Larry Guth - Decoupling Theory

Ao Sun

July 24, 2018

This is the course note of 18.118 at MIT taught by Larry Guth in 2017 Fall. Because I am lazy so this note only contains the notes of the first semester. Please go to Larry’s homepage to see the notes of the whole semester.

Contents

1 Background and Motivation 1

2 Decoupling 5

2.1 Simplest Decoupling Problem .......................... 5
2.2 Decoupling for Paraboloid ............................ 9
2.3 Multi-Linear Kakeya Conjecture ................... 13

2.3.1 Loomis-Whitney Inequality ........................ 13
2.3.2 Tubes in Different Directions ....................... 14
2.3.3 Proof of Multilinear Kakeya Conjecture .......... 15

2.4 Application: Restriction Theory ..................... 17

2.4.1 Another Version of Restriction Theory .......... 22
2.4.2 $n = 2$ Restriction conjecture .................... 23

2.5 Bourgain’s Decoupling Theorem ...................... 24

2.5.1 Multiscale Tools ................................. 25
2.5.2 Multilinear Decoupling ............................ 26
2.5.3 Main Induction Lemma ............................ 27

1 Background and Motivation

Strichartz estimate on torus

Let us start with linear Schrödinger equation

$$\begin{cases} \partial_t u = i \Delta u \\ u(x, 0) = u_0(x) \end{cases} \tag{1}$$

The equation can be defined on $\mathbb{R}^d$, $\mathbb{T}^d = \mathbb{R}^d/(2\pi)^d$ or even more general manifold.
The Schrödinger equation describes the behavior of quantum particles. Suppose $u(x, t)$ is a solution, then

$$\int_{A} |u(x, t)|^2 dx = \text{Probability(Particle} \in A \text{ at time) t}$$

Either calculate directly or from physics meaning we can see that the total $L^2$ norm $\int_{\mathbb{R}^d} |u(x, t)|^2 dx$ is conserved. By Heisenberg principle, $u$ cannot concentrate highly somewhere, otherwise, we will know the position of the particle very precise. Mathematically, we have the following famous estimate

**Theorem 1.1** (Strichartz 1970's). If $u$ solve Schrödinger equation on $\mathbb{R}^n$, let $p = 2 \frac{d+2}{d}$, then

$$\|u\|_{L^p(\mathbb{R}^d \times \mathbb{R})} \lesssim \|u_0\|_{L^2(\mathbb{R}^d)}$$

**Notation remark:** We will always use $\lesssim$ to denote $\leq C_d$, where $C_d$ is some constant only depends on dimension.

**Corollary 1.1.** If $\|u_0\|_{L^2(\mathbb{R}^d)} = 1$, then

$$|\{(x, t) : |u(x, t)| > \lambda\}| \lesssim \frac{1}{\lambda^p}$$

This corollary somewhat quantitatively illustrate that $|u|$ can not concentrate somewhere.

On torus the similar problem is much harder. Let us first get some intuition why the problem is much harder. Let us consider a wave packet in $\mathbb{R}^d$ (see picture). The domain where $|u|$ concentrate will move in certain directions. Then intuitively two different wave packets may have their concentrate domain overlap, however that would happens at most once. On $\mathbb{T}^d$, however, the domain of two different wave packets may overlap infinite many times (see picture). As a result, the solutions to Schrödinger equation on $\mathbb{T}^d$ behave much wilder.

Let us assume the initial data satisfies

$$u_0 = \sum_{n \in \mathbb{Z}^d, |n| \leq N} a_n e^{inx}$$

we say $u_0$ has frequency no more than $N$. It is easy to check the solution is

$$u(x, t) = \sum_{n \in \mathbb{Z}^d, |n| \leq N} a_n e^{i(nx + |n|^2 t)}$$

which is $2\pi$-periodic in $x$ and $t$.

**Theorem 1.2.** (Bourgain-Demeter) If $u_0$ has frequency no more than $N$ which is a solution to Schrödinger equation, let $p = 2 \frac{d+2}{d}$, then

$$\|u\|_{L^p(\mathbb{T}^d \times [0, 2\pi])} \leq C_{d, N} \|u_0\|_{L^2(\mathbb{T}^d)}$$

for any $\epsilon > 0$.

In 1990s Bourgain conjectured theorem 1.2. He also proved the case for $d = 1, 2$. 
Number theory problem

Let us sketch the proof for the case \( d = 2 \) of theorem 1.2.

Proof. \( d = 2 \) case is very special because in this case \( p = 4 \). Then what we want to estimate is

\[
\|u\|_{L^4_{x,t}}^4 = \|u^2\|_{L^2_{x,t}}^2
\]

and \( L^2 \) norm is easy to analyze because we can use the Plancherel theorem.

\[
u^2 = \sum_{m,n} a_m a_n e^{i[(m+n)x+(|m|^2+|n|^2)t]}
\]

Let us assume the following lemma

Lemma 1.1. For all \( l \in \mathbb{Z}^3 \),

\[
\{ (m,n) \in \mathbb{Z}^2 \times \mathbb{Z}^2 \text{ such that } (m+n,|m|^2+|n|^2) = l \} \lesssim N^\epsilon
\]

So by Plancherel theorem we have

\[
\|u^2\|_{L^2_{x,t}}^2 \lesssim N^\epsilon \sum_{m,n} |a_m|^2|a_n|^2 = N^\epsilon (\sum_n |a_n|^2)^2 = N^\epsilon \|u_0\|_{L^4_x}^4
\]

Note the key step here it to apply a number theory lemma. This number theory lemma is closed related to another number theory problem.

Lemma 1.2.

\[
\{ a_1^2 + a_2^2 = M \ : a_1, a_2 \in \mathbb{N} \} \lesssim M^\epsilon
\]

Let us sketch the proof of this lemma: we can factorize

\[
M = a_1^2 + a_2^2 = (a_1 + ia_2)(a_1 - ia_2)
\]

in the ring \( \mathbb{Z}[i] \). Note \( \mathbb{Z}[i] \) has unique factorization property, so we can write \( M = p_{1}^{e_1} \cdots p_{n}^{e_n} \) for \( p_i \) are Gauss primes. Then \( a_1 + ia_2 = p_1^{e_1'} \cdots p_n^{e_n'} \) for \( 0 \leq e_i' \leq e_i \). Then we get the number of possible \( a_1 + ia_2 \) is at most \( \lesssim M^\epsilon \) for any \( \epsilon > 0 \).

The proof of \( d = 2 \) case suggests that number theory can deduce the periodic Strichartz estimate. On the other hand, periodic Strichartz estimate can also deduce some number theory result. Let us consider \( d = 1 \) case. Suppose \( u_0(x) = \sum_{1 \leq a \leq N} e^{i a x} \), then \( u(x,t) = \sum_{1 \leq a \leq N} e^{i a x + a^2 t} \).

Suppose \( d = 1 \) case of theorem 1.2 holds, we have

\[
\int \|u\|^6 \lesssim N^{3+\epsilon}
\]
We can write
\[
\int |u|^3 = \int u^3 \overline{u^3} = \int \sum_{1 \leq a_1, a_2, a_3 \leq N} \sum_{1 \leq b_1, b_2, b_3 \leq N} e^{i(a_1 + a_2 + a_3 - b_1 - b_2 - b_3)x + (a_1^2 + a_2^2 + a_3^2 - b_1^2 - b_2^2 - b_3^2)x}
\]
\[
= \# \{1 \leq a_i, b_i \leq N, a_i, b_i \in \mathbb{Z} : a_1 + a_2 + a_3 = b_1 + b_2 + b_3, a_1^2 + a_2^2 + a_3^2 = b_1^2 + b_2^2 + b_3^2\}
\]
So we get
\[
\# \{1 \leq a_i, b_i \leq N, a_i, b_i \in \mathbb{Z} : a_1 + a_2 + a_3 = b_1 + b_2 + b_3, a_1^2 + a_2^2 + a_3^2 = b_1^2 + b_2^2 + b_3^2\} \lesssim N^{3 + \epsilon}
\]

This is a very good bound for the number of solutions to the algebraic equations. There are two clues to show this. First, note that there are \(N^3\) diagonal solutions \(a_i = b_i\). So \# must be at least \(N^3\); second, we can use some probability intuition: the probability to solve \(a_1 + a_2 + a_3 = b_1 + b_2 + b_3\) is roughly \(1/N\), the probability to solve \(a_1^2 + a_2^2 + a_3^2 = b_1^2 + b_2^2 + b_3^2\) is roughly \(1/N^2\). So the probability to solve the algebraic equation is roughly \(1/N^3\). As a result, we roughly have \(N^3\) solutions, which again indicate this bound is optimal.

This problem is first well-settled by Vinogradov. He defined the following function
\[
J_{s,k}(N) := \# \{a_1^j + \cdots + a_s^j = b_1^j + \cdots + b_s^j, j = 1, \ldots, k, 1 \leq a_i, b_i \leq N\}
\]

Then previous result is just the following corollary:

**Corollary 1.2.**
\[
J_{3,2}(N) \lesssim N^{3 + \epsilon}
\]

In general, via our previous intuition about the number of diagonal solutions and the number obtained from probability, Vinogradov conjectured that

**Conjecture 1.**
\[
J_{s,k}(N) \lesssim N^s(N^s + N^{2s - \frac{k(k+1)}{2}})
\]

In particular, this conjecture is most interesting when \(s = \frac{k(k+1)}{2}\). When \(k = 2\), this is the old result as we shown previously. Vinogradov showed the conjecture is true if \(s \geq 10k^2 \log k\). Intuitively, when \(s\) is large, the problem is easier because as \(s\) grows, the counting function becomes smoother.

**Decoupling**

The main tool we are going to study in the course is decoupling.

Suppose \(\Omega \subset \mathbb{R}^n\). Assume \(\Omega = \cup \theta\). If the support of the Fourier transform of a function \(\hat{f}\) lies in \(\Omega\), then we can write
\[
f = \sum_{\theta} f_{\theta}
\]
where \(f_{\theta} = \int_{\theta} \hat{f}(\omega) e^{i\omega \cdot x} d\omega\).
Definition 1.1. We define \( D_p = D_p(\Omega = \cup \theta) \) to be the smallest constant such that

\[
\|f\|_{L^p(\mathbb{R}^n)} \leq D_p(\Omega = \cup \theta)(\sum_{\theta} \|f_\theta\|_{L^p(\mathbb{R}^n)})^{1/2}
\]

Intuitively, decoupling gives a way to analyze different frequency parts of the function we want to study. It also provides the tool to study the space in different scales. We can see this during this course.

2 Decoupling

2.1 Simplest Decoupling Problem

Let us consider the simplest decoupling problem: Let \( \Omega = [0, N] = \cup_{j=1}^N [j-1, j] \)

Let \( \theta_j = [j-1, j] \).

If the support \( \text{supp} \hat{f} \subset \Omega \), then we can write

\[
f_j(x) = \int_{\theta_j} e^{2\pi i \omega x} \hat{f}(\omega) d\omega
\]

\( D = D_p(N) \) is defined to be the smallest constant such that for any \( f \) as above,

\[
\|f\|_{L^p(\mathbb{R})} \leq D_p(N)(\sum_{j=1}^N \|f_j\|_{L^p(\mathbb{R})})^{1/2}
\]

Let us first analyze the building block \( f_1 \). All we know is \( \text{supp} f_1 \subset [0, 1] \).

Let us consider \( \eta \) is a function such that \( \eta = 1 \) on \( [0, 1] \) and \( \eta \in \mathcal{S} \) (Schwartz space). Then

\[
\hat{f}_1 = \hat{f}_1 \eta
\]

As a result,

\[
f_1 = f_1 \ast \tilde{\eta}
\]

where \( \tilde{\eta} \in \mathcal{S} \). Then we have the following global bound:

Corollary 2.1.

\[
\|f_1\|_{L^\infty(\mathbb{R})} \lesssim \|f_1\|_{L^1(\mathbb{R})}
\]

Proof.

\[
\|f_1\|_{L^\infty(\mathbb{R})} = \|f_1 \ast \tilde{\eta}\|_{L^\infty(\mathbb{R})} \leq \|f_1\|_{L^1(\mathbb{R})} \|\tilde{\eta}\|_{L^\infty(\mathbb{R})}
\]

This is a very naive bound because we only use the \( L^\infty \) bound of \( \tilde{\eta} \). However we know even more, \( \tilde{\eta} \) is not only bounded, but also rapidly decaying. Consider this we can get a local bound. For a (good enough) function \( \omega \), we will use \( \|f\|_{L^\infty(\omega)} \) to denote the weighted integral

\[
\|f_1\|_{L^\infty(\omega)} := \int_{\mathbb{R}} |f_1| \omega
\]
Lemma 2.1. If \( \text{supp} \hat{f}_1 \subset [0,1], \) let \( I \) be the unit interval, then
\[
\| f_1 \|_{L^\infty(I)} \lesssim \| f_1 \|_{L^1(\Omega_1)}
\]
where \( \Omega_1 \) is roughly 1 on \( I \), and rapidly decaying off \( I \).

Proof. Suppose \( x \in I \), we have
\[
|f(x)| \leq \int |f_1(y)||\hat{\eta}(x-y)|dy 
\leq \int |f_1(y)|\sup_{x \in I} |\hat{\eta}(x-y)|dy
\]
Then we let \( \omega_I(y) = \sup_{x \in I} |\hat{\eta}(x-y)| \) to conclude the lemma.

Remark: We could fix \( \omega_I(y) = \left( \frac{1}{1+\text{dist}(y,I)} \right)^{50} \), in particular in practice.

From the lemma above we can see \( |f_1| \) is roughly locally constant. i.e. we should not see peaks.

Since \( \text{supp} \hat{f}_j \subset [j-1,j] \) indicates \( \text{supp}(e^{-2\pi i (j-1)x} f_j) \subset [0,1] \), so the locally constant property holds for all \( f_j \).

Let us see an example. Consider \( f_1 \) is a bump height 1 function concentrate on \([-1,1]\) such that \( f_1(0) = 1 \) (see picture). Let \( f_j x = e^{2\pi i (j-1)x} f_1(x) \). Then we have \( f(0) = N \). Moreover, if \( |x| < \frac{1}{10N} < \frac{1}{10N} \) small, we have \( |f_j(x) - 1| < 1/4 \), so \( |f(x)| \sim N \) close to 0. So we have
\[
\| f \|_{L^p(\mathbb{R})} \gtrsim N \cdot N^{-\frac{1}{p}} = N^{1-\frac{1}{p}}
\]
and
\[
\| f_j \|_{L^p(\mathbb{R})} \sim 1
\]
so
\[
\left( \sum_{j=1}^{N} \| f_j \|_{L^p(\mathbb{R})} \right)^{1/2} \sim N^{1/2}
\]
By decoupling, we should get
\[
D_p(N) \gtrsim N^{\frac{1}{2}-\frac{1}{p}}
\]
This is in fact a very sharp estimate. The main proposition we want to show in this section is

**Proposition 2.1.** If \( 2 \leq p \leq \infty \), \( D_p(N) \lesssim N^{\frac{1}{2}-\frac{1}{p}} \)

Remark: This result is easy for two boundary case \( p = 2, \infty \).
If \( p = 2 \), \( \| f \|_{L^2}^2 = \sum_j \| f_j \|_{L^2}^2 \) directly from orthogonality and Plancherel formula;
If \( p = \infty \), the result comes from Cauchy-Schwartz inequality: \( \| f \|_{L^\infty} \leq \sum_{j=1}^{N} \| f_j \|_{L^\infty} \leq N^{1/2} \left( \sum_{j=1}^{N} \| f_j \|_{L^\infty}^2 \right)^{1/2} \)
As a result, the proposition is somewhat like an interpolation theorem. Just like some other interpolation theorem, the main issue is the different norm of \( |f_j| \) may be dominated by different terms for different \( p \).
Example 2.1. Let us see the case $p = 4$ for example. Suppose for all $j$, $|f_j|$ looks like constant $1$ in $[0, 1]$ and $1/N$ on interval $[1, N^3]$. Then we have $\|f_j\|_{L^2}^2 \sim N$, which is dominated by the short wide piece; but $\|f_j\|_{L^3} \sim 1$, which is dominated by $[0, 1]$ part, just like $\|f_j\|_{L^\infty}$.

So let us carefully study the different reasons for decoupling inequality holds for $p = 2$ and $p = \infty$ case.

- For $L^2$ case, the result holds because of orthogonality, as a result $\|f\|_{L^2} \sim N$.
- For $L^\infty$ case, the result holds because of triangle inequality, as a result $\|f\|_{L^\infty} \leq N$.

So for the cases between 2 and $\infty$, we may hope both phenomenon appear. Triangle inequality is obvious a local inequality. So what we need is a local version of orthogonality.

Lemma 2.2 (Orthogonality). Suppose $I$ is an unit interval. Let $f = \sum_{j=1}^N f_j$ and $\text{supp } \hat{f}_j \subset [j-1, j]$, then

$$\|f\|_{L^2(I)}^2 \lesssim \sum_j \|f_j\|_{L^2(\omega I)}^2$$

where $\omega_I$ is a rapidly decay function as before, for example $(\frac{1}{1+\text{dist}(y, I)})^{50}$.

Proof. Again choose a rapidly decay $\eta$ such that $|\eta| \sim 1$ on $I$ and $\text{supp } \hat{\eta} \subset [-1, 1]$. For example, if $I = [0, 1]$ we can choose $\eta = \check{\psi}$ where $\psi$ is a bump function on $[-1, 1]$. In general if $I = [a-1, a]$ we choose $\eta = (e^{2\pi i (a-1) \omega} \psi)^\vee$.

So

$$\int_I |f|^2 \lesssim \int_\mathbb{R} |\eta f|^2 = \int_\mathbb{R} |\hat{\eta} \ast \hat{f}|^2 = \int_\mathbb{R} \sum_j |\hat{\eta} \ast \hat{f}_j|^2$$

Note $\hat{\eta} \ast \hat{f}_j$ is supported in $[j-2, j+1]$, so any frequency lies in at most three of the interval. Hence previous term satisfies

$$\lesssim \sum_j \int_\mathbb{R} |\hat{\eta} \ast \hat{f}_j|^2 = \sum_j \int_\mathbb{R} |\hat{f}_j|^2 |\eta|^2$$

and $|\eta|$ is the weight function.

With this local orthogonality lemma, we can prove a local decoupling theorem.
Proposition 2.2 (Local decoupling). Let $I$ be an unit interval, $2 \leq p \leq \infty$, $f, f_j$ as before. Then
\[
\|f\|_{L^p(I)} \lesssim N^{\frac{1}{2} - \frac{1}{p}} \left( \sum_j \|f_j\|^2_{L^p(\omega_I)} \right)^{1/2}
\]

Proof.
\[
\int_I |f|^p \leq \left( \int_I |f|^2 \right)^{p-2} \|f\|_{L^\infty(I)}^p
\]

Apply local orthogonality on first term and triangle inequality on second term we get the above term
\[
\lesssim \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{1/2} \left( \sum_j \|f_j\|_{L^\infty(\omega_I)}^p \right)^{1/2}
\]

(By Hölder inequality)
\[
\lesssim \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{1/2} \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{p-2}
\]

(By Cauchy-Schwartz)
\[
\lesssim \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{1/2} \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{p-2}
\]

\[
\lesssim N^{\frac{1}{2} - 1} \left( \sum_j \|f_j\|^2_{L^2(\omega_I)} \right)^{1/2}
\]

Next step is to generalize local case to global case. For future purpose, we prove a little more strong lemma.

Lemma 2.3 (Parallel decoupling lemma). If $g_j$ are functions, $g = \sum g_j$. Suppose $p \geq 2$ and
\[
\|g\|_{L^p(\mu)} \leq D \left( \sum_j \|g_j\|^2_{L^2(\omega_I)} \right)^{1/2}
\]

for all $i$, where $\mu_i$’s are disjoint intervals and $\omega_i$’s are weight function. Let $\mu = \sum_i \mu_i$ and $\omega = \sum_i \omega_i$, then
\[
\|g\|_{L^p(\mu)} \leq D \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{1/2}
\]

Intuition: Let us consider one example.

Example 2.2. Suppose all $\|g_j\|_{L^p(\mu_i)} = 1$ for all $i,j$. Then
\[
\|g\|_{L^p} \leq \left( \#i \right)^{1/p} \sup_i \|g\|_{L^p(\mu_i)} \leq \left( \#i \right)^{1/p} D \left( \sum_j \|g_j\|^2_{L^p(\mu_i)} \right)^{1/2} = \left( \#i \right)^{1/p} D \left( \#j \right)^{1/2}
\]

\[
= D \left( \sum_j \|g_j\|^2_{L^p(\mu)} \right)^{1/2}
\]
So more or less the idea is similar to Minkowski inequality, where we change the order of summation and power.

Proof.

\[
\int |g|^p \mu = \sum_i \int |g|^p \mu_i = \sum_i \|g\|^p_{L^p(\mu_i)} \\
\leq D^p \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

(By Minkowski inequality) \leq D^p \left( \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

\[
\int |g|^p \mu = \sum_i \int |g|^p \mu_i = \sum_i \|g\|^p_{L^p(\mu_i)} \\
\leq D^p \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

\[
\int |g|^p \mu = \sum_i \int |g|^p \mu_i = \sum_i \|g\|^p_{L^p(\mu_i)} \\
\leq D^p \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

\[
(\text{By Minkowski inequality}) \leq D^p \left( \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

\[
\int |g|^p \mu = \sum_i \int |g|^p \mu_i = \sum_i \|g\|^p_{L^p(\mu_i)} \\
\leq D^p \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

\[
\int |g|^p \mu = \sum_i \int |g|^p \mu_i = \sum_i \|g\|^p_{L^p(\mu_i)} \\
\leq D^p \sum_i \left( \sum_j \|g_j\|^2_{L^p(\omega_i)} \right)^{p/2} = D^p \left( \sum_i \left( \sum_j \|g_j\|^p_{L^p(\omega_i)} \right)^{2/p} \right)^{p/2}
\]

\[
= D^p \left( \sum_j \|g_j\|^2_{L^p(\omega)} \right)^{p/2}
\]

Summary: the simplest decoupling illustrate the core idea of decoupling. We can view it as a interpolation argument. And the boundary cases \( p = 2 \) uses local orthogonality, and \( p = \infty \) uses local boundedness argument. In between, we use both argument. This would be the guiding principle of other decoupling theory.

2.2 Decoupling for Paraboloid

In this section, we start studying the decoupling of paraboloid \( P = \{ w \in \mathbb{R}^n : w_n = w_1^2 + \cdots + w_{n-1}^2, |w| \leq 1 \} \). Let \( \Omega = N\frac{1}{2}P \) the \( \frac{1}{2} \)-tubular neighborhood of \( P \), where \( R \) a large parameter. Decompose \( \Omega = \bigcup \theta \), where each \( \theta \) is roughly rectangle \( R^{-1/2} \times \cdots \times R^{-1/2} \). (see picture) Let us define \( D_p(R) \) decoupling constant for this situation.

Theorem 2.1 (Bourgain-Demeter). If \( 2 \leq p \leq p_s = \frac{2(n+1)}{n-1} \), then \( D_p(R) \lesssim R^\epsilon \) (means \( \leq C(\epsilon, n, p) R^\epsilon \)), i.e.

\[
\|f\|_{L^p(\mathbb{R}^n)} \lesssim R^\epsilon \left( \sum_{\theta} \|f_{\theta}\|^2_{L^p} \right)^{1/2}
\]

We first study the building blocks: each \( f_\theta \). Let us study some basic examples.

Case 1

Just like the simplest case, let \( \psi_\theta \) be smooth bump function supported in \( \theta \) with height 1.

Lemma 2.4. \( |\psi_\theta(x)| \sim |\theta| \) on \( \theta^* \), and is rapidly decay.

Here \( \theta^* \) is the dual domain of \( \theta \). (see picture)

Proof. Let \( \omega_\theta \) be the center of \( \theta \).
\[ \hat{\psi}_\theta(x) = \int_\theta e^{2\pi i \omega x} \psi_\theta(\omega) d\omega = e^{2\pi i \omega \theta x} \int_\theta e^{2\pi i (\omega - \omega \theta) x} \psi_\theta d\omega \]

If \( x \in \frac{1}{m^2} \theta^* \), \( |(\omega - \omega \theta) x| < \frac{1}{m} \) for any \( \omega \in \theta \). So we have

\[ |\hat{\psi}_\theta(x)| \sim \int |\psi_\theta| \sim |\theta| \]

If \( x \) is far from \( \theta^* \), the integration by part several times we know \( |\hat{\psi}_\theta| \) is rapidly decay.

**Remark:** Since the scale of \( \theta \) is very small, \( \hat{\psi}_\theta \) is oscillatory \( \sim e^{2\pi i \omega y x} |\theta| \psi_\theta \).

**Further Case 1**

Consider \( \sum_k a_k \hat{\psi}_\theta(x - x_k) \) be sum of wave packets. Then on each domain, we hope the same estimates holds.

**Lemma 2.5** (Locally constant lemma). If \( \text{supp} \hat{f}_\theta \subset \theta \), \( T \) be a translated \( \theta^* \). Then

\[ \|f_\theta\|_{L^\infty(T)} \lesssim \|f_\theta\|_{L^1(\omega T)} \] (4)

Here \( \|f_\theta\|_{L^1(\omega T)} \) is the weighted average

\[ \|f_\theta\|_{L^1(\omega T)} = \int |f_\theta| \omega_T \int \omega_T \]

and \( \omega \) is roughly 1 on \( T \) and rapidly decaying.

We just do the estimate again as in case 1. Note now the test function is almost like \( \psi_\theta \).

As a result, we see that for any building block \( f_\theta \) defined on single \( \theta \), locally boundedness still holds.

Next let us study a very special example.

**Example 1:**

Consider each \( f_\theta \) relates to a single wave packet through 0 (see picture). Let \( f_\theta(0) = 1 \). Then

\[ |f(0)| = \# \theta \sim (R^{1/2})^{n-1} \]

So if \( |x| < \frac{1}{m} \) small, note in most of theta the wave length is 1, we have

\[ |f(x)| \sim R^{\frac{n-1}{2}} \]

Thus we have

\[ \|f\|_{L^p} \geq \|f\|_{L^p(B_1)} \gtrsim R^{\frac{n-1}{2}} \]

On the other hand

\[ \|f_\theta\|_{L^p} \sim |\theta^*|^\frac{1}{p} \sim (R \cdot R^{-\frac{n-1}{2}})^\frac{1}{p} = R^{\frac{n+1}{2p}} \]

10
Hence
\[
\left( \sum_{\theta} \left\| f_{\theta} \right\|_{L^p}^2 \right)^{1/2} = R^{\frac{n+1}{4}} + \frac{n+1}{4p}
\]

Then we have an estimate
\[
D_p(R) \gtrsim R^{\frac{n+1}{4} - \frac{n+1}{2p}}
\]

**Remark:** The exponential greater than 0 if and only if \( p > p_s \). So this result partially explain why we need \( p_s \).

**Remark:** The estimate here does not rely on the shape paraboloid. In fact, any \( \sim R^{\frac{n+1}{4}} \) number of packet could give us such bound.

Example 1 is an heuristic example. It gives us the sharp bound for paraboloid decoupling.

**Theorem 2.2** (Bourgain-Demeter).
\[
D_p(R) \lesssim \begin{cases} 
R^\epsilon & 2 \leq p \leq p_s \\
R^\epsilon R^{\frac{n+1}{4} - \frac{n+1}{2p}} & p_s \leq p \leq \infty
\end{cases}
\]

**Remark:** Again, the case \( p = 2, p = \infty \) are easy. In fact the proof of these two cases is exactly the same as the simplest case. (see picture)

Note Example 1 shows that this bound is sharp for \( p_s \leq p \leq \infty \).

**Example 2**
Let us consider now we have disjoint copies of Example 1. i.e. suppose \( f, f_\theta \) are just like in Example 1, and \( g = \sum f, g_\theta = \sum f_\theta \). Then we have
\[
\left\| g \right\|_{L^p} = N^{\frac{1}{p}} \left\| f \right\|_{L^p}, \left\| g_\theta \right\|_{L^p} = N^{\frac{1}{p}} \left\| f_\theta \right\|_{L^p}
\]

Then an easy computation indicates we get the same \( D_p \). So the bound is again sharp for \( p_s \leq p \leq \infty \).

So if all the packets do not overlap too much, then the situation does not change too much.

**Example 3**
Let us now consider \( n = 2 \) case. For all \( \theta \), let
\[
h_\theta = \sum N \text{ wave packets in a row}
\]
where \( N \sim R^{\frac{1}{2}} \). see picture. Note the scale of a side of \( \theta^* \) is \( \sim R^{\frac{1}{2}} \), so we can assume \( \left| h_\theta \right| \sim \chi_{B_R} \), i.e. it is almost 1 on each wave packets as a result of locally constant property.

Now we have a natural question: How many constructive interference can happen?

When \( n = 2 \), By previous example we have \( |h(x)| \lesssim R^{1/2} \). Now we consider Red to be the set \( \{ x : |h(x)| \gtrsim R^{1/2} \} \).

Then by the decoupling inequality, we get
\[
|\text{red}|(R^{1/4})^6 \leq \left\| h \right\|_{L^6}^6 \lesssim R' \left( \sum \left\| h_\theta \right\|_{L^6}^2 \right)^{6/2}
\]
Hence

\[ |\text{red}| \lesssim R^* R^{1/2} = R^* N \]

We see that this volume is not significant bigger than what in Example 2.

**Example 4**

Consider

\[ f = \sum_{j=1}^{N} e^{2\pi i \left( \frac{j}{N} x_1 + \frac{j^2}{N^2} x_2 \right)} \eta_{B_R} \]

where \( N = R^{1/2} \) as in previous example and \( \eta_{B_R} \) just like characterize function.

Then \( f(0) = N = R^{1/2} \). Moreover, \( f(aN) = N = R^{1/2} \) for \( a \in \mathbb{Z}, |a| \leq R^{1/2} \).

So we again get the size of "red" points.

Summary: The decoupling for intervals and decoupling for paraboloid have many similar features.

<table>
<thead>
<tr>
<th>Decoupling for Intervals</th>
<th>Decoupling for Paraboloid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local orthogonality</td>
<td>Local orthogonality (more precise later)</td>
</tr>
<tr>
<td>Local constant</td>
<td>Local constant (for translates of ( \theta^* ))</td>
</tr>
</tbody>
</table>

The only difference is the tiling. In intervals case, \( \mathbb{R} \) can be tiled by any unit intervals, hence for any \( j \), \( |f_j| \sim \text{constant} \) on each interval. But for paraboloid case, each \( \theta \) has its own tiling. Which means local constant holds on each tile, but may varies a lot when combined all tiling together.

For example (see picture), consider \( |f_{\theta_j}(x)| \sim \chi_{B_R} \) like in Example 3. Let \( f_{\tau_1} = \sum_{\theta \subset \tau_1} f_\theta \), and suppose \( |f_{\tau_1}| \) big in sparse set of translates of \( \tau_1^* \). Same for \( \tau_2, \tau_2, \cdots \). Then \( |f_{\tau_j}| \) in general big in different places.

If \( |f| \) is very focused somewhere, we want \( f_{\tau_j} \)'s are also very focused in the same places. Then it is reasonable to talk about locally constant.

Then a natural question arise: how to evaluate the overlap of the cylinders? Say \( T_j \)'s are cylinders in \( \mathbb{R}^n \) with radius 1 and length \( L \). Suppose \( T_j \) points in direction \( \theta_j \in S^{n-1} \), and all these directions \( \{\theta_j\} \) form a \( 1/L \)-separated \( 1/L \)-net in \( S^{n-1} \). Then \( \#\theta_j \sim L^{n-1} \). We can ask how big is \( \| \sum \chi_{T_j} \|_{L^p} \)?

This is the **Kakeya maximal function conjecture**.

**Example 2.3.** Suppose each \( T_j \) centered at \( 0 \). Let \( g_1(x) = \sum \chi_{T_j} \). This seems like the most overlapping case. Intuitively \( g_1(x) \sim L^{n-1} \) for \( x \in B_1 \), and \( g_1(x) \gtrsim 1 \) for \( x \in B_L \). In general by some interpolation argument

\[ g_1(x) \sim \left( \frac{L}{|x|} \right)^{n-1}, 1 \leq |x| \leq L \]

The **Kakeya conjecture** is

\[ \| \sum \chi_{T_j} \|_{L^p} \lesssim L' \| g_1 \|_{L^p} \]
This conjecture is true in $n = 2$, and still open in $n \geq 3$.

Remark: Besicovitch gave examples where $| \cup T_j | \sim \frac{|B^n(L)|}{\log L}$. Hence we need $\log L$ term in the conjecture for $1 \leq p < \frac{n}{n-1}$.

2.3 Multi-Linear Kakeya Conjecture

2.3.1 Loomis-Whitney Inequality

Let us start with a warm up problem. Suppose $U \subset \mathbb{R}^3$ is a nice enough set, such that

- The area of projection of $U$ to $xy$ plane $\leq A$
- The area of projection of $U$ to $yz$ plane $\leq B$
- The area of projection of $U$ to $zx$ plane $\leq C$

We can ask a question: how large can $|U|$ be?

Let

$$U_z := \{(x, y) : (x, y, z) \in U\}$$

which is a section of $U$. By Fubini theorem we know

$$|u| = \int_{\mathbb{R}} |U_z| dz$$

Let us estimate the area of this piece of section. On one hand, $|U_z| \leq A$ which is the projection of $U$ to $xy$-plane. On the other hand, suppose $X(z)$ is the length of the projection of $U_z$ to $x$-axis, and $Y(z)$ is the length of the projection of $U_z$ to $y$-axis. Then $|U_z| \leq X(z)Y(z)$. Moreover

$$\int X(z)dz \leq B, \int Y(z)dz \leq C$$

Then

$$|U| = \int |U_z| \leq \int A^{1/2}X^{1/2}(z)Y^{1/2}(z) \leq A^{1/2}(\int X(z)dz)^{1/2}(\int Y(z)dz)^{1/2} \leq (ABC)^{1/2}$$

This bound is sharp if $U$ is a rectangle.

Corollary 2.2 (Isoperimetric inequality). $|U| \leq |\partial U|^{3/2}$ if $U$ is a bounded region.

Proof. Just note the area of the projection $\leq$ the area of the boundary. □

In 1951, Loomis and Whitney generalize the above fact to the following theorem:
Theorem 2.3 (Loomis-Whitney 1951). If \( f_j : \mathbb{R}^{n-1} \to \mathbb{R}, j = 1, \ldots, n \) are (regularity good enough) functions and \( \pi_j : \mathbb{R}^n \to \mathbb{R}^{n-1} \) are projections which forgets the \( j \)th coordinate, then

\[
\int_{\mathbb{R}^n} \prod_{j=1}^{n} (f_j \circ \pi_j)^\frac{1}{n} \leq \prod_{j=1}^{n} \left( \int_{\mathbb{R}^{n-1}} f_j \right)^\frac{1}{n-1}
\]

We can see this is a generalization as follows: Let \( f_j = \chi_{\pi_j(U)} \) for \( U \) be a subset of \( \mathbb{R}^n \). Then

\[
|U| \leq \text{LHS} \leq \prod_{j=1}^{n} (|\pi_j(U)|)^{\frac{1}{n-1}}
\]

In particular when \( n = 3 \) we get our previous result.

The proof of Loomis-Whitney theorem uses induction on \( n \). We should cleverly use Hölder. Let us skip the proof here. (Maybe complete it later)

2.3.2 Tubes in Different Directions

Loomis-Whitney theorem can be used to study the overlaps of the tubes. Let us first study the simplest case.

**Setup 0:**

Let \( l_{j,a} \) be lines in \( \mathbb{R}^n \) parallel to \( x_j \)-axis for \( j = 1, \ldots, n \) and \( a = 1, \ldots, N_j \).

Let \( T_{j,a} \) be the characteristic function of 1-neighborhood of \( l_{j,a} \).

**Corollary 2.3.**

\[
\int_{\mathbb{R}^n} \prod_{j=1}^{n} \left( \sum_{a=1}^{N_j} T_{j,a} \right)^{\frac{1}{n-1}} \lesssim \prod_{j=1}^{n} N_j^{\frac{1}{n-1}}
\]  \hspace{1cm} (7)

The left hand side in fact characterize the overlap of these tubes in different directions.

**Example 2.4.** Suppose \( Q_s \) is the cube of side \( S \). Consider it is covered by disjoint tubes for each direction.

Then \( N_j \sim s^{n-1} \), \( LHS \sim \int_{Q_s} 1 = s^n \), and \( \text{RHS} \) also \( \sim s^n \).

**Example 2.5.** Let us consider \( n = 2 \), and there are multiplicity \( N_1 \) tubes in \( x_1 \) direction, and multiplicity \( N_2 \) tubes in \( x_2 \) direction.

Then \( LHS \sim \int_{B_1} \prod_{j=1}^{n} N_j^{\frac{1}{n-1}} \sim \prod_{j=1}^{n} N_j^{\frac{1}{n-1}} \).

In either case the estimate is sharp.

**proof of Corollary.** Just take \( f_j = \sum_a \chi_{D_{j,a}} \). Here \( D_{j,a} \) is the unit disk in \( \mathbb{R}^{n-1} \) with center is the projection of line \( l_{j,a} \). Then \( \sum_a T_{j,a} = f_j \circ \pi_j \).

Then the corollary is a result of Loomis-Whitney theorem. \( \square \)
Setup 1:
Now let us consider the tubes can tilt a little bit. Suppose $l_{j,a}$ are lines in $\mathbb{R}^n$, which are no longer parallel to the axis but the angle between $l_{j,a}$ and $x_j$-axis is less than $\frac{1}{100n}$. Again $T_{j,a}$ is the characteristic function of 1-neighborhood of $l_{j,a}$. Does inequality (7) still true?

When $n = 2$, the answer is yes. In fact, LHS of (7) is

$$\int_{\mathbb{R}^2} (\sum_a T_{1,a}) (\sum_b T_{1,b}) = \sum_{a,b} \int_{\mathbb{R}^2} T_{1,a} T_{2,b}$$

So we only need to check two tubes case. Note since each of them has very tiny angle between the axis and itself, thus the overlap is $\sim 1$. Hence

$$\sum_{a,b} \int_{\mathbb{R}^2} T_{1,a} T_{2,b} \lesssim N_1 N_2$$

Setup 2:
In the paraboloid problem we want to study, the tubes are not only tilting, but also bending. Let us consider the banding setup.

Let $\gamma_{j,a}$ be a curve such that for any $x \in \gamma_{j,a}$, the angle between the tangent line of $\gamma_{j,a}$ at $x$ and $x_j$-axis is less than $\delta = \frac{1}{100n}$. Let $T_{j,a}$ be the characteristic function of 1-neighborhood of $\gamma_{j,a}$.

If $n = 2$, then just like setup 1 the inequality (7) is still true. However for higher $n$, we have scary example:

**Example 2.6** (Csornyei). If $n = 3$, there exists $\gamma > 0$ depending on $\delta$ and an example such that

$$\int_{\mathbb{R}^3} \prod_{j=1}^{3} (\sum_{a=1}^{N_j} T_{j,a})^{\frac{1}{2}} \geq N^{\gamma} (\prod_{j=1}^{3} N_j^{\frac{1}{2}})$$

So in general we can not hope inequality (7) holds for bending case.

2.3.3 Proof of Multilinear Kakeya Conjecture

**Theorem 2.4** (Benett-Carbery-Tao 05'). Consider Setup 1. Suppose $Q_s$ is the cube of side $s$. Then

$$\int_{Q_s} \prod_j (\sum_a T_{j,a})^{\frac{1}{s}} \lesssim s^c \prod_{j=1}^{n} N_j^{\frac{1}{s} - 1}$$

**(Remark)**: Even $s^c$ can be removed, but the proof of this is a relevant idea.

The main idea is to estimate in different scale. In very tiny scale, the tubes can even ”cover” the whole cube, and the result is easy. Then we pass to a different scale, and again and again to get the estimate.

The first step is reduce to the case for the angle between $l_{j,a}$ and $x_j$-axis less than some $\delta$ small.
Lemma 2.6 (Main Lemma). For any $\epsilon > 0$, there exists $\delta > 0$ such that if the angle between $l_{j,a}$ and $x_j$-axis less than $\delta$, then inequality (8) holds.

Let us show this lemma implies the main theorem.

Proof. Let $S_j = \{ \theta \in S^{n-1} : \text{angle (} \theta, e_j \text{)} < \frac{1}{100n} \}$. Here $e_j$ is the $j$-axis direction. Then we can decompose $S_j$ into set of smaller tilting angles’ set $S_j = \bigcup_b S_{j,b}$, where each $S_{j,a}$ has diameter less then $\delta/10$.

Let $g_j = \sum_{T_{j,a}}$, and $g_{j,b} = \sum_{\text{directions in } S_{j,b}} T_{j,a}$. We have $g_j = \sum_b g_{j,b}$.

Then

$$\hat{Q} \prod_j \left( g_{j,b} \right)^{1/n} \leq \prod_j N_j(Q) \prod_j \left( g_{j,b} \right)^{1/n}$$

By main lemma each term $\int \prod_j (g_{j,b})^{1/n}$ is controlled. In fact, if $S_{j,b}$ lies in $\delta$-neighborhood of $e_j$, then we can directly apply the main lemma; otherwise, we correct by linear change of variable such that the center of $S_{j,b}$ be $e_j$. Since the angle has a uniform bound $1/100n$, the Jacobian of the change of variable is uniformly bounded. As a result we can bound RHS of the above inequality.

Then we get the main theorem. \qed

Now we only need to prove the main lemma.

proof of Main Lemma. Let us start with scale $\delta^{-1}$.

Lemma 2.7. If $Q$ has side length $\leq \delta^{-1}$, then

$$\int_Q \prod_j \left( \sum_a T_{j,a} \right)^{1/n} \leq \prod_j N_j(Q)$$

where $N_j(Q) = \# \{ T_{j,a} \text{ intersects } Q \}$

Proof. Since the angle between every $l_{j,a}$ and $x_j$-axis is small, we have $T_{j,a} \cap Q \subset \tilde{T}_{j,a}$, here $\tilde{T}_{j,a}$ is parallel to $x_j$-axis with radius $\lesssim 1$. Then

$$\text{LHS} \leq \int_Q \prod_j \left( \tilde{T}_{j,a} \right)^{1/n} \leq \text{RHS}$$

by Loomis-Whitney inequality. \qed

Now we are going to next scale. Consider cube $Q_{\delta^{-2}}$ to be the cube of side length $\delta^{-2}$. We sub-divide it into cubes of side length $\delta^{-1}$. In order to apply lemma 2.7 on each $Q_{\delta^{-1}}$, we need to know $N_j(Q_{\delta^{-1}})$ for each $j$.

Denote $T_{j,a,w}$ to be the characteristic function of $w$-neighborhood of $l_{j,a}$. Then $T_{j,a}$ is just $T_{j,a,1}$.

An observation of different scales: if side length of $Q \leq \frac{1}{100n} \delta^{-1}$, and $T_{j,a}$ intersects $Q$, then $T_{j,a,\delta^{-1}} \equiv 1$ on $Q$. 16
Lemma 2.8. If $\frac{1}{20n}\delta^{-1} \leq \text{side } Q \leq \frac{1}{10n}\delta^{-1}$, then
\[
\int_Q \prod_{j=a=1}^{N_j} \left( \sum_{T_{j,a}} \right)^\frac{1}{n-1} \leq C(n)\delta^n \int_Q \prod_{j=a=1}^{N_j} \left( \sum_{T_{j,a,\delta^{-1}}} \right)^\frac{1}{n-1}
\]

Proof. Given lemma 2.7, it is only necessary to check $\text{RHS} \gtrsim \prod_{j} N_{j}(Q)^\frac{1}{n-1}$. By the observation above, $\sum_{j} T_{j,a,\delta^{-1}}(x) \geq N_{j}(Q)$ for $x \in Q$. Note $|Q| \gtrsim \delta^{-n}$, we proved lemma 2.8.

Lemma 2.9.
\[
\int_{Q_s} \prod_{j} \left( \sum_{T_{j,a}} \right)^\frac{1}{n-1} \leq C(n)\delta^n \int_{Q_s} \prod_{j} \left( \sum_{T_{j,a,\delta^{-1}}} \right)^\frac{1}{n-1}
\]

Proof. Just sub-divide $Q_s$ into small cubes as in lemma 2.8, then add them up.

Lemma 2.10. For any $k$, we have
\[
\int_{Q_s} \prod_{j} \left( \sum_{T_{j,a,\delta^{-k}}} \right)^\frac{1}{n-1} \leq C(n)\delta^n \int_{Q_s} \prod_{j} \left( \sum_{T_{j,a,\delta^{-k-1}}} \right)^\frac{1}{n-1}
\]

Proof. We just rescale lemma 2.9.

Now say $s = \delta^{-m}$. Iterate lemma 2.10 we get
\[
\int_{Q_s} \prod_{j} \left( \sum_{T_{j,a}} \right)^\frac{1}{n-1} \leq C(n)^m \delta^{-mn} \int_{Q_s} \prod_{j} \left( \sum_{T_{j,a,s}} \right)^\frac{1}{n-1}
\]
\[
= C(n)^m \delta^{-n} \int_{Q_s} \prod_{j} N_{j}^\frac{1}{n-1}
\]
\[
= C(n)^m \prod_{j=1}^{n} N_{j}^\frac{1}{n-1}
\]

Note $s = \delta^{-m}$, so $m = \frac{\log s}{\log(1/\delta)}$. Hence $C(n)^m = s^{\frac{\log C(n)}{\log(1/\delta)}}$. If we choose $b \to 0$ small enough, then $\frac{\log C(n)}{\log(1/\delta)} \leq \epsilon$ small enough.

Then we finished the proof.

2.4 Application: Restriction Theory

In this section, we will see one application of multilinear Kakeya conjecture, which is the restriction theory.

The Theme of this section is the following problem: Suppose $\text{supp } \hat{f} \subset \Omega$. How does the geometry of $\Omega$ relate to the "geometry" of $f$?
Example 2.7. Let $\Sigma$ be a hypersurface in $\mathbb{R}^n$, for instance $S^{n-1}$. If

$$f = E_\Sigma \phi = \int_{\Sigma} e^{2\pi i \omega \cdot \phi(\omega)} d\omega$$

Then $\text{supp} \hat{f} \subset \Omega$.

Example 2.8. Consider the eigenfunction $f$ of Laplacian operator:

$$\Delta f = -f$$

Then this is equivalent to $\text{sup} \hat{f} \subset S^{n-1}$. Moreover, if $f$ satisfies some little decay property, then $f = E_{S^{n-1}} \phi$.

However, even we can write down the integral form of the solution in previous example, it is hard to read things we care about from this solution. What we want is like a dictionary between the properties of $\phi$ and properties of $f$. For example, what do we know about $f$ if $\|\phi\|_{L^2} \leq 1$? More generally, what do we know about $f$ if $\|\phi\|_{L^q} = 1$?

In 1960s-1970s, Stein asked the following question: Given $\|\phi\|_{L^q} = 1$, what do we know about $\text{sup} \|E\phi\|_{L^p(B_R)}$?

Let us see some basic examples.

Example 2.9. Let us consider the case of single wave map. Let $\theta \subset S^{n-1}$ which is a $R^{-\frac{1}{2}}$-cap in the sphere.

Let $\phi_\theta$ to be the smooth bump on $\theta$ such that it is almost 1 inside $\theta$. Then

$$|E\phi_\theta(x)| \sim \text{vol}_{n-1}(\theta) \sim R^{-\frac{n-1}{2}}, \text{ if } x \in \theta^*$$

Then we know

$$\|E\phi_\theta\|_{L^p(B_R)} \sim R^{-\frac{n-1}{2}} |\theta^*| \sim R^{-\frac{n-1}{2} + \frac{n+1}{2p}}$$

Note

$$\|\phi_\theta\|_{L^p} \sim \text{Vol}_{n-1}(\theta)^{\frac{1}{p}} \sim R^{-\frac{n-1}{2p}}$$

So we have an intuition about the estimate.

Example 2.10. Let us consider a more complicated case. Suppose $S^{n-1} = \sqcup \theta$, each with radius less than $R^{-\frac{1}{2}}$, and $\text{Vol}_{n-1} \theta \sim R^{-\frac{n-1}{2}}$. The number of $\theta \sim R^{2-n}$.

Let $\phi_2$ be constant 1 on the sphere, or (more or less equivalently) $\sum \phi_\theta$, where each $\phi_\theta$ is like a single wave packet in the previous example.

Then in the place $B_1$ where every $\theta$ intersect and have small frequency change, we have $|E\phi_2(x)| \sim 1$ on $B_1$. Outside the domain of many intersection, we have $|E\phi_2(x)| \sim \text{Vol}_{n-1}(\theta)$ on $B_R \setminus B_{R/2}$.

By interpolation argument, we can see

$$\|E\phi_2\|_{L^p(B_R)} \sim \begin{cases} \text{if } p > \frac{2n}{n-1} & B_1 \text{ dominates, } \|E\phi_2\|_{L^p} \sim 1 \\ \text{if } p < \frac{2n}{n-1} & \text{outer part dominates, } \|E\phi_2\|_{L^p} \sim R^{-\frac{n-1}{2} - \frac{n}{p}} \end{cases}$$
From the above heuristic examples, the following conjecture arise:

**General Conjecture:**

\[
\frac{\|E\phi\|_{L^p(B_R)}}{\|\phi\|_{L^p}} \lesssim R^\epsilon \max(\text{Constant in previous examples})
\]

Here \(R^\epsilon\) is only necessary for very few \(p,q\).

When \(q = 2\), some result is known:

**Theorem 2.5** (Tomas-Stein, Strichartz). If \(p \geq \frac{2(n+1)}{n-1}\), \(\|E\phi\|_{L^p(B_R)} \lesssim \|\phi\|_{L^2}\).

This can be seen as a one end sharp case of the estimate. So Stein conjecture the following result:

**Conjecture** (Stein) \(\|E\phi\|_{L^p(B_R)} \lesssim \|\phi\|_{L^\infty}\) if \(p > \frac{2n}{n-1}\).

This result is proved by Fefferman when \(n = 2\). It is still open in higher dimensions.

There are several reasons that why \(n = 2\) case is simpler. Dimension 2 means the geometry is much simpler. For example, 2-dimensional Kakeya conjecture is much more well understood than higher dimensions. Also \(p = 4\) if \(n = 2\), which makes many analysis simpler (recall this case for the similar problem we dealt in previous sections).

One approach to this problem is the wave packet approach. Intuitively, let \(S = \sqcup \theta\) where \(\theta\) are \(R^{-\frac{1}{2}}\) radius caps of \(S^{n-1}\). Let \(\phi = \sum_{\theta} \phi_{\theta}\). Then intuitively we have locally constant property

\[
|E\phi(x)| \sim \text{constant on translates of } \theta^*
\]

Then

\[
|E\phi(x)| \sim \sum_{T \text{ parallel to } \theta^* \text{ disjoint}} a_T \chi_T
\]

We can check by Plancherel identity

\[
\sum a_T^2 \sim R^\text{some power} \|\phi_{\theta}\|_{L^2}^2
\]

Next thing to do is to study how tubes overlap each other. For this purpose, we define for each \(B_{R^{1/2}}\)

\[
\mu(B_{R^{1/2}}) = \sum_{T \cap B_{R^{1/2}} \neq \emptyset} a_T^2
\]

We can ask how many \(B_{R^{1/2}}\) such that \(\mu \sim 2^k\) for some \(k\)? This is closely relate to Kakeya problem.

**Remark:** On \(B_{R^{1/2}}\), even we know \(\mu(B_{R^{1/2}})\), \(\|E\phi\|_{L^p(B_{R^{1/2}})}\) is still hard to understand because the wave packets oscillate.

For the case \(q \neq 2\), the problem is hard because the Kakeya problem is hard. However even we know the result of Kakeya problem, this restriction problem is still very hard because the different scales would not match up. So what we hope to do is under some situation that we know both the Kakeya problem well and how scale match up well.
Theorem 2.6 (Multilinear Restriction, Bennett-Carbery-Tao 2015). For special case \( \Sigma \subset S^{n-1} \) which are the spherical caps \( \frac{1}{100n} \)-neighborhood of \( e_j \). Let \( \phi_j : \Sigma_j \to \mathbb{C} \) and \( f_j = E \phi_j \). Then

\[
\| \prod_j |f_j|^\frac{1}{n} \|_{L^{\frac{2n}{n-1}}(B_R)} \lesssim R^c \prod_{j=1}^n \| \phi_j \|_{L^2}^{\frac{1}{n}} \tag{9}
\]

Remark: Suppose \( \phi = \sum \phi_j \), the the RHS \( \lesssim R^c \| \phi \|_{L^2} \lesssim \| \phi \|_{L^\infty} \) by locally constant property. Then we get Stein conjecture for the case of \( p = \frac{2n}{n-1} \) with a \( R^c \) revision version. So people conjecture as an extension to Stein conjecture we should have

\[
\| f_j \|_{L^\frac{2n}{n-1}(B_R)} \lesssim R^c \| \phi_j \|_{L^\infty}
\]

However, another result for \( \lesssim \| \phi_j \|_{L^2} \) is not true in general even for single wave packet. In fact we have example such that

\[
\| f_j \|_{L^\frac{2n}{n-1}(B_R)} \gg \| \phi_j \|_{L^2}
\]

Let us sketch the core idea of the proof. Instead of proving the original Bennett-Carbery-Tao version, let us state the following version of multilinear restriction problem:

Theorem 2.7 (Multilinear Restriction). Suppose \( \Sigma_1, \cdots, \Sigma_n \) are \( C^2 \) hypersurface in \( \mathbb{R}^n \), with diameter \( \leq 1 \) and \( |\text{curvature}| \lesssim 1 \). Moreover, if \( \omega_j \in \Sigma_j \), we assume angle between the normal of \( \Sigma_j \) at \( \omega_j \) and \( e_j \leq \frac{1}{100n} \). Let \( f_j = E \Sigma_j \phi_j \) we have supp \( f_j \subset N_{\frac{1}{100}} \Sigma_j \). Then

\[
\| \prod_j |f_j|^\frac{1}{n} \|_{L^{\frac{2n}{n-1}}(B_R)} \lesssim R^c \prod_{j=1}^n \| f_j \|_{L^2(\omega_B R)} \tag{10}
\]

here

\[
\| g \|_{L^2(\omega)} := \left( \frac{\int |g|^p \omega}{\int \omega} \right)^{\frac{1}{p}}
\]

Tools and Sketch

We will use the tools we are familiar with from previous sections. Since we only want to sketch the key idea of the proof here, we would not want to be in tangle with analysis details, though they are very important. So here we will introduce a White lie version of the tools for the sketch.

Local Orthogonality:

\[
\| f_j \|_{L^2(B(R^{1/2}))}^2 \lesssim \sum_\theta \| f_j, \theta \|_{L^2(\omega_{B(R^{1/2})})}^2 \tag{11}
\]

Local Orthogonality (White Lie Version):

\[
\| f_j \|_{L^2(B(R^{1/2}))}^2 \sim \sum_\theta \| f_j, \theta \|_{L^2(B(R^{1/2}))}^2 \tag{12}
\]
Locally Constant: If $T$ is a translation of $\theta^*$

$$\|f_{j, \theta}\|_{L^\infty(T)} \lesssim \|f_{j, \theta}\|_{L^1(\omega_T)}$$  \hspace{1cm} (13)

Locally Constant (White Lie Version): If $T$ is a translation of $\theta^*$, then for $x_1, x_2 \in T$

$$|f_{j, \theta}(x_1)| \sim |f_{j, \theta}(x_2)|$$  \hspace{1cm} (14)

Multilinear Kakeya:

If $T_{j,a}$ is the characteristic function of 1-neighborhood of line $l_{j,a}$ and the angle between $l_{j,a}$ and the $x_j$-axis is no more than $\frac{1}{100}$, then

$$\hat{Q}_s \prod_{j} (\sum_{\sum_{a=1}^{N_j} T_{j,a}})^{\frac{1}{n}} \lesssim s^n \prod_{j=1}^{n} N_j^{\frac{1}{n-1}}$$  \hspace{1cm} (15)

Consider $g_j = \sum T_{j,a}$, we may also use the version connecting different scales:

$$\int_{Q_s} \prod_{j} (g_j)^{\frac{1}{n}} \lesssim s^n \prod_{j=1}^{n} (\int_{Q_s} g_j)^{\frac{1}{n}}$$  \hspace{1cm} (16)

Now we can sketch the proof scheme.

Proof:

$$\int_{B_R} \prod_{j} |f_j|^{\frac{1}{n}} \sim \text{Average}_{B_{R^{1/2}} \subset B_R} \int_{B_{R^{1/2}}} \prod_{j} |f_j|^{\frac{1}{n}}$$

(By induction) $\lesssim R^{2n} \text{Average}_{B_{R^{1/2}} \subset B_R} \prod_{j=1}^{n} \|f_j\|_{L(B_{R^{1/2}})}^{\frac{1}{n}}$.

(Local Orthogonality) $\sim R^n \text{Average}_{B_{R^{1/2}} \subset B_R} \prod_{j=1}^{n} (\int_{B_{R^{1/2}}} \sum_{\theta} |f_{j, \theta}|^2)^{\frac{1}{n}}$.

(Locally Constant) $\sim R^n \text{Average}_{B_{R^{1/2}} \subset B_R} \int_{B_{R^{1/2}}} \prod_{j=1}^{n} (\sum_{\theta} |f_{j, \theta}|^2)^{\frac{1}{n}}$

$$= R^n \int_{B_R} \prod_{j=1}^{n} (\sum_{\theta} |f_{j, \theta}|^2)^{\frac{1}{n}}$$

(Multilinear Kakeya) $\lesssim R^{2n} \prod_{j=1}^{n} (\int_{B_R} \sum_{\theta} |f_{j, \theta}|^2)^{\frac{1}{n}}$.

(Local Orthogonality) $\sim \prod_{j=1}^{n} \|f_j\|_{L^2(B_R)^{\frac{1}{n}}}$.

The proof is based on induction. So we should have a base of induction. In fact, that is rather simple like for $R = 10$. \hfill \square

21
2.4.1 Another Version of Restriction Theory

In last section, we proved a version of multi-linear restriction theorem: Let 

\[ E_{\Sigma_j} \phi(x) = \int_{\Sigma_j} e^{2\pi i \omega \cdot x} \phi(\omega) dVol_{\Sigma_j} \]

where \( \Sigma_j \) is surface where normal vectors has angles less than \( \frac{1}{100n} \) with the \( j \)-th axis. Then

\[
\| \prod_j |E_{\Sigma_j} \phi_j|^{\frac{1}{n}} \|_{L^p(B_R)} \lesssim R \prod_j \| \phi_j \|_{L^2(\Sigma_j)}^{\frac{1}{n}}
\]

Here \( p = \frac{2n}{n-1} \).

Note we also have another version of restriction conjecture which is known for some special case, i.e. Theorem 2.5. Now we want to recover this version from the version we proved.

**Lemma 2.11.** If \( \pi \) is a hyperplane which is perpendicular to \( e_j \), then

\[
\int_\pi |E_j \phi|^2 \sim \int_{\Sigma_j} |\phi|^2
\]

**Proof.** WLOG let \( j = n \). For \( x \in \pi \) we could write it in coordinate \( x = (x_1, \cdots, x_{n-1}, t) \).

We can also write \( \Sigma_n \) as a graph \( w_n = (h(w_1, \cdots, w_{n-1})) \). Then \( dVol_{\Sigma_j} = J dw_1 \cdots dw_{n-1} \). Hence

\[
E\phi(x) = \int_\pi e^{2\pi i (x_1 w_1 + \cdots + x_{n-1} w_{n-1})} e^{2\pi i h(w_1, \cdots, w_{n-1})} \phi(w_1, \cdots, w_{n-1}) J dw_1 \cdots dw_{n-1}
\]

\[
= \hat{g}(x)
\]

If we let \( g(w) = e^{2\pi i h(w_1, \cdots, w_{n-1})} \phi(w_1, \cdots, w_{n-1}) J \)

Then we have

\[
\int_\pi |E\phi|^2 = \int |\hat{g}|^2
\]

(Plancheral) \[
= \int |g|^2 = \int J^2 |\phi|^2 \sim \int |\phi|^2
\]

**Corollary 2.4.** Let \( E \) defined as above, then

\[
\| E\phi \|_{L^2(B_R)} \lesssim R^2 \| \phi \|_{L^2}
\]

**Proof.**

\[
LHS^2 \leq \int_{-R}^R \int_{\mathbb{R}^{n-1}} |E\phi|^2 \sim \int_{-R}^R \| \phi \|_{L^2}^2 \sim R \| \phi \|_{L^2}^2
\]
2.4.2 $n = 2$ Restriction conjecture

Theorem 2.8.

$$\|E\phi\|_{L^4(B_R^2)} \lesssim R^\epsilon \|\phi\|_{L^4(S^1)}$$

Before the proof, let us sketch the idea of the proof. First let us consider $S^1 = \cup r$ which are $K^{-1}$-arcs of $S^1$, and the number of $\tau \sim k$. Then we know for any $x \in B_R^2$, there are two possible cases: either the for those $\tau$ such that $R\phi_\tau$ supported mainly in the ball satisfy the requirement of multilinear restriction, which is good because we know how to prove multilinear restriction; or those $\tau$ such that $R\phi_\tau$ supported mainly in the ball do not satisfy the requirement of multilinear restriction, but if we can control the size of them, we win again. So the idea of the proof is to analyze both cases.

We define the following concept for general dimensions.

Definition 2.1. We say $\tau_1, \cdots, \tau_n$ are transversal if there exists a linear change of coordinate $L$ with $|L| \lesssim K O(1)$ such that $L\tau_j$ obey the hypothesis of multilinear restriction.

In particular on $S^{n-1}$, $\tau_1, \cdots, \tau_n$ are not transversal $\implies \tau_1, \cdots, \tau_n$ lie in the $O(K^{-1})$-neighborhood of an equator.

Sketch of the proof of $n = 2$ case. Define

$$S(x) := \{\tau : |E\phi_\tau(x)| \geq \frac{1}{100K} |E\phi(x)|\}$$

Intuitively $S(x)$ contains those caps contribute most of $E\phi(x)$. In fact we have the following observation:

$$\sum_{\tau \in S(x)} |E\phi_\tau(x)| \leq \frac{1}{10} |E\phi(x)|$$

$$|\sum_{\tau \in S(x)} E\phi_\tau(x)| \sim |E\phi(x)|$$

Let us call $x$ broad if there exists $\tau_1, \tau_2 \in S(x)$ such that $\tau_1, \tau_2$ transversal; we call $x$ narrow is else.

Let us estimate each part.

$$\int_{B_R \cap \text{broad}} |E\phi|^4 \lesssim K O(1) \sum_{\tau_1, \tau_2 \text{trans}} \int_{B_R} |E\phi_{\tau_1}|^2 |E\phi_{\tau_2}|^2$$

Multilinear restriction $\lesssim R^\epsilon K O(1) \sum \|\phi_{\tau_1}\|_{L^2}^2 \|\phi_{\tau_2}\|_{L^2}^2$

$$\lesssim R^\epsilon K O(1) \|\phi\|_{L^2(S^1)}^4$$

So the broad case is good.

For $x$ narrow, let us first analyze the size of $S(x)$. By the definition of transversal, all element of $S(x)$ must lie in the $O(K^{-1})$-neighborhood of a pair
of antipodal points of $S^1$, hence $|S(x)| \lesssim 1$. As a result, by Hölder inequality for $x$ narrow

$$|E\phi(x)|^4 \sim |\sum_{\tau \in S(x)} E\phi_\tau(x)|^4 \lesssim \sum_{\tau \in S(x)} |E\phi_\tau(x)|^4$$

Hence

$$\int_{B_R \cap \text{narrow}} |E\phi|^4 \lesssim \sum_{\tau} \int_{B_R} |E\phi_\tau|^4$$

Let $C(R)$ to be the best constant such that

$$\|E\phi\|_{L^4(B_R)} \leq C(R)\|\phi\|_{L^4(S^1)}$$

We want to show $C(R) \lesssim R^\epsilon$, but first we need to claim that $C(R)$ is well-defined, i.e. it would not be infinity. This only need to note the following estimate:

$$\int_{B_R} |E\phi|^4 \leq |B_R||E\phi||_{L^\infty}^4 \leq |B_R|\|\phi\|^4_{L^4}$$

So $C(R) \lesssim R^{O(1)}$. In order to show $C(R) \lesssim R^\epsilon$, we show the following lemma:

**Lemma 2.12.**

$$\|E\phi_\tau\|_{L^4} \lesssim C\left(\frac{R}{K}\right)\|\phi_\tau\|$$

Then we combine the broad and narrow case to get

$$C(R) \lesssim R^\epsilon K^{O(1)} + C\left(\frac{R}{K}\right)$$

By induction we get $C(R) \lesssim R^\epsilon$, then we finish the proof of $n = 2$ case.

What happens if $n \geq 3$? Now we are unable to control the size of $S(x)$ when $x$ narrow. The broad case is still nice but the narrow case is no longer nice.

### 2.5 Bourgain’s Decoupling Theorem

In this section we discuss the first decoupling theorem we are going to prove in this notes.

The set up is the following: Let $P$ be the truncate parabola $\{w_n = w_1^2 + \cdots + w_{n-1}^2\}$ in $\mathbb{R}^n$ for $|w| \leq 1$. Let $\Omega = N_{1/2}P$, and $\Omega = \sqcup \theta$ where $\theta$’s are $R^{-1/2}$ caps.

As before, we define $D_{p,n}(R)$ to be the decoupling constant for the above decoupling, i.e.

$$\|f\|_{L^p} \leq D_{p,n}(R)\left(\sum_{\theta} \|f_{\theta}\|^2_{L^{2p}}\right)^{\frac{1}{2}}$$
Theorem 2.9 (Bourgain). If \( 2 \leq p \leq \frac{2n}{n-1} \), then
\[
D_{p,n}(R) \leq C(n,\epsilon)R^\epsilon
\]
Later, Bourgain-Guth generalize the above result to the case \( 2 \leq p \leq \frac{2(n+1)}{n-1} \).

2.5.1 Multiscale Tools

Before we prove the theorem, let us introduce some tools which will be used in the proof. These tools are mainly used to do analysis between different scales.

**Lemma 2.13.** If \( L \) is an isomorphism of \( \mathbb{R}^n \), then
\[
D_{p}(\Omega = \sqcup \theta) = D_{p}(L\Omega = \sqcup L\theta)
\]

*Sketch of the proof.* Suppose \( \text{supp} \hat{f} \subset \Omega \) and \( f = \sum f_\theta \). Let \( \hat{f}(x) = f((L^*)^{-1}x) \) we have \( \text{supp} \hat{f} \subset L\Omega \), and \( \text{supp} \hat{f}_\theta \subset L\theta \). We can compute that all the integrals will change by a Jacobian defined of \( L \), and since this appear in both the integral of \( \hat{f} \) and \( \hat{f}_\theta \), both side’s Jacobian cancel and we are done. \( \square \)

This lemma help us compare the decoupling of the whole parabola and a part of it. In fact, assume \( \theta \) are \( R^{-1/2} \) caps and \( \tau \) are \( R_1^{-1/2} \), \( = R_1R_2 \). Then we can always do a linear change of variable \( L \) such that
\[
L\tau = P, L\theta = R_2^{-1/2} \text{ caps}
\]

In other word, we can use the lemma to compare the decoupling constant in different scales. A direct corollary is
\[
D_{p}(\tau = \theta) = D_{p,n}(R_2)
\]
in the above setting after the lemma.

**Lemma 2.14.** If \( R = R_1R_2 \), then
\[
D_{p,n}(R) \lesssim D_{p,n}(R_1)D_{p,n}(R_2)
\]

*Sketch of the proof.* Consider \( \tau s \) are \( R_3^{-1/2} \) caps and \( \theta s \) are \( R^{-1/2} \) caps. Then
\[
\|f\|_{L^p} \lesssim D_{p}(R_1)(\sum_{\tau} \|f_\tau\|_{L^p}^2)^{1/2}
\]
Also
\[
\|f_\tau\|_{L^p} \lesssim D_{p}(R_2)(\sum_{\theta \subset \tau} \|f_\theta\|_{L^p}^2)^{1/2}
\]
Combining together we have
\[
\|f\|_{L^p} \lesssim D_{p}(R_1)D_{p}(R_2)(\sum_{\theta} \|f_\theta\|_{L^p}^2)^{1/2}
\]

\( \square \)
With these multiscale tools we now give a heuristic wrong proof here.

**Wrong proof.** We want to show $D_{p,n}(R) \lesssim R^\epsilon$, then we use induction. By previous lemma, we have

$$D_{p,n}(R) \lesssim (D_{p,n}(R^{1/2}))^2$$

By induction, we get

$$((D_{p,n}(R^{1/2}))^2 \lesssim (R^{\epsilon/2})^2 = R^\epsilon$$

Then we finish the proof. □

The main issue for this proof is that we do not have a base case. Moreover, since we have examples we know $2 \leq p \leq \frac{2n}{n-1}$ is necessary, however in this wrong proof this bound is not used.

### 2.5.2 Multilinear Decoupling

In previous sections we develop the theory of multilinear Kakeya problem and multilinear restriction theory. Here we can also develop the multilinear theory for decoupling.

Let us consider $P$ is the paraboloid, suppose $P_j \subset P$ for $j = 1, \cdots, n$ are transversal. Consider $\Omega_j = \mathcal{N}_j \setminus P_j$ and assume $\Omega_j = \cup \theta$ are union of $R^{-1/2}$ caps.

We define the multilinear decoupling constant $MD_{p,n}(R)$ to be the best constant such that

$$\| \prod_{j=1}^n (f_j)^{1/n} \|_{L^p} \leq MD_{p,n}(R) \prod_{j=1}^n (\sum_{\theta \subset \Omega_j} \| f_{j,\theta} \|_{L^p}^2)^{1/2}$$

for all $f_j = \sum f_{j,\theta}$ such that the supp $f_{j,\theta} \subset \theta \subset \Omega_j$.

We first observe that $MD_{p,n}(R) \leq D_{p,n}(R)$.

**Proof.**

$$\| \prod_j |f_j|^{1/n} \|_{L^p} \leq \prod_j \|f_j\|_{L^p} \leq D_{p,n}(R) \prod_j \left( \sum_{\theta} \| f_{j,\theta} \|_{L^p}^2 \right)^{1/2}$$

Let us first show $MD \lesssim R^\epsilon$. This is the requirement for $D \lesssim R^\epsilon$. 

26
Proof.

\[ \left\| \prod_j |f_j|^{\frac{1}{n}} \right\|_{L^p_{avg}(B_R)} \leq \left\| \prod_j |f_j|^{\frac{1}{n}} \right\|_{L^2_{avg}(B_R)} \]

(by multilinear restriction) \[ \lesssim R^n \prod_j \left\| f_j \right\|_{L^2_{avg}(B_R)}^{\frac{1}{n}} \]

(by locally constancy) \[ R^n \left( \sum_\theta \|f_{j,\theta}\|_{L^2_{avg}(B_R)}^2 \right)^{\frac{1}{2}} \]

(by H"older) \[ \leq R^n \prod_j \left( \sum_\theta \|f_{j,\theta}\|_{L^p_{avg}(B_R)}^2 \right)^{\frac{1}{2}} \]

2.5.3 Main Induction Lemma

Lemma 2.15 (Main Lemma). For any $K \geq 1$, we have

\[ D_{p,n}(R) \lesssim K^{O(1)} MD_{p,n}(R) + D_{p,n-1}(K^2)D_{p,n}(\frac{R}{K^2}) \]

Once we have the main lemma, we can easily show the decoupling theorem by induction on dimensions. We will choose $K = \log R$ or $K = K(\epsilon)$ for some large constant $K$.

RHS of main lemma consists two parts: the first term is based on the analysis of broad set, and the second term is based on the analysis of narrow set, just like what we did for restriction theory in $n=2$ case.

Proof of main lemma. Recall the set up: Let $P$ be the truncate paraboloid in $\mathbb{R}^n$, $\Omega = \cup \tau$ be the $K^{-1}$-caps, and $f = \sum_\tau f_\tau$ where each $f_\tau$ has Fourier transformation supported on $\tau$. We know $|f_\tau|$ ~ constant on $K \times K^2$ tubes.

Now we tile $\mathbb{R}^n$ with balls $B(x_0, K^2) = B$. Define

\[ S(B) := \{ \tau : \|f_\tau\|_{L^p(B)} \geq \frac{1}{100^\#\tau} \|f\|_{L^p(B)} \} \]

Just like the proof of $n=2$ restriction conjecture before, we have the observation

\[ \| \sum_{\tau \in S(B)} f_\tau \|_{L^p(B)} \sim \|f\|_{L^p(B)} \]

Again like before, we define $B$ is broad if there exists $\tau_1, \cdots, \tau_n \in S(B)$ transversal; otherwise we define $B$ is narrow. Recall that $B$ narrow means there exists a hyperplane $\pi^*$ such that for any $\tau \in S(B)$ the angle between $\tau, \pi^*$ \lesssim K^{-1}. We define

\[ Broad = \cup_B broadB, Narrow = \cup_B narrowB \]

Now we break the proof into two pieces:
• Broad estimate:

\[ \|f\|_{L^p(Broad)} \lesssim K^{o(1)} M D_{p,n}(R) \left( \sum_{\theta} \|f_\theta\|_{L^p(R^n)}^2 \right)^{\frac{1}{2}} \]

• Narrow estimate:

\[ \|f\|_{L^p(Narrow)} \lesssim D_{p,n-1}(K^2) \left( \sum_{\tau} \|f_\tau\|_{L^p(R^n)}^2 \right)^{\frac{1}{2}} \]

Note by the multiscale tools we have

\[ \left( \sum_{\tau} \|f_\tau\|_{L^p(R^n)}^2 \right)^{\frac{1}{2}} \leq D_{p,n} \left( \frac{R}{K^2} \right) \left( \sum_{\theta} \|f_\theta\|_{L^p(R^n)}^2 \right)^{\frac{1}{2}} \]

Hence combine the broad estimate and the narrow estimate together we will get the main lemma.

\[
\square
\]