THE CIRCULAR MAXIMAL OPERATOR ON HEISENBERG RADIAL FUNCTIONS

DAVID BELTRAN   SHAOMING GUO
JONATHAN HICKMAN   ANDREAS SEEGER

ABSTRACT. Lebesgue space estimates are obtained for the circular maximal function on the Heisenberg group $\mathbb{H}^1$ restricted to a class of Heisenberg radial functions. Under this assumption, the problem reduces to studying a maximal operator on the Euclidean plane. This operator has a number of interesting features: it is associated to a non-smooth curve distribution and, furthermore, fails both the usual rotational curvature and cinematic curvature conditions.

1. INTRODUCTION

Let $\mathbb{H}^n$ denote the Heisenberg group given by endowing $\mathbb{R} \times \mathbb{R}^{2n}$ with the non-commutative group operation

$$(u,x) \cdot (v,y) := \big(u + v + x^\top By, x + y\big) \qquad \text{for all } (u,x),(v,y) \in \mathbb{R} \times \mathbb{R}^{2n}$$

where $B = bJ$ with $J := \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$ the matrix associated to the standard symplectic form on $\mathbb{R}^{2n}$ and $b \neq 0$ (usually one takes $b = 1/2$).

Let $\mu \equiv \mu_1$ denote the normalised surface measure on the sphere

$$\{0\} \times S^{2n-1} := \{(0,y) \in \mathbb{R} \times \mathbb{R}^{2n} : |y| = 1\}.$$

If $\mathrm{Dil}_t(u,x) := (t^2u, tx)$ are the automorphic dilations on $\mathbb{H}^n$, then the normalised surface measure μ_t supported on tS^{2n-1} can be viewed as a dilate of μ_1 in the sense that $\langle f, \mu_t\rangle = \langle f(\mathrm{Dil}_t\,\cdot\,), \mu\rangle$.

Given a function f on $\mathbb{H}^n$ belonging to a suitable *a priori* class consider the spherical means

$$f * \mu_t(u,x) := \int_{S^{2n-1}} f(u - tx^\top By, x - ty)\,\mathrm{d}\mu(y) \quad \text{for } (u,x) \in \mathbb{H}^n \text{ and } t > 0.$$

For smooth functions f one has $f * \mu_t(u,x) \to f(u,x)$ pointwise as $t \to 0$. It is of interest to extend this convergence result to an almost everywhere convergence result for functions on $L^p(\mathbb{H}^n)$, in a suitable range of p. Such a result follows from L^p bounds for the associated spherical maximal function

$$Mf(u,x) := \sup_{t>0} |f * \mu_t(u,x)|. \tag{1.1}$$

The operator M can be understood as a Heisenberg analogue of the classical (Euclidean) spherical maximal function of Stein [31] and Bourgain [5] (see also [19, 29, 28]). The maximal function (1.1) was introduced by Nevo and Thangavelu

Date: January 11, 2021.
2010 *Mathematics Subject Classification.* 42B25, 22E25, 43A80, 35S30.

in [23] where L^p estimates were proven in dimensions $n \geq 2$ for p belonging to a non-sharp range. By choosing f to be the standard example

$$f(u,x) := \big(|x| \log(1/|x|)\big)^{1-2n} \chi(u,x)$$

for an appropriate choice of cutoff function χ, it follows that $L^p \to L^p$ estimates can only hold for $p > \frac{2n}{2n-1}$. For $n \geq 2$ the sufficiency of this condition was established independently by Müller and the fourth author [21] and by Narayanan and Thangavelu [22]; the work in [21] also treats a wider class of operators defined on Métivier groups. Results in a more general variable coefficient setting can be found in a recent paper by Kim [15]. Related to these investigations the L^p results of [21, 22] were extended in [1] to deal with variants of the operator (1.1) where the original sphere, centred at the origin, does not lie in the subspace $\{0\} \times \mathbb{R}^{2n}$ (that is, the corresponding dilates of μ are no longer supported in a fixed hyperplane). The latter paper is closely related to [26], [27] which establish sharp L^p-Sobolev bounds for certain Radon-type operators associated to curves in three-dimensional manifolds; in particular [27] covers the averages $f \mapsto f * \mu_t$ in $\mathbb{H}^1$, and perturbations of these operators, when acting on compactly supported functions. Mapping properties and sparse domination for a lacunary version of M have been recently studied in [2], also under the assumption $n \geq 2$. We note that for the proofs of the positive results on the Heisenberg spherical maximal functions mentioned above it was essential that a boundedness result holds for $p = 2$, which leads to the restriction $n \geq 2$. Such an L^2 result fails to hold on $\mathbb{H}^1$, and it is currently not known whether the circular maximal operator (1.1) on the Heisenberg group $\mathbb{H}^1$ is bounded on $L^p(\mathbb{H}^1)$ for *any* $p < \infty$.

In this paper we consider the problem of estimating the maximal function (1.1) on the sub-algebra of *Heisenberg-radial* (or *$\mathbb{H}$-radial*) functions on $\mathbb{H}^1$. Here a function $f \colon \mathbb{H}^1 \to \mathbb{C}$ is said to be $\mathbb{H}$-radial if $f(u, Rx) = f(u,x)$ for all $R \in \mathrm{SO}(2)$. Given the underlying symmetries of the maximal operator, this is a natural condition to impose on the input function: indeed, if f is $\mathbb{H}$-radial then, Mf is also $\mathbb{H}$-radial. Our main theorem characterises the L^p mapping properties of M acting on $\mathbb{H}$-radial functions.

Theorem 1.1. *For $2 < p \leq \infty$ the a priori estimate*

$$\|Mf\|_{L^p(\mathbb{H}^1)} \leq C_p \|f\|_{L^p(\mathbb{H}^1)}$$

holds for $\mathbb{H}$-radial functions on $\mathbb{H}^1$. Here C_p is a constant depending only on p.

We shall reduce Theorem 1.1 to bounding a maximal function $\sup_{t>0} |A_t f|$ where the A_t are non-convolution averaging operators in two dimensions. We aim to follow the strategy used in [19, 20] to study the Euclidean circular maximal function and its relatives. However, in comparison with [20], substantial new difficulties arise. First, we need to consider a distribution of curves which is not smooth. Moreover, the *rotational curvature* and *cinematic curvature* conditions (as formulated in [30, 20]) fail to hold, and hence $\sup_{t>0} |A_t f|$ does not belong to the classes of variable coefficient maximal functions considered in [20]. Significant technical challenges are encountered when dealing with the various singularities of the operator, and our arguments are based on the analysis of a class of oscillatory integral operators with 2-sided fold singularities which extends the work in [25] and [8]. A more detailed discussion of the proof strategy can be found in §2 below.

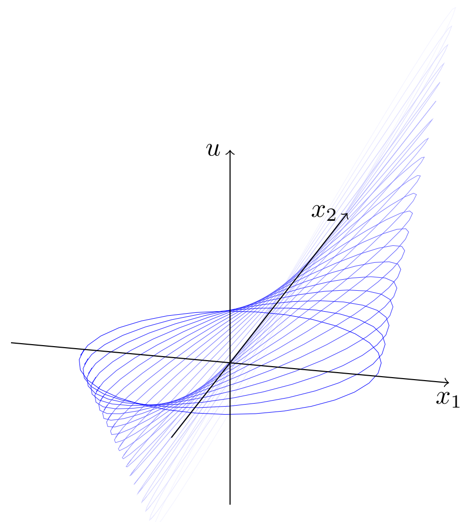

FIGURE 1. The unit circle tilts and stretches as it is translated along the x_2-axis under the Heisenberg operation.

Structure of the paper. Section 2 reviews the strategy for bounding the Euclidean circular maximal function based on local smoothing estimates. The difficulties encountered in our particular situation are also described. In Sections 3 – 8 we prove bounds for a local variant of M, where the supremum is restricted to $1 \leq t \leq 2$. In particular, Section 3 reduces Theorem 1.1 to a bound for a maximal function in two dimensions. Section 4 describes notions of curvature which feature in the analysis of M. In Section 5 the maximal function is decomposed into different pieces according to curvature considerations. In Section 6 we consider classes of oscillatory integral operators depending on two parameters which are crucial for the relevant L^2-theory, mainly based on a 'fixed-time' analysis. In Section 7 we apply these L^2 estimates to the problem on the Heisenberg group. In Section 8 we discuss the L^p theory, based on L^p space-time ('local smoothing') estimates. Finally, in Section 9 the bounds for the local maximal function are extended to bounds for M. Two appendices are included for the reader's convenience, providing useful integration-by-parts lemmata and many explicit computations helpful to the analysis.

Notational conventions. Given a (possibly empty) list of objects L, for real numbers $A_p, B_p \geq 0$ depending on some Lebesgue exponent p the notation $A_p \lesssim_L B_p$, $A_p = O_L(B_p)$ or $B_p \gtrsim_L A_p$ signifies that $A_p \leq CB_p$ for some constant $C = C_{L,p} \geq 0$ depending on the objects in the list and p. In addition, $A_p \sim_L B_p$ is used to signify that both $A_p \lesssim_L B_p$ and $A_p \gtrsim_L B_p$ hold. Given a, $b \in \mathbb{R}$ we write $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. Given $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ we will often write $x = (x_1, x'') \in \mathbb{R} \times \mathbb{R}^2$ or $x = (x', x_3) \in \mathbb{R}^2 \times \mathbb{R}$. Given $x \in \mathbb{R}^2$ and $t \in \mathbb{R}$ we will also often write $\vec{x} = (x, t) \in \mathbb{R}^2 \times \mathbb{R}$. Throughout the article N denotes some fixed large integer, chosen so as to satisfy the forthcoming requirements of the proofs. The choice of $N = 10^{1000}$ is permissible (and in the d-dimensional version of estimates in Sections 6 and 7, it never needs to exceed

10^{100d}). For a phase function $\varphi(x;z)$ the notation $\partial^2_{xz}\varphi$ refers to the matrix A with entries $A_{ij} = \partial^2_{x_i z_j}\varphi$ while the notation $\partial^2_{zx}\varphi$ refers to its transpose. The length of a multiindex $\alpha \in \mathbb{N}_0^d$ is given by $|\alpha| = \sum_{i=1}^d \alpha_i$. The C^N norm of $(x;z) \mapsto a(x;z)$ is given by $\max_{|\alpha|+|\beta|\le N} \|\partial_x^\alpha \partial_z^\beta a\|_\infty$. We also use the notation $\|a\|_{C_z^N}$ for $\sup_x \|a(x;\cdot)\|_{C^N}$. For a linear operator T bounded from L^p to L^q we use both $\|T\|_{L^p\to L^q}$, $\|T\|_{p\to q}$ as a notation for the operator norm. For a one-parameter family of linear operators $\{T_t\}_{t\in E}$, $\|\sup_{t\in E}|T_t|\|_{p\to q}$ denotes the $L^p \to L^q$ operator norm of the sublinear operator $f \mapsto \sup_{t\in E}|T_t f|$.

Acknowledgements

The authors thank the American Institute of Mathematics for funding their collaboration through the SQuaRE program, also supported in part by the National Science Foundation. D.B. was partially supported by ERCEA Advanced Grant 2014 669689 - HADE, by the MINECO project MTM2014-53850-P, by Basque Government project IT-641-13 and also by the Basque Government through the BERC 2018-2021 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2017-0718. S.G. was partially supported by NSF grant DMS-1800274. A.S. was partially supported by NSF grant DMS-1764295 and by a Simons fellowship. This material is partly based upon work supported by the National Science Foundation under Grant No. DMS-1440140 while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. The authors would also like to thank the anonymous referee for a careful reading and valuable suggestions.

2. Proof strategy

Theorem 1.1 easily reduces to bounding a maximal function $\sup_{t>0}|A_t f|$ where the A_t are averaging operators on the Euclidean plane. We aim to follow the broad strategy introduced in [19] to study the Euclidean circular maximal function, which we now recall. Define $A_t^{\mathrm{eucl}} f$ by taking $A_t^{\mathrm{eucl}} f(x)$ to be the average of f over the circle $\Sigma_{x,t}^{\mathrm{eucl}}$ in the plane centred at x with radius t. Note that the associated curve distribution is described by the defining function

$$\Phi^{\mathrm{eucl}}(x,t;y) := |x-y|^2 - t^2 \qquad \text{for } (x,t;y) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2;$$

in particular, $\Sigma_{x,t}^{\mathrm{eucl}} = \{y \in \mathbb{R}^2 : \Phi(x,t;y) = 0\}$. The associated maximal function

$$M^{\mathrm{eucl}} f(x) := \sup_{t>0} |A_t^{\mathrm{eucl}} f(x)|$$

is the classical circular maximal function studied by Bourgain [5] and also in [19]. A Littlewood–Paley argument reduces the problem of bounding $M^{\mathrm{eucl}} f$ to bounding the local maximal function

$$\sup_{1\le t\le 2} |A_t^{\mathrm{eucl}} f(x)|.$$

Decompose the averaging operator $A_t^{\mathrm{eucl}} f$ as a sum of pieces $A_t^{\mathrm{eucl},j} f$ localised at frequency scale 2^j. The sum of the low frequency pieces ($j \le 0$) can be bounded in one go via comparison with the Hardy–Littlewood maximal operator and it remains to bound the high frequency pieces. There are two steps in the argument:

i) The first step is to show that the inequality
$$\big\| \sup_{1\le t\le 2} |A_t^{\mathrm{eucl},j} f|\big\|_{L^2(\mathbb{R}^2)} \le C\|f\|_{L^2(\mathbb{R}^2)} \tag{2.1}$$
holds uniformly in j. An elementary Sobolev embedding reduces (2.1) to proving L^2 estimates for certain oscillatory integral operators. A T^*T argument further reduces (2.1) to bounding the corresponding kernels, which are then amenable to stationary phase analysis.

ii) Interpolating (2.1) with the trivial L^∞ estimate,
$$\big\| \sup_{1\le t\le 2} |A_t^{\mathrm{eucl},j} f|\big\|_{L^p(\mathbb{R}^2)} \le C\|f\|_{L^p(\mathbb{R}^2)} \qquad \text{for all } 2\le p\le\infty. \tag{2.2}$$
The problem here is that (2.2) does not sum in j. If, however, there exists some $2<p_\circ<\infty$ and $\varepsilon(p_\circ)>0$ such that
$$\big\| \sup_{1\le t\le 2} |A_t^{\mathrm{eucl},j} f|\big\|_{L^{p_\circ}(\mathbb{R}^2)} \le C2^{-j\varepsilon(p_\circ)}\|f\|_{L^{p_\circ}(\mathbb{R}^2)}, \tag{2.3}$$
then one may interpolate (2.2) and (2.3) to obtain favourable j dependence for all $2<p<\infty$, concluding the proof. The strategy in [19] is to prove a bound of the form (2.3) via local in time L^p space-time bounds (so-called *local smoothing estimates*) for the wave equation.

There are two key properties of the circular maximal function which allow the above analysis to be carried out, both of which can be expressed in terms of the defining function Φ^{eucl}. The first is the standard decay properties of the Fourier transform of surface carried measure which correspond to nonvanishing of the *Phong–Stein rotational curvature* (see, for instance, [32, Chapter IX, §3.1],[1]). This is used to prove the oscillatory integral estimates i). The second is that the *cinematic curvature* (see, [30]) is non-vanishing, which features in the proof of the local smoothing estimates used in ii). The analysis can be generalised to variable coefficient maximal functions formed by averaging operators on the plane associated to defining functions Φ which satisfy these two conditions [30].

Now suppose $A_t f$ denote the averaging operators on $\mathbb{R}^2$ which arises in the study of our maximal operator acting on $\mathbb{H}$-radial functions. This family of operators has an associated defining function Φ, which is described in (3.2) below. As before, one may decompose $A_t f$ as a sum of pieces $A_t^j f$ localised at a frequency scale 2^j. Significant issues arise, however, when it comes to implementing either of the above steps to analyse the $A_t^j f$ in this case:

i′) The defining function Φ has vanishing rotational curvature. Indeed, the oscillatory integral estimates in the above proof sketch of (2.1) do not hold in this case.

ii′) The defining function Φ also has vanishing cinematic curvature. This precludes direct application of local smoothing estimates in the proof of (2.3).

In order to deal with these issues it is necessary decompose the operator A_t with respect to the various curvatures and to prove bounds of the form (2.1), (2.2) and (2.3) for each of the localised pieces.

In bounding the localised pieces of A_t, the main difficulty is caused by the vanishing of the rotational curvature. In particular, here the L^2 theory relies on certain

[1]The definitions of the rotational curvature and other concepts featured in this discussion are also reviewed in §4 below.

two parameter variants of estimates for oscillatory integral operators with two-sided fold singularities. Our arguments build on the techniques in [8, 11]. This is in contrast with the analysis of the Euclidean maximal function, where the classical estimates for non-degenerate oscillatory integral operators of Hörmander [13] suffice. The presence of a two-sided fold incurs a (necessary) loss in the oscillatory integral estimates (compared with the non-vanishing rotational curvature case), but special properties of the Heisenberg maximal function allow one to compensate for this. A similar phenomenon was previously observed in the analysis of the spherical maximal function in $\mathbb{H}^n$ for $n > 1$ in [21].

The vanishing of the cinematic curvature presents less of a problem, essentially because the desired bound (2.3) is non-quantitative: all that is required is for (2.3) to hold for *some* $p_\circ$ and *some* $\varepsilon(p_\circ) > 0$. Roughly speaking, the strategy is to decompose the operator into two parts: one piece supported on the δ-neighbourhood of the variety where the cinematic curvature vanishes and a complementary piece. The former is dealt using a variant of (2.2) which includes a gain in δ arising from the additional localisation. The latter piece has non-vanishing cinematic curvature and can be dealt with using local smoothing estimates. Choosing δ appropriately, one obtains the desired bound. Similar ideas were used by Kung [17] to treat a family of Fourier integral operators with vanishing cinematic curvature.

3. Reduction to a maximal operator in the plane

3.1. **Singular support of the Schwartz kernel and implicit definition.** A computation shows that $f * \mu_t(u, x)$ corresponds to an average of f over the ellipse in $\mathbb{R}^3$ given by

$$S_{u,x,t} := \big\{(v, z) \in \mathbb{R} \times \mathbb{R}^2 : v - u + b(x_1 z_2 - x_2 z_1) = 0, \quad |x - z|^2 - t^2 = 0\big\}.$$

Furthermore, using the identity $(x_1z_1 + x_2z_2)^2 + (x_1z_2 - x_2z_1)^2 = |x|^2|z|^2$, one checks that $(v, z) \in S_{u,x,t}$ satisfies

$$\Phi_t(u, |x|; v, |z|) = 0 \tag{3.1}$$

where $\Phi_t(u, r; v, \rho) := \Phi(u, r, t; v, \rho)$ and

$$\Phi(u, r, t; v, \rho) := (u - v)^2 - \Big(\frac{b}{2}\Big)^2 \big(4r^2\rho^2 - (r^2 + \rho^2 - t^2)^2\big). \tag{3.2}$$

Below we relate explicitly $f * \mu_t$ to an operator acting on functions of the two variables (v, ρ), with a Schwartz kernel $\delta \circ \Phi$ which will define this integral operator as a weakly singular Radon transform.

In the forthcoming sections it will be necessary to carry out many computations involving Φ. For the reader's convenience, a dictionary of derivatives of this function is provided in Appendix B.1.

3.2. **Properties of $\mathbb{H}$-radial functions.** A function $f : \mathbb{H}^1 \to \mathbb{C}$ is $\mathbb{H}$-radial if and only if there exists some function $f_0 : \mathbb{R} \times [0, \infty) \to \mathbb{C}$ such that

$$f(u, x) = f_0(u, |x|). \tag{3.3}$$

Using the fact that $R^\top BR = B$ for $R \in \mathrm{SO}(2)$, if f and g are $\mathbb{H}$-radial, then $f * g$ is $\mathbb{H}$-radial, and we have

$$(f*g)_0(u, r) = \int_0^{2\pi}\int_{\mathbb{R}}\int_0^\infty f_0(v, \rho)g_0(u-v-br\rho \sin\vartheta, \sqrt{r^2 + \rho^2 - 2r\rho\cos\vartheta})\, \rho \,\mathrm{d}\rho\mathrm{d}v\mathrm{d}\vartheta.$$

This observation extends to $\mathbb{H}$-radial measures and, in particular, if f is $\mathbb{H}$-radial, then $f * \mu_t$ is $\mathbb{H}$-radial, and we get

$$\begin{aligned}(f*\mu_t)_0(u,r) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_0(u-btr\sin\vartheta, \sqrt{r^2+t^2-2rt\cos\vartheta})\,\mathrm{d}\vartheta \\ &= \sum_{\pm}\frac{1}{2\pi}\int_0^{\pi} f_0(u\pm btr\sin\vartheta, \sqrt{r^2+t^2-2rt\cos\vartheta})\,\mathrm{d}\vartheta. \qquad (3.4)\end{aligned}$$

Applying polar coordinates in the planar slices $\{u\}\times\mathbb{R}^2$, given $p>2$ and f as in (3.3), the goal is to establish the inequality

$$\Big(\int_{-\infty}^{\infty}\int_0^{\infty}\big|(Mf)_0(u,r)\big|^p r\,\mathrm{d}r\mathrm{d}u\Big)^{1/p} \lesssim \Big(\int_{-\infty}^{\infty}\int_0^{\infty}\big|f_0(v,\rho)|^p\rho\,\mathrm{d}\rho\,\mathrm{d}v\Big)^{1/p}. \quad (3.5)$$

3.3. **A weakly singular Radon-type operator on $\mathbb{R}^2$.** By the implicit definition of the circle $S_{u,x,t}$ from (3.1), the function $(f*\mu_t)_0$ corresponds to an integral operator associated with the curve

$$\Sigma_{u,r,t} := \big\{(v,\rho)\in\mathbb{R}\times[0,\infty) : \Phi_t(u,r;v,\rho)=0\big\}.$$

It is easy to see that $\Sigma_{u,r,t}$ is smooth whenever $r\neq t>0$. If $r=t>0$, then there is a unique singular point on the curve at the point where it touches the v axis. See Figure 2. Furthermore, any $(v,\rho)\in\Sigma_{u,r,t}$ satisfies

$$|r-t|\le\rho\le r+t \quad\text{and}\quad |u-v|\le|b|\min\{r\rho, rt, t\rho\}; \qquad (3.6)$$

these bounds follow since for $(v,\rho)\in\Sigma_{u,r,t}$

$$\begin{aligned}0\le(b/2)^{-2}(u-v)^2 &= 4r^2\rho^2-(r^2+\rho^2-t^2)^2\\ &= 4r^2t^2-(r^2+t^2-\rho^2)^2\\ &= 4t^2\rho^2-(t^2+\rho^2-r^2)^2. \qquad (3.7)\end{aligned}$$

Consider the integral operator in two dimensions defined on functions of the variables (v,ρ) by

$$A_tf(u,r)\equiv A_{p,t}f(u,r) := \int_{-\infty}^{\infty}\int_0^{\infty} f(v,\rho)r^{1/p}\rho^{1-1/p}\delta\big(\Phi_t(u,r;v,\rho)\big)\mathrm{d}v\mathrm{d}\rho. \quad (3.8)$$

In view of (3.5), Theorem 1.1 will be a consequence of the following maximal estimate in the Euclidean plane.

Theorem 3.1. *For all $p>2$,*

$$\Big(\int_0^{\infty}\int_{-\infty}^{\infty}\big(\sup_{t>0}|A_tf(u,r)|\big)^p\,\mathrm{d}u\,\mathrm{d}r\Big)^{1/p}\lesssim\Big(\int_0^{\infty}\int_{-\infty}^{\infty}|f(v,\rho)|^p\,\mathrm{d}v\,\mathrm{d}\rho\Big)^{1/p}.$$

Note that the $r^{1/p}\rho^{-1/p}$ factor featured in the averaging operator in (3.8) arises from the weights induced by the polar coordinates in (3.5). In order to relate Theorem 1.1 to Theorem 3.1 we have to write for $\mathbb{H}$-radial test functions the expression $(f*\mu_t)_0(u,r)$ in terms of the distribution $\delta\circ\Phi_t$ which is understood as a weak limit of $\chi_\varepsilon\circ\Phi_t$ as $\varepsilon\to 0$. The calculation, which is given in the proof of Lemma 3.2 below, is standard, (*cf.* [32, p.498] which provides a proof for a local version). For the sake of convenience we include below a direct proof for our example.

In what follows we shall use, for a continuous compactly supported function g the integral notation $g(c)=\int g(v)\delta(c-v)\,\mathrm{d}v$ for the pairing of g with the Dirac measure

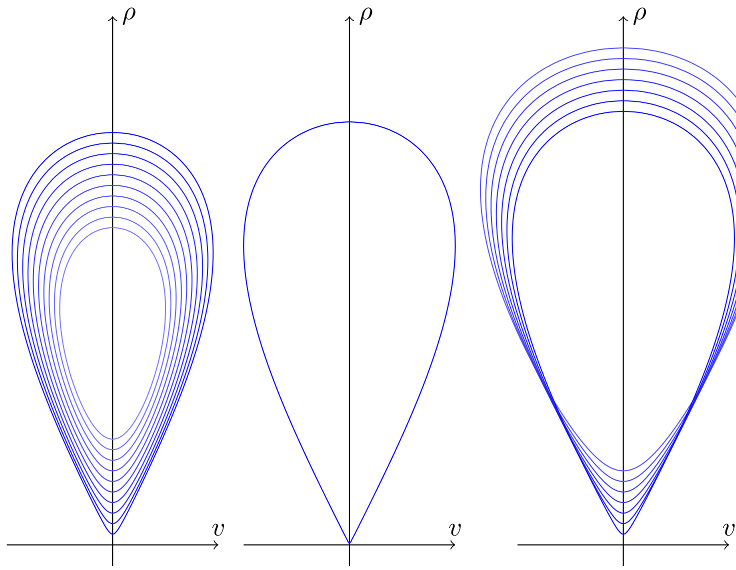

FIGURE 2. The curves $\Sigma_{0,r,t}$ for t fixed and $r < t$ (left), $r = t$ (centre) and $r > t$ (right). When $r = t$ the curve has a unique singular point on the v axis.

at c. We also let $\chi_\varepsilon(s) = \varepsilon^{-1}\chi(\varepsilon^{-1}s)$ with χ even and supported in $(-1/2, 1/2)$ such that $\int \chi \mathrm{d}s = 1$. We shall prove the following.

Lemma 3.2. *Let $f \in C^1(\mathbb{H}^1)$ be $\mathbb{H}$-radial and compactly supported in $\{(v,\rho) \in \mathbb{R}^2 : \rho > 0\}$. Then, for any $r > 0$,*

$$\begin{aligned}(f * \mu_t)_0(u,r) &= \frac{|b|}{\pi}\lim_{\varepsilon\to 0}\int_0^\infty\int_{\mathbb{R}} \chi_\varepsilon(\Phi_t(u,r;v,\rho)) f_0(v,\rho)\rho \,\mathrm{d}v\,\mathrm{d}\rho \\ &=: \frac{|b|}{\pi}\int_0^\infty\int_{\mathbb{R}} \delta(\Phi_t(u,r;v,\rho)) f_0(v,\rho)\rho\,\mathrm{d}v\,\mathrm{d}\rho.\end{aligned}$$

With the above lemma in hand, Theorem 3.1 immediately implies Theorem 1.1.

Proof that Theorem 3.1 implies Theorem 1.1. We prove the a priori inequality for smooth $\mathbb{H}$-radial functions which are compactly supported in $\{(u,y) \in \mathbb{R}^3 : |y| \neq 0\}$. By Lemma 3.2

$$r^{1/p}(Mf)_0(u,r) = \frac{|b|}{\pi}\sup_{t>0} A_t[\rho^{1/p}f_0](u,r),$$

and the assertion follows. $\square$

Proof of Lemma 3.2. We use (3.4) and make a change of variable by setting

$$\rho = \rho(\vartheta) = \sqrt{r^2 + t^2 - 2rt\cos\vartheta}, \quad 0 < \vartheta < \pi.$$

Observe that the condition $0 < \vartheta < \pi$ is equivalent with $|r - t| < \rho < r + t$. Then

$$u \pm btr\sin\vartheta = u \pm btr\sqrt{1 - \Big(\frac{r^2+t^2-\rho^2}{2rt}\Big)^2} = u \pm \frac{b}{2}G(r,t,\rho)$$

where

$$G(r,t,\rho) := \sqrt{4r^2t^2 - (r^2+t^2-\rho^2)^2}.$$

For the relevant range $|r-t| < \rho < r+t$ the root is well defined (as $\sin\vartheta > 0$), and we have the factorisation

$$G(r,t,\rho) = \Big((r+t+\rho)(r+t-\rho)(r-t+\rho)(t-r+\rho)\Big)^{1/2}. \tag{3.9}$$

We calculate

$$\frac{d\rho}{d\vartheta} = \rho^{-1} rt \sin(\vartheta) = (2\rho)^{-1} G(r,t,\rho)$$

and thus

$$\begin{aligned}\pi(f*\mu_t)_0(u,r) &= \sum_{\pm} \int_{|r-t|}^{r+t} f_0(u \pm \tfrac{b}{2} G(r,t,\rho), \rho) \frac{\rho}{G(r,t,\rho)}\, \mathrm{d}\rho \\ &= \sum_{\pm} \lim_{\varepsilon\to 0} \int_{|r-t|}^{r+t} \int_{\mathbb{R}} \rho f_0(v,\rho) \chi_\varepsilon(u \pm \tfrac{b}{2} G(r,t,\rho) - v)\, \mathrm{d}v\, \frac{1}{G(r,t,\rho)}\, \mathrm{d}\rho.\end{aligned}$$

Let U be an open interval with compact closure contained in $(0,\infty)$ such that $\mathrm{supp}\,(f_0(u,\cdot)) \subset U$ for all $u \in \mathbb{R}$. Let $U(r,t) = \{\rho \in U : |r-t| < \rho < r+t\}$. We observe from (3.9) that for fixed r,t with $r \neq t$, the function $\rho \mapsto |G(r,t,\rho)|^{-1}$ satisfies

$$\int_{U(r,t)} |G(r,t,\rho)|^{-p} \mathrm{d}\rho \le C(r,t) < \infty \quad \text{ for } 1 \le p < 2, \tag{3.10}$$

which we use for $p > 1$. Let $E_\varepsilon(r,t) = \{\rho \in U(r,t) : G(r,t,\rho) \le \varepsilon^{1/2}\}$ and $F_\varepsilon(r,t) = U(r,t) \setminus E_\varepsilon(r,t)$. We use Hölder's inequality to bound

$$\begin{aligned}&\int_{E_\varepsilon(r,t)} \int_{\mathbb{R}} \rho\, |f_0(v,\rho)| |\chi_\varepsilon(u \pm \tfrac{b}{2} G(r,t,\rho) - v)|\, \mathrm{d}v\, \frac{1}{G(r,t,\rho)}\, \mathrm{d}\rho \\ &\qquad\qquad \lesssim_{r,t,f} |E_\varepsilon(r,t)|^{1/p'} C(r,t)^{1/p} = O(\varepsilon^{(p-1)/2}),\end{aligned}$$

noting that (3.10) implies $|E_\varepsilon| \lesssim_{r,t} \varepsilon^{p/2}$. For $\rho \in F_\varepsilon(r,t,\rho)$ we use the change of variable

$$w \to v_\pm(w) = u \pm \frac{b}{2} G(r,t,\rho) - (u-w)^2 + \frac{b^2}{4} G(r,t,\rho)^2$$

which is one-to-one on (u,∞) and on $(-\infty,u)$ and satisfies

$$u - v_\pm(w) \pm \frac{b}{2} G(r,t,\rho) = (u-w)^2 - \frac{b^2}{4} G(r,t,\rho)^2.$$

We have $|v'(w)| = 2|u-w|$, and $|v(w)-w| = O(\varepsilon)$ on the support of the integrand, and therefore also $|u-w| = G(r,t,\rho)|b|/2 + O(\varepsilon)$. Hence, by Taylor expansion of $f(v,\rho)$ around (w,ρ),

$$\begin{aligned}&\int_{F_\varepsilon(r,t)} \int_{\mathbb{R}} \rho f_0(v,\rho) \chi_\varepsilon(u \pm \tfrac{b}{2} G(r,t,\rho) - v)\, \mathrm{d}v\, \frac{1}{G(r,t,\rho)}\, \mathrm{d}\rho \\ &\quad = \frac{1}{2} \int_{F_\varepsilon(r,t)} \int_{\mathbb{R}} \rho f_0(v(w),\rho) \chi_\varepsilon((u-w)^2 - (\tfrac{b}{2} G(r,t,\rho))^2) \frac{2|u-w|}{G(r,t,\rho)}\, \mathrm{d}w\, \mathrm{d}\rho \\ &\quad = \frac{|b|}{2} \int_{F_\varepsilon(r,t)} \int_{\mathbb{R}} \rho f_0(w,\rho) \chi_\varepsilon((u-w)^2 - (\tfrac{b}{2} G(r,t,\rho))^2)\, \mathrm{d}w\, \mathrm{d}\rho + O(\varepsilon^{1/2})\end{aligned}$$

and by using the estimate $|E_\varepsilon(r,t)| \lesssim_{r,t} \varepsilon^{p/2}$ the last displayed expression is equal to

$$\frac{|b|}{2}\int_{|r-t|}^{r+t}\int_{\mathbb{R}} \rho f_0(w,\rho)\chi_\varepsilon((u-w)^2-(\tfrac{b}{2}G(r,t,\rho))^2)\,\mathrm{d}w\,\mathrm{d}\rho + O(\varepsilon^{1/2}),$$

for both choices of $\pm$. We sum in $\pm$ and, using (3.7), obtain, for $r\neq t$,

$$(f*\mu_t)_0(u,r) = \frac{|b|}{\pi}\int_{|r-t|}^{r+t}\int_{\mathbb{R}} \rho f_0(w,\rho)\chi_\varepsilon(\Phi_t(u,r;w,\rho))\mathrm{d}w\,\mathrm{d}\rho + O(\varepsilon^{(p-1)/2}).$$

Letting $\varepsilon\to 0$ concludes the proof. $\square$

3.4. **A local variant of the maximal operator.** The main work in proving Theorem 3.1 will be to establish the following local variant.

Theorem 3.3. *For all $p>2$,*

$$\big\|\sup_{1\le t\le 2}|A_t f|\big\|_{L^p(\mathbb{R}\times(0,\infty))} \lesssim \|f\|_{L^p(\mathbb{R}\times(0,\infty))}.$$

This will be established in §4 – §8. The passage from Theorem 3.3 to the global result in Theorem 3.1 is postponed until §9.

4. Curvature considerations

As indicated in the introduction and Section 2, various ‘curvatures’, which feature extensively in the analysis of generalised Radon transforms, are fundamental to the proof of Theorem 3.3. In this section these concepts are reviewed and some calculations are carried out in relation to the operator A_t introduced above.

Definition 4.1. A *smooth family of defining pairs* $[\Phi;\mathfrak{a}]$ consists of a pair of functions $\mathfrak{a}\in C^\infty(\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}^2)$ and $\Phi\in C^\infty$ defined on a neighbourhood of supp $\mathfrak{a}$ satisfying

$$\nabla_{(x,z)}\Phi(x,t;z)\neq 0 \qquad \text{for } (x,t;z)\in \operatorname{supp}\mathfrak{a}.$$

The t variable will play a preferred rôle in the forthcoming analysis. For any fixed $t\in\mathbb{R}$ let $\Phi_t(x;z):=\Phi(x,t;z)$ and $\mathfrak{a}_t(x;z):=\mathfrak{a}(x,t;z)$; then $[\Phi_t;\mathfrak{a}_t]$ is referred to as a *defining pair* . The Schwartz kernel $\mathfrak{a}\,\delta\circ\Phi$ is then well defined, and the corresponding integral operator $A[\Phi_t;\mathfrak{a}_t]f(x)$ mapping test functions to distributions is given by the pairing

$$\big\langle A[\Phi_t;\mathfrak{a}_t]f,g\big\rangle := \iint_{\mathbb{R}^2\times\mathbb{R}^2} g(x)f(z)\mathfrak{a}_t(x;z)\delta\big(\Phi_t(x;z)\big)\,\mathrm{d}z\,\mathrm{d}x. \tag{4.1}$$

Key example. For the defining function Φ_t in (3.2), where $t\sim 1$, with the identification of coordinates $(u,r)=(x_1,x_2)$, $(v,\rho)=(z_1,z_2)$, the distribution $\delta\circ\Phi$ is defined when paired with $g(u,r)f(v,\rho)$ where g and f are compactly supported C^∞ functions with support away from $\{r=0\}$ and $\{\rho=0\}$ respectively. The calculations in Lemma 3.2 show that in this case $A[\Phi_t;\mathfrak{a}_t]f(x)$ is pointwise defined for $x_2\neq 0$, as long as $f\in C_0^\infty(\mathbb{R}^2)$ with supp $f\subset\{y\in\mathbb{R}^2: y_2\neq 0\}$.

4.1. **Rotational curvature.** Given a defining pair $[\Phi_t; \mathfrak{a}_t]$ the *rotational curvature* $\mathrm{Rot}(\Phi_t)$ is defined to be the function of $(x; z) \in \mathbb{R}^2 \times \mathbb{R}^2$ given by the determinant of the Monge–Ampère matrix

$$\mathfrak{M}(\Phi_t) := \begin{bmatrix} \Phi_t & (\partial_z \Phi_t)^\top \\ \partial_x \Phi_t & \partial^2_{xz}\Phi_t \end{bmatrix}.$$

Note that $\mathfrak{M}(\Phi_t)$ is the mixed Hessian $D^2_{(\theta,x),(s,z)}\Psi_t|_{\theta=s=1}$ of the function

$$(\theta, x, s, z) \mapsto \Psi_t(\theta, x; s, z) := \theta s \Phi_t(x; z)$$

and, more generally,

$$D^2_{(\theta,x),(s,z)}\Psi_t = \begin{bmatrix} \Phi_t & s\partial_z \Phi_t^\top \\ \theta\partial_x \Phi_t & \theta s \partial_{xz}\Phi_t. \end{bmatrix}.$$

It is well-known (see, for instance, [32, Chapter XI, §3]) that the behaviour of $\mathrm{Rot}(\Phi_t)$ on the incidence relation $\{\Phi = 0\}$ plays an important rôle in determining the mapping properties of averaging operators $A[\Phi_t; \mathfrak{a}_t]$ on L^2-Sobolev spaces as well as the L^p theory of their maximal variants. It is of particular interest to identify points where the rotational curvature vanishes together with the defining function.

Key example. For the defining function Φ_t in question, as introduced in (3.2), we now have $(x_1, x_2) \equiv (u, r)$ and $(z_1, z_2) = (v, \rho)$ and

$$\begin{aligned}\mathfrak{M}(\Phi_t) &= \begin{bmatrix} \Phi_t & \partial_v \Phi_t & \partial_\rho \Phi_t \\ \partial_u \Phi_t & \partial^2_{uv}\Phi_t & \partial^2_{u\rho}\Phi_t \\ \partial_r \Phi_t & \partial^2_{rv}\Phi_t & \partial^2_{r\rho}\Phi_t \end{bmatrix} \\ &= \begin{bmatrix} \Phi_t & -2(u-v) & -b^2\rho(r^2-\rho^2+t^2) \\ 2(u-v) & -2 & 0 \\ -b^2 r(\rho^2 - r^2 + t^2) & 0 & -2b^2 r\rho \end{bmatrix}.\end{aligned}$$

Then, one computes that

$$\begin{aligned}\det \mathfrak{M}(\Phi_t) &= 2b^2 r\rho\big(b^2(\rho^2 - r^2 + t^2)(r^2 - \rho^2 + t^2) - 4(u-v)^2 + 2\Phi\big) \\ &= 2b^4 r\rho\big((\rho^2 - r^2 + t^2)(r^2 - \rho^2 + t^2) - 4r^2\rho^2 + (r^2 + \rho^2 - t^2)^2\big) + 4b^2 r\rho\Phi_t.\end{aligned}$$

Setting $\Phi_t = 0$ one obtains after further computation

$$\mathrm{Rot}(\Phi_t)(u, r; v, \rho) = 4b^4 r t^2 \rho(t^2 - r^2 - \rho^2) \qquad \text{for } (v, \rho) \in \Sigma_{u,r,t}. \tag{4.2}$$

Thus, $\mathrm{Rot}(\Phi_t)$ vanishes along the co-ordinate hyperplanes $r = 0$, $t = 0$ and $\rho = 0$ and also, more significantly, along the hypersurface $t^2 = r^2 + \rho^2$.

Continuing with Φ_t as in (3.2), the rotational curvature and t-derivative of the defining function are related via the identity

$$\mathrm{Rot}(\Phi_t)(u, r; v, \rho) = 4b^2 r t\rho(\partial_t \Phi_t)(u, r; v, \rho). \tag{4.3}$$

A relationship of this kind was previously noted in [21] in the context of the spherical maximal operator on $\mathbb{H}^n$ for $n \geq 2$. Here, in close analogy with [21], the identity (4.3) will be important in the analysis near the singular hypersurface $t^2 = r^2 + \rho^2$.

Rather than freezing t for the computation of the rotational curvature, it is sometimes useful to freeze r and set

$$\Phi^\star_r(u, t; v, \rho) := \Phi_t(u, r; v, \rho).$$

A similar computation to the one above yields in this case

$$\mathrm{Rot}(\Phi_r^\star)(u,t;v,\rho) = 4b^4r^2t\rho(r^2-t^2-\rho^2) \qquad \text{for } (v,\rho)\in\Sigma_{u,r,t}. \tag{4.4}$$

4.2. **The fold conditions.** For the defining function from (3.2), the vanishing of the rotational curvature along the hypersurface $t^2 = r^2+\rho^2$ corresponds to a *two-sided fold singularity.*

Definition 4.2. A defining function Φ_{t_0} satisfies the *two-sided fold condition* at $(x_0;z_0)\in\mathbb{R}^2\times\mathbb{R}^2$ if the following hold:

i) $\Phi_{t_0}(x_0;z_0)=0$ and $\mathrm{Rank}\,\mathfrak{M}(\Phi_{t_0})(x_0;z_0)=2$.
ii) If $U=(u_1,u_2,u_3)$ and $V=(v_1,v_2,v_3)\in\mathbb{R}^3$ span the cokernel and kernel of $\mathfrak{M}(\Phi_{t_0})(x_0;z_0)$, respectively, then

$$\Big\langle \partial^2_{zz}\Big\langle U, \begin{bmatrix}\Phi_{t_0}\\ \partial_x\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}V'', V''\Big\rangle \neq 0,$$
$$\Big\langle \partial^2_{xx}\Big\langle V, \begin{bmatrix}\Phi_{t_0}\\ \partial_z\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}U'', U''\Big\rangle \neq 0,$$

where $U''=(u_2,u_3)$ and $V''=(v_2,v_3)$.

As a consequence of the fold condition, $\mathfrak{M}(\Phi_{t_0})(x_0;z_0)$ may be transformed into a 'normal form'. In particular, there exist $\mathbf{X},\mathbf{Z}\in\mathrm{GL}(3,\mathbb{R})$ satisfying

- $\mathbf{X}e_3=U$ and $\mathbf{X}e_1$, $\mathbf{X}e_2$ are orthogonal to
$$\Big(0, \partial^2_{xx}\Big\langle V, \begin{bmatrix}\Phi_{t_0}\\ \partial_z\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}U''\Big),$$
- $\mathbf{Z}e_3=V$ and $\mathbf{Z}e_1$, $\mathbf{Z}e_2$ are orthogonal to
$$\Big(0, \partial^2_{zz}\Big\langle U, \begin{bmatrix}\Phi_{t_0}\\ \partial_x\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}V''\Big),$$

where e_j denote the standard basis vectors in $\mathbb{R}^3$, and therefore

$$\mathbf{X}^\top\circ\mathfrak{M}(\Phi_{t_0})(x_0;z_0)\circ\mathbf{Z} = \begin{bmatrix}\mathbf{M}(x_0,t_0;z_0) & 0\\ 0 & 0\end{bmatrix}$$

for $\mathbf{M}(x_0,t_0;z_0)$ a non-singular 2×2 matrix.

Key example. For the defining function Φ_t from (3.2), if Φ_{t_0} and $\mathrm{Rot}(\Phi_{t_0})$ both vanish at $(x_0;z_0)=(u_0,r_0;v_0,\rho_0)$ and $r_0t_0\rho_0\neq 0$, then

$$U := \begin{bmatrix}1\\ -(u_0-v_0)\\ -r_0\end{bmatrix} \qquad\text{and}\qquad V := \begin{bmatrix}1\\ u_0-v_0\\ -\rho_0\end{bmatrix} \tag{4.5}$$

span the cokernel and kernel of $\mathfrak{M}(\Phi_{t_0})(x_0;z_0)$, respectively. Moreover,

$$\Big\langle \partial^2_{zz}\Big\langle U, \begin{bmatrix}\Phi_{t_0}\\ \partial_x\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}V'', V''\Big\rangle = 2b^2\rho_0^2(r_0^2+\rho_0^2)>0,$$
$$\Big\langle \partial^2_{xx}\Big\langle V, \begin{bmatrix}\Phi_{t_0}\\ \partial_z\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}U'', U''\Big\rangle = 2b^2r_0^2(r_0^2+\rho_0^2)>0$$

and the matrices $\mathbf{X}$ and $\mathbf{Z}$ can be taken to be

$$\mathbf{X} := \begin{bmatrix}1 & 0 & 1\\ 0 & -b^2r_0^3 & -(u_0-v_0)\\ 0 & u_0-v_0 & -r_0\end{bmatrix}, \qquad \mathbf{Z} := \begin{bmatrix}1 & 0 & 1\\ 0 & b^2\rho_0^3 & u_0-v_0\\ 0 & u_0-v_0 & -\rho_0\end{bmatrix}. \tag{4.6}$$

Remark. For standard incidence relations $\mathcal{M} \subset \mathbb{R}^2_L \times \mathbb{R}^2_R$, where $\mathbb{R}^2_L \equiv \mathbb{R}^2_R \equiv \mathbb{R}^2$ and $\mathcal{M} = \{\Phi = 0\}$ with $\nabla\Phi$ bounded below, the two sided fold condition is equivalent to the more common assumption ([18], [25]) that the projections π_L, π_R mapping the conormal bundle $N^*\mathcal{M}$ to $T^*\mathbb{R}^2_L$, $T^*\mathbb{R}^2_R$ have fold singularities.

4.3. **Individual curves.** It is also useful to consider the curvatures of the individual curves in the curve distribution induced by a defining family Φ. In particular, for fixed (x,t) the non-vanishing of the curvature of $\Sigma_{x,t} := \{z \in \mathbb{R}^2 : \Phi_t(x;z) = 0\}$ is equivalent to the non-vanishing (on $\Sigma_{x,t}$) of

$$\kappa(\Phi_t)(x;z) := \det \begin{bmatrix} 0 & (\partial_z \Phi_t)^\top(x;z) \\ \partial_z \Phi_t(x;z) & \partial^2_{zz}\Phi_t(x;z) \end{bmatrix}. \tag{4.7}$$

Key example. For the defining family Φ as introduced in (3.2), the curves have non-vanishing curvature whenever $r \neq t$. To see this, note that

$$\begin{aligned}\kappa(\Phi_t) &= \det \begin{bmatrix} \Phi_t & \partial_v \Phi_t & \partial_\rho \Phi_t \\ \partial_v \Phi_t & \partial^2_{vv}\Phi_t & \partial^2_{v\rho}\Phi_t \\ \partial_\rho \Phi_t & \partial^2_{\rho v}\Phi_t & \partial^2_{\rho\rho}\Phi_t \end{bmatrix} \\ &= \det \begin{bmatrix} \Phi_t & -2(u-v) & -b^2\rho(t^2+r^2-\rho^2) \\ -2(u-v) & 2 & 0 \\ -b^2\rho(t^2+r^2-\rho^2) & 0 & -b^2(t^2+r^2-3\rho^2) \end{bmatrix},\end{aligned}$$

which after a computation reduces, for $(v,\rho) \in \Sigma_{u,r,t}$, i.e. $\Phi_t = 0$, to

$$\kappa(\Phi_t)(u,r,t;v,\rho) = b^4\big(\rho^6 - 3(r^2+t^2)\rho^4 + 3(r^2-t^2)^2\rho^2 - (r^2-t^2)^2(r^2+t^2)\big).$$

Thus, $\kappa(\Phi_t)(u,r,t;v,\rho) = \wp_{r,t}(\rho^2)$, where $\wp_{r,t}$ is a cubic polynomial with coefficients depending on r,t. We first calculate $\wp((r-t)^2) = -8b^4r^2t^2(r-t)^2$. One may verify that $\wp_{r,t}$ is a decreasing function on the interval $[(r-t)^2, (r+t)^2]$, leading to the lower bound

$$|\kappa(\Phi)(u,r,t;v,\rho)| \geq 8b^4r^2t^2(r-t)^2 \qquad \text{for all } (v,\rho) \in \Sigma_{u,r,t} \tag{4.8}$$

Thus, the curves have non-vanishing curvature if $r \neq t$, as claimed.

4.4. **Cinematic curvature.** It is also necessary to analyse the average operator from the perspective of the *cinematic curvature condition* of [30].

Definition 4.3. A smooth family of defining pairs $[\Phi;\mathfrak{a}]$ is said to satisfy the *projection condition* if

$$\operatorname{Proj}(\Phi) := \det \begin{bmatrix} \partial_{\vec{x}}\Phi & \partial^2_{\vec{x}z}\Phi \end{bmatrix}$$

is non-vanishing on an open neighbourhood U of $\operatorname{supp}\mathfrak{a}$. Here $\vec{x} = (x,t) \in \mathbb{R}^2 \times \mathbb{R}$.

Fixing $\vec{x} \in \mathbb{R}^2 \times \mathbb{R}$, the projection condition implies that the map

$$(U \cap \Sigma_{\vec{x}}) \times \mathbb{R} \to \mathbb{R}^3; \quad (z;\theta) \mapsto \theta\partial_{\vec{x}}\Phi(\vec{x};z)$$

is a diffeomorphism and therefore its image $\Gamma_{\vec{x}}$ is an immersed submanifold of $\mathbb{R}^3$. If $\zeta := \theta\partial_{\vec{x}}\Phi(\vec{x};z) \in \Gamma_{\vec{x}}$, then a basis for $T_\zeta\Gamma_{\vec{x}}$ is given by the vector fields

$$\mathbf{T}^1 := \partial_{\vec{x}}\Phi, \qquad \mathbf{T}^2 := (\mathbf{T}^2_1, \mathbf{T}^2_2, \mathbf{T}^2_3) \quad \text{where } \mathbf{T}^2_j := \det\begin{bmatrix}\partial_z\Phi & \partial_z\partial_{x_j}\Phi\end{bmatrix} \tag{4.9}$$

evaluated at $(\vec{x};z)$; this may be seen computing the tangent vectors of the parametrisation $\sigma_{\vec{x}}$ below. Note that $\Gamma_{\vec{x}}$ is clearly a cone and therefore has everywhere vanishing Gaussian curvature. If at every point on $\Gamma_{\vec{x}}$ there is a non-zero principal curvature, then $[\Phi;\mathfrak{a}]$ is said to satisfy the *cinematic curvature condition* (see [30] or [20] for further details).

Definition 4.4. For any defining family Φ let

$$\mathrm{Cin}(\Phi) := \det \begin{bmatrix} \mathbf{S} & \mathbf{T}^1 & \mathbf{T}^2 \end{bmatrix}$$

where $\mathbf{S} = \mathbf{S}^1 - \mathbf{S}^2$ where $\mathbf{S}^i = (\mathbf{S}^i_1, \mathbf{S}^i_2, \mathbf{S}^i_3)$ for

$$\mathbf{S}^1_j := \det \begin{bmatrix} 0 & (\partial_z \Phi)^\top \\ \partial_z \Phi & \partial^2_{zz}\partial_{x_j}\Phi \end{bmatrix}, \quad \mathbf{S}^2_j := \det \begin{bmatrix} 0 & (\partial_z \partial_{x_j}\Phi)^\top \\ \partial_z \Phi & \partial^2_{zz}\Phi \end{bmatrix}.$$

If $[\Phi; \mathfrak{a}]$ satisfies the projection condition, then the cinematic curvature condition is equivalent to the non-vanishing of $\mathrm{Cin}(\Phi)(\vec{x}; z)$ whenever $z \in \Sigma_{\vec{x}}$. Indeed, fix $\vec{x}$ and let $\gamma_{\vec{x}} \colon [0,1] \to \Sigma_{\vec{x}}$ denote a unit speed parametrisation of $\Sigma_{\vec{x}}$; this induces a parametrisation $\sigma_{\vec{x}} \colon (\theta, s) \mapsto \theta \partial_{\vec{x}} \Phi(\vec{x}; \gamma_{\vec{x}}(s))$ of the cone $\Gamma_{\vec{x}}$. The cinematic curvature condition is then equivalent to the non-vanishing of

$$\det \begin{bmatrix} \partial_{ss}\sigma_{\vec{x}}(\theta, s) & \partial_\theta \sigma_{\vec{x}}(\theta, s) & \partial_s \sigma_{\vec{x}}(\theta, s) \end{bmatrix} \tag{4.10}$$

and a computation shows that (4.10) is equal to $-|\theta|^2 |\partial_z \Phi|^{-3} \mathrm{Cin}(\Phi)$.

Key example. For the defining family Φ as introduced in (3.2) one has

$$\mathrm{Proj}(\Phi)(u, r, t; v, \rho) = -8b^4 r t \rho (r^2 - t^2), \tag{4.11}$$

$$\mathrm{Cin}(\Phi)(u, r, t; v, \rho) = 64 b^8 r^3 t^3 \rho^3 (r^2 - t^2). \tag{4.12}$$

Thus, $[\Phi; \mathfrak{a}]$ satisfies the cinematic curvature condition whenever $\operatorname{supp} \mathfrak{a}$ avoids the hyperplanes $r = 0$, $t = 0$ and $r = t$.[2] For reference, Appendix B.1 contains the formulæ for the various derivatives featured in these computations.

5. The initial decomposition

For Φ as defined in (3.2) both the rotational and cinematic curvature conditions fail. In this section, the operator A_t is decomposed in order to isolate the singularities corresponding to the failure of these curvature conditions.

5.1. **Spatial decomposition.** The operator A_t is first decomposed dyadically with respect to the r variable. To this end, fix a nonnegative $\eta \in C^\infty_c(\mathbb{R})$ such that

$$\eta(r) = 1 \quad \text{if } r \in [-1, 1] \quad \text{and} \quad \operatorname{supp} \eta \subseteq [-2, 2] \tag{5.1a}$$

and define $\beta \in C^\infty_c(\mathbb{R})$ and $\eta^m, \beta^m \in C^\infty_c(\mathbb{R})$ by

$$\beta(r) := \mathbb{1}_{(0,\infty)}(r)(\eta(r) - \eta(2r)) \tag{5.1b}$$

and, for each $m \in \mathbb{Z}$,

$$\eta^m(r) := \eta(2^{-m} r) \quad \text{and} \quad \beta^m(r) := \beta(2^{-m} r). \tag{5.2}$$

One may then decompose

$$A_t f(u, r) = \sum_{m \in \mathbb{Z}} \beta^m(r) A_t f(u, r) \qquad \text{for } (u, r) \in \mathbb{R} \times (0, \infty).$$

The r-localisation induces various spatial orthogonality relations via (3.6). In particular, if $r \in \operatorname{supp} \beta^m$, then $r \sim 2^m$ and it follows from (3.6) that

$$|u - v| \lesssim 2^m, \quad |r - \rho| \lesssim 1 \quad \text{and} \quad |t - \rho| \lesssim 2^m \quad \text{for } (v, \rho) \in \Sigma_{u,r,t}. \tag{5.3}$$

[2] In this case, one may further deduce that $\Gamma_{u,r,t}$ is the cone defined implicitly by the equation

$$\zeta_1^2 - \frac{\zeta_2^2}{b^2(t^2 - r^2)} + \frac{\zeta_3^2}{b^2(t^2 - r^2)} = 0.$$

To exploit this, given $m, \sigma \in \mathbb{Z}$ define

$$\eta^{m,\sigma}(u,v) := \eta(2^{-m}u - \sigma)\eta(C^{-1}(2^{-m}v - \sigma)),$$

where $C \geq 1$ is an absolute constant which is chosen to be sufficiently large for the purposes of the forthcoming arguments. We define

$$\begin{aligned} \mathfrak{a}^0(u,r,t;v,\rho) &:= \beta(r)r^{1/p}\rho^{1-1/p}, \\ \mathfrak{a}^{m,\vec{\sigma}}(u,r,t;v,\rho) &:= \beta^m(r)\eta^{m,\sigma_1}(u,v)\eta^{0,\sigma_2}(r,\rho), \quad \text{if } m > 0, \\ \mathfrak{a}^{m,\vec{\sigma}}(u,r,t;v,\rho) &:= \beta^m(r)\eta^{m,\sigma_1}(u,v)\eta^{m,\sigma_2}(t,\rho), \quad \text{if } m < 0, \end{aligned} \tag{5.4}$$

so that for $m > 0$, $\mathfrak{a}^{m,\vec{\sigma}}$ is supported where

$$r \sim 2^m, \quad |u - 2^m\sigma_1| \lesssim 2^m, \quad |v - 2^m\sigma_1| \lesssim 2^m, \quad |r - \sigma_2| \lesssim 1, \quad |\rho - \sigma_2| \lesssim 1.$$

Moreover for $m < 0$, $\mathfrak{a}^{m,\vec{\sigma}}$ is supported where

$$r \sim 2^m, \quad |u - 2^m\sigma_1| \lesssim 2^m, \quad |v - 2^m\sigma_1| \lesssim 2^m, \quad |t - \sigma_2| \lesssim 2^m, \quad |\rho - \sigma_2| \lesssim 2^m.$$

In view of (5.3), one may bound (using the notation in (4.1))

$$A_t f \lesssim A[\Phi_t; \mathfrak{a}_t^0]f + \sum_{\vec{\sigma}\in\mathbb{Z}^2}\sum_{m>0} 2^m A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f + \sum_{\vec{\sigma}\in\mathbb{Z}^2}\sum_{m<0} 2^{m/p} A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f, \tag{5.5}$$

whenever f is a (say) continuous, non-negative function.

The unit scale piece $\mathfrak{a}_t^0$ is supported where $r \sim 1$ and it is now further dyadically decomposed with respect to both the ρ variable and $|r - t|$. The rationale behind this decomposition is to quantify the value of $\mathrm{Rot}(\Phi_t)$: in view of (4.2), the function $\mathrm{Rot}(\Phi_t)$ can vanish on supp $\mathfrak{a}_t^0$. If $r \sim 1$ and $\rho \sim 2^{-k}$, then it follows from (3.6) that $|u - v| \lesssim 2^{-k}$ for $(v,\rho) \in \Sigma_{u,r,t}$. Thus, given a function $k \mapsto \ell(k)$ on $\mathbb{Z}$ to be defined momentarily we set

$$\begin{aligned} \mathfrak{a}^{k,\ell,\vec{\sigma}}(u,r,t;v,\rho) &:= \beta(r)\beta^{-k}(\rho)\eta^{-k,\sigma_1}(u,v)\beta^{-\ell}(|r-t|)\eta^{-\ell,\sigma_2}(r,t), \qquad \ell < \ell(k), \\ \mathfrak{c}^{k,\vec{\sigma}}(u,r,t;v,\rho) &:= \beta(r)\beta^{-k}(\rho)\eta^{-k,\sigma_1}(u,v)\eta^{-\ell(k)}(|r-t|)\eta^{-\ell(k),\sigma_2}(r,t), \end{aligned}$$

so that on the support of $\mathfrak{a}^{k,\ell,\vec{\sigma}}$, $\mathfrak{c}^{k,\vec{\sigma}}$ we have

$$r \sim 1, \quad \rho \sim 2^{-k}, \quad |u - 2^{-k}\sigma_1| \lesssim 2^{-k} \quad \text{and} \quad |v - 2^{-k}\sigma_1| \lesssim 2^{-k};$$

moreover

$$|r - t| \sim 2^{-\ell}, \quad |r - 2^{-\ell}\sigma_2| \lesssim 2^{-\ell}, \quad |t - 2^{-\ell}\sigma_2| \lesssim 2^{-\ell}$$

on supp $\mathfrak{a}^{k,\ell,\vec{\sigma}}$, and

$$|r - t| \lesssim 2^{-\ell(k)}, \quad |r - 2^{-\ell(k)}\sigma_2| \lesssim 2^{-\ell(k)}, \quad |t - 2^{-\ell(k)}\sigma_2| \lesssim 2^{-\ell(k)}$$

on supp $\mathfrak{c}^{k,\vec{\sigma}}$. One may bound

$$A[\Phi_t; \mathfrak{a}_t^0]f \lesssim \sum_{\vec{\sigma}\in\mathbb{Z}^2}\sum_{\substack{(k,\ell)\in\mathbb{Z}^2 \\ \ell<\ell(k)}} 2^{-k(1-1/p)} A[\Phi_t; \mathfrak{a}_t^{k,\ell,\vec{\sigma}}]f \tag{5.6}$$

$$+ \sum_{\vec{\sigma}\in\mathbb{Z}^2}\sum_{k\in\mathbb{Z}} 2^{-k(1-1/p)} A[\Phi_t; \mathfrak{c}_t^{k,\vec{\sigma}}]f. \tag{5.7}$$

For the purposes of our proof, we let

$$\ell(k) := 2k + C_{\mathrm{rot}}$$

for some (absolute) constant $C_{\mathrm{rot}} \geq 1$, suitably chosen so as to satisfy the forthcoming requirements. Furthermore, by the first inequality in (3.6), one may in fact restrict the range of the k summation in the above expression to $k \geq -4$ and of the (k, ℓ) summation to the parameter set

$$\mathfrak{P} := \big\{(k, \ell) \in \mathbb{Z} \times \mathbb{Z} : k \geq -4 \text{ and } k - 3 \leq \ell < \ell(k)\big\}.$$

We show presently that the following bounds imply Theorem 3.3.

Theorem 5.1. *For all $2 < p < \infty$ there exists some $\varepsilon_p > 0$ such that*

i) $\big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{k,\ell,\vec{\sigma}}]f|\big\|_p \lesssim 2^{-\ell/p - k\varepsilon_p} 2^{k(1-1/p)} \|f\|_p$ *for $(k, \ell) \in \mathfrak{P}$*

ii) $\big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{c}_t^{k,\vec{\sigma}}]f|\big\|_p \lesssim 2^{-\ell(k)/p - k\varepsilon_p} 2^{(1-1/p)k} \|f\|_p$ *for all $k \geq -4$*

iii) $\big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\big\|_p \lesssim 2^{-m} \|f\|_p$ *for $m > 0$*

iv) $\big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\big\|_p \lesssim 2^{m\varepsilon_p} 2^{-m/p} \|f\|_p$ *for $m < 0$*

uniformly in $\vec{\sigma} \in \mathbb{Z}^2$. The above a priori estimates hold for all $f \in C_0^\infty(\mathbb{R}^2)$ with support in $\{y \in \mathbb{R}^2 : y_2 \neq 0\}$.

Proof of Theorem 3.3 assuming Theorem 5.1 holds. Consider the second and third terms on the right-hand side of (5.5).

When $m > 0$ there is spatial orthogonality among the pieces of the decomposition in both $\vec{\sigma}$ and m. This observation combined with Theorem 5.1 iii) above yields

$$\begin{aligned}&\Big\| \sum_{\vec{\sigma} \in \mathbb{Z}^2} \sum_{m > 0} 2^m \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\Big\|_p \\ &\qquad \lesssim \Big(\sum_{\vec{\sigma} \in \mathbb{Z}^2} \sum_{m > 0} 2^{mp} \big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\big\|_p^p \Big)^{1/p} \lesssim \|f\|_p,\end{aligned}$$

as desired.

When $m < 0$, note that by the support properties of $\mathfrak{a}_t^{m,\vec{\sigma}}$,

$$\begin{aligned}\sup_{1 \leq t \leq 2} \sum_{\sigma_2 \in \mathbb{Z}} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f| &\lesssim \sup_{\sigma_2 \in \mathbb{Z}} \sup_{\substack{1 \leq t \leq 2 \\ |t - 2^m \sigma_2| \lesssim 2^m}} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f| \\ &\leq \Big(\sum_{\sigma_2 \in \mathbb{Z}} \sup_{\substack{1 \leq t \leq 2 \\ |t - 2^m \sigma_2| \lesssim 2^m}} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|^p \Big)^{1/p}.\end{aligned}$$

Furthermore, applying spatial orthogonality in the $\vec{\sigma}$ parameter, the triangle inequality to the sum in m and Theorem 5.1 iv), one deduces that

$$\begin{aligned}&\Big\| \sum_{m < 0} 2^{m/p} \sup_{1 \leq t \leq 2} \sum_{\vec{\sigma} \in \mathbb{Z}^2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\Big\|_p \\ &\qquad \lesssim \sum_{m < 0} 2^{m/p} \Big(\sum_{\vec{\sigma} \in \mathbb{Z}^2} \big\| \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{m,\vec{\sigma}}]f|\big\|_p^p \Big)^{1/p} \lesssim_p \|f\|_p,\end{aligned}$$

where the last step uses the exponential decay $2^{m\varepsilon_p}$ to sum in m.

Next, consider the sums in (5.6). Again, there is spatial orthogonality in the σ_1 parameter. This fact and Theorem 5.1 i) yield

$$\Big\| \sum_{\sigma_1 \in \mathbb{Z}} 2^{-k(1-1/p)} \sup_{1 \leq t \leq 2} |A[\Phi_t; \mathfrak{a}_t^{k,\ell,\vec{\sigma}}]f|\Big\|_p \lesssim 2^{-\ell/p} 2^{-k\varepsilon_p} \|f\|_p$$

uniformly in σ_2. As the parameter σ_2 corresponds to a decomposition of the r spatial variable,

$$\begin{aligned}
&\Big\| \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{-k(1-1/p)} \sup_{1\le t\le 2} |A[\Phi_t;\mathfrak{a}_t^{k,\ell,\vec{\sigma}}]f| \Big\|_p \\
&\qquad \lesssim \Big(\sum_{\substack{\sigma_2\in\mathbb{Z}\\ |\sigma_2|\lesssim 2^\ell}} \Big\| \sum_{\sigma_1\in\mathbb{Z}} 2^{-k(1-1/p)} \sup_{1\le t\le 2} |A[\Phi_t;\mathfrak{a}_t^{k,\ell,\vec{\sigma}}]f| \Big\|_p^p \Big)^{1/p} \\
&\qquad \lesssim 2^{-k\varepsilon_p} \Big(\sum_{\substack{\sigma_2\in\mathbb{Z}\\ |\sigma_2|\lesssim 2^\ell}} 2^{-\ell} \|f\|_p^p \Big)^{1/p} \lesssim 2^{-k\varepsilon_p}\|f\|_p.
\end{aligned}$$

The desired result then follows from the triangle inequality in (k,ℓ), using the exponential decay $2^{-k\varepsilon_p}$ to sum over k and $\ell \le \ell(k)$. The sum in (5.7) is bounded in a similar manner. $\square$

5.2. **Rescaling.** Each piece of the decomposition is appropriately rescaled in order to obtain, wherever possible, favourable bounds on the various curvatures. For the reader's convenience, Appendix B.2 describes the behaviour of the functions Φ, $\mathrm{Rot}(\Phi)$, $\mathrm{Cin}(\Phi)$, etc under general rescalings. These rescalings lead to phase functions satisfying certain nonisotropic conditions which will require extensions of some classical results on oscillatory integral operators (see §6 below).

5.2.1. *The case $m = 0$.* For $(k,\ell)\in\mathfrak{P}$ we define the dilations

$$D^{k,\ell}(u,r,t;v,\rho) := (2^{-k}u, 2^{-\ell}r, 2^{-\ell}t; 2^{-k}v, 2^{-k}\rho).$$

Let

$$e(k,\ell) := \ell - 2k + \ell\wedge 2k = \begin{cases} \ell, & \text{if } \ell \ge 2k, \\ 2\ell - 2k, & \text{if } \ell \le 2k, \end{cases}$$

and define

$$\begin{aligned} \Phi^{k,\ell} &:= 2^{2k+e(k,\ell)/3}\Phi\circ D^{k,\ell}, \quad \tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}} := \mathfrak{a}^{k,\ell,\vec{\sigma}}\circ D^{k,\ell}, \\ \Phi^k &:= \Phi^{k,\ell(k)}, \quad \tilde{\mathfrak{c}}^{k,\vec{\sigma}} := \mathfrak{c}^{k,\vec{\sigma}}\circ D^{k,\ell(k)}. \end{aligned} \tag{5.8}$$

Note that $\tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}$ is supported where

$$\rho\sim 1, \quad r\sim 2^\ell, \quad |r-t|\sim 1, \quad |u-\sigma_1|\lesssim 1, \quad |v-\sigma_1|\lesssim 1, \quad |r-\sigma_2|\lesssim 1 \quad |t-\sigma_2|\lesssim 1.$$

The support of $\tilde{\mathfrak{c}}^{k,\vec{\sigma}}$ has similar properties, with $\ell(k)$ in place of ℓ and $|r-t|\lesssim 1$.

The appearance of the factor $2^{2k+e(k,\ell)/3}$ is motivated by the fact that

$$\mathrm{Rot}(\Phi_t^{k,\ell}) \sim 1 \quad \text{on } \operatorname{supp}\tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}} \quad \text{if } |\ell-2k|\ge C_{\mathrm{rot}}, \tag{5.9}$$

$$\mathrm{Rot}(\Phi_t^{k}) \sim 1 \quad \text{on } \operatorname{supp}\tilde{\mathfrak{c}}^{k,\vec{\sigma}}, \tag{5.10}$$

$$\mathrm{Rot}((\Phi^{k})_r^\star) \sim 1 \quad \text{on } \operatorname{supp}\tilde{\mathfrak{c}}^{k,\vec{\sigma}} \tag{5.11}$$

where $(\Phi^k)_r^\star(u,t;v,\rho) := \Phi_t^k(u,r;v,\rho)$. Note, however, that $\mathrm{Rot}(\Phi_t^{k,\ell})$ may vanish on $\operatorname{supp}\tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}$ if $|\ell-2k|\le C_{\mathrm{rot}}$.

Setting $f_k(v,\rho) = f(2^{-k}v, 2^{-k}\rho)$, and using that δ is homogeneous of degree -1 one has

$$\begin{aligned} A[\Phi_{2^{-\ell}t};\mathfrak{a}_{2^{-\ell}t}^{k,\ell,\vec{\sigma}}]f(2^{-k}u,2^{-\ell}r) &= 2^{e(k,\ell)/3}A[\Phi_t^{k,\ell},\tilde{\mathfrak{a}}_t^{k,\ell,\vec{\sigma}}]f_k(u,r), \\ A[\Phi_{2^{-\ell(k)}t};\mathfrak{c}_{2^{-\ell(k)}t}^{k,\vec{\sigma}}]f(2^{-k}u,2^{-\ell}r) &= 2^{\ell(k)/3}A[\Phi_t^{k,\ell},\tilde{\mathfrak{c}}_t^{k,\vec{\sigma}}]f_k(u,r). \end{aligned}$$

Thus, by rescaling to prove Theorem 5.1 i) and ii) it suffices to show that

$$\big\| \sup_t |A[\Phi_t^{k,\ell}; \tilde{\mathfrak{a}}_t^{k,\ell,\vec{\sigma}}]| \big\|_{L^p \to L^p} \lesssim 2^{-e(k,\ell)/3 + (1-2/p)k - k\varepsilon_p}, \tag{5.12}$$

$$\big\| \sup_t |A[\Phi_t^{k}; \tilde{\mathfrak{c}}_t^{k,\vec{\sigma}}]| \big\|_{L^p \to L^p} \lesssim 2^{-\ell(k)/3 + (1-2/p)k - k\varepsilon_p}; \tag{5.13}$$

where (by a slight abuse of notation) we indicate the operator norms of the maximal operators on the left-hand side. We note that in view of the support properties of $\tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}$, $\tilde{\mathfrak{c}}^{k,\sigma}$ the global supremum in the definition of the maximal operator reduces to a supremum over an interval I of length $|I| \sim 1$ centered at σ_2.

It is helpful to isolate the key features of the rescaled averaging operators used to prove the above inequality. As a first step in this direction, note that each $[\Phi_t^{k,\ell}; \tilde{\mathfrak{a}}_t^{k,\ell,\vec{\sigma}}]$ belongs to the class in the following definition. We use coordinates $(x; z)$ for the rescaled phase functions where (x_1, x_2) corresponds to a scaled version of (u, r) and (z_1, z_2) to a scaled version of (v, ρ).

We define collections $\mathfrak{A}^{k,\ell}$ of defining pairs $[\Phi; \mathfrak{a}]$ involving inequalities and support assumptions that are uniform in k, ℓ.

Definition 5.2. Let $\mathfrak{A}^{k,\ell}$ denote the set of all smooth families of defining pairs $[\Phi; \mathfrak{a}]$ for which the following conditions hold:

a)$_{k,\ell}$ $\quad |\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \mathfrak{a}(x,t,z)| \lesssim 1$ and $\operatorname{diam} \operatorname{supp} \mathfrak{a} \lesssim 1$

$\Phi_1)_{k,\ell}$ $\quad |\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \Phi_t(x;z)| \lesssim \begin{cases} 2^{-2e(k,\ell)/3} & \text{if } \alpha_2 \text{ or } \gamma \neq 0 \\ 2^{e(k,\ell)/3} & \text{otherwise} \end{cases}$, $|\partial_z \Phi_t(x;z)| \sim 2^{e(k,\ell)/3}$,

$\Phi_2)_{k,\ell}$ $\quad \partial_t \Phi_t(x;z) = 2^{-2e(k,\ell)/3} c(x,t;z) \mathrm{Rot}(\Phi_t)(x;z)$ for some $c \in C^\infty$ depending on $[\Phi; \mathfrak{a}]$ and with uniform C^∞ bounds on $\operatorname{supp} \mathfrak{a}$.

These estimates are understood to hold on $\operatorname{supp} \mathfrak{a}$, with the constants only depending on the multiindices $\alpha, \beta, \gamma \in \mathbb{N}_0^2$. That is, if we fix a large N then we get uniform estimates for $|\alpha|$, $|\beta|$, $|\gamma| \leq N$.

For $[\Phi^{k,\ell}; \tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}]$ it is easy to see that a)$_{k,\ell}$ and $\Phi_1)_{k,\ell}$ hold via a direct computation (the lower bound in $\Phi_1)_{k,\ell}$ is a little trickier and uses (3.7)). The remaining condition $\Phi_2)_{k,\ell}$ follows from an appropriately rescaled variant of the key identity (4.3). Indeed, note that

$$\partial_t \Phi_t^{k,\ell}(u,r;v,\rho) = 2^{-2e(k,\ell)/3} \frac{1}{4b^2(2^{-\ell}r)(2^{-\ell}t)\rho} \mathrm{Rot}(\Phi_t^{k,\ell})(u,r;v,\rho)$$

where $r \sim t \sim 2^\ell$ and $\rho \sim 1$ on supp $\tilde{a}_t^{k,\ell,\vec{\sigma}}$. Similarly, each $[\Phi^k; \tilde{\mathfrak{c}}^{k,\vec{\sigma}}]$ belongs to $\mathfrak{A}^{k,\ell(k)} =: \mathfrak{C}^k$.

5.2.2. *The cases $m \neq 0$.* For $m \in \mathbb{Z} \setminus \{0\}$, define[3] (recalling $m \wedge 0 = \min\{m, 0\}$)

$$D^m(u,r,t;v,\rho) := (2^m u, r, 2^{m\wedge 0} t; 2^m v, 2^{m \wedge 0}\rho),$$

$$\Phi^m := 2^{-2m} \Phi \circ D^m, \quad \tilde{\mathfrak{a}}^{m,\vec{\sigma}} := \mathfrak{a}^{m,\vec{\sigma}} \circ D^m, \tag{5.14}$$

and let $(\Phi^m)_r^\star(u,t;v,\rho) := \Phi^m(u,r,t;v,\rho)$. It follows from (4.2) and (4.4) that

$$\mathrm{Rot}(\Phi_t^m) \sim 1 \quad \text{if } m > 0 \quad \text{and} \quad \mathrm{Rot}\big((\Phi^m)_r^\star\big) \sim 1 \quad \text{if } m < 0 \quad \text{on supp}\, \tilde{\mathfrak{a}}^{m,\vec{\sigma}}; \tag{5.15}$$

[3]The Φ^m notation in (5.14) conflicts with the Φ^k notation introduced in (5.8). Nevertheless, it shall always be clear from the context which definition is intended.

this observation motivates the choice of normalising factor 2^{-2m}.

Note that for $m > 0$ the new amplitude $\tilde{\mathfrak{a}}^{m,\vec{\sigma}}$ is supported where

$$r \sim 2^m, \quad |u - \sigma_1| \lesssim 1, \quad |v - \sigma_1| \lesssim 1, \quad |r - \sigma_2| \lesssim 1, \quad |\rho - \sigma_2| \lesssim 1,$$

and if $m < 0$ then $\tilde{\mathfrak{a}}^{m,\vec{\sigma}}$ is supported where

$$r \sim 2^m, \quad |u - \sigma_1| \lesssim 1, \quad |v - \sigma_1| \lesssim 1, \quad |t - \sigma_2| \lesssim 1, \quad |\rho - \sigma_2| \lesssim 1.$$

Setting $f^m(v, \rho) = f(2^m v, 2^{m\wedge 0}\rho)$ a computation shows

$$A[\Phi_{2^{m\wedge 0}t}; \mathfrak{a}^{m,\vec{\sigma}}_{2^{m\wedge 0}t}]f(2^m u, r) = 2^{-m}2^{m\wedge 0}A[\Phi^m_t, \tilde{\mathfrak{a}}^{m,\vec{\sigma}}_t]f^m(u, r).$$

Thus by rescaling, to prove Theorem 5.1 iii) and iv) it suffices to show that

$$\big\| \sup_t |A[\Phi^m_t; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}_t]| \big\|_{L^p \to L^p} \lesssim 2^{(m\wedge 0)\varepsilon_p}. \tag{5.16}$$

Note that in view of the support properties of $\tilde{\mathfrak{a}}^{m,\vec{\sigma}}_t$, the global supremum in the definition of the maximal operator reduces to a supremum over an interval I which equals $[1, 2]$ if $m > 0$ and has length $|I| \sim 1$ and it is centered at σ_2 if $m < 0$; in the case $m > 0$ we abuse of notation and assume that $\tilde{\mathfrak{a}}^{m,\vec{\sigma}}_t$ is supported on $t \sim 1$, adding a cut-off function if necessary.

If $m > 0$, then a simple computation shows that $[\Phi^m; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}] \in \mathfrak{A}^{0,0} =: \mathfrak{A}^0$. On the other hand, if $m < 0$, then $[\Phi^m; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}]$ belongs to the following class classes $\mathfrak{A}_m$ in the following definition where the implicit constants are uniform in m.

Definition 5.3. For $m < 0$ let $\mathfrak{A}^m$ denote the set of all smooth families of defining pairs $[\Phi; \mathfrak{a}]$ satisfying:

a)$_m$ $\quad |\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \mathfrak{a}(x, t; z)| \lesssim \begin{cases} 2^{-m\alpha_2} & \text{if } \alpha_2 \neq 0 \\ 1 & \text{otherwise} \end{cases}$, $\quad \operatorname{diam} \operatorname{supp} \mathfrak{a} \lesssim 1$, and the projection of $\operatorname{supp} \mathfrak{a}$ in the x_2-variable lies in an interval of length $\lesssim 2^m$;

$\Phi_1)_m$ $\quad |\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \Phi_t(x; z)| \lesssim \begin{cases} 2^{-2m} & \text{if } \alpha_2 \neq 0 \\ 1 & \text{otherwise} \end{cases}$, $\quad |\partial_z \Phi_t(x; z)| \sim 1$

on $\operatorname{supp} \mathfrak{a}$ for all $\alpha, \beta, \gamma \in \mathbb{N}_0^2$ with $|\alpha|$, $|\beta|$, $|\gamma| \leq N$.

The derivative bounds on the amplitude for $\alpha_2 = 0$, which are uniformly bounded, are used for the L^2-estimates in Section 7. The bounds for $\alpha_2 \neq 0$ are used for the L^p-estimates in Section 8, although they do not introduce any loss for the purposes of the desired inequality (8.1).

5.3. **Cinematic curvature decomposition.** The decomposition described in §5.1 automatically isolates the region where the cinematic curvature vanishes.

5.3.1. *The case* $m = 0$. By (4.8), (4.11) and (4.12), each $[\Phi^{k,\ell}; \tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}]$ belongs to the following class.

Definition 5.4. Let $\mathfrak{A}^{k,\ell}_{\mathrm{Cin}}$ denote the set of all $[\Phi; \mathfrak{a}] \in \mathfrak{A}^{k,\ell}$ satisfying:

C)$_k$ $\quad |\kappa(\Phi)(\vec{x}; z)|$, $|\mathrm{Proj}(\Phi)(\vec{x}; z)|$, $|\mathrm{Cin}(\Phi)(\vec{x}; z)| \gtrsim 2^{-Mk} \quad$ for $(\vec{x}; z) \in \operatorname{supp} \mathfrak{a}$.

Here $M \geq 1$ is an appropriate chosen absolute constant.

Observe, however, that the $[\Phi^k;\tilde{\mathfrak{c}}^{k,\vec{\sigma}}]$ lie in $\mathfrak{A}^{k,\ell(k)}$ but do **not** belong to $\mathfrak{A}^{k,\ell(k)}_{\mathrm{Cin}}$; it is for this reason that this part of the operator is isolated in the analysis. Indeed, the amplitude $\tilde{\mathfrak{c}}^{k,\vec{\sigma}}$ is supported on the region $|r-t|\lesssim 2^{-\ell(k)}$ and therefore $\kappa(\Phi)$, $\mathrm{Proj}(\Phi)$ and $\mathrm{Cin}(\Phi)$ can vanish on $\mathrm{supp}\,\tilde{\mathfrak{c}}^{k,\vec{\sigma}}$. Nevertheless, these quantities only vanish on a small set and, in particular, $[\Phi^k;\tilde{\mathfrak{c}}^{k,\vec{\sigma}}]$ belongs to the following class.

Definition 5.5. Let $\mathfrak{C}^k_{\mathrm{Cin}}$ denote the set of all $[\Phi;\mathfrak{c}]\in\mathfrak{A}^{k,\ell(k)}$ such that, for all $\delta>0$, if $(x,t;z)\in\mathrm{supp}\,\mathfrak{c}$ with $|t-x_2|>\delta$, then

$\mathbf{C}_\delta)_k$ $\quad |\kappa(\Phi)(x,t;z)|,\ \ |\mathrm{Proj}(\Phi)(x,t;z)|,\ \ |\mathrm{Cin}(\Phi)(x,t;z)|\gtrsim \delta 2^{-Mk}.$

As before, $M\geq 1$ is an appropriately chosen absolute constant.

5.3.2. *The cases $m\neq 0$.* If $m>0$, then (4.8), (4.11) and (4.12) show that $[\Phi^m;\tilde{\mathfrak{a}}^{m,\vec{\sigma}}]$ belongs to $\mathfrak{A}^{0,0}_{\mathrm{Cin}}=:\mathfrak{A}^0_{\mathrm{Cin}}$. On the other hand, if $m<0$, then $[\Phi^m;\tilde{\mathfrak{a}}^{m,\vec{\sigma}}]$ belongs to the following class.

Definition 5.6. For $m<0$ let $\mathfrak{A}^m_{\mathrm{Cin}}$ denote the set of $[\Phi;\mathfrak{c}]\in\mathfrak{A}^m$ satisfying $\mathbf{C})_{-m}$.

5.4. **Rotational curvature decomposition.** Further decomposition is required in order to isolate the regions where the rotational curvature vanishes.

5.4.1. *The case $m=0$.* Let $\varepsilon_\circ>0$ be a fixed constant, chosen small enough to satisfy the requirements of the forthcoming proof, and define

$$\mathfrak{b}^{k,\ell,\vec{\sigma}}(u,r,t;v,\rho):=\tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}}(u,r,t;v,\rho)\eta\big(\varepsilon_\circ^{-1}\mathrm{Rot}(\Phi^{k,\ell}_t)(u,r;v,\rho)\big)\big).$$

In view of (5.9), one may readily verify that $\mathfrak{b}^{k,\ell,\vec{\sigma}}$ is identically zero unless $|\ell-2k|\lesssim 1$, in which case $[\Phi^{k,\ell};\mathfrak{b}^{k,\ell,\vec{\sigma}}]\in\mathfrak{A}^{k,2k}_{\mathrm{Cin}}=:\mathfrak{B}^k_{\mathrm{Cin}}$.

Vanishing rotational curvature. To analyse the operators $A[\Phi^{k,\ell}_t;\mathfrak{b}^{k,\ell,\vec{\sigma}}_t]$ it is necessary to exploit the fold conditions discussed in §4.1. The observations of §4.1 imply that $[\Phi^{k,\ell};\mathfrak{b}^{k,\ell,\vec{\sigma}}]$ belongs to the following class.

Definition 5.7. Let $\mathfrak{B}^k_{\mathrm{Rot}}$ denote the set of all smooth families of defining pairs $[\Phi;\mathfrak{b}]\in\mathfrak{A}^{k,2k}$ that, in addition to $\mathrm{a})_{k,2k}$, $\Phi_1)_{k,2k}$, $\Phi_2)_{k,2k}$, satisfy:

The *support condition*:

- $\mathrm{b})_k$ $\mathrm{supp}\,\mathfrak{b}_t$ is contained in an $O(\varepsilon_\circ)$-neighbourhood of $\mathrm{supp}\,\mathfrak{b}_t\cap\mathcal{Z}_t$ where $\mathcal{Z}_t$ denotes the *fold surface*

$$\mathcal{Z}_t:=\big\{(x;z)\in\mathbb{R}^2\times\mathbb{R}^2:\Phi_t(x;z)=\mathrm{Rot}(\Phi_t)(x;z)=0\big\}. \tag{5.17}$$

The *fold conditions*: For every $(x_0;z_0)\in\mathrm{supp}\,\mathfrak{b}_{t_0}\cap\mathcal{Z}_{t_0}$ there exist:

- $\mathrm{F}_1)_k$ Vectors $U=(u_1,u_2,u_3), V=(v_1,v_2,v_3)\in\mathbb{R}^3$ satisfying

$$\Big|\Big\langle \partial^2_{zz}\Big\langle U,\begin{bmatrix}\Phi_{t_0}\\ \partial_x\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}V'',\,V''\Big\rangle\Big|\sim 2^{-4k/3},$$

$$\Big|\Big\langle \partial^2_{xx}\Big\langle V,\begin{bmatrix}\Phi_{t_0}\\ \partial_z\Phi_{t_0}\end{bmatrix}\Big\rangle\Big|_{(x_0;z_0)}U'',\,U''\Big\rangle\Big|\sim 2^{-4k/3},$$

  where $U''=(u_2,u_3)$ and $V''=(v_2,v_3)$.
- $\mathrm{F}_2)_k$ 3×3 real matrices $\mathbf{X}$ and $\mathbf{Z}$ such that:
  - i) If $\mathbf{X}^{ij}$ and $\mathbf{Z}^{ij}$ denote the (i,j) entry of $\mathbf{X}$ and $\mathbf{Z}$, respectively, then

$$|\mathbf{X}^{ij}|\lesssim\begin{cases}2^{-2k} & \text{if } (i,j)\in\{(1,3),(2,3)\}\\ 1 & \text{otherwise}\end{cases},\qquad |\mathbf{Z}^{ij}|\lesssim 1.$$

ii) $\mathbf{X}e_3 = U$, $\mathbf{Z}e_3 = V$ and $|\det \mathbf{X}| \sim |\det \mathbf{Z}| \sim 1$.

iii) The matrices $\mathbf{X}, \mathbf{Z}$ transform $\mathfrak{M}(\Phi_{t_0})(x_0; z_0)$ into the normal form

$$\mathbf{X}^\top \circ \mathfrak{M}(\Phi_{t_0})(x_0; z_0) \circ \mathbf{Z} = \begin{bmatrix} \mathbf{M}_{t_0}(x_0; z_0) & 0 \\ 0 & 0 \end{bmatrix},$$

where the 2×2 principal minor satisfies $|\det \mathbf{M}_{t_0}(x_0; z_0)| \sim 2^{4k/3}$.

For $[\Phi^{k,\ell}; \mathfrak{b}^{k,\ell,\vec{\sigma}}]$ the support condition is satisfied owing to the choice of localisation whilst, for the fold conditions, U, V and $\mathbf{X}$, $\mathbf{Z}$ can be taken to be suitably rescaled versions of the vectors in (4.5) and the matrices in (4.6), respectively.

Nonvanishing rotational curvature. By (5.9), each $[\Phi^{k,\ell}; \tilde{\mathfrak{a}}^{k,\ell,\vec{\sigma}} - \mathfrak{b}^{k,\ell,\vec{\sigma}}]$ belongs to the following class.

Definition 5.8. Let $\mathfrak{A}^{k,\ell}_{\mathrm{Rot}}$ denote the set of all $[\Phi; \mathfrak{a}] \in \mathfrak{A}^{k,\ell}$ that satisfy

$\mathbf{R})_{k,\ell}$ $\quad \mathrm{Rot}(\Phi_t) \sim 1$ on $\mathrm{supp}\, \mathfrak{a}_t$.

Recalling (5.12), to prove Lemma 5.1 i) it therefore suffices to show:

$$\big\| \sup_{t \in I} |A[\Phi_t; \mathfrak{b}_t]| \big\|_{L^p \to L^p} \lesssim 2^{-\frac{2k}{3} + (1-\frac{2}{p})k - k\varepsilon_p} \|\mathfrak{b}\|_{C^N} \qquad \text{if } [\Phi; \mathfrak{b}] \in \mathfrak{B}^k_{\mathrm{Cin}} \cap \mathfrak{B}^k_{\mathrm{Rot}},$$
$$\big\| \sup_{t \in I} |A[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p \to L^p} \lesssim 2^{-\frac{e(k,\ell)}{3} + (1-\frac{2}{p})k - k\varepsilon_p} \|\mathfrak{a}\|_{C^N} \qquad \text{if } [\Phi; \mathfrak{a}] \in \mathfrak{A}^{k,\ell}_{\mathrm{Cin}} \cap \mathfrak{A}^{k,\ell}_{\mathrm{Rot}},$$

where I is an interval of length $|I| \sim 1$ containing the t-support of $\mathfrak{a}$ or $\mathfrak{b}$.

Similarly, by (5.10) and (5.11), each $[\Phi^k; \tilde{\mathfrak{c}}^{k,\vec{\sigma}}]$ belongs to the following class.

Definition 5.9. Let $\mathfrak{C}^k_{\mathrm{Rot}}$ denote the set of all $[\Phi; \mathfrak{c}] \in \mathfrak{C}^k$ that satisfy $\mathbf{R})_{k,\ell(k)}$ and

$\mathbf{R}^\star)_k$ $\quad \mathrm{Rot}(\Phi^\star_{x_2}) \sim 1$ on $\mathrm{supp}\, \mathfrak{c}^\star_{x_2}$

where $\Phi^\star_{x_2}(x_1, t; z) := \Phi_t(x_1, x_2; z)$ and $\mathfrak{c}^\star_{x_2}(x_1, t; z) := \mathfrak{c}_t(x_1, x_2; z)$.

Thus, recalling (5.13), to prove Lemma 5.1 ii) it suffices to show:

$$\big\| \sup_{t \in I} |A[\Phi_t; \mathfrak{c}_t]| \big\|_{L^p \to L^p} \lesssim 2^{-\frac{2k}{3} + (1-\frac{2}{p})k - k\varepsilon_p} \|\mathfrak{c}\|_{C^N} \qquad \text{if } [\Phi; \mathfrak{c}] \in \mathfrak{C}^k_{\mathrm{Cin}} \cap \mathfrak{C}^k_{\mathrm{Rot}},$$

where I is an interval of length $|I| \sim 1$ containing the t-support of $\mathfrak{c}$.

5.4.2. *The cases $m \neq 0$.* If $m > 0$, then it follows from (5.15) that $[\Phi^m; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}] \in \mathfrak{A}^{0,0}_{\mathrm{Rot}} =: \mathfrak{A}^0_{\mathrm{Rot}}$. On the other hand, if $m < 0$, then (5.15) implies that $[\Phi^m; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}]$ has favourable rotational curvature properties once the rôles of the r and t variables are interchanged. In particular, in this case $[\Phi^m; \tilde{\mathfrak{a}}^{m,\vec{\sigma}}]$ belongs to the following class.

Definition 5.10. For $m < 0$ let $\mathfrak{A}^m_{\mathrm{Rot}}$ denote the set of all $[\Phi; \mathfrak{a}] \in \mathfrak{A}^m$ that satisfy

$\mathbf{R}^\star)_m$ $\quad \mathrm{Rot}(\Phi^\star_{x_2}) \sim 1$ on $\mathrm{supp}\, \mathfrak{a}^\star_{x_2}$

where $\Phi^\star_{x_2}(x_1, t; z) := \Phi_t(x_1, x_2; z)$ and $\mathfrak{a}^\star_{x_2}(x_1, t; z) := \mathfrak{a}_t(x_1, x_2; z)$.

Thus, recalling (5.16), to prove Lemma 5.1 iii) and iv) it suffices to show that

$$\big\| \sup_{t \in I} |A[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p \to L^p} \lesssim \|\mathfrak{a}\|_{C^N} \qquad \text{if } [\Phi; \mathfrak{a}] \in \mathfrak{A}^0_{\mathrm{Cin}} \cap \mathfrak{A}^0_{\mathrm{Rot}},$$
$$\big\| \sup_{t \in I} |A[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p \to L^p} \lesssim 2^{m\varepsilon_p} \sup_{x_2} \|\mathfrak{a}\|_{C^N_{x_1,z,t}} \qquad \text{if } [\Phi; \mathfrak{a}] \in \mathfrak{A}^m_{\mathrm{Cin}} \cap \mathfrak{A}^m_{\mathrm{Rot}},\ m < 0,$$

where I is an interval of length $|I| \sim 1$ containing the t-support of $\mathfrak{a}$.

5.5. **Frequency decomposition.** Given a smooth family of defining pairs $[\Phi;\mathfrak{a}]$ note that, since the inverse Fourier transform $\check{\eta}$ of the cutoff η from (5.1a) has unit mean,

$$A[\Phi_t;\mathfrak{a}_t]f(x) = \lim_{j\to\infty} 2^j \int_{\mathbb{R}^2} \check{\eta}(2^j\Phi_t(x;z))\mathfrak{a}_t(x;z)f(z)\,\mathrm{d}z,$$

where η is a bump function as in (5.1a). The integral formula for $\check{\eta}$ then yields

$$A[\Phi_t;\mathfrak{a}_t] = A_{\leq J}[\Phi_t;\mathfrak{a}_t] + \sum_{j=J}^{\infty} A_j[\Phi_t;\mathfrak{a}_t]$$

for any $J\in\mathbb{Z}$ where

$$\begin{aligned} A_{\leq J}[\Phi_t;\mathfrak{a}_t]f(x) &:= \frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}} e^{i\theta\Phi_t(x;z)}\mathfrak{a}_t(x;z)\eta^J(\theta)\,\mathrm{d}\theta\, f(z)\,\mathrm{d}z,\\ A_j[\Phi_t;\mathfrak{a}_t]f(x) &:= \frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}} e^{i\theta\Phi_t(x;z)}\mathfrak{a}_t(x;z)\beta^j(|\theta|)\,\mathrm{d}\theta\, f(z)\,\mathrm{d}z. \end{aligned} \tag{5.18}$$

This provides a frequency decomposition of (4.1). The low frequency part of the operator (corresponding to $A_{\leq J}[\Phi_t;\mathfrak{a}_t]$ for a suitable choice of J) can be dealt with via pointwise comparison with the Hardy–Littlewood maximal operator, and so the remainder of the article will focus on the high frequency parts. In view of this and the observations of the preceding subsection, Theorem 5.1 is a consequence of the following proposition, which will be proved in §7 and §8 using the theory developed in §6.

Proposition 5.11. *There exists $N\in\mathbb{N}$, $\varepsilon_p>0$ such that for all $k\geq -4$, $(k,\ell)\in\mathfrak{P}$, $j\geq -e(k,\ell)/3$ and $2<p<\infty$, the following bounds hold, with the implicit constants depending on p. In each inequality, I denotes an interval of length $|I|\sim 1$ containing the t-support of the amplitude.*

(i) For $[\Phi;\mathfrak{b}]\in\mathfrak{B}^k_{\mathrm{Cin}}\cap\mathfrak{B}^k_{\mathrm{Rot}}$,

$$\big\|\sup_{t\in I}|A_j[\Phi_t;\mathfrak{b}_t]f|\big\|_p \lesssim 2^{-(j\vee 0)\varepsilon_p}2^{-\frac{2k}{3}+(1-\frac{2}{p})k-k\varepsilon_p}\|\mathfrak{b}\|_{C^N}\|f\|_p.$$

(ii) For $[\Phi;\mathfrak{a}]\in\mathfrak{A}^{k,\ell}_{\mathrm{Cin}}\cap\mathfrak{A}^{k,\ell}_{\mathrm{Rot}}$,

$$\big\|\sup_{t\in I}|A_j[\Phi_t;\mathfrak{a}_t]f|\big\|_p \lesssim 2^{-(j\vee 0)\varepsilon_p}2^{-\frac{e(k,\ell)}{3}+(1-\frac{2}{p})k-k\varepsilon_p}\|\mathfrak{a}\|_{C^N}\|f\|_p.$$

(iii) For $[\Phi;\mathfrak{c}]\in\mathfrak{C}^k_{\mathrm{Cin}}\cap\mathfrak{C}^k_{\mathrm{Rot}}$,

$$\big\|\sup_{t\in I}|A_j[\Phi_t;\mathfrak{c}_t]f|\big\|_p \lesssim 2^{-(j\vee 0)\varepsilon_p}2^{-\frac{2k}{3}+k(1-\frac{2}{p})-k\varepsilon_p}\|\mathfrak{c}\|_{C^N}\|f\|_p.$$

(iv) For $[\Phi;\mathfrak{a}]\in\mathfrak{A}^0_{\mathrm{Cin}}\cap\mathfrak{A}^0_{\mathrm{Rot}}$,

$$\big\|\sup_{t\in I}|A_j[\Phi_t;\mathfrak{a}_t]f|\big\|_p \lesssim 2^{-j\varepsilon_p}\|\mathfrak{a}\|_{C^N}\|f\|_p.$$

(v) For $m<0$ and $[\Phi;\mathfrak{a}]\in\mathfrak{A}^m_{\mathrm{Cin}}\cap\mathfrak{A}^m_{\mathrm{Rot}}$,

$$\big\|\sup_{t\in I}|A_j[\Phi_t;\mathfrak{a}_t]f|\big\|_p \lesssim 2^{-j\varepsilon_p}2^{m\varepsilon_p}\sup_{x_2}\|\mathfrak{a}\|_{C^N_{x_1,z,t}}\|f\|_p.$$

Remark. Here cases i), iii), iv) and v) are understood to hold for $\ell=2k$ so that j ranges over values $j\geq -2k/3$, with $k=0$ in the cases iv) and v). In each case, similar estimates hold for $A_{\leq -e(k,\ell)/3}[\Phi_t;\mathfrak{a}_t]$ (corresponding to the low frequency part), which can be proved by elementary means.

6. L^2 BOUNDS FOR TWO PARAMETER OSCILLATORY INTEGRAL OPERATORS

The first step towards establishing Proposition 5.11 is to obtain L^2 bounds for the frequency localised pieces with favourable dependence in the parameters k and ℓ. This will follow from certain estimates for maximal functions associated to two parameter oscillatory integrals, which will be proven in this section.

To this end, let $U \subset \mathbb{R}^d \times \mathbb{R}^d$ be an open set, $\Psi : U \to \mathbb{R}$ be a smooth phase function and $a \in C_0^\infty(U)$. Consider, for $\lambda > 1$, the oscillatory integral operator associated to the phase/amplitude pair $[\Psi; a]$,

$$T^\lambda f(x) \equiv T^\lambda[\Psi; a]f(x) := \int_{\mathbb{R}^d} e^{i\lambda\Psi(x;z)} a(x; z) f(z)\, \mathrm{d}z. \tag{6.1}$$

We now let $0 < \delta_\circ \le 1$ and we shall assume that the following nonisotropic derivative estimates

$$|\partial_x^\alpha \partial_z^\beta \Psi(x; z)| + \delta_\circ^{-1} |\partial_x^\alpha \partial_z^\beta \partial_{x_d} \Psi(x; z)| \le C_{\alpha,\beta} \tag{6.2}$$

hold for all $(x; z) \in U$ and all $\alpha, \beta \in \mathbb{N}_0^d$. We shall then derive estimates in terms of the two parameters $\lambda > 1$ and $\delta_\circ \le 1$. Our results could be rewritten as a two parameter oscillatory integral estimates with phase $\lambda(\varphi(x'; z) + \delta_\circ \psi(x; z))$, where $x = (x', x_d)$, and uniform upper bound derivative estimates on φ and ψ.

6.1. **The nondegenerate case.** We first formulate a variant of the classical L^2 result of Hörmander in [13] under the assumption (6.2).

Proposition 6.1. *Let $\lambda \ge 1$, $0 < \delta_\circ \le 1$, Ψ be as in* (6.2) *and suppose that there is $c > 0$ such that $|\det \partial^2_{zx} \Psi(x_0; z_0)| \ge c\delta_\circ$ for some $(x_0; z_0) \in U$. Then there exist $\varepsilon_\circ > 0$ and $N > 0$, independent of λ and $\delta_\circ$, such that for all smooth a with* supp $a \subset B_{\varepsilon_\circ}(x_0; z_0)$,

$$\|T^\lambda\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \min\{(\lambda\delta_\circ)^{-1/2}, 1\} \|a\|_{C^N}.$$

Proof. After applying translation operators we may assume $(x_0; z_0) = (0; 0)$. The kernel of $T^\lambda(T^\lambda)^*$ is given by

$$K^\lambda(x, y) := \int_{\mathbb{R}^d} e^{i\lambda(\Psi(x;z)-\Psi(y;z))} a(x; z)\overline{a(y; z)}\, \mathrm{d}z,$$

and by the Schur's test, the desired estimate follows from the bounds

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} |K^\lambda(x, y)|\, \mathrm{d}y, \quad \sup_{y\in\mathbb{R}^d} \int_{\mathbb{R}^d} |K^\lambda(x, y)|\, \mathrm{d}x \;\lesssim \lambda^{-(d-1)} \min\{(\lambda\delta_\circ)^{-1}, 1\} \|a\|^2_{C^N}. \tag{6.3}$$

We have

$$\nabla_z(\Psi(x; z) - \Psi(y; z)) = A_{\delta_\circ}(x, y; z) \begin{bmatrix} x' - y' \\ \delta_\circ(x_d - y_d) \end{bmatrix}$$

where $x = (x', x_d), y = (y', y_d)$ and

$$A_{\delta_\circ}(x, y; z) = \int_0^1 \begin{bmatrix} \partial^2_{z'x'}\Psi & \delta_\circ^{-1}\partial^2_{z'x_d}\Psi \\ \partial^2_{z_d x'}\Psi & \delta_\circ^{-1}\partial^2_{z_d x_d}\Psi \end{bmatrix} \Big|_{(y+s(x-y);z)} \mathrm{d}s.$$

By (6.2) we have $\|A_{\delta_\circ}\|_{C^N} \lesssim_N 1$. Also clearly $|\det A_{\delta_\circ}(0, 0; 0)| \ge c$ and thus there is an $\varepsilon_\circ > 0$ such that for $|(x, y; z)| \le \varepsilon_\circ$ the matrix A_δ is invertible and we obtain the estimate $\|\partial^\alpha_{x,y,z} A^{-1}_{\delta_\circ}(x,y;z)\| \le C_\alpha$ for all $\alpha \in \mathbb{N}_0^{3d}$ for the matrix norms of the derivatives of $A^{-1}_{\delta_\circ}$. Hence for $|x|, |y|, |z| \le \varepsilon_\circ$

$$|\nabla_z(\Psi(x;z) - \Psi(y;z))| \ge c(|x'-y'| + \delta_\circ|x_d - y_d|).$$

By (6.2) we have

$$|\partial^\alpha_z(\Psi(x;z) - \Psi(y;z))| \le C(|x'-y'| + \delta_\circ|x_d - y_d|)$$

for all $\alpha \in \mathbb{N}_0^d$. By repeated integration-by-parts in the form of Corollary A.2, with the choices of $\rho(x,y) = |x'-y'| + \delta_\circ|x_d - y_d|$ and $R_2(x,y) = 1$, one obtains

$$|K^\lambda(x,y)| \lesssim_N \|a\|^2_{C^N}(1 + \lambda|x'-y'| + \lambda\delta_\circ|x_d - y_d|)^{-N}.$$

In view of the compact support of a, the desired bounds (6.3) follow from integrating in $x \in \operatorname{supp} a$ for fixed $y \in \operatorname{supp} a$, and in $y \in \operatorname{supp} a$ for fixed $x \in \operatorname{supp} a$ respectively. □

6.2. **A two parameter oscillatory integral estimate under two-sided fold conditions.** We shall also formulate a variant of the L^2 estimates for oscillatory integral operators with fold singularities of Pan and Sogge [24], which are based on the previous work on Fourier integral operators by Melrose and Taylor [18], under the assumption (6.2). We will instead follow the approach in the works of Phong and Stein [25], Cuccagna [8] and Greenleaf and the fourth author [11].

Proposition 6.2. *Let $\lambda \ge 1$, $0 < \delta_\circ < 1$, Ψ be as in (6.2) and suppose that for some $(x_0;z_0) \in U$ there is $c > 0$ such that*

$$|\det \partial^2_{z'x'}\Psi(x_0;z_0)| \ge c, \tag{6.4a}$$

$$\partial^2_{zx_d}\Psi(x_0;z_0) = 0, \qquad \partial^2_{z_d x}\Psi(x_0;z_0) = 0, \tag{6.4b}$$

$$|\partial^3_{x_d z_d z_d}\Psi(x_0;z_0)| \ge c\delta_\circ, \qquad |\partial^3_{z_d x_d x_d}\Psi(x_0;z_0)| \ge c\delta_\circ. \tag{6.4c}$$

Then there exist $\varepsilon_\circ > 0$ and $N > 0$, independent of λ and $\delta_\circ$, such that for all smooth a with $\operatorname{supp} a \subset B_{\varepsilon_\circ}(x_0;z_0)$*,*

$$\|T^\lambda\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \min\{(\lambda\delta_\circ)^{-1/3}, 1\}\|a\|_{C^N}.$$

Following [25, 8, 11], we decompose dyadically our operator according to the size of $\det \partial^2_{xz}\Psi$. It is useful to consider the auxiliary quantity

$$\sigma \equiv \sigma(\Psi) = \partial^2_{x_d z_d}\Psi - \partial^2_{x_d z'}\Psi[(\partial^2_{x'z'}\Psi)^\top]^{-1}\partial^2_{z_d x'}\Psi, \tag{6.5}$$

which measures the size of the mixed Hessian. In fact, note that if A is an invertible $(d-1)\times(d-1)$ matrix, $b, c \in \mathbb{R}^{d-1}$ and $d \in \mathbb{R}$, one has the identity

$$\begin{bmatrix} I & 0 \\ -c^\top A^{-1} & 1 \end{bmatrix} \begin{bmatrix} A & b \\ c^\top & d \end{bmatrix} = \begin{bmatrix} A & b \\ 0^\top & d - c^\top A^{-1} d \end{bmatrix}$$

and therefore

$$\det \partial^2_{xz}\Psi(x;z) = \sigma(x;z)\det \partial^2_{x'z'}\Psi(x;z) \tag{6.6}$$

for $(x;z)$ near $(x_0;z_0)$. Hence we get, assuming that $\varepsilon_\circ$ is small enough,

$$|\sigma(x;z)| \sim |\det \partial^2_{xz}\Psi(x;z)|.$$

The fold conditions (6.4c) together with (6.4b) imply that

$$|\partial_{x_d}\sigma(x;z)| = |\partial^3_{x_d x_d z_d}\Psi(x;z)| + O(\varepsilon_\circ\delta_\circ),$$
$$|\partial_{z_d}\sigma(x;z)| = |\partial^3_{x_d z_d z_d}\Psi(x;z)| + O(\varepsilon_\circ\delta_\circ),$$

and using (6.4c) we get

$$|\partial_{x_d}\sigma(x;z)| \sim \delta_\circ, \qquad |\partial_{z_d}\sigma(x;z)| \sim \delta_\circ. \tag{6.7}$$

Finally, note that the assumption (6.2) implies

$$|\partial_x^\alpha \partial_z^\beta \sigma(x;z)| \lesssim_{\alpha,\beta} \delta_\circ \tag{6.8}$$

for all $\alpha, \beta \in \mathbb{N}_0^d$.

Let η_0, η_1 be C^∞ functions on the real line with

$$\operatorname{supp} \eta_0 \subset [-2,2], \qquad \operatorname{supp} \eta_1 \subset [-2,-1/2] \cup [1/2,2].$$

For $\lambda \geq 1$, set

$$M := \max\{\lfloor \log_2(\lambda^{1/2}) \rfloor, 0\} \tag{6.9}$$

and define

$$T^{\lambda,m} f(x) := \int_{\mathbb{R}^d} e^{i\lambda\Psi(x;z)} a(x;z) \eta_1(2^m \delta_\circ^{-1} |\sigma(x;z)|) f(z)\, \mathrm{d}z, \quad 0 \leq m < M, \tag{6.10}$$

$$T^{\lambda,M} f(x) := \int_{\mathbb{R}^d} e^{i\lambda\Psi(x;z)} a(x;z) \eta_0(2^M \delta_\circ^{-1} \sigma(x;z)) f(z)\, \mathrm{d}z. \tag{6.11}$$

By (6.6) and (6.8) we have $|\det \partial^2_{zx}\Psi| \sim 2^{-m}\delta_\circ$ on the support of the amplitude in $T^{\lambda,m}$ if $0 \leq m < M$ and $|\det \partial^2_{zx}\Psi| \lesssim 2^{-M}\delta_\circ \lesssim \lambda^{-1/2}\delta_\circ$ on the support of the amplitude in $T^{\lambda,M}$.

Proposition 6.3. *Let $\lambda \geq 1$, $\delta_\circ < 1$, $[\Psi; a]$ be as in Proposition 6.2 and M as in (6.9).*

(i) If $\lambda \geq \delta_\circ^{-1}$ then, for $0 \leq m < M$,

$$\|T^{\lambda,m}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \min\{(2^m/(\lambda\delta_\circ))^{1/2}, 2^{-m}\} \|a\|_{C^N}.$$

Moreover,

$$\|T^{\lambda,M}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d}{2}} \|a\|_{C^N}.$$

(ii) If $1 \leq \lambda \leq \delta_\circ^{-1}$ then, for $0 \leq m < M$

$$\|T^{\lambda,m}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim 2^{-m} \lambda^{-\frac{d-1}{2}} \|a\|_{C^N}.$$

Moreover,

$$\|T^{\lambda,M}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d}{2}} \|a\|_{C^N}.$$

We first note that the bounds in Proposition 6.3 imply Proposition 6.2 by summing in the m-parameter.

Proof of Proposition 6.2, assuming Proposition 6.3. Let η, β be defined as in (5.1a), (5.1b). Taking $\eta_1 = \beta(|\cdot|)$ and $\eta_0 = \eta$ in the definitions (6.10), (6.11), we have $T^\lambda = \sum_{m=0}^M T^{\lambda,m}$.

If $\lambda\delta_\circ \leq 1$, the bound trivially follows from summing in m the estimates in (ii) in Proposition 6.3.

If $\lambda\delta_\circ \geq 1$, note that the bounds in (i) in Proposition 6.3 imply

$$\begin{aligned} \|T^\lambda\|_{L^2\to L^2} &\lesssim \lambda^{-\frac{d-1}{2}} \Big(\sum_{1 \leq 2^m \leq (\lambda\delta_\circ)^{1/3}} 2^{m/2} (\lambda\delta_\circ)^{-1/2} + \sum_{(\lambda\delta_\circ)^{1/3} < 2^m \leq \lambda^{1/2}} 2^{-m} \Big) \|a\|_{C^N} \\ &\lesssim \lambda^{-\frac{d-1}{2}} (\lambda\delta_\circ)^{-1/3} \|a\|_{C^N}, \end{aligned}$$

as desired. □

6.3. *Proof of Proposition 6.3.* We fix $N \geq 100d$. As the operators depend linearly on a we may assume $\|a\|_{C_N} \leq 1$. The proof is based on a variant of the arguments in [25], [8], [11]; the latter two are themselves inspired by the Calderón–Vaillancourt theorem on the L^2 boundedness of pseudo-differential operators [6]. Again, by performing translations we may take $(x_0; z_0) = (0; 0)$.

Recall that, by hypothesis, $\sigma(0;0) = 0$ and by (6.8) and (6.7) we have that $|\partial_{x_d}\sigma| \sim \delta_\circ$, $|\partial_{z_d}\sigma| \sim \delta_\circ$ and $|\partial_x^\alpha \partial_z^\beta \sigma| \lesssim_{\alpha,\beta} \delta_\circ$ in $B_{\varepsilon_\circ}(0;0)$ for some small $\varepsilon_\circ > 0$. By an application of a quantitative version of the implicit function theorem (see for example [7, §8]) there exist smooth functions

$$(x'; z) \mapsto \mathfrak{u}(x'; z) \qquad \text{and} \qquad (x; z') \mapsto \mathfrak{v}(x; z'),$$

defined for $|x'| \leq 2\varepsilon_\circ$, $|z| \leq 2\varepsilon_\circ$ and $|x| \leq 2\varepsilon_\circ$, $|z'| \leq 2\varepsilon_\circ$ respectively, such that

$$\sigma(x', \mathfrak{u}(x'; z); z) = 0 \qquad \text{and} \qquad \sigma(x; z', \mathfrak{v}(x; z')) = 0.$$

Furthermore, by (6.7)

$$|\mathfrak{u}(x'; z) - x_d|,\ |\mathfrak{v}(x; z') - z_d| \sim \delta_\circ^{-1} |\sigma(x; z)|.$$

We may expand $|x_d - y_d| \leq |x_d - \mathfrak{u}(x'; z)| + |\mathfrak{u}(x'; z) - \mathfrak{u}(y'; z)| + |\mathfrak{u}(y'; z) - y_d|$ and obtain the crucial estimate

$$|\sigma(x; z)| \sim 2^{-m}\delta_\circ,\ |\sigma(y; z)| \sim 2^{-m}\delta_\circ \implies |x_d - y_d| \lesssim 2^{-m} + |x' - y'| \tag{6.12}$$

and similarly (using $\mathfrak{v}$)

$$|\sigma(x; w)| \sim 2^{-m}\delta_\circ,\ |\sigma(x; z)| \sim 2^{-m}\delta_\circ \implies |w_d - z_d| \lesssim 2^{-m} + |w' - z'|.$$

These observations suggest further decomposing the amplitude into functions supported essentially on $C\varepsilon_\circ 2^{-m}$ cubes. Let $\zeta \in C_0^\infty(\mathbb{R})$ supported in $(-1, 1)$ such that $\sum_{n \in \mathbb{Z}} \zeta(\cdot - n) \equiv 1$. Set

$$b^m_{\mu\nu}(x; z) := a(x; z)\eta_1(2^m \delta_\circ^{-1} \sigma(x; z)) \Big(\prod_{j=1}^d \zeta(\varepsilon_\circ^{-1} 2^m x_j - \mu_j) \zeta(\varepsilon_\circ^{-1} 2^m z_j - \nu_j) \Big)$$

and write the corresponding decomposition

$$T^{\lambda,m} = \sum_{(\mu,\nu) \in \mathbb{Z}^d \times \mathbb{Z}^d} T^{\lambda,m}_{\mu\nu}$$

where $T^{\lambda,m}_{\mu\nu} f := T^\lambda[\Psi; b^m_{\mu\nu}] f$. Observe that

$$|\partial_x^\alpha \partial_z^\beta b^m_{\mu\nu}(x; z)| \lesssim 2^{m(|\alpha| + |\beta|)}$$

for all $\alpha, \beta \in \mathbb{N}_0^d$. By the Cotlar–Stein lemma (see, for instance, [32, Chapter VII, §2.1]), the proof of the proposition reduces to showing the estimates

$$\|T^{\lambda,m}_{\mu\nu} (T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}})^*\|_{2\to 2} + \|(T^{\lambda,m}_{\mu\nu})^* T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}}\|_{2\to 2} \lesssim \frac{\lambda^{-(d-1)} \min\{2^m/(\lambda\delta_\circ), 2^{-2m}\}}{(1 + |\mu - \tilde{\mu}| + |\nu - \tilde{\nu}|)^{3d}} \tag{6.13}$$

for all $(\mu, \nu), (\tilde{\mu}, \tilde{\nu}) \in \mathbb{Z}^d \times \mathbb{Z}^d$. The proof of (6.13) is divided in two cases.

Off-diagonal estimates. The first step is to establish (6.13) in the *off-diagonal* case where

$$\max\{|\mu - \tilde{\mu}|, |\nu - \tilde{\nu}|\} \geq C_{\mathrm{diag}} \varepsilon_\circ^{-1} \tag{6.14}$$

for a large absolute constant $C_{\mathrm{diag}} \geq 2$, chosen independently of $\varepsilon_\circ$. To this end, it is convenient to introduce the kernels associated to the operators of the type TT^* and T^*T. The Schwartz kernel of $T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}})^*$ is given by

$$K^{\lambda,m}_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y) := \int_{\mathbb{R}^d} e^{i\lambda(\Psi(x;z)-\Psi(y;z))} b^m_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y;z)\,\mathrm{d}z, \tag{6.15}$$

and the Schwartz kernel of $(T^{\lambda,m}_{\mu\nu})^* T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}}$ is given by

$$\tilde{K}^{\lambda,m}_{\mu\nu,\tilde{\mu}\tilde{\nu}}(z,w) := \int_{\mathbb{R}^d} e^{-i\lambda(\Psi(x;z)-\Psi(x;w))} \tilde{b}^m_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x;z,w)\,\mathrm{d}x;$$

here the symbols are given by

$$b^m_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y;z) := b^m_{\mu\nu}(x;z)\overline{b^m_{\tilde{\mu}\tilde{\nu}}(y;z)}, \qquad \tilde{b}^m_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x;z,w) := b^m_{\mu\nu}(x;z)\overline{b^m_{\tilde{\mu}\tilde{\nu}}(x;w)}.$$

Lemma 6.4 (Off-diagonal estimate). *Let $1 \leq 2^m \leq \lambda^{1/2}$ and suppose that* (6.14) *holds.*

i) If $|\mu - \tilde{\mu}| \geq C_{\mathrm{diag}}\varepsilon_\circ^{-1}$, then $(T^{\lambda,m}_{\mu\nu})^* T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}} \equiv 0$ and

$$\|T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}})^*\|_{2\to 2} \lesssim_N 2^{-2dm}(\lambda 2^{-2m}|\mu - \tilde{\mu}|)^{-N}.$$

ii) If $|\nu - \tilde{\nu}| \geq C_{\mathrm{diag}}\varepsilon_\circ^{-1}$, then $T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}})^* \equiv 0$ and

$$\|(T^{\lambda,m}_{\mu\nu})^* T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}}\|_{2\to 2} \lesssim_N 2^{-2dm}(\lambda 2^{-2m}|\nu - \tilde{\nu}|)^{-N}.$$

Proof. Only the proof of i) is given; the same argument can be applied to ii) *mutatis mutandis* (the asymmetry of assumptions regarding the x_d dependence does not make a difference for the current proof). Furthermore, if $|\mu - \tilde{\mu}| \geq 2$, then it immediately follows from the support properties of the symbols that $(T^{\lambda,m}_{\mu\nu})^* T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}} \equiv 0$ and it only remains to consider the Schwartz kernel $K^{\lambda,m}_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y)$ of of $T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\tilde{\mu}\tilde{\nu}})^*$. By Schur's test, the desired estimate follows from

$$\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} |K^{\lambda,m}_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y)|\,\mathrm{d}y,\ \sup_{y\in\mathbb{R}^d}\int_{\mathbb{R}^d} |K^{\lambda,m}_{\mu\nu,\tilde{\mu}\tilde{\nu}}(x,y)|\,\mathrm{d}x \lesssim \frac{2^{-2dm}(2^{-2m}\lambda)^{-N}}{|\mu - \tilde{\mu}|^N}. \tag{6.16}$$

First note that, provided C_{diag} is suitably chosen, combining the hypothesis $|\mu - \tilde{\mu}| \geq C_{\mathrm{diag}}\varepsilon_\circ^{-1}$ with (6.12) yields

$$|x_d - y_d| \lesssim |x' - y'| \qquad \text{on } \operatorname{supp} b^m_{\mu\nu,\tilde{\mu}\tilde{\nu}}. \tag{6.17}$$

Thus, by Taylor's theorem and (6.17)

$$\big|\partial^\alpha_{z'}(\Psi(x;z) - \Psi(y;z))\big| \lesssim_\alpha |x' - y'|. \tag{6.18a}$$

For the lower bounds we use (6.4a) and, from (6.4b), $\partial^2_{z'x_d}\Psi(0;0) = 0$, to deduce

$$\partial_{z'}\Psi(x;z) - \partial_{z'}\Psi(y;z) = \int_0^1 \partial^2_{z'x'}\Psi(y + s(x-y);z)\,\mathrm{d}s\,(x'-y') + O(\varepsilon_\circ|x_d - y_d|).$$

Thus, from (6.17) we obtain that, for $(x,y;z)$ near $(0,0;0)$,

$$\big|\partial_{z'}(\Psi(x;z) - \Psi(y;z))\big| \geq c|x' - y'|. \tag{6.18b}$$

Finally, $|\partial_z^\alpha b^m_{\mu\nu,\tilde\mu\tilde\nu}| \lesssim_\alpha 2^{m|\alpha|}$, and the z-integration is extended over a set of diameter $O(2^{-m})$. By (6.18b) and (6.18a), we may use repeated integration-by-parts in the form of Corollary A.2, with the choices of $\rho(x,y) := |x'-y'|$ and $R(x,y) := 1$, to obtain

$$|K^{\lambda,m}_{\mu\nu,\tilde\mu\tilde\nu}(x,y)| \lesssim 2^{-dm}(2^{-m}\lambda|x'-y'|)^{-N}.$$

By (6.12), the kernel is identically zero unless $|\mu_3-\tilde\mu_3| \lesssim \max\{1, |\mu'-\tilde\mu'|\}$. Provided C_{diag} is sufficiently large, $|\mu'-\tilde\mu'| \sim |\mu-\tilde\mu|$ and, furthermore, $|\mu'-\tilde\mu'| \geq 2$. Consequently, $\varepsilon_\circ^{-1}2^m|x'-y'| \sim |\mu-\tilde\mu|$ and so

$$|K^{\lambda,m}_{\mu\nu,\tilde\mu\tilde\nu}(x,y)| \lesssim 2^{-dm}(2^{-2m}\lambda|\mu-\tilde\mu|)^{-N}.$$

For fixed x, the support of $y \mapsto K^{\lambda,m}_{\mu\nu,\tilde\mu\tilde\nu}(x,y)$ is a set of measure $O(2^{-dm})$ and likewise, for fixed y the support of $x \mapsto K^{\lambda,m}_{\mu\nu,\tilde\mu\tilde\nu}(x,y)$, and (6.16) follows. □

Diagonal estimates. The proof of (6.13) has now been reduced to the following two lemmata.

Lemma 6.5. *Suppose that $\lambda \geq 1$ and $1 \leq 2^m \lesssim \lambda^{1/2}$. Then, for all $(\mu,\nu) \in \mathbb{Z}^d \times \mathbb{Z}^d$,*

$$\|T^{\lambda,m}_{\mu\nu}\|_{2\to2} \lesssim 2^{-m}\lambda^{-(d-1)/2}.$$

Furthermore,

$$\|T^{\lambda,M}_{\mu\nu}\|_{2\to2} \lesssim \lambda^{-d/2}.$$

Lemma 6.6. *Suppose that $\lambda\delta_\circ \geq 1$ and $1 \leq 2^m \leq (\lambda\delta_\circ)^{1/3}$. Then for all $(\mu,\nu) \in \mathbb{Z}^d \times \mathbb{Z}^d$,*

$$\|T^{\lambda,m}_{\mu\nu}\|_{2\to2} \lesssim 2^{m/2}\delta_\circ^{-1/2}\lambda^{-d/2}.$$

Note that the estimate in Lemma 6.6 is better than the estimate in Lemma 6.5 in the range $\lambda\delta_\circ \geq 1$, $1 \leq 2^m \leq (\lambda\delta_\circ)^{1/3}$.

Proof of Lemma 6.5. Let I_{μ_d}, J_{ν_d} denote the intervals of length $\varepsilon_\circ 2^{1-m}$ centered at $x_{\mu_d} = \varepsilon_\circ 2^{-m}\mu_d$, $z_{\nu_d} = \varepsilon_\circ 2^{-m}\nu_d$, respectively. For $g \in L^2(\mathbb{R}^{d-1})$ define

$$T^{\lambda,m,x_d,z_d}_{\mu\nu} g(x') = \int_{\mathbb{R}^{d-1}} e^{i\lambda\Psi(x',x_d;z',z_d)} b^m_{\mu,\nu}(x;z)g(z')\,\mathrm{d}z'$$

and observe that

$$T^{\lambda,m}_{\mu\nu} f(x) = \mathbb{1}_{I_{\mu_d}}(x_d)\int_{J_{\nu_d}} T^{\lambda,m,x_d,z_d}_{\mu\nu}[f(\cdot,x_d)]\,\mathrm{d}z_d.$$

The Schwartz kernel $K^{\lambda,m,x_d,z_d}_{\mu\nu}(x',y')$ of $T^{\lambda,m,x_d,z_d}_{\mu\nu}(T^{\lambda,M,x_d,z_d}_{\mu\nu})^*$ is equal to

$$\int_{\mathbb{R}^{d-1}} e^{i\lambda(\Psi(x',x_d;z',z_d)-\Psi(y',x_d;z',z_d))} b^m_{\mu\nu}(x',x_d;z',z_d)\overline{b^m_{\mu\nu}(y',x_d;z',z_d)}\,\mathrm{d}z'.$$

We use integration-by-parts based on (6.4a); that is, we use the $(d-1)$-dimensional case of Corollary A.2 with the choices $\rho(x',y') := |x'-y'|$, $R(x,y) := 1$ and the fact that $\partial_{z'}^\alpha$ applied to the amplitude yields a term which is $O(2^{m|\alpha|})$. This implies

$$|K^{\lambda,m,x_d,z_d}_{\mu\nu}(x',y')| \lesssim_N 2^{-m(d-1)}(1+\lambda 2^{-m}|x'-y'|)^{-N}$$

uniformly in x_d, z_d, and by the Schur's test one has

$$\|T^{\lambda,m,x_d,z_d}_{\mu,\nu}\|_{L^2(\mathbb{R}^{d-1})\to L^2(\mathbb{R}^{d-1})} \lesssim \lambda^{-(d-1)/2}.$$

Consequently,

$$\begin{aligned}\|T^{\lambda,m}_{\mu\nu} f\|_{L^2(\mathbb{R}^d)} &\lesssim \int_{J_{\nu_d}} \Big(\int_{I_{\mu_d}} \|T^{\lambda,m,x_d,z_d}_{\mu\nu}[f(\cdot,z_d)]\|^2_{L^2(\mathbb{R}^{d-1})} \mathrm{d}x_d \Big)^{1/2} \mathrm{d}z_d \\ &\lesssim 2^{-m/2}\lambda^{-(d-1)/2} \int_{J_{\nu_d}} \|[f(\cdot,z_d)]\|_{L^2(\mathbb{R}^{d-1})}\,\mathrm{d}z_d \\ &\lesssim 2^{-m}\lambda^{-(d-1)/2}\|f\|_{L^2(\mathbb{R}^d)}\end{aligned}$$

and hence $\|T^{\lambda,m}_{\mu\nu}\|_{2\to 2} \lesssim 2^{-m}\lambda^{-(d-1)/2}$ as desired. The arguments for $T^{\lambda,M}_{\mu\nu}$ is analogous. □

Proof of Lemma 6.6. Let $K^{\lambda,m}_{\mu\nu} := K^{\lambda,m}_{\mu\nu,\mu\nu}$ denote the kernel of $T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\mu\nu})^*$, as given by the formula in (6.15). It will also be useful to write $b^m_{\mu\nu}$ for the symbol $b^m_{\mu\nu,\mu\nu}$. By the Schur test, the problem is reduced to showing

$$\sup_{y\in\mathbb{R}^d} \int_{\mathbb{R}^d} |K^{\lambda,m}_{\mu\nu}(x,y)|\,\mathrm{d}x \lesssim 2^m\delta_\circ^{-1}\lambda^{-d}, \tag{6.19a}$$

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} |K^{\lambda,m}_{\mu\nu}(x,y)|\,\mathrm{d}y \lesssim 2^m\delta_\circ^{-1}\lambda^{-d}. \tag{6.19b}$$

Since $T^{\lambda,m}_{\mu\nu}(T^{\lambda,m}_{\mu\nu})^*$ is self-adjoint (6.19b) follows from (6.19a). We proceed to show (6.19a).

Since the partial mixed Hessian $\partial^2_{z'x'}\Psi$ is non-singular, there exist local solutions in x' to the implicit equation $\nabla_{z'}\Psi(x;z) = \nabla_{z'}\Psi(y;z)$. In particular, by applying a quantitative version of the implicit function theorem (see, for instance, [7, §8]), provided $\varepsilon_\circ$ is chosen suitably small, there exists a smooth $\mathbb{R}^{d-1}$-valued function $(x_d, y, z) \mapsto X(x_d;y;z)$ defined by

$$\partial_{z'}\Psi(X(x_d;y;z), x_d; z) = \partial_{z'}\Psi(y;z), \tag{6.20}$$

$$X(y_d;y;z) = y'. \tag{6.21}$$

Implicit differentiation yields

$$\partial_{x_d} X(x_d;y;z) = -(\partial^2_{x'z'}\Psi)^{-1}\partial^2_{z'x_d}\Psi\Big|_{(X(x_d;y;z),x_d;z)}. \tag{6.22}$$

From this formula, the chain rule and the definition of σ one deduces that

$$\partial_{x_d}\big[\partial_{z_d}\Psi(X(x_d;y;z), x_d;z))\big] = \sigma(X(x_d;y;z), x_d;z). \tag{6.23}$$

Notice that the right hand side of (6.22) vanishes at $(x_d;y;z) = (0;0;0)$ and that $\partial^\alpha_{x_d} X(x_d;y;z) = O(\delta_\circ)$. Hence we get

$$|\partial_{x_d} X(x_d;y;z)| \lesssim \varepsilon_\circ\delta_\circ. \tag{6.24}$$

Moreover, implicit differentiation of (6.20) with respect to z yields

$$\begin{aligned}\partial^2_{z'x'}\Psi(X(x_d;y;z), y_d;z)\partial_z X(x_d;y;z) &= \partial^2_{z'z}\Psi(y', y_d;z) - \partial^2_{z'z}\Psi(X(x_d;y;z), x_d;z) \\ &\lesssim |y' - X(x_d;y;z)| + \delta_\circ|x_d - y_d| \\ &= O(\delta_\circ|x_d - y_d|),\end{aligned}$$

where we have used (6.21) and (6.24). This gives

$$|\partial_z X(x_d;y;z)| \lesssim \delta_\circ|x_d - y_d|. \tag{6.25}$$

We shall now state the inequalities for the integration-by-parts argument which will allow us to prove (6.19a). In what follows we write $X := X(x_d;y;z)$ and

$X_\nu := X(x_d; y; z_\nu)$ where $z_\nu := \varepsilon_\circ 2^m \nu$, noting that the z-support of $b^m_{\mu\nu}$ lies in a ball of radius $O(\varepsilon_\circ 2^{-m})$ about this point. We claim that

$$\big|\partial_z^\alpha \Psi(x;z) - \partial_z^\alpha \Psi(y;z)\big| \le C_\alpha\big(|x' - X_\nu| + \delta_\circ |x_d - y_d|\big) \tag{6.26}$$

and

$$|\nabla_z \Psi(x;z) - \nabla_z \Psi(y;z)| \ge c\big(|x' - X_\nu| + \delta_\circ 2^{-m}|x_d - y_d|\big). \tag{6.27}$$

To see (6.26), by Taylor expansion the left-hand side is dominated by a constant times $|x' - y'| + \delta_\circ |x_d - y_d|$. We then bound $|x' - y'| \le |x' - X_\nu| + |y' - X_\nu|$ and, using (6.21), by the mean value theorem, (6.24) and (6.25) one has

$$\begin{aligned} |y' - X_\nu| &\le |X(x_d; y; z) - X(y_d; y; z)| + |X(y_d; y; z) - X(y_d; y; z_\nu)| \\ &\lesssim \delta_\circ |x_d - y_d|. \end{aligned}$$

Now (6.26) easily follows.

We turn to (6.27). Taking a Taylor expansion in the x' variables,

$$\begin{aligned} \partial_{z'}\Psi(x;z) - \partial_{z'}\Psi(y;z) &= \partial_{z'}\Psi(x;z) - \partial_{z'}\Psi(X, x_d; z) \\ &= \partial^2_{z'x'}\Psi(X, x_d; z)(x' - X) + O(|x' - X|^2) \end{aligned} \tag{6.28}$$

whilst, by a Taylor expansion in the z-variables, the last expression is equal to

$$\partial^2_{z'x'}\Psi(X, x_d; z)(x' - X_\nu) + O\big(|x' - X_\nu|^2 + \varepsilon_\circ 2^{-m}\delta_\circ |x_d - y_d|\big). \tag{6.29}$$

Here the additional error term arises by applying the mean value theorem to $|X - X_\nu|$ together with (6.25).

On the other hand, one may write $\partial_{z_d}\Psi(x;z) - \partial_{z_d}\Psi(y;z) = \mathrm{I} + \mathrm{II}$ where

$$\mathrm{I} := \partial_{z_d}\Psi(X, x_d; z) - \partial_{z_d}\Psi(y;z), \quad \mathrm{II} := \partial_{z_d}\Psi(x;z) - \partial_{z_d}\Psi(X, x_d; z).$$

To estimate I, take a Taylor expansion first in the x_d variable and then in the z variable to obtain

$$\begin{aligned} \mathrm{I} &= \sigma(y;z)(x_d - y_d) + O(\delta_\circ |x_d - y_d|^2) \\ &= \sigma(y;z_\nu)(x_d - y_d) + O\big(\varepsilon_\circ 2^{-m}\delta_\circ |x_d - y_d|\big). \end{aligned} \tag{6.30}$$

Here σ appears owing to (6.23) and (6.21). The second estimate holds due to (6.8) and the localisation of the $(x, y; z)$-support of $b^m_{\mu\nu}$. To estimate the II term, arguing as in (6.28), take a Taylor expansion in the x' variable and then in the z variable to obtain

$$\begin{aligned} \mathrm{II} &= \partial^2_{z_d x'}\Psi(X, x_d; z)(x' - X) + O(|x' - X|^2) \\ &= \partial^2_{z_d x'}\Psi(X, x_d; z)(x' - X_\nu) + O\big(\varepsilon_\circ |x' - X_\nu| + \varepsilon_\circ 2^{-m}\delta_\circ |x_d - y_d|\big). \end{aligned} \tag{6.31}$$

In the last step we applied (6.25). From (6.29), (6.30) and (6.31) we get (assuming $\varepsilon_\circ$ is chosen sufficiently small) that

$$|\partial_{z'}\Psi(x;z) - \partial_{z'}\Psi(y;z)| \ \ge\ c_1 |x' - X_\nu| \quad \text{if } |x' - X_\nu| \ \ge\ C_1 \varepsilon_\circ 2^{-m}\delta_\circ |x_d - y_d|$$

and

$$|\partial_{z_d}\Psi(x;z) - \partial_{z_d}\Psi(y;z)| \ge (\delta_\circ/2) 2^{-m} |x_d - y_d| \qquad \text{if } |x' - X_\nu| \le C_1 \varepsilon_\circ 2^{-m}\delta_\circ |x_d - y_d|,$$

and these inequalities imply (6.27).

We now estimate $K^{\lambda,m}_{\mu\nu}(x,y)$. Using just the size and support of the integrand we get

$$|K^{\lambda,m}_{\mu\nu}(x,y)| \lesssim 2^{-md} \tag{6.32}$$

which we use for $|x'-X_\nu|+2^{-m}\delta_\circ|x_d-y_d| \le \lambda^{-1}$.

Now assume $|x'-X_\nu|+2^{-m}\delta_\circ|x_d-y_d| \ge \lambda^{-1}$; we use integration-by-parts to improve on (6.32). By (6.26), (6.27) we can apply Corollary A.2 with the choices $R(x,y) := 2^m$ and $\rho(x,y) := |x'-X_\nu(x_d,y;z_\nu)|+2^{-m}\delta_\circ|x_d-y_d|$. We also use that for fixed x,y the amplitude is supported in a set of diameter 2^{-m} and the estimates

$$|\partial_z^\alpha[b^m_{\mu\nu}(x,z)b^m_{\mu\nu}(y,z)]| \lesssim 2^{m|\alpha|}.$$

Altogether, Corollary A.2 yields, for $x \neq y$,

$$|K^{\lambda,m}_{\mu\nu}(x,y)| \lesssim 2^{-md}\lambda^{-N}2^{mN}\big(|x'-X_\nu|+\lambda 2^{-m}\delta_\circ|x_d-y_d|)\big)^{-N}.$$

Combining this with (6.32) we obtain

$$|K^{\lambda,m}_{\mu\nu}(x,y)| \lesssim 2^{-md}\big(1+\lambda 2^{-m}|x'-X_\nu|+\lambda 2^{-2m}\delta_\circ|x_d-y_d|)\big)^{-N}.$$

Fixing y and integrating in x yields

$$\int_{\mathbb{R}^d}|K^{\lambda,m}_{\mu\nu}(x,y)|\,\mathrm{d}x \lesssim 2^{-md}(2^m\lambda^{-1})^{d-1}2^{2m}\lambda^{-1}\delta_\circ^{-1} \lesssim 2^m\delta_\circ^{-1}\lambda^{-d},$$

which is the desired estimate for the first term in (6.19a). This finishes the proof of (6.19a) and the proof of the lemma. □

6.4. **Uniform estimates depending on a t-variable.** The estimates obtained in Propositions 6.1, 6.2 and 6.3 will be used to obtain L^2-bounds for the operators $A_j[\Phi_t;\mathfrak{a}_t]$. To this end, we shall allow a t-dependence in our operator and obtain uniform estimates in t. Consider now an open set $U \subset \mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d$, a phase function $\Psi : U \to \mathbb{R}$ and an amplitude $a \in C_0^\infty(U)$, and define

$$\Psi_t(x;z) = \Psi(x;t;z) \quad \text{ and } \quad a_t(x;z) = a(x;t;z). \tag{6.33}$$

Given $\lambda \ge 1$, let T_t^λ denote the oscillatory integral associated to the pair $[\Psi_t;a_t]$ as in (6.1), given by $T_t^\lambda \equiv T^\lambda[\Psi_t;a_t]$. For $0<\delta_\circ\le 1$, we assume that the condition (6.2) continues to hold under t-derivatives. That is, the estimates

$$|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\Psi_t(x;z)|+\delta_\circ^{-1}|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\partial_{x_d}\Psi_t(x;z)| \le C_{\alpha,\beta,\gamma} \tag{6.34}$$

hold for all $(x;t;z) \in U$ and all $\alpha,\beta \in \mathbb{N}_0^d$, $\gamma \in \mathbb{N}_0$. Thus, if the condition $|\det\partial^2_{zx}\Psi_{t_0}(x_0;z_0)| \ge c\delta_\circ$ holds for some $(x_0;t_0;z_0) \in U$, Proposition 6.1 in conjunction with (6.34) immediately extends to a uniform estimate for the operators T_t^λ for all $|t-t_0| \le \varepsilon_\circ$, for suitable $\varepsilon_\circ$. Likewise if (6.34) holds and the conditions (6.4a), (6.4b) and (6.4c) are satisfied at a certain $(x_0;t_0;z_0) \in U$, Propositions 6.2 and 6.3 also extend to the operators T_t^λ for all $|t-t_0| \le \varepsilon_\circ$, with uniform bounds on t; note that (6.34) implies that the quantity $\sigma_t(x;z) \equiv \sigma(x;t;z)$ defined as in (6.5), also satisfies the derivative bounds (6.8) under t-differentiation, that is,

$$|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\sigma_t(x;z)| \lesssim_{\alpha,\beta,\gamma} \delta_\circ \tag{6.35}$$

holds for all $(x;t;z) \in U$ and all $\alpha,\beta \in \mathbb{N}_0^d$, $\gamma \in \mathbb{N}_0$.

6.5. **Estimates for maximal oscillatory integrals.** We now state the version of the estimates in Propositions 6.1 and 6.2 for the maximal functions associated to the oscillatory integral operators T_t^λ.

To obtain such maximal estimates we will assume that (6.34) holds and that, in addition, there is $\delta_\circ$-smallness when we differentiate with respect to the t-variable; more precisely we assume that

$$|\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \Psi_t(x;z)| \lesssim_\gamma \delta_\circ \tag{6.36}$$

holds for all $(x;t;z) \in U$ and all $\alpha, \beta \in \mathbb{N}_0^d$, $\gamma > 0$.

Proposition 6.7. *Let* $[\Psi;\, a]$ *be as in* (6.33)*. Suppose* Ψ *satisfies* (6.34)*,* (6.36) *and* $|\det \partial_{zx}^2 \Psi_{t_0}(x_0; z_0)| \geq c\delta_\circ$ *for some* $(x_0; t_0; z_0) \in U$*. Then there is* $\varepsilon_\circ > 0$ *and* $N > 0$ *such that, under the assumption of* a_t *supported in* $B_{\varepsilon_\circ}(x_0, z_0)$*,*

$$\big\| \sup_{|t-t_0| \leq \varepsilon_\circ} |T^\lambda[\Psi_t; a_t]| \big\|_{L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \|a\|_{C^N}.$$

Proposition 6.8. *Let* $[\Psi;\, a]$ *be as in* (6.33)*. Assume that* Ψ *satisfies* (6.4a)*,* (6.4b) (6.4c) *at a certain* $(x_0; t_0; z_0) \in U$*, the estimates* (6.34) *and* (6.36) *and, in addition, assume that*

$$\partial_t \Psi_t(x;z) = c_t(x;z) \det \partial_{zx}^2 \Psi_t(x;z) \tag{6.37}$$

for some $c \in C^\infty$ *in a neighbourhood of* $\operatorname{supp} a$*. Then there is* $\varepsilon_\circ > 0$ *and* $N > 0$ *such that, under the assumption of* a_t *supported in* $B_{\varepsilon_\circ}(x_0, z_0)$

$$\big\| \sup_{|t-t_0| \leq \varepsilon_\circ} |T^\lambda[\Psi_t; a_t]| \big\|_{L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \log(2 + \lambda\delta_\circ) \|a\|_{C^N}.$$

The proofs rely on a standard Sobolev embedding inequality (see for instance [32, Chapter XI, §3.2]). Namely, for a C^1 function $t \mapsto g(t)$ supported on an interval I, with $t_0 \in I$, we have, for $1 \leq p < \infty$,

$$\sup_{t \in I} |g(t)|^p \leq |g(t_0)|^p + p\|g\|_{L^p(I)}^{p-1} \|g'\|_p \tag{6.38}$$

which follows by the fundamental theorem of calculus applied to $|g|^p$ and Hölder's inequality. We can apply this to $F(x,t)$ with $F \in L^p(\mathbb{R}^d; C^1)$, and after integrating in x and another application of Hölder's inequality, (6.38) gives

$$\|\sup_{t \in I} |F(\cdot,t)|\|_{L^p(\mathbb{R}^d)}^p \leq \inf_{t_0 \in I} \|F(\cdot,t_0)\|_{L^p(\mathbb{R}^d)}^p + p\|F\|_{L^p(\mathbb{R}^d \times I)}^{p-1} \|\partial_t F\|_{L^p(\mathbb{R}^d \times I)}. \tag{6.39}$$

Proof of Proposition 6.7. Note that if $T_t^\lambda := T^\lambda[\Psi_t; a_t]$, then $\partial_t T_t^\lambda = T^\lambda[\Psi_t; d_t]$, where $d_t := (i\lambda \partial_t \Psi_t) a_t + \partial_t a_t$. By (6.36) one has $\|d_t\|_{C^N} \lesssim (1 + \lambda\delta_\circ)\|a_t\|_{C^{N+1}}$. Thus, by the hypothesis and Proposition 6.1 applied to T_t^λ and $\partial_t T_t^\lambda$ (as discussed in §6.4), there exist $\varepsilon_\circ$ and $N > 0$ such that, if a_t is supported in $B_{\varepsilon_\circ}(x_0; z_0)$, the bounds

$$(1 + \lambda\delta_\circ)^{1/2} \|T_t^\lambda f\|_{L^2(\mathbb{R}^d)} + (1 + \lambda\delta_\circ)^{-1/2} \|\partial_t T_t^\lambda f\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \|a\|_{C^{N+1}} \|f\|_{L^2(\mathbb{R}^d)}$$

hold uniformly in $|t - t_0| \leq \varepsilon_\circ$. Now the assertion follows immediately by the Sobolev inequality (6.39) for the exponent $p = 2$. $\square$

Proof of Proposition 6.8. Given $0 \le m \le M$, let

$$T_t^{\lambda,m} f(x) := \int_{\mathbb{R}^d} e^{i\lambda \Psi_t(x;z)} a_t(x;z) \beta(2^m \delta_\circ^{-1} |\sigma_t(x;z)|) f(z)\, \mathrm{d}z, \quad 0 \le m < M,$$
$$T_t^{\lambda,M} f(x) := \int_{\mathbb{R}^d} e^{i\lambda \Psi_t(x;z)} a_t(x;z) \eta(2^M \delta_\circ^{-1} \sigma_t(x;z)) f(z)\, \mathrm{d}z;$$

that is, (6.10) and (6.11) with $\beta(|\cdot|)$, η in place of η_1, η_0 and with the phase-amplitude pair $[\Psi_t; a_t]$.

Using (6.37) and (6.6) we compute

$$\begin{aligned}\partial_t T_t^{\lambda,m} f(x) &= \lambda\delta_\circ 2^{-m} \int_{\mathbb{R}^d} e^{i\lambda\Psi_t(x;z)} \widetilde{c}_t(x;z) a_t(x;z) \eta_1(2^m \delta_\circ^{-1} \sigma_t(x;z)) f(z)\, \mathrm{d}z \\ &\quad + 2^m \int_{\mathbb{R}^d} e^{i\lambda\Psi_t(x;z)} 2^{-m} \partial_t a_t(x;z) \beta(2^m \delta_\circ^{-1} |\sigma_t(x;z)|) f(z)\, \mathrm{d}z \\ &\quad + 2^m \int_{\mathbb{R}^d} e^{i\lambda\Psi_t(x;z)} \delta_\circ^{-1} \partial_t \sigma_t(x;z) a_t(x;z) \widetilde{\eta}_1(2^m \delta_\circ^{-1} \sigma_t(x;z)) f(z)\, \mathrm{d}z\end{aligned}$$

where $\widetilde{c}_t = c_t \partial^2_{x'z'} \Psi_t$, is smooth and $\eta_1(s) = s\beta(|s|)$, $\widetilde{\eta}_1(s) = \frac{d}{ds}(\beta(|s|))$. For $m = M$ we have a similar formula with β replaced by η. Note that in view of (6.35) we have $|\partial_x^\alpha \partial_z^\beta [\delta_0^{-1} \partial_t \sigma_t a_t]| \lesssim C_{\alpha,\beta}$.

Assume $1 \le 2^m \le (\lambda\delta_\circ)^{1/3}$. By the hypothesis and Proposition 6.3 applied to $T_t^{\lambda,m}$ and $\partial_t T_t^{\lambda,m}$, there exist $\varepsilon_\circ$ and $N > 0$ such that, if a_t is supported in $B_{\varepsilon_\circ}(x_0; z_0)$, one has the bounds

$$\|T_t^{\lambda,m} f\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \Big(\frac{2^m}{\lambda\delta_\circ}\Big)^{1/2} \|f\|_{L^2(\mathbb{R}^d)}$$

and

$$\begin{aligned}\|\partial_t T_t^{\lambda,m} f\|_{L^2(\mathbb{R}^d)} &\lesssim \lambda^{-\frac{d-1}{2}} \Big(\frac{2^m}{\lambda\delta_\circ}\Big)^{1/2} \big(\lambda\delta_\circ 2^{-m} + 2^m\big) \|f\|_{L^2(\mathbb{R}^d)} \\ &\lesssim \lambda^{-\frac{d-1}{2}} \Big(\frac{2^m}{\lambda\delta_\circ}\Big)^{-1/2} \|f\|_{L^2(\mathbb{R}^d)}\end{aligned}$$

uniformly in $|t - t_0| \le \varepsilon_\circ$, where the last inequality follows because we are under the assumption $1 \le 2^m \le (\lambda\delta_\circ)^{1/3} \le (\lambda\delta_\circ)^{1/2}$. Therefore, the above estimates combined with (6.39) yield

$$\sum_{0 \le m \le \lfloor \log_2 (\lambda\delta_\circ)^{1/3} \rfloor} \big\| \sup_{|t-t_0| \le \varepsilon_\circ} |T_t^{\lambda,m} f| \big\|_{L^2(\mathbb{R}^d)} \lesssim \log(2 + \lambda\delta_\circ) \lambda^{-\frac{d-1}{2}} \|f\|_{L^2(\mathbb{R}^d)}.$$

Similarly, if $\lambda^{1/2} \ge 2^m \ge \min\{(\lambda\delta_\circ)^{1/3}, 1\}$, Proposition 6.3 implies

$$\|T_t^{\lambda,m} f\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} 2^{-m} \|f\|_{L^2(\mathbb{R}^d)}$$

and

$$\|\partial_t T_t^{\lambda,m} f\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} 2^{-m} \big(\lambda\delta_\circ 2^{-m} + 2^m\big) \|f\|_{L^2(\mathbb{R}^d)}$$

uniformly in $|t - t_0| \le \varepsilon_\circ$. The above bounds imply, by (6.39), that

$$\big\| \sup_{|t-t_0| \le \varepsilon_\circ} |T_t^{\lambda,m} f| \big\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} 2^{-m} (\lambda\delta_\circ 2^{-m} + 2^m)^{1/2} \|f\|_{L^2(\mathbb{R}^d)},$$

and thus

$$\sum_{\lfloor \log_2(\lambda\delta_\circ)^{1/3}\rfloor \wedge 1 \le m \le M} \Big\| \sup_{|t-t_0|\le\varepsilon_\circ} |T_t^{\lambda,m} f| \Big\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-\frac{d-1}{2}} \|f\|_{L^2(\mathbb{R}^d)}$$

follows from summing a geometric series, as $\lambda\delta_\circ 2^{-m} \le 2^{2m}$ in the range of summation. Combining both sums one obtains the desired bound by the triangle inequality, which concludes the proof of the proposition. □

6.6. **Radon-type operators in d dimensions versus oscillatory integral operators in $d+1$ dimensions.** In this section we use variables $(x;z) \in \mathbb{R}^{d+1}\times\mathbb{R}^{d+1}$ and split $x = (x_1, x'')$, $z = (z_1, z'')$ with $x'' \in \mathbb{R}^d$, $z'' \in \mathbb{R}^d$. Recall that the frequency localised Radon-type operators in (5.18) are of the form (with $d = 2$)

$$\begin{aligned} A_j[\Phi_t; \mathfrak{a}_t] f(x'') &= \int_{\mathbb{R}^d} \mathfrak{a}_t(x''; z'') \int_{\mathbb{R}} \beta(2^{-j}|\theta|) e^{i\theta\Phi_t(x'';z'')} f(z'')\,\mathrm{d}\theta\,\mathrm{d}z'' \\ &= 2^j \int_{\mathbb{R}\times\mathbb{R}^d} \mathfrak{a}_t(x''; z'')\beta(|\omega|) e^{i2^j\omega\Phi_t(x'';z'')} f(z'')\,\mathrm{d}\omega\,\mathrm{d}z'', \qquad (6.40)\end{aligned}$$

We rely on an idea in [32, Chapter XI, §3.2.1] to show that a $L^p(\mathbb{R}^d)$ estimate for $\sup_{t\in I} |A_j[\Phi_t; a_t]f|$ is implied by a L^p-estimate for a maximal function associated with a closely related family of oscillatory integral operators acting on functions on $\mathbb{R}^{d+1}$ which we will presently define.

Recall that β is supported in $[1/2, 2]$. Let $\widetilde{\beta}$ be supported in $(1/4, 4)$ such that also $\widetilde{\beta}(s) = 1$ for $s \in [1/3, 3]$. Notice that $\widetilde{\beta}(s)\beta(us) = \beta(us)$ for $2/3 < u < 3/2$. Now let $\chi_1 \in C_0^\infty(\mathbb{R})$ so that $\chi_1(r) = 1$ on $J := [2/3, 3/2]$. Consider the family of oscillatory integral operators $T^{2^j}[\phi_t; a_t]$, as defined in (6.1) but acting on functions g on $\mathbb{R}^{d+1}$, where

$$\phi_t(x; z) = x_1 z_1 \Phi_t(x''; z''), \quad \text{and} \quad a_t(x; z) = \chi_1(x_1) x_1 \mathfrak{a}_t(x''; z'')\beta(x_1|z_1|). \quad (6.41)$$

Lemma 6.9. *Let $E \subset (0,\infty)$, Φ, ϕ, $\mathfrak{a}$, a as in* (6.41)*, and define*

$$M_j[\Phi; \mathfrak{a}]f := \sup_{t\in E} |A_j[\Phi_t; \mathfrak{a}_t]f|, \qquad \mathcal{M}_j[\phi; a]g = \sup_{t\in E} |T^{2^j}[\phi_t; a_t]g|.$$

Then

$$\|M_j[\Phi; \mathfrak{a}]\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \le 2^j (6/5)^{1/p} \|\widetilde{\beta}\|_{L^p(\mathbb{R})} \|\mathcal{M}_j[\phi; a]\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})}.$$

Proof. For fixed x_1 we change variables $\omega = x_1 z_1$ in (6.40). We use that $\chi(x_1) = 1$ for $x_1 \in J$ and that $\widetilde{\beta}(|z_1|)\beta(x_1|z_1|) = \beta(x_1|z_1|)$ for $(x_1, z_1) \in J \times \mathbb{R}$ to obtain the identity

$$A_j[\Phi_t; \mathfrak{a}_t] f(x'') = 2^j T^{2^j}[\phi_t; a_t](\widetilde{\beta}\otimes f)(x_1, x'') \quad \text{for all } x_1 \in J.$$

This identity implies that

$$\begin{aligned} 2^{-j}\|M_j[\Phi; \mathfrak{a}]f|\|_{L^p(\mathbb{R}^d)} &\le |J|^{-1/p} \|\mathcal{M}_j[\phi; a](\widetilde{\beta}\otimes f)\|_{L^p(J\times\mathbb{R}^d)} \\ &\le (3/2 - 2/3)^{1/p} \|\mathcal{M}_j[\phi; a](\widetilde{\beta}\otimes f)\|_{L^p(\mathbb{R}^{d+1})} \\ &\le (6/5)^{1/p} \|\mathcal{M}_j[\phi; a]\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})} \|\widetilde{\beta}\|_{L^p(\mathbb{R})} \|f\|_{L^p(\mathbb{R}^d)} \end{aligned}$$

which implies the assertion. □

7. Proof of Proposition 5.11: L^2 bounds

In this section we apply the maximal function results in §6 to deduce favourable L^2 bounds which will feature in the proof of Proposition 5.11.

Proposition 7.1. *For all $m < 0$, $k \geq -4$, $(k,\ell) \in \mathfrak{P}$ and $j \geq -e(k,\ell)/3$, the following bounds hold, where in each inequality I denotes an interval of length $|I| \sim 1$ containing the t-support of the amplitude.*

i) $\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{b}_t]| \big\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim (j\vee 1)2^{-2k/3}\|\mathfrak{b}\|_{C^N}$ *if* $[\Phi;\mathfrak{b}] \in \mathfrak{B}^k_{\mathrm{Rot}}$;

ii) $\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{a}_t]| \big\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim 2^{-e(k,\ell)/3}\|\mathfrak{a}\|_{C^N}$ *if* $[\Phi;\mathfrak{a}] \in \mathfrak{A}^{k,\ell}_{\mathrm{Rot}}$;

iii) $\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{c}_t]| \big\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim 2^{-2k/3}\|\mathfrak{c}\|_{C^N}$ *if* $[\Phi;\mathfrak{c}] \in \mathfrak{C}^k_{\mathrm{Rot}}$;

iv) $\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{a}_t]| \big\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim \|\mathfrak{a}\|_{C^N}$ *if* $[\Phi;\mathfrak{a}] \in \mathfrak{A}^0_{\mathrm{Rot}}$;

v) $\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{a}_t]| \big\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim 2^{m/2}\sup_{x_2}\|\mathfrak{a}\|_{C^N_{x_1,z,t}}$, *if* $[\Phi;\mathfrak{a}] \in \mathfrak{A}^m_{\mathrm{Rot}}$.

As in Section 5, the cases i), iii), iv) and v) are understood to hold for $\ell = 2k$, with $k = 0$ in the cases iv) and v).

The proof of Proposition 7.1 is presented in what follows. Observe that, by the definition of the classes, iii) and iv) are both just special cases of ii). Thus, it will suffice to prove i) , ii) and v) only.

Remark. Only rotational curvature considerations are required to establish the above L^2 bounds. The cinematic curvature is used in §8 to deduce local smoothing estimates in order to obtain summable bounds in the j parameter.

Using Lemma 6.9 the estimates in Proposition 7.1 may be deduced from estimates on oscillatory integral operators acting on functions in $\mathbb{R}^3$; in particular, our assumptions on the phase/amplitude pairs allow direct applications of Propositions 6.7 and 6.8 with suitable choices of the parameters λ and $\delta_\circ$.

7.1. *Proof of Proposition 7.1 (i).*

By Lemma 6.9, it suffices to show that

$$\| \sup_{t\in I} |T^{2^j}[\phi_t;b_t]| \|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim 2^{-j}(j\vee 1)2^{-2k/3}\|\mathfrak{b}\|_{C^N},$$

where $\phi_t(x;z) = x_1 z_1 \Phi(x'',t;z'')$ and $b_t(x;z) = \chi(x_1)x_1\mathfrak{b}_t(x'';z'')\beta(x_1|z_1|)$.

First we use the fold conditions, inherent in the hypotheses $\mathrm{F}_1)_k$ and $\mathrm{F}_2)_k$ in the definition of $\mathfrak{B}^k_{\mathrm{Rot}}$, to place the operator in a normal form. By assumption $\mathrm{b})_k$, one may assume without loss of generality, decomposing b_t into at most $O(1)$ pieces, that $\operatorname{supp}\mathfrak{b}$ is contained in an $\varepsilon_\circ$-ball centred at some point $(x_0'';t_0;z_0'')$ with $(x_0'';z_0'') \in \mathcal{Z}_{t_0}$. Here $\mathcal{Z}_{t_0}$ is as defined in (5.17). Fix a pair of 3×3 matrices $\mathbf{X}$ and $\mathbf{Z}$ satisfying the properties enumerated in property $\mathrm{F}_2)_k$. Since $|\det\mathbf{X}| \sim |\det\mathbf{Z}| \sim 1$, by a change of variables it suffices to show the L^2 bound for the maximal function $\sup_{|t-t_0|<\varepsilon_\circ} |T^{2^j}[\tilde{\phi}_t;\tilde{b}_t]f(x)|$ in $\mathbb{R}^3$, where

$$\tilde{\phi}_t(x;z) := \phi_t(\mathbf{X}x;\mathbf{Z}z), \quad \tilde{b}_t(x;z) := b_t(\mathbf{X}x;\mathbf{Z}z).$$

Now the assumption $[\Phi;\mathfrak{b}] \in \mathfrak{B}^k_{\mathrm{Rot}}$ implies that the support of $\tilde{b}_t$ is contained in a $\varepsilon_\circ$-ball centred at $(x_0,t_0;z_0) = (0,x_0'',t_0;1,z_0'') \in \mathbb{R}^3\times\mathbb{R}\times\mathbb{R}^3$; moreover we have the following conditions on the derivatives of $\tilde{\phi}$:

$$|\partial_x^\alpha \partial_z^\beta \tilde{\phi}_t(x;z)| \lesssim \begin{cases} 2^{-4k/3} & \text{if } \alpha_3 \neq 0, \\ 2^{2k/3} & \text{otherwise,} \end{cases} \tag{7.1a}$$

$$\partial^2_{x_3 z}\tilde{\phi}_{t_0}(x_0;z_0) = \partial^2_{x z_3}\tilde{\phi}_{t_0}(z_0;z_0) = 0 \text{ and } |\det \partial^2_{x'z'}\tilde{\phi}_{t_0}(x_0;z_0)| \sim 2^{4k/3}, \tag{7.1b}$$

$$|\partial^3_{x_3 z_3 z_3}\tilde{\phi}_{t_0}(x_0;z_0)|, |\partial^3_{x_3 x_3 z_3}\tilde{\phi}_{t_0}(x_0;z_0)| \sim 2^{-4k/3}, \tag{7.1c}$$

$$|\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \tilde{\phi}_t(x;z)| \lesssim 2^{-4k/3} \text{ for } \gamma > 0, \tag{7.1d}$$

$$\partial_t \tilde{\phi}_t(x;z) = \tilde{c}(x,t;z) 2^{-4k/3} \det \partial^2_{xz}\tilde{\phi}_t(x;z) \text{ for some } \tilde{c} \in C^\infty \text{ with} \tag{7.1e}$$

uniform C^∞ bounds on supp $\tilde{b}_t$.

The following table shows which conditions for the class $\mathfrak{B}^k_{\mathrm{rot}}$ of defining functions imply the conditions in (7.1).

(7.1a)	$\Phi_1)_{k,2k}$ and $\mathrm{F}_2)_k$ i)	(7.1d)	$\Phi_1)_{k,2k}$ and $\mathrm{F}_2)_k$ i)
(7.1b)	$\mathrm{F}_2)_k$ iii)	(7.1e)	$\Phi_2)_{k,2k}$ and $\mathrm{F}_2)_k$ ii)
(7.1c)	$\mathrm{F}_1)_k$ and $\mathrm{F}_2)_k$ ii)		

One now checks that the phase function

$$\Psi_t(x;z) = 2^{-2k/3}\tilde{\phi}_t(x;z)$$

satisfies the assumptions in Proposition 6.8 with $d = 3$ and $\delta_\circ = 2^{-2k}$ via (7.1). If we put $\lambda = 2^{j+2k/3}$, then $\lambda\Psi = 2^j\,\tilde{\phi}$ and we can apply Proposition 6.8 to obtain

$$\|\sup_{t\in I}|T^{2^j}[\phi_t;b_t]|\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim \lambda^{-1}\log(2+\lambda\delta_\circ)\|b\|_{C^N} \lesssim 2^{-j-2k/3}(j\vee 1)\|\mathfrak{b}\|_{C^N},$$

as desired □.

7.2. *Proof of Proposition 7.1 (ii).* We again use the reduction in §6.6 so that it suffices to show

$$\|\sup_{t\in I}|T^{2^j}[\phi_t;a_t]|\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim 2^{-j}2^{-e(k,\ell)/3}\|\mathfrak{a}\|_{C^N},$$

where $\phi_t(x;z) = x_1 z_1 \Phi(x'',t;z'')$ and $a_t(x;z) = \chi(x_1)x_1\mathfrak{a}_t(x'';z'')\beta(x_1|z_1|)$. The condition $[\Phi;\mathfrak{a}] \in \mathfrak{A}^{k,\ell}_{\mathrm{Rot}}$ implies that the phase function $\phi_t(x;z) = x_1 z_1 \Phi_t(x'';z'')$ satisfies the inequalities

$$|\partial_x^\alpha \partial_z^\beta \phi_t(x;z)| \lesssim \begin{cases} 2^{-2e(k,\ell)/3} & \text{if } \alpha_3 \neq 0 \\ 2^{e(k,\ell)/3} & \text{otherwise} \end{cases} \tag{7.2}$$

$$|\det \partial^2_{xz}\phi_t(x;z)| \sim 1 \tag{7.3}$$

$$|\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \phi_t(x;z)| \lesssim 2^{-2e(k,\ell)/3} \quad \text{for } \gamma > 0. \tag{7.4}$$

These estimates are understood to hold on $\operatorname{supp} a_t$ (which has diameter $\lesssim 1$) for all $\alpha, \beta \in \mathbb{N}_0^3$, $\gamma \in \mathbb{N}$ with implicit constants depending on the multiindices. One checks that (7.2) and (7.4) are implied by $\Phi_1)_{k,\ell}$ in the definition of $\mathfrak{A}^{k,\ell}$ while (7.3) is implied by the additional rotational curvature condition in Definition 5.8.

We can now verify that the phase function

$$\Psi_t(x;z) = 2^{-e(k,\ell)/3}\phi_t(x;z)$$

satisfies the assumptions in Proposition 6.7 with $d = 3$ and $\delta_\circ = 2^{-e(k,\ell)}$. If we put $\lambda = 2^{j+e(k,\ell)/3}$, then $\lambda\Psi = 2^j\phi$ and by Proposition 6.7 we get

$$\|\sup_{t\in I}|T^{2^j}[\phi_t; a_t]|\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim \lambda^{-1}\|a\|_{C^N} \lesssim 2^{-j-e(k,\ell)/3}\|\mathfrak{a}\|_{C^N},$$

as desired. □.

7.3. *Proof of Proposition 7.1 (v).* Again, by Lemma 6.9, it suffices to show that

$$\|\sup_{t\in I}|T^{2^j}[\Psi_t; a_t]|\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim 2^{-j}2^{m/2}\sup_{x_3}\|a\|_{C^N_{x_1,x_2,z,t}}, \tag{7.5}$$

where $\Psi(x;t;z) = x_1z_1\Phi(x'',t;z'')$ and $a_t(x;z) = \chi(x_1)x_1\mathfrak{a}_t(x'';z'')\beta(x_1|z_1|)$. By the condition $[\Phi;\mathfrak{a}] \in \mathfrak{A}^m_{\mathrm{Rot}}$, the diameter of the support of a is $O(1)$ and moreover the following conditions hold (see Definitions 5.3 and 5.10). First, there exists an interval I_m of length $\lesssim 2^m$ so that $a(x;t;z) = 0$ when $x_3 \notin I_m$. Next, if $\Psi^\star_{x_3}(x_1,x_2,t;z) := x_1z_1\Phi(x'',t;z'')$ then $\Psi^\star_{x_3}$ satisfies

$$|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\Psi^\star_{x_3}(x_1,x_2,t;z)| \lesssim \begin{cases} 2^{-2m} & \text{if } \alpha_3 \neq 0 \\ 1 & \text{otherwise} \end{cases}, \tag{7.6a}$$

$$|\det\partial^2_{(x_1,x_2,t),(z_1,z_2,z_3)}\Psi^\star_{x_3}| \sim 1. \tag{7.6b}$$

In what follows we will freeze x_3, so the derivatives with respect to x_3 in (7.6a) will be irrelevant for our purposes.

To establish (7.5) we show that if

$$S^{2^j}_{x_3}f(x_1,x_2,t) \equiv T^{2^j}[\Psi^\star_{x_3}; a^\star_{x_3}]f(x_1,x_2,t) \equiv T^{2^j}[\Psi_t; a_t]f(x_1,x_2,x_3),$$

where $a^\star_{x_3}(x_1,x_2,t;z) = a_t(x;z)$, then we have, for all $x_3 \in I_m$,

$$\Big(\int_{\mathbb{R}^3}|S^{2^j}_{x_3}f(x_1,x_2,t)|^2\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}t\Big)^{1/2} + 2^{-j}\Big(\int|\partial_tS^{2^j}_{x_3}f(x_1,x_2,t)|^2\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}t\Big)^{1/2} \lesssim 2^{-3j/2}\sup_{x_3}\|a\|_{C^N_{x_1,x_2,z,t}}\|f\|_2. \tag{7.7}$$

Indeed, note that $\partial_tS^{2^j}_{x_3}f(x_1,x_2,t) = T^{2^j}[\Psi^\star_{x_3}; d^\star_{x_3}]f(x_1,x_2,t)$, where

$$d^\star_{x_3} := (i2^j\partial_t\Psi^\star_{x_3})\,a + \partial_ta^\star_{x_3},$$

and, in view of (7.6a) and (7.6b), the estimate (7.7) is now an immediate consequence of the oscillatory integral estimate in Proposition 6.1 with $\delta_\circ = 1$, which holds uniformly in $x_3 \in I_m$. Note that, in this case, our application of Proposition 6.1 corresponds to the classical Hrmander L^2-estimate for oscillatory integrals [13]. Integrating the square of the left-hand side of (7.7) over $x_3 \in I_m$ and using $|I_m| \lesssim 2^m$, we get

$$\Big(\int_{I_m}\int_{\mathbb{R}^3}|T^{2^j}[\Psi_t; a_t]f(x)|^2 + 2^{-2j}|\partial_tT^{2^j}[\Psi_t; a_t]f(x)|^2\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}t\,\mathrm{d}x_3\Big)^{1/2} \lesssim 2^{-3j/2}2^{m/2}\sup_{x_3}\|a\|_{C^N_{x_1,x_2,z,t}}\|f\|_2.$$

By the Sobolev inequality (6.39) and Fubini's theorem, the desired estimate (7.5) immediately follows. □

8. Proof of Proposition 5.11: L^p theory

This section deals with the remainder of the proof of Proposition 5.11. Local space-time L^p estimates are used to establish L^p bounds with favourable j dependence when $p > 2$. These bounds can be combined with the L^2 estimates from Proposition 7.1 and L^∞ estimates to yield the desired results.

8.1. **L^p bounds.** It is first noted that the L^2 bounds of the previous section imply L^p estimates via interpolation with straightforward L^∞ bounds.

Corollary 8.1. *For all $m < 0$, $(k,\ell) \in \mathfrak{P}$, $j \geq -e(k,\ell)/3$ and $2 \leq p \leq \infty$, there exists $N \in \mathbb{N}$ such that the following bounds hold. In each inequality, I denotes an interval of length ~ 1 containing the t-support of the amplitude.*

$$\begin{aligned}
&i) && \big\| \sup_{t\in I} |A_j[\Phi_t; \mathfrak{b}_t]| \big\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim (j \vee 1)^{2/p} 2^{-2k/3} \|\mathfrak{b}\|_{C^N} && \text{if } [\Phi;\mathfrak{b}] \in \mathfrak{B}^k_{\mathrm{Rot}},\\
&ii) && \big\| \sup_{t\in I} |A_j[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim 2^{-e(k,\ell)/3} \|\mathfrak{a}\|_{C^N} && \text{if } [\Phi;\mathfrak{a}] \in \mathfrak{A}^{k,\ell}_{\mathrm{Rot}},\\
&iii) && \big\| \sup_{t\in I} |A_j[\Phi_t; \mathfrak{c}_t]| \big\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim 2^{-2k/3} \|\mathfrak{c}\|_{C^N} && \text{if } [\Phi;\mathfrak{a}] \in \mathfrak{C}^k_{\mathrm{Rot}},\\
&iv) && \big\| \sup_{t\in I} |A_j[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim \|\mathfrak{a}\|_{C^N} && \text{if } [\Phi;\mathfrak{a}] \in \mathfrak{A}^0_{\mathrm{Rot}},\\
&v) && \big\| \sup_{t\in I} |A_j[\Phi_t; \mathfrak{a}_t]| \big\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim 2^{m/p} \sup_{x_2} \|\mathfrak{a}\|_{C^N_{x_1,z,t}}, && \text{if } [\Phi;\mathfrak{a}] \in \mathfrak{A}^m_{\mathrm{Rot}}.
\end{aligned}$$

Remark. The estimates from Corollary 8.1 are not summable in the j parameter, so alone they do not imply Proposition 5.11. However, i), ii) and iii) have better k dependence than what is required in Proposition 5.11 (by a factor of $2^{(1-\frac{2}{p})k - k\varepsilon_p}$) and, similarly, v) has a better m dependence (by a factor of $2^{m/p - m\varepsilon_p}$). This observation is used below to mitigate losses in k and m in Proposition 8.2.

Proof of Corollary 8.1. We will only consider i) since the proofs of the remaining cases are similar. For $p = 2$ the desired bound is precisely Proposition 7.1 i). By interpolation, it suffices to verify the bound for $p = \infty$.

Let $[\Phi;\mathfrak{b}] \in \mathfrak{B}^k_{\mathrm{Rot}}$ and recall from (5.18) that

$$A_j[\Phi_t;\mathfrak{b}_t]f(x) = \int_{\mathbb{R}^2} f(z)\mathfrak{b}_t(x;z) 2^j \check{\beta}\big(2^j \Phi_t(x;z)\big)\mathrm{d}z.$$

Further recall that Φ_t satisfies Definition 5.7 and, in particular, the condition $\Phi_1)_{k,2k}$ as stated in Definition 5.2. Thus, on the support of $\mathfrak{b}_t$ we have

$$|\partial_z \Phi_t(x;z)| \gtrsim 2^{2k/3}$$

and so the desired L^∞ estimate follows. □

The following proposition provides the crucial j summability for $j > 0$.

Proposition 8.2. *There exist $N, M \in \mathbb{N}$ and $\varepsilon_\circ > 0$ such that for all $(k,\ell) \in \mathfrak{P}$, the inequality*

$$\big\| \sup_{t\in I} |A_j[\Phi_t;\mathfrak{a}_t]| \big\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^2)} \lesssim 2^{Mk} 2^{-j\varepsilon_\circ} \|\mathfrak{a}\|_{C^N} \tag{8.1}$$

holds if $[\Phi;\mathfrak{a}]$ belongs to any one of the following classes:

i) $\mathfrak{B}^k_{\mathrm{Cin}}$, *iv)* $\mathfrak{A}^0_{\mathrm{Cin}}$ *taking $k = 0$ in (8.1),*
ii) $\mathfrak{A}^{k,\ell}_{\mathrm{Cin}}$, *v)* $\mathfrak{A}^m_{\mathrm{Cin}}$, *$m < 0$, taking $k = -m$ in (8.1).*
iii) $\mathfrak{C}^k_{\mathrm{Cin}} \cap \mathfrak{C}^k_{\mathrm{Rot}}$.

In (8.1)*, I denotes an interval of length* ~ 1 *containing the t-support of* $\mathfrak{a}$*.*

Remark. The exponent $p = 6$ does not play a significant rôle and is used merely for convenience (one could equally work with other p values). See the comments after Theorem 8.5 below.

Assuming this result, Proposition 5.11 easily follows by interpolation with the estimates in Corollary 8.1.

Proof of Proposition 5.11 assuming Proposition 8.2 holds. For $-e(k,\ell)/3 \leq j \leq 0$ the asserted bounds are an immediate consequence of Corollary 8.1. For $j > 0$ it suffices, by Corollary 8.1, to show each of the five estimates in Proposition 5.11 hold *for a single value* $2 < p_* < \infty$: indeed, once this is established, one may interpolate the p_* estimates with the $p = 2$ and $p = \infty$ cases of Corollary 8.1 to obtain Proposition 5.11 for all $2 < p < \infty$.

We interpolate the inequalities from Proposition 8.2 with the corresponding L^∞ estimates of Corollary 8.1, or the L^2 estimate in case v). In the case v), note that $\|\mathfrak{a}\|_{C^N} \lesssim 2^{-O(N)m}$, which is harmless in view of the 2^{-Mm} loss in (8.1). Therefore, it follows that Proposition 5.11 holds for some p_* in the range $6 < p_* < \infty$ for the cases i) to iv), or in the range $2 < p_* < 6$ for case v), concluding the proof. □

It remains to prove Proposition 8.2. By the definition of the classes, Proposition 8.2 i) and iv) automatically follow from ii). Furthermore, for the purposes of the argument, the cases ii) and v) are essentially simplified variants of case iii). In particular, the main difficulties occur in the proof of iii).

8.2. **Reduction to Fourier integral estimates.** Following the strategy of [19, 20], Proposition 8.2 is derived from local smoothing estimates for Fourier integral operators. In order to invoke the local smoothing inequalities, it desirable to express $A_j[\Phi_t;\mathfrak{a}_t]$ as a Fourier integral operator with two Fourier variables. That such a representation is possible is a standard result, referred to as the *equivalence of phase theorem* (see, for instance, [12] or [9]). Since here, however, the estimates are required to be quantitative, at least in some weak sense, basic stationary phase techniques are instead applied to obtain an explicit two Fourier variable representation of the frequency localised averaging operators.

Fourier integral representation. Fix a smooth family of defining pairs $[\Phi;\mathfrak{a}]$ and, for the purposes of this subsection, assume that

$$|\kappa(\Phi)(\vec{x};z)|,\ |\mathrm{Proj}(\Phi)(\vec{x};z)|,\ |\mathrm{Cin}(\Phi)(\vec{x};z)| \geq \varepsilon_{\mathrm{Cin}} > 0 \qquad \text{for all } (\vec{x};z) \in \operatorname{supp} \mathfrak{a};$$

moreover, assume that upper bounds for the derivatives of Φ depend polynomially on 2^k, where k is as in Proposition 8.2. Here $\vec{x} = (x,t) \in \mathbb{R}^2 \times \mathbb{R}$. Owing to the nature of the estimates in Proposition 8.2, here one does not need to be very precise about dependencies involving various derivatives of Φ and $\mathfrak{a}$ and the bounds on the curvatures (as opposed to the situation in §7). For instance, the constant $\varepsilon_{\mathrm{Cin}}$ may depend on the parameters k,ℓ and m. In what follows, we will not determine the precise dependence of our estimates on these parameters but will only be concerned with showing that it is not worse than 2^{Mk} for some large constant $M \geq 1$.

Given a phase/amplitude pair $[\Phi_t;\mathfrak{a}_t]$, from (5.18) and the Fourier inversion formula,

$$A_j[\Phi_t;\mathfrak{a}_t]f(x) = \int_{\hat{\mathbb{R}}^2} \widetilde{K}^{2^j}(\vec{x};\xi)\hat{f}(\xi)\,\mathrm{d}\xi$$

where

$$\widetilde{K}^{\lambda}(\vec{x};\xi) := \frac{1}{(2\pi)^3}\int_{\mathbb{R}^2}\int_{\mathbb{R}} e^{i(\theta\Phi(\vec{x};z)+\langle z,\xi\rangle)}\mathfrak{a}(\vec{x};z)\beta(\theta/\lambda)\,\mathrm{d}\theta\,\mathrm{d}z. \tag{8.2}$$

This function can be analysed via stationary phase arguments. The critical points (z_{cr},θ_{cr}) of the phase function

$$(z,\theta)\mapsto \Psi(z,\theta;\vec{x};\xi) = \theta\Phi(\vec{x};z) + \langle z,\xi\rangle \tag{8.3}$$

satisfy $\Phi(\vec{x};z_{cr}) = 0$ and $\theta_{cr}\partial_z\Phi(\vec{x};z_{cr}) + \xi = 0$. The former condition implies that $z_{cr}\in\Sigma_{\vec{x}} = \{z\in\mathbb{R}^2 : \Phi_t(x;z) = 0\}$ while the latter implies that the normal to $\Sigma_{\vec{x}}$ at z_{cr} is parallel to $\pm\xi$. We also have $|\det\partial^2_{(z,\theta)}\Psi| = |\theta|^d\kappa(\Phi)$ so that the critical points are nondegenerate.

Let $C_\circ \geq 1$ satisfy

$$(C_\circ/10)^{-1} \leq |\partial_z\Phi(\vec{x};z)| \leq C_\circ/10 \qquad \text{for all } (\vec{x};z)\in\operatorname{supp}\mathfrak{a}.$$

There are no critical points for the phase if $|\xi|\geq 4C_\circ\lambda$ or $|\xi|\leq\lambda/4C_\circ$. Thus, by repeated integration-by-parts

$$\widetilde{K}^{\lambda}(\vec{x};\xi) = \widetilde{K}^{\lambda}(\vec{x};\xi)\tilde{\beta}(|\xi|/\lambda) + E^{\lambda}(\vec{x};\xi) \tag{8.4}$$

where $\tilde{\beta}(r) := \eta(C_\circ^{-1}r) - \eta(C_\circ r)$ and the error E^λ satisfies

$$|\partial_\xi^\alpha[e^{-i\langle x,\xi\rangle}E^\lambda(\vec{x};\xi)]| \lesssim C_\circ^N\lambda^{-N/2}(1+|\xi|)^{-N/2} \quad \text{for all } |\alpha|\leq N, \tag{8.5}$$

with implicit bounds depending on $\|\mathfrak{a}\|_{C^N}$. Note that the value of $C_\circ$ will generally depend on k or m for the classes considered in Proposition 8.2, but this dependence is admissible in our forthcoming analysis.

Key example. Let $[\Phi;\mathfrak{a}]\in\mathfrak{C}^k_{\mathrm{Cin}}\cap\mathfrak{C}^k_{\mathrm{Rot}}$. The condition $\Phi_1)_{k,2k}$ ensures that $|\partial_z\Phi(\vec{x};z)|\sim 2^{2k/3}$ and so $C_\circ\sim 2^{2k/3}$ in this case.

We further analyse $\widetilde{K}^\lambda(\vec{x};\xi)$ for $C_\circ^{-1}/4 < |\xi| < 4C_\circ|\xi|$. Decompose

$$\widetilde{K}^\lambda(\vec{x};\xi)\tilde{\beta}(|\xi|/\lambda) = \sum_{i\in\mathcal{J}}\widetilde{K}^\lambda_i(\vec{x};\xi)$$

where the cardinality of the index set $\mathcal{J}$ is polynomial in $\varepsilon_{\mathrm{C}in}$ and $C_\circ$ and each $\widetilde{K}^\lambda_i$ is of the form

$$\mathcal{K}^\lambda(\vec{x};\xi) := \chi(\vec{x};\xi)\frac{1}{(2\pi)^3}\int_{\mathbb{R}^2}\int_{\mathbb{R}} e^{i(\theta\Phi(\vec{x};z)+\langle z,\xi\rangle)}\mathfrak{a}(\vec{x};z)\beta(\theta/\lambda)\,\mathrm{d}\theta\,\mathrm{d}z, \tag{8.6}$$

with χ, β, $\mathfrak{a}$ supported on sets of diameter ε with $\varepsilon^C_{\mathrm{Cin}}\lesssim\varepsilon\lesssim\varepsilon_{\mathrm{Cin}}$.

It suffices to consider the kernel (8.6). If $\Phi(\vec{x};z)$ does not vanish in the neighborhood of the support then the integral represents a smooth function with derivative bounds polynomial in $\varepsilon^{-1}_{\mathrm{Cin}}$, $C_\circ$. Otherwise we may use the method of stationary phase, using that the critical points of (8.3) are nondegenerate. In a neighborhood of the support of the symbol, we can then solve the equation $\nabla_{\theta,z}\Psi(z,\theta;\vec{x},\xi) = 0$ in (z,θ) with $z = \nu(\vec{x};\xi)$, $\theta = \Theta(\vec{x};\xi)$ denoting the solutions; moreover ν is homogeneous of degree 0 in ξ and Θ is homogeneous of degree 1 in ξ. Hence

$$\begin{cases} \Phi(\vec{x};\nu(\vec{x};\xi)) &= 0 \\ \Theta(\vec{x};\xi)\partial_z\Phi(\vec{x};\nu(\vec{x};\xi)) + \xi &= 0 \end{cases}. \tag{8.7}$$

Furthermore, if

$$\varphi(\vec{x};\xi) := \Psi(\nu(\vec{x},\xi),\Theta(\vec{x};\xi);\vec{x};\xi),$$

then (8.7) implies that

$$\varphi(\vec{x};\xi) = \langle \nu(\vec{x};\xi), \xi\rangle. \tag{8.8}$$

By rescaling and applying the method of stationary phase [14, Theorem 7.7.5], one deduces that

$$\mathcal{K}^{\lambda}(\vec{x};\xi) = e^{i\lambda\varphi(\vec{x};\xi/\lambda)}\frac{\mathbb{a}(\vec{x};\xi/\lambda)}{(1+|\xi|^2)^{1/4}} + E(\vec{x};\xi/\lambda) \tag{8.9}$$

where, for some $M_N > 0$:

- The symbol $\mathbb{a}$ is supported in $\{C_\circ^{-1} \lesssim |\xi| \lesssim C_\circ\}$ and satisfies
  $$|\partial_{\vec{x}}^{\alpha}\partial_{\xi}^{\beta}\mathbb{a}(\vec{x};\xi)| \lesssim (\varepsilon_{\mathrm{Cin}}^{-1} + C_\circ + \|\Phi\|_{C^{3N+2}} + \|\mathfrak{a}\|_{C^{3N}})^{M_N}$$
  and all $(\alpha,\beta) \in \mathbb{N}_0^3 \times \mathbb{N}_0^2$ with $|\alpha|, |\beta| \le N$.
- The error term E is rapidly decaying in the sense that
  $$|\partial_{\xi}^{\alpha}[e^{-i\langle x,\xi\rangle}E(\vec{x};\xi/\lambda)]| \lesssim (\varepsilon_{\mathrm{Cin}}^{-1} + \|\Phi\|_{C^{3N+2}} + \|\mathfrak{a}\|_{C^{3N}})^{M_N}\lambda^{-N} \tag{8.10}$$
  for any $\alpha \in \mathbb{N}_0^2$ with $|\alpha| \le N$.

One is therefore led to consider operators belonging to the following class.

Definition 8.3. An FIO pair $[\varphi;\mathbb{a}]$ consists of a pair of functions $\varphi, \mathbb{a} \in C^\infty(\mathbb{R}^3 \times \hat{\mathbb{R}}^2)$ with $\mathbb{a}$ supported in a compact set of diameter 1. For any such pair $[\varphi;\mathbb{a}]$ and $\mu \in \mathbb{R}$ define Fourier integral operators of order μ by

$$\mathcal{F}_{\mu}^{\lambda}[\varphi;\mathbb{a}]f(\vec{x}\,) := \int_{\hat{\mathbb{R}}^2} e^{i\lambda\varphi(\vec{x};\xi/\lambda)}\frac{\mathbb{a}(\vec{x};\xi/\lambda)}{(1+|\xi|^2)^{-\mu/2}}\beta(|\xi|/\lambda)\hat{f}(\xi)\,\mathrm{d}\xi \qquad \text{for } \lambda \ge 1. \tag{8.11}$$

Local smoothing estimates. Under certain 'geometric' hypotheses on the phase, $L^p_x \to L^p_{x,t}$ estimates are known for the operators (8.11) with good λ decay (indeed, the best possible decay (up to ε losses) for $6 \le p \le \infty$). Here the relevant hypotheses are stated in a weakly quantitative form. In what follows we use the notation $\bigwedge_{k=1}^{2}\vec{v}_k$ for the standard vector product $\vec{v}_1 \times \vec{v}_2$ for vectors in $\mathbb{R}^3$.

Definition 8.4. For $R \ge 1$ let $\mathbb{A}(R)$ denote the class of all $[\varphi;\mathbb{a}]$ satisfying

H0) $|\partial_{\vec{x}}^{\alpha}\partial_{\xi}^{\beta}\varphi(\vec{x};\xi)| \lesssim R \quad$ for $|\alpha| \le N$ and $0 < |\beta| \le N$,

H1) $\Big|\bigwedge_{k=1}^{2}\partial_{\xi_k}\partial_{\vec{x}}\varphi(\vec{x};\xi)\Big| \ge R^{-1}$,

H2) $\max_{1\le i,j\le 2}\Big|\Big\langle \partial_{\xi_i\xi_j}^2\partial_{\vec{x}}\,\varphi(\vec{x};\xi)\,,\ \bigwedge_{k=1}^{2}\partial_{\xi_k}\partial_{\vec{x}}\varphi(\vec{x};\xi)\Big\rangle\Big| \ge R^{-1}$

for all $(\vec{x};\xi) \in \operatorname{supp}\mathbb{a}$.

The following theorem is the key ingredient in the proof of Proposition 8.2.

Theorem 8.5 ([3])**.** *There exist $N, M \in \mathbb{N}$ such that*

$$\|\mathcal{F}_{\mu}^{\lambda}[\varphi;\mathbb{a}]\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^3)} \lesssim_{\varepsilon} R^M\lambda^{1/6+\mu+\varepsilon}\|\mathbb{a}\|_{C^N} \qquad \textit{for all } [\varphi;\mathbb{a}] \in \mathbb{A}(R).$$

This weakly quantitative statement is not explicit in [3] or the corresponding survey [4] but it may be extracted from the proof. It is remarked that Theorem 8.5 is more than enough for the purposes of this article and, indeed, *any* non-trivial local smoothing estimate (that is, a gain of an epsilon derivative over the fixed term estimate) would suffice. Thus one could equally appeal to the older results of [20] (see also the related work [16, Chapter 3], or the more recent work [10]).

Relating the phase functions. In order to apply Theorem 8.5 we analyse the hypotheses H0), H1) and H2) for the specific case of the phase φ arising from the averaging operators $A[\Phi_t; \mathfrak{a}_t]f$.

Let φ be of the form (8.8), induced by some defining function Φ. Implicit differentiation of (8.7) yields

$$\begin{bmatrix} \partial_\xi \Theta \\ \partial_\xi \nu \end{bmatrix} = - \begin{bmatrix} 0 & (\partial_z \Phi)^\top \\ \partial_z \Phi & \Theta \partial_{zz}^2 \Phi \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathrm{Id}_2 \end{bmatrix}, \tag{8.12}$$

$$\begin{bmatrix} \partial_{\vec{x}} \Theta \\ \partial_{\vec{x}} \nu \end{bmatrix} = - \begin{bmatrix} 0 & (\partial_z \Phi)^\top \\ \partial_z \Phi & \Theta \partial_{zz}^2 \Phi \end{bmatrix}^{-1} \begin{bmatrix} \partial_{\vec{x}} \Phi & \Theta \partial_{\vec{x}z}^2 \Phi \end{bmatrix}^\top, \tag{8.13}$$

where the right-hand matrices are evaluated at $z = \nu(\vec{x}; \xi)$. In particular, (8.12) implies that $\partial_{\xi_2} \nu_1 = \partial_{\xi_1} \nu_2$ and combining this with Euler's homogeneity relation $\varphi(\vec{x}; \xi) = \langle \partial_\xi \varphi(\vec{x}; \xi), \xi \rangle$ yields

$$\partial_\xi \varphi = \nu. \tag{8.14}$$

Consequently, one can check that if $(\alpha, \beta) \in \mathbb{N}_0^3 \times \mathbb{N}_0^2$ satisfies $|\alpha|$, $|\beta| \leq N$, then

$$|\partial_{\vec{x}}^\alpha \partial_\xi^\beta \varphi(\vec{x}; \xi)| \lesssim_N (\|\Phi\|_{C^N} + \varepsilon_{\mathrm{Cin}}^{-1})^{M_N}. \tag{8.15}$$

for a certain $M_N > 0$.

Furthermore, (8.13) and (8.14) also imply that

$$\big| \bigwedge_{j=1}^2 \partial_{\xi_j} \partial_{\vec{x}} \varphi(\vec{x}; \xi) \big| \gtrsim \frac{1}{|\partial_{\vec{x}} \Theta(\vec{x}; \xi)|} \det \begin{bmatrix} \partial_{\vec{x}} \Theta(\vec{x}; \xi) \\ \partial_{\vec{x}} \nu(\vec{x}; \xi) \end{bmatrix} \geq \mathrm{Proj}(\Phi)(\vec{x}; \xi) \cdot \|\Phi\|_{C^2}^{-3}. \tag{8.16}$$

These inequalities allow one to deduce H0) and H1); the condition H2) requires a slightly more involved analysis.

Letting $\sigma_1 := 2$ and $\sigma_2 := 1$, the identities in (8.12) and (8.13) give

$$\partial_{\xi_j} \nu_i = \frac{(-1)^{i+j+1} \xi_{\sigma_i} \xi_{\sigma_j}}{\Theta\big(\Theta^2 \kappa(\Phi)\big)}, \tag{8.17}$$

$$\partial_{\vec{x}} \nu_i = \frac{(-1)^i}{\kappa(\Phi)} \left(\det \begin{bmatrix} \partial_{z_1} \Phi & \partial_{z_1 z_{\sigma_i}}^2 \Phi \\ \partial_{z_2} \Phi & \partial_{z_2 z_{\sigma_i}}^2 \Phi \end{bmatrix} \mathbf{T}^1 - (\partial_{z_{\sigma_i}} \Phi) \mathbf{T}^2 \right), \tag{8.18}$$

where $\kappa(\Phi)$ is as defined in (4.7) and the $\mathbf{T}^i$ are the tangent vector fields from (4.9). Recalling (8.14), the condition H2) for the phase function (8.8) involves mixed second order derivatives of ν; by (8.17), computing these derivatives boils down to differentiating $\big(\Theta^2 \kappa(\Phi)\big)^{-1} \Theta^{-1}$ with respect to $\vec{x}$. Recalling the definition of Θ and ν from (8.7) and the identities of (8.13),

$$\partial_{\vec{x}} \big(\Theta^2 \kappa(\Phi)\big) = \partial_{\vec{x}} \det \begin{bmatrix} 0 & \xi^\top \\ \xi & \partial_{zz}^2 \Phi \end{bmatrix} = \Theta^2 \mathbf{S}^1, \quad \partial_{\vec{x}} \Theta = - \frac{\Theta}{\kappa(\Phi)} \Big(\det \big[\partial_{zz}^2 \Phi\big] \mathbf{T}^1 + \mathbf{S}^2 \Big)$$

where the $\mathbf{S}^i$ are as in Definition 4.4. The product rule then yields

$$\partial_{\vec{x}} \Big(\big(\Theta^2 \kappa(\Phi)\big)^{-1} \Theta^{-1} \Big) = -\Theta^{-3} \kappa(\Phi)^{-2} \big(\mathbf{S} - \det \big[\partial_{zz}^2 \Phi\big] \mathbf{T}^1 \big). \tag{8.19}$$

Combining (8.14), (8.17), (8.18) and (8.19), one deduces that

$$\det \begin{bmatrix} \partial_{\xi_i \xi_j}^2 \partial_{\vec{x}} \varphi(\vec{x}; \xi) \\ \partial_{\xi_1} \partial_{\vec{x}} \varphi(\vec{x}; \xi) \\ \partial_{\xi_2} \partial_{\vec{x}} \varphi(\vec{x}; \xi) \end{bmatrix} = (-1)^{i+j+1} \frac{\mathrm{Cin}(\Phi)(\vec{x}; \nu(\vec{x}; \xi))}{\Theta(\vec{x}; \xi)^3 \kappa(\Phi)(\vec{x}; \nu(\vec{x}; \xi))^3} \xi_{\sigma_i} \xi_{\sigma_j}. \tag{8.20}$$

The identities (8.15), (8.16) and (8.20) allow one to relate the conditions H0), H1) and H2) of the phase φ to properties of the underlying defining function (and, in particular, bounds on $\|\Phi\|_{C^N}$, $\kappa(\Phi)$, $\mathrm{Proj}(\Phi)$ and $\mathrm{Cin}(\Phi)$).

8.3. **Application of local smoothing.** Theorem 8.5 can now be applied to yield Proposition 8.2.

Proof of Proposition 8.2. The main difficulty is to prove iii). Fix $[\Phi; \mathfrak{c}] \in \mathfrak{C}^k_{\mathrm{Cin}} \cap \mathfrak{C}^k_{\mathrm{Rot}}$ and $\delta > 0$; it Let I denote an interval of length $|I| \sim 1$. The Sobolev embedding argument used to prove (6.39) yields

$$\|\sup_{t \in I} |A_j[\Phi_t; \mathfrak{c}_t] f|\|^6_{L^6(\mathbb{R}^2)} \leq |I|^{-1} \|A_j[\Phi_t; \mathfrak{c}_t] f\|^6_{L^6(\mathbb{R}^2 \times I)} + 6\|A_j[\Phi_t; \mathfrak{c}_t] f\|^5_{L^6(\mathbb{R}^2 \times I)} \|A_j[\Phi_t; \mathfrak{d}_t] f\|_{L^6(\mathbb{R}^2 \times I)}, \tag{8.21}$$

where $\mathfrak{d}_t := 2\pi i 2^j (\partial_t \Phi_t) \mathfrak{c}_t + \partial_t \mathfrak{c}_t$.[4] By the definition of the class $\mathfrak{C}^k_{\mathrm{Cin}}$,

$$|\kappa(\Phi)(x, t; z)|, \ \ |\mathrm{Proj}(\Phi)(x, t; z)|, \ \ |\mathrm{Cin}(\Phi)(x, t; z)| \gtrsim 2^{-Mk} \delta \tag{8.22}$$

whenever $(x, t; z) \in \operatorname{supp} \mathfrak{c}$ and $|t - x_2| \gtrsim \delta$. Decompose $\mathfrak{c} := \mathfrak{c}^{(\delta)} + \mathfrak{c}^{\dagger}$ where

$$\mathfrak{c}^{(\delta)}(x, t; z) := \mathfrak{c}(x, t; z) \eta((t - x_2)/10\delta)$$

so that the estimates (8.22) hold on the support of $\mathfrak{c}^{\dagger}$.

The piece corresponding to $\mathfrak{c}^{(\delta)}$ can be bounded using the theory from Sections 6.1 and 6.6. Indeed, let $\mathcal{G}(x_1, x_2, t, z) := (x_1, x_2 + t, t; z)$ and define

$$\tilde{\Phi} := \Phi \circ \mathcal{G}, \qquad \tilde{\mathfrak{c}}^{(\delta)} := \mathfrak{c}^{(\delta)} \circ \mathcal{G}.$$

Note that $|x_2| \leq \delta$ in $\operatorname{supp} \tilde{\mathfrak{c}}^{(\delta)}$. Performing the above change of variables, by Fubini's theorem

$$\|A_j[\Phi_t; \mathfrak{c}^{(\delta)}_t] f\|^6_{L^6(\mathbb{R}^2 \times I)} = \int_{-\delta}^{\delta} \|A_j[(\tilde{\Phi})^{\star}_r; (\tilde{\mathfrak{c}}^{(\delta)})^{\star}_r] f\|^6_{L^6(\mathbb{R}^2)} \,\mathrm{d}r$$

where

$$(\tilde{\Phi})^{\star}(u, t, r; v, \rho) := \tilde{\Phi}(u, r, t; v, \rho) \quad \text{and} \quad (\tilde{\mathfrak{c}}^{(\delta)})^{\star}(u, t, r; v, \rho) := \tilde{\mathfrak{c}}^{(\delta)}(u, r, t; v, \rho).$$

Since $[\Phi; \mathfrak{c}] \in \mathfrak{C}^k_{\mathrm{Rot}}$, it follows that $[(\tilde{\Phi})^{\star}; (\tilde{\mathfrak{c}}^{(\delta)})^{\star}] \in \mathfrak{A}^{k, \ell(k)}_{\mathrm{Rot}}$. Combining Proposition 6.1 with Lemma 6.9 we get an $L^2(\mathbb{R}^2)$ estimate for fixed r,

$$\|A_j[(\tilde{\Phi})^{\star}_r; \tilde{\mathfrak{c}}^{(\delta)}_r\|_{L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)} \lesssim (2^{-j/2} \wedge 2^{-2k/3}) \|\mathfrak{c}^{\delta}\|_{C^N_z}.$$

Interpolating this bound with the L^{∞} estimate from Corollary 8.1 iii) one gets

$$\|A_j[(\tilde{\Phi})^{\star}_r; \tilde{\mathfrak{c}}^{(\delta)}_r\|_{L^6(\mathbb{R}^2) \to L^6(\mathbb{R}^2)} \lesssim (2^{-j/6} \wedge 2^{-2k/3}) \|\mathfrak{c}^{\delta}\|_{C^N_z}$$

and therefore

$$\|A_j[\Phi_t; \mathfrak{c}^{(\delta)}_t]\|_{L^6(\mathbb{R}^2) \to L^6(\mathbb{R}^2 \times I)} \lesssim \delta^{1/6} (2^{-j/6} \wedge 2^{-2k/3}) \|\mathfrak{c}\|_{C^N} \lesssim \delta^{1/6} 2^{-j/6} \|\mathfrak{c}\|_{C^N}. \tag{8.23}$$

On the other hand, Theorem 8.5 can be used to show that

$$\big\|A_j[\Phi_t; \mathfrak{c}^{\dagger}_t]\big\|_{L^6(\mathbb{R}^2) \to L^6(\mathbb{R}^2 \times I)} \lesssim_{\varepsilon} \delta^{-M} 2^{Mk} 2^{-j(1/3 - \varepsilon)} \|\mathfrak{c}\|_{C^N}. \tag{8.24}$$

---

[4] To be more precise, one may write $A_j[\Phi_t; \mathfrak{d}_t] = A_j[\Phi_t; \mathfrak{d}^1_t] + A_j[\Phi_t; \mathfrak{d}^2_t]$ where $\mathfrak{d}^1_t := 2\pi i 2^j (\partial_t \Phi_t) \mathfrak{c}_t$ and the average corresponding to $\mathfrak{d}^2_t$ is defined with the frequency cut-off $\theta \mapsto \theta\beta(\theta)$, rather than just β. It is remarked that this ambiguity in the definition has no bearing on the analysis.

Temporarily assuming (8.24), by taking $\delta := 2^{-j/(24M)}$ and $\varepsilon := 1/12$, we get

$$\|A_j[\Phi_t; \mathfrak{c}_t^{\dagger}]\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^2\times I)} \lesssim 2^{Mk} 2^{-j(1/3-1/12-1/24)}\|\mathfrak{c}\|_{C^N}$$

and hence combining this with (8.23) we obtain

$$\begin{aligned}\|A_j[\Phi_t; \mathfrak{c}_t]\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^2\times I)} &\lesssim_\varepsilon 2^{Mk}2^{-j/6}(2^{-j/24} + 2^{-j/(144M)})\|\mathfrak{c}\|_{C^N} \\ &\lesssim 2^{Mk}2^{-j/6-j\varepsilon_0}\|\mathfrak{c}\|_{C^N} \qquad (8.25)\end{aligned}$$

for some $\varepsilon_0 > 0$ (indeed $\varepsilon_0 = (144M)^{-1}$). This gives a favourable bound for the terms on the right-hand side of (8.21) involving $\mathfrak{c}_t$. For the amplitude $\mathfrak{d}_t$ it suffices to note that $\|\mathfrak{d}\| \lesssim 2^j\|\mathfrak{c}\|$ and that $[\Phi; \mathfrak{d}] \in \mathfrak{C}^k_{\mathrm{Cin}} \cap \mathfrak{C}^k_{\mathrm{Rot}}$. Therefore

$$\|A_j[\Phi_t; \mathfrak{d}_t]\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^2\times I)} \lesssim_\varepsilon 2^{Mk}2^{j(5/6-\varepsilon_0)}\|\mathfrak{c}\|_{C^N}. \qquad (8.26)$$

Combining (8.25) and (8.26) in (8.21) concludes the argument of Proposition 8.2 for $[\Phi; \mathfrak{c}] \in \mathfrak{C}^k_{\mathrm{Cin}} \cap \mathfrak{C}^k_{\mathrm{Rot}}$.

It remains to prove (8.24). Let $[\varphi; \mathbb{c}]$ be the FIO pair associated to $[\Phi; \mathfrak{c}^{\dagger}] \in \mathfrak{C}^k_{\mathrm{Cin}}$, defined as in (8.8) and (8.9). Thus,

$$A_j[\Phi_t; \mathfrak{c}_t^{\dagger}]f(x) = \mathcal{F}^{2^j}_{-1/2}[\varphi; \mathbb{c}]f(\vec{x}) + \mathcal{E}_j f(\vec{x}),$$

where the operator $\mathcal{E}_j$ arises from the errors in (8.4) and (8.9). The smoothing term $\mathcal{E}_j$ can be easily estimated using repeated integration-by-parts and the rapid decay from (8.5) and (8.10).

Turning to the main term $\mathcal{F}^{2^j}_{-1/2}[\varphi; \mathbb{c}]f$, the condition $\mathbf{C}_\delta)_k$ together with (8.15), (8.16) and (8.20) imply that $[\varphi; \mathbb{c}] \in \mathbb{A}^k := \mathbb{A}(\delta^{-M_\circ}2^{M_\circ k})$ (in the sense of Definition 8.4) for some absolute constant $M_\circ \geq 1$. Thus, Theorem 8.5 implies that

$$\|\mathcal{F}^{\lambda}_{\mu}[\varphi; \mathbb{c}]\|_{L^6(\mathbb{R}^2)\to L^6(\mathbb{R}^3)} \lesssim_\varepsilon \delta^{-M}2^{Mk}\lambda^{1/6+\mu+\varepsilon}\|\mathbb{c}\|_{C^N}.$$

The case of interest is given by $\mu = -1/2$; note that for this value the λ exponent is $-1/3 + \varepsilon$, corresponding to the 2^j exponent in (8.24).

For the remaining cases i), ii), iv) and v) of the proposition the argument is similar but somewhat easier. Indeed, here the condition $\mathbf{C})_k$ provides favourable lower bounds for the various curvatures and this obviates the need to form any decomposition $\mathfrak{a} = \mathfrak{a}^{(\delta)} + \mathfrak{a}^{\dagger}$ (one may bound $A_j[\Phi; \mathfrak{a}]$ directly using Theorem 8.5). □

9. The global maximal function

It remains to extend the bound for the local maximal function from Theorem 3.3 to the bound on the 'global' maximal function from Theorem 3.1. This is the last step in the proof of Theorem 1.1.

Proof of Theorem 3.1. Break the operator according to the relative size of r with respect to t, thus:

$$\sup_{t>0}|A_t f(u,r)| = \sup_{T\in\mathbb{Z}}\sup_{2^T\leq t<2^{T+1}}\Big(\sum_{m\geq 10} + \sum_{m\leq -10} + \sum_{|m|<10}\Big)\beta^{m+T}(r)|A_t f(u,r)|.$$

Each of the three terms is estimated separately. Of these, the first case (corresponding to $t \ll r$) presents the most interesting features.

The first term: $t \ll r$. The orthogonality relation (3.6) induces spatial orthogonality and it therefore suffices to show that

$$\Big\| \sup_{T\in\mathbb{Z}} \sup_{2^T\le t\le 2^{T+1}} \sum_{m\ge 10} \beta^{m+T}\cdot |A_t f|\Big\|_{L^p(\mathbb{R}\times[2^W,2^{W+1}])} \lesssim \|f\chi_{\mathbb{R}\times[2^{W-1},2^{W+2}]}\|_p,$$

uniformly in $W\in\mathbb{Z}$. By the rescaling $(u,r,t;v,\rho)\mapsto(2^{2W}u,2^W r,2^W t;2^{2W}v,2^W\rho)$, the problem reduces to the case $W=0$, and therefore one needs to only show that

$$\Big\| \sup_{T\le -5} \sup_{2^T\le t\le 2^{T+1}} \beta^0\cdot |A_t f|\Big\|_{L^p(\mathbb{R}\times[1,2])} \lesssim \|f\|_p.$$

For fixed $T\le -5$, decompose f into frequency localised pieces

$$f = P_{\le -T} f + \sum_{k=1}^{\infty} P_{-T+k} f,$$

where $(P_{\le m} f)\,\widehat{}\,(\xi) := \eta^m(|\xi|)\widehat{f}(\xi)$ and $(P_m f)\,\widehat{}\,(\xi) := \beta^m(|\xi|)\widehat{f}(\xi)$ for the functions η^m and β^m defined in (5.2). A routine computation shows that the precomposition of the above maximal operator with $P_{\le -T}$ is pointwise dominated by the Hardy–Littlewood maximal function. Consequently, for $p>2$ it suffices to show that

$$\Big\| \sup_{T\le -5} \sup_{2^T\le t\le 2^{T+1}} \beta^0\cdot |A_t P_{-T+k} f|\Big\|_{p} \lesssim 2^{-\varepsilon_p k}\|f\|_p$$

and Littlewood–Paley theory further reduces the problem to proving

$$\Big\| \sup_{2^T\le t\le 2^{T+1}} \beta^0\cdot |A_t P_{-T+k} f|\Big\|_{p} \lesssim 2^{-\varepsilon_p k}\|f\|_p, \tag{9.1}$$

uniformly in $T\le -5$. The rescaling $(u,r,t;v,\rho)\mapsto(2^{2T}u,2^T r,2^T t;2^{2T}v,2^T\rho)$ transforms (9.1) into

$$\Big\| \sup_{1\le t\le 2} \beta^{-T}\cdot |A_t P_k^T f|\Big\|_{p} \lesssim 2^{-\varepsilon_p k}\|f\|_p,$$

where P_k^T denotes the anisotropic frequency projection associated to the multiplier $\beta^k\big(|(2^{-T}\xi_1,\xi_2)|\big)$.

The situation in the last display is close to the case $m=-T>0$ in the decomposition (5.5), although a direct application of Theorem 5.1 iii) will not give the desired decay in j. Instead, we decompose the operator A as a sum of frequency localised operators A_j as in (5.18) and appeal to Proposition 5.11 iv). First, for fixed $T\le -5$, write

$$\beta^{-T}(r)\cdot A_t P_k^T f(u,r) = \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{-T} A[\Phi_t;\mathfrak{a}_t^{-T,\vec{\sigma}}]P_k^T f(u,r),$$

where $\mathfrak{a}_t^{-T,\vec{\sigma}}$ is as in (5.4). The relations (3.6) ensure that $|r-\rho|\lesssim 1$ and $|u-v|\lesssim 2^{-T}$, so by spatial orthogonality it suffices to prove

$$\Big\| \sup_{1\le t\le 2} |A[\Phi_t;\mathfrak{a}_t^{-T,\vec{\sigma}}]P_k^T f|\Big\|_{p} \lesssim 2^T 2^{-\varepsilon_p k}\|f\|_p$$

uniformly in $\vec{\sigma}\in\mathbb{Z}^2$. A further rescaling $(u,v)=(2^{-T}u,2^{-T}v)$ transforms the above estimate into

$$\big\| \sup_{1\le t\le 2} |A[\Phi_t^{-T};\tilde{\mathfrak{a}}_t^{-T,\vec{\sigma}}]P_k f|\big\|_{p} \lesssim 2^{-\varepsilon_p k}\|f\|_p, \tag{9.2}$$

where now P_k is the usual dyadic frequency projection at scale 2^k and Φ^{-T} and $\tilde{\mathfrak{a}}^{-T,\vec{\sigma}}$ are defined as in (5.14); in particular, $[\Phi^{-T};\mathfrak{a}^{-T,\vec{\sigma}}]\in\mathfrak{A}^0_{\mathrm{Cin}}\cap\mathfrak{A}^0_{\mathrm{Rot}}$. Decompose $A[\Phi_t^{-T};\mathfrak{a}_t^{-T,\vec{\sigma}}]=\sum_{j\geq 0}A_j[\Phi_t^{-T};\mathfrak{a}_t^{-T,\vec{\sigma}}]$ as in (5.18). Then, for fixed $k>0$, one needs to understand

$$A_j[\Phi_t^{-T};\mathfrak{a}_t^{-T,\vec{\sigma}}]P_kf(u,r)=\int_{\hat{\mathbb{R}}^2}\widetilde{K}^{2^j}(u,r,t;\xi)\beta^k(\xi)\hat{f}(\xi)\,\mathrm{d}\xi \tag{9.3}$$

for $j\geq 0$, where $\widetilde{K}^{2^j}$ is as in (8.2).

The main contribution arises from the terms with $|j-k|\leq 5$. Here we appeal to Proposition 5.11 iv), which yields

$$\big\|\sup_{1\leq t\leq 2}|A_j[\Phi_t^{-T};\tilde{\mathfrak{a}}_t^{-T,\vec{\sigma}}]P_kf|\big\|_p\lesssim 2^{-k\varepsilon_p}\|f\|_p,$$

with some $\varepsilon_p>0$ when $p>2$.

Now consider the case $|j-k|>5$ in (9.3). In our present rescaled situation we have $|\partial_{(v,\rho)}\Phi^{-T}|\sim 1$ and also favourable upper bounds for the higher (v,ρ)-derivatives. Hence, arguing as in §8.2, using repeated integration-by-parts, we obtain

$$|\partial_\xi^\alpha[e^{-2\pi i\langle(u,r),\xi\rangle}\widetilde{K}^{2^j}(u,r,t;\xi)]|\lesssim\min\{2^{-jN/2},2^{-kN/2}\}\,(1+|\xi|)^{-N/2}$$

for all $(u,r,t)\in\operatorname{supp}\tilde{\mathfrak{a}}^{-T,\vec{\sigma}}$, $\xi\in\operatorname{supp}\beta^k$, $\alpha\in\mathbb{N}_0^2$ such that $|\alpha|\leq N$. This yields, via another integration-by-parts,

$$|A_j[\Phi_t^{-T};\mathfrak{a}_t^{-T,\vec{\sigma}}]P_kf(u,r)|\lesssim\big(2^{-jN/2}\wedge 2^{-kN/2}\big)\int_{\mathbb{R}^2}\frac{f(v,\rho)}{(1+|(u,r)-(v,\rho)|)^{N/2}}\,\mathrm{d}v\,\mathrm{d}\rho,$$

which readily implies that

$$\big\|\sup_{1\leq t\leq 2}|A_j[\Phi_t^{-T};\tilde{\mathfrak{a}}_t^{-T,\vec{\sigma}}]P_kf|\big\|_p\lesssim\big(2^{-jN/2}\wedge 2^{-kN/2}\big)\|f\|_p$$

for $1\leq p\leq\infty$, whenever $|j-k|>5$. Combining the above observations, one obtains the desired estimate (9.2).

The second term: $t\gg r$. By the triangle inequality, for all $p>2$ it suffices to show that

$$\big\|\sup_{T\in\mathbb{Z}}\sup_{2^T\leq t<2^{T+1}}\beta^{m+T}\cdot|A_tf|\big\|_p\lesssim 2^{\varepsilon_p m}\|f\|_p$$

holds uniformly in m for some $\varepsilon_p>0$. The orthogonality relation (3.6) ensures that $|t-\rho|\leq r\sim 2^{m+T}\lesssim 2^T$. This induces spatial orthogonality between the t and ρ variables and reduces the analysis to proving

$$\big\|\sup_{2^T\leq t<2^{T+1}}\beta^{m+T}\cdot|A_tf|\big\|_p\lesssim 2^{\varepsilon_p m}\|f\|_p$$

uniformly in $T\in\mathbb{Z}$. By the rescaling $(u,r,t;v,\rho)\mapsto(2^{2T}u,2^Tr,2^Tt;2^{2T}v,2^T\rho)$, it suffices to consider the case $T=0$. The resulting term corresponds to

$$\sum_{\vec{\sigma}\in\mathbb{Z}^2}2^{m/p}A[\Phi_t;\mathfrak{a}_t^{m,\vec{\sigma}}]f(u,r)$$

in (5.5), whose L^p norm is bounded by $2^{m\varepsilon_p}$ for some $\varepsilon_p>0$ if $p>2$ via Theorem 5.1 iv), using the orthogonality arguments in the proof of Theorem 3.3.

The third term: $t \sim r$. Without loss of generality, by replacing β with a cutoff function with slightly larger support, it suffices to bound the term corresponding to $m = 0$. Assuming f is non-negative, for each fixed T perform a decomposition of the operator similar to that in (5.6) and (5.7) by dominating

$$\begin{aligned}\beta^T(r) \cdot A_t f(u,r) \lesssim & \sum_{\substack{(k,\ell)\in\mathbb{Z}^2\\ k\geq -4\\ k-3\leq \ell<\ell(k)}} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{a}_T^{k,\ell,\vec{\sigma}})_t] f \\ & + \sum_{\substack{k\in\mathbb{Z}\\ k\geq -4}} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{c}_T^{k,\ell,\vec{\sigma}})_t] f\end{aligned}$$

where

$$\begin{aligned}\mathfrak{a}_T^{k,\ell,\vec{\sigma}}(u,r,t;v,\rho) &:= \mathfrak{a}^{k,\ell,\vec{\sigma}}(2^{-2T}u, 2^{-T}r, 2^{-T}t; 2^{-2T}v, 2^{-T}\rho), \qquad \ell<\ell(k),\\ \mathfrak{c}_T^{k,\vec{\sigma}}(u,r,t;v,\rho) &:= \mathfrak{c}^{k,\vec{\sigma}}(2^{-2T}u, 2^{-T}r, 2^{-T}t; 2^{-2T}v, 2^{-T}\rho).\end{aligned}$$

By the triangle inequality, for all $p > 2$ it suffices to prove

$$\begin{aligned}\Big\| \sup_{T\in\mathbb{Z}} \sup_{2^T\leq t\leq 2^{T+1}} \sum_{k-3\leq\ell<\ell(k)} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{a}_T^{k,\ell,\vec{\sigma}})_t] f \Big\|_p &\lesssim 2^{-\varepsilon_p k}\|f\|_p,\\ \Big\| \sup_{T\in\mathbb{Z}} \sup_{2^T\leq t\leq 2^{T+1}} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{c}_T^{k,\ell,\vec{\sigma}})_t] f \Big\|_p &\lesssim 2^{-\varepsilon_p k}\|f\|_p\end{aligned}$$

for some $\varepsilon_p > 0$. After fixing k, spatial orthogonality becomes available: the variable ρ is localised at $\rho \sim 2^{-k+T}$. Therefore, in order to show the above estimates, it suffices to prove

$$\begin{aligned}\Big\| \sup_{2^T\leq t\leq 2^{T+1}} \sum_{k-3\leq\ell<\ell(k)} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{a}_T^{k,\ell,\vec{\sigma}})_t] f \Big\|_p &\lesssim 2^{-\varepsilon_p k}\|f\|_p,\\ \Big\| \sup_{2^T\leq t\leq 2^{T+1}} \sum_{\vec{\sigma}\in\mathbb{Z}^2} 2^{k(1-1/p)+T} A[\Phi_t; (\mathfrak{c}_T^{k,\ell,\vec{\sigma}})_t] f \Big\|_p &\lesssim 2^{-\varepsilon_p k}\|f\|_p,\end{aligned}$$

uniformly in T. By the rescaling $(u,r,t;v,\rho) \mapsto (2^{2T}u, 2^T r, 2^T t; 2^{2T}v, 2^T\rho)$, it suffices to only consider the case $T = 0$. This follows from Theorem 5.1 i) and ii) using the arguments in the proof of Theorem 3.3 (following the statement of Theorem 5.1). $\square$

Appendix A. Lemmata on integration-by-parts

The proofs on oscillatory integrals in §6 use a lemma which keeps track of the terms that occur in the repeated integration-by-parts arguments. Assume that $z \mapsto h(z) \in C_c^\infty$ (and keep track of the C^N-norms of h), and that $\nabla\Theta \neq 0$ on supp (h). Define a differential operator $\mathcal{L}$ by

$$\mathcal{L}h = \operatorname{div}\Big(\frac{h\nabla\Theta}{|\nabla\Theta|^2}\Big).$$

Then, by integration by parts,

$$\int_{\mathbb{R}^d} e^{i\lambda\Theta(z)} h(z)\,\mathrm{d}z = i^N \lambda^{-N} \int_{\mathbb{R}^d} e^{i\lambda\Theta(z)} \mathcal{L}^N h(z)\,\mathrm{d}z$$

and thus

$$\begin{aligned}\Big|\int_{\mathbb{R}^d} e^{i\lambda\Theta(z)} h(z)\,\mathrm{d}z\Big| &\le \lambda^{-N}\int_{\mathbb{R}^d} |\mathcal{L}^N h(z)|\,\mathrm{d}z\\ &\le \lambda^{-N}\mathrm{meas}(\mathrm{supp}\ \chi)\sup_{z\in\mathbb{R}^d}|\mathcal{L}^N h(z)|. \qquad (A.1)\end{aligned}$$

A careful analysis of the term $\mathcal{L}^N h$ is needed for various integration-by-parts arguments in this paper and elsewhere in the literature, but a detailed analysis is often left to the reader. For an explicit reference, a straightforward induction proof of the following lemma is contained e.g. in the appendix of [1] (and probably elsewhere).

We shall introduce the following notation. We say that a term is of *type* (A, j) if it is of the form $h_j/|\nabla_z\Theta|^j$ where h_j is a z-derivative of order j of h. A term of *type* $(B, 0)$ is equal to 1. A term is of *type* (B, j) for some $j \ge 1$ if it is of the form $\Theta_{j+1}/|\nabla_z\Theta|^{j+1}$ where Θ_{j+1} is a z-derivative of order $j+1$ of Θ.

Lemma A.1. *Let $N = 0, 1, 2, \dots$. Then*

$$\mathcal{L}^N h = \sum_{\nu=1}^{K(N,d)} c_{N,\nu} h_{N,\nu}$$

where $K(N, d) > 0$, $c_{N,\nu}$ are absolute constants independent of h and Θ, and each function $h_{N,\nu}$ is of the form [5]

$$P_\nu(\tfrac{\nabla\Theta}{|\nabla\Theta|})\beta_{A,\nu}\prod_{\ell=1}^{M_\nu}\gamma_{\ell,\nu} \qquad (A.2)$$

such that each P_ν is a polynomial of d variables (independent of h and Θ), $\beta_{A,\nu}$ is of type $(A, j_{A,\nu})$ for some $j_{A,\nu} \in \{0, \dots, N\}$ and the terms $\gamma_{\ell,\nu}$ are of type $(B, \kappa_{\ell,\nu})$ for some $\kappa_{\ell,\nu} \in \{1, \dots, N\}$, so that for $\nu = 1, \dots, K(N, d)$

$$j_{A,\nu} + \sum_{\ell=1}^{M_\nu}\kappa_{\ell,\nu} = N. \qquad (A.3)$$

Example. In §6 we use the Lemma A.1 in the form of Corollary A.2 below, choosing

$$\Theta(z) = \Psi(x; z) - \Psi(y; z), \qquad (A.4)$$

for fixed $x = (x', x_d)$, $y = (y', y_d) \in \mathbb{R}^d$. Our differential operator $\mathcal{L} = \mathcal{L}_{x,y}$ depends then on x, y.

Corollary A.2. *Let $h \in C^N(\mathbb{R}^d)$ be compactly supported. Let $\rho(x, y) > 0$, and assume that for all z in a neighborhood of* supp h

$$|\nabla_z\Psi(x; z) - \nabla_z\Psi(y; z)| \gtrsim \rho(x, y). \qquad (A.5a)$$

Let $R(x, y) \ge 1$ and assume that for all z-derivatives up to order $|\alpha| \le N + 1$,

$$|\partial_z^\alpha\Psi(x; z) - \partial_z^\alpha\Psi(y; z)| \lesssim_N R(x, y)\rho(x, y). \qquad (A.5b)$$

Then

$$\Big|\int_{\mathbb{R}^d} e^{i\lambda(\Psi(x;z)-\Psi(y;z))}h(z)\,\mathrm{d}z\Big| \lesssim_N \lambda^{-N}\mathrm{meas}(\mathrm{supp}\ h)\max_{j=0,\dots,N}\frac{\|h\|_{C^j}R(x, y)^{N-j}}{\rho(x, y)^N}.$$

[5] The product $\prod_{\ell=1}^{M_\nu}$ is interpreted to be 1 if $M_\nu = 0$, i.e. $j_{A,\nu} = N$.

Proof. We use (A.1) and the assertion follows from

$$|\mathcal{L}_{x,y}^N h(z)| \lesssim_N \max_{j=0,\dots,N} \frac{\|h\|_{C^j} R(x,y)^{N-j}}{\rho(x,y)^N}. \tag{A.6}$$

To see this use Lemma A.1 with the choice (A.4). Observe that by (A.5a) an expression of type (A,j) is bounded by a constant times $\|h\|_{C^j}(\rho(x,y))^{-j}$. By (A.5a) and (A.5b) an expression of type (B,κ) is bounded by a constant times $R(x,y)(\rho(x,y))^{-\kappa}$. We use (A.3) to see that the expression corresponding to (A.2) is bounded by

$$C_N \frac{\|h\|_{C^{j_{A,\nu}}} R(x,y)^{M_\nu}}{(\rho(x,y))^{j_{A,\nu}+\sum_{\ell=1}^{M_\nu}\kappa_{\ell,\nu}}} \lesssim_N \frac{\|h\|_{C^{j_{A,\nu}}} R(x,y)^{N-j_{A,\nu}}}{\rho(x,y)^N}$$

and hence we get (A.6). □

Applications of Corollary A.2. Here $0<\delta_\circ\le 1$ and $m>0$.

- In the proof of Proposition 6.1, Corollary A.2 is applied with the choice of $\rho(x,y) := |x'-y'| + \delta_\circ|x_d-y_d|$, $R(x,y)\lesssim 1$ and the C^N norm of the amplitude is $O(1)$.
- In the proof of Lemma 6.4 Corollary A.2 is applied with $\rho(x,y) := |x'-y'|$ and $R(x,y)\lesssim 1$, and the C^N norm of the amplitude is $O(2^{mN})$.
- In the proof of Lemma 6.5 the $d-1$-dimensional version of Corollary A.2 is applied with $\rho(x',y') := |x'-y'|$ and $R(x',y')\lesssim 1$, and the C^N norm of the amplitude is $O(2^{mN})$.
- In the proof of Lemma 6.6 Corollary A.2 is applied with the choices of $\rho(x,y) := |x' - X_\nu(x_d, y; z_\nu)| + \delta_\circ 2^{-m}|x_d-y_d|$, and $R(x,y)\lesssim 2^m$, and the C^N norm of the amplitude is $O(2^{mN})$.

Appendix B. Computations related to the defining function

B.1. **Derivative dictionary.** For reference, here some derivatives are computed for the specific defining function Φ_t in (3.2). Recall,

$$\Phi(u,r,t;v,\rho) := (u-v)^2 - \Big(\frac{b}{2}\Big)^2\big(4r^2\rho^2 - (r^2+\rho^2-t^2)^2\big)$$

so that the first order derivatives are

$$\partial_u\Phi_t = 2(u-v), \qquad \partial_r\Phi_t = -b^2 r(t^2-r^2+\rho^2)$$

and

$$\partial_v\Phi_t = -2(u-v), \qquad \partial_\rho\Phi_t = -b^2\rho(t^2+r^2-\rho^2)$$

together with the time derivative

$$\partial_t\Phi_t = b^2 t(t^2-r^2-\rho^2).$$

Of course $\partial^2_{ru}\Phi_t = \partial^2_{\rho u}\Phi_t = \partial^2_{rv}\Phi_t = \partial^2_{\rho v}\Phi_t = 0$ whilst the non-vanishing second order derivatives are

$$\partial^2_{uu}\Phi_t = \partial^2_{vv}\Phi_t = 2, \qquad \partial^2_{uv}\Phi_t = -2,$$

$$\partial^2_{rr}\Phi_t = -b^2(t^2-3r^2+\rho^2), \qquad \partial^2_{r\rho}\Phi_t = -2b^2 r\rho, \qquad \partial^2_{\rho\rho}\Phi_t = -b^2(t^2+r^2-3\rho^2)$$

and the time derivatives

$$\partial^2_{tr}\Phi = -2b^2 tr \qquad \text{and} \qquad \partial^2_{t\rho}\Phi = -2b^2 t\rho.$$

Finally, the third order derivative relevant to the argument are

$$\partial^3_{\rho r r}\Phi_t = -2b^2\rho \qquad \text{and} \qquad \partial^3_{\rho\rho r} = -2b^2 r.$$

With these formulæ in hand, it is a simple computation to obtain the expressions (4.2) and (4.4) for the rotational curvature,

$$\mathrm{Rot}(\Phi_t)(u,r;v,\rho) = 4b^4 r t^2 \rho(t^2 - r^2 - \rho^2),$$
$$\mathrm{Rot}(\Phi_r^\star)(u,t;v,\rho) = 4b^4 r^2 t \rho(r^2 - t^2 - \rho^2),$$

as well as the key identity (4.3),

$$\mathrm{Rot}(\Phi_t)(u,r;v,\rho) = 4b^2 r t\rho(\partial_t\Phi_t)(u,r;v,\rho),$$

and expressions (4.11) and (4.12) related to the cinematic curvature

$$\mathrm{Proj}(\Phi)(u,r,t;v,\rho) = -8b^4 r t\rho(r^2 - t^2),$$
$$\mathrm{Cin}(\Phi)(u,r,t;v,\rho) = 64 b^8 r^3 t^3 \rho^3 (r^2 - t^2)$$

for $(v,\rho) \in \Sigma_{u,r,t}$.

B.2. **Rescaling.** It is useful to note how the expressions in the previous subsection behave under rescaling. Given $k, \tau \in \mathbb{Z}$ and $\varepsilon, \delta \in \mathbb{Z}^2$, let $\Phi^{k,\varepsilon,\tau,\delta} := 2^k\,\Phi \circ D^{\varepsilon,\tau,\delta}$ where

$$D^{\varepsilon,\tau,\delta}(u,r,t;v,\rho) := (2^{\varepsilon_1}u, 2^{\varepsilon_2}r, 2^{\tau}t; 2^{\delta_1}v, 2^{\delta_2}\rho).$$

Then

$$\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\Phi^{k,\varepsilon,\tau,\delta}(x,t;z) = 2^k 2^{\varepsilon\cdot\alpha}2^{\delta\cdot\beta}2^{\tau\gamma}(\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\Phi)\circ D^{\varepsilon,\tau,\delta}(x,t;z)$$

for all $\alpha, \beta \in \mathbb{N}_0^2$, $\gamma \in \mathbb{N}_0$. In particular,

$$\mathrm{Rot}(\Phi_t^{k,\varepsilon,\tau,\delta})(x;z) = 2^{3k}2^{|\varepsilon|+|\delta|}\mathrm{Rot}(\Phi_{2^\tau t})\circ D^{\varepsilon,\delta}(x;z)$$

where $D^{\varepsilon,\delta}(x;z) := (2^\varepsilon x; 2^\delta z)$, and the rescaled key identity becomes

$$\mathrm{Rot}(\Phi_t^{k,\varepsilon,\tau,\delta})(x;z) = 4b^2 r\rho t 2^{\varepsilon_2+\delta_2}2^{2k}2^{|\varepsilon|+|\delta|}\partial_t\Phi^{k,\varepsilon,\tau,\delta}(x,t;z).$$

Furthermore,

$$\kappa(\Phi^{k,\varepsilon,\tau,\delta})(\vec{x};z) = 2^{3k}2^{2|\delta|}\kappa(\Phi)\circ D^{\varepsilon,\tau,\delta}(\vec{x};z),$$
$$\mathrm{Proj}(\Phi^{k,\varepsilon,\tau,\delta})(\vec{x};z) = 2^{3k}2^{|\varepsilon|+\tau+|\delta|}\mathrm{Proj}(\Phi)\circ D^{\varepsilon,\tau,\delta}(\vec{x};z),$$
$$\mathrm{Cin}(\Phi^{k,\varepsilon,\tau,\delta})(\vec{x};z) = 2^{6k}2^{|\varepsilon|+\tau+3|\delta|}\mathrm{Cin}(\Phi)\circ D^{\varepsilon,\tau,\delta}(\vec{x};z).$$

References

[1] Theresa C. Anderson, Laura Cladek, Malabika Pramanik, and Andreas Seeger, *Spherical means on the Heisenberg group: Stability of a maximal function estimate,* To appear *J. Anal. Math.* Preprint: arXiv:1801.06981.

[2] Sayan Bagchi, Sourav Hait, Luz Roncal, and Sundaram Thangavelu, *On the maximal function associated to the spherical means on the Heisenberg group,* Dedicated to the memory of Eli Stein. Preprint: arXiv:1812.11926v2.

[3] David Beltran, Jonathan Hickman, and Christopher D. Sogge, *Variable coefficient Wolff-type inequalities and sharp local smoothing estimates for wave equations on manifolds,* Anal. PDE **13** (2020), no. 2, 403433, MR 4078231.

[4] ———, *Sharp local smoothing estimates for Fourier integral operators,* Preprint: arXiv:1812.11616.

[5] J. Bourgain, *Averages in the plane over convex curves and maximal operators,* J. Analyse Math. **47** (1986), 69–85. MR 874045

[6] Alberto-P. Calderón and Rémi Vaillancourt, *On the boundedness of pseudo-differential operators,* J. Math. Soc. Japan **23** (1971), 374–378. MR 0284872

[7] Michael Christ, *Hilbert transforms along curves. I. Nilpotent groups,* Ann. of Math. (2) **122** (1985), no. 3, 575–596. MR 819558

[8] Scipio Cuccagna, L^2 *estimates for averaging operators along curves with two-sided k-fold singularities,* Duke Math. J. **89** (1997), no. 2, 203–216. MR 1460620

[9] J. J. Duistermaat, *Fourier integral operators,* Modern Birkhäuser Classics, Birkhäuser Springer, New York, 2011, Reprint of the 1996 edition, based on the original lecture notes published in 1973. MR 2741911

[10] Chuanwei Gao, Changxing Miao, and Jianwei-Urbain Yang, *Improved variable coefficient square functions and local smoothing of Fourier integral operators,* Preprint: arXiv:1901.01487.

[11] Allan Greenleaf and Andreas Seeger, *On oscillatory integral operators with folding canonical relations,* Studia Math. **132** (1999), no. 2, 125–139. MR 1669698

[12] Lars Hörmander, *Fourier integral operators. I,* Acta Math. **127** (1971), no. 1-2, 79–183. MR 0388463

[13] ______, *Oscillatory integrals and multipliers on* FL^p, Ark. Mat. **11** (1973), 1–11. MR 0340924

[14] ______, *The analysis of linear partial differential operators. I,* Classics in Mathematics, Springer-Verlag, Berlin, 2003, Distribution theory and Fourier analysis, Reprint of the second (1990) edition. MR 1996773

[15] Joonil Kim, *Annulus maximal averages on variable hyperplanes,* Preprint: arXiv:1906.03797.

[16] David T. S. Kung, *Local smoothing phenomena for operators failing the cinematic curvature condition,* ProQuest LLC, Ann Arbor, MI, 2000, Thesis (Ph.D.)–The University of Wisconsin - Madison. MR 2701188

[17] ______, *Local smoothing for operators failing the cinematic curvature condition,* J. Math. Anal. Appl. **324** (2006), no. 1, 460–471. MR 2262484

[18] Richard B. Melrose and Michael E. Taylor, *Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle.* Adv. in Math. 55 (1985), no. 3, 242–315. MR 0778964.

[19] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge, *Wave front sets, local smoothing and Bourgain's circular maximal theorem,* Ann. of Math. (2) **136** (1992), no. 1, 207–218. MR 1173929

[20] ______, *Local smoothing of Fourier integral operators and Carleson-Sjölin estimates,* J. Amer. Math. Soc. **6** (1993), no. 1, 65–130. MR 1168960

[21] Detlef Müller and Andreas Seeger, *Singular spherical maximal operators on a class of two step nilpotent Lie groups,* Israel J. Math. **141** (2004), 315–340. MR 2063040

[22] E. K. Narayanan and S. Thangavelu, *An optimal theorem for the spherical maximal operator on the Heisenberg group,* Israel J. Math. **144** (2004), 211–219. MR 2121541

[23] Amos Nevo and Sundaram Thangavelu, *Pointwise ergodic theorems for radial averages on the Heisenberg group,* Adv. Math. **127** (1997), no. 2, 307–334. MR 1448717

[24] Yibiao Pan and Christopher D. Sogge, *Oscillatory integrals associated to folding canonical relations.* Colloq. Math. 60/61 (1990), no. 2, 413–419. MR 1096387

[25] Duong H. Phong and Elias M. Stein, *Radon transforms and torsion.* Internat. Math. Res. Notices **1991**, no. 4, 49–60. MR 1121165.

[26] Malabika Pramanik and Andreas Seeger, L^p *regularity of averages over curves and bounds for associated maximal operators,* Amer. J. Math. **129** (2007), no. 1, 61–103. MR 2288738

[27] ______, L^p*-Sobolev estimates for a class of integral operators with folding canonical relations,* Journal of Geometric Analysis, DOI 10.1007/s12220-020-00388-0. arXiv:1909.04173

[28] W. Schlag, *A geometric proof of the circular maximal theorem,* Duke Math. J. **93** (1998), no. 3, 505–533. MR 1626711

[29] Wilhelm Schlag and Christopher D. Sogge, *Local smoothing estimates related to the circular maximal theorem,* Math. Res. Lett. **4** (1997), no. 1, 1–15. MR 1432805

[30] Christopher D. Sogge, *Propagation of singularities and maximal functions in the plane,* Invent. Math. **104** (1991), no. 2, 349–376. MR 1098614

[31] Elias M. Stein, *Maximal functions. I. Spherical means* , Proc. Nat. Acad. Sci. U.S.A. **73** (1976), no. 7, 2174–2175. MR 0420116

[32] ______, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals* , Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192

David Beltran: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.
E-mail address: dbeltran@math.wisc.edu

Shaoming Guo: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.
E-mail address: shaomingguo@math.wisc.edu

Jonathan Hickman: School of Mathematics, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK.
E-mail address: jonathan.hickman@ed.ac.uk

Andreas Seeger: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.
E-mail address: seeger@math.wisc.edu