*

$$\|Ef\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_2.$$

There is an equivalent way to rephrase this theorem, using a rather standard local to global mechanism. The resulting inequality is an example of *discrete restriction estimate.*

Corollary 2.2. *Let $\mathfrak{M}$ be either $\mathbb{S}^{n-1}$ or $\mathbb{P}^{n-1}$. For each $R \geq 1$, each collection $\Lambda \subset \mathfrak{M}$ consisting of $\frac{1}{R}$-separated points, each sequence $a_\lambda \in \mathbb{C}$ and each ball B_R of radius R in $\mathbb{R}^n$ we have*

$$\|\sum_{\lambda\in\Lambda} a_\lambda e(\lambda \cdot x)\|_{L^{\frac{2(n+1)}{n-1}}(B_R)} \lesssim R^{\frac{n-1}{2}} \|a_\lambda\|_{l^2}. \tag{2}$$

If we introduce the normalized L^p norms

$$\|F\|_{L^p_\sharp(B)} := (\frac{1}{|B|}\int_B |F|^p)^{1/p},$$

then (2) says that

$$\|\sum_{\lambda\in\Lambda} a_\lambda e(\lambda \cdot x)\|_{L_\sharp^{\frac{2(n+1)}{n-1}}(B_R)} \lesssim R^{\frac{n-1}{2(n+1)}} \|a_\lambda\|_{l^2}. \tag{3}$$

The exponent of R is sharp, as can be seen by taking $a_\lambda \equiv 1$ and Λ a maximal $\frac{1}{R}$-separated set.

We will call the scale R of the spatial balls B_R the *uncertainty principle scale*, as it is the reciprocal of the scale that separates the frequency points λ. Since averages over large balls are controlled by averages over smaller balls, inequality (3) persists if B_R is replaced with $B_{R'}$ for $R' \geq R$. However, we will see that averaging the exponential sums over larger spatial balls will lead to improved estimates. This will be a direct consequence of the new decoupling phenomenon. In short, the waves $e(\lambda \cdot x)$ oscillate in different directions, and annihilate each other better if they are given more room to interact.

3 A first look at decouplings

Let $(f_j)_{j=1}^N$ be N elements of a Banach space X. The triangle inequality

$$\|\sum_{j=1}^N f_j\|_X \leq \sum_{j=1}^N \|f_j\|_X$$

is universal, it does not incorporate any possible cancelations between the f_j. When combined with the Cauchy–Schwarz inequality it leads to

$$\|\sum_{j=1}^N f_j\|_X \leq N^{\frac12}(\sum_{j=1}^N \|f_j\|_X^2)^{1/2}.$$

But if X is a Hilbert space (e.g. $X = L^2(\mathbb{T})$) and if f_j are pairwise orthogonal (e.g. $f_j(x) = e(xj)$) then we have a stronger inequality (in fact an equality)

$$\|\sum_{j=1}^{N} f_j\|_X \le (\sum_{j=1}^{N} \|f_j\|_X^2)^{1/2}.$$

We will call such an inequality $l^2(X)$ *decoupling*. It is natural to ask if there is something analogous in $L^p(\mathbb{R}^n)$ when $p \neq 2$, in the absence of Hilbert space orthogonality.

The answer is yes. Our first example ($X = L^4[0,1]$) is due to the "bi-orthogonality" of the squares. Note that we lose N^ϵ (and some loss in N is in fact necessary in this case), but this will be acceptable in our definition of decoupling.

Theorem 3.1 (Discrete $l^2(L^4)$ decoupling for squares). *For each $\epsilon > 0$, the following decoupling holds*

$$\|\sum_{j=1}^{N} a_j e(j^2 x)\|_{L^4[0,1]} \lesssim_\epsilon N^\epsilon (\sum_{j=1}^{N} \|a_j e(j^2 x)\|_{L^4[0,1]}^2)^{1/2} = N^\epsilon \|a_j\|_{l^2}.$$

Proof. We present the argument in the case $a_j = 1$, the general case requires only minor modifications. By raising to the fourth power, the left hand side equals

$$\int_0^1 \sum_{1 \le j_i \le N} e((j_1^2 + j_2^2 - j_3^2 - j_4^2)x)dx = \sum_{1 \le j_1, j_2 \le N} |\{(j_3, j_4) : \ j_3^2 + j_4^2 = j_1^2 + j_2^2\}| \lesssim_\epsilon N^{2+\epsilon}.$$

The last inequality follows since the equation

$$j_3^2 + j_4^2 = A$$

has $\lesssim_\epsilon A^\epsilon$ solutions, Grosswald [1985].

□

The second well known example relies on the "multi-orthogonality" of the sequence 2^j.

Theorem 3.2 (Discrete L^p decoupling for lacunary exponential sums). *For $1 \le p < \infty$ and $a_j \in \mathbb{C}$*

$$\|\sum_{j=1}^{N} a_j e(2^j x)\|_{L^p[0,1]} \sim_p (\sum_{j=1}^{N} \|a_j e(2^j x)\|_{L^p[0,1]}^2)^{1/2} = \|a_j\|_{l^2}.$$

These easy examples are of arithmetic structure. We will develop tools that do not depend on this restriction. The reader will notice that what lies behind both examples is the fact that there is increasing level of separation between higher frequencies (squares, powers of 2). We will see that in higher dimensions quasi-uniform separation will suffice, as long as the frequencies lie on a curved manifold.

We close this section with one of the most important results in classical harmonic analysis, perhaps in the entire mathematics. It is a consequence and a continuous reformulation of Theorem 3.2.

Theorem 3.3 (Littlewood–Paley theorem). *Given $f:\mathbb{R}\to\mathbb{C}$, let*

$$P_j f(x)=\int_{I_j}\widehat{f}(\xi)e(x\xi)d\xi$$

be its Fourier projection on $I_j=[2^j,2^{j+1}]\cup[-2^{j+1},-2^j]$, for $j\in\mathbb{Z}$. Then for each $1<p<\infty$

$$\|f\|_{L^p(\mathbb{R})}\sim_p\|(\sum_j|P_jf|^2)^{\frac12}\|_{L^p(\mathbb{R})}.$$

For our purposes, it suffices to note that, when combined with Minkowski's inequality, the Littlewood–Paley theorem leads to the following $l^2(L^p)$ decoupling on the real line, for $p\ge 2$

$$\|f\|_{L^p(\mathbb{R})}\lesssim(\sum_j\|P_jf\|^2_{L^p(\mathbb{R})})^{\frac12}.$$

4 Fourier analytic decouplings

For a ball (or cube) B_R in $\mathbb{R}^n$ with center c and radius (side length) R, we will denote by $w_{B_R}(x)$ a weight of the form $(1+\frac{|x-c|}{R})^{-C}$, for some large unspecified C. This can be thought of as being a smooth approximation of 1_{B_R}.

Fix a manifold $\mathfrak{M}=\mathfrak{M}^\psi$ as in (1) and let $f:D\to\mathbb{C}$. If we partition D into sets τ then we may write

$$E^{\mathfrak{M}}f=\sum_\tau E^{\mathfrak{M}}_\tau f.$$

Roughly speaking, $E^{\mathfrak{M}}_\tau f(x)$ has the oscillatory phase $e(x\cdot(\xi_\tau,\psi(\xi_\tau)))$, where ξ_τ is a point in τ. If $\mathfrak{M}$ has some curvature, which is the same as saying that ψ is "far" from being affine, then there will be lots of cancellations between the components $E^{\mathfrak{M}}_\tau f(x)$. This will be formalized by a Fourier decoupling, which (for now) takes the following rather vague conjectural form.

Conjecture 4.1 (Fourier decoupling). *Let $\mathfrak{M}$ be sufficiently curved. Then there is a critical index $p_c > 2$ and some $q \geq 1$ so that for each partition $\mathcal{P}_\delta$ of the domain D into N "caps" τ of "size" δ we have*

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_{\tau \in \mathcal{P}_\delta} \|E_\tau^{\mathfrak{M}} f\|_{L^p(w_{B_R})}^2\Big)^{1/2} \quad (l^2(L^p) \text{ decoupling})$$

or (alternatively)

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon} N^{\frac{1}{2}-\frac{1}{p}}\Big(\sum_{\tau \in \mathcal{P}_\delta} \|E_\tau^{\mathfrak{M}} f\|_{L^p(w_{B_R})}^p\Big)^{1/p} \quad (l^p(L^p) \text{ decoupling})$$

for each ball B_R with radius $R \geq \delta^{-q}$ and each $2 \leq p \leq p_c$.

The presence of w_{B_R} on the right hand side is probably necessary, but completely harmless for applications.

Note that an l^2 decoupling always implies an l^p decoupling, due to the Cauchy-Schwarz inequality. However, sometimes the former is false and the latter is true. In most applications, an l^p decoupling is as good as an l^2 decoupling.

The shape of the "caps", the precise meaning of "size" as well as the values of q and of the critical exponent p_c depend on the manifold $\mathfrak{M}$. Due to orthogonality considerations, there is always a decoupling for $p = 2$, even for flat manifolds (hyperplanes). In this latter case however, considering $f \equiv 1$ shows that there is no decoupling outside L^2, so $p_c = 2$.

The formulation of Conjecture 4.1 is vague in many ways. There are interesting examples of manifolds that are known to simultaneously host different decoupling phenomena, corresponding to different values of p_c, q and for different types of caps. While, as observed in Section 2, restriction inequalities are associated with the uncertainty principle scale ($q = 1$), the most genuine decouplings will happen at spatial scales of magnitude $q \geq 2$.

The first to consider a Fourier decoupling was Wolff [2000]. He proved an $l^p(L^p)$ decoupling for the cone $\mathbb{C}o^2$ when $p > 74$, by masterfully combining Fourier analytic and incidence geometric arguments. In his theorem the caps are thin annular sectors of dimensions $\sim \delta$ and 1. Wolff also showed that his decoupling has consequences for the local smoothing of the solutions to the wave equation. Subsequent developments for the higher dimensional cone and other manifolds, prior to the work we are about the describe here, have appeared in Łaba and Wolff [2002], Laba and Pramanik [2005], Garrigós and Seeger [2009], Garrigós and Seeger [2010], Pramanik and Seeger [2007], Bourgain [2013] and Demeter [n.d.].

The first full range results for any manifold came in our joint work Bourgain and Demeter [2015] with Jean Bourgain.

Theorem 4.2. *Assume $\mathfrak{M}$ has positive definite second fundamental form (e.g. $\mathbb{S}^{n-1}$, $\mathbb{P}^{n-1}$). For any partition of the domain D into square-like caps τ with diameter δ we have*

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E_\tau^{\mathfrak{M}} f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-2}$ and $2 \leq p \leq \frac{2(n+1)}{n-1}$.

The proof of this for $\mathbb{P}^1$ will be presented in Section 6. Quite surprisingly, we were able to use this result for $\mathbb{P}^{n-1}$ as a black box, in order to derive the sharp result for the cone in all dimensions, thus closing the program initiated by Wolff. Let us get a glimpse into our argument for $\mathbb{C}o^2$. After a rotation, the equation of $\mathbb{C}o^2$ can be rewritten as $z = \frac{x^2}{y}$. The y-slices are parabolas with roughly the same curvature. This forces small pieces of the cone to be close to parabolic cylinders. One may combine the decoupling for the parabola with Fubini in the zero curvature direction to gradually separate the cone into smaller pieces. This argument is very different from Wolff's, in that it does not require any incidence geometry.

Theorem 4.3 (Bourgain and Demeter [2015]). *Let $\mathfrak{M} = \mathbb{C}o^{n-1}$ be the cone. For any partition of the domain $D = \{|\xi| \sim 1\}$ into sectors τ with angular width δ we have*

$$\|E^{\mathbb{C}o^{n-1}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E_\tau^{\mathbb{C}o^{n-1}} f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-2}$ and $2 \leq p \leq \frac{2n}{n-2}$. The range for p is sharp.

Another milestone of decoupling theory was the resolution of curves with torsion, in collaboration with Bourgain and Guth. More precisely, consider $\Phi : [0,1] \to \mathbb{R}^n$,

$$\Phi(\xi) = (\phi_1(\xi), \dots, \phi_n(\xi))$$

with $\phi_i \in C^n([0,1])$ and such that the Wronskian $W(\phi_1', \dots, \phi_n')(\xi)$ is nonzero on $[0,1]$. One example is the moment curve Γ_n. Let E^Φ be the associated extension operator.

Theorem 4.4 (Bourgain, Demeter, and Guth [2016]). *Partition $[0,1]$ into intervals τ of length $\sim \delta$. Then*

$$\|E^{\Phi} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E_\tau^{\Phi} f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-n}$ and $2 \leq p \leq n(n+1)$. The range for p is sharp.

The proof for $\Gamma_2 = \mathbb{P}^1$ appeared in Bourgain and Demeter [2015], while for $n \geq 3$ in Bourgain, Demeter, and Guth [2016]. The extension to the arbitrary Φ with torsion is explained in Section 4 of Bourgain and Demeter [2017].

Decouplings for a wide variety of other manifolds have been proved in Bourgain and Demeter [ibid.], Bourgain and Demeter [2016a], Bourgain and Demeter [2016b], Bourgain [2017], Bourgain, Demeter, and Guo [2017], Demeter, Guo, and Shi [n.d.], Bourgain and Watt [2017], Bourgain and Watt [n.d.], Guo and Oh [n.d.], and the list is rapidly growing.

5 Applications: Exponential sums on large balls

It turns out that there is a very simple mechanism that allows decouplings to imply exponential sum estimates that are often sharp. Essentially, one applies the decoupling to a weighted combination of (approximations of) Dirac deltas. In this regard each decoupling seems to be stronger than the exponential sum estimate it implies, the author is not aware of any argument that reverses the implication.

Theorem 5.1. *Let $\mathfrak{M} = \mathfrak{M}^{\psi}$. Consider a partition $\mathcal{P}_\delta$ as in Conjecture 4.1, with $N = |\mathcal{P}_\delta|$. Let $\xi_\tau \in \tau$ for each $\tau \in \mathcal{P}_\delta$ and let $\lambda_\tau = (\xi_\tau, \psi(\xi_\tau))$ be the corresponding point on $\mathfrak{M}$. Then for each $2 \leq p \leq p_c$, $a_\tau \in \mathbb{C}$ and each $R \geq \delta^{-q}$ we have*

$$\|\sum_{\tau \in \mathcal{P}_\delta} a_\tau e(\lambda_\tau \cdot x)\|_{L^p_\sharp(B_R)} \lesssim_\epsilon \delta^{-\epsilon} \|a_\tau\|_{l^2},$$

if the l^2 version of the decoupling in Conjecture 4.1 holds true, and

$$\|\sum_{\tau \in \mathcal{P}_\delta} a_\tau e(\lambda_\tau \cdot x)\|_{L^p_\sharp(B_R)} \lesssim_\epsilon \delta^{-\epsilon} N^{\frac{1}{2}-\frac{1}{p}} \|a_\tau\|_{l^p},$$

if the l^p version of the conjecture holds instead.

Proof. Apply the conjecture to functions of the form $f(\xi) = \sum_{\tau \in \mathcal{P}_\delta} a_\tau 1_{B(\xi_\tau, r)}(\xi)$ and let $r \to 0$. The computation is straightforward. □

One notable feature of this estimate is that it does not assume anything else about ξ_τ other than the separation guaranteed by the pairwise disjointness of the caps. In particular, the points need not belong to a rescaled lattice. This is indicative of the fact that our methods do not involve number theory, and that in fact sometimes they transcend the barrier that is currently accessible using number theoretic methods.

Let us now consider a few particular cases of interest.

5.1 Stricharz estimates. An application of Theorem 5.1 to $\mathbb{S}^{n-1}$ and $\mathbb{P}^{n-1}$ leads to the following corollary.

Corollary 5.2. *For each $R \geq 1$, each collection Λ consisting of $\frac{1}{R}$-separated points on either $\mathbb{S}^{n-1}$ or $\mathbb{P}^{n-1}$ and each ball $B_{R'}$ of radius $R' \geq R^2$ in $\mathbb{R}^n$ we have*

$$\|\sum_{\lambda \in \Lambda} a_\lambda e(\lambda \cdot x)\|_{L_\sharp^{\frac{2(n+1)}{n-1}}(B_{R'})} \lesssim_\epsilon R^\epsilon \|a_\lambda\|_{l^2}. \tag{4}$$

Comparing this with (3) shows that large ball averages get smaller.

This corollary leads to sharp Strichartz estimates in the periodic and quasi-periodic case. More precisely, fix $\frac{1}{2} < \theta_1, \dots, \theta_{n-1} < 2$ either rational or irrational. For $\phi \in L^2(\mathbb{T}^{n-1})$ consider its Laplacian

$$\Delta\phi(x_1, \dots, x_{n-1}) =$$

$$\sum_{(\xi_1,\dots,\xi_{n-1}) \in \mathbb{Z}^{n-1}} (\xi_1^2\theta_1 + \dots + \xi_{n-1}^2\theta_{n-1})\hat{\phi}(\xi_1, \dots, \xi_{n-1}) e(\xi_1 x_1 + \dots + \xi_{n-1}x_{n-1})$$

on the torus $\prod_{i=1}^{n-1} \mathbb{R}/(\theta_i\mathbb{Z})$. Let also

$$e^{it\Delta}\phi(x_1, \dots, x_{n-1}, t) =$$

$$\sum_{(\xi_1,\dots,\xi_{n-1}) \in \mathbb{Z}^{n-1}} \hat{\phi}(\xi_1, \dots, \xi_{n-1}) e(x_1\xi_1 + \dots + x_{n-1}\xi_{n-1} + t(\xi_1^2\theta_1 + \dots + \xi_{n-1}^2\theta_{n-1}))$$

be the solution of the Schrödinger equation in this context. We have the following result. When $p > \frac{2(n+1)}{n-1}$, the N^ϵ loss can be removed, see Bourgain and Demeter [2015] and Killip and Vişan [2016].

Theorem 5.3 (Strichartz estimates for rational and irrational tori, Bourgain and Demeter [2015]). *Let $\phi \in L^2(\mathbb{T}^{n-1})$ with $\operatorname{supp}(\hat{\phi}) \subset [-N, N]^{n-1}$. Then for each $\epsilon > 0$ and $p \geq \frac{2(n+1)}{n-1}$ we have*

$$\|e^{it\Delta}\phi\|_{L^p(\mathbb{T}^{n-1}\times[0,1])} \lesssim_\epsilon N^{\frac{n-1}{2} - \frac{n+1}{p} + \epsilon}\|\phi\|_2, \tag{5}$$

and the implicit constant does not depend on N.

Proof. It suffices to consider the case $p = \frac{2(n+1)}{n-1}$. For $-N \leq \xi_1, \dots, \xi_{n-1} \leq N$ define $\eta_i = \frac{\theta_i^{1/2}\xi_i}{4N}$ and $a_\eta = \hat{\phi}(\xi)$. A simple change of variables shows that

$$\int_{\mathbb{T}^{n-1}\times[0,1]} |e^{it\Delta}\phi|^p \lesssim$$

$$\frac{1}{N^{n+1}}\int_{\substack{|y_1|,\dots,|y_{n-1}|\le 8N\\ y_n\in I_{N^2}}}|\sum_{\eta_1,\dots,\eta_{n-1}}a_\eta e(y_1\eta_1+\ldots+y_{n-1}\eta_{n-1}+y_n(\eta_1^2+\ldots\eta_{n-1}^2))|^p dy_1\ldots dy_n,$$

where I_{N^2} is an interval of length $\sim N^2$. By periodicity in the $y_1,\dots,y_{n-1}$ variables we bound the above by

$$\frac{1}{N^{n+1}N^{n-1}}\int_{B_{N^2}}|\sum_{\eta_1,\dots,\eta_{n-1}}a_\eta e(y_1\eta_1+\ldots+y_{n-1}\eta_{n-1}+y_n(\eta_1^2+\ldots\eta_{n-1}^2))|^p dy_1\ldots dy_n,$$

for some ball B_{N^2} of radius $\sim N^2$. Our result will follow once we note that the points

$$(\eta_1,\dots,\eta_{n-1},\eta_1^2+\ldots\eta_{n-1}^2)$$

are $\sim\frac{1}{N}$ separated on $\mathbb{P}^{n-1}$ and then apply Corollary 5.2 with $R'\sim N^2$.

$\square$

5.2 Diophantine inequalities and the Vinogradov Mean Value Theorem. An application of Theorem 5.1 to the moment curve gives the following exponential sum estimate.

Corollary 5.4. *For each $1\le i\le N$, let t_i be a point in $(\frac{i-1}{N},\frac{i}{N}]$. Then for each $R\gtrsim N^n$ and each $p\ge 2$ we have*

$$\Big(\frac{1}{|B_R|}\int|\sum_{i=1}^N a_i e(x_1t_i+x_2t_i^2+\ldots+x_nt_i^n)|^p w_{B_R}(x)dx_1\ldots dx_n\Big)^{\frac1p}\lesssim$$

$$\lesssim_\epsilon (N^\epsilon+N^{\frac12(1-\frac{n(n+1)}{p})+\epsilon})\|a_i\|_{l^2(\{1,\dots,N\})},\tag{6}$$

and the implicit constant does not depend on N, R and a_i.

For each $1\le i\le N$ consider some real numbers $i-1<X_i\le i$. We do not insist that X_i are integers. Let $S_X=\{X_1,\dots,X_N\}$. For each $s\ge 1$, denote by $J_{s,n}(S_X)$ the number of solutions of the following system of inequalities

$$|X_1^i+\ldots+X_s^i-(X_{s+1}^i+\ldots+X_{2s}^i)|\le N^{i-n},\quad 1\le i\le n\tag{7}$$

with $X_i\in S_X$.

Corollary 5.5. *For each integer $s\ge 1$ and each S_X as above we have that*

$$J_{s,n}(S_X)\lesssim_\epsilon N^{s+\epsilon}+N^{2s-\frac{n(n+1)}{2}+\epsilon},$$

where the implicit constant does not depend on S_X.

Proof. Let $\phi : \mathbb{R}^n \to [0, \infty)$ be a positive Schwartz function with positive Fourier transform satisfying $\widehat{\phi}(\xi) \geq 1$ for $|\xi| \lesssim 1$. Define $\phi_N(x) = \phi(\frac{x}{N})$. Using the Schwartz decay, (6) with $a_i = 1$ implies that for each $s \geq 1$

$$(\frac{1}{|B_{N^n}|}\int_{\mathbb{R}^n} \phi_{N^n}(x)|\sum_{i=1}^{N} e(x_1 t_i + \ldots + x_n t_i^n)|^{2s} dx_1 \ldots dx_n)^{\frac{1}{2s}}$$

$$\lesssim_\epsilon N^{\frac{1}{2}+\epsilon} + N^{1-\frac{n(n+1)}{4s}+\epsilon}, \tag{8}$$

whenever $t_i \in [\frac{i-1}{N}, \frac{i}{N})$. Apply (8) to $t_i = \frac{X_i}{N}$. Let now

$$\phi_{N,1}(x_1, x_2, \ldots, x_n) = \phi(\frac{x_1}{N^{n-1}}, \frac{x_2}{N^{n-2}}, \ldots, x_n).$$

It suffices to consider the case $s = \frac{n(n+1)}{2}$. After making a change of variables and expanding the product, the term

$$\int_{\mathbb{R}^n} \phi_{N^n}(x)|\sum_{i=1}^{N} e(x_1 t_i + \ldots + x_n t_i^n)|^{2s} dx_1 \ldots dx_n$$

can be written as the sum over all $X_i \in S_X$ of

$$N^{\frac{n(n+1)}{2}} \int_{\mathbb{R}^n} \phi_{N,1}(x) e(x_1 Z_1 + \ldots + x_n Z_n) dx_1 \ldots dx_n,$$

where

$$Z_i = X_1^i + \ldots + X_s^i - (X_{s+1}^i + \ldots + X_{2s}^i).$$

Each such term is equal to

$$N^{n^2} \widehat{\phi}(N^{n-1} Z_1, N^{n-2} Z_2, \ldots, Z_n).$$

Recall that this is always positive, and in fact greater than N^{n^2} at least $J_{s,n}(S_X)$ times. It now suffices to use (8).

□

The special case of Corollary 5.5 when $X_i = i$ and the inequalities (7) are replaced with equalities

$$X_1^i + \ldots + X_s^i = X_{s+1}^i + \ldots + X_{2s}^i, \quad 1 \leq i \leq n$$

was known as the Main Conjecture in the Vinogradov Mean Value Theorem. The case $n = 2$ is very easy, while the case $n = 3$ was only recently proved by Wooley using the efficient congruencing method, Wooley [2016]. The case $n \geq 4$ was proved for the first time in Bourgain, Demeter, and Guth [2016].

6 The proof of the decoupling theorem for the parabola

In this section we prove Theorem 4.2 for the parabola $\mathbb{P}^1$. The argument in higher dimensions is very similar, though technically slightly more complicated. We will denote by E the extension operator associated with $\mathbb{P}^1$.

It will be more convenient to think of B_R as being an arbitrary square (rather than ball) with side length R in $\mathbb{R}^2$. We will often partition big squares into smaller ones.

For $m \geq 0$ let $\mathbb{I}_m$ be the collection of the 2^m dyadic subintervals of $[0,1]$ of length 2^{-m}. Thus $\mathbb{I}_0$ consists of only $[0,1]$. Note that each $I \in \mathbb{I}_{m+1}$ is inside some $I' \in \mathbb{I}_m$ and each $I' \in \mathbb{I}_m$ has two "children" in $\mathbb{I}_{m+1}$, adjacent to each other.

Define $\mathrm{Dec}(n,p)$ to be the smallest constant such that the inequality

$$\|Ef\|_{L^p(w_{B_{4^n}})} \leq \mathrm{Dec}(n,p)\Big(\sum_{I\in\mathbb{I}_n}\|E_I f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/2}$$

holds true for each $f : [0,1] \to \mathbb{C}$. Minkowski's inequality shows that $\mathrm{Dec}(n,p)$ controls the decoupling on larger squares, too. By that we mean that the inequality

$$\|Ef\|_{L^p(w_{B_R})} \lesssim \mathrm{Dec}(n,p)\Big(\sum_{I\in\mathbb{I}_n}\|E_I f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

holds for each $R \geq 4^n$, with the implicit constant in $\lesssim$ independent of R.

It will suffice to prove that

$$\mathrm{Dec}(n,p) \lesssim_{\epsilon,p} 2^{n\epsilon},\ 2 \leq p \leq 6. \tag{9}$$

The inequality $\mathrm{Dec}(n,2) \lesssim 1$ follows from simple orthogonality reasons. We start by explaining why the proof of (9) is not quite as immediate as one would wish, when $p > 2$. Let A_p be the smallest constant that governs the decoupling into two intervals. In precise terms, assume

$$\|Ef\|_{L^p(B)} \leq A_p\big(\|E_{J_1}f\|^2_{L^p(w_B)} + \|E_{J_2}f\|^2_{L^p(w_B)}\big)^{1/2}$$

holds for each disjoint intervals $J_1, J_2 \subset \mathbb{R}$ of arbitrary length $L \leq 1$ that are adjacent to each other, each $f : J_1 \cup J_2 \to \mathbb{C}$ and for each square $B \subset \mathbb{R}^2$ with side length at least L^{-2}. One can check that $A_p > 1$ for $p > 2$.

Since each $I' \in \mathbb{I}_m$ has two children in $\mathbb{I}_{m+1}$, it is easy to see that

$$\mathrm{Dec}(m+1,p) \leq A_p \mathrm{Dec}(m,p).$$

Iterating this, we get the very unfavorable estimate $\mathrm{Dec}(n,p)\le A_p^{n-1}\mathrm{Dec}(1,p)$. Indeed, note that $A_p>1$ forces $A_p^{n-1}\gg 2^{n\epsilon_0}$, for some $\epsilon_0>0$. This shows that we can not afford to lose A_p each time we go one level up (call this a *step*). Instead, we are going to make huge *leaps*, and rather than going from level m to level $m+1$ at a time, we will instead go from m to $2m$. The choice for the size of this leap is motivated by the fact that intervals in $\mathbb{I}_{2m}$ have length equal to those in $\mathbb{I}_m$ squared. We have a very efficient mechanism to decouple from scale δ to δ^2, namely the bilinear Kakeya inequality.

Each leap will combine two inequalities. One is a consequence of the bilinear Kakeya, the other one is a form of L^2 orthogonality. The loss for each application of the bilinear Kakeya is rather tiny, at most n^C (compare this with the loss A_p^m accumulated if instead we went from level m to $2m$ in m steps). From $\mathbb{I}_0$ to $\mathbb{I}_n$, we need $\log n$ such leaps, so the overall loss from the repeated use of bilinear Kakeya amounts to $n^{O(\log n)}$. This is easily seen to be $O(2^{n\epsilon})$ for each $\epsilon>0$, as desired. There is however a price we pay in our approach: in each leap we only decouple a $1-\kappa_p$ fraction of the operator. See Proposition 6.5 for a precise statement.

Here is a sketch of how we put things together, and we will limit attention to the hardest case $p=6$. First, we will do a trivial decoupling (Cauchy-Schwarz) to get from $\mathbb{I}_0$ to $\mathbb{I}_{\frac{n}{2^s}}$ by loosing only $2^{O(\frac{n}{2^s})}$. We will be able to choose s as large as we wish, so this loss will end up being controlled by $2^{n\epsilon}$. The transition from $\mathbb{I}_{\frac{n}{2^s}}$ to $\mathbb{I}_n$ will then be done in s leaps, by each time applying Proposition 6.5. Collecting all contributions, an a priori bound of the form

$$\mathrm{Dec}(n,6)\lesssim 2^{nA},\ \text{ for some } A>0$$

will get upgraded to a stronger (assuming s is large enough) bound

$$\mathrm{Dec}(n,6)\lesssim 2^{n(A(1-\frac{s+1}{2^{s+1}})+\frac{1}{2^s-1})}.$$

Applying this bootstrapping argument will force A to get smaller and smaller, arbitrarily close to 0.

The leaps are performed using bilinear decouplings, in order to take advantage of the bilinear Kakeya phenomenon. The fact that there is no serious loss in bilinearization is proved in Proposition 6.2.

6.1 Parabolic rescaling and linear vs. bilinear decoupling. One of our main tools will be the following parabolic rescaling, that takes advantage of the affine invariance of the parabola.

Proposition 6.1. *Let $I=[t,t+2^{-l}]\subset\mathbb{R}$ be an interval of length 2^{-l} and for $n>l$ let the collection $\mathbb{I}_n(I)$ consist of all subintervals of I of the form $[t+j2^{-n},t+(j+1)2^{-n}]$,* *with $j \in \mathbb{N}$. Then for each f supported on I*

$$\|Ef\|_{L^p(w_{B_{4^n}})} \lesssim \mathrm{Dec}(n-l,p)\Big(\sum_{J\in\mathbb{I}_n(I)} \|E_J f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/2}$$

Note that the upper bound $\mathrm{Dec}(n-l,p)$ is morally stronger than the trivial upper bound $\mathrm{Dec}(n,p)$.

Proof. The proof is a simple applications of affine change of variables. Indeed $L_I(\xi) = 2^l(\xi - t)$ maps $\mathbb{I}_n(I)$ to $\mathbb{I}_{n-l}$ and the square B_{4^n} to a parallelepiped that can be covered efficiently with squares $B_{4^{n-l}}$.

□

For the rest of the argument let $I_1 = [0, \frac{1}{4}]$, $I_2 = [\frac{1}{2}, 1]$. Define $\mathrm{BilDec}(n,p)$ to be the smallest constant such that the inequality

$$\||Ef_1 Ef_2|^{1/2}\|_{L^p(w_{B_{4^n}})} \le \mathrm{BilDec}(n,p)\Big(\sum_{I\in\mathbb{I}_n(I_1)} \|E_I f_1\|^2_{L^p(w_{B_{4^n}})} \sum_{I\in\mathbb{I}_n(I_2)} \|E_I f_2\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/}$$

holds true for all f_1, f_2 supported on I_1 and I_2, respectively.

It is immediate that

$$\mathrm{BilDec}(n,p) \le \mathrm{Dec}(n,p).$$

The next result is some sort of a converse.

Proposition 6.2. *For each $\epsilon > 0$*

$$\mathrm{Dec}(n,p) \lesssim_\epsilon 2^{n\epsilon}(1 + \max_{m\le n} \mathrm{BilDec}(m,p)).$$

Proof. It will suffice to prove that for each $k < n$

$$\mathrm{Dec}(n,p) \lesssim C^{\frac{n}{k}}(1 + C_k n \max_{m\le n} \mathrm{BilDec}(m,p)).$$

This will instead follow by iterating the inequality

$$\mathrm{Dec}(n,p) \le C\,\mathrm{Dec}(n-k,p) + C_k \max_{m\le n} \mathrm{BilDec}(m,p), \tag{10}$$

with C independent of n, k. Let us next prove this inequality.

Fix k and let f be supported on $[0,1]$. Since

$$Ef(x) = \sum_{I\in\mathbb{I}_k} E_I f(x),$$

it is not difficult to see that

$$|Ef(x)| \le 4 \max_{I \in \mathbb{I}_k} |E_I f(x)| + 2^{O(k)} \sum_{\substack{J_1, J_2 \in \mathbb{I}_k \\ 2J_1 \cap 2J_2 = \emptyset}} |E_{J_1} f(x) E_{J_2} f(x)|^{1/2}, \tag{11}$$

where the sum on the right is taken over all pairs of intervals $J_1, J_2 \in \mathbb{I}_k$ which are not neighbors. Fix such a pair $J_1 = [a, a + 2^{-k}]$, $J_2 = [b, b + 2^{-k}]$, and let m be a positive integer satisfying $2^{-m} \le b - a < 2^{-m+1}$. Since J_1, J_2 are not adjacent to each other, we must have $m \le k - 1$. It follows that the affine function $T(\xi) = \frac{\xi - a}{2^{-m+1}}$ maps J_1 to a dyadic subinterval of $[0, \frac{1}{4}]$ and J_2 to a dyadic subinterval of $[\frac{1}{2}, 1]$. Thus, parabolic rescaling shows that

$$\| |E_{J_1} f E_{J_2} f|^{1/2} \|_{L^p(w_{B_{4^n}})}$$

$$\lesssim \mathrm{BilDec}(n - m + 1, p) \Big(\sum_{I \in \mathbb{I}_n(J_1)} \|E_I f\|^2_{L^p(w_{B_{4^n}})} \sum_{I \in \mathbb{I}_n(J_2)} \|E_I f\|^2_{L^p(w_{B_{4^n}})} \Big)^{1/4}$$

$$\le \mathrm{BilDec}(n - m + 1, p) \Big(\sum_{I \in \mathbb{I}_n} \|E_I f\|^2_{L^p(w_{B_{4^n}})} \Big)^{1/2}. \tag{12}$$

Finally, invoking again Proposition 6.1 we get

$$\| \max_{I \in \mathbb{I}_k} |E_I f| \|_{L^p(w_{B_{4^n}})} \le \Big(\sum_{I \in \mathbb{I}_k} \|E_I f\|^2_{L^p(w_{B_{4^n}})} \Big)^{\frac{1}{2}}$$

$$\lesssim \mathrm{Dec}(n - k, p) \Big(\sum_{I \in \mathbb{I}_k} \sum_{I' \in \mathbb{I}_n(I)} \|E_{I'} f\|^2_{L^p(w_{B_{4^n}})} \Big)^{\frac{1}{2}}$$

$$= \mathrm{Dec}(n - k, p) \Big(\sum_{I' \in \mathbb{I}_n} \|E_{I'} f\|^2_{L^p(w_{B_{4^n}})} \Big)^{\frac{1}{2}}. \tag{13}$$

Now (10) follows by combining (11), (12) and (13).

□

6.2 A consequence of bilinear Kakeya. We start by recalling the following bilinear Kakeya inequality. While this inequality is rather trivial in two dimensions, its higher dimensional analogs that are needed in order to prove decouplings for the paraboloid $\mathbb{P}^{n-1}$, $n \ge 3$, are more complicated. The multilinear Kakeya inequality was first proved in Bennett, Carbery, and Tao [2006], and an easier proof appeared in Guth [2015].

Theorem 6.3. *Consider two families $\mathbb{T}_1, \mathbb{T}_2$ consisting of rectangles T in $\mathbb{R}^2$ having the following properties*

(i) each T has the short side of length $R^{1/2}$ and the long side of length equal to R pointing in the direction of the unit vector v_T

(ii) $v_{T_1} \wedge v_{T_2} \geq \frac{1}{100}$ for each $T_i \in \mathbb{T}_i$.

We have the following inequality

$$\int_{\mathbb{R}^2} \prod_{i=1}^{2} F_i \lesssim \frac{1}{R^2} \prod_{i=1}^{2} \int_{\mathbb{R}^2} F_i \tag{14}$$

for all functions F_i of the form

$$F_i = \sum_{T \in \mathbb{T}_i} c_T 1_T,\ c_T \in [0, \infty).$$

The implicit constant will not depend on $R, c_T, \mathbb{T}_i$.

Proof. The verification is immediate using the fact that $|T_1 \cap T_2| \lesssim R$ whenever $T_i \in \mathbb{T}_i$. □

If $I \subset \mathbb{R}$ is an interval of length 2^{-l} and $\delta = 2^{-k}$ with $k \geq l$, we will denote by $\mathrm{Part}_\delta(I)$ the partition of I into intervals of length δ. Recall also that $I_1 = [0, \frac{1}{4}]$, $I_2 = [\frac{1}{2}, 1]$ and that $L^q_\sharp$ denotes the average integral in L^q.

The following result is part of a two-stage process. Note that, strictly speaking, this inequality is not a decoupling, since the size of the frequency intervals $I_{i,1}$ remains unchanged. However, the side length of the spatial squares increases from δ^{-1} to δ^{-2}. This will facilitate a subsequent decoupling, as we shall later see in Proposition 6.5.

Proposition 6.4. *Let $q \geq 2$ and $\delta < 1$. Let B be an arbitrary square in $\mathbb{R}^2$ with side length δ^{-2}, and let $\mathcal{B}$ be the unique partition of B into squares Δ of side length δ^{-1}. Then for each $g : [0, 1] \to \mathbb{C}$ we have*

$$\frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \left[\prod_{i=1}^{2} \Big(\sum_{I_{i,1} \in \mathrm{Part}_\delta(I_i)} \|E_{I_{i,1}} g\|^2_{L^q_\sharp(w_\Delta)} \Big)^{\frac{1}{2}} \right]^q \tag{15}$$

$$\lesssim (\log(\frac{1}{\delta}))^{O(1)} \left[\prod_{i=1}^{2} \Big(\sum_{I_{i,1} \in \mathrm{Part}_\delta(I_i)} \|E_{I_{i,1}} g\|^2_{L^q_\sharp(w_B)} \Big)^{\frac{1}{2}} \right]^q. \tag{16}$$

Moreover, the implicit constant is independent of g, δ, B.

Proof. We will reduce the proof to an application of Theorem 6.3. Indeed, for each interval J of length δ, the Fourier transform of $E_J g$ is supported inside a $2\delta \times 2\delta^2$-rectangle. This in turn suggests that $|E_J g|$ is essentially constant on $\delta^{-1} \times \delta^{-2}$-rectangles dual to this rectangle. Note that due to the separation of I_1 and I_2, the rectangles corresponding to intervals $I_{1,1} \subset I_1$, $I_{2,1} \subset I_2$ satisfy the requirements in Theorem 6.3 with $R = \delta^{-2}$.

Since we can afford logarithmic losses in δ, it suffices to prove the inequality with the summation on both sides restricted to families of $I_{i,1}$ for which $\|E_{I_{i,1}} g\|_{L^q_\sharp(w_B)}$ have comparable size (within a multiplicative factor of 2), for each i. Indeed, the intervals $I'_{i,1}$ satisfying (for some large enough $C = O(1)$)

$$\|E_{I'_{i,1}} g\|_{L^q_\sharp(w_B)} \le \delta^C \max_{I_{i,1} \in \mathrm{Part}_\delta(I_i)} \|E_{I_{i,1}} g\|_{L^q_\sharp(w_B)}$$

can be easily dealt with by using the triangle inequality, since we automatically have

$$\max_{\Delta \in \mathcal{B}} \|E_{I'_{i,1}} g\|_{L^q_\sharp(w_\Delta)} \le \delta^C \max_{I_{i,1} \in \mathrm{Part}_\delta(I_i)} \|E_{I_{i,1}} g\|_{L^q_\sharp(w_B)}.$$

This leaves only $\log_2(\delta^{-O(1)})$ sizes to consider.

Let us now assume that we have N_i intervals $I_{i,1}$, with $\|E_{I_{i,1}} g\|_{L^q_\sharp(w_B)}$ of comparable size. Since $q \ge 2$, by Hölder's inequality (15) is at most

$$(\prod_{i=1}^{2} N_i^{\frac{1}{2}-\frac{1}{q}})^q \frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} (\prod_{i=1}^{2} (\sum_{I_{i,1}} \|E_{I_{i,1}} g\|^q_{L^q_\sharp(w_\Delta)})). \tag{17}$$

For each $I = I_{i,1}$ centered at c_I, consider the family $\mathfrak{F}_I$ of pairwise disjoint, mutually parallel rectangles T_I. They have the short side of length δ^{-1} and the longer side of length δ^{-2}, pointing in the direction of the normal $N(c_I)$ to the paraboloid $\mathbb{P}^1$ at c_I.

The function

$$F_I(x) := \|E_I g\|^q_{L^q_\sharp(w_{B(x,\delta^{-1})})}$$

can be thought of as being essentially constant on rectangles in $\mathfrak{F}_I$. This can be made precise, but we will sacrifice a bit of the rigor for the sake of keeping the argument simple enough. Thus we may write

$$\frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \prod_i (\sum_{I_{i,1}} \|E_{I_{i,1}} g\|^q_{L^q_\sharp(w_\Delta)}) \approx \frac{1}{|B|} \int \prod_{i=1}^{2} F_i,$$

with $F_i(x) = \sum_{I_{i,1}} F_{I_{i,1}}(x)$.

Applying Theorem 6.3 we can dominate the term on the right by

$$\frac{1}{|B|^2}\prod_i \int F_i.$$

Note also that

$$\frac{1}{|B|}\int F_i \approx \sum_{I_{i,1}} \|E_{I_{i,1}} g\|^q_{L^q_\sharp(w_B)}.$$

It follows that (17) is dominated by

$$(18)\qquad (\prod_{i=1}^{2} N_i^{\frac{1}{2}-\frac{1}{q}})^q \prod_{i=1}^{2}(\sum_{I_{i,1}} \|E_{I_{i,1}} g\|^q_{L^q_\sharp(w_B)}).$$

Recalling the restriction we have made on $I_{i,1}$, (18) is comparable to

$$\left[\prod_{i=1}^{2}(\sum_{I_{i,1}} \|E_{I_{i,1}} g\|^2_{L^q_\sharp(w_B)})^{1/2}\right]^q,$$

as desired.

□

6.3 The leap: decoupling from scale δ to δ^2. To simplify notation, we denote by B^r an arbitrary square in $\mathbb{R}^2$ with side length 2^r. Given $q, r \in \mathbb{N}$, $t \geq 2$ and g supported in $I_1 \cup I_2$ write

$$D_t(q, B^r, g) = [(\sum_{I\in\mathbb{I}_q(I_1)} \|E_I g\|^2_{L^t_\sharp(w_{B^r})})(\sum_{I\in\mathbb{I}_q(I_2)} \|E_I g\|^2_{L^t_\sharp(w_{B^r})})]^{\frac{1}{4}}.$$

For $s \leq r$ we will denote by $\mathcal{B}_s(B^r)$ the partition of B^r into squares B^s. Define

$$A_p(q, B^r, s, g) = \Big(\frac{1}{|\mathcal{B}_s(B^r)|}\sum_{B^s\in\mathcal{B}_s(B^r)} D_2(q, B^s, g)^p\Big)^{\frac{1}{p}}.$$

Note that when $r = s$,

$$A_p(q, B^r, r, g) = D_2(q, B^r, g).$$

For $p \geq 4$, let $0 \leq \kappa_p \leq 1$ satisfy

$$\frac{2}{p} = \frac{1-\kappa_p}{2} + \frac{\kappa_p}{p},$$

that is

$$\kappa_p = \frac{p-4}{p-2}.$$

The following result shows how to decouple from scale $\delta = 2^{-q}$ to scale δ^2. Note also that only a $1-\kappa_p$ fraction gets decoupled.

Proposition 6.5. *We have for each $p \geq 4$, $r \geq q$ and each g supported in $I_1 \cup I_2$*

$$A_p(q, B^{2r}, q, g) \lesssim_\epsilon \delta^{-\epsilon} A_p(2q, B^{2r}, 2q, g)^{1-\kappa_p} D_p(q, B^{2r})^{\kappa_p}$$

Proof. By using elementary inequalities, it suffices to prove the proposition for $r = q$. By Hölder's inequality,

$$\|E_I g\|_{L^2_\sharp(w_{B^q})} \lesssim \|E_I g\|_{L^{\frac{p}{2}}_\sharp(w_{B^q})}.$$

Using this and Proposition 6.4 with $\delta = 2^{-q}$ we can write

(19)
$$A_p(q, B^{2q}, q, g) \lesssim_\epsilon \delta^{-\epsilon}[(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^{\frac{p}{2}}_\sharp(w_{B^{2q}})})(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^{\frac{p}{2}}_\sharp(w_{B^{2q}})})]^{\frac{1}{4}}.$$

Using Hölder's inequality again, we can dominate this by

$$\delta^{-\epsilon}[(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})})(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})})]^{\frac{1-\kappa_p}{4}}$$
$$\times[(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^p_\sharp(w_{B^{2q}})})(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^p_\sharp(w_{B^{2q}})})]^{\frac{\kappa_p}{4}}.$$

To further process the first term we invoke L^2 orthogonality for each $I \in \mathbb{I}_q$

$$\|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})} \lesssim \sum_{J \in \mathbb{I}_{2q}(I)} \|E_J g\|^2_{L^2_\sharp(w_{B^{2q}})}.$$

□

6.4 Putting everything together. We will now prove inequality (9). It will suffice to work with $n = 2^u$, $u \in \mathbb{N}$.

Iterating Proposition 6.5 s times leads to the following multi-scale inequality, for each $p \geq 4$.

Proposition 6.6. *For each g supported on $I_1 \cup I_2$ and $s \leq u$ we have*

$$A_p(\frac{n}{2^s}, B^{2n}, \frac{n}{2^s}, g) \lesssim_{s,\epsilon} 2^{\epsilon s n} A_p(n, B^{2n}, n, g)^{(1-\kappa_p)^s} \prod_{l=1}^{s} D_p(\frac{n}{2^l}, B^{2n}, g)^{\kappa_p(1-\kappa_p)^{s-l}}.$$

Via one application of Cauchy–Schwarz we see that

$$\||Ef_1Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} = \Big(\frac{1}{|\mathcal{B}_{\frac{n}{2^s}}(B^{2n})|}\sum_{B\in\mathcal{B}_{\frac{n}{2^s}}(B^{2n})}\||Ef_1Ef_2|^{1/2}\|^p_{L^p_\sharp(B)}\Big)^{1/p}$$

$$\leq 2^{\frac{n}{2^{s+1}}}\Big[\frac{1}{|\mathcal{B}_{\frac{n}{2^s}}(B^{2n})|}\sum_{B\in\mathcal{B}_{\frac{n}{2^s}}(B^{2n})}\Big(\sum_{I\in\mathbb{I}_{\frac{n}{2^s}}(I_1)}\|E_I f_1\|^2_{L^p_\sharp(B)}\sum_{I\in\mathbb{I}_{\frac{n}{2^s}}(I_2)}\|E_I f_2\|^2_{L^p_\sharp(B)}\Big)^{p/4}\Big]^{1/p}$$

holds true for all f_1, f_2 supported on I_1 and I_2, respectively.

At this point we need to invoke the following reverse Hölder's inequality

$$\|E_I f_i\|_{L^p_\sharp(B)} \lesssim \|E_I f_i\|_{L^2_\sharp(B)}, \tag{20}$$

for each square B with side length $2^{\frac{n}{2^s}}$ and each I of length $2^{-\frac{n}{2^s}}$. This is a consequence of the fact that $|E_I f_i|$ is essentially constant on B.

We conclude as follows.

Proposition 6.7. *The inequality*

$$\||Ef_1Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} \leq 2^{\frac{n}{2^{s+1}}} A_p\Big(\frac{n}{2^s}, B^{2n}, \frac{n}{2^s}, g\Big)$$

holds true when f_1, f_2 are the restrictions of g to I_1, I_2, respectively.

To combine the last two propositions, we need one more inequality, a consequence of Proposition 6.1.

Proposition 6.8. *For each $l \leq u$*

$$D_p\Big(\frac{n}{2^l}, B^{2n}, g\Big) \lesssim \operatorname{Dec}\Big(n-\frac{n}{2^l}, p\Big) D_p(n, B^{2n}, g).$$

The following result is now rather immediate.

Theorem 6.9. *Assume f_1, f_2 are the restrictions of g to I_1 and I_2, respectively. Then for each $s \leq u$*

$$\||Ef_1Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} \lesssim_{s,\epsilon} 2^{\epsilon s n} 2^{\frac{n}{2^{s+1}}} D_p(n, B^{2n}, g)\prod_{l=1}^{s}\operatorname{Dec}\Big(n-\frac{n}{2^l}, p\Big)^{\kappa_p(1-\kappa_p)^{s-l}}$$

Proof. Using Hölder and Minkowski's inequality in $l^{\frac{p}{2}}$ we find that

$$A_p(n, B^{2n}, n, g) \lesssim D_p(n, B^{2n}, g).$$

Combine this with the previous three propositions.

□

Since this inequality holds for arbitrary g we can take the supremum to get the following inequality.

Corollary 6.10. *For each $s \le u$*

$$\mathrm{BilDec}(n, p) \lesssim_{s,\epsilon} 2^{\epsilon s n} 2^{\frac{n}{2^{s+1}}} \prod_{l=1}^{s} \mathrm{Dec}(n - \frac{n}{2^l}, p)^{\kappa_p (1-\kappa_p)^{s-l}}.$$

We are now ready to finalize the proof of inequality (9). Take $p = 6$, and note that $\kappa_6 = \frac{1}{2}$. The case $p < 6$ would follow very similarly since $\kappa_p < \frac{1}{2}$.

We will use a bootstrapping argument. Assume $\mathrm{Dec}(n, 6) \lesssim 2^{nA}$ holds for some A and all n. For example, it is easy to see that $A = \frac{1}{2}$ works. We will show that a smaller value of A always works, too. Corollary 6.10 implies that for each s and each n

$$\mathrm{BilDec}(n, 6) \lesssim_s 2^{n(Ac_s + \frac{1}{2^s})},$$

where

$$c_s = \sum_{l=1}^{s} (1 - \frac{1}{2^l}) \frac{1}{2^{s-l+1}} = 1 - \frac{s+1}{2^{s+1}}.$$

Combining this with Proposition 6.2 we may write

$$\mathrm{Dec}(n, 6) \lesssim_s 2^{n(Ac_s + \frac{1}{2^{s-1}})}. \tag{21}$$

Define

$$\mathcal{A} := \{A > 0 : \ \mathrm{Dec}(n, 6) \lesssim 2^{nA}\}$$

and let $A_0 = \inf \mathcal{A}$. Note that $\mathcal{A}$ is either (A_0, ∞) or $[A_0, \infty)$. We claim that A_0 must be zero, which will finish the proof of our theorem. Indeed, if $A_0 > 0$ then

$$Ac_s + \frac{1}{2^{s-1}} < A_0$$

for some $A \in \mathcal{A}$ sufficiently close to A_0 and s sufficiently large. This combined with (21) contradicts the definition of A_0.

References

Jonathan Bennett, Anthony Carbery, and Terence Tao (2006). "On the multilinear restriction and Kakeya conjectures". *Acta Math.* 196.2, pp. 261–302. MR: 2275834 (cit. on pp. 1536, 1550).

J. Bourgain (2013). "Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces". *Israel J. Math.* 193.1, pp. 441–458. MR: 3038558 (cit. on p. 1541).

– (2017). "Decoupling, exponential sums and the Riemann zeta function". *J. Amer. Math. Soc.* 30.1, pp. 205–224. MR: 3556291 (cit. on p. 1543).

Jean Bourgain and Ciprian Demeter (2015). "The proof of the l^2 decoupling conjecture". *Ann. of Math. (2)* 182.1, pp. 351–389. MR: 3374964 (cit. on pp. 1541–1544).

– (2016a). "Decouplings for surfaces in $\mathbb{R}^4$". *J. Funct. Anal.* 270.4, pp. 1299–1318. MR: 3447712 (cit. on p. 1543).

– (2016b). "Mean value estimates for Weyl sums in two dimensions". *J. Lond. Math. Soc. (2)* 94.3, pp. 814–838. MR: 3614930 (cit. on p. 1543).

– (2017). "Decouplings for curves and hypersurfaces with nonzero Gaussian curvature". *J. Anal. Math.* 133, pp. 279–311. MR: 3736493 (cit. on p. 1543).

Jean Bourgain, Ciprian Demeter, and Shaoming Guo (2017). "Sharp bounds for the cubic Parsell-Vinogradov system in two dimensions". *Adv. Math.* 320, pp. 827–875. MR: 3709122 (cit. on p. 1543).

Jean Bourgain, Ciprian Demeter, and Larry Guth (2016). "Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three". *Ann. of Math. (2)* 184.2, pp. 633–682. MR: 3548534 (cit. on pp. 1542, 1543, 1546).

Jean Bourgain and Larry Guth (2011). "Bounds on oscillatory integral operators based on multilinear estimates". *Geom. Funct. Anal.* 21.6, pp. 1239–1295. MR: 2860188 (cit. on p. 1536).

Jean Bourgain and Nigel Watt (n.d.). "Mean square of zeta function, Gauss circle problem and divisor problem revisited". arXiv: 1709.04340 (cit. on p. 1543).

– (2017). "Decoupling for perturbed cones and mean square of $\zeta(\frac{1}{2}+it)$". *International Mathematics Research Notices*, rnx009 (cit. on p. 1543).

C. Demeter, S. Guo, and F. Shi (n.d.). "Sharp decouplings for three dimensional manifolds in $\mathbb{R}^5$". To appear in Revista Matematica Iberoamericana (cit. on p. 1543).

Ciprian Demeter (n.d.). "Incidence theory and restriction estimates". arXiv: 1401.1873 (cit. on p. 1541).

Gustavo Garrigós and Andreas Seeger (2009). "On plate decompositions of cone multipliers". *Proc. Edinb. Math. Soc. (2)* 52.3, pp. 631–651. MR: 2546636 (cit. on p. 1541).

– (2010). "A mixed norm variant of Wolff's inequality for paraboloids". In: *Harmonic analysis and partial differential equations*. Vol. 505. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 179–197. MR: 2664568 (cit. on p. 1541).

Emil Grosswald (1985). *Representations of integers as sums of squares*. Springer-Verlag, New York, pp. xi+251. MR: 803155 (cit. on p. 1539).

Shaoming Guo and Changkeun Oh (n.d.). "Remarks on Wolff's inequality for hypersurfaces". arXiv: 1602.05861 (cit. on p. 1543).

Larry Guth (2015). "A short proof of the multilinear Kakeya inequality". *Math. Proc. Cambridge Philos. Soc.* 158.1, pp. 147–153. MR: 3300318 (cit. on p. 1550).

Rowan Killip and Monica Vişan (2016). "Scale invariant Strichartz estimates on tori and applications". *Math. Res. Lett.* 23.2, pp. 445–472. MR: 3512894 (cit. on p. 1544).

Izabella Laba and Malabika Pramanik (2005). *Wolff's inequality for hypersurfaces* (cit. on p. 1541).

Izabella Łaba and Thomas Wolff (2002). "A local smoothing estimate in higher dimensions". *J. Anal. Math.* 88. Dedicated to the memory of Tom Wolff, pp. 149–171. MR: 1956533 (cit. on p. 1541).

Malabika Pramanik and Andreas Seeger (2007). "L^p regularity of averages over curves and bounds for associated maximal operators". *Amer. J. Math.* 129.1, pp. 61–103. MR: 2288738 (cit. on p. 1541).

Elias M. Stein (1993). *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, pp. xiv+695. MR: 1232192 (cit. on p. 1537).

Robert S. Strichartz (1977). "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations". *Duke Math. J.* 44.3, pp. 705–714. MR: 0512086 (cit. on p. 1537).

Peter A. Tomas (1975). "A restriction theorem for the Fourier transform". *Bull. Amer. Math. Soc.* 81, pp. 477–478. MR: 0358216 (cit. on p. 1537).

T. Wolff (2000). "Local smoothing type estimates on L^p for large p". *Geom. Funct. Anal.* 10.5, pp. 1237–1288. MR: 1800068 (cit. on p. 1541).

Trevor D. Wooley (2016). "The cubic case of the main conjecture in Vinogradov's mean value theorem". *Adv. Math.* 294, pp. 532–561. MR: 3479572 (cit. on p. 1546).

Received 2017-11-02.

Ciprian Demeter
Department of Mathematics
Indiana University
Bloomington IN
demeterc@indiana.edu