Lecture 20 : Decoupling for the moment curve I

Scaling considerations.

Throughout this lecture we'll work with the moment curve

$$\gamma_0(t) := \left(t, \frac{t^2}{2}, \dots, \frac{t^k}{k!}\right).$$

This is a slightly different definition from the previous lecture; the choice of the $1/j!$ coefficients here is merely a matter of convenience.

A key observation is that γ_0 admits 'affine self-symmetry' - similar to the parabolic rescaling for the parabola (which corresponds, of course, to the $k=2$ case of γ_0) and the Lorentz rescaling of the light cone.

Let $J \subseteq \mathbb{R}$ be a compact subinterval of $[0,1]$ and consider the portion of the curve over J :-

$$\gamma_0(J) = \{ \gamma_0(t) : t \in J \}.$$

We can find an affine transformation of the ambient space $\mathbb{R}^k$ which diffeomorphically maps the portion $\gamma_0(J)$ to the "whole" curve $\gamma_0([0,1])$.

Write $J = [a, a+h]$ so every $s \in J$ can be represented by $s = a + th$ for some $t \in [0,1]$. Taylor expanding

$$\gamma_0(a+th) = \gamma_0(a) + \sum_{j=1}^{k} \gamma_0^{(j)}(a) \frac{(ht)^j}{j!} \qquad (1)$$

Now, $\frac{(ht)^j}{j!}$ is the jth component of

$$\gamma(ht) = D_h \gamma(t) \quad \text{where}$$

$$D_h = \begin{bmatrix} h & & & \\ & h^2 & & \\ & & \ddots & \\ & & & h^k \end{bmatrix} \in Mat(k, \mathbb{R})$$

so that (1) can be expressed succinctly as

$$\gamma_0(a + th) = \gamma_0(a) + M_a D_h \gamma(t) \qquad (2)$$

where

$$M_a = \begin{bmatrix} | & & | \\ \gamma_0^{(1)}(a) & \cdots & \gamma_0^{(k)}(a) \\ | & & | \end{bmatrix} \in Mat(k, \mathbb{R})$$

is the matrix whose jth column is $\gamma^{(j)}(a)$. It is not difficult to see

$$\det M_a = 1. \qquad (3)$$

The formula (2) realises $\gamma_0(J)$ as an affine image of $\gamma_0([0,1])$. Indeed, from (2) we see

$$\begin{aligned} \gamma_0(J) &= \{ \gamma_0(a + th) : t \in [0,1] \} \\ &= \gamma_0(a) + M_a D_h \gamma_0([0,1]) \\ &= \mathcal{A}_{a,h} \gamma_0([0,1]) \end{aligned}$$

where $\mathcal{A}_{a,h} : \xi \mapsto \gamma_0(a) + M_a D_h \xi$. Thus, $\gamma_0(J)$ can be mapped back to the whole curve via the inverse transformation

$$\mathcal{A}_{a,h}^{-1} \gamma_0(J) = \gamma_0([0,1]).$$

This scaling is often referred to as 'generalised parabolic rescaling' since it generalises the scaling structure of the parabola used in our investigations of Bochner-Riesz multipliers.

Decoupling constants :-

We recall the decoupling setup from the previous lecture, in a slightly more general context.

Def^{n} :- A curve $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ has non-vanishing torsion if the matrix

$$M_{\gamma,a} := \begin{bmatrix} | & & | \\ \gamma^{(1)}(a) & \cdots & \gamma^{(n)}(a) \\ | & & | \end{bmatrix}$$

has non-vanishing determinant for all $a \in [0,1]$.

By (3) we see the moment curve is an example of a curve of non-vanishing torsion.

For any such curve one may define the Frenet frame and associated curvature functions $\kappa_1, \dots, \kappa_{n-1}$. The non-vanishing torsion condition guarantees all these curvature functions are bounded away from 0 (this gives a parameterisation-invariant characterisation of the non-vanishing torsion condition).

- For $0 < \delta < 1$ let $\mathcal{P}(\delta)$ denote a partition of $[0,1]$ into intervals of length $\sim \delta$.
- Given $J \in \mathcal{P}(\delta)$ with centre c_J let $\Theta_J = \Theta_{\gamma,J}$ denote the parallelepiped with
  - centre $\gamma(c_J)$
  - dimensions $\delta \times \delta^2 \times \dots \times \delta^n$
  - sides parallel to $\gamma'(c_J), \dots, \gamma^{(n)}(c_J)$.

The non-vanishing torsion condition is precisely what is needed to ensure the $\Theta_{\gamma,J}$ are non-degenerate parallelepipeds.

Def^{n}:- For $0 < \delta < 1$, the $\ell^2 L^p$ decoupling constant $D_{k,p,\gamma}(\delta)$ of a curve $\gamma : [0,1] \to \mathbb{R}^k$ of non-vanishing torsion is the infimum over all $C \geq 1$ for which the inequality

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

holds for all tuples $(f_J)_{J \in \mathcal{P}(\delta)}$ of Schwartz functions satisfying

$$\operatorname{supp} \hat{f}_J \subseteq \Theta_{\gamma, J}.$$

Remark:- • In the case of the moment curve γ_0 we will often drop the subscript γ_0 and write $D_{k,p,\gamma_0}(\delta) = D_{k,p}(\delta)$.

• Similarly, in the case of the critical exponent

$$p = p_k = k(k+1)$$

we will often drop the subscript p_k.

With the above notation, the Bourgain–Demeter–Guth decoupling theorem can be stated succinctly as follows:-

Theorem (Bourgain–Demeter–Guth):- For all $k \in \mathbb{N}$, $\varepsilon, \delta > 0$ we have

$$D_{k,p}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon} \quad \text{for } 2 \leq p \leq p_k.$$

Rescaling and decoupling inequalities:-

We will work with many different scales in order to prove decoupling estimates.

Def^{n}:- Given $0 < \delta < \eta < 1$ and $I \in \mathcal{P}(\eta)$, let

$$\mathcal{P}(I, \delta)$$

denote a partition of I into intervals of length $\sim J$.

Lemma (Rescaling) Let $0 < \delta < \eta < 1$ and $I \in \mathcal{D}(\eta)$. If $(f_J)_{J \in \mathcal{D}(I,\delta)}$ is a tuple of functions satisfying

$$\operatorname{supp} \hat{f}_J \subseteq \Theta_J \qquad \text{for all} \quad J \in \mathcal{D}(I,\delta),$$

then, for all $2 \leq p < \infty$, we have

$$\Big\| \sum_{J \in \mathcal{D}(I,\delta)} f_J \Big\|_{L^p(\mathbb{R}^n)} \lesssim D_{k,p}(\delta/\eta) \Big(\sum_{J \in \mathcal{D}(I,\delta)} \|f_J\|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2}$$

Before proving the lemma, we note a consequence.

Corollary (Self-similarity) If $0 < \delta < \eta < 1$ and $2 \leq p < \infty$, then

$$D_{k,p}(\delta) \lesssim D_{k,p}(\eta) \cdot D_{k,p}(\delta/\eta).$$

Proof:- Take $(f_J)_{J \in \mathcal{D}(\delta)}$ with $\operatorname{supp} \hat{f}_J \subseteq \Theta_J$ and write

$$f_I := \sum_{\substack{J \in \mathcal{D}(\delta) \\ J \cap I \neq \emptyset}} f_J \qquad \text{for all} \quad I \in \mathcal{D}(\eta)$$

By definition,

$$\Big\| \sum_{J \in \mathcal{D}(\delta)} f_J \Big\|_{L^p(\mathbb{R}^n)} \lesssim \Big\| \sum_{I \in \mathcal{D}(\eta)} f_I \Big\|_{L^p(\mathbb{R}^n)}$$

$$\leq D_{k,p}(\eta) \Big(\sum_{I \in \mathcal{D}(\eta)} \| f_I \|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2} \quad (4)$$

Applying the rescaling lemma to each term in the sum,

$$\| f_I \|_{L^p(\mathbb{R}^n)} \lesssim D_{k,p}(\delta/\eta) \Big(\sum_{\substack{J \in \mathcal{D}(\delta) \\ J \cap I \neq \emptyset}} \| f_J \|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2} \quad (5)$$

Putting (4) and (5) together

$$\Big\| \sum_{J \in \mathcal{D}(\delta)} f_J \Big\|_{L^p(\mathbb{R}^n)} \lesssim D_{k,p}(\eta) \cdot D_{k,p}(\delta/\eta) \quad \blacksquare$$

$$\left[\sum_{I \in \mathcal{P}(\eta)} \left(\sum_{\substack{J \in \mathcal{P}(\delta) \\ J \cap I \neq \emptyset}} \|f_J\|^2_{L^p(\mathbb{R}^n)} \right)^{p/2} \right]^{1/2}$$

$$\lesssim D_{n,p}(\eta) \cdot D_{n,p}(\delta/\eta) \cdot \left(\sum_{J \in \mathcal{P}(\delta)} \|f_J\|^2_{L^p(\mathbb{R}^n)} \right)^{1/2}$$

So the corollary follows from the definition of $D_{n,p}(\delta)$. □

<u>Remark:-</u> The self-similarity property should be compared with the (moral) bound

$$S(r_1, r_3) \leq S(r_1, r_2)\, S(r_2, r_3)$$

for the square function constant, which we proved using Lorentz rescaling in earlier lectures.

- Whilst we had to spend considerable effort in the square function problem to obtain a 'self-similar' setup, in the decoupling case we get this property for free!
- Furthermore, the decoupling constants $D_{n,p}(\delta)$ are, in some ways, easier to work with than the $S(r_1, r_2)$ since they involve only a single 'essential' parameter δ.

<u>Remark:-</u> We can use the self-similarity to give a 'false proof' of the decoupling theorem via induction. In particular, it is trivial to show

$$D_{k,p}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon} \quad \text{if} \quad \delta \sim 1$$

which serves as our base case. Assume:

<u>Induction hypothesis</u>: For all $2\delta' < \delta < 1$

$$D_{n,p}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}.$$

Now take $\delta' < \delta \leq 2\delta'$ so that, by self-similarity,

$$D_{n,p}(\delta) \lesssim D_{n,p}(1/2) \cdot D_{n,p}(2\delta)$$
$$\lesssim \delta^{-\varepsilon}$$

where $D_{n,p}(1/2) \lesssim 1$ by the base case
and
$D_{k,p}(2\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$ by the induction
hypothesis.

Of course, things aren't this easy and the above proof is not correct: again, we need to control the constants in our induction hypothesis!

Nevertheless, there is the kernel of a good idea here, and induction-on-scale techniques will be very useful in what follows.

Proof (of rescaling lemma):- Let $(f_J)_{J \in \mathcal{P}(I,\delta)}$ satisfy the hypotheses of the lemma. If $I = [a, a+h]$, then recall

$$\mathcal{A}_{a,h}^{-1}\, \gamma_0(I) = \gamma_0([0,1]).$$

The key observation is that for all $J \in \mathcal{P}(I,\delta)$

$$(\mathcal{A}_{a,h}^{-1}\, \Theta_J)_{J \in \mathcal{P}(I,\delta)} = (\Theta_{\tilde{J}})_{\tilde{J} \in \mathcal{P}(\delta/\eta)} \qquad (6)$$

corresponds to a decomposition of a neighbourhood of the whole curve at the larger scale δ/η

Indeed, once we have this the lemma follows by noting the decoupling inequality is invariant under an affine transformation of the frequency space. We give the details below.

To see (6), note that $\xi \in \Theta_J$ can be written as

$$\xi = \gamma_0(c_J) + \sum_{j=1}^{k} r_j\, \gamma_0^{(j)}(c_J)$$

for $|r_j| \leq \delta^j$, $j = 1, \dots, k$.

If $I = [a, a+\eta]'$, then since $J \in \mathcal{D}(I, \delta)$ we can write

$$c_J = a + s_J \eta \quad \text{where the}$$

s_J form a δ/η-net on $[0,1]$.
Thus,

$$\xi = \gamma(a + s_J\eta) + \sum_{j=1}^{k} r_j \gamma^{(j)}(a + s_J \eta).$$

We now study the action of $A_{a,\eta} := M_a D_\eta$, the linear part of $\mathcal{A}_{a,\eta}$, on the derivatives of $\gamma^{(j)}$. In particular, by Taylor

$$\gamma^{(j)}(a + s_J\eta) = \sum_{i=j}^{k} \gamma^{(i)}(a) \frac{(s_J\eta)^{i-j}}{(i-j)!}$$

$$= \eta^{-j} A_{a,\eta} \gamma^{(j)}(s_J).$$

Thus,
$$\xi = \gamma(a + s_J\eta) + \sum_{j=1}^{k} e_j A_{a,\eta} \gamma^{(i)}(s_J)$$
where $e_j := r_j/\eta^j$ satisfies $|e_j| \leq (\delta/\eta)^j$, $1 \leq j \leq k$.
Applying $\mathcal{A}_{a,\eta}^{-1}$, we obtain

$$\mathcal{A}_{a,\eta}^{-1}\xi = \mathcal{A}_{a,\eta}^{-1}\gamma(a + s_J\eta) + \sum_{j=1}^{k} e_j A_{a,\eta}^{-1} A_{a,\eta} \gamma^{(j)}(s_J)$$

$$= \gamma(s_J) + \sum_{j=1}^{k} e_j \gamma^{(j)}(s_J).$$

Thus, (6) follows by taking

$$\Theta_{\tilde{J}} = \left\{ \gamma(s_J) + \sum_{j=1}^{k} e_j \gamma^{(j)}(s_J) : |e_j| \leq (\delta/\eta)^j \right\}$$
$\square$

Finally, let

$$\hat{g}_{\tilde{J}} := |\det \mathcal{A}_{a,h}| \hat{f}_J \circ \mathcal{A}_{a,h}$$

for all $J \in \mathcal{D}(I, \delta)$ so that

$$g_{\tilde{J}}(x) = e^{2\pi i \langle x, \gamma(a) \rangle} f_J \circ \mathcal{A}_{a,h}^{-1}$$

and $\quad \operatorname{supp} \hat{g}_{\tilde{J}} \subseteq \Theta_{\tilde{J}}$.
By a change of variable,

$$\Big\|\sum_{J\in\mathcal{P}(I,\delta)} f_J\Big\|_{L^p(\mathbb{R}^n)} = |\det A_{a,h}|^{-\frac{1}{p}} \Big\|\sum_{\tilde{J}\in\mathcal{P}(\delta/\eta)} g_{\tilde{J}}\Big\|_{L^p(\mathbb{R}^n)}$$

$$\le D_{k,p}(\delta/\eta)\, |\det A_{a,h}|^{-1/p}\Big(\sum_{\tilde{J}\in\mathcal{P}(\delta/\eta)} \|g_{\tilde{J}}\|^2_{L^p(\mathbb{R}^n)}\Big)^{1/2}$$

$$= D_{n,p}(\delta/\eta)\Big(\sum_{J\in\mathcal{P}(I,\delta)} \|f_J\|^2_{L^p(\mathbb{R}^n)}\Big)^{1/2},$$

as required. $\square$.

Stability of decoupling estimates: the Pramanik-Seeger argument.

Here we illustrate a 'stability' property of decoupling inequalities which is a consequence of their self-similar structure. In particular, we will show decoupling for γ_0 implies decoupling for any curve given by a small perturbation of γ_0. This in turn allows us to prove decoupling for all curves of non-vanishing torsion.

Lemma:- If $D_{k,p}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$ for all $\delta\in(0,1)$, and $\gamma:[0,1]\to\mathbb{R}^k$ is a curve of non-vanishing torsion, then

$$D_{n,p,\gamma}(\delta) \lesssim_{\varepsilon,\gamma} \delta^{-\varepsilon} \quad \text{for all } \delta\in(0,1).$$

More precisely, the constant in the above display depends only on ε, k and

$$\inf_{a\in[0,1]} |\det M_{\gamma,a}|, \qquad \max_{1\le j\le n+1} \|\gamma^{(i)}\|_{L^\infty([0,1])}.$$

Geometric preliminaries: Before giving the proof of the above stability lemma we note some basic properties of curves $\gamma:[0,1]\to\mathbb{R}^n$ of non-vanishing torsion.

If $I = [a, a+h]$ for $0 < h < 1$ small, then one may find an affine transformation of $\mathbb{R}^k$, say $\mathcal{A}_{\gamma,a,h}$, such that

$\mathcal{A}_{\gamma,a,h}^{-1}\,\gamma(I)$ is a small perturbation of $\gamma_0([0,1])$.

More precisely, by Taylor, (7)

$$\gamma(a+th) = \gamma(a) + M_{\gamma,a} D_h \gamma_0(t) + \text{higher order}$$

so we take $\mathcal{A}_{\gamma,a,h} : \xi \longmapsto \gamma(a) + M_{\gamma,a} D_h \xi$.

With this definition,

$$\mathcal{A}_{\gamma,a,h}^{-1}\,\gamma(a+th) = \gamma_0(t) + \mathcal{E}_{\gamma,a,h}(t) \qquad (8)$$

where

$$|\mathcal{E}_{\gamma,a,h}(t)| \lesssim_\gamma h \cdot |t|^{k+1}.$$

By decomposing $[0,1]$ into small intervals and applying affine transformations, using the fact that the decoupling inequalities are invariant under affine transformation of the frequency space as in the proof of the rescaling lemma, we may assume any curve of non-vanishing torsion is of the form

$$\gamma(t) = \gamma_0(t) + \mathcal{E}(t)$$

where γ_0 is the moment curve and $\mathcal{E}$ is order $k+1$:

$$|\mathcal{E}(t)| \leq c|t|^{k+1}$$

Furthermore, we can choose c to be any prescribed value (independent of δ).

To prove the stability lemma we use an inductive procedure essentially due to Pramanik - Seeger.

<u>Proof</u>: We use induction-on-scale. Fix $\varepsilon > 0$. As always, the estimate

$$D_{n,p,\gamma}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$$

trivially holds for $\delta \sim_\varepsilon 1$, which serves as the base case.

<u>Induction hypothesis</u>: If $2\delta' \leq \delta < 1$, then

$$D_{n,p,\gamma}(\delta) \leq \bar{C}_{\gamma,\varepsilon}\, \delta^{-\varepsilon}$$

Here $\bar{C}_{\gamma,\varepsilon}$ is a fixed constant, chosen sufficiently large to satisfy the forthcoming requirements of the proof.

Let $(f_J)_{J \in \mathcal{P}(\delta)}$ satisfy $\operatorname{supp} \hat{f}_J \subseteq \Theta_J$ and for $I \in \mathcal{P}(\delta^{k/k+1})$ let

$$f_I := \sum_{\substack{J \in \mathcal{P}(\delta) \\ J \cap I \neq \emptyset}} f_J$$

Thus, by definition,

$$\Big\| \sum_{J \in \mathcal{P}(\delta)} f_J \Big\|_{L^p(\mathbb{R}^n)} \lesssim \Big\| \sum_{I \in \mathcal{P}(\delta^{k/k+1})} f_I \Big\|_{L^p(\mathbb{R}^n)} \qquad (9)$$

$$\leq D_{k,p,\gamma}(\delta^{k/k+1}) \Big(\sum_{I \in \mathcal{P}(\delta^{k/k+1})} \|f_I\|^2_{L^p(\mathbb{R}^n)} \Big)^{\frac{1}{2}}$$

By the transformations used in (7) and (8), for $I = [a, a + \delta^{k/k+1}] \in \mathcal{P}(\delta^{k/k+1})$ we can map $\gamma(I)$ to a small perturbation of the moment curve. In particular, by (7) we have

$$\gamma(a + t\delta^{k/k+1}) = A_{\gamma,a,h}\, \gamma_0(t) + O((\delta^{k/k+1})^{k+1})$$

$$= A_{\gamma,a,h}\, \gamma_0(t) + O(\delta^k)$$

Thus,

$$\Theta_{\gamma,J} \subseteq 2 \cdot \Theta_{\gamma_{a,h},J}, \text{ say},$$

where $\gamma_{a,h} := A_{\gamma,a,h}\, \gamma_0$.

since the decoupling inequalities are invariant under affine transformation of the frequency space, the hypothesis

$$D_{k,p}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$$

implies that

$$\|f_I\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon_0} \delta^{-\varepsilon_0} \Big(\sum_{\substack{J \in \mathcal{P}(\delta) \\ J \cap I \neq \emptyset}} \|f_J\|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2}. \quad (10)$$

Combining (9) and (10),

$$D_{k,p,\gamma}(\delta) \lesssim_{\varepsilon_0} \delta^{-\varepsilon_0} D_{k,p,\gamma}(\delta^{k/k+1})$$

Choosing $\varepsilon_0 := \frac{1}{2(k+1)}\varepsilon$ and applying the induction hypothesis,

$$D_{k,p,\gamma}(\delta) \lesssim C_\varepsilon \delta^{\frac{1}{2(k+1)}\varepsilon} \bar{C}_\varepsilon \delta^{-\varepsilon}$$

$$\leq \bar{C}_\varepsilon \delta^{-\varepsilon}$$

provided δ is sufficiently small, depending on ε and k (which we may assume since otherwise we are in the base case). $\square$