A survey on ℓ^2 decoupling

Po-Lam Yung [1]

The Chinese University of Hong Kong

January 31, 2018

[1] Research partially supported by HKRGC grant 14313716, and by CUHK direct grants 4053220, 4441563

Introduction

- Today we will discuss *decoupling inequalities*.
- They capture certain "interference patterns" that occur when we add up functions whose Fourier transforms are supported in different regions of a curved submanifold of $\mathbb{R}^n$.
  (Think of a paraboloid, or a curve, in $\mathbb{R}^n$.)
- Decoupling first appeared in the work of Wolff, and has been further developed by Łaba, Pramanik, Seeger and Garrigós.
- Recent breakthrough came from the work of Bourgain and Demeter, who established ℓ^2 decoupling for the paraboloid for the optimal range of exponents (up to ϵ losses).
- This has applications to PDE, additive combinatorics, and number theory.
- Shortly afterwards, Bourgain, Demeter and Guth proved Vinogradov's main conjecture in number theory, via ℓ^2 decoupling for a monomial curve in $\mathbb{R}^n$.
- We will survey some of these developments below.

Outline of the talk

- Part I: The case of the paraboloid
- Part II: The case of the moment curve
- Part III: The case of the truncated cone
- Part IV: Other applications of decoupling
- Part V: A few words about proofs

Part I: The case of the paraboloid

- Let $n \geq 2$, $Q = [-1,1]^{n-1}$ and $\Phi \colon Q \to \mathbb{R}^n$ be a parametrization of the paraboloid given by

$$\Phi(\xi) = (\xi, |\xi|^2), \quad \xi \in Q.$$

- If f is a (say C^∞) function defined on Q, define the Fourier extension operator

$$Ef(x) = \int_Q f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- If $f = \widehat{g}$ for some function g on $\mathbb{R}^{n-1}$ and the support of $\widehat{g}$ is in Q, then by writing $x = (x', x_n)$ where $x' \in \mathbb{R}^{n-1}$ and $x_n \in \mathbb{R}$, and thinking of x' as the space variable, x_n as the time variable, one can interpret $Ef(x)$ as a solution to the Schrödinger equation on $\mathbb{R}^{n-1} \times \mathbb{R}$ with initial data g.
- On the other hand, one can interpret $f(\xi)d\xi$ as a measure on the paraboloid, and $Ef(x)$ is simply the inverse Fourier transform of this measure. Hence E is also called the Fourier extension operator associated to the paraboloid.

$$Ef(x) = \int_Q f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- Clearly E maps L^∞ to L^∞ regardless of what Φ is.
- When $\Phi(\xi) = (\xi, |\xi|^2)$, E.M. Stein (1967) observed that E maps L^∞ to L^p for some $p < \infty$.
- This is only possible because the image of Φ (i.e. the paraboloid) has non-vanishing Gaussian curvature.
- It is this fundamental observation that started a fruitful investigation about the (Fourier) restriction problem for the past 50 years.

The restriction problem

- The question is what is the smallest value of p, for which E maps L^∞ to L^p; when say $n = 3$, this is known to hold when

$p \geq 4$	Tomas, Stein 1976
$p > 4 - 2/15$	Bourgain 1991
$p > 4 - 2/11$	Wolff 1995; Moyua, Vargas, Vega 1996
$p > 4 - 2/9$	Tao, Vargas, Vega 1998
$p > 4 - 2/7$	Tao, Vargas 2000
$p > 4 - 2/3 = 10/3$	Tao 2003
$p > 3.3$	Bourgain, Guth 2011
$p > 3.25$	Guth 2016

  and is conjectured to hold when $p > 3$.
- We remark in passing that the restriction problem is named as such, because the adjoint of E is given by the restriction of the Fourier transform of a function to the paraboloid.

The decoupling problem

$$Ef(x) = \int_Q f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- The restriction problem seeks to bound $\|Ef\|_{L^p(\mathbb{R}^n)}$ by $\|f\|_{L^\infty}$.
- On the other hand, let's partition Q into N^{n-1} disjoint cubes $Q_1, \dots, Q_{N^{n-1}}$ of equal sizes. For $j = 1, \dots, N^{n-1}$, let

$$E_j f(x) = \int_{Q_j} f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- Clearly

$$Ef = \sum_{j=1}^{N^{n-1}} E_j f.$$

Hence by Cauchy-Schwarz, for any $1 \leq p \leq \infty$, we have

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq N^{\frac{n-1}{2}} \left(\sum_{j=1}^{N^{n-1}} \|E_j f\|_{L^p(\mathbb{R}^n)}^2 \right)^{1/2}.$$

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq N^{\frac{n-1}{2}} \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^2}.$$

- This is the best one can say if $p = \infty$.
- On the other hand, if $p < \infty$, then one may do better, and the constant $N^{\frac{n-1}{2}}$ can be reduced.
- This is because of the curvature present on the paraboloid; indeed the $E_j f$'s have frequency supports on caps on the paraboloid whose normal vectors are transverse to each other. (Think of $E_j f$ as waves travelling in the direction normal to its frequency support; since these directions are separated for different j's, there is a lot of destructive interference.)
- The decoupling problem (for a given p) is the problem of determining the smallest constant for which the above inequality holds for all $N \in \mathbb{N}$.

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq (\text{best constant}) \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^2}.$$

- Bourgain and Demeter (2014) showed that when $p \geq \frac{2(n+1)}{n-1}$, the best constant is $\lesssim_\varepsilon N^{\alpha(n,p)+\varepsilon}$ for all $\varepsilon > 0$, where
$$\alpha(n,p) := \frac{n-1}{2} - \frac{n+1}{p};$$
here the range of p is optimal. (Also the power of N is optimal up to the N^ε loss for the indicated range of p.)
- Indeed, Bourgain and Demeter proved a localized version, which (roughly) says that whenever $p \geq \frac{2(n+1)}{n-1}$ and $\varepsilon > 0$,
$$\|Ef\|_{L^p(B_{N^2})} \lesssim_\varepsilon N^{\alpha(n,p)+\varepsilon} \left\| \|E_j f\|_{L^p(B_{N^2})} \right\|_{\ell^2},$$
where B_{N^2} is any ball of radius N^2 in $\mathbb{R}^n$.
(Fine print: actually the $L^p(B_{N^2})$ norm on the right hand side should be replaced by a suitable weighted L^p norm.)
- This has applications to PDE, additive combinatorics and number theory; we give only one application to PDE below.

Strichartz inequalities on tori

- Bourgain and Demeter used their decoupling inequality to derive certain Strichartz estimates for the Schrödinger equation on tori.
- They proved that if $\Delta = \sum_{j=1}^{n-1} \partial_{x_j}^2$ is the Laplacian on $\mathbb{T}^{n-1} = (\mathbb{R}/\mathbb{Z})^{n-1}$, then for any f on $\mathbb{T}^{n-1}$ whose Fourier transform is supported on $[-N, N]^{n-1}$, we have

$$\|e^{-\frac{it\Delta}{2\pi}} f\|_{L^p(\mathbb{T}^n)} \lesssim_\varepsilon N^{\alpha(n,p)+\varepsilon} \|f\|_{L^2(\mathbb{T}^{n-1})}$$

  whenever $p \geq \frac{2(n+1)}{n-1}$ and $\varepsilon > 0$ (range of p is optimal here).
- Earlier this was known only in dimensions $n = 2$ and 3. When $n \geq 4$, this was known only when $p \geq \frac{2(n+2)}{n-1}$, via the circle method in analytic number theory.
- Similarly, one can get a Strichartz estimate for all irrational tori $(\mathbb{R}/\alpha_1\mathbb{Z}) \times \cdots \times (\mathbb{R}/\alpha_{n-1}\mathbb{Z})$ (over a unit time interval), for which number theoretic arguments do not work as well. (See also Deng, Germain and Guth (2017) for the corresponding estimates over long time intervals.)

Part II: The case of the moment curve

- Next we move on to decoupling for the moment curve.
- Let $n \geq 2$, and $\Phi \colon [0,1] \to \mathbb{R}^n$ be a parametrization of the moment curve given by

$$\Phi(\xi) = (\xi, \xi^2, \dots, \xi^n), \quad \xi \in [0,1].$$

- If f is a function defined on $[0,1]$, define

$$Ef(x) = \int_0^1 f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- Let

$$f(\xi) = \sum_{k=1}^{N} \delta_{k/N}(\xi)$$

be the sum of N evenly spaced Dirac-delta functions on $[0,1]$.

- Then $Ef(x)$ is the exponential sum

$$Ef(x) = \sum_{k=1}^{N} e^{2\pi i \left(k \frac{x_1}{N} + k^2 \frac{x_2}{N^2} + \dots + k^n \frac{x_n}{N^n} \right)}.$$

- So if $s \in \mathbb{N}$, then the averaged integral

$$N^{-n^2} \int_{[0,N^n]^n} |Ef(x)|^{2s} dx$$

is equal to the number of solutions

$$(k_1, \ldots, k_{2s}) \in \{1, \ldots, N\}^{2s}$$

to the translation-dilation invariant Diophantine system

$$\begin{cases} k_1 + \cdots + k_s = k_{s+1} + \cdots + k_{2s} \\ k_1^2 + \cdots + k_s^2 = k_{s+1}^2 + \cdots + k_{2s}^2 \\ \vdots \\ k_1^n + \cdots + k_s^n = k_{s+1}^n + \cdots + k_{2s}^n, \end{cases} \tag{1}$$

which is the main object of study for the Vinogradov mean value theorem.

Vinogradov's mean value theorem

- Wooley (2014) succeeded in counting the number of solutions to (1) in $\{1, \ldots, N\}^{2s}$ when $n = 3$, using a method called efficient congruencing.
- Bourgain, Demeter and Guth (2016) approached this problem using decoupling, and this works for all n.
- Let's partition $[0, 1]$ into N disjoint intervals $\mathcal{I}_1, \ldots, \mathcal{I}_N$ of equal lengths. For $j = 1, \ldots, N$, let

$$E_j f(x) = \int_{\mathcal{I}_j} f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n$$

  so that $Ef = \sum_{j=1}^N E_j f$.
- Bourgain, Demeter and Guth (2016) proved (morally) that

$$\|Ef\|_{L^p[0,N^n]^n} \leq C_\varepsilon N^{\frac{1}{2} - \frac{n(n+1)}{2p} + \varepsilon} \left\| \|E_j f\|_{L^p[0,N^n]^n} \right\|_{\ell^2}$$

  whenever $p \geq n(n+1)$ and $\varepsilon > 0$ (range of p is optimal here; power of N is sharp up to N^ε loss).

- Thus they showed that the number of solution $(k_1, \ldots, k_{2s}) \in \{1, \ldots, N\}^{2s}$ to the Diophantine system

$$\begin{cases} k_1 + \cdots + k_s = k_{s+1} + \cdots + k_{2s} \\ k_1^2 + \cdots + k_s^2 = k_{s+1}^2 + \cdots + k_{2s}^2 \\ \vdots \\ k_1^n + \cdots + k_s^n = k_{s+1}^n + \cdots + k_{2s}^n, \end{cases}$$

is at most $C_\varepsilon N^{2s - \frac{n(n+1)}{2} + \varepsilon}$ for all $\varepsilon > 0$, whenever $s \geq \frac{n(n+1)}{2}$. (This is sharp up to the N^ε loss.)
- This bound is reasonable because among the N^{2s} choices of $(k_1, \ldots, k_{2s}) \in \{1, \ldots, N\}^{2s}$, the probability that it solves the degree j equation above is heuristically N^{-j} (think of $k_1^j + \cdots + k_s^j$ and $k_{s+1}^j + \cdots + k_{2s}^j$ as random integers in $[1, sN^j]$; the probability that they are equal is $\simeq 1/(sN^j)$).
- So if these probabilities were independent, then the number of solutions to the system would be $\simeq N^{2s} N^{-1-2-\cdots-n}$.

Part III: The case of the truncated cone

- Next we consider decoupling for the cone in $\mathbb{R}^n$, which is the setting where decoupling inequalities were first formulated.
- Let $n \geq 3$, $A \subset \mathbb{R}^{n-1}$ be the unit annulus $\{1 \leq |\xi| \leq 2\}$, and $\Phi\colon A \to \mathbb{R}^n$ be a parametrization of the truncated cone:

$$\Phi(\xi) = (\xi, |\xi|), \quad \xi \in A.$$

- If f is a function defined on A, define

$$Ef(x) = \int_A f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n.$$

- Let's partition A into N^{n-2} sectors $A_1, \dots, A_{N^{n-2}}$ of equal sizes. For $j = 1, \dots, N^{n-2}$, let

$$E_j f(x) = \int_{A_j} f(\xi) e^{2\pi i x \cdot \Phi(\xi)} d\xi \quad \text{for } x \in \mathbb{R}^n$$

so that $Ef = \sum_{j=1}^{N^{n-2}} E_j f$.

- Wolff (1999) showed that

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq C_\varepsilon N^{n-2-\frac{2(n-1)}{p}+\varepsilon} \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^p}$$

whenever $n = 3$ and $p > 74$; Wolff and Łaba (2000) showed that the same holds if $n \geq 4$ and $p > 2 + \min\{\frac{8}{n-4}, \frac{32}{3n-10}\}$.

- Note that this is ℓ^p decoupling instead of ℓ^2.
- The exponent of N is sharp up to N^ε loss; the range of exponents was not sharp.
- Bourgain and Demeter (2014) proved ℓ^2 decoupling for the optimal range of p in this context: they showed

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq C_\varepsilon N^{\frac{n-2}{2}-\frac{n}{p}+\varepsilon} \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^2}$$

whenever $n \geq 3$ and $p \geq \frac{2n}{n-2}$.
(This readily implies the results of Wolff and Łaba; one just passes from ℓ^2 to ℓ^p using Hölder.)

ℓ^2 decoupling vs ℓ^p decoupling

- So now we have ℓ^2 decoupling for the cone:

$$\|Ef\|_{L^p(\mathbb{R}^n)} \le C_\varepsilon N^{\frac{n-2}{2}-\frac{n}{p}+\varepsilon} \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^2}, \quad p \ge \frac{2n}{n-2},$$

which implies ℓ^p decoupling for the cone:

$$\|Ef\|_{L^p(\mathbb{R}^n)} \le C_\varepsilon N^{n-2-\frac{2(n-1)}{p}+\varepsilon} \left\| \|E_j f\|_{L^p(\mathbb{R}^n)} \right\|_{\ell^p}, \quad p \ge \frac{2n}{n-2}.$$

- Wolff (1999) and Łaba and Wolff (2000) already noticed that decoupling inequalities for the cone can be used to establish the local smoothing estimates for the wave equation.
- In particular, they observed that if the ℓ^p decoupling inequality for the cone holds for a certain exponent p, then

$$\left\| e^{it\sqrt{-\Delta}} f \right\|_{L^p(\mathbb{R}^{n-1}\times[1,2])} \lesssim \|f\|_{L^p_\alpha(\mathbb{R}^{n-1})}, \quad \alpha > (n-2)\left(\frac{1}{2}-\frac{1}{p}\right)-\frac{1}{p}.$$

(Here L^p_α is the Sobolev space of functions that has α derivatives in L^p.)

Local smoothing estimates

- Local smoothing estimate for the wave equation again:

$$\left\| e^{it\sqrt{-\Delta}} f \right\|_{L^p(\mathbb{R}^{n-1}\times[1,2])} \lesssim \|f\|_{L^p_\alpha(\mathbb{R}^{n-1})}, \quad \alpha > (n-2)\left(\frac{1}{2}-\frac{1}{p}\right)-\frac{1}{p}$$

- This is an improvement over the best fixed time estimate for the wave equation, which says

$$\sup_{t\in[1,2]} \left\| e^{it\sqrt{-\Delta}} f \right\|_{L^p(\mathbb{R}^{n-1})} \lesssim \|f\|_{L^p_\alpha(\mathbb{R}^{n-1})}, \quad \alpha \geq (n-2)\left(\frac{1}{2}-\frac{1}{p}\right).$$

  (One gains $1/p$ in the regularity α when one takes L^p norm in t in place of a sup norm in t.)

- The decoupling inequality of Bourgain-Demeter establishes local smoothing for the wave equation for $p \geq \frac{2n}{n-2}$.
- Conjecture: The local smoothing for the wave equation holds for all $p > \frac{2n}{n-1}$.

- Local smoothing estimates of the kind on the previous slide were first discovered by Sogge (1991).
- Shortly after that, Mockenhaupt, Seeger and Sogge (1992) observed that such inequalities (with $n = 3$) can be used to give a simple and conceptual proof of the circular maximal function theorem of Bourgain (1986) on $\mathbb{R}^2$, namely

  $$\left\| \sup_{t>0} |A_t f| \right\|_{L^p(\mathbb{R}^2)} \lesssim \|f\|_{L^p(\mathbb{R}^2)}, \quad p > 2,$$

  where $A_t f(x)$ is the average of f on a circle of radius t with center $x \in \mathbb{R}^2$.
  (See also Pramanik and Seeger (2007), who used such ideas to establish the boundedness of the maximal operator along any smooth curves of finite type in $\mathbb{R}^3$. Also see Beltran, Hickman and Sogge (2018) for local smoothing on manifolds using a variable coefficient variant of decoupling.)

Part IV: Other applications of decoupling

- One can still consider decoupling for other submanifolds, and in the past few years we saw a number of important breakthroughs from such considerations.
- Among them there's Bourgain's new record (2017) on the Lindelöf hypothesis.
- Lindelöf (1908) showed that the Riemann zeta function satisfies the bound $\left|\zeta\left(\frac{1}{2}+it\right)\right| \lesssim t^{1/4}$ as $t \to \infty$.
- He conjectured that $\left|\zeta\left(\frac{1}{2}+it\right)\right| \leq C_\varepsilon t^\varepsilon$ for any $\varepsilon > 0$.
- The power $1/4$ has been lowered by Hardy and Littlewood to $1/6$, and by Bombieri and Iwaniec (1986) to $9/56$; see also Huxley (1993, 2005), who improved the bound to $32/205$.
- Bourgain (2017) improved this exponent to $13/84$, doubling the saving over $1/6$ from the exponent $9/56$ of Bombieri and Iwaniec.
- The main new estimate (which Bourgain used to feed into the machinery of Huxley) is a bilinear decoupling inequality for certain curves in $\mathbb{R}^4$.

- Another application of decoupling is in the study of the maximal Schrödinger operator on $\mathbb{R}^{2+1}$.
- Du, Guth and Li (2017) showed that if $u(x,t)$ is the solution to the Schrödinger equation in (2+1) dimensions with initial data $f(x) \in H^s(\mathbb{R}^2)$, where $s > 1/3$, then $u(x,t)$ converges to $f(x)$ for a.e. $x \in \mathbb{R}^2$ as $t \to 0^+$.
  (Here H^s is the Sobolev space L^2_s.)
- In light of the recent example of Bourgain (2016), this is sharp up to the endpoint $1/3$.
- The result of Du, Guth and Li are obtained by estimating a maximal Schrödinger operator: they showed that if $s > 1/3$, then
$$\left\| \sup_{0<t\leq 1} |e^{it\Delta} f| \right\|_{L^3(B_1)} \lesssim \|f\|_{H^s(\mathbb{R}^2)} \tag{2}$$
  where B_1 is the unit ball in $\mathbb{R}^2$.
- One step of the proof involves a refined Strichartz inequality, that they proved using the Bourgain-Demeter decoupling inequality for the parabola.

Part V: A few words about proofs

- Let's take a brief look at how ℓ^2 decoupling is proved for the parabola in $\mathbb{R}^2$.
- Below are a few ingredients that go into the proof:

1. Wave packet decompositions
2. Kakeya type estimates
3. Multilinear reduction
4. Bourgain-Guth induction on scales

- Recall the extension operator for the parabola:

$$Ef(x) = \int_{-1}^{1} f(\xi) e^{2\pi i (x_1 \xi + x_2 \xi^2)} d\xi, \quad x \in \mathbb{R}^2.$$

- We decompose $[-1, 1]$ into the disjoint union of N intervals $\mathcal{I}_1, \ldots, \mathcal{I}_N$ of equal lengths, and let

$$E_j f(x) = \int_{\mathcal{I}_j} f(\xi) e^{2\pi i (x_1 \xi + x_2 \xi^2)} d\xi, \quad x \in \mathbb{R}^2,$$

so that

$$Ef = \sum_{j=1}^{N} E_j f.$$

- Morally, the goal is to understand the best constant $D(N)$ in the inequality

$$\|Ef\|_{L^p(B_{N^2})} \leq D(N) \left\| \|E_j f\|_{L^p(B_{N^2})} \right\|_{\ell^2}$$

for a given p.

Wave packet decompositions

- The Fourier transform of $E_j f$ is supported on a short arc C_j of length N^{-1} on the parabola, which is contained (because of the curvature of the parabola) in a rectangular slab R_j of dimensions $N^{-1} \times N^{-2}$ (think thin slabs since N is big).
- This rectangular slab R_j is oriented so that the short sides are morally parallel to the normal vector to the parabola at any point on C_j.
- Hence heuristically, we think of the modulus of $E_j f$ to be constant on boxes of dimensions $N \times N^2$ dual to R_j (think of these as long thin tubes).
- The localization of $E_j f$ to such a tube is called a wave packet.
- Each $E_j f$ is thus the sum of wave packets whose physical supports are parallel tubes, and as j varies, the directions of these tubes gradually varies.
- Ef is the sum of $E_j f$'s, so Ef is also a sum of wave packets, except that the wave packets can now be supported in tubes in different directions.

Connection to Kakeya and multilinear Kakeya

- Decoupling inequalities capture in some sense how these wave packets interfere with each other.
- Since wave packets are supported on long thin tubes, it will be helpful to know how much these thin tubes can overlap.
- Such overlaps have been studied in connection with the Kakeya conjecture (which states that every set in $\mathbb{R}^n$ that contains a unit line segment in every possible direction has Hausdorff dimension n).
- While the Kakeya conjecture is still open in dimensions 3 or above, its multilinear counterpart has been understood, thanks to the breakthroughs by Bennett, Carbery, Tao (2006) and Guth (2010) (see also Carbery and Valdimarsson (2013)).

Bourgain-Guth induction on scales

- Recall that in proving decoupling inequalities, we wanted to estimate Ef, and that E is a linear operator.
- To reduce the estimate of Ef to multilinear quantities, Bourgain and Demeter used an iteration scheme, first devised by Bourgain and Guth (2011) to study the (linear) restriction conjecture based on advances on its multilinear counterpart.
- It is at this step that the curvature of the parabola comes in: the arcs C_j on the parabola have normal vectors transverse to each other as j varies, and transversality allows one to apply multilinear Kakeya estimates.
- At the back of all this are multiple applications of *induction on scales* (which is somewhat reminiscent of the *induction on energy* in the study of dispersive PDEs); indeed multilinear Kakeya estimates are used to prove *ball inflation lemmas*, that allows one to pass from smaller spatial balls to larger ones.

$$\|Ef\|_{L^p(B_{N^2})} \leq D(N) \left\| \|E_j f\|_{L^p(B_{N^2})} \right\|_{\ell^2}$$

- The simplest manifestation of such induction on scales in decoupling is the following observation: if $D(N)$ is the best constant to the decoupling inequality (for a certain fixed $p \geq 2$), then
$$D(N_1 N_2) \lesssim D(N_1) D(N_2)$$
for all $N_1, N_2 \geq 1$.
- If $D(N) \leq N^{\varepsilon}$ for all $1 \leq N \leq 2$ (with implicit constant 1), then by induction, clearly $D(N) \lesssim (\log N)^C N^{\varepsilon}$ for all $N > 2$ (write $N = 2^{2^m}$ and iterate the previous inequality m times).
- Of course this does not work because the implicit constant is not really 1 as $\varepsilon \to 0$.
- But this suggests that induction on scales is useful; indeed induction on scales came up useful many times in the actual proof of the decoupling inequality of the parabola.