2001

Bilinear Operators with Non-Smooth Symbol, I

John E. Gilbert

Andrea R. Nahmod
university of massachusetts amherst, nahmod@math.umass.edu

BILINEAR OPERATORS WITH NON-SMOOTH SYMBOL, I

John E. Gilbert and Andrea R. Nahmod*

In memory of A. P. Calderón

Abstract. This paper proves the L^p-boundedness of general bilinear operators associated to a symbol or multiplier which need not be smooth. The Main Theorem establishes a general result for multipliers that are allowed to have singularities along the edges of a cone as well as possibly at its vertex. It thus unifies ealier results of Coifman-Meyer for smooth multipliers and ones, such the Bilinear Hilbert transform of Lacey-Thiele, where the multiplier is not smooth. Using a Whitney decomposition in the Fourier plane a general bilinear operator is represented as infinite discrete sums of time-frequency paraproducts obtained by associating wave-packets with tiles in phase-plane. Boundedness for the general bilinear operator then follows once the corresponding L^p-boundedness of time-frequency paraproducts has been established. The latter result is the main theorem proved in Part II, our subsequent paper [11], using phase-plane analysis.

1. Introduction

Let $\mathcal{B} : \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ be a continuous bilinear operator which commutes with simultaneous translations. Then there exists m in $\mathcal{S}'(\mathbb{R} \times \mathbb{R})$, the *symbol* or *multiplier*, such that

$$\mathcal{B}(f, g)(x) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} m(\xi, \eta)\widehat{f}(\xi)\; \widehat{g}(\eta)\; e^{2\pi i x(\xi+\eta)}\; d\xi d\eta, \tag{1.1}$$

and $\mathcal{B}$ commutes also with simultaneous dilations if m is homogeneous of degree 0. It is easy to see that $f,\, g \longrightarrow \mathcal{B}(f,\, g)$ is continuous as a mapping from $\mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R})$ into $L^2(\mathbb{R})$ when m is in $L^\infty(\mathbb{R}^2)$, and that $\mathcal{B}(f,\, g)$ lies in the complex Hardy space $H^2_{\mathbb{C}}(\mathbb{R})$ if in addition the support of m lies in the half-plane $\xi + \eta \geq 0$. The basic L^p-boundedness problem is to prescribe conditions on $m = m(\xi,\, \eta)$ so that $\mathcal{B}$ extends to a bounded operator from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$ for $p,\, q > 1$. This two-part series of papers establishes such L^p-boundedness when m is not necessarily smooth, unifying previous results of Coifman-Meyer for smooth multipliers with ones for the non-smooth case, including the recent results of Lacey-Thiele for the Bilinear Hilbert transform, as suggested by Remark V.3 in [2].

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.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Main Theorem I. *Let Γ be a closed one-sided cone with vertex at the origin and $m = m(\xi,\,\eta)$ a function having derivatives of all orders inside Γ such that*

$$|D^\alpha m(\xi,\,\eta)| \;\le\; \text{const.}\left(\frac{1}{\text{dist}((\xi,\,\eta),\;\partial\Gamma)}\right)^{|\alpha|}, \qquad |\alpha| \ge 0. \tag{1.2}$$

Then the bi-linear operator

$$\mathcal{C}_\Gamma : f,\, g \;\longrightarrow\; \int_\Gamma m(\xi,\,\eta)\,\widehat{f}(\xi)\,\widehat{g}(\eta)\,e^{2\pi i x(\xi+\eta)}\,d\xi d\eta$$

is bounded from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 3/2$, so long as no edge of Γ lies on the diagonal $\xi+\eta = 0$ or on a coordinate axis. Furthermore, when Γ lies in the half-plane $\xi+\eta > 0$ and $r \ge 1$, the operator $\mathcal{C}_\Gamma$ has range in the complex Hardy space $H^r_{\mathbb{C}}(\mathbb{R})$.

There is a corresponding Hardy space result when Γ lies in the half-plane $\xi + \eta < 0$. By changing variables $\eta \longrightarrow -\eta$ we also obtain an equivalent result for sesqui-linear operators that is often useful.

Corollary 1. *Under the hypotheses of Main Theorem I the sesqui-linear operator*

$$\overline{\mathcal{C}}_\Gamma : f,\, g \;\longrightarrow\; \int_\Gamma m(\xi,\,\eta)\,\widehat{f}(\xi)\,\overline{\widehat{g}(\eta)}\,e^{2\pi i x(\xi-\eta)}\,d\xi d\eta$$

is bounded from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 3/2$, so long as no edge of Γ lies on the diagonal $\xi-\eta = 0$ or on a coordinate axis. Furthermore, when Γ lies in the half-plane $\xi-\eta > 0$ and $r \ge 1$, the operator $\overline{\mathcal{C}}_\Gamma$ has range in the complex Hardy space $H^r_{\mathbb{C}}(\mathbb{R})$.

Strictly speaking, in these two results the multiplier m need only be smooth up to some sufficiently high order, but no attempt is made to quantify the necessary smoothness. If m is C^∞ everywhere in the plane except possibly at the origin its restriction to any cone Γ will satisfy (1.2) automatically provided

$$|D^\alpha m(\xi,\,\eta)| \;\le\; \text{const.}\,\frac{1}{(|\xi| + |\eta|)^{|\alpha|}}, \qquad |\alpha| \ge 0; \tag{1.3}$$

in particular, (1.3) will be satisfied whenever m is C^∞ and homogeneous of degree 0. For such multipliers the edges of the cone could be allowed to lie on one or more of the coordinate axes since $\widehat{f} \longrightarrow \widehat{f}|_{(0,\,\infty)}$ is bounded on $L^p(\mathbb{R})$. Thus an easy corollary of Main Theorem I is the following result confirming a conjecture made in [10].

Corollary 2. *Let $m_0 = m_0(\xi,\eta)$ be a piecewise C^∞-function on Σ_1 which is C^∞ in a neighborhood of the points $(\xi,-\xi)$ and let $\mathcal{B}_m$ be the bilinear operator whose symbol is the degree zero homogeneous extension $m = m(\xi,\,\eta)$ of m_0. Then $\mathcal{B}_m : L^p(\mathbb{R}) \times L^{p'}(\mathbb{R}) \to L^1(\mathbb{R})$; in addition,*

(a) if $m_0(\xi,-\xi) = 0$, then $\mathcal{B}_m : L^p(\mathbb{R}) \times L^{p'}(\mathbb{R}) \to H^1(\mathbb{R})$, while

(b) if $m_0(\xi,\eta) = 0$ when $\xi + \eta \le 0$, then $\mathcal{B}_m : L^p(\mathbb{R}) \times L^{p'}(\mathbb{R}) \to H^1_{\mathbb{C}}(\mathbb{R})$.

The formulation of Main Theorem I also arises naturally from the Fourier plane geometry of cone operators as well as time-frequency analysis. For when f, g are replaced by their wave-packet expansions, a cone lying in the half-plane $\xi + \eta > 0$ will eliminate from $\mathcal{C}_\Gamma(f,\, g)$ all wave-packets except those having vanishing moments and frequency in a fixed half-line. Consequently, translations in frequency take place in one direction only, and the wave packets not eliminated all belong to a *complex* Hardy space.

The proof of Main Theorem I proceeds via special cases. For a given θ let

$$\mathcal{C}_{P_\theta} : f,\, g \;\longrightarrow\; \int_{P_\theta} m(\xi,\, \eta)\, \widehat{f}(\xi)\, \widehat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi d\eta$$

be the cone operator associated with the half-plane $P_\theta = \big\{(\xi,\, \eta) :\ \xi \tan\theta - \eta > 0\,\big\}$ and $\overline{\mathcal{C}}_{P_\theta}$ the corresponding sesqui-linear version.

(1.4) Theorem. *Let $m = m(\xi,\, \eta)$ be a function having derivatives of all orders in the half-plane P_θ such that*

$$|D^\alpha m(\xi,\, \eta)| \;\le\; \text{const.} \left(\frac{1}{\text{dist}((\xi,\, \eta),\ \partial P_\theta)}\right)^{|\alpha|}, \qquad |\alpha| \ge 0.$$

Then, if ∂P_θ is not one of the coordinate axes, $\mathcal{C}_{P_\theta}$ and $\overline{\mathcal{C}}_{P_\theta}$ are bounded from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 3/2$, whenever $\theta \ne -\pi/4$ and $\pi/4$ respectively.

Again the coordinate axes can be allowed if m satisfies (1.3) everywhere away from the origin in the plane. By taking $m(\xi,\, \eta) \equiv 1$ we thus obtain all the Bilinear Hilbert transform results of Lacey-Thiele (*cf.* [16], [17]). Actually, one could attempt to establish Main Theorem I based on these results of Lacey-Thiele, but we do not do so because our goal is to develop techniques that will be readily applicable in other contexts. While these techniques have certain commonalities with those used by Lacey-Thiele, they are significantly different. Save for the restriction $r > 2/3$, theorem (1.4) also includes the well-known result of Coifman-Meyer establishing the boundedness of

$$\mathcal{C}_{\mathbb{R}^2}(f,\, g) \;=\; \int_{\mathbb{R}^2} m(\xi,\, \eta)\, \widehat{f}(\xi)\, \widehat{g}(\eta)\, e^{2\pi i x(\xi-\eta)}\, d\xi d\eta$$

from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $r > 1/2$, for any C^∞-function m satisfying (1.3) (*cf.*, [3, 4]). In fact, it is enough to write $\mathcal{C}_{\mathbb{R}^2}$ as the sum $\mathcal{C}_{P_\theta} + \mathcal{C}_{\mathbb{R}^2\backslash P_\theta}$ for any allowed choice of θ. What is interesting, however, is that a natural 'miniaturization' of the proof of Main Theorem I actually provides a proof of the L^p-boundedness of $\mathcal{C}_{\mathbb{R}^2}$ for the full range of r as well as the reason for the failure to obtain the lower value of r in Main Theorem I. Indeed, in (1.3) the only singularity in the multiplier is at the origin - there is a preferred point in frequency, in other words - so that wave packets have only to contain translations in time and dilation. By contrast, in Main Theorem I there is no such preferred point because the singularities can lie on the full boundary of Γ. As a result wave packets now have to contain translation in frequency as well, *i.e.*, modulation. Even after including modulations, however, there is only one point in the proof of Main Theorem

I, an application of the Hausdorff-Young inequality, at which it becomes essential to impose the condition $r > 2/3$. Save for this, the proof of Main Theorem I would be valid without restriction on r. For the reader's convenience we have included the 'miniaturized' proof for $\mathcal{C}_{\mathbb{R}^2}$ in an Appendix (see also [12] [14] for other recent and independent proofs of the latter and more).

Although L^p-boundedness of $\overline{\mathcal{C}}_{P_{-\theta}}$ in (1.4) follows from that of $\mathcal{C}_{P_\theta}$ by a change of variable it is geometrically more convenient to deal independently with $\mathcal{C}_{P_\theta}$ and $\overline{\mathcal{C}}_{P_\theta}$ for a restricted range of θ. To be precise, we shall prove the following results.

(1.5) Theorem. *Let $m = m(\xi, \eta)$ be a function having derivatives of all orders in the half-plane P_θ such that*

$$|D^\alpha m(\xi, \eta)| \;\le\; \text{const.} \left(\frac{1}{\operatorname{dist}((\xi, \eta),\ \partial P_\theta)}\right)^{|\alpha|}, \qquad |\alpha| \ge 0.$$

Then $\mathcal{C}_{P_\theta}$ is bounded from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 3/2$, so long as $0 < \theta \le \pi/4$ while $\overline{\mathcal{C}}_{P_\theta}$ is bounded if $0 < \theta < \pi/4$.

Granted (1.5), Main Theorem I and theorem (1.4) follow easily.

Proof of (1.4). After a change of variable $\xi \to -\xi, \eta \to -\eta$ and $x \to -x$ if necessary, we can assume that $-\pi/4 < \theta < 3\pi/4,\ \ \theta \ne 0$. On the other hand, by interchanging the roles of f and g if necessary, we can further assume that $-\pi/4 < \theta < 0$ or $0 < \theta \le \pi/4$. Now the L^p-boundedness of $\mathcal{C}_{P_\theta}$ established in (1.5) takes care of this last range of θ, leaving just the case $-\pi/4 < \theta < 0$. But this follows from the boundedness of $\overline{\mathcal{C}}_{P_\theta}$ established in (1.5), changing variables $\eta \to -\eta$. □

Proof of Main Theorem I. That $\mathcal{C}_\Gamma$ has range in the complex Hardy space $H^r_{\mathbb{C}}(\mathbb{R})$, $r \ge 1$, when Γ lies in the half-plane $\xi + \eta > 0$ is clear once L^p-boundedness has been established. To deduce the boundedness of $\mathcal{C}_\Gamma$ from (1.4) choose half-planes P_{θ_1} and P_{θ_2} so that Γ is one half of the two-sided cone $P_{\theta_1} \cap P_{\theta_2}$. Then there exist C^∞-functions σ_0, σ_1, and σ_2 so that the support cone of σ_0 lies strictly inside Γ and

$$\mathcal{C}_\Gamma(f, g) \;=\; \left(\int_{\mathbb{R}^2} \sigma_0(\xi, \eta) + \int_{P_{\theta_1}} \sigma_1(\xi, \eta) \;+\; \int_{P_{\theta_2}} \sigma_2(\xi, \eta)\right) \times m(\xi, \eta)\, \widehat{f}(\xi)\, \widehat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi d\eta.$$

The L^p-boundness of $\mathcal{C}_\Gamma$ thus follows immediately from (1.4) and the Coifman-Meyer result for $\mathcal{C}_{\mathbb{R}^2}$ which itself is a consequence of (1.4). □

Thus we shall concentrate on proving theorem (1.5). There are two fundamental ideas. The first is to represent $\mathcal{C}_{P_\theta}$ in terms of a doubly-infinite sum of 'discrete' bilinear operators, and then secondly to establish L^p-boundedness for these discretizations. Given positive numbers a_j, a positive rational ρ, and $\mathcal{M}_\mu$-test functions $\phi^{(j)}$, let

$$\phi^{(j)}_{k\ell n}(x) \;=\; \phi^{(j)}_Q(x) \;=\; s^{k/2}\phi_j(s^k x - a_j \ell)\, e^{2\pi i s^k x n}, \qquad s = 2^\rho, \tag{1.6}$$

be the corresponding wave packet associated with a tile $Q \sim \{k, \ell, n\}$ in phase plane, incorporating translation in time, scaling, and modulation; *cf.* section 2 for terminology and notation. By analogy with 'standard' paraproducts we form the sum

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

over all tiles $Q \sim \{k, \ell, n\}$ in phase plane, the coefficients $c_{k\ell n}$ being in ℓ^∞. In 'standard' paraproducts there are no modulations and boundedness from $\ell^\infty \times L^\infty(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^q(\mathbb{R})$ is well-known under the assumption that at least two of the 'mother wave functions' have vanishing moment (and more generally). Since modulation need not preserve vanishing moments, however, stronger conditions will have to be imposed to secure analogous L^p-boundedness results for $\mathcal{D}(f,\, g)$. Let $w^{(j)}$ be finite intervals such that

$$\text{(1.7)} \qquad \operatorname{supp} \widehat{\phi}^{(1)} \;\subseteq\; w^{(1)}, \qquad \operatorname{supp} \widehat{\phi}^{(2)} \;\subseteq\; w^{(2)}, \qquad \operatorname{supp} \widehat{\phi}^{(3)} \;\subseteq\; w^{(3)};$$

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.

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

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

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

There is a corresponding sesqui-linear version

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

These definitions do not allow any of the $\phi^{(j)}$ to be the traditional choice of a gaussian, of course, since we have maximized the localization in frequency. The series

$$\sum_{k=-\infty}^{\infty} \Big(\sum_{\ell,\, n=-\infty}^{\infty} s^{k/2}\, c_{k\ell n}\, \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n} \Big)$$

converges in $L^2(\mathbb{R})$ whenever g is an L^2-function and f is a band-limited Schwartz function (*cf.* (2.5)). So by restricting to such functions we can always assume that a time-frequency paraproduct is measurable, finite almost everywhere, and locally integrable. Part II of this series [11] is devoted to a proof of the following result. It requires a delicate 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 [1] [7].

Main Theorem II. *Let $\phi^{(j)}$ be $\mathcal{M}_\mu(\mathbb{R})$-test functions whose Fourier support intervals $w^{(j)}$ have pairwise-disjoint closure. Then the time-frequency paraproduct*

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

is bounded from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p+1/q = 1/r < 3/2$. Furthermore, the operator norm of $\mathcal{D}$ satisfies the inequality

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

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

Examples show that the restriction $r > 2/3$ in Main Theorem II is sharp (*cf.* [15]). Since

$$\overline{\langle\, g,\, \phi^{(2)}_{k\ell n} \,\rangle} \;=\; \lambda_{k\ell n} \,\langle\, g,\, \phi^{(2)}_{k\ell n} \,\rangle, \qquad |\lambda_{k\ell n}| \;=\; 1,$$

corresponding L^p-boundedness results for $\overline{\mathcal{D}}$ follow immediately from those for $\mathcal{D}$.

To represent $\mathcal{C}_{P_\theta}$ in terms of doubly-infinite sum of time-frequency paraproduct

$$\mathcal{D}_{\lambda_1\lambda_2}(f,\, g) \;=\; \sum_{k,\ell,n=-\infty}^{\infty} c_{kn}(\lambda_1, \lambda_2)\, s^{k/2} \,\langle\, f,\, \phi^{(1)}_{k\ell n} \,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n} \,\rangle\, \phi^{(3)}_{k\ell n}$$

with respect to 'mother' wave functions $\phi^{(j)}$ varying with $\lambda_1,\, \lambda_2 \in \mathbb{Z}^2$. The coefficients c_{kn} will be defined in terms of the multiplier $m = m(\xi,\, \eta)$; smoothness of m guarantees decay of the c_{kn}. There is an entirely analogous decomposition of $\overline{\mathcal{C}}_{P_\theta}$. This will done in section 3 by first constructing a Whitney covering of P_θ by squares and then taking Short Time Fourier transform expansions on each square. The key requirements of the $\mathcal{D}_{\lambda_1\lambda_2}$ are readily apparent. For by the triangle inequality (taking $r \geq 1$, for example), Main Theorem II ensures that

$$\|\mathcal{C}_{P_\theta}(f,\, g)\|_r \;\leq\; \text{const.} \Bigg(\sum_{\lambda_1,\lambda_2} \big(\sup_{k,n} |c_{kn}(\lambda_1, \lambda_2)|\big)\, \|\mathcal{D}_{\lambda_1\lambda_2}\|_{op} \Bigg) \|f\|_p\, \|g\|_q.$$

Now (1.2) will guarantee that $\sup_{k,n} |c_{kn}(\lambda_1, \lambda_2)|$ decays as fast as any polynomial in $\lambda_1,\, \lambda_2$, while Main Theorem II controls $\|\mathcal{D}_{\lambda_1\lambda_2}\|_{op}$. In diagonalizing $\mathcal{C}_{P_\theta}$, therefore, it will be crucial to ensure that $\|\mathcal{D}_{\lambda_1\lambda_2}\|_{op}$ increases no faster than some *fixed* polynomial in $\lambda_1,\, \lambda_2$. It is here that translation in time plays a key role. The proof of Main Theorem II proceeds by reducing a general time-frequency paraproduct into ever more simple cases.

In section 4 we establish L^p-boundedness for each fixed frequency n, extending previous results ([3][18, p.287] [5] [20, p. 274]) to the full range $r > 1/2$. More precisely, given $\mathcal{M}_\mu$-test functions $\psi^{(i)}$, their translates and dilates will be defined by

$$\psi^{(i)}_{k\ell}(x) \;=\; s^{k/2}\, \psi^{(i)}(s^k x - a_i \ell), \qquad s = 2^\rho,$$

for fixed positive a_i and rational ρ.

(1.9) Theorem. *Let $\psi^{(i)}$ be $\mathcal{M}_\mu$-test functions at least two of which have vanishing moment. Then the 'standard' paraproduct*

$$\mathcal{P} : f,\, g \longrightarrow \sum_{k,\,\ell\,=\,-\infty}^{\infty} c_{k\ell}\, s^{k/2} \langle\, f,\, \psi^{(1)}_{k\ell}\,\rangle\, \langle\, g,\, \psi^{(2)}_{k\ell}\,\rangle\, \psi^{(3)}_{k\ell}$$

is a bounded operator from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/r = 1/p + 1/q < 2$, whose norm satisfies the inequality

$$\|\mathcal{P}\|_{op} \;\le\; \text{const.}\ \|\pi(a_1)\psi^{(1)}\|\,\|\pi(a_2)\psi^{(2)}\|\,\|\pi(a_3)\psi^{(3)}\|$$

uniformly in the $\psi^{(i)}$.

Underlying a time-frequency paraproduct is an essential structural invariance in translation, modulation and dilation coming from the Schrödinger representation of the so-called Affine-Weyl-Heisenberg group, an extension of the Heisenberg group. This invariance is fundamental to the proof both of Main Theorem I and Main Theorem II. In section 2 by applying the same affine transformation in frequency to all the $\phi^{(j)}$, hence preserving disjointness of their Fourier support intervals, a general time-frequency paraproduct is represented as a finite sum of ones in which

(i) $s = 2^K$ *for some* K *which we are free to specify, and*

(ii) *the* $w^{(j)}$ *all lie in some interval* $(\alpha,\, \alpha + \frac{1}{2})$, $|\alpha| < \frac{1}{2}$, *which either contains the origin or is contained in* $(0,\, 1)$.

Thus the three $w^{(j)}$ can be assumed to lie inside one of the basic intervals

$$(1.10) \qquad (0,\ 1), \quad (M = 1); \qquad \left(-\frac{2^{M-1} - 1}{2^M - 1},\ \frac{2^{M-1} - 1}{2^M - 1}\right), \quad (M > 1)$$

which generate respective grids $\mathcal{W}_M$ in $\mathbb{R}$ via affine transformations $\xi \longrightarrow 2^{Mk}\xi + n$ (*cf.* section 5). In the case $M = 1$ this is just the usual dyadic grid, of course. The value of K is specified in terms of the separation of the $w^{(j)}$; more precisely, $s = 2^{MN}$ where N is chosen so large that in case $M = 1$ there is at least one interval in $\mathcal{W}_1$ of length $1/2^N$ between adjacent $w^{(j)}$ as well as one between each end-point of $[0,\ 1)$ and the nearest $w^{(j)}$, while in case $M > 1$ there are corresponding intervals in $\mathcal{W}_M$ of length $\sim 1/2^{MN}$. Hence in proving Main Theorm II it is enough to begin with time-frequency paraproduct

$$(1.11) \qquad \mathcal{D}(f,\, g) \;=\; \sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, 2^{MNk/2}\, \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}$$

where M is determined by which of the intervals in (1.10) contains all the Fourier support intervals $w^{(j)}$ and

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

Such a time-frequency paraproduct will be said to be $(M,\, N)$-*canonical form.* Fuller details of this reduction and its implications are given in section 5.

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 [19].

2. Preliminaries, Time-Frequency Paraproducts

As norm estimates involving smoothness and decay are needed frequently, it is convenient to work within the setting of test functions and molecules as in [9], for instance. In this terminology a function $\phi = \phi(x)$ is said to be an $\mathcal{M}_\mu$-*test function* when it has continuous derivatives up to order $[\mu]$ such that

$$|D^k \phi(x)| \;\le\; \frac{\text{const.}}{(1+|x|)^{1+\mu+k}}, \qquad 0 \le k \le [\mu]$$

and the inequality

$$|D^{[\mu]}\phi(x+y) - D^{[\mu]}\phi(x)| \;\le\; \text{const.}\, \frac{|y|^{\mu-[\mu]}}{(1+|x|)^{1+2\mu}}$$

holds for all $x, y \in \mathbb{R}$ with $|y| \le \frac{1}{2}(1+|x|)$. The set of all such functions becomes a Banach space, denoted by $\mathcal{M}_\mu(\mathbb{R})$, under the natural weighted supremum norm, where μ is very large but fixed. Given an $\mathcal{M}_\mu$-function ϕ we shall always denote its norm by $\|\phi\|$. Furthermore, wherever this notation occurs, ϕ is to be interpreted as an $\mathcal{M}_\mu$-function since an appropriate subscript will be added to $\|(\cdot)\|$ to indicate the norm on any Banach space other than $\mathcal{M}_\mu(\mathbb{R})$. A function ϕ in $\mathcal{M}_\mu(\mathbb{R})$ is said to be an $\mathcal{M}_\mu$-*molecule* when it has vanishing moments up to order $[\mu]$. The advantages that vanishing moments create will play a key role in this series of papers.

Like each $L^p(\mathbb{R})$-space, $\mathcal{M}_\mu(\mathbb{R})$ is invariant under the respective operations

$$\phi(x) \longrightarrow \phi(x-\lambda), \qquad \phi(x) \longrightarrow s^{1/2}\phi(sx), \qquad \phi(x) \longrightarrow \phi(x)\, e^{2\pi i x\xi} \tag{2.1}$$

of translation, dilation and modulation. Each is bounded, but not uniformly bounded, on $\mathcal{M}_\mu(\mathbb{R})$; in the case of translation, for instance, the inequality

$$\|\phi(\cdot - \lambda)\| \;\le\; \text{const.}(1+|\lambda|)^{1+\mu}\|\phi\| \tag{2.2}$$

holds uniformly in ϕ and λ, and the same inequality holds for dilation and modulation. Together these representations generate the Schrödinger representation of the so-called Affine-Weyl-

Heisenberg group on $\mathcal{M}_\mu(\mathbb{R})$ (*cf.* [8, ch.3], [13]). Given $\mathcal{M}_\mu$-test functions $\varphi^{(j)}$, the set of wave-packets

$$\text{(2.3)(i)} \qquad \varphi^{(j)}_{k\ell n}(x) \;=\; \varphi^{(j)}_Q(x) \;=\; s^{k/2}\varphi^{(j)}(s^k x - a_j\ell)\, e^{2\pi i s^k x b_j n}, \qquad s = 2^\rho,$$

is then the representation of a lattice $\{(s^k,\, a_j\ell,\, b_j n) :\, k,\, \ell,\, n \in \mathbb{Z}\}$ in this group having *mesh size* $\{s,\, a_j,\, b_j\}$ without restriction on a_j and b_j, while

$$\text{(2.3)(ii)} \qquad \sum_{k,\ell,n} c_{k\ell n}\, s^{k/2} \,\langle\, f,\, \varphi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \varphi^{(2)}_{k\ell n}\,\rangle\, \varphi^{(1)}_{k\ell n}, \qquad s = 2^\rho$$

is the sum of matrix coefficients of these representations over lattices with three possibly different mesh sizes. Although we shall avoid the language of representation theory, group-invariance in the form of coordinate changes and modulations will used repeatedly to simplify time-frequency paraproduct. When doing so,

$$\text{(2.4)} \qquad \pi(a) : f(x) \longrightarrow a^{1/2}\, f(ax), \qquad a > 0$$

will denote the unitary action of dilation on $L^2(\mathbb{R})$.

Convergence of the series defining $\mathcal{D}(f,\, g)$ is easily established.

(2.5) Proposition. *Let f be a band-limited Schwartz function. Then the series*

$$\sum_{k=-\infty}^{\infty} \Big(\sum_{\ell,\, n=-\infty}^{\infty} s^{k/2}\, c_{k\ell n} \,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n} \Big)$$

defining a time-frequency paraproduct $\mathcal{D}(f,\, g)$ converges in $L^2(\mathbb{R})$ whenever g is an L^2-function; furthermore, the inequality

$$\Big(\int_{-\infty}^{\infty} |\mathcal{D}(f,\, g)(x)|^2\, dx \Big)^{1/2} \;\le\; \text{const.}\; \|\{c_{k\ell n}\}\|_\infty \Big(\int_{-\infty}^{\infty} |g(x)|^2\, dx \Big)^{1/2}$$

holds uniformly in g and $\{c_{k\ell n}\}$ with constant depending on f and the $\phi^{(j)}$.

By restricting the constants $c_{k\ell n}$ we obtain corresponding results for the time-frequency paraproduct in which the frequencies lie on a half-line, for instance.

Proof. Write

$$\mathcal{D}(f,\, g) \;=\; \Big(\sum_{k\le K} + \sum_{k>K} \Big) \Big(\sum_{\ell,n=-\infty}^{\infty} s^{k/2}\, c_{k\ell n} \,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n} \Big)$$

with K a large positive integer. The first sum can be estimated using well-known L^2-boundedness properties of Gabor frames (*cf.* [6, p.440]). To this end, fix f, g and h in $L^2(\mathbb{R})$. Then

$$\Big|\int_{-\infty}^{\infty}\Big(\sum_{k\le K}\sum_{\ell,n=-\infty}^{\infty} s^{k/2}\, c_{k\ell n}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\,\phi^{(3)}_{k\ell n}(x)\Big)\,\overline{h(x)}\,dx\Big|$$

$$\le\ \text{const.}\ \|\{c_{k\ell n}\}\|_\infty\,\|f\|_2\Big\{\sum_{k\le K} s^{k/2}\Big(\sum_{\ell,n=-\infty}^{\infty}|\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle|\,|\langle\, h,\, \phi^{(3)}_{k\ell n}\,\rangle|\Big)\Big\}.$$

Now after a change of variable

$$\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\ =\ \int_{-\infty}^{\infty}(\pi(s^{-k})f)(y)\,\overline{\phi^{(2)}(y-a_2\ell)}\,e^{-2\pi i y n}\,dy,$$

and correspondingly for h. The result of Daubechies *et al.* [6] thus ensures that the inequality

$$\sum_{\ell,n=-\infty}^{\infty}|\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle|\,|\langle\, h,\, \phi^{(3)}_{k\ell n}\,\rangle|\ \le\ \text{const.}\,\|g\|_2\|h\|_2$$

holds uniformly in k with constant depending on $\phi^{(2)}$ and $\phi^{(3)}$. Hence

$$\Big|\int_{-\infty}^{\infty}\Big(\sum_{k\le K}\sum_{\ell,n=-\infty}^{\infty} s^{k/2}\, c_{k\ell n}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\,\phi^{(3)}_{k\ell n}(x)\Big)\,\overline{h(x)}\,dx\Big|$$

$$\le\ \text{const.}\ s^{K/2}\|\{c_{k\ell n}\}\|_\infty\,\|f\|_2\|g\|_2\|h\|_2.$$

It is in dealing with large values of k that the hypothesis on f is needed since

$$\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\ =\ s^{-k/2}\int_{-\infty}^{\infty}\widehat{f}(\xi)\,\overline{\widehat{\phi}^{(1)}(s^{-k}\xi-n)}\,e^{-2\pi i a_1(s^{-k}\xi-n)\ell}\,d\xi.$$

For if supp $\widehat{f}\subseteq[a,\ b]$ and $w^{(1)}=[\alpha,\beta]$, the wave packet coefficient $\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle$ will vanish unless

$$[a,\ b]\cap[s^k(n+\alpha),\ s^k(n+\beta)]\ \ne\ \emptyset.$$

Consequently, given K large, there exists n_0 (depending on f and $w^{(1)}$) so that this coefficient will be zero for all $k\ge K$ unless $|n|\le n_0$. Thus the sum over large k reduces to one of the form

$$\sum_{|n|\le n_0}\Big(\sum_{k,\ell=-\infty}^{\infty} s^{k/2}\, c_{k\ell n}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\,\phi^{(3)}_{k\ell n}\Big),$$

in other words to a finite sum of 'standard' paraproducts. But each of these will be bounded as a mapping from $\ell^\infty\times L^2(\mathbb{R})$ into $L^2(\mathbb{R})$ since each modulate $\phi^{(j)}(x)e^{2\pi i n x}$ remains an $\mathcal{M}_\mu$-test function with norm depending on n, while disjointness of the Fourier supports ensures that at least two of these modulates has vanishing moment for each fixed n. This completes the proof of the proposition. $\square$

The essential invariance under translation, dilation and modulation is exploited in a number of ways. Firstly, it enables a general time-frequency paraproduct $\mathcal{D}(f,\, g)$ to be written in terms of a finite number of ones having a *canonical form*, and so reduce the proof of L^p-boundedness to these special forms. Let

$$\mathcal{D}(f,\, g) \;=\; \sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, s^{k/2}\, \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}, \qquad s = 2^{\rho} \tag{2.6}$$

be a general time-frequency paraproduct in which $\rho = L_1/L$ and no restrictions are placed on the Fourier support intervals $w^{(j)}$ of the $\phi^{(j)}$ other than the fact they have disjoint closure. After padding by zeros if necessary we can obviously assume that $L_1 = 1$. In addition, since every integer k can be written as $k = \kappa L + \lambda$ with $0 \le \lambda < L$,

$$\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \;=\; 2^{\kappa/2} \int_{-\infty}^{\infty} (\pi(1/2^{\lambda})f)(y)\, \overline{\phi^{(1)}(2^{\kappa}y - a_1 y)}\, e^{2\pi i 2^{\kappa} yn}\, dy.$$

Thus

$$\pi(1/2^{\lambda})\big(\mathcal{D}(f,\, g)\big) \;=\; \sum_{\lambda} \mathcal{D}_{\lambda}(\pi(1/2^{\lambda})f, \pi(1/2^{\lambda})g)$$

where

$$\mathcal{D}_{\lambda}(f,\, g) \;=\; \sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, 2^{k/2}\, \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n} \tag{2.7}$$

and

$$\phi^{(j)}(x) \;=\; 2^{k/2}\phi^{(j)}(2^k x - a_j \ell)\, e^{2\pi i 2^k xn}.$$

On the other hand, after replacing each $\phi^{(j)}$ in $\mathcal{D}_{\lambda}$ by the same modulate $\phi^{(j)}(x)\, e^{2\pi i m x}$ if necessary, we can also assume that the Fourier support intervals of the $\phi^{(j)}$ in (2.7) are disjoint intervals in $(0, 2^d)$ for some sufficiently large choice of integer d. But $n = \nu 2^{d+1} + \gamma$ with $-2^d < \gamma \le 2^d$. Hence if we set

$$\psi^{(j)}_{\gamma} \;=\; \pi\big(1/2^{d+1}\big)\big(\phi^{(j)}(\cdot)\, e^{2\pi i \gamma(\cdot)}\big),$$

then the $\psi^{(j)}_{\gamma}$ have disjoint Fourier support intervals $w^{(j)}_{\gamma}$ such that

$$w^{(j)}_{\gamma} \;\subseteq\; \big(\gamma/2^{d+1},\ \gamma/2^{d+1} + \tfrac{1}{2}\big) \;\subseteq\; \big(-\tfrac{1}{2},\ 1\big);$$

furthermore,

$$\mathcal{D}_{\lambda}(f,\, g) \;=\; \sum_{\gamma} \mathcal{D}_{\lambda\gamma}(f,\, g)$$

where $\mathcal{D}_{\lambda\gamma}$ is the time frequency paraproduct associated with the $\psi^{(j)}_{\gamma}$. Finally, the same coset argument used to pass from $\rho = 1/L$ to $\rho = 1$ can be used again on each $\mathcal{D}_{\lambda\gamma}$ to pass from $\rho = 1$ to $\rho = K$ for any choice of positive integer K. As dilation and modulation are bounded on all L^p-spaces we thus obtain the following result, taking $\alpha = \gamma/2^{d+1}$.

(2.8) Theorem. *In proving Main Theorem II it is enough to assume that $\mathcal{D}(f,\,g)$ is a time-frequency paraproduct in which $\rho = K$, where K is a positive integer which can be chosen freely, and the Fourier support intervals of the $\phi^{(j)}$ all lie in an interval $(\alpha,\,\alpha+\frac{1}{2})$ containing the origin or lying inside* $(0,1)$.

The reason for restricting to the particular time frequency paraproduct in (2.8) is that we shall then be able to link the Fourier support intervals with grid structures in frequency (*cf.* also section 5). This link will become crucial in Part II (*cf.* [11]).

Invariance also allows Main Theorem II to be applied on occasion to sums such as those in (2.3)(ii) even though the $\varphi^{(j)}_{k\ell n}$ may contain modulations in which $b_j \neq 1$. Set

$$\mathcal{D}^{(\varphi)}(f,\,g) \;=\; \sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, s^{k/2}\,\langle\, f,\, \varphi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \varphi^{(2)}_{k\ell n}\,\rangle\, \varphi^{(3)}_{k\ell n}\,, \qquad s = 2^{\rho} \tag{2.9}$$

where the $\varphi^{(j)}_{k\ell n}$ are defined by (2.3)(i), allowing $b_j \neq 1$. Dilation eliminates the b_j from the wave packets in $\mathcal{D}^{(\varphi)}$. For after a change of variable,

$$\begin{aligned} s^{k/2}\int_{-\infty}^{\infty} (\pi(b_1)f)(x)\,\overline{\varphi^{(1)}(s^k x - a_1\ell)}\,e^{-2\pi i s^k b_1 n x}dx \\ =\; s^{k/2}\int_{-\infty}^{\infty} f(x)\,\overline{(\pi(1/b_1)\varphi^{(1)})(s^k x - a_1 b_1\,\ell)}\,e^{-2\pi i s^k n x}dx. \end{aligned}$$

Consequently, if we define $\phi^{(j)}$ by

$$\phi^{(j)}(x) \;=\; (\pi(1/b_j)\varphi^{(j)})(x) \;=\; b_j^{-1/2}\varphi^{(j)}(x/b_j)$$

and corresponding wave packets $\phi^{(j)}_{k\ell n}$ by

$$\phi^{(j)}_{k\ell n}(x) \;=\; s^{k/2}\phi^{(j)}(s^k x - a_j b_j \ell)\, e^{2\pi i s^k n x},$$

then

$$D^{(\varphi)}(\,\pi(b_1)f,\; \pi(b_2)g) \;=\; \pi(b_3)\Bigl(\sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, s^{k/2}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}\Bigr).$$

On the other hand, dilation ensures that

$$\mathrm{supp}\,\widehat{\varphi}^{(1)} \;\subseteq\; [\xi_0,\;\xi_1] \quad\Longrightarrow\quad \mathrm{supp}\;\widehat{\phi}^{(1)} \;\subseteq\; [\xi_0\,/b_1,\;\xi_1\,/b_1],$$

and correspondingly for $\widehat{\varphi}^{(2)}$. As dilation is bounded on all L^p-spaces, the following result is an immediate corollary of Main Theorem II.

(2.10) Theorem. *The operator $\mathcal{D}^{(\varphi)}$ in (2.9) associated with wave packets*

$$\varphi^{(j)}_{k\ell n}(x) \;=\; s^{k/2}\varphi^{(j)}(s^k x - a_j\ell)\, e^{2\pi i s^k b_j n x}$$

is bounded from $\ell^\infty \times L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$, $1/p + 1/q = 1/r < 3/2$ provided the Fourier support intervals $w^{(j)}$ of the dilates $\pi(1/b_j)\varphi^{(j)}$ have pairwise-disjoint closure. Furthermore, the operator norm of $\mathcal{D}^{(\varphi)}$ satisfies the inequality

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

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

Specific applications of (2.10) will arise in the next section.

3. Diagonalization of Cone Operators

In this section $\mathcal{M}_\mu$-test functions $\psi^{(j)}$ will be chosen so that $\mathcal{C}_{P_\theta}$ can be represented as a doubly-infinite sum

(3.1) $$\mathcal{C}_{P_\theta}(f,\, g) \;=\; \sum_{\lambda,\,\lambda_2\,=\,-\infty}^{\infty} \mathcal{D}^{(\varphi)}_{\lambda_1\lambda_2}(f,\, g)$$

of functions

$$\mathcal{D}^{(\varphi)}_{\lambda_1\lambda_2}(f,\, g) \;=\; \sum_{k,\ell,n\,=-\infty}^{\infty} c_{kn}(\lambda_1,\lambda_2)\, s^{k/2}\,\langle\, f,\, \varphi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \varphi^{(2)}_{k\ell n}\,\rangle\, \varphi^{(3)}_{k\ell n}$$

in which

(3.2) $$\varphi^{(j)}(x) \;=\; \psi^{(j)}(x + a_j\lambda_j), \quad (j = 1,\, 2); \qquad \varphi^{(3)}(x) \;=\; \psi^{(3)},$$

and the wave packets $\varphi^{(j)}_{k\ell n}$ are defined by

$$\varphi^{(j)}_{k\ell n}(x) \;=\; s^{k/2}\varphi^{(j)}(s^k x - a\,\ell)\, e^{2\pi i s^k b_j n x}$$

for a fixed choice of positive constants a_j, b_j and a independently of λ_1, λ_2. The coefficients c_{kn} will satisfy the inequality

(3.3) $$|c_{kn}(\lambda_1,\lambda_2)| \;\le\; \text{const.}\left(\frac{1}{1 + |\lambda_1| + |\lambda_2|}\right)^{|\alpha|}$$

uniformly in k, n for each multi-index α because of smoothness condition (1.2). There are two crucial points to note.

- *Property (3.2) forces the $\varphi^{(j)}$ to have the same Fourier support interval as $\psi^{(j)}$ for each j, independently of λ_1, λ_2. In turn this guarantees that the dilates $\pi(1/b_j)\varphi^{(j)}$ the $\phi^{(j)}$ too have Fourier support intervals independent of λ_1, λ_2* for each j.

- *The construction also ensures that the $\varphi^{(j)}$ have disjoint Fourier support intervals which remain disjoint after dilation $\varphi^{(j)} \longrightarrow \phi^{(j)} = \pi(1/b_j)\varphi^{(j)}$, guaranteeing that (2.10) can be applied to each $\mathcal{D}^{(\varphi)}_{\lambda_1\lambda_2}$.*

There is a corresponding representation of $\overline{\mathcal{C}}_{P_\theta}$. Granted these, theorem (1.5) follows quickly.

Proof of Theorem (1.5). By the triangle inequality

$$\|\mathcal{C}_{P_\theta}\|_r \ \le\ \text{const.}\Big(\sum_{\lambda_1,\lambda_2=-\infty}^{\infty} \sup_{k,n} |c_{kn}(\lambda_1,\lambda_2)|\, \|\mathcal{D}_{\lambda_1\lambda_2}\|_{op}\Big)\|f\|_p\,\|g\|_q.$$

when $r \ge 1$, while for $r < 1$

$$\|\mathcal{C}_{P_\theta}\|_r \ \le\ \text{const.}\Big(\sum_{\lambda_1,\lambda_2=-\infty}^{\infty} \big(\sup_{k,n} |c_{kn}(\lambda_1,\lambda_2)|\, \|\mathcal{D}_{\lambda_1\lambda_2}\|_{op}\big)^r\Big)^{1/r}\|f\|_p\,\|g\|_q.$$

On the other hand, by (2.10) and (3.2),

$$\|\mathcal{D}_{\lambda_1\lambda_2}\|_{op} \ \le\ \text{const.}\ P\big(\|\psi^{(1)}(\cdot + a_1\lambda_1)\|,\, \|\psi^{(2)}(\cdot + a_2\lambda_2)\|,\, \|\psi^{(3)}\|\big)$$

uniformly in λ_1, λ_2, so

$$\|\mathcal{C}_{P_\theta}(f,\, g)\|_r \ \le\ \text{const.}\ \|f\|_p\,\|g\|_q,$$

using (2.2) and (3.3). An entirely analogous argument takes care of $\overline{\mathcal{C}}_{P_\theta}$, completing the proof of theorem (1.5). $\square$

To 'diagonalize' $\mathcal{C}_{P_\theta}$ fix $\theta \in (0,\, \pi/4]$ and recall that P_θ is the half-plane $\{(\xi,\,\eta) : \xi\tan\theta - \eta > 0\}$. The basic idea is to generate a Whitney covering $\{R_{kn}\}$ of P_θ by translating and dilating a single square R; the $\psi^{(j)}$ then arise as smooth bump functions associated with R. Choose $L \ge 8$, and let $\Theta = \Theta(\xi)$ be a $\mathcal{C}^\infty$-bump function which generates a partition of unity for $(0,\,\infty)$ in the sense that

$$\operatorname{supp}\Theta \subseteq \big(L-1,\ L+1\big); \qquad \sum_{k=-\infty}^{\infty} \Theta(s^{-k}\,\xi) \ =\ \chi_{(0,\infty)}(\xi), \quad s = 2^{1/L}.$$

Then $\Theta(\xi\tan\theta - \eta)$ has support in the strip

$$S_\theta \ =\ \big\{(\xi,\,\eta) :\ L-1 \ <\ \xi\tan\theta - \eta \ <\ L+1\big\}$$

lying inside P_θ, while

$$\sum_{k=-\infty}^{\infty} \Theta(s^{-k}(\xi\tan\theta - \eta)) \ \equiv\ 1, \qquad (\xi,\,\eta) \in P_\theta,$$

provides a partition of unity of P_θ. Consequently, the decomposition

$$\mathcal{C}_{P_\theta}(f,\,g)(x) \ =\ \sum_{k=-\infty}^{\infty}\Big(\int_{P_\theta} \Theta(s^{-k}(\xi\tan\theta - \eta))\, m(\xi,\,\eta)\, \widehat{f}(\xi)\,\widehat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi d\eta\Big)$$

localizes $\mathcal{C}_{P_\theta}$ smoothly to the strip S_θ and its dilates. To construct the Whitney covering of P_θ set

$$w^{(1)} \;=\; \left\{\xi : \left|\xi - \frac{L}{1+\tan\theta}\right| \;\le\; 2\right\}, \qquad w^{(2)} \;=\; \left\{\xi : \left|\xi + \frac{L}{1+\tan\theta}\right| \;\le\; 2\right\}$$

and $w^{(3)} \;=\; \{\xi \;:\; |\xi| \;\le\; 1\}$. The conditions $L \ge 8$ and $0 < \theta \le \pi/4$ ensure that the $w^{(j)}$ are pairwise-disjoint. Now set $R = w^{(1)} \times w^{(2)}$; translations and dilations of R will provide the required covering of P_θ. Indeed, set

$$b_1 \;=\; \frac{1}{1+\tan\theta}, \qquad b_2 \;=\; \frac{\tan\theta}{1+\tan\theta}, \qquad b_3 = b_1 + b_2 = 1, \tag{3.4}$$

so that $(\xi,\,\eta) \longrightarrow (\xi + b_1 n,\, \eta + b_2 n)$ fixes both S_θ and P_θ. Then the translates and dilates

$$R_{kn} \;=\; \{(s^k(\xi + b_1 n),\, s^k(\eta + b_2 n)) :\; (\xi,\,\eta) \in R\}, \qquad -\infty < k, n < \infty,$$

of R cover P_θ; consequently,

$$\bigcup_{k,n=-\infty}^{\infty} R_{kn} \;=\; P_\theta, \qquad \mathrm{dist}(R_{kn},\, \partial P_\theta) \;\sim\; s^k.$$

To exploit this geometry first choose a function $\psi^{(3)}$ whose Fourier transform is a C^∞-bump function such that $\mathrm{supp}\,\widehat{\psi}^{(3)} \;\subseteq\; w^{(3)}$ and

$$\sum_{n=-\infty}^{\infty} |\widehat{\psi}^{(3)}(\xi - n)|^2 \;\equiv\; 1, \qquad \xi \in \mathbb{R}.$$

The function

$$\sigma(\xi,\,\eta) \;=\; \Theta(\xi\tan\theta - \eta)\,\overline{\widehat{\psi}^{(3)}(\xi + \eta)}$$

thus has support in the parallelogram $S_\theta \cap \{(\xi,\,\eta) : \xi + \eta \in w^{(3)}\}$; in particular, σ has support in the square $R = w^{(1)} \times w^{(2)}$. Consequently,

$$\mathcal{C}_{P_\theta}(f,\,g)(x) \;=\; \sum_{k,n=-\infty}^{\infty} s^k\, \mathcal{C}_{kn}(f,\,g)(x)\, e^{2\pi i s^k n x}$$

is a smooth localization of $\mathcal{C}_{P_\theta}$ to the squares R_{kn}, setting

$$\mathcal{C}_{kn}(f,\,g)(x) = \;\frac{1}{s^k}\int_{R_{kn}} m(\xi,\,\eta)\,\sigma(s^{-k}\xi - b_1 n,\, s^{-k}\eta - b_2 n)$$
$$\times\;\; \widehat{\psi}^{(3)}(s^{-k}\xi + s^{-k}\eta - n)\,\widehat{f}(\xi)\,\widehat{g}(\eta)\,e^{2\pi i x(\xi+\eta-s^k n)}\,d\xi d\eta.$$

By changing variables we can also express $\mathcal{C}_{kn}$ as an integral

$$\mathcal{C}_{kn}(f,\,g)(x) = \;s^k\Bigg(\int_R m(s^k(\xi + b_1 n),\, s^k(\eta + b_2 n))\,\sigma(\xi,\,\eta)\,\widehat{\psi}^{(3)}(\xi + \eta)$$
$$\times\;\; \widehat{f}(s^k(\xi + b_1 n))\,\widehat{g}(s^k(\eta + b_2 n))\,e^{2\pi i s^k x(\xi+\eta)}\,d\xi d\eta\Bigg)$$

over R. The still finer decomposition of $\mathcal{C}_{P_\theta}$ stems from Short Time Fourier expansions on R. Choose functions $\psi^{(1)}$, $\psi^{(2)}$ whose Fourier transforms are C^∞-bump functions such that $\operatorname{supp}\widehat{\psi}^{(j)} \subseteq w^{(j)}$, $j = 1, 2$, and

$$\overline{\widehat{\psi}^{(1)}(\xi)}\,\overline{\widehat{\psi}^{(2)}(\eta)}\,\sigma(\xi, \eta) \;=\; \sigma(\xi, \eta).$$

Set $a_j \;=\; 1/|w^{(j)}|$. Then on R

$$\widehat{f}(s^k(\xi + b_1 n))\,\overline{\widehat{\psi}^{(1)}(\xi)} \;=\; \frac{1}{s^{k/2}|w^{(1)}|}\bigg(\sum_{\ell_1=-\infty}^{\infty} \langle\, f,\, \psi^{(1)}_{k\ell_1 n}\,\rangle\, e^{-2\pi i a_1 \ell_1 \xi}\bigg),$$

while

$$\widehat{g}(s^k(\eta + b_2 n))\,\overline{\widehat{\psi}^{(2)}(\eta)} \;=\; \frac{1}{s^{k/2}|w^{(2)}|}\bigg(\sum_{\ell_2=-\infty}^{\infty} \langle\, g,\, \psi^{(2)}_{k\ell_2 n}\,\rangle\; e^{-2\pi i a_2 \ell_2 \eta}\bigg)$$

where

$$\psi^{(j)}_{k\ell_j n}(x) \;=\; s^{k/2}\psi^{(j)}(s^k x - a_j\,\ell_j)\,e^{2\pi i s^k x b_j n}, \qquad j = 1, 2. \tag{3.5}$$

Substituting these expansions into the integral for $\mathcal{C}_{kn}$ we see that

$$\mathcal{C}_{kn}(f, g)(x) \;=\; \sum_{\ell_1,\ell_2=-\infty}^{\infty} \langle\, f,\, \psi^{(1)}_{k\ell_1 n}\,\rangle\,\langle\, g,\, \psi^{(2)}_{k\ell_2 n}\,\rangle\, C_{\ell_1\ell_2}(x)$$

with

$$C_{\ell_1\ell_2}(x) \;=\frac{1}{|R|}\int_R m(s^k(\xi + b_1 n),\, s^k(\eta + b_2 n))\,\sigma(\xi, \eta)\,\widehat{\psi}^{(3)}(\xi + \eta)$$
$$\times\; e^{-2\pi i a_1 \ell_1 \xi}\, e^{-2\pi i a_2\,\ell_2\,\eta}\, e^{2\pi i s^k x(\xi+\eta)}\, d\xi d\eta.$$

But

$$\widehat{\psi}^{(3)}(\xi)\,e^{2\pi i s^k x \xi} \;=\; \frac{1}{|w^{(3)}|}\bigg(\sum_{\ell_3=-\infty}^{\infty} \psi_3(s^k x - a_3\ell_3)\,e^{2\pi i a_3 \ell_3 \xi}\bigg)$$

on $w^{(3)}$, setting $a_3 = 1/|w^{(3)}|$. To complete the decomposition of $\mathcal{C}_{kn}$ it remains to compute the inverse Fourier transform of the smooth 'localizations' of the multiplier m; smoothness condition (1.2) then controls the decay of these inverse Fourier transforms. More precisely, when $M_{kn} = M_{kn}(x, y)$ is defined by

$$M_{kn}(x, y) \;=\; \frac{1}{|R|}\bigg(\int_R m(s^k(\xi + b_1 n),\, s^k(\eta + b_2 n))\,\sigma(\xi, \eta)\,e^{2\pi i(x\xi + y\eta)}\,d\xi d\eta\bigg)$$

a simple Fourier transform argument together with (1.2) shows that the inequality

$$|M_{kn}(x, y)| \;\le\; \text{const.}\bigg(\frac{1}{1 + |x| + |y|}\bigg)^{|\alpha|} \tag{3.6}$$

holds uniformly in k and n for each multi-index α. Furthermore,

$$\mathcal{C}_{kn}(f,\, g)(x) \;=\; \frac{1}{|w^{(3)}|}\Bigg(\sum_{\ell_1,\ell_2,\ell_3=-\infty}^{\infty} M_{kn}(a_3\ell_3 - a_1\ell_1,\, a_3\ell_3 - a_2\ell_2) \times \langle\, f,\, \psi^{(1)}_{k\ell_1 n}\,\rangle\, \langle\, g,\, \psi^{(2)}_{k\ell_2 n}\,\rangle\, \psi^{(3)}(s^k x - a_3\ell_3)\Bigg).$$

All that remains is to replace the sum over all integers ℓ_1, ℓ_2, and ℓ_3 with a sum over integers $\ell, \lambda_1, \lambda_2$. Set

$$\ell_3 \;=\; \ell, \qquad \lambda_1 = \frac{a_3}{a_1}\,\ell - \ell_1, \qquad \lambda_2 = \frac{a_3}{a_2}\,\ell - \ell_2.$$

Then $a_1/a_2 = a_3/a_1 = 2$; furthermore, in view of (3.5),

$$\psi^{(j)}_{k\ell_j n}(x) \;=\; s^{k/2}\psi^{(j)}(s^k x - a_3\ell + a_j\lambda_j)\, e^{2\pi i s^k x b_j n}, \qquad j = 1,\, 2. \tag{3.7}$$

We have now assembled all the ingredients necessary for (3.1). Set

$$\varphi^{(j)}(x) \;=\; \psi^{(j)}(x + a_j\lambda_j), \quad (j = 1,\, 2); \qquad \varphi^{(3)}(x) \;=\; \psi^{(3)}(x)$$

with $a_j = 1/|w^{(j)}|$; clearly $\operatorname{supp}\widehat{\varphi}^{(j)} = \operatorname{supp}\widehat{\psi}^{(j)} \subseteq w^{(j)}$ whatever the value of λ_1, λ_2 or j. Next, in view of (3.5) and (3.7), take $a = a_3$ and define wave packets $\varphi^{(j)}_{k\ell n}$ by

$$\varphi^{(j)}_{k\ell n}(x) \;=\; s^{k/2}\varphi^{(j)}(s^k x - a\,\ell)\, e^{2\pi i s^k b_j n x}, \qquad s = 2^{1/L},$$

with the b_j being specified in (3.4). Then

$$\mathcal{C}_{P_\theta}(f,\, g) \;=\; \sum_{k,n=-\infty}^{\infty} s^k\, \mathcal{C}_{kn}(f,\, g)\, e^{2\pi i s^k n x} \;=\; \sum_{\lambda_1,\, \lambda_2} \mathcal{D}^{(\varphi)}_{\lambda_1\lambda_2}(f,\, g)$$

setting

$$\mathcal{D}^{(\varphi)}_{\lambda_1\,\lambda_2}(f,\, g) \;=\; \sum_{k,\,\ell,\,n\,=-\infty}^{\infty} \frac{1}{|w^{(3)}|} M_{kn}(a_1\lambda_1, a_2\lambda_2)\, s^{k/2}\, \langle\, f,\, \varphi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \varphi^{(2)}_{k\ell n}\,\rangle\, \varphi^{(3)}_{k\ell n}.$$

Finally, to check that the Fourier support intervals remain disjoint after dilation $\varphi^{(j)} \to \pi(1/b_j)\varphi^{(j)}$, note that $0 < b_1,\, b_2 < 1$, while $b_3 = 1$. Thus $w^{(3)}$ is unchanged; on the other hand, $w^{(1)}$ lies in $(0,\, \infty)$, say $w^{(1)} \;=\; [\xi_0,\, \xi_1]$, and lies to the right of $w^{(3)}$ because the latter contains the origin. Since

$$\operatorname{supp}(\pi(1/b_1)\varphi^{(1)})\widehat{\ } \subseteq\; [\xi_0\,/b_1,\, \xi_1\,/b_1], \qquad 0 < b_1 < 1,$$

the Fourier support of $\pi(1/b_1)\varphi^{(1)}$ will be disjoint from that of $\pi(1/b_3)\varphi^{(3)}\,(= \varphi^{(3)})$. The same argument applies to $\pi(1/b_2)\varphi^{(2)}$ because $w^{(2)}$ lies in $(-\infty,\, 0)$ to the left of $w^{(3)}$.

The corresponding representation for $\overline{\mathcal{C}}_{P_\theta}$ is obtained in exactly the same way except for changes in the geometry made necessary by the presence of the term $\xi - \eta$ in $\overline{\mathcal{C}}_{P_\theta}$ instead of the corresponding

$\xi + \eta$ in $\mathcal{C}_{P_\theta}$. In fact, this is why this θ has to be restricted to the range $0 < \theta < \pi/4$. Fix such a θ and choose any integer L with $L = 2^K$ for some integer K large enough so that $L > 4/(1 - \tan\theta)$; in particular, L becomes increasingly large as $\theta \to \pi/4$. As before, let $\Theta = \Theta(\xi)$ be a C^∞-bump function so that $\operatorname{supp}\Theta \subseteq \big(L-1,\ L+1\big)$ and

$$\sum_{k=-\infty}^{\infty} \Theta(s^{-k}\,\xi) \;=\; \chi_{(0,\infty)}(\xi), \qquad s = 2^{1/L}.$$

Then again

$$\overline{\mathcal{C}}_{P_\theta}(f,\,g)(x) \;=\; \sum_{k=-\infty}^{\infty}\left(\int_{P_\theta} \Theta(s^{-k}(\xi\tan\theta - \eta))\,m(\xi,\,\eta)\,\widehat{f}(\xi)\,\overline{\widehat{g}(\eta)}\,e^{2\pi i x(\xi-\eta)}\,d\xi d\eta\right)$$

localizes $\overline{\mathcal{C}}_{P_\theta}$ smoothly to the strip S_θ and its dilates. Now set

$$w^{(1)} \;=\; \left\{\xi : \left|\xi - \frac{L}{1+\tan\theta}\right| \;\le\; \tfrac{1}{2}L\right\}, \qquad w^{(2)} \;=\; \left\{\xi : \left|\xi + \frac{L}{1+\tan\theta}\right| \;\le\; \tfrac{1}{2}L\right\}$$

and

$$w^{(3)} \;=\; \left\{\xi : \left|\xi - \frac{2L}{1+\tan\theta}\right| \;\le\; 1\right\}.$$

The conditions on L and θ ensure that the $w^{(j)}$ are pairwise disjoint intervals. Again we let $R = w^{(1)} \times w^{(2)}$, but now the geometry becomes fundamentally different because $\xi + \eta$ has been replaced by $\xi - \eta$. Set

$$b_1 \;=\; \frac{1}{1-\tan\theta}, \qquad b_2 \;=\; \frac{\tan\theta}{1-\tan\theta}, \qquad b_3 = b_1 - b_2 = 1,$$

so that $(\xi,\,\eta) \longrightarrow (\xi + b_1 n,\, \eta + b_2 n)$ fixes S_θ and P_θ. Then the squares

$$R_{kn} \;=\; \big\{(s^k(\xi + b_1 n),\, s^k(\eta + b_2 n)) :\ (\xi,\,\eta) \in R\big\}, \qquad -\infty < k, n < \infty,$$

provide a Whitney covering of P_θ in the sense that

$$\bigcup_{k,n=-\infty}^{\infty} R_{kn} \;=\; P_\theta, \qquad \operatorname{dist}(\partial P_\theta,\, R_{kn}) \;\sim\; s^k.$$

To exploit this new geometry let $\psi^{(3)}$ be a function whose Fourier transform is a C^∞-bump function such that $\operatorname{supp}\widehat{\psi}^{(3)} \;\subseteq\; w^{(3)}$, while

$$\sum_{n=-\infty}^{\infty} |\widehat{\psi_3}(\xi - n)|^2 \;\equiv\; 1, \quad (\xi \in \mathbb{R}).$$

Then the function

$$\sigma(\xi,\,\eta) \;=\; \Theta(\xi\tan\theta - \eta)\,\overline{\widehat{\psi}^{(3)}(\xi - \eta)}$$

has support in the parallelogram $S_\theta \cap \{(\xi, \eta) : \xi - \eta \in w^{(3)}\}$, and hence also in the rectangle $R = w^{(1)} \times w^{(2)}$. From here-on with these new definitions the construction is exactly the same as for the previous case save for the fact that $\widehat{g}$ is now replaced by $\overline{\widehat{g}}$. Thus $\overline{\mathcal{C}}_{P_\theta}$ admits a doubly-infinite sum representation

$$\overline{\mathcal{C}}_{P_\theta}(f, g) \;=\; \sum_{\lambda_1, \lambda_2 = -\infty}^{\infty} \overline{\mathcal{D}}^{(\varphi)}_{\lambda_1\lambda_2}(f, g)$$

of time-frequency paraproduct

$$\overline{\mathcal{D}}_{\lambda_1\lambda_2}(f, g) \;=\; \sum_{k,\ell,n=-\infty}^{\infty} c_{kn}(\lambda_1, \lambda_2)\, s^{k/2} \langle f, \varphi^{(1)}_{k\ell n} \rangle \overline{\langle g, \varphi^{(2)}_{k\ell n} \rangle}\, \varphi^{(3)}_{k\ell n}$$

with $s = 2^{1/L}$. We omit the details.

4. L^p-boundedness for 'Standard' Paraproducts

In this section we start down the path to Main Theorem II by proving the preliminary theorem (1.9) establishing L^p-boundedness of a 'standard' paraproduct

$$\mathcal{P} : f, g \longrightarrow \sum_{k,\, \ell = -\infty}^{\infty} c_{k\ell}\, s^{k/2} \langle f, \psi^{(1)}_{k\ell} \rangle \langle g, \psi^{(2)}_{k\ell} \rangle\, \psi^{(3)}_{k\ell}$$

associated with functions

$$\psi^{(i)}_{k\ell}(x) \;=\; s^{k/2}\, \psi^{(i)}(s^k x - a_i \ell) \;=\; \psi^{(i)}_I(x), \qquad s = 2^\rho$$

in which at least two $\psi^{(j)}$ have vanishing moment; thus extending previous results for 'standard' paraproducts (*cf.* [3][18, p.287] [5] [20, p. 274]). In view of the reduction arguments in section 2 it is enough to establish the corresponding weak type estimate

$$|\{x : \; |\mathcal{P}(f, g)(x)| \,\ge\, 2\gamma\}| \;\le\; C_\psi \left(\frac{\|f\|_p \|g\|_q}{\gamma} \right)^r \tag{4.1}$$

for each $\gamma > 0$ assuming $s = 2$, $a_j = 1$, and $I = [2^{-k}\ell, 2^{-k}(\ell + 1))$ a dyadic interval; here C_ψ will denote a constant satisfying

$$C_\psi \;\le\; \text{const.}\,(1 \,+\, \|\psi^{(1)}\|\|\psi^{(2)}\|\|\psi^{(3)}\|)$$

where the constant on the right may change but it will always be independent of the ψ. There are three steps in the proof.

Step 1. Choose $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ and $\{c_I\} \in \ell^\infty$, $\|\{c_I\}\|_\infty = 1$. The first step is reminiscent of the familiar Calderón-Zygmund decomposition. Fix $\gamma > 0$ and set

$$E_{bad} \;= \{x : M_p(Mf)(x) > \kappa_p\} \,\cup\, \{x : M_q(Mg)(x) > \kappa_q\}$$

where

$$\kappa_p \;=\; \left(\frac{\|f\|_p^{1/q}\,\gamma^{1/p}}{\|g\|_q^{1/p}} \right)^r, \qquad \kappa_q \;=\; \left(\frac{\|g\|_q^{1/p}\,\gamma^{1/q}}{\|f\|_p^{1/q}} \right)^r.$$

With these choices

$$|E_{bad}| \;\le\; \text{const.}\left(\frac{\|f\|_p\,\|g\|_q}{\gamma} \right)^r \tag{4.2}$$

since

$$\left(\frac{\|f\|_p}{\kappa_p} \right)^p \;=\; \left(\frac{\|f\|_p\,\|g\|_q}{\gamma} \right)^r \;=\; \left(\frac{\|g\|_q}{\kappa_q} \right)^q, \qquad \kappa_p\kappa_q \;=\; \gamma.$$

As a *function*, $\mathcal{P}(f,g)$ can now be decomposed into 'bad' and 'good' functions

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

by setting

$$\mathcal{P}_{bad}(f,\,g) \;=\; \sum_{I\,\subseteq\,E_{bad}} c_{k\ell}\,2^{k/2}\,\langle\, f,\,\psi^{(1)}_{k\ell}\,\rangle\,\langle\, g,\,\psi^{(2)}_{k\ell}\,\rangle\,\psi^{(3)}_{k\ell}$$

i.e., summing over dyadic intervals contained wholly within E_{bad}. The $\phi_I^{(i)}$ appearing in $\mathcal{P}_{bad}(f,\,g)$ are 'concentrated' inside E_{bad}, so the bad function can be estimated sufficiently far away from E_{bad} using solely decay estimates. Set

$$E_1 \;=\; \bigcup_{I\,\subseteq\,E_{bad}} 4I$$

where AI denotes the interval centered at I of length $A|I|$.

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

$$\frac{1}{\gamma}\int_{\mathbb{R}\setminus E_1} |\mathcal{P}_{bad}(f,\,g)(x)|\,dx \;\le\; C_\psi\,|E_{bad}|$$

hold uniformly in f,g and γ, provided $\mu > 1$.

Granted (4.3) it follows that

$$|\{x : |\mathcal{P}_{bad}(f,\,g)(x)| \;>\; \gamma\,\}| \;\le\; C_\psi\left(\frac{\|f\|_p\,\|g\|_q}{\gamma} \right)^r,$$

leaving only the proof of the corresponding estimate for the good function

$$\mathcal{P}_{good}(f,\,g) \;=\; \sum_{I\,\subseteq\,E_{bad}} c_{k\ell}\,2^{k/2}\,\langle\, f,\,\psi^{(1)}_{k\ell}\,\rangle\,\langle\, g,\,\psi^{(2)}_{k\ell}\,\rangle\,\psi^{(3)}_{k\ell}.$$

Proof of (4.3).To estimate $|E_1|$ let $J_1,\, J_2,\, \ldots$ be maximal, hence disjoint, dyadic intervals in E_{bad}. Then

$$|E_1| \;\le\; 4\sum_j |J_j| \;\le\; \text{const.}\, |E_{bad}|.$$

To estimate $\mathcal{P}_{bad}$ let I be an arbitrary dyadic interval, not necessarily contained in E_{bad} for the moment. Then the inequality

$$\begin{aligned}\int_{\mathbb{R}\setminus 2^m I} \frac{1}{\sqrt{|I|}} |\langle\, f,\, \psi_I^{(1)}\,\rangle \,\langle\, g,\, \psi_I^{(2)}\,\rangle\, \psi_I^{(3)}(x)|\, dx \\ \le\; C_\psi\, \frac{|I|}{2^{m\mu}} \Big(\inf_{x\,\in\, I} Mf(x)\Big)\Big(\inf_{x\,\in\, I} Mg(x)\Big)\end{aligned} \tag{4.4}$$

holds uniformly in $m,\ m > 1$. Indeed, the Hardy-Littlewood maximal function controls the coefficients in the sense that

$$\frac{1}{\sqrt{|I|}}\, |\langle\, f,\, \psi^{(1)}\,\rangle| \;\le\; \text{const.}\, \|\psi^{(1)}\| \Big(\inf_{x\,\in\, I} M(f)(x)\Big), \tag{4.5}$$

and similarly for g, irrespective of vanishing moments. On the other hand,

$$\frac{1}{\sqrt{|I|}} \int_{\mathbb{R}\setminus 2^m I} |\psi_I^{(3)}(x)|\, dx \;\le\; \text{const.}\, \frac{|I|}{2^{m\mu}} \|\psi^{(3)}\|.$$

This establishes (4.4). The presence of the factor 2^m allows (4.4) to be extended to all dyadic intervals I in a given dyadic interval J. Fix $k \ge 0$ and let I be any interval in J with $|I| = 2^{-k}|J|$. Then

$$|I| \;=\; 2^{-k}|J| \quad\Longrightarrow\quad J \;\subseteq\; 2^{k+2} I \;\subseteq\; 4J.$$

Because of the last of these inclusions,

$$\begin{aligned}\int_{\mathbb{R}\setminus 4J} \Big|\, \frac{1}{\sqrt{|I|}} |\langle\, f,\, \psi_I^{(1)}\,\rangle \,\langle\, g,\, \psi_I^{(2)}\,\rangle\, \psi_I^{(3)}(x)\big|\, dx \\ \le\; C_\psi\, \frac{|I|}{2^{k\mu}} \Big(\inf_{x\,\in\, I} Mf(x)\Big)\Big(\inf_{x\,\in\, I} Mg(x)\Big).\end{aligned}$$

But by the first of these inclusions, the inequality

$$\inf_{x\,\in\, I} Mf(x) \;\le\; \text{const.}\, 2^k \inf_{x\,\in\, J} Mf(x)$$

together with the corresponding one for g always holds. Thus, summing first over all $I \subseteq J$, $|I| = 2^{-k}|J|$, and then over all $k \ge 0$, we see that

$$\begin{aligned}\int_{\mathbb{R}\setminus 4J} \Big|\sum_{I\,\subseteq\, J} c_I\, \frac{1}{\sqrt{|I|}} |\langle\, f,\, \psi_I^{(1)}\,\rangle \,\langle\, g,\, \psi_I^{(2)}\,\rangle\, \psi_I^{(3)}(x)\big|\, dx \\ \le\; C_\psi\, |J| \Big(\inf_{x\,\in\, \mathbb{J}} Mf(x)\Big)\Big(\inf_{x\,\in\, \mathbb{J}} Mg(x)\Big).\end{aligned} \tag{4.6}$$

Now, finally, let $J_1, J_2, \ldots$ be the same maximal dyadic intervals in E_{bad} as before. Maximality ensures that the next larger dyadic interval to J_j is not contained in E_{bad}, which in turn ensures that $4J_j \not\subseteq E_{bad}$. Consequently, for such J_j

$$\inf_{x\in J_j} Mf(x) \ \leq \ \text{const.}\left(\inf_{x\in 4J_j} M_p(Mf)(x)\right) \ \leq \ \text{const.}\, \kappa_p$$

and similarly for g. Since the J_j are disjoint, theorem (4.3) thus follows immediately from (4.6). □

Step 2. Estimates for $\mathcal{P}_{good}$ are needed. Denote by $\mathbb{I}_0$ all dyadic intervals I for which $I \not\subseteq E_{bad}$. Then

$$\mathcal{P}_{good}(f,\, g) \ = \ \sum_{I\in\mathbb{I}_0} c_I \,\frac{1}{\sqrt{|I|}}\langle\, f,\, \psi_I^{(1)}\,\rangle\,\langle\, g,\, \psi_I^{(2)}\,\rangle\,\psi_I^{(3)},$$

and, in view of (4.5), all the coefficients in $\mathcal{P}_{good}(f, g)$ have bounds

$$\frac{1}{\sqrt{|I|}}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle| \ \leq \ \text{const.}\,\|\psi^{(1)}\|\,\kappa_p, \qquad \frac{1}{\sqrt{|I|}}\,|\langle\, g,\, \psi_I^{(2)}\,\rangle| \ \leq \ \text{const.}\,\|\psi^{(2)}\|\,\kappa_q$$

irrespective of vanishing moments. We have to show that

$$|\{x:\ |\mathcal{P}_{good}(f,\, g)(x) \ \geq \ \gamma\,\}| \ \leq \ C_\psi\left(\frac{\|f\|_p\,\|g\|_q}{\gamma}\right)^r. \tag{4.7}$$

When $\psi^{(2)}$ and $\psi^{(3)}$ have vanishing moment this is straightforward. For then

$$\{d_I\}\times h \longrightarrow \sum_{I\in\mathbb{I}_0} d_I\,\langle\, h,\, \psi_I^{(2)}\,\rangle\,\psi_I^{(3)},$$

is bounded from $\ell^\infty \times L^q(\mathbb{R}^n)$ into $L^q(\mathbb{R})$, and so

$$\left(\int_{\mathbb{R}^n} |\mathcal{P}_{good}(f,\, h)(x)|^q\,dx\right)^{1/q} \ \leq \ C_\psi\,\kappa_p\,\|h\|_q.$$

Taking $h = g$ we obtain (4.7) because of the choice of κ_p, κ_q, completing the proof of theorem (1.9) for all $r > 1/2$ when $\psi^{(2)}$, $\psi^{(3)}$ have vanishing moment. A reversal of the roles of f and g establishes the same result when $\psi^{(1)}$, $\psi^{(3)}$ have vanishing moment. But the adjoints of a 'standard' paraproduct are well-defined when $r \geq 1$. Consequently, theorem (1.9) also remains true, at least for $r \geq 1$, irrespective of which two of the $\psi^{(j)}$ have vanishing moment.

Step 3. All that remains is to establish (4.7) for $r < 1$ when $\psi^{(1)}$, $\psi^{(2)}$ have vanishing moment and $\phi^{(3)}$ does not. We will actually prove that

$$\frac{1}{\gamma^2}\int_{-\infty}^{\infty} |\mathcal{P}_{good}(f,\, g)(x)|^2\,dx \ \leq \ C_\psi\left(\frac{\|f\|_p\|g\|_q}{\gamma}\right)^r \tag{4.8}$$

using a Tent space argument. Given a dyadic interval J let Δ_J be the square of side-length $|J|$ sitting above J in the dyadic tiling of the upper half-plane, and let χ_{Δ_J} be its characteristic function. Now set

$$F(z) \ = \ \sum_{I\in\mathbb{I}_0} \langle\, f,\, \psi_I^{(1)}\,\rangle\,\chi_{\Delta_I}(z), \qquad G(z) \ = \ \sum_{I\in\mathbb{I}_0} \langle\, g,\, \psi_I^{(1)}\,\rangle\,\chi_{\Delta_I}(z)$$

and

$$H(z) \;=\; \sum_{I\,\in\,\mathbb{I}_0} c_I\,\frac{1}{\sqrt{|I|}}\,\langle\, h,\, \psi_I^{(3)}\,\rangle\,\chi_{\Delta_I}(z)$$

where h is an arbitrary function in $L^2(\mathbb{R})$. Since

$$\int_{-\infty}^{\infty} \mathcal{P}_{good}(f,\, g)(x)\,\overline{h(x)}\,dx \;=\; \text{const.}\int_0^{\infty}\int_{-\infty}^{\infty} F(v,\, t)\,G(v,\, t)\overline{H(v,\, t)}\,\frac{dvdt}{t^2},$$

the proof of (4.8) becomes one of using Carleson measure arguments in the upper half plane. As $\psi^{(3)}$ does not have vanishing moments, however, the only estimates available for H are those of its non-tangential maximal function; Carleson measures have to come from F or G.

Recall that $\mathbb{I}_0$ denotes all dyadic intervals I for which $I \not\subseteq E_{bad}$ and that $\phi^{(1)}$, $\phi^{(2)}$ have vanishing moment. Then by the localized 'Lusin Area' result for frames established in the Appendix to Part II ([11]), the inequality

$$\text{(4.9)(i)} \qquad \left(\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle|^2\chi_I(x)\right)^{p/2} dx\right)^{1/p} \;\le\; \text{const}_\psi\,\kappa_p$$

holds for all J in $\mathbb{I}_0$ together with

$$\text{(4.9)(ii)} \qquad \left(\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, g,\, \psi_I^{(2)}\,\rangle|^2\chi_I(x)\right)^{q/2} dx\right)^{1/q} \;\le\; \text{const}_\psi\,\kappa_q$$

for g. At the expense of using a possibly larger constant we shall assume that the *same* constant appears in (4.9). We use these to begin an iterative choice of families of intervals exhausting $\mathbb{I}_0$. Choose a dyadic interval J in $\mathbb{I}_0$ such that

$$\text{(4.10)} \qquad \frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle|^2\chi_I(x)\right)^{1/2} dx \;\ge\; \text{const}_\psi\, 2^{-1/p}\,\kappa_p.$$

Since

$$\left(\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle|^2\chi_I(x)\right)^{1/2} dx\right)^{p}$$

$$\le\; \frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle|^2\chi_I(x)\right)^{p/2} dx \;\le\; \text{const.}\;\frac{1}{|J|}\int_{-\infty}^{\infty} |f(x)|^p\,dx,$$

the intervals J satisfying (4.10) will be uniformly bounded in length. Thus we can choose J_1 for which $|J_1|$ is *maximal* among all dyadic intervals J in $\mathbb{I}_0$ satisfying (4.10); set $\mathbb{J}_1 = \{I \in \mathbb{I}_0 : I \subseteq J_1\}$ and $I_{\mathbb{J}_1} = J_1$. Now choose J_2 for which $|J_2|$ is maximal among all dyadic intervals in $\mathbb{I}_0\backslash\mathbb{J}_1$ satisfying (4.10); set $\mathbb{J}_2 = \{I \in \mathbb{I}_0 \setminus \mathbb{J}_1 : I \subseteq J_2\}$ and $I_{\mathbb{J}_2} = J_2$. Continuing in this way until no further intervals exist we obtain disjoint dyadic intervals $J_1,\, J_2,\, \ldots$ and corresponding disjoint families $\mathbb{J}_1,\, \mathbb{J}_2,\, \ldots$. Set $\mathcal{J}_0^{(1)} \;=\; \{\mathbb{J}_1,\, \mathbb{J}_2,\, \ldots\,\}$ and

$$\mathbb{I}_0^{(1)} \;=\; \bigcup_{\mathbb{J}\,\in\,\mathcal{J}_0^{(1)}} \{I : I\,\in \mathbb{J}\}.$$

The same construction can be carried out beginning with maximal intervals J in $\mathbb{I}_0 \setminus \mathbb{I}_0^{(1)}$ for which the inequality

$$\frac{1}{|J|}\int_J \Big(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, g,\, \psi_I^{(2)}\,\rangle|^2 \chi_{_I}(x)\Big)^{1/2} dx \;\geq\; \mathrm{const}_\psi\, 2^{-1/q}\,\kappa_q \tag{4.11}$$

holds. Denote by $\mathcal{J}_0^{(2)} \;=\; \{\mathbb{J}_1,\, \mathbb{J}_2,\, \dots\,\}$ the disjoint dyadic intervals $J_1,\, J_2,\, \dots$ so obtained, and set

$$\mathbb{I}_0^{(2)} \;=\; \bigcup_{\mathbb{J}\,\in\,\mathcal{J}_0^{(2)}} \{I : I\,\in \mathbb{J}\}, \qquad \mathbb{I}_1 \;=\; \mathbb{I}_0 \setminus (\mathbb{I}_0^{(1)} \,\cup\, \mathbb{I}_0^{(2)}), \qquad \mathcal{J}_0 \;=\; \mathcal{J}_0^{(1)} \,\cup\, \mathcal{J}_0^{(2)}.$$

We continue inductively. Suppose a family $\mathbb{I}_\nu$ of intervals in $\mathbb{I}_0$ remains. Choose an interval J_1 for which $|J_1|$ is maximal among all dyadic intervals in $\mathbb{I}_\nu$ satisfying

$$\frac{1}{|J|}\int_J \Big(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, f,\, \psi_I^{(1)}\,\rangle|^2 \chi_{_I}(x)\Big)^{1/2} dx \;\geq\; \mathrm{const}_\psi\, 2^{-(\nu+1)/p}\,\kappa_p, \tag{4.12}$$

and set $\mathbb{J}_1 \,=\, \{I\,\in\, \mathbb{I}_\nu : I\,\subseteq\, J_1\}$, $I_{\mathbb{J}_1} \;=\; J_1$. Now continue as before, first with f until no further intervals satisfying (4.12) exist, producing $\mathcal{J}_\nu^{(1)}$ as well as the associated family of intervals

$$\mathbb{I}_\nu^{(1)} \;=\; \bigcup_{\mathbb{J}\,\in\,\mathcal{J}_\nu^{(1)}} \{I : I\,\in \mathbb{J}\};$$

then with intervals for which

$$\frac{1}{|J|}\int_J \Big(\sum_{I\subseteq J} \frac{1}{|I|}\,|\langle\, g,\, \psi_I^{(2)}\,\rangle|^2 \chi_{_I}(x)\Big)^{1/2} dx \;\geq\; \mathrm{const}_\psi\, 2^{-(\nu+1)/q}\,\kappa_q, \tag{4.13}$$

holds until they too have been exhausted, producing $\mathcal{J}_\nu^{(2)}$ as well as the associated family $\mathbb{I}_\nu^{(2)}$ of intervals. Setting

$$\mathbb{I}_{\nu+1} \;=\; \mathbb{I}_\nu \setminus (\mathbb{I}_\nu^{(1)} \,\cup\, \mathbb{I}_\nu^{(2)}), \qquad \mathcal{J}_\nu \;=\; \mathcal{J}_\nu^{(1)} \,\cup\, \mathcal{J}_\nu^{(2)},$$

completes the inductive construction. Since any interval in $\mathbb{I}_0$ is always a candidate for one of the J's, every I in $\mathbb{I}_0$ for which $\langle\, f,\, \psi^{(1)}\,\rangle \neq 0$ and $\langle\, g,\, \psi^{(2)}\,\rangle \neq 0$ will belong to one of the $\mathbb{J}$ in some $\mathcal{J}_\nu$. Thus

$$\mathcal{P}_{good}(f,\, g) \;=\; \sum_{\nu\,=\,0}^{\infty}\Big(\sum_{\mathbb{J}\,\in\,\mathcal{J}_\nu} \mathcal{P}_{\mathbb{J}}(f,\, g)\Big)$$

where

$$\mathcal{P}_{\mathbb{J}}(f,\, g) \;=\; \sum_{I\,\in\,\mathbb{J}} c_I\, \frac{1}{\sqrt{|I|}}\,\langle\, f,\, \psi^{(1)}\,\rangle\,\langle\, g,\, \psi^{(2)}\,\rangle\, \psi_I^{(3)}.$$

The previous construction enables both the L^2-norm of individual $\mathcal{P}_{\mathbb{J}}(f,\, g)$ to be estimated as well as that of (4.8). Indeed, for any $\mathbb{J}$ in $\mathcal{J}_\nu^{(i)}$ set

$$F_{\mathbb{J}}^{(i)}(z) \;=\; \sum_{I\,\in\,\mathbb{J}} \langle\, f,\, \psi_I^{(1)}\,\rangle\, \chi_{\Delta_I}(z), \qquad G_{\mathbb{J}}^{(i)}(z) \;=\; \sum_{I\,\in\,\mathbb{J}} \langle\, g,\, \psi_I^{(2)}\,\rangle\, \chi_{\Delta_I}(z)$$

and

$$H_{\mathbb{J}}^{(i)} \;=\; \sum_{I\,\in\,\mathbb{J}} c_I \frac{1}{\sqrt{|I|}} \langle\, h,\, \psi^{(3)}\,\rangle\, \chi_{\Delta_I}(z).m$$

Then

$$\int_{-\infty}^{\infty} \mathcal{P}_{\mathbb{J}}(f,\, g)(x)\, \overline{h(x)}\, dx = \text{ const.}\Big(\int_0^\infty \int_{-\infty}^\infty F_{\mathbb{J}}^{(i)}(v,\, t)\, G_{\mathbb{J}}^{(i)}(v,\, t)\overline{H_{\mathbb{J}}^{(i)}(v,\, t)}\, \frac{dvdt}{t^2}\Big).$$

Now, by (4.5), the non-tangential maximal function of $H_{\mathbb{J}}^{(1)}$ satisfies the inequality

$$N(H_{\mathbb{J}}^{(1)})(x) \;\leq\; \text{const. } \|\psi^{(3)}\| \left(\inf_{v\,\in\, I_{\mathbb{J}}} Mh(v)\right) \chi_{4I_{\mathbb{J}}}(x).$$

On the other hand, the tent space norms of $F_{\mathbb{J}}^{(i)}$, $G_{\mathbb{J}}^{(i)}$ are controlled by the choice of the $\mathbb{J}$.

(4.14) Theorem. *The function $F_{\mathbb{J}}^{(1)}$ belongs to the Tent space $\mathfrak{N}^\infty$ for each $\mathbb{J}$ in $\mathcal{J}_\nu^{(1)}$, while $G_{\mathbb{J}}^{(1)}$ belongs to $\mathfrak{N}^2$; more precisely, the inequalities*

$$\|F_{\mathbb{J}}^{(1)}\|_{\mathfrak{N}^\infty} \;\leq\; \text{const. } 2^{-\nu/p}\, \kappa_p, \qquad \|G_{\mathbb{J}}^{(1)}\|_{\mathfrak{N}^2} \;\leq\; \text{const. } 2^{-\nu/q}\, \kappa_q\, |I_{\mathbb{J}}|^{1/2}$$

hold uniformly in $\mathbb{J}$ and ν.

There are corresponding results for the intervals in $\mathcal{J}_\nu^{(2)}$, reversing the roles of F and G in (4.14). Granted (4.14), the L^2-norm of $\mathcal{P}_{\mathbb{J}}(f,\, g)$ is easily estimated. For when $\mathbb{J}$ belongs to $\mathcal{J}_\nu^{(1)}$,

$$\left|\int_{-\infty}^{\infty} \mathcal{P}_{\mathbb{J}}(f,\, g)(x)\, \overline{h(x)}\, dx\right| \leq \left(\int_0^\infty \int_{-\infty}^\infty |G_{\mathbb{J}}^{(1)}(v,\, t)|^2\, \frac{dvdt}{t^2}\right)^{1/2}$$
$$\times \left(\int_0^\infty \int_{-\infty}^\infty |F_{\mathbb{J}}^{(1)}(v,\, t)|^2\, |H_{\mathbb{J}}^{(1)}(v,\, t)|^2\, \frac{dvdt}{t^2}\right)^{1/2}.$$

Consequently,

$$\left|\int_{-\infty}^{\infty} \mathcal{P}_{\mathbb{J}}(f,\, g)(x)\, \overline{h(x)}\, dx\right| \leq \text{ const. } 2^{-\nu/r} \kappa_p\, \kappa_q \left(\int_{I_{\mathbb{J}}} Mh(x)^2\, dx\right)^{1/2} |I_{\mathbb{J}}|^{1/2},$$

and so

$$\sum_{\mathbb{J}\,\in\,\mathcal{J}_\nu^{(1)}} \|\mathcal{P}_{\mathbb{J}}(f,\, g)\|_2 \;\leq\; \text{const. } 2^{-\nu/r} \kappa_p\, \kappa_q \left(\sum_{\mathbb{J}\,\in\,\mathcal{J}_\nu^{(1)}} |I_{\mathbb{J}}|\right)$$

since the $I_{\mathbb{J}}$ are disjoint. Once again the construction of the $\mathbb{J}$ provides an estimate for this last sum.

(4.15) Theorem. *The* $\mathbb{J}$ *in* $\mathcal{J}_\nu^{(1)}$ *satisfy the counting estimate*

$$\sum_{\mathbb{J}\,\in\,\mathcal{J}_\nu^{(1)}} |I_{\mathbb{J}}| \;\le\; \text{const.}\; 2^\nu \left(\frac{\|f\|_p}{\kappa_p}\right)^p$$

uniformly in ν.

The analogous result for $\mathbb{J}$ in $\mathcal{J}_\nu^{(2)}$ is obtained by reversing the roles of f, g and interchanging p, q. Hence

$$\sum_{\nu\,=\,0}^{\infty}\Biggl(\sum_{\mathbb{J}\,\in\,\mathcal{J}_\nu^{(1)}} \|\mathcal{P}_{\mathbb{J}}(f,\,g)\|_2\Biggr) \;\le\; \text{const.}\; \kappa_p\,\kappa_q \left(\frac{\|f\|_p}{\kappa_p}\right)^p,$$

because $r<1$. This together with the corresponding result for $\mathbb{J}$ in $\mathcal{J}_\nu^{(2)}$, $\nu\ge 0$, completes the proof of (4.8), and hence that of (1.9) also, once (4.14) and (4.15) have been proved.

Proof of (4.14). Since

$$\begin{aligned}\|F_\nu^{(1)}\|_{\mathfrak{N}^\infty} \;&=\; \sup_{J\,\subseteq\, I_{\mathbb{J}}}\left(\frac{1}{|J|}\int_{C(J)} |F_\nu^{(1)}(v,\,t)|^2\,\frac{dvdt}{t^2}\right)^{1/2} \\ &=\; \text{const.}\;\sup_{J\,\subseteq\, I_{\mathbb{J}}}\left(\frac{1}{|J|}\int_J \Bigl\{\sum_{I\,\subseteq\, J}\frac{1}{|I|}\,|\langle\, f,\,\phi_I^{(1)}\,\rangle|^2\,\chi_{{}_I}(x)\Bigr\}\,dx\right)^{1/2},\end{aligned}$$

we have to show that the L^1-norm used in (4.9) and thereafter can be replaced by an L^2-norm at the expense possibly of introducing an extra constant factor in the right hand side.

Fix a dyadic interval $J\subseteq I_{\mathbb{J}}$ and define an $\ell^2(\mathbb{J})$-valued function on J by

$$\Phi_{\mathbb{J}}(x) \;=\; \left\{\frac{1}{\sqrt{|I|}}\,\langle f,\,\phi_I^{(1)}\,\rangle\,\chi_{{}_I}(x)\right\}_{I\in\mathbb{J}} \qquad (x\in J).$$

Then

$$\begin{aligned}&\left(\frac{1}{|J|}\int_J\Bigl\{\sum_{I\,\subseteq\, J}\frac{1}{|I|}\,\bigl|\langle\, f,\,\phi_I^{(1)}\,\rangle\bigr|^2\chi_{{}_I}(x)\Bigr\}\,dx\right)^{1/2} \\ &\qquad\qquad =\;\left(\frac{1}{|J|}\int_J \|\Phi_{\mathbb{J}}(x)\|_{\ell^2}^2\,dx\right)^{1/2} \;\le\; \text{const.}\;\|\Phi_{\mathbb{J}}\|_{BMO(J)}\end{aligned}$$

where $BMO(J)$ is understood with respect to the dyadic structure. But because of this structure,

$$\begin{aligned}&\frac{1}{|J_0|}\int_{J_0}\|\Phi_{\mathbb{J}}(y)-\frac{1}{|J_0|}\int_{J_0}\Phi_{\mathbb{J}}\|_{\ell^2}\,dy \\ &\qquad\qquad \le\;\frac{2}{|J_0|}\int_{J_0}\Bigl(\sum_{I\,\subseteq\, J_0}\frac{1}{|I|}\,\bigl|\langle g,\,\phi_I^{(1)}\rangle\bigr|^2\chi_{{}_I}(x)\Bigr)^{1/2}\,dx\end{aligned}$$

for any dyadic interval $J_0 \subseteq J$. Consequently,

$$\|\Phi_{\mathbb{J}}\|_{BMO(J)} \;\le\; \text{const.}\ \kappa_p\, 2^{-\nu/p}.$$

Together these establish the $\mathfrak{N}^\infty$-estimate for $F^{(1)}_{\mathbb{J}}$. There is a correponding $\mathfrak{N}^\infty$-estimate for $G^{(1)}_{\mathbb{J}}$. From this the $\mathfrak{N}^2$-norm estimate follows since $BMO(I_{\mathbb{J}}) \subseteq L^2(I_{\mathbb{J}})$ on finite intervals. □

Proof of (4.15). Because the intervals $I_{\mathbb{J}},\ \mathbb{J} \in \mathcal{J}^{(1)}_\nu$, are pairwise disjoint,

$$\sum_{\mathbb{J}\,\in\,\mathcal{J}^{(1)}_\nu} \Big(\sum_{I\,\in\,\mathbb{J}} \frac{1}{|I|}\,|\langle\, f,\, \phi^{(1)}\,\rangle|^2\,\chi_{_I}(x)\Big)^{p/2} \;\le\; \Big(\sum_{I\,\in\,\mathcal{I}^{(1)}_\nu} \frac{1}{|I|}\,|\langle\, f,\, \phi^{(1)}\,\rangle|^2\,\chi_{_I}(x)\Big)^{p/2}.$$

Thus by (A.7) in the Appendix to Part II ([11]),

$$\sum_{\mathbb{J}\,\in\,\mathcal{J}^{(1)}_\nu} \Big\{\int_{I_{\mathbb{J}}} \Big(\sum_{I\,\in\,\mathbb{J}} \frac{1}{|I|}\,|\langle\, f,\, \phi^{(1)}\,\rangle|^2\,\chi_{_I}(x)\Big)^{p/2} dx\Big\} \;\le\; \text{const.} \int_{-\infty}^{\infty} |f(x)|^p\, dx.$$

On the other hand, the construction of the $\mathbb{J}$ ensures that

$$\frac{1}{|I_{\mathbb{J}}|}\int_{I_{\mathbb{J}}} \Big(\sum_{I\,\in\,\mathbb{J}} \frac{1}{|I|}\,|\langle\, f,\, \phi^{(1)}\,\rangle|^2\,\chi_{_I}(x)\Big)^{p/2} dx \;\ge\; \text{const.}\,(2^{-\nu/p}\,\kappa_p)^p.$$

Theorem (4.15) follows immediately. □

5. Grid sructures

A family $\mathcal{W}$ of intervals in $\mathbb{R}$ is said to form a *grid* provided

(5.1)
$$w \cap w' \ne \emptyset \quad\Longrightarrow\quad w \subseteq w',\ \text{or}\ \ w' \subseteq w,$$
$$w' \subset w \quad\Longrightarrow\quad 2|w'| \;\le\; |w|$$

hold for all pairs $w, w' \in \mathcal{W}$. One such example is the grid $\mathcal{W}_1$ consisting of all dyadic intervals $[2^k n,\ 2^k(n+1))$; more generally, the family $\mathcal{W}_{1,\rho}$ of all intervals

(5.2)
$$w_{kn} \;=\; [2^{\rho k} n,\ 2^{\rho k}(n+1)), \qquad -\infty < k, n < \infty$$

is a grid for each positive integer ρ. Similarly, the family $\mathcal{I}_1$ of all dyadic intervals $I_{k\ell} = [2^{-k}\ell,\, 2^{-k}(\ell+1))$ is a grid as is the family $\mathcal{I}_{1,\rho}$ of intervals

(5.3)
$$I_{k\ell} \;=\; [2^{-\rho k}\ell,\ 2^{-\rho k}(\ell+1)), \qquad -\infty < k, \ell < \infty.$$

More generally still, to each integer $M \ge 2$ there corresponds a family $\mathcal{W}_M$ of intervals satisfying (5.1).

(5.4) Theorem. *The family $\mathcal{W}_M$ of intervals*

$$w_{kn} \ = \ \big[\, 2^{Mk}(n-\alpha_M),\ 2^{Mk}(n+\alpha_M)\big), \qquad \alpha_M \ = \ \frac{2^{M-1}-1}{2^M-1}$$

is a grid for each integer $M \ge 2$.

When $M=3$, for instance, the intervals

$$w_{kn} \ = \ [8^k(n-\tfrac{3}{7}),\ 8^k(n+\tfrac{3}{7})), \qquad -\infty < k, n < \infty,$$

thus form a grid of intervals of length $\frac{6}{7}8^k$.

Proof of Theorem (5.4). Suppose $M \ge 2$. To establish the first of the properties in (5.1) for $\mathcal{W}_M$ we can assume without loss of generality that

$$w_{kn} \ = \ [2^{Mk}(n-\alpha_M),\ 2^{Mk}(n+\alpha_M)), \qquad w_{k'n'} \ = \ [n'-\alpha_M,\ n'+\alpha_M)$$

with $k \ge 0$. If $w_{kn} \cap w_{k'n'} \ne \emptyset$, then at least one of

$$\begin{aligned} &2^{Mk}(n-\alpha_M) \ \le \ n'-\alpha_M \ < \ 2^{Mk}(n+\alpha_M), \\ &2^{Mk}(n-\alpha_M) \ < \ n'+\alpha_M \ \le \ 2^{Mk}(n+\alpha_M) \end{aligned} \tag{5.5}$$

must hold - say the first one. But $2^{Mk}-1 = A(2^M-1)$ with A a positive integer. Consequently,

$$n' \ < \ 2^{Mk}n + A\alpha_M(2^M-1) + 2\alpha_M \ = \ 2^{Mk}n + A(2^{M-1}-1) + \frac{2^M-2}{2^M-1}.$$

Thus

$$n' \ \le \ 2^{Mk}n + A(2^{M-1}-1),$$

and so

$$n'+\alpha_M \ \le \ 2^{Mk}n \ + \ A(2^{M-1}-1) + \alpha_M \ = \ 2^{Mk}(n+\alpha_M).$$

Hence $w_{k'n'} \subseteq w_{kn}$. An entirely analogous argument shows that $w_{k'n'} \subseteq w_{kn}$ continues to hold for the second inclusion, completing the proof since the second of the conditions in (5.1) is obviously satisfied. $\square$

To link grids with paraproduct let

$$\mathcal{D}(f,\,g) \ = \ \sum_{k,\,\ell,\,n\,=\,-\infty}^{\infty} c_{k\ell n}\, 2^{Kk/2}\,\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n} \tag{5.6}$$

be a time-frequency paraproduct in which

$$\phi^{(j)}_{k\ell n}(x) \ = \ 2^{Kk/2}\phi^{(j)}(2^{Kk}x - a_j\ell)\, e^{2\pi i 2^K knx}$$

where K will be specified in a moment and the Fourier support intervals of the $\phi^{(j)}$ all lie in interval $(\alpha,\ \alpha+\frac{1}{2})$, $|\alpha| < 1/2$ containing the origin or contained in $(0,\ 1)$ (*cf.* (2.8)). Then clearly the $w^{(j)}$ all lie in $(0,\,1)$ when $\alpha > 0$ ($M=1$ in this case) or there exists $M \ge 2$ so that all lie in $(-\alpha_M,\,\alpha_M)$. Geometrically, the following result is clear.

(5.7) Theorem. *When the support intervals $w^{(j)}$ all lie in $(0,\ 1)$ there is an integer N so that without loss of generality we can take*

$$w^{(j)} \;=\; [\alpha_j/2^N,\ \beta_j/2^N), \qquad 0<\alpha_j<\beta_j<2^N$$

for a suitable choice of integers α_j and 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 $w^{(j)}$.

There are analogous results for the case $M\geq 2$, but the geometry becomes more complicated because at each generation k the intervals

$$\text{(5.8)(i)} \qquad [2^{-Mk}(n-\alpha_M),\ 2^{-Mk}(n+\alpha_M))$$

leave gaps in $(-\alpha_M,\ \alpha_M)$; in fact, between every adjacent pair of intervals at generation k of $\mathcal{W}_M$ there is an interval

$$\text{(5.8)(ii)} \qquad [2^{-Mk}(n+\alpha_M),\ 2^{-Mk}(n+1-\alpha_M))$$

which does not belong to $\mathcal{W}_M$. Nonetheless, for each $k\geq 1$

$$[-\alpha_M,\ \alpha_M) \;=\; \left(\bigcup_{n=n_\ell}^{n_r-1}[2^{-Mk}(n-\alpha_m),\ 2^{-Mk}(n+1-\alpha_M))\right)\cup[2^{-Mk}(n_r-\alpha_M),\ \alpha_M)$$

where

$$n_\ell \;=\; -\alpha_M\,(2^{Mk}-1), \qquad n_r \;=\; \alpha_M(2^{Mk}-1)$$

and each interval in the first union is itself the union of an interval (5.8)(i) in $\mathcal{W}_M$ and a gap (5.8)(ii) which does not belong to $\mathcal{W}_M$, of course. Note that there are no such gaps adjacent to the endpoints of $[-\alpha_M,\ \alpha_M)$.

(5.9) Theorem. *When the Fourier support intervals all lie in $[-\alpha_M,\ \alpha_M)$ there is an integer N so that without loss of generality we can assume*

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

for a suitable choice of integers α_j, β_j; in addition, we can assume that there is an interval in $\mathcal{W}_M$ 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 $\mathcal{W}_M$ of length $2\alpha_M 2^{-MN}$ just as in the dyadic case (5.7). Families of tiles in phase plane can be introduced using these grids. Let N be the integer in (5.7) or (5.9) according as $M=1$ or $M>1$, 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; where $\mathcal{I}_{M,N}$ and $\mathcal{W}_{M,N}$ denote the family of intervals $I_{k\ell} = [2^{-MNk}\,\ell\,,\, 2^{-MNk}(\ell + 1))$ and $w_{kn} = \big[\,2^{MNk}(n-\alpha_M),\ 2^{MNk}(n+\alpha_M)\big)$ respectively. The intervals $I_Q = I_{k\ell}$ and $w_Q = w_{kn}$ will be called respectively the *time* and *frequency* intervals of Q. By taking $K = MN$ in (5.6) we thus arrive at the fundamental link between tiles in phase space and paraproduct: for each $Q \sim \{k,\,\ell,\,n\}$ in $\mathbb{Q}_{M,N}$ set

$$\phi_Q^{(j)}(x) \ = \ s^{k/2}\phi^{(j)}(s^k - a_j\ell)\, e^{2\pi s^k nx}, \qquad s = 2^{MN}. \tag{5.10}$$

(5.11) Definition. *We say that a time-frequency paraproduct* $\mathcal{D}(f,\,g)$ *is in* (M,N)*-canonical form when*

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

where the wave packets are defined by (5.10) and the Fourier support intervals $w^{(j)}$ *of the* $\phi^{(j)}$ *satisfy (5.7) in case* $M = 1$ *and (5.9) in case* $M > 1$.

The notation in (5.6) and (5.11) for a paraproduct in $(M,\,N)$-form will be used interchangably. These paraproducts have a number of special properties. 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. The intervals $w_Q^{(j)} = \tau_Q(w^{(j)})$ are then the Fourier support intervals of the wave packets $\phi_Q^{(j)}$.

(5.12) Theorem. *The family* $\{w_Q^{(j)} : Q \in \mathbb{Q}_{M,N}\}$ *is a grid for each* j.

Proof. It is clearly enough to check the first condition

$$w_P^{(j)} \cap w_Q^{(j)} \neq \emptyset \quad \Longrightarrow \quad w_P^{(j)} \subseteq w_Q^{(j)},\ \textit{or},\ \ w_Q^{(j)} \subseteq w_P^{(j)}$$

in (5.1). So suppose $w_P^{(j)} \cap w_Q^{(j)} \neq \emptyset$. If $|w_P^{(j)}| = |w_Q^{(j)}|$, then $w_P^{(j)} = w_Q^{(j)}$ and there is nothing to prove. Consequently, we can assume that $|w_P^{(j)}| < |w_Q^{(j)}|$. Since

$$w_P^{(j)} \cap w_Q^{(j)} \neq \emptyset \quad \Longrightarrow \quad w_P \cap w_Q^{(j)} \neq \emptyset,$$

either $w_P \subseteq w_Q^{(j)}$, or w_P overlaps $w_Q^{(j)}$ at one edge. In the latter case, theorems (5.7) and (5.9) ensure that there is an interval $d_Q^{(j)}$ in $\mathcal{W}_{M,N}$ such that

$$|d_Q^{(j)}| \ = \ 2^{-MN}|w_Q|, \quad d_Q^{(j)} \cap w_P \neq \emptyset, \quad d_Q^{(j)} \subseteq w_Q^{(j)}. \tag{5.13}$$

But then

$$|w_P| \ = \ 2^{MN}\,\frac{|w_P|}{|w_Q|}\,|d_Q^{(j)}| \ \leq \ |d_Q^{(j)}|,$$

and so $w_P \subseteq d_Q^{(j)} \subseteq w_Q^{(j)}$ because of the grid structure on $\mathcal{W}_{M,N}$. ☐

It will be useful to reformulate the main step in the previous proof in a slightly different way.

(5.14) Corollary. *Let P and Q be tiles in $\mathbb{Q}_{M,N}$ such that $w_P \cap w_Q^{(j)} \neq \emptyset$ and $|I_Q| < |I_P|$. Then, $w_P \subseteq w_Q^{(j)}$.*

One further consequence of (5.11) will be important in Part II (*cf.* [11]).

(5.15) Theorem. *Fix frequencies λ_1, λ_2 with $\lambda_1 < \lambda_2$ and let*

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

be a tile 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 in the case $M = 1$ for a tile in $\mathbb{Q}_{1,N}$

Proof. With the notation of (5.9), λ_1 belongs to $w_Q^{(i)}$ if and only if

$$2^{MNk}(2^{-MN}(\alpha_i - \alpha_M) + m) \;\leq\; \lambda_1 \;<\; 2^{MNk}(2^{-MN}(\beta_i + \alpha_M) + m),$$

while λ_2 belongs to $w_Q^{(j)}$ if and only if

$$2^{MNk}(2^{-MN}(\alpha_j - \alpha_M) + m) \;\leq\; \lambda_2 \;<\; 2^{MNk}(2^{-MN}(\beta_j + \alpha_M) + m).$$

Consequently,

$$2^{-MN}(\alpha_j \,-\, \beta_i \,- 2\alpha_M) \;\leq\; 2^{-MNk}(\lambda_2 \,-\, \lambda_1) \;<\; 2^{-MN}(\beta_j \,-\, \alpha_i + 2\alpha_M).$$

But

$$1 \;\leq\; \alpha_j \,-\, \beta_i - 2\alpha_M$$

since there is an interval of length 2^{-MN} between $w^{(i)}$ and $w^{(j)}$; on the other hand,

$$\beta_j - \alpha_i \;<\; n_r \,-\, n_\ell \;\leq\; (2\alpha_M)(2^{MN} - 1).$$

Thus

$$2^{MN(k-1)} \;\leq\; \lambda_2 \,-\, \lambda_1 \;<\; 2^{MNk},$$

completing the proof of (5.15) since these can be satisfied for just one value of k. $\square$

The grid structure built into an $(M,\,N)$-canonical form can also be exploited to establish pointwise convergence results sharpening the norm convergence results in (2.5).

(5.16) Theorem. *The infinite series defining a time-frequency paraproduct* $\mathcal{D}(f,\, g)(x)$ *in* $(M,\, N)$*-canonical form converges absolutely for each* x *when* $f,\, g$ *are band-limited* $\mathcal{M}_\mu$*-molecules; more precisely, the inequality*

$$\sum_{k\,\ell,\, n\,=\,-\infty}^{\infty} |c_{k\ell n}\, s^{k/2} \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}(x)| \;\le\; C\, \|\{c_{k\ell n}\}\|_\infty$$

holds uniformly in x *for each such* f *and* g*.*

Thus so long as we restrict to band-limited $\mathcal{M}_\mu$-molecules any time-frequency paraproduct in $(M,\, N)$-canonical form can be manipulated freely. Such unconditionality will become crucial in Part II (*cf.* [11]).

Proof of (5.16). We give the proof in the case $M \ge 2$, leaving the reader to make the necessary changes for $M = 1$. Recall that if f and ϕ are $\mathcal{M}_\mu$-molecules then the vanishing moment property ensures that

$$|\langle\, f,\, \phi_{k\ell}\,\rangle| \;\le\; \text{const.}\, \|f\| \|\phi\|\, s^{-k/2}\, \frac{\min(1,\, s^{-k})^\mu}{(1 + s^{-k} + s^{-k}|\ell|\,)^{\mu+1}}$$

when $\phi_{k\ell}(x) \;=\; s^{k/2}\phi(s^k - a\ell)$ (*cf.*, for instance, [9]). On the other hand,

$$|\phi_{k\ell}(x)| \;\le\; \|\phi\|\, \frac{s^{k/2}}{(1 + |s^k x - a\ell|)^{1+\mu}}$$

whether or not ϕ has vanishing moment. Now let $f,\, g$ be band-limited $\mathcal{M}_\mu$-molecules. By dilating if necessary, we can assume that $\widehat{f},\, \widehat{g}$ have support in $(-\alpha_M,\, \alpha_M)$. For each n the modulate

$$\phi^{(j)}_n(x) \;=\; \phi^{(j)}(x)\, e^{2\pi i n x}$$

is an $\mathcal{M}_\mu$-test function such that

$$\|\phi^{(i)}_n\| \;\le\; \text{const.}\, (1 + |n|^\mu) \|\phi^{(i)}\|_{\mathcal{M}_\mu},$$

and the restriction on the supports of the Fourier transforms of the $\phi^{(i)}$ ensures that at least one of $\phi^{(1)}_n$, $\phi^{(2)}_n$ has vanishing moment for a given n. Fix n. Then the sum

$$\sum_{k,\, \ell\,=\,-\infty}^{\infty} |c_{k\ell n}\, s^{k/2} \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \,\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}(x)|$$

converges for each x. Indeed, suppose that $\phi^{(1)}_n$ has vanishing moment. By the earlier basic estimate for $\mathcal{M}_\mu$-molecules,

$$|\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle| \;\le\; \text{const.}\, \|f\| \|\phi^{(1)}\|\, (1 + |n|)^\mu \left(s^{-k/2}\, \frac{\min(1,\, s^{-k})^\mu}{(1 + s^{-k})^{\mu+1}} \right)$$

uniformly in ℓ and n, while

$$|\phi^{(3)}_{k\ell n}(x)| \;\le\; \|\phi^{(3)}\| \frac{s^{k/2}}{(1 + |s^k x - a\ell|)^{1+\mu}}$$

for all x. On the other hand,

$$|\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle| \;\le\; \text{const.}\, \|\phi^{(2)}\|\, \|g\|$$

for all k, ℓ and n irrespective of vanishing moment. (If $\phi^{(2)}_n$ has vanishing moment we reverse the roles of f and g.) Consequently, the inequality

$$\begin{aligned} &\sum_{k,\,\ell\,=\,-\infty}^{\infty} |c_{k\ell n}\, s^{k/2} \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}(x)| \\ &\qquad \le\; C\, \|\{c_{k\ell n}\}\|_\infty (1+|n|)^\mu \Big(\sum_{k\,=\,-\infty}^{\infty} s^{k/2} \frac{\min(1,\, s^{-k})^\mu}{(1+s^{-k})^{1+\mu}} \Big) \end{aligned} \tag{5.17}$$

holds uniformly in x for each fixed n. So far, the band-limited assumption has not been been needed. Its purpose is to restrict which k, n can occur. Again fix n for the moment. Then $\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle = 0$ unless the interval $[s^k(n-\alpha_M),\, s^k(n+\alpha_M))$ containing the support of $\widehat{\phi}^{(1)}_{k\ell n}$ intersects $[-\alpha_M,\, \alpha_M)$. Hence by the grid property,

$$\langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle \ne 0 \quad \Longrightarrow \quad \begin{cases} [s^k(n-\alpha_M),\, s^k(n+\alpha_M)) \subseteq [-\alpha_M,\, \alpha_M), \ \textit{or} \\ \quad [s^k(n-\alpha_M),\, s^k(n+\alpha_M)) \supset [-\alpha_M,\, \alpha_M). \end{cases}$$

In the first of these cases the only possible values of (k, n) are $k \le 0$ and $|n| \le s^{-k}\alpha_M$, while in the second, $n = 0$ and $k > 0$. There are similar results for $\langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle$. As the case of a single value of n has been dealt with already in (5.17), we are left with the first case. But here

$$\begin{aligned} &\sum_{k\,=\,-\infty}^{0} \Big(\sum_{|n|\,\le\, s^{-k}\alpha_M} \Big\{ \sum_{\ell\,=\,-\infty}^{\infty} |c_{k\ell n}\, s^{k/2} \langle\, f,\, \phi^{(1)}_{k\ell n}\,\rangle\, \langle\, g,\, \phi^{(2)}_{k\ell n}\,\rangle\, \phi^{(3)}_{k\ell n}(x)| \Big\} \Big) \\ &\qquad \le\; C\, \|\{c_{k\ell n}\}\|_\infty \sum_{k\,=\,-\infty}^{0} \Big(\sum_{|n|\,\le\, s^{-k}\alpha_M} s^{k/2}\, (1+|n|)^\mu\, \frac{\min(1,\, s^{-k})^\mu}{(1+s^{-k})^{1+\mu}} \Big), \end{aligned}$$

completing the proof. □

Appendix: L^p-boundedness of $\mathcal{C}_{\mathbb{R}^2}$

In this appendix a proof of the L^p-boundedness of $\mathcal{C}_{\mathbb{R}^2}$ for the full range of r, *i.e.*, $1/r = 1/p + 1/q < 2$, is given (see also [12][14]). Yet again it is enough to consider cone operators. In fact, by using a finite partition of unity on the unit circle we can write $\mathcal{C}_{\mathbb{R}^2}$ is a finite sum of bilinear cone operators

$$\mathcal{C}_\Gamma(f,\, g) \;=\; \int_\Gamma m(\xi,\, \eta)\, \widehat{f}(\xi)\, \widehat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi d\eta$$

where the cone Γ either lies wholly inside a quadrant of $\mathbb{R}^2$ or inside a cone straddling a coordinate axis. Without loss of generality, therefore, it is enough to consider the case of

$$\Gamma_0 \;=\; \{(\xi,\, \eta) : 0 \,<\, \tfrac{1}{16}\xi \,\le\, \eta \,\le\, 16\xi\,\}, \qquad \Gamma_+ \;=\; \{(\xi,\, \eta) : 0 \,<\, |\eta| \,\le\, \tfrac{1}{8}\xi\,\}.$$

Most crucially of all, however, the support of the symbol $m = m(\xi, \eta)$ can be assumed to lie inside Γ. Thus, unlike a $\mathcal{C}_\Gamma$ in the Main Theorem, the singularity occurs only at the origin, not along the edges of the cone. The basic idea is essentially the same as in section 3: we choose $\mathcal{M}_\mu$-test functions $\psi^{(j)}$ so that each $\mathcal{C}_\Gamma$ can be written as a doubly-infinite sum

$$\mathcal{C}_\Gamma(f, g) = \sum_{\lambda_1, \lambda_2 = -\infty}^{\infty} \mathcal{P}_{\lambda_1 \lambda_2}(f, g) \tag{A.1}$$

of *standard* paraproduct

$$\mathcal{P}_{\lambda_1 \lambda_2}(f, g) \ = \ \sum_{k,\, \ell = -\infty}^{\infty} 2^{k/2} c_k(\lambda_1, \lambda_2) \, \langle\, f, \, \phi^{(1)}_{k\ell} \,\rangle \, \langle\, g, \, \phi^{(2)}_{k\ell} \,\rangle \, \phi^{(3)}_{k\ell}$$

in which

$$\phi^{(j)}(x) \ = \ \psi^{(j)}(x + a_j\lambda_j) \quad (j = 1, 2), \qquad \phi^{(3)}(x) \ = \ \psi^{(3)}(x) \tag{A.2}$$

and the $\phi^{(j)}_{k\ell}$ are defined by

$$\phi^{(j)}(x) \ = \ 2^{k/2} \phi^{(j)}(2^k x - a\ell)$$

for a fixed choice of positive constant a. Smoothness of the multiplier m ensures that the coefficients satisfy an inequality

$$|c_k(\lambda_1, \lambda_2)| \ \le \ \text{const.} \left(\frac{1}{1 + |\lambda_1| + |\lambda_2|} \right)^m \tag{A.3}$$

uniformly in k for every integer $m \ge 0$. One crucial point of the construction is that $\phi^{(3)}$ will always have vanishing moment, as will at least one $\phi^{(1)}$ or $\phi^{(2)}$, whether $\Gamma = \Gamma_0$ or Γ_+ so that Theorem (1.9) can be applied to every paraproduct in (A.1). Granted such a representation, therefore, the L^p-boundedness of $\mathcal{C}_\Gamma$ for the full range of r follows easily.

Suppose $r < 1$. Then

$$\|\mathcal{C}_\Gamma(f, g)\|_r \ \le \ \text{const.} \left(\sum_{\lambda_1, \lambda_2 = -\infty}^{\infty} \big(\sup_k \, |c_k(\lambda_1, \lambda_2)| \, \|\mathcal{P}_{\lambda_1 \lambda_2}\|_{op} \big)^r \right)^{1/r} \|f\|_p \, \|g\|_q.$$

On the other hand, by (1.9) and (A.2),

$$\|\mathcal{P}_{\lambda_1 \lambda_2}\|_{op} \ \le \ \text{const.} \left\{ 1 + \big(\prod_{j=1}^{2} \|\psi^{(j)}(\cdot + a_j\lambda_j)\| \big) \|\psi^{(3)}\| \right\}$$

uniformly in λ_1, λ_2, so

$$\|\mathcal{C}_\Gamma(f, g)\|_r \ \le \ \text{const.} \ \|f\|_p \, \|g\|_q,$$

using (2.2) and (A.3), hence completing proof of the L^p-boundedness of $\mathcal{C}_{\mathbb{R}^2}$. It is perhaps worth noting that only the simpler version of (1.9) was used in this proof above since we could guarantee that $\phi^{(3)}$ had vanishing moment and hence avoid the intricacies of step 3 in section 4.

To establish the representation of $\mathcal{C}_\Gamma$, suppose first that $\Gamma = \Gamma_0$ and let R be the square $w^{(1)} \times w^{(2)}$ where $w^{(1)} = w^{(2)} = [2,\ 66]$. Then the dilates $R_k = \{(2^k\xi,\ 2^k\eta) : (\xi,\ \eta) \in R\}$ of R provides a covering

$$\Gamma_0 \subset \bigcup_{k=-\infty}^{\infty} R_k, \qquad \mathrm{dist}(R_k, 0) \sim 2^k,$$

of Γ_0. On the other hand, when $w^{(3)} = [1, 257]$, then

$$R \subset \{ (\xi,\ \eta) : \xi + \eta \in w^{(3)} \}.$$

To exploit this geometry, first choose functions $\psi^{(j)}$ whose Fourier transforms are C^∞-bump functions such that $\mathrm{supp}\,\widehat{\psi}^{(j)} \subseteq w^{(j)}$, $j = 1,\ 2$, and

$$\sum_{k=-\infty}^{\infty} \widehat{\psi}^{(1)}(2^{-k}\xi)\,\widehat{\psi}^{(2)}(2^{-k}\eta) = 1, \qquad (\xi,\ \eta) \in \Gamma_0;$$

next choose a C^∞-function $\sigma = \sigma(\xi,\ \eta)$ so that $\mathrm{supp}\ \sigma \subset R$ and

$$\overline{\sigma(\xi,\ \eta)}\,\widehat{\psi}^{(1)}(\xi)\,\widehat{\psi}^{(2)}(\eta) = \widehat{\psi}^{(1)}(\xi)\,\widehat{\psi}^{(2)}(\eta).$$

Finally, choose a $\psi^{(3)}$ whose Fourier transform is a C^∞-function such that $\mathrm{supp}\,\widehat{\psi}^{(3)} \subset w^{(3)}$ and

$$\overline{\widehat{\psi}^{(3)}(\xi+\eta)}\,\widehat{\psi}^{(1)}(\xi)\,\widehat{\psi}^{(2)}(\eta) = \widehat{\psi}^{(1)}(\xi)\,\widehat{\psi}^{(2)}(\eta).$$

Then

$$\begin{aligned}\mathcal{C}_{\Gamma_0}(f,g)(x) &= \sum_k \Big(\int_{R_k} m(\xi,\ \eta)\,\sigma(2^{-k}\xi,\ 2^{-k}\eta) \\ &\times\ \widehat{f}(\xi)\,\widehat{g}(\eta)\,\overline{\widehat{\psi}^{(1)}(2^{-k}\xi)}\,\overline{\widehat{\psi}^{(2)}(2^{-k}\eta)}\,\widehat{\psi}^{(3)}(2^{-k}(\xi+\eta))\,e^{2\pi i x(\xi+\eta)}\,d\xi d\eta\end{aligned}$$

provides a smooth localization of $\mathcal{C}_{\Gamma_0}$ to R_k. After a change of variable, therefore,

$$\mathcal{C}_{\Gamma_0}(f,\ g)(x) = \sum_k 2^k \mathcal{C}_{R_k}(f,\ g)(x)$$

where

$$\begin{aligned}\mathcal{C}_{R_k}(f,g)(x) &= \frac{1}{2^k}\int_R m(2^k\xi,\ 2^k\eta)\,\sigma(\xi,\ \eta) \\ &\times\ \widehat{f}(2^k\xi)\,\widehat{g}(2^k\eta)\,\overline{\widehat{\psi}^{(1)}(\xi)}\,\overline{\widehat{\psi}^{(2)}(\eta)}\,\widehat{\psi}^{(3)}(\xi+\eta)\,e^{2\pi i 2^k x(\xi+\eta)}\,d\xi d\eta.\end{aligned}$$

The 'standard' paraproduct decomposition (A.1) now follows taking Fourier series expansions on the $w^{(i)}$. Indeed, on R

$$\widehat{f}(2^k\xi)\,\overline{\widehat{\psi}(\xi)} \;=\; \frac{1}{2^{k/2}|w^{(1)}|}\Big(\sum_{\ell_1=-\infty}^{\infty} \langle\, f,\, \psi^{(1)}_{k\ell_1}\,\rangle\, e^{-2\pi i a_1\ell_1\xi}\Big),$$

while

$$\widehat{g}(2^k\eta)\,\overline{\widehat{\phi}(\eta)} \;=\; \frac{1}{2^{k/2}|w^{(2)}|}\Big(\sum_{\ell_2=-\infty}^{\infty} \langle\, g,\, \psi^{(1)}_{k\ell_2}\,\rangle\, e^{-2\pi i a_2\ell_2\eta}\Big),$$

setting

$$\psi^{(j)}_{k\ell_j}(x) \;=\; 2^{k/2}\psi^{(j)}(2^k x - a_j\,\ell_j), \qquad a_j = 1/|w^{(j)}|.$$

On the other hand,

$$\widehat{\psi}^{(3)}(\xi)\, e^{2\pi i 2^k x\xi} \;=\; \frac{1}{|w^{(3)}|}\Big(\sum_{\ell_3=-\infty}^{\infty} \psi^{(3)}(2^k x - a_3\ell_3)\, e^{2\pi i a_3\ell_3\xi}\Big).$$

After substituting these in $\mathcal{C}_{R_k}(f,\, g)$ the representation (A.1) of $\mathcal{C}_\Gamma(f,\, g)$ then follows exactly as in section 3. We omit the details. Notice that in this this localization all three $\phi^{(j)}$ have vanishing moments.

There is an entirely analogous localization of $\mathcal{C}_{\Gamma_+}$ taking

$$w^{(1)} \;=\; [2,\; 10], \qquad w^{(2)} \;=\; [-1,\; 1], \qquad w^{(3)} \;=\; (0,\; 16).$$

Here $\phi^{(1)}$ and $\phi^{(3)}$ still have vanishing moment but $\phi^{(2)}$ does not. Again we omit the details.

References

1. L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135-157.
2. R. R. Coifman, P. L. Lions, Y. Meyer and S. Semmes, *Compensated compactness and Hardy spaces*, J. Math Pures Appl. **72** (1993), 247–286.
3. R. R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-differentiels*, Astérisque, **57**, Société Mathématique de France (1978).
4. R. R. Coifman and Y. Meyer, *Fourier analysis of multilinear convolutions, Calderón's theorem and analysis of Lipschitz curves*, Lecture Notes in Math., vol. 779, Springer-Verlag, Berlin-Heidelberg-New York, 1979, pp. 104–122.
5. R. R. Coifman and Y. Meyer, *Ondelettes et Opérateurs III, Opérateurs Multilinéaires*, Hermann ed., Paris, 1991.
6. I. Daubechies, H.J. Landau, and Z. Landau, *Gabor time-frequency lattices and the Wexler-Raz identity*, J. Fourier Anal. Appl. **1** (1995), 437–478.
7. C. Fefferman, *Pointwise convergence of Fourier series*, Annals of Math. **98** (1973), 551–571.
8. G.B. Folland, *Harmonic Analysis in Phase Space*, Princeton Univ. Press, Princeton NJ, 1989.

9. 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).

10. 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.

11. J.E. Gilbert and A.R. Nahmod, *L^p-Boundedness of Time-Frequency Paraproducts, II*, preprint (1999).

12. L. Grafakos, and R.H. Torres, *Multilinear Calderón-Zygmund Theory*, preprint (1999).

13. J.A. Hogan and J.D. Lakey, *Extensions of the Heisenberg group by dilations and frames*, Appl. Comp. Harm. Anal. **2** (1995), 174–199.

14. C.E. Kenig and E.M. Stein, *Multilinear estimates and fractional integration*, Math. Research Letters **6** (1999), 1–15.

15. M. Lacey, *The Bilinear Maximal function maps in L^p for $2/3 < p \leq 1$*, Ann. of Math (2) **151** (2000), 35–57.

16. M. Lacey, and C. Thiele, *L^p estimates on the bilinear Hilbert transform,* $2 < p < \infty$, Annals of Math. **146** (1997), 683–724.

17. M. Lacey, and C. Thiele, *On Calderón's conjecture*, Annals of Math. **149** (1999), 475–496.

18. Y. Meyer, *Ondelettes et Opérateurs. II, Opérateurs de Calderón-Zygmund*, Hermann ed. Paris, 1990.

19. C. Muscalu, T. Tao and C. Thiele, *Multilinear operators given by singular multipliers*, preprint (1999).

20. E.M. Stein, *Harmonic Analysis: Real variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, N.J., 1993..

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`