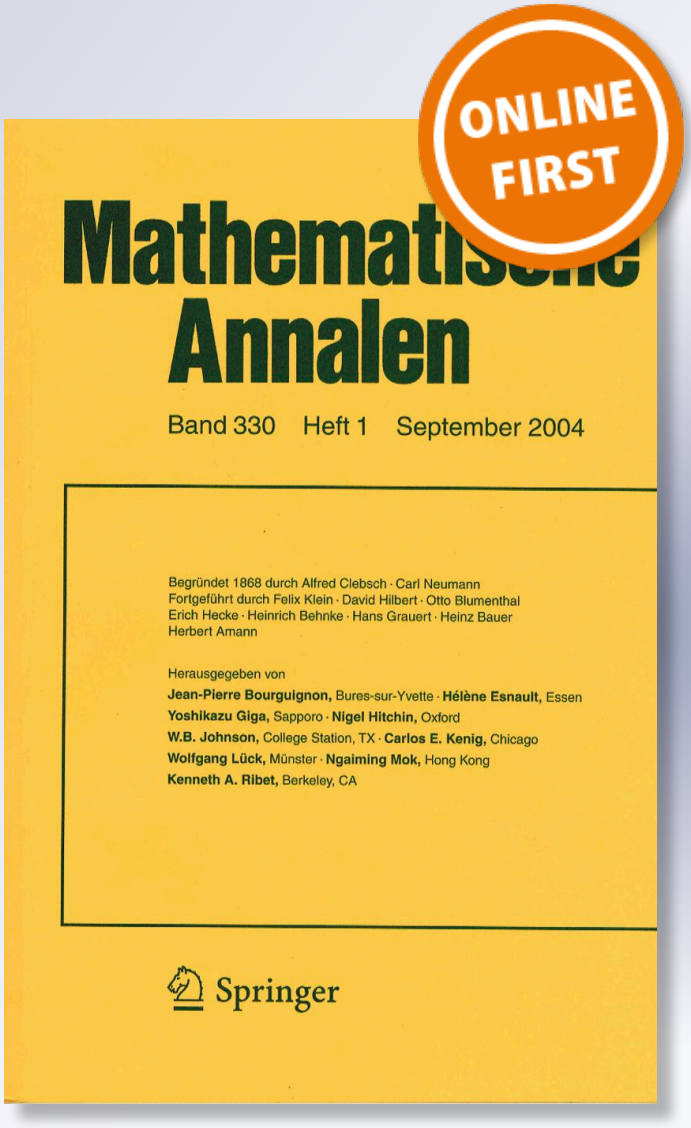
A maximal function for families of Hilbert transforms along homogeneous curves

Shaoming Guo, Joris Roos, Andreas Seeger & Po-Lam Yung

Mathematische Annalen

ISSN 0025-5831

Math. Ann.
DOI 10.1007/s00208-019-01915-3



Springer

Mathematische Annalen
https://doi.org/10.1007/s00208-019-01915-3

Mathematische Annalen

A maximal function for families of Hilbert transforms along homogeneous curves

Shaoming Guo[1] · Joris Roos[1] · Andreas Seeger[1] · Po-Lam Yung[2]

Received: 4 March 2019 / Revised: 17 September 2019

Abstract
Let $H^{(u)}$ be the Hilbert transform along the parabola (t, ut^2) where $u \in \mathbb{R}$. For a set U of positive numbers consider the maximal function $\mathcal{H}^U f = \sup\{|H^{(u)} f| : u \in U\}$. We obtain an (essentially) optimal result for the L^p operator norm of $\mathcal{H}^U$ when $2 < p < \infty$. The results are proved for families of Hilbert transforms along more general nonflat homogeneous curves.

1 Introduction and statement of results

Given $b > 1$, $u > 0$, consider the curve

$$\Gamma_{u,b}(t) = (t, u\gamma_b(t)), \quad t \in \mathbb{R},$$

where γ_b is homogeneous of degree b, with $\gamma_b(\pm 1) \neq 0$.

In memory of Eli Stein.

Communicated by Loukas Grafakos.

✉ Shaoming Guo
shaomingguo@math.wisc.edu

Joris Roos
jroos@math.wisc.edu

Andreas Seeger
seeger@math.wisc.edu

Po-Lam Yung
plyung@math.cuhk.edu.hk

[1] Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Dr, Madison, WI 53706, USA

[2] Department of Mathematics, The Chinese University of Hong Kong, Ma Liu Shui, Shatin, Hong Kong

Published online: 17 October 2019

That is, there are $c_+ \neq 0$, $c_- \neq 0$ such that

$$\gamma_b(t) = \begin{cases} c_+ t^b, & t > 0, \\ c_-(-t)^b, & t < 0. \end{cases} \tag{1.1}$$

For $f \in \mathcal{S}(\mathbb{R}^2)$ the Hilbert transform along $\Gamma_{u,b}$ is defined by

$$H^{(u)} f(x) = p.v. \int_{\mathbb{R}} f(x_1 - t, x_2 - u\gamma_b(t)) \frac{dt}{t}.$$

For an arbitrary nonempty $U \subset \mathbb{R}$ consider the maximal function

$$\mathcal{H}^U f(x) = \sup_{u \in U} |H^{(u)} f(x)|. \tag{1.2}$$

The individual operators $H^{(u)}$ extend to bounded operators on $L^p(\mathbb{R}^2)$ for $1 < p < \infty$ (see [10,27]). The purpose of this paper is to prove, for $p > 2$, optimal L^p bounds for the maximal operator $\mathcal{H}^U$ in terms of suitable properties of U.

Our maximal function is motivated by a similar one involving directional Hilbert transforms which correspond to the limiting case $b = 1$, $c_+ = -c_-$ not covered here. This maximal function for Hilbert transforms along lines was considered by Karagulyan [18] who proved that in this case the $L^2 \to L^{2,\infty}$ operator norm is bounded below by $c\sqrt{\log(\#U)}$; the lower bound was extended to all L^p by Łaba, Marinelli and Pramanik [19]. Demeter and Di Plinio [7] showed the upper bound $O(\log(\#U))$ for $p > 2$ (see also [6] for the sharp L^2 result with bound $O(\log(\#U)))$. Moreover there is a sharp bound $\approx \sqrt{\log(\#U)}$ for lacunary sets of directions (see also Di Plinio and Parissis [9]) and there are other improvements for direction sets of Vargas type. Another motivation for our work comes from the recent papers [8,16] which take up the curved cases and analyze the linear operator $f \mapsto H^{(u(\cdot))} f$ for special classes of measurable functions $x \mapsto u(x)$. [16] covers the case when $u(x)$ depends only on x_1 and [8] covers the case where u is Lipschitz. The analogous questions for variable lines are still not completely resolved (*cf.* [1,2] for partial L^p ranges in the one-variable case, and [15] and the references therein for partial results related to the Lipschitz case).

For our curved variant we seek to get sharp results about the dependence of the operator norm

$$\|\mathcal{H}^U\|_{L^p \to L^p} = \sup\{\|\mathcal{H}^U f\|_p : \|f\|_p \leq 1\}$$

on U. Unlike in the case for lines we obtain for $b > 1$ an optimal bound when $p > 2$ and also observe a different type of dependence on U; namely it is not the cardinality of U that determines the size of the operator norm for the maximal operator but rather the minimal number of intervals of the form $(R, 2R)$ that is needed to cover U. This number is comparable to

$$\mathfrak{N}(U) := 1 + \#\{n \in \mathbb{Z} : [2^n, 2^{n+1}] \cap U \neq \emptyset\}. \tag{1.3}$$

Theorem 1.1 *For every $p \in (2, \infty)$, the operator $\mathcal{H}^U$ is bounded on L^p if and only if $\mathfrak{N}(U) < \infty$. Moreover,*

$$\|\mathcal{H}^U\|_{L^p \to L^p} \approx \sqrt{\log(\mathfrak{N}(U))}\,.$$

The constants implicit in this equivalence depend only on p, b and $|c_+/c_-|$.

Remarks (i) The lower bound $c\sqrt{\log(\mathfrak{N}(U))}$ can be extended to all $p > 1$. Indeed, if we had a smaller operator norm for some $p_0 < 2$ we could, by interpolation, also deduce a better upper bound for $p > 2$ which is not possible. The lower bound for $p < 2$ is generally not efficient, see however some results for lacunary sets in Sect. 7.

(ii) Concerning upper bounds there is no endpoint result for general U with $\mathfrak{N}(U) < \infty$ when $p = 2$. In fact one can show using the Besicovitch set that for $U = [1, 2]$ the operator $\mathcal{H}^U$ even fails to be of restricted weak type $(2, 2)$. Cf. [24, §8.3] for the details of a similar argument in the context of maximal functions for circular means.

(iii) In our theorem we avoid the cases $c_\pm = 0$, for the following reasons. For the case $c_+ = 0 = c_-$ in (1.1) the operators $H^{(u)}$ are equal to the Hilbert transform along a fixed line and the problems on $\mathcal{H}^U$ become trivial. For the choices $c_+ \neq 0$, $c_- = 0$ and $c_+ = 0$, $c_- \neq 0$ the curves are unbalanced and by [5, §6] the individual operators $\mathcal{H}^u$ are not bounded on L^p.

(iv) The operators $\mathcal{H}^U$ are invariant under conjugation with dilation operators with respect to the second variable; i.e. if $\delta_v^{(2)} f(x) = f(x_1, vx_2)$ then we have $\mathcal{H}^{vU} = \delta_{v^{-1}}^{(2)} \mathcal{H}^U \delta_v^{(2)}$ and thus the L^p operator norm of $\mathcal{H}^U$ and $\mathcal{H}^{vU}$ are the same. This shows that any dependence of c_+, c_- in the operator norms can always be reduced to a dependence on just $|c_+/c_-|$ as one can assume that $c_+ = 1$. The implicit constants in the above theorems depend on $c_\pm$, b, p but are uniform as long as $|c_+/c_-|$ is taken in a compact subset of $(0, \infty)$, and b and p are taken in compact subsets of $(1, \infty)$. Thus implicit constants in all inequalities in this paper will be allowed to depend on $c_\pm$, b, with the above understanding of boundedness on compact sets.

This paper

In Sect. 2 we describe the basic decomposition (2.8) of the Hilbert transform $H^{(u)}$ into a standard nonisotropic singular integral operator S^u and two operators $T_\pm^u$ which can be viewed as singular Fourier integral operators with favorable frequency localizations. The growth condition in terms of $\sqrt{\log \mathfrak{N}(U)}$ is only relevant for the maximal function $\sup_{u \in U} |S^u f|$ for which we prove L^p bounds for all $1 < p < \infty$. Here we use the Chang–Wilson–Wolff inequality, together with a variant of an approximation argument in [16]. It turns out that the full maximal operators associated to the $T_\pm^u$ are bounded in $L^p(\mathbb{R}^2)$ for $2 < p < \infty$. This is related to space-time L^p inequalities (so-called local smoothing estimates) for Fourier integral operators in [21]. This connection has already been used by Marletta and Ricci in their work [20] on families of maximal

functions along homogeneous curves. The results for S^u, $T^u_\pm$ are formulated in Sect. 2 as Theorems 2.2 and 2.3.

Section 3 contains several auxiliary results. A version of our maximal function for Mikhlin multipliers (dilated in the second variable) is given in Sect. 4; this is used to prove Theorem 2.2 in Sect. 5. Theorem 2.3 is proved in Sect. 6. In Sect. 7 we prove some results about upper bounds for the maximal functions $\sup_{u\in U}|T^u_\pm f|$ when U is a lacunary set; one of these results will be helpful in the proof of lower bounds for the operator norm.

The proof of lower bounds is given in Sect. 8. The arguments for the lower bounds in L^2 are based on ideas of Karagulyan [18]. "Appendix A" contains a Cotlar type inequality which is used in the proof of Theorem 2.2.

2 Decomposition of the Hilbert transforms

Let χ_+ be supported in $(1/2, 2)$ such that $\sum_{j\in\mathbb{Z}}\chi_+(2^j t) = 1$ for $t > 0$. Let $\chi_-(t) = \chi_+(-t)$ and $\chi = \chi_+ + \chi_-$. We define measures σ_+ and σ_- by

$$\langle\sigma_\pm, f\rangle = \int f(t, \gamma_b(t))\chi_\pm(t)\frac{dt}{t}. \tag{2.1}$$

Let, for $j \in \mathbb{Z}$, the measure σ_j be defined by

$$\langle\sigma_j, f\rangle = \int f(t, \gamma_b(t))\chi(2^j t)\frac{dt}{t}.$$

By homogeneity of γ_b we see that (in the sense of distributions) $\sigma_j = 2^{j(1+b)}\sigma_0(\delta^b_{2^j}\cdot)$ with $\delta^b_t x = (tx_1, t^b x_2)$. Observe that $\sigma_0 = \sigma_+ + \sigma_-$ satisfies the cancellation condition $\widehat{\sigma_0}(0) = 0$ (where $\widehat{\sigma}(\xi) \equiv \mathcal{F}[\sigma](\xi) = \int e^{-i\langle x,\xi\rangle}d\sigma(x)$ denotes the Fourier transform). For Schwartz functions f the Hilbert transform along Γ_b is then given by

$$Hf = \sum_{j\in\mathbb{Z}}\sigma_j * f.$$

2.1 Asymptotics for the Fourier transform of σ_0

We analyze $\widehat{\sigma_\pm}(\xi)$ for large ξ. We have

$$\widehat{\sigma_\pm}(\xi) = \int e^{-i\psi_\pm(t,\xi)}\chi_\pm(t)\frac{dt}{t}$$

with

$$\psi_+(t,\xi) = t\xi_1 + c_+ t^b\xi_2,$$
$$\psi_-(t,\xi) = t\xi_1 + c_-(-t)^b\xi_2.$$

Observe that

$$\begin{aligned}\partial_t \psi_+(t,\xi) &= \xi_1 + c_+ b t^{b-1}\xi_2,\\ \partial_t \psi_-(t,\xi) &= \xi_1 - c_- b(-t)^{b-1}\xi_2.\end{aligned} \tag{2.2}$$

Thus ψ_+ has a critical point $t_+(\xi) > 0$ when $\xi_1/(c_+\xi_2) < 0$, and ψ_- has a critical point $t_-(\xi) < 0$ when $\xi_1/(c_-\xi_2) > 0$, and $t_\pm(\xi)$ are given by

$$t_+(\xi) = \left(\frac{-\xi_1}{bc_+\xi_2}\right)^{\frac{1}{b-1}}, \quad t_-(\xi) = -\left(\frac{\xi_1}{bc_-\xi_2}\right)^{\frac{1}{b-1}}.$$

These critical points are nondegenerate as we have

$$\partial_{tt}\psi_\pm(t,\xi) = c_\pm b(b-1)(\pm t)^{b-2}\xi_2.$$

Setting $\Psi_\pm(\xi) = -\psi_\pm(t_\pm(\xi),\xi)$ we get

$$\Psi_+(\xi) = (b-1)c_+\xi_2\left(-\frac{\xi_1}{bc_+\xi_2}\right)^{\frac{b}{b-1}},$$

$$\Psi_-(\xi) = (b-1)c_-\xi_2\left(\frac{\xi_1}{bc_-\xi_2}\right)^{\frac{b}{b-1}}.$$

The functions $\Psi_\pm$ are homogeneous of degree one and putting $\xi_2 = \pm 1$ we have the crucial lower bounds for the second derivatives of $\xi_1 \mapsto \Psi(\xi_1, \pm 1)$ needed for the application of the space time estimate in Sect. 3.4.

Assume $|\xi| > 1$. We observe that then

$$\inf_{1/3\le t\le 3}\left|\partial_t\psi_+(t,\xi)\right| \gtrsim |\xi| \tag{2.3a}$$

if $\xi_1/c_+\xi_2$ does not belong to the interval $[-b(7/2)^{b-1}, -b(2/7)^{b-1}]$.

Likewise, again for $|\xi| > 1$ we observe that

$$\inf_{-3\le t\le -1/3}\left|\partial_t\psi_-(t,\xi)\right| \gtrsim |\xi| \tag{2.3b}$$

if $\xi_1/c_-\xi_2$ does not belong to the interval $[b(2/7)^{b-1}, b(7/2)^{b-1}]$. These observations suggest the following decomposition of σ_0.

Let η_0 be supported in $\{|\xi| \le 100\}$ and equal to 1 for $|\xi| \le 50$. Let ς_+ be a $C_c^\infty(\mathbb{R})$ function supported on $(b(1/4)^{b-1}, b4^{b-1})$ which is equal to 1 on $[b(2/7)^{b-1}, b(7/2)^{b-1}]$. Let ς_- be a $C_c^\infty(\mathbb{R})$ function supported on $(-b4^{b-1}, -b(1/4)^{b-1})$ which is equal to 1 on $[-b(7/2)^{b-1}, -b(2/7)^{b-1}]$. Then we decompose

$$\sigma_0 = \phi_0 + \mu_{0,+} + \mu_{0,-} \tag{2.4a}$$

where ϕ_0 is given by

$$\widehat{\phi_0}(\xi) = \eta_0(\xi)\widehat{\sigma_0}(\xi) + (1-\eta_0(\xi))\left(1-\varsigma_-\left(\tfrac{\xi_1}{c_+\xi_2}\right)\right)\widehat{\sigma}_+(\xi) + (1-\eta_0(\xi))\left(1-\varsigma_+\left(\tfrac{\xi_1}{c_-\xi_2}\right)\right)\widehat{\sigma}_-(\xi) \tag{2.4b}$$

and $\mu_{0,\pm}$ are given by

$$\widehat{\mu}_{0,+}(\xi) = (1-\eta_0(\xi))\varsigma_-\left(\tfrac{\xi_1}{c_+\xi_2}\right)\widehat{\sigma}_+(\xi), \tag{2.4c}$$

$$\widehat{\mu}_{0,-}(\xi) = (1-\eta_0(\xi))\varsigma_+\left(\tfrac{\xi_1}{c_-\xi_2}\right)\widehat{\sigma}_-(\xi). \tag{2.4d}$$

Lemma 2.1 (i) *ϕ_0 is a Schwartz function with $\widehat{\phi_0}(0) = 0$.*
(ii) *The function $\widehat{\mu}_{0,+}$ is supported on*

$$\mathrm{Sect}_+ = \left\{\xi : |\xi| > 50,\ -b4^{b-1} < \frac{\xi_1}{c_+\xi_2} < -\frac{b}{4^{b-1}}\right\} \tag{2.5a}$$

and satisfies

$$\widehat{\mu}_{0,+}(\xi) = \omega_+(\xi)e^{i\Psi_+(\xi)} + E_+(\xi)$$

where ω_+ is a standard symbol of order $-1/2$, and $E_+(\xi)$ is a Schwartz function, both supported on Sect_+.
(iii) *The function $\widehat{\mu}_{0,-}$ is supported on*

$$\mathrm{Sect}_- = \left\{\xi : |\xi| > 50,\ \frac{b}{4^{b-1}} < \frac{\xi_1}{c_-\xi_2} < b4^{b-1}\right\} \tag{2.5b}$$

and satisfies

$$\widehat{\mu}_{0,-}(\xi) = \omega_-(\xi)e^{i\Psi_-(\xi)} + E_-(\xi)$$

where ω_- is a standard symbol of order $-1/2$, and $E_-(\xi)$ is a Schwartz function, both supported on Sect_-.

Proof. In view of the lower bounds for $\partial_t\psi_\pm$ stated in (2.3a), (2.3b) under their respective assumptions we see that ϕ_0 is a Schwartz function. We have that $\widehat{\sigma}_+(0) = -\widehat{\sigma}_-(0)$ and it follows that $\widehat{\phi_0}(0) = 0$. The formulas for $\widehat{\mu}_{0,\pm}(\xi)$ follow by the method of stationary phase. □

We now define Φ_0 by $\widehat{\Phi}_0 = \widehat{\phi_0} + E_+ + E_-$ so that Φ_0 is a Schwartz function with $\widehat{\Phi}_0(0) = 0$. Define Φ_j , $\kappa_{j,\pm}$ by

$$\widehat{\Phi}_j(\xi) = \widehat{\Phi}_0(2^{-j}\xi_1, 2^{-jb}\xi_2)$$

and

$$\widehat{\kappa_{j,\pm}}(\xi) = \omega_{\pm}(2^{-j}\xi_1, 2^{-jb}\xi_2)e^{i\Psi_{\pm}(2^{-j}\xi_1, 2^{-jb}\xi_2)}.$$

Define operators S^u and $T^u_{\pm}$ by

$$\widehat{S^u f}(\xi) = \sum_{j\in\mathbb{Z}} \widehat{\Phi}_j(\xi_1, u\xi_2)\widehat{f}(\xi) \tag{2.6}$$

$$\widehat{T^u_{\pm} f}(\xi) = \sum_{j\in\mathbb{Z}} \widehat{\kappa_{j,\pm}}(\xi_1, u\xi_2)\widehat{f}(\xi) \tag{2.7}$$

These expressions are at least well defined if f is a Schwartz function whose Fourier transform is compactly supported in $\mathbb{R}^2\backslash\{0\}$. For these functions we have then decomposed our Hilbert transform as

$$H^{(u)}f = S^u f + T^u_+ f + T^u_- f. \tag{2.8}$$

For the upper bound in Theorem 1.1 we shall prove

Theorem 2.2 *For* $1 < p < \infty$,

$$\left\| \sup_{u\in U} |S^u f| \right\|_p \lesssim \sqrt{\log(\mathfrak{N}(U))}\|f\|_p. \tag{2.9}$$

Theorem 2.3 *For* $2 < p < \infty$,

$$\left\| \sup_{u>0} |T^u_{\pm} f| \right\|_p \lesssim \|f\|_p. \tag{2.10}$$

3 Auxiliary results

3.1 The Chang–Wilson–Wolff inequality

We consider the conditional expectation operators $\mathbb{E}_j$ generated by dyadic cubes of length 2^{-j}, i.e. intervals of the form $\prod_{i=1}^d [n_i 2^{-j}, (n_i+1)2^{-j})$ with $n \in \mathbb{Z}^d$. Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$. For each $j \in \mathbb{N}\cup\{0\}$, $\mathbb{E}_j$ is given by

$$\mathbb{E}_j f(x) = \frac{1}{2^{-jd}} \int_{I_j(x)} f(y)dy$$

where $I_j(x)$ is the unique dyadic cube of side length 2^{-j} that contains x. Let

$$\mathbb{D}_j = \mathbb{E}_{j+1} - \mathbb{E}_j$$

be the martingale difference operator. Let $\mathfrak{S} f$ be the dyadic square function, defined by

$$\mathfrak{S} f(x) = \left(\sum_{j \in \mathbb{Z}} |\mathbb{D}_j f(x)|^2 \right)^{1/2}.$$

Also let $\mathcal{M}$ be the dyadic maximal function, given by

$$\mathcal{M} f(x) = \sup_{j \in \mathbb{Z}} |\mathbb{E}_j f(x)|.$$

The following is a slight variant of an inequality due to Chang et al. [4]:

Proposition 3.1 *Suppose that $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ for some $p < \infty$. Then there exist two universal constants c_1 and c_2 such that*

$$\begin{aligned} &\operatorname{meas}\Big(\Big\{ x \in \mathbb{R}^d : |f(x)| > 4\lambda \text{ and } \mathfrak{S} f(x) \le \varepsilon \lambda \Big\} \Big) \\ &\qquad \le c_2 \exp(-c_1 \varepsilon^{-2}) \operatorname{meas}\Big(\Big\{ x \in \mathbb{R}^d : \mathcal{M} f(x) > \lambda \Big\} \Big) \end{aligned} \tag{3.1}$$

for all $\lambda > 0$ and $0 < \varepsilon < 1/2$.

This is a scaling invariant version of the Chang–Wilson–Wolff inequality. For a detailed proof we refer to the arXiv version of our paper (arXiv:1902.00096, "Appendix B").

We shall apply the one-dimensional version of this theorem for the vertical slices in $\mathbb{R}^2$. Let f be a measurable function in $L^p(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$, and for $j \ge 0$, let $\mathbb{E}_j^{(2)}$ be the conditional expectation operator acting on the second variable, i.e.

$$\mathbb{E}_j^{(2)} f(x) = \frac{1}{2^{-j}} \int_{I_j(x_2)} f(x_1, y) dy$$

where $I_j(x_2)$ is the unique dyadic interval of length 2^{-j} that contains x_2. Let $\mathbb{D}_j^{(2)} = \mathbb{E}_{j+1}^{(2)} - \mathbb{E}_j^{(2)}$, and

$$\mathfrak{S}^{(2)} f(x) = \left(\sum_{j \in \mathbb{Z}} |\mathbb{D}_j^{(2)} f(x)|^2 \right)^{1/2}.$$

Then from the above proposition, we clearly have

$$\begin{aligned} &\operatorname{meas}\Big(\Big\{ x \in \mathbb{R}^2 : |f(x)| > 4\lambda \text{ and } \mathfrak{S}^{(2)} f(x) \le \varepsilon \lambda \Big\} \Big) \\ &\quad \le c_2 e^{-c_1 \varepsilon^{-2}} \operatorname{meas}\Big(\Big\{ x \in \mathbb{R}^2 : \mathcal{M}^{(2)} f(x) > \lambda \Big\} \Big) \end{aligned} \tag{3.2}$$

for all $\lambda > 0$ and $0 < \varepsilon < \frac{1}{2}$, where $\mathcal{M}^{(2)}$ is the dyadic maximal function in the second variable, i.e. $\mathcal{M}^{(2)} f(x) = \sup_{j\in\mathbb{Z}} |\mathbb{E}_j^{(2)} f(x)|$.

3.2 Martingale difference operators and Littlewood-Paley projections

We need some computations from [14] which are summarized in the following lemma. Let M denote the Hardy-Littlewood maximal operator acting on functions in $L^p(\mathbb{R})$. Let ϕ be supported in $(c^{-1}, c) \cup (-c, -c^{-1})$ for some $c > 1$.

Lemma 3.2 *Assume that $f \in L^1 + L^\infty(\mathbb{R})$. Then*

(i) *For $q \geq 1$, $n \geq 0$,*

$$\mathbb{E}_k(\mathcal{F}^{-1}[\phi(2^{-k-n}\cdot)\widehat{f}])(x) \lesssim 2^{-n(1-\frac{1}{q})}\Big(M(|f|(x)^q)\Big)^{1/q}$$

(ii) *For $n \geq 0$*

$$\mathbb{D}_k(\mathcal{F}^{-1}[\phi(2^{-k+n}\cdot)\widehat{f}])(x) \lesssim 2^{-n} Mf(x)$$

almost everywhere.

Proof of Lemma 3.2 Cf. Sublemma 4.2 in [14]. □

Given a function on $\mathbb{R}^2$ we shall apply this lemma to $y_2 \mapsto f(y_1, y_2)$ and relate the square function $\mathfrak{S}^{(2)}$ to Littlewood-Paley square functions in the second variable.

Let χ_b be an even C^∞ function supported in $(2^{-b}, 2^b) \cup (-2^b, -2^{-b})$ such that $\sum_{k\in\mathbb{Z}} \chi_b(2^{-kb}t) = 1$ for all $t \neq 0$. Define the Littlewood-Paley projection type operators $P_k^{(1)}$, $P_{k,b}^{(2)}$ acting on Schwartz functions on $\mathbb{R}^2$ by

$$\widehat{P_k^{(1)} f}(\xi) = \chi_1(2^{-k}\xi_1)\widehat{f}(\xi) \tag{3.3}$$

$$\widehat{P_{k,b}^{(2)} f}(\xi) = \chi_b(2^{-kb}\xi_2)\widehat{f}(\xi) \tag{3.4}$$

Lemma 3.3 *Let $q > 1$, $b > 0$, and let $g \in L^1 + L^\infty$. Then the pointwise inequality*

$$\mathfrak{S}^{(2)} g \leq C_{b,q}\left(\sum_{k\in\mathbb{Z}}\Big[M^{(2)}\Big(\Big|P_{k,b}^{(2)} g\Big|^q\Big)\Big]^{2/q}\right)^{1/2}$$

holds almost everywhere. Here $M^{(2)}$ denotes the Hardy-Littlewood maximal operator in the second variable.

Proof of Lemma 3.3 Let ϕ_b be a C^∞ function with

$$\operatorname{supp}(\phi_b) \subset (2^{-b}, 2^b) \cup (-2^b, -2^{-b})$$

which equals 1 on the support of χ_b. Define $\widehat{\tilde{P}^{(2)}_{k,b} f}(\xi) = \phi_b(2^{-kb}\xi_2)\widehat{f}(\xi)$. We write

$$\mathbb{D}^{(2)}_k = \sum_{n\in\mathbb{Z}} \sum_{\substack{l\in\mathbb{Z}:\\ n\le k-lb<n+1}} \mathbb{D}^{(2)}_k \tilde{P}^{(2)}_{l,b} P^{(2)}_{l,b}$$

and use Minkowski's inequality and Lemma 3.2 to estimate, with $\varepsilon < 1 - 1/q$,

$$\begin{aligned}\mathfrak{S}^{(2)} f &\lesssim \sum_{n\in\mathbb{Z}} 2^{-|n|\varepsilon} \left(\sum_{k=0}^{\infty} \left[\sum_{\substack{l\in\mathbb{Z}:\\ n\le k-lb<n+1}} M^{(2)}(|P^{(2)}_{l,b} f|^q) \right]^{2/q} \right)^{1/2} \\ &\lesssim \left(\sum_{l\in\mathbb{Z}} \left[M^{(2)}(|P^{(2)}_{l,b} f|^q) \right]^{2/q} \right)^{1/2}.\end{aligned}$$

This finishes the proof of Lemma 3.3. □

3.3 A variant of Cotlar's inequality

Recall that $\chi_+ \in C^\infty_c(\mathbb{R})$ be supported in $(1/2, 2)$ such that $\sum_{j=-\infty}^{\infty} \chi_+(2^j t) = 1$ for $t > 0$ and let $\eta = \chi_+(|\cdot|)$.

Consider a Mikhlin–Hörmander multiplier m on $\mathbb{R}^d$ satisfying the assumption

$$\sup_{t>0} \|\eta\, m(t\cdot)\|_{\mathscr{L}^1_\alpha} =: B(m) < \infty, \quad \alpha > d; \tag{3.5}$$

here $\mathscr{L}^1_\alpha$ is the potential space of functions g with $(I-\Delta)^{\frac{\alpha}{2}} g \in L^1$. Let $Sf = \mathcal{F}^{-1}[m\widehat{f}]$, and for $n \in \mathbb{Z}$ let S_n be defined by

$$\widehat{S_n f}(\xi) = \sum_{j\le n} \eta(2^{-j}\xi) m(\xi) \widehat{f}(\xi).$$

Then both S and the S_n are of weak type $(1,1)$ and bounded on L^p for $p \in (1,\infty)$ with uniform operator norms $\lesssim_p B(m)$. We are interested in bounds for the maximal function

$$S_* f(x) = \sup_{n\in\mathbb{Z}} |S_n f(x)| \tag{3.6}$$

Proposition 3.4 *Let $\alpha > d$, $r > 0$ and $B(m)$ as in (3.5). For $f \in L^p(\mathbb{R}^d)$, we have, for almost every x, and for $0 < \delta \le 1/2$*

$$S_* f(x) \le \frac{1}{(1-\delta)^{1/r}} \big(M(|Sf|^r)(x)\big)^{1/r} + C_{d,\alpha}\delta^{-1} B(m) Mf(x). \tag{3.7}$$

Proposition 3.4 is a variant of the standard Cotlar inequality regarding truncations of singular integrals. A proof is included in "Appendix A".

3.4 An L^p space time estimate for Fourier integral operators of convolution type and vector valued extensions

Let $S(a_0, a_1)$ be the sectorial region in $\mathbb{R}^2$

$$S(a_0, a_1) = \{(\xi_1, \xi_2) : a_0 < |\xi_1|/|\xi_2| < a_1,\ \xi_2 > 0\}$$

and let η_{sect} be C^∞ and compactly supported in $S_{\text{ann}} := S(a_0, a_1) \cap \{\xi : 1 < |\xi| < 2\}$. Let $q \in C^\infty$ be defined in $S(a_0, a_1)$ and homogeneous of degree one, satisfying

$$q_{\xi\xi} \neq 0 \quad \text{on } S(a_0, a_1)$$

i.e. the Hessian $q_{\xi\xi}$ has rank one on the sector $S(a_0, a_1)$. Model cases for $q(\xi)$ are given by $|\xi|$, or ξ_1^2/ξ_2 in the sector $\{|\xi_1| \le c|\xi_2|\}$. Define

$$F_R f(x,t) = \int e^{i(\langle x,\xi\rangle + tq(\xi))} \eta_{\text{sect}}(\xi/R)\widehat{f}(\xi)d\xi.$$

We need a so-called local smoothing estimate from [21] (the terminology is supposed to indicate that the integration over a compact time interval improves on the fixed time estimate $\|F_R f(\cdot,t)\|_p \lesssim R^{\frac{1}{2}-\frac{1}{p}}\|f\|_p$, $2 \le p < \infty$).

Theorem [21] *If I is a compact interval then*

$$\left(\int_I \int_{\mathbb{R}^2} |F_R f(x,t)|^p dx\, dt\right)^{1/p} \lesssim C_I R^{\frac{1}{2}-\frac{1}{p}-\varepsilon(p)}\|f\|_p, \tag{3.8}$$

with $\varepsilon(p) > 0$ if $2 < p < \infty$. The estimates are uniform as η_{sect} ranges over a bounded subset of C^∞ functions supported in S_{ann}.

In this paper we shall need a square-function extension of (3.8) which involves nonisotropic dilations of the associated multipliers of the form $\xi \mapsto (2^{-j}\xi_1, 2^{-bj}\xi_2)$ with $b \ge 1$, $j \in \mathbb{Z}$ (the strict inequality $b > 1$ assumed in the introduction is not used here); see (6.8) below. We rely on a variant of a theorem in [23], for families of smooth multipliers $\xi \mapsto m(\xi,t)$ on $\mathbb{R}^d$ depending continuously on the parameter $t \in I$, where I is a compact interval. Let $\mathcal{P}$ be a real matrix whose eigenvalues have positive real parts and consider the dilations $\delta_s = \exp(s \log \mathcal{P})$.

Proposition 3.5 *Let $2 < p < \infty$ and $I \subset \mathbb{R}$ be a compact interval. Recall that η is a radial non-trivial C^∞ function with support in $\{\xi : 1/2 < |\xi| < 2\}$. Suppose*

$$\sup_{t\in I}\sup_{\xi} |m(\xi,t)| \le A,$$

and assume that for all $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\sup_{s>0}\left(\frac{1}{|I|}\int_I \left\|\mathcal{F}^{-1}[\eta m(\delta_s\cdot,t)\widehat{f}]\right\|_p^p dt\right)^{1/p} \le A\|f\|_p.$$

Moreover, suppose that for all multiindices α with $|\alpha_1| + |\alpha_2| \le d+1$,

$$\left|\partial_\xi^\alpha [\eta(\xi) m(\delta_s \xi, t)]\right| \le B, \quad t \in I, s > 0.$$

Then there is a constant $C_p > 0$ such that

$$\left(\frac{1}{|I|}\int_I \left\|\mathcal{F}^{-1}[m(\cdot,t)\widehat{f}]\right\|_p^p dt\right)^{1/p} \le C_p A \log(2+B/A)^{1/2-1/p}\|f\|_p. \tag{3.9}$$

The proof is exactly the same as the proof for standard multipliers in [23]. We shall use the following consequence for a square function inequality to derive (6.8).

Corollary 3.6 *Let $2 < p < \infty$ and $I \subset \mathbb{R}$ be a compact interval. Suppose that there is a compact subset $K \subset \mathbb{R}^2\backslash\{0\}$ such that $m_0(\xi,t) = 0$ if $\xi \in K^{\complement}$ or $t \in I^{\complement}$. Suppose that for all multiindices α with $|\alpha_1| + |\alpha_2| \le 10$,*

$$|\partial_\xi^\alpha m_0(\xi,t)| \le B, \quad t \in I,$$

and that

$$\sup_{t\in I}\sup_{\xi} |m_0(\xi,t)| \le A.$$

Moreover, suppose that for all $f \in \mathcal{S}(\mathbb{R}^2)$ the inequality

$$\left(\frac{1}{|I|}\int_I \left\|\mathcal{F}^{-1}[m_0(\cdot,t)\widehat{f}]\right\|_p^p dt\right)^{1/p} \le A\|f\|_p$$

holds. Define $T_j f(x,t)$ by $\widehat{T_j f}(\xi,t) = m_0(\delta_{2^{-j}}\xi, t)\widehat{f}(\xi)$. Then there is a constant $C(K,p)$ such that for all $\{f_j\} \in L^p(\ell^2)$ we also have

$$\left(\frac{1}{|I|}\int_I \left\|\left(\sum_{j\in\mathbb{Z}} |T_j f_j(\cdot,t)|^2\right)^{1/2}\right\|_p^p dt\right)^{1/p}$$
$$\le C(K,p) A \log(2+B/A)^{1/2-1/p} \left\|\left(\sum_j |f_j|^2\right)^{1/2}\right\|_p. \tag{3.10}$$

Proof of Corollary 3.6. This is a straightforward consequence of Proposition 3.5 (alternatively one can adapt the proof of Proposition 3.5 to a vector-valued setting). Let $\widetilde{\phi} \in C_c^\infty(\mathbb{R}^d\backslash\{0\})$ such that $\widetilde{\phi}(\xi) = 1$ for $\xi \in K$. Let $\mathcal{J}$ be a subset of integers with the property that the supports of $\widetilde{\phi}(\delta_{2^{-j}}\cdot)$, $j \in \mathcal{J}$ are disjoint. We may write $\mathbb{Z}$ as union over C_K such families. It is sufficient to show the analogue of (3.10) with the

j-summation extended over $\mathcal{J}$. It will be convenient to work with an enumeration $\{j_1, j_2, \dots\}$ of $\mathcal{J}$.

Let L_j be defined by $\widehat{L_j f} = \widetilde{\phi}(\delta_{2^{-j}}\xi)\widehat{f}(\xi)$. Let $g = \sum_i L_{j_i} f_{j_i}$; then by the adjoint version of the Littlewood-Paley inequality we have

$$\|g\|_p \lesssim \left\| \left(\sum_i |f_{j_i}|^2 \right)^{1/2} \right\|_p . \tag{3.11}$$

Notice that

$$T_j g = T_j f_j \tag{3.12}$$

by the disjointness condition on the supports of $\phi(\delta_{2^{-j_i}}\cdot)$. Let $\{r_i\}_{i=1}^{\infty}$ denote the sequence of Rademacher functions. Applying Proposition 3.5 to the multipliers

$$m_\alpha(\xi) = \sum_{i=1}^{\infty} r_i(\alpha) m_0(\delta_{2^{-j_i}}\xi, t)$$

and the function $g = \sum_{i=1}^{\infty} \mathcal{F}[\widetilde{\phi}(\delta_{2^{-j_i}}\cdot)\widehat{f_{j_i}}]$ we get

$$\left(\int_0^1 \frac{1}{|I|} \int_I \left\| \mathcal{F}^{-1}[m_\alpha(\cdot,t)\widehat{g}] \right\|_p^p dt\, d\alpha \right)^{1/p} \lesssim A \log(2+B/A)^{1/2-1/p} \|g\|_p . \tag{3.13}$$

By interchanging the α-integral and the (x,t)-integral and applying Khintchine's inequality we obtain

$$\left(\frac{1}{|I|} \int_I \left\| \left(\sum_{j\in\mathbb{Z}} |T_j g(\cdot,t)|^2 \right)^{1/2} \right\|_p^p dt \right)^{1/p} \lesssim A \log(2+B/A)^{1/2-1/p} \|g\|_p$$

and the proof is completed by applying (3.11) and (3.12). □

3.5 A version of the Marcinkiewicz multiplier theorem

In the proof of Proposition 7.1 we shall use a well known version of the Marcinkiewicz multiplier theorem with minimal assumptions on the number of derivatives. Let η_{pr} be a nontrivial C_c^∞ function which is even in all variables and supported in $\{\xi : 1/2 < |\xi_i| \le 2, i = 1, 2\}$. Let $\mathscr{L}^2_{\alpha,\alpha}$ the Sobolev space with mixed dominating smoothness consisting of $g \in L^2$ such that

$$\|g\|_{\mathscr{L}^2_{\alpha,\alpha}} = \left(\int (1+|\xi_1|^2)^\alpha (1+|\xi_2|^2)^\alpha |\widehat{g}(\xi)|^2 d\xi \right)^{1/2}$$

is finite. Let $\alpha > 1/2$ and m be a bounded function such that

$$\sup_{t_1>0,t_2>0} \|\eta_{\mathrm{pr}}\, m(t_1\cdot, t_2\cdot)\|_{\mathcal{L}^2_{\alpha,\alpha}} \le B. \tag{3.14}$$

Then we have, for $1 < p < \infty$,

$$\|\mathcal{F}^{-1}[m\widehat{f}]\|_p \le c_p B \|f\|_p. \tag{3.15}$$

One can prove this using a straightforward product-type modification of Stein's proof of the Mikhlin–Hörmander multiplier theorem in [25, §3]. One can also deduce it from Fefferman's theorem [12], cf. [3,13].

4 Some maximal function estimates for families of Mikhlin type multipliers on $\mathbb{R}^2$

In this section we consider Mikhlin–Hörmander multipliers with respect to the dilation group δ_t^b, $b > 0$, with $\delta_t^b(\xi) = (t\xi_1, t^b\xi_2)$.

Theorem 4.1 *Suppose that*

$$\sup_{t>0} \sum_{|\alpha|\le 4} \left\| \partial^\alpha \left(\eta(\cdot)a(\delta_t^b\cdot)\right)\right\|_{L^1(\mathbb{R}^2)} \le 1 \tag{4.1}$$

Define, for $n \in \mathbb{Z}$ the operator T_n by

$$\widehat{T_n f}(\xi) = a(\xi_1, 2^{bn}\xi_2)\widehat{f}(\xi). \tag{4.2}$$

Let $\mathcal{N}$ be a subset of $\mathbb{Z}$ with $\#\mathcal{N} = N$. Then for $1 < p < \infty$,

$$\left\| \sup_{n\in\mathcal{N}} |T_n f| \right\|_p \le C_p \sqrt{\log(1+N)}\, \|f\|_p. \tag{4.3}$$

By the Marcinkiewicz interpolation theorem it suffices to show that there is $A = A(p)$ such that the inequality

$$\mathrm{meas}\left(\{x : \sup_{n\in\mathcal{N}} |T_n f| > 4\lambda\}\right) \le \left(A\sqrt{\log(1+N)}\lambda^{-1}\|f\|_p\right)^p \tag{4.4}$$

holds for all Schwartz functions f whose Fourier transform is compactly supported in $\mathbb{R}^2\backslash\{0\}$, all $\lambda > 0$ and all $\mathcal{N}$ with $\#\mathcal{N} \le N$.

One can decompose

$$a(\xi_1, \xi_2) = \sum_{j\in\mathbb{Z}} a_j(2^{-j}\xi_1, 2^{-bj}\xi_2) \tag{4.5}$$

where each a_j is supported in $\{(\xi_1,\xi_2) : 1/2 < |\xi_1| + |\xi_2|^{1/b} < 2\}$ and

$$\sup_j \int \left|\partial_\xi^\alpha a_j(\xi)\right| d\xi \le C_\alpha, \quad |\alpha| \le 4.$$

We shall repeatedly use that the operators T_n are bounded on $L^p(\mathbb{R}^2)$ with norm independent of n. This follows by the Mikhlin-Hörmander multiplier theorem and rescaling in the second variable.

Let $\mathcal{T}_{\mathcal{N}} f := \sup_{n\in\mathcal{N}} |T_n f|$ and set

$$\varepsilon_N := (\log(C_1 N))^{-1/2} \tag{4.6}$$

where $C_1 > c_1^{-1}$ with c_1 as in (3.2), also $\varepsilon_N < 1/2$. Since f is a Schwartz function, with $\widehat{f}$ compactly supported in $\mathbb{R}^2\setminus\{0\}$ the function $\mathcal{T}_{\mathcal{N}} f$ is in $L^\infty \cap L^2$ which allows us to apply the Chang-Wilson-Wolff inequality.

We have that

$$\begin{aligned}
&\operatorname{meas}\big(\{x\in\mathbb{R}^2 : \mathcal{T}_{\mathcal{N}} f(x) > 4\lambda\}\big)\\
&\quad\le \sum_{n\in\mathcal{N}} \operatorname{meas}\big(\{x\in\mathbb{R}^2 : |T_n f(x)| > 4\lambda,\ \mathfrak{S}^{(2)} T_n f(x) \le \varepsilon_N \lambda\}\big)\\
&\qquad + \operatorname{meas}\left(\left\{x\in\mathbb{R}^2 : \sup_{n\in\mathcal{N}} |\mathfrak{S}^{(2)}[T_n f](x)| > \varepsilon_N\lambda\right\}\right).
\end{aligned} \tag{4.7}$$

By the Chang–Wilson–Wolff inequality (3.2), the first term on the right hand side of (4.7) is bounded by

$$\begin{aligned}
&c_2 N e^{-c_1\varepsilon_N^{-2}} \max_{n\in\mathcal{N}} \operatorname{meas}\big(\{x\in\mathbb{R}^2 : \mathcal{M}^{(2)}[T_n f] > \lambda\}\big)\\
&\quad\le c_2 N e^{-c_1\varepsilon_N^{-2}} \max_{n\in\mathcal{N}} \lambda^{-p} \|\mathcal{M}^{(2)}[T_n f]\|_p^p \lesssim N e^{-c_1\varepsilon_N^{-2}} \lambda^{-p}\|f\|_p^p \lesssim \lambda^{-p}\|f\|_p^p
\end{aligned}$$

where we used that $N e^{-c_1\varepsilon_N^{-2}} \le 1$ (by (4.6)) and that the operators T_n are uniformly bounded.

By Chebyshev's inequality the second term on the right hand side of (4.7) is bounded by

$$\begin{aligned}
&\varepsilon_N^{-p}\lambda^{-p} \left\| \sup_{n\in\mathcal{N}} \mathfrak{S}^{(2)}[T_n f]\right\|_{L^p}^p\\
&\quad\lesssim \varepsilon_N^{-p}\lambda^{-p} \left\| \sup_{n\in\mathcal{N}} \left(\sum_{k\in\mathbb{Z}} \left[M^{(2)}(|T_n P_{k,b}^{(2)} f|^q)\right]^{2/q}\right)^{1/2}\right\|_p^p.
\end{aligned}$$

Here we have used Lemma 3.3 with $g = T_n f$ and the fact that the operators T_n and $P_{k,b}^{(2)}$ commute; q will be chosen so that $1 < q < p$.

We shall now use an idea in [16] and approximate the operators T_n by a convolution operator acting in the first variable. Define $T^{(1)}$ by

$$\widehat{T^{(1)}f}(\xi_1,\xi_2)=\sum_{j\in\mathbb{Z}}a_j(2^{-j}\xi_1,0)\widehat{f}(\xi_1,\xi_2).$$

Recall the definition of χ_b in Lemma 3.3. Notice also that

$$a_j(2^{-j}\xi_1,2^{(n-j)b}\xi_2)\chi_b(2^{-kb}\xi_2)\equiv 0$$

if $j<n+k-1$ and therefore we have

$$\begin{aligned}T_nP^{(2)}_{k,b}f&=\sum_{j\geq n+k-1}\mathcal{F}^{-1}[a_j(2^{-j}\cdot,2^{(n-j)b}\cdot)]*P^{(2)}_{k,b}f\\&=\sum_{j\geq n+k-1}\mathcal{F}^{-1}[a_j(2^{-j}\cdot,0)]*P^{(2)}_{k,b}f \qquad (4.8a)\\&\quad+\sum_{j\geq n+k-1}\mathcal{F}^{-1}[a_j(2^{-j}\cdot,2^{(n-j)b}\cdot)-a_j(2^{-j}\cdot,0)]*P^{(2)}_{k,b}f. \qquad (4.8b)\end{aligned}$$

For the first term (4.8a) we use the one-dimensional version of Proposition 3.4 to get

$$\left|\sum_{j\geq n+k-1}\mathcal{F}^{-1}[a_j(2^{-j}\cdot,0)]*P^{(2)}_{k,b}f\right|\lesssim M^{(1)}\left(P^{(2)}_{k,b}f\right)+M^{(1)}\left(T^{(1)}P^{(2)}_{k,b}f\right). \qquad (4.9)$$

Here $M^{(1)}$ denotes the Hardy–Littlewood maximal operator acting on the first variable.

Now consider the second term (4.8b). Let $\tilde{\phi}$ be an appropriately chosen non-negative bump function supported in $(1/4,3)\cup(-3,-1/4)$ and let $K_{j,k,n}$ be the convolution kernel with multiplier

$$\widehat{K_{j,k,n}}(\xi)=\tilde{\phi}(2^{-kb}\xi_2)\big(a_j(2^{-j}\xi_1,2^{(n-j)b}\xi_2)-a_j(2^{-j}\xi_1,0)\big).$$

Then

$$\widehat{K_{j,k,n}}(2^j\xi_1,2^{kb}\xi_2)=2^{(k+n-j)b}\tilde{\phi}(\xi_2)\xi_2\int_0^1\partial_2a_j(\xi_1,2^{(k+n-j)b}s\xi_2)\,ds$$

and we have $\big\|\partial^\alpha\big(\widehat{K_{j,k,n}}(2^j\cdot,2^{kb}\cdot)\big)\big\|_1\lesssim 2^{(k+n-j)b}$ for multiindices $|\alpha|\leq 3$. This implies

$$|K_{j,k,n}(x)|\lesssim 2^{(k+n-j)b}\frac{2^{j+kb}}{(1+2^j|x_1|+2^{kb}|x_2|)^3}$$

and hence

$$\sum_{j\geq n+k-1}\left|K_{j,k,n}*P^{(2)}_{k,b}f(x)\right|\lesssim M_{\rm str}\left(P^{(2)}_{k,b}f\right)(x)$$

where $M_{\rm str}$ is the strong maximal operator which is controlled by $M^{(2)}\circ M^{(1)}$.

Combining the estimates we thus see that the second term on the right hand side of (4.7) is bounded by

$$\varepsilon_N^{-p}\lambda^{-p}\left(\left\|\left(\sum_{k\in\mathbb{Z}}\left[M^{(2)}(|M^{(2)}M^{(1)}P^{(2)}_{k,b}f|^q)\right]^{2/q}\right)^{1/2}\right\|_p\right.$$
$$\left.+\left\|\left(\sum_{k\in\mathbb{Z}}\left[M^{(2)}(|M^{(2)}M^{(1)}T^{(1)}P^{(2)}_{k,b}f|^q)\right]^{2/q}\right)^{1/2}\right\|_p\right)^p.$$

We use this with $1<q<p$ and apply Fefferman-Stein estimates for the vector-valued versions of $M^{(1)}$ and $M^{(2)}$ and the Marcinkiewicz-Zygmund theorem on $L^p(\ell^2)$ boundedness applied to the operator $T^{(1)}$. Consequently the last expression can be bounded by

$$C_p^p\varepsilon_N^{-p}\lambda^{-p}\|f\|_p^p\lesssim C_p^p(\log(1+N))^{p/2}\lambda^{-p}\|f\|_p^p,$$

by the definition of ε_N. This finishes the proof of (4.4) and thus the proof of Theorem 4.1. □

5 Proof of Theorem 2.2

We decompose $\Phi_0=\sum_{l\in\mathbb{Z}}\Phi_{0,l}$ where $\widehat{\Phi}_{0,l}(\xi)=\chi_+(2^{-l}|\xi|)\widehat{\Phi}_0(\xi)$. Define

$$a_{0,l}(\xi)=\widehat{\Phi}_{0,l}(2^l\xi),$$
$$\widetilde{a}_{0,l,s}(\xi)=s^b\xi_2\frac{\partial\widehat{\Phi}_{0,l}}{\partial\xi_2}(2^l\xi_1,2^ls^b\xi_2).$$

Then the functions $a_{0,l}$ and $\widetilde{a}_{0,l,s}$, for every $s\in(1/2,2)$, are supported in $\{\xi:10^{-b}<|\xi|<10^b\}$ and satisfy the estimates

$$\int\left|\partial_\xi^\alpha a_{0,l}(\xi)\right|d\xi+\int\left|\partial_\xi^\alpha\widetilde{a}_{0,l,s}(\xi)\right|d\xi\leq C2^{-|l|}$$

for all multiindices α with $|\alpha_1|+|\alpha_2|\leq 10$. This means that there is a $c>0$ such that the multipliers

$$\begin{aligned} a_l(\xi) &= c2^{|l|}\sum_{j\in\mathbb{Z}} a_{0,l}(2^{-j}\xi_1, 2^{-jb}\xi_2), \\ \widetilde{a}_{l,s}(\xi) &= c2^{|l|}\sum_{j\in\mathbb{Z}} a_{0,l,s}(2^{-j}\xi_1, 2^{-jb}\xi_2) \end{aligned} \tag{5.1}$$

satisfy the conditions (4.1) in Theorem 4.1. Now define operators S_l^u and R_l^u

$$\begin{aligned} \widehat{S_l^u f}(\xi) &= \sum_{j\in\mathbb{Z}} \widehat{\Phi_{0,l}}(2^{-j}\xi_1, 2^{-jb}u\xi_2)\widehat{f}(\xi), \\ \widehat{R_l^u f}(\xi) &= \sum_{j\in\mathbb{Z}} \widehat{\Phi_{0,l}}(2^{l-j}\xi_1, 2^{l-jb}u\xi_2)\widehat{f}(\xi). \end{aligned}$$

The assertion of the theorem follows if we can prove

$$\left\| \sup_{u\in U} |S_l^u f| \right\|_p \lesssim 2^{-|l|}\sqrt{\log \mathfrak{N}(U)}\|f\|_p$$

which follows by isotropic rescaling from

$$\left\| \sup_{u\in U} |R_l^u f| \right\|_p \lesssim 2^{-|l|}\sqrt{\log \mathfrak{N}(U)}\|f\|_p. \tag{5.2}$$

Now let

$$\mathcal{N} = \{n \in \mathbb{Z} : \exists s \in (1/2, 2) \text{ such that } (2^n s)^b \in U\}.$$

Observe that $\#\mathcal{N} \le C(b)\mathfrak{N}(U)$. The inequality (5.2) follows from

$$\left\| \sup_{n\in\mathcal{N}} \sup_{1/2<s<2} |R_l^{(2^n s)^b} f| \right\|_p \lesssim 2^{-|l|}\sqrt{\log(1+\#\mathcal{N})}\|f\|_p$$

which is a consequence of

$$\left\| \sup_{n\in\mathcal{N}} |R_l^{2^{nb}} f| \right\|_p \lesssim 2^{-|l|}\sqrt{\log(1+\#\mathcal{N})}\|f\|_p \tag{5.3}$$

and

$$\int_{1/2}^{2} \left\| \sup_{n\in\mathcal{N}} \left| \frac{\partial}{\partial s} R_l^{(2^n s)^b} f \right| \right\|_p ds \lesssim 2^{-|l|}\sqrt{\log(1+\#\mathcal{N})}\|f\|_p . \tag{5.4}$$

Since

$$\mathcal{F}[R_l^{2^{nb}} f](\xi) = \sum_j a_{0,l}(2^{-j}\xi_1, 2^{nb-jb}\xi_2)\widehat{f}(\xi),$$
$$\mathcal{F}[\partial_s R_l^{(2^n s)^b} f](\xi) = \frac{b}{s}\sum_j a_{0,l,s}(2^{-j}\xi_1, 2^{nb-jb}\xi_2)\widehat{f}(\xi),$$

the inequalities (5.3) and (5.4) follow by applying Theorem 4.1 to the multipliers in (5.1). □

6 Proof of Theorem 2.3

We only consider the maximal function for the operator T_+^u, since the analogous problem for T_-^u can be reduced to the former one by a change of variable (with a different curve). We omit the subscript and set $T^u = T_+^u$.

Decompose $\kappa_{0,+} = \sum_{\ell=0}^{\infty} \kappa_{0,\ell}$ where

$$\widehat{\kappa_{0,\ell}}(\xi) = \chi_+(2^{-\ell}|\xi|)\omega_+(\xi)e^{i\Psi_+(\xi)}.$$

Notice that, by Lemma 2.1, $|\xi_1| \approx |\xi_2| \approx 2^\ell$ for $\xi \in \operatorname{supp}(\widehat{\kappa_{0,\ell}})$. Define $\kappa_{j,\ell}$ by $\widehat{\kappa_{j,\ell}}(\xi) = \widehat{\kappa_{0,\ell}}(2^{-j}\xi_1, 2^{-jb}\xi_2)$ and define $T_{j,\ell}^u$ by

$$\widehat{T_{j,\ell}^u f}(\xi) = \widehat{\kappa_{j,\ell}}(\xi_1, u\xi_2)\widehat{f}(\xi). \tag{6.1}$$

Then we have $T^u = \sum_{\ell\geq 0}\sum_{j\in\mathbb{Z}} T_{j,\ell}^u$.

The assertion of the theorem follows if we can show, for $2 < p < \infty$, that there exists some $\varepsilon = \varepsilon(p) > 0$ with

$$\left\| \sup_{n\in\mathbb{Z}} \sup_{1/2<s<2} \left| \sum_{j\in\mathbb{Z}} T_{j,\ell}^{(2^n s)^b} f \right| \right\|_p \lesssim 2^{-\ell\varepsilon}\|f\|_p. \tag{6.2}$$

Define $\mathcal{R}_{j,\ell}^u$ by

$$\widehat{\mathcal{R}_{j,\ell}^u f}(\xi) = \widehat{\kappa_{0,\ell}}(2^{\ell-j}\xi_1, 2^{\ell-jb}u\xi_2)\widehat{f}(\xi).$$

By isotropic rescaling inequality (6.2) is equivalent with

$$\left\| \sup_{n\in\mathbb{Z}} \sup_{1/2<s<2} \left| \sum_{j\in\mathbb{Z}} \mathcal{R}_{j,\ell}^{(2^n s)^b} f \right| \right\|_p \lesssim 2^{-\ell\varepsilon}\|f\|_p. \tag{6.3}$$

This inequality follows, by the embedding $\ell^p \subset \ell^\infty$ and Fubini's theorem from

$$\left(\sum_{n\in\mathbb{Z}}\left\|\sup_{1/2<s<2}\left|\sum_{j\in\mathbb{Z}}\mathcal{R}_{j,\ell}^{(2^n s)^b}f\right|\right\|_p^p\right)^{1/p}\lesssim 2^{-\ell\varepsilon}\|f\|_p \tag{6.4}$$

Fix n, x and set $G(s)=\sum_j \mathcal{R}_{j,\ell}^{(2^n s)^b}f(x)$. We use the standard argument of applying the fundamental theorem of calculus to $|G(s)|^p$ and then Hölder's inequality which gives

$$|G(s)|^p \le |G(1)|^p + p\left(\int_{1/2}^2 |G(s)|^p ds\right)^{1/p'}\left(\int_{1/2}^2 |G'(s)|^p ds\right)^{1/p}.$$

This inequality and another application of Hölder's inequality in $\mathbb{R}^2$ shows that (6.4) follows from

$$\left(\sum_{n\in\mathbb{Z}}\int_{1/2}^2\left\|\sum_j \mathcal{R}_{j,\ell}^{(2^n s)^b}f\right\|_p^p ds\right)^{1/p}\lesssim 2^{-\ell(\varepsilon+1/p)}\|f\|_p, \tag{6.5a}$$

$$\left(\sum_{n\in\mathbb{Z}}\int_{1/2}^2\left\|\frac{\partial}{\partial s}\Big(\sum_j \mathcal{R}_{j,\ell}^{(2^n s)^b}f\Big)\right\|_p^p ds\right)^{1/p}\lesssim 2^{\ell-\ell(\varepsilon+1/p)}\|f\|_p \tag{6.5b}$$

and

$$\left(\sum_{n\in\mathbb{Z}}\left\|\sum_j \mathcal{R}_{j,\ell}^{2^{nb}}f\right\|_p^p\right)^{1/p}\lesssim 2^{-\ell/p}\|f\|_p \tag{6.5c}$$

for $2<p<\infty$.

We focus on the derivation of the inequality (6.5a). Note that for $s\in[1/2,2]$

$$\begin{aligned}\widehat{\kappa_{0,\ell}}(\xi_1, s^b\xi_2) &= \omega_+(\xi_1, s^b\xi_2)\chi_+(2^{-\ell}|(\xi_1, s^b\xi_2)|)e^{i\Psi_+(\xi_1, s^b\xi_2)}\\ &= 2^{-\ell/2}\eta_{\ell,s}(2^{-\ell}\xi)e^{-is^{\frac{b}{b-1}}\Psi_+(\xi_1,\xi_2)}\end{aligned}$$

where

$$\eta_{\ell,s}(\xi_1,\xi_2)=2^{\ell/2}\omega_+(2^\ell\xi_1, 2^\ell s^b\xi_2)\chi_+(|(\xi_1, s^b\xi_2)|)$$

and taking into account that ω_+ is a symbol of order $-1/2$ we see that the $\eta_{\ell,s}$ belong to a bounded set of C^∞ functions supported in an annulus $\{\xi : a_0\le|\xi|\le a_0^{-1}\}$, for fixed $a_0=a_0(b)<1$.

After changing variables $t = s^{-\frac{b}{b-1}}$, with $t \in (2^{-\frac{b}{b-1}}, 2^{\frac{b}{b-1}})$ this puts us in the position to apply (3.8) with $R = 2^\ell$ and we obtain, with suitable $\varepsilon' = \varepsilon'(p) > 0$

$$\left(\int_{1/2}^2 \left\| \mathcal{F}^{-1}[\widehat{\kappa_{0,\ell}}(\xi_1, s^b\xi_2)\widehat{f}]\right\|_p^p ds\right)^{1/p} \lesssim 2^{-\ell(\varepsilon'+1/p)}\|f\|_p.$$

By isotropic scaling, replacing $\widehat{\kappa_{0,\ell}}(\xi_1, s^b\xi_2)$ with $\widehat{\kappa_{0,\ell}}(2^\ell\xi_1, s^b 2^\ell\xi_2)$, we also have

$$\left(\int_{1/2}^2 \left\|\mathcal{R}_{0,\ell}^{s^b} f\right\|_p^p ds\right)^{1/p} \le C_\varepsilon 2^{-\ell(\varepsilon'+1/p)}\|f\|_p. \tag{6.6}$$

Let

$$m_{j,\ell}(\xi, s) = \widehat{\kappa_{0,\ell}}\left(2^{\ell-j}\xi_1, s^b 2^{\ell-jb}\xi_2\right)$$

and observe $\widehat{\mathcal{R}_{j,\ell}^{s^b} f}(\xi) = m_{j,\ell}(\xi, s)\widehat{f}(\xi)$. The functions $\xi \mapsto m_{0,\ell}(\xi, s)$ are supported in a fixed annulus and satisfy

$$\left|\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2} m_{0,\ell}(\xi, s)\right| \lesssim 2^{\ell(\alpha_1+\alpha_2)}. \tag{6.7}$$

By Corollary 3.6 we get the inequality

$$\begin{aligned}&\left(\int_{1/2}^2 \left\|\left(\sum_{j\in\mathbb{Z}} |\mathcal{R}_{j,\ell}^{s^b} f_j|^2\right)^{1/2}\right\|_p^p ds\right)^{1/p}\\ &\lesssim 2^{-\ell(\varepsilon'+1/p)}(1+\ell)^{1/2-1/p}\left\|\left(\sum_j |f_j|^2\right)^{1/2}\right\|_p.\end{aligned} \tag{6.8}$$

We can replace the multipliers $m_{j,\ell}(\xi_1, \xi_2, s)$ by $m_{j,\ell}(\xi_1, 2^{nb}\xi_2, s)$, after scaling in the second variable. This means that for every fixed n we have proved, for $\varepsilon < \varepsilon'$,

$$\left(\int_{1/2}^2 \left\|\left(\sum_j |\mathcal{R}_{j,\ell}^{(2^n s)^b} f_j|^2\right)^{1/2}\right\|_p^p ds\right)^{1/p} \lesssim 2^{-\ell(\varepsilon+1/p)}\left\|\left(\sum_j |f_j|^2\right)^{1/2}\right\|_p, \tag{6.9}$$

with the implicit constant independent of n.

We now combine this with Littlewood-Paley inequalities to prove (6.5a). Let $\widetilde{\chi}^{(1)}$ be an even C^∞ function supported on $\{\xi_1 : |c_+|b2^{-3b-1} \le |\xi_1| \le |c_+|b2^{3b+1}\}$ and equal to 1 for $|c_+|b2^{-3b} \le |\xi_1| \le |c_+|b2^{3b}$. Let $\widetilde{\chi}_b^{(2)}$ be an even C^∞ function supported on

$\{\xi_2 : 2^{-2b-1} \le |\xi_2| \le 2^{2b+1}\}$ and equal to 1 for $2^{-2b} \le |\xi_2| \le 2^{2b}$. Define $\tilde{P}_j^{(1)}$, $\tilde{P}_{j,b}^{(2)}$ by

$$\widehat{\tilde{P}_j^{(1)} f}(\xi) = \widetilde{\chi}^{(1)}(2^{-j}\xi_1)\widehat{f}(\xi)$$
$$\widehat{\tilde{P}_{j,b}^{(2)} f}(\xi) = \widetilde{\chi}^{(2)}(2^{-jb}\xi_2)\widehat{f}(\xi)$$

Then by the support properties of $\widehat{\kappa_{0,\ell}}(2^\ell \cdot)$ we get for $1/2 \le s \le 2$

$$\mathcal{R}_{j,\ell}^{(2^n s)^b} = \tilde{P}_j^{(1)} \tilde{P}_{j-n,b}^{(2)} \mathcal{R}_{j,\ell}^{(2^n s)^b} \tilde{P}_{j-n,b}^{(2)} \tilde{P}_j^{(1)}. \tag{6.10}$$

Hence, by Littlewood-Paley theory

$$\left(\sum_{n\in\mathbb{Z}} \int_{1/2}^2 \left\| \sum_j \mathcal{R}_{j,\ell}^{(2^n s)^b} f \right\|_p^p ds\right)^{1/p}$$
$$\lesssim \left(\sum_{n\in\mathbb{Z}} \int_{1/2}^2 \left\| \left(\sum_j |\mathcal{R}_{j,\ell}^{(2^n s)^b} \tilde{P}_{j-n,b}^{(2)} \tilde{P}_j^{(1)} f|^2\right)^{1/2} \right\|_p^p ds\right)^{1/p}$$

and by (6.9) this is controlled by

$$2^{-\ell(\varepsilon(p)+1/p)} \left(\sum_{n\in\mathbb{Z}} \left\| \left(\sum_{j\in\mathbb{Z}} |\tilde{P}_{j-n,b}^{(2)} \tilde{P}_j^{(1)} f|^2\right)^{1/2} \right\|_p^p\right)^{1/p}$$

for some $\varepsilon(p) > 0$ when $2 < p < \infty$. We finish the proof of (6.5a) by observing that

$$\left(\sum_{n\in\mathbb{Z}} \left\| \left(\sum_{j\in\mathbb{Z}} |\tilde{P}_{j-n,b}^{(2)} \tilde{P}_j^{(1)} f|^2\right)^{1/2} \right\|_p^p\right)^{1/p} \le \left\| \left(\sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} |\tilde{P}_{j-n,b}^{(2)} \tilde{P}_j^{(1)} f|^2\right)^{1/2} \right\|_p$$
$$= \left\| \left(\sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} |\tilde{P}_{k_2,b}^{(2)} \tilde{P}_{k_1}^{(1)} f|^2\right)^{1/2} \right\|_p \lesssim \|f\|_p$$

where we have used the embedding $\ell^2 \hookrightarrow \ell^p$ for $p > 2$, and applied a two-parameter Littlewood-Paley inequality.

We now turn to the estimate (6.5b). A computation shows

$$2^{-\ell}\frac{\partial}{\partial s}\left(\sum_j \mathcal{F}[\mathcal{R}_{j,\ell}^{(2^n s)^b} f](\xi)\right)$$
$$= \widehat{f}(\xi)\frac{b}{s}\sum_j \upsilon_\ell\left(2^{-j}\xi_1, s^b 2^{(n-j)b}\xi_2\right) e^{i2^\ell \Psi_+(2^{-j}\xi_1, s^b 2^{(n-j)b}\xi_2)} \tag{6.11a}$$

where

$$\upsilon_\ell(\xi) = 2^{-\ell}\chi'_+(|\xi|)\frac{\xi_2^2}{|\xi|}\omega_+(2^\ell\xi_1, 2^\ell\xi_2) + \chi_+(|\xi|)\xi_2\frac{\partial\omega_+}{\partial\xi_2}(2^\ell\xi_1, 2^\ell\xi_2)$$
$$+\chi_+(\xi)\omega_+(2^\ell\xi_1, 2^\ell\xi_2) i\xi_2\frac{\partial\Psi_+}{\partial\xi_2}(\xi_1,\xi_2). \tag{6.11b}$$

Here the main contribution in (6.11b) comes from the third term (the others are similar but better by a factor of about $2^{-\ell}$).

It is now straightforward to check that in the proof of (6.5a) the term $\mathcal{R}_{j,\ell}^{(2^n s)^b} f$ can be replaced with $2^{-\ell}\partial_s(\mathcal{R}_{j,\ell}^{(2^n s)^b} f)$ and one obtains (6.5b).

Finally, a simple modification of the proof of (6.5a) would also prove (6.5c): in place of (3.8), one would use a fixed time estimate, as stated immediately before (3.8). This finishes the proof of Theorem 2.3.

7 Maximal functions for lacunary sets

We shall prove some upper bounds for the operator norm of $\mathcal{H}^U$ for *lacunary* sets.

Definition Let $\kappa > 1$. A finite set U is called κ-lacunary if it can be arranged in a sequence $U = \{u_1 < u_2 < \cdots < u_M\}$ where $u_{j+1} \le u_j/\kappa$ for $j = 1, \ldots, M-1$. U is lacunary if U is κ-lacunary for some $\kappa > 1$.

Note that for lacunary sets we have $\#U \approx \mathfrak{N}(U)$ (with the implicit constant depending on κ).

Proposition 7.1 *Let U be a lacunary set. Then, for $4/3 < p < \infty$*

$$\|\mathcal{H}^U\|_{L^p\to L^p} \lesssim \sqrt{\log(1+(\#U))}\,. \tag{7.1}$$

Proposition 7.1 will be used in the proof of lower bounds in Sect. 8. For this application it is important that (7.1) just holds for some $p < 2$. We do not know at this time whether the result extends to all $p > 1$.[1] For special lacunary sequences it does:

[1] Added in September 2019: After the submission of this paper the authors showed the bound of Proposition 7.1 for general lacunary sets U, in the full range $1 < p < \infty$. This result can be found in the paper [17] which also contains L^p results, $p < 2$, for more general sets U, under suitable dimension assumptions.

Proposition 7.2 *Let U be a subset of $\{2^{nb} : n \in \mathbb{Z}\}$. Then, for $1 < p < \infty$*

$$\|\mathcal{H}^U\|_{L^p \to L^p} \lesssim \sqrt{\log(1 + (\#U))}\,.$$

Here b is as in the definition of the curve γ_b in (1.1).

7.1 Proof of Proposition 7.1

We may assume that for every interval $I_n := [2^{nb}, 2^{(n+1)b})$, $n \in \mathbb{Z}$, there is at most one $u \in U \cap I_n$. This is because of the lacunarity assumption we can split U in $O(1)$ many sets with this assumption.

We order $U = \{u_\nu\}$ such that $u_\nu < u_{\nu+1}$ and let $n(\nu)$ be the unique integer for which $u_\nu \in I_n$.

We split $H^{(u)} = S^u + T^u$ as in (2.8). In view of Theorems 2.2, 2.3 it suffices to prove the inequality

$$\Big\| \sup_{u \in U} |T^u_\pm f| \Big\|_p \lesssim \|f\|_p \tag{7.2}$$

for $4/3 < p \le 2$. By the reduction in Sect. 6 this can be accomplished if

$$\Big\| \sup_\nu \Big| \sum_j \mathcal{R}^{u_\nu}_{j,\ell} f \Big| \Big\|_p \lesssim 2^{-\ell\epsilon(p)} \|f\|_p \tag{7.3}$$

can be proved for $\epsilon(p) > 0$, in our case in the range $4/3 < p \le 2$.

Replacing the sup by an ℓ^2 norm we see that (7.3) follows from

$$\Big\| \Big(\sum_\nu \Big| \sum_j \mathcal{R}^{u_\nu}_{j,\ell} f \Big|^2 \Big)^{1/2} \Big\|_p \lesssim 2^{-\ell\epsilon(p)} \|f\|_p \tag{7.4}$$

Analogously to (6.10) we have

$$\mathcal{R}^{u_\nu}_{j,\ell} = \tilde{P}^{(1)}_j \tilde{P}^{(2)}_{j-n(\nu),b} \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_j$$

and thus, by Littlewood–Paley theory, (7.4) is a consequence of

$$\Big\| \Big(\sum_\nu \sum_{j \in \mathbb{Z}} \Big| \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_j f \Big|^2 \Big)^{1/2} \Big\|_p \lesssim 2^{-\ell\epsilon(p)} \|f\|_p. \tag{7.5}$$

By a standard application of Khintchine's inequality this estimate follows if we can prove

$$\left\| \sum_{\nu} \sum_{j\in\mathbb{Z}} c(\nu, j) \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_{j} f \right\|_p \lesssim 2^{-\ell\epsilon(p)} \|f\|_p. \tag{7.6}$$

for an arbitrary choice of $\{c(\nu, j)\}$ with $\sup_{j,\nu} |c(\nu, j)| \le 1$. Let

$$\omega_\ell(\xi) = \omega_+(2^\ell \xi)\chi_+(|\xi|)$$

then ω_ℓ and its derivatives are $O(2^{-\ell/2})$, by the symbol property of ω_+, and are supported on a common annulus. We see that the L^2 operator norms of the individual operators $\mathcal{R}^{u_\nu}_{j,\ell}$ are $O(2^{-\ell/2})$, and that the function

$$\begin{aligned} m_\ell(\xi) = \sum_{\nu} \sum_{j} & \tilde{\chi}^{(1)}(2^{-j}\xi_1)\tilde{\chi}^{(2)}(2^{-jb+n(\nu)b}\xi_2) \\ & \times \omega_\ell(2^{-j}\xi_1, 2^{(n(\nu)-j)b}\xi_2) e^{i2^\ell \Psi_+(2^{-j}\xi_1, 2^{(n(\nu)-j)b}\xi_2)} \end{aligned}$$

has L^∞ norm $\lesssim 2^{-\ell/2}$. This implies

$$\left\| \sum_{\nu} \sum_{j\in\mathbb{Z}} c(\nu, j) \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_{j} f \right\|_2 \lesssim 2^{-\ell/2} \|f\|_2. \tag{7.7}$$

For p near 1 we apply the Marcinkiewicz multiplier theorem in the form described in Sect. 3.5. It is not hard to check that the multiplier m_ℓ satisfies the condition (3.14) with constant $B \le C_\alpha 2^{\ell(2\alpha-1/2)}$. Hence we get

$$\left\| \sum_{\nu} \sum_{j\in\mathbb{Z}} c(\nu, j) \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_{j} f \right\|_p \lesssim 2^{\ell(2\alpha-\frac{1}{2})} \|f\|_p, \quad \alpha > 1/2. \tag{7.8}$$

We interpolate between (7.7) and (7.8). By choosing α very close to $1/2$, we obtain (7.6) for any $p \in (4/3, 2]$. □

7.2 Proof of Proposition 7.2

We argue as in the proof of Proposition 7.1. The desired conclusion follows if under our present conditions (7.8) can be upgraded to

$$\left\| \sum_{\nu} \sum_{j\in\mathbb{Z}} c(\nu, j) \mathcal{R}^{u_\nu}_{j,\ell} \tilde{P}^{(2)}_{j-n(\nu),b} \tilde{P}^{(1)}_{j} f \right\|_p \le c_p(1+\ell^4) \|f\|_p, \quad 1 < p \le 2. \tag{7.9}$$

As now $u_\nu = 2^{n(\nu)b}$ for a strictly increasing sequence $\{n(\nu)\}$ we see by another application of Littlewood-Paley theory that (7.9) is a consequence of the inequality

$$\left\| \left(\sum_{n\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \left| \mathcal{R}^{2^{nb}}_{j,\ell} f_{j,n} \right|^2 \right)^{1/2} \right\|_p \lesssim (1+\ell^4) \left\| \left(\sum_{j,n} |f_{j,n}|^2 \right)^{1/2} \right\|_p . \tag{7.10}$$

This is proved as in [16] by using a superposition of shifted maximal operators, in a vector-valued setting. To analyze the situation we recall how $\mathcal{R}^u_{j,\ell}$ was formed (namely by rescaling $T^u_{j,\ell}$, then see Sect. 2).

Let σ_+ be as in (2.1). Then there is a Schwartz function ς such that

$$\begin{aligned} \widehat{\mathcal{R}^{2^{nb}}_{j,\ell} f}(\xi) = {} & \chi_+(|(2^{-j}\xi_1, 2^{nb-jb}\xi_2)|)\widehat{\sigma_+}(2^{\ell-j}\xi_1, 2^{\ell+nb-jb}\xi_2)\widehat{f}(\xi) \\ & + \chi_+(|(2^{-j}\xi_1, 2^{nb-jb}\xi_2)|)\widehat{\varsigma}(2^{\ell-j}\xi_1, 2^{\ell+nb-jb}\xi_2)\widehat{f}(\xi). \end{aligned}$$

Consider the second (error) term. It is easy to see that

$$\left| \mathcal{F}[\chi_+(|(2^{-j}\xi_1, 2^{nb-jb}\xi_2)|)\widehat{\varsigma}(2^{\ell-j}\xi_1, 2^{\ell+nb-jb}\xi_2)\widehat{f}(\xi)](x) \right| \lesssim 2^{-\ell} M_{\mathrm{str}} f(x)$$

so that these terms are taken care of by an application of the Fefferman-Stein inequality for the vector-valued strong maximal function.

We concentrate on the main term. We write $\sigma_+ = \sum_{m=2^{\ell-1}}^{2^{\ell+1}} \mu_m$ where the measure μ_m is given by

$$\langle \mu_m, f \rangle = \int_{m2^{-\ell}}^{(m+1)2^{-\ell}} f(t, \gamma_b(t))\chi_+(t)\frac{dt}{t}.$$

Define $\mathcal{R}^u_{j,\ell,m} f$ by

$$\widehat{\mathcal{R}^u_{j,\ell,m} f}(\xi) = \chi(|(2^{-j}\xi_1, 2^{-jb}\xi_2)|)\widehat{\mu_m}(2^{\ell-j}\xi_1, 2^{\ell-jb}u\xi_2)\widehat{f}(\xi).$$

Then by the above discussion we have

$$\left| \mathcal{R}^{2^{nb}}_{j,\ell} f(x) - \sum_{m=2^{\ell-1}}^{2^{\ell+1}} \mathcal{R}^{2^{nb}}_{j,\ell,m} f(x) \right| \lesssim 2^{-\ell} M_{\mathrm{str}} f(x)$$

and hence, by Minkowski's inequality, it suffices to show that

$$\left\| \left(\sum_{n,j\in\mathbb{Z}} \left| \mathcal{R}^{2^{nb}}_{j,\ell,m} f_{j,n} \right|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\ell}(1+\ell)^4 \left\| \left(\sum_{j,n\in\mathbb{Z}} |f_{j,n}|^2 \right)^{1/2} \right\|_p \tag{7.11}$$

for $2^{\ell-1} \le m \le 2^{\ell+1}$. Notice that

$$\begin{aligned}&\left|\mu_m * \mathcal{F}^{-1}[\chi_+(|\cdot|2^{-\ell})](y)\right| \\ &\lesssim 2^{-\ell} \frac{2^\ell}{(1+2^\ell|y_1 - m2^{-\ell}|)^{10}} \frac{2^\ell}{(1+2^\ell|y_2 - m^b 2^{-\ell b}|)^{10}}\end{aligned}$$

Now define

$$\begin{aligned}\rho^{(1)}_{m,k_1}(y_1) &= 2^{k_1}\left(1+|2^{k_1}y_1 - m|\right)^{-10} \\ \rho^{(2)}_{m,k_2}(y_2) &= 2^{bk_2}\left(1+2^{bk_2}|y_2 - m^b 2^{-\ell(b-1)}|\right)^{-10}\end{aligned}$$

We then have the pointwise estimate

$$\left|\mathcal{R}^{2^{nb}}_{j,\ell} f(x)\right| \lesssim 2^{-\ell}\left(\rho^{(1)}_{m,j} \otimes \rho^{(2)}_{m,j-n}\right) * |f|. \tag{7.12}$$

By an application of inequalities for the shifted maximal operators (see [16, Theorem 3.1]) we see that the expressions

$$\left(\int \left|\left(\sum_{k_1,k_2}\left[\int \rho^{(1)}_{m,k_1}(x_1 - y_1)|g_{k_1,k_2}(y_1,x_2)|dy_2\right]^2\right)^{p/2} dx\right)^{1/p},$$

$$\left(\int \left|\left(\sum_{k_1,k_2}\left[\int \rho^{(2)}_{m,k_2}(x_2 - y_2)|g_{k_1,k_2}(x_1,y_2)|dy_2\right]^2\right)^{p/2} dx\right)^{1/p}$$

are both bounded by a constant times

$$(\log m)^2 \left\|\left(\sum_{k_1,k_2}|g_{k_1,k_2}|^2\right)^{1/2}\right\|_p.$$

Applying both estimates iteratively we get

$$\left\|\left(\sum_{k_1,k_2}\left[\left(\rho^{(1)}_{m,k_1}\otimes\rho^{(2)}_{m,k_2}\right) * |g_{k_1,k_2}|\right]^2\right)^{1/2}\right\|_p \lesssim (\log m)^4 \left\|\left(\sum_{k_1,k_2}|g_{k_1,k_2}|^2\right)^{1/2}\right\|_p.$$

We apply this with $g_{k_1,k_2} = f_{k_1,k_1-k_2}$ and use (7.12) to obtain (7.11). □

8 Lower bounds

8.1 The main lower bound and some consequences

The purpose of this section is to prove the lower bound

Theorem 8.1 *Let $U \subset (0,\infty)$ and $1 < p < \infty$. Then there is a constant c_p such that*

$$\|\mathcal{H}^U\|_{L^p\to L^p} \geq c_p\sqrt{\log(\mathfrak{N}(U))}.$$

8.1.1 Some consequences

(i) First, Theorem 8.1 in combination with the already proven upper bounds in Theorems 2.2 and 2.3 yields the equivalence (with constants depending on p)

$$\|\mathcal{H}^U\|_{L^p\to L^p} \approx \sqrt{\log(\mathfrak{N}(U))} \tag{8.1}$$

for $2 < p < \infty$, stated as Theorem 1.1.

(ii) We also immediately get an equivalence in Propositions 7.1 and 7.2 which we formulate as

Corollary 8.2 *Let U be a lacunary set. Then (8.1) holds for $4/3 < p < \infty$. If U is contained in $\{2^{nb} : n \in \mathbb{Z}\}$ then (8.1) holds for $1 < p < \infty$.*

8.1.2 Reduction to the case $p = 2$

Let U_* be a maximal subset of U with the property that each interval $[2^n, 2^{n+1}]$ contains at most one point in U. Then $\#(U_*) \approx \mathfrak{N}(U)$. Let $\widetilde{U}$ be any finite subset of U_* with the understanding that $\widetilde{U} = U_*$ if U_* is already finite. Clearly

$$\|\mathcal{H}^U\|_{L^p\to L^p} \geq \|\mathcal{H}^{U_*}\|_{L^p\to L^p} \geq \|\mathcal{H}^{\tilde{U}}\|_{L^p\to L^p}$$

and thus it suffices to prove the inequality

$$\|\mathcal{H}^{\tilde{U}}\|_{L^p\to L^p} \gtrsim A_p\sqrt{\log(\#\widetilde{U})}. \tag{8.2}$$

We show that it suffices to prove (8.2) for $p = 2$: Since $\widetilde{U}$ is a disjoint union of two lacunary sets we have the inequality

$$\|\mathcal{H}^{\tilde{U}}\|_{L^q\to L^q} \leq C_q\sqrt{\log(\#\widetilde{U})}, \quad \text{for } 4/3 < q < \infty,$$

by Proposition 7.1.

If $1 < p < 2$ we pick q such that $2 < q < \infty$, and if $2 < p < \infty$ we pick q such that $4/3 < q < 2$. Let $\theta \in (0,1)$ such that $(1-\theta)/p + \theta/q = 1/2$. We have

$$A_2\big(\log(\#\widetilde{U})\big)^{1/2} \le \|\mathcal{H}^{\tilde U}\|_{L^2\to L^2} \le \|\mathcal{H}^{\tilde U}\|_{L^p\to L^p}^{1-\theta}\|\mathcal{H}^{\tilde U}\|_{L^q\to L^q}^{\theta}$$
$$\le \big(c_q(\log(\#\widetilde{U}))^{1/2}\big)^{\theta}\|\mathcal{H}^{\tilde U}\|_{L^p\to L^p}^{1-\theta}$$

which implies

$$\|\mathcal{H}^{\tilde U}\|_{L^p\to L^p} \ge A_2^{\frac{1}{1-\theta}} c_q^{-\frac{\theta}{1-\theta}}\sqrt{\log(\#\widetilde{U})}.$$

For the remainder of this section we shall verify the lower bound in (8.2) for $p = 2$. We shall need to skim the set $\widetilde{U}$ a bit more. To prepare for this we first study in more detail the multipliers of the Hilbert transforms.

8.2 Observations on the multipliers for the Hilbert transforms

We may assume $c_+ > 0$. We write $\widehat{H^{(u)}f}(\xi) = m(\xi_1, u\xi_2)\widehat{f}(\xi)$ where

$$m(\xi_1,\xi_2) = \lim_{\substack{\varepsilon\to 0+\\ R\to\infty}}\left(\int_{\varepsilon<t\le R} e^{-i(t\xi_1+c_+t^b\xi_2)}\frac{dt}{t} + \int_{-R<t<-\varepsilon} e^{-i(t\xi_1+c_-(-t)^b\xi_2)}\frac{dt}{t}\right).$$

By the homogeneity of the curve Γ_b with respect to the dilations $(\xi_1,\xi_2)\mapsto(\lambda\xi_1,\lambda^b\xi_2)$, we see that $m(\lambda\xi_1,\lambda^b\xi_2) = m(\xi_1,\xi_2)$ for $\lambda > 0$. Moreover one can check that m is continuous on $\mathbb{R}^2\backslash\{0\}$,

$$m(\xi_1,0) = -\pi i\,\mathrm{sign}\xi_1, \quad \xi_1\neq 0, \tag{8.3a}$$

and if $\xi_2 > 0$, then

$$m(0,\xi_2) = \begin{cases} -\frac{1}{b}\log(c_+/c_-) & \text{if } c_- > 0\\ -\frac{1}{b}\log(-c_+/c_-) - \frac{1}{b}\pi i & \text{if } c_- < 0.\end{cases} \tag{8.3b}$$

We shall need the following Hölder continuity condition at the axes.

Lemma 8.3 *There is $C_\circ = C_\circ(b,c_\pm) \ge 1$ such that we have the estimates*

$$|m(\xi_1,\xi_2) - m(\xi_1,0)| \le C_\circ\left(\frac{|\xi_2|}{|\xi_1|^b}\right)^{\frac{1}{2b}}, \tag{8.4a}$$

$$|m(\xi_1,\xi_2) - m(0,\xi_2)| \le C_\circ\left(\frac{|\xi_1|^b}{|\xi_2|}\right)^{\frac{1}{2b}}. \tag{8.4b}$$

Proof of Lemma 8.3 We have $|m(\xi_1,\xi_2)| \le C_\circ(b,c_\pm)$ and therefore it suffices to show that (8.4a) holds for $|\xi_2| \ll |\xi_1|^b$ and (8.4b) holds for $|\xi_1|^b \ll |\xi_2|$.

For the proof of (8.4a) it suffices to check, by homogeneity and boundedness of m,

$$|m(\pm1,\xi_2) - m(\pm1,0)| \lesssim |\xi_2|^\beta, \quad |\xi_2|\le 1, \tag{8.5}$$

for some $\beta \geq (2b)^{-1}$. Let

$$A = A(\eta) = \frac{1}{2}|\eta|^{-\frac{1}{b+1}}. \tag{8.6}$$

We have

$$m(1,\xi_2) - m(1,0) = \sum_{j=1}^{3}(I_{j,+}(c_+ b\xi_2) - I_{j,-}(c_- b\xi_2))$$

where

$$\begin{aligned} I_{1,\pm}(\eta) &= \int_0^{A(\eta)} e^{\mp it}(e^{-it^b\eta/b} - 1)\frac{dt}{t}, \\ I_{2,\pm}(\eta) &= \int_{A(\eta)}^{\infty} e^{\mp it - it^b\eta/b}\frac{dt}{t}, \\ I_{3,\pm}(\eta) &= -\int_{A(\eta)}^{\infty} e^{\mp it}\frac{dt}{t}. \end{aligned}$$

Clearly

$$|I_{1,\pm}(\eta)| \leq \int_0^A t^{b-1}|\eta| b^{-1} dt = A^b b^{-2}|\eta|.$$

By integration by parts,

$$|I_{3,\pm}| \leq 2A^{-1}.$$

By our choice (8.6)

$$|I_{1,\pm}(\eta)| + |I_{3,\pm}(\eta)| \lesssim |\eta|^{\frac{1}{b+1}}$$

We may assume $|\eta| < 1$. Let $B_1 = B_1(\eta) = |\eta^{-1/(b-1)}|/2$ and $B_2 = B_2(\eta) = 2|\eta^{-1/(b-1)}|$. Then $B_1(\eta) \geq A(\eta)$ and we split

$$I_{2,\pm}(\eta) = \int_A^{B_1} + \int_{B_1}^{B_2} + \int_{B_2}^{\infty} e^{i\psi(t)} t^{-1} dt$$

with $\psi(t) = \mp t - t^b\eta/b$.

Note that for $|t| \leq B_1$ we have $1/2 < |\psi'(t)| \leq 2$ and thus, by van der Corput's lemma with first derivative we have $|\int_A^{B_1}(...)dt| \lesssim A^{-1}$.

Note that $|\psi''(t)| = |\eta|(b-1)t^{b-2}$. For the second integral we apply van der Corput's lemma with second derivatives and get $|\int_{B_1}^{B_2}(...)dt| \lesssim |B_1|^{-1}|\eta|^{-1/2}(b-1)^{-1/2}|B_1|^{-(b-2)/2} \lesssim (b-1)^{-1/2}|\eta|^{1/(2b-2)}$.

Finally for the third integral we use that $|\psi'(t)| \approx |\eta| t^{b-1}$ and $|\psi''(t)| \approx |\eta|(b-1)t^{b-2}$ and a straightforward integration by parts argument yields the bound $O(|\eta|^{-1} B_2^{-b}) = O(|\eta|^{\frac{1}{b-1}})$.

The estimate for $m(-1, \xi_2) - m(-1, 0)$ is analogous. Altogether we obtain (8.5) with $\beta = \min\{(b+1)^{-1}, (2b-2)^{-1}\}$, and we have $\beta \geq (2b)^{-1}$.

We now turn to the proof of (8.4b). It suffices to check, by homogeneity and boundedness of m,

$$|m(\xi_1, \pm 1) - m(0, \pm 1)| \lesssim |\xi_1|^{1/2}, \quad |\xi_1| \leq 1. \tag{8.7}$$

Let

$$B = B(\xi_1) = (a|\xi_1|)^{-1/2} \quad \text{where } a = \min_{\pm}(bc_{\pm}/2)^{\frac{2}{b-1}}. \tag{8.8}$$

We have

$$m(\xi_1, 1) - m(0, 1) = \sum_{j=1}^{3} \big(II_{j,+}(\xi_1) - II_{j,-}(\xi_1) \big)$$

where

$$\begin{aligned} II_{1,\pm}(\xi_1) &= \int_0^{B(\xi_1)} (e^{\mp it\xi_1} - 1) e^{-ic_{\pm}t^b} \frac{dt}{t}, \\ II_{2,\pm}(\xi_1) &= \int_{B(\xi_1)}^{\infty} e^{\mp it\xi_1 - ic_{\pm}t^b} \frac{dt}{t}, \\ II_{3,\pm}(\xi_1) &= -\int_{B(\xi_1)}^{\infty} e^{\mp it} \frac{dt}{t}. \end{aligned}$$

The estimation of these terms is straightforward; we get

$$|II_{1,\pm}(\xi_1)| \lesssim |\xi_1| B(\xi_1)$$

and

$$|II_{3,\pm}(\xi_1)| \lesssim B(\xi_1)^{-1}$$

and both terms are $O(|\xi|^{1/2})$, by our choice (8.8). By this choice we also have $2 \leq |c_{\pm}| b t^{b-1}$ for $t \geq B(\xi_1)$ which implies that for $|\xi_1| \leq 1$

$$\frac{1}{2}|c_{\pm}| b t^{b-1} \leq |\partial_t(\mp t\xi_1 - c_{\pm}t^b)| \leq 2|c_{\pm}| b t^{b-1} \text{ for } t \geq B(\xi_1).$$

Integration by parts now shows that

$$|II_{2,\pm}(\xi_1)| \lesssim B(\xi_1)^{-b}$$

which is $O(|\xi_1|^{b/2})$, hence also $O(|\xi_1|^{1/2})$. The term $m(\xi_1, -1) - m(0, -1)$ is similarly estimated. This completes the proof of (8.7). □

8.3 Reduction to a lower bound for a lacunary maximal operator

Recall that $\tilde{U} \subset U$ with $\mathfrak{N}(\tilde{U}) < \infty$. Let $\mathfrak{J}$ be the collection of all integers n such that $[2^n, 2^{n+1}]$ has nonempty intersection with $\tilde{U}$, thus $\mathfrak{N}(\tilde{U}) = 1 + \#\mathfrak{I}$. Let

$$K = K(\tilde{U}) = (C_\circ \mathfrak{N}(\tilde{U}))^{2b} \tag{8.9}$$

where $C_\circ$ is as in (8.4a), (8.4b). Let $\mathfrak{I}'$ be a *maximal* subfamily of $\mathfrak{I}$ with the condition

$$n_1 \in \mathfrak{I}',\ n_2 \in \mathfrak{I}',\ n_1 < n_2 \implies n_2 - n_1 + 1 \geq \log_2(8K^2). \tag{8.10}$$

Pick an integer M such that $M + 1$ is of the form 2^μ with $\mu \in \mathbb{N}$ and such that

$$\frac{\mathfrak{N}(\tilde{U})}{\log_2(16K^2)} = \frac{\mathfrak{N}(\tilde{U})}{4 + 4b\log_2(C_\circ \mathfrak{N}(\tilde{U}))} \in [M, 2M).$$

We may assume that the displayed quantity is $\geq e^{100}$, so that the logarithm of this quantity is comparable to $\log M$ (otherwise the desired lower bound for $\|\mathcal{H}^U\|_{L^2 \to L^2}$ just follows from the trivial lower bound for the Hilbert transform along a fixed curve).

We may now pick an increasing sequence $\{u_j\}_{j=1}^M$ such that each u_j belongs to $\tilde{U}$ and to exactly one interval determined by the collection $\mathfrak{I}'$. Hence we have

$$\frac{u_{j+1}}{u_j} \geq 16K^2. \tag{8.11}$$

Given the reduction in Sect. 8.1.2 the lower bound $\sqrt{\log(\mathfrak{N}(U))}$ in Theorem 8.1 follows from

Proposition 8.4 *Let $\tilde{U}$ and $\{u_j\}_{j=1}^M$ be as above. Then there is $c > 0$ such that*

$$\sup_{\|f\|_2 = 1} \Big\| \sup_{1 \leq j \leq M} |\mathcal{H}^{(u_j)} f| \Big\|_2 \geq c\sqrt{\log M}\,.$$

The proof of this proposition is based on a construction by Karagulyan [18].

8.4 A theorem of Karagulyan

We will invoke the following proposition, which is a small generalization of the main theorem of Karagulyan [18] (see also [19]). For $\mu \in \mathbb{N}$, let

$$W_\mu = \{\emptyset\} \cup \bigcup_{\ell=1}^{\mu-1} \{0, 1\}^\ell$$

be the set of binary words of length at most $\mu - 1$, and let

$$\tau : W_\mu \to \{1, \dots, 2^\mu - 1\}$$

be the bijection given by $\tau(\emptyset) = 2^{\mu-1}$ and

$$\tau(w) = w_1 2^{\mu-1} + w_2 2^{\mu-2} + \cdots + w_\ell 2^{\mu-\ell} + 2^{\mu-\ell-1}$$

if $w = w_1 w_2 \dots w_\ell$ for some $\ell \in \{1, \dots, \mu - 1\}$, and each $w_1, \dots, w_\ell \in \{0, 1\}$. Observe that for a word w of length ℓ, $\tau(w)$ is divisible by $2^{\mu-\ell-1}$ but not by $2^{\mu-\ell}$.

Proposition 8.5 *Let μ be any positive integer, $M = 2^\mu - 1$, and let $S_1, \dots, S_M$ be pairwise disjoint subsets of the (frequency) plane $\mathbb{R}^2$, so that every S_j contains balls of arbitrarily large radii (in other words, for every $1 \le j \le M$ and every $R > 0$, S_j contains some ball of radius R). Then there exists an L^2 function f on $\mathbb{R}^2$, that admits an orthogonal decomposition*

$$f = \sum_{w \in W_\mu} f_w,$$

where

$$\operatorname{supp} \widehat{f_w} \subset S_{\tau(w)} \quad \text{for all } w \in W_\mu, \text{ and} \tag{8.12}$$

$$\|f\|_{L^2}^2 = \sum_{w \in W_\mu} \|f_w\|_{L^2}^2 \le 2; \tag{8.13}$$

in addition,

$$\left\| \sup_{1 \le j \le M} \left| \sum_{w \in W_\mu :\, \tau(w) \ge j} f_w \right| \right\|_{L^2} \ge \frac{\sqrt{\mu}}{100} \|f\|_{L^2}. \tag{8.14}$$

Accepting this for the moment, we prove Proposition 8.4.

8.5 Proof of Proposition 8.4

As before, suppose $c_+ > 0$. Let

$$\rho = \begin{cases} -\frac{1}{b} \log(c_+/c_-) & \text{if } c_- > 0, \\ -\frac{1}{b} \log(-c_+/c_-) - \frac{1}{b}\pi i & \text{if } c_- < 0. \end{cases}$$

Then $m(0, \xi_2) = \rho$ for $\xi_2 > 0$ and $m(\xi_1, 0) = -\pi i$ for $\xi_1 > 0$ (cf. (8.3b), (8.3a)). Let K as in (8.9), then

$$C_\circ K^{-\frac{1}{2b}} \le (\mathfrak{N}(\tilde{U}))^{-1} \le M^{-1}.$$

From (8.4a) and (8.4b) we see, for $\xi_1 > 0, \xi_2 > 0$

$$\xi_2/\xi_1^b \le K^{-1} \implies |m(\xi_1,\xi_2)+\pi i| \le C_\circ K^{-\frac{1}{2b}} \le M^{-1}. \tag{8.15a}$$

$$\xi_2/\xi_1^b \ge K \implies |m(\xi_1,\xi_2)-\rho| \le C_\circ K^{-\frac{1}{2b}} \le M^{-1} \tag{8.15b}$$

For $1 \le j \le M$, define

$$S_j = \left\{(\xi_1,\xi_2) : \xi_1 > 0,\ \xi_2 > 0,\ \frac{1}{2Ku_j} < \frac{\xi_2}{\xi_1^b} < \frac{1}{Ku_j}\right\}, \tag{8.16}$$

so that the S_j are pairwise disjoint, and contain balls of arbitrarily large radii. By Proposition 8.5, there exists an L^2 function $f = \sum_{w\in W_\mu} f_w$ on $\mathbb{R}^2$, such that (8.12), (8.13) and (8.14) hold. Now for $1 \le j \le M$,

$$\begin{aligned}|\mathcal{H}^{(u_j)} f(x) - \rho f(x)| \ge &\left|\sum_{\substack{w\in W_\mu:\\ \tau(w)\ge j}} (\pi i+\rho)\, f_w(x)\right| \\ &- \left|\sum_{\substack{w\in W_\mu:\\ \tau(w)\ge j}} \left(\mathcal{H}^{(u_j)} f_w(x) + \pi i f_w(x)\right)\right| - \left|\sum_{\substack{w\in W_\mu:\\ \tau(w)< j}} \left(\mathcal{H}^{(u_j)} f_w(x) - \rho f_w(x)\right)\right|,\end{aligned}$$

and thus, with $c_0 = \pi(1-\frac{1}{b})$,

$$\begin{aligned}\sup_{1\le j\le M} |\mathcal{H}^{(u_j)} f(x) - \rho f(x)| \ge c_0 &\sup_{1\le j\le M}\left|\sum_{\substack{w\in W_\mu:\\ \tau(w)\ge j}} f_w(x)\right| \\ &- \sup_{1\le j\le M}\left|\sum_{\substack{w\in W_\mu:\\ \tau(w)\ge j}} (\mathcal{H}^{(u_j)}+\pi i) f_w(x)\right| - \sup_{1\le j\le M}\left|\sum_{\substack{w\in W_\mu:\\ \tau(w)< j}} (\mathcal{H}^{(u_j)}-\rho) f_w(x)\right|.\end{aligned} \tag{8.17}$$

Now $\operatorname{supp}\widehat{f_w} \in S_{\tau(w)}$. If $\tau(w) \ge j$, then for $\xi \in \operatorname{supp}\widehat{f_w}$, we have $u_j\xi_2/\xi_1^b < u_{\tau(w)}\xi_2/\xi_1^b < K^{-1}$ and therefore, by (8.15a), we have $|m(\xi_1, u_j\xi_2)+\pi i| \le M^{-1}$ for $\xi \in \operatorname{supp}\widehat{f_w}$. Hence

$$\left\|(\mathcal{H}^{(u_j)}+\pi i) f_w\right\|_2 \le M^{-1}\|f_w\|_2 \quad \text{if } \tau(w) \ge j. \tag{8.18}$$

Moreover if $\tau(w) < j$ we have, for $\xi \in \operatorname{supp} \widehat{f_w}$,

$$u_j \frac{\xi_2}{\xi_1^b} = \frac{u_j}{u_{\tau(w)}} u_{\tau(w)} \frac{\xi_2}{\xi_1^b} \geq 16K^2 \frac{1}{2K} = 8K$$

and hence, by (8.15b), $|m(\xi_1, u_j\xi_2) - \rho| \leq M^{-1}$ for $\xi \in \operatorname{supp} \widehat{f_w}$. Thus

$$\left\| (\mathcal{H}^{(u_j)} - \rho) f_w \right\|_2 \leq M^{-1} \|f_w\|_2 \quad \text{if } \tau(w) < j. \tag{8.19}$$

Statements (8.18) and (8.19) imply

$$\left\| \sup_{1 \leq j \leq M} \left| \sum_{\substack{w \in W_\mu : \\ \tau(w) < j}} (\mathcal{H}^{(u_j)} - \rho) f_w \right| \right\|_2 \lesssim \|f\|_2 \tag{8.20}$$

$$\left\| \sup_{1 \leq j \leq M} \left| \sum_{\substack{w \in W_\mu : \\ \tau(w) \geq j}} (\mathcal{H}^{(u_j)} + \pi i) f_w \right| \right\|_2 \lesssim \|f\|_2. \tag{8.21}$$

Indeed, to obtain (8.21) we use the Cauchy-Schwarz inequality in the w sum and replace a sup in j by an ℓ^2 norm, then interchange integrals and sums and apply (8.19) to get

$$\begin{aligned}
&\left\| \sup_{1 \leq j \leq M} \left| \sum_{\substack{w \in W_\mu : \\ \tau(w) < j}} (\mathcal{H}^{(u_j)} - \rho) f_w \right| \right\|_2 \\
&\leq M^{1/2} \left\| \left(\sum_{j=1}^{M} \sum_{\tau(w) < j} |(\mathcal{H}^{(u_j)} - \rho) f_w|^2 \right)^{1/2} \right\|_2 \\
&= M^{1/2} \left(\sum_{j=1}^{M} \sum_{\tau(w) < j} \left\| (\mathcal{H}^{(u_j)} - \rho) f_w \right\|_2^2 \right)^{1/2} \\
&\leq M^{1/2} \left(\sum_{j=1}^{M} M^{-2} \sum_{w} \|f_w\|_2^2 \right)^{1/2} \lesssim \|f\|_2
\end{aligned}$$

(the last line following from (8.13)). Inequality (8.20) is proved in exactly the same way (relying on (8.18)).

Now we go back to (8.17), use (8.14) for the main part and (8.20), (8.21) for the two error terms. Then we get

$$\left\| \sup_{1\le j\le M} |(\mathcal{H}^{(u_j)} - \rho) f| \right\|_2 \ge c\sqrt{\mu}\|f\|_2$$

for some constant $c = c(b, c_\pm) > 0$. If $\sqrt{\mu} \ge 2|\rho|/c$ this also implies

$$\left\| \sup_{1\le j\le M} |\mathcal{H}^{(u_j)} f| \right\|_2 \ge (c/2)\sqrt{\mu}\|f\|_2.$$

This completes the proof of Proposition 8.4, except for Proposition 8.5. □

8.6 Proof of Proposition 8.5

Fix a non-negative Schwartz function ϕ on $\mathbb{R}^2$ with $\int_{\mathbb{R}^2} \phi(x)dx = 1$, such that $\widehat{\phi}$ is supported in the unit ball $B(0, 1)$ centered at the origin. Define the frequency cutoff ϕ_ρ by

$$\phi_\rho(x) := \rho^2\phi(\rho x).$$

Then $\widehat{\phi_\rho}$ is supported on $B(0, \rho)$.

The following lemma explains what we actually construct, in order to prove Proposition 8.5:

Lemma 8.6 *Let $\mu \in \mathbb{N}$, $M = 2^\mu - 1$, and let $S_1, \dots, S_M$ be as given in Proposition 8.5. Then there exist a sequence of sets $\{E_w\}_{w\in W_\mu}$, modulation frequencies $\{\xi_w\}_{w\in W_\mu} \subset \mathbb{R}^2$, and radii $\{\rho_w\}_{w\in W_\mu}$ such that the following holds:*

(a) *For every $w \in W_\mu$, $E_w \subset [0, 1]^2$, and for every $w \in W_{\mu-1}$, E_w is the disjoint union of E_{w0} and E_{w1} Also, $E_\emptyset = [0, 1]^2$. For $\ell = 0, \dots, \mu - 1$, $[0, 1]^2$ is a disjoint union of the E_w with length$(w) = \ell$, and*

$$\sum_{w\in W_\mu} \mathbb{1}_{E_w}(x) = \mu. \tag{8.22}$$

for every $x \in [0, 1]^2$.

(b) *For every $w \in W_\mu$,*

$$\|\mathbb{1}_{E_w} * \phi_{\rho_w} - \mathbb{1}_{E_w}\|_{L^2} \le 2^{-\mu-10}, \tag{8.23}$$

$$\int_{E_w} |\cos(\langle \xi_w, x\rangle)|dx \ge \frac{|E_w|}{3}, \tag{8.24}$$

$$B(\xi_w, \rho_w) \subset S_{\tau(w)}. \tag{8.25}$$

(c) *For every* $w \in W_{\mu-1}$, *we have*

$$\begin{aligned} \cos(\langle \xi_w, x\rangle) &\geq 0 \quad \text{if } x \in E_{w0}, \\ \cos(\langle \xi_w, x\rangle) &< 0 \quad \text{if } x \in E_{w1}. \end{aligned} \tag{8.26}$$

With this lemma we can prove Proposition 8.5 as follows.

Proof of Proposition 8.5 For every $w \in W_\mu$, let E_w, ρ_w and ξ_w be as in Lemma 8.6. We set

$$f_w(x) := \mu^{-1/2} e^{i\langle \xi_w, x\rangle} \mathbb{1}_{E_w} * \phi_{\rho_w}(x), \tag{8.27a}$$

and let

$$f := \sum_{w \in W_\mu} f_w. \tag{8.27b}$$

Then the support of $\widehat{f_w}$ is contained inside $B(\xi_w, \rho_w)$, so (8.12) follows from (8.25). Also, the $\widehat{f_w}$'s are supported in the sets $S_{\tau(w)}$ which are disjoint and thus by orthogonality we have

$$\|f\|_2 = \left\| \left(\sum_{w \in W_\mu} |f_w|^2 \right)^{1/2} \right\|_2.$$

But, from (8.23), we have

$$\left\| f_w - \mu^{-1/2} e^{i\langle \xi_w, x\rangle} \mathbb{1}_{E_w} \right\|_2 \leq 2^{-\mu-10}. \tag{8.28}$$

Observe

$$\begin{aligned} &\left(\sum_{w \in W_\mu} |f_w|^2 \right)^{1/2} \\ &\leq \left(\sum_{w \in W_\mu} \left| f_w - \mu^{-1/2} e^{i\langle \xi_w, x\rangle} \mathbb{1}_{E_w} \right|^2 \right)^{1/2} + \left(\sum_{w \in W_\mu} \left| \mu^{-1/2} e^{i\langle \xi_w, x\rangle} \mathbb{1}_{E_w} \right|^2 \right)^{1/2}, \end{aligned}$$

and using (8.22) to simplify the second term we get

$$\left(\sum_{w \in W_\mu} |f_w|^2 \right)^{1/2} \leq \left(\sum_{w \in W_\mu} \left| f_w - \mu^{-1/2} e^{i\langle \xi_w, x\rangle} \mathbb{1}_{E_w} \right|^2 \right)^{1/2} + \mathbb{1}_{[0,1]^2}$$

for almost every $x \in \mathbb{R}^2$. Taking L^2 norms of both sides, and using (8.28), we have

$$\left\| \left(\sum_{w \in W_\mu} |f_w|^2 \right)^{1/2} \right\|_2 \le 2^{-5-\mu/2} + 1 < 2.$$

Thus (8.13) follows.

Lastly we have to verify (8.14). To do so, we first introduce an auxiliary family of functions $\{F_w\}_{w \in W_\mu}$, where

$$F_w := \operatorname{Re} f_w \, \mathbb{1}_{E_w}. \tag{8.29}$$

These F_w's satisfy three key properties, namely

$$\sum_{w \in W_\mu} \|F_w - \operatorname{Re} f_w\|_{L^2} \le 2^{-10}, \tag{8.30}$$

$$\frac{1}{3} \le \frac{\sup_{1 \le j \le M} \left| \sum_{w \in W_\mu : \, \tau(w) \ge j} F_w(x) \right|}{\sum_{w \in W_\mu} |F_w(x)|} \le 1 \quad \text{for a.e. } x \in [0,1]^2, \tag{8.31}$$

and

$$\frac{\sqrt{\mu}}{4} \le \left\| \sum_{w \in W_\mu} |F_w| \right\|_1 \le \left\| \sum_{w \in W_\mu} |F_w| \right\|_2 \le \left\| \sum_{w \in W_\mu} |F_w| \right\|_\infty \le \sqrt{\mu}. \tag{8.32}$$

Indeed, (8.30) will be a consequence of

$$\|F_w - \operatorname{Re} f_w\|_{L^2} \lesssim 2^{-\mu-10} \quad \text{for all } w \in W_\mu. \tag{8.33}$$

Since $F_w - \operatorname{Re} f_w = \operatorname{Re} f_w \mathbb{1}_{\mathbb{R}^2 \setminus E_w}$, heuristically, (8.33) says that the real part of each f_w is essentially supported on E_w: the L^2 norm of $\operatorname{Re} f_w$ outside E_w is small. Furthermore, (8.31) says that there isn't much cancellation, if we first order the F_w's according to the value of $\tau(w)$, and then sum successively; this will be achieved by showing that $\{F_w\}_{w \in W_\mu}$ form a *tree system* in the sense of Karagulyan [18] (who credits the idea to Nikišin and Ul'janov [22]).

Let us now establish the three key properties of the F_w's, namely (8.30), (8.31) and (8.32). Since $F_w - \operatorname{Re} f_w = \operatorname{Re} f_w \mathbb{1}_{(E_w)^\complement}$, and since

$$\operatorname{Re} f_w(x) = \frac{1}{\sqrt{\mu}} \cos(\langle \xi_w, x \rangle) \mathbb{1}_{E_w} * \phi_{\ell_w}(x), \tag{8.34}$$

we have

$$\begin{aligned} \|F_w - \operatorname{Re} f_w\|_{L^2(\mathbb{R}^2)} &= \left\| \mu^{-1/2} \cos(\langle \xi_w, x \rangle) \mathbb{1}_{E_w} * \phi_{\ell_w} \right\|_{L^2(\mathbb{R}^2 \setminus E_w)} \\ &\le \left\| \mathbb{1}_{E_w} - \mathbb{1}_{E_w} * \phi_{\ell_w} \right\|_{L^2(\mathbb{R}^2 \setminus E_w)} \le 2^{-\mu-10} \end{aligned}$$

by (8.23). This establishes (8.33), and (8.30) follows by summing over $w \in W_\mu$.

Next we verify (8.31). The second inequality in (8.31) is immediate by the triangle inequality. For the first, we observe from (8.34) that if $x \in E_w$, then $F_w(x)$ has the same sign as $\cos(\langle \xi_w, x\rangle)$ since $\mathbb{1}_{E_w} * \phi_{\ell_w}$ is everywhere positive. We claim that for almost every $x \in [0,1]^2$, there exists $j = j(x)$ such that $F_w(x) \geq 0$ for every $w \in W_\mu$ with $\tau(w) \geq j$, and $F_w(x) < 0$ for every $w \in W_\mu$ with $\tau(w) < j$. This is because for almost every $x \in [0,1]^2$, there exists a unique word $w(x) = w_1 \dots w_{\mu-1}$ of length $\mu - 1$ such that $x \in E_{w(x)}$. By (8.26), it follows that, for every $\ell = 0, 1, \dots, \mu - 2$,

$$\begin{aligned} F_{w_1 \dots w_\ell}(x) &> 0 \quad \text{if } w_{\ell+1} = 0, \\ F_{w_1 \dots w_\ell}(x) &< 0 \quad \text{if } w_{\ell+1} = 1, \end{aligned}$$

and that $F_{w'}(x) = 0$ if $w' \in W_\mu \backslash \{\emptyset, w_1, w_1 w_2, \dots, w_1 \cdots w_{\mu-1}\}$. But

$$\tau(w_1 \dots w_\ell) = w_1 2^{\mu-1} + \cdots + w_\ell 2^{\mu-\ell} + 2^{\mu-\ell-1},$$

while

$$\tau(w(x)) = w_1 2^{\mu-1} + \cdots + w_\ell 2^{\mu-\ell} + w_{\ell+1} 2^{\mu-\ell-1} + \cdots + w_{\mu-1} 2^1 + 2^0.$$

This shows that for every $\ell = 0, 1, \dots, \mu - 2$,

$$\begin{aligned} \tau(w_1 \dots w_\ell) &> \tau(w(x)) \quad \text{if } w_{\ell+1} = 0, \\ \tau(w_1 \dots w_\ell) &< \tau(w(x)) \quad \text{if } w_{\ell+1} = 1. \end{aligned}$$

Thus for any $w' \in W_\mu$, one has

$$\begin{aligned} F_{w'}(x) &\geq 0 \quad \text{if } \tau(w') > \tau(w(x)), \\ F_{w'}(x) &\leq 0 \quad \text{if } \tau(w') < \tau(w(x)). \end{aligned}$$

If $F_{w(x)}(x) \geq 0$, we set $j(x) = \tau(w(x))$; if $F_{w(x)}(x) < 0$, we set $j(x) = \tau(w(x)) + 1$. It follows that that $F_w(x) \geq 0$ whenever $\tau(w) \geq j(x)$, and $F_w(x) \leq 0$ whenever $\tau(w) < j(x)$. We distinguish two cases now. In the first case we have

$$\left| \sum_{w \in W_\mu \colon \tau(w) \geq j(x)} F_w(x) \right| \geq \frac{1}{3} \sum_{w \in W_\mu} |F_w(x)|.$$

In the opposite case, we have $|\sum_{w \in W_\mu \colon \tau(w) \geq j(x)} F_w(x)| < \frac{1}{3} \sum_{w \in W_\mu} |F_w(x)|$, so $|\sum_{w \in W_\mu \colon \tau(w) < j(x)} F_w(x)| \geq \frac{2}{3} \sum_{w \in W_\mu} |F_w(x)|$. Then

$$\left| \sum_{w \in W_\mu} F_w(x) \right| \geq \left| \sum_{\substack{w \in W_\mu \colon \\ \tau(w) < j(x)}} F_w(x) \right| - \left| \sum_{\substack{w \in W_\mu \colon \\ \tau(w) \geq j(x)}} F_w(x) \right|$$

$$\geq \sum_{\substack{w\in W_\mu:\\ \tau(w)<j(x)}} |F_w(x)| - \frac{1}{3}\sum_{w\in W_\mu} |F_w(x)| \geq \frac{1}{3}\sum_{w\in W_\mu} |F_w(x)|.$$

Hence in both cases

$$\sup_{1\leq j\leq M} \left| \sum_{w\in W_\mu:\, \tau(w)\geq j} F_w(x) \right| \geq \frac{1}{3}\sum_{w\in W_\mu} |F_w(x)|$$

for every $x \in [0,1]^2$. This completes the proof of (8.31).

Finally, we have to verify (8.32). Note that F_w is supported on $[0,1]^2$ for every $w \in W_\mu$, and for almost every $x \in [0,1]^2$, there exists at most μ words $w \in W_\mu$ for which $F_w(x) \neq 0$. Furthermore, $|F_w(x)| \leq \mu^{-1/2}$ for every $x \in [0,1]^2$ and every $w \in W_\mu$. Thus, we have

$$\left\| \sum_{w\in W_\mu} |F_w| \right\|_1 \leq \left\| \sum_{w\in W_\mu} |F_w| \right\|_2 \leq \left\| \sum_{w\in W_\mu} |F_w| \right\|_\infty \leq \sqrt{\mu}.$$

Next, for the lower bound,

$$\left\| \sum_{w\in W_\mu} |F_w| \right\|_1 = \sum_{w\in W_\mu} \int_{E_w} \mu^{-1/2} |\cos(\langle \xi_w, x\rangle) \mathbb{1}_{E_w} * \phi_{\ell_w}(x)| dx$$

which is

$$\begin{aligned}
&\geq \frac{1}{\sqrt{\mu}} \sum_{w\in W_\mu} \int_{E_w} \Big(|\cos(\langle \xi_w, x\rangle)| - \big|\cos(\langle \xi_w, x\rangle)[\mathbb{1}_{E_w} - \mathbb{1}_{E_w} * \phi_{\ell_w}]\big| \Big) dx \\
&\geq \frac{1}{\sqrt{\mu}} \sum_{w\in W_\mu} \left(\int_{E_w} |\cos(\langle \xi_w, x\rangle)| dx - \|\mathbb{1}_{E_w} - \mathbb{1}_{E_w} * \phi_{\ell_w}\|_{L^2} |E_w|^{1/2} \right) \\
&\geq \frac{1}{\sqrt{\mu}} \sum_{w\in W_\mu} \left(\frac{|E_w|}{3} - 2^{-\mu-10} \right) \geq \frac{\sqrt{\mu}}{3} - 2^{-\mu-10}\sqrt{\mu} \geq \frac{\sqrt{\mu}}{4},
\end{aligned}$$

where for the last line we have used (8.24), (8.23) and (8.22). This completes the proof of (8.32).

We will now return to the proof of (8.14). First,

$$\sup_{1\leq j\leq M} \left| \sum_{w\in W_\mu:\, \tau(w)\geq j} f_w(x) \right| \geq \sup_{1\leq j\leq M} \left| \sum_{w\in W_\mu:\, \tau(w)\geq j} \mathrm{Re}\, f_w(x) \right|$$

$$\geq \sup_{1\leq j\leq M}\left|\sum_{w\in W_\mu:\,\tau(w)\geq j} F_w(x)\right| - \sum_{w\in W_\mu}|F_w(x)-\operatorname{Re} f_w(x)|,$$

which by (8.31) is

$$\geq \frac{1}{3}\sum_{w\in W_\mu}|F_w(x)| - \sum_{w\in W_\mu}|F_w(x)-\operatorname{Re} f_w(x)|.$$

From (8.30) and (8.32), we then have

$$\left\|\sup_{1\leq j\leq M}\left|\sum_{w\in W_\mu:\,\tau(w)\geq j} f_w\right|\right\|_{L^2} \geq \frac{\sqrt{\mu}}{12} - 2^{-10} \geq \frac{\sqrt{\mu}}{50}.$$

Hence (8.14) follows from (8.13). This finishes the proof of Proposition 8.5, except for the proof of Lemma 8.6. □

The proof of Lemma 8.6 is done by induction over the length of words. The basic step is contained in

Lemma 8.7 *Given $\varepsilon > 0$, a set E of finite measure and a set S in frequency space that contains balls of arbitrary large radii, there exist $\rho_0 > 0$, a frequency ξ_0 and a ball $B = B(\xi_0, \rho_0) \subset S$ such that $\|\phi_{\rho_0} * \mathbb{1}_E - \mathbb{1}_E\|_2 < \varepsilon$ and $\int_E |\cos(\langle \xi_0, x\rangle)|\,dx \geq |E|/3$.*

Proof. Since $\{\phi_\rho\}_{\rho>0}$ form an approximation of the identity there is $R_1 = R_1(S, E, \varepsilon)$ such that

$$\|\phi_\rho * \mathbb{1}_E - \mathbb{1}_E\|_2 < \varepsilon \tag{8.35}$$

for $\rho > R_1$. Also observe that

$$\begin{aligned}\liminf_{|\xi|\to+\infty}\int_E |\cos(\langle\xi, x\rangle)|dx &\geq \liminf_{|\xi|\to+\infty}\int_E \cos^2(\langle\xi, x\rangle)dx\\ &= \lim_{|\xi|\to+\infty}\int_E \frac{1+\cos(2\langle\xi, x\rangle)}{2}dx = \frac{|E|}{2},\end{aligned}$$

by the Riemann–Lebesgue lemma. Hence we find $R_2 = R_2(S, E, \varepsilon)$ such that

$$\int_E |\cos(\langle\xi, x\rangle)|dx \geq |E|/3, \tag{8.36}$$

for $|\xi| \geq R_2$.

By assumption on S we can find a ball B_0 of radius $R_0 > 10\max\{R_1, R_2\}$, centered at some Ξ_0 such that $B_0 \subset S$. There is a point $\xi_0 \in B(\Xi_0, R_0/2)$ that satisfies $|\xi_0| \geq R_0/4$. Set $\rho_0 = R_0/4$. The ball $B(\xi_0, \rho_0)$ is contained in B_0 and thus in S. Also since $\rho_0 \geq R_1$ we have (8.35) for $\rho = \rho_0$ and since $|\xi_0| > R_2$ we have (8.36) for $\xi = \xi_0$. □

Proof of Lemma 8.6 We will construct a sequence of sets $\{E_w\}$, radii ρ_w and modulation frequencies ξ_w using induction on the length of words. We use $\varepsilon = 2^{-\mu-10}$ in Lemma 8.7.

First let $E_\emptyset = [0,1]^2$. We apply Lemma 8.7 with $E = E_\emptyset$ and $S = S_{\tau(\emptyset)}$. We thus find $\xi_\emptyset$, $\rho_\emptyset$ such that (8.23), (8.24), (8.25) hold for $w = \emptyset$. We consider the two words of length one, i.e. 0 and 1 and let

$$E_0 := \{x \in E_\emptyset \colon \cos(\langle \xi_w, x\rangle) \geq 0\}$$
$$E_1 := \{x \in E_\emptyset \colon \cos(\langle \xi_w, x\rangle) < 0\}$$

so that $E_\emptyset$ is a disjoint union of E_0 and E_1, and (8.26) holds for $w = \emptyset$. Clearly $[0,1]^2$ is a disjoint union of the E_w with words w of length 1.

Suppose E_w, ρ_w, ξ_w are defined for all words of length $\ell < \mu - 1$. Take any word of length $\ell + 1$, of the form $w0$ or w_1 where w is of length ℓ, and where E_w, ρ_w, ξ_w satisfy (8.23), (8.24), (8.25), and where $[0,1]^2$ is a disjoint union of the E_w with $\text{length}(w) = \ell$. We let

$$E_{w0} := \{x \in E_w \colon \cos(\langle \xi_w, x\rangle) \geq 0\}$$
$$E_{w1} := \{x \in E_w \colon \cos(\langle \xi_w, x\rangle) < 0\}$$

so that (8.26) holds, E_w is a disjoint union of E_{w0} and E_{w1}, and thus $[0,1]^2$ is a disjoint union of all $E_{\overline{w}}$ where $\overline{w}$ runs over all words of length $\ell + 1$.

We now use Lemma 8.7 to find ρ_{w0}, ξ_{w0} so that (8.23), (8.24) and (8.25) hold for $w0$ in place of w. Then we use Lemma 8.7 again to find ρ_{w1}, ξ_{w1} so that (8.23), (8.24) and (8.25) hold for $w1$ in place of w.

At step $\ell = \mu - 1$ this completes our construction of E_w, ρ_w and ξ_w for all $w \in W_\mu$, and all the properties stated in Lemma 8.6 are satisfied at every stage of the construction. Note that the balls $B(\xi_w, \rho_w)$, $B(\xi_{\tilde{w}}, \rho_{\tilde{w}})$ are disjoint for different w, $\tilde{w}$ because these balls belong to the disjoint sets $S_{\tau(w)}$, $S_{\tau(\tilde{w})}$, respectively.

Finally we have by our construction, for $\ell = 0, \ldots, \mu - 1$,

$$\sum_{w:\text{length}(w)=\ell} \mathbb{1}_{E_w} = \mathbb{1}_{[0,1]^2},$$

and we obtain (8.22) by summing in ℓ. □

Acknowledgements S.G. was supported in part by a direct grant for research from the Chinese University of Hong Kong (4053295). A.S. was supported in part by National Science Foundation grants DMS 1500162 and 1764295. He would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the program Approximation, Sampling and Compression in Data Science where some work on this paper was undertaken. This work was supported by EPSRC grant no EP/K032208/1. P.Y. was supported in part by a General Research Fund CUHK14303817 from the Hong Kong Research Grant Council, and direct grants for research from the Chinese University of Hong Kong (4441563 and 4441651).
The authors thank Rajula Srivastava for reading a draft of this paper and for providing useful comments.

Appendix A: Proof of Proposition 3.4

The proof is a modification of the argument for the standard Cotlar inequality regarding truncations of singular integrals, *cf.* [26, §I.7].

Let $m_j(\xi) = \eta(2^{-j}\xi)m(\xi)$ and let $a_j(\xi) = m_j(2^j\xi)$. We pick $0 < \varepsilon < \min\{\alpha - d, 1\}$. Then by assumption

$$\sup_{j\in\mathbb{Z}} \|a_j\|_{\mathscr{L}^1_\alpha} \le B < \infty \tag{A.1}$$

which implies that $|\mathcal{F}^{-1}[a_j](x)| \le CB(1+|x|)^{-d-\varepsilon}$, and thus, with $K_j = \mathcal{F}^{-1}[m_j]$,

$$|K_j(x)| + 2^{-j}|\nabla K_j(x)| \le CB2^{jd}(1+2^j|x|)^{-d-\varepsilon}.$$

For Schwartz functions f we have $Sf = \sum_{j\in\mathbb{Z}} K_j * f$ and $S_n f = \sum_{j\le n} K_j * f$.

Lemma A.1 *Fix $\tilde{x} \in \mathbb{R}^d$ and $n \in \mathbb{Z}$, and let $g(y) = f(y)\mathbb{1}_{B(\tilde{x},2^{-n})}(y)$ and $h = f - g$. Then*

(i) $|S_n g(\tilde{x})| \lesssim B\, M[f](\tilde{x})$.
(ii) $|S_n h(\tilde{x}) - Sh(\tilde{x})| \lesssim B\, M[f](\tilde{x})$.
(iii) *For $|w - \tilde{x}| \le 2^{-n-1}$ we have* $|Sh(\tilde{x}) - Sh(w)| \lesssim B\, M[f](\tilde{x})$.

Proof. By appropriate normalization of the multiplier we may assume $B = 1$.

(i) is immediate since for $j \le n$

$$|K_j * g(\tilde{x})| \lesssim 2^{jd}\int_{|\tilde{x}-y|\le 2^{-n}} |g(y)|dy \lesssim 2^{(j-n)d} M[g](\tilde{x})$$

and the assertion follows since $|g| \le |f|$.

For (ii) notice that $|S_n h(\tilde{x}) - Sh(\tilde{x})| \le \sum_{j>n} |K_j * h(\tilde{x})|$. For $j > n$ we estimate

$$\begin{aligned} |K_j * h(\tilde{x})| &\lesssim 2^{-j\varepsilon}\int_{|\tilde{x}-y|\ge 2^{-n}} |\tilde{x}-y|^{-d-\varepsilon}|h(y)|dy \\ &\lesssim 2^{-j\varepsilon}\sum_{l\ge 0} 2^{(n-l)\varepsilon} \fint_{B(\tilde{x},2^{l-n})} |h(y)|dy \end{aligned}$$

where the slashed integral denotes the average. Thus we get

$$\sum_{j\ge n} |K_j * h(\tilde{x})| \lesssim M[h](\tilde{x})$$

and, since $|h| \le |f|$, the assertion follows.

Concerning (iii) we consider the terms $K_j * h(\tilde{x}) - K_j * h(w)$ separately for $j \le n$ and $j > n$. The term $\sum_{j>n} |K_j * h(\tilde{x})|$ was already dealt with in (ii). Since

$|w-\tilde{x}| \le 2^{-n-1}$ we have $|w-y| \approx |\tilde{x}-y|$ for $|\tilde{x}-y| \ge 2^{-n}$ and thus the previous calculation also yields

$$\sum_{j>n} |K_j * h(w)| \lesssim M[h](\tilde{x}) \lesssim Mf(\tilde{x}).$$

It remains to consider the terms for $j \le n$. In that range we write

$$K_j * h(\tilde{x}) - K_j * h(w) = \int_0^1 \int_{|\tilde{x}-y|\ge 2^{-n}} \langle \tilde{x}-w, \nabla K_j(w+s(\tilde{x}-w)-y)\rangle h(y)dy\,ds.$$

Since $|w-\tilde{x}| \le 2^{-n-1}$ we can replace $|w+s(\tilde{x}-w)-y|$ in the integrand with $|\tilde{x}-y|$ and estimate the displayed expression by $C\sum_{l\ge 0} A_{l,j,n}$ where

$$\begin{aligned} A_{l,j,n} &= 2^j|\tilde{x}-w| \int_{2^{-n+l-1}\le|\tilde{x}-y|\le 2^{-n+l}} \frac{2^{jd}}{(1+2^j|\tilde{x}-y|)^{d+\varepsilon}}|h(y)|dy \\ &\lesssim 2^{(j-n)(1-\varepsilon)}2^{-l\varepsilon} \int_{B(\tilde{x},2^{l-n})} |h(y)|dy. \end{aligned}$$

Summing in $l>0$ and then $j \le n$ yields

$$\sum_{j\le n} |K_j * h(\tilde{x}) - K_j * h(w)| \lesssim Mh(\tilde{x}) \lesssim Mf(\tilde{x}). \tag{A.2}$$

□

Proof of (3.7) We proceed arguing as in [26, §I.7]. Fix $\tilde{x} \in \mathbb{R}^d$ and $n \in \mathbb{Z}$ and define g and h as in the lemma. For (suitable) w with $|w-\tilde{x}| \le 2^{-n-1}$ we write

$$\begin{aligned} S_n f(\tilde{x}) &= S_n g(\tilde{x}) + (S_n - S)h(\tilde{x}) + Sh(\tilde{x}) \\ &= S_n g(\tilde{x}) + (S_n - S)h(\tilde{x}) + Sh(\tilde{x}) - Sh(w) + Sf(w) - Sg(w). \end{aligned} \tag{A.3}$$

By Lemma A.1

$$|S_n g(\tilde{x})| + |(S_n - S)h(\tilde{x})| + |Sh(\tilde{x}) - Sh(w)| \lesssim B\, M[f](\tilde{x})$$

and it remains to consider the term $Sf(w) - Sg(w)$ for a substantial set of w with $|w-\tilde{x}| \le 2^{-n-1}$.

By the Mikhlin-Hörmander theorem we have for all $f \in L^1(\mathbb{R}^d)$ and all $\lambda > 0$

$$\text{meas}(\{x : |Sf(x)| > \lambda\}) \le A\lambda^{-1}\|f\|_1$$

where $A \le C_{\alpha,d} B$.

Now let $\delta \in (0, 1/2)$ and consider the set

$$\Omega_n(\tilde{x}, \delta) = \big\{w : |w - \tilde{x}| < 2^{-n-1}, \quad |Sg(w)| > 2^d \delta^{-1} A\, M[f](\tilde{x})\big\}.$$

In (A.3) we can estimate the term $|Sg(w)|$ by $2^d\delta^{-1}A\,M[f](\tilde{x})$ when $w \in B(\tilde{x}, 2^{-n-1})\backslash\Omega_n(\tilde{x}, \delta)$. Hence we obtain

$$|S_n f(\tilde{x})| \le \inf_{w\in B(\tilde{x},2^{-n-1})\backslash\Omega_n(\tilde{x},\delta)} |Sf(w)| + C(\alpha, d)B(1 + \delta^{-1})M[f](\tilde{x}). \quad \text{(A.4)}$$

By the weak type inequality for S we have

$$\begin{aligned}\operatorname{meas}(\Omega_n(\tilde{x}, \delta)) &\le \frac{A\|g\|_1}{2^d\delta^{-1}AM[f](\tilde{x})} = \frac{\delta}{2^d M[f](\tilde{x})}\int_{|\tilde{x}-y|\le 2^{-n}} |f(y)|dy\\ &\le \delta\, 2^{-d}\operatorname{meas}(B(\tilde{x}, 2^{-n})) = \delta\operatorname{meas}(B(\tilde{x}, 2^{-n-1})).\end{aligned}$$

Hence $\operatorname{meas}(B(\tilde{x}, 2^{-n-1})\backslash\Omega_n(\tilde{x}, \delta)) \ge (1 - \delta)\operatorname{meas}(B(\tilde{x}, 2^{-n-1}))$ and thus for all $r > 0$

$$\begin{aligned}&\inf_{w\in B(\tilde{x},2^{-n-1})\backslash\Omega_n(\tilde{x},\delta)} |Sf(w)|\\ &\le \left(\frac{1}{\operatorname{meas}(B(\tilde{x}, 2^{-n-1})\backslash\Omega_n(\tilde{x}, \delta))}\int_{B(\tilde{x},2^{-n-1})} |Sf(w)|^r dw\right)^{1/r}\\ &\le \left(\frac{1}{(1-\delta)|B(\tilde{x}, 2^{-n-1})|}\int_{B(\tilde{x},2^{-n-1})} |Sf(w)|^r dw\right)^{1/r}.\end{aligned}$$

We obtain

$$|S_n f(\tilde{x})| \le (1 - \delta)^{-1/r}(M[|Sw|^r](\tilde{x}))^{1/r} + C(\alpha, d)(1 + \delta^{-1})B\, M[f](\tilde{x})$$

uniformly in n. This implies (3.7). □

References

1. Bateman, M.: Single annulus L^p estimates for Hilbert transforms along vector fields. Rev. Mat. Iberoam. **29**(3), 1021–1069 (2013)
2. Bateman, M., Thiele, C.: L^p estimates for the Hilbert transforms along a one-variable vector field. Anal. PDE **6**(7), 1577–1600 (2013)
3. Carbery, A., Seeger, A.: H^p- and L^p-variants of multiparameter Calderón-Zygmund theory. Trans. Am. Math. Soc. **334**(2), 719–747 (1992)
4. Chang, S.-Y.A., Wilson, J.M., Wolff, T.H.: Some weighted norm inequalities concerning the Schrödinger operators. Comment. Math. Helv. **60**(2), 217–246 (1985)
5. Córdoba, A., Rubio de Francia, J.L.: Estimates for Wainger's singular integrals along curves. Rev. Mat. Iberoam. **2**(1–2), 105–117 (1986)
6. Demeter, C.: Singular integrals along N directions in $\mathbb{R}^2$. Proc. Am. Math. Soc. **138**(12), 4433–4442 (2010)

7. Demeter, C., Di Plinio, F.: Logarithmic L^p bounds for maximal directional singular integrals in the plane. J. Geom. Anal. **24**(1), 375–416 (2014)
8. Di Plinio, F., Guo, S., Thiele, C., Zorin-Kranich, P.: Square functions for bi-Lipschitz maps and directional operators. J. Funct. Anal. **275**(8), 2015–2058 (2018)
9. Di Plinio, F., Parissis, I.: A sharp estimate for the Hilbert transform along finite order lacunary sets of directions. Israel J. Math. **227**(1), 189–214 (2018)
10. Duoandikoetxea, J., Rubio de Francia, J.L.: Maximal and singular integral operators via Fourier transform estimates. Invent. Math. **84**(3), 541–561 (1986)
11. Fefferman, C., Stein, E.M.: Some maximal inequalities. Am. J. Math. **93**(1), 107–115 (1971)
12. Fefferman, R.: Harmonic analysis on product spaces. Ann. Math. **126**(1), 109–130 (1987)
13. Fefferman, R., Lin, K.C.: A sharp Marcinkiewicz multiplier theorem **(unpublished)**
14. Grafakos, L., Honzík, P., Seeger, A.: On maximal functions for Mikhlin–Hörmander multipliers. Adv. Math. **204**(2), 363–378 (2006)
15. Guo, S.: Hilbert transform along measurable vector fields constant on Lipschitz curves: L^2 boundedness. Anal. PDE **8**(5), 1263–1288 (2015)
16. Guo, S., Hickman, J., Lie, V., Roos, J.: Maximal operators and Hilbert transforms along variable non-flat homogeneous curves. Proc. Lond. Math. Soc. **115**(1), 177–219 (2017)
17. Guo, S., Roos, J., Seeger, A., Yung, P.L.: Maximal functions associated with families of homogeneous curves: L^p bounds for $p \leq 2$ (2019). arXiv:1906.05997 **(Preprint)**
18. Karagulyan, G.A.: On unboundedness of maximal operators for directional Hilbert transforms. Proc. Am. Math. Soc. **135**(10), 3133–3141 (2007)
19. Łaba, I., Marinelli, A., Pramanik, M.: On the maximal directional Hilbert transform (2017). arXiv:1707.01061
20. Marletta, G., Ricci, F.: Two-parameter maximal functions associated with homogeneous surfaces in $\mathbb{R}^n$. Stud. Math. **130**(1), 53–65 (1998)
21. Mockenhaupt, G., Seeger, A., Sogge, C.D.: Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. J. Am. Math. Soc. **6**(1), 65–130 (1993)
22. Nikišin, E.M., Ul'janov, P.L.: On absolute and unconditional convergence. Uspehi Mat. Nauk **22**(3), 240–242 (1967) **(Russian)**
23. Seeger, A.: Some inequalities for singular convolution operators in l^p-spaces. Trans. Am. Math. Soc. **308**(1), 259–272 (1988)
24. Seeger, A., Tao, T., Wright, J.: Endpoint mapping properties of spherical maximal operators. J. Inst. Math. Jussieu **2**(1), 109–144 (2003)
25. Stein, E.M.: Singular integrals and differentiability properties of functions. Princeton Mathematical Series, no. 30, pp. xiv+290. Princeton University Press, Princeton (1970)
26. Stein, E.M.: Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, vol. 43. Monographs in Harmonic Analysis, III, pp. xiv+695. Princeton University Press, Princeton (1993)
27. Stein, E.M., Wainger, S.: Problems in harmonic analysis related to curvature. Bull. Am. Math. Soc. **84**(6), 1239–1295 (1978)