Decoupling seminar notes

In the last two lectures, we discuss the proof of the l^2-decoupling conjecture. I thought it would be helpful to have some notes to look over between the two lectures. There are some exercises folded into the notes.

Throughout these notes $S \subset \mathbb{R}^n$ denotes a compact positively curved C^3 hypersurface, such as the sphere or the truncated paraboloid.

Theorem 1. *(Bourgain-Demeter) Suppose that* $\operatorname{supp} \hat{f} \subset N_{1/R}S$. *Let* $N_{1/R}S$ *be the disjoint union of blocks* θ *of dimensions* $R^{-1/2} \times ... \times R^{-1/2} \times R^{-1}$. *Let* $s = \frac{2(n+1)}{n-1}$, *the critical exponent in the Strichartz inequality. Then for any ball* B_R *of radius* R *in* $\mathbb{R}^n$ *(in physical space), and any* $2 \le p \le s$,

$$\|f\|_{L^p_{avg}(B_R)} \lesssim R^{\epsilon} \left(\sum_{\theta} \|f_\theta\|^2_{L^p_{avg}(\mu_{B_R})} \right)^{1/2}.$$

Throughout these notes, we ignore weights. In particular, we will treat the μ_{B_R} on the right-hand side as a B_R. We use the white lies of orthogonality and locally constant explained on the previous problem sets, and when one corrects these white lies then weights appear.

1. Decoupling norms and the decoupling constants

Let us introduce a notation for the decoupling sum that appears on the right-hand side of the main theorem.

If $\operatorname{supp} \hat{f} \subset N_\delta S$, and $\Omega \subset \mathbb{R}^n$ is any domain, then define

$$\|f\|_{L^{p,\delta}_{avg}(\Omega)} := \left(\sum_{\theta \text{ a } \delta^{1/2} \text{ cap} \subset N_\delta S} \|f_\theta\|^2_{L^p_{avg}(\Omega)} \right)^{1/2}.$$

The $L^{p,\delta}$ are norms, and they share some of the properties of L^p norms. Here are a few exercises to get used to them.

Exercise 1. *Check that* $\|f+g\|_{L^{p,\delta}_{avg}(\Omega)} \le \|f\|_{L^{p,\delta}_{avg}(\Omega)} + \|g\|_{L^{p,\delta}_{avg}(\Omega)}$.

Exercise 2. *By orthogonality, check that for any* $\rho \ge \delta^{-1/2}$, $\|f\|_{L^{2,\delta}_{avg}(B_\rho)} \sim \|f\|_{L^2_{avg}(B_\rho)}$.

Exercise 3. *(Holder-type inequality) If* $1 \le q, q_1, q_2 \le \infty$ *and* $\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha \frac{1}{q_2}$, *then check that*

$$(1) \qquad \|f\|_{L^{q,\delta}_{avg}(B_R)} \le \|f\|^{1-\alpha}_{L^{q_1,\delta}_{avg}(B_R)} \|f\|^{\alpha}_{L^{q_2,\delta}_{avg}(B_R)}.$$

Suppose that Ω is a disjoint union of Ω_j. How does the nrom $\|f\|_{L^{p,\delta}}(\Omega)$ relate to the norms $\|f\|_{L^{p,\delta}}(\Omega_j)$? For regular L^p norms, we would have $\sum_j \|f\|^p_{L^p(\Omega_j)} = \|f\|^p_{L^p(\Omega)}$. For the decoupling norms, we have an inequality in one direction:

Lemma 2. *If Ω is a disjoint union of Ω_j, and if $p \geq 2$, then for any δ and any f with* $\operatorname{supp} \hat{f} \subset N_\delta S$*, we have*

$$\sum_j \|f\|^p_{L^{p,\delta}(\Omega_j)} \leq \|f\|^p_{L^{p,\delta}(\Omega)}.$$

Proof. The left-hand side is

$$\sum_j \left(\sum_\theta \|f_\theta\|^2_{L^p(\Omega_j)} \right)^{\frac{p}{2}} = \left\| \sum_\theta \|f_\theta\|^2_{L^p(\Omega_j)} \right\|^{\frac{p}{2}}_{l_j^{p/2}}.$$

Using the Minkowski inequality for the $l_j^{p/2}$ norm, this expression is

$$\leq \left(\sum_\theta \left\| \|f_\theta\|^2_{L^p(\Omega_j)} \right\|_{l_j^{p/2}} \right)^{\frac{p}{2}} = \left(\sum_\theta \|f_\theta\|^2_{L^p(\Omega)} \right)^{\frac{p}{2}} = \|f\|^p_{L^{p,\delta}(\Omega)}.$$

□

As a corollary, we see that if we have a decoupling inequality on each Ω_j, then we get a decoupling inequality on their union. We state this precisely.

Proposition 3. *(Parallel decoupling) Suppose that Ω is a disjoint union of Ω_j, and that $p \geq 2$ and* $\operatorname{supp} \hat{f} \subset N_\delta S$*. Suppose that for each j, we have the inequality*

$$\|f\|_{L^p(\Omega_j)} \leq M \|f\|_{L^{p,\delta}(\Omega_j)}.$$

Then we also have the inequality

$$\|f\|_{L^p(\Omega)} \leq M \|f\|_{L^{p,\delta}(\Omega)}.$$

Proof. We write

$$\|f\|^p_{L^p(\Omega)} = \sum_j \|f\|^p_{L^p(\Omega_j)} \leq M^p \sum_j \|f\|^p_{L^{p,\delta}(\Omega_j)}.$$

By Lemma 2, this is bounded by $M^p \|f\|^p_{L^{p,\delta}(\Omega)}$. Taking p^{th} roots finishes the proof. □

The decoupling problem is about comparing standard norms L^p_{avg} and decoupled norms $L^{p,\delta}_{avg}$. We define the decoupling constant $D_p(R)$ as the smallest constant so that for all f with $\operatorname{supp} \hat{f} \subset N_{1/R} S$,

$$\|f\|_{L^p_{avg}(B_R)} \leq D_p(R) \|f\|_{L^{p,1/R}_{avg}(B_R)}.$$

(The constant $D_p(R)$ also depends in a mild way on the surface S. We assume that the second fundamental form of S is $\geq c > 0$ and the third derivatives of the functions locally defining S are $\leq C$. Then the constant $D_p(R)$ depends in a polynomial way on c^{-1}, C, but we ignore this small point in the notes...)

Theorem 1 says that $D_p(R) \lesssim R^\epsilon$ for $2 \leq p \leq s = \frac{2(n+1)}{n-1}$.

In the next couple sections, we review two of the fundamental ideas we've been working with. The first idea is to look at a problem at multiple scales in Fourier space. Using parabolic rescaling, we see that the problem of breaking a medium cap τ into smaller blocks θ is basically equivalent to the original decoupling problem. The second idea is to look at multilinear vs. linear versions of the decoupling problem. We will see that for the decoupling problem, the multilinear version and the linear version are essentially equivalent! This is a crucial point, which makes decoupling more accessible than restriction or Kakeya.

2. Parabolic rescaling

We can consider the decoupling problem at many scales in Fourier space. Instead of just breaking all of $N_{1/R} S$ into $R^{-1/2}$ caps θ, what happens if we start with a function supported on a cap $\tau \subset N_{1/R} S$ and break τ into caps θ?

Proposition 4. *If $\tau \subset N_{1/R} S$ is a $r^{-1/2}$ cap for some $r \leq R$, and* $\operatorname{supp} \hat{f} \subset \tau$, *and $\theta \subset N_{1/R} S$ are $R^{-1/2}$-caps as above, then*

$$\|f\|_{L^p_{avg}(B_R)} \lesssim D_p(R/r) \left(\sum_{\theta \subset \tau} \|f_\theta\|^2_{L^p_{avg}(B_R)} \right)^{1/2}.$$

We remark that $D_p(R/r)$ involves decoupling a cap of scale 1 into blocks at scale $(R/r)^{-1/2}$. The setup of the Proposition involves a decoupling a cap of scale $r^{-1/2}$ into blocks at scale $R^{-1/2} = (R/r)^{-1/2} r^{-1/2}$. The Proposition is saying that these are essentially the same problem in different coordinates.

The proof is by parabolic rescaling. We discussed it in lecture in connection with the narrow sets.

Proof. After rotating and translating, we can choose coordinates so that τ is contained in a region of the form $0 < \omega_n < r^{-1}$, and $|\omega_i| < r^{-1/2}$ for $i \le n-1$. Now we do a parabolic rescaling. We define new coordinates

$$\bar{\omega}_n = r\omega_n, \bar{\omega}_i = r^{1/2}\omega_i.$$

We let $\bar{\tau}$ denote the image of τ in the new coordinates. In the new coordinates, $\bar{\tau}$ has diameter 1. It is the $(R/r)^{-1}$-neighborhood of a surface $\bar{S}$ obeying all the good properties of the original S. We let $\bar{\theta}$ be the image of θ in the new coordinates. Each $\bar{\theta}$ is a $(R/r)^{-1/2} \times ... \times (R/r)^{-1/2} \times (R/r)^{-1}$ block.

There is a corresponding coordinate change in physical space. We have

$$\bar{x}_n = r^{-1}x_n, \bar{x}_i = r^{-1/2}x_i.$$

We define $g(\bar{x}) = f(x)$. Since $\operatorname{supp} \hat{f} \subset \tau$, $\operatorname{supp} \hat{g} \subset \bar{\tau}$. We also have $f_\theta(x) = g_{\bar{\theta}}(\bar{x})$.

In the new coordinates, the ball B_R becomes an ellipsoid E, shaped roughly like a pancake, with a short principal axis of length R/r and $n-1$ long principal axes of length $R/r^{-1/2}$. The ellipsoid E can be divided into disjoint shapes that are essentially balls of radius R/r. (The number of such balls is $r^{(n-1)/2}$, but this number won't be important in our computations.)

On each ball $B_{R/r}$, the function g obeys the estimate

$$\|g\|_{L^p_{avg}(B_{R/r})} \le D_p(R/r) \left(\sum_{\bar{\theta}} \|g_{\bar{\theta}}\|^2_{L^p_{avg}(B_{R/r})} \right)^{1/2}.$$

We can cover the ellipsoid E with essentially disjoint balls $B_{R/r}$. By parallel decouping, Proposition 3, a decoupling inequality on each $B_{R/r}$ gives us a decoupling inequality on E. Therefore,

$$\|g\|_{L^p_{avg}(E)} \lesssim D_p(R/r) \left(\sum_{\bar{\theta}} \|g_{\bar{\theta}}\|^2_{L^p_{avg}(E)} \right)^{1/2}.$$

Now we change coordinates back to the original coordinates. Since all the norms are averaged, there are no Jacobian factors, and we just get

$$\|f\|_{L^p_{avg}(B_R)} \lesssim D_p(R/r) \left(\sum_{\theta} \|f_\theta\|^2_{L^p_{avg}(B_R)} \right)^{1/2} = D_p(R/r)\|f\|_{L^{p,1/R}_{avg}(B_R)}.$$

□

As a corollary, we get the following estimate:

Proposition 5. *For any radii $R_1, R_2 \ge 1$, we have $D_p(R_1R_2) \lesssim D_p(R_1)D_p(R_2)$.*

Proof. We let $R = R_1 R_2$. We let f be a function with $\operatorname{supp} \hat{f} \subset N_{1/R}S$. We let $\tau \subset N_{1/R}S$ be $R_1^{-1/2}$ caps. First, since $R \geq R_1$, we can cover B_R with disjoint balls of radius R_1 to get the estimate:

$$\|f\|_{L^p_{avg}(B_R)} \leq D_p(R_1) \left(\sum_\tau \|f_\tau\|^2_{L^p_{avg}(B_R)} \right)^{1/2}.$$

We let $\theta \subset N_{1/R}S$ be $R^{-1/2}$ blocks. Next, we use the last Proposition to bound $\|f_\tau\|_{L^p_{avg}(B_R)} \lesssim D_p(R_2)(\sum_{\theta\subset\tau} \|f_\theta\|^2_{L^p_{avg}(B_R)})^{1/2}$. Plugging this estimate into the last equation, we get

$$\|f\|_{L^p_{avg}(B_R)} \lesssim D_p(R_1) D_p(R_2) \left(\sum_\theta \|f_\theta\|^2_{L^p_{avg}(B_R)} \right)^{1/2}.$$

□

In particular, we see that there is a unique power $\gamma = \gamma(n, p)$ so that for all R, ϵ,

$$R^{\gamma-\epsilon} \lesssim D_p(R) \lesssim R^{\gamma+\epsilon}.$$

We write $D_p(R) \approx R^\gamma$. We want to prove that $\gamma = 0$ (for $2 \leq p \leq s = \frac{2(n+1)}{n-1}$).

3. Multilinear vs. linear decoupling

A second main idea we consider is looking at the multilinear version of a problem. We have seen that the multilinear versions of Kakeya and restriction are much more approachable than the original problems and have useful applications. We formulate a multilinear version of the decoupling problem.

We say that functions $f_1, ..., f_n$ on $\mathbb{R}^n$ obey the multilinear decoupling setup (MDS) if

- For $i = 1, ..., n$, $\operatorname{supp} \hat{f}_i \subset N_{1/R}S_i$
- As usual, $S_i \subset \mathbb{R}^n$ are compact positively curved C^3 hypersurfaces. (More precisely, the second fundamental form of S_i is $\gtrsim 1$ and the third derivatives are $\lesssim 1$.)
- (Transversality) For any point $\omega \in S_i$, the normal vector $\nu(\omega)$ obeys

$$\text{Angle}(\nu(\omega), i^{th} \text{ coordinate axis}) \leq (10n)^{-1}.$$

We define $\tilde{D}_{n,p}(R)$ to be the smallest constant so that whenever f_i obey (MDS),

$$\left\| \prod_{i=1}^n |f_i|^{1/n} \right\|_{L^p_{avg}(B_R)} \leq \tilde{D}_{n,p}(R) \prod_{i=1}^n \|f_i\|^{1/n}_{L^{p,1/R}_{avg}(B_R)}.$$

Bourgain and Demeter proved the following result connecting linear decoupling and multilinear decoupling.

Theorem 6. *(Bourgain-Demeter) Suppose that in dimesion $n-1$, the decoupling constant $D_{n-1,p}(R) \lesssim R^\epsilon$ for any $\epsilon > 0$. Then for any $\epsilon > 0$,*

$$D_{n,p}(R) \lesssim R^\epsilon \tilde{D}_{n,p}(R).$$

Proof sketch. We cover $N_{1/R}S$ by K^{-1} caps τ, where K is a parameter that we can select below. Each $|f_\tau|$ is morally constant on cubes Q_K (of side length K) in B_R. We cover B_R with cubes Q_K, and we classify the cubes as broad or narrow depending on which τ make a significant contribution to $f|_{Q_K}$.

We control the contribution of the broad cubes using the multilinear decoupling inequality.

$$\|f\chi_{Broad}\|_{L^p_{avg}(B_R)} \lesssim K^c \tilde{D}_{n,p}(R) \|f\|_{L^{p,1/R}_{avg}(B_R)}.$$

We control the contribution of each narrow cube using $D_{n-1,p}(K)$. If Q_K is a narrow cube, we get

$$\|f\|_{L^p_{avg}(Q_K)} \lesssim D_{n-1,p}(K) \left(\sum_{\alpha \text{ a } K^{-1/2} \text{ cap}} \|f_\alpha\|^2_{L^p_{avg}(Q_K)} \right)^{1/2}.$$

By assumption, for any $\bar{\epsilon} > 0$, we get

$$\|f\|_{L^p_{avg}(Q_K)} \lesssim K^{\bar{\epsilon}} \left(\sum_{\alpha \text{ a } K^{-1/2} \text{ cap}} \|f_\alpha\|^2_{L^p_{avg}(Q_K)} \right)^{1/2}.$$

Using parallel decoupling, Proposition 3, we can combine the estimate from each narrow cube to give a decoupling estimate on their union. We get

$$\|f\chi_{Narrow}\|_{L^p_{avg}(B_R)} \lesssim K^{\bar{\epsilon}} \left(\sum_{\alpha \text{ a } K^{-1/2} \text{ cap}} \|f_\alpha\|^2_{L^p_{avg}(B_R)} \right)^{1/2}.$$

Now we apply the parabolic rescaling proposition, Proposition 4, to each cap α. We then get

$$\|f\chi_{Narrow}\|_{L^p_{avg}(B_R)} \lesssim C_{\bar{\epsilon}} K^{\bar{\epsilon}} D_{n,p}(R/K) \|f\|_{L^{p,1/R}_{avg}(B_R)}.$$

Combining the broad and narrow pieces, we see that

$$D_{n,p}(R) \lesssim_{\bar{\epsilon}} K^c \tilde{D}_{n,p}(R) + K^{\bar{\epsilon}} D_{n,p}(R/K).$$

We can now choose K and iterate this inequality. Choosing either $K = R^\epsilon$ or K a large constant or $K = \log R$ all work fine. The result is that $D_{n,p}(R) \lesssim R^\epsilon \tilde{D}_{n,p}(R)$. This finishes our sketch of the proof of Theorem 6.

We also remark that the multilinear decoupling constant is bounded by the linear one. We make this an exercise.

Exercise 4. *Prove that $\tilde{D}_{n,p}(R) \le D_{n,p}(R)$ for any n, p, R.*

We will prove the decoupling theorem by induction on the dimension n. If the decoupling theorem holds in dimension $n-1$, and if $2 \le p \le s = \frac{2(n+1)}{n-1}$, then Theorem 6 and Exercise 4 show that

$$\tilde{D}_{n,p}(R) \approx D_{n,p}(R) \approx R^\gamma. \tag{2}$$

In summary, we see that the decoupling problem is essentially equivalent to the multilinear decoupling problem. This situation is quite different from the Kakeya problem and the restriction problem. The original Kakeya problem is still open. The multilinear Kakeya inequality may be a useful tool, but it currently seems very hard to prove Kakeya using multilinear Kakeya. The same holds for the restriction problem. But in the decoupling problem, the linear version can be reduced to the multilinear version!

We will study multilinear decoupling using our multilinear toolbox. We will apply multilinear Kakeya (and/or restriction), and we will also adapt ideas from the proof of multilinear restriction.

4. Multilinear restriction revisited

We revisit the proof of multilinear restriction, presenting it in a slightly different way to parallel the proof of the decoupling theorem. In this organization, the multilinear restriction theorem follows from a main lemma which we use at many scales.

Main Lemma 1. *If $\operatorname{supp} f_i \subset N_{1/R} S_i$, and S_i are smooth compact transverse hypersurfaces, and if $2 \le p \le \frac{2n}{n-1}$, then*

$$\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{p}{n}} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_R)}^{\frac{p}{n}}.$$

Exercise 5. *Fill in the proof using orthogonality, the locally constant property, and multilinear Kakeya.*

Multilinear restriction can be proven by using Main Lemma 1 at many scales. We consider a radius $r \ll R$, and we use main lemma 1 at scale $r^2, r^4, r^8, ..., r^{2^A} = R$. We define $r_a := r^{2^a}$.

Suppose that f_i are as in Main Lemma 1. By Bernstein's inequality, if $\operatorname{supp} \hat{g} \subset B_{10}$, then there is a constant C_n so that $\|g\|_{L^\infty(B_r)} \lesssim r^{C_n} \|g\|_{L^2_{avg}(B_r)}$. Applying this crude inequality to each f_i on each $B_r \subset B_R$, we get:

$$\fint_{B_R} \prod_{i=1}^n |f_i|^{\frac{p}{n}} \lesssim r^C \operatorname{Avg}_{B_r \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{p}{n}}_{L^2_{avg}(B_r)}.$$

This first step is not sharp. We have lost a factor of $r^C = R^{\frac{C}{2^M}}$. We will choose $M \to \infty$, and so this loss is acceptable. Now we can bring Main Lemma 1 into play. Using Main Lemma 1 at each step, we see that

$$\operatorname{Avg}_{B_r \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{p}{n}}_{L^2_{avg}(B_r)} \lesssim R^\epsilon \operatorname{Avg}_{B_{r_1} \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{p}{n}}_{L^2_{avg}(B_{r_1})} \lesssim$$

$$\lesssim R^\epsilon \operatorname{Avg}_{B_{r_2} \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{p}{n}}_{L^2_{avg}(B_{r_2})} \lesssim ...(M \text{ times}) \lesssim$$

$$\leq C(M, \epsilon) R^\epsilon \prod_{i=1}^n \|f_i\|^{\frac{p}{n}}_{L^2_{avg}(B_R)}.$$

This proves multilinear restriction.

Multilinear restriction quickly implies multilinear decoupling for $2 \leq p \leq \frac{2n}{n-1}$. Suppose that f_i obey the multilinear decoupling setup (MDS) - or just suppose that f_i obey the weaker assumptions in Main Lemma 1.

$$\left\| \prod_{i=1}^n |f_i|^{1/n} \right\|_{L^p_{avg}(B_R)} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|^{\frac{1}{n}}_{L^2_{avg}(B_R)} \sim$$

Taking $\theta \subset N_{1/R} S$ to be $R^{-1/2}$ caps,

$$\sim \prod_{i=1}^n \left(\sum_\theta \|f_{i,\theta}\|^2_{L^2_{avg}(B_R)} \right)^{\frac{1}{2} \cdot \frac{1}{n}} \leq \prod_{i=1}^n \left(\sum_\theta \|f_{i,\theta}\|^2_{L^p_{avg}(B_R)} \right)^{\frac{1}{2} \cdot \frac{1}{n}} = \prod_{i=1}^n \|f_i\|^{\frac{1}{n}}_{L^{p,1/R}_{avg}(B_R)}.$$

Putting together all the arguments so far, we get a decoupling estimate $D_p(R) \lesssim R^\epsilon$ for $2 \leq p \leq \frac{2n}{n-1}$. This material is essentially what appears in Bourgain's first decoupling paper around 2011.

Notice that when we use multilinear restriction to prove multilinear decoupling, in the last step we use the inequality $\|f_{i,\theta}\|_{L^2_{avg}(B_R)} \leq \|f_{i,\theta}\|_{L^p_{avg}(B_R)}$. Potentially, one can lose a lot in this inequality, and yet the argument still works. In other words, multilinear restriction is much stronger than multilinear decoupling for $p = \frac{2n}{n-1}$. In hindsight, this may have been a clue that it is possible to push the argument further and get decoupling for some exponent greater than $\frac{2n}{n-1}$.

5. The second main lemma

Now we turn to the new ideas in the more recent Bourgain-Demeter paper. The goal is to prove Theorem 1, getting decoupling in the sharp range $2 \leq p \leq s = \frac{2(n+1)}{n-1}$. We will focus on the critical exponent s, which is the most interesting. The proof uses a variation of the main lemma, called Main Lemma 2, and applies Main Lemma 2 at many scales following the outline of the proof of multilinear restrition.

Recall that Main Lemma 1 says that for $p = \frac{2n}{n-1}$, we have

$$\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{p}{n}} \lesssim R^{\epsilon} \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_R)}^{\frac{p}{n}}.$$

Since we want to prove decoupling with the exponent s instead of $p = \frac{2n}{n-1}$, it's natural to try to estimate $\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{s}{n}}$. We first give an estimate with a naive argument, and then we explain how Bourgain and Demeter improved it.

We prove our first upper bound by interpolating Main Lemma 1 with a trivial L^{∞} estimate. To see this we rewrite Main Lemma 1 in the form

$$\left\| \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_{R^{1/2}}))}^{\frac{1}{n}} \right\|_{l^p_{avg}} \lesssim R^{\epsilon} \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_R)}^{\frac{1}{n}}. \tag{3}$$

In this inequality, the l^p_{avg} norm on the left-hand side is over the copies of $B_{R^{1/2}} \subset B_R$. On the other hand we have an easy L^{∞} bound:

$$\left\| \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_{R^{1/2}}))}^{\frac{1}{n}} \right\|_{l^{\infty}_{avg}} \leq \prod_{i=1}^{n} \|f_i\|_{L^{\infty}_{avg}(B_R)}^{\frac{1}{n}}. \tag{4}$$

Interpolating between these last two inequalities, we get

$$\left\| \prod_{i=1}^{n} \|f_i\|_{L^2_{avg}(B_{R^{1/2}}))}^{\frac{1}{n}} \right\|_{l^s_{avg}} \lesssim R^{\epsilon} \prod_{i=1}^{n} \|f_i\|_{L_{avg}^{\frac{2(n+1)}{n}}(B_R)}^{\frac{1}{n}}.$$

The norms $L^{\frac{2(n+1)}{n}}$ haven't appeared in our story before, but we can use Holder to bound $\|f_i\|_{L_{avg}^{\frac{2(n+1)}{n}}(B_R)} \leq \|f_i\|_{L^2_{avg}(B_R)}^{1/2} \|f_i\|_{L^s_{avg}(B_R)}^{1/2}$. (It's a nice algebraic feature of the exponent s that we get 1/2 and 1/2 in the last expression.) Plugging in, we get all together

$$\text{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{s}{n}} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_R)}^{\frac{1}{2}\cdot\frac{s}{n}} \prod_{i=1}^n \|f_i\|_{L^s_{avg}(B_R)}^{\frac{1}{2}\cdot\frac{s}{n}}. \tag{5}$$

This inequality is not good enough to prove decoupling. Bourgain and Demeter observed that we can improve it by replacing the L^s-norm on the right-hand side by a decoupled norm $L^{s,1/R}$. Here is the statement of their main lemma.

Main Lemma 2. *If* $\text{supp}\, f_i \subset N_{1/R} S_i$, *and* S_i *are compact positively curved transverse hypersurfaces, and if* $s = \frac{2(n+1)}{n-1}$, *and* $\delta = R^{-1}$, *then*

$$\text{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{s}{n}} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_R)}^{\frac{1}{2}\cdot\frac{s}{n}} \prod_{i=1}^n \|f_i\|_{L^{s,\delta}_{avg}(B_R)}^{\frac{1}{2}\cdot\frac{s}{n}}.$$

The output is a geometric average of a product of L^2 norms and a product of decoupled norms. The L^2 norms look similar to the output of Main Lemma 1, so they are very good. The decoupled norms represent some progress towards decoupling, so they are also good.

To see how the decoupling may come in, let us observe that there is a decoupled version of the trivial L^∞ estimate in Equation 4. If we let $\theta \subset N_{1/R} S$ be disjoint $R^{-1/2}$ caps as usual, then

$$\left\| \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}}))}^{\frac{1}{n}} \right\|_{l^\infty_{avg}} \sim \left\| \prod_{i=1}^n \left(\oint_{B_{R^{1/2}}} \sum_\theta |f_{i,\theta}|^2 \right)^{\frac{1}{2}\cdot\frac{1}{n}} \right\|_{l^\infty_{avg}} \leq \prod_{i=1}^n \left\| \sum_\theta |f_{i,\theta}|^2 \right\|_{L^\infty(B_R)}^{\frac{1}{2}\cdot\frac{1}{n}}.$$

But $\| \sum_\theta |f_{i,\theta}|^2 \|_{L^\infty(B_R)}^{\frac{1}{2}} = \|f_i\|_{L^{\infty,1/R}_{avg}(B_R)}$ and so we get

$$\left\| \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}}))}^{\frac{1}{n}} \right\|_{l^\infty_{avg}} \lesssim \prod_{i=1}^n \|f_i\|_{L^{\infty,1/R}_{avg}(B_R)}^{\frac{1}{n}}. \tag{6}$$

If we could do interpolation with the $L^{p,\delta}$ norms, then Main Lemma 2 would follow from interpolating between Main Lemma 1 and the $L^{\infty,1/R}$ bound in Equation 6. I don't know how generally it is possible to do such interpolation, but Bourgain and Demeter prove it in this particular case. Now we turn to their proof.

Proof. Let $\theta \subset N_{1/R}S$ be $R^{-1/2}$ caps as usual.

$$\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{R^{1/2}})}^{\frac{s}{n}} \sim \mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \left(\fint_{B_{R^{1/2}}} \sum_\theta |f_{i,\theta}|^2 \right)^{\frac{1}{2} \cdot \frac{s}{n}}.$$

Let $p = \frac{2n}{n-1}$, and bound the last expression by

$$\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \left(\fint_{B_{R^{1/2}}} \sum_\theta |f_{i,\theta}|^2 \right)^{\frac{1}{2} \cdot \frac{p}{n}} \cdot \prod_{i=1}^n \left(\sum_\theta \|f_{i,\theta}\|^2_{L^\infty(B_R)} \right)^{\frac{1}{2} \cdot \frac{s-p}{n}}.$$

The first factor is exactly what appears in Main Lemma 1. Applying Main Lemma 1 to the first factor we see that the whole expression is

$$\lesssim R^\epsilon \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_R)}^{\frac{p}{n}} \cdot \left(\sum_\theta \|f_{i,\theta}\|^2_{L^\infty} \right)^{\frac{1}{2} \cdot \frac{s-p}{n}}.$$

If we rewrite this expression in terms of the decoupling norms, we get

$$R^\epsilon \prod_{i=1}^n \|f_i\|_{L^{2,\delta}_{avg}(B_R)}^{\frac{p}{n}} \cdot \|f_i\|_{L^{\infty,\delta}_{avg}(B_R)}^{\frac{s-p}{n}}.$$

At this point, we would like combine some of the $L^{2,\delta}_{avg}$ and the $L^{\infty,\delta}_{avg}$ norms to create some $L^{s,\delta}_{avg}$ norm. This has a similar feel to a reverse Holder inequality. In Exercise 3, we proved a version of the Holder inequality for these decoupling norms, which we now recall.

If $1 \le q, q_1, q_2 \le \infty$ and $\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha \frac{1}{q_2}$, then

$$\|f\|_{L^{q,\delta}_{avg}(B_R)} \le \|f\|^{1-\alpha}_{L^{q_1,\delta}_{avg}(B_R)} \|f\|^{\alpha}_{L^{q_2,\delta}_{avg}(B_R)}. \tag{7}$$

We would like to use the opposite inequality, so we pose the question, when do we have the inequality

$$\|f\|^{1-\alpha}_{L^{q_1,\delta}_{avg}(B_R)} \|f\|^{\alpha}_{L^{q_2,\delta}_{avg}(B_R)} \lesssim \|f\|_{L^{q,\delta}_{avg}(B_R)}? \tag{8}$$

The reverse Holder inequality, Equation 8, does not hold for an arbitrary function f_i. But Bourgain and Demeter observed that an arbitrary function can be broken into a small number of 'balanced' pieces so that the reverse Holder inequality holds on each piece. We state this as a lemma.

Lemma 7. *Suppose that* $\operatorname{supp} \hat{f}_i \subset N_{1/R} S_i$. *We can write* $f_i = \sum_k f_{i,k} + e_i$, *where*

- *the number of terms in the sum is at most* R^ϵ*,*
- *for each* k*,* $\operatorname{supp} \hat{f}_{i,k} \subset N_{1/R} S_i$*,*
- $\|e_i\|_{L^\infty(B_R)} \le R^{-100n} \|f_i\|_{L^\infty(B_R)}$*,*
- *for each* k*, and any exponent* p*,* $\|f_{i,k}\|_{L^{p,\delta}_{avg}(B_R)} \le \|f_i\|_{L^{p,\delta}_{avg}(B_R)}$*, and*
- *each* $f_{i,k}$ *obeys a reverse Holder inequality. If* $1 \le q, q_1, q_2 \le \infty$ *and if*

$$\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha \frac{1}{q_2},$$

then

$$\|f_{i,k}\|^{1-\alpha}_{L^{q_1,\delta}_{avg}(B_R)} \|f_{i,k}\|^{\alpha}_{L^{q_2,\delta}_{avg}(B_R)} \sim \|f_{i,k}\|_{L^{q,\delta}_{avg}(B_R)}. \tag{9}$$

Now we return to the start of the argument and break up our estimate into contributions from the different $f_{i,k}$. Since e_i is really tiny compared to f_i, we can neglect it. Since the number of k is $\lesssim R^\epsilon$, we get

$$\operatorname{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^2_{avg}(B_{R^{1/2}})} \lesssim R^\epsilon \max_{k_1,\dots,k_n} \operatorname{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_{i,k_i}\|^{\frac{s}{n}}_{L^2_{avg}(B_{R^{1/2}})}.$$

The functions $f_{i,k}$ obey all the hypotheses of the f_i, and so by our previous analysis, the last expression is bounded by

$$\lesssim R^\epsilon \prod_{i=1}^n \|f_{i,k_i}\|^{\frac{p}{n}}_{L^{2,\delta}_{avg}(B_R)} \cdot \|f_{i,k_i}\|^{\frac{1}{2}\cdot\frac{s-p}{n}}_{L^{\infty,\delta}_{avg}(B_R)}.$$

Now Lemma 7 tells us that we can apply the reverse Holder inequality, giving

$$\sim R^\epsilon \prod_{i=1}^n \|f_{i,k_i}\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^2_{avg}(B_R)} \prod_{i=1}^n \|f_{i,k_i}\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^{s,\delta}_{avg}(B_R)} \lesssim$$

$$\lesssim R^\epsilon \prod_{i=1}^n \|f_i\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^2_{avg}(B_R)} \prod_{i=1}^n \|f_i\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^{s,\delta}_{avg}(B_R)}.$$

This finishes the proof of Main Lemma 2, except for the proof of the reverse Holder inequality, Lemma 7. □

Before we prove Lemma 7 for the fancy norms $L^{p,\delta}_{avg}(B_R)$, let us begin with a similar inequality for the regular L^p norms.

Lemma 8. *Suppose that* $\operatorname{supp} \hat{f} \subset B(1)$*. Then we can write* $f = \sum_k f_k + e$*, where*

- *the sum has* $\lesssim R^\epsilon$ *terms,*
- $\|e\|_{L^p(B_R)} \le R^{-10n} \|f\|_{L^p(B_R)}$ *for any* $1 \le p \le \infty$*, and*

- *each f_k obeys a reverse Holder inequality. If $1 \le q, q_1, q_2 \le \infty$ and if*

$$\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha\frac{1}{q_2},$$

then

$$\|f_k\|^{1-\alpha}_{L^{q_1}_{avg}(B_R)} \|f_k\|^{\alpha}_{L^{q_2}_{avg}(B_R)} \sim \|f_k\|_{L^q_{avg}(B_R)}. \tag{10}$$

Proof. First suppose that a function g has the form $h \cdot \chi_A$ for some $A \subset B_R$. Then $\|g\|_{L^q} = h|A|^{1/q}$. Therefore, if $\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha\frac{1}{q_2}$, then

$$\|g\|_{L^q(B_R)} = \|g\|^{1-\alpha}_{L^{q_1}(B_R)} \|g\|^{\alpha}_{L^{q_2}(B_R)}.$$

More generally, g obeys a reverse Holder inequality if $|g|$ is essentially constant on $\operatorname{supp} g$. An arbitrary function f can now be broken into dyadic pieces $f = \sum_k \chi_{\{|f| \sim 2^k\}} f = \sum_k f_k$, and each of these pieces obeys a reverse Holder inequality.

Suppose that $2^{\bar{k}} \sim \sup_{B_R} |f|$. We write f as $\sum_{k=\bar{k}-100n\log_2 R}^{\bar{k}} f_k + e$. We immediately get that $\|e\|_{L^\infty(B_R)} \le R^{-100n}\|f\|_{L^\infty(B_R)}$. Because of Bernstein's theorem, $\|f\|_{L^p(B_R)} \gtrsim \|f\|_{L^\infty(B_R)}$ for any $1 \le p \le \infty$, and then it follows that $\|e\|_{L^p(B_R)} \lesssim R^{-10n}\|f\|_{L^p(B_R)}$ for any $1 \le p \le \infty$. $\square$

Now we turn to the proof of Lemma 7.

Proof. Consider a function f with $\operatorname{supp} \hat{f} \subset N_{1/R}S$. First we recall the wave packet decomposition of f. We have $f = \sum_\theta f_\theta$. Each f_θ can be written as a sum of wave packets:

$$f_\theta \sim \sum_{T \in T(\theta)} a_T \phi_T.$$

Here T is a translate of θ^*, a tube of length R and radius $R^{1/2}$. The coefficient a_T is a complex number. And ϕ_T is a "wave packet", a function essentially supported on T, with $\hat{\phi}_T \subset \theta$. We also normalize ϕ_T so that $|\phi_T| \sim 1$ on T.

First we decompose f by the amplitude of the wave packets: $f = \sum f_k$, where

$$f_k = \sum_\theta \sum_{T, |a_T| \sim 2^k} a_T \phi_T.$$

Now we subdivide f_k further. Fix k. For each θ, let $T_k(\theta) = \{T \text{ so that } |a_T| \sim 2^k\}$. We decompose f_k by the order of magnitude of $|T_k(\theta)|$. We have $f_k = \sum_l f_{k,l}$ where

$$f_{k,l} = \sum_{\theta, |T_k(\theta)| \sim 2^l} \sum_{T \in T_k(\theta)} a_T \phi_T.$$

For each $\theta \in T_k(\theta)$, we have

$$\|f_{k,l,\theta}\|_{L^q_{avg}} \sim 2^k \left(\frac{2^l |T|}{|B_R|} \right)^{1/q}.$$

Using this formula for each θ, it follows that $f_{k,l}$ obeys the reverse Holder inequality in Equation 9.

We also note that for any q, $\|f_{k,l}\|_{L^{q,\delta}_{avg}(B_R)} \lesssim \|f\|_{L^{q,\delta}_{avg}(B_R)}$.

For each $T \in T(\theta)$, $\operatorname{supp} \hat{\phi}_T \subset \theta$. Therefore, $\operatorname{supp} \hat{f}_{k,l} \subset \operatorname{supp} \hat{f}$.

If we use R^ϵ terms $f_{k,l}$ the remaining low amplitude terms lead to an error term e with $\|e\|_{L^\infty(B_R)} \leq R^{-1000n} \|f\|_{L^\infty(B_R)}$.

$\square$

To help digest the upcoming argument, it might help to consider the following question. What would happen in Main Lemma 2 if we tried to use an exponent $q > \frac{2(n+1)}{n-1}$ in place of $s = \frac{2(n+1)}{n-1}$?

For any q, the proof of Main Lemma 2 can be modified to yield an inequality of the following form:

$$\mathrm{Avg}_{B_{R^{1/2}} \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{q}{n}}_{L^2_{avg}(B_{R^{1/2}})} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|^{\alpha(q) \cdot \frac{q}{n}}_{L^2_{avg}(B_R)} \prod_{i=1}^n \|f_i\|^{(1-\alpha(q)) \cdot \frac{q}{n}}_{L^{q,\delta}_{avg}(B_R)}.$$

If $q = \frac{2n}{n-1}$, then $\alpha(q) = 1$. If $q = \frac{2(n+1)}{n-1}$, then $\alpha(q) = 1/2$. If $q > \frac{2(n+1)}{n-1}$, then $\alpha(q) < 1/2$. The iteration argument below proves decoupling with an R^ϵ loss if and only if $\alpha(q) \geq 1/2$.

To see why $1/2$ might be a natural barrier here, consider the following. The left-hand side involves a mix of L^2 norms (at scale up to $R^{1/2}$) and l^q norms (at scales from $R^{1/2}$ to R). In some sense it is a half and half mix of L^2 and L^q. (Note that the functions f_i are essentially constant at scale 1.) The right-hand side involves a mix of L^2 and $L^{q,\delta}$, where the L^2 part is weighted by α and the $L^{q,\delta}$ part is weighted by $(1-\alpha)$. If $\alpha = 1/2$, then Main Lemma 2 essentially decouples the L^q part and turns the L^2 part into a different L^2 part. If $\alpha < 1/2$, then the inequality above turns a little bit of L^2 plus some L^q into decoupled L^q. Since L^2 is smaller than L^q, this is not as good as turning L^q into $L^{q,\delta}$.

6. The last iteration

In this section, we prove the main decoupling theorem by using Main Lemma 2 at many scales. The argument follows the same outline as the proof of multilinear restriction in Section 4, but there is an extra twist.

As above, let $s = \frac{2(n+1)}{n-1}$. We will prove that $D_{n,p}(R) \lesssim R^\epsilon$ for $2 \le p \le s$. The proof is by induction on the dimension, so we can assume that $D_{n-1,p}(R) \lesssim R^\epsilon$. As we saw in Section 3, $D_{n,p}(R) \approx \tilde{D}_{n,p}(R) \approx R^{\gamma(n,p)}$. We just have to show that $\gamma(n,p) = 0$. To understand $\tilde{D}_{n,p}(R)$, we work in the multilinear decoupling setup. Suppose that f_i obey the multilinear decoupling setup (MDS) as in Section 3.

We focus on the most interesting case $p = s$. At the end we discuss how to slightly modify the proof to handle the cases $2 \le p \le s$.

We want to prove a bound of the form

$$\fint_{B_R} \prod_{i=1}^n |f_i|^{\frac{s}{n}} \lesssim R^\epsilon \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^{s,1/R}_{avg}(B_R)}.$$

As in Section 4, we consider a sequence of scales. We pick a large integer M and we let $r = R^{2^{-M}}$. We will use Main Lemma 2 at scales $r, r^2, r^4, ..., r^{2^M} = R$. We abbreviate $r_a = r^{2^a}$.

We begin with a crude inequality at the small scale r:

$$\fint_{B_R} \prod_{i=1}^n |f_i|^{\frac{s}{n}} \lesssim r^C \operatorname{Avg}_{B_r \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^2_{avg}(B_r)}. \tag{11}$$

This inequality was not sharp, so we lost a factor of $r^C = R^{C 2^{-M}}$. Now we can apply Main Lemma 2 on each ball of radius $r^2 = r_1$. We get

$$r^C \operatorname{Avg}_{B_r \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^2_{avg}(B_r)} \lesssim r^C \operatorname{Avg}_{B_{r_1} \subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^2_{avg}(B_{r_1})} \|f_i\|^{\frac{1}{2}\cdot\frac{s}{n}}_{L^{s,1/r_1}_{avg}(B_{r_1})}.$$

(Remark: When we use Main Lemma 2, there is also a factor of r^ϵ which we absorb into r^C.)

At this point, we are not exactly in a position to iterate Main Lemma 2, because of the terms of the form $\|f_i\|_{L^{s,1/r_1}_{avg}(B_{r_1})}$. These terms are partly decoupled. We use Holder's inequality to separate them from the L^2 terms, and then we will decouple them the rest of the way in terms of $D_{n,s}(R/r_1)$.

Recall that the multilinear Holder inequality says that if $b_j > 0$ and $\sum b_j = 1$, then

$$\operatorname{Avg} \prod_j A_j^{b_j} \le \prod_j \left(\operatorname{Avg} A_j\right)^{b_j}.$$

We will apply the $(n+1)$-linear Holder inequality with exponents $\frac{1}{2} + \frac{1}{2n} + ... + \frac{1}{2n} = 1$.

$$r^C \operatorname{Avg}_{B_{r_1}\subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{r_1})}^{\frac{1}{2}\cdot\frac{s}{n}} \|f_i\|_{L^{s,1/r_1}_{avg}(B_{r_1})}^{\frac{1}{2}\cdot\frac{s}{n}} =$$

$$= r^C \operatorname{Avg}_{B_{r_1}\subset B_R} \left(\prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{r_1})}^{\frac{s}{n}}\right)^{\frac{1}{2}} \prod_{i=1}^n \left(\|f_i\|^s_{L^{s,1/r_1}_{avg}(B_{r_1})}\right)^{\frac{1}{2n}} \leq (\text{ by Holder })$$

$$\leq r^C \left(\operatorname{Avg}_{B_{r_1}\subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{r_1})}^{\frac{s}{n}}\right)^{\frac{1}{2}} \prod_{i=1}^n \left(\operatorname{Avg}_{B_{r_1}\subset B_R} \|f_i\|^s_{L^{s,1/r_1}_{avg}(B_{r_1})}\right)^{\frac{1}{2n}}. \tag{12}$$

The first factor is ready to apply Main Lemma 2 again at the next scale. The second factor is slightly decoupled, and now we explain how to decouple it the rest of the way.

Using Minkowski's inequality, Lemma 2, we can bound

$$Avg_{B_{r_1}\subset B_R}\|f_i\|^s_{L^{s,1/r_1}_{avg}(B_{r_1})} \leq \|f_i\|^s_{L^{s,1/r_1}_{avg}(B_R)}.$$

This expression involves decoupling f_i into contributions from caps of size $r_1^{-1/2}$. We want to decouple f_i into finer caps of size $R^{-1/2}$. To do so, we use parabolic rescaling, Proposition 4, to decouple f_i further, bringing in a factor of $D_{n,s}(R/r_1)$:

$$\|f_i\|_{L^{s,1/r_1}_{avg}(B_R)} \leq D_{n,s}(R/r_1)\|f_i\|_{L^{s,1/R}_{avg}(B_R)} \lesssim (R/r_1)^{\gamma}\|f_i\|_{L^{s,1/R}_{avg}(B_R)} = R^{\gamma(1-\frac{2}{2^M})}\|f_i\|_{L^{s,1/R}_{avg}(B_R)}.$$

All together, we see that the second factor of Equation 12 is bounded as follows:

$$\prod_{i=1}^n \left(\operatorname{Avg}_{B_{r_1}\subset B_R} \|f_i\|^s_{L^{s,1/r_1}_{avg}(B_{r_1})}\right)^{\frac{1}{2n}} \leq R^{s\gamma(\frac{1}{2}-\frac{1}{2^M})} \left(\prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^{s,1/R}_{avg}(B_R)}\right)^{1/2}.$$

Putting together the whole argument so far, we have proven that:

$$\fint_{B_R} \prod_{i=1}^n |f_i|^{\frac{s}{n}} \lesssim r^C \left(\operatorname{Avg}_{B_r\subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^2_{avg}(B_r)}\right) \leq$$

$$\leq r^C \left(\operatorname{Avg}_{B_{r_1}\subset B_R} \prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^2_{avg}(B_{r_1})}\right)^{\frac{1}{2}} R^{s\gamma(\frac{1}{2}-\frac{1}{2^M})} \left(\prod_{i=1}^n \|f_i\|^{\frac{s}{n}}_{L^{s,1/R}_{avg}(B_R)}\right)^{1/2}.$$

Now we can iterate this computation. Repeating the computation one more time, we get:

$$\leq r^C \left(\mathrm{Avg}_{B_{r_2} \subset B_R} \prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_{r_2})}^{\frac{s}{n}} \right)^{\frac{1}{4}} R^{s\gamma(\frac{3}{4} - \frac{2}{2^M})} \left(\prod_{i=1}^n \|f_i\|_{L^{s,1/R}_{avg}(B_R)}^{\frac{s}{n}} \right)^{3/4}.$$

After iterating all the way to scale R, we get that this is

$$\leq r^C \left(\prod_{i=1}^n \|f_i\|_{L^2_{avg}(B_R)}^{\frac{s}{n}} \right)^{\frac{1}{2^M}} R^{s\gamma(1 - \frac{M}{2^M})} \left(\prod_{i=1}^n \|f_i\|_{L^{s,1/R}_{avg}(B_R)}^{\frac{s}{n}} \right)^{1 - \frac{1}{2^M}}.$$

Using that $\|f_i\|_{L^2_{avg}(B_R)} = \|f_i\|_{L^{2,1/R}_{avg}(B_R)} \leq \|f_i\|_{L^{s,1/R}_{avg}(B_R)}$, we see that this is

$$\leq r^C R^{s\gamma(1 - \frac{M}{2^M})} \prod_{i=1}^n \|f_i\|_{L^{s,1/R}_{avg}(B_R)}^{\frac{s}{n}}.$$

In other words, we have shown that

$$\fint_{B_R} \prod_{i=1}^n |f_i|^{\frac{s}{n}} \lesssim r^C R^{s\gamma(1 - \frac{M}{2^M})} \prod_{i=1}^n \|f_i\|_{L^{s,1/R}_{avg}(B_R)}^{\frac{s}{n}}.$$

Since f_i were arbitrary functions obeying the multilinear decoupling setup, we see that

$$R^{s\gamma} \approx \tilde{D}_{n,s}(R) \lesssim r^C R^{s\gamma(1 - \frac{M}{2^M})}.$$

Rearranging, we see that

$$R^{s\gamma \frac{M}{2^M}} \lesssim r^C = R^{\frac{C}{2^M}}.$$

Since C in independent of M, we can now take $M \to \infty$, showing that $\gamma = 0$. This proves decoupling at the sharp exponent $s = \frac{2(n+1)}{n-1}$.

Let us summarize and review the argument. We proved a multilinear decoupling estimate by a combination of the crude inequality 11, Main Lemma 2, and induction. The crude inequality dealt with scale 1 to scale $r = R^{2^{-M}}$, and so in some sense it was used $\frac{1}{2^M}$ of the time. Main Lemma 2 accomplished a fraction $\frac{M}{2^M}$ of the decoupling. The remaining decoupling was done by induction. The argument works because the crude step is negligible compared to the part done by Main Lemma 2.