Lecture 21 : Tools for proving decoupling inequalities

In the previous lectures we saw some of the basic features of decoupling inequalities for the moment curve, namely

- Self-similarity / rescaling.

We also saw how these scaling properties can be effectively combined in

- Induction-on-scale methods,

particularly in the context of the Pramanik-Seeger argument.

An underlying philosophy behind induction-on-scale is that once we find a small scale at which the inequality behaves well or there is some kind of gain, we can propagate this gain throughout all scales via the induction procedure.

Take the Pramanik-Seeger argument of the last lecture:

- On an interval I of length $\sim \delta^{k/k+1}$ we can prove the decoupling for γ by direct comparison with $\gamma_{a,h}$ $(I = [a-\frac{h}{2}, a+h/2])$

- This gives us "good behaviour" at the $\delta^{k/k+1}$ scale; we then ramp this up to global good behaviour using the induction.

The Bourgain-Demeter-Guth decoupling theorem will be proved via an induction-on-scale using the same philosophy. In particular, the goal will be to isolate a "good scale"

The key tool here is L^2-orthogonality:-

$$\Big\| \sum_{j=1}^{K} f_j \Big\|_{L^2(\mathbb{R}^n)}^2 = \sum_{j=1}^{K} \|f_j\|_{L^2(\mathbb{R}^n)}^2 \qquad (1)$$

if $\operatorname{supp} \hat{f}_j$ are disjoint.

This will provide the desired gain at suitable scales. Of course, since we're interested in L^p-based estimates (1) cannot be applied directly to the problem.

In order to make use of L^2-orthogonality it will be useful to work with bilinear decoupling inequalities.

The purpose of this lecture is to describe these two additional tools,

- L^2-orthogonality.
- Bilinear estimates

In detail and indicate how they can be combined to study L^p-based decoupling.

L^2-orthogonality: Take 1.

For $p = 2$, the Bourgain-Demeter-Guth theorem is a consequence of Plancherel's theorem:-

$$\Big\| \sum_{J \in \mathcal{P}(\delta)} f_J \Big\|_{L^2(\mathbb{R}^k)} \lesssim \Big(\sum_{J \in \mathcal{P}(\delta)} \|f_J\|_{L^2(\mathbb{R}^k)}^2 \Big)^{1/2} \qquad (2)$$

whenever $(f_J)_{J \in \mathcal{P}(\delta)}$ is a tuple of functions satisfying

$$\operatorname{supp} \hat{f}_J \subseteq \Theta_J .$$

Indeed, this follows simply as the Θ_J are finitely-overlapping. (ie essentially disjoint).

It will be helpful to recall the local version of the L^2 orthogonality inequality :-

Lemma (Local L^2-orthogonality). For $R \geq 1$, the inequality

$$\| \sum_{j=1}^{K} f_j \|_{L^2(B_R)} \lesssim \left(\sum_{j=1}^{K} \| f_j \|_{L^2(w_{B_R})}^2 \right)^{1/2} \quad (3)$$

holds whenever

$\mathcal{N}_{1/R}(\mathrm{supp}\, \hat{f}_j)$ are finitely-overlapping.

Here B_R is an arbitrary ball of radius R and w_{B_R} is a rapidly decreasing weight;

$$w_{B_R}(x) := (1 + R^{-1}|x - \bar{x}|)^{-N}$$

for $\bar{x} \in \mathbb{R}^n$ the centre of B_R and $N \in \mathbb{N}$ a large integer. (We can choose N as large as we please, however the implied constant in (3) depends on N).

We can use the local L^2-orthogonality lemma to prove a local version of (2), for instance:-

$$\| \sum_{J \in \mathcal{P}(\delta)} f_J \|_{L^2(B_{\delta^{-1}})} \lesssim \left(\sum_{J \in \mathcal{P}(\delta)} \| f_J \|_{L^2(w_{B_{\delta^{-1}}})}^2 \right)^{1/2}$$

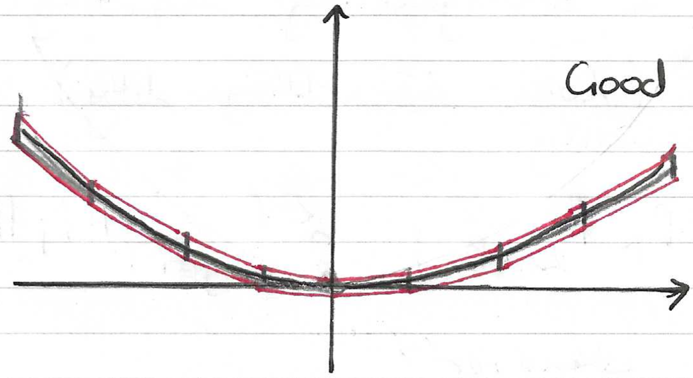
whenever $(f_J)_{J \in \mathcal{P}(\delta)}$ satisfies $\mathrm{supp}\, \hat{f}_J \subseteq \Theta_J$.

It is important to note the

$\mathcal{N}_\delta\, \Theta_J$ are finitely-overlapping

where the scale δ corresponds to the largest sidelength of the θ_J.

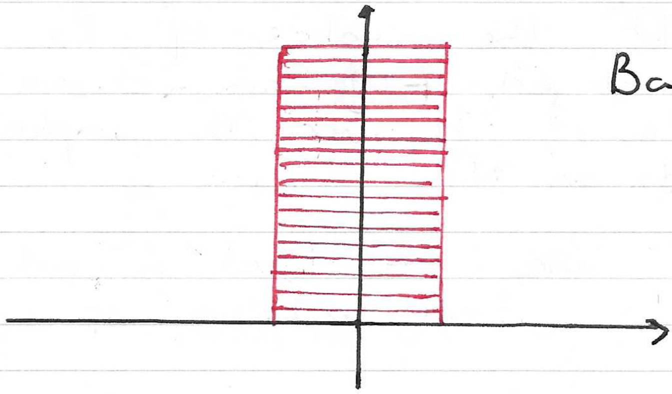
This is because the θ_J are arranged "tangentially" along a curve.



If boxes of the same dimensions were arranged arbitrarily, then the best we could say would be

$$\mathcal{N}_{\delta^k}\theta_J \text{ are finitely-overlapping,}$$

which leads to much weaker local estimates.



This will play an important rôle in what follows.

L^2-orthogonality : Take 2

The key difficulty is to understand how to utilise L^2-orthogonality in the setting of L^p-decoupling for $p > 2$.

Let's focus on the $k = 2$ case for the purpose of this discussion so that the critical exponent is $p_2 = 2 \cdot 3 = 6$.

For want of a better idea, we take an L^6 norm and try to write it in terms of L^2 so as to be able to apply L^2 orthogonality:

$$\int_{\mathbb{R}^2} |f|^6 = \int_{\mathbb{R}^2} |f|^2 \cdot |f|^4 . \qquad (4)$$

Of course, the $|f|^4$ prevents us from going any further here. As an experiment, however, let's consider "bilinearising" the set up. In particular, consider

$$\int_{\mathbb{R}^2} |f|^2 |g|^4 \qquad (5)$$

which corresponds to (4) when $f = g$. If the functions f and g are sufficiently "independent" of one another, then we could hope

$$\int_{\mathbb{R}^2} |f|^2 |g|^4 \sim \left(\int_{\mathbb{R}^2} |f|^2\right) \cdot \left(\int_{\mathbb{R}^2} |g|^4\right)$$

which would allow us to apply L^2 orthogonality to f.

Of course, for the case $f = g$ of interest there is no chance such independence will hold (this is the antithesis of an independent situation!). However, we shall be able, in practice, to run a <u>bilinear reduction</u>, similar to that used to study Bochner-Riesz multipliers in earlier lectures, to reduce considerations to expressions of the form (5) for f and g "independent".

In particular, there is a correspondence

transversality in bilinear reduction $\rightsquigarrow$ "independence" between f and g.

The above discussion is somewhat vague and so to conclude this lecture we will describe concrete situations which lead to "independent" f and g.

In the context of decoupling, hypotheses on the input functions are always framed in terms of Fourier support conditions. We will therefore frame our hypotheses on f and g in a similar manner.

• First f: let's suppose $f = \sum_{j=1}^{K} f_j$

where
$$\mathcal{N}_{1/R}(\operatorname{supp} \hat{f}_j) \quad \text{are finitely overlapping.}$$

Thus, local L^2 orthogonality holds at spatial scale R and it makes sense to decompose (5) by writing

$$\int_{\mathbb{R}^2} |f|^2 |g|^4 \approx \sum_{B_R \in \mathcal{B}_R} \int_{B_R} |f|^2 |g|^4$$

where $\mathcal{B}_R$ is a finitely-overlapping collection of R-balls which cover $\mathbb{R}^2$.

• Now we deal with g. What would be really great would be if g is constant on each B_R. In this case, we could write

$$\sum_{B_R \in \mathcal{B}_R} \int_{B_R} |f|^2 |g|^4$$
$$= \sum_{B_R \in \mathcal{B}_R} \Big(\int_{B_R} |f|^2\Big) \cdot |g(x_{B_R})|^4$$
$$\lesssim \sum_{B_R \in \mathcal{B}_R} \sum_{j=1}^{K} \Big(\int_{B_R} |f_j|^2\Big) \cdot |g(x_{B_R})|^4$$
$$= \sum_{B_R \in \mathcal{B}_R} \int_{B_R} \Big(\sum_{j=1}^{K} |f_j|^2\Big) \cdot |g|^4$$

by local L^2-orthogonality. Here each x_{B_R} is an arbitrary choice of point in B_R. We've cheated a little by replacing the weight function w_{B_R} in the local L^2 lemma with a sharp cut off χ_{B_R}, but this is just a technicality.

Thus, in the end we can conclude

$$\int_{\mathbb{R}^2} |f|^2 |g|^4 \lesssim \int_{\mathbb{R}^2} \Big(\sum_{j=1}^{K} |f_j|^2\Big) \cdot |g|^4$$

which successfully applies the L^2 orthogonality to f, provided g is constant on the B_R.

However, we can frame this property in terms of Fourier localisation, using the uncertainty principle. In particular, if

$$\operatorname{supp} \hat{g} \subseteq B_{1/R} \quad \text{for some } 1/R\text{-ball in } \hat{\mathbb{R}}^2,$$

then the uncertainty principle dictates $|g|$ is "essentially constant" on R-balls B_R.

Thus, we have the following heuristic lemma:-

Lemma (Heuristic) Suppose $f, g \in L^1(\mathbb{R}^2)$ satisfy

- $f = \sum_{j=1}^{K} f_j$ where $f_j \in L^1(\mathbb{R}^2)$ with

  $\mathcal{N}_{1/R}(\operatorname{supp} \hat{f}_j)$ finitely overlapping

- $\operatorname{supp} \hat{g} \subseteq B_{1/R}$, some $1/R$-ball in $\hat{\mathbb{R}}^2$

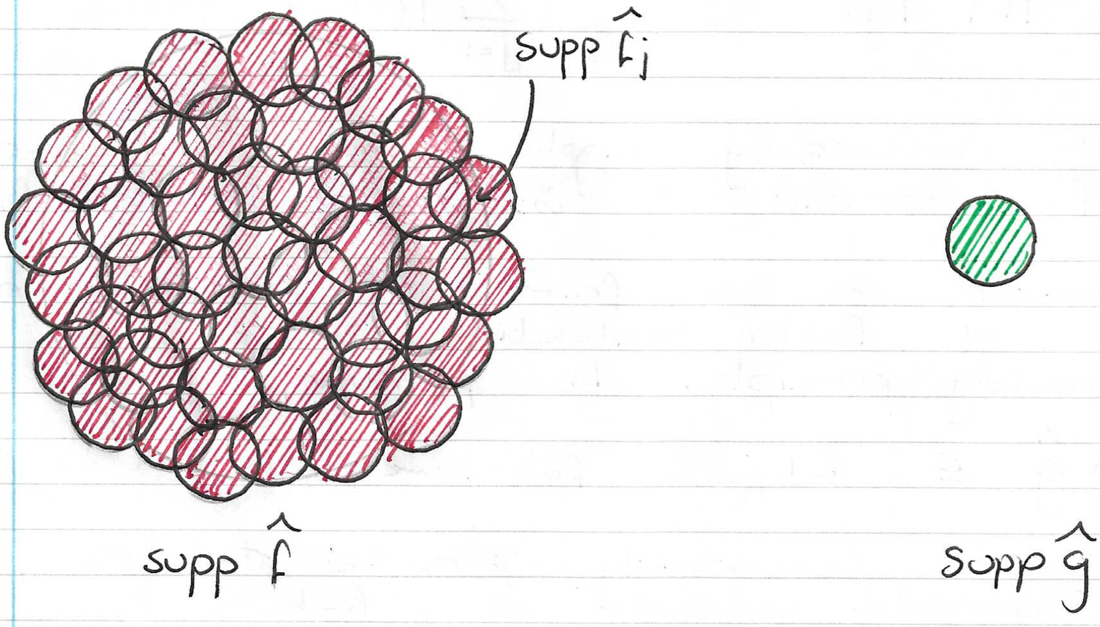
Then

$$\int_{\mathbb{R}^2} |f|^2 |g|^4 \lesssim \int_{\mathbb{R}^2} \Big(\sum_{j=1}^{K} |f_j|^2\Big) \cdot |g|^4$$

The arguments above can be made precise

using standard rigorous implementations of the uncertainty principle, working with weight functions etc.

Note that the lemma requires the Fourier support of f to be much larger than g (at least in interesting cases where $\overline{K}$ is large).



This can be thought of as analogous to the condition $h \sim \delta^{k/k+1}$ in the Pramanik–Seeger argument – we need a favourable choice of scales for the argument to work.

As before, induction-on-scale arguments will be used to propagate this 'gain' to all scales, but the setup is substantially more complicated owing to the need to bilinearise.