JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 15, Number 2, Pages 469–496
S 0894-0347(01)00379-4
Article electronically published on December 10, 2001

MULTI-LINEAR OPERATORS GIVEN BY SINGULAR MULTIPLIERS

CAMIL MUSCALU, TERENCE TAO, AND CHRISTOPH THIELE

1. Introduction

Let $n > 1$ be an integer, and let $m(\xi_1, \ldots, \xi_n)$ be a function on the $(n-1)$-dimensional vector space

$$\Gamma = \{\xi \in \mathbb{R}^n : \xi_1 + \ldots + \xi_n = 0\}.$$

For any m, we associate the multi-linear operator $T = T_m$ on $n-1$ functions on $\mathbb{R}$ by

$$\begin{aligned} & T_m(f_1, \ldots, f_{n-1})\hat{\ }(-\xi_n) \\ & \qquad = \int \delta(\xi_1 + \ldots + \xi_n) m(\xi) \hat{f}_1(\xi_1) \ldots \hat{f}_{n-1}(\xi_{n-1})\, d\xi_1 \ldots d\xi_{n-1}, \end{aligned} \tag{1}$$

where $\xi = (\xi_1, \ldots, \xi_n)$. We may write this operator more symmetrically as an n-linear form $\Lambda = \Lambda_m$ given by

$$\Lambda_m(f_1, \ldots, f_n) = \int \delta(\xi_1 + \ldots + \xi_n) m(\xi) \hat{f}_1(\xi_1) \ldots \hat{f}_n(\xi_n)\, d\xi;$$

the relationship between T and Λ is given by

$$\Lambda(f_1, \ldots, f_n) = \int T(f_1, \ldots, f_{n-1})(x) f_n(x)\, dx. \tag{2}$$

When $n = 2$, T is a Fourier multiplier, and it is well known that such operators are bounded on L^p, $1 < p < \infty$, if m is a symbol of order 0. Coifman and Meyer [2]–[7], Kenig and Stein [14], and Grafakos and Torres [12] extended this result to the $n > 2$ case, showing that one had the mapping properties

$$T : L^{p_1} \times \ldots \times L^{p_{n-1}} \to L^{p_n'} \tag{3}$$

whenever

$$1 < p_i \leq \infty \tag{4}$$

for $i = 1, \ldots, n-1$,

$$1/(n-1) < p_n' < \infty, \tag{5}$$

Received by the editors November 30, 1999 and, in revised form, May 31, 2001.
1991 *Mathematics Subject Classification.* Primary 42A45, 47H60; Secondary 45P05.
Key words and phrases. Fourier analysis, multi-linear operators.
The second author was supported by NSF Grant #9706764.
The third author was supported by NSF Grant #9970469.

and

$$\frac{1}{p_1} + \ldots + \frac{1}{p_n} = 1, \tag{6}$$

and m satisfies the symbol estimates

$$|\partial_\xi^\alpha m(\xi)| \lesssim |\xi|^{-|\alpha|} \tag{7}$$

for all partial derivatives ∂_ξ^α on Γ up to some finite order. We interpret estimate (3) in the way that T is originally defined on the product of suitable subspaces of the L^{p_i} and then extends to the product of the closures of these subspaces. In case $p_i \neq \infty$ the subspace is simply the test function space which is dense in L^{p_i}. If any estimate of the type (3) holds with $p_i \neq \infty$ for all i, then we can use this to unambiguously define T on the product of $n-1$ copies of $L^1 \cap L^\infty$. Once this is done, we can choose $\overline{L^1 \cap L^\infty}$ as a subspace of L^∞ whenever $p_i = \infty$ in some other exponent tuple. If $p_n' \geq 1$, we can use a duality argument to extend the operator from $\overline{L^1 \cap L^\infty}$ to L^∞. If $p_n' < 1$ we shall be satisfied with replacing L^∞ by $\overline{L^1 \cap L^\infty}$ in (3) where applicable.

The interesting observation that p_n' can be smaller than or equal to 1 traces back (at least) to papers by C. Calderon [1] and Coifman and Meyer [2], where special multi-linear operators are discussed.

When m is identically one, then T_m is the pointwise product operator

$$T(f_1, \ldots, f_{n-1}) = f_1 \ldots f_{n-1},$$

so estimate (3) may be viewed as a generalization of Hölder's inequality, where products are replaced by paraproducts.

The bilinear Hilbert transform

$$T(f_1, f_2) = \int f_1(x-t) f_2(x+t) \frac{dt}{t}$$

can also be viewed as an operator of the form (1), with symbol

$$m(\xi_1, \xi_2, \xi_3) = \pi i \operatorname{sgn}(\xi_2 - \xi_1).$$

This multiplier does not satisfy the estimates (7). Nevertheless, Lacey and Thiele [15], [16] showed that (3) continues to hold, provided that one makes the additional assumption $p_3' > 2/3$.

The purpose of this paper is to unify these results, allowing us to prove (3) for a class of multipliers which are singular on a subspace of Γ. More precisely, we have

Theorem 1.1. *Let Γ' be a subspace of Γ of dimension k where*

$$0 \leq k < n/2. \tag{8}$$

Assume that Γ' is non-degenerate in the sense that for every $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, the space Γ' is a graph over the variables $\xi_{i_1}, \ldots, \xi_{i_k}$. Suppose that m satisfies the estimates

$$|\partial_\xi^\alpha m(\xi)| \lesssim \operatorname{dist}(\xi, \Gamma')^{-|\alpha|} \tag{9}$$

for all partial derivatives ∂_ξ^α on Γ up to some finite order. Then (3) *holds whenever* (4), (5), (6) *hold and*

$$\frac{1}{p_{i_1}} + \ldots + \frac{1}{p_{i_r}} < \frac{n-2k+r}{2} \tag{10}$$

for all $1 \leq i_1 < \ldots < i_r \leq n$ and $1 \leq r \leq n$.

In particular, (3) holds whenever $1 < p_i \leq \infty$ for $i = 1, \dots, n$ and (6) holds.

As discussed above the case $k = 0$ is well known. The case $n = 3$, $k = 1$ follows from the work by Gilbert and Nahmod [9], [10], [11]. Unfortunately Theorem 1.1 does not quite cover the trilinear Hilbert transform

$$T(f_1, f_2, f_3) = \int f_1(x-t) f_2(x+t) f_3(x+2t) \frac{dt}{t}$$

since one has $n = 4$, $k = 2$ in this case, which does not satisfy (8). To obtain the analogue of this theorem when (8) fails would probably require radically different techniques than the ones developed to date. However, an elementary argument can be used to handle this case if enough functions are in the Wiener algebra $A = \mathcal{F}^{-1} L^1$; see Section 13.

In the $k = 0$ case the origin $\xi = 0$ has special significance, and this is reflected in the tools used to handle this case, namely Littlewood-Paley theory and/or wavelets. However, when $k > 0$ there is no preferred frequency origin, and the tools used should be invariant under frequency translations along Γ'. This necessitates the employment of "tiles" in the time-frequency plane which have arbitrary frequency location, spatial location, and scale.

If the multiplier m of Theorem 1.1 is invariant under translations in the direction of Γ', then we can write the $(n-1)$-linear operator T as

$$T(f_1, \dots, f_{n-1})(x) = \int_{\Gamma'' \cap \Gamma} f_1(x + \gamma_1) \dots f_{n-1}(x + \gamma_{n-1}) K(\gamma)\, d\gamma,$$

where Γ'' is the orthogonal complement of Γ', $d\gamma$ is Lebesgue measure on $\Gamma'' \cap \Gamma$, γ_i is the i-th coordinate of γ as an element of $\mathbb{R}^n$, and K is a Calderon-Zygmund kernel on the space $\Gamma'' \cap \Gamma$. Thus we obtain L^p bounds for such operators provided $n - 1 \leq 2d$, where d is the dimension of $\Gamma'' \cap \Gamma$. This gives a partial answer to question (2) in [14] raised by Kenig and Stein.

It would be interesting to study the behavior of the bounds in (3) as the space Γ' degenerates in the sense of Theorem 1.1; see [18] for some results in this direction in the special case of the bilinear Hilbert transform. We do not discuss this issue here.

This paper is organized as follows. In Section 3 we introduce some multi-linear interpolation theory, which allows us to reduce (3) to a "restricted type" estimate on the n-form Λ. In Section 4 we remove an exceptional set, and reduce matters to estimating Λ on functions whose Hardy-Littlewood maximal function is under control. In Section 5 we then decompose the multiplier m using a Whitney decomposition, which allows us to replace Λ by a discretized analogue which involves the size of the f_i on various tiles in the time-frequency plane; roughly speaking, we only need consider those tiles that lie outside the exceptional set. To handle these tiles we first consider the case $k = 1$. This is done by subdividing the tiles into essentially disjoint trees, using Littlewood-Paley theory to estimate the contribution of each tree individually, and then using orthogonality arguments to control the total number of trees. Finally, in Section 12, we induct on k to obtain the general case.

The first author wishes to express his gratitude to the UCLA Mathematics Department for its hospitality during his visit to Los Angeles and to Jill Pipher for her financial and moral support. The second author thanks Michael Lacey and Jim Wright for many helpful discussions during a delightful semester at the University of New South Wales. The third author acknowledges an enjoyable stay at the Erwin Schrödinger Institute in Vienna, during which part of this work was done.

2. Preliminaries

We use $A \lesssim B$ to denote the statement that $A \leq CB$ for some large constant C, and $A \ll B$ to denote the statement that $A \leq C^{-1}B$ for some large constant C.

Throughout the paper we shall assume $k > 0$, the case $k = 0$ being well known.

We make the a priori assumption that the symbol m is smooth and compactly supported; this makes Λ bounded on every product of n Lebesgue spaces. Our final estimates will not depend on the smoothness or support bounds that m satisfies, and the general case can be handled by the usual limiting argument.

If I is an interval, then CI denotes the interval with the same center but C times the length. Let χ_I denote the characteristic function of I. We define the approximate cutoff function $\tilde{\chi}_I$ as

$$\tilde{\chi}_I(x) = (1 + \frac{\mathrm{dist}(x, I)}{|I|})^{-1}.$$

We use $\mathrm{M}f$ to denote the Hardy-Littlewood maximal function.

3. Interpolation

In this section we develop some multi-linear interpolation theory which allows us to reduce (3) to a certain "restricted type" estimate on Λ. We find it convenient to work with the quantity $\alpha_i = 1/p_i$ when p_i is the exponent of L^{p_i}. We fix n throughout this section.

Definition 3.1. A tuple $(\alpha_1, \ldots, \alpha_n)$ is called admissible if

$$-\infty < \alpha_i < 1 \tag{11}$$

for all $1 \leq i \leq n$,

$$\sum_i \alpha_i = 1, \tag{12}$$

and there is at most one index j such that $\alpha_j < 0$. We call an index i *good* if $\alpha_i \geq 0$, and we call it *bad* if $\alpha_i < 0$. A *good* tuple is an admissible tuple without bad index, and a *bad* tuple is an admissible tuple with a bad index.

Definition 3.2. Let E, E' be sets of finite measure. We say that E' is a *major subset* of E if $E' \subset E$ and $|E'| \geq \frac{1}{2}|E|$.

Definition 3.3. If E is a set of finite measure, we let $X(E)$ denote the space of all functions F supported on E such that $\|F\|_\infty \leq 1$.

Definition 3.4. If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an admissible tuple, we say that an n-linear form Λ is of restricted type α if for every sequence $E_1, \ldots, E_n$ of subsets of $\mathbb{R}$ with finite measure, there exists a major subset E'_j of E_j for each bad index j (one or none) such that

$$|\Lambda(F_1, \ldots, F_n)| \lesssim |E|^\alpha \tag{13}$$

for all functions $F_i \in X(E'_i)$, $i = 1, \ldots, n$, where we adopt the convention $E'_i = E_i$ for good indices i, and $|E|^\alpha$ is shorthand for

$$|E|^\alpha = |E_1|^{\alpha_1} \ldots |E_n|^{\alpha_n}.$$

The restricted type result we will prove directly is:

Theorem 3.5. *The form Λ as in Theorem 1.1 is of restricted type α for all bad tuples α such that*

$$1/2 < \alpha_i < 1 \tag{14}$$

for all all good indices i and

$$k - n + 2 > \alpha_j > k - n + \frac{3}{2} \tag{15}$$

for the bad index j.

Once Theorem 3.5 is granted, which we shall assume throughout this section, the issue of proving Theorem 1.1 is to pass to the (admissible part of the) convex hull of tuples described in Theorem 3.5 and convert restricted type estimates to strong type estimates.

We need the following easy lemma on permutahedrons:

Lemma 3.6. *Let $a_1 > \ldots > a_n$ be numbers. Then the convex hull of all permutations of $(a_1, \ldots, a_n)$ consists of those points $(x_1, \ldots, x_n)$ such that $x_1 + \ldots + x_n = a_1 + \ldots + a_n$ and $x_{i_1} + \ldots + x_{i_r} \leq a_1 + \ldots + a_r$ for all $1 \leq i_1 < \ldots < i_r \leq n$ and $1 \leq r \leq n$.*

Proof. It is clear that the convex hull belongs to the set described above. It thus suffices to show that the only extreme points of the above set are the permutations of $(a_1, \ldots, a_n)$.

Let $(x_1, \ldots, x_n)$ be an extreme point; by symmetry we may assume that $x_1 \geq \ldots \geq x_n$. If $x_1 + \ldots + x_r < a_1 + \ldots + a_r$ for some $1 \leq r < n$, then we may modify x by a small multiple of $e_r - e_{r+1}$ in either direction without leaving the set, contradicting the extremality of x. Thus $x_1 + \ldots + x_r = a_1 + \ldots + a_r$ for all r, so that $(x_1, \ldots, x_n) = (a_1, \ldots, a_n)$, as desired. $\square$

Now let P denote the set of all admissible tuples described by Theorem 3.5 and let Q denote the set of all admissible tuples α such that

$$\alpha_{i_1} + \cdots + \alpha_{i_r} < \frac{n - 2k + r}{2} \tag{16}$$

for all $1 \leq i_1 < \cdots < i_r \leq n$ and $1 \leq r \leq n$.

We have

Lemma 3.7. *The set Q is contained in the convex hull of P. The set Q contains all good tuples. If $\alpha \in Q$ has bad index j, then there is an $\tilde{\alpha} \in P$ with $\tilde{\alpha}_i > \alpha_i$ for all $i \neq j$ and such that α is in the convex hull of $\tilde{\alpha}$ and the elements in Q whose bad index is not equal to j.*

Proof. For the first statement it suffices to prove that all tuples satisfying (11), (12), and (16) are contained in the convex hull of $\overline{P}$. This in turn follows immediately from Lemma 3.6 and the observation that $\overline{P}$ contains all tuples which have $n - 2k$ elements equal to 1, $2k-1$ elements equal to $1/2$, and the remaining elements equal to $3/2 + k - n$.

The second statement follows immediately from the observation that the right-hand side of (16) is greater than or equal to one with strict inequality in case $r > 1$.

To see the third statement, assume by symmetry that $\alpha \in Q$ has bad index n. Define $\tilde{\alpha}_i = \max(\alpha_i, 1/2)$ if $i \neq n$ and $\tilde{\alpha}_n = 3/2 + k - n$. Then (16) shows that $\sum_i \tilde{\alpha}_i < 1$, so we can enlarge the entries of $\tilde{\alpha}$ so that $\tilde{\alpha} \in P$ and $\tilde{\alpha}_i > \alpha_i$ for $i \neq n$. We can write $\alpha = \theta\tilde{\alpha} + (1-\theta)\alpha'$, where α' is some tuple with $\alpha'_n = 1/2$. Then a similar application of Lemma 3.6 as before implies that α' is in the convex hull of those elements $\alpha'' \in \overline{P}$ for which $\alpha''_n = 1/2$. This implies the third statement of Lemma 3.7. □

We first discuss good exponent n-tuples:

Lemma 3.8. *Let the assumptions and notation be as in Theorem 1.1. Then Λ is of restricted type α for all good tuples α.*

Proof. Let α be a good tuple. By symmetry we can assume that $\alpha_n = \max_i \alpha_i$. Then we have $\alpha_n > 0$ and $\alpha_i \leq 1/2$ for all $i \neq n$. By Lemma 3.7 we find $\theta_j \geq 0$ such that

$$\alpha = \sum_{j=1}^{n} \theta_j \alpha^{(j)}, \quad \sum_{i=1}^{n} \theta_j = 1, \tag{17}$$

where $\alpha^{(j)} = (\alpha_i^{(j)})_{i=1}^n$ is an admissible tuple in P with bad index j. We can arrange that $\theta_n > 0$.

For $\lambda > 0$ let $A(\lambda)$ be the best constant such that

$$|\Lambda(F_1, \dots, F_n)| \leq A(\lambda)|E|^\alpha$$

for all sets $E_1, \dots, E_n$ of finite measure with

$$|E|^{\alpha^{(n)}} < \lambda |E|^\alpha \tag{18}$$

and functions $F_i \in X(E_i)$. Let $A(\infty)$ be the supremum of all $A(\lambda)$. By the a priori smoothness and support assumptions on m, $A(\infty)$ is finite and the point is to prove that it is bounded.

By splitting

$$|\Lambda(F_1, \dots, F_n)| \leq |\Lambda(F_1, \dots, F_n \chi_{E'_n})| + |\Lambda(F_1, \dots, F_n \chi_{E_n \setminus E'_n})|$$

appropriately and using restricted type $\alpha^{(n)}$ from Theorem 3.5 we obtain

$$A(\lambda) \leq C\lambda + 2^{-\alpha_n} A(\infty). \tag{19}$$

On the other hand, if

$$\frac{\lambda}{2}|E|^\alpha \leq |E|^{\alpha^{(n)}} < \lambda |E|^\alpha,$$

then we can use (17) to find an index $j \neq n$ such that

$$|E|^{\alpha^{(j)}} \leq C\lambda^{-\frac{\theta_n}{1-\theta_n}} |E|^\alpha.$$

By splitting

$$|\Lambda(F_1, \dots, F_n)| \leq |\Lambda(F_1, \dots, F_j \chi_{E'_j}, \dots, F_n)| + |\Lambda(F_1, \dots, F_j \chi_{E_j \setminus E'_j}, \dots, F_n)|$$

appropriately and using restricted type $\alpha^{(j)}$ from Theorem 3.5 we obtain

$$A(\lambda) \leq \max(A(\lambda/2), C\lambda^{-\frac{\theta_n}{1-\theta_n}} + A(\max_{j \neq n} 2^{\alpha_j - \alpha_j^{(n)}} \lambda)).$$

Since $\max_{j \neq n}(\alpha_j - \alpha_j^{(n)})$ is negative, we can iterate the previous inequality to obtain for sufficiently large λ:

$$A(\infty) \leq 1 + A(\lambda).$$

Combining this with (19) gives

$$A(\infty) \leq C + 2^{-\alpha_n} A(\infty),$$

which proves boundedness of $A(\infty)$. $\square$

Lemma 3.9. *Let $1 < p_i \leq \infty$ for $1 \leq i \leq n$ such that (6) holds. Then*

$$\Lambda(f_1, \dots, f_n) \leq C \|f_1\|_{p_1} \dots \|f_n\|_{p_n}$$

for all functions f_i supported on a set of finite measure.

Proof. By symmetry we can assume that $p_i \neq \infty$ for $i \leq j$ and $p_i = \infty$ for $i > j$ for a certain j. Lemma 3.8 implies

$$\Lambda(f_1, \dots, f_n) \leq C \|f_1\|_{L^{q_1,1}} \dots \|f_j\|_{L^{q_j,1}} \|f_{j+1}\|_\infty \|f_n\|_\infty$$

for all $q_1, \dots, q_j$ in a small neighborhood of $p_1, \dots, p_j$ satisfying

$$1/q_1 + \dots + 1/q_j = 1.$$

Fix functions $f_{j+1}, \dots, f_n$. Then Marcinkiewicz interpolation as in [13] implies

$$\Lambda(f_1, \dots, f_n) \leq C \|f_1\|_{L^{p_1}} \dots \|f_j\|_{L^{p_j}} \|f_{j+1}\|_\infty \|f_n\|_\infty$$

for all functions $f_1, \dots, f_j$. $\square$

We turn to bad tuples α.

Lemma 3.10. *Let the assumptions and notation be as in Theorem 1.1. Then Λ is of restricted type α for all bad tuples α satisfying* (16).

Proof. Fix $\alpha = (\alpha_1, \dots, \alpha_n)$; by symmetry we may assume that α has bad index n. By Lemma 3.7 we find θ_j such that

$$\alpha = \sum_{j=1}^n \theta_j \alpha^{(j)}, \quad \sum_{i=1}^n \theta_j = 1, \tag{20}$$

where $\alpha^{(j)} = (\alpha_i^{(j)})_{i=1}^n$ is an admissible tuple in P with bad index j. We have $\theta_n > 0$ and by the last statement of Lemma 3.7 we can assume that $\alpha_j - \alpha_j^{(n)}$ is negative for all $j \neq n$.

For $\lambda > 0$ let $A(\lambda)$ be the best constant such that for all sets $E_1, \dots, E_n$ of finite measure with

$$|E|^{\alpha^{(n)}} < \lambda |E|^\alpha \tag{21}$$

there is a major subset E_n' of E_n such that

$$|\Lambda(F_1, \dots, F_n)| \leq A(\lambda) |E|^\alpha$$

for all $F_i \in X(E_i')$. Let $A(\infty)$ be the supremum of all $A(\lambda)$.

Using restricted type $\alpha^{(n)}$ from Theorem 3.5 we obtain

$$A(\lambda) \leq C\lambda. \tag{22}$$

On the other hand let

$$\frac{\lambda}{2} |E|^\alpha \leq |E|^{\alpha^{(n)}} < \lambda |E|^\alpha.$$

Then we can find an index $j \neq n$ such that

$$|E|^{\alpha_j} \leq C \lambda^{-\frac{\theta_n}{1-\theta_n}} |E|^\alpha,$$

and we can use restricted type $\alpha^{(j)}$ from Theorem 3.5 to conclude

$$A(\lambda) \leq \max(A(\lambda/2), C\lambda^{-\frac{\theta_n}{1-\theta_n}} + A(\max_{j\neq n} 2^{\alpha_j - \alpha_j^{(n)}}\lambda)).$$

Since $\max_{j\neq n}(\alpha_j - \alpha_j^{(n)})$ is negative, we can iterate the previous inequality to obtain for sufficiently large λ:

$$A(\infty) \leq 1 + A(\lambda).$$

Together with (22) this proves the boundedness of $A(\infty)$. $\square$

Finally, we convert restricted type estimates for bad tuples α into strong type estimates by proving a Marcinkiewicz interpolation result in the spirit of [13].

Lemma 3.11. *Let α be a bad tuple satisfying* (16) *and assume that n is the bad index. Set $p_i = 1/\alpha_i$ for $1 \leq i \leq n$. Then*

$$\|T(f_1, \dots, f_{n-1})\|_{p_n'} \leq C\|f_1\|_{p_1} \dots \|f_{n-1}\|_{p_{n-1}}$$

for all functions f_i supported on a set of finite measure.

Proof. We assume for simplicity that $p_i \neq \infty$ for all i. If this was not the case, we could freeze the function f_i and the exponent p_i whenever $p_i = \infty$ and run the argument on the remaining functions only, as done in the proof of Lemma 3.9.

Let $f_1, \dots, f_{n-1}$ be functions such that $\|f_i\|_{p_i} = 1$ for $1 \leq i < n$. We have to show that

$$\|T(f_1, \dots, f_{n-1})\|_{p_n'} \lesssim 1.$$

We may assume that the f_i are non-negative. By a measure-preserving rearrangement, we may assume that the f_i are supported on the half-line $(0, \infty)$ and are monotone non-increasing on this half-line.

Let χ_k denote the function $\chi_k = \chi_{(2^k, 2^{k+1}]}$. We can expand the desired estimate as

$$\|\sum_{k_1,\dots,k_{n-1}} T(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}})\|_{p_n'} \lesssim 1.$$

Since $p_n' \leq 1$, we have the elementary inequality

$$\|\sum_\beta F_\beta\|_{p_n'}^{p_n'} \leq \sum_\beta \|F_\beta\|_{p_n'}^{p_n'},$$

so it suffices to show that

$$\sum_{k_1,\dots,k_{n-1}} \|T(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}})\|_{p_n'}^{p_n'} \lesssim 1. \tag{23}$$

By symmetry we may restrict the summation to the region

$$k_1 \geq k_2 \geq \dots \geq k_{n-1}.$$

Fix $k_1, \dots, k_{n-1}$. Let $\lambda > 0$ be arbitrary, and consider the set

$$E_n = \{\Re T(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}}) > \lambda\}.$$

Let α be an admissible tuple close to $1/p$; we may thus assume α has bad index n. Since Λ is of restricted type α, and $f_i\chi_{k_i} \in f_i(2^{k_i})X((2^k, 2^{k+1}])$, we may thus find a major subset E_n' of E_n such that

$$|\Lambda(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}}, \chi_{E_n'})| \lesssim |E_n|^{\alpha_n} \prod_{i=1}^{n-1} f_i(2^{k_i})2^{k_i\alpha_i}.$$

By definition of E_n, we thus have

$$\lambda |E_n| \lesssim |E_n|^{\alpha_n} \prod_{i=1}^{n-1} f_i(2^{k_i}) 2^{k_i \alpha_i}.$$

Solving for $|E_n|$ and optimizing in α, one obtains

$$|E_n| \lesssim \lambda^{-p_n'} 2^{-\varepsilon(k_1 - k_{n-1})} \min\left(\frac{F}{\lambda}, \frac{\lambda}{F}\right)^{\varepsilon} \left(\prod_{i=1}^{n-1} f_i(2^{k_i}) 2^{k_i/p_i}\right)^{p_n'}$$

for some $\varepsilon > 0$, where $F = \prod_{i=1}^{n-1} f_i(2^{k_i})$. By symmetry one may obtain the same bound when E_n is replaced by

$$\{|T(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}})| > \lambda\}.$$

Integrating this over all λ, one then obtains

$$\left\| T(f_1\chi_{k_1}, \dots, f_{n-1}\chi_{k_{n-1}}) \right\|_{p_n'} \lesssim 2^{-\varepsilon(k_1 - k_{n-1})} \prod_{i=1}^{n-1} f_i(2^{k_i}) 2^{k_i/p_i}.$$

To prove (23), it thus suffices to show

$$(24) \qquad \Big(\sum_{k_1 \geq \dots \geq k_{n-1}} 2^{-\varepsilon(k_1 - k_{n-1})} \Big(\prod_{i=1}^{n-1} f_i(2^{k_i}) 2^{k_i/p_i}\Big)^{p_n'} \Big)^{1/p_n'} \lesssim 1.$$

Write $s = k_1 - k_{n-1}$. For fixed s and k_1 there are at most $(1+s)^C$ choices of k_i. Fixing s, and then applying Hölder's inequality using (6), we can estimate the left-hand side of (24) by

$$\sum_{s \geq 0} (1+s)^C 2^{-\varepsilon s} \prod_{i=1}^{n-1} \Big(\sum_k (f_i(2^k) 2^{k/p_i})^{p_i}\Big)^{1/p_i}.$$

The s sum is convergent, and the expression inside the product is essentially $\|f_i\|_{p_i} = 1$. The claim is thus proved. □

Theorem 1.1 now follows from Lemma 3.9 and Lemma 3.11.

4. Exceptional set

It remains to prove Theorem 3.5. Let p satisfy the hypotheses of the theorem; by symmetry we may assume that the bad index of p is n. We have to show that for any $E_1, \dots, E_n$ one can find a major subset E_n' of E_n such that (13) holds for all $F_i \in X(E_i')$. By (6) and a scaling argument one may take $|E_n| = 1$.

We shall define E_n' explicitly as

$$(25) \qquad E_n' = \{x \in E_n : \mathrm{M}\chi_{E_i}(x) < C|E_i| \text{ for all } 1 \leq i \leq n\}.$$

From the Hardy-Littlewood maximal inequality we see that $|E_n \backslash E_n'| \leq \frac{1}{2}$ if C is chosen sufficiently large. Thus we have $|E_n'| \geq \frac{1}{2}|E_n|$ as desired.

Let F_i be arbitrary elements of $X(E_i')$. Define the normalized functions $f_1, \dots, f_n$ by

$$f_i = \frac{F_i \chi_{E_i'}}{|E_i'|^{1/2}}, \quad i = 1, \dots, n;$$

note that

$$(26) \qquad \|f_i\|_2 \lesssim 1 \text{ for all } i = 1, \dots, n.$$

Also define the numbers $a_i = |E_i|^{1/2}$. We may rewrite (13) as

$$|\Lambda(f_1, \dots, f_n)| \lesssim \prod_{i=1}^{n} a_i^{\theta_i}, \tag{27}$$

where $\theta_i = \frac{2}{p_i} - 1$ for $1 \leq i \leq n-1$. Since $a_n = 1$, the value of θ_n is arbitrary, but we shall set it so that

$$\theta_1 + \ldots + \theta_n = n - 2k. \tag{28}$$

From (6), (14) and (15) we see that

$$0 < \theta_i < 1 \text{ for all } i = 1, \dots, n. \tag{29}$$

Let $N \gg 1$ be a large constant to be chosen later. For any interval I and $1 \leq i \leq n$, define the normalized averages $\lambda_i(I)$ by

$$\lambda_i(I) = \frac{1}{|I||E_i'|} \int_{E_i'} \tilde{\chi}_I^N.$$

Clearly we have the estimates

$$\|f_i \tilde{\chi}_I^N\|_1 \lesssim a_i \lambda_i(I)|I| \tag{30}$$

and

$$\|f_i \tilde{\chi}_I^{N/2}\|_2 \lesssim \lambda_i(I)^{1/2} |I|^{1/2} \tag{31}$$

for all I and i.

From the construction of E_n' we see that the λ_i cannot simultaneously be large. More precisely, we have

Lemma 4.1. *For any interval I we have*

$$\lambda_n(I) \lesssim (1 + \lambda_1(I) + \ldots + \lambda_{n-1}(I))^{1-N}. \tag{32}$$

Proof. Suppose first that $2I$ intersected E_i'. Then there exists $x \in 2I$ such that $\mathrm{M}\chi_{E_i}(x) \leq |E_i|$ for all $1 \leq i \leq n$. This implies that $\lambda_i(I) \lesssim 1$ for all $1 \leq i \leq n$, which implies (32).

Now suppose that $j \geq 1$ was such that $2^j I$ was disjoint from E_i', but $2^{j+1} I$ intersected E_i'. By arguing as before we see that $\lambda_i(I) \lesssim 2^j$ for all $1 \leq i \leq n-1$, and $\lambda_n(I) \lesssim 2^{-j(1-N)}$, which again implies (32). $\square$

As we shall see, the dominant contribution to (27) shall come from those intervals I for which $\lambda_i(I) \sim 1$.

To prove Theorem 3.5 it thus suffices to prove the following estimate.

Theorem 4.2. *Let Λ be as above, let $f_1, \dots, f_n$ be functions satisfying* (26)*, and let $a_1, \dots, a_n$ be positive numbers. For each interval I and $1 \leq i \leq n$ we let $\lambda_i(I)$ be a non-negative number such that* (30)*,* (31)*,* (32) *hold for all I and i. Then for any θ_i satisfying* (28) *and* (29) *we have* (27)*, provided that N is chosen sufficiently large depending on θ.*

We have thus reduced the problem to that of estimating Λ on functions which are L^2-normalized, and whose L^1 and L^2 averages on intervals are somewhat under control.

5. Discretization

Let f_i, a_i, $\lambda_i(I)$ be as in Theorem 4.2. We now decompose the multiplier m using a Whitney decomposition, and replace Λ with a discretized variant.

We may extend m from the $(n-1)$-dimensional hyperplane Γ to the entire space $\mathbb{R}^n$ in such a way that (9) holds for all $\xi \in \mathbb{R}^n \backslash \Gamma'$ and all derivatives α up to a sufficiently large order.

Define a *shifted n-dyadic mesh* $D = D^n_\alpha$ to be a collection of cubes of the form

$$D^n_\alpha = \{2^j(k + (0,1)^n + (-1)^j\alpha) :: j \in \mathbb{Z}, \quad k \in \mathbb{Z}^n\},$$

where $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^n$. We define a *shifted dyadic cube* to be any member of a shifted n-dyadic mesh.

Observe that for every cube Q, there exists a shifted dyadic cube[1] Q' such that $Q \subseteq \frac{9}{10}Q'$ and $|Q'| \sim |Q|$; this is best seen by first verifying the $n = 1$ case.

Consider the collection $\mathbf{Q}$ of all shifted dyadic cubes Q such that

$$\operatorname{dist}(Q, \Gamma') \sim C_0 \operatorname{diam}(Q);$$

here C_0 is a large constant to be chosen later. From the above observation we see that the cubes $\{\frac{9}{10}Q : Q \in \mathbf{Q}\}$ form a finitely overlapping cover of $\mathbb{R}^n \backslash \Gamma'$. This implies that we may partition

$$m = \sum_{Q \in \mathbf{Q}} m_Q, \tag{33}$$

where each m_Q is supported in $Q \cap \Gamma$ and satisfies the bounds

$$|\partial^\alpha_\xi m_Q(\xi)| \lesssim \operatorname{diam}(Q)^{-|\alpha|} \tag{34}$$

for all derivatives ∂^α_ξ on Γ up to some sufficiently large order.

From (33) we have

$$\Lambda = \sum_{Q \in \mathbf{Q}} \Lambda_{m_Q}.$$

Of course Λ_{m_Q} vanishes unless Q intersects Γ. Since there are only a finite number of shifted dyadic meshes, we see that (27) will follow from

$$\sum_{Q \in \mathbf{Q} \cap D, Q \cap \Gamma \neq \emptyset} |\Lambda_{m_Q}(f_1, \dots, f_n)| \lesssim \prod_{i=1}^n a_i^{\theta_i},$$

where $D = D^n_\alpha$ is any shifted dyadic mesh. Henceforth $\alpha = (\alpha_1, \dots, \alpha_n)$ will be fixed.

To estimate the contribution of each Λ_Q we introduce tiles in the time-frequency plane $\mathbb{R} \times \mathbb{R}$.

Definition 5.1. Let $1 \leq i \leq n$. An i-tile is a rectangle $P = I_P \times w_P$ with area 1 and with $I_P \in D^1_0$, $w_P \in D^1_{\alpha_i}$. A multi-tile is an n-tuple $\vec{P} = (P_1, \dots, P_n)$ such that each P_i is an i-tile, and the $I_{P_i} = I_{\vec{P}}$ are independent of i. The frequency cube $Q_{\vec{P}}$ of a multi-tile is defined to be $\prod_{i=1}^n w_{P_i}$.

If $\vec{P}$ appears in an expression, we shall always adopt the convention that P_i denotes the i-th component of $\vec{P}$.

[1] This observation is due to Michael Christ.

Definition 5.2. Let $1 \leq i \leq n$, and let P be an i-tile. The semi-norm $\|f\|_P$ is defined by

$$\|f\|_P = \frac{1}{|I_P|} \|(\Delta_{w_P} f)\tilde{\chi}_{I_P}^{2N}\|_1,$$

where Δ_{w_P} is a Fourier multiplier whose symbol ψ_{w_P} is a bump function adapted to w_P and which equals 1 on $\frac{9}{10}w_P$.

The quantity $\|f\|_P$ can be viewed as an average value of f on the time-frequency tile P. From the rapid decay of $\Delta_{w_P} f$ we observe the crude estimate

Lemma 5.3. *For any P, we have*

$$\|f\|_P \lesssim \frac{1}{|I_P|} \|f\tilde{\chi}_{I_P}^{2N}\|_1.$$

The relationship between these semi-norms and the Λ_{m_Q} is given by

Lemma 5.4. *For any $Q \in \mathbf{Q} \cap D$, we have*

$$|\Lambda_{m_Q}(f_1, \dots, f_n)| \lesssim \sum_{\vec{P}: Q_{\vec{P}} = Q} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i},$$

where $\vec{P}$ runs over all multi-tiles with frequency cube Q.

Proof. By translation and scale invariance we may make Q the unit cube $[0,1]^n$.

We may write $m_Q(\xi) = \tilde{m}(\xi) \prod_{i=1}^n \psi_{w_{P_i}}(\xi_i)$, where $\tilde{m}$ is supported on $[0,1]^n \cap \Gamma$ and satisfies the same bounds (34) as m_Q; in other words, $\tilde{m}$ is a bump function on Γ. Since

$$\Lambda_{m_Q}(f_1, \dots, f_n) = \Lambda_{\tilde{m}}(\Delta_{w_{P_1}} f_1, \dots, \Delta_{w_{P_n}} f_n),$$

it suffices to show the estimate

$$|\Lambda_{\tilde{m}}(g_1, \dots, g_n)| \lesssim \sum_l \prod_{i=1}^n \|g_i \tilde{\chi}_{[l,l+1]}^N\|_1.$$

From Plancherel's theorem and (34) one sees that

$$\Lambda_{\tilde{m}}(g_1, \dots, g_n) = \int K(x) \prod_{i=1}^n g_i(x_i)\, dx,$$

where $x = (x_1, \dots, x_n)$ and the kernel K satisfies the estimate

$$|K(x)| \lesssim (1 + \sum_{i,j} |x_i - x_j|)^{-M}$$

for arbitrarily large M. In particular, we have

$$|K(x)| \lesssim \sum_l \prod_{i=1}^n \tilde{\chi}_{[l,l+1]}^{2N}(x_i)$$

and the claim follows. □

Let $\vec{\mathbf{P}}$ denote the set of all multi-tiles $\vec{P}$ such that $Q_{\vec{P}} \in \mathbf{Q} \cap D$ and $Q_{\vec{P}}$ intersects Γ. From the above lemma, it suffices to show that

$$\sum_{\vec{P} \in \vec{\mathbf{P}}} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i} \lesssim \prod_{i=1}^n a_i^{\theta_i}. \tag{35}$$

Note that the multiplier m no longer plays a role.

6. Rank

The tiles in $\vec{\mathbf{P}}$ have essentially k independent frequency parameters. To make this more precise we need some notation.

Definition 6.1. Let P and P' be tiles. We write $P' < P$ if $I_{P'} \subsetneq I_P$ and $w_P \subseteq 3w_{P'}$, and $P' \leq P$ if $P' < P$ or $P' = P$. We write $P' \lesssim P$ if $I_{P'} \subseteq I_P$ and $w_P \subseteq CC_0 w_{P'}$. We write $P' \lesssim' P$ if $P' \lesssim P$ and $P' \not\leq P$.

Note that the ordering $<$ is slightly different from the one in Fefferman [8] or Lacey and Thiele [15], [16], [18] as P' and P do not quite have to intersect. This slightly less strict ordering is more convenient for technical purposes.

If C_0 is sufficiently large, then we have

Lemma 6.2. *Let $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ be integers, and let $\vec{P}$, $\vec{P}'$ be multi-tiles in $\vec{\mathbf{P}}$. If $P'_{i_s} \leq P_{i_s}$ for all $s = 1, \ldots, k$, then $P'_i \lesssim P_i$ for all $1 \leq i \leq n$. If we further assume that $|I_{\vec{P}'}| \ll |I_{\vec{P}}|$, then we have $P'_i \lesssim' P_i$ for at least two choices of i.*

Proof. Since Γ' is non-degenerate, we can write it as a graph

$$\{\xi : \xi = h(\xi_{i_1}, \ldots, \xi_{i_k})\},$$

where h is a linear map from $\mathbb{R}^k$ to Γ.

Let ξ, ξ' denote the centers of $Q_{\vec{P}}$ and $Q_{\vec{P}'}$, respectively. From the definition of $\vec{\mathbf{P}}$ we have

$$|\xi - h(\xi_{i_1}, \ldots, \xi_{i_k})| \sim C_0 |I_{\vec{P}}|^{-1} \tag{36}$$

and

$$|\xi_1 + \ldots + \xi_n| \lesssim |I_{\vec{P}}|^{-1}, \tag{37}$$

and similarly for ξ'. Since $3w_{P'_{i_s}}$ contains $w_{P_{i_s}}$, we have

$$\xi_{i_s} = \xi'_{i_s} + O(|I_{\vec{P}'}|^{-1}).$$

Combining this with (36) we see that

$$\xi = \xi' + O(C_0 |I_{\vec{P}'}|^{-1}),$$

which implies that $P'_i \lesssim P_i$ for all I as desired.

Now suppose $|I_{\vec{P}'}| \ll |I_{\vec{P}}|$. By subtracting (36) for ξ and ξ' we thus have

$$|(\xi - \xi') - h((\xi - \xi')_{i_1}, \ldots, (\xi - \xi')_{i_k})| \sim C_0 |I_{\vec{P}'}|^{-1},$$

which implies that

$$|\xi - \xi'| \gtrsim C_0 |I_{\vec{P}'}|^{-1}.$$

On the other hand, from (37) we have

$$|(\xi - \xi')_1 + \ldots + (\xi - \xi')_n| \sim |I_{\vec{P}'}|^{-1}.$$

If C_0 is sufficiently large, this guarantees that there exist $1 \leq i < i' \leq n$ such that

$$|(\xi - \xi')_i|, |(\xi - \xi')_{i'}| \geq 3|I_{\vec{P}'}|^{-1},$$

which combined with the previous observations gives $P'_i \lesssim' P_i$ and $P'_{i'} \lesssim' P_{i'}$ as desired. $\square$

Definition 6.3. If $\vec{\mathbf{P}}$ is a collection of tiles, we define the norm $\|f_i\|_{\vec{\mathbf{P}},i}$ by

$$\|f_i\|_{\vec{\mathbf{P}},i} = \sup_{\vec{P}\in\vec{\mathbf{P}}} \|f_i\|_{P_i}.$$

We now claim that Theorem 4.2 follows from

Theorem 6.4. *Let $f_1, \ldots, f_n$ be functions obeying (26), and let $a_1, \ldots, a_n$, $\lambda_1, \ldots, \lambda_n$ be positive numbers. Let $\vec{\mathbf{P}}$ be a finite collection of multi-tiles such that Lemma 6.2 holds, and such that*

$$\|f_i \tilde{\chi}^N_{I_{\vec{P}}}\|_1 \lesssim a_i \lambda_i |I_{\vec{P}}|, \tag{38}$$

$$\|f_i \tilde{\chi}^{N/2}_{I_{\vec{P}}}\|_2 \lesssim \lambda_i^{1/2} |I_{\vec{P}}|^{1/2}, \tag{39}$$

for all $\vec{P} \in \vec{\mathbf{P}}$ and $1 \le i \le n$. Let I_0 be an interval such that $I_{\vec{P}} \subseteq I_0$ for all $\vec{P} \in \vec{\mathbf{P}}$, and

$$\|f_i \tilde{\chi}^{N/2}_{I_0}\|_2 \lesssim \lambda_i^{1/2} |I_0|^{1/2}, \tag{40}$$

for all $1 \le i \le n$. Then one has

(41)
$$\sum_{\vec{P}\in\vec{\mathbf{P}}} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i} \lesssim A^{n-2k-\theta_1-\ldots-\theta_n} \min(1, |I_0|) \prod_{i=1}^n (\lambda_i a_i)^{\theta_i} (1+\lambda_i)$$

for any θ_i satisfying (29) and

$$\theta_1 + \ldots + \theta_n \le n - k, \tag{42}$$

where A is the quantity

$$A = \sup_{1\le i\le n} \|f_i\|_{\vec{\mathbf{P}},i}. \tag{43}$$

Theorem 6.4 contains some rather technical assumptions which are convenient for induction purposes. In applications, we would only use the following corollary:

Corollary 6.5. *Let $f_1, \ldots, f_n$, $a_1, \ldots, a_n$, $\lambda_1, \ldots, \lambda_n$, and $\vec{\mathbf{P}}$ be as in the previous theorem. Then*

$$\sum_{\vec{P}\in\vec{\mathbf{P}}} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i} \lesssim \prod_{i=1}^n (\lambda_i a_i)^{\theta_i} (1+\lambda_i)$$

for any θ_i satisfying (28) and (29).

Now let $\lambda_1, \ldots, \lambda_n$ be dyadic numbers such that

$$\lambda_n \lesssim (1 + \lambda_1 + \ldots + \lambda_{n-1})^{1-N},$$

and apply the corollary to those multi-tiles $\vec{P}$ such that $\lambda_i(P_i) \sim \lambda_i$ for $1 \le i \le n$. The estimate (35) then follows by summing in λ_n and then in each of the λ_i, $1 \le i \le n-1$.

It remains to prove Theorem 6.4. This shall be done in two stages. First we shall handle the case $k = 1$ by arguments similar to those in Lacey and Thiele [15], [16], [18]; this is the longest part of the proof, occupying Sections 7–11. Then, in Section 12, we induct on k to obtain the general case.

7. Trees

Let $k = 1$. Fix the f_i, a_i, λ_i, $\vec{\mathbf{P}}$, and I_0.

In order to estimate (41) we shall have to organize $\vec{\mathbf{P}}$ into trees as in [8], [15], [16], [18].

Definition 7.1. For any $1 \leq j \leq n$ and a multi-tile $\vec{P}_T \in \vec{\mathbf{P}}$, define a j-tree with top $\vec{P}_T$ to be a collection of multi-tiles $T \subseteq \vec{\mathbf{P}}$ such that

$$P_j \leq P_{T,j} \text{ for all } \vec{P} \in T,$$

where $P_{T,j}$ is the j-th component of $\vec{P}_T$. We write I_T and $w_{T,j}$ for $I_{\vec{P}_T}$ and $w_{P_{T,j}}$, respectively. We say that T is a tree if it is a j-tree for some $1 \leq j \leq n$.

Note that T does not necessarily have to contain its top $\vec{P}_T$.

Definition 7.2. For any tree T, define the i-size $\mathrm{size}_i(T)$ of T to be the quantity

$$\mathrm{size}_i(T) = \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T : P_i \lesssim' P_{T,i}} |I_{\vec{P}}| \|f_i\|_{P_i}^2 \right)^{1/2} + \|f_i\|_{T,i}. \tag{44}$$

The relationship between the i-size to (41) is given by

Lemma 7.3. *If T is a tree, then*

(45)
$$\sum_{\vec{P} \in T} |I_{\vec{P}}| \prod_{i=1}^{n} \|f_i\|_{P_i} \lesssim |I_T| \sup_{1 \leq i_1 < i_2 \leq n} \mathrm{size}_{i_1}(T)\, \mathrm{size}_{i_2}(T) \prod_{i \neq i_1, i_2} \|f_i\|_{T,i}.$$

Proof. We first deal with the contribution of those multi-tiles $\vec{P}$ such that $|I_{\vec{P}}| \sim |I_T|$. From Lemma 6.2 there are only $O(1)$ of these multi-tiles, and the contribution can be handled by the estimate

$$\|f_i\|_{P_i} \leq \|f_i\|_{T,i} \leq \mathrm{size}_i(T). \tag{46}$$

Now let us consider those multi-tiles for which $|I_{\vec{P}}| \ll |I_T|$. From Lemma 6.2 there exist i_1, i_2 such that $P_{i_s} \lesssim' P_{T,i_s}$ for $s = 1, 2$; by pigeonholing we may make i_1, i_2 independent of $\vec{P}$. If one then uses (46) for all $i \neq i_1, i_2$, one reduces to showing that

$$\sum_{\vec{P} \in T} |I_{\vec{P}}| \|f_{i_1}\|_{P_{i_1}} \|f_{i_2}\|_{P_{i_2}} \lesssim |I_T|\, \mathrm{size}_{i_1}(T)\, \mathrm{size}_{i_2}(T).$$

But this follows from Cauchy-Schwarz. $\square$

To apply Lemma 7.3 we need to partition $\vec{\mathbf{P}}$ into trees T in such a way that we have good control on the i-sizes $\mathrm{size}_i(T)$ and the spatial sizes $|I_T|$. This shall be done in four stages.

First, in Section 8, we control the number of trees of a certain size by Lemma 7.5 below.

Definition 7.4. Let $1 \leq i \leq n$. Two trees T, T' are said to be *strongly i-disjoint* if

- $P_i \neq P'_i$ for all $\vec{P} \in T$, $\vec{P}' \in T'$.
- Whenever $\vec{P} \in T$, $\vec{P}' \in T'$ are such that $w_{P_i} \subsetneq w_{P'_i}$, then one has $I_{\vec{P}'} \cap I_T = \emptyset$, and similarly with T and T' reversed.

Note that if T and T' are strongly i-disjoint, then $P_i \cap P_i' = \emptyset$ for all $\vec{P} \in T$, $\vec{P}' \in T'$.

Lemma 7.5. *Let $1 \leq i \leq n$, $m \in \mathbb{Z}$, and let $\mathbf{T}$ be a collection of trees in $\vec{\mathbf{P}}$ which are mutually strongly i-disjoint and such that*

$$\mathrm{size}_i(T) \sim 2^{-m} \text{ for all } T \in \mathbf{T}. \tag{47}$$

Let I_0 be an interval such that $I_T \subseteq I_0$ for all $T \in \mathbf{T}$. Then we have

$$\sum_{T \in \mathbf{T}} |I_T| \lesssim 2^{2m} \| f_i \tilde{\chi}_{I_0}^{N/2} \|_2^2 \lesssim 2^{2m} \min(1, \lambda_i |I_0|). \tag{48}$$

By applying Lemma 7.5 to singleton trees and $n = k = 1$, one obtains

Corollary 7.6. *Let f be a function, $m \in \mathbb{Z}$, I_0 an interval, $\mathbf{P}$ a collection of disjoint tiles such that $I_P \subseteq I_0$ and $\|f\|_P \sim 2^{-m}$ for all $P \in \mathbf{P}$. Then we have*

$$\sum_{P \in \mathbf{P}} |I_P| \lesssim 2^{2m} \| f \tilde{\chi}_{I_0}^{N/2} \|_2^2.$$

In Section 9, we use Lemma 7.5 to obtain the following tree selection algorithm.

Lemma 7.7. *Let $1 \leq i \leq n$, $m \in \mathbb{Z}$, and suppose that one has*

$$\mathrm{size}_i(T) \leq 2^{-m} \tag{49}$$

for all trees T in $\vec{\mathbf{P}}$. Then there exists a collection $\mathbf{T}$ of trees in $\vec{\mathbf{P}}$ such that

$$\sum_{T \in \mathbf{T}} |I_T| \lesssim 2^{2m} \min(1, \lambda_i |I_0|) \tag{50}$$

and

$$\mathrm{size}_i(T') \leq 2^{-m-1} \tag{51}$$

for all trees T' in $\vec{\mathbf{P}} - \bigcup_{T \in \mathbf{T}} T$.

In Section 10, we shall bound the i-size by

Lemma 7.8. *For any tree T in $\vec{\mathbf{P}}$ and $1 \leq i \leq n$, we have*

$$\mathrm{size}_i(T) \lesssim a_i \lambda_i.$$

Finally, in Section 11 we combine Lemma 7.7 and Lemma 7.8 with Lemma 7.3 to prove (41) in the $k = 1$ case.

8. Proof of Lemma 7.5

The second inequality in (48) follows from (39) and the L^2-normalization of f_i, so it suffices to prove the first inequality.

Fix $1 \leq i \leq n$. By refining the trees T, we may assume that the tiles $\{P_i : \vec{P} \in T\}$ are all disjoint, and that

$$\sum_{\vec{P} \in T} |I_{\vec{P}}| \|f_i\|_{P_i}^2 \sim 2^{-2m} |I_T|.$$

In particular, we have

$$\sum_{\vec{P} \in \bigcup_T T} |I_{\vec{P}}| \|f_i\|_{P_i}^2 \sim 2^{-2m} \sum_T |I_T|. \tag{52}$$

Also, from (47) we have

$$\|f_i\|_{P_i} \lesssim 2^{-m} \tag{53}$$

for all $\vec{P} \in \bigcup_T T$.

We shall shortly prove the estimate

$$\sum_{\vec{P} \in \bigcup_T T} |I_{\vec{P}}| \|f_i\|_{P_i}^2 \lesssim 2^{-m} \|f_i \tilde{\chi}_{I_0}^{N/2}\|_2 (\sum_T |I_T|)^{1/2}; \tag{54}$$

the claim then follows by combining (52) and (54).

The estimate (54) is somewhat reminiscent of an orthogonality estimate. Accordingly, we shall use TT^* methods and similar techniques in the proof.

By duality we may find a function $\phi_{\vec{P}}$ for each $\vec{P} \in \bigcup_T T$ such that $|\phi_{\vec{P}}(x)| \lesssim \tilde{\chi}_{I_{\vec{P}}}^{2N}(x)$ for all $x \in \mathbb{R}$, and

$$\|f_i\|_{P_i} = \frac{1}{|I_{\vec{P}}|} \langle \Delta_{w_{P_i}}^* \phi_{\vec{P}}, f_i \rangle.$$

We can thus write the left-hand side of (54) as

$$\langle \sum_{\vec{P} \in \bigcup_T T} \|f_i\|_{P_i} \Delta_{w_{P_i}}^* \phi_{\vec{P}}, f_i \rangle.$$

From the Cauchy-Schwarz inequality, the inequality (54) will follow from the estimate

$$\| \sum_{\vec{P} \in \bigcup_T T} \|f_i\|_{P_i} \Delta_{w_{P_i}}^* \phi_{\vec{P}} \tilde{\chi}_{I_0}^{-N/2} \|_2^2 \lesssim 2^{-2m} \sum_T |I_T|. \tag{55}$$

Let us first consider the portion of the L^2 norm in (55) outside of $2I_0$. From the triangle inequality, it will suffice to show that

$$\| \sum_{\vec{P} \in \bigcup_T T : I_{\vec{P}} = I} \|f_i\|_{P_i} \Delta_{w_{P_i}}^* \phi_{\vec{P}} \tilde{\chi}_{I_0}^{-N/2} \|_{L^2(\mathbb{R} \setminus 2I_0)}^2 \lesssim \frac{|I|^3}{|I_0|^3} 2^{-2m} \sum_T |I_T| \tag{56}$$

for all $I \subseteq I_0$.

Fix I. The left-hand side of (56) can be rewritten as

$$\sum_{\vec{P}} \sum_{\vec{P}'} \|f_i\|_{P_i} \|f_i\|_{P_i'} \int_{\mathbb{R} \setminus 2I_0} \Delta_{w_{P_i}}^* \phi_{\vec{P}}(x) \overline{\Delta_{w_{P_i'}}^* \phi_{\vec{P}'}(x)} \tilde{\chi}_{I_0}^{-N}(x)\, dx,$$

where $\vec{P}$, $\vec{P}'$ are constrained by $I_{\vec{P}} = I_{\vec{P}'} = I$.

From the decay of $\phi_{\vec{P}}$ and the kernel of $\Delta_{w_{P_i}}$, we may estimate the integral by $O(|I|^{N+1}|I_0|^{-N})$. By translating w_{P_i} to be centered at the origin, and integrating by parts repeatedly, one can also obtain the bound of $|I| O(1 + |I| \operatorname{dist}(w_{P_i}, w_{P_i'}))^{-N}$. Taking the geometric mean of these estimates, we can bound the left-hand side of (56) by

$$|I|^{N/2+1} |I_0|^{-N/2} \sum_{\vec{P}} \sum_{\vec{P}'} \|f_i\|_{P_i} \|f_i\|_{P_i'} (1 + |I| \operatorname{dist}(w_{P_i}, w_{P_i'}))^{-N/2}.$$

By Schur's test (or Young's inequality), this is bounded by

$$|I|^{N/2+1} |I_0|^{-N/2} \sum_{\vec{P}} \|f_i\|_{P_i}^2.$$

Thus it is only left to show that

$$\sum_{\vec{P}} |I| \|f_i\|_{P_i}^2 \lesssim 2^{-2m} \sum_T |I_T|.$$

But this follows from (53) and the observation that each tree T contributes at most $O(1)$ multi-tiles $\vec{P}$ to the left-hand sum.

It thus remains to show that

$$\| \sum_{\vec{P} \in \bigcup_T T} \|f_i\|_{P_i} \Delta^*_{w_{P_i}} \phi_{\vec{P}} \|_2^2 \lesssim 2^{-2m} \sum_T |I_T|. \tag{57}$$

We estimate the left-hand side of (57) as

$$\sum_{\vec{P}, \vec{P}' \in \bigcup_T T} \|f_i\|_{P_i} \|f_i\|_{P_i'} |\langle \Delta^*_{w_{P_i}} \phi_{\vec{P}}, \Delta^*_{w_{P_i'}} \phi_{\vec{P}'} \rangle|.$$

The inner product vanishes unless w_{P_i} and $w_{P_i'}$ intersect; by the nesting property of dyadic intervals this means that one of these intervals is a subset of the other. By symmetry it suffices to consider the case $w_{P_i} \subseteq w_{P_i'}$.

One can easily verify that $\Delta^*_{w_{P_i}} \phi_{\vec{P}} \lesssim \tilde{\chi}^{2N}_{I_{\vec{P}}}$, and similarly with $\vec{P}$ replaced by $\vec{P}'$. Thus we may estimate the inner product as

$$|\langle \Delta^*_{w_{P_i}} \phi_{\vec{P}}, \Delta^*_{w_{P_i'}} \phi_{\vec{P}'} \rangle| \lesssim |I_{\vec{P}'}| (1 + \frac{\mathrm{dist}(I_{\vec{P}'}, I_{\vec{P}})}{|I_{\vec{P}}|})^{-2N}.$$

To show (57) it thus suffices to show that

$$\sum_{\vec{P}, \vec{P}' \in \bigcup_T T : w_{P_i} \subseteq w_{P_i'}} \|f_i\|_{P_i} \|f_i\|_{P_i'} |I_{\vec{P}'}| (1 + \frac{\mathrm{dist}(I_{\vec{P}'}, I_{\vec{P}})}{|I_{\vec{P}}|})^{-2N} \lesssim 2^{-2m} \sum_T |I_T|. \tag{58}$$

Let us first deal with the portion of the sum where $|I_{\vec{P}}| \sim |I_{\vec{P}'}|$. In this case we use the estimate

$$\|f_i\|_{P_i} \|f_i\|_{P_i'} \lesssim \|f_i\|_{P_i}^2 + \|f_i\|_{P_i'}^2.$$

We treat the first term, as the second is similar. For each $\vec{P}$, the associated $\vec{P}'$ have disjoint spatial intervals $I_{\vec{P}'}$. Thus one may compute the $\vec{P}'$ summation, and estimate this contribution to (58) as

$$\sum_{\vec{P} \in \bigcup_T T} \|f_i\|_{P_i}^2 |I_{\vec{P}}|.$$

But this is acceptable by (52).

Now suppose $|I_{\vec{P}}| \gg |I_{\vec{P}'}|$. By (53) we may estimate the contribution to (58) by

$$2^{-2m} \sum_T \sum_{\vec{P} \in T} \sum_{\vec{P}' \in \bigcup_{T'} T' : w_{P_i} \subseteq w_{P_i'}, |I_{\vec{P}}| \gg |I_{\vec{P}'}|} |I_{\vec{P}'}| (1 + \frac{\mathrm{dist}(I_{\vec{P}'}, I_{\vec{P}})}{|I_{\vec{P}}|})^{-2N}.$$

From the assumptions on $\vec{P}$ and $\vec{P}'$ we see that $\vec{P}'$ must belong to a tree other than T; since the trees are strongly i-disjoint we thus have $I_{\vec{P}'} \cap I_T = \emptyset$, and that the $I_{\vec{P}'}$ are disjoint. We may thus estimate the contribution to (58) by

$$2^{-2m} \sum_T \sum_{\vec{P} \in T} \int_{\mathbb{R} \setminus I_T} (1 + \frac{\mathrm{dist}(x, I_{\vec{P}})}{|I_{\vec{P}}|})^{-2N} \, dx.$$

The integral in this expression is bounded by

$$(1 + \frac{\operatorname{dist}(\mathbb{R}\backslash I_T, I_{\vec{P}})}{|I_{\vec{P}}|})^{-3}.$$

Inserting this into the previous estimate and computing the inner sum, we obtain (58) as desired. This completes the proof of Lemma 7.5.

9. Proof of Lemma 7.7

Fix i, m. The idea will be to remove trees T from $\vec{\mathbf{P}}$ one at a time until (51) is satisfied.

By refining the tree by a finite factor we may assume (using Lemma 6.2) that for each dyadic interval I there is at most one multi-tile $\vec{P} \in T$ such that $I_{\vec{P}} = I$. We may assume that for any $\vec{P}, \vec{P}' \in \vec{\mathbf{P}}$, $|I_{\vec{P}}|/|I'_{\vec{P}}|$ is an integer power of 2^{C_1}, where C_1 is a large constant to be chosen shortly. By Lemma 6.2 and a further refinement we can ensure that if w_{P_i} is fixed, then w_{P_j} is also fixed for every $1 \leq j \leq n$.

Let $\vec{\mathbf{P}}^*$ consist of those multi-tiles $\vec{P}$ in $\vec{\mathbf{P}}$ such that

$$\|f_i\|_{P_i} \geq 2^{-m-2};$$

for these tiles we thus have

$$\|f_i\|_{P_i} \sim 2^{-m} \tag{59}$$

by (49). We place a partial order $<$ on the multi-tiles in $\vec{\mathbf{P}}^*$ by defining $\vec{P}' < \vec{P}$ if $P'_i < P_i$. Let $\vec{\mathbf{P}}^{**}$ be those tiles which are maximal with respect to this ordering.

By construction, the tiles $\{P_i : \vec{P} \in \vec{\mathbf{P}}^{**}\}$ are disjoint. From this, (59), and Corollary 7.6 we see that

$$\sum_{\vec{P} \in \vec{\mathbf{P}}^{**}} |I_{\vec{P}}| \lesssim 2^{2m} \|f_i \tilde{\chi}_{I_0}^{N/2}\|_2^2 \lesssim 2^{2m} \min(1, \lambda_1 |I_0|). \tag{60}$$

For each $\vec{P} \in \vec{\mathbf{P}}^{**}$ we associate the i-tree

$$T = \{\vec{P}' \in \vec{\mathbf{P}}^* : \vec{P}' \leq \vec{P}\}.$$

From (60) we see that one can remove these trees T from $\vec{\mathbf{P}}$ and place them into $\mathbf{T}$ while respecting (50). After removing these trees, we have eliminated all elements of $\vec{\mathbf{P}}^*$, so that we have

$$\|f_i\|_{P_i} < 2^{-m-2} \tag{61}$$

for all remaining multi-tiles $\vec{P}$.

If P is a tile, let ξ_P denote the center of w_P. If P and P' are tiles, we write $P' \lesssim^+ P$ if $P' \lesssim' P$ and $\xi_{P'} > \xi_P$, and $P' \lesssim^-$ if $P' \lesssim' P$ and $\xi_{P'} < \xi_P$. If T is a tree, write $\xi_{T,i}$ for $\xi_{P_{T,i}}$.

We now perform the following algorithm. We consider the set of all trees T in $\vec{\mathbf{P}}$ such that

$$P_i \lesssim^+ P_{T,i} \text{ for all } \vec{P} \in T \tag{62}$$

and

$$\sum_{\vec{P} \in T} |I_{\vec{P}}| \|f_j\|_{P_i}^2 \geq 2^{-2m-5} |I_T|. \tag{63}$$

If there are no trees obeying (62) and (63), we terminate the algorithm. Otherwise, we choose T among all such trees so that $\xi_{T,i}$ is maximal, and that T is maximal with respect to set inclusion. Let T' denote the i-tree

$$T' = \{\vec{P} \in \vec{\mathbf{P}} : P_i \leq P_{T,i}\}.$$

We remove both T and T' from $\vec{\mathbf{P}}$, and add them to $\mathbf{T}$. (These two trees are allowed to overlap.) Then one repeats the algorithm until we run out of trees obeying (62) and (63).

Since $\vec{\mathbf{P}}$ is finite, this algorithm terminates in a finite number of steps, producing trees $T_1, T_1', T_2, T_2', \ldots, T_M, T_M'$. We claim that the trees $T_1, \ldots, T_M$ produced in this manner are strongly disjoint. It is clear from construction that $T_s \cap T_{s'} = \emptyset$ for all $s \neq s'$; by our assumptions on the multi-tiles we thus see that $P_i \neq P_i'$ for all $\vec{P} \in T_s$, $\vec{P'} \in T_{s'}$, $s \neq s'$.

Now suppose for contradiction that we had multi-tiles $\vec{P} \in T_s$, $\vec{P'} \in T_{s'}$ such that $w_{P_i} \subsetneq w_{P_i'}$ and $I_{P_i'} \subseteq I_{T_s}$. From our assumptions on the multi-tiles we thus have $|w_{P_i'}| \geq 2^{C_1}|w_{P_i}|$. Since $P_i \lesssim P_{T_s,i}$ and $P_i' \lesssim^+ P_{T_{s'},i}$, we thus see that $\xi_{T_{s'},i} < \xi_{T_s,i}$ if C_1 is sufficiently large. By our selection algorithm this implies that $s < s'$.

Also, since $|w_{P_i'}| \geq 2^{C_1}|w_{P_i}|$, $I_{P_i'} \subseteq I_{T_s}$, and $P_i \lesssim P_{T_s,i}$ we see that $P_i' \leq P_{T_s,i}$ if C_1 is sufficiently large. Since $s < s'$, this means that $\vec{P'} \in T_s'$. But T_s' and $T_{s'}$ are disjoint by construction, which is a contradiction. Thus the trees T_s are strongly disjoint. From (49) and (63) we see that these trees obey (47), and thus we have

$$\sum_{s=1}^{M} |I_{T_s}| \lesssim 2^{2m} \min(1, \lambda_i |I_0|).$$

Since T_s' has the same top as T_s, we may thus add all the T_s and T_s' to $\mathbf{T}$ while respecting (50).

Now consider the set $\vec{\mathbf{P}}$ of remaining multi-tiles. We note that

$$\sum_{\vec{P} \in T : P_i \lesssim^+ P_{T,i}} |I_{\vec{P}}| \|f_j\|_{P_i}^2 < 2^{-2m-5} |I_T| \tag{64}$$

for all trees T in $\vec{\mathbf{P}}$, since otherwise the portion of T which obeyed (62) would be eligible for selection by the above algorithm.

We now repeat the previous algorithm, but replace $\lesssim^+$ by $\lesssim^-$ and select the trees T so that $\xi_{T,i}$ is *minimized* rather than maximized. This yields a further collection of trees to add to $\mathbf{T}$ while still respecting (50), and the remaining collection of tiles $\vec{\mathbf{P}}$ has the property that

$$\sum_{\vec{P} \in T : P_i \lesssim^- P_{T,i}} |I_{\vec{P}}| \|f_j\|_{P_i}^2 < 2^{-2m-5} |I_T| \tag{65}$$

for all trees T in $\vec{\mathbf{P}}$. Combining (61), (64), and (65) we see that

$$\mathrm{size}_i(T) \leq 2^{-m-1}$$

for all trees T in $\vec{\mathbf{P}}$, and we are done.

10. Proof of Lemma 7.8

Fix $1 \leq i \leq n$. We may refine the collection T of tiles as in the previous section.

Let P be a tile. Since the convolution kernel of Δ_{w_P} is rapidly decreasing for $|x| \gg |I_P|$, we see from the definition of $\|f\|_P$ that

$$\|f\|_P \lesssim \frac{1}{|I_P|} \|f \tilde{\chi}_{I_P}^N\|_1.$$

From (38) we thus have

$$\|f_i\|_{P_i} \lesssim a_i \lambda_i$$

for all $\vec{P} \in \vec{\mathbf{P}}$. In particular we have

$$\|f_i\|_{T,i} \lesssim a_i \lambda_i$$

for all trees T in $\vec{\mathbf{P}}$.

Let B denote the best constant such that

$$\sum_{\vec{P} \in T : P_i \lesssim' P_{T,i}} |I_{\vec{P}}| \|f_j\|_{P_i}^2 \leq B |I_T| \tag{66}$$

for all trees T in $\vec{\mathbf{P}}$; to finish the proof of Lemma 7.8 we must show that $B \lesssim a_i^2 \lambda_i^2$.

To achieve this we first need to prove an apparently weaker estimate.

Lemma 10.1. *For any tree T and function f, we have*

$$\|(\sum_{\vec{P} \in T : P_i \lesssim' P_{T,i}} \|f\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2}\|_{L^{1,\infty}} \lesssim \|f \tilde{\chi}_{I_T}^N\|_1.$$

Proof. The expression in the norm is a variant of a Littlewood-Paley square function. Thus, we shall use Calderón-Zygmund techniques to prove this estimate.

By frequency translation invariance we may assume that $w_{T,i}$ contains the origin.

Let us first assume that f is supported outside of $2I_T$. From Lemma 5.3 we have

$$\|f\|_{P_i} \lesssim \frac{|I_{\vec{P}}|^{N-1}}{|I_T|^N} \|f \tilde{\chi}_{I_T}^N\|_1. \tag{67}$$

Applying this estimate, we obtain

$$\|(\sum_{\vec{P} \in T : P_i \lesssim' P_{T,i}} \|f\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2}\|_2 \lesssim |I_T|^{-1/2} \|f \tilde{\chi}_{I_T}^N\|_1,$$

and the claim follows from Hölder.

It thus remains to show that

$$\{(\sum_{\vec{P} \in T : P_i \lesssim' P_{T,i}} \|f\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2} \gtrsim \alpha\} \lesssim \alpha^{-1} \|f\|_1 \tag{68}$$

for all $\alpha > 0$.

Fix α. Perform a Calderón-Zygmund decomposition at level α,

$$f = g + \sum_I b_I,$$

where $\|g\|_2 \lesssim \alpha^{1/2} \|f\|_1^{1/2}$, the I are intervals such that

$$\sum_I |I| \lesssim \alpha^{-1} \|f\|_1, \tag{69}$$

and the b_I are supported on I and satisfy $\int_I b_I \sim \alpha |I|$ and $\int b_I = 0$.

To control the contribution of g, it suffices from Chebyshev to verify the L^2 bound

$$\|(\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} \|g\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2}\|_2 \lesssim \|g\|_2.$$

The left-hand side of this is

$$(\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} |I_{\vec{P}}| \|g\|_{P_i}^2)^{1/2}. \tag{70}$$

However, from Hölder and the definition of $\|g\|_{P_i}$ we have

$$|I_{\vec{P}}| \|g\|_{P_i}^2 \lesssim \|\Delta_{w_{P_i}}(g) \tilde{\chi}_{I_{\vec{P}}}^{-1}\|_2^2.$$

Thus we may bound (70) by

$$(\sum_w \|\Delta_w(g)\|_2^2)^{1/2},$$

where w ranges over the set $\{w_{P_i} : \vec{P} \in T\}$. But the desired bound of $\|g\|_2$ then follows from Plancherel and the lacunary nature of the w.

To deal with the b_I, it suffices from the triangle inequality, Chebyshev, and by showing that

$$\|(\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} \|b_I\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2}\|_{L^1(\mathbb{R}\setminus 2I)} \lesssim \alpha|I|$$

for all I. In fact we prove the stronger

$$\|\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} \|b_I\|_{P_i} \chi_{I_{\vec{P}}}\|_{L^1(\mathbb{R}\setminus 2I)} \lesssim \alpha|I|. \tag{71}$$

Fix I. We may restrict the summation to those $\vec{P}$ such that $I_{\vec{P}} \not\subseteq 2I$.

From Lemma 5.3 we have

$$\|b_I\|_{P_i} \lesssim \alpha \frac{|I|}{|I_{\vec{P}}|}(1 + \frac{\operatorname{dist}(I_{\vec{P}}, I)}{|I_{\vec{P}}|})^{-N};$$

in particular, from the hypothesis $I_{\vec{P}} \not\subseteq 2I$ we have

$$\|b_I\|_{P_i} \lesssim \alpha \frac{|I_{\vec{P}}|^{N-1}}{|I|^{N-1}}.$$

Also, by playing off the moment condition on b_I against the smoothness of $\Delta_{w_{P_i}}$, we have

$$\|b_I\|_{P_i} \lesssim \alpha \frac{|I|^2}{|I_{\vec{P}}|^2}.$$

Combining all these estimates, we obtain

$$\|b_I\|_{P_i} \lesssim \alpha \frac{|I|}{|I_{\vec{P}}|}(1 + \frac{\operatorname{dist}(I_{\vec{P}}, I)}{|I_{\vec{P}}|})^{-N/2} \min(\frac{|I_{\vec{P}}|}{|I|}, \frac{|I|}{|I_{\vec{P}}|})^{1/2}.$$

Inserting this into (71) we obtain the result. □

To bootstrap Lemma 10.1 to Lemma 7.8 we shall employ a variant of arguments used to prove the John-Nirenberg inequality.

By construction of B, there exists a tree T such that

$$\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} |I_{\vec{P}}| \|f_j\|_{P_i}^2 = B|I_T|. \tag{72}$$

Fix this tree. From Lemma 10.1 and (38) we have

$$\|(\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} \|f\|_{P_i}^2 \chi_{I_{\vec{P}}})^{1/2}\|_{L^{1,\infty}} \lesssim |I_T| a_i \lambda_i.$$

We thus have $|E| \leq \frac{1}{2}|I_T|$, where

$$E = \{\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}} \|f\|_{P_i}^2 \chi_{I_{\vec{P}}} \geq C a_i^2 \lambda_i^2\}$$

and C is a sufficiently large constant.

From the nesting properties of dyadic intervals we see that there must exist a subset T^* of T such that the intervals $\{I_{\vec{P}} : \vec{P} \in T^*\}$ form a partition of E. In particular we have

$$\sum_{\vec{P}\in T^*} |I_{\vec{P}}| \leq \frac{1}{2}|I_T|. \tag{73}$$

If $\vec{P} \in \vec{\mathbf{P}}$ is such that $I_{\vec{P}} \not\subseteq E$, then we must have $P_i \lesssim' P_i'$ for some $\vec{P'} \in T^*$, if C_1 is chosen sufficiently large. We can thus decompose the left-hand side of (72) as

$$\|\sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}, I_{\vec{P}} \not\subseteq E} \|f_j\|_{P_i}^2 \chi_{I_{\vec{P}}}\|_1 + \sum_{\vec{P'}\in T^*} \sum_{\vec{P}\in T: P_i \lesssim' P_{T,i}, I_{\vec{P}} \subseteq E} |I_{\vec{P}}| \|f_j\|_{P_i}^2.$$

Consider the former term. From the definition of E and the nesting properties of dyadic intervals we see that the expression in the norm is $O(a_i^2\lambda_i^2)$. Thus the former term is $O(|I_T|a_i^2\lambda_i^2)$.

Now consider the latter summation. For each $\vec{P'} \in T^*$ the inner sum is $O(B|I_{\vec{P'}}|)$ from (66). Inserting these estimates back into (72) and using (73) we obtain

$$B|I_T| \lesssim |I_T| a_i^2 \lambda_i^2 + \frac{1}{2} B |I_T|,$$

and the claim follows. This concludes the proof of Lemma 7.8.

11. Conclusion of the $k = 1$ case

We now prove (41). We first observe from iterating Lemma 7.7 and using Lemma 7.8 that

Corollary 11.1. *Let $1 \leq i \leq n$. Then there exists a partition*

$$\vec{\mathbf{P}} = \bigcup_{m: 2^{-m} \lesssim \lambda_i a_i} \vec{\mathbf{P}}^{m,i},$$

where one has (47) for all trees T in $\vec{\mathbf{P}}^{m,i}$, and such that $\vec{\mathbf{P}}^{m,i}$ can be covered as

$$\vec{\mathbf{P}}^{m,i} = \bigcup_{T\in \mathbf{T}^{m,i}} T, \tag{74}$$

where $\mathbf{T}^{m,i}$ is a collection of trees such that

$$\sum_{T\in\mathbf{T}^{m,i}} |I_T| \lesssim 2^{2m} \min(1, \lambda_i |I_0|). \tag{75}$$

Write the left-hand side of (41) as

$$\sum_{m_1,\dots,m_n} \sum_{\vec{P} \in \vec{\mathbf{P}}^{m_1,1} \cap \dots \cap \vec{\mathbf{P}}^{m_n,n}} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i},$$

where we implicitly assume

$$2^{-m_i} \leq \lambda_i a_i. \tag{76}$$

By symmetry we may restrict the summation to the case

$$m_1 \leq m_2 \leq \dots \leq m_n. \tag{77}$$

We then estimate the sum by

$$\sum_{m_1 \leq \dots \leq m_n} \sum_{T \in \mathbf{T}^{m_1,1}} \sum_{\vec{P} \in T'} |I_{\vec{P}}| \prod_{i=1}^n \|f_i\|_{P_i}, \tag{78}$$

where $T' = T'(T, m_2, \dots, m_n)$ denotes the tree

$$T' = T \cap \vec{\mathbf{P}}^{m_2,2} \cap \dots \cap \vec{\mathbf{P}}^{m_n,n}.$$

By Lemma 7.3 we may estimate (78) by

$$\sum_{m_1 \leq \dots \leq m_n} \sum_{T \in \mathbf{T}^{m_1,1}} |I_T| \sup_{1 \leq i_1 < i_2 \leq n} \mathrm{size}_{i_1}(T') \,\mathrm{size}_{i_2}(T') \prod_{i \neq i_1, i_2} \|f_i\|_{T',i}. \tag{79}$$

From (47) we have

$$\mathrm{size}_i(T') \lesssim 2^{-m_i},$$

which implies with (43) that

$$\|f_i\|_{T',i} \lesssim \min(2^{-m_i}, A).$$

Thus we may estimate (79) by

$$\sum_{m_1 \leq \dots \leq m_n} \sum_{T \in \mathbf{T}^{m_1,1}} |I_T| \sup_{1 \leq i_1 < i_2 \leq n} 2^{-m_{i_1}} 2^{-m_{i_2}} \prod_{i \neq i_1, i_2} \min(2^{-m_i}, A).$$

It is clear that the supremum is attained when $i_1 = 1$, $i_2 = 2$. Applying (75) we can thus estimate the previous by

$$\min(1, \lambda_1 |I_0|) \sum_{m_1 \leq \dots \leq m_n} 2^{2m_1} 2^{-m_1} 2^{-m_2} \prod_{2 < i \leq n} \min(2^{-m_i}, A).$$

Clearly we have the estimate

$$\min(1, \lambda_1 |I_0|) \leq \min(1, |I_0|) \prod_{i=1}^n (1 + \lambda_i).$$

To show (41), it thus suffices to show

$$\sum_{m_1 \leq \dots \leq m_n} 2^{2m_1} 2^{-m_1} 2^{-m_2} \prod_{2 < i \leq n} \min(2^{-m_i}, A) \lesssim A^{n-2-\theta_1-\dots-\theta_n} \prod_{i=1}^n (a_i \lambda_i)^{\theta_i}. \tag{80}$$

We first consider the case when (42) holds with equality (i.e. (29) holds). In this case we need only show that

$$\sum_{m_1 \leq \dots \leq m_n} 2^{2m_1} \prod_{i=1}^n 2^{-m_i} \lesssim \prod_{i=1}^n (a_i \lambda_i)^{\theta_i}. \tag{81}$$

From (42) we may write

$$2^{2m_1} = \prod_{i=1}^{n} 2^{(1-\theta_i)m_1} \leq \prod_{i=1}^{n} 2^{(1-\theta_i)m_i} \tag{82}$$

by (28) and (77). Thus (81) reduces to

$$\sum_{m_1,\dots,m_n} \prod_{i=1}^{n} 2^{-\theta_i m_i} \lesssim \prod_{i=1}^{n} (a_i\lambda_i)^{\theta_i}. \tag{83}$$

But this follows from (76) and (28).

Now suppose that (42) holds with strict inequality. We may then find θ_i' satisfying (28) and (29) such that $\theta_i = \theta_i'$ for $i = 1, 2$ and $\theta_i' > \theta_i$ for $i > 2$; note how one needs (28) and (42) for $k = 1$ to ensure that θ_i' exists. Using the estimate

$$\min(2^{-m_i}, A) \leq A^{\theta_i' - \theta_i} 2^{-m_i(1-\theta_i'+\theta_i)}$$

and canceling the A factors, we reduce to

$$\sum_{m_1\leq\dots\leq m_n} 2^{2m_1} \prod_{i=1}^{n} 2^{-m_i(1-\theta_i'+\theta_i)} \lesssim \prod_{i=1}^{n} (a_i\lambda_i)^{\theta_i}.$$

Applying (82) with the θ_i replaced by θ_i', we reduce to (83) as before. Thus in either case (41) is proven.

12. The induction on k

We have just proven Theorem 6.4 when $k = 1$. Now suppose inductively that $k > 1$, and the claim has already been proven for $k - 1$.

We need to show (41). By symmetry it suffices to consider those tiles $\vec{P}$ for which

$$\|f_1\|_{P_1} \geq \|f_2\|_{P_2} \geq \dots \geq \|f_n\|_{P_n}; \tag{84}$$

we shall implicitly assume this in the sequel.

From Lemma 5.3, (38), and (43) we have

$$\|f_1\|_{P_1} \lesssim \min(a_1\lambda_1, A).$$

Thus (41) reduces to showing that

$$\begin{aligned} &\sum_{m:2^{-m}\lesssim\min(a_1\lambda_1,A)} \sum_{\vec{P}\in\vec{\mathbf{P}}^m} |I_{\vec{P}}| \prod_{i=1}^{n} \|f_i\|_{P_i} \\ &\qquad \lesssim A^{n-2k-\theta_1-\dots-\theta_n} \min(1, |I_0|) \prod_{i=1}^{n} (\lambda_i a_i)^{\theta_i}(1+\lambda_i), \end{aligned} \tag{85}$$

where

$$\vec{\mathbf{P}}^m = \{\vec{P} \in \vec{\mathbf{P}} : \|f_1\|_{P_1} \sim 2^{-m}\}.$$

Fix m. We order the multi-tiles in $\vec{\mathbf{P}}^m$ by setting $\vec{P}' < \vec{P}$ if $P_1' < P_1$. Let $\vec{\mathbf{P}}^{m,*}$ be the tiles in $\vec{\mathbf{P}}^m$ which are maximal with respect to this ordering. By applying Corollary 7.6 as in the proof of Lemma 7.7, we see that

$$\sum_{\vec{P}\in\vec{\mathbf{P}}^{m,*}} |I_{\vec{P}}| \lesssim 2^{2m} \min(1, \lambda_1|I_0|). \tag{86}$$

We may estimate the left-hand side of (85) as

$$\sum_{m:2^{-m}\lesssim\min(a_1\lambda_1,A)}\ \sum_{\vec{P}'\in\vec{\mathbf{P}}^{m,*}}\ \sum_{\vec{P}\in\vec{\mathbf{P}}^m:P_1\leq P_1'} |I_{\vec{P}}|2^{-m}\prod_{i=2}^n \|f_i\|_{P_i}. \tag{87}$$

For fixed $\vec{P}'$, the collection of multi-tiles $\{\vec{P}\in\vec{\mathbf{P}}^m : P_1\leq P_1'\}$ satisfies the conditions of Lemma 6.2 with (n,k) replaced by $(n-1,k-1)$, if we forget the first tile P_1 from each multi-tile $\vec{P}$. Thus we may apply the induction hypothesis, with I_0 replaced by $I_{\vec{P}'}$ and A estimated by 2^{-m} (thanks to (84)), and estimate (87) by

$$\sum_{m:2^{-m}\lesssim\min(a_1\lambda_1,A)}\ \sum_{\vec{P}'\in\vec{\mathbf{P}}^{m,*}} 2^{-m}\min(1,|I_{\vec{P}'}|)2^{-m((n-1)-2(k-1)-\theta_2-\ldots-\theta_n)}$$
$$\times\prod_{i=2}^n(\lambda_i a_i)^{\theta_i}(1+\lambda_i).$$

Estimating $\min(1,|I_{\vec{P}'}|)$ by $|I_{\vec{P}'}|$ and applying (86), and then gathering the powers of 2^m, this can be estimated by

$$\sum_{m:2^{-m}\lesssim\min(a_1\lambda_1,A)} 2^{-m(n-2k-\theta_2-\ldots-\theta_n)}\min(1,\lambda_1|I_0|)\prod_{i=2}^n(\lambda_i a_i)^{\theta_i}(1+\lambda_i).$$

Evaluating the m summation and applying the elementary inequalities

$$\min(a_1\lambda_1,A)^{n-2k-\theta_2-\ldots-\theta_n}\lesssim A^{n-2k-\theta_2-\ldots-\theta_n}(a_1\lambda_1)^{\theta_1}$$

(which follows from (28) and (42)) and

$$\min(1,\lambda_1|I_0|)\leq\min(1,|I_0|)(1+\lambda_1)$$

we see that (85) follows. This concludes the proof of Theorem 6.4, and thus Theorem 1.1, for general k.

13. Remarks

Let A denote the Wiener algebra, that is, the space of functions whose Fourier transform is in L^1. The purpose of this section is to extend Theorem 1.1 slightly to

Theorem 13.1. *Let $0\leq s<n-1$, and let Γ' be a subspace of Γ of dimension k where*

$$0\leq k-s<(n-s)/2.$$

Assume that Γ' is non-degenerate in the sense of Theorem 1.1, and that m satisfies (9)*. Then one has*

$$T: L^{p_1}\times\ldots\times L^{p_{n-s-1}}\times A\times\ldots\times A\to L^{p'_{n-s}}, \tag{88}$$

$1<p_i\leq\infty$ *for* $i=1,\ldots,n-s-1$,

$$\frac{1}{p_1}+\ldots+\frac{1}{p_{n-s}}=1,$$

and

$$\frac{1}{p_{i_1}}+\ldots+\frac{1}{p_{i_r}}<\frac{(n-s)-2(k-s)+r}{2}$$

for all $1\leq i_1<\ldots<i_r\leq n-s$ *and* $1\leq r\leq n-s$.

Thus, for instance, the trilinear Hilbert transform maps $L^p\times L^q\times A$ to L^r whenever $1<p,q\leq\infty$, $1/p+1/q=1/r$, and $2/3<r<\infty$.

Proof. Let $g_1, \ldots, g_s$ be elements in the unit ball of A, let T_g denote the $(n-s-1)$-linear operator

$$T_g(f_1, \ldots, f_{n-s-1}) = T(f_1, \ldots, f_{n-s-1}, g_1, \ldots, g_s),$$

and let Λ_g be the associated $(n-s)$-form as in (2). We need to show that

$$T_g : L^{p_1} \times \ldots \times L^{p_{n-s-1}} \to L^{p'_{n-s}}.$$

By the reductions in Section 3 it suffices to show that Λ_g is of restricted type p for all exponent $(n-s)$-tuples p such that $1 < p_i < 2$ for all indices i which are not equal to the bad index j of p, and

$$(k-s) - (n-s) + 2 > \frac{1}{p_j} > (k-s) - (n-s) + \frac{3}{2}.$$

Fix p; by symmetry we may assume that p has bad index $n-s$. Let $E_1, \ldots, E_{n-s}$ be sets of finite measure. We have to find a major subset E'_{n-s} of E_{n-s} such that

$$|\Lambda_g(F_1, \ldots, F_{n-s})| \lesssim |E|^{1/p} \tag{89}$$

for all $F_i \in X(E'_i)$. By scaling we may take $|E_{n-s}| = 1$.

We choose E'_{n-s} to be the set defined by (25), with n replaced by $n-s$ throughout. Since (89) is sub-additive in g, and the unit ball of A is the convex hull of the plane waves, we may assume that each g_j is a plane wave $g_j(x) = e^{2\pi i x \xi_j}$ for some constants ξ_j. By modulating the F_i suitably, and translating the symbol m by a direction in Γ', one may set $\xi_j = 0$. The functions g are now completely harmless, and the claim follows from Theorem 3.5 with n replaced by $n-s$. □

References

[1] Calderon, C. *On commutators of singular integrals*, Studia Math. **53** (1975), no. 2, 139–174. MR **52:**1418

[2] Coifman, R. R and Meyer, Y. *On commutators of singular integrals and bilinear singular integrals*, Trans. Amer. Math. Soc **212** (1975), 315–331. MR **52:**1144

[3] Coifman, R. R and Meyer, Y. *Commutateurs d'integrales singulières et opérateurs multilinéaires*, Ann. Inst. Fourier (Grenoble) **28** (1978), no. 3, xi, 177–202. MR **80a:**47076

[4] Coifman, R. R and Meyer, Y. *Fourier analysis of multilinear convolutions, Calderón's theorem, and analysis of Lipschitz curves*, Euclidean harmonic analysis (Proc. Sem., Univ. Maryland, College Park, Md.), pp. 104–122, Lecture Notes in Math., 779 [1979]. MR **81g:**42021

[5] Coifman, R. R and Meyer, Y. *Au delà des opérateurs pseudo-différentiels*, Astérisque 57, Société Mathématique de France, Paris 1978. MR **81b:**47061

[6] Coifman, R. R and Meyer, Y. *Non-linear harmonic analysis, operator theory and P.D.E.*, Beijing lectures in harmonic analysis (Beijing 1984), 3–45, Annals of Math. Studies 112, Princeton Univ. Press, Princeton NJ 1986. MR **88e:**42030

[7] Coifman, R. R and Meyer, Y. *Ondelettes et opérateurs III, Opérateurs multilinéaires*, Actualités Mathématiques, Hermann, Paris 1991. MR **93i:**42004

[8] Fefferman, C. *Pointwise convergence of Fourier series*, Ann. of Math. (2) **98**, no.3, 551–571. MR **49:**5676

[9] Gilbert, J. and Nahmod, A. *Boundedness of bilinear operators with non-smooth symbols,* Math. Res. Lett. **7** (2000), no. 5-6, 767–778, CMP 2001:07

[10] Gilbert, J. and Nahmod, A. *Bilinear Operators with Non-Smooth Symbols I.* J. Fourier Anal. and Appl. **7** (2001), 437–469.

[11] Gilbert, J. and Nahmod, A. *L^p - Boundedness of Time-Frequency Paraproducts*, to appear in J. Fourier Anal. and Appl.

[12] Grafakos, L. and Torres, R. *On multilinear singular integrals of Calderon-Zygmund type*, to appear in the Proceedings of the El Escorial Conference held in El Escorial, Spain, July 3–7, 2000.

[13] Janson, S., *On interpolation of multilinear operators*, Function spaces and applications (Lund 1986), 290–302, Lecture Notes in Math. 1302, Springer, Berlin-New York, 1988. MR **89g:**46114

[14] Kenig, C. and Stein, E. *Multilinear estimates and fractional integration*, Math. Res. Lett. **6** (1999), no. 1, 1–15. MR **2000k:**42023a

[15] Lacey, M. and Thiele, C. *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$,* Ann. of Math. (2) **146** (1997), no. 3, 693–724. MR **99b:**42014

[16] Lacey, M. and Thiele, C. *On Calderon's conjecture*, Ann. of Math. (2) **149** (1999), no. 2, 475–196. MR **2000d:**42003

[17] Stein, E. *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals,* Princeton University Press, Princeton NJ 1993. MR **95c:**42002

[18] Thiele, C. *On the Bilinear Hilbert transform*, Universität Kiel, Habilitationsschrift [1998].

Department of Mathematics, Brown University, Providence, Rhode Island 02912

E-mail address: `camil@math.brown.edu`

Current address: Department of Mathematics, University of California at Los Angeles, Los Angeles, California 90095-1555

E-mail address: `camil@math.ucla.edu`

School of Mathematics, University of New South Wales, Sydney, New South Wales 2052, Australia

Current address: Department of Mathematics, University of California at Los Angeles, Los Angeles, California 90095-1555

E-mail address: `tao@math.ucla.edu`

Department of Mathematics, University of California at Los Angeles, Los Angeles, California 90095-1555

E-mail address: `thiele@math.ucla.edu`