

ON PERMUTATIONS OF LACUNARY INTERVALS

KATHRYN E. HARE AND IVO KLEMES

ABSTRACT. Let $\{I_j\}$ be an interval partition of the integers and consider the Littlewood-Paley type square function $S(f) = (\sum |f_j|^2)^{1/2}$ where $\hat{f}_j = \hat{f}\chi_{I_j}$. We prove that if the lengths $\ell(I_j)$ of the intervals I_j satisfy $\ell(I_{j+1})/\ell(I_j) \rightarrow \infty$, then $\|S(f)\|_p \approx \|f\|_p$ for $1 < p < \infty$. As these intervals need not be adjacent, such partitions can be thought of as permutations of lacunary intervals. This work generalizes the specific partition considered in a previous paper [H2]. We conjecture that it suffices to assume $\ell(I_{j+1})/\ell(I_j) \geq \lambda > 1$, and we also conjecture a necessary and sufficient condition.

1. INTRODUCTION

Let $\{I_j\}_{j=1}^\infty$ be disjoint intervals in \mathbf{Z} whose union is \mathbf{Z} , i.e. a partition of \mathbf{Z} into intervals. For a fixed partition $\mathcal{P} = \{I_j\}$ we define an operator $S_{\mathcal{P}}$ on $L^p(\mathbf{T})$, \mathbf{T} = the circle group, by

$$S_{\mathcal{P}}(f) = \left(\sum |f_j|^2 \right)^{1/2}, \quad f \in L^p(\mathbf{T}),$$

where $\hat{f}_j = \hat{f}\chi_{I_j}$. We call $S_{\mathcal{P}}(f)$ the Littlewood-Paley square function determined by \mathcal{P} . One general (unsolved) problem concerning $S_{\mathcal{P}}$ is to determine which interval partitions \mathcal{P} have the property that the norms $\|S_{\mathcal{P}}(f)\|_p$ and $\|f\|_p$ are equivalent for all $1 < p < \infty$, i.e.

$$(1) \quad A_p \|f\|_p \leq \|S_{\mathcal{P}}(f)\|_p \leq B_p \|f\|_p, \quad f \in L^p(\mathbf{T}),$$

where $A_p, B_p > 0$ depend only on p and \mathcal{P} . We will say that \mathcal{P} is a Littlewood-Paley (L-P) partition if (1) holds for all $1 < p < \infty$.

The motivation for this problem is the Littlewood-Paley theorem, which states that this does hold when \mathcal{P} consists of $(-\infty, 0)$ and intervals of the form $[n_j, n_{j+1})$ with $n_{j+1} \geq \lambda n_j \geq 0$ for a fixed $\lambda > 1$ ("lacunary intervals"). Other examples of L-P partitions can be generated with lacunary intervals by using vector-valued operators and the Marcinkiewicz-Zygmund theorem. For example, Marcinkiewicz [M] showed that the partition \mathcal{P} determined by the integers $2^j + 2^k$ is L-P. This was later generalized as in [G] and [S]. A comprehensive account can be found in [E], where the above general problem is mentioned [p. 162]. In two previous papers, we identified a type of partition \mathcal{P} not covered by the previous results [H1], and then proved that it was L-P

Received by the editors December 17, 1993 and, in revised form, November 22, 1994.

1991 *Mathematics Subject Classification.* Primary 42B25; Secondary 42A45.

Both authors were supported by NSERC Canada grants.

[H2]. Although this set was not obtained by iterating lacunary partitions, there was a relationship to lacunarity, namely that the set of interval lengths, when rearranged in increasing order, was a lacunary set. This suggested the conjecture that any permutation of a lacunary partition is L-P. In this paper we prove the following special case of this conjecture:

Theorem 1. *Fix a sequence of real numbers $\lambda_j \rightarrow \infty$. Then any interval partition $\mathcal{P} = \{I_j\}$ such that the lengths $\ell(I_j)$ satisfy*

$$\ell(I_{j+1}) \geq \lambda_j \ell(I_j)$$

has the Littlewood-Paley property. Moreover, we can take A_p and B_p to depend only on the exponent p and the sequence $\{\lambda_j\}$.

At first sight, Theorem 1 may appear to be a special case of the Littlewood-Paley theorem, since we require $\lambda_j \rightarrow \infty$ instead of just $\lambda_j \geq \lambda > 1$. However, the main point of the theorem is that the positions of the I_j are not specified; I_{j+1} need not be adjacent to I_j . One may think of $\{I_j\}$ as an arbitrary rearrangement or permutation of lacunary intervals $[n_j, n_{j+1})$ (with $n_{j+1}/n_j \rightarrow \infty$) so as to form a partition of \mathbf{Z} . Hence the theorem is not a special case of the Littlewood-Paley theorem. In fact, in our first paper [H1], we proved in a formal sense that such partitions are not covered by the classical results in all cases. On the other hand, we do suspect that Theorem 1 is still far from the most general picture. In particular we still suspect that any permutation of a lacunary partition has the L-P property, in other words, that $\lambda_j \rightarrow \infty$ can be replaced by $\lambda_j \geq \lambda > 1$. More generally we conjecture that: \mathcal{P} is L-P iff \mathcal{P} “does not separate arbitrarily large parallelepipeds”. Further details are given at the end of the paper. We should also mention that our proof of Theorem 1 goes via the following essentially stronger result:

Theorem A. *Fix $p \in (1, \infty)$. Then there exist real $\lambda_p, A_p, B_p > 0$ such that for any interval partition $\mathcal{P} = \{I_j\}$ satisfying $\ell(I_{j+1}) \geq \lambda_p \ell(I_j)$ we have*

$$A_p \|f\|_p \leq \|S_{\mathcal{P}}(f)\|_p \leq B_p \|f\|_p, \quad f \in L^p.$$

Thus, we take $\lambda_j \rightarrow \infty$ in Theorem 1 in order to get $\lambda_j \geq \lambda_p$ eventually for all p . However, we will not explicitly prove Theorem A in this paper; it is implied by our Theorem 2 below.

Lastly, let us point out that most of the paper will concentrate on the proof of only “one quarter” of the inequalities contained in (1). Namely, it will suffice to prove

$$(2) \quad \|f\|_p \leq B_p \|S_{\mathcal{P}}(f)\|_p \quad \text{for } 2 \leq p < \infty.$$

The reason for this is twofold: First, it is known that

$$\|f\|_p \leq C_p \|S_{\mathcal{P}} f\|_p, \quad f \in L^p,$$

if and only if

$$\|S_{\mathcal{P}} f\|_q \leq C_q \|f\|_q, \quad f \in L^q,$$

where $1 < p, q < \infty, 1/p + 1/q = 1$. This is a consequence of a duality argument, the M. Riesz theorem and the Marcinkiewicz-Zygmund theorem. And secondly, the remarkable theorem of Rubio de Francia [R] (see also [B]) states that, for any interval partition \mathcal{P} and any $2 \leq p < \infty$,

$$\|S_{\mathcal{P}} f\|_p \leq C_p \|f\|_p, \quad f \in L^p.$$

Hence we only need to check the inequalities (2) above. We shall do this by working with the even integers $p = 2n$, $n = 1, 2, \dots$, and invoking interpolation results afterwards.

2. TREES

We now define trees of intervals with some special properties (see Figure 0). Let the root node be a bounded interval in \mathbf{R} . Let each node have either one or two subnodes, which are disjoint subintervals. Moreover, let all node intervals in a given generation (henceforth level) have the same length, and let these lengths decrease rapidly. On each level, we allow at most one node to branch into two subnodes; the others have just one subnode. Lastly, a subnode must sit essentially in the middle of its parent node. In the case of two subnodes, both should be near the middle, but not too close to each other. This will be made precise shortly.

The intervals of a tree T generate an interval partition (of \mathbf{R} say) and hence a square function operator S_T . (Two of the intervals are unbounded, but this is irrelevant.) The special properties above will allow us to expand and estimate $\int |f|^{2n}$ in terms of $\int S_T(f)^{2n}$ (or in terms of a finite refinement of $S_T(f)$). This will be stated as Theorem 2, and it will be used to prove Theorem 1. We now list precise forms of the above tree properties.

Let $\lambda \geq 100$ and $\sigma > 0$. A tree with the properties below will be called a (λ, σ) -tree.

- (a) On any level j , all node intervals have the same length, which will be denoted by ℓ_j . Furthermore: $\ell_j \geq \lambda \ell_{j+1}$, $j = 0, 1, 2, \dots$
- (b) All nodes have either one subnode or two subnodes. These will be called nodes of type 1 and type 2, respectively. Each level contains at most one type 2 node.

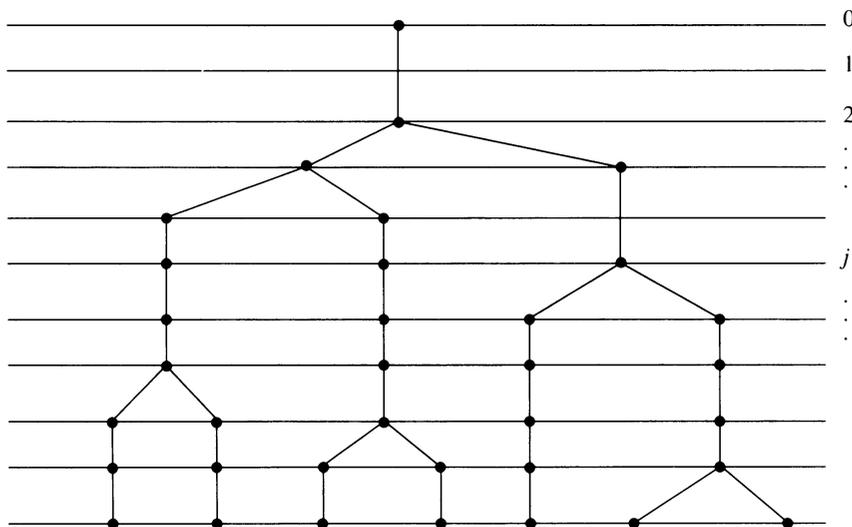


FIGURE 0. Schematic picture of the type of tree being considered. Each dot is a node, and the nodes on the same horizontal line constitute a level.

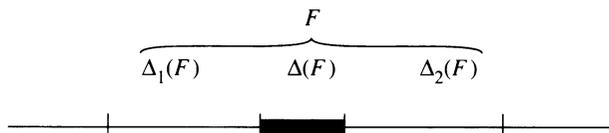


FIGURE 1

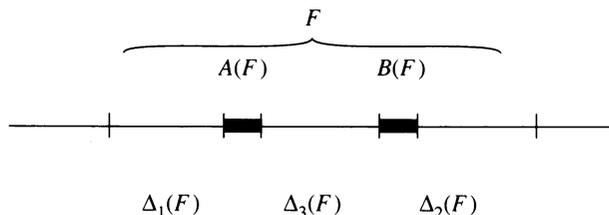


FIGURE 2

(c) The two types of nodes have the following structures:

Type 1. Let F denote the node, $A(F)$ the subnode, and let $\Delta(F) = F \setminus A(F) = \Delta_1(F) \cup \Delta_2(F)$ consist of the two intervals $\Delta_1(F)$ and $\Delta_2(F)$, as shown in Figure 1. Then

$$(T1) \quad 1/2 \leq \ell(\Delta_1)/\ell(\Delta_2) \leq 2,$$

where $\ell(\cdot)$ denotes the diameter of a set.

Type 2. Let $A(F)$ and $B(F)$ denote the two subnodes of the node F , and let

$$\Delta(F) = F \setminus (A(F) \cup B(F)) = \Delta_1(F) \cup \Delta_2(F) \cup \Delta_3(F)$$

consist of the three intervals $\Delta_i(F)$, as shown in Figure 2.

Then, for F on level j (so $\ell(F) = \ell_j$, and $\ell(A) = \ell(B) = \ell_{j+1}$),

$$(T2) \quad 1/2 \leq \ell(\Delta_1)/\ell(\Delta_2) \leq 2,$$

$$(T3.1) \quad \ell_j \geq \lambda \ell(\Delta_3),$$

$$(T3.2) \quad \ell(\Delta_3) \geq \lambda \ell_{j+1},$$

$$(T3.3) \quad \ell(\Delta_3) \geq \sigma \ell_j.$$

We will always use trees with a finite number of nontrivial levels. But for convenience the index j will run to ∞ . By convention this will mean that, on the last nontrivial level, we declare all F to be of type 1 with $A(F) = \emptyset$, $\Delta(F) = F$, and so for all subsequent levels the nodes are empty.

Theorem 2. Let $n \in \mathbb{N}$ and $\sigma > 0$. Then there exists a constant $C = C_{n,\sigma}$ such that, for any (λ, σ) -tree T with $\lambda \geq 100n$, we have

$$\|f\|_{2n} \leq C \|S_T(f)\|_{2n}$$

for all $f \in L^{2n}(\mathbb{T})$. Here S_T is the square function corresponding to the partition of \mathbb{Z} generated by T .

Remarks. When there are only type 1 nodes, there is in effect no parameter σ , and the “tree” T reduces to a classical lacunary partition. When there is a type

2 node, then the “nonvacuous” range for σ is $0 < \sigma \leq 1/\lambda$, since $\lambda\sigma \leq 1$ by properties (T3.1) and (T3.3). Also, recently we have been able to prove the theorem with $C = C_n$ not depending on σ , thus making property (T3.3) unnecessary. But the present theorem will suffice for the proof of Theorem 1, because there we use a specific tree with $\sigma = \sigma(n)$, $\lambda = \lambda(n)$, and hence $C = C(n)$.

The proof of Theorem 2 begins in section 5. At this point we derive some consequences of the tree properties. Some of our assertions and formulae will be easier to write if the subnodes of type 1 and type 2 nodes are denoted the same way. For this purpose, let $B(F) = \emptyset = \Delta_3(F)$ whenever F is a type 1 node. Let $n \in \mathbb{N}$ and $\lambda \geq 100n$. We claim that, for each fixed level j and F on level j ,

(T4) $\text{dist}(A(F) \cup B(F), \mathbf{R} \setminus F) > 2n \max\{\ell(A(G) \cup \Delta_3(G) \cup B(G)) : G \text{ on level } j\},$

(T5) $\text{dist}(A(F), B(F)) > 2n\ell_{j+1},$

(T6) $\min(1/10, \sigma/2)\ell_j + (n - 2)\ell_{j+1} < \ell(\Delta_i(F)), \quad i = 1, 2, 3.$

Proof of (T4). For each G on level j ,

$$\ell(A(F) \cup \Delta_3(G) \cup B(G)) \leq 2\ell_{j+1} + \ell(\Delta_3(G)) \leq 2\ell_{j+1} + \frac{1}{\lambda}\ell_j \leq (3/\lambda)\ell_j.$$

On the other hand,

$$\text{dist}(A(F) \cup B(F), \mathbf{R} \setminus F) = \min\{\ell(\Delta_1(F)), \ell(\Delta_2(F))\}.$$

But $\ell(\Delta_1) + \ell(\Delta_2) \geq (1 - 3/\lambda)\ell_j \geq \frac{97}{100}\ell_j$. And, by (T1) or (T2),

$$\begin{aligned} \min\{\ell(\Delta_1), \ell(\Delta_2)\} &\geq \frac{1}{3}(\ell(\Delta_1) + \ell(\Delta_2)) \\ &\geq \frac{97}{300} \cdot \frac{\lambda}{3} \ell(A \cup \Delta_3 \cup B). \end{aligned}$$

But $97\lambda/900 \geq 97n/9 > 2n$; so we are done.

Proof of (T5). This follows from (T3.2) since $\text{dist}(A, B) = \ell(\Delta_3)$.

Proof of (T6). There are two cases, according to when there is or is not a type 2 node on level j . If there is none, there is no Δ_3 to consider. Then

$$(n - 2)\ell_{j+1} \leq ((n - 2)/\lambda)\ell_j \leq (1/100)\ell_j$$

and

$$(1/10)\ell_j + (n - 2)\ell_{j+1} \leq (0.11)\ell_j.$$

On the other hand,

$$\ell(\Delta_i) \geq \frac{1}{3} \left(1 - \frac{1}{\lambda}\right) \ell_j \geq \frac{99}{300} \ell_j > (0.11)\ell_j.$$

If there is a type 2 node, we have

$$(\sigma/2)\ell_j < (1/2)\ell(\Delta_3) \quad \text{by (T3.3)}$$

and

$$(n - 2)\ell_{j+1} < (1/100)\ell(\Delta_3) \quad \text{by (T3.2)}.$$

On the other hand, it is clear that this $\ell(\Delta_3)$ is the minimum $\ell(\Delta_i(F))$ over all F on level j . So we are done, since $1/2 + 1/100 < 1$.

3. REDUCING THEOREM 1 TO THEOREM 2

Here we explain informally how to obtain trees from the partition in the statement of Theorem 1. The essential case of the construction can be stated as follows:

Proposition. *Let $n \in \mathbf{N}$. Then there are $\alpha(n), \lambda(n) \geq 100n$ and $\sigma(n) > 0$ such that if J_1, J_2, \dots are disjoint intervals whose union is a bounded interval J and whose lengths satisfy $\ell(J_{k+1}) \geq \alpha(n)\ell(J_k)$, $k = 1, 2, \dots$, then there is a $(\lambda(n), \sigma(n))$ -tree T which refines the partition of \mathbf{R} generated by the J_k . Moreover, T can be chosen so that each J_k (and each of the two components of $\mathbf{R} \setminus J$) is refined (partitioned) into lacunary blocks with uniform ratio ≥ 2 .*

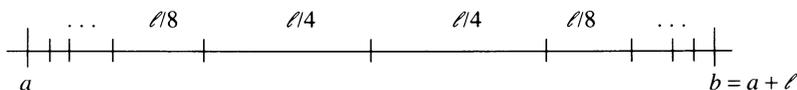


FIGURE 3. Example of an interval $[a, b]$ refined into lacunary blocks with ratio 2.

The proof of the proposition is not difficult. We omit the details and give only an outline. Let the root node of the tree be an interval F with the same centre as $J = \bigcup J_k$ but much larger than J . The main idea is that if α is large enough, then the largest of the J_k is much larger than the union of the remaining J_k , which split into at most two components, J^A and J^B . Thus, in general, we can define either one or two subnodes of F , namely $A(F)$ and $B(F)$, centred on J^A and J^B respectively, with appropriate lengths. There is also the possibility that we still want only one subnode $A(F)$ centred on J again, with length $\ell(F)/\lambda(n)$. That is, we must wait for the length of F to become comparable to the largest of the J_k (i.e. the distance between J^A and J^B) before deciding to make the split into $A(F) \supset J^A$ and $B(F) \supset J^B$. Then we iterate the procedure. We leave it to the reader to think of a precise procedure that will yield the (λ, σ) -tree properties. To see that there will be at most one splitting (type 2 node) on a given level, observe that there is always only one largest J_k remaining (in the union of the nodes on the given level), and the next largest is not comparable to it if α is large enough.

Remarks. Let S_T be the square function determined by the tree T of the proposition. Assume that Theorem 2 holds. Then $\|f\|_{2n} \leq C(n)\|S_T(f)\|_{2n}$ for $f \in L^{2n}$. Also $\|S_T(f)\|_{2n} \leq B(n)\|f\|_{2n}$ for $f \in L^{2n}$ by Rubio de Francia's theorem. Hence the $2n$ -norms are comparable. By considering the equivalent multiplier formulation (see [E]) it can be seen that $S_A(f)$ and f have comparable $2n$ -norms for any partition A which T is a refinement of. In particular $\|f\|_{2n} \leq C_A(n)\|S_A f\|_{2n}$ when A is the partition of the proposition determined by J_1, J_2, \dots .

This latter inequality can be proven by an alternative argument which does not invoke Rubio de Francia's theorem, as follows: From the proposition, we have that T is in fact a lacunary refinement of A . By standard arguments involving the Littlewood-Paley theorem and Marcinkiewicz-Zygmund theorem,

$S_A f$ and $S_T f$ have equivalent L^p -norms for any $1 < p < \infty$. Hence, assuming Theorem 2 again,

$$\|f\|_{2n} \leq C(n)\|S_T(f)\|_{2n} \leq C_A(n)\|S_A(f)\|_{2n}, \quad f \in L^{2n}.$$

The Proof That Theorem 2 Implies Theorem 1. First consider the case $p = 2n$ of Theorem 1. It suffices to prove the inequality for all trigonometric polynomials f . Given f , there is an interval I which is a union of some partition intervals I_j and contains the support of \hat{f} . Also, there is a j_0 such that $\lambda_j \geq \alpha_0(n)$ whenever $j \geq j_0$. Moreover, for a fixed sequence $\{\lambda_j\}$, we can choose $j_0 = j_0(n)$ depending on n only. Removing the intervals I_1, \dots, I_{j_0} from I leaves at most $j_0 + 1$ component intervals J , each of which is a union of intervals I_j for $j > j_0$. For a fixed J , we can label these I_j as J_k with $\ell(J_{k+1}) \geq \alpha_0(n)\ell(J_k)$, and apply the proposition above. Then apply Theorem 2 to the tree T . By the remarks following the proposition, it follows that

$$\|f_J\|_{2n} \leq C_n \|S(f_J)\|_{2n}$$

where $\hat{f}_J = \hat{f}\chi_J$ and S corresponds to the original partition $\{I_j\}$ (which however is in effect restricted to J for the moment).

Finally, combine these inequalities over the different J 's (and I_1, \dots, I_{j_0}) to obtain

$$\|f\|_{2n} \leq C_{n,j_0} \|S(f)\|_{2n} = C_n \|S(f)\|_{2n},$$

since j_0 depends only on n and the sequence $\{\lambda_j\}$ which is associated with the partition $\{I_j\}$. Here we first used the triangle inequality

$$\|f\|_{2n} \leq \sum_{j=1}^{j_0} \|f_{I_j}\|_{2n} + \sum_J \|f_J\|_{2n},$$

then Theorem 2 and the above result for each J , and then the fact that

$$\sum_{j=1}^{j_0} |f_{I_j}|^2 + \sum_J |S(f_J)|^2 = |S(f)|^2$$

to recombine the results.

This gives the $p = 2n$ case of Theorem 1; we obtain the general case by "interpolation". To be specific, we now have that S is a bounded sublinear operator on L_{q_n} for $1/q_n + 1/2n = 1$ by the remarks in section 1. Hence, by the Marcinkiewicz interpolation theorem it is bounded for all $q \in (1, 2]$. This again is equivalent to $\|f\|_p \leq B_p \|Sf\|_p$, $p \in [2, \infty)$, so we are done.

4. PRELIMINARIES TO THE PROOF OF THEOREM 2

The proof of Theorem 2 is rather computational but completely elementary. Therefore it will be a good idea to single out some identities and tricks which are used in the proof. We also recommend that the reader first consider the case $n = 2, p = 2n = 4$ in the sequel, especially for the proof of Lemma 5.3.

4.1. Summations over nodes. For an arbitrary partition $\{\Delta_j\}$ it is customary to write $\hat{f}_{\Delta_j} = \hat{f}\chi_{\Delta_j}$,

$$f = \sum_j f_{\Delta_j}, \quad \text{and} \quad S(f)^2 = \sum_j |f_{\Delta_j}|^2.$$

In our situation it is more useful to replace the arbitrary index j by the “tree structure”. Thus, if F ranges over the nodes of the tree T , we have

$$f = \sum_F f_{\Delta(F)},$$

provided \hat{f} is supported in the tree T , where $\Delta(F) = F \setminus (A(F) \cup B(F))$ is the interval F without its subnode intervals, $A(F)$ and $B(F)$. Since $\Delta(F) = \Delta_1(F) \cup \Delta_2(F) \cup \Delta_3(F)$ as in Figures 1 and 2, we have also

$$f = \sum_F f_{\Delta(F)} = \sum_F \sum_{i=1,2,3} f_{\Delta_i(F)},$$

and

$$S_T(f)^2 = \sum_F \sum_{i=1,2,3} |f_{\Delta_i(F)}|^2.$$

4.2. Relevance of the tree properties. To estimate $\|f\|_{2n}$, we multiply out $|f|^{2n}$ in an organized fashion (see the section on “difference series” below). Then we encounter various terms, one of which looks like

$$\int f_X \bar{f}_Y |f_{J_1}|^2 \cdots |f_{J_{n-1}}|^2$$

where $X = A(F)$ for some node F , $Y \subset \mathbf{R} \setminus F$, and J_1, \dots, J_{n-1} are some intervals with lengths $\ell(J_i) < \text{dist}(X, Y)/n$. But this implies that the integral is in fact zero, since

$$(X - Y) \cap ((J_1 - J_1) + (J_2 - J_2) + \cdots + (J_{n-1} - J_{n-1})) = \varnothing.$$

For example, this situation occurs when J_1, \dots, J_{n-1} are nodes on the same level as $A(F)$, by property (T4). Thus we see that the tree properties ensure various orthogonality relations when we compute $\int |f|^{2n}$.

4.3. The notation $f_E = E$. Since f shall be fixed, it will do no harm to omit writing it. So from this point on, we write $f_E = E$. For example, we have

$$\sum_F f_{\Delta(F)} = \sum_F \Delta(F) = \sum_F \sum_{i=1,2,3} \Delta_i(F),$$

and also

$$f_X \bar{f}_Y |f_{J_1}|^2 \cdots |f_{J_{n-1}}|^2 = X \bar{Y} |J_1|^2 \cdots |J_{n-1}|^2.$$

It will be implied by the context when these same symbols denote the original intervals and sets, as in

$$(X - Y) \cap ((J_1 - J_1) + \cdots + (J_{n-1} - J_{n-1})) = \varnothing,$$

or in the notations $\ell(J_i)$ and $\text{dist}(X, Y)$.

4.4. Estimates. When an integral is not zero for any canonical reason, we estimate it instead. The inequality

$$|FG| \leq |F|^2 + |G|^2$$

is often useful because $|F|^2$ and $|G|^2$ are more “manageable” than $|FG|$. We also use

$$|FG| \leq \varepsilon |F|^2 + \frac{1}{\varepsilon} |G|^2$$

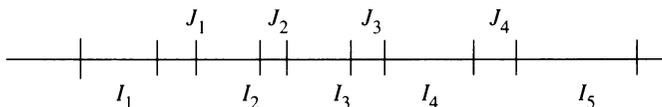


FIGURE 4. The case $N = 4$

for small $\varepsilon > 0$, when the $|F|^2$ term should have been on the other side of the inequality in some sense. This will be seen in the proof of the next lemma. The purpose of this lemma is to replace “partial square functions”, $\sum |F_k|^2$, by the whole function $|f|^2$. Here the F_k may be quite arbitrary blocks of the Fourier series of f . The lemma is reproduced from [H2] for completeness:

Lemma 4.1. *Let $I_1, J_1, I_2, J_2, I_3, \dots, J_N, I_{N+1}$ be adjacent intervals in \mathbf{Z} (see Figure 4). Let f be a trigonometric polynomial with $\text{supp } \hat{f} \subset I_1 \cup J_1 \cup I_2 \cup \dots \cup J_N \cup I_{N+1}$, and write $f = F_1 + \Delta_1 + F_2 + \dots + \Delta_N + F_{N+1}$, where $\hat{F}_k = \hat{f}\chi_{I_k}, \hat{\Delta}_k = \hat{f}\chi_{J_k}$.*

Let $v \geq 0$ be a trigonometric polynomial with

$$\text{deg } v < \min_k \ell(J_k).$$

Then

$$(*) \quad \int v \cdot \sum_{k=1}^{N+1} |F_k|^2 \leq 2 \int v |f|^2 + 70 \int v \cdot \sum_{k=1}^N |\Delta_k|^2.$$

Proof.

$$\begin{aligned} \int v |f|^2 &= \int v f \bar{f} \\ &= \int v \sum (|F_k|^2 + |\Delta_k|^2) + 2 \text{Re} \int v \cdot \sum (F_{k+1} \bar{\Delta}_k + \Delta_k \bar{F}_k + \Delta_k \bar{\Delta}_{k+1}), \end{aligned}$$

(since all other terms in the expansion of $f \bar{f}$ are orthogonal to v).

But

$$\begin{aligned} |F_{k+1} \bar{\Delta}_k| &\leq \varepsilon |F_{k+1}|^2 + \frac{1}{\varepsilon} |\Delta_k|^2, \\ |\Delta_k \bar{F}_k| &\leq \varepsilon |F_k|^2 + \frac{1}{\varepsilon} |\Delta_k|^2, \\ |\Delta_k \bar{\Delta}_{k+1}| &\leq |\Delta_k|^2 + |\Delta_{k+1}|^2, \end{aligned}$$

for any $\varepsilon > 0$. Therefore,

$$\int v |f|^2 \geq (1 - 4\varepsilon) \int v \sum |F_k|^2 + \left(1 - \frac{4}{\varepsilon} - 4\right) \int v \sum |\Delta_k|^2.$$

So choosing $\varepsilon = \frac{1}{8}$ gives

$$\int v |f|^2 \geq \frac{1}{2} \int v \sum |F_k|^2 - 35 \int v \sum |\Delta_k|^2,$$

which gives (*).

4.5. **Difference series.**

4.5.1. Our strategy is to multiply out $|f|^{2n}$ one node at a time. For each node F define

$$D_F = |F|^{2n} - |A(F)|^{2n} - |B(F)|^{2n},$$

recalling that $B(F) = 0$ when F is of type 1.

Then it is easy to see that

$$|f|^{2n} = \sum_F D_F.$$

We call this the “difference series” of $|f|^{2n}$. (We are assuming that \hat{f} is supported in the tree T , and that T has a finite number of nontrivial levels, so that the series is finite.)

4.5.2. When estimating $|\int D_F|$, we shall encounter products of the form $|F_1|^2 \cdots |F_n|^2$ where F_1, \dots, F_n are nodes on the same level. In the applications there will be at least two distinct F_i , but some of them may repeat. We then need to multiply out $|F_1|^2 \cdots |F_n|^2$ one node at a time also. For this purpose we now define another kind of “difference series”. Write

$$F_i = A(F_i) + B(F_i) + \Delta(F_i).$$

Then multiply out $|F_1|^2 \cdots |F_n|^2$ in terms of A 's, B 's, and Δ 's. Define

$$D_{F_1, \dots, F_n} = |F_1|^2 \cdots |F_n|^2 - (\text{terms of the form } |G_1|^2 \cdots |G_n|^2)$$

where $G_i = A(F_i)$ or $B(F_i)$. It should be noted that some terms arise in more than one way; we subtract all terms with the correct multiplicity, $C_{G_1 \dots G_n}^{F_1 \dots F_n}$. The definition of $D_{F_1 \dots F_n}$ can then be restated as the identity:

$$(**) \quad |F_1|^2 \cdots |F_n|^2 = \sum_{G_1, \dots, G_n} C_{G_1 \dots G_n}^{F_1 \dots F_n} |G_1|^2 \cdots |G_n|^2 + D_{F_1 \dots F_n}$$

where by convention the summation is over the nodes $G_i = A(F_i)$ or $B(F_i)$ such that terms giving the same product $|G_1|^2 \cdots |G_n|^2$ are collected (this occurs when some F_i repeats and is of type 2) and $C_{G_1 \dots G_n}^{F_1 \dots F_n}$ is the appropriate coefficient.

Example 1. Let $n = 3$, and let

$$F_1 = A(F_1) + \Delta(F_1) \equiv G + \Delta(F_1) \quad (\text{type 1}),$$

$$F_2 = F_3 = A(F_3) + B(F_3) + \Delta(F_3) \equiv A + B + \Delta(F_3) \quad (\text{type 2}).$$

Then

$$\begin{aligned} |F_1|^2 |F_2|^2 |F_3|^2 &= (G + \Delta(F_1)) \overline{(G + \Delta(F_1))} (A + B + \Delta(F_3))^2 \overline{(A + B + \Delta(F_3))^2} \\ &= |G|^2 |A|^4 + 4|G|^2 |A|^2 |B|^2 + |G|^2 |B|^4 + D_{F_1 F_2 F_3}. \end{aligned}$$

Thus in this example, $C_{GAB}^{F_1 F_2 F_3} = 4$, and $C_{GAA}^{F_1 F_2 F_3} = C_{GBB}^{F_1 F_2 F_3} = 1$.

Example 2. Suppose each F_i is a type 1 node. Then

$$\begin{aligned} |F_1|^2 \cdots |F_n|^2 &= \prod_{i=1}^n (A(F_i) + \Delta(F_i)) \overline{(A(F_i) + \Delta(F_i))} \\ &\equiv |A(F_1)|^2 \cdots |A(F_n)|^2 + D_{F_1 \dots F_n}. \end{aligned}$$

Remark. In general

$$0 \leq C_{G_1 \dots G_n}^{F_1 \dots F_n} \leq (n!)^2.$$

We are now in a position to define the “difference series” of $|F_1|^2 \dots |F_n|^2$. To define it, we begin with the identity (**) and then iterate it. That is, expand each $|G_1|^2 \dots |G_n|^2$ down to the next level, and keep doing this until a level is reached where all nodes are empty. The resulting difference series has the form

$$|F_1|^2 \dots |F_n|^2 = \sum_{\ell=j}^{\infty} \sum_{\substack{G_1, \dots, G_n \\ \text{on level } \ell}} C_{G_1 \dots G_n}^{F_1 \dots F_n} D_{G_1 \dots G_n}$$

where the F_1, \dots, F_n are on level j , and $G_i \subset F_i$ for each i . Note that we have implicitly extended the notation $C_{G_1 \dots G_n}^{F_1 \dots F_n}$ to each level $\ell \geq j$ (with $C_{F_1 \dots F_n}^{F_1 \dots F_n} = 1$). Although it may seem that iterating (**) will cause these coefficients to increase without bound, they actually remain bounded by $(n!)^2$ on all levels ℓ . This is because $C_{G_1 \dots G_n}^{F_1 \dots F_n}$ is in fact the coefficient that $|G_1|^2 \dots |G_n|^2$ would have if we directly multiplied out $|F_1|^2 \dots |F_n|^2$ down to level ℓ instead of one level at a time.

4.5.3. We now introduce a certain sum of products of the type $|F_1|^2 \dots |F_n|^2$. We call it X , for “cross terms”, and define it as follows. For each level j containing a type 2 node $F(j)$, let $A_j = A(F(j))$ and $B_j = B(F(j))$ be the two subnodes of $F(j)$ on level $j + 1$, and let

$$X_j = \sum_{\{F_1, \dots, F_n\} \supset \{A_j, B_j\}} |F_1|^2 \dots |F_n|^2.$$

The summation is over groups of nodes on level $j + 1$ which contain both A_j and B_j . (By convention F_1, \dots, F_n need not be all distinct, and they are not ordered.) If level j does not contain a type 2 node, let $X_j = 0$. Finally, define

$$X = \sum_{j=0}^{\infty} X_j.$$

It will be useful to also define

$$\begin{aligned} Y_j &= |A_j|^2 |B_j|^2 \left(\sum_{F \text{ on level } j+1} |F|^2 \right)^{n-2} \\ &= |A(F(j))|^2 |B(F(j))|^2 \left(\sum_{F \text{ on level } j} |A(F)|^2 + |B(F)|^2 \right)^{n-2}. \end{aligned}$$

Clearly,

$$X_j \leq Y_j \leq (n - 2)! X_j,$$

so that X_j and Y_j will be interchangeable when making estimates.

Now expand X by expanding each product $|F_1|^2 \dots |F_n|^2$ as a difference series of the type defined in the previous section. Then collect coefficients of identical $D_{G_1 \dots G_n}$ terms to obtain a series of the form

$$X = \sum_{j=1}^{\infty} \sum_{\substack{G_1, \dots, G_n \\ \text{on level } j}} C_{G_1 \dots G_n} D_{G_1 \dots G_n}.$$

We call this the difference series of X . By definition, the coefficients $C_{G_1 \dots G_n}$ are given by

$$C_{G_1 \dots G_n} = \sum_{F_1, \dots, F_n} C_{G_1 \dots G_n}^{F_1 \dots F_n},$$

where the summation is over all F_1, \dots, F_n which appear in some X_j and satisfy $F_i \supset G_i, i = 1, \dots, n$. It follows that:

1. If $G_1 = \dots = G_n$, then $C_{G_1 \dots G_n} = 0$,
2. $C_{G_1 \dots G_n} \leq (n - 1)(n!)^2$.

The first property holds because, in each X_j , the F_1, \dots, F_n include the two distinct nodes A_j and B_j , and hence at least two of the G_1, \dots, G_n are always distinct. For the second property, observe that A_j and B_j collapse into the same node at the previous level, j . So if G_1, \dots, G_n are fixed, there are at most $(n - 1)$ different (F_1, \dots, F_n) in the summation defining $C_{G_1 \dots G_n}$. But each term in the summation is bounded by $(n!)^2$.

5. THE PROOF OF THEOREM 2

Without loss of generality \hat{f} is supported in the tree T . Moreover, we can write $f = f_1 + f_2$ such that each \hat{f}_i is zero “between” alternative levels of the tree. That is, $\hat{f}_1 = 0$ on $F \setminus (A \cup B) = \Delta(F)$ whenever F is a node on an even-numbered level, and $\hat{f}_2 = 0$ on $\Delta(F)$ for every F on an odd-numbered level. Thus, we may assume that $f = f_i$ for either $i = 1$ or 2 . We now list the main steps of the proof as Lemmas 5.1 to 5.3. Recall the difference series (4.5.1)

$$|f|^{2n} = \sum_{\text{all } F} D_F$$

where

$$D_F = |F|^{2n} - |A(F)|^{2n} - |B(F)|^{2n}.$$

(We defined $B(F) = 0$ when F is a type 1 node.) In the following lemmas we also assume that $n \geq 2$.

Lemma 5.1. *For each node F define*

$$X_F = |A(F)|^2 |B(F)|^{2n-2} + |A(F)|^{2n-2} |B(F)|^2$$

and

$$\delta_F = |\Delta(F)|^2 |A(F)|^{2n-2} + |\Delta(F)|^2 |B(F)|^{2n-2} + |\Delta(F)|^{2n}$$

recalling that by convention $B(F) = 0$ for type 1 nodes F .

Then

$$\left| \int D_F \right| \leq C_n \int (X_F + \delta_F).$$

Lemma 5.2.

$$\int \sum_F \delta_F \leq C_{n, \sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n})$$

where S_1 is an m -refinement of S_T with $m = m(\sigma)$. (This means that the intervals associated with S_1 are obtained by dividing each interval of T into at most m subintervals.)

To state the next lemma, note that the above X_F is nonzero only if F is a type 2 node. Let $F = F(j)$ be the unique type 2 node on a level j . Recall the

definitions (4.5.3)

$$X_j = \sum_{\{F_1, \dots, F_n\} \supset \{A(F), B(F)\}} |F_1|^2 \cdots |F_n|^2$$

and

$$X = \sum_j X_j,$$

where the F_1, \dots, F_n are nodes on level $j+1$ and they may contain repetitions. Clearly $X_F \leq 2X_j$ and

$$\sum_F X_F = \sum_j X_{F(j)} \leq 2 \sum_j X_j = 2X.$$

Lemma 5.3. For any $0 < \varepsilon \leq 1$,

$$\int X \leq \varepsilon C_n \int X + \frac{1}{\varepsilon} C_{n,\sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n})$$

where S_1 is another m -refinement of S_T with $m = m(\sigma)$.

Remark. Our reason for introducing $X = \sum X_j$ was that we could not prove an estimate for $\sum X_F$ directly (except when $n = 2$ and one has $X_F = 2X_j$, so that $\sum X_F = 2X$).

Proof of Theorem 2. Taking $\varepsilon = 1/2C_n$ in Lemma 5.3 we have

$$\left(1 - \frac{1}{2}\right) \int X \leq 2C_n C_{n,\sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n}).$$

Now S_1 may denote the same (common) m -refinement in both Lemmas 5.2 and 5.3. Therefore

$$\begin{aligned} \int |f|^{2n} &\leq \sum_F \left| \int D_F \right| \leq C_n \sum_F \int (X_F + \delta_F) \\ &\leq C_n \left(\int X + \int \sum_F \delta_F \right) \\ &\leq C_{n,\sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n}) \\ &\leq \varepsilon \int |f|^{2n} + C_{n,\sigma,\varepsilon} \int S_1(f)^{2n} \end{aligned}$$

where the constants change from one line to the next. Taking $\varepsilon = 1/2$ and rearranging gives

$$\int |f|^{2n} \leq C_{n,\sigma} \int S_1(f)^{2n} \leq C_{n,\sigma} \int S_T(f)^{2n}$$

as required.

6. PROOFS OF LEMMAS 5.1, 5.2 AND 5.3

The proofs of Lemmas 5.1 and 5.2 will be seen to be similar to the proofs of Lemmas 3 and 4 (first part) in [H2].

Proof of Lemma 5.1.

$$\begin{aligned}
 |F|^{2n} &= F^n \overline{F^n} = (A + B + \Delta)^n \overline{(A + B + \Delta)^n} \\
 &= \sum A^a B^b \Delta^c \cdot \overline{A^\alpha B^\beta \Delta^\gamma} \cdot C_{a,b,c,\alpha,\beta,\gamma}
 \end{aligned}$$

where the $C_{a,b,c,\alpha,\beta,\gamma}$ are integers whose sum is $(3^n)^2$, and where we have written $A(F)$, $B(F)$, and $\Delta(F)$ as A , B and Δ respectively. Since D_F is the sum of all terms t other than $A^n \overline{A^n}$ or $B^n \overline{B^n}$, it suffices to show $|\int t| \leq C_n \int (X_F + \delta_F)$ for each such term.

When F is a type 1 node, the $B = 0$ so that $t = A^a \Delta^c \overline{A^\alpha \Delta^\gamma}$. When no Δ is present ($c = \gamma = 0$), we must have $A^n \overline{A^n}$, which was excluded. When there is a $\Delta = \Delta(F)$ present, we may assume that all Δ 's on the next level are zero, in particular $\Delta(A) = 0$. Regardless of whether the A is type 1 or type 2, the nonzero part of A (i.e. in effect $A = A \setminus \Delta(A)$) satisfies $\text{dist}(A, \Delta(F)) \geq n\mathcal{L}(A)$ by property (T4).

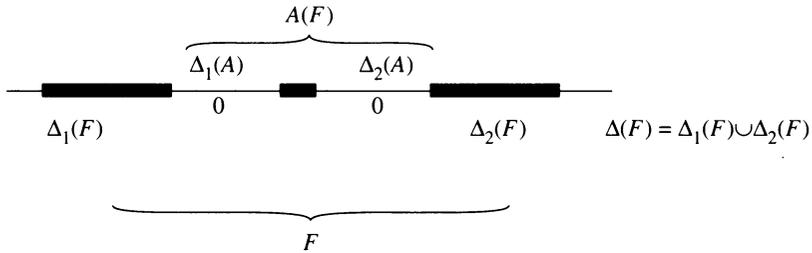


FIGURE 5

Therefore

$$\int A^{n-1} \Delta \overline{A^n} = \int (\overline{A} \Delta) (A \overline{A})^{n-1} = 0,$$

so that there is no need to consider $t = A^{n-1} \Delta \overline{A^n}$ or its complex conjugate $A^n \overline{A^{n-1} \Delta}$. This means that we are left with $t = A^a \Delta^c \overline{A^\alpha \Delta^\gamma}$ where $c + \gamma \geq 2$. But then

$$|t| = |A|^{a+\alpha} |\Delta|^{c+\gamma} \leq |A|^{2n-2} |\Delta|^2 + |\Delta|^{2n} = \delta_F$$

since $a + c = \alpha + \gamma = n$. So we have proved the lemma when F is a type 1 node.

Remark. If we had just a classical lacunary partition $\{\Delta_j\}$, and $A_j = \Delta_1 \cup \dots \cup \Delta_{j-1}$, then the proof of $\|f\|_{2n} \leq C_n \|Sf\|_{2n}$ would essentially consist of the previous paragraph plus Lemma 5.2.

Next, let F be a type 2 node, and consider the terms t with no Δ factor ($c = \gamma = 0$), other than $A^n \overline{A^n}$ or $B^n \overline{B^n}$. When $a = \alpha$ and $b = \beta$ we have, since $a + b = n$ and $1 \leq a, b \leq n - 1$,

$$\begin{aligned}
 t &= A^a B^b \overline{A^a B^b} = |A|^{2a} |B|^{2b} \\
 &= |A|^2 |B|^{2n-2} + |A|^{2n-2} |B|^2 = X_F.
 \end{aligned}$$

All other terms with no Δ factor are of the form

$$t = (A \overline{B})^k |A|^{2l} |B|^{2m}, \quad k \geq 1,$$

with $k + l + m = n$, or the complex conjugate of this. But the gap between A and B satisfies

$$\text{dist}(A, B) > n\ell(A) = n\ell(B) > l \cdot \ell(A) + m \cdot \ell(B)$$

by property (T5). Also, A is completely on one side of B (the left) so we can conclude that

$$\int t = \int (A\bar{B})^k |A|^{2l} |B|^{2m} = 0.$$

Remark. We have used the “total ordering” of \mathbf{Z} by saying A is to the left of B , but it was not really necessary. This is because the case $k = 1$ does not require it (then $\text{dist}(A, B) > l \cdot \ell(A) + m \cdot \ell(B)$ is enough), and when $k \geq 2$ we can instead estimate $|t|$, by

$$|t| \leq |A|^2 |B|^{2n-2} + |A|^{2n-2} |B|^2 = X_F.$$

Now consider terms with one Δ factor ($c + \gamma = 1$). If there are no B factors present ($b = \beta = 0$), then

$$t = A^{n-1} \Delta \bar{A}^n$$

or the conjugate, and $\int t = 0$ as in the first paragraph of the proof. The same holds for $t = B^{n-1} \Delta \bar{B}^n$ or its conjugate. So we may assume that both A and B are present ($a + \alpha \geq 1$ and $b + \beta \geq 1$). Then we claim that

$$|t| \leq X_F + \delta_F.$$

If $|t| = |A|^{a+\alpha} |B|^{b+\beta} |\Delta|$ and precisely one of the exponents $a + \alpha$ and $b + \beta$ is odd, say $a + \alpha$, then

$$\begin{aligned} |t| &= |A|^{2l+1} |B|^{2m} |\Delta| \quad (m \geq 1) \\ &\leq |A|^{2l} |B|^{2m} (|A|^2 + |\Delta|^2) \\ &\leq |A|^2 |B|^{2n-2} + |A|^{2n-2} |B|^2 + |\Delta|^2 (|A|^{2n-2} + |B|^{2n-2}) \\ &\leq X_F + \delta_F \end{aligned}$$

since $l + m = n - 1$.

Otherwise, we have terms with two or more Δ factors ($c + \gamma \geq 2$). For these we estimate

$$|t| = |A|^{a+\alpha} |B|^{b+\beta} |\Delta|^{c+\gamma} \leq (|A|^{2n-2} + |B|^{2n-2} + |\Delta|^{2n-2}) |\Delta|^2 = \delta_F,$$

thus completing the proof of the lemma.

Proof of Lemma 5.2. Observe that whenever a Δ is divided (refined) into several pieces, say $\Delta = \Delta_1 + \Delta_2 + \dots + \Delta_M$, we have

$$|\Delta|^2 \leq M \left(\sum_{i=1}^M |\Delta_i|^2 \right),$$

so that as long as $M \leq m$ is bounded over all the Δ 's we will have

$$S(f)^2 \equiv \sum_{\Delta} |\Delta|^2 \leq m \sum_{\Delta} \sum_i |\Delta_i|^2 \equiv m S_1(f)^2$$

where S_1 is an m -refinement of S .

So first of all we have

$$|\Delta(F)|^2 \leq 3(|\Delta_1(F)|^2 + |\Delta_2(F)|^2 + |\Delta_3(F)|^2)$$

for each $\Delta(F)$ in the definition of δ_F , and we can dispose of the $|\Delta(F)|^{2n}$ term by

$$\sum_F |\Delta(F)|^{2n} \leq \left(\sum_F |\Delta(F)|^2 \right)^n \leq 3^n \left(\sum_{F, i=1,2,3} |\Delta_i|^2 \right)^n \equiv 3^n S_T(f)^{2n}.$$

Later, S_T will be refined further to some S_1 . Now consider the term $|\Delta(F)|^2 |A(F)|^{2n-2}$, when F is type 2 for example. (The $|\Delta(F)|^2 |B(F)|^{2n-2}$ and the type 1 cases are both similar to the following, so we omit them.) The first step is to divide $\Delta(F)$ into enough subintervals Δ_{ij} , $i = 1, 2, 3$, $j = 1, \dots, M$, such that each has length $\ell(\Delta_{ij})$ satisfying

$$(n - 2)\ell(A) + \ell(\Delta_{ij}) < \ell(\Delta_3)$$

where Δ_3 is the middle piece of Δ (see Figure 2). This is possible with a uniform $M \leq m(\sigma)$, depending on σ , because of the tree property (T6). Then write

$$|\Delta_{ij}|^2 |A|^{2n-2} = |\Delta_{ij}|^2 |A|^{2n-4} |A|^2 \equiv v|A|^2$$

and apply Lemma 4.1, obtaining

$$\begin{aligned} \int v|A|^2 &\leq 2 \int v|f|^2 + 70 \int v(|\Delta_1|^2 + |\Delta_3|^2) \\ &\leq 2 \int v|f|^2 + 70 \int vS_T(f)^2. \end{aligned}$$

(The lemma applies with $N = 2$, $J_1 = \Delta_1$, $J_2 = \Delta_3$, $I =$ everything to the left of J_1 , $I_2 = A$, $I_3 =$ everything to the right of J_2 .) Now use the $\varepsilon, \frac{1}{\varepsilon}$ trick:

$$v|f|^2 = |\Delta_{ij}|^2 |A|^{2n-4} |f|^2 \leq \varepsilon |\Delta_{ij}|^2 |A|^{2n-2} + C(n, \varepsilon) |\Delta_{ij}|^2 |f|^{2n-2}$$

and

$$vS_T(f)^2 \leq \varepsilon |\Delta_{ij}|^2 |A|^{2n-2} + C(n, \varepsilon) |\Delta_{ij}|^2 S_T(f)^{2n-2},$$

so that with $\varepsilon = 1/140$ we get

$$\int |\Delta_{ij}|^2 |A|^{2n-2} \leq C_n \int |\Delta_{ij}|^2 (|f|^{2n-2} + S_T(f)^{2n-2}).$$

The point of this was that we have replaced the ‘‘partial sum’’ A by the whole function f , at the cost of an $S_T(f)$ term, which is harmless. Finally, sum the last inequality over i, j and then over all nodes F to get

$$\int \sum_F |\Delta(F)|^2 |A(F)|^{2n-2} \leq C_{n,m} \int S_1(f)^2 (|f|^{2n-2} + S_T(f)^{2n-2}).$$

This concludes the proof since the S_T terms can be refined into S_1 terms and $m = m(\sigma)$.

Remark. We have found a way to avoid property (T3.3), and the rather inelegant chopping of Δ_i . Instead, one can tolerate a few of the $F_k \bar{F}_{k+1}$ terms in the proof of Lemma 4.1, which are just like the X_F terms of Lemma 5.1 and will be handled in Lemma 5.3.

Proof of Lemma 5.3. The proof begins with more general versions of the proofs of Lemmas 5.1 and 5.2. We will therefore divide the proof into steps analogous to these lemmas. Step 1 is analogous to Lemma 5.1 and Step 2 is analogous to Lemma 5.2. The final step, Step 3, allows us to conclude the proof. Recall that X has a difference series $X = \sum_j \sum C_{F_1 \dots F_n} D_{F_1 \dots F_n}$ (4.5.3) with coefficients bounded in terms of n ;

$$C_{F_1 \dots F_n} \leq C_n.$$

Step 1. Let F_1, \dots, F_n be nodes on a level j . If they are all type 1 nodes, define $X_{F_1 \dots F_n} = 0$. If at least one is the (unique) type 2 node $F(j)$, say $F_i = F(j)$ (it may be repeated), define

$$X_{F_1 \dots F_n} = |A(F_i)| |B(F_i)| \sum_{p=1}^n |\Delta(F_p)| (|A(F_p)| + |B(F_p)|) \prod_{q \neq p, i} (|A(F_q)|^2 + |B(F_q)|^2)$$

recalling that $B(F) = 0$ when F is of type 1.

Also define $\delta_{F_1 \dots F_n}$ in all cases by

$$\delta_{F_1 \dots F_n} = \sum_{p=1}^n |\Delta(F_p)|^2 \prod_{q \neq p} (|A(F_q)|^2 + |B(F_q)|^2 + |\Delta(F_q)|^2).$$

Then we have

$$\left| \int D_{F_1 \dots F_n} \right| \leq C_n \int (X_{F_1 \dots F_n} + \delta_{F_1 \dots F_n}).$$

Step 2. For each level j define δ_j by

$$\delta_j = \left(\sum_{F \text{ on level } j} |\Delta(F)|^2 \right) \left(\sum_{F \text{ on level } j} |A(F)|^2 + |B(F)|^2 + |\Delta(F)|^2 \right)^{n-1}.$$

Then

$$\sum_{F_1 \dots F_n \text{ on level } j} \delta_{F_1 \dots F_n} \leq C_n \delta_j,$$

and

$$\int \sum_j \delta_j \leq C_{n, \sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n})$$

where S_1 is an m -refinement of S_T with $m = m(\sigma)$.

Step 3. Let j be a level containing a type 2 node $F(j)$. Recall the definition (4.5.3)

$$Y_j = |A(F(j))|^2 |B(F(j))|^2 \left(\sum_{F \text{ on level } j} |A(F)|^2 + |B(F)|^2 \right)^{n-2}.$$

Then for any $0 < \varepsilon \leq 1$,

$$\sum_{F_1, \dots, F_n \text{ on level } j} X_{F_1 \dots F_n} \leq \varepsilon Y_j + \frac{1}{\varepsilon} C_n \delta_j.$$

(Note that δ_j was defined in Step 2.)

Conclusion of the proof of Lemma 5.3. Recall that $Y_j \leq C_n X_j$. Therefore

$$\begin{aligned} \int X &\leq \sum_{j=1}^{\infty} \left(\sum_{F_1 \dots F_n \text{ on level } j} C_n \left| \int D_{F_1 \dots F_n} \right| \right) \\ &\leq \sum_j \sum C_n \int (X_{F_1 \dots F_n} + \delta_{F_1 \dots F_n}) \quad (\text{Step 1}) \\ &\leq C_n \sum_j \int \left(\varepsilon Y_j + \frac{1}{\varepsilon} C_n \delta_j + C_n \delta_j \right) \quad (\text{Steps 2 and 3}) \\ &\leq \varepsilon C_n \int \left(\sum_j X_j \right) + \frac{1}{\varepsilon} C_{n, \sigma} \int (S_1(f)^2 |f|^{2n-2} + S_1(f)^{2n}) \quad (\text{Step 2}). \end{aligned}$$

This concludes the proof since $X = \sum X_j$.

Proof of Step 1. Suppose that F_1, \dots, F_n are all type 1 nodes. Write $A(F_q) = A_q, \Delta(F_q) = \Delta_q, q = 1, \dots, n$.

Then

$$\begin{aligned} |F_1|^2 \dots |F_n|^2 &= \prod_{q=1}^n (A_q + \Delta_q) \overline{(A_q + \Delta_q)} \\ &= \sum \prod_{q=1}^n A_q^{a_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} \Delta_q^{\gamma_q}} \end{aligned}$$

where $a_q + c_q = 1 = \alpha_q + \gamma_q$. Now $D_{F_1 \dots F_n}$ is the sum of all terms t ,

$$t = \prod_{q=1}^n A_q^{a_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} \Delta_q^{\gamma_q}},$$

except those of the form $t = \prod |A_q|^2$, of which there is only one, given by $c_q = \gamma_q = 0$ for all q . When there is a Δ factor present in t , we may assume that all Δ 's on the next level are zero, in particular all $\Delta(A_q) = 0$, and thus $A_q = A_q \setminus \Delta(A_q)$. Suppose that there is exactly one Δ factor present, say Δ_p , so that $c_p = 1, \gamma_p = 0$ and $c_q = \gamma_q = 0, q \neq p$, and

$$t = \Delta_p \overline{A_p} \prod_{q \neq p} A_q \overline{A_q}.$$

Then $\int t = 0$ since

$$\text{dist}(A_p, \Delta_p) \geq 2n \max \ell(A_q) > \sum_{q=1}^n \ell(A_q)$$

by tree property (T4). So we are left with the case $\sum_{q=1}^n c_q + \gamma_q \geq 2$ (two or more Δ factors). In this case it is clear that $|t| \leq \delta_{F_1 \dots F_n}$, since

$$|A_q \Delta_q A_p \Delta_p| \leq |\Delta_q|^2 |A_p|^2 + |\Delta_p|^2 |A_q|^2,$$

and

$$|A_q^{a_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} \Delta_q^{\gamma_q}}| \leq |A_q|^2 + |\Delta_q|^2.$$

So we have proved Step 1 when all of the nodes are type 1.

Next, suppose that there is a type 2 node present, and write $F_q = A_q + B_q + \Delta_q$, so that

$$|F_1|^2 \cdots |F_n|^2 = \sum \prod_{q=1}^n A_q^{a_q} B_q^{b_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} B_q^{\beta_q} \Delta_q^{\gamma_q}}$$

where $a_q + b_q + c_q = 1 = \alpha_q + \beta_q + \gamma_q$.

Now, there is actually only one type 2 node, say F_i , repeated r times with $1 \leq r \leq n$. So we can write the terms of the sum in the form

$$t = A_i^a B_i^b \Delta_i^c \overline{A_i^\alpha B_i^\beta \Delta_i^\gamma} \prod_{F_q \text{ type 1}} A_q^{a_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} \Delta_q^{\gamma_q}}$$

with $a + b + c = r = \alpha + \beta + \gamma$ and $a_q + c_q = 1 = \alpha_q + \gamma_q$. Recall that $D_{F_1 \dots F_n}$ in this case consists of all terms other than $A_i^a B_i^b \overline{A_i^\alpha B_i^\beta} \prod A_q \overline{A_q}$. So first, if there is no Δ factor present in t , we must have

$$t = (A_i \overline{B_i})^k |A_i|^{2l} |B_i|^{2m} \prod |A_q|^2, \quad k \geq 1,$$

with $k + l + m = r$ or the complex conjugate of this. The gap between A_i and B_i satisfies

$$\text{dist}(A_i, B_i) \geq 2n\ell_{j+1} > l \cdot \ell(A_i) + m \cdot \ell(B_i) + \sum \ell(A_q)$$

by tree property (T5). Thus $\int t = 0$.

Now suppose there is exactly one Δ factor, and without loss of generality all $\Delta(A_q) = \Delta(B_q) = 0$ on the next level. Suppose that in addition, both A_i and B_i are present ($a + \alpha \geq 1$ and $b + \beta \geq 1$). Then clearly $|t| \leq X_{F_1 \dots F_n}$. For, if Δ_i is the one Δ factor, we have

$$|A_i^a B_i^b \Delta_i^c \overline{A_i^\alpha B_i^\beta \Delta_i^\gamma}| = |A_i| |B_i| |\Delta_i| |A_i|^x |B_i|^y$$

where $x + y = 2r - 3$, and

$$|A_i|^x |B_i|^y \leq (|A_i| + |B_i|)(|A_i|^2 + |B_i|^2)^z$$

where $x + y = 1 + 2z$. Or, if the one Δ factor is $\Delta_p = \Delta(F_p)$ for some F_p of type 1, then

$$|A_i^a B_i^b \overline{A_i^\alpha B_i^\beta}| \leq |A_i| |B_i| (|A_i|^2 + |B_i|^2)^z$$

where $a + b + \alpha + \beta = 2r = 2 + 2z$, and

$$\left| \prod_{q \neq p} A_q^{a_q} \Delta_q^{c_q} \overline{A_q^{\alpha_q} \Delta_q^{\gamma_q}} \right| \leq |\Delta_p| |A_p| \prod_{q \neq p} |A_q|^2.$$

Now suppose A_i and B_i are not both present; for example suppose that A_i is present, and there is only one Δ factor still. Then either

$$t = (A_i \overline{\Delta_i})(A_i \overline{A_i})^{r-1} \prod |A_q|^2$$

when Δ_i is present, or

$$t = |A_i|^{2r} (A_p \overline{\Delta_p}) \prod_{q \neq p} |A_q|^2$$

when some other $\Delta = \Delta_p$ is present. In both cases $\int t = 0$ by tree property (T4), just as in the case when F_1, \dots, F_n were all of type 1.

Finally, we have the case of two or more Δ factors present. Then $|t| \leq \delta_{F_1 \dots F_n}$ by the estimates already used in the type 1 case:

$$|A_q^w B_q^x \Delta_q A_p^y B_p^z \Delta_p| \leq |\Delta_q|^2 (|A_p|^2 + |B_p|^2) + |\Delta_p|^2 (|A_q|^2 + |B_q|^2)$$

(where $w + x = 1 = y + z$), and

$$|A_q^{\alpha_q} B_q^{\beta_q} \Delta_q^{c_q} A_q^{\alpha_q} B_q^{\beta_q} \Delta_q^{\gamma_q}| \leq |A_q|^2 + |B_q|^2 + |\Delta_q|^2.$$

Thus, the proof of Step 1 is complete.

Proof of Step 2. The first inequality is just algebra: if S is a set of indeterminates $S = \{s_1, \dots, s_k\}$, then

$$\sum_{(t_1, \dots, t_{n-1}) \in S^{n-1}} t_1 \cdots t_{n-1} = \left(\sum_{s_i \in S} s_i \right)^{n-1}$$

where $S^{n-1} = S \times S \times \dots \times S$ is the $(n - 1)$ -fold Cartesian product.

For the second part, fix j . We can first note that

$$\begin{aligned} & \left(\sum |A(F)|^2 + |B(F)|^2 + |\Delta(F)|^2 \right)^{n-1} \\ & \leq c_n \left(\sum |A(F)|^2 + |B(F)|^2 \right)^{n-1} + c_n \left(\sum |\Delta(F)|^2 \right)^{n-1}. \end{aligned}$$

The latter of these two terms can be disposed of since

$$\sum_j \left(\sum_{F \text{ on level } j} |\Delta(F)|^2 \right)^n \leq \left(\sum_{\text{all } F} |\Delta(F)|^2 \right)^n \leq 3^n S_T(f)^{2n}$$

(recall the proof of Lemma 5.2). This leaves us with

$$\sum_j \int \left(\sum_{F \text{ on level } j} |\Delta(F)|^2 \right) \left(\sum_{F \text{ on level } j} |A(F)|^2 + |B(F)|^2 \right)^{n-1}.$$

Here we would like to replace the second factor by $|f|^{2n-2}$, and this is precisely what Lemma 4.1 allows us to do. We fix a level j . Then divide each $\Delta(F)$ into enough intervals $\Delta_{ik}(F)$, $i = 1, 2, 3$, $k = 1, \dots, M$, such that their lengths satisfy

$$\ell(\Delta_{ik}(F)) + (n - 2)\ell_{j+1} < d_j \equiv \min_{\substack{F, \\ i=1, 2, 3}} \ell(\Delta_i(F)).$$

This d_j is the minimum distance between the components of

$$\bigcup_F (A(F) \cup B(F)) \cup \left(\mathbf{R} \setminus \bigcup_F F \right) = \mathbf{R} \setminus \bigcup \Delta_i(F)$$

(where the F ranges over level j only). This is possible with a uniform $M \leq m(\sigma)$ by tree property (T6). (Recall also that $\ell_{j+1} = \ell(A(F)) = \ell(B(F))$ for all F on level j .) Now let

$$\left(\sum_{F \text{ on level } j} |\Delta_{ik}(F)|^2 \right) \left(\sum_{F \text{ on level } j} |A(F)|^2 + |B(F)|^2 \right)^{n-2} = v,$$

so that $\deg v < d_j$, and apply Lemma 4.1.

In applying Lemma 4.1, we must introduce $\{H_k\}$, the component intervals of $\mathbf{R} \setminus \bigcup_{F \text{ on level } j} F$. Then

$$f = \sum_k H_k + \sum_{F \text{ on level } j} (A(F) + B(F) + \Delta(F)).$$

But the $|H_k|^2$ terms only make things larger; in other words

$$\begin{aligned} & \int v \cdot \sum_{F \text{ on level } j} (|A(F)|^2 + |B(F)|^2) \\ & \leq \int v \left(\sum_{F \text{ on level } j} (|A(F)|^2 + |B(F)|^2) + \sum_k |H_k|^2 \right) \\ & \leq 2 \int v|f|^2 + 70 \int v \sum_{F, i=1,2,3} |\Delta_i(F)|^2 \\ & \leq 2 \int v|f|^2 + 70 \int v S_T(f)^2. \end{aligned}$$

Now use the $\varepsilon, \frac{1}{\varepsilon}$ trick as in Lemma 5.2 and sum over j to finish the proof. The $S_1(f)$ is the $m(\sigma)$ -refinement defined by the $\Delta_{ik}(F)$.

Proof of Step 3. The nonzero $X_{F_1 \dots F_n}$ occur only when the n -tuple $F_1 \dots F_n$ contains the type 2 node $F(j)$. As in Step 2, we again have an algebraic identity

$$\sum_{F_1 \dots F_{n-2}} \prod_{q=1}^{n-2} (|A(F_q)|^2 + |B(F_q)|^2) = \left(\sum_{F \text{ on level } j} (|A(F)|^2 + |B(F)|^2) \right)^{n-2},$$

where $F_1 \dots F_{n-2}$ ranges over $(n-2)$ -tuples on level j . Therefore

$$\begin{aligned} & \sum_{F_1 \dots F_n \text{ on level } j} X_{F_1 \dots F_n} \\ & \leq C_n |A(F(j))| |B(F(j))| \left(\sum_{F \text{ on level } j} |\Delta(F)| (|A(F)| + |B(F)|) \right) \\ & \quad \cdot \left(\sum_{F \text{ on level } j} (|A(F)|^2 + |B(F)|^2) \right)^{n-2} \\ & \leq \varepsilon |A(F(j))|^2 |B(F(j))|^2 \left(\sum |A(F)|^2 + |B(F)|^2 \right)^{n-2} \\ & \quad + \frac{1}{\varepsilon} C_n^2 \left(\sum |\Delta(F)| (|A(F)| + |B(F)|) \right)^2 \left(\sum |A(F)|^2 + |B(F)|^2 \right)^{n-2} \\ & \leq \varepsilon Y_j + \frac{1}{\varepsilon} C_n^2 \left(\sum |\Delta(F)|^2 \right) \left(\sum (|A(F)| + |B(F)|)^2 \right) \\ & \quad \cdot \left(\sum |A(F)|^2 + |B(F)|^2 \right)^{n-2} \text{ (by the Cauchy-Schwarz inequality)} \\ & \leq \varepsilon Y_j + \frac{1}{\varepsilon} \cdot 2C_n^2 \left(\sum |\Delta(F)|^2 \right) \left(\sum |A(F)|^2 + |B(F)|^2 \right)^{n-1} \\ & \leq \varepsilon Y_j + \frac{1}{\varepsilon} 2C_n^2 \delta_j, \end{aligned}$$

which is the desired estimate.

7. DISCUSSION

An obvious conjecture is that Theorem 1 should hold under the weaker assumption that the lengths of the I_j are “lacunary”;

$$\ell(I_{j+1}) \geq \lambda \ell(I_j), \quad \lambda > 1.$$

This conjecture was already proposed in [H2, Remark 4] but we have not settled it.

There is some evidence for a more general conjecture. Define a parallelepiped of dimension N to be a set E of 2^N integers which is the sum of N two element sets. We observed in [H1] that if the interval partition $\mathcal{P} = \{I_j\}$ has the Littlewood-Paley property, then there is an integer N such that there is no parallelepiped E of dimension N which is partitioned by \mathcal{P} into its elements (i.e. $|E \cap I_j| \leq 1$ for all j). We conjecture that this necessary condition is also sufficient. It is not difficult to show that this conjecture implies the previous conjecture, because the points of a large parallelepiped define many disjoint intervals of equal length.

There is a curious similarity between our tree structure and parallelepipeds. Also, our restriction on the amount of branching (type 2 nodes) appears to be of the same nature as the restriction on the dimension N of the parallelepiped.

We have evidence that our definition of a type 2 node can be generalized; there are trees with an unbounded number of type 2 nodes per level which still have the Littlewood-Paley property. The reason is roughly as follows: We really used only the “arithmetic” independence of the intervals in our proof. Thus, it is possible to have many branching nodes on the same level (length scale) as long as their subnodes $(A_1, B_1), \dots, (A_k, B_k)$ are “arithmetically” independent. For example if one is trying to prove the $p = 4$ inequality,

$$\|f\|_4 \leq C \|S(f)\|_4,$$

then one needs:

$$\int A_i \overline{B_i} A_j B_j = 0$$

for $i \neq j$. In other words, “length scales” are too crude to characterize the most general conditions under which the proof may still work.

It would be nice to find a less algebraic proof of our result, perhaps along the lines of [B] and [R]. However, let us remark that an algebraic method was also used by S. K. Pichorides ([P1] and [P2]) to obtain results on the behaviour of the constants in the classical lacunary case.

Finally, we note that our results can be transferred to the real line \mathbf{R} in a natural way: We let $\{I_j\}_{j=-\infty}^{\infty}$ be disjoint intervals in \mathbf{R} such that $\mathbf{R} \setminus (\cup I_j)$ has measure 0. Let their lengths $\ell(I_j)$ satisfy

$$\ell(I_{j+1}) \geq \lambda_j \ell(I_j)$$

for some $\lambda_j \geq 0$, $j \in \mathbf{Z}$. If $\lambda_j \rightarrow \infty$ as $|j| \rightarrow \infty$, then the conclusion of Theorem 1 holds (on $L^p(\mathbf{R})$). We omit the proof. Essentially, given p , one can reduce the proof to the case of a finite interval $J \subset \mathbf{R}$ partitioned into finitely many intervals J_k with $\ell(J_{k+1}) \geq \alpha \ell(J_k)$ for a sufficiently large α depending on p . Then one can invoke our results on \mathbf{Z} via an approximation argument (or just adapt the proof of the latter case to \mathbf{R}). Since one can construct such

intervals $\{I_j\}$ with $\mathbf{R} \setminus (\bigcup I_j)$ an uncountable set, our results imply the existence of new Littlewood-Paley sets on the line, which are fundamentally different from those in \mathbf{Z} .

REFERENCES

- [B] J. Bourgain, *On square functions on the trigonometric system*, Bull. Soc. Math. Belg. Sér. B **37** (1985), 20–26.
- [E] R. E. Edwards and G. I. Gaudry, *Littlewood-Paley and multiplier theory*, Springer-Verlag, Berlin and Heidelberg, 1977.
- [G] G. I. Gaudry, *Littlewood-Paley theorems for sum and difference sets*, Math. Proc. Cambridge Philos. Soc. **83** (1978), 65–71.
- [H1] K. E. Hare and I. Klemes, *Properties of Littlewood-Paley sets*, Math. Proc. Cambridge Philos. Soc. **105** (1989), 485–494.
- [H2] ———, *A new type of Littlewood-Paley partition*, Ark. Math. **30** (1992), 297–309.
- [M] J. Marcinkiewicz, *Sur les multiplicateurs des series de Fourier*, Studia Math. **8** (1939), 78–91.
- [P1] S. K. Pichorides, *A note on the Littlewood-Paley square function inequality*, Collect. Math. **60/61** (1990), 687–691.
- [P2] ———, *A remark on the constants of the Littlewood-Paley inequality*, Proc. Amer. Math. Soc. **114** (1992), 787–789.
- [R] J. L. Rubio de Francia, *A Littlewood-Paley inequality for arbitrary intervals*, Rev. Mat. Iberoamericana **1** (1985), 1–14.
- [S] P. Sjögren and P. Sjölin, *Littlewood-Paley decompositions and Fourier multipliers with singularities on certain sets*, Ann. Inst. Fourier (Grenoble) **31** (1981), 157–175.

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF WATERLOO, WATERLOO, ONTARIO N2L 3G1, CANADA

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY, MONTRÉAL, QUÉBEC H3A 2K6, CANADA