From nahmod@math.umass.edu Tue Feb 6 17:22:51 2001 Date: Tue, 30 Jan 2001 15:45:02 -0500 (EST) From: Andrea Nahmod ¡nahmod@math.umass.edu¿ To: gilbert@mail.ma.utexas.edu Subject: TeX file paper II

L^p-BOUNDEDNESS FOR TIME-FREQUENCY PARAPRODUCTS, II

John E. Gilbert and Andrea R. Nahmod*

In memory of A. P. Calderón

Abstract. This paper completes the proof of the L^p-boundedness of bilinear operators associated to nonsmooth symbols or multipliers begun in Part I, our companion paper [8], by establishing the corresponding L^p-boundedness of time-frequency paraproducts associated with tiles in phase plane. The affine invariant structure of such operators in conjunction with the geometric properties of the associated phase-plane decompositions allow Littlewood-Paley techniques to be applied locally, ie. on trees. Boundedness of the full time-frequency paraproduct then follows using 'almost orthogonality' type arguments relying on estimates for tree-counting functions together with decay estimates.

Introduction

This paper completes the study of bilinear operators with non-smooth symbols begun in [8] (referred to hereafter as Part I) by establishing the L^p-boundedness of a time-frequency paraproduct operator

$$\mathcal{D} : f,\, g \longrightarrow \sum_{Q\,\in\,\mathbb{Q}} \frac{1}{\sqrt{|I_Q|}}\, c_Q \,\langle\, f,\, \phi^{(1)}_Q \,\rangle \,\langle\, g,\, \phi^{(2)}_Q \,\rangle\, \phi^{(3)}_Q . \tag{1.1}$$

The proof uses phase-plane analysis in the spirit of C. Fefferman's proof of Carleson's theorem on the almost everywhere convergence of Fourier series of L^2-functions [2] [4] (*cf.* also [10]). These paraproducts are formed with special building blocks, wave-packets, which are simultaneously 'localized' in the space or 'time' variable x and in the Fourier variable ξ. This localization is better understood, however, when viewed as a localization in the (x, ξ)-plane, phase-plane, where it provides a powerful tool for organizing transformations efficiently in terms of proximity in time or frequency on a given scale and interactions among neighboring scales. The geometry of the operator is then reflected in the geometry of the associated phase-plane decomposition, while the behaviour of the operator can be understood through the delicate interplay between properties of the operator and the geometric structure of its phase-plane decomposition. Such ideas have appeared in many forms in diverse situations throughout the literature. They were used, for instance, by C. Fefferman to provide another proof of Carleson's theorem [4], and related ideas appear also in his work with A. Córdoba [3], as well as with D.H. Phong (*cf.* [5] and the references therein).

Wave-packets incorporate translation in time, scaling and modulation reflecting the action of the affine groups in time and frequency on phase-plane. By associating time-frequency paraproducts with affine-invariant families $\mathbb{Q} = \{Q \sim \{k,\, \ell,\, n\} : k,\, \ell,\, n\, \in \mathbb{Z}\}$ of tiles the $\mathcal{D}(f,\, g)$ thus acquire

1991 *Mathematics Subject Classification.* Primary 42B15, 42C15 Secondary 42B20 42B25 40A20.

Key words and phrases. bilinear operator, singular multiplier, wave-packet, paraproduct.

* The second author was partially supported by NSF grant DMS 9971159 and acknowledges the support of MSRI, Berkeley (NSF grant DMS 9701755) and of IAS at Princeton (NSF grant DMS 9304580) during 1997-1998. She warmly thanks both Institutes for their hospitality while part of this research was being carried out.

a crucial structural invariance. Given positive numbers a_j, a positive rational ρ, and $\mathcal{M}_\mu$-test functions $\phi^{(j)}$, let

$$\phi^{(j)}_{k\ell n}(x) \;=\; \phi^{(j)}_Q(x) \;=\; s^{k/2}\phi_j(s^k x - a_j\ell)\, e^{2\pi i s^k x n}, \qquad s = 2^\rho, \tag{1.2}$$

be the corresponding wave packet associated with a tile $Q \sim \{k, \ell, n\}$ in phase plane (*cf.* section 2 in Part I for terminology and notation). The time-frequency paraproduct in (1.1) can then be written as

$$\mathcal{D}(f,\, g)(x) \;=\; \sum_{k,\ell,n} s^{k/2}\, c_{k\ell n} \,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}, \tag{1.3}$$

summing over all tiles $Q \sim \{k, \ell, n\}$ in the family $\mathbb{Q}$; the notation in (1.1) and (1.3) will be used interchangably for $\mathcal{D}(f,\, g)$. In 'standard' paraproducts there are no modulations and in Part I boundedness from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 2$, was established for $p,\, q > 1$ whenever at least two of the $\phi^{(j)}$ have vanishing moment, extending well-known results for $r \geq 1$. Since modulation need not preserve vanishing moments, however, stronger conditions have to be imposed to secure analogous L^p-boundedness results for $\mathcal{D}(f,\, g)$. Let $w^{(j)}$ be finite intervals such that

$$\operatorname{supp}\widehat{\phi}^{(1)} \;\subseteq\; w^{(1)}, \qquad \operatorname{supp}\widehat{\phi}^{(2)} \;\subseteq\; w^{(2)}, \qquad \operatorname{supp}\widehat{\phi}^{(3)} \;\subseteq\; w^{(3)}; \tag{1.4}$$

these $w^{(j)}$ will be referred to as the *Fourier support intervals* of the $\phi^{(j)}$ though the actual support may well be a subset of $w^{(j)}$. The substitute for vanishing moments is the requirement that the $w^{(j)}$ have pairwise-disjoint closure; at least two of the $\phi^{(j)}$ will then have vanishing moment as will the modulates $\phi^{(j)}(x)\, e^{2\pi i n x}$ for each fixed n.

(1.5) Definition. *Fix positive constants a_j, a positive rational ρ, and $\mathcal{M}_\mu$-test functions $\phi^{(j)}$. Then the operator*

$$\mathcal{D} : \{c_{k\ell n}\},\, f,\, g \longrightarrow \sum_{k,\ell,n} s^{k/2}\, c_{k\ell n}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}, \qquad s = 2^\rho,$$

will be called a time-frequency paraproduct if the Fourier support intervals $w^{(j)}$ of the $\phi^{(j)}$ have pairwise-disjoint closure.

The following result completes the proof of Main Theorem I in Part I of this series of papers.

Main Theorem II. *Let $\phi^{(j)}$ be $\mathcal{M}_\mu$-test functions whose Fourier support intervals have pairwise-disjoint closure. Then a time-frequency paraproduct $\mathcal{D}$ is bounded as an operator from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$ whenever $1/p + 1/q = 1/r < 3/2$ and $p, q > 1$; in addition, its operator norm satisfies the inequality*

$$\|\mathcal{D}\|_{op} \;\leq\; \text{const.}\, P\big(\|\phi^{(1)}\|_{\mathcal{M}_\mu},\, \|\phi^{(2)}\|_{\mathcal{M}_\mu},\, \|\phi^{(3)}\|_{\mathcal{M}_\mu}\big)$$

for some polynomial P depending only on a_j, ρ and the Fourier support intervals $w^{(j)}$.

Examples show that the restriction $r > 2/3$ is sharp (*cf.* [11]). Several simplifications established in Part I facilitate the proof. Because of structural invariance $\mathcal{D}(f,\, g)$ can be written as a finite sum of time-frequency paraproducts in which the Fourier support intervals are closely tied to affine-invariant grids. It is this that allows the delicate phase-plane analysis to be carried through. Given a positive integer ρ, the families of intervals

$$w_{kn} \;=\; [2^{\rho k}n,\ 2^{\rho k}(n+1)), \quad I_{k\ell} \;=\; [2^{-\rho k}\ell,\ 2^{-\rho k}(\ell+1)), \qquad -\infty < k, \ell, n < \infty$$

will be denoted by $\mathcal{W}_{1,\rho}$ and $\mathcal{I}_{1,\rho}$ respectively. Each of these is an affine-invariant grid generated by a lattice of affine transformations from the basic interval $[0,\,1)$. More generally, given positive integers M, N the family $\mathcal{W}_{M,N}$ of intervals

$$w_{kn} \;=\; \big[\,2^{MNk}(n-\alpha_M),\ 2^{MNk}(n+\alpha_M)\big), \quad \alpha_M = \frac{2^{M-1}-1}{2^M-1}.$$

is an affine-invariant grid for each integer $M \geq 2$ generated now from a basic interval $(-\alpha_M,\,\alpha_M)$ which is symmetric about the origin. Then any time-frequency paraproduct in (1.5) can be written as a finite sum of ones in which ρ is a positive integer, which can be chosen freely, and the Fourier support intervals satisfy what we will call the *Fourier Support Condition*; namely, the $w^{(j)}$ have pairwise-disjoint closure and all lie in $(0,\,1)$ or in $(-\alpha_M,\,\alpha_M)$ for some $M \geq 2$ (*cf.* (2.8) and section 5 in Part I for details). Even more is true.

(1.6) Theorem. *There is a positive integer N so that*

(*a*) *when the Fourier support intervals all lie in* $(0,\,1)$ *we can take*

$$w^{(j)} \;=\; [\alpha_j/2^N,\ \beta_j/2^N)$$

for a suitable choice of integers α_j, β_j; furthermore, it can assumed that there is a dyadic interval of length 2^{-N} between adjacent $w^{(j)}$ as well as one between each end-point of $[0,\;1)$ *and the nearest $w^{(j)}$;*

(*b*) *when the Fourier support intervals all lie in* $(-\alpha_M,\,\alpha_M)$ *we can take*

$$w^{(j)} \;=\; [2^{-MN}(\alpha_j-\alpha_M),\ 2^{-MN}(\beta_j+\alpha_M))$$

for a suitable choice of integers α_j, β_j; furthermore, it can assumed that there is an interval in $\mathcal{W}_{M,N}$ of length $2\alpha_M 2^{-MN}$ between adjacent $w^{(j)}$ as well as one between each end-point of $[-\alpha_M,\,\alpha_M)$ *and the nearest $w^{(j)}$.*

Notice that each $w^{(j)}$ begins and ends with an interval in $\mathcal{W}_{M,N}$ of length $2\alpha_M 2^{-MN}$ just as in the dyadic case. Now let $\mathcal{I}_{M,N}$ be the family of dyadic intervals $I_{k\ell} = [2^{-MNk}\,\ell\,,\,2^{-MNk}(\ell+1))$ and let $\mathbb{Q}_{M,N}$ be the family of tiles

$$Q \sim \{k,\ell,n\} \;=\; I_{k\ell}\times w_{kn}, \qquad I_{k\ell}\in\mathcal{I}_{M,N},\ w_{kn}\in\mathcal{W}_{M,N}$$

in phase plane with N specified by (1.6). The respective intervals $I_Q = I_{k\ell}$ and $w_Q = w_{kn}$ will be called the *time* and *frequency* intervals of Q. By taking $\rho = MN$ in (1.5) we thus arrive at the fundamental link between tiles in phase space and paraproducts: set

$$\phi_Q^{(j)}(x) \;=\; s^{k/2}\phi^{(j)}(s^k x - a_j\ell)\, e^{2\pi s^k n x}, \qquad s = 2^{MN} \tag{1.7}$$

for each $Q \sim \{k,\,\ell,\,n\}$ in $\mathbb{Q}_{M,N}$.

(1.8) Definition. *A time-frequency paraproduct* $\mathcal{D}(f,\, g)$ *is said to be in* (M, N)*-canonical form when*

$$\mathcal{D}(f,\, g) \;=\; \sum_{Q\,\in\,\mathbb{Q}_{M,N}} c_Q\,\frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi^{(1)}_Q\,\rangle\,\langle\, g,\, \phi^{(2)}_Q\,\rangle\,\phi^{(3)}_Q$$

and the Fourier support intervals $w^{(j)}$ *of the* $\phi^{(j)}$ *are specified by* (1.6)*. By relabelling, if necessary, we take*

$$w^{(1)} \;<\; w^{(2)} \;<\; w^{(3)},$$

meaning that $w^{(1)}$ *lies to the left of* $w^{(2)}$*, which in turn lies to the left of* $w^{(3)}$*.*

These paraproducts have a number of special properties. For instance, when $f,\, g$ are band-limited $\mathcal{M}_\mu$-molecules, as we shall assume from now on, the series

$$\sum_{Q\,\in\,\mathbb{Q}_{M,N}} |c_Q|\,\frac{1}{\sqrt{|I_Q|}}\,|\langle\, f,\, \phi^{(1)}_Q\,\rangle\,\langle\, g,\, \phi^{(2)}_Q\,\rangle\,\phi^{(3)}_Q(x)|$$

converges pointwise for each x; in this case the series defining $\mathcal{D}(f,\, g)(x)$ can be manipulated freely. This unconditionality will be crucial to the proof of the following result whose proof occupies the remainder of this paper. We shall denote by $\mathbb{Q}^{(+)}_{M,N}$ those tiles $Q \sim \{k,\, \ell,\, n\}$ in which $n > 0$.

(1.9) Theorem. *A time-frequency paraproduct*

$$\mathcal{D}(f,\, g) \;=\; \sum_{Q\,\in\,\mathbb{Q}^{(+)}_{M,N}} c_Q\,\frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi^{(1)}_Q\,\rangle\,\langle\, g,\, \phi^{(2)}_Q\,\rangle\,\phi^{(3)}_Q$$

in $(M,\, N)$*-canonical form is bounded as an operator from* $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ *into* $L^r(\mathbb{R})$*, whenever* $1/p + 1/q = 1/r < 3/2$ *and* $p,\, q > 1$*. Its operator norm satisfies the inequality*

$$\|\mathcal{D}\|_{op} \;\le\; \text{const.}\, P\big(\|\phi^{(1)}\|,\, \|\phi^{(2)}\|,\, \|\phi^{(3)}\|\big)$$

for some polynomial P *depending only on* a_j, ρ *and the Fourier support intervals* $w^{(j)}$*.*

In (1.9) as well as throughout the rest of the paper we adopt the same convention as in Part I in which $\|\phi\|$ will always mean that ϕ is an $\mathcal{M}_\mu$-test function and $\|\phi\|$ is its $\mathcal{M}_\mu$-norm, unless the contrary is explicitly indicated. Because of the need in section 10 to use interpolation along with symmetry and adjoint properties of $\mathcal{D}$ a slightly stronger result than (1.9) will actually proved.

(1.10) Remark. *For any permutation* $\varepsilon \;=\; \{\varepsilon_1,\, \varepsilon_2,\, \varepsilon_3\}$ *of* $\{1,\, 2,\, 3\}$ *set*

$$\mathcal{D}^{(\varepsilon)}(f,\, g) \;=\; \sum_{Q\,\in\,\mathbb{Q}_{M,N}} c_Q\,\frac{1}{\sqrt{|I_Q|}}\langle\, f,\, \phi^{(\varepsilon_1)}_Q\,\rangle\,\langle\, g,\, \phi^{(\varepsilon_2)}_Q\,\rangle\,\phi^{(\varepsilon_3)}_Q.$$

Then $\mathcal{D}^{(\varepsilon)}$ *maps* $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ *boundedly into* $L^r(\mathbb{R})$ *for each permutation* ε *under the same restrictions on* $p,\, q,\, r$ *as above.*

In other words, having fixed the ordering $w^{(1)} < w^{(2)} < w^{(3)}$ of the Fourier support intervals in (1.8), we can form the coefficients in $\mathcal{D}$ with respect to any ordering of the wave packets so long as that ordering remains fixed throughout the proof. For notational simplicity, however, we take $\varepsilon_i = i$, leaving the necessary changes for other ε to the reader. In the particular case where the smooth wave-packets are replaced by Walsh wave-packets, such a result, and more, was established in [7]. In this case the time-frequency analysis of the canonical operator simplifies greatly due to the sharp localization in the phase-plane (see also [17]).

Main Theorem II follows very easily from theorem (1.9). For if we decompose $\mathbb{Q}_{M,N}$ by partitioning it into tiles $Q \sim \{k,\, \ell,\, n\}$ with $n < 0,\ n = 0$, and $n > 0$ so that

$$\mathbb{Q}_{M,N} \;=\; \mathbb{Q}^{(-)}_{M,N} \,\cup\, \mathbb{Q}^{(0)}_{M,N} \,\cup\, \mathbb{Q}^{(+)}_{M,N}$$

with obvious notation, then any time-frequency paraproduct in $(M,\, N)$ canonical form can be written as

$$\Big(\sum_{Q\,\in\,\mathbb{Q}^{(-)}_{M,N}} + \sum_{Q\,\in\,\mathbb{Q}^{(0)}_{M,N}} + \sum_{Q\,\in\,\mathbb{Q}^{(+)}_{M,N}}\Big)\, c_Q\, \frac{1}{\sqrt{|I_Q|}}\, \langle\, f,\, \phi^{(1)}_Q\,\rangle\, \langle\, g,\, \phi^{(2)}_Q\,\rangle\, \phi^{(3)}_Q.$$

Theorem (1.9) establishes L^p-boundedness of the third operator on the right, while theorem (1.9) in [8] applies to the second one since it is a 'standard' paraproduct in which modulation is absent. On the other hand, after a change of variable $n \to -n$, $\ell \to -\ell$ and $x \to -x$, the individual coefficients in the first operator become

$$\langle\, f,\, \phi^{(1)}_{k,-\ell,-n}\,\rangle \;=\; \int_{-\infty}^{\infty} f(-x)\,\overline{\varphi^{(1)}(s^k x - a_1\ell)}\, e^{-2\pi i s^k x n}\, dx,$$

setting $\varphi^{(1)}(x) \;=\; \phi^{(1)}(-x)$. When $M > 1$ the Fourier support interval of the $\varphi^{(j)}$ will still lie in $(-\alpha_M,\, \alpha_M)$ but the order of the these intervals will be reversed. Nonetheless, in view of (1.10), the L^p-boundedness of the first operator is established. In the case $M = 1$, however, the supports of the $\varphi^{(j)}$ will again be reversed but they will now lie in $(-1, 0)$. To remedy this we need to define the $\varphi^{(j)}$ by $\varphi^{(j)}(x) \;=\; \phi^{(j)}(-x)\, e^{2\pi i x}$ so that their Fourier support intervals still lie in $(0,\, 1)$ though they are again reversed in order. In this case the first operator above becomes

$$\Big(\sum_{Q\,\in\,\mathbb{Q}^{(+)}_{M,N}} + \sum_{Q\,\in\,\mathbb{Q}^{(0)}_{M,N}}\Big)\, c_Q\, \frac{1}{\sqrt{|I_Q|}}\, \langle\, f,\, \varphi^{(1)}_Q\,\rangle\, \langle\, g,\, \varphi^{(2)}_Q\,\rangle\, \varphi^{(3)}_Q, \qquad (M = 1).$$

As before, it will be bounded, completing the proof of Main Theorem II.

Finally, for each $Q \in \mathbb{Q}_{M,N}$ let

$$\tau_Q : [0,\, 1) \longrightarrow w_Q, \quad (M = 1); \qquad \tau_Q : [-\alpha_M, \alpha_M) \longrightarrow w_Q, \quad (M > 1),$$

be the affine transformation in frequency mapping $[0,\, 1)$ and $[-\alpha_M,\, \alpha_M)$ respectively onto the frequency interval w_Q of Q; ie. $\tau_Q(\xi) = s^k(\xi + n)$, $Q \sim \{k, l, n\}$. The intervals $w^{(j)}_Q \;=\; \tau_Q(w^{(j)})$ are then the Fourier support intervals of the wave packets $\phi^{(j)}_Q$. The following properties established in Part I (*cf.* [8, section 5]) will become fundamental to the restriction to time-frequency paraproducts in (M, N)-canonical form:

(1.11) *the family* $\{w^{(j)}_Q : Q \,\in\, \mathbb{Q}_{M,N}\,\}$ *is a grid for each* j*;*

(1.12) *if P, Q are tiles in $\mathbb{Q}_{M,N}$ with $w_P \cap w_Q^{(j)} \neq \emptyset$ and $|I_Q| < |I_P|$, then $w_P \subseteq w_Q^{(j)}$;*

(1.13) *given λ_1, λ_2 with $\lambda_1 < \lambda_2$ and a tile*

$$Q \;=\; I_Q \times w_Q, \qquad w_Q = [2^{MNk}(m - \alpha_M),\, 2^{MNk}(m + \alpha_M))$$

in $\mathbb{Q}_{M,N}$ such that $w_Q^{(i)} < w_Q^{(j)}$, then $\lambda_1 \in w_Q^{(i)}$ and $\lambda_2 \in w_Q^{(j)}$ hold simultaneously for at most one choice of k. There is a corresponding result for a tile in $\mathbb{Q}_{1,N}$.

From hereon

$$\pi(a) : f(x) \longrightarrow a^{1/2} f(ax), \qquad a > 0$$

will denote the unitary action of dilation on $L^2(\mathbb{R})$. The appearance of $\pi(a_j)$ - or evaluation of functions at $a_j x$ which amounts to the same thing - will occur many times because of the need to estimate wave packets $\phi_Q^{(j)}$. Indeed,

$$\begin{aligned} a_j^{1/2}\phi_Q^{(j)}(a_j x) \;&=\; (\pi(a_j)\phi_Q^{(j)})(x) \;=\; a_j^{1/2} s^{k/2}\phi^{(j)}(s^k a_j x - a_j \ell)\, e^{2\pi i s^k a_j n x} \\ &=\; s^{k/2}(\pi(a_j)\phi^{(j)})(s^k x - \ell)\, e^{2\pi i s^k a_j n x}, \end{aligned}$$

and so

$$|a_j^{1/2}\,\phi_Q^{(j)}(a_j x)| \;\le\; \frac{1}{\sqrt{|I_Q|}}\|\pi(a_j)\phi^{(j)}\|\left(\frac{1}{1 + |s^k x - \ell|}\right)^{\mu+1}. \tag{1.14}$$

The importance of this is that estimates in time can now be made *independently* of j; in other words, with respect to the same s-adic grid in time for each j.

This paper has had a gestation period of several years with the final written version being completed in the summer of 1999. During that time period different aspects of this paper and most of the ideas have been presented by the authors at various lectures, including those in 1997 at Georgia Tech (AMS meeting), the University of New Mexico (AMS meeting), Rutgers University, and MSRI at Berkeley (Special semester in Harmonic Analysis); in 1998 at IAS in Princeton, Temple University (AMS meeting), the University of Texas at Austin and Brown University and in 1999 at Georgia Tech.

As the final edition of this paper was being completed we learned that C. Muscalu, C. Thiele and T.Tao were able to extend our bilinear result to certain multilinear operators. Their approach is somewhat different in that they exploit the idea of using restricted-type estimates to do an induction argument to pass from symbols having one dimensional singularities -as in the bilinear case - to certain multilinear operators associated to symbols with higher dimensional singularities but of codimension strictly larger than one. In the process of doing so they provide a different proof of our bilinear result [15].

2. Outline of the proof of Theorem (1.9)

The proof of Theorem (1.9) relies on a careful study of the phase plane associated with $\mathcal{D}$. Given $\delta > 0$, δ small, choose p, $q > 1$ so that

$$\text{(2.1)(i)} \qquad \frac{1}{2} + 2\delta \ < \ 1/p + 1/q \ < \ \frac{3}{2} - 2\delta, \qquad \big|1/p - 1/q\big| \ < \ \frac{1}{2} - 2\delta.$$

The lower bound is needed to secure convergence of various geometric series occuring in the proof and will be removed using interpolation in section 10, exploiting the symmetry and adjoint properties of the family of all $\mathcal{D}^{(\varepsilon)}$ as in [7]. But the upper bound is needed *solely* to prove the error estimate in section 3 (*cf.* (3.4)). Set

$$\text{(2.1)(ii)} \qquad p_0 \ = \ \max\{p, p'\}, \quad q_0 \ = \ \max\{q, q'\}, \quad 1/r_0 \ = \ 1/p_0 + 1/q_0,$$

so that

$$\text{(2.1)(iii)} \qquad \frac{1}{2} + 2\delta \ < \ 1/p_0 + 1/q_0 \ < \ \frac{3}{2} - 2\delta.$$

Now fix $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ and $\{c_Q\} \in l^\infty$; without loss of generality we assume $\|\{c_Q\}\|_\infty = 1$. The goal is to establish the weak type estimate

$$\text{(2.2)} \qquad |\{x : |\mathcal{D}(f,g)(a_3 x)| \ \geq \ 2\gamma\}| \ \leq \ \text{const.}\Big(\frac{\|\pi(a_1)f\|_p \,\|\pi(a_2)g\|_q}{\gamma}\Big)^r, \quad \gamma > 0$$

with $1/r = 1/p + 1/q$ as usual. The first step in the proof is reminiscent of the familiar Calderón-Zygmund decomposition. Fix a small $\eta > 0$ to be specified later depending on the earlier choice of δ and r_0. Set

$$\text{(2.3)(i)} \qquad \begin{aligned} E_{bad} \ &= \big\{x : M_p(M(\pi(a_1)f)(x) > s^{-1/\eta}\,\kappa_p\,\big\} \\ &\qquad \bigcup \big\{x : M_q(M(\pi(a_2)g)(x) > s^{-1/\eta}\,\kappa_q\,\big\}, \end{aligned}$$

where

$$\text{(2.3)(ii)} \qquad \kappa_p \ = \ \Big(\frac{\|\pi(a_1)f\|_p^{1/q}\,\gamma^{1/p}}{\|\pi(a_2)g\|_q^{1/p}}\Big)^r, \qquad \kappa_q \ = \ \Big(\frac{\|\pi(a_2)g\|_q^{1/p}\,\gamma^{1/q}}{\|\pi(a_1)f\|_p^{1/q}}\Big)^r.$$

With these choices

$$\text{(2.4)(i)} \qquad |E_{bad}| \ \leq \ \text{const.}\Big(\frac{\|\pi(a_1)f\|_p \,\|\pi(a_2)g\|_q}{\gamma}\Big)^r$$

since $\kappa_p\kappa_q \ = \ \gamma$ and

$$\text{(2.4)(ii)} \qquad \Big(\frac{\|\pi(a_1)f\|_p}{\kappa_p}\Big)^p \ = \ \Big(\frac{\|\pi(a_1)f\|_p \,\|\pi(a_2)g\|_q}{\gamma}\Big)^r \ = \ \Big(\frac{\|\pi(a_2)g\|_q}{\kappa_q}\Big)^q.$$

The constant in (2.4) will depend on η, of course. As a *function*,

$$\mathcal{D}(f,\, g) \;=\; \mathcal{D}_{bad}(f,\, g) \;+\; \mathcal{D}_{good}(f,\, g) \tag{2.5}$$

decomposes $\mathcal{D}(f,g)$ into 'bad' and 'good' functions setting

$$\mathcal{D}_{bad}(f,\, g) \;=\; \sum_{I_Q \,\subseteq\, E_{bad}} \frac{1}{\sqrt{|I_Q|}}\, c_Q \,\big\langle\, f,\, \phi^{(1)}_Q \,\big\rangle \big\langle\, g,\, \phi^{(2)}_Q \,\big\rangle\, \phi^{(3)}_Q,$$

i.e., by summing only over tiles with $I_Q \subseteq E_{bad}$. The Hardy-Littlewood maximal function controls wave packet coefficients of f in the sense that

$$\frac{1}{\sqrt{|I_Q|}}\big|\big\langle\, f,\, \phi^{(1)}_Q \,\big\rangle\big| \;\le\; \text{const.}\, \|\pi(a_1)\phi^{(1)}\|\, \Big(\inf_{x\in I_Q}\; M(\pi(a_1)f)(x)\Big), \tag{2.6}$$

holds uniformly in f and $\phi^{(1)}$, and correspondingly for g. Thus removal of all tiles with $I_Q \subseteq E_{bad}$ ensures that the coefficients in

$$\mathcal{D}_{good}(f,\, g) \;=\; \sum_{I_Q \,\not\subseteq\, E_{bad}} \frac{1}{\sqrt{|I_Q|}}\, c_Q \,\big\langle\, f,\, \phi^{(1)}_Q \,\big\rangle \big\langle\, g,\, \phi^{(2)}_Q \,\big\rangle\, \phi^{(3)}_Q \tag{2.7}$$

satisfy uniform bounds

$$\frac{1}{\sqrt{|I_Q|}}\big|\big\langle\, f,\, \phi^{(1)}_Q \,\big\rangle\big| \;\le\; \text{const}_\phi\, s^{-1/\eta}\kappa_p, \qquad \frac{1}{\sqrt{|I_Q|}}\big|\big\langle\, g,\, \phi^{(2)}_Q \,\big\rangle\big| \;\le\; \text{const}_\phi\, s^{-1/\eta}\kappa_q \tag{2.8(i)}$$

where

$$\text{const}_\phi \;\le\; \text{const.}\ \max\big\{\|\pi(a_1)\phi^{(1)}\|,\, \|\pi(a_2)\phi^{(2)}\|\big\}. \tag{2.8(ii)}$$

On the other hand, the $\phi^{(i)}$ appearing in $\mathcal{D}_{bad}(f,\, g)$ are 'concentrated' inside E_{bad}, so the bad function can be estimated sufficiently far away from E_{bad} using solely decay estimates on the $\phi^{(i)}$ and Hausdorff-Young inequalities. Set

$$E_1 \;=\; \bigcup_{I_Q \subseteq\, E_{bad}} s^2 I_Q \tag{2.9}$$

where for an interval J, AJ will always denote the interval of length $A|J|$ having the same center as J. In section 3 the following estimates will be established.

(2.10) Theorem. *The inequalities* $|E_1| \;\le\; \text{const.}\, |E_{bad}|$ *and*

$$\frac{1}{\gamma}\int_{\mathbb{R}\setminus E_1} |\mathcal{D}_{bad}(f,\, g)(a_3 x)|\, dx \;\le\; \text{const.}\; \frac{|E_{bad}|}{a_3}$$

hold uniformly in f, g and γ as well as the a_j.

Clearly then

$$|\{x : |\mathcal{D}_{bad}(f,g)(a_3x)| \geq \gamma\}| \leq \text{const.}\left(\frac{\|\pi(a_1)f\|_p \, \|\pi(a_2)g\|_q}{\gamma}\right)^r,$$

leaving only the proof of the corresponding estimate for $\mathcal{D}_{good}(f, g)$. This requires a very delicate decomposition of the 'good' function into the sum of functions associated with 'trees' of tiles defined using the partial order

$$Q \leq Q' \qquad \Longleftrightarrow \qquad I_Q \subseteq I'_Q, \quad w_Q \supseteq w_{Q'} \tag{2.11}$$

on $\mathbb{Q}^{(+)}_{M,N}$. Notice that two tiles are *comparable*, meaning that $Q \leq Q'$ or $Q' \leq Q$, if and only if $Q \cap Q' \neq \emptyset$.

A *tree* $\mathbb{T}$ is set of tiles containing a tile $\overline{Q}$ which is maximal in the sense that $Q \in \mathbb{T} \implies Q \leq \overline{Q}$. This maximal tile will be called the *tree-top* of $\mathbb{T}$ and will often be denoted by $I_\mathbb{T} \times w_\mathbb{T}$ to emphasize its dependence on $\mathbb{T}$. To each tree there corresponds a Carleson box or a tent in the usual upper half-plane and so there are intimate connections between trees and Tent spaces. These we exploit later. The role of a tree, however, is to control in an efficient manner the oscillatory behaviour that an otherwise random group of tiles in phase-plane has. To illustrate this consider the *tree operator*

$$f,\, g \longrightarrow \mathcal{D}_\mathbb{T}(f,\, g) \;=\; \sum_{Q \,\in\, \mathbb{T}} c_Q \, \frac{1}{\sqrt{|I_Q|}} \, \langle\, f,\, \phi^{(1)}_Q \,\rangle \, \langle\, g,\, \phi^{(2)}_Q \,\rangle \, \phi^{(3)}_Q$$

obtained by summing only over tiles in $\mathbb{T}$, and suppose $M = 1$. For each tile $Q \sim \{k,\, \ell,\, n\}$ in $\mathbb{T}$ the tree structure ensures that $n = [s^{-k}\lambda_\mathbb{T}]$, where $\lambda_\mathbb{T}$ is the left-hand endpoint of $w_\mathbb{T}$ (*cf.* (4.1)). After suitable conjugations by $e^{2\pi i x \lambda_\mathbb{T}}$, therefore, $\mathcal{D}_\mathbb{T}$ can be rewritten in terms of modulated wave-packets all having roughly the *same oscillation* and hence $\mathcal{M}_\mu$-norm which is uniform in $\mathbb{T}$. To be precise their frequency satisfy the inequality

$$0 \leq s^{-k}\lambda_\mathbb{T} - [s^{-k}\lambda_\mathbb{T}] < 1.$$

A tree operator is thus a 'standard' paraproduct modulated by a single exponential $e^{2\pi i x \lambda_\mathbb{T}}$. Familiar techniques now produce L^2-norm estimates for $\mathcal{D}_\mathbb{T}$ which are independent of $\lambda_\mathbb{T}$ provided at least two of the modulated wave-packets $\psi^{(i)}(x) = \phi^{(i)}(x) e^{2\pi i x (s^{-k}\lambda_\mathbb{T} - [s^{-k}\lambda_\mathbb{T}])}$ have vanishing moments. But

$$s^{-k}\lambda_\mathbb{T} - [s^{-k}\lambda_\mathbb{T}] \notin w^{(i)} \quad \Longrightarrow \quad \int_\infty^\infty \psi^{(i)}(x)\, dx \;=\; \widehat{\phi}^{(i)}(s^{-k}\lambda_\mathbb{T} - [s^{-k}\lambda_\mathbb{T}]) \;=\; 0,$$

so we need to know that $s^{-k}\lambda_\mathbb{T} - [s^{-k}\lambda_\mathbb{T}]$ fails to belong to at least two of the $w^{(i)}$. This, however, is exactly what disjointness of the Fourier support intervals guarantees. A corresponding argument applies in the case $M \geq 2$, setting $\lambda_\mathbb{T} = \tau_{\overline{Q}}(0)$. Hence we can also view this grouping of tiles into trees as an 'efficient localization' in phase plane, for which the origin becomes once again a 'distinguished' point in frequency, in the sense that locally, *i.e.*, on each tree, Littlewood-Paley theory applies (*cf.* Appendix A.).

The idea now is to choose families of trees. In section 4 a family $\mathcal{F}_\nu$ of trees will be constructed for each $\nu \geq 0$ so that

$$\mathcal{D}_{good}(f,\, g) \;=\; \sum_{\nu=0}^{\infty}\Big(\sum_{\mathbb{T}\in\mathcal{F}_\nu} \mathcal{D}_{\mathbb{T}}(f,\, g) \Big). \tag{2.12}$$

There will be three different classes of trees, each specifying which two of the three wave-packets $\phi_Q^{(i)}$, $i = 1,\, 2,\, 3$, have vanishing moments *uniformly* for tiles Q in that tree. All the difficulty comes in establishing L^2-estimates for (2.12). Ideally, what one really wants is that each $\mathcal{D}_{\mathbb{T}}(f, g)$ be an L^2-function and that pairs of such functions be 'almost orthogonal'. Armed with the Fourier support condition and the vanishing moment conditions available for each tree we prove:

(1) *an L^2-norm estimate*

$$\Big(\frac{1}{\gamma^2}\int_{-\infty}^{\infty} \big|\mathcal{D}_{\mathbb{T}}(f,\, g)(x)\big|^2\, dx\Big)^{1/2} \;\leq\; \text{const.}\; s^{-\nu/r_0}\, |I_{\mathbb{T}}|^{1/2}$$

for each tree $\mathbb{T}$ in $\mathcal{F}_\nu$ (cf. (5.5)), and

(2) *an L^σ-norm estimate*

$$\Big(\int_{-\infty}^{\infty} N_{\overline{\mathcal{F}}_\nu}(x)^\sigma\, dx\Big)^{1/\sigma} \;\leq\; \text{const.}\, s^{(1+2\delta)\nu}\, \Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/\sigma}$$

for the function $N_{\overline{\mathcal{F}}_\nu} = N_{\overline{\mathcal{F}}_\nu}(x)$ counting the number of trees in $\overline{\mathcal{F}}_\nu$ above x where $\overline{\mathcal{F}}_\nu$ is a suitable truncation of $\mathcal{F}_\nu$ (cf. (6.1) and (7.1)).

This counting function

$$N_{\overline{\mathcal{F}}_\nu}(x) \;=\; \sum_{\mathbb{T}\in\overline{\mathcal{F}}_\nu} \chi_{I_{\mathbb{T}}}(x)$$

controls most aspects of the rest of the proof as it captures the interactions among trees. It enables us to sum 'almost orthogonal' tree *functions* in much the same spirit as almost orthogonal operators are summed in the Cotlar-Stein lemma. In the case of just one tree, for instance, it provides the L^2-bound

$$\frac{1}{\gamma^2}\int_{-\infty}^{\infty} \big|\mathcal{D}_{\mathbb{T}}(f,\, g)(x)\big|^2\, dx \;\leq\; \text{const.}\; s^{-2\delta\nu}\, \Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r} \tag{2.13}$$

for each tree function $\mathcal{D}_{\mathbb{T}}(f,\, g)$ since

$$|I_{\mathbb{T}}| \;=\; \int_{-\infty}^{\infty} \chi_{I_{\mathbb{T}}}(x)\, dx \;\leq\; \int_{-\infty}^{\infty} N_{\overline{\mathcal{F}}_\nu}(x)\, dx.$$

If the estimate in (2.13) for a single tree could be replaced by the sum over trees then the companion estimate

$$|\{x : |\mathcal{D}_{good}(f,\, g)(a_3 x)| \, > \gamma \,\}| \;\le\; \text{const.} \left(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma} \right)^r$$

to the one for the 'bad' function would follow immediately. Our approach has to be less direct, however, though it is the same in principle. We adopt the strategy Fefferman used at the corresponding point of his pointwise convergence proof ([4]):

(a) *'thin out' the trees in* $\mathcal{F}_\nu$*, and seek families of new trees to be called forests*;

(b) *decompose the 'thinned'* $\mathcal{F}_\nu$ *into* $O(\nu)$ *forests whose trees still satisfy* (1) *and whose counting function satisfies the same* L^σ*-estimate* (2) ;

(c) *'trim' the new trees in each forest so that an estimate like* (2.13) *holds now for the sum of trees in a forest* ;

(d) *estimate the error terms created by this double pruning process.*

To 'thin out' the trees set

$$\text{(2.14)(i)} \qquad E^{(\nu)}_{dense} \;=\; \{\, x : \; N_{\mathcal{F}_\nu}(x) \,>\, s^{2\nu/r_0} \,\}, \quad E_{dense} \;=\; \bigcup_{\nu=0}^{\infty} E^{(\nu)}_{dense};$$

the measure

$$\text{(2.14)(ii)} \qquad |E^{(\nu)}_{dense}| \;\le\; \text{const.}\; s^{-2\delta\nu} \left(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma} \right)^r$$

of $E^{(\nu)}_{dense}$, hence of E_{dense}, is controlled by the L^1-norm of the counting estimate for $\mathcal{F}_\nu$. Taking the wave packets 'concentrated' in E_{dense} leads to an error term

$$\mathcal{D}_{dense}(f,\, g) \;=\; \sum_{I_Q \,\subseteq\, E_{dense}} c_Q \, \frac{1}{\sqrt{|I_Q|}} \, \langle\, f,\, \phi^{(1)}_Q \,\rangle \, \langle\, g,\, \phi^{(2)}_Q \,\rangle \, \phi^{(3)}_Q$$

for which an estimate entirely analogous to (2.10) is established after introducing a second exceptional set, E_2, defined from E_{dense} in the same manner E_1 was from E_{bad}.

Thinning the trees in $\mathcal{F}_\nu$ allows us to write the remaining trees as a union of at most $O(\nu)$ 'forests' $\mathcal{W}^{(\nu)}_n$ using what Fefferman calls 'an elementary combinatorial argument' ([4, p.554]). Each forest $\mathcal{W}^{(\nu)}_n$ consists of new trees $\mathbb{S}$ which themselves are subtrees of the original trees in $\mathcal{F}_\nu$. Thus

$$\mathcal{D}_{good}(f,\, g) \;=\; \mathcal{D}_{dense}(f,\, g) \;+\; \sum_{\nu=0}^{\infty} \Big[\sum_{n=1}^{O(\nu)} \Big(\sum_{\mathbb{S} \,\in\, \mathcal{W}^{(\nu)}_n} \mathcal{D}_{\mathbb{S}}(f,\, g) \Big) \Big],$$

thereby reducing the proof to establishing estimates for 'forest' operators. This is accomplished by first 'trimming' the new trees in $\mathcal{W}^{(\nu)}_n$, removing all tiles which are 'too close' to the edges of the time interval of their tree-top. Set

$$\text{(2.15)(i)} \qquad E^{(\nu)}_{edge} \;=\; \bigcup_{n} \Big(\bigcup_{\mathbb{S} \in \mathcal{W}^{(\nu)}_n} \{\, x : \; \text{dist}(x, \partial I_{\mathbb{S}}) \;\le\; s^{-2\nu/r_0} |I_{\mathbb{S}}| \,\} \Big)$$

and $E_{edge} = \bigcup_\nu E^{(\nu)}_{edge}$. The counting function estimate ensures that

$$|E^{(\nu)}_{edge}| \;\le\; \text{const.}\ \nu\, s^{-2\delta\nu} \left(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma}\right)^r, \tag{2.15)(ii}$$

and once again there is an estimate entirely analogous to (2.10) for the error term

$$\mathcal{D}_{edge}(f,\, g) \;=\; \sum_{I_Q \subseteq E_{edge}} c_Q\, \frac{1}{\sqrt{|I_Q|}}\, \langle\, f,\, \phi^{(1)}_Q\,\rangle\, \langle\, g,\, \phi^{(2)}_Q\,\rangle\, \phi^{(3)}_Q$$

off a third and last exceptional set E_3 defined now from E_{edge} as before. Consequently, if we denote by $\mathbb{S}^{trim}$ the tree left after trimming its edges by (2.15)(i), then

$$\mathcal{D}_{good}(f,\, g) \;=\; \mathcal{D}_{dense}(f,\, g) \;+\; \mathcal{D}_{edge}(f,\, g) \;+\; \sum_{\nu=0}^{\infty} \left[\sum_{n=1}^{O(\nu)} \left(\sum_{\mathbb{S}\, \in\, \mathcal{W}^{(\nu)}_n} \mathcal{D}_{\mathbb{S}^{trim}}(f,\, g) \right) \right],$$

and so the proof has been reduced to establishing the following 'forest' estimate for each $\mathcal{W}^{(\nu)}_n$.

Theorem. (Forest Estimate) *The inequality*

$$\frac{1}{\gamma^2} \int_{-\infty}^{\infty} \Big| \sum_{\mathbb{S}\, \in\, \mathcal{W}^{(\nu)}_n} \mathcal{D}_{\mathbb{S}^{trim}}(f,\, g)(x) \Big|^2 dx \;\le\; \text{const.}\ s^{-2\delta\nu} \left(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma}\right)^r$$

holds uniformly in f, g, γ and forest $\mathcal{W}^{(\nu)}_n$.

Combining all the previous estimates we finally deduce that

$$|\{x : |\mathcal{D}_{good}(f,g)(a_3 x)| \;\ge\; \gamma\,\}| \;\le\; \text{const.}\left(\frac{\|\pi(a_1)f\|_p\, \|\pi(a_2)g\|_q}{\gamma}\right)^r,$$

thereby completing the proof of (2.2), hence of theorem (1.9) and of Main Theorem II.

3. Bad function and Error estimates

This section is devoted to the proof of theorem (2.10). It is interesting to note that only the decay

$$|\phi(x)| \;\le\; \|\phi\| \left(\frac{1}{1+|x|^{\mu+1}}\right), \qquad |\widehat{\phi}(\xi)| \;\le\; \text{const.}\, \|\phi\| \left(\frac{1}{1+|\xi|^{\mu}}\right). \tag{3.1}$$

in time and frequency enter; no vanishing moments or Fourier support condition is required because the modulation component of the wave packets

$$\phi^{(j)}_Q(x) \;=\; s^{k/2} \phi^{(j)}(s^k x - a_j \ell)\, e^{2\pi i s^k n x} \;=\; \phi^{(j)}_I(x)\, e^{2\pi i s^k n x}, \qquad (Q = I \times w)$$

provides enough orthonormality to allow use of Hausdorff-Young arguments for Fourier series. It is also interesting to observe that this is the one point in the proof that makes essential use of the restriction $1/p + 1/q < 3/2$ (*cf.* [8], [11]). To simplify notation we proceed with a generic wave packet

$$\phi_{I\times w}(x) \;=\; s^{k/2} \phi(s^k x - a\ell)\, e^{2\pi i s^k n x} \;=\; \phi_I(x)\, e^{2\pi i s^k n x}$$

for the moment.

(3.2) Proposition. *Let ϕ be an $\mathcal{M}_\mu$-test function. Then for each τ, $1 \le \tau \le 2$, and interval I in $\mathcal{I}_{M,N}$ the inequality*

$$\Big(\frac{1}{\sqrt{|I|}}\int_{-\infty}^{\infty}\big|\sum_n d_n\, e^{2\pi i s^k xn}\big|^{\tau'}\,|\phi_I(x)|dx\Big)^{1/\tau'} \;\le\; \text{const.}\,\|\phi\|\Big(\sum_n |d_n|^\tau\Big)^{1/\tau}$$

holds uniformly in $\{d_n\}$ and I; here the left hand side is understood as an L^∞-norm with respect to the finite measure $\frac{1}{\sqrt{|I|}}\,|\phi_I(x)|/\|\phi\|$ when $\tau = 1$.

Proof. In explicit terms the integral in the left hand side is given by

$$\begin{aligned}\int_{-\infty}^{\infty}\Big(\big|\sum_n d_n\, e^{2\pi i s^k xn}\big|^{\tau'}\Big)\big|s^k\phi(s^k x - a\ell)\big|dx \\ = \int_{-\infty}^{\infty}\Big(\big|\sum_n\; d_n\, e^{2\pi i n(x+a\ell)}\big|^{\tau'}\Big)\big|\phi(x)\big|dx \\ \le\; \text{const.}\,\|\phi\|\int_0^1\big|\sum_n\; d_n\, e^{2\pi i n(x+a\ell)}\big|^{\tau'}dx.\end{aligned}$$

The result now follows immediately from the classical Hausdorff-Young inequality for Fourier series. □

There is a corresponding inequality in the reverse direction.

(3.3) Proposition. *Let ϕ be an $\mathcal{M}_\mu$-function. Then for each I in $\mathcal{I}_{M,N}$ the inequality*

$$\frac{1}{\sqrt{|I|}}\Big(\sum_w \big|\langle\, f,\, \phi_{I\times w}\,\rangle\big|^{\tau'}\Big)^{1/\tau'} \;\le\; C_\tau\,\Big(\inf_{x\in I}\; M_p(\pi(a)f)(x)\Big)$$

holds uniformly in f and I with constant

$$C_\tau \;\le\; \text{const.}\,\|\phi\|^{(2-\tau)/\tau}\,\|\pi(a)\phi\|^{2(\tau-1)/\tau}$$

for each τ, $1 < \tau < 2$, and all $p \ge \tau$.

Proof. As usual the proof proceeds by interpolation from end-point results together with a standard maximal function inequality. The case $\tau = 1$ takes the simple form

$$\sup_w \big|\langle\, f,\, \phi_{I\times w}\,\rangle\big| \;\le\; \int_{-\infty}^{\infty}|f(x)|\,|\phi_I(x)|\,dx.$$

On the other hand, in the case $\tau = 2$,

$$\sum_w\; |\langle\, f,\, \phi_{I\times w}\,\rangle|^2 \;\le\; \text{const.}\,\|\phi\|^2\Big(\int_{-\infty}^{\infty}\,|f(x)|^2\,dx\Big).$$

Indeed, after periodization,

$$\begin{aligned}\langle f, \phi_{I\times w}\rangle &= s^{-k/2}\int_{-\infty}^{\infty} f(s^{-k}x)\,\overline{\phi(x-a\ell)}\,e^{-2\pi ixn}\,dx\\ &= \int_0^1 \Big(\sum_m (\pi(s^{-k})f)(x+m)\,\overline{\phi(x+m-a\ell)}\Big)\,e^{-2\pi ixn}\,dx,\end{aligned}$$

and so

$$\begin{aligned}\sum_w |\langle f, \phi_{I\times w}\rangle|^2 &= \int_0^1 |\sum_m (\pi(s^{-k})f)(x+m)\,\overline{\phi(x+m-a\ell)}|^2\,dx\\ &\le \text{ const. } \|\phi\|^2 \int_0^1 |\sum_m (\pi(s^{-k})f)(x+m)|^2\,dx\end{aligned}$$

uniformly in a and ℓ. The case $\tau = 2$ follows. Hence by interpolation

$$\begin{aligned}\Big(\sum_w |\langle f,\, \phi_{I\times w}\rangle|^{\tau'}\Big)^{1/\tau'} &\le \text{ const. } \|\phi\|^{2(\tau-1)/\tau}\Big(\int_{-\infty}^{\infty}|f(x)|^{\tau}\,|\phi_I(x)|^{2-\tau}\,dx\Big)^{1/\tau}\\ &\le \text{ const. } \|\phi\|^{2(\tau-1)/\tau}\Big(\int_{-\infty}^{\infty}|\phi_I(x)|^{2-\tau}\,dx\Big)^{1/\tau-1/p}\Big(\int_{-\infty}^{\infty}|f(x)|^{p}\,|\phi_I(x)|^{2-\tau}\,dx\Big)^{1/p}\end{aligned}$$

for each τ, $1 \le \tau \le 2$, and all $p \ge \tau$. But

$$\int_{-\infty}^{\infty} |\phi_I(ax)|^{2-\tau}\,dx \;\le\; \text{const.}\Big(\frac{\|\pi(a)\phi\|}{a^{1/2}}\Big)^{2-\tau}|I|^{\tau/2},$$

while

$$\int_{-\infty}^{\infty}|f(ax)|^p\,|\phi_I(ax)|^{2-\tau}\,dx \;\le\; \text{const.}\|\frac{1}{a^{1/2}}\pi(a)\phi\|^{2-\tau}\Big(\frac{1}{a^{1/2}}\inf_{x\in I} M_p(\pi(a)f)(x)\Big)^p|I|^{\tau/2}.$$

The proposition now follows easily by interpolation after a change of variable. □

Combining these inequalities we obtain a key decay estimate.

(3.4) Proposition. *Let $\phi^{(i)}$, $1 \le i \le 3$, be $\mathcal{M}_\mu$-test functions. Then the inequality*

$$\begin{aligned}&\int_{\mathbb{R}\setminus s^m I} |\sum_w c_{I\times w}\,\frac{1}{\sqrt{|I|}}\langle f,\, \phi^{(1)}_{I\times w}\rangle\,\langle g,\, \phi^{(2)}_{I\times w}\rangle\;\phi^{(3)}_{I\times w}(a_3x)\,|\,dx\\ &\qquad\le\; \frac{\text{const.}}{s^{m\mu/2}}\,\|\{c_{I\times w}\}\|_\infty\Big(\inf_{x\in I} M_p(\pi(a_1)f)(x)\Big)\Big(\inf_{x\in I} M_q(\pi(a_2)g)(x)\Big)\frac{|I|}{a_3}\end{aligned}$$

holds uniformly in $f, g, c_{I\times w}$ and $m > 0$ for all I in $\mathcal{I}_{M,N}$ whenever $p, q > 1$ and $1/p + 1/q < 3/2$.

Proof. Choose σ and ρ so that

$$1 < \sigma,\, \rho < 2, \qquad \sigma \le p, \qquad \rho \le q, \qquad \frac{1}{\sigma}+\frac{1}{\rho} < \frac{3}{2},$$

and then set $1/\tau = 1/\sigma' + 1/\rho'$; this is possible because of the restrictions on p and q. By Hölder's inequality,

$$\int_{\mathbb{R}\setminus s^m I} \big|\sum_w d_w\, \phi^{(3)}_{I\times w}(x)\big|\, dx \;\le\; \Big(\int_{\mathbb{R}\setminus s^m I} |\phi^{(3)}_I(a_3x)|\, dx\Big)^{1/\tau}$$

$$\times\; \Big(\frac{1}{a_3}\int_{-\infty}^{\infty} \big|\sum_n d_n\, e^{2\pi i s^k x n}\big|^{\tau'}\, |\phi^{(3)}_I(x)|\, dx\Big)^{1/\tau'},$$

writing

$$d_n \;=\; d_w \;=\; c_{I\times w}\,\langle\, f,\, \phi^{(1)}_{I\times w}\,\rangle\,\langle\, g,\, \phi^{(2)}_{I\times w}\,\rangle.$$

Because of decay in time

$$\frac{1}{\sqrt{|I|}}\int_{\mathbb{R}\setminus s^m I} |\phi^{(3)}_I(a_3x)|\, dx \;\le \text{const.}\, \frac{1}{s^{m\mu}}\, \|\pi(a_3)\phi^{(3)}\|,$$

while the restrictions on σ and ρ ensure that $1 < \tau < 2$ so that (3.3) can be applied to the second integral. In this case

$$\Big(\frac{1}{\sqrt{|I|}}\int_{-\infty}^{\infty} \big|\sum_n d_n\, e^{2\pi i n x}\big|^{\tau'} |\phi_I(x)|\, dx\Big)^{1/\tau'} \;\le\; \text{const.}\Big(\sum_n |d_n|^{\tau}\Big)^{1/\tau}$$

$$\le\; \text{const.}\; \|\{c_{I\times w}\}\|_\infty \Big(\sum_w |\langle\, f,\, \phi^{(1)}_{I\times w}\,\rangle|^{\sigma'}\Big)^{1/\sigma'} \Big(\sum_w |\langle\, g,\, \phi^{(2)}_{I\times w}\,\rangle|^{s'}\Big)^{1/s'}$$

because of the way τ was chosen. But now we can apply (3.3) to f, g because of the way σ and s were chosen. The proposition follows. □

The inequality in (3.4) is the crucial special case of the following fundamental decay estimate.

(3.5) Proposition. *Let $\phi^{(i)}$, $1 \le i \le 3$, be $\mathcal{M}_\mu$-test functions and let J be an arbitrary interval in $\mathcal{I}_{M,N}$. Then the inequality*

$$\int_{\mathbb{R}\setminus s^2 J} \Big|\sum_{I_Q\subseteq J} \frac{c_Q}{\sqrt{|I_Q|}}\langle\, f,\, \phi^{(1)}_Q\,\rangle\,\langle\, g,\, \phi^{(2)}_Q\,\rangle\, \phi^{(3)}_Q(a_3x)\Big|\, dx$$

$$\le\; \text{const.}\, \frac{|J|}{a_3}\Big(\inf_{x\in J} M_p(\pi(a_1)f)(x)\, \inf_{x\in J} M_q(\pi(a_2)g)(x)\Big)$$

holds uniformly in f, g and J whenever $p, q > 1$ and $1/p + 1/q < 3/2$.

Proof. It is the presence of the dilation factor s^2 that allows us to extend the sum from tiles with $I_Q = J$ to all those for which $I_Q \subseteq J$. Fix $k > 0$ and let I be an interval in $\mathcal{I}_{M,N}$ with $I \subseteq J$, $|I| = s^{-k}|J|$. Then

$$|I| \;=\; s^{-k}|J| \quad\Longrightarrow\quad J \;\subseteq\; s^{k+2}I \;\subseteq\; s^2 J.$$

Because of the last of these inclusions,

$$\int_{\mathbb{R}\setminus s^2 J} \Big|\sum_{I_Q=I} \frac{c_Q}{\sqrt{|I_Q|}} \langle\, f,\, \phi_Q^{(1)}\,\rangle\,\langle\, g,\, \phi_Q^{(2)}\,\rangle\, \phi_Q^{(3)}(a_3x)\Big|\, dx$$

$$\leq\ \text{const.}\ \frac{|I|}{s^{k\mu/2}} \left(\inf_{x\in I} M_p(\pi(a_1)f)(x)\ \inf_{x\in I} M_q(\pi(a_2)g)(x)\right).$$

On the other hand, because of the first of the inclusions, the estimate

$$\inf_{x\in I} M_p(\pi(a_1)f)(x)\ \leq\ \text{const.}\ s^k \inf_{x\in J} M_p(\pi(a_1)f)(x)$$

holds together with a corresponding one for g. Summing first over all $I \subseteq J$, $|I| = s^{-k}|J|$, and then over all $k > 0$, we finally obtain (3.5). $\square$

Proof of theorem (2.10). Let $I_1, I_2, \ldots$, be the maximal intervals in $\mathcal{I}_{M,N}$ such that $I_j \subseteq E_{bad}$. Maximality ensures that the I_j are disjoint, but it also ensures that the next larger interval to I_j in $\mathcal{I}_{M,N}$ is not contained in E_{bad} which in turn ensures that $s^2 I_j \not\subseteq E_{bad}$. Maximality of the I_j thus ensures that

$$\inf_{x\in I_j} M_p(M(\pi(a_1)f)(x)\ \leq\ \text{const.}\ \inf_{x\in s^2 I_j} M_p(M(\pi(a_1)f)(x)\ \leq\ \text{const.}\ \kappa_p,$$

and

$$\inf_{x\in I_j} M_q(Mg)(x)\ \leq\ \text{const.}\ \inf_{x\in s^2 I_j} M_q(Mg)(x)\ \leq\ \text{const.}\ \kappa_q,$$

both hold uniformly in j. By Proposition (3.5), therefore,

$$\int_{\mathbb{R}\setminus E_1} \big|\mathcal{D}_{bad}(f,g)(a_3x)\big|\, dx$$

$$\leq\ \sum_j \left(\int_{\mathbb{R}\setminus s^2 I_j} \Big|\sum_{I_Q\subseteq I_j} \frac{c_Q}{\sqrt{|I_Q|}}\, \langle\, f,\, \phi_Q^{(1)}\,\rangle\,\langle\, g,\, \phi_Q^{(2)}\,\rangle\, \phi_Q^{(3)}(x)\Big|\, dx\right)$$

$$\leq\ \text{const.}\ \sum_j \frac{|I_j|}{a_3} \left(\inf_{x\in I_j} M_p(\pi(a_1)f)(x)\ \inf_{x\in I_j} M_q(\pi(a_2)g)(x)\right)$$

from which the theorem follows since $\kappa_p\kappa_q = \gamma$ (*cf.* (2.9)). $\square$

The proof of theorem (2.10) applies more generally to produce a generic decay estimate. Let Ω be a measurable subset of $\mathbb{R}$ and let

$$E_\Omega\ =\ \bigcup \{s^2 I_Q : I_Q\ \subseteq\ \Omega,\ I_Q\ \not\subseteq\ E_{bad}\}.$$

Now let

$$\mathcal{D}_\Omega(f,\, g)\ =\ \sum_{I_Q\subseteq\Omega,\, I_Q\not\subseteq E_{bad}} \frac{1}{\sqrt{|I_Q|}}\, c_Q\,\big\langle\, f,\, \phi_Q^{(1)}\,\big\rangle\,\big\langle\, g,\, \phi_Q^{(2)}\,\big\rangle\, \phi_Q^{(3)}$$

be the operator obtained by summing over those tiles whose time interval lies in Ω but 'pokes outside' the bad set E_{bad}; in particular, therefore, $\mathcal{D}_\Omega$ is just part of $\mathcal{D}_{good}$. For instance, the operators $\mathcal{D}_{dense}$ and $\mathcal{D}_{edge}$ defined in the previous section have this form. Then $\mathcal{D}_\Omega$ can be estimated away from E_Ω in the same way as $\mathcal{D}_{bad}$ was.

(3.6) Theorem. *With the notation above the inequality*

$$\frac{1}{\gamma}\int_{\mathbb{R}\setminus E_\Omega} |\mathcal{D}_\Omega(f,\, g)(a_3 x)|\, dx \;\le\; \text{const.}\ \frac{|\Omega|}{a_3}$$

holds uniformly in Ω *and* γ.

Proof. The proof follows that of (2.10). Choose intervals $I_1, I_2, \ldots$ which are maximal in $\mathcal{I}_{M,N}$ and such that $I_j \subseteq \Omega$. By (3.4)

$$\int_{\mathbb{R}\setminus E_\Omega} \big|\mathcal{D}_\Omega(f,g)(a_3x)\big|\, dx$$
$$\le \sum_j \Big(\int_{\mathbb{R}\setminus s^2 I_j} \big|\sum_{I_Q\subseteq I_j} \frac{1}{\sqrt{|I_Q|}}\, c_Q \,\langle\, f,\, \phi_Q^{(1)}\,\rangle \,\langle\, g,\, \phi_Q^{(2)}\,\rangle\, \phi_Q^{(3)}(a_3x)\big|\, dx\Big)$$
$$\le \ \text{const.} \sum_j \frac{|I_j|}{a_3}\Big(\inf_{x\in I_j} M_p(\pi(a_1)f)(x)\ \inf_{x\in I_j} M_q(\pi(a_2)g)(x)\Big).$$

From this the theorem follows immediately since $I_j \not\subseteq E_{bad}$ and maximality ensures that the I_j are disjoint. □

4. Trees

In this section the families $\mathcal{F}_\nu$ of trees appearing in section 2 are chosen. Recall first that a tree with top $\overline{Q}$ consists of tiles Q in $\mathbb{Q}^{(+)}_{M,N}$ such that $Q \le \overline{Q}$ with respect to the partial order introduced in (2.11) and the tree top is one of the tiles in the tree. This maximality requirement imposes a crucial arithmetical-combinatorial structure on a tree.

(4.1) Proposition. *Let* $Q \sim \{k,\ell,m\}$, $\overline{Q} \sim \{j,\overline{\ell},\overline{m}\}$ *be tiles in* $\mathbb{Q}_s$ *such that* $Q \le \overline{Q}$. *Then*

$$j \,\le\, k \,\le\, \sigma+1+j, \qquad s^{k-j}\overline{\ell} \,\le\, \ell \,<\, s^{k-j}(\overline{\ell}+1)$$

and

$$m \;=\; \begin{cases} [s^{j-k}\overline{m}], & M=1,\\ [s^{j-k}\overline{m}] \ \text{or}\ [s^{j-k}\overline{m}]+1, & M\ge 2,\end{cases}$$

where $\sigma = [\log_s \overline{m}]$.

Proof. If $I_Q \subseteq I_{\overline{Q}}$, then

$$s^{-k} \,\le\, s^{-j}, \qquad s^{-j}\overline{\ell} \,\le\, s^{-k}\ell, \qquad s^{-k}(\ell+1) \,\le\, s^{-j}(\overline{\ell}+1)$$

so it is clear that $k \ge j$ and $s^{k-j}\overline{\ell} \le \ell < s^{k-j}(\overline{\ell}+1)$. On the other hand, control on m depends on the value of M. If $M=1$ and $w_Q \supseteq w_{\overline{Q}}$ then

$$s^k m \,\le\, s^j\overline{m}, \qquad s^j(\overline{m}+1) \,\le\, s^k(m+1),$$

so $0 \le s^{j-k}\overline{m} - m \le (1 - s^{j-k})$; in addition, $m < s^{\sigma+j-k+1}$. Thus $m = [s^{j-k}\overline{m}]$ and $k - j \le \sigma + 1$. But if $M \ge 2$ and $w_Q \supseteq w_{\overline{Q}}$, then

$$s^k(m - \alpha_M) \le s^j(\overline{m} - \alpha_M), \qquad s^j(\overline{m} + \alpha_M) \le s^k(m + \alpha_M).$$

In this case $|s^{j-k}\overline{m} - m| \le \alpha_M(1 - s^{j-k}) < \frac{1}{2}$. Hence, either $m = [s^{j-k}\overline{m}]$ or $m = [s^{j-k}\overline{m}] + 1$, but not both. This completes the proof. $\square$

The proof above brings out very clearly that the condition $Q \le \overline{Q}$ specifies the frequency component m of Q *uniquely* in terms of the frequency component of $\overline{Q}$. In particular, therefore, the oscillation of wave packets associated with a tree is completely determined by the tree-top of that tree. This control of oscillation can be quantified in a very important and useful way.

(4.2) Definition. *If $\overline{Q} \sim \{\, j,\, \overline{\ell},\, \overline{m}\,\}$ is the tree-top of a tree $\mathbb{T}$ of tiles in $\mathbb{Q}^{(+)}_{M,N}$ set $\lambda_{\mathbb{T}} = s^j\overline{m}$. In other words, $\lambda_{\mathbb{T}}$ is the left hand end-point of the frequency interval $w_{\mathbb{T}}$ of the tree-top of $\mathbb{T}$ when $M = 1$, whereas $\lambda_{\mathbb{T}}$ is the mid-point of $w_{\mathbb{T}}$ when $M \ge 2$.*

In addition to the numerical and geometric definitions, $\lambda_{\mathbb{T}}$ has an interesting group interpretation. For $Q \sim \{\, j,\, \ell,\, m\,\}$ let $\tau_Q : \xi \longrightarrow s^k(\xi + m)$ denote the affine transformation of $\mathbb{R}$ taking the basic interval onto w_Q; then $\lambda_{\mathbb{T}} = \tau_{\overline{Q}}(0)$. In other words, $\lambda_{\mathbb{T}}$ is the image of the origin under $\tau_{\overline{Q}}$. Together with (4.1) this value $\lambda_{\mathbb{T}}$ provides the crucial link between a tree and the Fourier support condition. Indeed, for a tile $Q \sim \{k,\, \ell,\, m\}$ the condition $Q \le \overline{Q}$ ensures that $m = [s^{-k}\lambda_{\mathbb{T}}]$ or $m = [s^{-k}\lambda_{\mathbb{T}}] + 1$, while a simple calculation shows that

$$s^{-k}\lambda_{\mathbb{T}} - m \in w^{(i)} \quad \Longleftrightarrow \quad \lambda_{\mathbb{T}} \in w_Q^{(i)}.$$

Thus, if the original $\mathcal{M}_\mu$-test function $\phi^{(i)}$ is replaced by its modulate

$$\psi^{(i)}(x) = \phi^{(i)}(x)\, e^{2\pi i(m - s^{-k}\lambda_{\mathbb{T}})x},$$

then

$$\psi^{(i)}_{I_Q}(x) = s^{k/2}\psi^{(i)}(s^k x - a_i\ell) = \gamma_{I_Q}\, \phi^{(i)}_Q(x)\, e^{2\pi i\lambda_{\mathbb{T}}x}, \qquad |\gamma_{I_Q}| = 1$$

and

$$\int_{-\infty}^{\infty} \psi^{(i)}(x)\, dx = \widehat{\phi}^{(i)}(s^{-k}\lambda_{\mathbb{T}} - m) = 0 \quad \Longleftrightarrow \quad \lambda_{\mathbb{T}} \notin w_Q^{(i)}.$$

Hence up to modulation, a unitary operator on $L^2(\mathbb{R})$, the time-frequency paraproduct

$$\mathcal{D}_{\mathbb{T}}(f,\, g) = \sum_{Q \in \mathbb{T}} \frac{1}{\sqrt{|I_Q|}}\, c_Q \,\langle\, f,\, \phi^{(1)}_Q \,\rangle \,\langle\, g,\, \phi^{(2)}_Q \,\rangle\, \phi^{(3)}_Q$$

obtained by summing over all tiles in the tree $\mathbb{T}$ reduces to a *standard* paraproduct

$$\mathcal{D}_{\mathbb{T}}(f,\, g) = \Big(\sum_{Q \in \mathbb{T}} \frac{1}{\sqrt{|I_Q|}}\, d_Q \,\langle F,\, \psi^{(1)}_{I_Q} \rangle \,\langle G,\, \psi^{(2)}_{I_Q} \rangle\, \psi^{(3)}_{I_Q}\Big)\, e^{-2\pi i\lambda_{\mathbb{T}}x}$$

where $F(x) = f(x)\, e^{-2\pi i \lambda_{\mathbb{T}} x}$ and $G(x) = g(x)\, e^{-2\pi i \lambda_{\mathbb{T}} x}$ are modulates of f and g, while $|d_{I_Q}| = |c_{I_Q}|$. Furthermore, the Fourier support condition ensures that at least two of the $\psi^{(j)}$ have vanishing moment (except possibly for the 'exceptional term' corresponding to $I_{\mathbb{T}} \times w_{\mathbb{T}}$). This will dictate how we choose trees.

First the specific structure imposed in (1.6)–(1.13) is used to introduce partial orders on $\mathbb{Q}^{(+)}_{M,N}$ stronger than (2.11).

(4.3) Definition. *Given tiles P, Q in $\mathbb{Q}^{(+)}_{M,N}$ we write*

$$P \prec_i Q \quad \Longleftrightarrow \quad I_P \subseteq I_Q, \quad w_Q^{(i)} \subseteq w_P^{(i)}.$$

Since $\{w_Q^{(i)} : Q \in \mathbb{Q}_{M,N}\}$ is a grid, this defines a partial order on $\mathbb{Q}^{(+)}_{M,N}$.

As $w_Q^{(i)} \subseteq w_Q$ and $\{w_Q : Q \in \mathbb{Q}^{(+)}_{M,N}\}$ is a grid, $\prec_i$ is stronger than $\leq$ in the sense that

$$P \prec_i Q \quad \Longrightarrow \quad P \leq Q;$$

in particular, a tree with respect to $\prec_i$ will automatically be a tree with respect to the partial order $\leq$ and each will have the same tree-top. This prompts the following definition.

(4.4) Definition. *A set $\mathbb{T}$ of tiles in $\mathbb{Q}^{(+)}_{M,N}$ is said to be a $\Lambda^{(i)}$-tree when it is a tree with respect to the partial order $\prec_i$.*

Notice that every tile Q by itself is a $\Lambda^{(i)}$-tree having itself as tree-top since $Q \prec_i Q$ for all i; this will be important shortly. The following result provides a key property of $\Lambda^{(i)}$-trees.

(4.5) Proposition. *Let $\mathbb{T}$ be a $\Lambda^{(i)}$-tree having tree-top $\overline{Q}$. Then $\lambda_{\mathbb{T}} \in w_Q^{(i)}$, but $\lambda_{\mathbb{T}} \notin w_Q^{(j)}$ for all $j \neq i$ whenever Q is a tile in $\mathbb{T}$ and $Q \neq \overline{Q}$.*

As the previous discussion indicates, the significance of (4.5) is to guarantee precisely when tiles in a $\Lambda^{(i)}$-tree have vanishing moment. Standard frame techniques can thus be used on functions associated with trees.

Proof. Fix Q in $\mathbb{T}$, $Q \neq \overline{Q}$; in particular, therefore, $|I_Q| < |I_{\mathbb{T}}|$ and $w_{\mathbb{T}} \cap w_Q^{(i)} \neq \emptyset$. But then property (1.12) ensures that $w_{\mathbb{T}} \subseteq w_Q^{(i)}$, and so $\lambda_{\mathbb{T}}$ must belong to $w_Q^{(i)}$. This completes the proof since the $w_Q^{(i)}$ are disjoint for a given Q. □

Trees can now be selected. We define nested families $\{\mathbb{Q}_\nu\}$,

$$\emptyset \subseteq \ldots \subseteq \mathbb{Q}_\nu \subseteq \mathbb{Q}_{\nu-1} \subseteq \ldots \subseteq \mathbb{Q}_{-1} = \left\{Q \in \mathbb{Q}^{(+)}_{M,N} : I_Q \not\subseteq E_{bad}\right\},$$

of tiles recursively by choosing families $\mathcal{F}_\nu = \bigcup_{i,j} \mathcal{F}_\nu^{(ij)}$ of trees so that

$$\mathbb{Q}_{\nu-1} \setminus \mathbb{Q}_\nu = \bigcup_{\mathbb{T} \in \mathcal{F}_\nu} \{Q : Q \in \mathbb{T}\}.$$

The choice will be made using size estimates whose connection with familiar ideas in Tent space theory will be brought out shortly. To begin, notice that the inequalities in (2.8) provide estimates for individual tiles in $\mathbb{Q}_{-1}$. On the other hand, by applying the 'Lusin area' result (A.6) in Appendix A we also have control on trees of tiles in $\mathbb{Q}_{-1}$.

(4.6) Proposition. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree in $\mathbb{Q}_{-1}$, $j \neq 1$. Then the inequality*

$$\frac{1}{|I_{\mathbb{T}}|}\int_{I_{\mathbb{T}}}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}\big|\langle\, f,\, \phi_Q^{(1)}\,\rangle\big|^2\chi_{I_Q}(x)\Big)^{1/2}dx \;\le\; \mathrm{const}_\phi\, s^{-1/\eta}\,\kappa_p$$

holds uniformly in $\mathbb{T}$ with constant

$$\mathrm{const}_\phi \;\le\; \mathrm{const.}\ \|\pi(a_1)\phi^{(1)}\|$$

for the function $f \in L^p(\mathbb{R})$ fixed earlier.

There is an entirely analogous result for the g in $L^q(\mathbb{R})$ fixed at the outset of section 2.

(4.7) Proposition. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree in $\mathbb{Q}_{-1}$, $j \neq 2$. Then the inequality*

$$\frac{1}{|I_{\mathbb{T}}|}\int_{I_{\mathbb{T}}}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}\big|\langle\, g,\, \phi_Q^{(1)}\,\rangle\big|^2\chi_{I_Q}(x)\Big)^{1/2}dx \;\le\; \mathrm{const}_\phi\, s^{-1/\eta}\,\kappa_p$$

holds uniformly in $\mathbb{T}$ with constant

$$\mathrm{const}_\phi \;\le\; \mathrm{const.}\ \|\pi(a_2)\phi^{(2)}\|$$

for the function $g \in L^q(\mathbb{R})$ chosen earlier.

At the expense of using a possibly larger choice of the constant in (2.8)(ii), we can (and shall) assume that the values of const_ϕ throughout (2.8), (4.6) and (4.7) are all the same. From now on, this choice of const_ϕ will fixed.

Suppose then that $\mathbb{Q}_{\nu-1}$ has been defined already for some $\nu \geq 0$. We single out families $\mathcal{F}_\nu^{(ij)}$ of $\Lambda^{(j)}$-trees in $\mathbb{Q}_{\nu-1}$, dealing first with the case $\mathcal{F}_\nu^{(23)}$. Denote by $\overline{\mathcal{Q}}$ those tiles $\overline{Q}$ in $\mathbb{Q}_{\nu-1}$ for which there is a $\Lambda^{(3)}$-tree $\mathbb{T}$ having $\overline{Q}$ as top and satisfying

$$\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2}dx \;\ge\; \mathrm{const}_\phi\kappa_q s^{-(\nu+1)/q_0} \tag{4.8}$$

where $q_0 = \max(q, q')$ as before. With $\lambda_{\overline{Q}} = \tau_{\overline{Q}}(0)$ as in (4.2) set

$$\lambda_1 \;=\; \min_{\overline{Q}\,\in\,\overline{\mathcal{Q}}} \lambda_{\overline{Q}}$$

and then select $\overline{Q}_1$ with maximal time interval among those tiles in $\overline{\mathcal{Q}}$ for which $\lambda_{\overline{Q}} = \lambda_1$. Now set

$$\mathbb{T}_1 \;=\; \{\, Q \in \mathbb{Q}_{\nu-1} : Q \underset{3}{\prec} \overline{Q}_1\,\},$$

thereby producing the largest $\Lambda^{(3)}$-tree in $\mathbb{Q}_{\nu-1}$ having $\overline{Q}_1$ as top. We repeat this construction starting with $\mathbb{Q}_{\nu-1} \setminus \{Q : Q \in \mathbb{T}_1\}$ to produce a tree $\mathbb{T}_2$ and so on until all $\Lambda^{(3)}$-trees in $\mathbb{Q}_{\nu-1}$ for which (4.8) holds have been selected. This produces a family $\mathcal{F}_\nu^{(23)}$ of trees $\mathbb{T}_\ell$ which are maximally ordered upwards in the sense that the points $\lambda_\ell = \tau_{\overline{Q}_\ell}(0)$ satisfy $\lambda_1 \le \lambda_2 \le \dots$. Moreover, when $m < \ell$,

$$Q \in \mathbb{T}_\ell, \quad I_Q \subsetneq I_{\mathbb{T}_\ell}, \quad I_Q \subseteq I_{\mathbb{T}_m} \quad \Longrightarrow \quad \lambda_m < w_Q^{(3)}, \tag{4.9}$$

meaning that λ_m lies to the left of $w_Q^{(3)}$. Indeed, since $\mathbb{T}_\ell$ is a $\Lambda^{(3)}$-tree, proposition (4.5) ensures $\lambda_\ell \in w_Q^{(3)}$. So if λ_m does not lie to the left of $w_Q^{(3)}$, then $\lambda_m \in w_Q^{(3)}$ because $\lambda_m \le \lambda_\ell \in w_Q^{(3)}$ implies $w_{\mathbb{T}_m} \cap w_Q^{(3)} \neq \emptyset$. But then we must have that $I_Q \subsetneq I_{\mathbb{T}_m}$ by maximality since $\mathbb{T}_\ell \neq \mathbb{T}_m$ and $Q \in \mathbb{T}_\ell$. Hence $|I_Q| < |I_{\mathbb{T}_m}|$, and so $w_{\mathbb{T}_m} \subseteq w_Q^{(3)}$ by (1.12). But then in particular we have that $w_{\mathbb{T}_m}^{(3)} \subseteq w_Q^{(3)}$ which implies that $Q \prec_3 \overline{Q}_m$. Thus by maximality Q should have been picked when $\mathbb{T}_m$ was selected. Therefore we must have that $\lambda_m < w_Q^{(3)}$. Denote by $\mathcal{Q}_\nu^{(23)}$ the set of all tiles in the trees in $\mathcal{F}_\nu^{(23)}$.

(4.10) Remark. The argument above together with the one in (4.5) prove in fact that: $\lambda_{\mathbb{T}} \in w_Q^{(j)}$ and $I_Q \subsetneq I_{\mathbb{T}}$ for some $\Lambda^{(j)}$-tree with maximal $\overline{Q} = I_{\mathbb{T}} \times w_{\mathbb{T}} \Longleftrightarrow Q \prec_j \overline{Q}$; i.e. $Q \prec_j \overline{Q} \Longleftrightarrow \lambda_{\mathbb{T}} \in w_Q^{(j)}$ and $I_Q \subsetneq I_{\mathbb{T}}$.

Next we choose a corresponding maximally ordered *downwards* collection of $\Lambda^{(1)}$-trees from $\mathbb{Q}_{\nu-1} \smallsetminus \mathcal{Q}_\nu^{(23)}$ such that

$$\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\ge\; \mathrm{const}_\phi \kappa_q s^{-(\nu+1)/q_0}. \tag{4.11}$$

The union of these $\Lambda^{(1)}$-trees form the family $\mathcal{F}_\nu^{(21)}$ and the set of all tiles in these $\Lambda^{(1)}$-trees will be denoted by $\mathcal{Q}_\nu^{(21)}$. To define the corresponding families $\mathcal{F}_\nu^{(1j)}$ and $\mathcal{Q}_\nu^{(1j)}$, $j > 1$, we repeat the construction with $\mathrm{const}_\phi \kappa_p s^{-(\nu+1)/p_0}$ as lower bound and f, $\phi_Q^{(1)}$ instead of g, $\phi_Q^{(2)}$. Since $j > 1$ in both cases we again maximally order upwards.

Finally, we repeat the above method of construction, with no prescribed order, to choose $\Lambda^{(2)}$ trees in

$$\mathbb{Q}_{\nu-1} \setminus \Big(\bigcup_{i=1}^{2}\Big\{\bigcup_{j\neq i}\mathcal{Q}_\nu^{(ij)}\Big\}\Big)$$

such that

$$\min_{Q\in\mathbb{T}}\Big(\frac{1}{|I_{\mathbb{T}}|^{1/2}}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|\Big) \;\ge\; \mathrm{const}_\phi \kappa_q s^{-(1+\eta)(\nu+1)/q_0}. \tag{4.12}$$

These $\Lambda^{(2)}$-trees together with the corresponding $\Lambda^{(1)}$-trees defined with respect to f form the families $\mathcal{F}_\nu^{(ii)}$; the tiles in these trees will be denoted by $\mathcal{Q}_\nu^{(ii)}$. Setting

$$\mathbb{Q}_\nu \;=\; \mathbb{Q}_{\nu-1} \setminus \Big(\bigcup_{i=1}^{2}\Big\{\bigcup_{j=1}^{3}\mathcal{Q}_\nu^{(ij)}\Big\}\Big) \tag{4.13}$$

completes the inductive construction.

(4.14) Remark. Although maximal intervals do not always exist in $\mathcal{I}_{M,N}$, they do in the context of the construction above. For by general frame results, the inequality

$$\Big(\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{q/2}\, dx\Big)^{1/q} \;\le\; \mathrm{const}_\phi\,\|g\|_q$$

holds uniformly in g and $\mathbb{T}$. Consequently,

$$\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle g, \phi_Q^{(2)}\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le \mathrm{const}_\phi\Big(\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}|g(x)|^q\, dx\Big)^{1/q}.$$

Hence (4.8) and (4.11) ensure that

$$|I_{\mathbb{T}}| \;\le\; \mathrm{const.}(\mathrm{const}_\phi)\Big(\frac{s^{\nu+1/q_0}}{\kappa_q}\Big)^q\int_{-\infty}^{\infty}|g(x)^q\, dx.$$

In other words, the Q in $\overline{\mathcal{Q}}$ have uniformly bounded time intervals once we fix q, g, s and ν, as we may. It thus makes sense to ask for a $\overline{Q}$ having maximal time interval. On the other hand,as the functions f and g fixed in section 2 are assumed to be band-limited the possible values of $\lambda_{\overline{Q}}$ will be bounded above and below, so it also makes sense to ask for minimum or maximum values λ_ℓ.

It will be helpful to summarize the properties of $\mathbb{Q}_\nu$ that follow immediately from the ν-th stage construction. They are refinements of *a priori* estimates (2.8), (4.6) and (4.7).

(4.15) Remark (i). The inequalities

$$\frac{1}{\sqrt{|I_Q|}}|\langle\, f,\, \phi_Q^{(1)}\,\rangle| \;\le\; \mathrm{const}_\phi s^{-(1+\eta)(1+\nu)/p_0}\, s^{-1/\eta}\kappa_p$$

and

$$\frac{1}{\sqrt{|I_Q|}}|\langle\, g,\, \phi_Q^{(2)}\,\rangle| \;\le\; \mathrm{const}_\phi s^{-(1+\eta)(1+\nu)/q_0}\, s^{-1/\eta}\kappa_q$$

hold for all Q in $\mathbb{Q}_\nu$ since by itself each Q is a $\Lambda^{(i)}$-tree.

(ii) The inequality

$$\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle\, f,\, \phi_Q^{(1)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le\; \mathrm{const}_\phi\kappa_p s^{-(\nu+1)/p_0}$$

holds for all $\Lambda^{(j)}$-trees in $\mathbb{Q}_\nu$, $j \ne 1$.

(iii) The inequality

$$\frac{1}{|I_{\mathbb{T}}|}\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le\; \mathrm{const}_\phi\kappa_q s^{-(\nu+1)/q_0}$$

holds for all $\Lambda^{(j)}$-trees in $\mathbb{Q}_\nu$, $j \neq 2$.

Because of (4.15)(i) the intersection of the $\mathbb{Q}_\nu$ can consist only of tiles Q for which $\langle f, \phi_Q^{(1)}\rangle = \langle g, \phi_Q^{(2)}\rangle = 0$, none of which contribute to $\mathcal{D}_{good}(f,g)$. Thus

$$\mathcal{D}_{good}(f,\, g) = \sum_{\nu=0}^{\infty}\Big(\sum_{\mathbb{T}\in\mathcal{F}_\nu}\mathcal{D}_{\mathbb{T}}(f,\, g)\Big), \qquad \mathcal{F}_\nu = \bigcup_{i=1}^{2}\Big(\bigcup_{j=1}^{3}\mathcal{F}_\nu^{(ij)}\Big)$$

provides the decomposition used in section 2. One remarkable consequence of this construction is that (4.15)(ii),(iii) remain valid for any interval J in $I_{\mathbb{T}}$, not just for $I_{\mathbb{T}}$ itself, leading in Section 5 to a *Carleson measure type estimate.*

(4.16) Proposition. *Let $\mathbb{T}$ be a tree in $\mathcal{F}_\nu^{(2j)}$ with $j \neq 2$ and J a subinterval of $I_{\mathbb{T}}$ which need not be an s-adic interval. Then the inequality*

$$\frac{1}{|J|}\int_J\Big(\sum_{I_Q\subseteq J}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2}dx \;\le\; \mathrm{const}_\phi\kappa_q s^{-(\nu+1)/q_0}$$

holds uniformly in $\mathbb{T}$, J and ν, the inner sum being taken over all Q in $\mathbb{T}$ with $I_Q\subseteq J$.

The proof of the corresponding result

$$\sup_{J\subseteq I_{\mathbb{T}}}\left[\frac{1}{|J|}\int_J\Big(\sum_{I_Q\,\subseteq\, J}\frac{1}{|I_Q|}|\langle f,\phi_Q^{(1)}\rangle|^2\chi_{I_Q}(x)\Big)^{1/2}dx\right] \;\le\; \mathrm{const}_\phi\kappa_p s^{-\nu/p_0} \tag{4.17}$$

for $\mathbb{T}$ in $\mathcal{F}_\nu^{(ij)}$, $j\neq i$, differs only in notation and details are left to the reader. In order to establish (4.16) a simple lemma guaranteeing that certain subsets of a $\Lambda^{(j)}$-tree are again $\Lambda^{(j)}$-trees is needed.

(4.18) Lemma. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree and Q' a tile in $\mathbb{T}$. Then the set*

$$\mathbb{S} \;=\; \{\, Q\in\mathbb{T}:\; I_Q\subseteq I_{Q'}\,\}$$

is a $\Lambda^{(j)}$-tree having Q' as its top. Of course, $\mathbb{S}=\mathbb{T}$ when $I_{Q'}=I_{\mathbb{T}}$.

Proof. We have to show that

$$I_Q\subseteq I_{Q'} \quad\Longrightarrow\quad w_{Q'}^{(j)}\subseteq w_Q^{(j)}\ .$$

But if $\overline{Q}$ is the tree top of $\mathbb{T}$, then $w_{Q'}^{(j)}\cap w_Q^{(j)}\supseteq w_{\overline{Q}}^{(j)}$ for any $Q\in\mathbb{T}$; on the other hand, $|w_{Q'}^{(j)}|\le|w_Q^{(j)}|$ for any $Q\in\mathbb{S}$. By (1.11) and (1.12), therefore, $w_{Q'}^{(j)}\subseteq w_Q^{(j)}$. □

Proof of (4.16). Fix an interval J and set $\mathbb{S}_J=\big\{Q\in\mathbb{T}: I_Q\subseteq J\big\}$. We shall express $\mathbb{S}_J$ as the union $\mathbb{S}_J=\bigcup_m\mathbb{S}_m$ of $\Lambda^{(j)}$-trees $\mathbb{S}_m$ each of whose tree top is a tile in $\mathbb{T}$. Let $\big\{I_{\overline{Q}_m}\big\}$ be the maximal time intervals in $\big\{I_Q: Q\in\mathbb{S}_J\big\}$ and set

$$\mathbb{S}_m \;=\; \big\{Q\in\mathbb{S}_J:\, I_Q\;\subseteq\; I_{\overline{Q}_m}\big\}.$$

In view of (4.10) and (4.15), each $\mathbb{S}_m$ is a $\Lambda^{(j)}$-tree for which

$$\int_{I_{\overline{Q}_m}} \Big(\sum_{Q\,\in\,\mathbb{S}_m} \frac{1}{|I_Q|}\, |\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x) \Big)^{1/2} dx \;\le\; \mathrm{const}_\phi (s^{-(\nu+1)/q_0}\,\kappa_q)\, |I_{\overline{Q}_m}|$$

since the $\mathbb{S}_m$ were not chosen prior to the ν^{th}-stage. But by maximality, the $\mathbb{S}_m$ are disjoint and their time intervals $I_{\overline{Q}_m}$ are *disjoint* in J; in addition, the union of the $\mathbb{S}_m$ is all of $\mathbb{S}_J$. The estimate in (4.16) thus follows after summing over m. □

We end this section with comments relating *trees* to *tent spaces*. Given a dyadic interval $I = [\ell_I, r_I)$ in $\mathbb{R}$, denote by Δ_I the usual dyadic square of side length $|I|$ sitting above I in the upper half plane and having $(\ell_I, |I|)$ as its lower left hand corner. Associated with I is a *Carleson box* $C(I) = \{(v,t) : v \in I,\, 0 \,<\, t \,\le\, 2|I|\,\}$ consisting of the points in Δ_{I_0}, $I_0 \,\subseteq\, I$. But by (4.1) the mapping $Q \longrightarrow \Delta_{I_Q}$ is 1-1 on a tree, so the mapping associates a subset of $C(I_{\mathbb{T}})$ to any tree $\mathbb{T}$. Now let $\chi_Q = \chi_Q(z)$, $z = (v,t)$, be the characteristic function of Δ_{I_Q}. Then for any function h and 'mother' wave function ϕ,

$$H(z) \;=\; \sum_{Q\in\mathbb{T}} \langle\, h,\, \phi_Q\,\rangle\, \chi_Q(z), \qquad (z = (v,t)), \tag{4.19}$$

defines a function in the upper half plane having support in the Carleson box $C(I_{\mathbb{T}})$. In view of (2.6), the non-tangential maximal function of H satisfies the inequality

$$\sup_{z\in\Gamma_x} \Big(\frac{1}{t^{1/2}} |H(z)| \Big) \;\le\; \mathrm{const.}\ M(h)(x) \tag{4.20}$$

(note that $(v,t) \in \Delta_Q \implies t \sim |I_Q|$), while the Lusin area type function is given by

$$A_x(H) \;=\; \Big(\int_{\Gamma_x} |H(v,t)|^2\, \frac{dvdt}{t^2} \Big)^{1/2} \;=\; \Big(\sum_{Q\in\mathbb{T}} \frac{1}{|I_Q|}\, |\langle\, h,\, \phi_Q\,\rangle|^2\, \chi_{I_Q}(x) \Big)^{1/2}.$$

Thus, as a function in the Tent space $\mathfrak{N}^p$, the function H has norm

$$\|H\|_{\mathfrak{N}^p} \;=\; \Big(\int_{I_{\mathbb{T}}} \Big[\sum_{Q\in\mathbb{T}} \frac{1}{|I_Q|}\, |\langle\, h,\, \phi_Q\,\rangle|^2\, \chi_{I_Q}(x) \Big]^{p/2} dx \Big)^{1/p} \tag{4.21}$$

where $\chi_{I_Q} = \chi_{I_Q}(x)$ is the characteristic function of I_Q. Clearly

$$\Big(\frac{1}{|I_{\mathbb{T}}|} \Big)^{1/p} \|H\|_{\mathfrak{N}^p} \;\le\; \Big(\frac{1}{|I_{\mathbb{T}}|} \Big)^{1/q} \|H\|_{\mathfrak{N}^q}, \qquad (p \le q). \tag{4.22}$$

Now let's specialize to the fixed functions f, g and corresponding functions

$$F(z) \;=\; \sum_{Q\in\mathbb{T}} \langle f,\, \phi_Q^{(1)}\,\rangle\, \chi_Q(z), \qquad G(z) \;=\; \sum_{Q\in\mathbb{T}} \langle g,\, \phi_Q^{(2)}\,\rangle\, \chi_Q(z). \tag{4.23}$$

associated to a tree $\mathbb{T}$. Property (i) of (4.15) ensures that the non-tangential maximal functions of F and G satisfy

$$\text{(4.24)(i)} \qquad \sup_{z\in\Gamma_x}\left(\frac{1}{t^{1/2}}|F(z)|\right) \;\leq\; \text{const}_\phi\,(s^{-(\nu+1)(1+\eta)/p_0}\,s^{-1/\eta}\,\kappa_p)\,\chi_{I_{4\mathbb{T}}}(x)$$

and

$$\text{(4.24)(ii)} \qquad \sup_{z\in\Gamma_x}\left(\frac{1}{t^{1/2}}|G(z)|\right) \;\leq\; \text{const}_\phi\,(s^{-(\nu+1)(1+\eta)/q_0}\,s^{-1/\eta}\,\kappa_q)\,\chi_{I_{4\mathbb{T}}}(x)$$

respectively for *any* tree $\mathbb{T}$ in $\mathbb{Q}_\nu$, while (4.15)(iii) says that

$$\text{(4.25)} \qquad \frac{1}{|I_{\mathbb{T}}|}\,\|G\|_{\mathfrak{N}^1} \;\leq\; \text{const}_\phi\,s^{-(\nu+1)/q_0}\,\kappa_q$$

for any $\Lambda^{(j)}$-tree in $\mathbb{Q}_\nu$, $j\neq 2$. There is an analogous interpretation of (4.15)(ii) for F.

5. Tree Estimates

In this section we establish an L^2-norm estimate for trees. More precisely, given $\mathbb{T}$ in $\mathcal{F}_\nu$, let $\mathbb{S}\subseteq\mathbb{T}$ be as in (4.18) and set

$$\text{(5.1)} \qquad \mathcal{D}_{\mathbb{S}}(f,g) \;=\; \sum_{Q\in\mathbb{S}} c_Q\frac{1}{\sqrt{|I_Q|}}\langle\, f,\,\phi_Q^{(1)}\,\rangle\langle\, g,\,\phi_Q^{(2)}\,\rangle\,\phi_Q^{(3)}.$$

Again in this section the functions f and g are those fixed earlier in section 2. The first step consists in establishing a Carleson measure type estimate improving on (4.25). Set

$$\text{(5.2)} \qquad F(z) \;=\; \sum_{Q\in\mathbb{S}}\big\langle f,\,\phi_Q^{(1)}\,\big\rangle\,\chi_Q(z), \qquad G(z) \;=\; \sum_{Q\in\mathbb{S}}\big\langle g,\,\phi_Q^{(2)}\,\big\rangle\,\chi_Q(z).$$

When $\mathbb{S}=\mathbb{T}$ these correspond to the functions defined in (4.23).

(5.3) Proposition. *Let $\mathbb{T}$ be a tree in $\mathcal{F}_\nu^{(2j)}$, $j\neq 2$, and $\mathbb{S}\subseteq\mathbb{T}$ as above. Then the function $G=G(z)$ defined by* (5.2) *belongs to the Tent space $\mathfrak{N}^\infty$; more precisely, the inequality*

$$\|G\|_{\mathfrak{N}^\infty} \;\leq\; \text{const. }s^{-\nu/q_0}\kappa_q$$

holds uniformly in $\mathbb{S}$, $\mathbb{T}$ and ν.

When $\mathbb{T}$ belongs to $\mathcal{F}_\nu^{(1j)}$, $j\neq 1$, there is a corresponding estimate

$$\text{(5.3)(i)} \qquad \|F\|_{\mathfrak{N}^\infty} \;\leq\; \text{const. }s^{-\nu/p_0}\kappa_p$$

for the function $F=F(z)$ defined by (5.2). The proof again differs only in notation, so details are left to the reader.

Proof of (5.3). Since

$$\|G\|_{\mathfrak{N}^\infty} \;=\; \sup_{J\,\subseteq\, I_{\mathbb{S}}} \left(\frac{1}{|J|}\int_J \Big\{ \sum_{I_Q\,\subseteq\, J} \frac{1}{|I_Q|}\,|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x)\Big\}\,dx\right)^{1/2},$$

it is enough to show that the L^1-norm in (4.15) can be replaced with the L^2-norm, at the expense possibly of introducing a extra constant factor in the right hand side. Fix an s-adic interval $J\subseteq I_{\mathbb{S}}$ and define an $\ell^2(\mathbb{S})$-valued function on J by

$$G_{\mathbb{S}}(x) \;=\; \Big\{ \frac{1}{\sqrt{|I_Q|}}\,\langle\, g,\, \phi_Q^{(2)}\,\rangle\,\chi_{I_Q}(x)\Big\}_{Q\in\mathbb{S}} \qquad (x\in J).$$

Then

$$\left(\frac{1}{|J|}\int_J \Big\{ \sum_{I_Q\,\subseteq\, J} \frac{1}{|I_Q|}\,|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x)\Big\}\,dx\right)^{1/2}$$

$$= \left(\frac{1}{|J|}\int_J \|G_{\mathbb{S}}(x)\|^2_{\ell^2}\,dx\right)^{1/2} \;\le\; \text{const.}\ \|G_{\mathbb{S}}\|_{BMO(J)}$$

where $BMO(J)$ is understood with respect to the s-grid. But because of grid structure,

$$\frac{1}{|J_0|}\int_{J_0} \|G_{\mathbb{S}}(y) - \frac{1}{|J_0|}\int_{J_0} G_{\mathbb{S}}\|_{\ell^2}\,dy$$

$$\le\; \frac{2}{|J_0|}\int_{J_0} \Big(\sum_{I_Q\,\subseteq\, J_0} \frac{1}{|I_Q|}\,|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x)\Big)^{1/2} dx$$

for every s-adic interval $J_0\subseteq J$. Consequently, by (4.16),

$$\|G_{\mathbb{S}}\|_{BMO(J)} \;\le\; \text{const.}\,(\text{const}_\phi\,\kappa_p)s^{-\nu/q_0}.$$

Together these estimates prove the proposition. $\square$

(5.4) Remark. For applications of Main Theorem II it is important to note that the constant in (5.2) and (5.3) has the form $C\|\phi^{(2)}\|$ where C depends on the constant in the John-Nirenberg inequality but not on the $\phi^{(i)}$.

(5.5) Theorem. *Let $\mathbb{T}$ be any tree in $\mathcal{F}_\nu$ and $\mathbb{S}\subseteq\mathbb{T}$ as above. Then*

$$\left(\frac{1}{\gamma^2}\int_{-\infty}^{\infty} |\mathcal{D}_{\mathbb{S}}(f,\, g)(x)|^2\,dx\right)^{1/2} \;\le\; \text{const.}(s^{-\nu/r_0})\,|I_{\mathbb{S}}|^{1/2}$$

uniformly in $\mathbb{S}$, $\mathbb{T}$ and ν with constant depending on r_0, κ_p, κ_q, const_ϕ and η.

(5.6) Remark. The reason for introducing $\mathbb{S}$ in (5.3) and (5.5) instead of working solely with $\mathbb{T}$ is to establish (5.5) for any subtree of $\mathbb{T}$, not just for a full tree $\mathbb{T}$ (*cf.* (4.18)). The usefulness of this will become clear in sections 8 and 9.

Proof. Fix h in $L^2(\mathbb{R})$. It is enough to establish the inequality

$$(5.7)\qquad \Big|\frac{1}{\gamma}\int_{-\infty}^{\infty} \mathcal{D}_{\mathbb{S}}(f,\,g)(x)\,\overline{h(x)}\,dx\Big| \;\le\; \text{const.}(s^{-\nu/r_0})|I_{\mathbb{S}}|^{1/2}\|h\|_2\ .$$

Let

$$H(z) \;=\; \sum_{Q\in\mathbb{S}} \langle\, h,\, \phi_Q^{(3)}\,\rangle\,\chi_{Q}(z), \qquad (z=(v,t))$$

be the function on the upper half plane associated with h (*cf.* (4.20), (5.2)). Then

$$(5.8)(\text{i})\qquad \Big|\int_{-\infty}^{\infty} \mathcal{D}_{\mathbb{S}}(f,\,g)(x)\,\overline{h(x)}\,dx\Big| \;\le\; \int_{-\infty}^{\infty}\int_0^{\infty} |\mathcal{F}(z)G(z)H(z)|\,\frac{dv\,dt}{t^{5/2}}$$

while

$$(5.8)(\text{ii})\qquad \Big|\int_{-\infty}^{\infty} \mathcal{D}_{\mathbb{S}}(f,\,g)(x)\,\overline{h(x)}\,dx\Big| \;\le\; \int_{-\infty}^{\infty}\int_0^{\infty} |F(z)\mathcal{G}(z)H(z)|\,\frac{dv\,dt}{t^{5/2}}$$

where $F(z)$ and $G(z)$ are defined by (5.2) but

$$(5.9)\qquad \mathcal{F}(z) \;=\; \sum_{Q\in\mathbb{S}} c_Q\langle\, f,\, \phi_Q^{(1)}\,\rangle\,\chi_{Q}(z), \qquad \mathcal{G}(z) \;=\; \sum_{Q\in\mathbb{S}} c_Q\langle\, g,\, \phi_Q^{(2)}\,\rangle\chi_{Q}(z).$$

Note that $\mathcal{F}$, $\mathcal{G}$ have support in the Carleson box $C(I_{\mathbb{S}}) = \{(v,t) : v\in I_{\mathbb{S}},\, 0<t\le 2|I_{\mathbb{S}}|\}$. The proof now breaks into different cases according to which family $\mathcal{F}_\nu^{(ij)}$ the tree $\mathbb{T}$ belongs.

(a) $i=j=2$. By Cauchy-Schwarz, the integral in the right hand side of (5.8)(ii) is dominated by

$$\Big(\int_{-\infty}^{\infty}\int_0^{\infty} \frac{1}{t}|F(z)\mathcal{G}(z)|^2\,\frac{dv\,dt}{t^2}\Big)^{1/2}\Big(\int_{-\infty}^{\infty}\int_0^{\infty} |H(z)|^2\,\frac{dv\,dt}{t^2}\Big)^{1/2}.$$

Since $\mathbb{S}$ is a $\Lambda^{(2)}$-tree, (5.3)(i) applies to F; on the other hand, by (4.24)(ii),

$$\sup_{z\in\Gamma_x}\Big(\frac{1}{t^{1/2}}\,|\mathcal{G}(z)|\Big) \;\le\; \text{const}_\phi\, s^{-(\nu+1)(1+\eta)/q_0}\,\kappa_q\,\chi_{4I_{\mathbb{S}}}(x).$$

Consequently,

$$\Big(\int_{-\infty}^{\infty}\int_0^{\infty} \Big|\frac{1}{t}F(z)\mathcal{G}(z)|^2\,\frac{dv\,dt}{t^2}\Big)^{1/2} \;\le\; \text{const.}\,s^{-\nu/r_0}|I_{\mathbb{S}}|^{1/2}.$$

But

$$(5.10)\qquad \Big(\int_{-\infty}^{\infty}\int_0^{\infty} |H(z)|^2\,\frac{dv\,dt}{t^2}\Big)^{1/2} \;\le\; \text{const.}\,\|h\|_2$$

by the Bessel inequality for a single tree. Inequality (5.5) thus follows.

(b) $i = 2$, $j = 3$. By Cauchy-Schwarz, the integral on the right hand side of (5.8)(i) is dominated by

$$\left(\int_{-\infty}^{\infty}\int_0^{\infty} |\mathcal{F}(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2}\left(\int_{-\infty}^{\infty}\int_0^{\infty} \frac{1}{t}|G(z)H(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2}.$$

Since $\mathbb{S}$ is now a $\Lambda^{(3)}$-tree, (5.3)(i) applies to G and (4.20) to H. Consequently,

$$\left(\int_{-\infty}^{\infty}\int_0^{\infty} \frac{1}{t}|G(z)H(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2} \le \text{ const.}\,\|G\|_{\mathfrak{N}^\infty}\,\|P^*\|_{L^1}^{1/2} \le s^{-\nu/q_0}\,|I_{\mathbb{S}}|^{1/2}\,\|h\|_2$$

where

$$\|P^*\|_{L^1} = \int_{I_{\mathbb{S}}}\left(\sup_{z\,\in\,\Gamma_x} \frac{|H(z)|^2}{t}\right) dx,$$

establishing (5.5) once again.

(c) $i = 2, j = 1$. This time Cauchy-Schwarz gives

$$\left(\int_{-\infty}^{\infty}\int_0^{\infty} \frac{1}{t}|\mathcal{F}(z)G(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2}\left(\int_{-\infty}^{\infty}\int_0^{\infty} |H(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2}$$

with $\mathbb{S}$ a $\Lambda^{(1)}$-tree. But now (5.3) applies to G, while

$$\sup_{z\in\Gamma_x}\left(\frac{1}{t^{1/2}}|\mathcal{F}(z)|\right) \le \text{ const}_\phi\, s^{-(\nu+1)(1+\eta)/p_0}\,\kappa_p\,\chi_{4I_{\mathbb{S}}}(x)$$

by (4.24)(i). So

$$\left(\int_{-\infty}^{\infty}\int_0^{\infty} \frac{1}{t}|\mathcal{F}(z)G(z)|^2 \frac{dv\,dt}{t^2}\right)^{1/2} \le \text{const}\, s^{-\nu/r_0}|I_{\mathbb{S}}|^{1/2}.$$

Since (5.10) holds, estimate (5.5) follows once more.

As the proof for $i = 1$ will be the same, interchanging the roles of F and G, all possible cases have now been covered, completing the proof. □

6. Counting function estimate $(i = j)$

In this section the first of the counting function estimates is established. We will also impose restrictions on the as yet unspecified $\eta > 0$ (*cf.* section 2). Firstly, in all that follows it will be assumed that $1/\eta$ is a positive *integer*.

(6.1) Theorem. *The function*

$$N_{\mathcal{F}_\nu^{(ii)}}(x) = \sum_{\mathbb{T}\,\in\,\mathcal{F}_\nu^{(ii)}} \chi_{I_{\mathbb{T}}}(x)$$

counting the number of trees in $\mathcal{F}_\nu^{(ii)}$ *above* x *satisfies the inequality*

$$\Big(\int_{-\infty}^{\infty}\Big(\min\{N_{\mathcal{F}_\nu^{(ii)}}(x), s^{\nu/\eta}\}\Big)^{\sigma}\,dx\Big)^{1/\sigma} \;\le\; \text{const.}\, s^{(1+2\delta)\nu}\Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/\sigma}$$

for each σ*,* $1 \le \sigma < \infty$*, provided* $1/\eta > \max\{4/\delta,\, 2/r_0,\, p_0,\, q_0\}$.

This will be achieved by thinking of the tree tops of the trees in $\mathcal{F}_\nu^{(ii)}$ as a family of mutually disjoint tiles. Once the corresponding $i \ne j$ version of (6.1) has been established, inequality (2.13) follows immediately since

$$\frac{2}{r_0} - (1+2\delta) \;\ge\; 2\,(\tfrac{1}{2}+2\delta) - 1 - 2\delta \;=\; 2\delta \;>\; 0.$$

(6.2) Remark. In proving (6.1) and its analogue (7.1) for $i \ne j$, it will be important to truncate $\mathcal{F}_\nu^{(ij)}$, replacing it with a subfamily $\overline{\mathcal{F}}_\nu^{(ij)}$ of trees if necessary so that

$$\|N_{\overline{\mathcal{F}}_\nu^{(ij)}}\|_\infty \;\le\; s^{\nu/\eta}, \qquad N_{\overline{\mathcal{F}}_\nu^{(ij)}}(x) \;=\; \min\{N_{\mathcal{F}_\nu^{(ij)}}(x),\; s^{\nu/\eta}\}.$$

The grid structure in time ensures that such a family exists (*cf.* [12, p.711]). Indeed, if $\overline{\mathcal{F}}_\nu^{(ij)}$ is a minimal subset of $\mathcal{F}_\nu^{(ij)}$ for which

$$N_{\overline{\mathcal{F}}_\nu^{(ij)}}(x) \;\ge\; \min\{N_{\mathcal{F}_\nu^{(ij)}}(x), s^{\nu/\eta}\},$$

then the equality

$$N_{\overline{\mathcal{F}}_\nu^{(ij)}}(x) \;=\; \min\{N_{\mathcal{F}_\nu^{(ij)}}(x),\; s^{\nu/\eta}\}$$

holds automatically.

To prove (6.1) it is enough, therefore, to show that

$$\|N_{\overline{\mathcal{F}}_\nu^{(ii)}}\|_\sigma \;\le\; \text{const.}\, s^{(1+2\delta)\nu}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^{r/\sigma}. \tag{6.3}$$

In fact, we prove a generic result containing (6.3). Let $\mathcal{P}_\nu$ be a family of mutually disjoint tiles such that $\|N_{\mathcal{P}_\nu}\|_\infty \le \text{const.}\, s^{\nu/\eta}$. Suppose further that h is a fixed function in $L^\gamma(\mathbb{R})$, $1 < \gamma < \infty$, and ϕ a wave function for which the inequalities

$$\text{const}\,\kappa_\gamma b \;\le\; \frac{1}{\sqrt{|I_P|}}|\langle h, \phi_P\rangle|, \qquad \inf_{x\in I_P} M_s(M(h))(x) \;\le\; s^{-1/\eta}\kappa_\gamma \tag{6.4}$$

hold for all $P \in \mathcal{P}_\nu$ where b and κ_γ are given constants.

(6.5) Theorem. *Let h be a function in $L^\gamma(\mathbb{R})$, $1<\gamma<\infty$, satisfying* (6.4) *with respect to ϕ and $\mathcal{P}_\nu$. Then the counting function $N_{\mathcal{P}_\nu}$ satisfies the inequality*

$$\left(\int_{-\infty}^{\infty} N_{\mathcal{P}_\nu}(x)^\sigma\right)^{1/\sigma} \le \text{const}\left(\frac{1}{b}\right)^{(\gamma_0+\delta_0)\theta}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^{\gamma/\sigma}$$

uniformly in h, b, and $\mathcal{P}_\nu$ for each $1\le\sigma<\infty$ and $\gamma_0=\max(\gamma,\gamma')$ provided $0<\delta_0\le\delta/2$ and $\theta=(1+2\delta)/(1+\delta)$.

To deduce (6.3) from (6.5) in the case $i=1$, say, let $\mathcal{P}_\nu$ be the tree tops of the trees in the family $\overline{\mathcal{F}}_\nu^{(11)}$ derived from $\mathcal{F}_\nu^{(11)}$ in (6.2). These will be mutually disjoint because the trees in $\mathcal{F}_\nu^{(11)}$ were constructed maximally. Now set

$$\gamma\;=\;p,\qquad h\;=\;\pi(a_1)f,\qquad \phi_P\;=\;\phi_P^{(1)},\qquad b\;=\;s^{-(1+\eta)\nu/p_0}.$$

Then

$$(\frac{1}{b})^{\theta(\gamma_0+\delta_0)}\;=\;(s^\nu)^{(1+\eta)\theta(1+\delta_0/p_0)}\;<\;s^{(1+2\delta)\nu}$$

provided $\eta<\delta/4$ since $\delta>0$ is very small, $p_0\ge 2$, $\theta=1+\delta/(1+\delta)$ and $\delta_0\le\delta/2$. When $i=2$ set

$$\gamma\;=\;q,\quad h\;=\;\pi(a_2)g,\quad \phi_p\;=\;\phi_p^{(2)},\quad b\;=\;s^{-(1+\eta)\nu/q_0}$$

and proceed as before.

Before beginning the proof of (6.5) let h be for the moment an arbitrary function in $L^2(\mathbb{R})$, condition (6.4) not being imposed yet. As no use of vanishing moments can be made in (6.5), some separation among tiles has to be made in order to be able to exploit decay.

(6.6) Definition. *Given an integer $A\ge 2$, a family $\mathcal{Q}$ of tiles is said to be A-separated when either $w_P\cap w_Q=\emptyset$ or $AI_P\cap AI_Q=\emptyset$ for each pair P,Q in $\mathcal{Q}$.*

In Appendix B we will show that the Bessel inequality

$$\text{(6.7)}\qquad \sum_{Q\in\mathcal{Q}}|\langle\, h,\,\phi_Q\,\rangle|^2\,dx\;\le\;\text{const}\left(1+\frac{\|N_{\mathcal{Q}}\|_\infty}{A^\mu}\right)\int_{-\infty}^{\infty}|h(x)|^2\,dx$$

holds uniformly in h and $\mathcal{Q}$ for any A-separated family $\mathcal{Q}$ whenever ϕ is an $\mathcal{M}_\mu$-test function. Here $N_{\mathcal{Q}}(x)$ is the function counting the number of tiles in $\mathcal{Q}$ above x.

Proof of (6.5). Fix $\lambda\ge 1$. By arguing as in (6.2) we can replace $\mathcal{P}_\nu$ with a subfamily $\overline{\mathcal{P}}_\nu$ such that

$$\text{(6.8)}\qquad \|N_{\overline{\mathcal{P}}_\nu}\|_{L^\infty}\;\le\;\lambda,\qquad \{x: N_{\overline{\mathcal{P}}_\nu}(x)\;\ge\;\lambda\}=\{x:N_{\mathcal{P}_\nu}(x)\ge\lambda\}.$$

In turn $\overline{\mathcal{P}}_\nu$ itself can be decomposed into a finite number of subfamilies

$$\text{(6.9)(i)}\qquad \overline{\mathcal{P}}_\nu=\left(\bigcup_{m=1}^{A^4}\mathcal{P}_m\right)\bigcup\mathcal{P}'_\nu$$

where each subfamily $\mathcal{P}_m$ is A-separated and the time intervals of the tiles in $\mathcal{P}'_\nu$ are controlled by

(6.9)(ii) $$\sum_{\mathcal{P}\in\mathcal{P}'_\nu} |I_P| \le \text{const}\, e^{-\,\text{const}\, A} \sum_{P\in\mathcal{P}_1} |I_P|$$

with constants uniform in A (*cf.* [12, section 4.3]).

Assume first that $\gamma = 2$ and h is an arbitrary function in $L^2(\mathbb{R})$. Then (6.7) applies to each $\mathcal{P}_m$:

$$\int_{-\infty}^{\infty}\Big(\sum_{\mathcal{P}\in\mathcal{P}_m} \frac{1}{|I_P|}|\langle\, h,\, \phi_P\,\rangle|^2 \chi_{I_P}(x)\Big)\, dx \ \le\ \text{const.}\Big(1 + \frac{\|N_{\mathcal{P}_m}\|_{L^\infty}}{A^\mu}\Big)\int_{-\infty}^{\infty} |h(x)|^2\, dx$$

where $N_{\mathcal{P}_m}$ is the counting function for $\mathcal{P}_m$. As $\mathcal{P}_m$ is a subfamily of $\overline{\mathcal{P}}_\nu$ the L^∞-norm of $N_{\mathcal{P}_m}$ is controlled by (6.8). On the other hand, since A can be specified arbitrarily we choose $A \sim \max\{2,\, [\lambda^{1/\mu}]\}$. In this case

(6.10) $$\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m} \frac{1}{|I_P|}|\langle\, h,\, \phi_P\,\rangle|^2 \chi_{I_P}(x)\Big)\, dx \ \le\ \text{const.}\int_{-\infty}^{\infty} |h(x)|^2\, dx.$$

To interpret (6.10) in a form suitable for interpolation, set

$$H_{\mathcal{P}_m}(x) = \Big\{\frac{1}{|I_P|^{1/2}}\langle\, h,\, \phi_P\,\rangle \chi_{I_P}(x)\Big\}_{P\in\mathcal{P}_m}.$$

Inequality (6.10) ensures that $h \to H_{\mathcal{P}_m}$ is bounded from $L^2(\mathbb{R})$ into $L^2(\mathbb{R}, \ell^2)$, while (6.4) ensures that $h \to H_{\mathcal{P}_m}$ extends to a bounded operator from $L^\rho(\mathbb{R})$ into $L^\rho(\mathbb{R}, \ell^\infty)$ for any ρ, $1 < \rho \le \infty$. At this point the path splits according as $\gamma < 2$ and $\gamma \ge 2$. Suppose first that $1 < \gamma < 2$. Then by interpolation $h \to H_{\mathcal{P}_m}$ is bounded from $L^\gamma(\mathbb{R})$ into $L^\gamma(\mathbb{R}, \ell^{\gamma'+\delta_0}(\mathcal{P}_m))$ for any $\delta_0 > 0$. 'Localizing' this as in Appendix A we thus obtain the inequality

$$\Big(\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m,\, I_P\subseteq J}\Big(\frac{1}{\sqrt{|I_P|}}|\langle\, h,\, \phi_P\,\rangle|\Big)^{\gamma'+\delta_0}\chi_{I_P}(y)\Big)^{\gamma/\gamma'+\delta_0} dy\Big)^{1/\gamma}$$
$$\le\ \text{const}\ \lambda^{1/\mu}\Big(\inf_{x\in J} M_\gamma(M(h))(x)\Big)$$

which holds uniformly in h, J and $\mathcal{P}_m$ for all J in $\mathcal{I}_s$. On the other hand, if $2 \le \gamma < \infty$ interpolation shows that $h \to H$ is bounded from $L^\gamma(\mathbb{R})$ into $L^\gamma(\mathbb{R}, \ell^\gamma)$ resulting in a localized estimate of the form

$$\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m,\, I_P\subseteq J}\Big(\frac{1}{\sqrt{|I_P|}}|\langle\, h,\, \phi_P\,\rangle|\Big)^{\gamma}\chi_{I_P}(y)\Big)^{1/\gamma} dy$$
$$\le\ \text{const}\ \lambda^{1/\mu}\Big(\inf_{x\in J} M_\gamma(Mh)(x)\Big).$$

For an h in $L^2(\mathbb{R})$ satisfying (6.4), therefore, these inequalities become

(6.11) $$\Big(\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m,\, I_P\subseteq J}\chi_{I_P}(y)\Big)^{\gamma/\gamma'+\delta_0} dy\Big)^{1/\gamma}$$
$$\le\ \text{const}\ \frac{\lambda^{1/\mu}}{\kappa_\gamma b}\Big(\min\Big\{\inf_{x\in J} M_\gamma(Mh)(x),\ \kappa_\gamma\Big\}\Big)$$

when $1 < \gamma < 2$, and

$$\begin{aligned}(6.12)\quad &\Big(\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m,\, I_P\subseteq J}\chi_{I_P}(y)\Big)dy\Big)^{1/\gamma}\\ &\qquad\qquad \le\ \text{const}\ \frac{\lambda^{1/\mu}}{\kappa_\gamma b}\Big(\min\Big\{\inf_{x\in J} M_\gamma(M(h))(x),\ \kappa_\gamma\Big\}\Big)\end{aligned}$$

when $2 \le \gamma < \infty$.

To utilize these estimates we shall think of the counting function for $\mathcal{P}_m$ temporarily as an ℓ^r-sequence valued function, $r \ge 1$,

$$N_{\mathcal{P}_m}(x)\ =\ \{\chi_{I_P}(x)\}_{P\in\mathcal{P}_m}$$

on $\mathbb{R}$ and denote by $N^{\#}_{\mathcal{P}_m}$ its s-adic sharp function

$$N^{\#}_{\mathcal{P}_m}(x)\ =\ \sup_{J\ni x}\Big(\frac{1}{|J|}\int_J\Big\|N_{\mathcal{P}_m}(y)-\frac{1}{|J|}\int_J N_{\mathcal{P}_m}\Big\|_{\ell^r}dy\Big).$$

Because of the grid structure on $\mathcal{I}_s$,

$$N^{\#}_{\mathcal{P}_m}(x)\ \le\ 2\sup_{j\ni x}\Big(\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{I_P\subseteq J}\chi_{I_P}(y)\Big)^{1/r}dy\Big).$$

Consequently,

$$\int_{-\infty}^{\infty}\Big(\sum_{P\in\mathcal{P}_m}\chi_{I_P}(x)\Big)^{t/r}dx\ \le\ \text{const.}\int_{-\infty}^{\infty}\Big[\sup_{J\ni x}\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{I_P\subseteq J}\chi_{I_P}(y)\Big)^{1/r}dy\Big]^t dx$$

since passage to the sharp function is bounded on $L^t(\mathbb{R})$ for any $t > 1$. In turn this guarantees that

$$\int_{-\infty}^{\infty}N_{\mathcal{P}_m}(x)^{t/r}\,dx \le \text{const.}\int_{-\infty}^{\infty}\Big[\sup_{J\ni x}\frac{1}{|J|}\int_{-\infty}^{\infty}\Big(\sum_{I_P\subseteq J}\chi_{I_P}(y)\Big)^{1/r}dy\Big]^t dx,$$

thinking now of the counting function as $N_{\mathcal{P}_m} = \sum_{P\in\mathcal{P}_m}\chi_{I_P}$.

We combine the discussion above with estimates (6.11) and (6.12). Fix $\rho > 1$, $1 < \gamma < 2$ and set $t = \rho r$ with $1/r = \gamma/(\gamma' + \delta_0)$ so that $\gamma t = \rho(\gamma' + \delta_0)$. Then

$$\begin{aligned}&\int_{-\infty}^{\infty}N_{\mathcal{P}_m}(x)^{\rho}\,dx\\ &\quad\le\ \text{const.}\Big(\frac{\lambda^{1/\mu}}{\kappa_\gamma b}\Big)^{\rho(\gamma'+\delta_0)}\int_{-\infty}^{\infty}\Big(\min\Big\{\inf_{x\in J}M_\gamma(M(h))(x),\kappa_\gamma\Big\}\Big)^{\rho(\gamma'+\delta_0)}dx\,.\\ &\quad=\ \text{const}\Big(\frac{\lambda^{1/\mu}}{\kappa_\gamma b}\Big)^{\rho(\gamma'+\delta_0)}\int_0^{\kappa_\gamma}\tau^{\rho(\gamma'+\delta_0)}|\{x : M_\gamma(M(h))(x)\ge\tau\}|\frac{d\tau}{\tau}\end{aligned}$$

Consequently, when $1 < \gamma < 2$, we see that

$$\int_{-\infty}^{\infty} N_{\mathcal{P}_m}(x)^\rho \, dx \ \le \ \text{const} \Big(\frac{\lambda^{1/\mu}}{b}\Big)^{\rho(\gamma'+\delta_0)} \Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^\gamma \tag{6.13}$$

since $\rho(\gamma' + \delta_0) - \gamma > 0$. The corresponding estimate for $2 \le \gamma < \infty$ is

$$\int_{-\infty}^{\infty} N_{\mathcal{P}_m}(x)^\rho \, dx \ \le \ \text{const} \Big(\frac{\lambda^{1/\mu}}{b}\Big)^{\rho\gamma} \Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^\gamma, \tag{6.14}$$

proceeding as before with $t = \rho$, $r = 1$ and $\gamma t = \rho\gamma > \gamma$.

The final step in the proof of (6.5) uses (6.8) and (6.9) to extend estimates for the individual $N_{\mathcal{P}_m}$ to all $N_{\mathcal{P}_\nu}$. Since counting functions are integer-valued, it is enough to establish the weak estimate

$$\lambda^{\sigma+\varepsilon_0} |\{x : N_{\overline{\mathcal{P}}_\nu}(x) \ \ge \ \lambda\}| \ \le \ \text{const} \Big(\frac{1}{b}\Big)^{\theta(\gamma_0+\delta_0)\sigma} \Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^\gamma, \qquad (\lambda > 0) \tag{6.15}$$

for some fixed $\varepsilon_0 > 0$. But by (6.9)

$$|\{x : N_{\overline{\mathcal{P}}_\nu}(x) \ \ge \ \lambda\}| \ \le \ \sum_{m=1}^{A^4} \Big|\Big\{x : N_{\mathcal{P}_m}(x) \ \ge \ \frac{\lambda}{A^4}\Big\}\Big| + |\{x : N_{\mathcal{P}'_\nu}(x) \ \ge \ 1\}|.$$

Suppose first that $1 < \gamma < 2$; then $\gamma' = \gamma_0$. In view of (6.13), therefore,

$$\sum_{m=1}^{A^4} \Big|\Big\{x : N_{\mathcal{P}_m}(x) \ge \frac{\lambda}{A^4}\Big\}\Big| \ \le \ \text{const} \frac{\lambda^{4(1+\rho)/\mu}}{\lambda^{\rho(1-(\gamma_0+\delta_0)/\mu)}} \Big(\frac{1}{b}\Big)^{\rho(\gamma_0+\delta_0)} \Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^\gamma \tag{6.16}$$

since $A \sim \lambda^{1/\mu}$.

To check that the right hand side has the right decay in λ we give specific values to ρ, δ_0, ε_0 and μ; namely, set

$$0 \ < \delta_0 \ \le \ \delta/2, \quad \rho \ = \ \Big(\frac{1+2\delta}{1+\delta}\Big)\sigma, \quad \varepsilon_0 \ \le \ \frac{\delta}{\mu}, \quad 0 < \frac{1}{\mu} < \frac{1}{50}\frac{\delta}{\gamma_0}. \tag{6.17}$$

Then

$$\rho\Big(1 + \frac{1}{\mu}(\gamma_0 + \delta_0) - \frac{4}{\mu}\Big) - \frac{4}{\mu} \ > \ \sigma + \varepsilon_0\,,$$

as required. On the other hand,

$$|\{x : N_{\mathcal{P}'_\nu}(x) \ge 1\}| \le \text{const}(\lambda^{1/\mu})^{\rho(\gamma_0+\delta_0)} e^{-\text{const}\, \lambda^{1/\mu}} \Big(\frac{1}{b}\Big)^{\rho(\gamma_0+\delta_0)} \Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^\gamma,$$

while

$$\lambda^{\sigma+\varepsilon_0} (\lambda^{1/\mu})^{\rho(\gamma_0+\delta_0)} e^{-\text{const}\, \lambda^{1/\mu}} \ \le \ \text{const}(\varepsilon_0, \gamma_0, \sigma, \delta)$$

because of the choices of ε_0, ρ and μ in (6.17). This establishes (6.5) for the case $1 < s < 2$, setting $\rho = \theta\sigma$.

Entirely analogous calculations using (6.14) and the same choice of parameters (6.17) establish the case $\gamma \ge 2$. Note that in this case $\gamma_0 = \gamma$ and δ_0 doesn't appear in the corresponding calculations. This completes the proof of (6.5). □

7. Counting function estimate $(i \neq j)$

In this section we complete the proof of counting estimates, studying here the case $i \neq j$. As before $1/r = 1/p + 1/q,\ 1/r_0 = 1/p_0 + 1/q_0$ and $1/\eta > \max\{4/\delta,\, 2/r_0,\, p_0,\, q_0\}$, where again it is assumed that $1/\eta$ is an integer.

(7.1) Theorem. *The function*

$$N_{\mathcal{F}_\nu^{(ij)}}(x) = \sum_{\mathbb{T}\in\mathcal{F}_\nu^{(ij)}} \chi_{I_\mathbb{T}}(x) \qquad (i \neq j)$$

counting the number of trees in $\mathcal{F}_\nu^{(ij)}$ *above* x *satisfies the inequality*

$$\Big(\int_{-\infty}^{\infty} (\min\{N_{\mathcal{F}_\nu^{(ij)}}(x), s^{\nu/\eta}\})^\sigma\, dx\Big)^{1/\sigma} \leq \ \text{const.}\, s^{(1+2\delta)\nu}\Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/\sigma}$$

for each $\sigma,\ 1 \leq \sigma < \infty$.

Again (6.2) reduces the proof to establishing the inequality

$$\|N_{\overline{\mathcal{F}}_\nu^{(ij)}}\|_{L^\sigma} \leq \text{const}\, s^{(1+2\delta)\nu}\Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/\sigma}. \tag{7.2}$$

This will be done under the assumption $i < j$ because only simple modifications are needed when $i > j$. We indicate those changes as the proof progresses.

The proof of (7.2) proceeds in exactly the same way as for (6.5) though the individual steps are necessarily more complicated because we are dealing with trees, not individual tiles. For trees $\mathbb{T}$ and $\mathbb{T}'$ whose respective tree tops $I_\mathbb{T}\times w_\mathbb{T}$ and $I_{\mathbb{T}'}\times w_{\mathbb{T}'}$ have disjoint time intervals this is not a serious problem because decay of the wave functions can be exploited so long as the time interval I_Q of tiles $\mathbb{T}$ do not 'crowd' too close to the boundary of $I_\mathbb{T}$; in other words, the tiles in $\mathbb{T}$ need to be separated in a controlled, explicit way from tiles in T'. But if $I_\mathbb{T}$ and $I_{\mathbb{T}'}$ are not disjoint, then a tile Q in $\mathbb{T}$ need not be disjoint from $I_{\mathbb{T}'}\times w_{\mathbb{T}'}$ and hence not be disjoint from tiles in $\mathbb{T}'$. Separation can thus fail in both time and frequency when tree-tops are not disjoint.

To overcome these difficulties let $\mathcal{F}$ be a family of $\Lambda^{(j)}$ trees $\mathbb{T}$ with tree tops $I_\mathbb{T}\times w_\mathbb{T}$ such that for $A \geq 2$ the following two conditions are satisfied when $i < j$:

(7.3)(i) *The inclusion* $AI_Q \subseteq I_\mathbb{T}$ *holds for all* $Q \in \mathbb{T}$, $I_Q \neq I_\mathbb{T}$, *and each* $\mathbb{T}\in\mathcal{F}$.

(7.3)(ii) *If* $\mathbb{T}, \mathbb{T}' \in \mathcal{F}$ *are trees for which* $\lambda_\mathbb{T} < \lambda_{\mathbb{T}'}$, *then* $w_P^{(i)} \cap w_Q^{(i)} = \emptyset$ *whenever* $P \in \mathbb{T},\ Q \in \mathbb{T}'$ *and* $I_Q \subseteq I_\mathbb{T}$.

Notice that (7.3)(ii) is vacuous unless $I_{\mathbb{T}'}\cap I_\mathbb{T} \neq \emptyset$ and so (7.3)(ii) is designed to handle the difficult case of tree tops with non-intersecting time intervals and tiles Q poking down. In analogy with [4] we will refer to trees satisfying (7.3)(i) as 'normal' trees - and (7.3)(i) as the 'normality' condition - while (7.3)(ii) will be referred to as the 'no V's' condition for the family $\mathcal{F}$ of trees. The corresponding statement for (7.3)(ii) when $i > j$ is obtained by writing $\lambda_\mathbb{T} < \lambda_{\mathbb{T}'}$ and leaving the rest unchanged; in this case the tile Q will 'poke up' instead of down.

(7.4) Definition. *For a tree* $\mathbb{T}$ *we denote by* $Q_{\mathbb{T}} = I_{\mathbb{T}} \times w_{\mathbb{T}}$ *its tree-top and for a tile* Q *in* $\mathbb{T}$ *write* $Q < Q_{\mathbb{T}}$ *if* $I_Q \neq I_{\mathbb{T}}$*; i.e., if* Q *is not the tree top of* $\mathbb{T}$.

Now let $N_{\mathcal{F}} = N_{\mathcal{F}}(x)$ be the function counting the number of trees in $\mathcal{F}$ above x. For such a family the Bessel inequality

$$\sum_{\mathbb{T}\in\mathcal{F}} \Bigg(\sum_{Q\in\mathbb{T},\, Q<Q_{\mathbb{T}}} |\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \Bigg) \;\leq\; \text{const}\left(1 + \frac{\|N_{\mathcal{F}}\|_\infty}{A^\mu}\right) \int_{-\infty}^{\infty} |g(x)|^2\, dx \tag{7.5}$$

holds uniformly in g and $\mathcal{F}$. This is the analogue of (6.7) in the previous section. A proof is given in Appendix C.

Guaranteeing that conditions (7.3)(i) and (7.3)(ii) can be applied to the family $\overline{\mathcal{F}}_\nu^{(ij)}$, at least after some 'tree-trimming', requires all the hypotheses built into the Fourier support condition and into the choice of trees. The constant η plays a critical role at this point.

(7.6) Definition. *For each* $\mathbb{T}$ *in* $\overline{\mathcal{F}}_\nu^{(ij)}$, $i < j$, *set*

$$\mathbb{T}^{fat} = \{Q \in \mathbb{T} : |I_Q| \geq s^{-2\eta\nu} |I_{\mathbb{T}}|\}$$

and

$$\mathbb{T}^{edge} = \{Q \in \mathbb{T} : \text{dist}(I_Q, I_{\mathbb{T}}^c) \leq \tfrac{1}{8} s^{-\eta\nu} |I_{\mathbb{T}}|\,\};$$

in addition, let $\mathbb{T}^{down}$ *be the family of all tiles* Q *in* $\mathbb{T}$ *for which there exists* $\mathbb{T}'$ *in* $\overline{\mathcal{F}}_\nu^{(ij)}$, *with* $\lambda_{\mathbb{T}'} < \lambda_{\mathbb{T}}$, *and* $P \in \mathbb{T}'$ *such that*

$$I_Q \subseteq I_{\mathbb{T}'}, \qquad w_P^{(i)} \cap w_Q^{(i)} \neq \emptyset$$

Finally, set

$$\mathbb{T}^{trim} = \mathbb{T} \setminus (\mathbb{T}^{fat} \cup \mathbb{T}^{edge} \cup \mathbb{T}^{down})\ .$$

We shall regard $\mathbb{T}^{trim}$ itself as a tree even though for $Q \in \mathbb{T}^{trim}$, $Q < Q_{\mathbb{T}}$; *i.e.*, we still view $\mathbb{T}^{trim}$ 'attached' to $Q_{\mathbb{T}}$ even though $\mathbb{T}^{trim}$ does not contain its tree top. Removal of 'fat' tiles amounts to removing tiles from the 'top' layers of a tree. By contrast, removing the 'down' tiles eliminates ones from the 'bottom' layer as the next result makes precise.

(7.7) Proposition. *Let* $\mathbb{T}$ *be a tree in* $\overline{\mathcal{F}}_\nu^{(ij)}$, $i < j$, *and* Q' *a tile in* $\mathbb{T}^{down}$. *Then* Q' *is minimal in the sense that* $Q' \leq Q$ *for all* Q *in* $\mathbb{T}$.

Proof. Suppose to the contrary that $Q < Q'$ for some Q in $\mathbb{T}$. Then $I_Q \subset I_{Q'}$ since $Q \neq Q'$ implies that $w_{Q'} \subseteq w_Q^{(j)}$ by (1.12). On the other hand since $Q' \in \mathbb{T}^{down}$, there exists a $\Lambda^{(j)}$-tree $\mathbb{T}'$ having $\lambda_{\mathbb{T}'} < \lambda_{\mathbb{T}}$ and a tile $P \in \mathbb{T}'$ such that

$$I_{Q'} \subseteq I_{\mathbb{T}'} \;\text{ and }\; w_{Q'}^{(i)} \cap w_P^{(i)} \neq \emptyset \tag{7.8}$$

Then $I_Q \subset I_{Q'} \subseteq I_{\mathbb{T}'}$ and $\lambda_{\mathbb{T}'} < w_Q^{(j)}$ by the maximally order upwards manner in which the trees in $\mathcal{F}_\nu^{(ij)}$ were selected.

Now since $i < j$ we have that for P as in (7.8) $w_P^{(i)} < w_P^{(j)} \ni \lambda_{\mathbb{T}'}$ by (4.10) since $\mathbb{T}'$ is a $\Lambda^{(j)}$-tree. Hence we should have $w_P^{(i)} < \lambda_{\mathbb{T}'} < w_Q^{(j)}$ which in particular implies that $w_P^{(i)} < \lambda_{\mathbb{T}'} < w_{Q'}^{(i)}$. But the latter implies that $w_P^{(i)} \cap w_{Q'}^{(i)} = \emptyset$ contradicting (7.8). Therefore we must have that there is no such $Q \in \mathbb{T}$; i.e. Q' is minimal. $\square$

It is easy to see that trimming off the sets in (7.6) from the trees in $\overline{\mathcal{F}}_\nu^{(ij)}$, $i < j$, leaves a family $\{\mathbb{T}^{trim} : \mathbb{T} \in \overline{\mathcal{F}}_\nu^{(ij)}\}$ of trees satisfying (7.3). Indeed, if a tile in $\mathbb{T}$ does not belong to $\mathbb{T}^{edge} \cup \mathbb{T}^{fat}$, then

$$\mathrm{dist}(I_Q,\, I_{\mathbb{T}}^c) \;>\; \tfrac{1}{8} s^{-\eta\nu} |I_{\mathbb{T}}| \;>\; \tfrac{1}{8} s^{\eta\nu} |I_Q|,$$

so that (7.3)(i) holds for any $A \le \frac{1}{8} s^{\eta\nu}$. On the other hand, if Q does not belong to $\mathbb{T}^{down}$ and P is a tile in some $\mathbb{T}' \in \overline{\mathcal{F}}_\nu^{(ij)}$, $i < j$, for which $\lambda_{\mathbb{T}'} < \lambda_{\mathbb{T}}$, then at least one of

$$w_P^{(i)} \cap w_Q^{(i)} = \emptyset \qquad \text{or} \qquad I_Q \cap I_{\mathbb{T}'} = \emptyset$$

must be satisfied. Consequently, 'trimming' the trees in $\overline{\mathcal{F}}_\nu^{(ij)}$ ensures that the Bessel inequality

$$\sum_{\mathbb{T} \in \overline{\mathcal{F}}_\nu^{(ij)}} \Big(\sum_{Q \in \mathbb{T}^{trim}} |\langle\, h,\, \phi_Q^{(i)}\,\rangle|^2 \Big) \;\le\; \mathrm{const} \Big(1 + \frac{\| N_{\overline{\mathcal{F}}_\nu^{(ij)}} \|_\infty}{A^\mu} \Big) \int_{-\infty}^{\infty} |h(x)|^2 \, dx \tag{7.9}$$

holds uniformly in h, A and ν so long as $A \le \frac{1}{8} s^{\eta\nu}$.

Now, the original trees in $\mathcal{F}_\nu^{(ij)}$ were chosen satisfying a lower bound of the form

$$\frac{1}{|I_{\mathbb{T}}|} \int_{-\infty}^{\infty} \Big(\sum_{Q \in \mathbb{T}} \frac{1}{|I_Q|} |\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x) \Big)^{1/2} dx \;\ge\; \mathrm{const}_\phi \, s^{-(\nu+1)/q_0} \kappa_q$$

in the case $i = 2$ and $j = 3$, but essentially the same estimate remains true even after trimming as the next result shows. This, together with

$$\frac{1}{\sqrt{|I_Q|}} |\langle\, g,\, \phi_Q^{(2)}\,\rangle| \;\le\; \mathrm{const}_\phi \, s^{-1/\eta} s^{-(1+\eta)(\nu+1)/q_0} \kappa_q \tag{7.10}$$

valid for all Q in $\mathbb{Q}_\nu$ will serve as the analogue of (6.4).

(7.11) Proposition. *After removal of the set $\mathbb{T}^{fat} \cup \mathbb{T}^{edge} \cup \mathbb{T}^{down}$ from a tree $\mathbb{T}$ in $\overline{\mathcal{F}}_\nu^{(23)}$, the set $\mathbb{T}^{trim}$ of remaining tiles satisfies the inequality*

$$\frac{1}{|I_{\mathbb{T}}|} \int_{-\infty}^{\infty} \Big(\sum_{Q \in \mathbb{T}^{trim}} \frac{1}{|I_Q|} |\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2 \chi_{I_Q}(x) \Big)^{1/2} dx \;\ge\; \tfrac{1}{4}\, \mathrm{const}_\phi \, \kappa_q s^{-(\nu+1)/q_0}$$

uniformly in $\mathbb{T}$.

Proof. As the tiles in $\mathbb{T}^{fat}$ belong to the top $2\eta\nu$ layers in $\mathbb{T}$; *i.e.*, the tiles with largest time intervals, the upper bound (7.10) on the coefficients ensures that

$$\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}^{fat}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le\; \mathrm{const}_\phi(2\eta\nu)s^{-1/\eta}\kappa_q s^{-(1+\eta)(\nu+1)/q_0}|I_{\mathbb{T}}|.$$

On the other hand, the time intervals of the tiles in $\mathbb{T}^{edge}$ all lie in the union of two subintervals of $\mathbb{T}^{edge}$ each of length at most $\frac{1}{8}s^{-\eta\nu}|I_{\mathbb{T}}|$. Thus, by (4.16),

$$\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}^{edge}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le\; \tfrac{1}{4}s^{-\eta\nu}|I_{\mathbb{T}}|\,\mathrm{const}_\phi\,\kappa_q s^{-(\nu+1)/q_0}.$$

Finally, in view of (7.7) the time intervals of tiles in $\mathbb{T}^{down}$ are all pairwise disjoint, so

$$\int_{-\infty}^{\infty}\Big(\sum_{Q\in\mathbb{T}^{down}}\frac{1}{|I_Q|}|\langle\, g,\, \phi_Q^{(2)}\,\rangle|^2\chi_{I_Q}(x)\Big)^{1/2} dx \;\le\; \mathrm{const}_\phi\, s^{-(1+\eta)(1+\nu)/q_0}s^{-1/\eta}\kappa_q|I_{\mathbb{T}}|,$$

using (7.10) once again. The choice of η thus ensures that

$$2\eta\nu s^{-\eta(\nu+1)/q_0}s^{-1/\eta}+\tfrac{1}{4}s^{-\eta\nu}+s^{-1/\eta}s^{-\eta(\nu+1)/q_0} \;\le\; \tfrac{3}{4}$$

holds for all $\nu\ge 0$. For such η the lower bound for the sum over tiles in $\mathbb{T}^{trim}$ then follows. □

(7.12) Remark. By (7.11)

$$N_{\overline{\mathcal{F}}_\nu^{(ij)}}(x) \;=\; \sum_{\mathbb{T}\,\in\,\overline{\mathcal{F}}_\nu^{(ij)}}\chi_{I_{\mathbb{T}}}(x) \;=\; \sum_{\mathbb{T}^{trim}\,\in\,\overline{\mathcal{F}}_\nu^{(ij)}}\chi_{I_{\mathbb{T}}}(x)$$

since $\mathbb{T}^{trim}$ is not empty.

From now until the end of this this section $\mathbb{T}$ means $\mathbb{T}^{trim}$. It will be convenient to bring out the relationship between (7.11) and tent spaces $\mathfrak{N}^s$, $1\le s\le\infty$. Given h in $L^\gamma(\mathbb{R})$, $1<\gamma<\infty$, let

$$H(z)=\sum_{Q\in\mathbb{T}}\langle\, h,\, \phi_Q\,\rangle\chi_Q(z),\qquad z=(v,t)\in\mathbb{R}^2_+ \tag{7.13}$$

be the function on the upper half-plane associated to h. Then the lower bound (7.11) can be rewritten as

$$\mathrm{const}\,\kappa_\gamma b \;\le\; \frac{1}{|I_{\mathbb{T}}|}\|H_{\mathbb{T}}\|_{\mathfrak{N}^1}\ ,\qquad \inf_{x\in I_Q}M_\gamma(Mh)(x)\;\le \mathrm{const}\,\kappa_\gamma \tag{7.14}$$

for all $\mathbb{T}\in\overline{\mathcal{F}}_\nu^{(ij)}$ and Q in $\mathbb{T}$ where b and κ_γ are given constants. The idea is to think of (7.14) as the replacement for (6.4) where the size of a wave packet coefficient is replaced by a tent space norm. Similarly, by setting $J=I_{\mathbb{T}}$ we can interpret (4.8)(ii) as saying that the inequality

$$\frac{1}{|I_{\mathbb{T}}|^{1/\gamma}}\|H_{\mathbb{T}}\|_{\mathfrak{N}^\gamma}\;\le\;\mathrm{const}\min_{x\in I_{\mathbb{T}}}M_\gamma(Mh)(x),\qquad 1<\gamma<\infty, \tag{7.15}$$

holds uniformly in $\mathbb{T}$ and in h. When $\gamma = \infty$ the inequality (7.15) becomes

(7.16)
$$\|H_{\mathbb{T}}\|_{\mathfrak{N}^\infty} \ \leq \ \text{const}\, \|h\|_{BMO}$$

(*cf.* [14, Vol.I, pp.150–152]). Recall that wave packets $|I_Q|^{-1/2}\phi_Q$ have L^1-norm uniform in Q, hence uniform in complex Hardy $H^1_{\mathbb{C}}(\mathbb{R})$-space norm. This is the analogue of the trivial estimate $|\langle\, h,\, \phi_Q\,\rangle| \ \leq \ \sqrt{|I_Q|}\, \|h\|_\infty$, making the analogy between wave packet coefficients and the function $H_{\mathbb{T}}$ closer still.

Just as (6.3) followed from the general result (6.5), theorem (7.2) will follow in the same manner from the following general theorem.

(7.17) Theorem. *Let h be a function in $L^\gamma(\mathbb{R})$, $1 < \gamma < \infty$, satisfying* (7.14) *with respect to ϕ and $\overline{\mathcal{F}}^{(ij)}_\nu$. Then the counting function $N_{\overline{\mathcal{F}}^{(ij)}_\nu}$ satisfies the inequality*

$$\left(\int_{-\infty}^{\infty} N_{\overline{\mathcal{F}}^{(ij)}_\nu}(x)^\sigma\right)^{1/\sigma} \leq \text{const}\left(\frac{1}{b}\right)^{(\gamma_0+\delta_0)\theta}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^{\gamma/\sigma}$$

uniformly in h, b, and $\overline{\mathcal{F}}^{(ij)}_\nu$ for each $1 \leq \sigma < \infty$ and $\gamma_0 = \max(\gamma, \gamma')$ provided $0 < \delta_0 \leq \delta/2$ and $\theta = (1+2\delta)/(1+\delta)$.

Indeed, (7.2) then follows from (7.17) by taking $h = \pi(a_1)f$, $\gamma = p$, $b = s^{-(\nu+1)/p_0}$ and $\phi_Q = \phi^{(1)}_Q$ when $i = 1$ and $h = \pi(a_2)g$, $\gamma = q$, $b = s^{-(\nu+1)/q_0}$ and $\phi_Q = \phi^{(2)}_Q$

Proof of (7.17). The proof follows the same path as the proof of (6.5). Fix $\lambda \geq 1$. Just as in (6.8), using the grid structure in time and reasoning as in (6.2) we can replace the family $\overline{\mathcal{F}}^{(ij)}_\nu$ by yet another subfamily $\mathfrak{F}$ such that

(7.18)
$$\|N_{\mathfrak{F}}\|_{L^\infty} \ \leq \ \lambda, \qquad \{x : N_{\overline{\mathcal{F}}^{(ij)}_\nu}(x) \ \geq \ \lambda\} = \{x : N_{\mathfrak{F}}(x) \geq \lambda\}$$

Since counting functions are positive integer-valued, it is enough to establish the weak estimate

(7.19)
$$\lambda^{\sigma+\varepsilon_0}|\{x : N_{\mathfrak{F}}(x) \geq \lambda\}| \leq \text{const}\left(\frac{1}{b}\right)^{\theta(\gamma_0+\delta_0)\sigma}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^{\gamma} \qquad 1 \leq \lambda \leq s^{\nu/\eta},$$

for some fixed $\varepsilon_0 > 0$.

Now for each fixed $\mathbb{T} \in \mathfrak{F}$ let us define $|H_{\mathbb{T}}|_\gamma \ = \ \dfrac{1}{|I_{\mathbb{T}}|^{1/\gamma}}\|H_{\mathbb{T}}\|_{\mathfrak{N}^\gamma}$ and say $H_{\mathbb{T}} \in \mathcal{N}^\gamma_{\mathbb{T}}$ if $|H_{\mathbb{T}}|_\gamma < \infty$. Obviously we have that for each $\mathbb{T}$ fixed the space $\mathcal{N}^\gamma_{\mathbb{T}}$ can be identified with $L^\gamma((I_{\mathbb{T}}, \frac{dx}{|I_{\mathbb{T}}|}), \ell^2(\mathbb{T}))$ and so $\mathcal{N}^\gamma_{\mathbb{T}}$ interpolates in the same manner. For $\gamma = \infty$ we let $|H_{\mathbb{T}}|_\infty = \|H_{\mathbb{T}}\|_{\mathfrak{N}^\infty}$.

Define

$$H(x) = \{H_{\mathbb{T}}\chi_{I_{\mathbb{T}}}(x)\}_{\mathbb{T}\in\mathfrak{F}}$$

as a function from $\mathbb{R} \to \ell^\infty(\mathfrak{F}, \mathcal{N}^\gamma_{\mathbb{T}})$ in $L^\rho(\mathbb{R}, \ell^\alpha(\mathfrak{F}, \mathcal{N}^\gamma_{\mathbb{T}}))$, $\rho > 1$. We have

(7.20)(i)
$$h \to H : \quad L^2 \to L^2(\mathbb{R}, \ell^2(\mathfrak{F}, \mathcal{N}^2_{\mathbb{T}}))$$

(7.20)(ii)
$$h \to H : \quad BMO \to L^\infty(\mathbb{R}, \ell^\infty(\mathfrak{F}, \mathcal{N}^\infty_{\mathbb{T}}))$$

(7.20)(iii)
$$h \to H : \quad L^{1+2\delta_0} \to L^{1+2\delta_0}(\mathbb{R}, \ell^\infty(\mathfrak{F}, \mathcal{N}^{1+\delta_0}_{\mathbb{T}}))$$

(7.20)(i) follows from the Bessel inequality (7.9) with $A \sim \|N_{\mathfrak{F}}\|_\infty^{1/\mu}$ and $\mu \geq \frac{1}{\eta^2}$ and constant uniform in h and in $\mathbb{T}$.

(7.20)(ii) follows from (7.16) with uniform constant in h and in $\mathbb{T}$

(7.20)(iii) follows from (4.16) since the localized inequality with $J = I_{\mathbb{T}}$ can be re-expressed as

$$\frac{1}{|I_{\mathbb{T}}|^{1/1+\delta_0}}\|H_{\mathbb{T}}\|_{\mathfrak{N}^{1+\delta_0}} \leq \text{const} \min_{x\in I_{\mathbb{T}}} M_{1+\delta_0}(M(h))(x)$$

uniformly in h and $\mathbb{T}$. Hence

$$\left\{ \frac{1}{|I_{\mathbb{T}}|^{1/1+\delta_0}}\|H_{\mathbb{T}}\|_{\mathfrak{N}^{1+\delta_0}}\chi_{I_{\mathbb{T}}}(x)\right\}_{\mathbb{T}\in\mathfrak{F}} \in \ell^\infty$$

with bound const. $M_{1+\delta_0}(M(h))(\cdot) \in L^{1+2\delta_0}(\mathbb{R})$. By complex interpolation, $h \to H$ thus extends to bounded linear mappings from $L^\gamma(\mathbb{R})$ into

$$L^\gamma(\mathbb{R}, \ell^\gamma(\mathfrak{F}, \mathcal{N}^1_{\mathbb{T}})), \quad 2 \leq \gamma < \infty; \qquad L^\gamma(\mathbb{R}, \ell^{\gamma'+\delta_0}(\mathfrak{F}, \mathcal{N}^1_{\mathbb{T}})), \quad 1 < \gamma < 2, \tag{7.21}$$

since $\mathcal{N}^\rho_{\mathbb{T}} \subseteq \mathcal{N}^1_{\mathbb{T}}$ for $\rho > 1$ in view of (4.22). Hence the localized inequality

$$\text{(7.22)(i)} \quad \begin{aligned}&\left(\frac{1}{|J|}\int_{-\infty}^\infty \left[\sum_{\mathbb{T}\in\mathfrak{F},\, I_{\mathbb{T}}\subseteq J}\left(\frac{1}{|I_{\mathbb{T}}|}\|H_{\mathbb{T}}\|_{\mathfrak{N}^1}\right)^{\gamma'+\delta_0}\chi_{I_{\mathbb{T}}}(y)\right]^{\gamma/\gamma'+\delta_0} dy\right)^{1/\gamma}\\ &\qquad\leq \text{ const}\,\lambda^{1/\mu}\Big(\inf_{x\in J} M_\gamma(M(h))(x)\Big)\end{aligned}$$

holds for each s-adic J when $1 < \gamma < 2$, while

$$\text{(7.22)(ii)} \quad \begin{aligned}&\left(\frac{1}{|J|}\int_{-\infty}^\infty \left[\sum_{\mathbb{T}\in\mathfrak{F},\, I_{\mathbb{T}}\subseteq J}\left(\frac{1}{|I_{\mathbb{T}}|}\|H_{\mathbb{T}}\|_{\mathfrak{N}^1}\right)^{\gamma}\chi_{I_{\mathbb{T}}}(y)\right] dy\right)^{1/\gamma}\\ &\qquad\leq \text{ const}\,\lambda^{1/\mu}\Big(\inf_{x\in J} M_\gamma(M(h))(x)\Big)\end{aligned}$$

holds when $2 \leq \gamma < \infty$; and both are valid uniformly in h. Now we proceed as in Section 6, substituting (7.14) for (6.4):

$$\text{(7.23)(i)} \quad \int_{-\infty}^\infty N_{\mathfrak{F}}(x)^\rho\, dx \;\leq\; \text{const}\left(\frac{\lambda^{1/\mu}}{b}\right)^{\rho(\gamma'+\delta_0)}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^\gamma, \qquad 1 < \gamma < 2,$$

since $\rho(\gamma' + \delta_0) - \gamma > 0$; on the other hand,

$$\text{(7.23)(ii)} \quad \int_{-\infty}^\infty N_{\mathfrak{F}}(x)^\rho\, dx \;\leq\; \text{const}\left(\frac{\lambda^{1/\mu}}{b}\right)^{\rho\gamma}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^\gamma, \qquad 2 \leq \gamma < \infty,$$

since $\rho\gamma > \gamma$. The final step of the proof is much the same as that for (6.5). For fixed $\sigma \geq 1$ we have to show

$$\lambda^{\sigma+\varepsilon_0}|\{x : N_{\mathfrak{F}}(x) \geq \lambda\}| \leq \text{const}\left(\frac{1}{b}\right)^{\theta(\gamma_0+\delta_0)\sigma}\left(\frac{\|h\|_\gamma}{\kappa_\gamma}\right)^\gamma, \qquad 1 \leq \lambda \leq s^{\nu/\eta}, \tag{7.24}$$

for some fixed $\varepsilon_0 > 0$. Suppose first $1 < \gamma < 2$. Then $\gamma' = \gamma_0$, and so by (7.22)(i)

$$|\{x : N_{\mathfrak{F}}(x) \geq \lambda\}| \;\leq\; \frac{1}{\lambda^{\rho(1-(\gamma_0+\delta_0)/\mu)}}\Big(\frac{1}{b}\Big)^{\rho(\gamma_0+\delta_0)}\Big(\frac{\|h\|_\gamma}{\kappa_\gamma}\Big)^{\gamma}.$$

To check that the right hand side has the right decay in λ choose once again

$$0 < \delta_0 \leq \delta/2, \quad \rho \,=\, \theta\sigma \,=\, \Big(\frac{1+2\delta}{1+\delta}\Big)\sigma, \quad \varepsilon_0 \leq \delta/\mu, \quad \mu = \max\Big\{\frac{1}{\eta^2},\, \frac{50\delta_0}{\delta}\Big\}.$$

Then

$$\rho\Big(1 - \frac{1}{\mu}(\gamma_0+\delta_0)\Big) > \sigma + \varepsilon_0\,.$$

Estimate (7.24) now follows immediately since $\lambda \geq 1$. Entirely analogous calculations using (7.22)(ii) and the same choice of parameters produce the corresponding estimate for $\gamma \geq 2$. This concludes the proof of (7.17). □

(7.25) Remark. The choice of θ, δ_0 ensure that (7.2) follows from (7.17). Indeed, whether $\gamma_0 = p_0$ or q_0,

$$\theta\Big(1 + \frac{\delta_0}{\gamma_0}\Big) \;=\; \Big(1 + \frac{\delta}{1+\delta}\Big)\Big(1 + \frac{\delta_0}{\gamma_0}\Big) \;\leq\; (1+2\delta),$$

in which case

$$\Big(\frac{1}{b}\Big)^{\theta(\gamma_0+\delta_0)} \;=\; s^{\nu\theta(1+\delta_0/\gamma_0)} \;\leq\; s^{\nu(1+2\delta)}$$

as required.

8. Forests

In this section we extract from the family $\mathcal{F}_\nu \,=\, \bigcup_{i\,j}\mathcal{F}_\nu^{(ij)}$ of trees the 'forests' $\mathcal{W}_n^{(\nu)}$ alluded to in section 2. Recall first

(8.1) $$\mathcal{Q}_\nu^{(ij)} = \{Q \in \mathbb{T} : \mathbb{T} \in \mathcal{F}_\nu^{(ij)}\}, \qquad \mathcal{Q}_\nu = \bigcup_{i,\,j} \mathcal{Q}_\nu^{(ij)};$$

these are precisely the tiles which contribute to $\mathcal{D}_{good}$ since $Q \in \mathcal{Q}_\nu \implies I_Q \not\subseteq E_{bad}$.

To 'thin' the trees in $\mathcal{F}_\nu$, set

$$E_{dense}^{(\nu,ij)} \;=\; \{\,x :\, N_{\mathcal{F}_\nu^{(ij)}}(x) \;>\; s^{2\nu/r_0}\}, \qquad E_{dense}^{(\nu)} \;=\; \bigcup_{i,\,j} E_{dense}^{(\nu,ij)}$$

and

(8.2) $$\mathcal{Q}_{dense}^{(\nu)} \;=\; \bigcup_{i,\,j}\{\,Q \in \mathcal{Q}_\nu^{(ij)} \,:\, I_Q \subseteq E_{dense}^{(\nu,ij)}\}, \qquad \mathcal{Q}_{sparse}^{(\nu)} = \mathcal{Q}_\nu \setminus \mathcal{Q}_{dense}^{(\nu)}$$

from now on we adopt the convention that if $2\nu/r_0$ is not an integer, then it is replaced by $[2\nu/r_0]+1$. In view of (6.1) and (7.1),

$$|E^{(\nu)}_{dense}| \;\le\; \text{const}\, s^{-2\delta\nu}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r, \tag{8.3}$$

since $1/\eta > 2/r_0$ and $2/r_0 - (1+2\delta) \ge 2\delta > 0$. The second of the exceptional sets to be removed and the error term created in doing so can now be introduced. Set

$$E_{dense} = \bigcup_{\nu\ge 0} E^{(\nu)}_{dense} \quad \text{and} \quad E_2 \;=\; \bigcup_{\nu\ge 0}\bigcup_{i,j}\Big\{x \in s^2 I_Q : Q \in \mathcal{Q}^{(\nu,ij)}_{dense}\Big\} \tag{8.4}$$

and

$$\mathcal{D}_{dense}(f,\,g)(x) \;=\; \sum_{\nu\ge 0}\Bigg(\sum_{Q\in\mathcal{Q}^{(\nu)}_{dense}} c_Q\frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi^{(1)}_Q\,\rangle\langle\, g,\, \phi^{(2)}_Q\,\rangle\phi^{(3)}_Q(x)\Bigg). \tag{8.5}$$

These can be estimated in complete analogy with (2.10).

(8.6) Theorem. *The inequalities* $|E_2| \;\le\; \text{const.}\,|E_{dense}|$ *and*

$$\frac{1}{\gamma}\int_{\mathbb{R}\setminus E_1} |\mathcal{D}_{dense}(f,\,g)(a_3x)|\,dx \;\le\; \text{const.}\;\frac{|E_{dense}|}{a_3}$$

hold uniformly in f, g *and* γ *as well as the* a_j.

Proof. If $I^{(\nu,ij)}_1$, $I^{(\nu,ij)}_2$, ... are maximal among the time intervals $\{I_Q : Q \in \mathcal{Q}^{(\nu,ij)}_{dense}\}$, then they are disjoint and

$$s^2 I^{(\nu,ij)}_1 \cup s^2 I^{(\nu,ij)}_2 \cup \ldots \;=\; \Big\{x \in s^2 I_Q : Q \in Q^{(\nu,ij)}_{dense}\Big\}.$$

Hence by (8.3),

$$|E_2| \;\le\; \text{const.}\;|E_{dense}| \le \text{const.}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r \tag{8.7}$$

after summing over ν. To estimate $\mathcal{D}_{dense}$ we apply Theorem (3.6) with $\Omega = E_{dense},\; D_\Omega(f,\,g) = D_{dense}(f,\,g)$, and $E_\Omega = E_2$. $\square$

Consequently, all that's left is to estimate

$$\sum_{\nu\ge 0}\Bigg(\sum_{Q\in\mathcal{Q}^{(\nu)}_{sparse}} c_Q\frac{1}{\sqrt{|I_Q|}}\langle\, f,\, \phi^{(1)}_Q\,\rangle\langle\, g,\, \phi^{(2)}_Q\,\rangle\phi^{(3)}_Q\Bigg). \tag{8.8}$$

To that end, the remaining tiles $\mathcal{Q}^{(\nu)}_{sparse}$ are reorganized into 'forests'. First for each pair i, j write

$$\mathcal{Q}^{(\nu,ij)}_{sparse} \;=\; \{Q \in \mathcal{Q}^{(ij)}_\nu : I_Q \not\subseteq E^{(\nu,ij)}_{dense}\}\;. \tag{8.9}$$

Following Fefferman [4] we introduce an integer-valued function B_{ij} on $\mathcal{Q}^{(\nu,ij)}_{sparse}$ for each fixed i, j by setting

$$B_{ij}(Q) \;=\; \textit{card}\,\{Q^{\max} : Q \le Q^{\max}\,\} \tag{8.10}$$

where $Q^{\max}$ are maximal in $\mathcal{Q}^{(\nu,ij)}_{sparse}$, with respect to the partial order $\prec_j$. Note that these maximal tiles are the tree tops of trees $\mathbb{T} \in \mathcal{F}^{(ij)}_\nu$, since the removal of $\mathcal{Q}^{(\nu)}_{dense}$ does not create new trees in the sense that, if the time interval of a tree top in $\mathcal{F}_\nu$ is removed so it is the whole tree. Finally, recall that $P \prec_j Q \implies P \le Q$ (*cf.* (4.3)). In fact, for $P, Q \in \mathbb{T}$ a fixed $\Lambda^{(j)}$-tree, $I_P, I_Q \subsetneq I_{\mathbb{T}}$ we have that $P \prec_j Q$ is equivalent to $P \le Q$. This follows from the grid properties (1.11) and (1.12).

(8.11) Lemma ([4]). *Let Q, Q' and Q'' be distinct tile in $\mathcal{Q}^{(\nu,ij)}_{sparse}$ such that $Q \le Q'$ and $Q \le Q''$. Then $B(Q) \ge B(Q') + B(Q'')$ whenever Q' and Q'' are not comparable with respect to the partial order $\le$ on tiles.*

Proof. Set $B(Q') = s,\ B(Q'') = t$. Then there exist maximal $P'_1, \ldots, P'_s$ and $P''_1, \ldots, P''_t$ such that $Q' \le P'_j$, $1 \le j \le s$, while $Q'' \le P''_\ell$ for all $1 \le \ell \le t$. If $P'_{j_0} = P''_{\ell_0}$, say, then $Q' \le P'_{j_0}$, $Q'' \le P'_{j_0}$. But

$$Q' \ \le\ P'_{j_0}, \quad Q'' \ \le\ P'_{j_0} \quad \Longrightarrow \quad w_{P'_{j_0}} \subseteq w_{Q'} \cap w_{Q''},$$

so Q', Q'' have overlapping frequent intervals. On the other hand,

$$Q \ \le\ Q', \quad Q \ \le\ Q'' \quad \Longrightarrow \quad I_Q \subseteq I_{Q'} \cap I_{Q''},$$

so Q', Q'' also have overlapping time intervals. In this case Q' and Q'' will overlap as tiles and so be comparable, a contradiction. Hence $B(Q) \ \ge\ s + t$. □

We use $B_{ij} = B_{ij}(Q)$ to partition $\mathcal{Q}^{(\nu,ij)}_{sparse}$: set

$$\mathcal{Q}^{(\nu,ij)}_n \ = \ \left\{ Q \in \mathcal{Q}^{(\nu,ij)}_{sparse} :\ s^{n-1}\ \le B_{ij}(Q)\ <\ s^n \right\}. \tag{8.12}$$

Removing $E^{(\nu)}_{dense}$ ensures that $\mathcal{Q}^{(\nu,ij)}_n = \emptyset$ unless $n \le 2\nu/r_0 + 1$. Now each tree $\mathbb{T} \in \mathcal{F}^{(ij)}_\nu$ determines a family

$$\mathbb{T}^{(n)} \ = \ \mathbb{T} \cap \mathcal{Q}^{(\nu,ij)}_n \tag{8.13}$$

of tiles in $\mathcal{Q}^{(\nu,ij)}_n$. If $\mathbb{T}^{(n)}$ is empty we discard it. Since the tree top of $\mathbb{T}$ will not belong to $\mathbb{T}^{(n)}$ unless $n = 1$, $\mathbb{T}^{(n)}$ will not in general be a tree in the sense it contains a unique maximal element, but it will be the union of $\Lambda^{(j)}$-subtrees. Indeed, let $I_{Q_1}, I_{Q_2}, \ldots$ be maximal among all time intervals I_Q, $Q \in \mathbb{T}^{(n)}$; set

$$\mathbb{S}_1 = \{Q \in \mathbb{T}^{(n)} : I_Q \subseteq I_{Q_1}\}, \qquad \mathbb{S}_2 = \{Q \in \mathbb{T}^{(n)} : I_Q \subseteq I_{Q_2}\}, \qquad \ldots \tag{8.14}$$

and so on. These are $\Lambda^{(j)}$-trees having the respective $Q_1, Q_2, \ldots$ as tree-tops (*cf.* (4.18)). By repeating this construction for each $\mathbb{T}$ in $\mathcal{F}^{(ij)}_\nu$ we thus partition $\mathcal{Q}^{(\nu,ij)}_n$ into a family $\mathcal{W}^{(\nu,ij)}_n$ of $\Lambda^{(j)}$-trees that inherit the crucial properties from those in $\mathcal{F}^{(ij)}_\nu$. Now set

$$\mathcal{W}^{(\nu)}_n \ = \ \bigcup_{i,\,j} \mathcal{W}^{(\nu,ij)}_n.$$

As before $\mathbb{T}$ will denote a generic tree in $\mathcal{F}^{(ij)}_\nu$ and, as (8.14) suggests, a generic tree in $\mathcal{W}^{(\nu,ij)}_n$ obtained as a subtree of such a $\mathbb{T}$ will be denoted by $\mathbb{S}$. By (8.8), therefore,

$$\mathcal{D}_{good}(f,\,g) \ = \ \mathcal{D}_{dense}(f,\,g) + \sum_{\nu \ge 0} \left[\sum_{n=1}^{2\nu/r_0+1} \Big(\sum_{\mathbb{S} \in \mathcal{W}^{(\nu)}_\nu} \mathcal{D}_{\mathbb{S}}(f,\,g) \Big) \right].$$

Before beginning to estimate this last operator, it will be useful to summarize properties of its constituent terms. Let $N_{\mathcal{W}^{(\nu,ij)}_n} = N_{\mathcal{W}^{(\nu,ij)}_n}(x)$ be the function counting the number of trees in $\mathcal{W}^{(\nu,ij)}_n$ above x.

(8.15) Theorem. *Forests have the following properties*

(a) *for each* Q *in* $\mathcal{Q}_n^{(\nu,ij)}$

$$\frac{1}{\sqrt{|I_Q|}}|\langle\, f,\, \phi_Q^{(1)}\,\rangle| \le \mathrm{const}_\phi\, s^{-\nu/p_0}\kappa_p, \qquad \frac{1}{\sqrt{|I_Q|}}|\langle\, g,\, \phi_Q^{(2)}\,\rangle| \le \mathrm{const}_\phi\, s^{-\nu/q_0}\kappa_q;$$

(b) *for each tree* $\mathbb{S}$ *in* $\mathcal{W}_n^{(\nu,ij)}$

$$\left(\frac{1}{\gamma^2}\int_{-\infty}^{\infty}|\mathcal{D}_{\mathbb{S}}(f,\,g)(x)|^2\,dx\right)^{1/2} \le \mathrm{const}\, s^{-\nu/r_0}|I_{\mathbb{S}}|^{1/2}$$

uniformly in ν *and* n*; where as usual* $1/r_0 = 1/p + 0 + 1/q_0$ *and* $I_{\mathbb{S}}$ *is the time interval of the tree top of* $\mathbb{S}$*.*

(c) *the counting function* $N_{\mathcal{W}_n^{(\nu,ij)}}$ *satisfies the inequality*

$$N_{\mathcal{W}_n^{(\nu,ij)}}(x) \;\le\; \min\{N_{\mathcal{F}_\nu^{(ij)}}(x),\, s^{2\nu/r_0}\}$$

uniformly in ν *and* n*.*

Proof. Property (a) is clear, while (b) follows from (5.5). We prove (c). Fix a tree $\mathbb{T}$ in $\mathcal{F}_\nu^{(ij)}$ and let $\mathbb{S}_1, \mathbb{S}_2, \ldots$ be the corresponding trees determined by maximal tiles $Q_1, Q_2, \ldots$ as in (8.14). Maximality ensures that their time intervals $I_{Q_1}, I_{Q_2}, \ldots$ are disjoint. Thus $\mathbb{T}$ gives rise to trees in $\mathcal{W}_n^{(\nu,ij)}$ whose tree-tops have non-overlapping time intervals. Consequently,

$$N_{\mathcal{W}_n^{(\nu,ij)}}(x) \;\le\; N_{\mathcal{F}_\nu^{(ij)}}(x).$$

Now suppose $N_{\mathcal{W}_n^{(\nu,ij)}}(x_0) > s^{2\nu/r_0}$ for some x_0. Then there exist trees $\mathbb{S}_k$ in $\mathcal{W}_n^{(\nu,ij)}$, $1 \le k \le s^{2\nu/r_0}+1$, with x_0 in every $I_{\mathbb{S}_k}$. By the previous argument the $\mathbb{S}_k$ must have come from distinct trees $\mathbb{T}_k$ in $\mathcal{F}_\nu^{(ij)}$. Hence $N_{\mathcal{F}_\nu^{(ij)}}(x_0) > s^{2\nu/r_0}$. On the other hand, because of the grid structure in time the $I_{\mathbb{S}_j}$ must be nested since they all contain x_0; in particular, $x_0 \in I_{\mathbb{S}_m} \subseteq I_{\mathbb{S}_k}$ for some m and all k. Thus the inequality $N_{\mathcal{F}_\nu^{(ij)}}(x) > s^{2\nu/r_0}$ holds throughout $I_{\mathbb{S}_m}$. But this forces the time interval of each tile in $\mathbb{S}_m$ to lie in $E_{dense}^{(\nu,ij)}$, which in turn forces the tiles in $\mathbb{S}_m$ to belong to $\mathcal{Q}_{dense}^{(\nu)}$ contrary the fact that they all belong to $\mathcal{Q}_{sparse}^{(\nu)}$. This completes the proof. □

Using (8.15)(c) in conjunction with (6.1) and (7.1) we obtain the crucial counting estimate for the forests $\mathcal{W}_n^{(\nu)}$.

(8.16) Corollary. *The function* $N_{\mathcal{W}_n^{(\nu,ij)}}$ *satisfies the norm inequality*

$$\left(\int_{-\infty}^{\infty} N_{\mathcal{W}_n^{(\nu,ij)}}(x)^\sigma\,dx\right)^{1/\sigma} \;\le\; \mathrm{const}\, s^{(1+2\delta)\nu}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r$$

uniformly in n*,* ν *and* γ *for each* $\sigma,\ 1 \le \sigma < \infty$*.*

Property (8.15)(b) provides an L^2-estimate for individual functions $\mathcal{D}_{\mathbb{S}}(f,\,g)$ when $\mathbb{S}$ is a tree in $\mathcal{W}_n^{(\nu,ij)}$, but to deal collectively with all trees in $\mathcal{W}_n^{(\nu,ij)}$ 'almost orthogonality' estimates for families of $\mathcal{D}_{\mathbb{S}}(f,\, g)$ will be needed. To facilitate this a third and final exceptional set is eliminated, creating the last error term that has to be estimated. Set

$$E^{(\nu,n,ij)}_{edge} \;=\; \bigcup_{\mathbb{S}\,\in\,\mathcal{W}^{(\nu,ij)}_n} \Big\{ x : \mathrm{dist}(x,\, \partial I_{\mathbb{S}}) \;\le\; s^{-2\nu/r_0}|I_{\mathbb{S}}| \Big\} \quad E^{(\nu,n)}_{edge} = \bigcup_{i,j} E^{(\nu,n,ij)}_{edge} \tag{8.17}$$

Then, in view of (8.16),

$$|E^{(\nu,n,ij)}_{edge}| \;\le \mathrm{const.}\;\, s^{-2\nu/r_0}\|N_{\mathcal{W}^{(\nu,ij)}_n}\|_1 \;\le \mathrm{const.}\;\, s^{-2\delta\nu}\bigg(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\bigg)^r \tag{8.18}$$

Now set

$$E_3 = \bigcup_{\nu,\,n,\,i,j} \bigg(\bigcup_{\mathbb{S}\,\in\,\mathcal{W}^{(\nu,ij)}_n} \Big\{ s^2 I_Q \,:\, Q \,\in\, \mathbb{S},\, I_Q \,\subseteq\, E^{(\nu,n,ij)}_{edge} \Big\}\bigg). \tag{8.19}$$

Since for each i,j fixed there are at most $2\nu/r_0+1$ possible forests $\mathcal{W}^{(\nu,ij)}_n$ for each ν, (8.17)-(8.19) ensure that

$$|E_3| \;\le\; \mathrm{const}\bigg(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\bigg)^r \tag{8.20}$$

by a proof similar to the one used for the previous two exceptional sets; the constant will depend on s, δ and r_0 of course. The corresponding error term is

$$\mathcal{D}_{edge}(f,\, g) \;=\; \sum_{\nu\,\ge\,0}\bigg[\sum_{n\,=\,1}^{2\nu/r_0+1}\bigg(\sum_{I_Q\,\subseteq\, E^{(\nu,n)}_{edge}} c_Q\frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi^{(1)}_Q\,\rangle\langle\, g,\, \phi^{(2)}_Q\,\rangle\phi^{(3)}_Q\bigg)\bigg],$$

where the tiles Q in the innermost sum belong to the trees in $\mathcal{W}^{(\nu)}_n$.

(8.21) Theorem. *The inequality*

$$\frac{1}{\gamma}\int_{\mathbb{R}\setminus E_3} |\mathcal{D}_{edge}(f,\, g)(a_3x)|\,dx \;\le\; \mathrm{const.}\,\frac{|E_{edge}|}{a_3}$$

holds uniformly in f, g and γ as well as the a_j.

Proof. Once again we apply (3.6) but now with $\Omega = E_{edge}$, $D_\Omega(f,\, g) = \mathcal{D}_{edge}(f,\, g)(x)$ and $E_\Omega = E_3$. Again note that for $Q \in \mathcal{Q}^{(\nu,ij)}_n \subseteq \mathcal{Q}_\nu$, $I_Q \not\subseteq E_{bad}$. $\square$

After this final 'trimming' of the trees in $\mathcal{W}^{(\nu)}_n$ the forest operator associated with $\mathcal{W}^{(\nu,ij)}_n$ can be introduced. Given a tree $\mathbb{S}$ in $\mathcal{W}^{(\nu,ij)}_n$, set

$$\mathbb{S}^{trim} \;=\; \Big\{ Q \,\in\, \mathbb{S} :\, I_Q \,\not\subseteq\, E^{(\nu,n,ij)}_{edge} \Big\}.$$

Although we may lose all of some trees and individual tiles from others in this 'trimming', what is important is that we have not created any new trees. Thus (8.15) and (8.16) still hold and

$$\mathcal{D}_{good}(f,\, g) \;=\; \mathcal{D}_{dense}(f,\, g) \,+\, \mathcal{D}_{edge}(f,\, g) \,+\, \sum_{\nu\,\geq\,0}\Big[\sum_{n=1}^{2\nu/r_0+1}\Big(\sum_{\mathbb{S}\,\in\,\mathcal{W}_n^{(\nu)}}\mathcal{D}_{\mathbb{S}^{trim}}(f,\, g)\Big)\Big].$$

In the next section we will prove an L^2-norm estimate for the 'forest operator'

$$f,\, g \;\longrightarrow\; \sum_{\mathbb{S}\,\in\,\mathcal{W}_n^{(\nu,ij)}}\mathcal{D}_{\mathbb{S}^{trim}}(f,\, g)$$

associated with $\mathcal{W}_n^{(\nu,ij)}$, leaving only a simple summation and interpolation argument before the proof of theorem (1.9) is complete.

9. Forest Estimate

The section is devoted to the proof of the following 'forest estimate' for the forest operator introduced in the previous section.

(9.1) Theorem. *The inequality*

$$\frac{1}{\gamma^2}\int_{-\infty}^{\infty}\Big|\sum_{\mathbb{S}\,\in\,\mathcal{W}_n^{(\nu,ij)}}\mathcal{D}_{\mathbb{S}^{trim}}(f,\, g)(a_3x)\Big|^2\,dx \;\leq\; \frac{\text{const.}}{a_3}s^{-2\delta\nu}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r$$

holds uniformly in f, g, and forest $\mathcal{W}_n^{(\nu,ij)}$, as well as γ and the a_j.

Summing now over ν, n as well as i and j, we see that

$$\tag{9.2} \begin{aligned}&\frac{1}{\gamma^2}\int_{-\infty}^{\infty}\Big|\sum_{\nu\,\geq\,0}\Big[\sum_{n\,=\,1}^{2\nu/r_0}\Big(\sum_{\mathbb{S}\,\in\,\mathcal{W}_n^{(\nu)}}\mathcal{D}_{\mathbb{S}^{trim}}(f,\, g)(a_3x)\Big)\Big]\Big|^2\,dx\\ &\qquad\qquad\leq\; \frac{\text{const.}}{a_3}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r,\end{aligned}$$

with constant depending on δ and r_0, of course. Combining (9.2) with the earlier estimates for the exceptional sets E_2, E_3 and the associated error terms $\mathcal{D}_{dense}(f,\, g)$, $\mathcal{D}_{edge}(f,\, g)$ we finally obtain the companion estimate

$$\tag{9.3} |\{x\in\mathbb{R} : \mathcal{D}_{good}(f,\, g)(a_3x) \,>\, \gamma\}| \;\leq\; \text{const}\left(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\right)^r$$

to that in (2.10) for $\mathcal{D}_{bad}(f,\, g)$. Before proceeding with the proof of (9.1) we need an auxiliary lemma providing an L^2-estimate for a generic family $\mathcal{P}$ of pairwise-disjoint tiles in terms of the counting function. Given an arbitrary $\mathcal{M}_\mu$-test function ϕ let

$$\phi_P(x) = s^{k/2}\phi(s^kx - a\ell)e^{2\pi is^kxn}$$

be the wave packet associated to a tile $P \sim \{k,\, \ell,\, n\}$.

(9.4) Lemma. *Let $\mathcal{P}$ be a family of mutually disjoint tiles whose counting function $N_{\mathcal{P}}$ belongs to $L^1(\mathbb{R})$. Then the inequality*

$$\int_{-\infty}^{\infty} \Big| \sum_{P\in\mathcal{P}} a_P\phi_P(x)\Big|^2 dx \ \le\ \text{const.} \sup_{P\in\mathcal{P}} \Big(\frac{1}{\sqrt{|I_P|}}|a_P|\Big)^2 \int_{-\infty}^{\infty} N_{\mathcal{P}}(x)\, dx \ .$$

holds uniformly in $\{a_P\}$ and $\mathcal{P}$.

(9.5) Remark. The proof of (9.4) depends solely on decay properties, not on cancellation, and so applies to any family $\mathcal{P}$ of mutually disjoint tiles in $\mathcal{W}_n^{(\nu,ij)}$ irrespective of vanishing moments. For such a family its counting function $N_{\mathcal{P}}$ will satisfy the same L^σ-norm estimates as $N_{\mathcal{W}_n^{(\nu,ij)}}$ since the inequality $N_{\mathcal{P}}(x) \le N_{\mathcal{W}_n^{(\nu,ij)}}(x)$ will hold everywhere.

Proof of (9.4). On expansion

$$\int_{-\infty}^{\infty} \Big| \sum_{P\in\mathcal{P}} a_P\phi_P(x)\Big|^2 dx \ =\ \sum_{P\in\mathcal{P}} |a_P|^2 \,+\, \sum_{P\neq Q} a_P\overline{a_Q}\,\langle\, \phi_P,\, \phi_Q\,\rangle.$$

The diagonal term is easy to estimate:

$$\sum_{P\in\mathcal{P}} |a_P|^2 \ \le\ \sup_{P\in\mathcal{P}} \Big(\frac{1}{\sqrt{|I_P|}} a_P\Big)^2 \sum_{P\in\mathcal{P}} |I_P|.$$

It is in dealing with the second term that mutual disjointness of the tiles is needed. Observe first that we can assume $w_P \cap w_Q \ne \emptyset$; hence for each fixed P, it can be assumed that the Q have mutually disjoint time intervals which themselves are all disjoint from I_P. Since we have no control over vanishing moments, we split the sum into two parts

$$\sum_{P\in\mathcal{P}} a_P\Big(\sum_{Q\neq P,\, w_P\subseteq w_Q} \overline{a_Q}\,\langle\, \phi_P,\, \phi_Q\,\rangle\Big) \,+\, \sum_{Q\in\mathcal{P}} \overline{a_Q}\Big(\sum_{P\neq Q,\, w_Q\subseteq w_P} a_P\langle\, \phi_P,\, \phi_Q\,\rangle\Big). \tag{9.6}$$

It is enough to estimate the first of these terms since the proof of the second will be essentially the same after reversing the roles of P and Q. So fix a tile P in the outer sum and split the inner sum into two parts, depending on whether $\mathrm{dist}(I_P,\, I_Q) \le s|I_P|$ or $\mathrm{dist}(I_P,\, I_Q) > s|I_P|$. The restriction $w_P \subseteq w_Q$ ensures that $|I_P| \ge |I_Q|$, and so

$$|\langle\, \phi_P,\, \phi_Q\,\rangle| \ \le\ \text{const.}\, \|\pi(a)\phi\|^2 \frac{\sqrt{|I_Q|}}{\sqrt{|I_P|}}\Big(\frac{|I_P|}{|I_P| + \mathrm{dist}(I_P, I_Q)}\Big)^{\mu+1}$$

(*cf.* Appendix (B.1)). Consequently,

$$\begin{aligned} \Big|\sum_{Q} \overline{a_Q}\,\langle\, \phi_P,\, \phi_Q\,\rangle\Big| \ &\le\ \text{const.} \sup_{Q\in\mathcal{P}} \Big(\frac{1}{\sqrt{|I_Q|}}|a_Q|\Big) \frac{1}{\sqrt{|I_P|}} \sum_{Q} |I_Q| \\ &\le\ \text{const.} \sup_{Q\in\mathcal{P}} \Big(\frac{1}{\sqrt{|I_Q|}}|a_Q|\Big)\sqrt{|I_P|} \end{aligned} \tag{9.7}$$

since the I_Q are mutually disjoint and lie 'close' to I_P. Thus

$$\Big|\sum_{P\in\mathcal{P}} a_P\Big(\sum_{Q\neq P,\, w_P\subseteq w_Q} \overline{a_Q}\,\langle\,\phi_P,\,\phi_Q\,\rangle\Big)\Big| \;\le\; \text{const.}\Big(\sup_{P\in\mathcal{P}}\frac{|a_P|}{\sqrt{|I_P|}}\Big)^2\sum_{P\in\mathcal{P}}|I_P|,$$

taking the inner sum on the left over those Q for which $\text{dist}(I_P,\,I_Q)\;\le\; s|I_P|$; the factor $\|\pi(a)\phi\|^2$ was incorporated into the constants. When $\text{dist}(I_P, I_Q)\;>\;s|I_P|$, however, the inequality

$$|\langle\,\phi_P,\,\phi_Q\,\rangle| \;\le\; \text{const.}\,\|\pi(a)\phi\|\frac{\sqrt{|I_Q|}}{\sqrt{|I_P|}}\Big(\frac{|I_P|}{|I_P|+|x-y|}\Big)^{1+\mu}$$

holds for all $x\in I_P$ and $y\in I_Q$. Thus

$$\Big|\sum_Q \overline{a_Q}\,\langle\,\phi_P,\,\phi_Q\,\rangle\Big| \;\le\; \text{const.}\,\frac{1}{\sqrt{|I_P|}}\sup_{Q\in\mathcal{P}}\Big(\frac{1}{\sqrt{|I_Q|}}|a_Q|\Big)$$
$$\times\sum_Q\Big[\inf_{x\in I_P}\int_{I_Q}\Big(\frac{|I_P|}{|I_P|+|x-y|}\Big)^{1+\mu}dy\Big]$$

where again $\|\pi(a)\phi\|^2$ was incorporated into the constant. In turn, this last sum is dominated by a constant multiple of

$$\inf_{x\in I_P}\int_{\mathbb{R}\setminus sI_P}\Big(\frac{|I_P|}{|I_P|+|x-y|}\Big)^{1+\mu}dy \;\le\; \text{const.}\,|I_P|$$

uniformly in P. This leaves us with an estimate for $\sum_{P\in\mathcal{P}}\sqrt{|I_P|}\,a_P$ exactly as in (9.7), completing the proof. $\square$

The proof of (9.1) can now begin. It will be useful to collect together all properties of the construction so far since will be needed for the forest estimate. Set $\mathcal{Q}=\{Q\in\mathbb{S}^{trim}:\mathbb{S}\in\mathcal{W}_n^{(\nu,ij)}\}$.

(9.8) Remark. The forests $\mathcal{W}_n^{(\nu,ij)}$ and tiles $\mathcal{Q}$ in them have the following properties:

(a) for each Q in $\mathcal{Q}$

$$\frac{1}{\sqrt{|I_Q|}}|\langle\,f,\,\phi_Q^{(1)}\,\rangle|\le\text{const}_\phi\,s^{-\nu/p_0}\kappa_p,\qquad \frac{1}{\sqrt{|I_Q|}}|\langle\,g,\,\phi_Q^{(2)}\,\rangle|\le\text{const}_\phi\,s^{-\nu/q_0}\kappa_p;$$

(b) for each $\mathbb{S}$ in $\mathcal{W}_n^{(\nu,ij)}$

$$\Big(\frac{1}{\gamma^2}\int_{-\infty}^{\infty}|\mathcal{D}_{\mathbb{S}^{trim}}(f,\,g)(x)|^2\,dx\Big)^{1/2} \;\le\; \text{const.}\,s^{-\nu/r_0}|I_{\mathbb{S}}|^{1/2}.$$

Note that for $\mathbb{S}^{trim}\neq\emptyset$ the tree top of $\mathbb{S}$ and $\mathbb{S}^{trim}$ are the same ;

(c) the function $N_{\mathcal{W}_n^{(\nu,ij)}}$ counting the number of trees in $\mathcal{W}_n^{(\nu,ij)}$ satisfies the norm inequality

$$\Big(\int_{-\infty}^{\infty}(N_{\mathcal{W}_n^{(\nu,ij)}}(x))^\sigma\Big)^{1/\sigma} \;\le\; \text{const.}\,s^{(1+2\delta)\nu}\Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/\sigma}$$

uniformly in n, ν and γ for each $1 \le \sigma < \infty$;

(d) for each $\mathbb{S} \in \mathcal{W}_n^{(\nu,ij)}$,

$$\mathrm{dist}(I_Q,\, I_{\mathbb{S}}^c) \;>\; s^{-2\nu/r_0}|I_{\mathbb{S}}|, \qquad Q \in \mathbb{S}^{trim}\,;$$

(e) for each pair $\mathbb{S}_m, \mathbb{S}_\ell$ in $\mathcal{W}_n^{(\nu,ij)}$ with $I_{\mathbb{S}_m} \subseteq I_{\mathbb{S}_\ell}$ there are 'no V's' with respect to the partial order $\le$, *i.e.*, given $P \in \mathbb{S}_m^{trim}$ and $Q \in \mathbb{S}_\ell^{trim}$, then

$$w_P \cap w_{\mathbb{S}_\ell} \;=\; \emptyset \quad \text{ and } \quad w_Q \cap w_{\mathbb{S}_m} \neq \emptyset \quad \Longrightarrow \quad I_Q \,\subset\, I_{\mathbb{S}_\ell} \smallsetminus I_{\mathbb{S}_m}.$$

(f) For each pair $\mathbb{S}_m, \mathbb{S}_\ell$ in $\mathcal{W}_n^{(\nu,ij)}$ with $I_{\mathbb{S}_m} \subseteq I_{\mathbb{S}_\ell}$ and each $Q \in {\mathbb{S}_\ell}^{trim}$,

$$I_Q \subset I_{\mathbb{S}_\ell} \setminus I_{\mathbb{S}_m}, \qquad |I_Q| \;<\; s^{-2\nu/r_0}|I_{\mathbb{S}_m}| \quad \Longrightarrow \quad \mathrm{dist}(I_Q, I_{\mathbb{S}_m}) \ge s^{-2\nu/r_0}|I_{\mathbb{S}_m}|;$$

this follows from (8.19) because $I_Q \not\subseteq E_{edge}^{(\nu,n,ij)}$ and the restriction on the size of $|I_Q|$.

Proof of (9.1). First following Fefferman's 'scale counting' argumente we start by 'trimming' the trees in $\mathcal{W}_n^{(\nu,ij)}$ still further to remove tiles that are 'too fat' ([4]). Denote by $\mathcal{Q}^{top}$ those Q in $\mathcal{Q}$ for which there is no ascending chain

$$Q = Q_1 < Q_2 < \cdots < Q_{\overline{m}+1}, \qquad (Q_j \in \mathcal{Q}) \tag{9.9}$$

of length $\overline{m}+1$ where $\overline{m} = 4\nu/r_0$. In view of (9.8)(e) this amounts to collecting in $\mathcal{Q}^{top}$ the tiles having the 'fattest' $\overline{m}$ scales within every tree $\mathbb{S}^{trim} \in \mathcal{W}_n^{(\nu,ij)}$. Then there exist families $\mathcal{Q}_j^{top}$, $j = 1, \ldots, \overline{m}+1$, consisting of mutually disjoint tiles so that

$$\mathcal{Q} = \mathcal{Q}^{top} \cup (\mathcal{Q} \setminus \mathcal{Q}^{top}), \qquad \mathcal{Q}^{top} = \mathcal{Q}_1^{top} \cup \mathcal{Q}_2^{top} \ldots \cup \mathcal{Q}_{\overline{m}+1}^{top} \tag{9.10}$$

and

$$Q \in \mathbb{S}^{trim} \cap (\mathcal{Q} \setminus \mathcal{Q}^{top}) \quad \Longrightarrow \quad |I_Q| \le s^{-\overline{m}}|I_{\mathbb{S}}|. \tag{9.11}$$

By (9.4) and (9.6), therefore,

$$\begin{aligned} &\Big(\int_{-\infty}^{\infty}\Big|\sum_{Q \,\in\, \mathcal{Q}^{top}} c_Q \frac{1}{\sqrt{|I_Q|}}\langle\, f,\, \phi_Q^{(1)}\,\rangle\langle\, g,\, \phi_Q^{(2)}\,\rangle\phi_Q^{(3)}(a_3 x)\Big|^2\, dx\Big)^{1/2} \\ &\qquad\qquad \le\; \overline{m}\frac{\text{const.}}{a_3^{1/2}}\, s^{-\delta\nu}\Big(\frac{\|\pi(a_1)f\|_p\|\pi(a_2)g\|_q}{\gamma}\Big)^{r/2}, \end{aligned} \tag{9.12}$$

using (9.8)(a) to estimate the coefficients. It remains to estimate the L^2-norm of

$$\begin{aligned} &\sum_{Q\in\mathcal{Q}\setminus\mathcal{Q}^{top}} c_Q\, \frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi_Q^{(1)}\,\rangle\langle\, g,\, \phi_Q^{(2)}\,\rangle\phi_Q^{(3)} \\ &\quad = \sum_{\mathbb{S}\,\in\,\mathcal{W}_n^{(\nu,ij)}}\Big(\sum_{Q\,\in\,\mathbb{S}^{trim}\,\cap(\mathcal{Q}\setminus\mathcal{Q}^{top})} c_Q\,\frac{1}{\sqrt{|I_Q|}}\,\langle\, f,\, \phi_Q^{(1)}\,\rangle\langle\, g,\, \phi_Q^{(2)}\,\rangle\phi_Q^{(3)}\Big) \end{aligned} \tag{9.13}$$

Although $\mathbb{S}^{trim} \cap (\mathcal{Q} \setminus \mathcal{Q}^{top})$ does not contain its tree top anymore, we still regard it as a subtree of $\mathbb{S}^{trim}$ 'attached' to the tree top of $\mathbb{S}$; ie. as a tree itself with the tree top of $\mathbb{S}$ as a 'ghost' tree top. Hence if we call $\mathbb{S}^{trim} \cap (\mathcal{Q} \setminus \mathcal{Q}^{top}) = \mathbb{S}^{bot}$, we have that $Q < I_{\mathbb{S}} \times w_{\mathbb{S}}$ for $Q \in \mathbb{S}^{bot}$ (*cf.* (7.4)) and $\mathbb{S}^{bot}$ satisfies all the same estimates and properties that $\mathbb{S}^{trim}$ does; including (9.8)(c) since the counting function for the $\mathbb{S}^{bot}$ is pointwise no bigger than $N_{\mathcal{W}_n^{(\nu,ij)}}(x)$. Write

$$\{\mathbb{S}^{trim} \cap (\mathcal{Q} \setminus \mathcal{Q}^{top}) : \mathbb{S} \in \mathcal{W}_n^{(\nu,ij)}\} \;=\; \{\mathbb{S}_m^{bot} \;:\; m = 1, 2, \dots\}.$$

Then, our goal is to prove the L^2-estimate:

$$\int_{-\infty}^{\infty} \Big|\sum_m \mathcal{D}_{\mathbb{S}_m^{bot}}(f,\, g)(x)\Big|^2 dx \;\le\; \text{const.}\, s^{-2\delta\nu}\bigg(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma}\bigg)^r. \tag{9.14}$$

With abuse of notation we shall use $\mathbb{S}_m$ instead $\mathbb{S}_m^{bot}$ until the end of this section. Thus it is to be understood that $\mathbb{S}_m$ and its tiles satisfy (9.8) and (9.11), in particular, $|I_Q| \le s^{-4\nu/r_0}|I_{\mathbb{S}_m}|$ holds for all Q in $\mathbb{S}_m$. By unitarity,

$$\begin{aligned}\int_{-\infty}^{\infty} \Big|\sum_m \mathcal{D}_{\mathbb{S}_m}(f,\, g)(x)\Big|^2 dx \;&=\; \sum_m \bigg(\int_{-\infty}^{\infty} |\mathcal{D}_{\mathbb{S}_m}(f,\, g)(x)|^2\, dx\bigg) \\ &\quad + \sum_{m\neq\ell}\bigg(a_3 \int_{-\infty}^{\infty} \mathcal{D}_{\mathbb{S}_m}(f,\, g)(a_3x)\overline{\mathcal{D}_{\mathbb{S}_\ell}(f,\, g)(a_3x)}\, dx\bigg),\end{aligned} \tag{9.15}$$

splitting the sum into a 'diagonal' part and an 'off diagonal' part. the required estimate (i.e., is bounded by right hand side of (9.14)) In view of (b) and (c) in (9.8) the 'diagonal' part satisfies

$$\int_{-\infty}^{\infty} \Big|\sum_m \mathcal{D}_{\mathbb{S}_m}(f,\, g)(x)\Big|^2 dx \;\le\; \text{const.}\, s^{-2\delta\nu}\bigg(\frac{\|\pi(a_1)f\|_p \|\pi(a_2)g\|_q}{\gamma}\bigg)^r.$$

since $2/r_0 - (1+2\delta) = 2\delta > 0$. On the other hand, the 'off diagonal' part is a measure of almost orthogonality; we deal separately with the case when the time intervals $I_{\mathbb{S}_m}$ and $I_{\mathbb{S}_\ell}$ are disjoint and the more complicated case when they are not disjoint.

For each Q in $\mathbb{S}_m$ localize $\pi(a_3)\phi_Q^{(3)}$ by setting

$$\phi_Q^{(in)}(x) \;=\; (\pi(a_3)\phi_Q^{(3)})(x)\chi_{I_{\mathbb{S}_m}}(x), \qquad \phi_Q^{(out)}(x) \;=\; (\pi(a_3)\phi_Q^{(3)})(x) \;-\; \phi_Q^{(in)}(x). \tag{9.16}$$

The omission of the suffix (3) should not cause confusion since it is decay not vanishing moments that will be crucial from here on (*cf.* (1.14)). Set

$$\mathcal{D}_{\mathbb{S}_m}^{(in)}(f,\, g) \;=\; \sum_{Q\in\mathbb{S}_m} c_Q \frac{1}{\sqrt{|I_Q|}}\langle\, f,\, \phi_Q^{(1)}\,\rangle\langle\, g,\, \phi_Q^{(2)}\,\rangle\phi_Q^{(in)} \tag{9.17)(i}$$

for the localization of the function $\pi(a_3)\mathcal{D}_{\mathbb{S}_m}(f,\, g)$ to $I_{\mathbb{S}_m}$ and

$$\mathcal{D}_{\mathbb{S}_m}^{(out)}(f,\, g) \;=\; \sum_{Q\in\mathbb{S}_m} c_Q \frac{1}{\sqrt{|I_Q|}}\langle\, f,\, \phi_Q^{(1)}\,\rangle\langle\, g,\, \phi_Q^{(2)}\,\rangle\phi_Q^{(out)} \tag{9.17)(ii}$$

for the error created in doing so. Trimming the tops and edges ensures that for $x \notin I_{\mathbb{S}_m}$:

$$\begin{aligned} \frac{1}{\gamma}|\mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\, g)(x)| \ &\leq\ \text{const.}\, s^{-\nu/r_0} \sum_{Q\in\mathbb{S}_m} \Big(\frac{|I_Q|}{|I_Q|+\text{dist}(x, I_Q)}\Big)^{1+\mu} \\ &\leq\ \text{const.}\, s^{-\nu/r_0}(s^{-2\nu/r_0})^{\mu}(M\chi_{I_{\mathbb{S}_m}})(x)^2 \end{aligned} \tag{9.18}$$

by (9.8)(a)(d) and (9.11) since $\overline{m} = 4\nu/r_0$. Hence

$$\frac{1}{\gamma^2}\int_{-\infty}^{\infty}\Big(\sum_m |\mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\, g)(x)|\Big)^2 dx \ \leq\ \text{const.}(s^{-2\nu/r_0})^{1+2\mu}\|N_{\mathcal{W}^{(\nu,ij)}_n}\|_2^2 \tag{9.19}$$

using the vector-valued form of the Hardy-Littlewood-maximal theorem. When $I_{\mathbb{S}_m}\cap I_{\mathbb{S}_\ell} = \emptyset$ write

$$\begin{aligned} \langle\, \mathcal{D}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}_{\mathbb{S}_\ell}(f,\, g)\,\rangle \ &=\ \langle\, \mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle \\ &\quad +\ \langle\, \mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(in)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle\ +\ \langle\, \mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle. \end{aligned} \tag{9.20}$$

Then (9.19) ensures that

$$\frac{1}{\gamma^2}\sum_{m\neq\ell}|\langle\, \mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle| \ \leq\ \frac{\text{const.}}{a_3}(s^{-2\nu/r_0})^{1+2\mu}\|N_{\mathcal{W}^{(\nu,ij)}_n}\|_2^2. \tag{9.21}$$

To estimate the remaining terms we use (9.8)(b), restricting one error term to the other time interval. More precisely,

$$\sum_\ell \langle\, \mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle \ =\ \int_{I_{\mathbb{S}_m}} \mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\, g)(x)\Big\{\sum_\ell \overline{\mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)(x)}\Big\}\, dx. \tag{9.22}$$

In view of (9.8)(b) and (9.18), therefore,

$$\begin{aligned} &\frac{1}{\gamma^2}\sum_\ell |\langle\, D^{(in)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle| \\ &\quad\leq\ \text{const.}(s^{-\nu/r_0})|I_{\mathbb{S}_m}|\Big\{\frac{1}{|I_{\mathbb{S}_m}|}\int_{I_{\mathbb{S}_m}} \frac{1}{\gamma^2}\Big(\sum_\ell |\mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)(x)|\Big)^2 dx\Big\}^{1/2} \\ &\quad\leq\ \text{const.}(s^{-2\nu/r_0})^{1+\mu}\int_{I_{\mathbb{S}_m}} M_2\Big(\sum_\ell |M\chi_{I_{\mathbb{S}_\ell}}|^2\Big)(x)\, dx. \end{aligned}$$

Consequently,

$$\begin{aligned} &\frac{1}{\gamma^2}\sum_{m\,\neq\,\ell} |\langle\, \mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\, g),\, \mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\, g)\,\rangle| \\ &\quad\leq\ \text{const.}(s^{-2\nu/r_0})^{1+\mu}\int_{-\infty}^{\infty} N_{\mathcal{W}^{(\nu,ij)}_n}(x) M_2\Big(\sum_\ell |M\chi_{I_{\mathbb{S}_\ell}}|^2\Big)(x)\, dx. \end{aligned} \tag{9.23}$$

Hence by Hölder's inequality,

$$\frac{1}{\gamma^2}\sum_{m\neq\ell}|\langle\,\mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\,g),\,\mathcal{D}^{(out)}_{\mathbb{S}_\ell}(f,\,g)\,\rangle|\;\leq\;\text{const.}(s^{-2\nu/r_0})^{1+\mu}\|N_{\mathcal{W}_n^{(\nu,ij)}}\|_{4/3}\|N_{\mathcal{W}_n^{(\nu,ij)}}\|_8^2. \tag{9.24}$$

Disjointness of $I_{\mathbb{S}_m}$ and $I_{\mathbb{S}_\ell}$ was used only in (9.20).

We turn now to the case of non-disjoint time intervals. Fix $\mathbb{S}_m$, $\mathbb{S}_\ell$ with $I_{\mathbb{S}_m}\subseteq I_{\mathbb{S}_\ell}$ and choose tiles $P\in\mathbb{S}_m$, $Q\in\mathbb{S}_\ell$. The construction has placed severe restrictions on the possibilities for P, Q. Now $\langle\,\phi^{(3)}_P,\,\phi^{(3)}_Q\,\rangle=0$ unless $w_P\cap w_Q\neq\emptyset$, so from the outset we can assume $w_P\cap w_Q\neq\emptyset$, in which case $w_P\cap w_{\mathbb{S}_\ell}=\emptyset$ by (9.8)(e). But then $w_P\subseteq w_Q$, which in turn ensures that $|I_Q|\leq|I_P|\leq s^{-4\nu/r_0}|I_{\mathbb{S}_m}|$ because of (9.11). On the other hand, (9.8)(e) also ensures that $I_Q\subset I_{\mathbb{S}_\ell}\setminus I_{\mathbb{S}_m}$. Consequently,

$$\langle\,\phi^{(3)}_P,\,\phi^{(3)}_Q\,\rangle\neq 0\quad\Longrightarrow\quad|I_Q|\;\leq\;\frac{1}{A^2}\,|I_{\mathbb{S}_m}|,\qquad I_Q\;\subseteq\;I_{\mathbb{S}_\ell}\setminus I_{\mathbb{S}_m}. \tag{9.25}$$

where, as throughout the remainder of this section, we have set $A=s^{2\nu/r_0}$ to minimize the algebra involved. Moreover,

$$\text{dist}(I_Q,\,I^c_{\mathbb{S}_\ell})\;\geq\;\frac{1}{A}|I_{\mathbb{S}_\ell}|,\qquad\text{dist}(I_Q,\,I_{\mathbb{S}_m})\;\geq\;\frac{1}{A}|I_{\mathbb{S}_m}| \tag{9.26}$$

in view of parts (d) and (f) of (9.8) as well as (9.25). Write

$$\begin{aligned}\langle\,\mathcal{D}_{\mathbb{S}_m}(f,\,g),\,D_{\mathbb{S}_\ell}(f,\,g)\,\rangle\;&=\;\langle\,\mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\,g),\,D^{(in)}_{\mathbb{S}_\ell}(f,\,g)\,\rangle\,+\,\langle\,\mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\,g),\,D^{(out)}_{\mathbb{S}_\ell}(f,\,g)\,\rangle\\&\quad+\,\langle\,\mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\,g),\,D^{(in)}_{\mathbb{S}_\ell}(f,\,g)\rangle\,+\,\langle\,\mathcal{D}^{(out)}_{\mathbb{S}_m}(f,\,g),\,D^{(out)}_{\mathbb{S}_\ell}(f,\,g)\,\rangle.\end{aligned}$$

We can concentrate on the first term since the last three terms involve functions $\phi^{(out)}_Q$ and so can be dealt with in the same way as when the time intervals $I_{\mathbb{S}_m}$ and $I_{\mathbb{S}_\ell}$ were disjoint. Hence the last term satisfies (9.21) while the second and third satisfy (9.24). Now

$$\begin{aligned}\langle\,\mathcal{D}^{(in)}_{\mathbb{S}_m}(f,\,g),\,D^{(in)}_{\mathbb{S}_\ell}(f,\,g)\,\rangle\;&=\;\int_{I_{\mathbb{S}_m}}\Big(\sum_{P\in\mathbb{S}_m}c_P\frac{1}{\sqrt{|I_P|}}\langle\,f,\,\phi^{(1)}_P\,\rangle\,\langle\,g,\,\phi^{(2)}_P\,\rangle\,\phi^{(in)}_P(a_3x)\Big)\\&\times\Big(\sum_{Q\in\mathbb{S}_\ell}\overline{c_Q}\,\frac{1}{\sqrt{|I_Q|}}\overline{\langle\,f,\,\phi^{(1)}_Q\,\rangle}\,\overline{\langle\,g,\,\phi^{(2)}_Q\,\rangle}\,\overline{\phi^{(in)}_Q(a_3x)}\Big)\,dx\end{aligned}$$

where P ranges over all of $\mathbb{S}_m$, while Q ranges only over those tiles in $\mathbb{S}_\ell$ satisfying (9.25) and (9.26). But by (1.14),

$$|I_Q|^{1/2}|\phi^{(in)}_Q(x)|\;\leq\;\text{const.}\;\|\pi(a_3)\phi\|\Big(\frac{|I_Q|}{|x\,-\,y|}\Big)^{1+\mu},\qquad x\in I_{\mathbb{S}_m},\;y\in I_{\mathbb{S}_\ell},$$

and so

$$\sum_{Q\,\in\,\mathbb{S}_\ell}|I_Q|^{1/2}|\phi^{(in)}_Q(x)|\;\leq\;\text{const.}\int_{|x-y|\,\geq\,\frac{1}{A}|I_{\mathbb{S}_m}|}\Big(\int_0^{\frac{1}{A^2}|I_{\mathbb{S}_m}|}\Big(\frac{t}{|x-y|}\Big)^{1+\mu}\frac{dt}{t^2}\Big)\chi_{I_{\mathbb{S}_\ell}}(y)\,dy$$

for all x in $I_{\mathbb{S}_m}$ because of the restrictions in (9.25) and (9.26). Summing now over all trees $\mathbb{S}_\ell$ with $I_{\mathbb{S}_\ell} \supseteq I_{\mathbb{S}_m}$, we see that

$$\sum_\ell \sum_{Q \in \mathbb{S}_\ell} |I_Q|^{1/2} |\phi_Q^{(in)}(x)| \;\le\; \text{const.} \left(\frac{|I_{\mathbb{S}_m}|}{A^2}\right)^\mu \int_{|x-y| \ge \frac{1}{A}|I_{\mathbb{S}_m}|} \frac{N_{\mathcal{W}_n^{(\nu,ij)}}(y)}{|x\,-\,y|^{1+\mu}}\, dy.$$

Together with Meyer's lemma ([14, p.242]) this ensures that

$$\sum_\ell \Bigl(\sum_{Q \in \mathbb{S}_\ell} |I_Q|^{1/2} |\phi_Q^{(in)}(x)| \Bigr) \;\le\; \text{const.}\, \frac{1}{A^\mu}\, M(N_{\mathcal{W}_n^{(\nu,ij)}})(x), \qquad x \in I_{\mathbb{S}_m}.$$

Hence by (9.8) (b)

$$\frac{1}{\gamma^2} \Bigl| \sum_\ell \langle\, \mathcal{D}_{\mathbb{S}_m}^{(in)}(f,\, g),\, \mathcal{D}_{\mathbb{S}_\ell}^{(in)}(f,\, g)\,\rangle \Bigr| \;\le\; \text{const.}\, \frac{1}{A^{1+\mu}} \int_{I_{\mathbb{S}_m}} M_2(M(N_{\mathcal{W}_n^{(\nu,ij)}}))(x)\, dx,$$

and so

$$\frac{1}{\gamma^2} \Bigl| \sum_{I_{\mathbb{S}_m} \cap I_{\mathbb{S}_\ell} \ne \emptyset} \langle\, \mathcal{D}_{\mathbb{S}_m}^{(in)}(f,\, g),\, \mathcal{D}_{\mathbb{S}_\ell}^{(in)}(f,\, g)\,\rangle \Bigr| \;\le\; \text{const.}(s^{-2\nu/r_0})^{1+\mu} \|N_{\mathcal{W}_n^{(\nu,ij)}}\|_{4/3} \|N_{\mathcal{W}_n^{(\nu,ij)}}\|_4,$$

substituting back in for $A = s^{2\nu/r_0}$. The estimate for the 'off-diagonal' part, hence (9.14) also, now follow applying (9.8) (c) to this last inequality as well as to (9.21) and (9.24) so long as $\mu \ge 3$. This together with (9.12) completes the proof of (9.1). $\square$

10. Interpolation

A simple combination of induction and interpolation now finishes the proof of theorem (1.9) and hence that of Main Theorem II as well as Main Theorem I in Part I. Fix $\delta > 0$, δ small. The proof given in the previous sections established the boundedness of

$$\mathcal{D}^{(\varepsilon)} : f,\, g \longrightarrow \sum_{Q \in \mathbb{Q}_s} c_Q\, \frac{1}{\sqrt{|I_Q|}} \langle\, f,\, \phi_Q^{(\varepsilon_1)}\,\rangle \langle\, g,\, \phi_Q^{(\varepsilon_2)}\,\rangle\, \phi_Q^{(\varepsilon_3)}$$

as a mapping from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$ for each permutation ε and all p, q subject to the restriction

$$\tfrac{1}{2} + 2\delta \;\le\; 1/p + 1/q \;\le\; \tfrac{3}{2} - 2\delta, \qquad \bigl|1/p - 1/q\bigr| \;\le\; \tfrac{1}{2} - 2\delta.$$

Letting $\delta \longrightarrow 0$ we thus obtain the following result.

(10.1) Theorem. *The canonical operator $\mathcal{D}^{(\varepsilon)}$ is bounded from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$ for all p, q in the region*

$$p,\, q > 1, \qquad \tfrac{1}{2} \;<\; 1/p + 1/q \;<\; \tfrac{3}{2}, \qquad \bigl|1/p - 1/q\bigr| \;<\; \tfrac{1}{2}.$$

and all permutations ε.

To extend (10.1) to the full range $0 < 1/p + 1/q < \frac{3}{2}$, denote by $\Omega_k,\ k \geq 1$, the region in the first quadrant of the $1/p$-$1/q$ plane defined by

$$\frac{1}{2k} < \frac{1}{p} + \frac{1}{q} < \frac{3}{2}, \qquad \frac{1}{p} - \frac{1}{kq} < \frac{2k-1}{2k}, \qquad \frac{1}{q} - \frac{1}{kp} < \frac{2k-1}{2k}. \tag{10.2}$$

This is the interior of the convex region having corners

$$(0,\, 1/2k), \quad (0,\, 1-1/2k), \quad (1/2,\, 1), \quad (1,\, 1/2), \quad (1-1/2k,\, 0), \quad (1/2k,\, 0).$$

Notice that Ω_1 coincides with the permitted values of $(p,\, q)$ in (10.1), while

$$\Omega_1 \subset \Omega_2 \subset \ldots \subset \Omega_k \subset \Omega_{k+1} \subset \ldots$$

and

$$\bigcup_k \Omega_k \;=\; \big\{(1/p,\, 1/q) : p,\, q > 1, \quad 1/p + 1/q < \tfrac{3}{2}\big\}.$$

In particular, therefore, $\mathcal{D}^{(\varepsilon)}$ is bounded for all $(p,\, q)$ in Ω_1 and all ε. Assume then that $\mathcal{D}^{(\varepsilon)}$ is bounded for all $(p,\, q)$ in Ω_k and all ε. For clarity we take $k = 1$. A segment of the line $1/p+1/q = 1$ lies inside Ω_1; in fact, all points between $(1/4,\, 3/4)$ and $(3/4,\, 1/4)$ lie in Ω_1. By hypothesis, for any such pair $\mathcal{D}^{(\epsilon)}$ is bounded from $L^p(\mathbb{R}) \times L^{p'}(\mathbb{R})$ into $L^1(\mathbb{R})$, and so one adjoint will be bounded from $L^p(\mathbb{R}) \times L^\infty(\mathbb{R})$ into $L^p(\mathbb{R})$ while the other adjoint will be bounded from $L^\infty(\mathbb{R}) \times L^{p'}(\mathbb{R})$ into $L^{p'}(\mathbb{R})$. Thus, $\mathcal{D}^{(\varepsilon)}$ will be bounded on $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ for all points between $(1/4,\, 0)$ and $(3/4,\, 0)$ in the $1/p$-$1/q$ plane as well as those between $(0,\, 1/4)$ and $(0,\, 3/4)$ and all points in Ω_1. But the convex hull of these points is Ω_2, so by interpolation, $\mathcal{D}^{(\varepsilon)}$ will be bounded for all $(p,\, q)$ in Ω_2 and all ε ([1], [9]). The proof for general k is exactly the same, beginning with the segment of the line $1/p + 1/q = 1$ lying inside Ω_k. This, finally, completes the proof of Main Theorem II, and hence that of Main Theorem I in [8].

Appendix: Square Function Estimates

In this section we prove (4.6)-(4.7) as well as the two Bessel inequalities (6.7) and (7.5) used in the course of establishing the counting estimates (6.1) and (7.1).

A. Proof of Inequalities (4.6) and (4.7).

In this section we state and prove the results underpinning the basic inequalities (4.6) and (4.7) of this paper. These follow from classical Littlewood-Paley theory thanks to the cancellation properties the wave-packets enjoy on each tree. Given families $\phi_{\{k\}}$ and $\psi_{\{k\}}$ of $\mathcal{M}_\mu$-molecules which are uniformly norm bounded in the sense that

$$C_{\phi\psi} \;=\; \sup_k \|\phi_{\{k\}}\| \|\psi_{\{k\}}\|$$

is finite let

$$\mathcal{T} : h \longrightarrow \sum_{k,\,\ell=-\infty}^{\infty} c_{k\ell}\, s^k \Big(\int_{-\infty}^{\infty} h(y)\; \overline{\phi_{\{k\}}(s^k y - \ell)}\, dy \Big)\, \psi_{\{k\}}(s^k x - \ell) \tag{A.1}$$

be the associated 'frame' operator. With some abuse of notation we shall often write this more concisely as

$$\mathcal{T}h(x) \;=\; \sum_{k,\,\ell=-\infty}^{\infty} c_{k\ell}\, \langle\, h,\, \phi_{I_{k\ell}}\, \rangle\, \psi_{I_{k\ell}}(x) \tag{A.2}$$

where $I_{k\ell} = [s^{-k}\ell,\ s^{-k}(\ell+1))$; thus $\phi_{I_{k\ell}}(x)$ really means $s^{k/2}\phi_{\{k\}}(s^k x - \ell)$ and so on. Now the kernel

$$K(x,\, y) \;=\; \sum_{k,\,\ell=-\infty}^{\infty} c_{k\ell}\, s^k \psi_{\{k\}}(s^k x - \ell)\, \overline{\phi_{\{k\}}(s^k y - \ell)}$$

of $\mathcal{T}$ satisfies standard estimates of Calderón-Zygmund

$$|K(x,\, y)| \;\le\; \text{const. } \|\{c_{k\ell}\}\|_\infty\, C_{\phi\psi} \frac{1}{|x-y|}, \qquad |\nabla K(x,\, y)| \;\le\; \text{const. } \|\{c_{k\ell}\}\|_\infty C_{\phi\psi} \frac{1}{|x-y|^2},$$

(assuming $\mu > 1$). Furthermore, because K has two-sided vanishing moments, the limit

$$\lim_{N\to\infty} \sum_{k=-N}^{N} \Big(\sum_{\ell=-\infty}^{\infty} c_{k\ell}\, s^k \Big(\int_{-\infty}^{\infty} h(y)\, \overline{\phi_{\{k\}}(s^k y - \ell)}\, dy \Big)\, \psi_{\{k\}}(s^k x - \ell) \Big)$$

converges both in norm in $L^2(\mathbb{R})$ and pointwise almost everywhere; in particular, $\mathcal{T}h$ is well-defined for each L^2-function h and

$$\Big(\int_{-\infty}^{\infty} |\mathcal{T}h(x)|^2\, dx \Big)^{1/2} \;\le\; \text{const. } \|\{c_{k\ell}\}\|_\infty C_{\phi\psi} \Big(\int_{-\infty}^{\infty} |h(x)|^2\, dx \Big)^{1/2}.$$

Thus, $\mathcal{T}$ extends to a bounded operator on $L^p(\mathbb{R})$, $1 < p < \infty$, with the same L^p-operator norm (*cf.* [6] for example). By taking $c_{\{k\ell\}} = \pm 1$ and applying Khintchine's inequality we also obtain a 'square function' result for $L^p(\mathbb{R})$.

(A.3) Theorem. *With the notation of* (*A*.2) *the inequality*

$$\Big(\int_{-\infty}^{\infty} \Big(\sum_{k,\,\ell=-\infty}^{\infty} \frac{1}{|I_{k\ell}|} \big|\big\langle\, h,\, \phi_{I_{k\ell}} \big\rangle\big|^2 \chi_{I_{k\ell}}(x) \Big)^{p/2} dx \Big)^{1/p} \;\le\; \text{const.}\, C_{\phi\psi} \Big(\int_{-\infty}^{\infty} |h(x)|^p \Big)^{1/p}$$

holds uniformly for all h in $L^p(\mathbb{R})$, $1 < p < \infty$.

It is interesting to interpret this result as a 'Lusin area' inequality for the function $H = H(z)$ defined in the upper half-plane $z = (x,\, t)$, $t > 0$, by setting

$$H(z) \;=\; \sum_{k,\,\ell=-\infty}^{\infty} \big\langle\, h,\, \phi_{I_{k\ell}} \big\rangle\, \chi_{\Delta_{k\ell}}(z)$$

where $\chi_{\Delta_{k\ell}}$ is the tile sitting above $I_{k\ell}$ in the s-adic tiling of the time-scale plane.

Proof of Theorem (A.3). Choose any $\mathcal{M}_\mu$-molecule ψ such that $|\psi(x)| \geq 1$ on $[0,\ 1]$ and set $\psi_{\{k\}} = \psi$. Then the L^p-boundedness of $\mathcal{T}$ ensures that

$$\left(\int_{-\infty}^{\infty}\left|\sum_{k,\,\ell=-\infty}^{\infty} c_{k\ell}\, s^k\left(\int_{-\infty}^{\infty} h(y)\,\overline{\phi_{\{k\}}(s^k y-\ell)}\,dy\right)\psi(s^k x-\ell)\right|^p dx\right)^{1/p}$$

$$\leq\ \text{const.}\ \|\{c_{k\ell}\}\|_\infty C_{\phi\psi}\left(\int_{-\infty}^{\infty}|h(x)|^p\,dx\right)^{1/p}.$$

Thus after taking $c_{k\ell} = \pm 1$ and applying Khintchine's inequality we see that

$$\left(\int_{-\infty}^{\infty}\left(\sum_{k,\,\ell=-\infty}^{\infty}|\langle\, h,\,\phi_{I_{k\ell}}\,\rangle|^2\,|\psi_{I_{k\ell}}(x)|^2\right)^{p/2} dx\right)^{1/p}\ \leq\ \text{const.}\ C_{\phi\psi}\left(\int_{-\infty}^{\infty}|h(x)|^p\,dx\right)^{1/p}$$

from which (A.3) follows since $\psi_{I_{k\ell}}(x)\ \geq\ s^{k/2}\chi_{I_{k\ell}}(x)$. □

There is also a localized version of this square function result.

(A.4) Theorem. *With the notation of* (*A.*2) *the inequality*

$$\left(\frac{1}{|J|}\int_J\left(\sum_{I_{k\ell}\subseteq J}\frac{1}{|I_{k\ell}|}\left|\langle\, h,\,\phi_{I_{k\ell}}\,\rangle\right|^2\chi_{I_{k\ell}}(x)\right)^{p/2} dx\right)^{1/p}$$

$$\leq\ \text{const.}\ \sup_k\ \|\phi_{\{k\}}\|\,C_{\phi\psi}\ \left(\inf_{x\in J}\ M_p(Mh)(x)\right)$$

holds uniformly for all h in $L^p(\mathbb{R})$, $1<p<\infty$, and all J in $\mathcal{I}_s$.

Proof. The proof is standard. Write $h = h\,\chi_{sJ} + h(1-\chi_{sJ})$ and apply the L^p-inequality to $h\,\chi_{sJ}$:

$$\left(\frac{1}{|J|}\int_J\left(\sum_{I_{k\ell}\subseteq J}\frac{1}{|I_{k\ell}|}\left|\langle\, h\,\chi_{sJ},\,\phi_{I_{k\ell}}\,\rangle\right|^2\chi_{I_{k\ell}}(x)\right)^{p/2} dx\right)^{1/p}$$

$$\leq\ \text{const.}\,C_{\phi\psi}\left(\frac{1}{|J|}\int_{sJ}|h(x)|^p\,dx\right)^{1/p}\ \leq\ \text{const.}\,C_{\phi\psi}\left(\inf_{x\in J}\ M_p(h)(x)\right).$$

On the other hand,

$$\frac{1}{\sqrt{|I_{k\ell}|}}|\langle\,(h-h\chi_{sJ}),\,\phi_{I_{k\ell}}\,\rangle|\ \leq\ \text{const.}\ \|\phi_{\{k\}}\|\int_{\mathbb{R}\setminus sJ}|h(y)|\,\frac{|I_{k\ell}|^\mu}{(|I_{k\ell}|+dist(y,\ |I_{k\ell}|)^{\mu+1}}\,dy,$$

and so in view of Meyer's lemma ([14, p.242]),

$$\frac{1}{\sqrt{|I_{k\ell}|}}|\langle\,(h-h\chi_{sJ}),\,\phi_{I_{k\ell}}\,\rangle|\ \leq\ \text{const.}\ \|\phi_{\{k\}}\|\left(\frac{|I_{k\ell}|}{|J|}\right)^\mu Mh(x)$$

for any x in $I_{k\ell}$. But

$$\sum_{I_{k\ell}\subseteq J} |I_{k\ell}|^{2\mu}\,\chi_{I_{k\ell}}(x) \;\le\; \text{const.}\;|J|^{2\mu}.$$

Hence on J

$$\Bigg(\sum_{I_{k\ell}\subseteq J}\frac{1}{|I_{k\ell}|}\,|\langle\, f - f\chi_{2J},\,\phi_{I_{k\ell}}\,\rangle|^2\chi_{I_{k\ell}}(x)\Bigg)^{1/2} \;\le\; \text{const.}\sup_k\,\|\phi_{\{k\}}\|\;(Mh)(x)$$

uniformly in h and J. From this and the previous estimate for $h\,\chi_{I_{k\ell}}$ the theorem now follows. □

To apply these ideas in the setting of this paper let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree and for each i let

$$\mathcal{T}_{\mathbb{T}}^{(i)} : h \;\longrightarrow\; \sum_{Q\,\in\,\mathbb{T}}\langle\, h,\,\phi_Q^{(i)}\,\rangle\,\phi_Q^{(i)}(x)$$

be the associated 'frame operator' in which the wave packets $\phi_Q^{(i)}$ are defined by (1.2). After modulation and dilation such an operator can be reduced to (A.1). Indeed, set

$$\varphi_{\{k\}}^{(i)}(x) \;=\; \phi^{(i)}(x)\,e^{2\pi i(n\,-\,s^{-k}\lambda_{\mathbb{T}})x}, \qquad Q\sim\{k,\ell,n\}\in\mathbb{T}.$$

Proposition (4.1) ensures that the frequency n is uniquely determined by k and $\lambda_{\mathbb{T}}$, but the dependence of $\varphi_{\{k\}}^{(i)}$ on $\lambda_{\mathbb{T}}$ has been suppressed; in addition,

$$\|\pi(a_i)\varphi_{\{k\}}^{(i)}\| \;\le\; \text{const.}\;\|\pi(a_i)\phi^{(i)}\|, \qquad (\{k,\,\ell,\,n\}\;\in\;\mathbb{T}).$$

On the other hand, the $\varphi_{\{k\}}^{(i)}$ also have vanishing moments whenever $i\neq j$ so long they do not correspond to the tree top itself (*cf.*(4.5) and (4.10)). Note that when vanishing moments can be exploited one has the following inequality

$$\frac{1}{(|I_P||I_Q|)^{1/2}}\,|\langle\,\phi_Q^{(i)},\,\phi_P^{(i)}\,\rangle| \;\le\; \text{const.}\,\|\pi(a_i)\phi^{(i)}\|\,\frac{\min\left(\,|I_P|^{\mu},\,|I_Q|^{\mu}\,\right)}{(|I_P|+|I_Q|+\text{dist}(I_P,\,I_Q))^{1+\mu}}$$

(*cf.* [6, Appendix] for instance).

With the above choice of $\varphi_{\{k\}}^{(i)}$, therefore, the modulated operator $(\pi(a_i)\mathcal{T}_{\mathbb{T}}^{(i)}h)(x)\,e^{-2\pi i\lambda_{\mathbb{T}}x}$ can be written as

$$\sum_{\{k,\,\ell,\,n\}\,\in\,\mathbb{T}} s^k\bigg(\int_{-\infty}^{\infty}(\pi(a_i)H)(y)\,\overline{(\pi(a_i)\varphi_{\{k\}}^{(i)})(s^k y-\ell)}\,dy\bigg)\,(\pi(a_i)\varphi_{\{k\}}^{(i)})(s^k x-\ell)$$

where $H(x) \;=\; h(x)\,e^{-2\pi i\lambda_{\mathbb{T}}x}$ is a modulate of h. Thus, up to modulation and dilation, the tree operator $\mathcal{T}_{\mathbb{T}}$ has the form of (A.1), except possibly for the 'exceptional term' corresponding to $Q = I_{\mathbb{T}}\times w_{\mathbb{T}}$ which we estimate separately in a trivial way. So its basic properties follow immediately from those of $\mathcal{T}$. In particular, there is a Bessel inequality for each tree with bound uniform over all trees.

(A.5) Theorem. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree. Then the inequality*

$$\sum_{Q\,\in\,\mathbb{T}} |\langle\, h,\, \phi^{(i)}_Q\,\rangle|^2 \;\leq\; \text{const.}\, \|\pi(a_i)\phi^{(i)}\|^2 \int_{-\infty}^{\infty} |h(x)|^2\, dx$$

holds uniformly in $h, \mathbb{T}$ and $\phi^{(i)}$ for each $i \neq j$.

Applying the 'square function' result to the case $J = I_{\mathbb{T}}$ we obtain the next result.

(A.6) Theorem. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree. Then the inequality*

$$\biggl(\int_{-\infty}^{\infty}\Bigl(\sum_{Q\in\mathbb{T}}\frac{1}{|I_Q|}\bigl|\langle\, h,\, \phi^{(i)}_Q\,\rangle\bigr|^2\chi_{I_Q}(x)\Bigr)^{p/2}dx\biggr)^{1/p}$$

$$\leq\; \text{const.}\; \|\pi(a_i)\phi^{(i)}\|\biggl(\int_{-\infty}^{\infty}|(\pi(a_i)h)(x)|^p\,dx\biggr)^{1/p}$$

holds uniformly in h and $\mathbb{T}$ for each p, $1 < p < \infty$.

In its localized form this becomes

(A.7) Theorem. *Let $\mathbb{T}$ be a $\Lambda^{(j)}$-tree. Then the inequality*

$$\biggl(\frac{1}{|J|}\int_J\Bigl(\sum_{I_Q\,\subseteq\,J}\frac{1}{|I_Q|}\bigl|\langle\, h,\, \phi^{(i)}_Q\rangle\bigr|^2\chi_{I_Q}(x)\Bigr)^{p/2}dx\biggr)^{1/p}$$

$$\leq\; \text{const.}\,\sup_k\, \|\pi(a_i)\phi^{(i)}\|^2\Bigl(\inf_{x\in J}\, M_p(M(\pi(a_i)h))(x)\Bigr)$$

holds uniformly in h and J for all $h \in L^p(\mathbb{R})$, $1 < p < \infty$, all $J \in \mathcal{I}_s$ and each $i \neq j$.

B. Proof of Bessel Inequality (6.7).

Recall that a family $\mathcal{Q}$ of tiles is said to be A-separated if either $w^{(i)}_P \cap w^{(i)}_Q = \emptyset$ or $AI_P \cap AI_Q = \emptyset$ for all pairs P, Q in $\mathcal{Q}$; in particular, the tiles in $\mathcal{Q}$ are mutually disjoint. In this section we prove that the Bessel inequality

$$\sum_{Q\in\mathcal{Q}} |\langle\, h,\, \phi^{(i)}_Q\,\rangle|^2 \;\leq\; \text{const.}\, \|\pi(a_i)\phi^{(i)}\|^2\Bigl(1+\frac{\|N_{\mathcal{Q}}\|_\infty}{A^\mu}\Bigr)\|h\|_2^2$$

in (6.7) holds uniformly in h and $\mathcal{Q}$ whenever $A \geq 2$. The essential ingredient is the classical inequality

$$\text{(B.1)}\qquad \frac{1}{(|I_P||I_Q|)^{1/2}}|\langle\, \phi^{(i)}_Q,\, \phi^{(i)}_P\,\rangle| \;\leq\; \text{const.}\, \|\pi(a_i)\phi^{(i)}\|^2\,\frac{\max\,(\,|I_P|^\mu,\, |I_Q|^\mu\,)}{\text{dist}(I_P, I_Q)^{1+\mu}}$$

whose proof depends only on the decay of $\phi^{(i)}$ (*cf.* [6., Appendix], for instance). Now let

$$\mathcal{R} : h \longrightarrow \mathcal{R}(h) \;=\; \sum_{Q\,\in\,\mathcal{Q}} \langle\, h,\, \phi^{(i)}_Q \,\rangle\, \phi^{(i)}_Q$$

be the 'frame' operator on $L^2(\mathbb{R})$ determined by $\mathcal{Q}$ and $\|\mathcal{R}\|$ its L^2-operator norm. It is clearly enough to show that

$$\|\mathcal{R}\| \;\le\; \text{const.}\ \|\pi(a_i)\phi^{(i)}\|^2 \Big(1 + \frac{\|N_{\mathcal{Q}}\|_\infty}{A^\mu}\Big). \tag{B.2}$$

Now

$$\begin{aligned}\langle\, \mathcal{R}(h),\, \mathcal{R}(h)\,\rangle \;&=\; \sum_{Q\in\mathcal{Q}} |\langle\, h,\, \phi^{(i)}_Q\,\rangle|^2 \;+\; \sum_{P\ne Q} \langle\, h,\, \phi^{(i)}_P\,\rangle \overline{\langle\, h,\, \phi^{(i)}_Q\,\rangle}\, \langle\, \phi^{(i)}_P,\, \phi^{(i)}_Q\,\rangle \\ &=\; \langle\, \mathcal{R}(h),\, h\,\rangle \;+\; \sum_{P\in\mathcal{Q}} \Big(\sum_{Q\ne P,\, |I_Q|\le|I_P|} \langle\, h,\, \phi^{(i)}_P\,\rangle \overline{\langle\, h,\, \phi^{(i)}_Q\,\rangle}\, \langle\, \phi^{(i)}_P,\, \phi^{(i)}_Q\,\rangle \Big) \\ &\qquad +\; \sum_{Q\in\mathcal{Q}} \Big(\sum_{P\ne Q,\, |I_P|<|I_Q|} \langle\, h,\, \phi^{(i)}_P\,\rangle \overline{\langle\, h,\, \phi^{(i)}_Q\,\rangle}\, \langle\, \phi^{(i)}_P,\, \phi^{(i)}_Q\,\rangle \Big).\end{aligned}$$

It is enough to estimate the first of the two sums on the right hand side immediately above since the other follows in the same way, reversing the roles of P and Q. So fix P in $\mathcal{Q}$. We can assume $w_P \cap w_Q \ne \emptyset$, for otherwise the inner product of the wave packets vanishes; in particular, therefore, $AI_P \cap AI_Q = \emptyset$. On the other hand, basic estimate (B.1) with $|I_Q| \le |I_P|$ ensures that

$$\begin{aligned}&\sum_{Q\in\mathcal{Q}} |\langle\, h,\, \phi^{(i)}_Q\,\rangle\, \langle\, \phi^{(i)}_P,\, \phi^{(i)}_Q\,\rangle| \\ &\qquad \le\; \text{const.}\ \|\pi(a_i)\phi^{(i)}\|^2\, \frac{1}{\sqrt{|I_P|}} \sum_{Q\in\mathcal{Q}} \Big\{ \int_{I_Q} \Big(\frac{|I_P|}{\text{dist}(I_P, I_Q)} \Big)^{1+\mu} M(\pi(a_i)h)(x)\, dx \Big\} \\ &\qquad\qquad \le\; \text{const.}\ \|\pi(a_i)\phi^{(i)}\|^2 \sqrt{|I_P|} \Big(\frac{1}{A^\mu} \inf_{x\in I_P} M(M(\pi(a_i)h))(x) \Big),\end{aligned}$$

and so

$$\begin{aligned}&\sum_{P\in\mathcal{Q}} \Big(\sum_{Q\,\ne\,P,\,|I_Q|\,\le\,|I_P|} |\langle\, h,\, \phi^{(i)}_P\,\rangle \overline{\langle\, h,\, \phi^{(i)}_Q\,\rangle}\, \langle \phi^{(i)}_P, \phi^{(i)}_Q \rangle| \Big) \\ &\qquad \le\; \text{const.}\ \|\pi(a_i)\phi^{(i)}\|^2 \frac{1}{A^\mu} \int_{-\infty}^{\infty} N_{\mathcal{Q}}(x) M(\pi(a_i)h)(x) M(M(\pi(a_i)h))(x)\, dx.\end{aligned}$$

Hence

$$\|\mathcal{R}\|^2 \;\le\; \|\mathcal{R}\| \;+\; \text{const.}\ \|\pi(a_i)\phi^{(i)}\|^2 \frac{\|N_{\mathcal{Q}}\|_\infty}{A^\mu},$$

from which (B.2) follows. $\square$

C. Proof of Bessel inequality (7.5).

As the family of individual tiles has now to be replaced by a family of trees vanishing moments as well as the following two hypotheses become crucial to the proof. Let $\mathcal{F} = \{\mathbb{T}\}_{\mathbb{T}\in\mathcal{F}}$ be a family of $\Lambda^{(j)}$-trees in $\overline{\mathcal{F}}^{(ij)}_{\nu}$, $i<j$, having respective tree tops $Q_{\mathbb{T}} = I_{\mathbb{T}} \times w_{\mathbb{T}}$, such that for some $A \geq 2$,

(NC) *the inclusion* $AI_Q \subseteq I_{\mathbb{T}}$ *holds for all* $Q \in \mathbb{T}$, $I_Q \neq I_{\mathbb{T}}$, *and each* $\mathbb{T} \in \mathcal{F}$;

(NV) *if* $\mathbb{T}, \mathbb{T}' \in \mathcal{F}$ *are trees with* $\lambda_{\mathbb{T}} < \lambda_{\mathbb{T}'}$, *then* $w^{(i)}_P \cap w^{(i)}_Q = \emptyset$ *when* $P \in \mathbb{T},\ Q \in \mathbb{T}'$ *and* $I_Q \subseteq I_{\mathbb{T}}$.

Under these conditions we establish the Bessel inequality

$$\sum_{\mathbb{T}\in\mathcal{F}} \Big(\sum_{Q\in\mathbb{T}} |\langle h,\, \phi^{(i)}_Q \rangle|^2 \Big) \ \leq\ \text{const.}\, \|\pi(a_i)\phi^{(i)}\| \Big(1 + \frac{\|N_{\mathcal{F}}\|_\infty}{A^\mu} \Big) \|h\|_2^2 \tag{C.1}$$

with constant uniform in h and $\mathcal{F}$. By maximality the tree-tops $Q_{\mathbb{T}}$ form a disjoint family so in proving (C.1) we use the Bessel inequality in Appendix B to deal separately with the $Q_{\mathbb{T}}$ and hence reduce the problem to establishing the 'trimmed version'

$$\sum_{\mathbb{T}\in\mathcal{F}} \Big(\sum_{Q<Q_{\mathbb{T}}} |\langle h,\, \phi^{(i)}_Q \rangle|^2 \Big) \ \leq\ \text{const.}\, \|\pi(a_i)\phi^{(i)}\| \Big(1 + \frac{\|N_{\mathcal{F}}\|_\infty}{A^\mu} \Big) \|h\|_2^2, \tag{C.2}$$

summing only over tiles $Q < Q_{\mathbb{T}}$ in $\mathbb{T}$.

It will be illuminating to bring out the role of conditions (NC) and (NV).

(C.3) Remark. Condition (NC) has a geometric interpretation in terms of tents. The usual Tent $T(I)$ above an interval I consists of all point in the upper half plane lying on or below the graph $\{(x,t) : t = \text{dist}(x, I^c)\}$ (*cf.* [16, p.58]). The restriction $Q \in \mathbb{T} \implies AI_Q \subseteq I_{\mathbb{T}}$ ensures that tents with smaller slopes can be used. More precisely, let

$$T_A(I_{\mathbb{T}}) \ =\ \Big\{ (x,\, t) :\ 0 \ <\ t\ \leq\ \frac{4\,\text{dist}(x, I^c)}{A-1} \Big\}$$

be the tent above $I_{\mathbb{T}}$ whose sides have slope $\pm 4/(A-1)$. Then the square Δ_{I_Q} above I_Q lies inside $T_A(I_{\mathbb{T}})$ whenever $AI_Q \subseteq I_{\mathbb{T}}$.

(C.4) Remark. The (NV) condition implies that if $\lambda_{\mathbb{T}} \leq \lambda_{\mathbb{T}'}$, $P \in \mathbb{T}$, $Q \in \mathbb{T}'$ then $\langle \phi^{(i)}_Q, \phi^{(i)}_P \rangle \neq 0$ only when $I_{\mathbb{T}} \cap I_{\mathbb{T}'} = \emptyset$; or when $I_Q \subseteq I_{\mathbb{T}'} \setminus I_{\mathbb{T}}$ if $I_{\mathbb{T}} \subseteq I_{\mathbb{T}'}$. Note also that in particular we must necessarily have that $w^{(i)}_P \cap w^{(i)}_Q \neq \emptyset$. When this is the case, only the following two situations may arise:

$$\text{(a)}\ \ w_P = w_Q \qquad \text{or} \qquad \text{(b)}\ \ w_P \subset w_Q\,.$$

This is because by (1.11), (1.12) and the ordering of the Fourier supports, $w_Q \subset w_P$ cannot occur. Otherwise we would have that $w^{(i)}_Q \subseteq w_Q \subseteq w^{(i)}_P$ by scale and (1.11) which would give $\lambda_{\mathbb{T}} \in w^{(j)}_P$, $\lambda_{\mathbb{T}'} \in w^{(i)}_P$ and $\lambda_{\mathbb{T}} \leq \lambda_{\mathbb{T}'}$; a contradiction since $w^{(i)}_P < w^{(j)}_P$ for $i<j$. In case (b), the crucial result we need to bear in mind is (1.13). Indeed, note that if $w_P \subset w_Q$, $w^{(i)}_P \cap w^{(i)}_Q \neq \emptyset$ then

$w_P^{(i)} \subseteq w_P \subseteq w_Q^{(i)}$ by (1.11) and (1.12). Hence $\lambda_{\mathbb{T}} \in w_Q^{(i)}$ and $\lambda_{\mathbb{T}'} \in w_Q^{(j)}$, $i < j$. So in particular $\lambda_{\mathbb{T}} < \lambda_{\mathbb{T}'}$ and now Property (1.13) guarantees that (b) can occur for no more than one scale, *i.e.*, for no more than one choice of $|w_Q|$ for each P in $\mathbb{T}$.

The proof of (C.2) proceeds in three steps.

STEP 1. *Single Tree:* Apply (A.5).

STEP 2. *Single Row:* A row of trees is a family $\{\mathbb{T}_m\}_m$ of trees whose tree tops 'sit side by side' meaning that their time intervals $I_{\mathbb{T}_m}$ are mutually disjoint. Let

$$\mathcal{S} : h \longrightarrow \sum_m \Big(\sum_{Q \in \mathbb{T}_m,\, Q < Q_{\mathbb{T}_m}} \langle\, h,\, \phi_Q^{(i)} \,\rangle\, \chi_Q(z) \Big)$$

be the associated operator mapping functions on $\mathbb{R}$ to functions on the half-plane. Then

$$\sum_m \Big(\sum_{Q \in \mathbb{T}_m,\, Q < Q_{\mathbb{T}_m}} |\langle\, h,\, \phi_Q^{(i)} \,\rangle|^2 \Big) \;=\; \text{const.} \int_{-\infty}^{\infty} \int_0^{\infty} |(\mathcal{S}h)(x,t)|^2 \,\frac{dxdt}{t^2},$$

so in the case of a single row (C.2) follows once the norm of $\mathcal{S} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}^2_+)$ has been estimated. To accomplish this localize $\mathcal{S}$ by setting

$$\begin{aligned} \mathcal{S}_m^{(in)}(h)(z) \;&=\; \sum_{Q \in \mathbb{T}_m} \langle\, \pi(a_i)h,\, \phi_Q^{(in)} \,\rangle\, \chi_Q(z), \qquad \mathcal{S}^{(in)} = \sum_m \mathcal{S}_m^{(in)}, \\ \mathcal{S}_m^{(out)}(h)(z) \;&=\; \sum_{Q \in \mathbb{T}_m} \langle\, \pi(a_i)h,\, \phi_Q^{(out)} \,\rangle\, \chi_Q(z), \qquad \mathcal{S}^{(out)} = \sum_m \mathcal{S}_m^{(out)} \end{aligned} ;$$

where $\phi_Q^{(in)}$ and $\phi_Q^{(out)}$ are defined as in (9.16), now for $Q \in \mathbb{T}_m$.

By Step 1 each $\mathcal{S}_m^{(in)}$ is uniformly bounded, but the crucial point is that $\mathcal{S}_m^{(in)}$ maps functions defined on $I_{\mathbb{T}_m}$ to ones defined on the tent $T_A(I_{\mathbb{T}_m})$. Since the $T_A(I_{\mathbb{T}_m})$ have mutually disjoint base the $\mathcal{S}_m^{(in)}$ are orthogonal in the sense

$$\mathcal{S}_m^{(in)} \circ (\mathcal{S}_\ell^{(in)})^* = 0, \qquad (\mathcal{S}_\ell^{(in)})^* \circ \mathcal{S}_m^{(in)} = 0, \qquad (\ell \neq m).$$

Consequently,

$$\mathcal{S}^{(in)} : L^2(\mathbb{R}) \to L^2(\mathbb{R}^2_+), \qquad \|\mathcal{S}^{(in)} h\|_2 \;\le\; \sup_m \|\mathcal{S}_m^{(in)} h\|_2 \;\le\; \text{const.}\, \|h\|_2$$

uniformly in $\mathbb{T}_m$. Note that disjointness of $I_{\mathbb{T}_m}$ not (NC) was all that was needed for $\mathcal{S}^{(in)}$, but (NC) does become essential in establishing boundedness of $\mathcal{S}^{(out)}$ because here only decay of $\phi^{(i)}$ is available. But by (1.14) and Meyer's lemma ([14, p.242]),

$$\begin{aligned} |\langle\, \pi(a_i)h,\, \phi_Q^{(in)} \,\rangle| \;&\le\; \text{const.}\, |I_Q|^{\frac{1}{2}+\mu} \int_{\mathbb{R}\setminus I_{\mathbb{T}_m}} |(\pi(a_i)h)(y)| \frac{1}{|y-x|^{1+\mu}}\, dy \\ &\le\; \text{const.}\, t^{1/2} \bigg(\frac{t}{\text{dist}(x, I_{\mathbb{T}_m}^c)} \bigg)^{\mu} M(\pi(a_i)h)(x) \end{aligned}$$

for all $(x,\, t)$ in Δ_Q. Thus

$$|\mathcal{S}_m^{(out)}(h)(x,t)| \;\le\; \text{const.}\; t^{1/2}\bigg(\frac{t}{\text{dist}(x, I^c_{\mathbb{T}_m})}\bigg)^{\mu} M(\pi(a_i)h)(x)$$

inside $T_A(I_{\mathbb{T}_m})$, while $\mathcal{S}_m^{(out)}(h) = 0$ outside $T_A(I_{\mathbb{T}_m})$. Hence, by orthogonality,

$$\Big|\Big\langle \sum_m \mathcal{S}_m^{(out)}(h),\, \sum_{m'} \mathcal{S}_{m'}^{(out)}(h)\Big\rangle\Big| \;=\; \Big|\sum_m \langle\, \mathcal{S}_m^{(out)}(h),\, \mathcal{S}_m^{(out)}(h)\,\rangle\Big|$$
$$\le\; \text{const.}\; \sum_m \bigg(\int_{T_A(I_{\mathbb{T}_m})}\bigg(\frac{t}{\text{dist}(x, I^c_{\mathbb{T}_m})}\bigg)^{2\mu} M^2(\pi(a_i)h)(x)\frac{dx\,dt}{t}\bigg).$$

Integrating over each tent we see that

$$\big\|\mathcal{S}^{(out)}(h)\big\|_2 \;\le\; \text{const.}\; \frac{1}{(A-1)^{\mu}}\bigg(\int_{-\infty}^{\infty} M^2(\pi(a_i)h)(x)\, dx\bigg)^{1/2} \;\le\; \frac{\text{const.}}{(A-1)^{\mu}}\|h\|_2,$$

again because the $I_{\mathbb{T}_m}$ are disjoint. Hence

$$\int_{-\infty}^{\infty}\int_0^{\infty} |\mathcal{S}(h)(x,\, t)|^2\,\frac{dxdt}{t^2} \;\le\; \text{const.}\bigg(1+\frac{1}{A^{\mu}}\bigg)\int_{-\infty}^{\infty} |h(x)|^2\, dx,$$

establishing (C.2) for a single row since $(A-1) \ge A/2$ when $A \ge 2$. $\square$

STEP 3. *Multiple Rows*: The family $\mathcal{F}$ can be decomposed into at most $\|N_{\mathcal{F}}\|_\infty$ rows of trees, say $\{\mathbb{T}_{jm}\}_m,\; j = 1, 2, \ldots$ (*cf.* [4, p. 567]); let

$$\mathcal{R}_j : h \;\longrightarrow\; \sum_m \bigg(\sum_{Q\,\in\,\mathbb{T}_{jm},\, Q\,<\,Q_{\mathbb{T}_{jm}}} \langle\, h,\, \phi_Q^{(i)}\,\rangle\, \phi_Q^{(i)}\bigg)$$

be the associated row operators. Since

$$\sum_{\mathbb{T}\,\in\,\mathcal{F}}\bigg(\sum_{Q\,\in\,\mathbb{T},\, Q\,<\,Q_{\mathbb{T}}} |\langle\, h,\, \phi^{(i)}\,\rangle|^2\bigg) \;=\; \sum_j \langle\, \mathcal{R}_j h,\, h\,\rangle,$$

Bessel inequality (C.2) follows once the estimate

$$\Big\|\sum_j \mathcal{R}_j\Big\|_{\mathcal{L}(L^2)} \;\le\; \text{const.}\bigg(1+\frac{\|N_{\mathcal{F}}\|_\infty}{A^{\mu}}\bigg) \tag{C.6}$$

for the norm of $\sum_j \mathcal{R}_j$ as an operator on $L^2(\mathbb{R})$ has been established. We apply Cotlar's Lemma to the $\{\mathcal{R}_j\}$. Step 2 takes care of individual R_j. Indeed, if $\mathcal{S}_j$ denotes the operator from $L^2(\mathbb{R})$ to $L^2(\mathbb{R}^2_+)$ determined by $\{\mathbb{T}_{jm}\}$, then $\mathcal{R}_j = \mathcal{S}_j^* \circ \mathcal{S}_j$; consequently, by Step 2,

$$\|\mathcal{R}_j\|_{\mathcal{L}(L^2)} \;\le\; \text{const.}\bigg(1+\frac{1}{A^{\mu}}\bigg)$$

uniformly in j. As the $\mathcal{R}_j$ are self-adjoint, it thus remains to estimate $\mathcal{R}_j \circ \mathcal{R}_k$ for $j \neq k$. Now,

$$\langle \mathcal{R}_j(g),\, \mathcal{R}_k(h) \rangle \;=\; \sum_{\ell, m} \Big(\sum_{P \in \mathbb{T}_{jm},\, Q \in \mathbb{T}_{k\ell}} \langle g,\, \phi_P^{(i)} \rangle \overline{\langle h,\, \phi_Q^{(i)} \rangle} \langle \phi_P^{(i)},\, \phi_Q^{(i)} \rangle \Big). \tag{C.7}$$

Observe that (NV) implies that the inner product $\langle \phi_P^{(i)}, \phi_Q^{(i)} \rangle$ does not contribute to the sums on the right hand side of (C.7) unless $I_{\mathbb{T}_{jm}} \cap I_{\mathbb{T}_{k\ell}} = \emptyset$ or

$$I_Q \subseteq I_{\mathbb{T}_{k\ell}} \setminus I_{\mathbb{T}_{jm}} \text{ when } \lambda_{jm} < \lambda_{k\ell}\,; \qquad I_P \subseteq I_{\mathbb{T}_{jm}} \setminus I_{\mathbb{T}_{k\ell}} \text{ when } \lambda_{k\ell} < \lambda_{jm}.$$

This brings us to the discussion in (C.4) and we divide the sums into these cases accordingly. First let us replace sums over tiles with integrals over tents. To simplify the notation we write I_m instead of $I_{\mathbb{T}_{jm}}$ and I_ℓ instead of $I_{\mathbb{T}_{k\ell}}$. Set

$$G_m(z) \;=\; \sum_{P \,\in\, \mathbb{T}_{jm}} \langle g,\, \phi_P^{(i)} \rangle\, \chi_P(z), \qquad H_\ell(w) \;=\; \sum_{Q \,\in\, \mathbb{T}_{k\ell}} \langle h,\, \phi_Q^{(i)} \rangle\, \chi_Q(w)$$

and

$$K_{m,\,\ell}(z,\, w) \;=\; \sum_{P \,\in\, \mathbb{T}_{jm},\, Q \,\in\, \mathbb{T}_{k\ell}} \langle \phi_P^{(i)},\, \phi_Q^{(i)} \rangle\, \chi_P(z)\, \chi_Q(w).$$

Then G_m and H_ℓ have support in the respective tents $T_A(I_m)$, $T_A(I_\ell)$, while

$$\frac{1}{t^{1/2}}\, |G_m(x,\, t)| \;\le\; \text{const.}\, M(\pi(a_i)g)(x), \qquad \frac{1}{v^{1/2}}\, |H_\ell(y,\, v)| \;\le\; \text{const.}\, M(\pi(a_i)h)(y).$$

In addition,

$$\sum_{P \,\in\, \mathbb{T}_{jm}} \sum_{Q \,\in\, \mathbb{T}_{k\ell}} \langle g,\, \phi_P^{(i)} \rangle\, \overline{\langle h,\, \phi_Q^{(i)} \rangle}\, \langle \phi_P^{(i)},\, \phi_Q^{(i)} \rangle$$
$$= \int_{T_A(I_m)} \int_{T_A(I_\ell)} K_{m,\,\ell}(x,\, t\,; y,\, v)\, G_m(x,\, t)\, \overline{H_\ell(y,\, v)}\, \frac{dydv}{v^2} \frac{dxdt}{t^2}.$$

Now suppose $\lambda_{jm} \le \lambda_{k\ell}$. Then for each fixed P in $\mathbb{T}_{jm}$, the discussion above plus (C.4) ensure that $K_{m,\,\ell}(x,\, t\,; y,\, v) \,=\, 0$ on $\Delta_P \times T_A(I_\ell)$ unless

$$y \,\in\, (\mathbb{R} \setminus I_m) \cap I_\ell, \qquad v \,\in\, \begin{cases} \big[\, |I_P|,\, s|I_P|\, \big), & \text{in case (a)}, \\ [v_0,\, sv_0), & \text{in case (b)}, \end{cases}$$

where v_0 depends on $\lambda_{k\ell} - \lambda_{jm}$ (*cf.* (1.13) and [8, (5.15)]). Furthermore, in view of (B.1),

$$|K_{\ell,\, m}(x,\, t\,; y,\, v)| \;\le\; \text{const.}\, (vt)^{\frac{1}{2}} \left(\frac{t^\mu}{|x-y|^{1+\mu}} \right)$$

on $\Delta_P \times T_A(I_\ell)$ whichever case (a) or (b) applies. Consequently,

$$\sum_{\{\ell:\lambda_{jm} \le \lambda_{k\ell}\}} \Big(\int_{T_A(I_\ell)} \big|K_{m,\ell}(x,\, t\,; y,\, v)\overline{H_\ell(y,\, v)}\,\big| \,\frac{dydv}{v^2}\Big)$$
$$\le \text{ const. } t^{1/2}\Big(\frac{t}{\operatorname{dist}(x,\, I_m^c)}\Big)^{\mu} M(M(\pi(a_i)h)(x)$$

on Δ_P because the base intervals I_ℓ are disjoint. Thus

$$\sum_m \sum_{\{\ell:\lambda_{jm} \le \lambda_{k\ell}\}} \Big(\int_{T_A(I_m)} \Big\{ \int_{T_A(I_\ell)} \big|K_{m,\ell}(x,\, t\,; y,\, v)\, G_m(x,\, t)\overline{H_\ell(y,\, v)}\,\big| \,\frac{dydv}{v^2}\Big\} \frac{dxdt}{t^2}\Big)$$
$$\le \text{ const. } \frac{1}{A^\mu} \int_{-\infty}^{\infty} M(\pi(a_i)g)(x)\, M(M(\pi(a_i)h)(x)\, dx$$

as the base intervals I_m also are disjoint. After reversing the roles of P and Q we obtain the corresponding inequality for the case of all $\lambda_{jm},\, \lambda_{k\ell}$ with $\lambda_{k\ell} < \lambda_{jm}$. Together the two estimates now ensure that

$$|\langle\, \mathcal{R}_j(g),\, \mathcal{R}_k(h)\,\rangle| \;\le\; \text{const. } \frac{1}{A^\mu} \int_{-\infty}^{\infty} \Big\{ M(\pi(a_i)g)(x) M(M(\pi(a_i)h))(x)$$
$$+\; M(M(\pi(a_i)h))(y)\, M(\pi(a_i)h)(x) \Big\}\, dx.$$

whenever $j \neq k$. Inequality (C.6), and hence (C.2) also, follows applying Cotlar's lemma. This, finally, completes the proof. □

References

1. A.P. Calderón and A. Zygmund, *A note on the interpolation of linear operators*, Studia Math. **12** (1951), 194-204.
2. L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135-157.
3. A. Córdoba, A. and C. Fefferman, *Wave packets and Fourier integral operators*, Comm. Partial Diff. Eqs. **3** (1978), 979-1005.
4. C. Fefferman, *Pointwise convergence of Fourier series*, Annals of Math. **98** (1973), 551–571.
5. C. Fefferman, *The Uncertainty Principle*, Bulletin A.M.S (New Series) **9** (1983), 129-206.
6. J.E. Gilbert, Y.S. Han, J.A. Hogan, J.D. Lakey, D. Weiland, and G. Weiss, *Calderón Reproducing Formula and Frames with Applications to Singular Integrals*, preprint (1998).
7. J.E. Gilbert, and A.R. Nahmod, *Hardy spaces and a Walsh model for Bilinear Cone Operators*, Trans. Amer. Math. Soc **351** (1999), 3267–3300.
8. J.E. Gilbert, and A.R. Nahmod, *Bilinear Operators with Non-Smooth Symbol, I*, preprint (1999).
9. S. Janson, *On interpolation of multi-linear operators*, Lecture Notes in Math., Springer-Verlag **1302** (1988), 290-302.
10. M. Lacey, *The bilinear Hilbert transform is locally finite*, Rev. Math. Iberoamericana **13** (1997), 411-469.
11. M. Lacey, *The Bilinear Maximal function maps in L^p for $2/3 < p \le 1$*, Ann. of Math (2) **151** (2000), 35–57.

12. M. Lacey, and C. Thiele, *L^p estimates on the bilinear Hilbert transform,* $2 < p < \infty$, Annals of Math. **146** (1997), 683–724.

13. M. Lacey, and C. Thiele, *On Calderón's conjecture*, Annals of Math. **149** (1999), 475–496.

14. Y. Meyer, *Ondelettes et Opérateurs. II, Opérateurs de Calderón-Zygmund*, Hermann ed. Paris, 1990.

15. C. Muscalu, T. Tao and C. Thiele, *Multilinear operators given by singular multipliers*, preprint (1999).

16. E.M. Stein, *Harmonic Analysis: Real variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, N.J., 1993..

17. C. Thiele, *Time-Frequency analysis in the Discrete Phase Plane, Thesis*, Yale University, 1995.

Department of Mathematics, The University of Texas at Austin, Austin, TX 78712-1082

E-mail address: gilbert@math.utexas.edu

Department of Mathematics and Statistics, Lederle GRT, University of Massachusetts, Amherst, MA 01003-4515

E-mail address: nahmod@math.umass.edu