(Revised VERSION 4/17/2019)

ESTIMATES FOR MAXIMAL FUNCTIONS ASSOCIATED TO HYPERSURFACES IN $\mathbb{R}^3$ WITH HEIGHT $h < 2$: PART I

STEFAN BUSCHENHENKE, SPYRIDON DENDRINOS, ISROIL A. IKROMOV, AND DETLEF MÜLLER

Abstract. In this article, we continue the study of the problem of L^p-boundedness of the maximal operator $\mathcal{M}$ associated to averages along isotropic dilates of a given, smooth hypersurface S of finite type in 3-dimensional Euclidean space. An essentially complete answer to this problem had been given about seven years ago by the last named two authors in joint work with M. Kempe [IKM10] for the case where the height h of the given surface is at least two. In the present article, we turn to the case $h < 2$. More precisely, in this Part I, we study the case where $h < 2$, assuming that S is contained in a sufficiently small neighborhood of a given point $x^0 \in S$ at which both principal curvatures of S vanish. Under these assumptions and a natural transversality assumption, we show that, as in the case $h \geq 2$, the critical Lebesgue exponent for the boundedness of $\mathcal{M}$ remains to be $p_c = h$, even though the proof of this result turns out to require new methods, some of which are inspired by the more recent work by the last named two authors on Fourier restriction to S. Results on the case where $h < 2$ and exactly one principal curvature of S does not vanish at x^0 will appear elsewhere.

Contents

2010 *Mathematical Subject Classification.* 42B25.
Key words and phrases. Maximal operator, hypersurface, oscillatory integral, Newton diagram.
We acknowledge the support for this work by the Deutsche Forschungsgemeinschaft under DFG-Grant MU 761/11-1. This material is based in parts also upon work supported by the National Science Foundation under Grant No. DMS-1440140 while the last author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester.
The first author was supported by the European Research Council grant No. 307617.

1. Introduction

Let S be a smooth hypersurface in $\mathbb{R}^n$ and let $\rho \in C_0^\infty(S)$ be a smooth non-negative function with compact support. Consider the associated averaging operators $A_t, t > 0$, given by

$$A_t f(x) := \int_S f(x - ty)\rho(y)\, d\sigma(y),$$

where $d\sigma$ denotes the surface measure on S. The associated maximal operator is given by

$$\mathcal{M} f(x) := \sup_{t>0} |A_t f(x)|, \quad (x \in \mathbb{R}^n). \tag{1.1}$$

We remark that by testing $\mathcal{M}$ on the characteristic function of the unit ball in $\mathbb{R}^n$, it is easy to see that a necessary condition for $\mathcal{M}$ to be bounded on $L^p(\mathbb{R}^n)$ is that $p > n/(n-1)$, provided the transversality assumption 1.1 below is satisfied.

In 1976, E. M. Stein [S76] proved that, conversely, if S is the Euclidean unit sphere in $\mathbb{R}^n$, $n \geq 3$, then the corresponding spherical maximal operator is bounded on $L^p(\mathbb{R}^n)$ for every $p > n/(n-1)$. The analogous result in dimension $n = 2$ was later proven by J. Bourgain [Bou85]. The key property of spheres which allows to prove such results is the non-vanishing of the Gaussian curvature on spheres. These results became the starting point for intensive studies of various classes of maximal operators associated to subvarieties. Stein's monograph [S93] is an excellent reference to many of these developments.

In the joint work [IKM10] of the last-named two authors with M. Kempe, maximal functions $\mathcal{M}$ associated to smooth hypersurfaces of finite type in $\mathbb{R}^3$ had been studied under the following transversality assumption on S.

Assumption 1.1 (Transversality)**.** *The affine tangent plane $x + T_xS$ to S through x does not pass through the origin in $\mathbb{R}^3$ for every $x \in S$. Equivalently, $x \notin T_xS$ for every $x \in S$, so that $0 \notin S$ and x is transversal to S for every point $x \in S$.*

Let us fix a point $x^0 \in S$. We recall that the transversality assumption allows us to find a linear change of coordinates in $\mathbb{R}^3$ so that in the new coordinates S can locally be represented as the graph of a function ϕ, and that the norm of $\mathcal{M}$ when acting on $L^p(\mathbb{R}^3)$ is invariant under such a linear change of coordinates. More precisely, after applying a suitable linear change of coordinates to $\mathbb{R}^3$ we may assume that $x^0 = (0, 0, 1)$, and that within the neighborhood U, S is given as the graph

$$U \cap S = \{(x_1, x_2, 1 + \phi(x_1, x_2)) : (x_1, x_2) \in \Omega\}$$

of a smooth function $1 + \phi$ defined on an open neighborhood Ω of $0 \in \mathbb{R}^2$ and satisfying the conditions

$$\phi(0, 0) = 0,\ \nabla\phi(0, 0) = 0. \tag{1.2}$$

The measure $\mu = \rho d\sigma$ is then explicitly given by

$$\int f\, d\mu = \int f(x, 1 + \phi(x))\eta(x)\, dx,$$

with a smooth, non-negative bump function $\eta \in C_0^\infty(\Omega)$, and we may write for $(y, y_3) \in \mathbb{R}^2 \times \mathbb{R}$

$$A_t f(y, y_3) = f * \mu_t(y, y_3) = \int_{\mathbb{R}^2} f(y - tx, y_3 - t(1 + \phi(x)))\eta(x)\, dx, \tag{1.3}$$

where μ_t denotes the norm preserving scaling of the measure μ given by $\int f\, d\mu_t = \int f\big(tx, t(1 + \phi(x))\big)\eta(x)\, dx$.

Recall also from [IKM10] that the *height* of S at the point x^0 is defined by $h(x^0, S) := h(\phi)$, where $h(\phi)$ is the height of ϕ in the sense of Varchenko (which can be computed by means of Newton polyhedra attached to ϕ). The height is invariant under affine linear changes of coordinates in the ambient space $\mathbb{R}^3$.

In [IKM10] the authors had given an essentially complete answer to the problem of L^p-boundedness of $\mathcal{M}$ when $h(x^0, S)$ or p are greater or equal to 2. More precisely, if $h(x^0, S) \geq 2$, and if the density ρ is supported in a sufficiently small neighborhood of x^0, then the condition $p > h(x^0, S)$ is sufficient for $\mathcal{M}$ to be L^p-bounded, and this result is sharp (with the possible exception of the endpoint $p_c = h(x^0, S)$, when S is non-analytic). For an alternative approach to some of these results based on "damping" techniques, see also [Gr13]. Matters change dramatically when the transversality assumption fails, as has been shown by E. Zimmermann in his doctoral thesis [Z5]. Zimmermann studied the case where the hypersurface passes through the origin and proved, among other things, that for analytic S and $\operatorname{supp} \rho$ sufficiently small, the condition $p > 2$ is always sufficient for the L^p- boundedness of $\mathcal{M}$.

In the present article, we return to this problem and look at the case where $h(x^0, S) < 2$. It then turns out that there is a big difference in the behaviour of the associated maximal operator, depending on how many of the principal curvatures of S do vanish at x^0 (instances of this phenomenon have already been observed in articles by Nagel, Seeger, Wainger [NSeW93], and Iosevich and Sawyer [ISa96], [ISa97], [ISaSe99]). The case where both principal curvatures do not vanish at x^0 is classical, and here the condition $p > 3/2$ is necessary and sufficient for the L^p-boundedness of $\mathcal{M}$, exactly as in the case of the 2-sphere (see Greenleaf [Gl81]). Notice that in this case $h(x^0, S) = 1 < 3/2$, so that, unlike the case where $h(x^0, S) \geq 2$, the height is not the controlling quantity for the maximal operator. Indeed, as mentioned before, the condition $p > 3/2$ is seen to be necessary by testing $\mathcal{M}$ on characteristic functions of small balls, whereas the notion of height is rather related to testing on characteristic functions of the intersection of a ball with a very thin neighborhood of some hyperplane.

We shall here mainly consider the case where both principal curvatures of S do vanish at x^0, i.e., when $D^2\phi(0,0) = 0$. In this case, it turns out that the height is still the controlling quantity. More precisely, our main theorem states the following:

Theorem 1.2. *Assume that S is a smooth, finite-type hypersurface in $\mathbb{R}^3$ satisfying the transversality assumption 1.1, and let $x^0 \in S$ be a given point at which $h(x^0, S) < 2$ and both principal curvatures of S do vanish.*

Then there exists a neighborhood $U \subset S$ of the point x^0 such that for every non-negative density $\rho \in C_0^\infty(U)$ the associated maximal operator $\mathcal{M}$ is bounded on $L^p(\mathbb{R}^3)$ whenever $p > h(x^0, S)$.

The condition $p > h(x^0, S)$ is indeed also necessary, as the following result shows, which does not require that both principal curvatures of S vanish at x^0.

Theorem 1.3. *Assume that the maximal operator $\mathcal{M}$ is L^p-bounded, and that S satisfies the transversality assumption 1.1. Then, for every point $x^0 \in S$ at which $h(x^0, S) < 2$ and $\rho(x_0) \neq 0$, we necessarily have $p > h(x^0, S)$.*

Notice an interesting difference between the cases $h(x^0, S) \geq 2$ and $h(x^0, S) < 2$: in the first case, studied in [IKM10], the necessity of the condition $p > h(x^0, S)$ when $\rho(x^0) \neq 0$ could be verified for analytic hypersurfaces, but not for all classes of smooth, finite type hypersurfaces (where the endpoint $p = h(x^0, S)$ remained open in certain situations). Indeed, problems with the L^p-boundedness of $\mathcal{M}$ at the endpoint $p = h(x^0, S)$ may arise, e.g., for ϕ of the form $\phi(x_1, x_2) = x_2^2 + f(x_1)$, where f is flat at the origin (we refer to the examples in [IKM10], Remark 12.3). Note, however, that in this example $h(\phi) = 2$, and we shall see that such kind of situation can never arise when $h(x^0, S) < 2$.

The case where only one principal curvature of S vanishes, and the other one not (which is the case of singularities of type A_n in the sense of Arnol'd - compare to Theorem 3.1) turns out to be the most difficult one to analyze, and we shall return to this problem in subsequent work. Indeed, it appears that in those situations where ϕ has a singularity of type $A_n, n \geq 3$, typically the optimal necessary conditions for L^p-boundedness of $\mathcal{M}$ rather seem to come from testing on characteristic functions of a very narrow tubular neighborhood of some line segment in $\mathbb{R}^3$. Here, we shall only look at one instance of singularities of type A_n, namely when $n = 2$ (see Theorem 7.1), since results on surfaces with A_2 - type singularity turn out to be of great relevance also to the study of D_4^+-type singularities, which are included in Theorem 1.2.

Before turning to the proof of Theorem 1.2, let us first briefly recall some relevant notation and results from [IKM10] (compare also the monograph [IM16]). These sources should also be consulted for further details and references to the topic. Our analysis in the present paper will indeed take advantage of several techniques and results developed in those articles.

2. Some background on Newton diagrams and adapted coordinates

We first recall some basic notions from [IM11], which essentially go back to A.N. Varchenko [V76]. Let ϕ be a smooth real-valued function defined on an open neighborhood Ω of the origin in $\mathbb{R}^2$ with $\phi(0,0) = 0$, $\nabla\phi(0,0) = 0$. Consider the associated Taylor series

$$\phi(x_1, x_2) \sim \sum_{\alpha_1,\alpha_2=0}^{\infty} c_{\alpha_1,\alpha_2} x_1^{\alpha_1} x_2^{\alpha_2}$$

of ϕ centered at the origin. The set

$$\mathcal{T}(\phi) := \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 : c_{\alpha_1,\alpha_2} = \frac{1}{\alpha_1!\alpha_2!}\partial_1^{\alpha_1}\partial_2^{\alpha_2}\phi(0,0) \neq 0\}$$

will be called the *Taylor support* of ϕ at $(0, 0)$. We shall always assume that the function ϕ is of *finite type* at every point, i.e., that the associated graph S of ϕ is of finite type. Since we are also assuming that $\phi(0,0) = 0$ and $\nabla\phi(0,0) = 0$, the finite type assumption at the origin just means that

$$\mathcal{T}(\phi) \neq \emptyset.$$

The *Newton polyhedron* $\mathcal{N}(\phi)$ of ϕ at the origin is defined to be the convex hull of the union of all the quadrants $(\alpha_1, \alpha_2) + \mathbb{R}_+^2$ in $\mathbb{R}^2$, with $(\alpha_1, \alpha_2) \in \mathcal{T}(\phi)$. The associated *Newton diagram* $\mathcal{N}_d(\phi)$ in the sense of Varchenko [V76] is the union of all compact faces of the Newton polyhedron; here, by a *face,* we shall mean an edge or a vertex.

We shall use coordinates (t_1, t_2) for points in the plane containing the Newton polyhedron, in order to distinguish this plane from the (x_1, x_2) - plane.

The *Newton distance* in the sense of Varchenko, or shorter *distance*, $d = d(\phi)$ between the Newton polyhedron and the origin is given by the coordinate d of the point (d, d) at which the bi-sectrix $t_1 = t_2$ intersects the boundary of the Newton polyhedron.

The principal face $\pi(\phi)$ of the Newton polyhedron of ϕ is the face of minimal dimension containing the point (d, d). We shall call the series

$$\phi_{\mathrm{pr}}(x_1, x_2) := \sum_{(\alpha_1,\alpha_2)\in\pi(\phi)} c_{\alpha_1,\alpha_2} x_1^{\alpha_1} x_2^{\alpha_2}$$

the *principal part* of ϕ. In case that $\pi(\phi)$ is compact, ϕ_{pr} is a mixed homogeneous polynomial; otherwise, we shall consider ϕ_{pr} as a formal power series.

Note that the distance between the Newton polyhedron and the origin depends on the chosen local coordinate system in which ϕ is expressed. By a *local coordinate system (at the origin)* we shall mean a smooth coordinate system defined near the origin which preserves 0. The *height* of the smooth function ϕ is defined by

$$h = h(\phi) := \sup\{d_y\},$$

where the supremum is taken over all local coordinate systems $y = (y_1, y_2)$ at the origin, and where d_y is the distance between the Newton polyhedron and the origin in the coordinates y.

A given coordinate system x is said to be *adapted* to ϕ if $h(\phi) = d_x$. In [IM11] we proved that one can always find an adapted local coordinate system in two dimensions, thus generalizing the fundamental work by Varchenko [V76] who worked in the setting of real-analytic functions ϕ (see also [PSS99]).

Notice that if the principal face of the Newton polyhedron $\mathcal{N}(\phi)$ is a compact edge, then it lies on a unique *principal line*

$$L := \{(t_1, t_2) \in \mathbb{R}^2 : \kappa_1 t_1 + \kappa_2 t_2 = 1\},$$

with $\kappa_1, \kappa_2 > 0$. By permuting the coordinates x_1 and x_2, if necessary, we shall always assume that $\kappa_1 \le \kappa_2$. The weight $\kappa = (\kappa_1, \kappa_2)$ will be called the *principal weight* associated to ϕ. It induces dilations $\delta_r(x_1, x_2) := (r^{\kappa_1} x_1, r^{\kappa_2} x_2),\ r > 0$, on $\mathbb{R}^2$, so that the principal part ϕ_{pr} of ϕ is κ-homogeneous of degree one with respect to these dilations, i.e., $\phi_{\mathrm{pr}}(\delta_r(x_1, x_2)) = r\phi_{\mathrm{pr}}(x_1, x_2)$ for every $r > 0$, and we find that

$$d = \frac{1}{\kappa_1 + \kappa_2} = \frac{1}{|\kappa|}.$$

Notice that $1/|\kappa| \le h$.

More generally, assume that $\kappa = (\kappa_1, \kappa_2)$ is any weight with $0 < \kappa_1 \le \kappa_2$ such that the line $L_\kappa := \{(t_1, t_2) \in \mathbb{R}^2 : \kappa_1 t_1 + \kappa_2 t_2 = 1\}$ is a supporting line to the Newton polyhedron $\mathcal{N}(\phi)$ of ϕ (recall that a *supporting line* to a convex set K in the plane is a line such that K is contained in one of the two closed half-planes into which the line divides the plane and such that this line intersects the boundary of K). Then $L_\kappa \cap \mathcal{N}(\phi)$ is a face of $\mathcal{N}(\phi)$, i.e., either a compact edge or a vertex, and the *κ-principal part* of ϕ

$$\phi_\kappa(x_1, x_2) := \sum_{(\alpha_1, \alpha_2) \in L_\kappa} c_{\alpha_1, \alpha_2} x_1^{\alpha_1} x_2^{\alpha_2}$$

is a non-trivial polynomial which is κ-homogeneous of degree 1 with respect to the dilations associated to this weight as before, which can be factorized in a similar way as in [IM11]. By definition, we then have

$$\phi(x_1, x_2) = \phi_\kappa(x_1, x_2) + \text{ terms of higher } \kappa\text{-degree}.$$

Adaptedness of a given coordinate system can be verified by means of the following proposition (see [IM11]):

If P is any given polynomial which is κ-homogeneous of degree one (such as $P = \phi_{\mathrm{pr}}$), then we denote by

$$n(P) := \operatorname{ord}_{S^1} P \tag{2.1}$$

the maximal order of vanishing of P along the unit circle S^1. Observe that by homogeneity, the Taylor support $\mathcal{T}(P)$ of P is contained in the face $L_\kappa \cap \mathcal{N}(P)$ of $\mathcal{N}(P)$. We therefore define the *homogeneous distance* of P by $d_h(P) := 1/(\kappa_1 + \kappa_2) = 1/|\kappa|$. Notice that $(d_h(P), d_h(P))$ is just the point of intersection of the line L_κ with the bi-sectrix $t_1 = t_2$, and that $d_h(P) = d(P)$ if and only if $L_\kappa \cap \mathcal{N}(P)$ intersects the bi-sectrix. We remark that the height of P can then easily be computed by means of the formula

$$h(P) = \max\{n(P), d_h(P)\} \tag{2.2}$$

(see Corollary 3.4 in [IM11]). Moreover, in [IM11] (Corollary 4.3 and Corollary 5.3), we also proved the following characterization of adaptedness of a given coordinate system:

Proposition 2.1. *The coordinates x are adapted to ϕ if and only if one of the following conditions is satisfied:*

(a) *The principal face $\pi(\phi)$ of the Newton polyhedron is a compact edge, and $n(\phi_{\mathrm{pr}}) \le d(\phi)$.*

(b) $\pi(\phi)$ *is a vertex.*
(c) $\pi(\phi)$ *is an unbounded edge.*

We also note that in case (a) we have $h(\phi) = h(\phi_{\rm pr}) = d_h(\phi_{\rm pr})$. Moreover, it can be shown that we are in case (a) whenever $\pi(\phi)$ is a compact edge and $\kappa_2/\kappa_1 \notin \mathbb{N}$; in this case we even have $n(\phi_{\rm pr}) < d(\phi)$ (cf. [IM11], Corollary 2.3).

3. Normal forms of ϕ under linear coordinate changes when $h(\phi) < 2$, and proof of Theorem 1.3

Our approach will be based on the description of normal forms of ϕ under linear coordinate changes given by the next theorem. The designation of the type of singularity that we list below corresponds to Arnol'd's classification of singularities (cf. [AGV88] and [Dui74]). However, Arnol'd's normal forms are achieved by means of non-linear coordinate changes, and since we shall be in need of very precise information on those coordinate changes, we shall present normal forms which give such precise information ("non-linear shears" will indeed always lead to adapted coordinates). Our normal forms will easily follow by expanding a bit on Proposition 2.11 in [IM16] in combination with the proof of Corollary 7.4 in [IKM10].

Notice that in Theorem 1.2, we are dealing with Case 2 of the following theorem.

Theorem 3.1. *Assume that $h(\phi) < 2$, and that ϕ has a degenerate critical point at the origin. Then, after applying a suitable linear change of coordinates, ϕ can be written on a sufficiently small neighborhood of the origin in one of the following forms:*

Case 1 (Type A). $\operatorname{rank} D^2\phi(0,0) = 1$.

$$\phi(x_1, x_2) = b(x_1, x_2)(x_2 - \psi(x_1))^2 + b_0(x_1), \tag{3.1}$$

where b, b_0 and ψ are smooth functions, and $b(0,0) \neq 0$.

Moreover, either ψ is flat at 0, or $\psi(x_1) = cx_1^m + O(x_1^{m+1})$, with $c \neq 0$ and $m \geq 2$, and
$b_0(x_1) = x_1^n\beta(x_1)$, *with* $\beta(0) \neq 0$ *and* $n \geq 3$ *(singularity of type A_{n-1}).*

The coordinates (x_1, x_2) are then adapted to ϕ if and only if $n \leq 2m$, with the understanding that $m := \infty$ if ψ is flat at the origin. If $n \leq 2m$, then the principal weight is given by $\kappa := (1/n, 1/2)$, and we have

$$d(\phi) = h(\phi) = \frac{2n}{n+2},$$

and if $n > 2m$, then the principal weight is given by $\kappa := (1/(2m), 1/2)$, and Newton distance and height are given by

$$d(\phi) = \frac{2m}{m+1}, \qquad h(\phi) = \frac{2n}{n+2}.$$

Case 2. $\operatorname{rank} D^2\phi(0,0) = 0$. *Here, we distinguish two subcases.*

(i) Type D. *ϕ is still of the form (3.1), with smooth functions b, b_0 and ψ as before, but now with $b(0,0) = 0$, and more precisely*

$$b(x_1, x_2) = x_1 b_1(x_1, x_2) + x_2^2 b_2(x_2), \tag{3.2}$$

where b_1 and b_2 are smooth functions, with $b_1(0,0) \neq 0$, and
$b_0(x_1) = x_1^n\beta(x_1)$, *where $\beta(0) \neq 0$ and $n \geq 3$ (singularity of type D_{n+1}).*

Here, the coordinates (x_1, x_2) are adapted to ϕ if and only if $n \leq 2m + 1$, with the understanding that $m := \infty$ if ψ is flat at the origin. If $n \leq 2m + 1$, then the principal weight is given by $\kappa := (1/n, (n-1)/(2n))$, and Newton distance and height are given by

$$d(\phi) = h(\phi) = \frac{2n}{n+1},$$

and if $n > 2m+1$, *then the principal weight is given by* $\kappa := (1/(2m+1), m/(2m+1))$, *and Newton distance and height are given by*

$$d(\phi) = \frac{2m+1}{m+1}, \qquad h(\phi) = \frac{2n}{n+1}.$$

(ii) Type E. ϕ *can be written as*

$$\phi(x) = \phi_{\rm pr}(x) + \phi_r(x),$$

where the remainder $\phi_r(x)$ *comprises all terms of* κ*-degree strictly bigger than* 1, *and where we may assume that the principal part* $\phi_{\rm pr}$ *of* ϕ *and the principal weight* κ *are from the following list:*

$\phi_{\rm pr}(x_1, x_2) = x_2^3 \pm x_1^4$, *with* $\kappa := (1/4, 1/3)$, *(singularity of type* E_6*),*
$\phi_{\rm pr}(x_1, x_2) = x_2^3 + x_1^3 x_2$, *with* $\kappa := (2/9, 1/3)$, *(singularity of type* E_7*),*
$\phi_{\rm pr}(x_1, x_2) = x_2^3 + x_1^5$, *with* $\kappa := (1/5, 1/3)$. *(singularity of type* E_8*).*

In all these cases, the coordinates (x_1, x_2) *are adapted to* ϕ. *Moreover,*

$$d(\phi) = h(\phi) = \begin{cases} 12/7 & \text{for type } E_6, \\ 9/5 & \text{for type } E_7, \\ 15/8 & \text{for type } E_8. \end{cases}$$

Notice that in all cases the principal face of the Newton polyhedron of ϕ is a compact edge, so that $d = d(\phi) = 1/|\kappa|$.

Remark 3.2. Following [IM16], we shall call the function ψ (respectively its graph) in (3.1) the *principal root jet* (compare [IM11] for the general definition of this notion). Notice that the coordinates $(y_1, y_2) := (x_1, x_2 - \psi(x_1))$ are adapted to ϕ for singularities of type A or D, and that in these coordinates ϕ is of the form

$$\phi^a(y_1, y_2) = b^a(y_1, y_2) y_2^2 + b_0(y_1), \tag{3.3}$$

where the function $b^a(y_1, y_2)$ has the same properties as the one described for $b(x_1, x_2)$.

Remark 3.3. For later purposes, let us observe that if ϕ is of type D_4, then the coordinates (x_1, x_2) are already adapted, and we may assume that the principal part of ϕ is of the form

$$\phi_{\rm pr}(x_1, x_2) = x_1 x_2^2 \pm x_1^3 = x_1(x_2^2 \pm x_1^2). \tag{3.4}$$

If $\phi_{\rm pr}(x_1, x_2) = x_1(x_2^2 + x_1^2)$, then we say that ϕ is of type D_4^+, and if $\phi_{\rm pr}(x_1, x_2) = x_1(x_2^2 - x_1^2)$, then we call ϕ of type D_4^-.

Proof. The statements in Case 1 follow immediately from the proof of Proposition 2.11 in [IM16]. Notice that if the function b_0 were flat at the origin, then we would have $h(\phi) = 2$. Thus singularities of type A_∞ are excluded here. Moreover, if we had $n = 2$, then the origin would be a non-degenerate critical point of ϕ, contrary to our assumption.

As for Case 2, following again the proof of Proposition 2.11 in [IM16], let us look at the polynomial P_3, which denotes the homogeneous part of degree 3 of the Taylor polynomial of ϕ with respect to the origin. Let us also denote by $n(P_3)$ the maximal multiplicity of real roots of P_3.

If $n(P_3) = 2$, then we can just follow the discussion in [IM16] in order to see that ϕ is of type D_{n+1}, with $n \geq 3$ (note that if $n = 2$, then we would be in Case 2). As before, the type D_∞ is excluded here, since in that case we would have $h(\phi) = 2$.

If $n(P_3) = 1$, then we have shown in [IM16] that P_3 must be of the form $P_3(x_1, x_2) = x_1(x_2 - \alpha x_1)(x_2 - \beta x_1)$, where either $\alpha \neq \beta$ are both real, or $\alpha = \overline{\beta}$ are non-real. It is easy to see that this case can easily be reduced to the form described by (3.4) by means of a linear change of coordinates.

Finally, if $n(P_3) = 3$, then as in the proof of Proposition 2.11 in [IM16] we may assume that $P_3(x_1, x_2) = x_2^3$. But then the proof of Corollary 7.4 in [IKM10] shows that ϕ is of type E_6, E_7 or E_8.

The remaining statements are easily verified using Proposition 2.1. Q.E.D.

Given these normal forms for ϕ, it is now easy to prove the necessary condition for L^p-boundedness of the maximal operator $\mathcal{M}$ in Theorem 1.3. Indeed, from our normal forms one deduces that the principal face of the Newton polyhedron of ϕ, when expressed as ϕ^a in adapted coordinates, is a compact edge (compare (3.3) in Remark 3.2 for singularities of type A and D). Then Theorem 1.3 is an immediate consequence of Proposition 12.1 in [IKM10] in combination with a result by Iosevich and Sawyer [ISa96], namely Theorem 1.5 in [IKM10].

In this paper, we shall mainly study maximal functions associated with Case 2 in Theorem 3.1 where $D^2\phi(0,0) = 0$. However, as we shall see, the study of D_4^+ - type singularity will also require the understanding of maximal functions associated to surfaces with A_2 - type singularities.

In the next section, we shall provide an auxiliary estimate for maximal functions associated to families of hypersurfaces depending and some perturbation parameter σ and some large translation parameter T (translations of the hypersurfaces in transversal directions). The corresponding estimates will become useful in many situations.

4. Auxiliary estimates for maximal operators

4.1. **An estimate for maximal operators depending on parameters.** Let us consider a smooth family $S^{\sigma,T}$ of hypersurfaces in $\mathbb{R}^{n+1}$ given as graphs

$$S^{\sigma,T} := \{(x, T + \phi(x,\sigma)) : x \in U\}, \tag{4.1}$$

where $\phi = \phi(x,\sigma)$ is a smooth, real-valued function defined on an open neighborhood $U \times V$ of a given point (x_0,σ_0) in $\mathbb{R}^n \times \mathbb{R}^m$, and T is a large real translation parameter. Denote furthermore by $\mu^{\sigma,T}$ the surface-carried measure on $S^{\sigma,T}$ defined by

$$\int f\, d\mu^{\sigma,T} := \int f(x, T + \phi(x,\sigma))\, a(x,\sigma)dx,$$

where $a(x,\sigma)$ is a nonnegative smooth function with compact support in $U \times V$. For $t > 0$ we denote by $\mu_t^{\sigma,T}$ the measure-preserving scaling of $\mu^{\sigma,T}$ given by

$$\int f\, d\mu_t^{\sigma,T} = \int f\big(tx, t(T + \phi(x,\sigma))\big)\, a(x,\sigma)dx,$$

and consider the averaging operator

$$A_t^{\sigma,T} f(y, y_{n+1}) := f * \mu_t^{\sigma,T}(y, y_{n+1}) = \int_{\mathbb{R}^n} f\big(y - tx, y_{n+1} - t(T + \phi(x,\sigma))\big)\, a(x,\sigma)dx. \tag{4.2}$$

Define the associated maximal operator by

$$\mathcal{M}^{\sigma,T} f(y, y_{n+1}) := \sup_{t>0} |A_t^{\sigma,T} f(y, y_{n+1})|, \qquad (y, y_{n+1}) \in \mathbb{R}^{n+1}. \tag{4.3}$$

Proposition 4.1. *Assume that the uniform estimate*

$$\left| \int e^{i(\xi\cdot x + \xi_{n+1}\phi(x,\sigma))} \eta(x) dx \right| \le C \frac{\|\eta\|_{C^{n+1}}}{(1 + |\xi_{n+1}|)^{\gamma}}, \qquad (\xi, \xi_{n+1}) \in \mathbb{R}^{n+1}, \tag{4.4}$$

holds true for every C^{n+1}- function η with compact support in U, where C is independent of η and σ, and that $n/2 \ge \gamma > 1/2$. Then for every $p > 1 + 1/(2\gamma)$ and $|T| \ge 1$ the maximal operator $\mathcal{M}^{\sigma,T}$ is bounded on $L^p(\mathbb{R}^{n+1})$, with norm

$$\|\mathcal{M}^{\sigma,T}\|_{p\to p} \le C_p |T|^{\frac{1}{p}},$$

where the constant C_p is independent of σ and T.

Proof. Since for the proof we can essentially follow a by now well-known pattern (see, e.g., [S93], Ch. XI. 3), we shall only sketch the argument. Let us assume without loss of generality that $(x_0, \sigma_0) = (0, 0)$. Moreover, in order to facilitate the notation, we shall drop superscripts σ, T and write μ for $\mu^{\sigma,T}$, etc..

We choose smooth non-negative bump functions χ_0 supported in $[-1, 1]$ and χ_1 supported in $[-2, -1/2] \cup [1/2, 2]$ such that

$$\sum_{l=0}^{\infty} \chi_l(s) = 1 \quad \text{for all} \quad s \in \mathbb{R},$$

where $\chi_l(x) := \chi_1(2^{1-l}x)$ for $l \geq 1$. Then we perform the inhomogeneous Littlewood-Paley decomposition

$$\mu = \sum_{l=0}^{\infty} \mu^l, \qquad \text{where} \qquad \widehat{\mu^l}(\xi, \xi_{n+1}) := \chi_l(\xi_{n+1})\hat{\mu}(\xi, \xi_{n+1}),$$

and denote by $\mathcal{M}^l$ the maximal operator associated to μ^l in place of μ, i.e., $\mathcal{M}^l f = \sup_{t>0} |A_t^l f|$, with $A_t^l f(y, y_{n+1}) := f * \mu_t^l(y, y_{n+1})$.

We first estimate $\mathcal{M}^l$ on L^2. By writing $t = t'2^{-j}$ with $t' \in [1, 2[$ and $j \in \mathbb{Z}$, it is easily seen that

$$\mathcal{M}^l f(y, y_{n+1}) \leq \Big(\sum_{j \in \mathbb{Z}} \sup_{t \in [1,2[} |A_{t2^{-j}}^l f(y, y_{n+1})|^2 \Big)^{1/2},$$

and since $\widehat{\mu^l}(t2^{-j}\,\cdot)$ is supported in the set where $|\xi_{n+1}| \sim 2^{l+j}$, we may replace f on the right-hand side of this inequality by $\sum_{k=-2}^{2} \Delta_{l+j+k} f$, with $\widehat{\Delta_m f}(\xi, \xi_{n+1}) := \chi_m(\xi_{n+1})\hat{f}(\xi, \xi_{n+1})$. Since the functions $\Delta_m f$ have almost disjoint Fourier supports, we easily see by means of Plancherel's theorem that it suffices to prove an estimate for the local maximal functions $\mathcal{M}_j^l f = \sup_{1 \leq t < 2} |A_{t2^{-j}}^l f|$, of the form

$$\|\mathcal{M}_j^l f\|_2 \leq C_l \|f\|_2, \qquad \text{with } C_l \text{ independent of } j, \tag{4.5}$$

in order to derive a corresponding estimate $\|\mathcal{M}^l f\|_2 \leq C C_l \|f\|_2$. But, notice that for $l \geq 1$

$$\widehat{\mu^l}(t2^{-j}(\xi, \xi_{n+1})) = \chi_1(t2^{-j-l}\xi_{n+1}) \int e^{it2^{-j}(\xi \cdot x + \xi_{n+1}(T + \phi(x,\sigma)))} a(x, \sigma)\, dx.$$

Our assumption (4.4) thus easily implies that, for $1 \leq t < 2$,

$$|\widehat{\mu^l}(t2^{-j}(\xi, \xi_{n+1}))| \leq C2^{-l\gamma},$$
$$|\partial_t \widehat{\mu^l}(t2^{-j}(\xi, \xi_{n+1}))| \leq C2^{-l\gamma}2^l|T|$$

(observe here that if $|\xi|/|\xi_{n+1}|$ is sufficiently large, then iterated integrations by parts lead to the stronger estimate $\left|\int e^{i(\xi \cdot x + \xi_{n+1}\phi(x,\sigma))}\psi(x)dx\right| \leq C\frac{\|\psi\|_{C^n}}{(1+|\xi|)^n}$.) By means of a variant of the Sobolev embedding theorem (compare [S93], Ch. XI. 3.2) and Plancherel's theorem we then find that the norm of $\mathcal{M}_j^l$ can be estimated by the geometric mean of the two right-hand sides of these estimates, so that we may choose $C_l = C2^{-l\gamma}2^{l/2}|T|^{1/2}$ in (4.5). Consequently,

$$\|\mathcal{M}^l\|_{2 \to 2} \leq C2^{l/2}2^{-l\gamma}|T|^{1/2}. \tag{4.6}$$

As for the estimation of $\mathcal{M}_l$ on L^1, observe that, except for some Schwartz tail, μ^l is essentially supported in a cuboid of dimensions comparable to $1 \times \cdots \times 1 \times |T|$, and that $\|\mu^l\|_\infty \leq C2^l$. This allows to dominate $\mathcal{M}^l f$ by $C2^l|T|M_{HL}(|f|)$, where M_{HL} denotes the Hardy-Littlewood maximal operator. Therefore we have the weak-type estimate

$$\|\mathcal{M}^l\|_{L^1 \to L^{1,\infty}} \leq C2^l|T|. \tag{4.7}$$

From these two estimates (4.6) and (4.7) we obtain by means of Marcinkiewicz' interpolation theorem (see, e.g., [G09]) that

$$\|\mathcal{M}^l\|_{L^p\to L^p} \le C_p 2^{l(\frac{1}{p}-2\gamma(1-\frac{1}{p}))}|T|^{1/p}.$$

Very similar arguments apply when $l=0$. Thus, if $p>1+1/(2\gamma)$, these estimates sum in l and we arrive at the desired estimate $\|\mathcal{M}\|_{L^p\to L^p}\le C_p|T|^{1/p}$ for this range of p's. Q.E.D.

4.2. **A variation on Hardy-Littlewood's maximal operator.** The following result, which may also be of independent interest, will become relevant to our "Airy-type" analysis in Section 7.

If A is a bounded Lebesgue measurable subset of $\mathbb{R}^n$, then we denote by $\mathcal{M}_A$ the corresponding maximal operator

$$\mathcal{M}_A f(x) := \sup_{t>0}\int_A |f(x+ty)|\,dy, \qquad f\in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

In particular, if $A=B_1(0)$ is the Euclidean unit ball, then $\mathcal{M}_A$ is the Hardy-Littlewood maximal operator $\mathcal{M}_{\mathrm{HL}}$.

Denote further by $\pi:\mathbb{R}^n\setminus\{0\}\to S^{n-1}$ the *spherical projection* onto the unit sphere S^{n-1}, given by $\pi(x):=x/|x|$, and by $|\pi(A)|$ the $n-1$-dimensional volume of this set with respect to the surface measure on the sphere.

Proposition 4.2. *Assume that A is an open subset of $\mathbb{R}^n$ contained in the annulus $R\le|x|<2R$, where $R>0$. Then, for $1<p\le\infty$, we have that*

$$\|\mathcal{M}_A f\|_{L^p\to L^p}\le C_pR^n|\pi(A)|. \tag{4.8}$$

Moreover, if $p=1$, then we have that $\|\mathcal{M}_Af\|_{L^1\to L^{1,\infty}}\ge c_1R^n|\pi(A)|$. Here, $c_1>0$ and C_p are positive constants which are independent of the set A.

Remarks: The estimate (4.8) gets only sharp as $p\to1$. For example, if $A=S_\delta,\delta>0$, is the δ-neighborhood of the sphere S^{n-1}, then the boundedness of the spherical maximal operator on L^p for $p>n/(n-1)$ implies that in this range of p's, $\|\mathcal{M}_{S_\delta}f\|_{L^p\to L^p}\le C\delta$, whereas $\pi(S_\delta)=S^{n-1}$.

We do not know if for $p=1$ a weak-type estimate of the form $\|\mathcal{M}_Af\|_{L^1\to L^{1,\infty}}\le C_1|\pi(A)|$ holds true. The argument that we shall use in the proof does not allow to show this, since weak-type estimates only sum with a kind of logarithmic loss (cf. [SW69])

Proof. Scaling by $1/R$, we can easily reduce considerations to the case $R=1$. By an ε-*tube* ($\varepsilon>0$) we shall mean in this proof any tube of length 6 and radius ε centered at the origin. By a standard Whitney decomposition of the open subset $\pi(A)$ of the sphere, we may find a sequence T_j of such tubes, where T_j is an ε_j-tube, such that the T_j cover the set $\pi(A)$ with bounded overlap. We can do this even in such a way that the T_j also cover the set A. Then

$$\mathcal{M}_Af(x)\le\sum_j\mathcal{M}_{T_j}f(x).$$

Moreover, a simple scaling argument allows to compare the operators $\mathcal{M}_{T_j}$ with the Hardy-Littlewood maximal operator, and we find that $\|\mathcal{M}_{T_j}f\|_{L^p\to L^p}\le C'_p|T_j|=C''_p\varepsilon_j^{n-1}\le C'''_p|T_j\cap S^{n-1}|$. This implies that $\|\mathcal{M}_Af\|_{L^p\to L^p}\le C_p|\pi(A)|$.

To prove an inverse estimate for $p=1$, let $\varepsilon>0$, and put $f:=\chi_{B_\varepsilon(0)}$. Denote by A_ε the set of all points in A whose ε-neighborhood is also contained in A. Then, for every point $a\in A_\varepsilon$, we see that $\mathcal{M}_Af(x)\ge C\varepsilon^n$ on 1/8'th of the ε-tube passing through $-a$. Therefore $\mathcal{M}_Af(x)\ge C\varepsilon^n=C'\|f\|_1$ on a set of measure $\ge c|\pi(A_\varepsilon)|$, where $c>0$ is a fixed constant. This implies that $\|\mathcal{M}_A\|_{L^1\to L^{1,\infty}}\ge c_1|\pi(A_\varepsilon)|$, for every $\varepsilon>0$. The asserted inverse estimate for $p=1$ follows, since $|\pi(A_\varepsilon)|\to|\pi(A)|$ as $\varepsilon\to0$. Q.E.D.

5. Estimation of the maximal operator $\mathcal{M}$ in the presence of a linear coordinate system which is adapted to ϕ

We begin the proof of Theorem 1.2 with the discussion of the cases where the coordinates (x_1, x_2) in Theorem 3.1 are adapted to ϕ, which strongly facilitates the arguments. Recall that in these cases

$$d = \frac{1}{|\kappa|} = h.$$

Following the approach in Section 7 of [IKM10], given the principal weight κ, we first perform a dyadic decomposition by means of the dilations $\delta_r(x_1, x_2) := (r^{\kappa_1} x_1, r^{\kappa_2} x_2)$, $r > 0$, associated to κ. To this end, we choose a smooth non-negative function χ_1 supported in the annulus $\mathcal{A} := \{1 \le |x| \le R\}$ satisfying

$$\sum_{k=k_0}^{\infty} \chi_1(\delta_{2^k} x) = 1 \quad \text{for} \quad 0 \ne x \in \Omega.$$

Notice that by choosing Ω small, we can choose $k_0 \in \mathbb{N}$ as large as we need. We then decompose the measure $\mu = \rho d\sigma$, which is explicitly given by

$$\int f\, d\mu = \int f(x, 1 + \phi(x))\eta(x)\, dx,$$

with a smooth, non-negative bump function $\eta \in C_0^\infty(\Omega)$, accordingly as

$$\mu = \sum_{k=k_0}^{\infty} \tilde{\mu}_k,$$

with

$$\int f\, d\tilde{\mu}_k = \int f(x, 1 + \phi(x))\, \eta(x)\chi_1(\delta_{2^k} x) dx.$$

It will then suffice to derive suitable L^p-estimates for the maximal operators $\sup_{t>0} |f*(\tilde{\mu}_k)_t|$. Applying a straight-forward L^p-isometric re-scaling to them by means of the dilations $\delta_{2^{-k}}$, we may assume that these are of the form $2^{-|\kappa|k}\mathcal{M}_k f$, where

$$\mathcal{M}_k f(y, y_3) := \sup_{t>0} |f * (\mu_k)_t(y, y_3)|$$

and

$$\int f\, d\mu_k := \int f(x, 2^k + \phi^k(x))\, \eta(\delta_{2^{-k}} x)\chi_1(x) dx.$$

Here we have set

$$\phi^k(x) := 2^k \phi(\delta_{2^{-k}}(x)) = \phi_{\rm pr}(x) + 2^k \phi_r(\delta_{2^{-k}}(x)).$$

Notice that the perturbation term $2^k\phi_r(\delta_{2^{-k}}(\cdot))$ is of order $O(2^{-\varepsilon k})$ for some $\varepsilon > 0$ in any C^M-norm. To express this fact, we shall for the remainder of this article use the short-hand notation $2^k\phi_r(\delta_{2^{-k}}(\cdot)) = O(2^{-\varepsilon k})$. To summarize, we shall then have

$$\|\mathcal{M}\|_{L^p \to L^p} \le \sum_{k=k_0}^{\infty} 2^{-|\kappa|k} \|\mathcal{M}_k\|_{L^p \to L^p}. \tag{5.1}$$

For the estimation of $\mathcal{M}_k$ we shall invoke Proposition 4.1, with $T := 2^k$ and $\sigma := 2^{-k}$. We then have to estimate oscillatory integrals of the form

$$J(\xi, \xi_3) := \int e^{i(\xi\cdot x + \xi_3 \phi^k(x))}\eta(x) dx,$$

where η is smooth with compact support in the annulus $\mathcal{A}$ on which $|x| \sim 1$. As in the proof of Proposition 4.1 we may and shall assume that $|\xi| \lesssim |\xi_3|$. This allows us to re-write

$$J(\xi, \xi_3) = \int e^{i|\xi_3|\phi(x,s,\sigma)} \eta(x) dx,$$

where the complete phase in this oscillatory integral is of the form

$$\phi(x, s, \sigma) = \phi(x, \sigma) + s \cdot x,$$

with $s = \xi/|\xi_3| \in \mathbb{R}^2$ satisfying $|s| \lesssim 1$, and where $\phi(x, 0) = \phi_{\text{pr}}(x)$.

Now *assume* that we can estimate $J(\xi, \xi_3)$ by

$$|J(\xi, \xi_3)| \lesssim \frac{\|\eta\|_{C^3}}{(1 + |\xi_3|)^\gamma} \tag{5.2}$$

uniformly in σ, with some γ such that

$$d \geq 1 + 1/(2\gamma). \tag{5.3}$$

Then Proposition 4.1 implies that

$$2^{-|\kappa|k} \|\mathcal{M}_k\|_{L^p \to L^p} \lesssim 2^{-|\kappa|k} (2^k)^{\frac{1}{p}} = 2^{-(\frac{1}{d} - \frac{1}{p})k}, \tag{5.4}$$

provided $p \geq d$, and if $p > d$, then the series in (5.1) converges and we see that $\|\mathcal{M}\|_{L^p \to L^p} < \infty$ whenever $p > d$. Thus, whenever we can verify (5.2) and (5.3), then we obtain the desired estimate in Theorem 1.2.

We begin with singularities of type E.

The case E_6. Here $\phi_{\text{pr}}(x_1, x_2) = x_2^3 \pm x_1^4$ and $\kappa := (1/4, 1/3)$, and we claim that estimate (5.2) holds with $\gamma = 3/4$.

To this end, fix x^0 in the support of η, s^0 and $\sigma^0 := 0$. We want to estimate the contribution of a small neighborhood of x^0 to the integral $J(\xi, \xi_3)$, uniformly for all (s, σ) in a small neighborhood of $(s^0, 0)$. If the complete phase $\phi(x, s^0, \sigma^0)$ has no critical point at x^0, then integrations by parts lead to even stronger estimates than required by (5.2). Let us therefore assume that x^0 is a critical point. Since $|x^0| \sim 1$, at least one of the two coordinates of x^0 is of size 1. This shows that in order to estimate the contribution of a small neighborhood of x^0 to the integral $J(\xi, \xi_3)$, we may first apply the method of stationary phase in one of the two variables of integration, and subsequently van der Corput's lemma of order either $2, 3$ or 4 in the remaining variable, leading to an estimate of order $O(|\xi_3|^{-1/2-1/4})$ for all (s, σ) sufficiently close to (s^0, σ^0), in the worst case scenario. By means of a partition of unity argument, this leads to (5.2), with $\gamma = 3/4$. Since here $d = 12/7 > 5/3 = 1 + 1/(2\gamma)$, we are done.

The case E_8. Here, a very similar argument applies and we obtain the estimate (5.2) with $\gamma = 7/10$. Since $d = 15/8 > 12/7 = 1 + 1/(2\gamma)$, we are again done.

The case E_7. In this case we have $\phi_{\text{pr}} = x_2^3 + x_1^3 x_2$, and we claim that here the estimate (5.2) holds with $\gamma = 5/6$.

Indeed, one computes that the Hessian determinant of ϕ_{pr} vanishes at $x^0 \in \operatorname{supp} \eta$ if and only if $x_0^1(4(x_2^0)^2 - (x_1^0)^3) = 0$, so that either $x_1^0 = 0$ and $x_2^0 \neq 0$, or $4(x_2^0)^2 = (x_1^0)^3$ and thus $x_1^0 \neq 0 \neq x_2^0$.

In these cases, we can follow the approach from [IM16] (p. 54): we perform a Taylor expansion of ϕ_{pr} around the point x^0 and collected the homogeneous terms of degree 2 and 3. Then, after applying a linear change of coordinates (altogether, this here amounts to the affine coordinate change $z_1 = x_1 - x_1^0,\ z_2 = x_2 - x_2^0 + \frac{(x_1^0)^2}{2x_2^0} z_1$), we see that in the new variables (z_1, z_2) we do have a singularity of type A_2, and thus we can first apply the method of stationary phase in z_2 and then van der Corput's lemma of order 3 in z_1 to arrive at the desired estimate. If the Hessian determinant of ϕ_{pr} does not

vanish at x^0, then we can apply the method of stationary phase in (x_1, x_2) to a small neighborhood of x^0 and find that we may even choose $\gamma = 1$ in (5.2). So, in all cases, (5.2) holds true with $\gamma = 5/6$.
Since $d = 9/5 > 8/5 = 1 + 1/(2\gamma)$, we are again done.

The case $D_{n+1} (n \geq 3)$. We next turn to singularities of type D_{n+1}, assuming that the coordinates (x_1, x_2) are adapted to ϕ, i.e., that $3 \leq n \leq 2m+1$. The case where $n = 2m+1$ is a bit more delicate, so let us first look at

a) The case $n \leq 2m$. Assuming without loss of generality that $b_1(0,0) = 1$, then under the condition $n \leq 2m$ the principal part of ϕ is given by

$$\phi_{\mathrm{pr}}\,(x_1, x_2) = x_1 x_2^2 + \beta(0) x_1^n, \quad \text{with } \beta(0) \neq 0.$$

We may then argue in a very similar way as in case E_7 to see that here $J(\xi, \xi_3)$ can again be estimated by (5.2) with $\gamma = 5/6$. Since for $n \geq 4$ we have $d = 2n/(n+1) \geq 8/5$, we are done in the case $n \geq 4$.

b) The case $n = 2m+1$. We claim that in this case $J(\xi, \xi_3)$ can be estimated by (5.2) with $\gamma = 3/4$.

In order to prove this, notice that we may assume without loss of generality that the principal part of ϕ is of the form

$$\phi_{\mathrm{pr}}\,(x_1, x_2) = x_1 (x_2 - c x_1^m)^2 + \beta(0) x_1^{2m+1}, \quad \text{with } c, \beta(0) \neq 0.$$

In order to estimate $J(\xi, \xi_3)$, let us fix as before x^0, s^0 and $\sigma^0 := 0$ with $|x^0| \sim 1$ and $|s^0| \lesssim 1$. Assume also that x^0 is the critical point of the phase function $\phi(x,\, s^0,\, 0)$. Then necessarily $|s^0| \sim 1$. Let us pass to new adapted coordinates $(y_1, y_2) := (x_1, x_2 - c x_1^m)$ in the integral defining $J(\xi, \xi_3)$. In these coordinates, the complete phase for $\sigma = 0$ is then given by $\phi^a(y_1, y_2, s^0, 0) := \phi(y_1, y_2 + c y_1^m, s^0, 0)$, i.e., by

$$\phi^a(y_1, y_2, s^0, 0) = y_1 y_2^2 + \beta(0) y_1^{2m+1} + s_1^0 y_1 + s_2^0 c y_1^m + s_2^0 y_2.$$

y^0 will denote the critical point of this phase corresponding to x^0. Obviously, we have $|y^0| \sim 1$. We distinguish now two cases:

Case 1. $y_1^0 = 0$. Then necessarily $y_2^0 \neq 0$ and $0 = \partial_2 \phi^a(y^0, s_2^0,\, 0) = s_2^0$. In this case $\mathrm{Hess}(\phi^a)(y^0, s^0, 0) = -4(y_2^0)^2 \neq 0$. So, y^0 is a non-degenerate critical point and we can use the method of stationary phase in two variables to obtain an estimate of order $O(|\xi_3|^{-1})$ for $J(\xi, \xi_3)$, which is stronger than what we need.

Case 2. $y_1^0 \neq 0$. Then, for given $y_1 \neq 0$ in a sufficiently small neighborhood of the point y_1^0, the critical point of the phase with respect to the variable y_2 is given by $y_2^c(y_1) := -s_2^0/(2y_1)$. Clearly it is non-degenerate. Applying the method of stationary phase method in y_2 we thus arrive at the new phase

$$\phi_1(y_1) := \phi^a(y_1, y_2^c(y_1), s^0, 0) = \beta(0) y_1^{2m+1} + s_1^0 y_1 + s_2^0 c y_1^m - \frac{(s_2^0)^2}{4y_1}.$$

Since the exponents of y_1 in this phase are all different, namely $2m+1, 1, m$ and -1 the equation

$$\phi_1'(y_1) = 0$$

can have at most a root of multiplicity 3 at the point y_1^0, and thus van der Corput's estimate (more precisely its variant given by Lemma 2.1 in [IM16]) implies that the remaining one-dimensional integral in y_1 admits a uniform estimation of order $O(|\xi_3|^{-1/4})$. In conclusion, by first applying the method of stationary phase in y_2, and then this van der Corput type estimate in y_1, we see that estimate (5.2) holds true, with $\gamma := 1/2 + 1/4 = 3/4$. Since for $n \geq 4$ we have $d = \frac{2m+1}{m+1} \geq 5/3 = 1 + 1/(2\gamma)$, we are done also in this case.

There remains the case $n = 3$, i.e., the case of D_4 - type singularities. Notice that in this case the coordinates (x_1, x_2) are always adapted, since $n = 3 \le 5 \le 2m + 1$. We may assume that

$$\phi_{\mathrm{pr}}(x_1, x_2) = x_1 x_2^2 \pm x_1^3,$$

i.e., that we have a D_4^+ or a D_4^- - type singularity (compare Remark 3.3). The case of a D_4^-- type singularity is easy, since in this case the Hessian determinant of ϕ_{pr} is given by $\mathrm{Hess}(\phi_{\mathrm{pr}})(x_1, x_2) = -12x_1^2 - 4x_2^2 \neq 0$ whenever $x \neq 0$. Thus, on the annulus D, there are only non-degenerate critical points of ϕ, and therefore the method of stationary phase implies that estimate (5.2) holds with $\gamma = 1$. Since here $d = 3/2 = 1 + 1/(2\gamma)$,we are again done.

There remains the D_4^+ case, in which the Hessian determinant $\mathrm{Hess}(\phi_{\mathrm{pr}})(x_1, x_2) = 12x_1^2 - 4x_2^2$ may vanish on the annulus. This case will require a more refined analysis, which will be carried out in the last two sections Sections 7, 8.

Remark 5.1. If ϕ has a singularity of type A_{n-1}, with $3 \le n \le 2m$, so that the coordinates (x_1, x_2) are adapted to ϕ, then a similar argument would show that estimate (5.2) holds with $\gamma = 1/2 + 1/n = 1/d = 1/h$. But then the condition (5.3) would just mean that $h \ge 2$, in contradiction to our assumption $h < 2$. This gives a first hint that the treatment of type A singularities will require much finer methods.

6. Estimation of the maximal operator $\mathcal{M}$ when there is no linear adapted coordinate system

Assuming that the coordinates (x_1, x_2) are not adapted to ϕ in Theorem 3.1 means that we are dealing with singularities of type $D_{n+1}, n \ge 3$, so that ϕ is of the form (3.1), with $b(x_1, x_2)$ given by (3.2), i.e.,

$$\phi(x) := (x_1 b_1(x_1, x_2) + x_2^2 b_2(x_2))(x_2 - x_1^m \omega(x_1))^2 + x_1^n \beta(x_1), \tag{6.1}$$

where b_1, b_2, β are smooth functions with $b_1(0, 0) \neq 0$, $\omega(0) \neq 0$ and $\beta(0) \neq 0$. Note also that $n \ge 2m + 2$ and $m \ge 2$, so that in particular $n \ge 6$ and hence

$$h = \frac{2n}{n + 1} \ge \frac{12}{7}. \tag{6.2}$$

As in [IKM10], the main problems will here arise from a sufficiently narrow neighborhood of the principal root jet, i.e., the curve $x_2 = \psi(x_1) = x_1^m \omega(x_1)$. More precisely, consider the function $\phi^a(y_1, y_2)$ in (3.3) of Remark 3.2, which describes ϕ in the adapted coordinates $(y_1, y_2) = (x_1, x_2 - \psi(x_1))$. Then one easily sees that

$$\phi^a(y) = (y_1 b_1^a(y_1, y_2) + y_2^2 b_2^a(y_2))y_2^2 + y_1^n \beta(y_1),$$

where $b_1^a(0, 0) = b_1(0, 0) \neq 0 \neq \beta(0)$, and therefore the principal part of ϕ^a is given by

$$\phi^a_{\mathrm{pr}}(y_1, y_2) = y_1 b_1(0, 0)y_2^2 + y_1^n \beta(0), \tag{6.3}$$

which is κ^a-homogeneous of degree one with respect to the weight

$$\kappa^a := \Big(\frac{1}{n}, \frac{n - 1}{2n}\Big).$$

Notice also that

$$h(\phi) = \frac{1}{|\kappa^a|} \quad \text{and } a := \kappa_2^a/\kappa_1^a = (n - 1)/2 > m.$$

Following our approach from [IKM10], we therefore decompose the domain Ω into three regions, an "exterior" region of the form

$$D_{\mathrm{ext}} := \{x \in \Omega : |x_2 - \psi(x_1)| \ge \varepsilon |x_1^m|\}, \tag{6.4}$$

the "principal region" close to the principal root jet, which is of the form

$$D_{\rm pr} := \{x \in \Omega : |x_2 - \psi(x_1)| \le N|x_1^a|\}, \tag{6.5}$$

and a "transition region" of the form

$$E := \{x \in \Omega : N|x_1^a| \le |x_2 - \psi(x_1)| \le \varepsilon |x_1^m|\}, \tag{6.6}$$

where $\varepsilon > 0$ is a sufficiently small and $N > 0$ a sufficiently large parameter to be chosen later.

Observe also that the region $D_{\rm ext}$ is essentially invariant under the dilations $\delta_r, r \ll 1$, associated to the weight κ, whereas the region $D_{\rm pr}$, when expressed in the adapted coordinates (y_1, y_2), is essentially invariant under the dilations $\delta^a_r, r \ll 1$, associated to the weight κ^a, i.e., the dilations defined by $\delta^a_r(y_1, y_2) := (r^{\kappa^a_1} y_1, r^{\kappa^a_2} y_2)$.

For the localizations to such regions, here and later on in this article, it will be useful to employ the following notation from [IKM10]: if A_t is the averaging operator from (1.3), and if χ is any integrable "cut-off" function defined on Ω, we shall denote by A^χ_t the correspondingly localized averaging operator

$$A^\chi_t f(y, y_3) = f * \mu^\chi_t(y, y_3) = \int_{\mathbb{R}^2} f(y - tx, y_3 - t(1 + \phi(x)))\eta(x)\chi(x)\, dx$$

corresponding to the measure $\mu^\chi := (\chi \otimes 1)\mu$, and by $\mathcal{M}^\chi$ the associated maximal operator

$$\mathcal{M}^\chi f(y, y_3) := \sup_{t>0} |A^\chi_t f(y, y_3)|.$$

6.1. **The contribution by the region away from the principal root jet.** Here we may essentially proceed as in Section 5, choosing again for κ the principal weight $\kappa := (1/(2m + 1), m/(2m + 1))$.

In order to localize to an exterior region $D_{\rm ext}$ in a smooth fashion, we fix a cut-off function $\rho \in C^\infty_0(\mathbb{R})$ supported in $[-2, 2]$ and identically 1 on $[-1, 1]$, and put

$$\rho_1(x_1, x_2) := \rho\Big(\frac{x_2 - \omega(0)x_1^m}{\varepsilon x_1^m}\Big), \qquad \varepsilon > 0.$$

Since $\psi(x_1) = \omega(0)x_1^m + o(|x_1|^m)$, we could as well have chosen $\psi(x_1)$ in place of the leading term $\omega(0)x_1^m$ of $\psi(x_1)$ in this definition, but the advantage of our choice is that $1 - \rho_1$ becomes κ-homogeneous of degree zero. Clearly, $1 - \rho_1$ is nevertheless supported in a region of the form $D_{\rm ext}$.

Proposition 6.1. *If $p > d = d(\phi)$, and if the neighborhood Ω of $(0, 0)$ is chosen sufficiently small, then the maximal operator $\mathcal{M}^{1-\rho_1}$ is bounded on L^p.*

Since $h \ge d$, this result implies in particular the L^p-boundedness of $\mathcal{M}^{1-\rho_1}$ when $p > h$.

Proof. We can proceed exactly as in Section 5, with the measure μ replaced by $\mu^{1-\rho_1}$; for κ we still choose the principal weight. The re-scaled measures μ_k are now given by

$$\int f\, d\mu_k := \int f(x, 2^k + \phi^k(x))\, \eta(\delta_{2^{-k}} x)\chi_1(x)(1 - \rho)\Big(\frac{x_2 - \omega(0)x_1^m}{\varepsilon x_1^m}\Big)\, dx,$$

and we have to estimate oscillatory integrals of the form

$$J(\xi, \xi_3) := \int e^{i(\xi\cdot x + \xi_3\phi^k(x))}\eta(x)(1 - \rho)\Big(\frac{x_2 - \omega(0)x_1^m}{\varepsilon x_1^m}\Big)\, dx,$$

where η is smooth with compact support in the annulus $\mathcal{A}$. Therefore the amplitude in this oscillatory integral $J(\xi, \xi_3)$ is supported in the intersection of the annulus $\mathcal{A}$, on which $|x| \sim 1$, and the region given by (6.4). This is exactly the kind of oscillatory integral (whose phase has at worst an Airy type A_2 singularity) that we had estimated in our discussion of singularities of type D in Chapter 3 of [IM16] (compare pp. 53–54), where we had shown that $J(\xi, \xi_3) = O((1 + |\xi_3|)^{-5/6})$, uniformly in k for k sufficiently large. This means that we may choose $\gamma := 5/6$ in (5.2).

But, since $m \geq 2$, we have $d = (2m+1)/(m+1) \geq 5/3 > 8/5 = 1 + 1/(2\gamma)$, and thus we see that $\mathcal{M}$ is L^p-bounded whenever $p > d$. Q.E.D.

6.2. **The contribution by the transition domain.** In order to localize to the transition domain in a smooth fashion, we put

$$\tau(x) := \rho\Big(\frac{x_2 - \psi(x_1)}{\varepsilon x_1^m}\Big)(1-\rho)\Big(\frac{x_2 - \psi(x_1)}{N x_1^a}\Big).$$

Then clearly τ is supported in a region of the form E. By means of the change to the adapted coordinates $(y_1, y_2) = (x_1, x_2 - \psi(x_1))$, we then see that the averaging operator A_t^τ can be written as

$$A_t^\tau f(z, z_3) = \int_{\mathbb{R}^2} f\big(z_1 - t y_1, z_2 - t(y_2 + \psi(y_1)), z_3 - t(1 + \phi^a(y))\big)\tau^a(y)\,\eta^a(y)\,dy,$$

where

$$\tau^a(y) = \rho\Big(\frac{y_2}{\varepsilon y_1^m}\Big)\,(1-\rho)\Big(\frac{y_2}{N y_1^a}\Big),$$

and η^a is a smooth function supported in a sufficiently small neighborhood of the origin.

Proposition 6.2. *If $p > h = h(\phi)$, and if the neighborhood Ω of $(0,0)$ is chosen sufficiently small, then the maximal operator $\mathcal{M}^\tau$ is bounded on L^p.*

Proof. Following our approach from [IKM10], which had been inspired by Phong and Stein's article [PS97], we decompose the domain E (which is a domain of transition between the homogeneities given be the weight κ and the weight κ^a) dyadically in each coordinate separately, and then re-scale each of the bi-dyadic pieces obtained in this way.

To this end, consider a dyadic partition of unity $\sum_{k=0}^\infty \chi_k(s) = 1$ on the interval $0 < s \leq 1$ with $\chi \in C_0^\infty(\mathbb{R})$ supported in the interval $[1/2, 4]$, where $\chi_k(s) := \chi(2^k s)$, and put

$$\chi_{j,k}(x) := \chi_j(x_1)\chi_k(x_2), \quad j, k \in \mathbb{N}.$$

We then decompose A_t^τ into the operators

$$A_t^{j,k} f(z) := \int_{\mathbb{R}^2} f\Big(z_1 - t y_1, z_2 - t(y_2 + \psi(y_1)), z_3 - t(1 + \phi^a(y))\Big)\tau^a(y)\,\eta^a(y)\,\chi_{j,k}(y)\,dy,$$

with associated maximal operators $\mathcal{M}^{j,k}$.

Notice that by choosing the neighborhood Ω of the origin sufficiently small, we need only consider sufficiently large j, k. Moreover, because of the localization imposed by τ^a, it suffices to consider only pairs (j,k) satisfying

$$mj + M \leq k \leq \frac{(n-1)j}{2} - M, \tag{6.7}$$

where M can still be chosen sufficiently large, because we had the freedom to choose ε sufficiently small and N sufficiently large. In particular, we have $j \sim k$, and clearly

$$\|\mathcal{M}^\tau\|_{p\to p} \leq \sum_{mj + M \leq k \leq \frac{(n-1)j}{2} - M} \|\mathcal{M}^{j,k}\|_{p\to p}. \tag{6.8}$$

By re-scaling in the integral, we have

$$\begin{aligned} A_t^{j,k} f(z) = 2^{-j-k}\int_{\mathbb{R}^2} \quad & f\Big(z_1 - t2^{-j}y_1, z_2 - t(2^{-k}y_2 + 2^{-mj}y_1^m\omega(2^{-j}y_1)),\\ & z_3 - t(1 + \phi^a(2^{-j}y_1, 2^{-k}y_2))\Big)\tilde\tau^{j,k}(y)\,\tilde\eta^{j,k}(y)\,\chi(y_1)\chi(y_2)\,dy, \end{aligned}$$

with

$$\tilde\tau^{j,k}(y) := \rho\Big(\frac{y_2}{\varepsilon 2^{k-mj} y_1^m}\Big)\,(1-\rho)\Big(\frac{y_2}{N 2^{k - \frac{(n-1)j}{2}} y_1^{\frac{n-1}{2}}}\Big), \quad \tilde\eta^{j,k}(y) := \eta^a(2^{-j}y_1, 2^{-k}y_2).$$

Notice that, by (6.7), all derivatives of $\tilde{\tau}^{j,k}$ are uniformly bounded in j,k.

The scaling operators

$$T^{j,k}f(z) := 2^{\frac{(m+2)j+2k}{p}} f(2^j z_1, 2^{mj} z_2, 2^{j+2k} z_3)$$

then transform these operators into the averaging operators $\tilde{A}_t^{j,k} := T^{-j,-k} A_t^{j,k} T^{j,k}$, i.e.,

$$\begin{aligned}\tilde{A}_t^{j,k} f(z) = 2^{-j-k} \quad & \int_{\mathbb{R}^2} f\Big(z_1 - ty_1, z_2 - t(2^{mj-k}y_2 + y_1^m \omega(2^{-j}y_1)), \\ & z_3 - t(2^{j+2k} + \tilde{\phi}^{j,k}(y))\Big) \tilde{\tau}^{j,k}(y)\, \tilde{\eta}^{j,k}(y)\, \chi(y_1)\chi(y_2)\, dy,\end{aligned}$$

where

$$\tilde{\phi}^{j,k}(y) := 2^{j+2k} \phi^a(2^{-j}y_1, 2^{-k}y_2).$$

Notice that by (6.7) we then have

$$\tilde{\phi}^{j,k}(y) = b_1(0,0) y_1 y_2^2 + O(2^{-j} + 2^{-k} + 2^{-M}).$$

In order to simplify notation let us put

$$\phi(x,\delta) := b(x,\delta)x_2^2 + \delta_3 x_1^n \beta(\delta_1 x_1), \tag{6.9}$$

where

$$b(x,\delta) := b_1^a(\delta_1 x_1, \delta_2 x_2)x_1 + \delta_4 x_2^2 b_2^a(\delta_2 x_2).$$

Here $\delta := (\delta_0, \dots, \delta_4)$ is supposed to be very small, i.e., $|\delta| \ll 1$. Then for the special values

$$\delta_0 := 2^{mj-k},\ \delta_1 := 2^{-j},\ \delta_2 := 2^{-k},\ \delta_3 := 2^{2k-(n-1)j},\ \delta_4 := 2^{j-2k}, \tag{6.10}$$

we find that $\tilde{\phi}^{j,k}(y) = \phi(y,\delta)$. Notice also that, due to the condition (6.7), for these values of δ we have indeed $|\delta| \ll 1$.

From now on we shall then consider the phase as well as the corresponding averaging operators as quantities depending on the non-negative small perturbation parameters δ_i of which the vector δ is composed, which are otherwise arbitrary. Moreover, we shall denote the variable y again by x.

Let us also introduce an additional parameter $T \geq 0$, which in our application to the averaging operators $\tilde{A}_t^{j,k}$ will become

$$T = 2^{j+2k}.$$

Then, for these more general sets of parameters δ and T, we introduce the measure $\nu_{\delta,T}$ by putting

$$\int f d\nu_{\delta,T} := \int f(x_1,\, \delta_0 x_2 + x_1^m \omega(\delta_1 x_1),\, T + \phi(x,\delta))\eta(x,\delta)\chi_1(x_1)\chi_1(x_2)dx,$$

where $\eta(x,\delta) := \eta^a(\delta_1 x_1, \delta_2 x_2)$, and set $A_{\delta,T} f := f * \nu_{\delta,T}$. By $\mathcal{M}_{\delta,T}$ we denote the maximal operator corresponding to the averaging operators $(A_{\delta,T})_t$.

Note that in the considered domain of integration $|x_1| \sim 1$, $|x_2| \sim 1$, and hence also $b(x,\delta) \sim 1$, if we assume without loss of generality that $b_1^a(0,0) = 1$, and that $\phi(x,0) = x_1 x_2^2$.

For the particular choice of δ given by (6.10) and $T = 2^{j+2k}$, we then have $\tilde{A}_t^{j,k} = 2^{-j-k}(A_{\delta,T})_t$, so that

$$\|\mathcal{M}^{j,k}\|_{p\to p} \leq 2^{-j-k} \|\mathcal{M}_{\delta,T}\|_{p\to p}. \tag{6.11}$$

It will thus suffice to estimate the maximal operator $\mathcal{M}_{\delta,T}$. Note, however, that in view of (6.7), in our application, where δ is given by (6.10) and $T = 2^{j+2k}$, we have that

$$T \geq \delta_0^{-2}, \quad \text{i.e.,} \quad \delta_0 \geq T^{-\frac{1}{2}}. \tag{6.12}$$

We shall therefore assume that this relation between δ_0 and T holds true also in the subsequent study of the maximal operator $\mathcal{M}_{\delta,T}$, as well as that $|\delta| \ll 1$, for otherwise general δ and T. Notice that that this does not effect our definition of $\nu_{\delta,T'}$ for $T' = 0$.

To begin with, notice that the function $\phi(x,\delta)$ is a small perturbation of the function $\phi(x,0)=x_1x_2^2$, so that in the limit as $\delta\to 0$ the limiting measure ν_0 is supported in the hypersurface given by all points $(x_1,\omega(0)x_1^m,T+x_1x_2^2)$ with $|x_1|\sim 1\sim |x_2|$. Choosing $y_1=x_1$ and $y_2=x_1x_2^2$ as new coordinates for this hypersurface, we see that it has exactly one non-vanishing principal curvature at every point, so that an application of Proposition 4.1 (with $\gamma=1/2$) would only allow to control the associated maximal operator for $p>2$. We therefore must apply a more refined analysis and shall invoke ideas as well as notation from [IM16] (see, e.g., Section 4.1) based on additional dyadic decompositions in every frequency variable. This analysis will allow us to take advantage of the lower bound for δ_0 given by (6.12).

To this end, we fix again suitable smooth cut-off functions $\chi_l\geq 0$ on $\mathbb{R}$ as in the proof of Proposition 4.1 such that for $l\geq 1$, $\chi_l(t)=\chi_1(2^{1-l}t)$ is supported where $|t|\sim 2^l$ and

$$\sum_{l=0}^{\infty}\chi_l(t)=1 \qquad \text{for all } t\in\mathbb{R},$$

and define for every multi-index $l=(l_1,l_2,l_3)\in\mathbb{N}^3$ the cut-off function

$$\chi_l(\xi):=\chi_{l_1}(\xi_1)\chi_{l_2}(\xi_2)\chi_{l_3}(\xi_3).$$

In what follows, we shall usually write $\lambda=(\lambda_1,\lambda_2,\lambda_3)$ in place of $(2^{l_1-1},2^{l_2-1},2^{l_3-1})$, and define accordingly the complex measures $\nu_{\delta,T}^{\lambda}$ by

$$\widehat{\nu_{\delta,T}^{\lambda}}(\xi):=\chi_l(\xi)\widehat{\nu_{\delta,T}}(\xi).$$

Notice that if $l_i\geq 1$ for $i=1,2,3$, then

$$\widehat{\nu_{\delta,T}^{\lambda}}(\xi)=\chi_1\left(\frac{\xi_1}{\lambda_1}\right)\chi_1\left(\frac{\xi_2}{\lambda_2}\right)\chi_1\left(\frac{\xi_3}{\lambda_3}\right)\widehat{\nu_{\delta,T}}(\xi),$$

and

$$|\xi_i|\sim\lambda_i \qquad \text{on } \operatorname{supp}\widehat{\nu_{\delta,T}^{\lambda}}. \tag{6.13}$$

We then find that, in the sense of distributions,

$$\nu_{\delta,T}=\sum_{\lambda}\nu_{\delta,T}^{\lambda}, \tag{6.14}$$

where the summation $\sum_\lambda$ will always mean summation over the set Λ of all *dyadic* λ with $\lambda_i\geq 2^{-1}$ for $i=1,2,3$. To simplify the subsequent discussion, we shall concentrate on those measures $\nu_{\delta,T}^{\lambda}$ for which none of its components λ_i equals 2^{-1}, i.e., $l_i\geq 1$, since the remaining cases where some $l_i=0$ can be dealt with in the same way as the corresponding cases where $l_i\geq 1$ is small. By $\mathcal{M}_{\delta,T}^{\lambda}$ we denote the maximal operator corresponding to the convolution operators $(A_{\delta,T}^{\lambda})_t f:=f*(\nu_{\delta,T}^{\lambda})_t,\ t>0$.

The Fourier transform of $\nu_{\delta,T}$ is explicitly given by

$$\widehat{\nu_{\delta,T}}(\xi)=\int e^{-i\Phi(x,\delta,T,\xi)}\eta(x,\delta)\chi_1(x_1)\chi_1(x_2)dx,$$

with complete phase

$$\Phi(x,\delta,T,\xi):=\xi_1x_1+\xi_2(\delta_0x_2+x_1^m\omega(\delta_1x_1))+\xi_3(T+\phi(x,\delta)).$$

The following lemma will be used frequently:

Lemma 6.3. (a) *The maximal operator $\mathcal{M}_{\delta,T}^{\lambda}$ is of weak-type $(1,1)$, with norm bounded by*

$$\|\mathcal{M}_{\delta,T}^{\lambda}\|_{L^1\to L^{1,\infty}}\leq CT\|\nu_{\delta,0}^{\lambda}\|_\infty.$$

where the constant C is independent of δ and T.

(b) *For $1 < p \le 2$, the maximal operator $\mathcal{M}^\lambda_{\delta,T}$ is bounded on L^p with norm controlled by*

$$\|\mathcal{M}^\lambda_{\delta,T}\|_{p\to p} \le C_p T^{\frac{2}{p}-1}(\lambda_3 T + \lambda_1 + \lambda_2)^{1-\frac{1}{p}} \|\nu^\lambda_{\delta,0}\|_\infty^{\frac{2}{p}-1} \|\widehat{\nu^\lambda_{\delta,0}}\|_\infty^{2-\frac{2}{p}}.$$

Proof. The proof follows the pattern of the proof of Proposition 4.1. Indeed, arguing in the same way as in that proof and observing that still the measures $\nu^\lambda_{\delta,T}$ are essentially supported in a cuboid of dimensions comparable to $1\times 1\times T$, we see that $\mathcal{M}^\lambda_{\delta,T} f$ is dominated by $CT\|\nu^\lambda_{\delta,0}\|_\infty M_{HL}(|f|)$, which implies (a).

Moreover, using again Littlewood-Paley theory, we also easily see that

$$\|\mathcal{M}^\lambda_{\delta,T}\|_{2\to 2} \le C_p(\lambda_3 T + \lambda_1 + \lambda_2)^{\frac{1}{2}} \|\widehat{\nu^\lambda_{\delta,0}}\|_\infty.$$

(b) follows then again by an application of Marcinkiewicz's interpolation theorem. Q.E.D.

In order to estimate $\|\nu^\lambda_{\delta,0}\|_\infty$, we write

$$\begin{aligned}\nu^\lambda_{\delta,0}(x) = \lambda_1\lambda_2\lambda_3 \quad &\int \check{\chi}_1\big(\lambda_1(x_1-y_1)\big)\,\check{\chi}_1\big(\lambda_2(x_2-\delta_0 y_2 - y_1^m\omega(\delta_1 y_1))\big)\\ &\check{\chi}_1\big(\lambda_3(x_3-\phi(y,\delta))\big)\,\eta(y,\delta)\,\chi(y_1)\chi(y_2)\,dy_1dy_2,\end{aligned} \tag{6.15}$$

where $\check{f}$ denotes the inverse Fourier transform of f. Observe that $|\partial_{y_2}\phi(y,\delta)| \sim 1$, so that $(y_1, \phi(y_1,y_2,\delta))$ can be used as coordinates in place of (y_1,y_2). We may therefore change coordinates from (y_1,y_2) to (u_1,u_2) in this integral, where $y_1 = u_1/\lambda_1$ and $\phi(y,\delta) = u_2/\lambda_3$, which easily leads to the uniform estimate $|\nu^\lambda_\delta(x)| \le C\lambda_2$, with C independent of x,δ and λ. Similarly, the change of coordinates $y_1 = u_1/\lambda_1$ and $y_2 = u_2/(\delta_0\lambda_2)$ leads to $|\nu^\lambda_\delta(x)| \le C\lambda_3/\delta_0$. Altogether, we arrive at the uniform estimate

$$\|\nu^\lambda_{\delta,0}\|_\infty \lesssim \min\{\lambda_2, \lambda_3\delta_0^{-1}\}. \tag{6.16}$$

Recall also that $\phi(x,\delta) = x_1x_2^2 + O(\delta)$, and that we are interested in exponents $2 > p > h \ge 3/2$.

We shall distinguish six cases depending on the relative sizes of $\lambda_1,\lambda_2,\lambda_3$ and δ_0, and shall accordingly decompose the set Λ of our dyadic λ's into subsets I_i, where I_i will correspond to Case i, $i = 1,\dots,6$. It will therefore be convenient to use the following notation: for any given subset $I\subset\Lambda$, we let $A^I_{\delta,T} := \sum_{\lambda\in I} A^\lambda_{\delta,T}$ denote the contribution to $A_{\delta,T}$ by the operators $A^\lambda_{\delta,T}$ with $\lambda\in I$, with associated maximal operator $\mathcal{M}^I_{\delta,T}$. Then clearly

$$\|\mathcal{M}^I_{\delta,T}\|_{p\to p} \le \sum_{\lambda\in I}\|\mathcal{M}^\lambda_{\delta,T}\|_{p\to p},$$

and moreover we shall have

$$\|\mathcal{M}_{\delta,T}\|_{p\to p} \le \sum_{i=1}^{6}\|\mathcal{M}^{I_i}_{\delta,T}\|_{p\to p}.$$

For each of the maximal operators $\mathcal{M}^{I_i}_{\delta,T}$ we shall prove the following estimate:

$$\|\mathcal{M}^{I_i}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}. \tag{6.17}$$

This will then imply that $\|\mathcal{M}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}} = 2^{\frac{j+2k}{p}}$, and combining this with (6.11) we see that

$$\|\mathcal{M}^{j,k}\|_{p\to p} \lesssim 2^{-j(1-\frac{1}{p})}2^{k(\frac{2}{p}-1)}.$$

By (6.8) we then obtain the estimate

$$\|\mathcal{M}^\tau\|_{p\to p} \lesssim \sum_{mj+M\le k\le\frac{(n-1)j}{2}-M} 2^{-j(1-\frac{1}{p})}2^{k(\frac{2}{p}-1)} = \sum_{j\ge 0} 2^{-j(\frac{n+1}{2}-\frac{n}{p})} < \infty,$$

since $p > h = 2n/(n+1)$, which will conclude the proof of Proposition 6.2.

Case 1: $\lambda_3 \gtrsim \max\{\lambda_1, \lambda_2\}$. Iterated integrations by parts in x_1 then lead to the estimate

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_3^{-N} \qquad \text{for every } N \in \mathbb{N}.$$

Moreover, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2$. Observe also that $\lambda_3 T$ is the dominant term in $\lambda_3 T+\lambda_1+\lambda_2$, and thus Lemma 6.3 implies

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} \lambda_2^{\frac{2}{p}-1} \lambda_3^{-N}$$

for every $N \in \mathbb{N}$. This easily implies that

$$\|\mathcal{M}_{\delta,T}^{I_1}\|_{p\to p} \le \sum_{\lambda_3 \gtrsim \max\{\lambda_1,\lambda_2\}} \|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}}. \tag{6.18}$$

The constants in these estimate do not depend on δ.

Case 2: $\lambda_1 \gg \max\{\lambda_2, \lambda_3\}$. Iterated integrations by parts in x_1 here lead to the estimate

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_1^{-N} \qquad \text{for every } N \in \mathbb{N}.$$

Moreover, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2$. Observe also that $\lambda_3 T + \lambda_1 + \lambda_2 \lesssim \lambda_1 T$. From here on we can proceed as in the previous case, with the roles of λ_3 and λ_1 interchanged, and in analogy with (6.18) arrive at

$$\|\mathcal{M}_{\delta,T}^{I_2}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 3: $\lambda_2 \gg \max\{\lambda_1, \lambda_3\}$. Then again iterated integrations by parts in x_1 lead to the estimate

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_2^{-N} \qquad \text{for every } N \in \mathbb{N}.$$

Also, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2$, and moreover clearly $\lambda_3 T + \lambda_1 + \lambda_2 \lesssim \lambda_2 T$. Thus Lemma 6.3 implies

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} \lambda_2^{-N}$$

for every $N \in \mathbb{N}$, so that

$$\|\mathcal{M}_{\delta,T}^{I_3}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

There remain those cases where $\lambda_1 \sim \lambda_2 \gg \lambda_3$.

Case 4: $\lambda_1 \sim \lambda_2$ **and** $\lambda_2\delta_0 \ll \lambda_3 \ll \lambda_2$. Here, we may first integrate by parts in x_2 N-times, and then either apply the method of stationary phase in x_1 or again integrate by parts in x_1 (in case that there is no critical point) in order to see that

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_3^{-N} \lambda_2^{-\frac{1}{2}} \qquad \text{for every } N \in \mathbb{N}.$$

Moreover, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2$. Observe also that $\lambda_3 T$ is the dominant term in $\lambda_3 T+\lambda_1+\lambda_2$, since, by (6.12),

$$\lambda_3 T \gg \lambda_2 T \delta_0 > \lambda_2 \delta_0^{-1} \gg \lambda_2.$$

Therefore Lemma 6.3 implies

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} \lambda_2^{-2+\frac{3}{p}} \lambda_3^{-N}$$

for every $N \in \mathbb{N}$. Since $p > 3/2$, this shows that we can sum these estimates over the $\lambda \in I_4$ and obtain

$$\|\mathcal{M}_{\delta,T}^{I_4}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 5: $\lambda_1 \sim \lambda_2$ **and** $\lambda_3 \ll \lambda_2\delta_0$. Arguing in the same way as in the previous case we here find that

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim (\lambda_2\delta_0)^{-N} \lambda_2^{-\frac{1}{2}} \qquad \text{for every } N \in \mathbb{N}.$$

Moreover, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_3\delta_0^{-1}$. Therefore Lemma 6.3 implies that

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{2}{p}-1}(\lambda_3 T+\lambda_2)^{1-\frac{1}{p}}(\lambda_3\delta_0^{-1})^{\frac{2}{p}-1}((\lambda_2\delta_0)^{-N}\lambda_2^{-\frac{1}{2}})^{2-\frac{2}{p}}$$

for every $N\in\mathbb{N}$. Since $\lambda_3 T \ll \lambda_2 T\delta_0$ and, as before, $\lambda_2 T\delta_0 \gg \lambda_2$, we may control $\lambda_3 T+\lambda_2$ by $\lambda_2 T\delta_0$, so that

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}}\delta_0^{2-\frac{3}{p}}\lambda_3^{\frac{2}{p}-1}(\lambda_2\delta_0)^{-N}$$

for every $N\in\mathbb{N}$. Summing first over all $\lambda_1\sim\lambda_2$, then all λ_2 such that $\lambda_2\delta_0\gg\lambda_3$ and finally over all λ_3, we then find that

$$\|\mathcal{M}_{\delta,T}^{I_5}\|_{p\to p} \lesssim T^{\frac{1}{p}}\delta_0^{2-\frac{3}{p}} \le T^{\frac{1}{p}},$$

the last inequality being true since $p>3/2$.

Case 6: $\lambda_1\sim\lambda_2$ **and** $\lambda_3\sim\lambda_2\delta_0$. In this case, there is possibly a (non-degenerate) critical point $x_2^c = x_2^c(x_1,\delta,\xi)$ of the phase with respect to x_2 in the region where $|x_2|\sim 1$. In that case, we apply the method of stationary phase to the integration in x_2, which leads to an oscillatory integral in x_1 whose phase is of the form

$$\xi_1 x_1 + \xi_2(\delta_0 x_2^c + x_1^m\omega(\delta_1 x_1)) + \xi_3(x_1(x_2^c)^2 + O(\delta)).$$

But, since $|\xi_3|\ll\max\{|\xi_1|,|\xi_2|\}$ under our assumptions, the last term can be viewed as a small error term, and thus we may apply van der Corput's estimate of order 2 to the remaining integral in x_1 and altogether arrive at the estimate

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_3^{-\frac{1}{2}}\lambda_2^{-\frac{1}{2}}.$$

If there is no critical point with respect to x_2, integrations by parts in x_2 in place of an application of the method of stationary phase lead to even better estimates. Moreover, by (6.16) we have $\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2$, and as in Case 4 we have $\lambda_3 T\gg\lambda_2$. Therefore Lemma 6.3 implies that

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}}\lambda_3^{1-\frac{1}{p}}\lambda_2^{\frac{2}{p}-1}(\lambda_3^{-\frac{1}{2}}\lambda_2^{-\frac{1}{2}})^{2-\frac{2}{p}} = T^{\frac{1}{p}}\lambda_2^{\frac{3}{p}-2}.$$

Since $p>3/2$, we see that the sum over all indices $\lambda\in I_6$ considered in this case is finite, and we obtain

$$\|\mathcal{M}_{\delta,T}^{I_6}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

This finishes the proof of Proposition 6.2. Q.E.D.

6.3. **The contribution by the region near the principal root jet.** Finally, we consider the contribution to the maximal operator $\mathcal{M}$ by the region

$$D_{\mathrm{pr}} := \{x\in\Omega : |x_2-\psi(x_1)|\le N|x_1|^a\}$$

close to the principal root jet (N is a fixed positive number). We localize to a region of this type by means of the cut-off function

$$\rho_2(x) := \rho\Big(\frac{x_2-\psi(x_1)}{Nx_1^a}\Big).$$

Proposition 6.4. *If $p>h=h(\phi)$, and if the neighborhood Ω of $(0,0)$ is chosen sufficiently small, the maximal operator $\mathcal{M}^{\rho_2}$ is bounded on L^p.*

In combination with Proposition 6.1 and 6.2 this will complete the proof of Theorem 1.2, with the exception of the case of D_4^+ - type singularities.

Proof. In the adapted coordinates $(y_1, y_2) = (x_1, x_2 - \psi(x_1))$ the measure μ^{ρ_2} can be expressed as

$$\int f\,d\mu = \int f(y_1, y_2 + \psi(y_1), 1 + \phi^a(y))\rho^a(y)\eta^a(y)\,dy,$$

with η^a a smooth function with support in a sufficiently small neighborhood Ω^a of the origin as before, and

$$\rho^a(y) := \rho\Big(\frac{y_2}{N y_1^a}\Big).$$

Notice that ρ^a is $\kappa^a = (1/n, (n-1)/(2n))$-homogeneous of degree 0. Working now in these adapted coordinates (y_1, y_2), we next proceed as in Section 5, only with the weight κ replaced by κ^a, and dilations δ_r replaced by the κ^a-dilations δ_r^a from Section 6. Recall to this end from (6.3) that the principal part of ϕ^a, which is κ^a-homogeneous of degree 1, is given by

$$\phi^a_{\mathrm{pr}}(y_1, y_2) = y_1 y_2^2 + y_1^n \beta(0),$$

if we assume again without loss of generality that $b_1(0,0) = 1$. We choose a smooth bump-function χ_1 supported in the annulus $\mathcal{A}$ such that

$$\sum_{k=k_0}^{\infty} \chi_1(\delta^a_{2^k} y) = 1 \quad \text{for} \quad 0 \neq y \in \Omega^a,$$

and decompose

$$\mu = \sum_{k=k_0}^{\infty} \tilde{\mu}_k,$$

with

$$\int f\,d\tilde{\mu}_k := \int f(y_1, y_2 + y_1^m \omega(y_1), 1 + \phi^a(y))\rho^a(y)\eta^a(y)\chi_1(\delta^a_{2^k} y)\,dy.$$

It will then suffice to derive suitable L^p-estimates for the maximal operators $\sup_{t>0} |f*(\tilde{\mu}_k)_t|$. Applying a straight-forward L^p-isometric re-scaling to them by means of the dilations $\delta^a_{2^{-k}}$, we may assume that these are of the form $2^{-|\kappa^a|k}\mathcal{M}_k f$, where

$$\mathcal{M}_k f(y, y_3) := \sup_{t>0} |f * (\mu_k)_t(y, y_3)|$$

and

$$\int f\,d\mu_k := \int f\big(y_1, 2^{-(\kappa_2^a - m\kappa_1^a)k} y_2 + y_1^m \omega(2^{-\kappa_1^a k} y_1), 2^k + \phi^k(y)\big)\rho\Big(\frac{y_2}{N y_1^a}\Big)\eta^a(\delta^a_{2^{-k}} y)\chi_1(y)\,dy.$$

Here we have set $\phi^k(y) := 2^k \phi^a(\delta^a_{2^{-k}}(y))$. Then clearly

$$\phi^k(y) = \phi^a_{\mathrm{pr}}(y) + O(2^{-\varepsilon k}) = y_1 y_2^2 + y_1^n \beta(0) + O(2^{-\varepsilon k})$$

for some $\varepsilon > 0$. Notice also that $|y_1| \sim 1$ and $|y_2| \lesssim 1$ on the support of μ_k.

Similarly as in Section 5, the following analogue of estimate (5.1) holds true:

$$\|\mathcal{M}^{\rho_2}\|_{L^p \to L^p} \leq \sum_{k=k_0}^{\infty} 2^{-|\kappa^a|k} \|\mathcal{M}_k\|_{L^p \to L^p}. \tag{6.19}$$

In order to simplify notation we shall here write

$$\delta_0 := 2^{-(\kappa_2^a - m\kappa_1^a)k},\ \delta_1 := 2^{-\kappa_1^a k},\ \delta_2 := 2^{-\kappa_2^a k}, \delta_3 := 2^{-\frac{k}{2n}} \quad \text{and} \quad T := 2^k, \tag{6.20}$$

and put $\delta := (\delta_0, \delta_1, \delta_2, \delta_3)$. Recall that $a = \kappa_2^a/\kappa_1^a > m$, so that $|\delta| \ll 1$.

Observe that as in the previous subsection the relation (6.12) is valid, i.e.,

$$T \geq \delta_0^{-2},$$

since $2\kappa_2^a = (n-1)/n < 1$, so that $\delta_0^{-2} \leq 2^{2\kappa_2^a k} \leq 2^k = T$.

We shall from now on consider the phase as well as the corresponding averaging operators as quantities depending on the non-negative perturbation parameters δ_i of which the vector δ is composed (the phase $\phi^k(y)$ will indeed be viewed as a function $\phi(y,\delta)$ depending only on y and the "dummy" parameter δ_3.) These are assumed to be sufficiently small. Accordingly, we shall re-write the measure μ_k as $\nu_{\delta,T}$, where $\nu_{\delta,T}$ is of the form

$$\int f\,d\nu_{\delta,T} := \int f\big(y_1, \delta_0 y_2 + y_1^m\omega(\delta_1 y_1), T+\phi(y,\delta)\big)\eta(y,\delta)\chi_0(y_2)\chi_1(y_1)\,dy,$$

where $\phi(y,\delta)$ and $\eta(y,\delta)$ are smooth functions in y and δ, and where χ_0 is smooth and supported in a compact neighborhood of 0, whereas as before $\chi_1(y_1)$ is supported where $|y_1|\sim 1$. Moreover,

$$\phi(y,0) = y_1 y_2^2 + y_1^n\beta(0).$$

As in the previous Subsection 6.2, the corresponding averaging operator will be denoted by $A_{\delta,T}$, with corresponding maximal operator $\mathcal{M}_{\delta,T}$. Then, in analogy with (6.11), we have

$$\|\mathcal{M}^k\|_{p\to p} \le \|\mathcal{M}_{\delta,T}\|_{p\to p},$$

if δ and T are given by (6.20). We shall prove the following uniform estimate

$$\|\mathcal{M}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}, \tag{6.21}$$

provided that $p>12/7$. In combination with (6.19) this will imply that

$$\|\mathcal{M}^{\rho_2}\|_{L^p\to L^p} \lesssim \sum_{k=k_0}^{\infty} 2^{-|\kappa^a|k}2^{\frac{k}{p}} < \infty, \tag{6.22}$$

if $p>h=1/|\kappa^a|$ and hence conclude the proof of Proposition 6.4.

For the proof of (6.21) we follow our approach from the previous subsection. Using also the same notation that we had introduced therein, we perform an additional dyadic frequency decomposition by putting

$$\widehat{\nu_{\delta,T}^{\lambda}}(\xi) = \chi_1\left(\frac{\xi_1}{\lambda_1}\right)\chi_1\left(\frac{\xi_2}{\lambda_2}\right)\chi_1\left(\frac{\xi_3}{\lambda_3}\right)\widehat{\nu_{\delta,T}}(\xi),$$

where

$$\widehat{\nu_{\delta,T}}(\xi) = \int e^{-i\Phi(y,\delta,T,\xi)}\eta(y,\delta)\chi_0(y_2)\chi_1(y_1)\,dy,$$

with complete phase

$$\begin{aligned}\Phi(y,\delta,T,\xi) &:= \xi_1y_1 + \xi_2(\delta_0y_2 + y_1^m\omega(\delta_1y_1)) + \xi_3(T+\phi(y,\delta))\\ &= \xi_1y_1 + \xi_2(\delta_0y_2 + y_1^m\omega(\delta_1y_1)) + \xi_3(T + y_1y_2^2 + y_1^n\beta(0) + O(\delta)).\end{aligned} \tag{6.23}$$

Notice, however, that in contrast to the previous subsection, in this integral we have

$$|y_1|\sim 1 \quad\text{and}\quad |y_2|\lesssim 1.$$

By $\mathcal{M}_{\delta,T}^{\lambda}$, we denote the maximal operator defined by the dilates of $\nu_{\delta,T}^{\lambda}$.

In analogy with (6.15), we have

$$\begin{aligned}|\nu_{\delta,0}^{\lambda}(x)| \le \lambda_1\lambda_2\lambda_3 \quad &\int \Big|\check\chi_1\big(\lambda_1(x_1-y_1)\big)\,\check\chi_1\big(\lambda_2(x_2-\delta_0y_2-y_1^m\omega(\delta_1y_1))\big)\\ &\check\chi_1\big(\lambda_3(x_3-\phi(y,\delta))\big)\,\eta(y,\delta)\,\chi_0(y_2)\chi_1(y_1)\Big|\,dy_1dy_2.\end{aligned}$$

We can estimate this by means of the same type of arguments that we used in [IM16]. Indeed, for y_1 fixed, we can first make use of the localization given by the third factor in this integral and apply the van der Corput-type Lemma 2.1 (b) of order $N=2$ in [IM16] to see that

$$\int |\check\chi_1\big(\lambda_3(x_3-\phi(y,\delta))\big)|\,dy_2 \lesssim \lambda_3^{-\frac{1}{2}},$$

uniformly in y_1, x and δ. Subsequently we may estimate the remaining integral in y_1 by performing the change of variables $y_1 \mapsto y_1/\lambda_1$, which gains another factor λ_1^{-1}. Altogether, this leads to the uniform estimate

$$\|\nu_{\delta,0}^\lambda\|_\infty \lesssim \lambda_2 \lambda_3^{\frac{1}{2}}. \tag{6.24}$$

Again we may and shall apply Lemma 6.3 and proceed by distinguishing here five cases, assuming always that $p > 12/7$.

Case 1: $\lambda_3 \gtrsim \max\{\lambda_1, \lambda_2\}$. For $\delta = 0$ and $T = 0$ the complete phase is given by

$$\Phi(y, 0, 0, \xi) := \xi_1 y_1 + \xi_2 y_1^m \omega(0) + \xi_3(y_1 y_2^2 + y_1^n \beta(0)).$$

We may here argue in a similar way as in Case 6 of the previous subsection. The complete phase has a non-degenerate critical point $y_2^c = 0$ at the origin, so that we can apply the method of stationary phase to the integration in y_2. This leads to an oscillatory integral in y_1 whose phase is given by

$$\xi_1 y_1 + \xi_2 y_1^m \omega(0) + \xi_3 y_1^n \beta(0).$$

To the remaining oscillatory integral in y_1 we may thus apply the version of Corput's estimate from Lemma 2.1 (a) in [IM16], of order $N = 3$, and altogether arrive at the estimate

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_3^{-\frac{1}{2}} \lambda_3^{-\frac{1}{3}} = \lambda_3^{-\frac{5}{6}}$$

for $\delta = 0$. The argument is stable under small perturbations, and thus this estimate remains valid for sufficiently small δ. In combination with (6.24) we may then conclude by means of Lemma 6.3 that

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} \lambda_3^{1-\frac{1}{p}} (\lambda_2 \lambda_3^{\frac{1}{2}})^{\frac{2}{p}-1} \lambda_3^{-\frac{5}{6}(2-\frac{2}{p})}.$$

Thus

$$\sum_{\lambda_1,\lambda_2 \lesssim \lambda_3} \|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} (\log \lambda_3) \lambda_3^{\frac{11}{3p}-\frac{13}{6}}.$$

We can sum the last inequality in λ_3 provided $p > 22/13$. In particular, since $p > 12/7 > 22/13$, we find that the sum over all indices $\lambda \in I_1$ considered in this case is finite, and we obtain

$$\|\mathcal{M}_{\delta,T}^{I_1}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 2: $\lambda_1 \gg \max\{\lambda_2, \lambda_3\}$. This case can be handled exactly as the corresponding case in the previous subsection by means of iterated integrations by parts in y_1, and we easily get for $p > 12/7$

$$\|\mathcal{M}_{\delta,T}^{I_2}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 3: $\lambda_2 \gg \max\{\lambda_1, \lambda_3\}$. Also this case can be handled exactly as the corresponding case in the previous subsection, and we get for $p > 12/7$

$$\|\mathcal{M}_{\delta,T}^{I_3}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 4: $\lambda_1 \sim \lambda_2$ **and** $\lambda_2 \delta_0 \lesssim \lambda_3 \ll \lambda_2$. In this case we have non-degenerate critical points in y_2 and y_1 as well. More precisely, applying first the method of stationary phase to the integration in y_2, and subsequently to the y_1-integration, we find that

$$\|\widehat{\nu_{\delta,0}^\lambda}\|_\infty \lesssim \lambda_1^{-\frac{1}{2}} \lambda_3^{-\frac{1}{2}}.$$

As in Case 4 of the previous subsection, $\lambda_3 T$ is the dominant term in $\lambda_3 T + \lambda_1 + \lambda_2$, and therefore Lemma 6.3 implies that

$$\|\mathcal{M}_{\delta,T}^\lambda\|_{p\to p} \lesssim T^{\frac{1}{p}} \lambda_3^{1-\frac{1}{p}} (\lambda_2 \lambda_3^{\frac{1}{2}})^{(\frac{2}{p}-1)} (\lambda_2^{-\frac{1}{2}} \lambda_3^{-\frac{1}{2}})^{(2-\frac{2}{p})} = T^{\frac{1}{p}} \lambda_3^{\frac{1}{p}-\frac{1}{2}} \lambda_2^{\frac{3}{p}-2}.$$

Since the exponent of λ_3 in this estimate is a positive real number, we can sum over all $\lambda_3 \ll \lambda_2$ and obtain

$$\sum_{\{\lambda_1,\lambda_3:\lambda_1\sim\lambda_2,\lambda_3\ll\lambda_2\}} \|\mathcal{M}^{\lambda}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}\lambda_2^{\frac{4}{p}-\frac{5}{2}}.$$

The expression on the right-hand side can be summed over all λ_2 provided $p > 8/5$. Since $12/7 > 8/5$, we conclude that for $p > 12/7$ we have

$$\|\mathcal{M}^{I_4}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}.$$

Case 5: $\lambda_1 \sim \lambda_2$ **and** $\lambda_3 \ll \lambda_2\delta_0$. Arguing in the same way as in the corresponding Case 5 of the previous subsection, we find that

$$\|\widehat{\nu^{\lambda}_{\delta,0}}\|_\infty \lesssim (\lambda_2\delta_0)^{-N}\lambda_2^{-\frac{1}{2}} \qquad \text{for every } N \in \mathbb{N}.$$

Therefore Lemma 6.3 implies that

$$\|\mathcal{M}^{\lambda}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{2}{p}-1}(\lambda_3 T+\lambda_2)^{1-\frac{1}{p}}(\lambda_2\lambda_3^{\frac{1}{2}})^{\frac{2}{p}-1}((\lambda_2\delta_0)^{-N}\lambda_2^{-\frac{1}{2}})^{2-\frac{2}{p}}$$

for every $N \in \mathbb{N}$. As before, we may control $\lambda_3 T + \lambda_2$ by $\lambda_2 T\delta_0$, so that

$$\|\mathcal{M}^{\lambda}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}\delta_0^{2-\frac{3}{p}}\lambda_3^{\frac{1}{p}-\frac{1}{2}}(\lambda_2\delta_0)^{-N}$$

for every $N \in \mathbb{N}$. Summing first over all $\lambda_1 \sim \lambda_2$, then all λ_2 such that $\lambda_2\delta_0 \gg \lambda_3$ and finally over all λ_3, we then find that

$$\|\mathcal{M}^{I_5}_{\delta,T}\|_{p\to p} \lesssim T^{\frac{1}{p}}\delta_0^{2-\frac{3}{p}} \le T^{\frac{1}{p}},$$

the last inequality being true if $p > 3/2$, hence in particular for $p > 12/7$. Q.E.D.

What remains is the study if D_4^+ - type singularities. As it turns out, this will indeed require an understanding also of maximal functions associated to surfaces with A_2 - type singularities, where exactly one of the principal curvatures of ϕ does not vanish at the origin. This (deeper) study will be carried out in the next section.

7. Maximal operators associated to families of surfaces with A_2 - type singularities depending on small parameters

Consider a smooth family of real-valued functions $(x_1, x_2) \mapsto \phi(x_1, x_2, \sigma)$ defined on a given open neighborhood U of the origin in $\mathbb{R}^2$ and depending smoothly on parameters σ from a given open neighborhood V of the origin in $\mathbb{R}^l$, such that

$$\phi(x_1, x_2, 0) = a_2x_2^2 + a_3x_1^3 + \phi_r(x_1, x_2), \tag{7.1}$$

where a_2, a_3 are non-zero real numbers and $\phi_r(x_1, x_2)$ is a smooth function whose Newton polyhedron satisfies $\mathcal{N}(\phi_r) \subset \{t_1/3 + t_2/2 > 1\}$ (so that $a_2x_2^2 + a_3x_1^3$ is the principal part of ϕ when $\sigma = 0$).

The goal for this section will be to prove the following

Theorem 7.1. *Denote by $\mathcal{M}^\sigma_T$ the maximal operator*

$$\mathcal{M}^\sigma_T f(y, y_3) := \sup_{t>0}\Big|\int_{\mathbb{R}^2} f(y - tx, y_3 - t(T + \phi(x,\sigma)))\eta(x,\sigma)\,dx\Big|, \tag{7.2}$$

where $T \ge 1$, and where η is a smooth, non-negative function supported in $U \times V$. Then, if we assume that the support of η is contained in a sufficiently small neighborhood of the origin, the maximal operators $\mathcal{M}^\sigma_T$ are uniformly bounded on $L^p(\mathbb{R}^3)$ in σ, for any given $p > 3/2$, with norm

$$\|\mathcal{M}^\sigma_T\|_{p\to p} \le C_{\delta,p}T^{\frac{1}{p}+\delta},$$

for every given $\delta > 0$.

Remark 7.2. As our proof will show, the same result holds true even if U is a small neighborhood of any point of distance $\lesssim 1$ to the origin and $T \gg 1$ sufficiently large. This will become important to our application of the theorem in the last Section 8.

Our proof will rely on the following result on normal forms, which is a parameter dependent version of the analogous result for singularities of type A_2 in Proposition 2.11 of [IM16]:

Lemma 7.3. *Assume that $\phi_r \equiv 0$ in (7.1). By restricting ourselves to sufficiently small open neighborhoods U of $(0,0)$ in $\mathbb{R}^2$ and V of the origin in the parameter space $\mathbb{R}^l$, we can find affine-linear coordinates depending smoothly on σ so that, in these new coordinates, the function $\phi(x_1,x_2,\sigma)$ can be written in the form*

$$\phi(x_1,x_2,\sigma) = b(x_1,x_2,\sigma)(x_2 - x_1^m\omega(x_1,\sigma))^2 + x_1^3\beta(x_1,\sigma) + \beta_1(\sigma)x_1 + \beta_0(\sigma), \tag{7.3}$$

where $b, \beta, \beta_0, \beta_1$ and ω are smooth functions and $m \geq 2$ is a positive integer, such that the following hold true:

$$b(x_1,x_2,0) = a_2 \neq 0,\ \beta(x_1,0) = a_3 \neq 0, \beta_0(0) = \beta_1(0) = 0,\ \text{and}\ \ \omega(x_1,0) = 0. \tag{7.4}$$

Proof. Following the proof of Proposition 2.11 in [IM16], we consider the equation

$$\partial_2\phi(x_1,x_2,\sigma) = 0. \tag{7.5}$$

Since $\partial_2^2\phi(0,0,0) \neq 0$, the implicit function theorem shows that locally near $(0,0,0)$, this equation has a unique, smooth solution $x_2 = \psi(x_1,\sigma)$ with $\psi(0,0) = 0$. In fact, since $\phi_r \equiv 0$, we even have $\psi(x_1,0) = 0$. A Taylor series expansion of the function $\phi(x_1,x_2,\sigma)$ with respect to the variable x_2 around $\psi(x_1,\sigma)$ then reveals that

$$\phi(x_1,x_2,\sigma) = b(x_1,x_2,\sigma)(x_2 - \psi(x_1,\sigma))^2 + b_0(x_1,\sigma), \tag{7.6}$$

where b, b_0 are smooth functions satisfying the conditions $b(x_1,x_2,0) = a_2 \neq 0$ and $b_0(0,0) = \partial_1 b_0(0,0) = \partial_1^2 b_0(0,0) = 0$ and $\partial_1^3 b_0(x_1,0) = 6a_3 \neq 0$.

Again by the implicit function theorem, we then see that the equation $\partial_1^2 b_0(x_1,\sigma) = 0$ locally near $\sigma = 0$ has a smooth solution $x_1 = x_1(\sigma)$ with $x_1(0) = 0$. Applying next the change of coordinates $x_1 \mapsto x_1 + x_1(\sigma)$, we thus see that we may assume that ϕ is of the form

$$\phi(x_1,x_2,\sigma) = b(x_1+x_1(\sigma),x_2,\sigma)(x_2 - \psi(x_1+x_1(\sigma),\sigma))^2 + x_1^3\beta(x_1,\sigma) + \beta_1(\sigma)x_1 + \beta_0(\sigma),$$

with smooth functions β, β_1, β_0 satisfying $\beta(x_1,0) = a_3 \neq 0, \beta_1(0) = 0, \beta_0(0) = 0$.

By means of a Taylor series expansion, we also see that we can write

$$\psi(x_1+x_1(\sigma),\sigma) = \psi(x_1(\sigma),\sigma) + \partial_1\psi(x_1(\sigma),\sigma)x_1 + x_1^m\omega(x_1,\sigma),$$

where ω is a smooth function and $m \geq 2$ is a positive integer. Observe that if all derivatives of $\psi(x_1+x_1(\sigma),\sigma)$ with respect to x_1 vanish at the origin, we may choose m as large as we wish. Notice also that since $\psi(x_1,0) = 0$, we have that $\omega(x_1,0) = 0$.

Finally, after applying the affine change of variables

$$(x_1, x_2 - \psi(x_1(\sigma),\sigma) - \partial_1\psi(x_1(\sigma),\sigma)x_1) \mapsto (x_1,x_2),$$

we arrive at the conclusion of the lemma. Q.E.D.

Proof of Theorem 7.1. By passing from the phase $\phi(x,\sigma)$ to $\tilde{\phi}(x,\sigma) := \phi(x,\sigma) - \phi(0,\sigma) - \nabla\phi(0,\sigma)\cdot x$ and applying a suitable linear change of coordinates to the ambient space $\mathbb{R}^3$, it is easily seen that we may assume without loss of generality that $\phi(0,\sigma) = 0$ and $\nabla\phi(0,\sigma) = (0,0)$.

As a next step, notice that the proof can be reduced to the case $\phi_r \equiv 0$ considered in Lemma 7.3, by adding further parameters to σ.

Indeed, observe that we may write

$$\phi(x_1,x_2,\sigma) = \alpha_2(\sigma)x_2^2 + \alpha_3(\sigma)x_1^3 + \phi_r(x,\sigma) + \beta_1(\sigma)x_1^2 + \beta_2(\sigma)x_1x_2,$$

where $\alpha_2(\sigma), \alpha_3(\sigma), \beta_1(\sigma)$ and $\beta_2(\sigma)$ are smooth functions of σ such that $\alpha_2(0) = a_2, \alpha_3(0) = a_3$ and $\beta_1(0) = \beta_2(0) = 0$, and where $\phi_r(x_1, x_2, \sigma)$ is smooth such that $\mathcal{N}(\phi_r) \subset \{t_1/3 + t_2/2 > 1\}$, for every σ.

Thus, if we put $\phi_{(s)}(x_1, x_2, \sigma) := \frac{1}{s}\phi(s^{1/3}x_1, s^{1/2}x_2, \sigma)$, $s > 0$, then

$$\phi_{(s)}(x_1, x_2, \sigma) = \alpha_2(\sigma)x_2^2 + \alpha_3(\sigma)x_1^3 + s^{\frac{1}{6}}\phi_r(x, \sigma, s^{\frac{1}{6}}) + \frac{\beta_1(\sigma)}{s^{\frac{1}{3}}}x_1^2 + \frac{\beta_2(\sigma)}{s^{\frac{1}{6}}}x_1x_2,$$

where also the new function ϕ_r is a smooth function of its arguments. Let us therefore view $s_1 = s^{1/6}$, $s_2 = \beta_1(\sigma)/s^{1/3}$ and $s_3 = \beta_2(\sigma)/s^{1/6}$ as new, small parameters and define a new parameter vector $\tilde{\sigma} := (s_1, s_2, s_3, \sigma)$. This allows to write

$$\phi_{(s)}(x_1, x_2, \sigma) = \alpha_2(\sigma)x_2^2 + \alpha_3(\sigma)x_1^3 + s_1\phi_r(x, \sigma, s_1) + s_2x_1^2 + s_3x_1x_2 =: \tilde{\phi}(x_1, x_2, \tilde{\sigma}).$$

Notice here that if we first choose s sufficiently small (the size of s will determine how much we have shrunk the support of the amplitude η), and then σ sufficiently small (depending on our chosen s), then we may indeed also assume that the new parameters are sufficiently small. This allows us to consider the components of $\tilde{\sigma}$ as independent, small parameters.

But then, $\tilde{\phi}(x_1, x_2, 0) = a_2x_2^2 + a_3x_1^3$, so that indeed $\tilde{\phi}_r \equiv 0$. Moreover, our discussion shows that it clearly suffices to prove the theorem for the phase $\tilde{\phi}$ in place of ϕ and $\tilde{\sigma}$ sufficiently small.

Thus, let us henceforth assume that $\phi_r \equiv 0$, so that Lemma 7.3 applies. Due to this lemma, after applying a suitable linear change of variables (depending possibly on σ) to the ambient space $\mathbb{R}^3$, we may assume that the maximal operator $\mathcal{M}_T^\sigma$ is of the following form:

$$\mathcal{M}_T^\sigma f(y, y_3) := \sup_{t>0} \left| \int f(y - t(x + \alpha(\sigma)), y_3 - t(T + E(\sigma) + \phi(x, \sigma)))\, a_0(x, \sigma)\, dx \right|, \tag{7.7}$$

where a_0 is again a smooth non-negative function supported in a sufficiently small neighborhood of the origin and

$$\phi(x_1, x_2, \sigma) = b(x_1, x_2, \sigma)(x_2 - x_1^m\omega(x_1, \sigma))^2 + x_1^3\beta(x_1, \sigma), \tag{7.8}$$

and where $\alpha(\sigma) = (\alpha_1(\sigma), \alpha_2(\sigma))$ and $E(\sigma)$ are smooth functions of σ such that $\alpha(0) = 0$ and $1/2 \leq E(\sigma) \leq 2$, and where $\omega(x_1, 0) = 0$.

For the proof of Theorem 7.1, we shall thus assume that $\mathcal{M}_T^\sigma$ is given by (7.7). The proof will be based on suitable dyadic frequency space decompositions, also with respect to the distance to certain "Airy cones" associated with our phase functions, and the study of the corresponding frequency-localized maximal operators. For the sake of simplicity of notation, we shall usually suppress the superscript σ in the proof.

7.1. **Dyadic decomposition with respect to the distance to the Airy cone.** Note that the maximal operator $\mathcal{M} = \mathcal{M}_T^\sigma$ is associated to the averaging operators given by convolutions with dilates of a measure μ whose Fourier transform at $\xi = (\xi_1, \xi_2, \xi_3)$ is given by

$$\hat{\mu}(\xi) := \int e^{-i\left(\xi_1(x_1+\alpha_1(\sigma))+\xi_2(x_2+\alpha_2(\sigma))+\xi_3(T+E(\sigma)+\phi(x_1,x_2,\sigma))\right)} a_0(x_1, x_2, \sigma)\, dx_1dx_2. \tag{7.9}$$

In order to estimate this maximal operator, we shall perform a dyadic decomposition with respect to the last variable ξ_3. The low frequency part is easily controlled by the Hardy-Littlewood maximal operator.

Let us thus denote by $\lambda \geq 2$ a sufficiently large dyadic number, and decompose $\hat{\mu}$ into

$$\widehat{\mu^\lambda}(\xi) := \chi_0\left(\frac{\xi_1}{\lambda}, \frac{\xi_2}{\lambda}\right)\chi_1\left(\frac{\xi_3}{\lambda}\right)\hat{\mu}(\xi).$$

and

$$\left(1-\chi_0\left(\frac{\xi_1}{\lambda},\frac{\xi_2}{\lambda}\right)\right)\chi_1\left(\frac{\xi_3}{\lambda}\right)\hat{\mu}(\xi),$$

where χ_0 and χ_1 are again smooth functions with sufficiently small compact supports, and χ_0 is identically 1 on a small neighborhood of the origin, whereas χ_1 vanishes near the origin and is identically one near 1.

Notice that if we choose the support of a_0 sufficiently small, then integrations by parts easily show that the latter contribution is of order $O(\lambda^{-N})$ for every $N\in\mathbb{N}$ as $\lambda\to+\infty$, so that the corresponding contribution to the maximal operators is under control.

It therefore suffices to control the contribution by μ^λ. Following [IM16] we write

$$\xi_3=\lambda s_3,\quad \xi_1=\lambda s_3 s_1,\quad \xi_2=\lambda s_3 s_2, \tag{7.10}$$

and put $s':=(s_1,s_2)$, $s:=(s',s_3)$. Then we have

$$|s_3|\sim 1\quad\text{and}\quad |s'|\ll 1$$

on the support of $\widehat{\mu^\lambda}$. Moreover, writing

$$\xi_1x_1+\xi_2x_2+\xi_3\phi(x_1,\,x_2,\sigma)=:\lambda s_3\Phi(x,s',\sigma),$$

where

$$\Phi(x,s',\sigma):=s_1x_1+s_2x_2+\phi(x,\sigma),$$

and putting

$$\Gamma(\sigma):=(\alpha_1(\sigma),\alpha_2(\sigma),T+E(\sigma)),$$

we may re-write

$$\widehat{\mu^\lambda}(\xi)=e^{-i\xi\cdot\Gamma(\sigma)}\chi_0(s_3s')\chi_1(s_3)J(\lambda,s,\sigma), \tag{7.11}$$

where $J(\lambda,s,\sigma)$ denotes the oscillatory integral

$$J(\lambda,s,\sigma):=\int_{\mathbb{R}^2}e^{-i\lambda s_3\Phi(x,s',\sigma)}a_0(x,\sigma)\,dx.$$

In view of (7.8), we perform the change of variables $x_2\mapsto x_2+x_1^m\omega(x_1,\sigma)$ and obtain

$$J(\lambda,s,\sigma)=\int_{\mathbb{R}^2}e^{-i\lambda s_3\Phi_1(x,s',\sigma)}a_0(x_1,\,x_2+x_1^m\omega(x_1,\sigma))\,dx, \tag{7.12}$$

where

$$\Phi_1(x,s',\sigma):=b(x_1,\,x_2+x_1^m\omega(x_1,\sigma),\sigma)x_2^2+s_2x_2+s_1x_1+s_2x_1^m\omega(x_1,\sigma)+x_1^3\beta(x_1,\sigma).$$

Applying the method of stationary phase to the integration in x_2, we next obtain that

$$J(\lambda,s,\sigma)=\lambda^{-1/2}\int_{\mathbb{R}}e^{-i\lambda s_3\tilde{\Psi}(x_1,s',\sigma)}\tilde{a}(x_1,\sigma)\,dx_1+r(\lambda,s),$$

where $\tilde{a}$ is another smooth bump function supported in a sufficiently small neighborhood of the origin, $r(\lambda,s)$ is a remainder term of order

$$r(\lambda,s)=O(\lambda^{-\frac{3}{2}})\quad\text{as}\quad\lambda\to+\infty,$$

and the new phase $\tilde{\Psi}$ is given by

$$\tilde{\Psi}(x_1,s',\sigma):=\Phi_1(x_1,x_2^c(x_1,s_2,\sigma),s',\sigma),$$

where $x_2^c(x_1,s_2,\sigma)$ denotes the unique (non-degenerate) critical point of the phase Φ_1 with respect to x_2.

The contribution of the error term $r(\lambda,s)$ to $\widehat{\mu^\lambda}$ and the corresponding maximal operator is easily estimated, and we shall therefore henceforth ignore it.

In order to understand $\tilde{\Psi}$, we need more information on the critical point $x_2^c(x_1, s_2, \sigma)$. Note first that the critical point must be of the form $x_2^c(x_1, s_2, \sigma) = s_2 w(x_1, s_2, \sigma)$, where, due to (7.4), the function w is smooth and satisfies the condition $w(x_1, s_2, 0) = -\frac{1}{2a_2} \neq 0$. Hence we have

$$\tilde{\Psi}(x_1, s', \sigma) = s_2^2 b_1(x_1, s_2, \sigma) + s_1 x_1 + s_2 x_1^m \omega(x_1, \sigma) + x_1^3 \beta(x_1, \sigma),$$

where b_1 is another smooth function, with $b_1(x_1, s_2, 0) = -\frac{1}{4a_2} \neq 0$.

Following [IM16] we consider the equation

$$\partial_{x_1}^2 \tilde{\Psi}(x_1, s', \sigma) = 0.$$

By the Implicit Function Theorem, we see in a similar way as before that it has a solution of the form $x_1^c(s_2, \sigma) = s_2 G_1(s_2, \sigma)$, where G_1 is a smooth function such that $G_1(s_2, 0) = 0$.

By performing finally the translation of the x_1-coordinate $x_1 \mapsto x_1 + x_1^c(s_2, \sigma)$, we see that we may write (up to an error term which, as mentioned before, can be ignored)

$$J(\lambda, s, \sigma) = \lambda^{-1/2} \int_{\mathbb{R}} e^{-i\lambda s_3 \Psi(x_1, s', \sigma)} a(x_1, s_2, \sigma)\, dx_1, \tag{7.13}$$

with a smooth amplitude a which has similar properties like $\tilde{a}$, and where the new phase Ψ is of the form

$$\Psi(x_1, s', \sigma) = B_0(s', \sigma) + B_1(s', \sigma) x_1 + x_1^3 B_3(x_1,\, s_2, \sigma), \tag{7.14}$$

where B_0, B_1, B_3 are smooth functions with $B_3(0, 0, 0) \neq 0$. One easily checks that the functions B_1, B_0 have the forms

$$B_0(s', \sigma) = s_2^2 \tilde{b}_1(s_2, \sigma) + s_1 s_2 G_1(s_2, \sigma), \quad B_1(s', \sigma) = s_1 - s_2^{m_1} G_3(s_2, \sigma), \tag{7.15}$$

where $\tilde{b}_1$ and G_3 are smooth functions with $\tilde{b}_1(0, 0) = b_1(0, 0, 0) \neq 0$ and $G_3(s_2, 0) = 0$, and where m_1 is an integer $m_1 \geq 2$.

Following [IM16] (cf. Chapter 5), we next perform a dyadic frequency decomposition of μ^λ with respect to the "Airy cone" given by $B_1(s', \sigma) = 0$.

More precisely, we choose smooth cut-off functions χ_0 and χ_1 such that $\chi_0 = 1$ on a sufficiently large neighborhood of the origin, and $\chi_1(t)$ is supported where $|t| \sim 1$ and $\sum_{k \in \mathbb{Z}} \chi_1(2^{-2k/3} t) = 1$ on $\mathbb{R} \setminus \{0\}$, and define the functions μ_{Ai}^λ and μ_k^λ by

$$\begin{aligned} \widehat{\mu_{Ai}^\lambda}(\xi) &:= \chi_0\Big(\lambda^{\frac{2}{3}} B_1(s', \sigma)\Big)\, \widehat{\mu^\lambda}(\xi), \\ \widehat{\mu_k^\lambda}(\xi) &:= \chi_1\Big((2^{-k}\lambda)^{\frac{2}{3}} B_1(s', \sigma)\Big)\, \widehat{\mu^\lambda}(\xi), \qquad M_0 \leq 2^k \leq \frac{\lambda}{M_1}, \end{aligned}$$

so that

$$\mu^\lambda = \mu_{Ai}^\lambda + \sum_{M_0 \leq 2^k \leq \frac{\lambda}{M_1}} \mu_k^\lambda. \tag{7.16}$$

Here, we may assume that M_0 and M_1 are sufficiently large positive numbers. We shall separately estimate the contributions $\mathcal{M}_{Ai}^\lambda$ by μ_{Ai}^λ and $\mathcal{M}_k^\lambda$ by the μ_k^λ to our maximal operator. More precisely, we shall prove the following lemma:

Lemma 7.4. *Let $1 < p \leq 2$. Then for every $\delta > 0$, we have*

$$\|\mathcal{M}_{Ai}^\lambda\|_{L^p \mapsto L^p} \leq C_{p,\delta}\, T^{\frac{1}{p}+\delta} \lambda^{2(\frac{1}{p}-\frac{2}{3})+\delta(\frac{2}{p}-1)} \tag{7.17}$$

and

$$\|\mathcal{M}_k^\lambda\|_{L^p \mapsto L^p} \leq C_{p,\delta}\, T^{\frac{1}{p}+\delta} 2^{k(\frac{1}{p}-\frac{2}{3})-\delta(\frac{2}{p}-1)}\, \lambda^{2(\frac{1}{p}-\frac{2}{3})+\delta(\frac{2}{p}-1)}, \tag{7.18}$$

where the constant $C_{p,\delta}$ does not depend on σ, T and λ.

If $p > 3/2$, we see that these estimates sum in k as well as over all dyadic $\lambda \gg 1$, provided we choose δ sufficiently small. This will then complete the proof of Theorem 7.1.

7.2. **The contribution by the region near the Airy cone.** In this subsection, we shall prove estimate (7.17). To this end, we write, according to (7.11),

$$\widehat{\mu^\lambda_{Ai}}(\xi) = e^{-i\xi\cdot\Gamma(\sigma)}\chi_0\Big(\lambda^{\frac23}B_1(s',\sigma)\Big)\,\chi_0(s_3s')\chi_1(s_3)J(\lambda,s).$$

Then, by (7.13), we may apply Lemma 2.2 (a) (with $B=3$) in Chapter 2 of [IM16] to $J(\lambda,s)$ and see that we may write

$$\widehat{\mu^\lambda_{Ai}}(\xi) = e^{-i\xi\cdot\Gamma(\sigma)}\,\lambda^{-\frac56}\chi_0\Big(\lambda^{\frac23}B_1(s',\sigma)\Big)\,g(\lambda^{\frac23}B_1(s',\sigma),\lambda,\sigma,s)\chi_0(s_3s')\chi_1(s_3)\,e^{-i\lambda s_3B_0(s',\sigma)},$$

where $g(u,\lambda,\sigma,s)$ is a smooth function of (u,λ,σ,s) whose derivatives of any order are uniformly bounded on its natural domain $|u|\lesssim 1, |\sigma|\lesssim 1, \lambda\ge 2$ and $|s_3|\sim 1$, $|s'|\ll 1$.

Notice also that a first order t-derivative of $e^{-it\xi\cdot\Gamma(\sigma)}$, as well as of $e^{-i\lambda ts_3B_0(s',\sigma)}$, produces additional factors of order $O(T\lambda)$, since $|\Gamma(\sigma)|\sim T$, and thus following again the arguments in the proof of Proposition 4.1 we easily see that

$$\|\mathcal{M}^\lambda_{Ai}\|_{L^2\to L^2}\lesssim T^{\frac12}\lambda^{\frac12-\frac56} = T^{\frac12}\lambda^{-\frac13}. \tag{7.19}$$

Next, we we shall control the function $\mu^\lambda_{Ai}(y)$. By Fourier inversion, changing variables as in (7.10), we may write

$$\begin{aligned}\mu^\lambda_{Ai}(y+\Gamma(\sigma)) \;&=\; \lambda^{\frac{13}{6}}\int\chi_0\Big(\lambda^{\frac23}B_1(s',\sigma)\Big)g(\lambda^{\frac23}B_1(s',\sigma),\lambda,\sigma,s)\chi_0(s_3s')\chi_1(s_3)\\ &\qquad\qquad e^{-i\lambda s_3\big(B_0(s',\sigma)-s_1y_1-s_2y_2-y_3\big)}s_3^2\,ds.\end{aligned}\tag{7.20}$$

Observe first that when $|y|\gg 1$, then we easily obtain by means of integrations by parts that

$$|\mu^\lambda_{Ai}(y)|\le C_N\lambda^{-N},\quad N\in\mathbb{N},\ \text{if } |y|\gg 1.$$

Indeed, when $|y_1|\gg 1$, then we integrate by parts repeatedly in s_1 to see this (in each integration by parts, we gain a factor λ^{-1} and loose at most a factor of $\lambda^{2/3}$), and a similar argument applies when $|y_2|\gg 1$, where we use the s_2-integration. Finally, when $|y_1|+|y_2|\lesssim 1$ and $|y_3|\gg 1$, then we can integrate by parts in s_3 in order to establish this estimate.

We may therefore assume from here on that $|y|\lesssim 1$. In view of (7.15), we then perform yet another change of coordinates from s_1 to

$$z := \lambda^{\frac23}B_1(s',\sigma) = \lambda^{\frac23}(s_1-s_2^{m_1}G_3(s_2,\sigma)),\ \text{i.e., } s_1 = s_2^{m_1}G_3(s_2,\sigma)+\lambda^{-\frac23}z.$$

Note that by (7.15), the function $B_0(s',\sigma)$ is then given by

$$B_0(s',\sigma) = s_2^2G_5(s_2,\sigma)+s_2G_1(s_2,\sigma)\lambda^{-\frac23}z,$$

where $G_5(s_2,\sigma) := \tilde b_1(s_2,\sigma)+s_2^{m_1-1}G_1(s_2,\sigma)G_3(s_2,\sigma)$ is smooth and where we may assume that $G_5(0,0)\neq 0$, since $m_1\ge 2$ and $G_1(s_2,0)=0$. We may thus re-write

$$\begin{aligned}\mu^\lambda_{A_i}(y+\Gamma(\sigma)) \;&=\; \lambda^{\frac32}\int\chi_0(z)\tilde g(z,\lambda,\sigma,s_2,s_3)\chi_0(z,s_2,s_3)\chi_1(s_3)\\ &\qquad\qquad e^{-i\lambda s_3\Phi_2(z,s_2,y,\sigma)}ds_2ds_3dz,\end{aligned}\tag{7.21}$$

where $\tilde g$ has similar properties to g, and where the phase Φ_2 is of the form

$$\begin{aligned}\Phi_2(z,s_2,y,\sigma) \;&=\; s_2^2G_5(s_2,\sigma)-s_2^{m_1}G_3(s_2,\sigma)y_1-s_2y_2-y_3\\ &\qquad +\lambda^{-\frac23}z(s_2G_1(s_2,\sigma)-y_1),\end{aligned}$$

with $G_5(0,0)\neq 0$ and $G_1(s_2,0)=G_3(s_2,0)=0$ (compare this with the analogous formula (5.21) in [IM16]). Recall also that $|z|\lesssim 1, |s_2|\ll 1$ and $|s_3|\sim 1$.

Decomposing

$$G_1(s_2,\sigma)=G_{1,1}(\sigma)+s_2G_{1,2}(s_2,\sigma), \tag{7.22}$$

where $G_{1,1}(0)=0$ and $G_{1,2}(s_2,0)=0$, we finally may re-write

$$\begin{aligned}\Phi_2(z,s_2,y,\sigma) &= s_2^2G_6(\lambda^{-\frac{2}{3}}z,s_2,\sigma)-s_2\big(y_2-\lambda^{-\frac{2}{3}}z\,G_{1,1}(\sigma)\big)\\ &\qquad -s_2^{m_1}G_3(s_2,\sigma)y_1-\lambda^{-\frac{2}{3}}zy_1-y_3,\end{aligned} \tag{7.23}$$

where $G_6(0,0,0)\neq 0$.

Assuming σ to be sufficiently small, we thus see that Φ_2 has a unique critical point s_2^c with respect to s_2, of size $|y_2|$, provided $|y_2|\ll 1$. Otherwise, iterated integrations by parts in s_2 will show that $|\mu^\lambda_{A_i}(y)|\lesssim\lambda^{-N}$ for every $N\in\mathbb{N}$, so that those $y's$ can be ignored. More precisely, from (7.23) we deduce that s_2^c is of the form

$$s_2^c=\big(y_2-\lambda^{-\frac{2}{3}}z\,G_{1,1}(\sigma)\big)H_1(\lambda^{-\frac{2}{3}}z,y_1,y_2,\sigma).$$

After applying the method of stationary phase in the s_2-variable, we thus find that

$$\mu^\lambda_{Ai}(y+\Gamma(\sigma))=\lambda\int e^{-i\lambda s_3\Phi_3(\lambda^{-\frac{2}{3}}z,y,\sigma)}\chi_0(z,s_3)\chi_1(s_3)a(\lambda^{-\frac{2}{3}}z,y,s_3,\sigma)\,dzds_3+O(1),$$

with a smooth amplitude a, and where the phase Φ_3 is of the form

$$\Phi_3(\lambda^{-\frac{2}{3}}z,y,\sigma)=\big(y_2-\lambda^{-\frac{2}{3}}z\,G_{1,1}(\sigma)\big)^2H_2(\lambda^{-\frac{2}{3}}z,y_1,y_2,\sigma)-\lambda^{-\frac{2}{3}}zy_1-y_3,$$

with a smooth function H_2 such that $H_2(0,0,0,0)\neq 0$. Notice that the functions χ_0 and χ_1 may possibly be different in different places.

A Taylor series expansion of H_2 with respect to $v=\lambda^{-\frac{2}{3}}z$ then allows to write

$$\begin{aligned}\Phi_3(\lambda^{-\frac{2}{3}}z,y,\sigma) &= y_2^2H_2(0,y_1,y_2,\sigma)+\lambda^{-\frac{2}{3}}z\big(-2y_2G_{1,1}(\sigma)H_2(0,y_1,y_2,\sigma)\\ &\quad + y_2^2\partial_vH_2(0,y_1,y_2,\sigma)-y_1\big)-y_3+O(\lambda^{-4/3}).\end{aligned} \tag{7.24}$$

The factor $e^{-i\lambda s_3O(\lambda^{-4/3})}$ corresponding to the term $O(\lambda^{-4/3})$ can be included into the amplitude, and thus we may assume that the complete phase is of the form

$$\lambda s_3\Phi_3=\lambda^{\frac{1}{3}}s_3z(-y_2H_3(y_1,y_2,\sigma)+y_2^2H_4(y_1,y_2,\sigma)-y_1)+\lambda s_3(y_2^2H_2(0,y_1,y_2,\sigma)-y_3).$$

By the implicit function theorem, we may re-write the first factor in parentheses in the form

$$-y_2H_3(y_1,y_2,\sigma)+y_2^2H_4(y_1,y_2,\sigma)-y_1=(y_1-\varphi(y_2,\sigma))H_5(y_1,y_2,\sigma),$$

where $\varphi(y_2,\sigma)$ and $H_5(y_1,y_2,\sigma)$ are smooth functions with $\varphi(0,0)=0$ and $H_5(0,0,0)\neq 0$. We shall also write $\psi(y_1,y_2,\sigma):=y_2^2H_2(0,y_1,y_2,\sigma)$. Then also ψ is smooth and $\psi(0,0,0)=0$.

Thus, eventually we may write

$$\mu^\lambda_{Ai}(y+\Gamma(\sigma))=\mu^\lambda_I(y+\Gamma(\sigma))+\mu^\lambda_{II}(y+\Gamma(\sigma)),$$

where $\mu^\lambda_{II}(y+\Gamma(\sigma))=O(1)$, and

$$\mu^\lambda_I(y+\Gamma(\sigma)):=\lambda\int e^{-i\lambda s_3\Phi_3(\lambda^{-\frac{2}{3}}z,y,\sigma)}\chi_0(z,s_3)\chi_1(s_3)a(\lambda^{-\frac{2}{3}}z,y,s_3,\sigma)\,dzds_3, \tag{7.25}$$

with

$$\lambda\Phi_3(\lambda^{-\frac{2}{3}}z,y,\sigma)=\lambda^{\frac{1}{3}}z\,(y_1-\varphi(y_2,\sigma))H_5(y_1,y_2,\sigma)-\lambda\big(y_3-\psi(y_1,y_2,\sigma)\big).$$

The contribution of μ^λ_{II} to the maximal operator is of order $O(T)$, as can easily be seen by comparison with the Hardy-Littlewood maximal operator (or by Poposition 4.2).

As for μ_I^λ, fix a sufficiently small number $\delta > 0$. If $\lambda^{\frac{1}{3}}|y_1 - \varphi(y_2,\sigma)| \geq c\lambda^\delta$, where $c > 0$ is assumed to be a sufficiently small constant, then we can use iterated integrations by parts in z to obtain that $|\mu_{Ai}^\lambda(y + \Gamma(\sigma))| \lesssim \lambda^{-N}$ for every N, uniformly in σ. Similarly, if $\lambda^{\frac{1}{3}}|y_1 - \varphi(y_2,\sigma)| < c\lambda^\delta$ and $\lambda|y_3 - \psi(y_1, y_2, \sigma)| \geq \lambda^\delta$, then we can use integrations by parts in s_3 to arrive at the same type of estimate.

The contributions by these regions to our maximal operator are thus negligible.

Let us next put

$$B_\delta := \{y \in \mathbb{R}^3 : |y| \lesssim 1,\ |y_1 - \varphi(y_2,\sigma)| < \lambda^{\delta-\frac{1}{3}},\ |y_3 - \psi(y_1, y_2, \sigma)| < \lambda^{\delta-1}\}.$$

Writing $x = y + \Gamma(\sigma)$, we are thus reduced to concentrating for $\mu_{Ai}^\lambda(x)$ on the small x- region $A_\delta := \Gamma(\sigma) + B_\delta$, on which we only have the estimate $|\mu_{Ai}^\lambda(x)| \lesssim \lambda$. But recall that $\Gamma(\sigma) = (\alpha_1(\sigma), \alpha_2(\sigma), T + E(\sigma))$, with $|\alpha_1(\sigma)|, |\alpha_2(\sigma)| \ll 1$, $1/2 \leq E(\sigma) \leq 2$ and $T \gg 1$. It is then obvious that the spherical projection $\pi(A_\delta) \subset S^2$ of A_δ has measure $|\pi(A_\delta)| \lesssim T^{-2}\lambda^{\delta-\frac{1}{3}}$, since $\lambda^{\delta-1} \ll \lambda^{\delta-\frac{1}{3}}$, so that Proposition 4.2 (with $R \sim T$) shows that, for every $\varepsilon > 0$, $\mathcal{M}_{Ai}^\lambda$ satisfies the estimate

$$\|\mathcal{M}_{Ai}^\lambda\|_{L^{1+\varepsilon}\mapsto L^{1+\varepsilon}} \leq C'_{\varepsilon,\delta} T^3 T^{-2}\lambda \cdot \lambda^{\delta-\frac{1}{3}} + O(T) \leq C_{\varepsilon,\delta} T \lambda^{\frac{2}{3}+\delta}, \tag{7.26}$$

for every $\delta > 0$.

The estimate (7.17) in Lemma 7.4 now follows from the estimates (7.19) and (7.26) by real interpolation (with a slightly bigger δ than the one considered here), if we choose ε sufficiently small.

7.3. **The contribution by the region away from the Airy cone.** According to (7.11), we have

$$\widehat{\mu_k^\lambda}(\xi) = e^{-i\xi\cdot\Gamma(\sigma)}\chi_1\Big((2^{-k}\lambda)^{\frac{2}{3}} B_1(s',\sigma)\Big)\,\chi_0(s_3 s')\chi_1(s_3) J(\lambda, s).$$

By (7.13) (ignoring again the error term $r(\lambda, s)$), this can be re-written as

$$\widehat{\mu_k^\lambda}(\xi) = \lambda^{-\frac{1}{2}} e^{-i\xi\cdot\Gamma(\sigma)}\chi_1\Big((2^{-k}\lambda)^{\frac{2}{3}} B_1(s',\sigma)\Big)\,\chi_0(s_3 s')\chi_1(s_3) e^{-i\lambda s_3 B_0(s',\sigma)} \tilde{J}(\lambda, s, \sigma), \tag{7.27}$$

where

$$\tilde{J}(\lambda, s, \sigma) := \int_{\mathbb{R}} e^{-i\lambda s_3 \tilde{\Psi}(x_1, s', \sigma)} a(x_1, s_2, \sigma)\, dx_1,$$

with phase function

$$\tilde{\Psi}(x_1,\ s', \sigma) := B_1(s',\sigma)x_1 + x_1^3 B_3(x_1,\ s_2, \sigma).$$

Recall also that the amplitude a is smooth and has a sufficiently small support in x_1.

Since $B_1(s',\sigma)$ is of size $(2^k\lambda^{-1})^{\frac{2}{3}}$, we scale by changing coordinates $x_1 = (2^k\lambda^{-1})^{\frac{1}{3}} u_1$ in the integral for $\tilde{J}(\lambda, s, \sigma)$ and obtain

$$\tilde{J}(\lambda, s, \sigma) = (2^k\lambda^{-1})^{\frac{1}{3}} \int e^{-i2^k s_3 \Psi_1(u_1,\, s', \sigma)} a((2^k\lambda^{-1})^{\frac{1}{3}} u_1, s_2, \sigma)\, du_1,$$

where

$$\Psi_1(u_1, s', \sigma) := (2^{-k}\lambda)^{\frac{2}{3}} B_1(s',\sigma)u_1 + B_3((2^k\lambda^{-1})^{\frac{1}{3}} u_1, s_2, \sigma)u_1^3. \tag{7.28}$$

Observe that the coefficients of u_1 and u_1^3 in Ψ_1 are of size 1, so that Ψ_1 has no critical point with respect to u_1 unless $|u_1| \sim 1$. Thus we may choose a smooth cut-off function $\chi_1 \in C_0^\infty(\mathbb{R})$ which vanishes near the origin so that Ψ_1 has no critical point on the support of the function $1 - \chi_1$, and decompose the integral

$$\tilde{J}(\lambda, s, \sigma) = J_1(\lambda, s, \sigma) + J_\infty(\lambda, s, \sigma),$$

where

$$J_1(\lambda, s, \sigma) := (2^k\lambda^{-1})^{\frac{1}{3}} \int e^{-i2^k s_3 \Psi_1(u_1,\, s', \sigma)} a((2^k\lambda^{-1})^{\frac{1}{3}} u_1, s_2, \sigma)\chi_1(u_1)\, du_1 \tag{7.29}$$

and

$$J_\infty(\lambda,s,\sigma):=(2^k\lambda^{-1})^{\frac13}\int e^{-i2^ks_3\Psi_1(u_1,s',\sigma)}a((2^k\lambda^{-1})^{\frac13}u_1,s_2,\sigma)(1-\chi_1(u_1))\,du_1. \tag{7.30}$$

Accordingly we decompose the measure $\mu_k^\lambda=\mu_{k,1}^\lambda+\mu_{k,\infty}^\lambda$, where the summands are given by

$$\widehat{\mu_{k,1}^\lambda}(\xi)=\lambda^{-\frac12}e^{-i\xi\cdot\Gamma(\sigma)}\chi_1\Big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma)\Big)\,\chi_0(s_3s')\chi_1(s_3)e^{i\lambda s_3B_0(s',\sigma)}J_1(\lambda,s,\sigma).$$

and

$$\widehat{\mu_{k,\infty}^\lambda}(\xi)=\lambda^{-\frac12}e^{-i\xi\cdot\Gamma(\sigma)}\chi_1\Big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma)\Big)\,\chi_0(s_3s')\chi_1(s_3)e^{i\lambda s_3B_0(s',\sigma)}J_\infty(\lambda,s,\sigma).$$

We denote by $\mathcal{M}_{k,1}^\lambda$ and $\mathcal{M}_{k,\infty}^\lambda$ the maximal operators defined by the functions $\mu_{k,1}^\lambda$ and $\mu_{k,\infty}^\lambda$, respectively.

7.3.1. *The contributions given by the $\mu_{k,\infty}^\lambda$.* Recall that $2^k\lambda^{-1}\ll1$. By means of integrations by parts, we then easily see that, given any $N\in\mathbb{N}$, we may write

$$J_\infty(\lambda,s,\sigma)=(2^k\lambda^{-1})^{\frac13}2^{-kN}g_N\big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma),2^k\lambda^{-1},s',\sigma\big), \tag{7.31}$$

where g_N is smooth. This implies in particular that

$$|J_\infty(\lambda,s,\sigma)|\lesssim(2^k\lambda^{-1})^{\frac13}2^{-kN},$$

hence

$$\|\widehat{\mu_{k,\infty}^\lambda}\|_\infty\lesssim\lambda^{-\frac12}(2^k\lambda^{-1})^{\frac13}2^{-kN}.$$

Arguing as before, we thus find that

$$\|\mathcal{M}_{k,\infty}^\lambda\|_{L^2\mapsto L^2}\lesssim T^{\frac12}(2^k\lambda^{-1})^{\frac13}2^{-kN} \tag{7.32}$$

for every $N\in\mathbb{N}$.

Next, we proceed in a similar way as in the previous subsection in order to obtain L^p-estimates for $\mathcal{M}_{k,\infty}^\lambda$ when p is close to 1. By Fourier inversion and (7.31), we find that

$$\begin{aligned}\mu_{k,\infty}^\lambda(y+\Gamma(\sigma))&=\lambda^3\int_{\mathbb{R}^3}e^{i\lambda s_3(s_1y_1+s_2y_2+y_3)}e^{i\xi\cdot\Gamma(\sigma)}\widehat{\mu_{k,\infty}^\lambda}(\xi)ds\\&=\lambda^{\frac52}\int\chi_1\Big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma)\Big)\,\chi_0(s_3s')\chi_1(s_3)e^{-i\lambda s_3\big(B_0(s',\sigma)-s_1y_1-s_2y_2-y_3\big)}J_\infty(\lambda,s,\sigma)\,ds,\\&=\lambda^{\frac52}(2^k\lambda^{-1})^{\frac13}2^{-kN}\int e^{-i\lambda s_3\big(B_0(s',\sigma)-s_1y_1-s_2y_2-y_3\big)}\chi_1\Big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma)\Big)\,\chi_0(s_3s')\chi_1(s_3)\\&\qquad g_N\big((2^{-k}\lambda)^{\frac23}B_1(s',\sigma),2^k\lambda^{-1},s',\sigma\big)\,ds,\end{aligned}$$

where we recall that $\xi=\lambda s_3(s_1,s_2,1)$.

Comparing this with (7.20) and arguing in a similar way is we did for the estimation of μ_{Ai}^λ, we again see that we may assume that $|y|\lesssim1$, since for $|y|\gg1$ integrations by parts in s show that, for every $N\in\mathbb{N}$,

$$|\mu_{k,\infty}^\lambda(y)|\le C_N2^{-kN}\lambda^{-N},\quad N\in\mathbb{N},\ \text{if } |y|\gg1.$$

Recalling (7.15), we next change coordinates from s_1 to

$$z:=(2^{-k}\lambda)^{\frac23}B_1(s_1,s_2,\sigma)=(2^{-k}\lambda)^{\frac23}(s_1-s_2^{m_1}G_3(s_2,\sigma)),$$

i.e.,

$$s_1=s_2^{m_1}G_3(s_2,\sigma)+(2^k\lambda^{-1})^{\frac23}z,\qquad|z|\sim1. \tag{7.33}$$

Note that the function $B_0(s',\sigma)$ is then given by

$$B_0(s',\sigma)=s_2^2G_5(s_2,\sigma)+s_2G_1(s_2,\sigma)(2^k\lambda^{-1})^{\frac23}z,$$

where $G_5(s_2,\sigma) := \tilde{b}_1(s_2,\sigma) + s_2^{m_1-1}G_1(s_2,\sigma)G_3(s_2,\sigma)$ is smooth and where we may assume that $G_5(0,0)\neq 0$, since $m_1\geq 2$ and $G_1(s_2,0)=0$. We then find that, in analogy with (7.21),

$$\mu^\lambda_{k,\infty}(y+\Gamma(\sigma)) = \lambda^{\frac{5}{2}}(2^k\lambda^{-1})2^{-kN}\int e^{-i\lambda s_3\Phi_2(z,s_2,y,\sigma)}\chi_1(z)\,\chi_0(s_3s')\chi_1(s_3)$$
$$a_N\big(z,(2^k\lambda^{-1})^{\frac{1}{3}},s_2,s_3,\sigma\big)\,ds_2ds_3dz, \tag{7.34}$$

where a_N is smooth, and where the phase function Φ_2 is given by

$$\begin{aligned}\Phi_2(z,s_2,y,\sigma) \quad &= \quad s_2^2G_5(s_2,\sigma) - s_2^{m_1}G_3(s_2,\sigma)y_1 - s_2y_2 - y_3\\ &\quad +(2^k\lambda^{-1})^{\frac{2}{3}}z(s_2G_1(s_2,\sigma)-y_1).\end{aligned} \tag{7.35}$$

Here, $G_5(0,0)\neq 0, G_1(s_2,0)=G_3(s_2,0)=0$, $2^k\lambda^{-1}\ll 1$, $|z|\sim 1, |s_2|\ll 1$ and $|s_3|\sim 1$.

Observe that, due to the presence of the factor 2^{-kN} in (7.34), we may concentrate primarily on gaining negative powers of λ in the estimation of $\mu^\lambda_{k,\infty}$. We may thus proceed very much in the same way as we did in the previous section, replacing $v=\lambda^{-2/3}z$ by $v:=(2^k\lambda^{-1})^{2/3}z$, and arrive at the following analogue of (7.25):

$$\begin{aligned}\mu^\lambda_{k,\infty}(y+\Gamma(\sigma)) \quad &= \quad 2^{-kN}\lambda\int e^{-i\lambda s_3\Phi_3((2^k\lambda^{-1})^{\frac{2}{3}}z,y,\sigma)}\\ &\qquad a_N(z,(2^k\lambda^{-1})^{\frac{2}{3}},s_3,y,\sigma)\chi_1(z)\chi_1(s_3)\,dzds_3+O(2^{-kN}),\end{aligned} \tag{7.36}$$

where here a_N is of the form

$$a_N(z,(2^k\lambda^{-1})^{\frac{2}{3}},s_3,y,\sigma) = \tilde{a}_N(z,(2^k\lambda^{-1})^{\frac{2}{3}},s_3,y,\sigma)\,e^{-i\lambda s_3(2^k\lambda^{-1})^{\frac{4}{3}}\eta\left((2^k\lambda^{-1})^{\frac{2}{3}}z,y,\sigma\right)},$$

with a smooth function $\tilde{a}_N$ and a smooth, real function η, and where $\lambda\Phi_3$ is of the form

$$\lambda\Phi_3((2^k\lambda^{-1})^{\frac{2}{3}}z,y,\sigma) = \lambda(2^k\lambda^{-1})^{\frac{2}{3}}z\,\big(y_1-\varphi(y_2,\sigma)\big)H_5(y_1,y_2,\sigma)\big) - \lambda\big(y_3-\psi(y_1,y_2,\sigma)\big),$$

with $H_5(0,0,0)\neq 0$. The oscillatory factor $e^{-i\lambda s_3(2^k\lambda^{-1})^{\frac{4}{3}}\eta((2^k\lambda^{-1})^{\frac{2}{3}}z,y,\sigma)}$ in this amplitude corresponds to the $O(\lambda^{-4/3})$ - term in (7.24). The contribution by the $O(2^{-kN})$ term is negligible, in a similar way as we saw this for the contribution of the term μ^λ_{II} before, so that we shall from now on ignore it.

Fix again a sufficiently small number $\delta>0$. If $2^{\frac{2k}{3}}\lambda^{\frac{1}{3}}|y_1-\varphi(y_2,\sigma)|\geq c\lambda^\delta$, where $c>0$ is assumed to be a sufficiently small constant, then we can use iterated integrations by parts in z to obtain that $|\mu^\lambda_{k,\infty}(y+\Gamma(\sigma))|\lesssim(2^k\lambda)^{-N}$ for every N, uniformly in σ. Indeed, since $\lambda(2^k\lambda^{-1})^{6/3}=2^{2k}\lambda^{-1}\lesssim 2^k$, in each integration by parts, we gain a factor $\lambda^{-\delta}$ and loose a factor 2^k from differentiating the amplitude, but this is acceptable. Similarly, if $2^{\frac{2k}{3}}\lambda^{\frac{1}{3}}|y_1-\varphi(y_2,\sigma)|<c\lambda^\delta$ and $\lambda|y_3-\psi(y_1,y_2,\sigma)|\geq\lambda^\delta$, then we can use integrations by parts in s_3 to arrive at the same type of estimate.

The contributions by these regions to our maximal operator are thus negligible.

Let us next put

$$B_{k,\delta} := \{y\in\mathbb{R}^3 : |y|\lesssim 1,\, |y_1-\varphi(y_2,\sigma)|<2^{-\frac{2k}{3}}\lambda^{\delta-\frac{1}{3}},\, |y_3-\psi(y_1,y_2,\sigma)|<\lambda^{\delta-1}\}.$$

As in the preceding subsection, let us write $x=y+\Gamma(\sigma)$. We are then reduced to concentrating for $\mu^\lambda_{k,\infty}(x)$ on the small x- region $A_{k,\delta}:=\Gamma(\sigma)+B_{k,\delta}$, on which we have the estimate $|\mu^\lambda_{k,\infty}(x)|\lesssim 2^{-kN}\lambda$. Recalling again that $\Gamma(\sigma)=(\alpha_1(\sigma),\alpha_2(\sigma),T+E(\sigma))$, with $|\alpha_1(\sigma)|,|\alpha_2(\sigma)|\ll 1$, $1/2\leq E(\sigma)\leq 2$ and $T\gg 1$, we then see that the spherical projection $\pi(A_{k,\delta})\subset S^2$ of $A_{k,\delta}$ has measure $|\pi(A_{k,\delta})|\lesssim T^{-2}2^{-\frac{2k}{3}}\lambda^{\delta-\frac{1}{3}}$, so that Proposition 4.2 shows that, for every $\varepsilon>0$, $\mathcal{M}^\lambda_{k,\infty}$ satisfies the estimates

$$\|\mathcal{M}^\lambda_{k,\infty}\|_{L^{1+\varepsilon}\mapsto L^{1+\varepsilon}}\leq C_{N,\varepsilon}T2^{-kN}\lambda\cdot\lambda^{\delta-\frac{1}{3}} = T2^{-kN}\lambda^{\frac{2}{3}+\delta}, \tag{7.37}$$

for every $\delta > 0$, uniformly in σ.

7.3.2. *The contributions given by the* $\mu^\lambda_{k,1}$. We next estimate the maximal operators $\mathcal{M}^\lambda_{k,1}$. First, by applying the method of stationary phase to the integral $J_1(\lambda, s)$, we easily see that

$$\|\widehat{\mu^\lambda_{k,1}}(\xi)\|_\infty \lesssim \lambda^{-\frac{1}{2}}(2^k\lambda^{-1})^{\frac{1}{3}}2^{-\frac{k}{2}} = 2^{-\frac{k}{6}}\lambda^{-\frac{5}{6}}.$$

This implies that

$$\|\mathcal{M}^\lambda_{k,1}\|_{L^2\mapsto L^2} \lesssim T^{\frac{1}{2}}\lambda^{\frac{1}{2}}2^{-\frac{k}{6}}\lambda^{-\frac{5}{6}} = T^{\frac{1}{2}}2^{-\frac{k}{6}}\lambda^{-\frac{1}{3}}. \tag{7.38}$$

In order to obtain L^p-estimates for $\mathcal{M}^\lambda_{k,\infty}$ when p is close to 1, we have to analyze more carefully which phase function arises from the application of the method of stationary phase to $J_1(\lambda, s)$. To this end, let us put again $z := (2^{-k}\lambda)^{\frac{2}{3}}B_1(s', \sigma)$, so that the phase (7.28) in $J_1(\lambda, s)$ can be written

$$\Psi_1(u_1, s_2, z, \sigma) = zu_1 + B_3((2^k\lambda^{-1})^{\frac{1}{3}}u_1, s_2, \sigma)u_1^3,$$

where $|u_1| \sim 1$. We may assume that it has a critical point $u_1^c = u_1^c(s_2, z, s_2)$ of size $|u_1^c| \sim 1$, for other wise we can integrate by parts in u_1 and can then proceed as we did for the $\mu^\lambda_{k,\infty}$. Assuming for instance that $z > 0$, by writing $u_1 = z^{1/2}w_1$, we see that u_1^c is of the form $u_1^c = z^{1/2}W((2^k\lambda^{-1})^{\frac{1}{3}}z^{\frac{1}{2}}, s_2, \sigma)$, where W is smooth and $|W| \sim 1$. Moreover, since $B_3(\cdot, 0, 0) \neq 0$ is constant, it is easy to see that by choosing σ and $|s'|$ sufficiently small, we get

$$\Psi_1(u_1^c, s_2, z, \sigma) = z^{\frac{3}{2}}H_1((2^k\lambda^{-1})^{\frac{1}{3}}z^{\frac{1}{2}}, s_2, \sigma),$$

with a smooth function H_1 such that $|H_1| \sim 1$. We may thus write

$$J_1(\lambda, s, \sigma) = 2^{-\frac{k}{6}}\lambda^{-\frac{1}{3}}\, e^{-i2^k s_3 z^{\frac{3}{2}} H_1((2^k\lambda^{-1})^{\frac{1}{3}}z^{\frac{1}{2}}, s_2, \sigma)} a((2^k\lambda^{-1})^{\frac{1}{3}}z^{\frac{1}{2}}, s_2, \sigma)s_3^{-\frac{1}{2}}, \tag{7.39}$$

again with a smooth amplitude a. More precisely, we would also pick up an error term of order $O(2^{-\frac{7k}{6}}\lambda^{-\frac{1}{3}})$, which we shall here ignore for the sake of simplicity of the presentation. As explained in [IM16], such an error term could be avoided by replacing the "gain" $2^{-k/2}$ in this application of the method of stationary phase by a symbol of order $-1/2$ in 2^k (which would, however, also depend on further variables). Changing then again variables from s_1 to z, in analogy with (7.34) and (7.35) we thus find that

$$\begin{aligned}\mu^\lambda_{k,1}(y + \Gamma(\sigma)) = \lambda^{\frac{5}{2}}(2^k\lambda^{-1})2^{-\frac{k}{2}} \int e^{-i\lambda s_3 \Phi_2(z, s_2, y, \sigma)}\chi_1(z)\,\chi_0(s_3 s')\chi_1(s_3)\\ a\big(z, (2^k\lambda^{-1})^{\frac{1}{3}}, s_2, \sigma\big)\, ds_2 ds_3 dz,\end{aligned} \tag{7.40}$$

where a is smooth, and where the phase function Φ_2 is given by

$$\begin{aligned}\Phi_2(z, s_2, y, \sigma) \;&=\; 2^k\lambda^{-1}z^{\frac{3}{2}}H_1((2^k\lambda^{-1})^{\frac{1}{3}}z^{\frac{1}{2}}, s_2, \sigma)\\ &+s_2^2G_5(s_2, \sigma) - s_2^{m_1}G_3(s_2, \sigma)y_1 - s_2y_2 - y_3 + (2^k\lambda^{-1})^{\frac{2}{3}}z(s_2G_1(s_2, \sigma) - y_1).\end{aligned}$$

Here, $G_5(0, 0) \neq 0, G_1(s_2, 0) = G_3(s_2, 0) = 0$, $2^k\lambda^{-1} \ll 1$, $|z| \sim 1, |s_2| \ll 1$ and $|s_3| \sim 1$.

In particular, we see that we can next apply the method of stationary phase to the integration in s_2 (assuming that there is a critical point s_2^c within the domain of integration; otherwise, we may pick up powers λ^{-N} by means of integrations by parts and are done). Let us also write

$$q := (2^k\lambda^{-1})^{\frac{1}{3}},$$

so that $q \ll 1$.

Let us also put $H_{1,1}(s_2,\sigma) := H_1(0,s_2,\sigma)$. Note that $|H_{1,1}| \sim 1$, since $|H_1| \sim 1$ (cf. (7.39)). Expanding in powers of q, we then find that

$$\begin{aligned} \Phi_2(z,s_2,y,\sigma) &= q^3 z^{\frac{3}{2}} H_{1,1}(s_2,\sigma) + q^2 z(s_2 G_1(s_2,\sigma) - y_1) \\ &\quad + s_2^2 G_5(s_2,\sigma) - s_2^{m_1} G_3(s_2,\sigma) y_1 - s_2 y_2 - y_3 + O(q^4). \end{aligned} \tag{7.41}$$

In order to get more precise information on the critical point s_2^c, let us again decompose

$$G_1(s_2,\sigma) = G_{1,1}(\sigma) + s_2 G_{1,2}(s_2,\sigma),$$

as in (7.22), and similarly

$$H_{1,1}(s_2,\sigma) = H_{1,2}(\sigma) + s_2 H_{1,3}(\sigma) + s_2^2 H_{1,4}(s_2,\sigma).$$

Then we may re-write

$$\begin{aligned} \Phi_2(z,s_2,y,\sigma) &= q^3 z^{\frac{3}{2}} H_{1,2}(\sigma) - q^2 z y_1 - y_3 - s_2\big(y_2 - q^2 z G_{1,1}(\sigma) - q^3 z^{\frac{3}{2}} H_{1,3}(\sigma)\big) \\ &\quad + s_2^2 G_6(q^2 z, q^3 z^{\frac{3}{2}}, s_2, y_1, \sigma) + O(q^4), \end{aligned} \tag{7.42}$$

where

$$G_6(q^2 z, q^3 z^{\frac{3}{2}}, s_2, y_1, \sigma) := G_5(s_2,\sigma) + q^3 z^{\frac{3}{2}} H_{1,4}(s_2,\sigma) - s_2^{m_1-2} G_3(s_2,\sigma) y_1 + q^2 z G_{1,2}(s_2,\sigma).$$

Notice that $|G_6| \sim 1$, since $|G_5| \sim 1$. We thus see that the critical point is of the form

$$s_2^c = \Big(y_2 - q^2 z G_{1,1}(\sigma) - q^3 z^{\frac{3}{2}} H_{1,3}(\sigma) + O(q^4)\Big) H_2(q^2 z, q^3 z^{\frac{3}{2}}, y_1, s_2, \sigma),$$

where H_2 is smooth and can be assumed to be very close to $1/(2G_6)$, so that in particular it is of size $|H_2| \sim 1$. It will become important later to observe that, since $|s_2^c| \ll 1, q \ll 1$, then also necessarily $|y_2| \ll 1$.

Plugging this into (7.42), we find that after applying the method of stationary phase to the integration in s_2, the new phase is of the form

$$\begin{aligned} \Phi_3(z,y,\sigma) &:= \Phi_2(z,s_2^c,y,\sigma) = q^3 z^{\frac{3}{2}} H_{1,2}(\sigma) - q^2 z y_1 - y_3 \\ &\quad + \Big(y_2 - q^2 z G_{1,1}(\sigma) - q^3 z^{\frac{3}{2}} H_{1,3}(\sigma)\Big)^2 H_3(q^2 z, q^3 z^{\frac{3}{2}}, y_1, \sigma) + O(q^4), \end{aligned}$$

where $H_3 = H_2(H_2 G_6 - 1)$ is smooth and very close to $-H_2/2$, so that $|H_3| \sim 1$. Decomposing also

$$H_3(q^2 z, q^3 z^{\frac{3}{2}}, y_1, \sigma) = H_{3,1}(y_1,\sigma) + q^2 z H_{3,2}(y_1,\sigma) + q^3 z^{\frac{3}{2}} H_{3,3}(y_1,\sigma) + O(q^4),$$

we see that

$$\begin{aligned} \Phi_3(z,y,\sigma) &= -(y_3 - y_2^2 H_{3,1}(y_1,\sigma)) \\ &\quad + q^2 z\Big(-y_1 + y_2^2 H_{3,2}(y_1,\sigma) - 2y_2 G_{1,1}(\sigma) H_{3,1}(y_1,\sigma)\Big) + q^3 z^{\frac{3}{2}} H_4(y_1,y_2,\sigma) + O(q^4), \end{aligned}$$

where

$$H_4(y_1,y_2,\sigma) := H_{1,2}(\sigma) - 2y_2 H_{1,3}(\sigma) H_{3,1}(y_1,\sigma) + y_2^2 H_{3,3}(y_1,\sigma).$$

Since $|y_2| \ll 1$, we see that also $|H_4| \sim 1$. Recall also that $|G_{1,1}| \ll 1$ is small. By the implicit function theorem, we may therefore write

$$-y_1 + y_2^2 H_{3,2}(y_1,\sigma) - 2y_2 G_{1,1}(\sigma) H_{3,1}(y_1,\sigma) = (y_1 - \varphi(y_2,\sigma)) H_5(y_1,y_2,\sigma),$$

with $|H_5| \sim 1$. We also put $\psi(y_1,y_2,\sigma) := y_2^2 H_{3,1}(y_1,\sigma)$. Then the phase takes on the form

$$\begin{aligned} \Phi_3(z,y,\sigma) &= -(y_3 - \psi(y_1,y_2,\sigma)) + q^2 z(y_1 - \varphi(y_2,\sigma)) H_5(y_1,y_2,\sigma) \\ &\quad + q^3 z^{\frac{3}{2}} H_4(y_1,y_2,\sigma) + O(q^4), \end{aligned} \tag{7.43}$$

where $|H_4| \sim 1$ and $|H_5| \sim 1$, and

$$\mu_{k,1}^{\lambda}(y+\Gamma(\sigma)) = \lambda 2^{\frac{k}{2}} \int e^{-i\lambda s_3 \Phi_3(z,y,\sigma)} \chi_1(z)\, \chi_0(s_3 s')\chi_1(s_3) \, a\big(z,q,s_2,\sigma\big)\, ds_3 dz, \tag{7.44}$$

where we recall that $q = (2^k\lambda^{-1})^{\frac{1}{3}}$.

Choose again a small positive number $\delta > 0$. If $q^2|y_1 - \varphi(y_2,\sigma)| > q^{3-\delta}$, then we can repeatedly integrate by parts in z to see that $\mu_{k,1}^{\lambda}(y+\Gamma(\sigma)) = O((\lambda 2^k)^{-N})$ for every N, since $\lambda q^{3-\delta} = 2^{k(1-\delta/3)}\lambda^{\delta/3}$. Similarly, if $q^2|y_1 - \varphi(y_2,\sigma)| \le q^{3-\delta}$ and $|y_3 - \psi(y_1,y_2,\sigma)| > q^{3-2\delta}$, then the term $-(y_3 - \psi(y_1,y_2,\sigma))$ becomes dominant in Φ_3, so that we can use integrations by parts in s_3 to arrive at the same kind of estimate.

Thus, the contributions of these regions to the maximal operator are well under control, and we are left with the control of the contribution by the region where $|y_3 - \psi(y_1,y_2,\sigma)| \le q^{3-2\delta}$ and $|y_1 - \varphi(y_2,\sigma)| \le q^{1-\delta}$.

To this end, observe that $|\partial_z^2 \Phi_3(z,y,\sigma)| \gtrsim q^3$. We may thus apply van der Corput's lemma to see that in this region, we have that

$$|\mu_{k,1}^{\lambda}(y+\Gamma(\sigma))| \lesssim \lambda 2^{\frac{k}{2}} 2^{-\frac{k}{2}} = \lambda$$

(note here that $\lambda q^3 = 2^k$).

Hence, by Proposition 4.2 we obtain similarly as before that for every $\delta > 0$,

$$\|\mathcal{M}_{k,1}^{\lambda}\|_{L^{1+\varepsilon}\mapsto L^{1+\varepsilon}} \le C_\varepsilon T \lambda^{\frac{2}{3}+\delta} 2^{k(\frac{1}{3}-\delta)}, \tag{7.45}$$

uniformly in k and σ. Combining the estimates (7.32) and (7.38) we find that

$$\|\mathcal{M}_k^{\lambda}\|_{L^2\mapsto L^2} \lesssim T^{\frac{1}{2}} 2^{-\frac{k}{6}} \lambda^{-\frac{1}{3}},$$

and similar from (7.37) and (7.45) we obtain that for every $\varepsilon > 0$ and $\delta > 0$ sufficiently small,

$$\|\mathcal{M}_k^{\lambda}\|_{L^{1+\varepsilon}\mapsto L^{1+\varepsilon}} \le C_\varepsilon T \lambda^{\frac{2}{3}+\delta} 2^{k(\frac{1}{3}-\delta)},$$

Interpolating these estimates leads to the estimate (7.18), which concludes the proof of Lemma 7.4, hence also that of Theorem 7.1.

8. Estimation of the maximal operator $\mathcal{M}$ when surface has a D_4^+ - type singularity

We assume now that ϕ has a singularity of type D_4^+, so that we may assume that its principal part has the form

$$\phi_{\rm pr}(x_1,x_2) = x_1x_2^2 + x_1^3$$

(compare Remark 3.3). The associated principal weight is given by $\kappa = (1/3, 1/3)$, and the corresponding dilations are given by $\delta_r(x_1,x_2) = (r^{\frac{1}{3}}x_1, r^{\frac{1}{3}}x_2),\ r>0$.

Following our discussion in Section 5, by means of a dyadic decomposition, we may reduce ourselves to the estimation of the maximal operators $\mathcal{M}_k$, given by

$$\mathcal{M}_k f(y,y_3) := \sup_{t>0} |f * (\mu_k)_t(y,y_3)|,$$

where the re-scaled measures μ_k are defined by

$$\int f\, d\mu_k := \int f(x, 2^k + \phi^k(x))\, \eta(\delta_{2^{-k}}x)\chi_1(x) dx.$$

Here we have set

$$\phi^k(x) := 2^k\phi(\delta_{2^{-k}}(x)) = \phi_{\rm pr}(x) + 2^k\phi_r(\delta_{2^{-k}}(x)).$$

Recall that the perturbation term $2^k\phi_r(\delta_{2^{-k}}(\cdot))$ is of order $O(2^{-\varepsilon k})$ for some $\varepsilon > 0$. By inequality (5.1), we then have

$$\|\mathcal{M}\|_{L^p\to L^p} \le \sum_{k=k_0}^{\infty} 2^{-\frac{2k}{3}} \|\mathcal{M}_k\|_{L^p\to L^p}, \tag{8.1}$$

where we may assume that k_0 is sufficiently large. We shall prove that if $p > h = 1/|\kappa| = 3/2$, then for every $\delta > 0$ and k sufficiently large k,

$$\|\mathcal{M}_k f\|_p \le C_\delta 2^{k(\frac{1}{p}+\delta)} \|f\|_p. \tag{8.2}$$

In combination with (8.1), this will imply that the maximal operator $\mathcal{M}$ is indeed L^p-bounded for every $p > 3/2$.

To this end, recall that the Hessian determinant of $\phi_{\rm pr}$ is given by $\mathrm{Hess}(\phi_{\rm pr})(x_1, x_2) = 12x_1^2 - 4x_2^2$. In order to prove (8.2), it is sufficient to show that, given any point x^0 in the support of χ_1 (which is contained in an annulus on which $|x| \sim 1$), there exist a smooth bump function $\alpha \ge 0$ supported on a sufficiently small neighborhood of x^0 with $\alpha(x^0) = 1$ so that the corresponding localized maximal operator $\mathcal{M}_k^\alpha$ satisfies the estimate (6.19).

If $\mathrm{Hess}(\phi_{\rm pr})(x^0) \neq 0$, then this has already been shown in Section 5 (compare estimate (5.4)). We shall therefore assume that $\mathrm{Hess}(\phi_{\rm pr})(x^0) = 0$, i.e.,

$$12(x_1^0)^2 - 4(x_2^0)^2 = 0.$$

But, since $|x^0| \sim 1$, this implies $x_1^0 \neq 0$ and $x_2^0 \neq 0$. By homogeneity, we may then even assume that, say, $x_1^0 = 1$, so that $x_2^0 = \sqrt{3}$. Then first we shift the point $x^0 = (1, \sqrt{3})$ to the origin in $\mathbb{R}^2$ by changing to the linear coordinates $z = (z_1, z_2)$ defined by $x_1 = 1 + z_1$, $x_2 = \sqrt{3} + z_2$. In these coordinates, the principal part of ϕ is given by

$$\phi_{\rm pr}(1 + z_1, \sqrt{3} + z_2) = 4 + 6z_1 + 2\sqrt{3}z_2 + (\sqrt{3}z_1 + z_2)^2 + z_1^3 + z_1 z_2^2.$$

Thus, by a linear change of variables in the ambient space the principal part can be reduced to the form

$$\widetilde{\phi_{\rm pr}}(y_1, y_2) := \phi_{\rm pr}(y_1, y_2 - \sqrt{3}y_1) = y_2^2 + y_1^3 + y_1(y_2 - \sqrt{3}y_1)^2 = y_2^2 + 4y_1^3 + y_1 y_2^2 - 2\sqrt{3}y_1^2 y_2.$$

Note that the principal part of the function $\widetilde{\phi_{\rm pr}}$ is $y_2^2 + 4y_1^3$. Thus, in these coordinates (y_1, y_2) near $(0, 0)$ (which corresponds to our original point x^0), $\tilde{\phi}$ has a singularity of type A_2 and depends smoothly on the small parameter $\sigma := 2^{-k/3}$, and consequently the estimate (8.2) for $\mathcal{M}_k^\alpha$ in place of $\mathcal{M}_k$ follows from Theorem 7.1.

This completes the proof of Theorem 1.3.

Bibliography

[AGV88] Arnol'd, V.I., Gusein-Zade, S.M. and Varchenko, A.N., Singularities of differentiable maps. Vol. II, Monodromy and asymptotics of integrals, *Monographs in Mathematics, 83.* Birkhäuser, Boston Inc., Boston, MA, 1988.

[Bou85] Bourgain, J., Estimations de certaines fonctions maximales. *C. R. Acad. Sci. Paris Sér. I Math.*, 301 (1985) no. 10, 499–502.

[Dui74] Duistermaat, J. J., Oscillatory integrals, Lagrange immersions and unfolding of singularities. *Comm. Pure Appl. Math.*, 27 (1974), 207–281.

[G09] Grafakos, L., Modern Fourier analysis. *Graduate Texts in Mathematics* 250. Springer, New York, 2009.

[Gr13] Greenblatt, M., L^p boundedness of maximal averages over hypersurfaces in $\mathbb{R}^3$. *Trans. Amer. Math. Soc.*, 365, (2013), 1875–1900.

[Gl81] Greenleaf, A., Principal curvature and harmonic analysis. *Indiana Univ. Math. J.*, 30(4) (1981), 519–537.

[IM11] Ikromov, I. A., Müller, D., On adapted coordinate systems. *Trans. Amer. Math. Soc.,* 363 (2011), no. 6, 2821–2848.

[IKM10] Ikromov, I. A., Kempe, M., Müller, D., *Estimates for maximal functions associated to hypersurfaces in* $\mathbb{R}^3$ *and related problems of harmonic analysis.* Acta Math. 204 (2010), 151–271.

[IM16] Ikromov, I. A., Müller, D., Fourier restriction for hypersurfaces in three dimensions and Newton polyhedra; *Annals of Mathematics Studies 194*, Princeton University Press, Princeton and Oxford 2016; 260 pp.

[ISa96] Iosevich, A. and Sawyer, E., Oscillatory integrals and maximal averages over homogeneous surfaces. *Duke Math. J.*, 82 (1996), no. 1, 103–141.

[ISaSe99] Iosevich, A., Sawyer, E. and Seeger, A., On averaging operators associated with convex hypersurfaces of finite type. *J. Anal. Math.*, 79 (1999), 159–187.

[ISa97] Iosevich, A., Sawyer, E., Maximal averages over surfaces. *Adv. Math.*, 132 (1997), 46–119.

[NSeW93] Nagel, A., Seeger, A. and Wainger, S., Averages over convex hypersurfaces. *Amer. J. Math.*, 115 (1993), no. 4, 903–927.

[PS97] Phong, D.H. and Stein, E.M., The Newton polyhedron and oscillatory integral operators. *Acta Math.*, 179 (1997), no. 1, 105–152.

[PSS99] Phong, D. H., Stein, E. M., Sturm, J. A., On the growth and stability of real-analytic functions. *Amer. J. Math.*, 121 (1999), no. 3, 519-554.

[S76] Stein, E.M., Maximal functions. I. Spherical means. *Proc. Nat. Acad. Sci. U.S.A.*, 73 (1976), no. 7, 2174–2175.

[S93] Stein, E. M., Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals. *Princeton Mathematical Series* 43. Princeton University Press, Princeton, NJ, 1993.

[SW69] Stein, E.M., Weiss, N. J., On the Convergence of Poisson Integrals. *Trans. Amer. Math. Soc.*, (1969), no. 140, 35–54.

[V76] Varchenko, A. N., Newton polyhedra and estimates of oscillating integrals. *Funkcional. Anal. i Priložen*, 10 (1976), 13–38 (Russian); English translation in *Funktional Anal. Appl.*, 18 (1976), 175–196.

[Z5] Zimmermann, E., On L^p-estimates for maximal averages over hypersurfaces not satisfying the transversality condition. *Doctoral thesis, Kiel 2014.*

Mathematisches Seminar, C.A.-Universität Kiel, Ludewig-Meyn-Strasse 4, D-24118 Kiel, Germany
E-mail address: buschenhenke@math.uni-kiel.de
URL: http://analysis.math.uni-kiel.de/buschenhenke/

School of Mathematical Sciences, University College Cork, Western gateway Building, Western Road, Cork, Ireland
E-mail address: sd@ucc.ie
URL: http://publish.ucc.ie/researchprofiles/D019/sd

Department of Mathematics, Samarkand State University, University Boulevard 15, 140104, Samarkand, Uzbekistan
E-mail address: ikromov1@rambler.ru

Mathematisches Seminar, C.A.-Universität Kiel, Ludewig-Meyn-Strasse 4, D-24118 Kiel, Germany
E-mail address: mueller@math.uni-kiel.de
URL: http://analysis.math.uni-kiel.de/mueller/