**LECTURE 5: FROM MULTI-LINEAR RESTRICTION TO MULTI-LINEAR DECOUPLING**

JONATHAN HICKMAN AND MARCO VITTURI

It remains to establish the multi-linear decoupling theorem:

**Theorem 1** (multi-linear $\ell^2$-decoupling [1]). Let $\tau_1, \ldots, \tau_n$ be $\omega$-transverse regions on $P^{n-1}$. For $p \geq 2$ the inequality

$$\left\| \prod_{j=1}^{n} |f_j|^{1/n} \right\|_{L^p_{\omega R}(B_R)} \lesssim \omega^{-O(1)} R^{\alpha(p)} \prod_{j=1}^{n} \left\| f_j \right\|^{1/n}_{L^p_{\omega R}^{-1}(w_{B_R})} \tag{1}$$

holds for Schwartz functions $f_j$ with Fourier support in $\tau_j$ for $1 \leq j \leq n$.

Recall the exponent $\alpha$ is given by

$$\alpha(p) := \begin{cases} 0 & \text{if } 2 \leq p \leq 2(n+1)/(n-1); \\ (n-1)/4 - (n+1)/2p & \text{if } 2(n+1)/(n-1) < p, \end{cases}$$

The proof of Theorem 1 will be presented in this set of notes. We will follow the arguments of the original paper [1], but here the presentation is also heavily influenced by the lecture notes of Guth [2].

1. **Multi-linear Restriction and partial results on the decoupling conjecture**

It was remarked at the end of the previous set of notes that multi-linear decoupling estimates follow directly from the multi-linear restriction theorem in the partial range $2 \leq p \leq 2n/(n-1)$. Here we presenting the details of this argument which relies on the formulation of the multi-linear restriction theorem for regions of $N_{1/R}(P^{n-1})$.

**Proof (of Theorem 1, $2 \leq p \leq 2n/(n-1)$).** The Bennett-Carbery-Tao theorem implies

$$\left\| \prod_{j=1}^{n} |f_j|^{1/n} \right\|_{L^p_{\omega R}(B_R)} \lesssim \omega^{-O(1)} \prod_{j=1}^{n} \left\| f_j \right\|_{L^p_{\omega R}^{-1}(w_{B_R})}$$

and so it suffices to show

$$\left\| f_j \right\|_{L^p_{\omega R}^{-1}(w_{B_R})} \lesssim \left\| f_j \right\|_{L^p_{\omega R}(w_{B_R})}$$

for $1 \leq j \leq n$. But this just follows from the H"older estimate

$$\left\| f_j \right\|_{L^p_{\omega R}^{-1}(w_{B_R})} \lesssim \left\| f_j \right\|_{L^p_{\omega R}(w_{B_R})} \tag{2}$$

together with an application of local orthogonality. □

**Remark 2.** As a consequence of the Bourgain-Guth method expounded in the previous notes, we conclude the (linear) $\ell^2$-decoupling inequality holds for the partial range of exponents $2 \leq p \leq 2n/(n-1)$.

The above argument involves the rather inefficient step (2) which bounds an $L^2$ norm by an $L^p$ norm via H"older’s inequality. This suggests that the multi-linear restriction is substantially stronger than multi-linear decoupling at the exponent
where 0 < \beta < 1 is defined by

\[ 1 - \beta = \frac{2}{(s - 2)(n - 1)}. \]

Notice if \( s = 2(n+1)/(n-1) \), then \( \beta = 1 - 1/2 \) and so the \( L^2 \) and \( L^s \)-norms in the right-hand side of (4) are in some sense “balanced”. The fact that the critical exponent \( 2(n+1)/(n-1) \) plays a special role in this inequality is perhaps some

---

1Recall, when we considered applying the Bourgain–Guth method to the (linear) Fourier restriction problem we encountered a similar inefficient step. Furthermore, this provided the motivation to apply the same techniques to the \( \ell^2 \)-decoupling inequalities in the hope that, in this new context, such inefficiencies could be circumvented.

2It is perhaps not obvious how one can perform such an interpolation since the functions we are dealing with must all satisfy certain Fourier support conditions. Whilst this step can be made rigorous using some of the theory we will introduce below, for now we’ll gloss over this point.
A quick computation verifies that (4) can act as a substitute for Proposition 3 in the $p > 2n/(n-1)$ regime.

It turns out that the new inequality (4) is not strong enough to prove decoupling inequalities, although it does provide motivation for the estimates we will consider in the next section.

3. Multi-linear decoupling via induction on scales

The approach adopted by Bourgain and Demeter was to prove a variant of (4) which can be applied iteratively in order to prove Theorem 1. Hence, broadly speaking the arguments mirror those of the proof of the multi-linear restriction theorem presented in Lecture 3, although we will see significant complications arise in the decoupling case. The new inductive step is identical to (4), but the $L^p$-norms are replaced with decoupling norms, which are, of course, more pertinent to the study of $\ell^2$-decoupling estimates.

**Proposition 4** (Inductive step - Decoupling). Let $\tau_1, \ldots, \tau_n$ be $\omega$-transverse regions of $\mathcal{N}_{R^{-1}}(P^{n-1})$, $B_R$ a fixed $R$-ball and $\mathcal{B}$ a covering of $B_R$ by $R^{1/2}$-balls with bounded-overlap. For $2n/(n-1) \leq s$ the inequality

$$\left\| \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(w_{B_{R/2}})} \right\|_{L^2_{\log}(\mathcal{B})} \lesssim \omega^{-O(1)} R^{-n^2/(n-1)} \prod_{j=1}^{n} \left\| f_j \right\|_{L^\gamma L^{\gamma-1}(\mathbb{R}^n)}^{(1-\beta)/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{\gamma/n} L^{\gamma/n-1}(\mathbb{R}^n)}^{\gamma/n}$$

holds for $f_j$ with Fourier support in $\tau_j$, $1 \leq j \leq n$.

Before presenting a full proof of Proposition 4 we will make some preliminary observations. For $2n/(n-1) \leq s \leq \infty$ define $0 \leq \gamma - \gamma(s) \leq 1$ by

$$s(1-\gamma) := \frac{2n}{n-1} \gamma(s) = 1$$

and observe, by local orthogonality,

$$\prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(w_{B_{R/2}})} \lesssim \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(w_{B_{R/2}})}^{(1-\gamma)/n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{\gamma/n} \lesssim \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(w_{B_{R/2}})}^{(1-\gamma)/n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{\gamma/n}.$$

Consequently, the left-hand side of (15) is dominated by

$$\prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(w_{B_{R/2}})} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{1-\gamma/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{\gamma/n}.$$

The first factor is a $(1-\gamma)$-power of precisely the expression appearing in the inductive step for the multi-linear restriction theorem. In particular, modulo sub-polynomial loss in $R$ we can bound (6) by

$$R^{-n^2/(n-1)s} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{(1-\gamma)/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{2\log}(\mathbb{R}^n)}^{\gamma/n}.$$

Now this is looks similar to the desired estimate, but here we have an $L^{\infty, R^{-1}}$-norm rather than an $L^{s, R^{-1}}$-norm. In particular, to prove Proposition 4 it suffices to show

$$\left\| f_j \right\|_{L^{\infty, R^{-1}}(\mathbb{R}^n)} \left\| f_j \right\|_{L^{\gamma/n} L^{\gamma/n-1}(\mathbb{R}^n)}^{\gamma/n} \leq \left\| f_j \right\|_{L^{s, R^{-1}}(\mathbb{R}^n)}^{2\beta}.$$

A quick computation verifies

$$\frac{1}{s} - \frac{1 - \gamma/\beta}{2} + \frac{\gamma/\beta}{\infty}$$
and so by Hölder’s inequality the reverse of (7) holds; namely,
\[ \|f}\|_{L^n\cdot n^{-1}(\mathbb{R}^n)} \leq \|f}\|_{L^n\cdot n^{-1}(\mathbb{R}^n)} \|f}\|_{L^n\cdot n^{-1}(\mathbb{R}^n)}. \]
Thus, the desired estimate (7) is a reverse Hölder inequality. Unfortunately, (7) does not hold for general \( f_j \), but Bourgain and Demeter observed it is essentially possible to decompose the functions into a small number of pieces for which the reverse Hölder inequality is true.

**Lemma 5 (Decomposition Lemma [1]).** Let \( f \) be a function with Fourier support in \( \tau_j \) which satisfied \( \|f\|_{L^2(\mathbb{R}^n)} = 1 \). For some \( N \leq 1 \) there exists a decomposition
\[ f = \sum_{k=1}^{N} f_k + e \]
for which the following hold:

a) The error \( e \) satisfies
\[ \|e\|_{L^2(w_{B_{N^{1/2}}})} \leq R^{-n^2}\|f\|_{L^2(\mathbb{R}^n)}. \]  

b) Each \( f_k \) has Fourier support in \( \tau_j \) and satisfies
\[ \|f_k\|_{L^n\cdot n^{-1}(\mathbb{R}^n)} \leq \|f\|_{L^n\cdot n^{-1}(\mathbb{R}^n)} \]  

for all \( 1 \leq p \leq \infty \).

c) Furthermore, each \( f_k \) satisfies a reverse Hölder inequality: if \( 1 \leq p, p_0, p_1 \leq \infty \) with \( 1/p - (1 - \alpha)/p_0 + \alpha/p_1 \) for some \( 0 \leq \alpha \leq 1 \), then
\[ \|f_k\|_{L^p\cdot n^{-1}(\mathbb{R}^n)}^{1-\alpha} \|f_k\|_{L^{p_1}\cdot n^{-1}(\mathbb{R}^n)}^\alpha \leq \|f_k\|_{L^p\cdot n^{-1}(\mathbb{R}^n)}. \]

**Remark 6.** The proof will provide substantially stronger control over the error than the \( L^2(w_{B_{N^{1/2}}}) \) estimate described above, but we choose to present the result in this form in view of future applications.

Temporarily assuming this decomposition lemma, the proof of the induction step is easily completed.

**Proof (of Proposition 4).** It suffices to prove the estimate (15) under the additional assumption \( \|f_j\|_{L^2(\mathbb{R}^n)} = 1 \) for \( 1 \leq j \leq n \). Apply the decomposition lemma to write
\[ f_j = \sum_{j=1}^{N_j} f_{j,k} + e_j \]
where, in particular, the \( f_{j,k} \) satisfy the reverse Hölder inequality and \( N_j \leq 1 \). We can now apply our earlier argument to conclude (15) holds with the \( f_{j,k} \) replaced with \( f_{j,k_j} \) for any choice of \( k_j \). The condition \( N_j \leq 1 \) and domination property (9) then imply
\[ \left\| \prod_{j=1}^{n} \sum_{k_j=1}^{N_j} f_{j,k_j} \right\|_{L^{2n}(w_{B_{N^{1/2}}})}^{1/n} \leq \omega^{-O(1)} R^{-n^2/(n-1)} \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(\mathbb{R}^n)}^{1-\beta/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L^{n\cdot n^{-1}(\mathbb{R}^n)}}, \]
where we note the right-hand side of this expression coincides with the right-hand side of (15). It therefore remains to verify the contributions arising from the error terms are negligible; in particular, without loss of generality it suffices to estimate
\[ \prod_{j=1}^{n} \left\| f_{j,k_j} \right\|_{L^{2n}(w_{B_{N^{1/2}}})}^{1/n} \prod_{j=n-m+1}^{n} \left\| e_j \right\|_{L^{2n}(w_{B_{N^{1/2}}})}^{1/n} \]
for any choice of $k_j$’s and $1 \leq m \leq n$. Applying a trivial estimate together with (9) we bound this expression by

$$\prod_{j=1}^{n-m} \|f_j\|_{L^1(\mathbb{R}^n)}^{1/n} \prod_{j=n-m+1}^{n} \|e_j\|_{L^1(\mathbb{R}^{n-1}/2)}^{1/n}$$

whilst property (8) of the error ensures

$$\prod_{j=n-m+1}^{n} \|e_j\|_{L^1(\mathbb{R}^{n-1}/2)} \leq R^{-n^2/(n-1)^2} \sum_{j=n-m+1}^{n} \|f_j\|_{L^1(\mathbb{R}^n)}.$$  

From this together with Hölder’s inequality, one deduces (13) is also bounded by the right-hand side of (15), completing the proof. 

It remains to prove the decomposition lemma. To this end we recall some of the properties of the wave packet decomposition introduced in the first lecture. In particular, fix $f$ with Fourier support in $\tau_j$ and recall

$$f = \sum_{\theta: R^{-1/2}=\text{slab}} f_\theta.$$ 

For any $R^{1/2}$-slab $\theta$ with normal $\omega$ we may apply the wave packet decomposition and write

$$f_\theta = c \sum_{T \in T(\theta)} a_T \psi_T$$

where $T(\theta)$ is a collection of $\sim R \times R^{1/2} \times \cdots \times R^{1/2}$ rectangles $T$ which have bounded overlap and are orientated in the direction of $\omega$. Indeed, for each $\theta$ take a smooth bump function $\hat{\psi}_\theta$ adapted to it; i.e. such that $\hat{\psi}_\theta = 1$ on $\theta$, has support in a small dilated of $\theta$, and that $\|\hat{\psi}_\theta\|_{L^1} \sim 1$ uniformly in $\theta$. As in lecture 1, we can use a Fourier expansion to decompose $f_\theta$ as follows. Let $A_\theta$ be the affine transformation that maps $[-1,1]^n$ to a minimal rectangle containing $\theta$, then $\hat{f}_\theta \circ A_\theta$ is a function that can be identified with a function on $\mathbb{T}^n$ and as such it can be expanded into a Fourier series,

$$\hat{f}_\theta(A_\theta \xi) = \sum_{k \in \mathbb{Z}^n} \langle \hat{f}_\theta \circ A_\theta, e^{2\pi ik \cdot \cdot \cdot} e^{2\pi i k \xi}.$$ 

Since $\hat{f}_\theta = \hat{f}_\theta \hat{\psi}_\theta$ and $|\det A_\theta| \sim |\theta|$, one may deduce that

$$f_\theta(x) = c \sum_{k \in \mathbb{Z}^n} \left\langle f_\theta, \frac{\overline{\psi}_\theta(c - (A_\theta^*)^k)}{|\theta|} \right\rangle \hat{\psi}_\theta(x - (A_\theta^*)^k).$$

We relabel the sum by the family $T(\theta)$, mapping $k \in \mathbb{Z}^n$ to the tube $\theta^* + (A_\theta^{-1})^k$, so that

$$f_\theta = c \sum_{T \in T(\theta)} a_T \psi_T,$$

with $a_T := \langle f_\theta, \overline{\psi}_T \rangle[T].$

The $\psi_T$ are wave packets adapted to the $T$ are easily seen to satisfy the follow properties:

i) ($\theta$ frequency localised) $\text{supp} \hat{\psi}_T \subseteq \theta$ for all $T \in T(\theta)$;

ii) (Almost disjoint spatial concentration) If $T \subseteq T(\theta)$ is any sub-collection of rectangles, then

$$\| \sum_{T \in T} a_T \psi_T \|_{L^p(\mathbb{R}^n)} \sim (\sum_{T \in T} |a_T|^2 |T|^{1/p})^{1/p}$$

$^{3}$In fact, the bound holds with an additional factor of $R^{-(n-m)/4}$, but the strength of the forthcoming estimates obviate the need for this additional decay.
for all $1 \leq p < \infty$, and the identity extends to the $p - \infty$ case in the obvious manner.

We will demonstrate why the right-hand term is dominated by the left in ii). Fixing $1 \leq p < \infty$ we observe that, by Hölder inequality,

$$|a_T|^p \leq |T|^p \int_{\mathbb{R}^n} |f_\theta(x)||\psi_T(x)| \, dx \left( \int_{\mathbb{R}^n} |\psi_T(x)| \, dx \right)^{p/p'}.$$ 

The bracketed term is $\sim 1$ and, therefore, given $T \subset T(\theta)$ it follows that

$$\left( \sum_{T \in \mathcal{T}} (|a_T||T|^{-1})^p |T| \right)^{1/p} \leq \left( \int_{\mathbb{R}^n} |f_\theta(x)||T| \sum_{T \in \mathcal{T}} |\psi_T(x)| \, dx \right)^{1/p} \leq \|f_\theta\|_{L^p(\mathbb{R}^n)}.$$ 

Here, in the last inequality, we have used the pointwise bound $\sum_{T \in \mathcal{T}} |\psi_T| \leq |T|^{-1}$ which is a consequence of the almost disjointness of the tubes in $T(\theta)$ and the rapid decay of $\psi_T$. For the remaining case when $p - \infty$ we simply note that $\langle f_\theta, \psi_T \rangle \leq \|f_\theta\|_{L^\infty(\mathbb{R}^n)}$.

We now perform a decomposition of $f$ according to the amplitudes of the various wave packets. In particular, for $k \in \mathbb{Z}$ define

$$T_k(\theta) := \{ T \in T(\theta) : |a_T||T|^{-1} \sim 2^k \}$$

and note $f = \sum_{k \in \mathbb{Z}} \tilde{f}_k$ where

$$\tilde{f}_k := \sum_{\theta : R^{1/2}-slab} \sum_{T_k(\theta)} a_T \psi_T.$$ 

We further decompose $f$ by writing $f = \sum_{k \in \mathbb{Z}} \sum_{t \in \mathbb{N}_0} f_{k,t}$ where

$$f_{k,t} := \sum_{\theta : \#T_k(\theta) \sim 2^t} a_T \psi_T.$$ 

The $f_k$ in the statement of Lemma 5 will be taken from these $f_{k,t}$; in particular, we notice they satisfy all the properties required by the decomposition.

**Lemma 7.** The $f_{k,t}$ defined above satisfy the following:

a) $\text{supp } \tilde{f}_k \subseteq \tau_j$;

b) $\|f_k\|_{L^p,n^i(\mathbb{R}^n)} \leq \|f\|_{L^p,n^i(\mathbb{R}^n)}$;

c) $\|f_k\|_{L^{p_0,n^{-i}(\mathbb{R}^n)}} \leq \|f_{k,t}\|_{L^{p_1,n^{-i}(\mathbb{R}^n)}} \leq \|f_{k,t}\|_{L^{p_2,n^{-i}(\mathbb{R}^n)}}$ for all $1 \leq p, p_0, p_1, p_2 \leq \infty$ with $1/p - (1 - \alpha)/p_0 + \alpha/p_1$ for some $0 \leq \alpha \leq 1$.

**Proof.** By property i) of the wave packets, the $f_{k,t}$ clearly have the desired Fourier support and, since the slabs are disjoint, for any $R^{1/2}$-slab $\theta$ we also have

$$f_{k,t,\theta} = \begin{cases} \sum_{T_k(\theta)} a_T \psi_T & \text{if } \#T_k(\theta) \sim 2^t \\ 0 & \text{otherwise} \end{cases}.$$ 

A similar formula holds for $f$ and consequently, by applying property ii), we deduce

$$\|f_{k,t}\|_{L^p(\mathbb{R}^n)} \leq \left( \sum_{T_k(\theta)} |a_T||T|^{-p+1} \right)^{1/p} \leq \left( \sum_{T \in \mathcal{T}(\theta)} |a_T||T|^{-p+1} \right)^{1/p} \sim \|f_\theta\|_{L^p(\mathbb{R}^n)}$$

and (9) immediately follows. Concerning the reverse Hölder inequality, by again applying property ii) we have

$$\|f_{k,t}\|_{L^{p,n^{-i}(\mathbb{R}^n)}} \sim \left( \sum_{\theta : \#T_k(\theta) \sim 2^t} \|f_{k,t,\theta}\|_{L^p(\mathbb{R}^n)}^2 \right)^{1/2} \sim 2^k (2^t|T|)^{1/p}.$$ 

\[4\text{In the language of Bourgain-Demeter, } \tilde{f}_k \text{ is an } R\text{-function and } f_{k,t} \text{ is a balanced } R\text{-function (see [1, §3]). We mention this only to aid comparison between these notes and the original paper [1] and, in particular, will not make use of the } R\text{-balanced terminology here.}\]
where in the right-most expression $|T| = R^{(n+1)/2}$ is the common measure of any $T \in \mathcal{T}(\theta)$. It is therefore immediate that
$$
\|f_k, t\|_{L^p, n-1(R^n)} \leq \|f_k, t\|_{L^{2p}, n-1(R^n)} \|f_k, t\|_{L^{12}, n-1(R^n)}
$$
holds for the relevant values of $p, p_0, p_1$ and $\alpha$.

To conclude the proof of Lemma 5 we will show $f$ is well-approximated by a sum of a very number of the $f_{k, t}$. The first observation is that for any fixed $k$ there are a very small number of non-zero $f_{k, t}$.

**Lemma 8.** For $f$ satisfying the hypotheses of Lemma 5 and $k \in \mathbb{Z}$ we have
$$
\# \bigcup_{\theta \in R^{(n-2)/2}-\text{slab}} \mathcal{T}_k(\theta) \lesssim 2^{-2k}|T|^{-1}. \tag{11}
$$
Consequently, for any $k \in \mathbb{Z}$ there are at most $O(\max(-k, 1))$ non-zero $f_{k, t}$.

**Proof.** By similar arguments to those applied above one observes
$$
\|\tilde{f}_{k, \theta}\|_{L^2(R^n)} \sim \left( \sum_{T \in \mathcal{T}_k(\theta)} (|\alpha_T||T|^{-1})^2 |T| \right)^{1/2} \sim 2^k (\# \mathcal{T}_k(\theta)|T|)^{1/2}
$$
from which it follows
$$
\|\tilde{f}_{k}\|_{L^2, n-1(R^n)} \lesssim 2^{2k}|T| \left( \sum_{\theta \in R^{(n-2)/2}-\text{slab}} \mathcal{T}_k(\theta) \right).
$$
Inequality (11) is now easily seen to be a consequence of the readily deduced estimate
$$
\|\tilde{f}_{k}\|_{L^2, n-1(R^n)} \lesssim \|f\|_{L^2, n-1(R^n)} - 1
$$
whilst the second part of the lemma is immediate (recall, the $f_{k, t}$ are formed by dyadically decomposing the sets $\mathcal{T}_k(\theta)$ according to cardinality).

**Proof (of Lemma 5).** Observe by the $p \to \infty$ case of Property ii) that
$$
|\alpha_T||T|^{-1} \lesssim \|f_0\|_{L^\infty(R^n)} \lesssim \|f\|_{L^\infty, n-1(R^n)} \lesssim R^{-|n+1)/4|} \|f\|_{L^2, n-1(R^n)}
$$
where the final estimate is due to Bernstein’s inequality. Consequently, by the $L^2$-normalisation of $f$ it follows $\tilde{f}_k = 0$ whenever $k > K_1$ where $K_1 := \log(CR^{-|n+1)/4|})$. For $K_0 := K_1 - n^2 \log R$ write\(^5\)
$$
f = \sum_{k=K_0}^{K_1} \tilde{f}_k + e - \sum_{k=K_0}^{K_1} \sum_{t \in \mathcal{T}_k(\theta)} f_{k, t} + e
$$
and let $\{f_k\}_{k=K_0}^{K_1}$ be an enumeration of the non-zero $f_{k, t}$ appearing in this decomposition, noting Lemma 8 forces $N \lesssim 1$. Finally, the usual arguments show
$$
\|e\|_{L^\infty(R^n)} \sim \sup_{k < K_0} \sup_{T \in \mathcal{T}_k(\theta)} |\alpha_T| \lesssim R^{-n^2} R^{-|n+1)/4|} \|f\|_{L^2, n-1(R^n)},
$$
where the inequality is due to the definition of $K_0$. Now, taking the square sum over all the $R^{-1/2}$ slabs (and recalling there are $\sim R^{-|n-1)/2}$ of them),
$$
\|e\|_{L^2, n-1(w_{n, R})} \lesssim \|e\|_{L^2, n-1(R^n)} \lesssim R^{-n^2-1/2} \|f\|_{L^2, n-1(R^n)}
$$
and local orthogonality concludes the proof.\(^6\)

\(^5\)Here, by slightly abusing notation, the sum in $k$ is understood to be over $[K_0, K_1] \cap \mathbb{Z}$.
\(^6\)It is important in the proof of Proposition 4 to have a power of $R^{-1}$ which is quadratic in $n$ appear on the right-hand side of the error bound. In the present argument we have ignored any larger factors involving linear powers of $R^{-1}$ (such as those arising from $|T|^{-1}$ in Lemma 8 and from Bernstein’s inequality), by bounding them by 1. This is not wasteful: for our purposes this additional decay is essentially irrelevant.
Remark 9. It is instructive to note there is a simple analogue of Lemma 5 for the classical Hölder inequality. In particular, given any \( f \) with (uniformly) bounded Fourier support, one may write \( f = \sum_{j=1}^{N} f_{j} + e \) where \( N \leq 1 \), \( e \) is a small (in some local \( L^{p} \)-sense) error and the \( f_{j} \) each satisfy a reverse Hölder inequality for \( L^{p} \)-norms over balls \( B_{R} \). We refer the reader to [2] for further details. The proof of this analogue follows the same general scheme as detailed above but relies on analysing the atomic decomposition

\[
f = \sum_{k \in \mathbb{Z}} f_{k} \chi_{|n|^{-2k}}
\]
of \( f \), rather than the wave-packet decomposition.

4. Concluding the argument

We are now in a position to simultaneously prove the full linear and multi-linear decoupling theorems, following the arguments of [1].\(^7\) To begin we apply the interpolation theory for decoupling norms developed in previous section. This reduce the problem to considering the inequalities at the critical exponent

\[
p_{c} := \frac{2(n + 1)}{n - 1}
\]
only.

Proposition 10. To prove the linear decoupling theorem in \( n \)-dimensions it suffices to show

\[
\|f\|_{L^{p,r}_{c}(B_{R})} \leq \|f\|_{L^{p,r,n-1}_{c}(w_{B_{R}})}
\]
holds for all \( f \) with Fourier support in \( \mathcal{N}_{R^{-1}}(\mathcal{P}^{n-1}) \).

Proof. Recall, the inequality (12) implies the global version

\[
\|f\|_{L^{p,r}(\mathbb{R}^{n})} \leq \|f\|_{L^{p,c,n-1}(\mathbb{R}^{n})}
\]
via the usual arguments.

The cases \( p - 2 \) and \( p - \infty \) of the decoupling theorem are immediate (with \( R \)-uniform constants), the former a consequence of the local orthogonality estimate whilst the latter is simply due to the Cauchy-Schwarz inequality. Explicitly, for \( f \) satisfying the hypothesis of the proposition we have

\[
\|f\|_{L^{2}(\mathbb{R}^{n})} \leq \|f\|_{L^{2,n-1}(\mathbb{R}^{n})} \quad \text{and} \quad \|f\|_{L^{\infty}(\mathbb{R}^{n})} \leq R^{(n-1)/2}\|f\|_{L^{\infty,n-1}(\mathbb{R}^{n})}.
\]
Now, fixing such an \( f \) we apply the decomposition lemma to write \( f = \sum_{j=1}^{N} f_{j} + e \) where \( \text{supp } f_{j} \subseteq \mathcal{N}_{R^{-1}}(\mathcal{P}^{n-1}) \); \( \|f_{j}\|_{L^{p,c,n-1}(\mathbb{R}^{n})} \leq \|f\|_{L^{p,c,n-1}(\mathbb{R}^{n})} \); for each \( f_{j} \) the reverse Hölder inequality (for decoupling norms) holds; \( e \) is an error term satisfying

\[
\|e\|_{L^{p,c,n-1}(\mathbb{R}^{n})} \leq R^{-\gamma}\|f\|_{L^{p,c,n-1}(\mathbb{R}^{n})}
\]
and \( N \leq 1 \). If \( p_{c} < p < \infty \) and \( 1 \leq j \leq N \), then

\[
\|f_{j}\|_{L^{p}(w_{B_{R}})} \leq \|f_{j}\|_{L^{p,c}(\mathbb{R}^{n})} \|f_{j}\|_{L^{p}(\mathbb{R}^{n})} \leq R^{(n-1)/2}\|f_{j}\|_{L^{p,c,n-1}(\mathbb{R}^{n})} \|f_{j}\|_{L^{p,c,n-1}(\mathbb{R}^{n})}^{\gamma}
\]
where \( 0 < \gamma < 1 \) is defined by \( p(1 - \gamma) = p_{c} \). One readily deduces

\[
\gamma(n - 1) = n + 1 - 2p
\]
\(^7\)As in the previous sections, our presentation is indebted to the expository lecture notes of Guth [2].
which is precisely the exponent featuring in the statement of the decoupling conjecture. By the properties of the decomposition,

$$
\|f_j\|_{L^p_{\mu,R}(\mathbb{R}^n)}^{1/n} \leq \|f_j\|_{L^{p,n-1}(\mathbb{R}^n)}
$$

and so, by the triangle inequality

$$
\|f\|_{L^p(\omega w_{B_R})} \leq R^{\frac{n-1}{p} + \frac{1}{n}} \|f\|_{L^{p,n-1}(\mathbb{R}^n)} + \|f\|_{L^p(\omega w_{B_R})}.
$$

To estimate the error, simply note

$$
\|e\|_{L^p(\omega w_{B_R})} \leq R^n \|e\|_{L^p(\mathbb{R}^n)}
$$

and apply the error bound.

One may apply a similar argument to show the decoupling inequality also holds for $2 \leq p < p_c$. □

It remains to prove (12), which is achieved by inducting on dimension. Recall the linear and multi-linear decoupling inequalities have already been established in dimension 2 and so we already have the base case for this induction.

**Definition 11.** For $2 \leq s$ and $R \geq 1$ let $\mathcal{D}_{s,n}(R)$ and $\mathcal{M}_{s,n}(R,\omega)$ denote the best constants in the $n$-dimensional linear and multi-linear decoupling inequalities, respectively.

For example, $\mathcal{M}_{s,n}(R,\omega)$ is the infimum over all $C$ for which, given $\omega$-transverse $1$-slabs $\tau_1, \ldots, \tau_n$, the inequality

$$
\left\| \prod_{j=1}^n |f_j|^{1/n} \right\|_{L^{\omega,n}(B_R)} \leq C \prod_{j=1}^n \|f_j\|_{L^{p,n-1}(\omega w_{B_R})}^{1/n}
$$

holds for all $f_j$ with $\mathrm{supp} \hat{f}_j \subseteq \tau_j$.

For $2 \leq s \leq 2(n+1)/(n-1)$ the linear decoupling theorem is the statement $\mathcal{D}_{s,n}(R) \leq 1$ for all $n \in \mathbb{N}$. We will fix $n \geq 3$ and assume, as an induction hypothesis, that

$$
\mathcal{D}_{s,n-1}(R) \leq 1.
$$

for all $2 \leq s \leq 2n/(n-2)$ (or, equivalently, the estimate holds for $s = 2n/(n-2)$ only).

**Remark 12.** Recall, by Hölder's inequality this implies $\mathcal{M}_{s,n-1}(R,\omega) \leq 1$ for $s$ belonging to the same range.

We may now invoke the proposition, proved in earlier notes, which essentially reduces the problem to considering multi-linear estimates. In particular, for any $0 < \omega \ll 1$, one has

$$
\mathcal{D}_{s,n}(R) \leq R^{\varepsilon(\omega)} \mathcal{M}_{s,n}(R,\omega)
$$

for all $R \geq 1$ and all $s$ as above, where $\varepsilon(\omega) \to 0$ as $\omega \to 0$. In particular, (13) holds for $s = p_c < 2n/(n-2)$ and it therefore suffices to show $\mathcal{M}_{p_c,n}(R,\omega) \leq 1$ for all sufficiently small $\omega$. Earlier we saw from the trivial decoupling lemma that $\mathcal{D}_{p_c,n}(R,\omega)$ grows at most polynomially and therefore the same is true for $\mathcal{M}_{p_c,n}(R,\omega)$. Hence,

$$
\mathcal{M}_{p_c,n}(R,\omega) \approx \omega^\gamma
$$

for some $\gamma \geq 0$ and it suffices to show $\gamma = 0$.

In the previous section we stated and proved the inductive step relevant to multi-linear decoupling inequalities:
Proposition 13 (Inductive step - Decoupling). Let \( \tau_1, \ldots, \tau_n \) be \( \omega \)-transverse \( \sim 1 \)-caps on \( \mathbb{R}^{n-1} \), \( B_{R} \) a fixed \( R \)-ball and \( \mathcal{B} \) a covering of \( B_{R} \) by boundedly-overlapping \( R^{1/2} \)-balls. For \( 2n/(n-1) \leq s < 1 + \beta := 2/(s-2)(n-1) \) the inequality

\[
\left\| \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R_1/2}}^\omega)} \right\|_{L^2(w_{B_{R}}^\omega)(\mathcal{B})} \leq \omega \left( \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R_1}}^\omega)} \right) \left\| \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R}}^\omega)} \right\|_{L^2(w_{B_{R}}^\omega)(\mathcal{B})}
\]

holds for \( f_j \) with Fourier support in \( \tau_j \), \( 1 \leq j \leq n \).

Remark 14. This is, in fact, a localised version of the inequality we proved earlier in the notes, which follows from the latter through standard arguments.

Henceforth fix \( 2n/(n-1) \leq s \leq 2(n+1)/(n-1) \); we'll show \( \mathfrak{U}_{s,n}(R, \omega) \leq 1 \).

Although we only need to prove the estimate for \( s = p_c \), it is instructive to consider more general exponents in order to observe how the argument breaks down as \( s \) approaches the critical exponent (or, equivalently, as \( \beta \) approaches 1/2). To prove uses an argument similar to that encountered in the context of the multi-linear restriction, although from the outset there are notable differences in the set up which will manifest in additional complications.\(^8\)

Fix \( R > 1 \) and apply the trivial version of the multi-linear restriction theorem to obtain

\[
\left\| \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R}}^\omega)} \right\|_{L^2(w_{B_{R}}^\omega)(\mathcal{B})} \leq r_0^{n/2} \left( \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R_1}}^\omega)} \right) \left\| \prod_{j=1}^{n} \left\| f_j \right\|_{L^2(w_{B_{R}}^\omega)} \right\|_{L^2(w_{B_{R}}^\omega)(\mathcal{B})}
\]

where \( r_0 \ll R \) and \( \mathcal{B}_{r_0} \) is boundedly-overlapping cover of \( B_{R} \) by \( r_0 \)-balls. The advantages of the trivial estimate are that it works for a wide range of exponents and it is flexible enough to be applied between vastly separated scales; in particular, we take \( r_0 := R^{1/2^N} \) for some large integer \( N \). We will now work up to the scale \( \bar{R} \) by passing through the intermediate scales

\[ r_{l+1} := r_l^{1/2} = R^{1/2^N-l} \]

for \( 0 \leq l \leq N \), applying the inductive step at each stage. To do this we will have to work with many collections of balls at different scales.

Definition 15. Let \( \mathcal{B}_{r_{N}} := \{ B_{R} \} \) and for each \( 0 \leq l \leq N-1 \) let \( \mathcal{B}_{r_{l+1}} \) be a boundedly-overlapping cover of \( B_{R} \) by \( r_{l+1} \)-balls. For any \( B_{r_{l+1}} \in \mathcal{B}_{r_{l+1}} \) define

\[ \mathcal{B}_{r_{l}}(B_{r_{l+1}}) := \{ B_{r_{l}} \in \mathcal{B}_{r_{l}} : B_{r_{l}} \cap B_{r_{l+1}} \neq \emptyset \}; \]

we assume, without loss of generality, that each of these sets is a covering for the \( B_{r_{l+1}} \).

We will use Proposition 13 to prove the following estimate, which is amenable to iterative application.

\(^8\)In particular:

- If we compare the inequalities featuring in the inductive steps for restriction and decoupling, then we see the former bounds \( L^2 \) (norms) by \( L^2 \), whilst the latter controls \( L^2 \) by a mix of \( L^2 \) and \( L^3 \). The appearance of \( L^3 \) in the decoupling case suggests the estimate seems less suited for iterative application.
- The multi-linear restriction estimate is \( L^2 = L^2 \) and so the domain matches up with the \( L^2 \)-expressions on the right-hand side of inductive step. In the decoupling case, the desired estimate is \( L^2 = L^4 \) and so it doesn’t fit so well with its corresponding inductive step.

As a result of these features, the argument used to prove multi-linear restriction does not carry over mutatis mutandis and some addition work is required (but not too much!).
Lemma 16. For each $0 \leq l \leq N - 1$ we have

$$\left\| \prod_{j=1}^{n} f_j \right\|_{L_{\text{avg}}^2(B_{r_{l+1}})}^{1/n} \lesssim \omega \mathcal{D}_{s,n}(R/r_{l+1})^{\beta} \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^2(w_{B_{r_{l+1}}})}^{1-\beta/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}^{\beta/n}.$$  

(17)

This lemma is very similar to Proposition 13, allowing us to pass from a small scale $r_l$ to a larger scale $r_{l+1}$. However, here there is a key difference in that, no matter which scale $r_l$ we are working with, the decoupling norms are always defined with respect to the small $R^{-1/2}$-slabs.

Proof (of Lemma 16). The left-hand side of the inequality in the statement of the lemma is comparable to

$$\left\| \prod_{j=1}^{n} f_j \right\|_{L_{\text{avg}}^2(w_{B_{r_{l+1}}})}^{1/n} \left\| f_j \right\|_{L_{\text{avg}}^2(B_{r_{l+1}})}^{1-\beta/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}^{\beta/n}.$$  

Applying Proposition 13 to pass from scale $r_l$ to scale $r_{l+1} - r_l^2$, we deduce

$$\left\| \prod_{j=1}^{n} f_j \right\|_{L_{\text{avg}}^2(w_{B_{r_{l+1}}})} \lesssim \omega \left\| \prod_{j=1}^{n} f_j \right\|_{L_{\text{avg}}^2(B_{r_{l+1}})}^{1-\beta/n} \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}^{\beta/n}.$$  

To isolate the $L^2$ contribution we apply the $(n + 1)$-linear Hölder inequality with exponents arising from the identity

$$\frac{1}{s} = \frac{1}{s} - \frac{\beta}{s} + \frac{\beta}{ns} + \cdots + \frac{\beta}{ns},$$

to bound the right-hand side of the above expression by

$$\left\| \prod_{j=1}^{n} f_j \right\|_{L_{\text{avg}}^2(B_{r_{l+1}})}^{1/n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}^{\beta/n}.$$  

The $L^2$ factor is now of the desired form. For the $L^2$ contribution, it is remarked that we have succeeded in decoupling the function to relatively large $r^{1/2}_{l+1}$-slabs, which represents some small progress towards full decoupling (i.e. to the small $R^{-1/2}$-slabs).

To conclude the proof of the lemma observe, as a simple consequence of Minkowski’s inequality,

$$\left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})} \lesssim \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}.$$  

Now we have a norm which decouples to (in general, relatively large) $r^{1/2}_{l+1}$-slabs, but is localised to scale $R$: this is precisely the situation to which the parabolic rescaling lemma of the earlier notes applies. Thus, by parabolic rescaling,

$$\left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})} \lesssim \mathcal{D}_{s,n}(R/r_{l+1}) \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_{r_{l+1}}})}.$$  

and the desired inequality follows by combining these observations.

We are now ready to prove the main theorem.

Proof (of the $L^2$-decoupling theorem for subcritical exponents). Recall, it suffices to show the exponent $\gamma$ defined above equals 0.

Starting with (16), apply Lemma 16 iteratively to conclude

$$\prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(B_R)} \lesssim \omega A_N(R) \left( \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_R})} \right)^{\beta N} \left[ \prod_{j=1}^{n} \left\| f_j \right\|_{L_{\text{avg}}^n(w_{B_R})} \right]^{1-\beta N}.$$  

(18)
where $\beta_l := (1 - \beta)l$ and
\[
A_N(R) := r_0^{n/2} \left[ \prod_{l=0}^{N-1} D_{s,n}(R/r_{l+1})^{\beta_l} \right]^{\beta}.
\]
Note the implied constant in (18) is independent of $N$. Moreover, for fixed $\varepsilon > 0$ we can take the constant to be $C_{\varepsilon, \beta, \omega}$ whenever (17) holds with $\leq \omega$ replaced with $\leq C_{\varepsilon, \beta, \omega} R^\beta$.

By local orthogonality we deduce
\[
\| f \|_{L^2_{w_\varepsilon}(w_{nR})} \leq \| f \|_{L^1_{w_\varepsilon}(w_{nR})} \leq \| f \|_{L^1_{w_\varepsilon}(w_{nR})}
\]
and therefore
\[
\| \prod_{j=1}^{n} |f_j|^{1/n} \|_{L^2_{w_\varepsilon}(B_R)} \leq \omega A_N(R) \prod_{j=1}^{n} \| f_j \|_{L^2_{w_\varepsilon}(w_{nR})}^{1/n}.
\]
Consequently we have an estimate
\[
\mathfrak{M}_{s,n}(R, \omega) \leq \omega A_N(R).
\]
Given $\varepsilon > 0$, by choosing $\omega$ to be sufficiently small (recalling (13)) and applying the definition of the $\tau_l$, it follows
\[
\mathfrak{M}_{s,n}(R, \omega) \leq \varepsilon \omega R^{\gamma N/2^{N+1}} \left[ \prod_{l=1}^{N} R^n_{\gamma (1 - 2^{N-l})} \beta_{l-1} \right]^{\beta}.
\]
We have therefore manage to compare $\mathfrak{M}_{s,n}(R, \omega)$ with the best constants at many smaller scales. Here it is convenient to recall (14) and thereby deduce
\[
R^\gamma \leq \varepsilon \omega R^{\gamma N/2^{N+1}} \left[ \prod_{l=1}^{N} R^n_{\gamma (1 - 2^{N-l})} \beta_{l-1} \right]^{\beta}.
\]
Now, estimating the exponent arising from the bracketed term we see
\[
\beta \sum_{l=1}^{N} \left( 1 - 1/2^{N-l} \right) \beta_{l-1} - 1 - \beta_N - \beta \sum_{l=1}^{N} \beta_{l-1}/2^{N-l} \leq 1 - 2\beta N/2^N
\]
where the inequality is (partly) due to the fact that the condition $s \leq 2(n+1)/(n-1)$ ensures $1/2 \leq \beta$. Hence
\[
R^{2\beta \gamma N/2^N} \leq \varepsilon \omega R^{\epsilon n/2^{N+1}}
\]
and, since this inequality holds for all $\varepsilon > 0$ and all $R \gg 1$, it follows
\[
\gamma N \leq 2\beta \gamma N \leq n/2
\]
for all $N \in \mathbb{N}$. Thus $\gamma = 0$, concluding the proof.

\[
\square
\]

References
\[\begin{align*}
\end{align*}\]

\[^{9}\text{Recall, this was the estimate used to derive multi-linear decoupling from multi-linear restriction in the } 2 \leq p \leq 2n/(n-1) \text{ range.}\]