SHARP $L^p$-ESTIMATES FOR MAXIMAL OPERATORS ASSOCIATED TO HYPERSURFACES IN $\mathbb{R}^3$ FOR $p > 2$.

DETLEF MÜLLER, ISROIL A.IKROMOV, AND MICHAEL KEMPE

Abstract. We study the boundedness problem for maximal operators $\mathcal{M}$ associated to smooth hypersurfaces $S$ in 3-dimensional Euclidean space. For $p > 2$, we prove that if no affine tangent plane to $S$ passes through the origin and $S$ is analytic, then the associated maximal operator is bounded on $L^p(\mathbb{R}^3)$ if and only if $p > h(S)$, where $h(S)$ denotes the so-called height of the surface $S$. For non-analytic finite type $S$ we obtain the same statement with the exception of the exponent $p = h(S)$. Our notion of height $h(S)$ is closely related to A. N. Varchenko’s notion of height $h(\phi)$ for functions $\phi$ such that $S$ can be locally represented as the graph of $\phi$ after a rotation of coordinates.

Several consequences of this result are discussed. In particular we verify a conjecture by E. M. Stein and its generalization by A. Iosevich and E. Sawyer on the connection between the decay rate of the Fourier transform of the surface measure on $S$ and the $L^p$-boundedness of the associated maximal operator $\mathcal{M}$, and a conjecture by Iosevich and Sawyer which relates the $L^p$-boundedness of $\mathcal{M}$ to an integrability condition on $S$ for the distance function to tangential hyperplanes, in dimension three.

In particular, we also give essentially sharp uniform estimates for the Fourier transform of the surface measure on $S$, thus extending a result by V. N. Karupshkin from the analytic to the smooth setting and implicitly verifying a conjecture by V. I. Arnol’d in our context.

Contents

1. Introduction 2
1.1. Outline of the proof of Theorem 1.2 and organization of the article 10
2. Newton diagrams and adapted coordinates 11
2.1. The principal part of $\phi$ associated to a supporting line of the Newton polyhedron as a mixed homogeneous polynomial 12
3. Uniform estimates for maximal operators associated to families of finite type curves and related surfaces 13
3.1. Finite type curves 13
3.2. Related results for families of surfaces 24
4. Auxiliary statements on the multiplicity of roots at a critical point of a mixed homogeneous polynomial function 28

2000 Mathematical Subject Classification. 35D05 35D10 35G05.
Key words and phrases. Maximal operator, Hypersurface, Oscillatory integral, Newton diagram, Oscillation index, Contact index.

We acknowledge the support for this work by the Deutsche Forschungsgemeinschaft.
1. Introduction

Let $S$ be a smooth hypersurface in $\mathbb{R}^n$ and let $\rho \in C^\infty_0(S)$ 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

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

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 \(p > n/(n-1)\).

In 1976, E. M. Stein [27] 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 proved by J. Bourgain [3]. These results became the starting point for intensive studies of various classes of maximal operators associated to subvarieties. Stein’s monography [28] is an excellent reference to many of these developments. From these early works, the influence of geometric properties on the validity of \(L^p\)-estimates of the maximal operator \(\mathcal{M}\) became evident. For instance, A. Greenleaf [9] proved that \(\mathcal{M}\) is bounded on \(L^p(\mathbb{R}^n)\) if \(n \geq 3\) and \(p > \frac{n}{n-1}\), provided \(S\) has everywhere non-vanishing Gaussian curvature and in addition \(S\) is starshaped with respect to the origin.

In contrast, the case where the Gaussian curvature vanishes at some points is still wide open, with the exception of the two-dimensional case \(n = 2\), i.e., the case of finite type curves in \(\mathbb{R}^2\) studied by A. Iosevich in [13]. As a partial result in higher dimensions, C. D. Sogge and E. M. Stein showed in [24] that if the Gaussian curvature of \(S\) does not vanish to infinite order at any point of \(S\), then \(\mathcal{M}\) is bounded on \(L^p(\mathbb{R}^n)\) in a certain range \(p > p(S)\). However, the exponent \(p(S)\) given in that article is in general far from being optimal, and in dimensions \(n \geq 3\), sharp results are known only for particular classes of hypersurfaces.

The perhaps best understood class in higher dimensions is the class of convex hypersurfaces of finite line type (see in particular the early work in this setting by M. Cowling and G. Mauceri in [6], [5], the work by A. Nagel, A. Seeger and S. Wainger in [20], and the articles [14], [15] and [16] by A. Iosevich, E. Sawyer and A. Seeger). In [20], sharp results were for instance obtained for convex hypersurfaces which are given as the graph of a mixed homogeneous convex function \(\phi\). Further results were based on a result due to Schulz [23](see also [31]), which states that, possibly after a rotation of coordinates, any smooth convex function \(\phi\) of finite line type can be written in the form \(\phi = Q + \phi_r\), where \(Q\) is a convex mixed homogeneous polynomial that vanishes only at the origin, and \(\phi_r\) is a remainder term consisting of terms of higher homogeneous degree than the polynomial \(Q\). By means of this result, Iosevich and Sawyer proved in [15] sharp \(L^p\)-estimates for the maximal operator \(\mathcal{M}\) for \(p > 2\). For further results in the case \(p \leq 2\), see also [28].

As is well-known since the early work of E. M. Stein on the spherical maximal operator, the estimates of the maximal operator \(\mathcal{M}\) on Lebesgue spaces are intimately
connected with the decay rate of the Fourier transform
\[ \hat{\rho \, d\sigma}(\xi) = \int_{S} e^{-ix \cdot x} \rho(x) \, d\sigma(x), \quad \xi \in \mathbb{R}^n, \]
of the superficial measure \( \rho \, d\sigma \), i.e., to estimates of oscillatory integrals. These in return are closely related to geometric properties of the surface \( S \), and have been considered by numerous authors ever since the early work by B. Riemann on this subject (see [28] for further information). Also the afore mentioned results for convex hypersurfaces of finite line type are based on such estimates. Indeed, sharp estimates for the Fourier transform of superficial measures on \( S \) have been obtained by J. Bruna, A. Nagel and S. Wainger in [4], improving on previous results by B. Randol [22] and I. Svensson [29]. They introduced a family of nonisotropic balls on \( S \), called ”caps”, by setting
\[ B(x, \delta) := \{ y \in S : \text{dist}(y, x + T_x S) < \delta \}, \quad \delta > 0. \]
Here \( T_x S \) denotes the tangent space to \( S \) at \( x \in S \). Suppose that \( \xi \) is normal to \( S \) at the point \( x_0 \). Then it was shown that
\[ |\hat{\rho \, d\sigma}(\xi)| \leq C |B(x_0, |\xi|^{-1})|, \]
where \( |B(x_0, \delta)| \) denotes the surface area of \( B(x_0, \delta) \). These estimate became fundamental also in the subsequent work on associated maximal operators.

However, such estimates fail to be true for non-convex hypersurfaces, which we shall be dealing with in this article. More precisely, we shall consider general smooth hypersurfaces in \( \mathbb{R}^3 \).

Assume that \( S \subset \mathbb{R}^3 \) is such a hypersurface, and let \( x^0 \in S \) be a fixed point in \( S \). We can then find a Euclidean motion of \( \mathbb{R}^3 \), so that in the new coordinates given by this motion, we can assume that \( x^0 = (0, 0, 1) \) and \( T_{x^0} = \{ x_3 = 0 \} \). Then, in a neighborhood \( U \) of the origin, the hypersurface \( S \) 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 \( \Omega \) of \( 0 \in \mathbb{R}^2 \) and satisfying the conditions
\[ \phi(0, 0) = 0, \quad \nabla \phi(0, 0) = 0. \tag{1.2} \]

To \( \phi \) we can then associate the so-called height \( h(\phi) \) in the sense of A. N. Varchenko [30] defined in terms of the Newton polyhedra of \( \phi \) when represented in smooth coordinate systems near the origin (see Section 2 for details). An important property of this height is that it is invariant under local smooth changes of coordinates fixing the origin. We then define the height of \( S \) at the point \( x^0 \) by \( h(x^0, S) := h(\phi) \). This notion can easily be seen to be invariant under affine linear changes of coordinates in the ambient space \( \mathbb{R}^3 \) (cf. Section 11) because of the invariance property of \( h(\phi) \) under local coordinate changes.

Now observe that unlike linear transformations, translations do not commute with dilations, which is why Euclidean motions are no admissible coordinate changes for the
study of the maximal operators \( \mathcal{M} \). We shall therefore study \( \mathcal{M} \) under the following transversality assumption on \( S \).

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

Notice that this 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 \( \phi \) as before, and that the norm of \( \mathcal{M} \) when acting on \( L^p(\mathbb{R}^3) \) is invariant under such a linear change of coordinates.

If \( \phi \) is flat, i.e., if all derivatives of \( \phi \) vanish at the origin, and if \( \rho(x^0) > 0 \), then it is well-known and easy to see that the maximal operator \( \mathcal{M} \) is \( L^p \)-bounded if and only if \( p = \infty \), so that this case is of no interest. Let us therefore always assume in the sequel that \( \phi \) is non-flat, i.e., of finite type. Correspondingly, we shall always assume without further mentioning that the hypersurface \( S \) is of finite type in the sense that every tangent plane has finite order of contact.

We can now state the main result of this article.

**Theorem 1.2.** Assume that \( S \) is a smooth hypersurface in \( \mathbb{R}^3 \) satisfying Assumption 1.1, and let \( x^0 \in S \) be a fixed point. Then there exists a neighborhood \( U \subset S \) of the point \( x^0 \) such that for any \( \rho \in C^\infty_0(U) \) the associated maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) whenever \( p > \max\{h(x^0, S), 2\} \).

Notice that even in the case where \( S \) is convex this result is stronger than the known results, which always assumed that \( S \) is of finite line type.

The following Theorem shows the sharpness of this theorem.

**Theorem 1.3.** Assume that the maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) for some \( p > 1 \), where \( S \) satisfies Assumption 1.1. Then, for any point \( x^0 \in S \) with \( \rho(x^0) > 0 \), we have \( h(x^0, S) \leq p \). Moreover, if \( S \) is analytic at such a point \( x^0 \), then \( h(x^0, S) < p \).

As an immediate consequence of these two results, we obtain

**Corollary 1.4.** Suppose \( S \) is a smooth hypersurface in \( \mathbb{R}^3 \) satisfying Assumption 1.1, and let \( x^0 \in S \) be a fixed point. Then there exists a neighborhood \( U \subset S \) of this point such that \( h(x, S) \leq h(x^0, S) \) for every \( x \in U \).

This shows in particular that if \( \Phi(x, s) = \phi(x_1, x_2) + s_1 x_1 + s_2 x_2 \) is a smooth deformation by linear terms of a smooth, finite type function \( \phi \) defined near the origin in \( \mathbb{R}^2 \) and satisfying (1.2), then the height of \( \Phi(\cdot, s) \) at any critical point of the function \( x \mapsto \Phi(x, s) \) is bounded by the height at \( h(\Phi(\cdot, 0)) = h(\phi) \) for sufficiently small perturbation parameters \( s_1 \) and \( s_2 \). This proves a conjecture by V.I. Arnol’d [2] in the smooth setting at least for linear perturbations. For analytic functions \( \phi \) of two variables, such a result has been proved for arbitrary analytic deformations by V. N. Karpushkin [17].
From these results, global results can be deduced easily. For instance, if $S$ is a compact hypersurface, then we define the height $h(S)$ of $S$ by $h(S) := \sup_{x \in S} h(x, S)$. Corollary 1.4 shows that in fact

$$h(S) := \max_{x \in S} h(x, S) < \infty,$$

and from Theorems 1.2, 1.3 we obtain

**Corollary 1.5.** Assume that $S$ is a smooth, compact hypersurface in $\mathbb{R}^3$ satisfying Assumption 1.1, that $\rho > 0$ on $S$ and that $p > 2$.

If $S$ is analytic, then the associated maximal operator $M$ is bounded on $L^p(\mathbb{R}^3)$ if and only if $p > h(S)$. If $S$ is only assumed to be smooth, then for $p \neq h(S)$ we still have that the maximal operator $M$ is bounded on $L^p(\mathbb{R}^3)$ if and only if $p > h(S)$.

Let $H$ be an affine hyperplane in $\mathbb{R}^3$. Following A. Iosevich and E. Sawyer [14], we consider the distance $d_H(x) := \text{dist} (H, x)$ from $x \in S$ to $H$. In particular, if $x^0 \in S$, then $d_{T,x^0}(x) := \text{dist} (x^0 + T_{x^0}S, x)$ will denote the distance from $x \in S$ to the affine tangent plane to $S$ at the point $x^0$. The following result has been proved in [14] in arbitrary dimensions $n \geq 2$ and without requiring Assumption 1.1.

**Theorem 1.6 (Iosevich-Sawyer).** If the maximal operator $M$ is bounded on $L^p(\mathbb{R}^n)$, where $p > 1$, then

$$\int_S d_H(x)^{-1/p} \rho(x) \, d\sigma(x) < \infty$$

for every affine hyperplane $H$ in $\mathbb{R}^n$ which does not pass through the origin.

Moreover, they conjectured that for $p > 2$ the condition (1.3) is indeed necessary and sufficient for the boundedness of the maximal operator $M$ on $L^p$, at least if for instance $S$ is compact and $\rho > 0$.

**Remark 1.7.** Notice that condition (1.3) is easily seen to be true for every affine hyperplane $H$ which is nowhere tangential to $S$, so that it is in fact a condition on affine tangent hyperplanes to $S$ only. Moreover, if Assumption 1.1 is satisfied, then there are no affine tangent hyperplanes which pass through the origin, so that in this case it is a condition on all affine tangent hyperplanes.

In Section 11, we shall prove

**Proposition 1.8.** Suppose $S$ is a smooth hypersurface in $\mathbb{R}^3$, and let $x^0 \in S$ be a fixed point. Then, for every $p < h(x^0, S)$, we have

$$\int_{S \cap U} d_{T,x^0}(x)^{-1/p} \, d\sigma(x) = \infty$$

for every neighborhood $U$ of $x^0$. Moreover, if $S$ is analytic near $x^0$, then (1.4) holds true also for $p = h(x^0, S)$. 
Notice that this result does not require Assumption 1.1.

As an immediate consequence of Theorem 1.2, Theorem 1.6 and Proposition 1.8 we obtain

**Corollary 1.9.** Assume that \( S \subset \mathbb{R}^3 \) satisfies Assumption 1.1, and let \( x^0 \in S \) be a fixed point. Moreover, let \( p > 2 \).

Then, if \( S \) is analytic near \( x^0 \), there exists a neighborhood \( U \subset S \) of the point \( x^0 \) such that for any \( p \in C_0^\infty(U) \) with \( \rho(x^0) > 0 \) the associated maximal operator \( M \) is bounded on \( L^p(\mathbb{R}^3) \) if and only if condition (1.3) holds for every affine hyperplane \( H \) in \( \mathbb{R}^3 \) which does not pass through the origin.

If \( S \) is only assumed to be smooth near \( x^0 \), then the same conclusion holds true, with the possible exception of the exponent \( p = h(x^0, S) \).

This confirms the conjecture by Iosevich and Sawyer in our setting for analytic \( S \), and for smooth \( S \) with the possible exception of the exponent \( p = h(x^0, S) \). For the critical exponent \( p = h(x^0, S) \), if \( S \) is not analytic near \( x^0 \), examples show that unlike in the analytic case it may happen that \( M \) is bounded on \( L^{h(x^0, S)}(\mathbb{R}^3) \) (see, e.g., [15]), and the conjecture remains open for this value of \( p \). For further details, we refer to Section 11.

As mentioned before, the estimates of the maximal operator \( M \) on Lebesgue spaces are intimately connected with the decay rate of the Fourier transform

\[
\widehat{\rho d\sigma}(\xi) = \int_S e^{-i\xi \cdot x} \rho(x) d\sigma(x), \quad \xi \in \mathbb{R}^n,
\]

of the superficial measure \( \rho d\sigma \). Estimates of such oscillatory integrals will naturally play a central role also in our proof Theorem 1.2. Indeed our proof of Theorem 1.2 will provide enough information that it will also be easy to derive from it the following uniform estimate for the Fourier transform of surface carried measures on \( S \).

**Theorem 1.10.** Let \( S \) be a smooth hypersurface of finite type in \( \mathbb{R}^3 \) and let \( x^0 \) be a fixed point in \( S \). Then there exists a neighborhood \( U \subset S \) of the point \( x^0 \) such that for every \( \rho \in C_0^\infty(U) \) the following estimate holds true:

\[
(1.5) \quad |\widehat{\rho d\sigma}(\xi)| \leq C ||\rho||_{C^3(S)} \log(2 + |\xi|)(1 + |\xi|)^{-1/h(x^0, S)} \quad \text{for every } \xi \in \mathbb{R}^3.
\]

This estimate generalizes Karpushkin’s estimates in [17] from the analytic to the finite type setting, at least for linear perturbations.

The next result establishes a direct link between the decay rate of \( \widehat{\rho d\sigma}(\xi) \) and Iosevich-Sawyer’s condition (1.3). In combination with Proposition 1.8 it shows in particular that the exponent \( -1/h(x^0, S) \) in estimate (1.5) is sharp (for the case of analytic hypersurfaces, the latter follows also from Varchenko’s asymptotic expansions of oscillatory integrals in [30]).
Theorem 1.11. Let $S$ be a smooth hypersurface in $\mathbb{R}^n$, and let $\rho \in C^\infty_0(S)$ be a smooth cut-off function $\rho \geq 0$, and assume that

$$|\rho d\sigma (\xi)| \leq C_\beta (1 + |\xi|)^{-\beta} \text{ for every } \xi \in \mathbb{R}^n,$$

for some $\beta > 0$. Then for every $p > 1$ such that $p > 1/\beta$,  

$$\int_S d_H(x)^{-1/p} \rho(x) d\sigma(x) < \infty,$$

for every affine hyperplane $H$ in $\mathbb{R}^n$.

In combination with Proposition 1.8 this result easily implies (see Section 11)

Corollary 1.12. Suppose $S$ is a smooth hypersurface in $\mathbb{R}^3$, let $x^0 \in S$ be a fixed point and assume that the estimate (1.6) holds true for some $\beta > 0$. If $\rho(x^0) > 0$, and if $\rho$ is supported in a sufficiently small neighborhood of $x^0$, then necessarily $\beta \leq 1/h(x^0, S)$.

Indeed, more is true. Let us introduce the following quantities. In analogy with V. I. Arnol’d’s notion of the ”singularity index” [2], we define the uniform oscillation index $\beta_u(x^0, S)$ of the hypersurface $S \subset \mathbb{R}^n$ at the point $x^0 \in S$ as follows:

Let $\mathfrak{B}_u(x^0, S)$ denote the set of all $\beta \geq 0$ for which there exists an open neighborhood $U_\beta$ of $x^0$ in $S$ such that estimate (1.6) holds true for every function $\rho \in C^\infty_0(U_\beta)$. Then

$$\beta_u(x^0, S) := \sup \{ \beta : \beta \in \mathfrak{B}_u(x^0, S) \}.$$

If we restrict our attention to the normal direction to $S$ at $x^0$ only, then we can define analogously the notion of oscillation index of the hypersurface $S$ at the point $x^0 \in S$. More precisely, if $n(x^0)$ is a unit normal to $S$ at $x^0$, then we let $\mathfrak{B}(x^0, S)$ denote the set of all $\beta \geq 0$ for which there exists an open neighborhood $U_\beta$ of $x^0$ in $S$ such that estimate (1.6) holds true for every function $\rho \in C^\infty_0(U_\beta)$, i.e.,

$$|\rho \bar{d}\sigma (\lambda n(x^0))| \leq C_\beta (1 + |\lambda|)^{-\gamma} \text{ for every } \lambda \in \mathbb{R}.$$

Then

$$\beta(x^0, S) := \sup \{ \beta : \beta \in \mathfrak{B}(x^0, S) \}.$$

If we regard $S$ locally as the graph of a function $\phi$, then we can introduce related notions $\beta_u(\phi)$ and $\beta(\phi)$ for $\phi$, regarded as the phase function of an oscillatory integral (cf. [11], and also Section 11).

We also define the uniform contact index $\gamma_u(x^0, S)$ of the hypersurface $S$ at the point $x^0 \in S$ as follows:

Let $\mathfrak{C}_u(x^0, S)$ denote the set of all $\gamma \geq 0$ for which there exists an open neighborhood $U_\gamma$ of $x^0$ in $S$ such that the estimate

$$\int_{U_\gamma} d_H(x)^{-\gamma} d\sigma(x) < \infty$$

holds true for every affine hyperplane $H$ in $\mathbb{R}^n$. Then we put

$$\gamma_u(x^0, S) := \sup \{ \gamma : \gamma \in \mathfrak{C}_u(x^0, S) \}.$$
Similarly, we let \( C(x^0, S) \) denote the set of all \( \gamma \geq 0 \) for which there exists an open neighborhood \( U_\gamma \) of \( x^0 \) in \( S \) such
\[
(1.10) \quad \int_{U_\gamma} d_{T,x^0}(x)^{-\gamma} \, d\sigma(x) < \infty,
\]
and call
\[
\gamma(x^0, S) := \sup\{ \gamma : \gamma \in C(x^0, S) \}
\]
the contact index \( \gamma(x^0, S) \) of the hypersurface \( S \) at the point \( x^0 \in S \). Then clearly
\[
(1.11) \quad \beta_u(x^0, S) \leq \beta(x^0, S), \quad \gamma_u(x^0, S) \leq \gamma(x^0, S).
\]

At least for hypersurfaces in \( \mathbb{R}^3 \), a lot more is true.

**Theorem 1.13.** Let \( S \) be a smooth, finite type hypersurface in \( \mathbb{R}^3 \), and let \( x^0 \in S \) be a fixed point. Then
\[
\beta_u(x^0, S) = \beta(x^0, S) = \gamma_u(x^0, S) = \gamma(x^0, S) = \frac{1}{h(x^0, S)}.
\]

Let us recall at this point a result by A. Greenleaf. In [9] he proved that if \( \hat{\rho}d\sigma(\xi) = O(|\xi|^{-\beta}) \) as \( |\xi| \to \infty \) and if \( \beta > 1/2 \), then the maximal operator is bounded on \( L^p \) whenever \( p > 1 + \frac{1}{2\beta} \). The case \( \beta \leq 1/2 \) remained open.

For \( \beta = 1/2 \) E. M. Stein and later for the full range \( \beta \leq 1/2 \) A. Iosevich and E. Sawyer [15] conjectured that if \( S \) is a smooth, compact hypersurface in \( \mathbb{R}^n \) such that
\[
|\hat{\rho}d\sigma(\xi)| = O(|\xi|^{-\beta}) \text{ for some } 0 < \beta \leq 1/2,
\]
then the maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^n) \) for every \( p > 1/\beta \), at least if we assume \( \rho > 0 \).

A partial confirmation of Stein’s conjecture has been given by C. D. Sogge [26] who proved that if the surface has at least one non-vanishing principal curvature everywhere, then the maximal operator is \( L^p \)-bounded for every \( p > 2 \). Certainly, if the surface has at least one non-vanishing principal curvature then the estimate above holds for \( \beta = 1/2 \).

Now, if \( n = 3 \), and if \( 0 < \beta \leq 1/2 \), then \( \beta_u(x^0, S) \geq \beta \) for every point \( x^0 \in S \), so that our Theorem 1.13 implies that \( 1/\beta \geq h(x^0, S) \). Then, if \( p > 1/\beta \), we have \( p > \max\{2, h(x^0, S)\} \). Therefore, by means of a partition of unity argument, we obtain from Theorem 1.2 the following confirmation of the Stein-Iosevich-Sawyer conjecture in this case.

**Corollary 1.14.** Let \( S \) be a smooth compact hypersurface in \( \mathbb{R}^3 \) satisfying Assumption 1.1, and let \( \rho > 0 \) be a smooth density on \( S \). We assume that there is some \( 0 < \beta \leq 1/2 \) such that
\[
|\hat{\rho}d\sigma(\xi)| = O(|\xi|^{-\beta}).
\]
Then the associated maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > 1/\beta \).
We finally remark that the case $p \leq 2$ behaves quite differently, and examples show that neither condition (1.3) nor the notation of height will be suitable to determine the range of exponents $p$ for which the maximal operator $\mathcal{M}$ is $L^p$-bounded (see, e.g., [16]). The study of this range for $p \leq 2$ is work in progress.

1.1. Outline of the proof of Theorem 1.2 and organization of the article.

The proof of our main result, Theorem 1.2, will strongly make use of the results in [11] on the existence of a so-called "adapted" coordinate system for a smooth, finite type function $\phi$ defined near the origin in $\mathbb{R}^2$ (see Section 2 for some basic notation). These results generalize the corresponding results for analytic $\phi$ by A. N. Varchenko [30], by means of a simplified approach inspired by the work of Phong and Stein [21]. According to these results, one can always find a change of coordinates of the form

$$ y_1 := x_1, \ y_2 := x_2 - \psi(x_1) $$

which leads to adapted coordinates $y$. The function $\psi$ can be constructed from the Pusieux series expansion of roots of $\phi$ (at least if $\phi$ is analytic) as the so-called principal root jet (cf.[11]). Somewhat simplifying, it agrees with a real-valued leading part of the (complex) root of $\phi$ near which $\phi$ is "small of highest order" in an averaged sense. One would preferably like to work in these adapted coordinates $y$, since the height of $\phi$ when expressed in these adapted coordinates, can be read off directly from the Newton polyhedron of $\phi$ as the so-called "distance." However, this change of coordinates leads to substantial problems, since it is in general non-linear.

Now, away from the curve $x_2 = \psi(x_1)$, it turns out that one can find some $k$ with $2 \leq k \leq h(\phi)$ such that $\partial^k_2 \phi \neq 0$. This suggests that one may apply the results on maximal functions on curves in [13]. Indeed this is possible, but we need estimates for such maximal operators along curves which are stable under small perturbations of the given curve. Such results, which will be based on the local smoothing estimates by G. Mockenhaupt, A. Seeger and C. Sogge in [18], and related estimates for maximal operators along surfaces, are derived in Section 3. The necessary control on partial derivatives $\partial^k_2 \phi$ will be obtained from the study of mixed homogeneous polynomials in Section 4. Indeed, in a similar way as the Schulz polynomial is used in the convex case to approximate the given function $\phi$, we shall approximate the function $\phi$ in domains close to a given root of $\phi$ by a suitable mixed homogeneous polynomial, following here some ideas in [21].

The case where our original coordinates $x$ are adapted or where the height $h(\phi)$ is strictly less than 2 is the simplest one, since we can here avoid non-linear changes of coordinates. This case is dealt with in Section 5.

We then concentrate on the situation where $h(\phi) \geq 2$ and where the coordinates are not adapted. The contributions to the maximal operator $\mathcal{M}$ by a suitable homogeneous domain away from the curve $x_2 = \psi(x_1)$ require a lot more effort and are estimated in Section 6 by means of the results in Sections 3 and 4.

There remains the domain near the curve $x_2 = \psi(x_1)$. For this domain, it is in general no longer possible to reduce its contribution to the maximal operator $\mathcal{M}$ to
maximal operators along curves, and we have to apply two-dimensional oscillatory integral technics. Indeed, we shall need estimates for certain classes of oscillatory integrals with small parameters, which will be given in Section 9. These results will be applied in Sections 7 and 8 in order to complete the proof of Theorem 1.2.

We remark that our proof does not make use of any damping technics, which had been crucial to many other approaches.

The proof of Theorem 1.10, which will be given in Section 10, can easily be obtained from the results established in the course of the proof of Theorem 1.2, except for the case $h(x^0, S) < 2$, which, however, has been studied in a complete way by Duistermaat [8]. The main difference is that we have to replace the estimates for maximal operators in Section 3 by van der Corput type estimates due to J. E. Björk and G. I. Arhipov.

In the last Section 11, we shall give proofs of all the other results stated above.

2. Newton diagrams and adapted coordinates

We recall here some basic notation (compare, e.g., [11] for further information). Let $\phi$ be a smooth real-valued function defined on a neighborhood of the origin in $\mathbb{R}^2$ with $\phi(0, 0) = 0$, $\nabla \phi(0, 0) = 0$, and consider the associated Taylor series

$$\phi(x_1, x_2) \sim \sum_{j,k=0}^{\infty} c_{jk} x_1^j x_2^k$$

of $\phi$ centered at the origin. The set

$$\mathcal{T}(\phi) := \{(j, k) \in \mathbb{N}^2 : c_{jk} = \frac{1}{j!k!} \partial_{x_1}^j \partial_{x_2}^k \phi(0, 0) \neq 0\}$$

will be called the Taylor support of $\phi$ at $(0, 0)$. We shall always assume that

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

i.e., that the function $\phi$ is of finite type at the origin. If $\phi$ is real analytic, so that the Taylor series converges to $\phi$ near the origin, this just means that $\phi \neq 0$. The Newton polyhedron $\mathcal{N}(\phi)$ of $\phi$ at the origin is defined to be the convex hull of the union of all the quadrants $(j, k) + \mathbb{R}_+^2$ in $\mathbb{R}^2$, with $(j, k) \in \mathcal{T}(\phi)$. The associated Newton diagram $\mathcal{N}_d(\phi)$ in the sense of Varchenko [30] 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 distance $d = d(\phi)$ between the Newton polyhedron and the origin in the sense of Varchenko is given by the coordinate $d$ of the point $(d, d)$ at which the bisectrix $t_1 = t_2$ intersects the boundary of the Newton polyhedron.
The principal face $\pi(\phi)$ of the Newton polyhedron of $\phi$ is the face of minimal dimension containing the point $(d, d)$. Deviating from the notation in [30], we shall call the series

$$\phi_p(x_1, x_2) := \sum_{(j,k) \in \pi(\phi)} c_{jk} x_1^j x_2^k$$

the principal part of $\phi$. In case that $\pi(\phi)$ is compact, $\phi_p$ is a mixed homogeneous polynomial; otherwise, we shall consider $\phi_p$ 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 $\phi$ is expressed. By a local analytic (respectively smooth) coordinate system at the origin we shall mean an analytic (respectively smooth) coordinate system defined near the origin which preserves 0. If we work in the category of smooth functions $\phi$, we shall always consider smooth coordinate systems, and if $\phi$ is analytic, then one usually restricts oneself to analytic coordinate systems (even though this will not really be necessary for the questions we are going to study, as we will see). The height of the analytic (respectively smooth) function $\phi$ is defined by

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

where the supremum is taken over all local analytic (respectively smooth) coordinate systems $x$ at the origin, and where $d_x$ is the distance between the Newton polyhedron and the origin in the coordinates $x$.

A given coordinate system $x$ is said to be adapted to $\phi$ if $h(\phi) = d_x$.

2.1. **The principal part of $\phi$ associated to a supporting line of the Newton polyhedron as a mixed homogeneous polynomial.** Let $\kappa = (\kappa_1, \kappa_2)$ with $\kappa_1, \kappa_2 > 0$ be a given weight, with associated one-parameter family of dilations $\delta_r(x_1, x_2) := (r^{\kappa_1} x_1, r^{\kappa_2} x_2)$, $r > 0$. A function $\phi$ on $\mathbb{R}^2$ is said to be $\kappa$-homogeneous of degree $a$, if $\phi(\delta_r x) = r^a \phi(x)$ for every $r > 0, x \in \mathbb{R}^2$. Such functions will also be called mixed homogeneous. The exponent $a$ will be denoted as the $\kappa$-degree of $\phi$. For instance, the monomial $x_1^j x_2^k$ has $\kappa$-degree $\kappa_1 j + \kappa_2 k$.

If $\phi$ is an arbitrary smooth function near the origin, consider its Taylor series

$$\sum_{j,k=0}^{\infty} c_{jk} x_1^j x_2^k$$

around the origin. We choose $a$ so that the line $L_\kappa := \{(t_1, t_2) \in \mathbb{R}^2 : \kappa_1 t_1 + \kappa_2 t_2 = a\}$ is the supporting line to the Newton polyhedron $\mathcal{N}(\phi)$ of $\phi$. Then the non-trivial polynomial

$$\phi_\kappa(x_1, x_2) := \sum_{(j,k) \in L_\kappa} c_{jk} x_1^j x_2^k$$

is $\kappa$-homogeneous of degree $a$; it will be called the $\kappa$-principal part of $\phi$. By definition, we then have

$$(2.1) \quad \phi(x_1, x_2) = \phi_\kappa(x_1, x_2) + \text{terms of higher } \kappa\text{-degree}.$$
above the line $L_\kappa$, i.e., we have $\kappa_1 j + \kappa_2 k > a$. Moreover, clearly
\[ \mathcal{N}_a(\phi_\kappa) \subset \mathcal{N}_a(\phi). \]

3. Uniform estimates for maximal operators associated to families of
finite type curves and related surfaces

3.1. Finite type curves. In this subsection, we shall prove an extension of some
results by Iosevich [Ios], which allows for uniform estimates for maximal operators
associated to families of curves which arise as small perturbations of a given curve.

We begin with a result whose proof is based on Iosevich’s approach in [Ios].

Proposition 3.1. Consider averaging operators along curves in the plane of the form
\[ A_t f(x) = A_t^{(\rho,\eta,\tau)} f(x) := \int_{\mathbb{R}} f \left( x_1 - t(\rho_1 s + \eta_1), x_2 - t(\eta_2 + \tau s + \rho_2 g(s)) \right) \psi(s) \, ds, \]
where $\rho = (\rho_1, \rho_2), \eta = (\eta_1, \eta_2) \in \mathbb{R}^2, \rho_1 > 0, \rho_2 > 0, \tau \in \mathbb{R}, \psi \in C_0^\infty(\mathbb{R})$ is supported in
a bounded interval $I$ containing the origin, and where
\[ g(s) = s^m \left( b(s) + R(s) \right), \quad s \in I, m \in \mathbb{N}, m \geq 2, \]
with $b \in C^\infty(I, \mathbb{R})$ satisfying $b(0) \neq 0$. Moreover, $R \in C^\infty(I, \mathbb{R})$ is a smooth perturba-
tion term.

By $\mathcal{M}^{(\rho,\eta,\tau)}$, we denote the associated maximal operator
\[ \mathcal{M}^{(\rho,\eta,\tau)} f(x) := \sup_{t > 0} |A_t^{(\rho,\eta,\tau)} f(x)|. \]

Then there exist a neighborhood $U$ of the origin in $I$ and $M \in \mathbb{N}, \delta > 0$, such that
for $p > m$,
\[ \| \mathcal{M}^{(\rho,\eta,\tau)} f \|_p \leq C_p \left( \frac{|\eta_1|}{\rho_1} + \frac{|\eta_2 - \tau \eta_1 / \rho_1|}{\rho_2} + 1 \right)^{1/p} \| f \|_p, \quad f \in \mathcal{S}(\mathbb{R}^2), \]
for every $\psi$ supported in $U$ and every $R$ with $\| R \|_{C^M} < \delta$, with a constant $C_p$ depending
only on $p$ and the $C^M$-norm of $\psi$ (such constants will be called “admissible”).

Proof. Consider the linear operator
\[ T f(x_1, x_2) = (\rho_1 \rho_2)^{-1/p} f \left( \rho_1^{-1} x_1, \rho_2^{-1} (x_2 - \frac{\tau}{\rho_1} x_1) \right). \]
Then $T$ is isometric on $L^p(\mathbb{R}^2)$, and one computes that $\tilde{A}_t := T^{-1} A_t T$ is given by
\[ \tilde{A}_t f(x) = \tilde{A}_t^\sigma f(x) = \int f \left( x_1 - t(s + \sigma_1), x_2 - t(\sigma_2 + g(s)) \right) \psi(s) \, ds, \]
where $\sigma = (\sigma_1, \sigma_2)$ is given by
\[ \sigma_1 = \frac{\eta_1}{\rho_1}, \quad \sigma_2 = \frac{\eta_2}{\rho_2} - \frac{\tau \eta_1}{\rho_2 \rho_1}. \]
Put 
\[ \tilde{M} f(x) = \sup_{t > 0} |\tilde{A}_t f(x)|. \]
Then (3.2) is equivalent to the following estimate for \( \tilde{M} \):

\[ \|\tilde{M} f\|_p \leq C_p (|\sigma| + 1)^{1/p} \|f\|_p, \quad f \in \mathcal{S}(\mathbb{R}^2), \]
for every \( \sigma \in \mathbb{R}^2 \), where \( C_p \) is an admissible constant.

a) We first consider the case \( m = 2 \).

By means of the Fourier inversion formula, we can write
\[ \tilde{A}_t f(x) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i(x - t\sigma) \cdot \xi} H(t\xi) \hat{f}(\xi) d\xi, \]
where
\[ H(\xi_1, \xi_2) := \int_{\mathbb{R}} e^{-i(\xi_1 s + \xi_2 g(s))} \psi(s) ds. \]
If \( s_0 \) is a critical point of the phase \( \xi_1 s + \xi_2 g(s) \) of \( H \), then \( g'(s_0) = -\xi_1/\xi_2 \), where by our assumptions on \( g \) we have \( |g'(s_0)| \sim |s_0| \).

This shows that we can choose a neighborhood \( U \) of \( s = 0 \) in \( \mathbb{R} \) and some \( \varepsilon_1 > 0 \) such that for any \( \xi \) with \( |\xi_1/\xi_2| < \varepsilon_1 \) the phase function has a unique non-degenerate critical point
\[ s_0(\xi_1/\xi_2) = -\frac{\xi_1}{\xi_2} \omega(\frac{\xi_1}{\xi_2}, R) \in U, \]
where \( \omega \) depends smoothly on \( \xi_1/\xi_2 \) and the error term \( R \), and \( \omega(\xi_1/\xi_2, 0) \neq 0 \). Moreover, if \( |\xi_1/\xi_2| \geq \varepsilon_1 \), we may assume that no critical point belongs to \( U \). In the last case, we may integrate by parts to see that
\[ |D_\xi^\alpha H(\xi)| \leq C_{\alpha,N} (1 + |\xi|)^{-N}, \quad |\xi_1/\xi_2| \geq \varepsilon_1, \]
for every \( \alpha \in \mathbb{N}^2 \) with \( |\alpha| \leq 3 \) and \( N = 0, \ldots, 3 \), where the constants \( C_{\alpha,N} \) are admissible.

A similar estimate holds obviously true for \( |\xi| \leq C \). Applying then the stationary phase method to the remaining frequency region, and combining these estimates, we get:
\[ H(\xi) = e^{i q(\xi)} \chi \left( \frac{\xi_1}{\xi_2} \right) A(\xi) \left( 1 + |\xi| \right)^{1/2} + B(\xi), \]
where \( \chi \) is a smooth function supported on a small neighborhood of the origin,
\[ q(\xi) = q(\xi, R) \]
is a smooth function of \( \xi \) and \( R \) which is homogenous of degree 1, and which can be considered as a small perturbation of \( q(\xi, 0) \), if \( R \) is contained in a sufficiently small neighborhood of 0 in \( C^\infty(I, \mathbb{R}) \). It is also important to notice that the Hessian \( D_\xi^2 q(\xi, 0) \)
has rank 1, so that the same applies to $D_2^2 q(\xi, R)$ for small perturbations $R$. Moreover, $A$ is a symbol of order zero such that

$$A(\xi) = 0, \quad \text{if } |\xi| \leq C,$$

and

$$|\xi^\alpha D_2^\xi A(\xi)| \leq C_\alpha, \quad \alpha \in \mathbb{N}^2, |\alpha| \leq 3,$$

where the $C_\alpha$ are admissible constants. Finally, $B$ is a remainder term satisfying

$$|D_2^\xi B(\xi)| \leq C_{\alpha,N}(1 + |\xi|)^{-N}, \quad |\alpha| \leq 3, \ 0 \leq N \leq 3,$$

again with admissible constants $C_{\alpha,N}$.

If we put

$$\tilde{A}_0^0 f(x) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i(x-t\sigma)\cdot \xi} B(t\xi) \hat{f}(\xi) \, d\xi,$$

then by (3.5),

$$\tilde{A}_t^0 f(x) = f \ast k^\sigma_t(x),$$

where

$$k^\sigma_t(x) = t^{-2} k \left( \frac{x}{t} \right)$$

and where $k^\sigma$ is the translate

$$k^\sigma(x) := k(x - \sigma)$$

of $k$ by the vector $\sigma$ of a fixed function $k$ satisfying an estimate of the form

$$|k(x)| \leq C(1 + |x|)^{-3}.$$

Let $\tilde{\mathcal{M}}^0 f(x) := \sup_{t > 0} |\tilde{A}_t^0(x)|$ denote the corresponding maximal operator. (3.6) and (3.7) show that

$$||\tilde{\mathcal{M}}^0||_{L^\infty \to L^\infty} \leq C,$$

with a constant $C$ which does not depend on $\sigma$.

Moreover, scaling by the factor $(|\sigma| + 1)^{-1}$ in direction of the vector $\sigma$, we see that

$$||\tilde{\mathcal{M}}^0||_{L^1 \to L^{1,\infty}} \leq C \left( |\sigma| + 1 \right),$$

since we then can compare with $(|\sigma| + 1)M$, where $M$ is the Hardy-Littlewood maximal operator. By interpolation, these estimates imply that

$$||\tilde{\mathcal{M}}^0||_{L^p \to L^p} \leq C_p (|\sigma| + 1)^{1/p},$$

if $p > 1$.

There remains the maximal operator $\tilde{\mathcal{M}}^1$ corresponding to the family of averaging operators

$$\tilde{A}_t^1(x) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i[x-(t\sigma+tq(\xi))] \cdot \xi} \chi(\xi_1/\xi_2) A(t\xi) \frac{\chi(t\xi_1/\xi_2)}{(1+t|\xi|^2)^{1/2}} \hat{f}(\xi) \, d\xi$$

(notice that $q(t\xi) = tq(\xi)$).
As usually we choose a non-negative function \( \beta \in C_0^\infty(\mathbb{R}) \) such that
\[
\text{supp} \beta \subset [1/2, 2], \quad \sum_{j=-\infty}^{\infty} \beta(2^{-j}r) = 1 \quad \text{for } r > 0,
\]
and put
\[
A_{j,t}f(x) := \int_{\mathbb{R}^2} e^{i[\xi \cdot x - t(\sigma \cdot \xi + q(\xi))]} \frac{\chi(\xi_1/\xi_2)}{(1+t|\xi|)^{1/2}} \hat{\beta}(2^{-j}t|\xi|) \hat{f}(\xi) d\xi.
\]
Since we may assume that \( A \) vanishes on a sufficiently large neighborhood of the origin, we have \( A_{j,t}f = 0 \) if \( j \leq 0 \), so that
\[
(3.8) \quad \tilde{A}^1_t f(x) = \sum_{j=1}^{\infty} A_{j,t}f(x).
\]
Denote by \( \mathcal{M}_j \) the maximal operator associated to the averages \( A_{j,t}, t > 0 \).

Since \( A_{j,t} \) is localized to frequencies \( |\xi| \sim 2^j t \), we can use Littlewood-Paley theory (see [28]) to see that
\[
(3.9) \quad ||\mathcal{M}_j||_{L^p \to L^p} \lesssim ||\mathcal{M}_{j,\text{loc}}||_{L^p \to L^p},
\]
where
\[
\mathcal{M}_{j,\text{loc}} f(x) := \sup_{1 \leq t \leq 2} |A_{j,t}f(x)|.
\]
Choose a bump function \( \rho \in C_0^\infty(\mathbb{R}) \) supported in \([1/2, 4] \) such that \( \rho(t) = 1 \), if \( 1 \leq t \leq 2 \). In order to estimate \( \mathcal{M}_{j,\text{loc}} \), we use the following well-known estimate (see, e.g. [13], Lemma 1.3)
\[
\sup_{t \in \mathbb{R}} |\rho(t)A_{j,t}f(x)|^p \leq p \left( \int_{-\infty}^{\infty} |\rho(t)A_{j,t}f(x)|^p dt \right)^{1/p'} \left( \int_{-\infty}^{\infty} |\frac{\partial}{\partial t} (\rho(t)A_{j,t}f(x))|^p dt \right)^{1/p},
\]
which follows by integration by parts. By Hölder’s inequality, this implies
\[
(3.10) \quad ||\mathcal{M}_{j,\text{loc}} f||_p^p \leq C \left( \int_{\mathbb{R}^2} \int_{1/2}^1 |A_{j,t}f(x)|^p dt \ dx \right)^{p-1} \left( \int_{\mathbb{R}^2} \int_{1/2}^1 \left| \frac{\partial}{\partial t} A_{j,t}f(x) \right|^p dt \ dx \right)^{1/p} + C \int_{\mathbb{R}} \int_{1/2}^4 |A_{j,t}f(x)|^p dt \ dx.
\]
Moreover,
\[
\frac{\partial}{\partial t} A_{j,t}f(x) = \int_{\mathbb{R}^2} e^{i[\xi \cdot x - t(\sigma \cdot \xi + q(\xi))]} \chi \left( \frac{\xi_1}{\xi_2} \right) h(t, j, \xi) d\xi,
\]
where
\[ h(t, j, \xi) = \frac{-i}{(1 + t|\xi|)^{1/2}} A(t\xi) \beta(2^{-j}t|\xi|) + \frac{\partial}{\partial t} \left[ \frac{A(t\xi)}{(1 + t|\xi|)^{1/2}} \right] \beta(2^{-j}t|\xi|) \]
\[ + \frac{A(t\xi)}{(1 + t|\xi|)^{1/2}} 2^{-j}|\xi| \beta'(2^{-j}t|\xi|). \]

Now, if \( t \sim 1 \), since \( A \) vanishes near the origin, it is easy to see that the amplitude of \( A_{j,t} \) can be written as \( 2^{-j/2}a_{j,t}(\xi) \), where \( a_{j,t} \) is a symbol of order 0 localized where \( |\xi| \sim 2^j \). Similarly, the amplitude of \( \frac{\partial}{\partial t}A_{j,t} \) can be written as \( 2^{j/2}(|\sigma| + 1)b_{j,t} \), where \( b_{j,t} \) is a symbol of order 0 localized where \( |\xi| \sim 2^j \), and \( a_{j,t}, b_{j,t} \) satisfy estimates of the form
\[ \left| (1 + |\xi|)^{\alpha} \left( |D^\alpha a_{j,t}(\xi)| + |D^\alpha b_{j,t}(\xi)| \right) \right| \leq C_\alpha, \]
with admissible constants \( C_\alpha \).

We can then apply the local smoothing estimates by Mockenhaupt, Seeger and Sogge from [18], [19] for operators of the form
\[ P_j f(x, t) = \int e^{i(\xi \cdot x - \alpha(\xi))} a(t, \xi)\beta(2^{-j}|\xi|) \hat{f}(\xi) d\xi, \]
where \( a(t, \xi) \) is a symbol of order 0 in \( \xi \), and the Hessian matrix of \( q \) has rank 1 everywhere. Their results imply in particular that for \( 2 < p < \infty \),
\[ \left( \int_{1/2}^4 \left( \int_{\mathbb{R}^2} |P_j f(x,t)|^p dx \right) dt \right)^{1/p} \leq C_p 2^j \left( \frac{1}{2} \frac{1}{p} - \delta(p) \right) \|f\|_{L^p(\mathbb{R}^2)}, \]
for some \( \delta(p) > 0 \).

Since \( 2^{j/2}A_{j,t}f(x) \) and \( 2^{-j/2}(|\sigma| + 1)^{-1}\frac{\partial}{\partial t}A_{j,t}f(x) \) are of the form \( P_j f(x - t\sigma) \), for suitable operators \( P_j \) of this type, we can apply (3.10) and (3.11) to obtain, if \( R = 0 \),
\[ \|\mathcal{M}_{j,loc}f\|_p \leq C_p 2^{j(\frac{1}{2} - \frac{1}{p} - \delta(p))} \cdot 2^{(j/2(p-1) + j/2)/p} (|\sigma| + 1)^{1/p} \|f\|_p, \]
i.e.,
\[ \|\mathcal{M}_{j,loc}f\|_p \leq C_p (|\sigma| + 1)^{1/p} 2^{-\delta(p)j} \|f\|_p, \]
if \( 2 < p < \infty \), where \( \delta(p) > 0 \).

However, as observed in [13], the estimate (3.11) remains valid under small, sufficiently smooth perturbations, and the constant \( C_p \) depends only on a finite number of derivatives of the phase function and the symbol of \( P_j \). Therefore, if \( \delta \) is sufficiently small and \( \|R\|_{C^M} < \delta \), then estimate (3.12) holds true also for \( R \neq 0 \), with an admissible constant \( C_p \).

Summing over all \( j \geq 1 \) (compare (3.8)), we thus get
\[ \|\tilde{\mathcal{M}}^1 f\|_p \leq C_p (|\sigma| + 1)^{1/p} \|f\|_p, \]
with an admissible constant \( C_p \).

This finishes the proof of the proposition in the case \( m = 2 \).
b) Let next \( m \in \mathbb{N}, m \geq 2, \) be arbitrary. Following [13], we shall reduce this case to the previous case \( m = 2 \) by means of a dyadic decomposition and scaling in \( s. \) Given \( K \in \mathbb{N} \) and \( d > 0, \) we choose a bump function \( \beta \in C^\infty_0(\mathbb{R}) \) supported in \((d/2, 2d)\) such that
\[
\sum_{k=K}^{\infty} \beta(2^k s) + \sum_{k=K}^{\infty} \beta(-2^k s) = 1 \quad \text{for every } s \in \text{supp } \psi \setminus \{0\}
\]
(this is possible, if \( \text{supp } \psi \) is assumed to be sufficiently small). Accordingly, we decompose the averaging operator \( \tilde{A}_t = \tilde{A}_t^k: \)
\[
(3.13) \quad \tilde{A}_t f(x) = \sum_{k=K}^{\infty} \tilde{A}_t^k f(x) + \sum_{k=K}^{\infty} \tilde{B}_t^k f(x),
\]
where
\[
\tilde{A}_t^k f(x) := \int_{\mathbb{R}} f \left( x_1 - t(s + \sigma_1), x_2 - t(\sigma_2 + g(s)) \right) \psi(s) \beta(2^k s) \, ds,
\]
and where \( \tilde{B}_t^k f(x) \) is defined in the same way, only with \( \beta(2^k s) \) replaced by \( \beta(2^{-k} s). \)

Let us consider the maximal operator
\[
\tilde{\mathcal{M}}^k f(x) := \sup_{t > 0} |\tilde{A}_t^k f(x)|, \quad k \geq K.
\]
Changing coordinates \( s \mapsto 2^{-k}(d + s), \) we obtain
\[
\tilde{A}_t^k f(x) = 2^{-k} \int_{\mathbb{R}} f \left( x_1 - t(2^{-k} s + 2^{-k} d + \sigma_1), x_2 - t(\sigma_2 + g(2^{-k}(d + s))) \right) \psi(2^{-k}(d + s)) \beta(d + s) \, ds,
\]
where \( s \mapsto \beta(d + s) \) is supported where \(-\frac{d}{2} < s < d.\) Observe that, by (3.1),
\[
g(2^{-k}(d + s)) = 2^{-mk}(d + s)^m \left( b(2^{-k}(d + s)) + R(2^{-k}(d + s)) \right),
\]
where
\[
(d + s)^m = (d^m + m d^{m-1} s + s^2 Q(s)),
\]
for some polynomial \( Q \) with \( Q(0) = \binom{m}{2} d^{m-2} \neq 0, \) whose coefficients are polynomials in \( d.\) Moreover, by Taylor’s formula
\[
b(2^{-k}(d + s)) = b(2^{-k} d) + 2^{-k} b'(2^{-k} d) s + \tilde{b}_k(s) s^2,
\]
where \( b(0) \neq 0 \) and \( ||\tilde{b}_k||_{C^M} \lesssim 2^{-2k}, \) if \( K \) is assumed to be sufficiently large. Similarly,
\[
R(2^{-k}(d + s)) = R(2^{-k} d) + 2^{-k} R'(2^{-k} d) s + \tilde{R}_k(s) s^2,
\]
where
\[
|R'(2^{-k} d)| \lesssim \delta \quad \text{and} \quad ||\tilde{R}_k||_{C^M} \lesssim 2^{-2k} \delta,
\]
if \( ||R||_{C^{M+2}} \leq \delta, \) say.

We may assume that \( 2^{-K} \leq \delta. \) Then we see that
\[
g(2^{-k}(d + s)) = \alpha + \beta s + 2^{-mk} s^2 g_k(s),
\]
where
\[
\alpha = 2^{-mk}d^m(b + R)(2^{-k}d),
\]
\[
\beta = 2^{-mk}md^{m-1}(b + R)(2^{-k}d) + d^m2^{-k}(b + R)'(2^{-k}d)
\]
\[
g_k(s) = Q(s)b(2^{-k}(d + s)) + d^m(\tilde{b}_k + \tilde{R}_k)(s) + md^{m-1}2^{-k}\left(b'(2^{-k}d) + R'(2^{-k}d)\right).
\]
Notice that, since \(2^{-k} \leq \delta\),
\[
g_k(s) = Q(s)b(0) + r_k(s),
\]
where \(\|r_k\|_{CM} \lesssim \delta\). This shows that \(g_k(s)\) is a perturbation of the fixed function \(b(0)Q(s)\), and thus we can apply the estimate (3.2) for the case \(m = 2\) to the maximal operator
\[
\tilde{M}_kf(x) := \sup_{t > 0} |\tilde{A}_t^k f(x)|,
\]
with
\[
\rho_1 = 2^{-k}, \quad \eta_1 = 2^{-k}d + \sigma_1,
\]
\[
\rho_2 = 2^{-mk}, \quad \eta_2 = \sigma_2 + \alpha, \quad \tau = \beta.
\]
I.e., if we fix some sufficiently small \(d > 0\), then for \(p > 2\),
\[
\|\tilde{M}_k\|_{L^p \to L^p} \leq C_p 2^{-k}\left(2^k|\sigma_1| + 2^{mk}(|\sigma_2| + |\alpha|) + 2^{mk}|d + 2^k\sigma_1| + 1\right)^{1/p}.
\]
Since clearly \(|\alpha| = O(2^{-mk})\), \(|\beta| = O(2^{-k})\), this implies
\[
\|\tilde{M}_k\|_{L^p \to L^p} \leq C_p(|\sigma_1| + |\sigma_2| + 1)^{1/p} 2^{-k(1 - \frac{m}{p})}.
\]
For \(p > m\), we can thus sum over all \(k \geq K\) and obtain (3.3) (notice that the maximal operators associated to the \(\tilde{B}_k^t\) can be estimated in the same way by means of the change of coordinates \(s \mapsto -s\)).

Q.E.D.

Consider now a smooth function \(a : I \to \mathbb{R}\), where \(I\) is a compact interval of positive length. We say that \(a\) is a function of polynomial type \(m \geq 2\) (\(m \in \mathbb{N}\)), if there is a positive constant \(c > 0\) such that

\[
(3.14) \quad c \leq \sum_{j=2}^{m} |a^{(j)}(s)| \quad \text{for every } s \in I,
\]
and if \(m\) is minimal with this property. Oscillatory integrals with phase functions \(a\) of this type have been studied, e.g., by J. E. Björk (see [7]) and G. I. Arhipov [1], and it is our goal here to estimate related maximal operators, allowing even for small perturbations of \(a\). More precisely, consider averaging operators
\[
A^\varepsilon_t f(x) := \int_\mathbb{R} f\left(x_1 - ts, x_2 - t(1 + \varepsilon(a(s) + r(s)))\right) \psi(s) ds, \quad f \in \mathcal{S}(\mathbb{R}^2),
\]
along dilates by factors \( t > 0 \) of the curve
\[
\gamma(s) := \left( s, 1 + \varepsilon(a(s) + r(s)) \right), \quad s \in I,
\]
where \( \varepsilon > 0, \psi \in C^\infty(I) \) is a smooth, non-negative density and \( r \in C^\infty(I) \) will be a sufficiently small perturbation term. By \( M^\varepsilon \) we denote the corresponding maximal operator
\[
M^\varepsilon f(x) := \sup_{t > 0} |A^\varepsilon_t f(x)|.
\]

**Theorem 3.2.** Let \( a \) be a function of polynomial type \( m \geq 2 \). Then there exist numbers \( M \in \mathbb{N}, \delta > 0, \) such that for every \( r \in C^\infty(I, \mathbb{R}) \) with \( ||r||_{CM} < \delta, 0 < \varepsilon << 1 \) and \( p > m \), the following a priori estimate is satisfied:
\[
(3.15) \quad ||M^\varepsilon f||_p \leq C_p \varepsilon^{-1/p} ||f||_p, \quad f \in \mathcal{S}(\mathbb{R}^2),
\]
with a constant \( C_p \) depending only on \( p \).

By means of an induction argument (based on an idea of J. J. Duistermaat [8]), we shall reduce this theorem to Proposition 3.1.

Let us fix a smooth function \( a : I \to \mathbb{R} \) of polynomial type \( m \geq 2 \). We shall proceed by induction on the type \( m \).

Observe first that it suffices to find for every fixed \( s_0 \in I \) a subinterval \( I_0 \subset I \) which is relatively open in \( I \) and contains \( s_0 \) such that (3.15) holds for every \( \psi \) supported in \( I_0 \). For, then we can cover \( I \) be a finite number of such subintervals \( I_j \), decompose \( \psi \) by means of a subordinate smooth partition of unity into \( \psi = \sum_j \psi_j \), where \( \psi_j \) is supported in \( I_j \), and apply the estimate (3.15) for each of the pieces.

So, fix \( s_0 \in I \). Extending the function \( a \) in a suitable way to a \( C^\infty \)-function beyond the boundary points of \( I \), we may assume that \( s_0 \) lies in the interior of \( I \). Translating by \( s_0 \), we may furthermore assume that \( s_0 = 0 \). Then, by (3.14), there is some \( k \in \mathbb{N}, 2 \leq k \leq m \), such that
\[
(3.16) \quad a^{(j)}(0) = 0 \text{ for } 2 \leq j \leq k - 1, \text{ and } a^{(k)}(0) \neq 0.
\]

Assume first that \( k = 2 \). Then we may write
\[
a(s) = \alpha_0 + \alpha_1 s + s^2 b(s) \quad \text{near } s = 0,
\]
where \( b \in C^\infty(I), b(0) \neq 0 \). Consequently, if \( r \in C^\infty(I) \) with \( ||r||_{CM+2} < \delta \), then, by Taylor’s formula,
\[
a(s) + b(s) = (\alpha_0 + r(0)) + (\alpha_1 + r'(0)) s + s^2 (b(s) + R(s)) \quad \text{near } s = 0,
\]
where \( ||R||_{CM} \lesssim \delta \). Estimate (3.15) thus follows from Proposition 3.1.

Let next \( k \geq 2 \).
Lemma 3.3. Assume \( a \) satisfies (3.16) with \( k \geq 3 \), and let \( N \in \mathbb{N} \). Then there is some \( \delta > 0 \), and for every function \( r \in C^\infty(I) \) with \( ||r||_{C^{k+N}(I)} < \delta \) a number \( \sigma(r) \in I \) with \( |\sigma(r)| \lesssim \delta \), depending smoothly on \( r \), such that

\[
(a + r)^{(k-1)}(\sigma(r)) = 0 .
\]

In particular, if we put \( I_r := -\sigma(r) + I \) and \( \mu := (\mu_0, \ldots, \mu_{m-2}) \), then

\[
(a + r)(s + \sigma(r)) = (b(s) + R(s))s^k + \mu_0 s + \cdots + \mu_{k-2}s^{k-2},
\]

where \( b \in C^\infty(I_r) \) with \( b(0) \neq 0 \), \( R \in C^\infty(I_r) \) with \( ||R||_{C^N} \lesssim \delta \) and \( |\mu| \lesssim \delta \).

Proof. (3.17) follows from the implicit function theorem, applied to the mapping

\[
f : I \times C^{k+N}(I) \to \mathbb{R}, f(s, r) := (a + r)^{(k-1)}(s),
\]

and (3.18) is then a consequence of Taylor’s formula.

Q.E.D.

The case \( k = 3 \) can now be treated by means of (3.18) in a similar way as the case \( k = 2 \) (notice that \( I_r \) and \( I \) overlap in a neighborhood \( U \) of 0 not depending on \( r \), if \( \delta \) is sufficiently small, so that we can again assume that \( \psi \) is supported in a fixed interval contained in \( U \)).

We may thus from now on assume that \( k \geq 4 \). Since we have seen that the cases \( m = 2 \) and \( m = 3 \) of Theorem 3.2 are true, we may assume that \( m \geq 4 \), and, by induction hypothesis, that the statement of Theorem 3.2 is true for all \( m' \leq m - 1 \). Then, we may also assume that \( k = m \) in (3.16), so that, by Lemma 3.3,

\[
(a + r)(s + \sigma(r)) = \tilde{b}(s)s^m + \mu_2 s^2 + \cdots + \mu_{m-2}s^{m-2}
\]
on \( I_r \), where \( m - 2 \geq 2 \) (the affine linear term \( \mu_0 + \mu_1 s \) can again be omitted by means of a linear change of coordinates). Here we have set \( \tilde{b} = b + R \), where, by Lemma 3.3, \( ||R||_{C^M} \lesssim \delta \).

Let us put now \( \mu = (\mu_2, \ldots, \mu_{m-2}) \). The case \( \mu = 0 \) can again be treated by Proposition 3.1, so assume \( \mu \neq 0 \).

If we scale in \( s \) by a factor \( \rho^{1/m} \), \( \rho > 0 \), we obtain

\[
(a + r) \left( \rho^{1/m}s + \sigma(r) \right) = \rho \left[ \tilde{b}(\rho^{1/m}s)s^m + \frac{\mu_2}{\rho^{m-2}}s^2 + \cdots + \frac{\mu_{m-2}}{\rho^{m-2}}s^{m-2} \right].
\]

This suggests to introduce a quasi-norm

\[
N(\mu) := \left[ \mu_2^{m-2} + \cdots + \mu_{m-2}^{m-2} \right]^{1/\nu},
\]
say with \( \nu := 2(m - 2)! \). For then \( N \) is smooth away from the origin, and if we put \( \rho := N(\mu) \), i.e., if we define \( \xi = (\xi_2, \ldots, \xi_{m-2}) \) by

\[
\xi_2 := \frac{\mu_2}{N(\mu)^{\frac{m-2}{m}}}, \ldots, \xi_{m-2} := \frac{\mu_{m-2}}{N(\mu)^{\frac{m}{m}}},
\]
then $N(\xi) = 1,$ and

$$g(s) = g(s, \rho, \xi) := \frac{1}{\rho} (a + r) (\rho^{1/m} s + \sigma(r)) = \bar{b}(\rho^{1/m} s) s^m + \xi_2 s^2 + \cdots + \xi_{m-2} s^{m-2}.$$ 

Then, putting $\eta := \sigma(r),$ we have

$$A_t f(x) = \rho^{1/m} \int_{\mathbb{R}} f \left( x_1 - t(\rho^{1/m} s + \eta), x_2 - t(1 + \varepsilon \rho g(s)) \psi(\rho^{1/m} s + \eta) \right) ds.$$ 

Recall at this point that $\eta \to 0$ and $\rho \to 0$ as $\delta \to 0$. In particular we may consider $g(s, \rho, \xi)$ as a $C^\infty$-perturbation of $g(s, 0, \xi)$, where

$$g(s, 0, \xi) = \bar{b}(0) s^m + \xi_2 s^2 + \cdots + \xi_{m-2} s^{m-2}.$$ 

Denote by $\Sigma$ the unit sphere

$$\Sigma := \{ \xi \in \mathbb{R}^{m-3} : N(\xi) = 1 \}$$

with respect to the quasi-norm $N$, and choose $B > 0$ so large that

$$|g''(s)| \geq c |s|^{m-2} \text{ whenever } |s| \geq B, \, \xi \in I, \, \rho < \delta,$$

where $c > 0$. This is possible, since $\bar{b}(0) \neq 0$, provided $\delta$ is sufficiently small. We then choose $\chi_0, \chi \in C_0^\infty(\mathbb{R})$ such that $\text{supp } \chi \subset (-2B, -B/2) \cup (B/2, 2B)$ and

$$1 = \chi_0(s) + \sum_{k=1}^{\infty} \chi \left( \frac{s}{2^k} \right) := \chi_0(s) + \sum_{k=1}^{\infty} \chi_k(s) \text{ for every } s \in \mathbb{R}.$$ 

Accordingly, we decompose

$$A_t = \sum_{k=0}^{\infty} A_t^\varepsilon k,$$

where

$$A_t^\varepsilon k f(x) := \rho^{1/m} \int_{\mathbb{R}} f \left( x_1 - t(\rho^{1/m} s + \eta), x_2 - t(1 + \varepsilon \rho g(s)) \psi(\rho^{1/m} s + \eta) \chi_k(s) \right) ds.$$ 

Assume first that $k \geq 1$. Then this can be re-written as

$$A_t^\varepsilon k f(x) = 2^k \rho^{1/m} \int_{\mathbb{R}} f \left( x_1 - t(\rho^{1/m} 2^k s + \eta), x_2 - t(1 + \varepsilon \rho 2^{mk} g_k(s)) \psi(\rho^{1/m} 2^k s + \eta) \chi(s) \right) ds,$$

where

$$g_k(s) = g_k(s, \rho, \xi) := 2^{-mk} g(2^k s, \rho, \xi).$$

And, by (3.19),

$$|g''(s)| \geq c > 0 \text{ for every } s \in \text{supp } \chi, \xi \in \Sigma, \rho < \delta.$$ 

More precisely, since

$$g_k(s) = \bar{b}(\rho^{1/m} 2^k s) s^m + \frac{\xi_2}{2^{(m-2)k}} s^2 + \cdots + \frac{\xi_{m-2}}{2^{2k}} s^{m-2},$$
where $|s| \sim B$, and where $\rho^{1/m}2^k \leq \delta$, unless $A^{\varepsilon,k}_t = 0$, if we choose $\text{supp } \psi$ sufficiently close to 0, we see that $g_k(s)$ is a small $\delta$-perturbation of $g_k(s,0,\xi)$.

Moreover, covering $\Sigma$ by a finite number of $\delta$-neighborhoods $\Sigma_j$ of points $\xi^{(j)} \in \Sigma$, for every $\xi \in \Sigma_j$ we may regard $g_k(s,0,\xi)$ as a $\delta$-perturbation of $g_k(s,0,\xi^{(j)})$. Thus, for $\xi \in \Sigma_j$, Proposition 3.1 can be applied for $n = 2$, in a similar way as in our discussion of the case $k = 2$, in order to estimate the maximal operator

$$
\mathcal{M}^{\varepsilon,k}_t f(x) = \sup_{t > 0} |A^{\varepsilon,k}_t f(x)|
$$

by

$$
||\mathcal{M}^{\varepsilon,k}_t f||_p \leq C_p \sum_{k=0}^{\infty} \rho^{1/m} \left( |\eta|^2 (\rho^{1/m})^{-1} + (\varepsilon \rho 2^m)^{-1} + 1 \right) \frac{1}{\rho} ||f||_p 
\leq C_p \left( (2^k \rho^{1/m})^{1-\frac{1}{p}} + \varepsilon^{-1/p} (2^k \rho^{1/m})^{1-\frac{1}{p}} \right) ||f||_p.
$$

Since $\mathcal{M}^{\varepsilon,k}_t = 0$ if $2^k \rho^{1/m} > \delta$, we then obtain for $p > m$

$$
\sum_{k \geq 1} ||\mathcal{M}^{\varepsilon,k}_t f||_p = \sum_{k \geq 1, 2^k \rho^{1/m} \leq \delta} ||\mathcal{M}^{\varepsilon,k}_t f||_p \leq C_p \varepsilon^{-1/p} ||f||_p.
$$

There remains the operator $\mathcal{M}^{\varepsilon,0}_t$. Conjugating $A^{\varepsilon,0}_t$ with the scaling operator

$$
T \rho f(x_1,x_2) := \rho^{-1/(mp)} f(\rho^{-1/m} x_1, x_2),
$$

which acts isometrically on $L^p(\mathbb{R}^2)$, we can reduce our considerations to the averaging operator

$$
T^{-1}_\rho A^{\varepsilon,0}_t T \rho f(x)
\quad := \quad \rho^{1/m} \int_{\mathbb{R}} f \left( x_1 - t(s + \rho^{-1/m} \eta), x_2 - t(1 + \varepsilon \rho g(s)) \right) \psi(\rho^{1/m} s + \eta) \chi_0(s) ds.
$$

Fixing again $\xi^0 \in \Sigma$, for $\xi$ in a $\delta$-neighborhood $\Sigma_0$ of $\xi^0$, we can consider $g(s,\rho,\xi)$ as a $\delta$-perturbation of the polynomial function

$$
P(s) := g(s,0,\xi^0) = \tilde{b}(0)s^m + \xi^0_2 s^2 + \cdots + \xi^0_{m-2} s^{m-2}.
$$

Since there is no term $\xi^0_{m-1} s^{m-1}$ in $P(s)$, and since $\xi^0 \neq 0$, it follows that for every $s_0$ one has

$$
\sum_{j=2}^{m-1} |(\partial_{s_0})^j g(s_0,0,\xi^0)| \neq 0,
$$

for otherwise we had

$$
P(s) - P(s_0) - P'(s_0)(s-s_0) = \tilde{b}(0)(s-s_0)^m = \tilde{b}(0)(s^m - ms_0 s^{m-1} + \cdots),
$$

hence $s_0 = 0$, and so $\xi^0 = 0$.

We can thus apply our induction hypothesis, and obtain for $p > m - 1$

$$
||T^{-1}_\rho \mathcal{M}^{\varepsilon,0}_t T \rho f||_p \leq C_p \rho^{1/m} \left( \rho^{-1/m} |\eta| + (\varepsilon \rho)^{-1} \right)^{1/p} ||f||_p.
$$
hence
\[ \| M^{\epsilon,0} f \|_p \leq C_p \epsilon^{-1/p} p^{1/m-1/p} \| f \|_p , \]
first for \( \xi \in \Sigma_0 \), and then, by covering \( \Sigma \) again by a finite number of \( \delta \)-neighborhoods of points \( \xi_j \), for every \( \xi \in \Sigma \). In particular, for \( p > m \) we get the uniform estimate
\[ \| M^{\epsilon,0} f \|_p \leq C_p \epsilon^{-1/p} \| f \|_p , \]
which concludes the proof of Theorem 3.2.

Q.E.D.

In the next subsection, we shall need a slight generalization of this theorem, namely for averaging operators of the form
\[ A_{t}^{\epsilon,\sigma_1} f(x) := \int_{\mathbb{R}} f \left( x - t(s + \sigma_1), x_2 - t(1 + \epsilon(a(s) + r(s))) \right) \psi(s) \, ds, \quad f \in \mathcal{S}(\mathbb{R}^2), \]
where \( \sigma_1 \) is a second real parameter which can be arbitrarily large. The corresponding maximal operator
\[ M^{\epsilon,\sigma_1} f(x) := \sup_{t>0} \| A_{t}^{\epsilon,\sigma_1} f(x) \|_p \]
can be estimated exactly as before, if we simply replace the shift term \( \eta \) in the proof of Theorem 3.2 by \( \eta + \sigma_1 \), and one easily obtains

**Corollary 3.4.** Let \( a \) be a function of polynomial type \( m \geq 2 \). Then there exist numbers \( M \in \mathbb{N}, \delta > 0 \), such that for every \( r \in C^\infty(I, \mathbb{R}) \) with \( \| r \|_{CM} < \delta \), \( 0 < \epsilon << 1 \) and \( p > m \), the following a priori estimate is satisfied:
\[ \| M^{\epsilon,\sigma_1} f \|_p \leq C_p (|\sigma_1| + \epsilon^{-1})^{1/p} \| f \|_p, \quad f \in \mathcal{S}(\mathbb{R}^2), \]
with a constant \( C_p \) depending only on \( p \).

3.2. Related results for families of surfaces. By decomposing a given surface in \( \mathbb{R}^3 \) by means of a “fan” of hyperplanes into a family of curves, we can easily derive suitable estimates for certain families of surfaces from the maximal estimates in the previous subsection.

Let \( U \) be an open neighborhood of the point \( x^0 \in \mathbb{R}^2 \), and let \( \phi_p \in C^\infty(U, \mathbb{R}) \) such that
\[ \partial^m_{x_1, x_2} \phi_p(x^0, x^0) \neq 0, \quad (3.20) \]
where \( m \geq 2 \). Let
\[ \phi = \phi_p + \phi_r, \]
where \( \phi_r \in C^\infty(U, \mathbb{R}) \) sufficiently small. Denote by \( S_\varepsilon \) the surface in \( \mathbb{R}^3 \) given by
\[ S_\varepsilon := \{(x_1, x_2, 1 + \varepsilon \phi(x_1, x_2)) : (x_1, x_2) \in U\}, \]
with \( \varepsilon > 0 \), and consider the averaging operators
\[ A_t f(x) = A_t^\varepsilon f(x) := \int_{S_\varepsilon} f(x - ty) \psi(y) \, d\sigma(y), \]
where $d\sigma$ denotes the surface measure and $\psi \in C_0^\infty(S_\epsilon)$ is a non-negative cut-off function. Define the associated maximal operator by

$$\mathcal{M}^\epsilon f(x) := \sup_{t>0} |A_t^\epsilon f(x)|.$$  

**Proposition 3.5.** Assume that $\phi_p$ satisfies (3.20) and that the neighborhood $U$ of the point $x^0$ is sufficiently small. Then there exist numbers $M \in \mathbb{N}$, $\delta > 0$, such that for every $\phi_\epsilon \in C^\infty(U, \mathbb{R})$ with $||\phi_\epsilon||_{CM} < \delta$ and any $p > m$ there exists a positive constant $C_p$ such that for $\epsilon > 0$ sufficiently small the maximal operator $\mathcal{M}^\epsilon$ satisfies the following a priori estimate:

$$(3.21) \quad ||\mathcal{M}^\epsilon f||_p \leq C_p \epsilon^{-1/p} ||f||_p, \quad f \in \mathcal{S}(\mathbb{R}^3).$$

**Proof.** Let us write the averaging operator $A_t$ in the form

$$A_t f(y) = \int_{\mathbb{R}^2} f \left( y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \epsilon \phi(x_1, x_2)) \right) \eta(x_1, x_2) \, dx,$$

where $\eta \in C_0^\infty(U)$. Choose $\theta_0$ such that $\sin(\theta_0) + \cos(\theta_0)x_1^0 = 0$ (notice that we may assume that $\cos(\theta_0) > 0$). For small $\theta$, consider the equation

$$(3.23) \quad \sin(\theta_0 + \theta)(1 + \epsilon \phi(x_1, x_2)) + \cos(\theta_0 + \theta)x_1 = 0$$

with respect to $x_1$. By the implicit function theorem, the last equation has a unique smooth solution $x_1(\theta, x_2, \epsilon)$ for $|\theta|, |x_2-x_2^0|$ and $\epsilon$ sufficiently small such that $x_1(0, x_2^0, 0) = x_1^0$. Moreover $\frac{\partial}{\partial \theta}x_1(0, x_2^0, 0) \neq 0$. In the integral (3.22) we can thus use the change of variables $(\theta, x_2) \mapsto (x_1(\theta, x_2, \epsilon), x_2)$ (assuming $U$ to be sufficiently small) and obtain

$$A_t f(y) = \int_{\mathbb{R}^2} f \left( y_1 - tx_1(\theta, x_2, \epsilon), y_2 - tx_2, y_3 - t(1 + \epsilon \phi(x_1(\theta, x_2, \epsilon), x_2)) \right) \psi(\theta, x_2, \epsilon) \, d\theta \, dx_2,$$

where $\psi(\theta, x_2, \epsilon) := \eta(x_1(\theta, x_2, \epsilon), x_2)|J(\theta, x_2, \epsilon)|$ and $J(\theta, x_2, \epsilon)$ denotes the Jacobian of this change of coordinates. Let us write the integral (3.24) as an iterated integral

$$A_t f(y) = \int_{-b}^b A_t^0 f(y) d\theta,$$

where $b$ is some positive number and $A_t^0$ denotes the following averaging operator along a curve:

$$A_t^0 f(y) := \int_{\mathbb{R}^2} f \left( y_1 - tx_1(\theta, s, \epsilon), y_2 - ts, y_3 - t(1 + \epsilon \phi(x_1(\theta, s, \epsilon), s)) \right) \psi(\theta, s, \epsilon) \, ds.$$  

Now, we define the rotation operator

$$R^\theta f(x) := f(x_1 \sin(\theta_0 + \theta) - x_3 \cos(\theta_0 + \theta), x_2, x_1 \cos(\theta_0 + \theta) + x_3 \sin(\theta_0 + \theta)),$$
which acts isometrically on every $L^p(\mathbb{R}^3)$. Then we have
\[
R^{-\theta}A^\theta_i R^\theta f(y) = \int_{\mathbb{R}^2} f(y_1 + t \frac{1}{\cos(\theta + \theta)}(1 + \varepsilon \phi(x_1(\theta, s, \varepsilon), s), y_2 - ts, y_3)) \psi(\theta, s, \varepsilon) \, ds.
\]

Observe that the last operator "acts" only on the first two variables. Moreover, for $\varepsilon = 0$, by (3.23), we have $x_1(\theta, x_2, 0) = -\tan(\theta + \theta)$, which is independent of $x_2$. This implies that
\[
\frac{d^m}{ds^m} \left. \left( \phi_p(x_1(0, s, 0), s) \right) \right|_{s=x_2^0} = \partial_2^m \phi_p(x_1^0, x_2^0) \neq 0.
\]
Notice also that for $\varepsilon, \delta$ and $U$ (hence also $\theta$) sufficiently small, $\phi(x_1(\theta, s, \varepsilon), s)$ can be regarded as a small perturbation of $\phi_p(x_1(0, s, 0), s)$. Therefore we can apply Theorem 3.2 (in the first two variables) and obtain that for $p > m$
\[
\left\| \sup_{t>0} |R^{-\theta}A^\theta_i R^\theta f| \right\|_p \leq C_p \varepsilon^{-1/p} \|f\|_p,
\]
hence
\[
\left\| \sup_{t>0} |A^\theta f| \right\|_p \leq C_p \varepsilon^{-1/p} \|f\|_p,
\]
where $C_p$ is independent of $\theta$ and $\varepsilon$. Integrating finally in the $\theta$ variable we obtain the required estimate.

Q.E.D.

In our later applications of this proposition, we shall also have to deal with functions $\phi$ which depend in fact also on the parameter $\varepsilon$ in such a way that they blow up as $\varepsilon \to 0$, however, in a particular way. More presisely, assume $\tilde{\phi} = \phi_p + \tilde{\phi}_r$ has the same properties as $\phi$ in the proposition, so that in particular (3.20) is satisfied by $\tilde{\phi}$. We assume for simplicity that $\tilde{\phi}$ is defined on $\mathbb{R}^2$ and supported in the neighborhood $V$ of the point $x^0$. Let further $\psi_\varepsilon \in C^\infty(V_1)$ be a smooth function depending on the parameter $\varepsilon$ so that there is some $0 \leq \delta < 1$ such that
\[
\psi_\varepsilon = O(\varepsilon^{-\delta}) \text{ in } C^\infty,
\]
in the sense that $\|\psi_\varepsilon\|_{C^m(V_1)} = O(\varepsilon^{-\delta})$ for every $m \in \mathbb{N}$, where $V_1$ denotes the orthogonal projection of the neighborhood $V$ onto the $x_1$-axis. Put then
\[
\phi_\varepsilon(x_1, x_2) := \tilde{\phi}(x_1, x_2 - \psi_\varepsilon(x_1)).
\]
Notice that then
\[
|\partial_1^j \partial_2^k \phi_\varepsilon(x)| = O(\varepsilon^{-j\delta}).
\]

This means that we cannot directly apply Proposition 3.5 to $\phi_\varepsilon$. We shall see that nevertheless the proof of this proposition can be extended to $\phi_\varepsilon$. To this end, observe first that $|\nabla(\varepsilon \phi_\varepsilon)(x)| \leq C \varepsilon^{1-\delta}$, uniformly in $x$. Therefore, again by the implicit function theorem, we can solve the equation
\[
\sin(\theta_0 + \theta)(1 + \varepsilon \phi_\varepsilon(x_1, x_2)) + \cos(\theta_0 + \theta)x_1 = 0
\]
in $x_1$ near the point $(x_1^0, x_2^0 + \psi_\varepsilon(x_1^0))$, and obtain a smooth solution $x_1(\theta, x_2, \varepsilon)$ for sufficiently small values of $|\theta|$, $|x_2 - (x_2^0 + \psi_\varepsilon(x_1^0))|$ and $\varepsilon > 0$, satisfying $x_1(0, x_2^0 + \psi_\varepsilon(x_1^0), 0) = x_1^0$.

Let us also define $x_1^0(\theta)$ as the solution of the equation

$$\sin(\theta_0 + \theta) + \cos(\theta_0 + \theta) x_1^0(\theta) = 0,$$

and put $g(\theta, x_2, \varepsilon) := x_1(\theta, x_2, \varepsilon) - x_1^0(\theta)$. Then $g$ satisfies the equation

$$\sin(\theta_0 + \theta) \varepsilon \phi_\varepsilon \left( x_1^0(\theta) + g(\theta, x_2, \varepsilon), x_2 \right) + \cos(\theta_0 + \theta) g(\theta, x_2, \varepsilon) = 0.$$

Implicit differentiation shows that

$$g_\varepsilon'(x_2) = -\varepsilon \frac{\sin(\theta_0 + \theta) \partial_2 \phi_\varepsilon \left( x_1^0(\theta) + g_\varepsilon(x_2), x_2 \right)}{\cos(\theta_0 + \theta) + \sin(\theta_0 + \theta) \varepsilon \partial_1 \phi_\varepsilon \left( x_1^0(\theta) + g_\varepsilon(x_2), x_2 \right)},$$

if we use the short-hand notation $g_\varepsilon(x_2) = g(\theta, x_2, \varepsilon)$. By (3.27), this implies that $|g_\varepsilon'(x_2)| = O(\varepsilon)$, and similarly $|g_\varepsilon^{(j)}(x_2)| = O(\varepsilon)$, for every $j \geq 1$, uniformly in $x_2$. But clearly this estimate is also true for $j = 0$, so that

$$g_\varepsilon = O(\varepsilon) \text{ in } C^\infty.$$

If put

$$\Phi_\varepsilon(\theta, s) := \phi_\varepsilon \left( x_1^0(\theta) + g_\varepsilon(s), s \right),$$

then (3.27), (3.28) show that $\Phi_\varepsilon(\theta, \cdot) = O(1)$ in $C^\infty$. The averaging operators associated to $\phi_\varepsilon$ will be of the form

$$(3.29) \quad A_t f(y) := \int_{\mathbb{R}^2} f \left( y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \varepsilon \phi_\varepsilon(x_1, x_2)) \right) \eta(x_1, x_2) \, dx,$$

where $\eta(x_1, x_2) = \tilde{\eta}(x_1, x_2 - \psi_\varepsilon(x_1))$, with $\tilde{\eta} \in C^\infty_0(\mathbb{R}^2)$ supported in a sufficiently small neighborhood $\tilde{U} \subset V$ of $x_0$. The corresponding operators $R^{-\theta} A_t^\theta R^\theta$ are then given by

$$R^{-\theta} A_t^\theta R^\theta f(y) = \int_{\mathbb{R}^2} f \left( y_1 + t \frac{1}{\cos(\theta_0 + \theta)} (1 + \varepsilon \Phi_\varepsilon(\theta, s)), y_2 - ts, y_3 \right) a(\theta, s, \varepsilon) \, ds,$$

with

$$a(\theta, s, \varepsilon) := \eta(x_1(\theta, s, \varepsilon), s) |J(\theta, s, \varepsilon)| = \tilde{\eta} \left( x_1^0(\theta) + g_\varepsilon(s), s - \psi_\varepsilon(x_1^0(\theta) + g_\varepsilon(s)) \right) |J(\theta, s, \varepsilon)|.$$

The substitution $s \mapsto s + \psi_\varepsilon(x_1^0(\theta))$ in the integral thus leads to

$$R^{-\theta} A_t^\theta R^\theta f(y) = \int_{\mathbb{R}^2} f \left( y_1 + t \frac{1}{\cos(\theta_0 + \theta)} (1 + \varepsilon \tilde{\Phi}_\varepsilon(\theta, s)), y_2 - t(s + \psi_\varepsilon(x_1^0(\theta))), y_3 \right) \tilde{a}(\theta, s, \varepsilon) \, ds,$$

with $\tilde{\Phi}_\varepsilon(\theta, s) := \tilde{\phi} \left( x_1^0(\theta) + \tilde{g}_\varepsilon(s), s + \psi_\varepsilon(x_1^0(\theta)) - \psi_\varepsilon(x_1^0(\theta) + \tilde{g}_\varepsilon(s)) \right)$ and

$$\tilde{a}(\theta, s, \varepsilon) := \tilde{\eta} \left( x_1^0(\theta) + \tilde{g}_\varepsilon(s), s + \psi_\varepsilon(x_1^0(\theta)) - \psi_\varepsilon(x_1^0(\theta) + \tilde{g}_\varepsilon(s)) \right) |J(\theta, s + x_1^0(\theta), \varepsilon)|,$$
where we have set $\tilde{g}_\varepsilon(s) := g_\varepsilon(s + \psi_\varepsilon(x^0_1(\theta)))$. From (3.25) and (3.28) it is clear that $\tilde{g}_\varepsilon = O(\varepsilon)$ in $C^\infty$ and $
abla \psi_\varepsilon(x^0_1(\theta)) - \psi_\varepsilon(x^0_1(\theta) + \tilde{g}_\varepsilon(s)) = O(\varepsilon^{1-\delta})$ in $C^\infty$.

Consequently, $\tilde{a}$ is supported in $V_1$, if $\varepsilon$ and $\theta$ are sufficiently small, and $\tilde{a} = O(1)$ in $C^\infty$. In a similar way, we see that

$$\tilde{\Phi}_\varepsilon(\theta, s) = \tilde{\phi}(x^0_1(\theta), s) + \tilde{\phi}_r(\theta, s, \varepsilon),$$

where the perturbation term $\tilde{\phi}_r(\theta, s, \varepsilon)$ can be made small in $C^\infty$ by choosing $\varepsilon$ and $\theta$ sufficiently small.

Notice finally that for $\varepsilon < 1$,

$$|\psi_\varepsilon(x^0_1(\theta))| \lesssim \varepsilon^{-\delta} \leq \varepsilon^{-1}.$$

We can therefore apply the maximal theorem for curves, Corollary 3.4, to each operator $R^{-\theta}A^\theta_R R^{-\theta}$ and obtain

**Corollary 3.6.** Let $V$ be an open neighborhood of the point $x^0 \in \mathbb{R}^2$, and let $\tilde{\phi}_r \in C^\infty(V, \mathbb{R})$ be such that

$$\partial^m_2 \tilde{\phi}_r(x^0_1, x^0_2) \neq 0,$$

where $m \geq 2$. Let

$$\tilde{\phi} := \tilde{\phi}_r + \tilde{\phi}_r,$$

where $\tilde{\phi}_r \in C^\infty(V, \mathbb{R})$ is sufficiently small, and assume that $\psi_\varepsilon \in C^\infty(V_1)$ satisfies (3.25) for some $0 \leq \delta < 1$. Put $\tilde{\phi}_\varepsilon(x_1, x_2) := \tilde{\phi}(x_1, x_2 - \psi_\varepsilon(x_1))$ and $\eta(x_1, x_2) = \tilde{\eta}(x_1, x_2 - \psi_\varepsilon(x_1))$, with $\tilde{\eta} \in C^\infty_0(\mathbb{R}^2)$ supported in a sufficiently small neighborhood $U \subset V$ of $x^0$, and consider the averaging operators $A_t$ given by (3.29), with associated maximal operator $M^\varepsilon$.

Assume that the neighborhood $\bar{U}$ of the point $x^0$ is sufficiently small. Then there exist numbers $M \in \mathbb{N}$, $\delta_1 > 0$, such that for every $\tilde{\phi}_r \in C^\infty(\bar{U}, \mathbb{R})$ with $\|	ilde{\phi}_r\|_{C^M} < \delta_1$ and any $p > m$ there exists a positive constant $C_p$ such that for $\varepsilon > 0$ sufficiently small the maximal operator $M^\varepsilon$ satisfies the following a priori estimate:

$$\|M^\varepsilon f\|_p \leq C_p \varepsilon^{-1/p}\|f\|_p, \quad f \in S(\mathbb{R}^3).$$

**4. Auxiliary statements on the multiplicity of roots at a critical point of a mixed homogeneous polynomial function**

We refer in this section to the definitions and results in [11]. We begin by recalling the following structural statements on mixed homogeneous polynomials.

Let $P \in \mathbb{R}[x_1, x_2]$ be a mixed homogeneous polynomial, and assume that $\nabla P(0, 0) = 0$. Following [12], we denote by

$$m(P) := \text{ord}_{S^1} P$$

the maximal order of vanishing of $P$ along the unit circle $S^1$ centered at the origin.

If $m_1, \ldots, m_n$ are positive integers, then we denote by $(m_1, \ldots, m_n)$ their greatest common divisor.
Proposition 4.1. Let $P$ be a $(\kappa_1, \kappa_2)$-homogeneous polynomial of degree one, and assume that $P$ is not of the form $P(x_1, x_2) = cx_1^{\nu_1} x_2^{\nu_2}$. Then $\kappa_1$ and $\kappa_2$ are uniquely determined by $P$, and $\kappa_1, \kappa_2 \in \mathbb{Q}$.

Let us assume that $\kappa_1 \leq \kappa_2$, and write $\kappa_1 = q/m$, $\kappa_2 = p/m$, $(p, q, m) = 1$, so that in particular $p \geq q$. Then $(p, q) = 1$, and there exist non-negative integers $\alpha_1, \alpha_2$ and a $(1, 1)$-homogeneous polynomial $Q$ such that the polynomial $P$ can be written as

$$P(x_1, x_2) = x_1^{\alpha_1} x_2^{\alpha_2} Q(x_1^{p/q}, x_2^{q/p}).$$

More precisely, $P$ can be written in the form

$$P(x_1, x_2) = cx_1^{\nu_1} x_2^{\nu_2} \prod_{l=1}^{M} (x_2^{q} - \lambda_l x_1^{p})^{n_l},$$

with $M \geq 1$, distinct $\lambda_l \in \mathbb{C} \setminus \{0\}$ and multiplicities $n_l \in \mathbb{N} \setminus \{0\}$, with $\nu_1, \nu_2 \in \mathbb{N}$ (possibly different from $\alpha_1, \alpha_2$ in (4.1)).

Let us put $n := \sum_{l=1}^{M} n_l$. The distance $d(P)$ of $P$ can then be read off from (4.2) as follows:

If the principal face of $N(P)$ is compact, then it lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$, and the distance is given by

$$d(P) = \frac{1}{\kappa_1 + \kappa_2} = \frac{\nu_1 q + \nu_2 p + pq n}{q + p}.$$

Otherwise, we have $d(P) = \max\{\nu_1, \nu_2\}$. In particular, in any case we have $d(P) = \max\{\nu_1, \nu_2, \frac{1}{\kappa_1 + \kappa_2}\}$.

The proposition shows that every zero (or “root”) $(x_1, x_2)$ of $P$ which does not lie on a coordinate axis is of the form $x_2 = \lambda_1^{1/p} x_1^{p/q}$. The quantity

$$d_h(P) = \frac{1}{\kappa_1 + \kappa_2}$$

will be called the homogeneous distance of the mixed homogeneous polynomial $P$. Recall that $(d_h(P), d_h(P))$ is just the point of intersection of the bisectrix with the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$ on which the Newton diagram $\mathcal{N}_d(P)$ lies. Moreover,

$$d_h(P) \leq d(P).$$

Notice also that

$$m(P) = \max\{\nu_1, \nu_2, \max_{l=1, \ldots, M} n_l\}.$$

In view of the homogeneity of $P$, we shall often restrict our considerations to roots lying on the unit circle. For the next result, compare [11], Corollary 2.3 and Corollary 3.4.
Corollary 4.2. Let $P$ be a $(\kappa_1, \kappa_2)$-homogeneous polynomial of degree one as in Proposition 4.1, and consider the representation (4.2) of $P$. We put again $n := \sum_{i=1}^{M} n_i$.

(a) If $\kappa_2/\kappa_1 \notin \mathbb{N}$, i.e., if $q \geq 2$, then $n < d_h(P)$. In particular, every real root $x_2 = \lambda_1^{1/q} x_1^{p/q}$ of $P$ has multiplicity $n_i < d_h(P)$.

(b) If $\kappa_2/\kappa_1 \in \mathbb{N}$, i.e., if $q = 1$, then there exists at most one real root of $P$ on the unit circle $S^1$ of multiplicity greater than $d_h(P)$. More precisely, if we put $n_0 := \nu_1, n_{M+1} := \nu_2$ and choose $l_0 \in \{0, \ldots, M + 1\}$ so that $n_0 = \max_{l=0,\ldots,M+1} n_l$, then $n_l \leq d_h(P)$ for every $l \neq l_0$.

(c) The height of the Newton polyhedron of $P$ is given by

$$h(P) = \max\{m(P), d_h(P)\}.$$  

In particular, we see that the multiplicity of every real root of $P$ not lying on a coordinate axis is bounded by the distance $d(P)$, unless $q = 1$, in which case there can at most be one real root $x_2 = \lambda_0 x_1^1$ with multiplicity exceeding $d(P)$. If such a root exists, we shall call it the principal root of $P$.

The next proposition will allow us to apply Proposition 3.5 respectively Corollary 3.6 in many situations.

Proposition 4.3. Let $P$ be a $(\kappa_1, \kappa_2)$-homogeneous polynomial of degree one such that $\nabla P(0) = 0$ and $\kappa_2/\kappa_1 > 2$, and assume that $\partial_x^2 P$ does not vanish identically. If $x^0 \in S^1$, then denote by $m_2(x^0)$ the order of vanishing of $\partial_x^2 P$ along $S^1$ in the point $x^0$. By $\mathcal{R}$ we shall denote the set of all roots of $\partial_x^2 P$ on the unit circle which do not lie on the $x_2$-axis.

(a) Assume that $p := \kappa_2/\kappa_1 \in \mathbb{N}$, so that $q = 1$ and $p \geq 3$, and that the set $\mathcal{R}$ is non-empty. Let then $x^m \in \mathcal{R}$ be a root of maximal multiplicity $m_2(x^m) \geq 1$ among all roots in $\mathcal{R}$. Then, for any other root $x^0 \neq x^m$ in $\mathcal{R}$, we have $m_2(x^0) \leq d_h(P) - 2$.

In particular, for every point $x \in S^1$ such that $x_1 \neq 0$ and $x \neq x^m$ there exists some $j$ with $2 \leq j \leq d_h(P)$ such that $\partial_x^2 P(x) \neq 0$.

(b) Assume that $p := \kappa_2/\kappa_1 \in \mathbb{N}$, so that $q = 1$ and $p \geq 3$, and that $P$ vanishes along $S^1$ of order $\nu_2 = d(P)$ in the point $e := (1,0)$ on the $x_1$-axis. Moreover, assume that $d(P) > d_h(P)$ and $d(P) > 2$. Then $m_2(x^0) \leq d_h(P) - 2$ for every $x^0 \in \mathcal{R}$ such that $x^0_{2} \neq 0$.

In particular, for every point $x \in S^1$ which does not lie on a coordinate axis, there exists some $j$ with $2 \leq j \leq d_h(P)$ such that $\partial_x^2 P(x) \neq 0$.

(c) If $\kappa_2/\kappa_1 \notin \mathbb{N}$, then $m_2(x^0) \leq d_h(P) - 2$ for every root with $x^0_1 \neq 0 \neq x^0_2$, unless the polynomial $P$ is of the form

$$(4.4) \quad P(x_1, x_2) = c(x_2^2 - \lambda_1 x_1^2)(x_2^2 - \lambda_2 x_1^2),$$

with $\lambda_1 + \lambda_2 \in \mathbb{R} \setminus \{0\}$ and $\lambda_1 \lambda_2 \in \mathbb{R}$.

In particular, for every point $x \in S^1$ which does not lie on a coordinate axis, there exists some $j$ with $2 \leq j \leq d_h(P)$ such that $\partial_x^2 P(x) \neq 0$, unless $P$ is of the form (4.4).
Remark. In case (a), if $m(P) > d(P)$, so that $P$ has a (unique) principal root $x^p \in S^1$, then $x^m = x^p$.

Proof. By our assumptions, $\partial_2^2 P(x)$ is a $\sigma$- homogeneous polynomial of degree one with respect to the weight

$$
\sigma_1 := \frac{\kappa_1}{1 - 2\kappa_2}; \quad \sigma_2 := \frac{\kappa_2}{1 - 2\kappa_2}.
$$

According to the Proposition 4.1, we can write the polynomial $\partial_2^2 P(x)$ in the form

$$
\partial_2^2 P(x_1, x_2) = x_1^{\nu_1} x_2^{\nu_2} Q_2(x_1^p, x_2^q),
$$

where $p$ and $q$ are coprime, $Q_2$ is a homogeneous polynomial of degree $n_2$, and

$$
\frac{p}{q} = \frac{\kappa_2}{\kappa_1} = \frac{\sigma_2}{\sigma_1} \geq 2.
$$

We shall also assume that no power of $x_2^q$ can be factored from $Q_2(x_1^p, x_2^q)$, so that we have

$$
(4.5) \quad \sigma_1 = \frac{q}{\nu_1 q + \nu_2 p + pqn_2}, \quad \sigma_2 = \frac{p}{\nu_1 q + \nu_2 p + pqn_2}.
$$

We begin with the case $\kappa_2/\kappa_1 \notin \mathbb{N}$, i.e., $q \geq 2$. Recall that we then assume that $x_1^0 \neq 0 \neq x_2^0$, so that in particular $m_2(x_0) \leq n_2$. Let us first consider the case $\nu_1 + \nu_2 \geq 1$. In this case we show that the assumption $m_2(x_0) > d_h(P) - 2$ cannot hold. For, otherwise we had

$$
\frac{1 + 2\sigma_2}{\sigma_1 + \sigma_2} = \frac{1}{\kappa_1 + \kappa_2} = d_h(P) < n_2 + 2,
$$

which by (4.5), and since $q \geq 2$, is equivalent to

$$
n_2 < \frac{2q - (\nu_1 q + \nu_2 p)}{(p - 1)q - p}.
$$

Since $\nu_1 q + \nu_2 p \geq (\nu_1 + \nu_2)q \geq q$, we then would get

$$
n_2 < \frac{q}{(p - 1)q - p}.
$$

And, straightforward computations show that for any $p \geq 2q$, $q \geq 2$, we have $\frac{q}{(p - 1)q - p} \leq 1$, so that necessarily $n_2 = 0$, i.e., $\partial_2^2 P(x_1, x_2) = cx_1^{\nu_1} x_2^{\nu_2}$, which would contradict the existence of a root away from the coordinates axes.

Assume next that $\nu_1 = \nu_2 = 0$, so that the assumption $m_2(x_0) > d_h(P) - 2$ implies the inequality

$$
(4.6) \quad 1 \leq n_2 < \frac{2q}{(p - 1)q - p}.
$$

Since $q \geq 2$, we get $p \geq 4$, and then $\frac{3q}{q} > p \geq 2q$, hence $2q < 5$, so that $q = 2$. Then (4.6) implies $n_2 = 1$ and $p = 4$, or $p = 5$. Since we assume that $p$ and $q$ are coprime,
the only possibility that remains is that \( q = 2, p = 5, n_2 = 1 \), so that \( \partial_2^2 P \) will be of the form

\[
\partial_2^2 P(x_1, x_2) = c(x_2^2 - ax_1^5).
\]

Integrating the last polynomial twice with respect to \( x_2 \), and observing that \( P \) must be \((\frac{1}{10}, \frac{1}{4})\)-homogeneous, we can apply Proposition 4.1 and obtain (4.4).

The remaining claim in case (c) is now evident.

Consider next the case \( q = 1 \). Let us put \( N := \nu_2 + n_2 \).

We first prove (a). If \( x^0 \) is a root different from \( x^m \) in \( \mathcal{R} \), then \( 1 \leq m_2(x^0) \leq m_2(x^m) \), and so we have

\[
2m_2(x^0) \leq m_2(x^0) + m_2(x^m) \leq N,
\]

hence in particular \( N \geq 2 \). Assume now that \( m_2(x^0) > d_h(P) - 2 \). Then

\[
\frac{1 + 2\sigma_2}{\sigma_1 + \sigma_2} = d_h(P) < \frac{N}{2} + 2,
\]

and in view of (4.5), one computes that \( N < 2\frac{\nu_1 - 1}{p-1} \). Because of \( N \geq 2 \), this implies \( p < 3 - \nu_1 \), contradicting our assumption \( p \geq 3 \).

Let us prove the statement of the Remark at this point. So, assume that \( m(P) > d(P) \), so that \( P \) has a (unique) principal root \( x^p \in S^1 \) of multiplicity \( m(x^p) = m(P) \).

If \( m(P) \geq 3 \), then \( x^p \in \mathcal{R} \), with multiplicity \( m_2(x^p) = m(P) - 2 > d(P) - 2 \geq d_h(P) - 2 \), so that, by (i), we must have \( x^p = x^m \), and the conclusion in (ii) is obvious.

Assume finally that \( m(P) \leq 2 \). Then \( d(P) < 2 \), hence \( \frac{1 + 2\sigma_2}{\sigma_1 + \sigma_2} < 2 \), which implies \( \nu_1 + pN \leq 1 \). Consequently, we have \( N = 0 \) and \( \nu_1 \leq 1 \), so that \( P \) would be a polynomial of degree at most one, hence \( \partial_2^2 P \) would vanish identically. This shows that this case actually cannot arise.

What remains to be proven is (b). So, assume that \( \nu_2 = d(P) > 2 \). Then \( \partial_2^2 P \) vanishes of order \( d(P) - 2 \geq 1 \) in the point \( e \), i.e., \( m_2(e) = d(P) - 2 \). Let \( x^0 \) be any root of \( \partial_2^2 P \) with \( x_1^0 \neq x_2^0 \). We want to show that \( m_2(x^0) \leq d_h(P) - 2 \).

Assume to the contrary that \( m_2(x^0) > d_h(P) - 2 \).

If \( m_2(x^0) < m_2(e) \), then

\[
2m_2(x^0) < m_2(e) + m_2(x^0) \leq N,
\]

and we obtain \( d_h(P) < \frac{N}{2} + 2 \).

If \( m_2(x^0) \geq m_2(e) \), then

\[
2m_2(e) \leq m_2(e) + m_2(x^0) \leq N,
\]

hence \( d(P) \leq \frac{N}{2} + 2 \). But \( d_h(P) < d(P) \), so that we have again \( d_h(P) < \frac{N}{2} + 2 \).

As in the proof of (a), this leads to a contradiction.

Q.E.D.
5. Estimation of the maximal operator $\mathcal{M}$ when the coordinates are adapted or the height is strictly less than 2

We now turn to the proof of our main result, Theorem 1.2. As observed in the Introduction, we may assume that $S$ is locally the graph $S = \{(x_1, x_2, 1 + \phi(x_1, x_2)) : (x_1, x_2) \in \Omega\}$ of a function $1 + \phi$. Here and in the subsequent sections, $\phi \in C^\infty(\Omega)$ will be a smooth real valued function of finite type defined on an open neighborhood $\Omega$ of the origin in $\mathbb{R}^2$ and satisfying

$$\phi(0, 0) = 0, \nabla \phi(0, 0) = 0.$$  

In this section we shall consider the easiest cases where the coordinates $x$ are adapted to $\phi$, or where $h(\phi) < 2$.

Recall that $A_t, t > 0$, denotes the corresponding family of averaging operators

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

where $d\sigma$ denotes the surface measure on $S$ and $\rho \in C^\infty_c(S)$ is a non-negative cut-off function. We shall assume that $\rho$ is supported in an open neighborhood $U$ the point $(0, 0, 1)$ which will be chosen sufficiently small. The associated maximal operator is given by

$$(5.1) \quad \mathcal{M} f(y) := \sup_{t > 0} |A_t f(y)|, \quad (y \in \mathbb{R}^3).$$

The averaging operator $A_t$ can be re-written in the form

$$A_t f(y) := \int_{\mathbb{R}^2} f\left(y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \phi(x_1, x_2))\right) \eta(x_1, x_2) \, dx,$$

where $\eta$ is a smooth function supported in $\Omega$. If $\chi$ is any integrable function defined on $\Omega$, we shall denote by $A^\chi_t$ the correspondingly localized averaging operator

$$A^\chi_t f(y) := \int_{\mathbb{R}^2} f\left(y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \phi(x_1, x_2))\right) \chi(x) \eta(x) \, dx,$$

and by $\mathcal{M}^\chi$ the associated maximal operator

$$\mathcal{M}^\chi f(y) := \sup_{t > 0} |A^\chi_t f(y)|, \quad (y \in \mathbb{R}^3).$$

**Proposition 5.1.** Let $\phi$ be as above, and assume that $\kappa = (\kappa_1, \kappa_2)$ is a given weight such that $0 < \kappa_1 \leq \kappa_2 < 1$. As in (2.1), we decompose

$$\phi = \phi_\kappa + \phi_r$$

into its $\kappa$-principal part $\phi_\kappa$ and the remainder term $\phi_r$ consisting of terms of $\kappa$-degree $> 1$. Then, if the neighborhood $\Omega$ of the point $(0, 0)$ is chosen sufficiently small, the maximal operator $\mathcal{M}$ is bounded on $L^p(\mathbb{R}^3)$ for every $p > \max\{2, h(\phi_\kappa)\}$. 

Proof. Let us modify our notation slightly and write points in $\mathbb{R}^3$ in the form $(x, x_3)$, with $x \in \mathbb{R}^2$ and $x_3 \in \mathbb{R}$. Recall from Corollary 4.2 the crucial fact that

$$h(\phi_\kappa) = \max\{m(\phi_\kappa), d_h(\phi_\kappa)\}.$$ 

In particular, the multiplicity of every real root of the $\kappa$-homogeneous polynomial $\phi_\kappa$ is bounded by $h(\phi_\kappa)$.

Consider then the dilations $\delta_r(x_1, x_2) := (r^{\kappa_1}x_1, r^{\kappa_2}x_2)$, $r > 0$. We choose a smooth non-negative function $\chi$ supported in the annulus $D := \{1 \leq |x| \leq R\}$ satisfying

$$\sum_{k=k_0}^{\infty} \chi_k(x) = 1 \quad \text{for} \quad 0 \neq x \in \Omega,$$

where $\chi_k(x) := \chi(\delta_{2^k}x)$. Notice that by choosing $\Omega$ small, we can choose $k_0 \in \mathbb{N}$ large. Assuming that $\Omega$ is sufficiently small, we can write $A_t$ as a sum of averaging operators

$$A_t f(y, y_3) = \sum_{k=0}^{\infty} A^k_t f(y, y_3),$$

where $A^k_t := A^1_t \chi_k$. If we apply the change of variables $x \mapsto \delta_{2^{-k}}(x)$ in the integral above, we obtain

$$A^k_t f(y, y_3) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(y - t\delta_{2^{-k}}(x), y_3 - t(1 + 2^{-k}\phi^k(x))\right) \eta(\delta_{2^{-k}}(x)) \chi(x) \, dx,$$

where

$$\phi^k(x) := \phi_\kappa(x) + 2^k \phi_\kappa(\delta_{2^{-k}}(x))$$

and where the perturbation term $2^k \phi_\kappa(\delta_{2^{-k}}(\cdot))$ is of order $O(2^{-\varepsilon k})$ for some $\varepsilon > 0$ in any $C^M$-norm. To express this fact, we shall in the sequel again use the short-hand notation

$$2^k \phi_\kappa(\delta_{2^{-k}}(\cdot)) = O(2^{-\varepsilon k}).$$

By $\mathcal{M}^k$ we shall denote the maximal operator $\mathcal{M}^{\chi_k}$ associated to the averaging operators $A^k_t$.

Assume now that $p > \max\{2, h(\phi_\kappa)\}$. We define the scaling operator $T^k$ by

$$T^k f(y, y_3) := 2^{\frac{|\kappa|}{p}} f(\delta_{2^k}(y), y_3).$$

Then $T^k$ acts isometrically on $L^p(\mathbb{R}^3)$, and

$$(T^{-k} A^k_t T^k) f(y, y_3) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(y - tx, y_3 - t(1 + 2^{-k}\phi^k(x))\right) \eta(\delta_{2^{-k}}(x)) \chi(x) \, dx.$$

Assuming that $\Omega$ is a sufficiently small neighborhood of the origin, we need to consider only the case when $k$ is sufficiently large.

Let $x^0 \in D$ be a fixed point.

If $\nabla \phi_\kappa(x^0) \neq 0$, then from by Euler’s homogeneity relation one easily derives that

$$\text{rank } (D^2 \phi_\kappa(x^0)) \geq 1 \quad (\text{see} \ [12], \ \text{Lemma} \ 3.3).$$

Therefore, we can find a unit vector $e \in \mathbb{R}^2$ such that $\partial^2_{\kappa e} \phi_\kappa(x^0) \neq 0$, where $\partial_{\kappa e}$ denotes the partial derivative in direction of $e$. 

If $\nabla \phi_k(x^0) = 0$, then by Euler’s homogeneity relation we have $\phi_k(x^0) = 0$ as well. Thus the function $\phi_k$ vanishes in $x^0$ at least of order two, so that $m(\phi_k) \geq 2$, hence $h(\phi_k) \geq 2$. On the other hand, by what we remarked earlier, it vanishes along the circle passing through $x^0$ and centered at the origin at most of order $h(\phi_k)$. Therefore, we can find a unit vector $e \in \mathbb{R}^2$ such that $\partial^m_x \phi_k(x^0) \neq 0$, for some $m$ with $2 \leq m \leq h(\phi_k)$.

Thus, in both cases, after rotating coordinates so that $e = (0,1)$, we may apply Proposition 3.5 to conclude that for $p > \max\{2, h(\phi_k)\}$ and sufficiently large $k$,

$$\|T^{-k} A^k T^k f\|_p \leq C 2^{k(\frac{1}{p} - |\kappa|)} \|f\|_p, \quad f \in \mathcal{S}(\mathbb{R}^3),$$

if we replace $\chi$ in the definition of $A^k_t$ by $\chi \eta$, where $\eta$ is a bump function supported in a sufficiently small neighborhood of $x^0$. This is equivalent to

$$\|M^k f\|_p \leq C 2^{k(\frac{1}{p} - |\kappa|)} \|f\|_p.$$

Decomposing $\chi$ and correspondingly $A^k_t$ by means of a suitable partition of unity into a finite number of such pieces, we see that the same estimate holds for the original operators $M^k$.

Since $\frac{1}{|\kappa|} = d_h(\phi_k) \leq h(\phi_k) < p$, we can sum over all $k \geq k_0$ and obtain the desired estimate for $M$.

Q.E.D.

Let us apply this result first to the case where the coordinates $x$ are adapted to $\phi$, possibly after a rotation of the coordinate system $(x_1, x_2)$. Observe first that a linear change of the coordinates $(x_1, x_2)$ induces a corresponding linear change of coordinates in $\mathbb{R}^3$ which fixes the coordinate $x_3$. This linear transformation is an automorphism of $\mathbb{R}^3$, so that it preserves the convolution product on $\mathbb{R}^3$ (up to a fixed factor), hence the norm of the maximal operator $M$. We may thus assume that the coordinates are adapted to $\phi$.

We shall also assume that non-negative numbers $\kappa_1, \kappa_2$ with $|\kappa| := \kappa_1 + \kappa_2 > 0$ are chosen so that the principal face $\pi(\phi)$ of the Newton polyhedron $\mathcal{N}(\phi)$ of $\phi$ lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. Notice that the weight $\kappa := (\kappa_1, \kappa_2)$ is then determined uniquely, unless $\pi(\phi)$ is a single point. Without loss of generality, as in [11] we shall assume $\kappa_2 \geq \kappa_1$.

Recall from Corollaries 4.3 and 5.2 in [11], that the coordinates $x$ are adapted to $\phi$ if and only if the principal face $\pi(\phi)$ of the Newton polyhedron $\mathcal{N}(\phi)$ satisfies one of the following conditions:

(a) $\pi(\phi)$ is a compact edge, and either $\frac{\kappa_2}{\kappa_1} \notin \mathbb{N}$, or $\frac{\kappa_2}{\kappa_1} \in \mathbb{N}$ and $m(\phi_p) \leq d(\phi)$.

(b) $\pi(\phi)$ consists of a vertex.

(c) $\pi(\phi)$ is unbounded.

Moreover, in this case we have $h(\phi) = h(\phi_p) = d(\phi_p)$.

In the sequel, we shall often refer to these cases as the cases (a) to (c) without further mentioning.
Corollary 5.2. Let \( \phi \) be as above, and assume that, possibly after a rotation of the coordinate system, the coordinates \( x \) are adapted to \( \phi \), i.e., that \( h(\phi) = d(\phi) \). Then, if the neighborhood \( \Omega \) of the point \((0,0)\) is chosen sufficiently small, the maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > \max\{2, h(\phi)\} \).

Proof. As mentioned before, we may assume that the coordinates are adapted to \( \phi \). We begin with case (a), in which \( \phi_p = \phi_\kappa \). In particular, we have \( h(\phi) = h(\phi_\kappa) \). Observe also that \( \kappa_j < 1 \) for \( j = 1, 2 \), since \( \nabla \phi(0) = 0 \), so that \( 0 < \kappa_1 \leq \kappa_2 < 1 \). The result is thus an immediate consequence of Proposition 5.1.

Consider next the case (b). If \( \pi(\phi) \) consists of a vertex \((N, N)\), then \( h(\phi) = N \geq 1 \). Moreover, by perturbing \( \kappa \) slightly, we may assume that the line \( \kappa_1 t_1 + \kappa_2 t_2 = 1 \) intersects \( \mathcal{N}(\phi) \) only in the point \((N, N)\) and that \( 0 < \kappa_1 \leq \kappa_2 < 1 \), so that again \( \phi_p = \phi_\kappa \). In particular, \( h(\phi) = h(\phi_\kappa) \), and we can now argue exactly as in the case (a).

There remains the case (c). Here, the principal face \( \pi(\phi) \) is a horizontal half-line, with left endpoint \((\nu_1, N)\), where \( \nu_1 < N = h(\phi) \). Notice that \( N \geq 2 \), since for \( N = 1 \) we had \( \nu_1 = 0 \), which is not possibly given our assumption \( \nabla \phi(0, 0) = 0 \). We can then choose \( \kappa \) with \( 0 < \kappa_1 < \kappa_2 \) so that the line \( \kappa_1 t_1 + \kappa_2 t_2 = 1 \) is a supporting line to the Newton polyhedron of \( \phi \) and that the point \((\nu_1, N)\) is the only point of \( \mathcal{N}(\phi) \) on this line (we just have to choose \( \kappa_2/\kappa_1 \) sufficiently large!). Then necessarily \( \kappa_2 < 1 \), and the \( \kappa \)-principal part \( \phi_\kappa \) of \( \phi \) is of the form \( \phi_\kappa(x) = c x_1^{\nu_1} x_2^{N} \), with \( c \neq 0 \). Since the coordinates are clearly also adapted to \( \phi_\kappa \), we find that \( h(\phi) = N = d(\phi_\kappa) = h(\phi_\kappa) \).

The result thus follows again from Proposition 5.1, with \( \kappa \) replaced by \( \kappa \).

Q.E.D.

Remark 5.3. One can easily extend Corollary 5.2 as follows:

If the neighborhood \( \Omega \) of the point \((0,0)\) is chosen sufficiently small, then the maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > \max\{2, h(\phi_p)\} \), no matter if the coordinates are adapted to \( \phi \) or not.

Proof. Indeed, if the coordinates are adapted, then \( h(\phi_p) = h(\phi) \). So, assume that the coordinates \((x_1, x_2)\) are not adapted to \( \phi \). Then the principal face of the Newton polyhedron is a compact edge, so that the principal part \( \phi_\kappa \) of \( \phi \) is \( \kappa \)-homogeneous, where \( \kappa \) satisfies the assumptions of Proposition 5.1. Since \( \phi_p = \phi_\kappa \), the result then follows this proposition.

Q.E.D.

The result above holds even when the coordinates are not adapted, but it will then in general not be sharp, since we have \( h(\phi_p) \geq h(\phi) \) (see [11], Corollary 4.3), and in general strict inequality holds.

For example, let \( \phi(x_1, x_2) := (x_2 - x_1^3)^2 + x_1^5 \). Then we have \( \phi_p(x) = (x_2 - x_1^3)^2 \). The coordinate system is not adapted to \( \phi \), because \( d(\phi) = 4/3 < 2 \), where 2 is the multiplicity of the root of \( \phi_p \). A coordinate system which is adapted to \( \phi \) is given by \( y_2 := x_2 - x_1^3 \) and \( y_1 := x_1 \). It is then easy to see that \( h(\phi_p) = 2 > \frac{10}{3} = h(\phi) \).
**Corollary 5.4.** If \( h(\phi) < 2 \), and if the neighborhood \( \Omega \) of the point \((0, 0)\) is chosen sufficiently small, then the maximal operator \( \mathcal{M} \) is bounded on \( L^p(\mathbb{R}^3) \) for any \( p > 2 \), also when the coordinates are not adapted to \( \phi \).

**Proof.** If \( D^2 \phi(0, 0) \neq 0 \) then we have at least one non-vanishing principal curvature at the origin, so that the result follows from C.D. Sogge’s main theorem in [26].

Next, we consider the case where \( D^2 \phi(0, 0) = 0 \). Then necessarily we have \( D^3 \phi(0, 0) \neq 0 \), for otherwise \( h(\phi) \geq d(\phi) \geq 2 \). In particular, \( h(\phi) > 1 \). Denote by \( P_3 \) the Taylor polynomial of degree 3 with base point 0 of the function \( \phi \). If \( h(P_3) \leq 2 \), then we obtain the desired estimate from Corollary 5.2, which \( \kappa := (\frac{1}{3}, \frac{1}{3}) \). Assume therefore that \( h(P_3) > 2 \). Then, by Corollary 4.2 (c), \( P_3 \) must have a root of order 3. Thus, possibly after rotating the coordinate system, we may assume that \( P_3(x_1, x_2) = cx_2^3 \) with \( c \neq 0 \).

Now, we consider the Taylor support \( T(\phi) \) of \( \phi \). Since \( T(\phi) \subset \{ \frac{t_1}{3} + \frac{t_2}{3} \geq 1 \} \), one checks easily that the subset
\[
\left\{ \frac{t_1}{6} + \frac{t_2}{3} < 1 \right\} \cap T(\phi)
\]
of \( T(\phi) \) contains at most 3 points, namely \((4, 0)\), \((5, 0)\), \((3, 1)\), all of them lying below the bisectrix \( t_1 = t_2 \).

Moreover, any line passing through the point \((0, 3)\) in \( T(\phi) \) corresponding to \( P_3 = cx_2^3 \) contains at most one of these points. Thus, if
\[
\left\{ \frac{t_1}{6} + \frac{t_2}{3} < 1 \right\} \cap T(\phi) \neq \emptyset,
\]
then the principal part \( \phi_p \) of \( \phi \) contains only two monomials, one corresponding to the point \((0, 3)\) above the bisectrix and the other one corresponding to one of the points listed above which lie below the bisectrix, i.e., \( \phi_p \) is of the form \( dx_1^3 + cx_2^3 \), \( dx_1^3 + cx_2^3 \) or \( dx_1^3 x_2 + cx_2^3 \), with \( d \neq 0 \) (note that these all satisfy \( d(\phi_p) < 2 \)). Therefore on the unit circle it has no root of multiplicity bigger than one, so that the coordinate system is adapted to \( \phi \), and thus \( h(\phi) < 2 \). The desired estimate for \( \mathcal{M} \) follows therefore in this case from Proposition 5.1.

Assume finally that
\[
\left\{ \frac{t_1}{6} + \frac{t_2}{3} < 1 \right\} \cap T(\phi) = \emptyset.
\]
Then \( T(\phi) \subset \{ \frac{t_1}{3} + \frac{t_2}{3} \geq 1 \} \), hence \( h(\phi) \geq d(\phi) \geq 2 \), which contradicts to our assumption.

**Q.E.D.**

In view of this result, we shall from now on always assume that
\[
h(\phi) \geq 2.
\]
6. Estimation of the maximal operator $M$ away from the principal root jet

Let $\phi \in C^\infty(\Omega)$ be as in Section 5, and assume now that the coordinates $x$ are not adapted to $\phi$. Recall from [30] in the analytic case (under some non-degeneracy condition), and from [11] in the general case that in this situation there exists a smooth function $\sigma$ which defines an adapted coordinate system

$$z_1 := x_1, \quad z_2 := x_2 - \sigma(x_1).$$

for the function $\phi$. In these coordinates, $\phi$ is given by

$$\phi^a(z) := \phi(z_1, z_2 + \sigma(z_1)).$$

Consider the Taylor expansion

$$\sigma(x_1) = \sum_{l=1}^{K} b_l x_1^{m_l} \quad (1 \leq K \leq \infty)$$

of $\sigma$, where we assume that $b_l \in \mathbb{R} \setminus \{0\}$ for every $l$, and where the $m_l \in \mathbb{N}$ form a strictly increasing sequence $1 \leq m_1 < m_2 < \cdots$.

Such a function $\sigma$ can be constructed by means of Varchenko’s algorithm [30] (see also [11]), and if $\phi$ is real-analytic, one obtains it in an explicit way from the Puiseux series expansion of the roots of $\phi$ as the principal root jet (see [11]). In the sequel we shall indeed assume that $\sigma$ is constructed by this algorithm. In particular, if this algorithm stops after finitely many steps, then $K$ coincides with this finite number of steps. This happens in particular if the principal face of the Newton polyhedron of $\phi^a(z)$ is compact.

The goal of this section is to prove that the main contribution to the maximal operator will be given by a small neighborhood a modified, polynomial curve $x_2 = \psi(x_1)$, of the form

$$|x_2 - \psi(x_1)| \leq \varepsilon_0 x_1^a,$$

where $\psi$ will be a suitable polynomial approximation to $\sigma$ of sufficiently high degree, and where $a \geq \deg \psi$.

We shall then often work in the coordinates $y$ given by

$$y_1 := x_1, \quad y_2 := x_2 - \psi(x_1)$$

for the function $\phi$. In these coordinates, $\phi$ is given by

$$\tilde{\phi}(y) := \phi(y_1, y_2 + \psi(y_1)).$$

To this end, we shall decompose $\Omega$ into various regions adapted to the roots of $\phi$ and estimate the contributions of these regions to the maximal operator $M$ separately.

We first make the simple observation that in case that $m_1 = 1$, the linear change of coordinates

$$y_1 := x_1, \quad y_2 := x_2 - b_1 x_1^{m_1}$$
allows to reduce to the case $m_1 \geq 2$. Observe to this end that the corresponding linear change of coordinates of $\mathbb{R}^3$, with $y_3 := x_3$, is an automorphism of $\mathbb{R}^3$, so that it preserves the convolution product on $\mathbb{R}^3$ (up to a fixed factor).

In the sequel, we shall therefore always assume that
\begin{equation}
2 \leq m_1 < m_2 < \cdots .
\end{equation}
We shall also only consider the region where $x_1 > 0$ in order to simplify the notation. The remaining half-plane can be treated in the same way.

In order to construct the polynomial $\psi$, notice that one of the cases (a) - (c) described after the proof of Proposition 5.1 applies to $\phi^a$ (in place of $\phi$), since the coordinates $z$ are adapted to $\phi^a$. We shall construct $\psi$ and at the same time a weight $\tilde{\kappa}$ satisfying
\begin{equation}
1/|\tilde{\kappa}| \leq h(\phi).
\end{equation}

Let us begin with case (a), in which the principal face $\pi(\tilde{\phi})$ is a compact edge, and the principal part $\phi^a_p$ is, say, $\tilde{\kappa} = (\tilde{\kappa}_1, \tilde{\kappa}_2)$-homogenous of degree one. Observe that, by Varchenko’s algorithm, $K < \infty$ and $a_p := \frac{\tilde{\kappa}_1}{\kappa_1} > m_K \geq m_1 \geq 2$. In this case, we shall put
\begin{equation}
\psi(x_1) := \sigma(x_1) + c_p x_1^{a_p},
\end{equation}
where the constant $c_p$ will be chosen as follows:

If $a_p \notin \mathbb{N}$, then we put $c_p := 0$. And, if $a_p \in \mathbb{N}$, then, according to Proposition 4.3, there exists a unique real constant $c_p$ such that $c_p z_1^{a_p}$ is a real root of maximal multiplicity of the $\tilde{\kappa}$-homogeneous polynomial $\partial^2 \phi^a_p(z)$.

Observe that the $\tilde{\kappa}$-homogeneous change of coordinates $y_1 := z_1, y_2 := z_2 - c_p z_1^{a_p}$ will again lead to adapted coordinates, and has the effect of modifying the coefficients of the roots of $\partial^2 \phi^a_p(z)$ in such a way that the root of maximal multiplicity will be given by $y_2 = 0$. We shall therefore define $\psi$ in case (a) by
\begin{equation}
\psi(x_1) := \sum_{l=1}^{K} b_l x_1^{m_l} + c_p x_1^{a_p}.
\end{equation}
Notice that the principal face $\pi(\tilde{\phi})$ is a compact edge in this case, and $h(\phi) = 1/|\tilde{\kappa}|$.

In case (b), the principal face $\pi(\phi^a)$ is a vertex, say $(N, N)$. Then $N = h(\phi) \geq 2$, and $\phi^a_p = c z_1^N z_2^N$. In this case, we choose for $\tilde{\kappa}$ any rational pair $0 < \tilde{\kappa}_1 < \tilde{\kappa}_2$ such that the line $\tilde{\kappa}_1 t_1 + \tilde{\kappa}_2 t_2 = 1$ is a supporting line to the Newton polyhedron $\mathcal{N}(\phi^a)$ of $\phi^a$ containing only the point $(N, N)$ from $\mathcal{N}(\phi^a)$. Then clearly $\phi^a_{\tilde{\kappa}} = \phi^a_p$ and $h(\phi) = \frac{1}{|\tilde{\kappa}|}$. Moreover, again $K < \infty$, and we define in this case
\begin{equation}
\psi(x_1) := \sum_{l=1}^{K} b_l x_1^{m_l} = \sigma(x_1).
\end{equation}
Notice that here $\tilde{\phi} = \phi^a$, and that the principal face $\pi(\tilde{\phi})$ is a vertex.
Consider finally case (c), in which the principal face \( \pi(\phi^a) \) is unbounded and possibly \( K = \infty \). As Varchenko’s algorithm shows, then \( \pi(\phi^a) \) is in fact a horizontal half-line, with left endpoint \((\nu_1, N)\), where \( \nu_1 < N = h(\phi) \). In this case, we shall put

\[
\psi(x_1) := \sum_{l=1}^{L} b_l x_1^{m_l},
\]

where \( L := K \), if \( K < \infty \), and where otherwise \( L \) will be chosen sufficiently large. Indeed, if \( K = \infty \), then the algorithm shows that there is a finite number of steps \( L_0 \) such that for every \( L \geq L_0 \), the principal part \( \tilde{\phi}_p \) of \( \phi \), when expressed in the coordinates \( y \) given by (6.3), is of the form

\[
\tilde{\phi}_p(y) = c_L y_1^{\nu_1}(y_2 - b_{L+1} y_1^{m_{L+1}})^N.
\]

The polynomial \( \tilde{\phi}_p \) is \( \tilde{\kappa} := (\frac{1}{\nu_1 + m_{L+1} N}, \frac{m_{L+1} N}{\nu_1 + m_{L+1} N}) \)-homogenous of degree one, where \( 1/|\tilde{\kappa}| \leq N = h(\phi) \).

Finally, if \( K < \infty \), we shall choose \( \tilde{\kappa} \) in the same way as in case (b) (for instance, we could choose it as for the case \( K = \infty \)), where we choose any sufficiently large number \( m_{L+1} \). Notice that in this case \( \tilde{\phi}_\kappa \) is of the form \( cy_1^{\nu_1} y_2^N \), and it may not coincide with the principal part \( \tilde{\phi}_p \).

In all three cases (a)-(c), we shall put

\[
a := \frac{\tilde{\kappa}_2}{\tilde{\kappa}_1} > m_1.
\]

Observe that then \( a > \deg \psi \), except for the case (a), when \( a = a_p \in \mathbb{N} \) and \( c_p \neq 0 \), where \( a = \deg \psi \). Moreover, in case (a) we have \( \tilde{\phi}_\kappa = \tilde{\phi}_p \), whereas in the cases (b) and (c) \( \tilde{\phi}_\kappa \) is of the form

\[
\tilde{\phi}_\kappa(y) = cy_1^{\nu_1}(y_2 - by_1^a)^N,
\]

with \( b \in \mathbb{R} \) and \( N \) as before. Finally, clearly (6.5) holds true in all three cases (a)-(c).

We next fix a cut-off function \( \rho \in C_0^\infty(\mathbb{R}) \) supported in a neighborhood of the origin such that \( \rho = 1 \) near the origin, and put

\[
\rho_0(x_1, x_2) := \rho\left(\frac{x_2 - \psi(x_1)}{\varepsilon_0 x_1^a}\right).
\]

Define averaging operators

\[
A_1^{1-\rho_0} f(y) := \int_{\mathbb{R}^2} \left( f(y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \phi(x_1, x_2)) \right) \left( 1 - \rho\left(\frac{x_2 - \psi(x_1)}{\varepsilon_0 x_1^a}\right) \right) \eta(x) \, dx,
\]

and consider the associated maximal operator \( M_1^{1-\rho_0} \). We shall then prove

**Proposition 6.1.** If the neighborhood \( \Omega \) of the point \((0, 0)\) is chosen sufficiently small, then the maximal operator \( M_1^{1-\rho_0} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > h(\phi) \).
6.1. Preliminary reduction to a $\kappa$-homogeneous neighborhood of the principal root $x_2 = b_1 x_1^{m_1}$ of $\phi_p$. Recall that since the coordinates $x$ are not adapted to $\phi$, the principal face $\pi(\phi)$ must be a compact edge of the Newton polyhedron of $\phi$, so that it lies on a unique line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. Again, we may assume that $\kappa_2 \geq \kappa_1$. Then, by the results in [11],
\[
\frac{\kappa_2}{\kappa_1} = m_1 \geq 2.
\]
Moreover, if $\phi_p = \phi_\kappa$ denotes the principal part of $\phi$, we must have $m(\phi_p) > d(\phi_p)$, and $m(\phi_p)$ is just the multiplicity of the principal root $b_1 x_1^{m_1}$ of the $\kappa$-homogeneous polynomial $\phi_p$. All other roots have multiplicity less or equal to $d(\phi_p)$.

This already indicates that the function $\phi$ will indeed be small of ”highest order” (in some averaged sense) near the curve $x_2 = \sigma(x_1)$ given by the principal root jet (even though $\phi$ need not vanish on this curve!), so that the region close to this curve should indeed give the main contribution to the maximal operator.

In order to localize to a $\kappa$-homogeneous region away from the principal root jet, put, in a first step,
\[
\rho_1(x_1, x_2) := \rho\left(\frac{x_2 - b_1 x_1^{m_1}}{\varepsilon_1 x_1^{m_1}}\right),
\]
where $\varepsilon_1 > 0$ is a small parameter to be determined later, and set
\[
A_t^{-\rho_1} f(y) := \int_{\mathbb{R}^2} f\left(y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \phi(x_1, x_2))\right) \left(1 - \rho\left(\frac{x_2 - b_1 x_1^{m_1}}{\varepsilon_1 x_1^{m_1}}\right)\right) \eta(x) \, dx.
\]
By $\mathcal{M}_t^{-\rho_1}$ we denote the associated maximal operator. We can now argue exactly as in the proof of Proposition 5.1. Using the dilations $\delta_r(x_1, x_2) = (r^{\kappa_1} x_1, r^{\kappa_2} x_2)$, $r > 0$, we can dyadically decompose the operators $A_t^{-\rho_1}$ into the sum of operators $A_t^k$, which, after re-scaling, are given by
\[
(T^{-k} A_t^k T^k) f(y, y_3) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(y_1 - tx_1, y_2 - tx_2, y_3 - t(1 + \phi^k(x_1, x_2))\right) \left(1 - \rho\left(\frac{x_2 - b_1 x_1^{m_1}}{\varepsilon_1 x_1^{m_1}}\right)\right) \eta(\delta_2^{-k} x) \chi(x) \, dx.
\]
All roots of $\phi_p$ lying in the domain of integration have a positive distance to the principal root $b_1 x_1^{m_1}$, hence have multiplicities bounded by the distance $d(\phi_p)$ (cf. Corollary 4.2), so that we can again estimate the associated maximal operators $\mathcal{M}_t^k$ by means of Proposition 3.5 (applied possibly in a rotated coordinate system) and obtain

**Lemma 6.2.** If the neighborhood $\Omega$ of the point $(0, 0)$ is chosen sufficiently small, then the maximal operator $\mathcal{M}_t^{-\rho_1}$ is bounded on $L^p(\mathbb{R}^2)$ for every $p > h(\phi)$.

We have thus reduced considerations to a narrow $\kappa$-homogeneous domain near the curve $x_2 = b_1 x_1^{m_1}$, of the form
\[
|x_2 - b_1 x_1^{m_1}| \leq \varepsilon_1 x_1^{m_1},
\]
where $\varepsilon_1 > 0$ can be chosen arbitrarily small.

6.2. The roots of $\tilde{\phi}$. For our further reduction, we need more information on $\tilde{\phi}$. Let us assume for a while that $\phi$ is real analytic (in Subsection 6.6 we shall explain how the general case can be reduced to the analytic setting). According to [11] and following [21], we may then write

$$\tilde{\phi}(y_1, y_2) = U(y_1, y_2)^{\nu_1} y_1^{\nu_2} \prod_{l=1}^{n} \Phi \left( \frac{\cdot}{l} \right)(y_1, y_2),$$

where $U(0, 0) \neq 0$ and

$$\Phi \left( \frac{\cdot}{l} \right)(y_1, y_2) := \prod_{r \in \left[ \frac{\cdot}{l} \right]} (y_2 - r(y_1)).$$

The roots $r(y_1)$ arising in this display can be expressed in a small neighborhood of 0 as Puiseux series

$$r(y_1) = c_{l_1}^{\alpha_1} y_1^{a_1} + c_{l_1 l_2}^{\alpha_1 \alpha_2} y_1^{a_1 a_2} + \cdots + c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_p} y_1^{a_1 \cdots a_p} + \cdots,$$

where

$$c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1} \beta} \neq c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1} \gamma} \quad \text{for} \quad \beta \neq \gamma,$$

with strictly positive exponents $a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}} > 0$ and non-zero complex coefficients $c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_p} \neq 0$, and where we have kept enough terms to distinguish between all the non-identical roots of $\tilde{\phi}$.

The cluster $\begin{bmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \cdots & l_p \end{bmatrix}$ designates all the roots $r(y_1)$, counted with their multiplicities, which satisfy

$$(6.11) \quad r(y_1) = c_{l_1}^{\alpha_1} y_1^{a_1} + c_{l_1 l_2}^{\alpha_1 \alpha_2} y_1^{a_1 a_2} + \cdots + c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_p} y_1^{a_1 \cdots a_p} = O(y_1^b)$$

for some exponent $b > a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}}$. We also introduce the clusters

$$\begin{bmatrix} \alpha_1 & \cdots & \alpha_{p-1} \\ l_1 & \cdots & l_{p-1} & l_p \end{bmatrix} := \bigcup_{\alpha_p} \begin{bmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \cdots & l_p \end{bmatrix}.$$

Each index $\alpha_p$ or $l_p$ varies in some finite range which we shall not specify here. We finally put

$$N \begin{bmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \cdots & l_p \end{bmatrix} := \text{number of roots in} \begin{bmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \cdots & l_p \end{bmatrix},$$

$$N \begin{bmatrix} \alpha_1 & \cdots & \alpha_{p-1} \\ l_1 & \cdots & l_{p-1} & l_p \end{bmatrix} := \text{number of roots in} \begin{bmatrix} \alpha_1 & \cdots & \alpha_{p-1} \\ l_1 & \cdots & l_{p-1} & l_p \end{bmatrix}.$$
Let $a_1 < \cdots < a_l < \cdots < a_n$ be the distinct leading exponents of all the roots $r$. Each exponent $a_l$ corresponds to the cluster $\text{l}$, so that the set of all roots $r$ can be divided as $\bigcup_{l=1}^{n} \text{l}$.

We introduce the following quantities:

\begin{align*}
A_l &= A \left[ \text{l} \right] := \nu_1 + \sum_{\mu \leq l} a_{\mu} N \left[ \text{\mu} \right], \\
B_l &= B \left[ \text{l} \right] := \nu_2 + \sum_{\mu \geq l+1} N \left[ \text{\mu} \right].
\end{align*}

Then the vertices of the Newton diagram $\mathcal{N}(\tilde{\phi})$ of $\tilde{\phi}$ are the points $(A_l, B_l)$, $l = 0, \ldots, n$, and the Newton polyhedron $\mathcal{N}(\tilde{\phi})$ is the convex hull of the set $\bigcup_{l} (A_l, B_l) + \mathbb{R}^2_+$. 

Let $L_l := \{(t_1, t_2) \in \mathbb{N}^2 : \kappa_l^1 t_1 + \kappa_l^2 t_2 = 1\}$ denote the line passing through the points $(A_{l-1}, B_{l-1})$ and $(A_l, B_l)$. Then

\[ \frac{\kappa_l^2}{\kappa_l^1} = a_l, \]

which in return is the reciprocal of the slope of the line $L_l$. The line $L_l$ intersects the bisectrix at the point $(d_l, d_l)$, where

\[ d_l := \frac{A_l + a_l B_l}{1 + a_l} = \frac{A_{l-1} + a_l B_{l-1}}{1 + a_l}. \]

Finally, the $\kappa_l$-principal part $\tilde{\phi}_{\kappa_l}$ of $\tilde{\phi}$ corresponding to the supporting line $L_l$ is given by

\[ \tilde{\phi}_{\kappa_l}(y) = c_l y_1^{A_l-1} y_2^{B_l} \prod_{\alpha} \left( y_2 - c_l \alpha y_1^{a_l} \right)^{N \left[ \alpha / \text{l} \right]} . \]

In view of this identity, we shall say that the edge $\gamma_l := [(A_{l-1}, B_{l-1}), (A_l, B_l)]$ is associated to the cluster of roots $\text{l}$.

Now, in case (a) where the principal face of $\tilde{\phi}$ is a compact edge, we choose $\lambda$ so that the edge $\gamma_{\lambda} = [(A_{\lambda-1}, B_{\lambda-1}), (A_{\lambda}, B_{\lambda})]$ is the principal face $\pi(\tilde{\phi})$ of the Newton polyhedron of $\tilde{\phi}$. Then

\[ \tilde{\kappa} = \kappa_{\lambda}, \quad a_{\lambda} = a = \frac{\tilde{\kappa}_2}{\tilde{\kappa}_1} . \]

In the case (b), where $\pi(\tilde{\phi})$ is a vertex $(A_{\lambda-1}, B_{\lambda-1})$ with $A_{\lambda-1} = B_{\lambda-1} = N = h(\phi)$ (which one may view as the limiting case of case (a) after shrinking the principal edge $\gamma_{\lambda}$ to this single point), and also in case (c), where $\pi(\tilde{\phi})$ is a horizontal half-line, or a
compact edge, with left endpoint \((A_{\lambda-1}, B_{\lambda-1})\) such that \(B_{\lambda-1} = N = h(\phi)\), we shall slightly abuse our previous notation and define

\[ \kappa_{\lambda} := \kappa, \quad a_{\lambda} := a = \frac{\kappa_{\lambda}}{\kappa_{1}}. \]

Note that \(a_{\lambda} = a\) may then possibly not be the reciprocal of the slope of an edge of \(N(\tilde{\phi})\).

In the sequel, it will be better to work with the \(\kappa\)-principal part \(\tilde{\phi}_{\kappa}\) of \(\tilde{\phi}\) in place of the principal part \(\tilde{\phi}_{p}\). The following observation will become useful.

**Lemma 6.3.** If \(h(\phi) \geq 2\), then \(\partial_2^2 \tilde{\phi}_{\kappa_{l}}\) does not vanish identically, and \(\kappa_{l} < 1\), for any \(l \leq \lambda\).

**Proof.** Consider first the situation where the principal face \(\pi(\tilde{\phi})\) is an edge. Here the statements will already follow from our general assumption \(\nabla \phi(0) = 0\). Indeed, write \(\tilde{\phi}_{\kappa_{l}}(y) = cy_{1}^{\nu_{1}}y_{2}^{\nu_{2}} \prod_{s=1}^{M}(y_{2} - \lambda_{s}y_{1}^{a_{s}})^{n_{s}}\).

with \(\lambda_{s} \neq 0\), where then \(M \geq 1\).

If we assume that \(\partial_2^2 \tilde{\phi}_{\kappa_{l}} = 0\), then clearly \(\nu_{2} + \sum_{s} n_{s} \leq 1\), so that there is only one, real root \(\lambda_{1}y_{1}^{a_{1}}\) of multiplicity one. This implies that \(\tilde{\phi}_{\kappa_{l}}(y) = cy_{1}^{\nu_{1}}(y_{2} - \lambda_{1}y_{1}^{a_{1}})\). Thus the Newton diagram \(\gamma_{l} = N_{d}(\tilde{\phi}_{\kappa_{l}})\) is the interval \([\nu_{1}, 1), (\nu_{1} + a_{1}, 0)]\). Since \(l \leq \lambda\), its left endpoint must lie above the bisectrix, so that \(\nu_{1} = 0\). But then \(\nabla \tilde{\phi}_{\kappa_{l}}(0) \neq 0\), hence \(\nabla \phi(0) \neq 0\), a contradiction.

A similar argument applies to show that \(\kappa_{l} < 1\). Indeed, since the polynomial \(\tilde{\phi}_{\kappa_{l}}\) is \(\kappa_{l}\)-homogeneous of degree one, and since \(M \geq 1\), \(\kappa_{2} \geq 1\) would imply that \(\nu_{1} = \nu_{2} = 0\) and \(\sum_{s} n_{s} = 1\), so that we could conclude as before that \(\nabla \phi(0) \neq 0\).

Finally, if \(\pi(\tilde{\phi})\) is a vertex or an unbounded edge, then the previous arguments still apply for \(l < \lambda\). And, for \(l = \lambda\), by (6.10) the polynomial \(\tilde{\phi}_{\kappa_{\lambda}}\) is of the form \(cy_{1}^{\nu_{1}}(y_{2} - by_{1}^{a_{1}})^{N}\), with \(\nu_{1} \leq N = h(\phi) \geq 2\), so that the statements are obvious.

Q.E.D.

6.3. **Further domain decompositions.** We have seen that we can control the maximal operator associated to sub-domains

\[ |x_{2} - b_{1}x_{1}^{m_{1}}| \geq \varepsilon_{1}x_{1}^{m_{1}} \]

of \(\Omega\), where \(\varepsilon_{1} > 0\) can be chosen arbitrarily small. Since \(m_{1}\) is the leading exponent of \(\psi\), choosing \(\Omega\) sufficiently small we see that we can reduce our considerations to a domain of the form

\[ |x_{2} - \psi(x_{1})| \leq \varepsilon_{1}x_{1}^{m_{1}}. \]
This domain, except for a small $\kappa^\lambda$-homogeneous neighborhood of the principal root jet of the form $|x_2 - \psi(x_1)| \leq \varepsilon x_1^{a_0}$, will be decomposed into domains $D_l$ of the form

$$D_l := \{ \varepsilon x_1^{a_l} < |x_2 - \psi(x_1)| \leq N_l x_1^{a_l} \}, \quad l = l_0, \ldots, \lambda,$$

which, when expressed in terms of the coordinates $y$, are $\kappa^l$-homogeneous, and the intermediate domains

$$E_l := \{ N_{l+1} x_1^{a_{l+1}} < |x_2 - \psi(x_1)| \leq \varepsilon x_1^{a_l} \}, \quad l = l_0, \ldots, \lambda - 1,$$

and

$$E_{l_0-1} := \{ N_{l_0} x_1^{a_{l_0}} < |x_2 - \psi(x_1)| \leq \varepsilon x_1^{m_1} \}.$$

Here, the $\varepsilon_l > 0$ are arbitrarily small and the $N_l > 0$ are arbitrarily large parameters, and $l_0 \geq 1$ is chosen such that

$$a_l \leq m_1 \text{ for } l < l_0 \text{ and } a_l > m_1 \text{ for } l \geq l_0. \quad (6.14)$$

To localize to domains of type $D_l$, we put

$$\rho_l(x_1, x_2) := \rho \left( \frac{x_2 - \psi(x_1)}{N_l x_1^{a_l}} \right) - \rho \left( \frac{x_2 - \psi(x_1)}{\varepsilon x_1^{a_l}} \right), \quad l = l_0, \ldots, \lambda,$$

and set

$$A_l^\rho f(z) := \int_{\mathbb{R}^2} f(z_1 - tx_1, z_2 - tx_2, z_3 - t(1 + \phi(x_1, x_2))) \rho_l(x) \eta(x) \, dx,$$

with associated maximal operator $M_l^\rho$.

Similarly, in order to localize to domains of type $E_l$, we put

$$\tau_l(x_1, x_2) := \rho \left( \frac{x_2 - \psi(x_1)}{\varepsilon x_1^{a_l}} \right) \left( 1 - \rho \right) \left( \frac{x_2 - \psi(x_1)}{N_{l+1} x_1^{a_{l+1}}} \right), \quad l = l_0, \ldots, \lambda - 1,$$

and

$$\tau_{l_0-1}(x_1, x_2) := \rho \left( \frac{x_2 - \psi(x_1)}{\varepsilon x_1^{m_1}} \right) \left( 1 - \rho \right) \left( \frac{x_2 - \psi(x_1)}{N_{l_0} x_1^{a_{l_0}}} \right),$$

and set

$$A_l^\tau f(z) := \int_{\mathbb{R}^2} f(z_1 - tx_1, z_2 - tx_2, z_3 - t(1 + \phi(x_1, x_2))) \tau_l(x) \eta(x) \, dx,$$

with associated maximal operator $M_l^\tau$.

Notice that it suffices to control all the maximal operators defined in this way in order to prove Proposition 6.1.
6.4. The maximal operators \( \mathcal{M}^p \).

**Lemma 6.4.** If the neighborhood \( \Omega \) of the point \((0,0)\) is chosen sufficiently small, then the maximal operator \( \mathcal{M}^p \) is bounded on \( L^p(\mathbb{R}^2) \) for every \( p > h(\phi) \).

**Proof.** 1) We begin with the special case \( l = \lambda \), where \( k_\lambda = \tilde{\kappa} \).

The change of variables (6.3) transforms the integral for \( A^p_\lambda f(z) \) into

\[
A^p_\lambda f(z) = \int_{\mathbb{R}^2} f\left(z_1 - ty_1, z_2 - t(y_2 + \psi(y_1)), z_3 - t(1 + \tilde{\phi}(y_1, y_2))\right) \tilde{\rho}_\lambda(y) \tilde{\eta}(y) \, dy,
\]

with

\[
\tilde{\rho}_\lambda(y) := \rho\left(\frac{y_2}{N_\lambda y_1^\alpha} - \rho\left(\frac{y_2 - c_\lambda^\alpha y_1^\alpha}{\varepsilon y_1^\alpha}\right)\right),
\]

and \( \tilde{\eta}(y) := \eta(y_1, y_2 + \psi(y_1)) \). Since \( \tilde{\phi}_{\tilde{\kappa}} \) is \( \tilde{\kappa} \)-homogeneous of degree one and \( \tilde{\rho}_\lambda \) is \( \tilde{\kappa} \)-homogeneous of degree zero with respect to the new dilations \( \tilde{\delta}_r(y_1, y_2) := \delta_{r\tilde{\kappa}}(y_1, y_2) := (r\tilde{\kappa}_1 y_1, r\tilde{\kappa}_2 y_2), \) \( r > 0 \), using these dilations we now dyadically decompose the operators \( A^p_\lambda \) into the sum of operators \( A^k \), with associated maximal operators \( \mathcal{M}^k \), given by

\[
A^k f(z) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(z_1 - t2^{-\tilde{\kappa}_1 k} y_1, z_2 - t(2^{-\tilde{\kappa}_2 k} y_2 + \psi(2^{-\tilde{\kappa}_1 k} y_1)), z_3 - t(1 + 2^{-k} \tilde{\phi}^k(y_1, y_2))\right) \tilde{\rho}_\lambda(y) \tilde{\eta}(\tilde{\delta}_{2^{-k}} y) \chi(y) \, dy,
\]

with

\[
\tilde{\phi}^k(y) := \tilde{\phi}_{\tilde{\kappa}}(y) + 2^k \tilde{\phi}_{\tilde{\delta}_{2^{-k}}} y.
\]

Notice that \( 2^k \tilde{\phi}_{\tilde{\delta}_{2^{-k}}} y = O(2^{-\varepsilon k}) \) in \( C^\infty \), for some \( \varepsilon > 0 \), so that this term can be considered as a perturbation term. Re-scaling by means of the operators

\[
\tilde{T}^k f(z, z_3) := 2^{\frac{k|\kappa|}{p}} f(\tilde{\delta}_{2^k}(z), z_3),
\]

we obtain

\[
(\tilde{T}^{-k} A^k \tilde{T}^k) f(z) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(z_1 - ty_1, z_2 - t(y_2 + \psi^k(y_1)), z_3 - t(1 + 2^{-k} \tilde{\phi}^k(y_1, y_2))\right) \tilde{\rho}_\lambda(y) \tilde{\eta}(\tilde{\delta}_{2^{-k}} y) \chi(y) \, dy,
\]

where by our construction of \( \psi(x_1) = b_1 x_1^{m_1} + \cdots \)

\[
\psi^k(y_1) := 2^{\tilde{\kappa}_2 k} \psi(2^{-\tilde{\kappa}_1 k} y_1) = O(2^{(\tilde{\kappa}_2 - \tilde{\kappa}_1 m_1)k}) \text{ in } C^\infty.
\]

Applying the change of variables \( x_1 := y_1, \) \( x_2 := y_2 + \psi^k(y_1) \) in this integral, we eventually arrive at

\[
(\tilde{T}^{-k} A^k \tilde{T}^k) f(z) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} f\left(z_1 - tx_1, z_2 - tx_2, z_3 - t(1 + 2^{-k} \tilde{\phi}^k(x_1, x_2 - \psi^k(x_1)))\right) \tilde{\rho}_\lambda^k(x) \chi^k(x) \eta(\tilde{\delta}_{2^{-k}} x) \, dx,
\]

(6.15)
with \( \chi^k(x) := \chi(x_1, x_2 - \psi^k(x_1)) \) and \( \rho^k(x) := \tilde{\rho}_\lambda(x_1, x_2 - \psi^k(x_1)) \). Since \( \tilde{\kappa}_2 - \tilde{\kappa}_1 m_1 = \tilde{\kappa}_1(a - m_1) > 0 \) (compare (6.9)), we can no longer argue with Proposition 3.5 as in the previous cases in order to estimate the corresponding maximal operators. However, we shall see that we can make use of Corollary 3.6 in combination with Proposition 4.3.

Notice that, according to Lemma 6.3, \( \partial^2_2 \tilde{\phi}_p \) does not vanish identically. We shall prove that for every point \( y^0 \) in the support of \( \tilde{\rho}_\lambda \chi \) the following holds true:

\[
(6.16) \quad \text{There exists some } j \text{ with } 2 \leq j \leq h(\phi) \text{ such that } \partial^j_2 \tilde{\phi}_\kappa(y^0) \neq 0.
\]

To prove this, consider first the case (a), where \( \tilde{\phi}_\kappa = \tilde{\phi}_p \). If \( a_p \in \mathbb{N} \), then by (6.9), we have \( \frac{\tilde{\kappa}_2}{\tilde{\kappa}_1} \geq 3 \). Recall also that we have changed coordinates in such a way that the root of maximal multiplicity of \( \partial^2_2 \tilde{\phi}_p(y) \) away from the \( y_2 \)-axis is given by \( y_2 = 0 \).

Our claim therefore follows from Proposition 4.3 (a) applied to \( \tilde{\phi}_p \), since \( d(\tilde{\phi}_p) \leq h(\phi) \) (notice that the support of \( \tilde{\rho}_\lambda \chi \) has positive distance to the root \( y_2 = 0 \) of \( \tilde{\phi}_p \)).

On the other hand, if \( a_p \notin \mathbb{N} \), then we can argue in the same way as before, by applying Proposition 4.3 (b) in place of Proposition 4.3 (a), unless \( \tilde{\phi}_p \) is one of the exceptional polynomials \( P \) given by (4.4).

In fact, these exceptional polynomials are \( \tilde{\kappa} \)-homogeneous of degree one, with \( \tilde{\kappa}_1 := \frac{1}{10} \) and \( \tilde{\kappa}_2 := \frac{1}{4} \), so that here \( h(\phi) = 1/|\tilde{\kappa}| = 20/7 \). Then, necessarily \( K = 1 \) and \( m_1 = 2 \), so that \( \psi(x_1) = b_1 x_1^2 \), and clearly \( a_p = 5/2 \). Moreover,

\[ \partial^2_2 P(y) = 12c(y_2^2 - \frac{\lambda_1 + \lambda_2}{6} y_1^5). \]

Thus, if \( \lambda_1 + \lambda_2 < 0 \), then we can argue as before and see that even the maximal operator associated to the domain \( |x_2 - \psi(x_1)| \leq N_\lambda x_1^{5/2} \), for any \( N_\lambda > 0 \), is bounded on \( L^p \) for \( p > h(\phi) \).

On the other hand, if \( \lambda_1 + \lambda_2 > 0 \), then \( \partial^2_2 \tilde{\phi}_p \) will have real roots given by \( y_2 = \pm \sqrt{\frac{\lambda_1 + \lambda_2}{6}} y_1^{5/2} \). In this case, with the same technics we can still reduce to small neighborhoods of these roots, which, in our original coordinates \( x \), are of the form

\[
(6.17) \quad \left| x_2 - (b_1 x_1^2 \pm \sqrt{\frac{\lambda_1 + \lambda_2}{6}} x_1^{5/2}) \right| \leq \varepsilon_0 x_1^{5/2},
\]

where \( \varepsilon_0 \) can be chosen as small as we wish.

These remaining domains will be treated in Subsection 7.2.

In the cases (b) and (c), according to (6.10) we can write \( \tilde{\phi}_\kappa \) in the form

\[ \tilde{\phi}_\kappa(y) = cy_1^{\nu_1}(y_2 - by_1^{m_1})^N, \]

with \( \nu_1 \leq N := h(\phi) \), where \( b = 0 \), \( \nu_1 = N \) in the case (b). It follows that (6.16) holds true with \( j = h(\phi) \).

Notice that in these cases, this remains true even for \( y^0 \) lying on the \( y_2 \)-axis!
Our claim is thus proved. Observe also that, if we put \( \varepsilon := 2^{-k} \) and \( \psi_\varepsilon(x_1) := \psi(x_1) \) in Corollary 3.6, then, by (6.15), \( \psi_\varepsilon = O(\varepsilon^{-\delta}) \) in \( C^\infty \), with \( 0 < \delta := \kappa_2 - \kappa_1 m_1 < 1 \).

We can therefore apply Corollary 3.6 to the maximal operators \( \mathcal{M}^k \) and obtain the estimate
\[
||\mathcal{M}^k f||_p \leq C 2^{-|k|^{1/k} + \frac{1}{p}} ||f||_p.
\]
Since \( p > h(\phi) \geq \frac{1}{|\kappa|} \) (compare (6.5)), these estimates sum in \( k \) and we obtain the desired estimate for the maximal operator \( \mathcal{M}^\lambda \). Notice here that the term \( \eta(\delta_{2-k} x) \) in the integral above causes no problem in the application of Corollary 3.6, since \( 2^{-\kappa_2 k} \psi^k(x_1) \) is small.

II) We now turn to the case \( l_0 \leq l \leq \lambda - 1 \). We can here follow the arguments in case i) up to formula (6.15) almost verbatim, if we replace \( \kappa \) by \( \kappa' \), the function \( \tilde{\rho}_\lambda \) by the \( \kappa' \)-homogeneous function
\[
\tilde{\rho}_l(y) := \rho\left(\frac{y_2}{N_l y_1^{\kappa'_1}}\right) - \rho\left(\frac{y_2}{\varepsilon_l y_1^{\kappa'_1}}\right),
\]
the dilations \( \tilde{\delta}_r \) by the dilations \( \delta'_r(x_1, x_2) := (r^{\kappa'_1} x_1, r^{\kappa'_2} x_2) \) and the function \( \tilde{\phi}_l \) by the \( \kappa' \)-homogeneous part
\[
\tilde{\phi}_l := \tilde{\phi}_{\kappa'}
\]
of \( \tilde{\phi} \). Notice that, again by Lemma 6.3, \( \partial^2_l \tilde{\phi}_l \) does not vanish identically and \( \kappa'_2 < 1 \). Moreover, because of (6.14) we then have
\[0 < \kappa'_1 (a_l - m_1) = \kappa'_2 - \kappa'_1 m_1 < 1 \text{ and } \frac{\kappa'_2}{\kappa'_1} = a_l > 2.\]
What remains to be shown in order to conclude as in the previous case \( l = \lambda \) is that given \( y^0 \) in the support of \( \tilde{\rho}_l \chi \), then there exists some \( 2 \leq j < h(\phi) \) such that \( \partial^j_l \tilde{\phi}_l(y^0) \neq 0 \). Notice that for roots in the support of \( \tilde{\rho}_l \chi \), we have \( y_2^0 \neq 0 \).

Now, from the geometry of the Newton polyhedron of \( \tilde{\phi} \), it is evident that \( d_h(\tilde{\phi}_l) \leq d(\tilde{\phi}_p) \leq h(\phi) \). It will therefore be sufficient to prove that
\[
(6.18) \quad \partial^j_l \tilde{\phi}_l(y^0) \neq 0 \text{ for some } 2 \leq j \leq d_h(\tilde{\phi}_l).
\]

i) Consider first the case where \( \frac{\kappa_1}{\kappa'_1} = a_l \notin \mathbb{N} \). Then, by Proposition 4.3 (c), (6.18) is true, unless \( \tilde{\phi}_l \) is the exceptional polynomial (4.4). But, in the latter case the Newton diagram \( N_d(\tilde{\phi}_l) \) of \( \tilde{\phi}_l \) would be the interval \([0, 4], [10, 0]\) which intersects the bisectrix, so that \( \tilde{\phi}_l \) would have to be the principal part of \( \tilde{\phi} \), contradiction our assumption \( l < \lambda \).

ii) Assume finally that \( \frac{\kappa_1}{\kappa'_1} = a_l \in \mathbb{N} \), so that \( a_l \geq 3 \). We first show that the root \( y_2 = 0 \) has maximal multiplicity \( B_l > d_h(\tilde{\phi}_l) \) among all real roots of \( \tilde{\phi}_l \) away from the \( y_2 \)-axis.

Indeed, since \( a_{l_0} > m_1 \), it is clear from Varchenko’s algorithm (see [11]) that the edge \( \tilde{\gamma}_l \) of the Newton polyhedron of \( \tilde{\phi} \) associated to the \( \kappa_l \)-homogeneous polynomial
\[ \tilde{\varphi}_l \] is an interval which is contained in an edge of the Newton polyhedrons arising in the course of this algorithm. More precisely, we must have \( a_i = m_k \) for some \( k < K \), if \( K < \infty \), or \( k < L \), if \( K = \infty \).

If then \( \phi^{(k-1)}(z_1, z_2) := \varphi(z_1, z_2 + \sum_{j=1}^{k-1} b_j z_1^{m_j}) \) is the function appearing in the \((k-1)\)st step of the algorithm, then \( b_k z_1^{m_k} = b_k z_1^{n_1} \) is the principal root of \((\varphi^{(k-1)})_p(z)\), which has multiplicity \( B_k > d_h(\phi^{(k-1)})_p \), since the coordinates \( z \) are not yet adapted. The next step in the algorithm, which changes the coordinates to \( y_z \), turns the root \( \gamma \) into the root \( \gamma \), still of multiplicity \( B_k \), of the \( \kappa_l \)-homogeneous polynomial \((\varphi^{(k-1)})_p(y_1, y_2 + b_k y_1^{m_k})\). But, in the subsequent steps of the algorithm, only terms of higher order than \( O(y_1^n) \) are added, so that clearly \( \tilde{\varphi}_l(y) = (\varphi^{(k-1)})_p(y_1, y_2 + b_k y_1^{m_k}) \), and \( y_2 = 0 \) is the real root of highest multiplicity \( B_k \) of \( \phi_l \).

Since the edge \( \gamma_l \) lies in the closed subspace above the bisectrix, we then conclude by means of Proposition 4.1 that in fact \( B_k = d(\tilde{\varphi}_l) > d_h(\tilde{\varphi}_l) \). Moreover, the left end point of this edge is of the form \((A_k, B_k)\), and since it belongs to the Newton diagram, but not to the principal face, of \( \tilde{\varphi} \), it is clear from the geometry of the Newton polyhedron that \( d(\tilde{\varphi}_l) = B_k > h(\tilde{\varphi}) = h(\tilde{\varphi}) \geq 2 \).

This shows that we can apply Proposition 4.3 (c) to \( \tilde{\varphi}_l \) and obtain (6.18).

Q.E.D.

6.5. The maximal operators \( \mathcal{M}^\tau \).

**Lemma 6.5.** If the neighborhood \( \Omega \) of the point \((0, 0)\) is chosen sufficiently small, then the maximal operator \( \mathcal{M}^\tau \) is bounded on \( L^p(\mathbb{R}^2) \) for every \( p > h(\tilde{\varphi}) \).

**Proof.**  1) We begin with the case \( l_0 \leq l \leq \lambda - 1 \). Since the domain \( E_l \), when viewed in \( y \)-coordinates, is a domain of transition between two different homogeneities, namely the ones given by the weights \( \kappa^l \) and \( \kappa^{l+1} \) (at least if \( l \geq 1 \)), we shall apply an idea from [21] and decompose it dyadically in each coordinate separately, and then re-scale each of the bi-dyadic pieces obtained in this way.

By the change of variables (6.3), we can write

\[ A^\tau_l f(z) = \int_{\mathbb{R}^2} f(z_1 - t y_1, z_2 - t (y_2 + \psi(y_1)), z_3 - t (1 + \tilde{\varphi}(y_1, y_2))) \tilde{\eta}(y) \tilde{\eta}(y) dy, \]

with

\[ \tilde{\eta}(y) := \rho \left( \frac{y_2}{\varepsilon_1 y_1^q} \right) (1 - \rho) \left( \frac{y_2}{N_{l+1} y_1^q} \right), \]

and \( \tilde{\eta}(y) := \eta(y_1, y_2 + \psi(y_1)) \).

Consider a dyadic partition of unity \( \sum_{k=0}^{\infty} \chi_k(s) = 1 \), \((0 < s < 1) \) on \( \mathbb{R} \), with \( \chi \in C^\infty_0 (\mathbb{R}) \) supported in the interval \([1/2, 4] \), where \( \chi_k(s) := \chi(2^k s) \), and put

\[ \chi_{j,k}(x) := \chi_j(x_1) \chi_k(x_2), \ j, k \in \mathbb{N}. \]
We then decompose $A^{j,k}_t$ into the operators

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

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

Notice that by choosing the neighborhood $\Omega$ of the origin sufficiently small, we need only consider sufficiently large $j, k$. Moreover, because of the localization imposed by $\tilde{\tau}$, it suffices to consider only pairs $(j, k)$ satisfying

$$a_t j + M \leq k \leq a_{t+1} j - M,$$

where $M$ can still be choosen sufficiently large, because we had the freedom to choose $\varepsilon_t$ sufficiently small and $N_{t+1}$ sufficiently large. In particular, we have $j \sim k$.

By re-scaling in the integral, we have

$$A^{j,k}_t f(z) = 2^{-j-k} \int_{\mathbb{R}^2} f \left( z_1 - t 2^{-j} y_1, z_2 - t(2^{-k} y_2 + \psi(2^{-j} y_1),
z_3 - t(1 + \tilde{\phi}(2^{-j} y_1, 2^{-k} y_2)) \right) \tilde{\tau}^{j,k}(y) \tilde{\eta}^{j,k}(y) \chi(y_1) \chi(y_2) \, dy,$$

with

$$\tilde{\tau}^{j,k}(y) := \rho \left( \frac{y_2}{\varepsilon_t 2^{-j} y_1 a_1} \right) (1 - \rho) \left( \frac{y_2}{N_{t+1} 2^{-k} y_1 a_1} \right), \quad \tilde{\eta}^{j,k}(y) := \tilde{\eta}(2^{-j} y_1, 2^{-k} y_2).$$

Notice that, by (6.19), all derivatives of $\tilde{\tau}^{j,k}$ are uniformly bounded in $j, k$.

The scaling operators

$$T^{j,k} f(z) := 2^{j+k} f(2^j z_1, 2^k z_2, z_3)$$

then transform these operators into

$$(T^{-j,-k} A^{j,k} T^{j,k}) f(z) = 2^{-j-k} \int_{\mathbb{R}^2} f \left( z_1 - t y_1, z_2 - t(y_2 + \psi^{j,k}(y_1),
\tilde{\phi}^{j,k}(y_1)) \right) \tilde{\tau}^{j,k}(y) \tilde{\eta}^{j,k}(y) \chi(y_1) \chi(y_2) \, dy,$$

where

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

Notice that

$$\psi^{j,k} = O(2^{-m_1 j}) \text{ in } C^\infty.$$

Applying the change of variables $x_1 := y_1, x_2 := y_2 + \psi^{j,k}(y_1)$ in this integral, we eventually arrive at

$$(T^{-j,-k} A^{j,k} T^{j,k}) f(z) = 2^{-j-k} \int_{\mathbb{R}^2} f \left( z_1 - t x_1, z_2 - t x_2,
\tilde{\phi}^{j,k}(x_1, x_2 - \psi^{j,k}(x_1)) \right) \tilde{\tau}^{j,k}(x) \tilde{\eta}^{j,k}(x) \chi^{j,k}(x) \, dx,$$

(6.20)
where
\[ \tau^{j,k}(x) := \tilde{\tau}^{j,k}(x_1, x_2 - \psi^{j,k}(x_1)), \quad \eta^{j,k}(x) := \eta(2^{-j}x_1, 2^{-k}x_2) \]
and
\[ \chi^{j,k}(x) := \chi(x_1) \chi(x_2 - \psi^{j,k}(x_1)). \]

We next determine \( \tilde{\phi}^{j,k} \), up to an error term. To this end, notice that if \( y_1 \sim 1 \) and \( y_2 \sim 1 \), and if \( r \in \begin{bmatrix} \cdot \mu \end{bmatrix} \), then
\[ r(2^{-j}y_1) = c_\mu^2 2^{-a_\mu j} y_1^{a_\mu} + O(2^{-\varepsilon(j+k)}) \text{ in } C^\infty, \text{ for some } \varepsilon > 0. \]
In view of (6.19), we thus get
\[ 2^{-k}y_2 - r(2^{-j}y_1) = \begin{cases} -c_\mu^2 2^{-a_\mu j} (y_1^{a_\mu} + O(2^{-\varepsilon(j+k)})), & \text{if } \mu < l, \\ -c_\mu^2 2^{-a_\mu j} (y_1^{a_\mu} + O(2^{-M})), & \text{if } \mu = l, \\ 2^{-k} (y_2 + O(2^{-\varepsilon(j+k)})), & \text{if } \mu \geq l + 1, \end{cases} \]
with \( M \) as in (6.19). Multiplying all these terms, we then see that
\[ \tilde{\phi}^{j,k}(y) = 2^{-(A_lj + B_lk)} \left(c_\mu y_1^{A_l} y_2^{B_l} + O(2^{-CM})\right), \]
for some constant \( C > 0 \), where \( A_l \) and \( B_l \) are given by (6.12) and \( M \) can still be chosen as large as we wish.

Observe that since \( l \leq \lambda - 1 \), we have \( B_l \geq B_{\lambda-1} \geq \nu_2 + N \left[ \frac{1}{\lambda} \right] \), and similarly as in the proof of Lemma 6.3, it is easy to see that we must have \( \nu_2 + N \left[ \frac{1}{\lambda} \right] \geq 2 \), hence \( B_l \geq 2 \). This implies that
\[ \partial^2_{y_2}(y_1^{A_l} y_2^{B_l}) \sim 1, \]
and that \( A_lj + B_lk \geq 2k \), so that
\[ 2^{k - m_l j} \leq C 2^{\frac{j}{2}(A_lj + B_lk)}. \]
We can therefore argue in a similar way as in the previous subsection and apply Corollary 3.6 to obtain
\[ ||M^{j,k}f||_p \leq C 2^{\frac{A_lj + B_lk}{p} - j - k} ||f||_p, \]
whenever \( p > 2 \), provided \( j + k \) is sufficiently large.

Summing all these estimates, we thus have
\[ ||M^n f||_p \leq C J ||f||_p, \]
where
\[ J := \sum_{(j,k):a_lj + M \leq k \leq a_{l+1}j - M} 2^{\frac{A_lj + B_lk}{p} - j - k}. \]
Assume now that \( p > h(\phi) \). Since \( h(\phi) \geq d(\tilde{\phi}_k) \geq d_h(\tilde{\phi}_{\kappa^{l+1}}) \), and since \( d_h(\tilde{\phi}_{\kappa^{l+1}}) = \frac{A_l + a_{l+1}B_l}{1 + a_{l+1}} \), we have
\[
p > \frac{A_l + a_{l+1}B_l}{1 + a_{l+1}}.
\]
This condition is equivalent to
\[
(1 - \frac{A_l}{p}) + a_{l+1}(1 - \frac{B_l}{p}) > 0.
\]
Similarly, since the mapping \( a \mapsto \frac{A_l + aB_l}{1 + a} \) is increasing, we may replace \( a_{l+1} \) by \( a_l \) in this estimate and also get
\[
(1 - \frac{A_l}{p}) + a_l(1 - \frac{B_l}{p}) > 0.
\]
In order to estimate \( J \), let us write \( k \) in the form \( k = \theta a_l j + (1 - \theta)a_{l+1}j + \omega \), with \( 0 \leq \theta \leq 1 \) and \( |\omega| \leq M \). Then
\[
j + k - \frac{A_l j + B_l k}{p} = (1 - \frac{A_l}{p}) j + (1 - \frac{B_l}{p}) k
\]
\[
= \left( \theta[(1 - \frac{A_l}{p}) + a_l(1 - \frac{B_l}{p})] + (1 - \theta)[(1 - \frac{A_l}{p}) + a_{l+1}(1 - \frac{B_l}{p})] \right) j + (1 - \frac{B_l}{p}) k.
\]
In view of (6.22) and (6.23), this shows that there exists a positive constant \( \varepsilon > 0 \) such that
\[
j + k - \frac{A_l j + B_l k}{p} > \varepsilon j,
\]
provided \( j \) is sufficiently big. It is now clear that \( J < \infty \), so that the maximal operator \( \mathcal{M}^n \) is bounded on \( L^p \) whenever \( p > h(\phi) \).

II) There remains the case \( l = l_0 - 1 \). This case can be treated in a very similar way (formally, it is like the previous case, only with \( a_{l_0 - 1} \) replaced by \( m_1 \geq a_{l_0 - 1} \)). Indeed, in this case (6.19) must be replaced by the inequalities
\[
m_1 j + M \leq k \leq a_{l_0}j - M,
\]
from which one derives that (6.21) remains valid, with \( l = l_0 - 1 \). From here, we can proceed exactly as before.

Q.E.D.

6.6. Reduction of the smooth case to the analytic setting. The estimates for the maximal operators in the preceding subsections hold true also for smooth functions \( \phi \). Indeed, denote by \( \phi(n) \) the Taylor polynomial of order \( n \) of \( \phi \) centered at the origin. For \( n \) sufficiently large, the Newton polyhedra of \( \phi \) and \( \phi(n) \) coincide, as do their faces and the corresponding principal parts (see [11]). It is then clear that the estimations of the operators \( \mathcal{M}^n \) work in the same way for smooth functions as in the analytic setting. Moreover, there exists a constant \( c > 0 \) such that if \( R_{j,k} \) is a dyadic rectangle on which \( x_1 \sim 2^{-j} \) and \( x_2 \sim 2^{-k} \), then the remainder term \( \phi - \phi(n) \) is of order \( O(2^{-c(j+k)n}) \) in
We may thus apply our previous approach to the polynomial \( \phi(n) \) in place of \( \phi \) and choose \( n \) so large that the contributions of the remainder term \( \phi - \phi(n) \) can be considered as negligible errors for the estimations of the operators \( M^\rho \) (compare the order \( O(2^{-c(j+k)n}) \) with the order of \( \tilde{\phi}^{j,k} \) in formula (6.21)).

7. Estimation of the maximal operator \( M \) near the principal root jet

We have reduced ourselves to the domain

\[
|x_2 - \psi(x_1)| \leq \varepsilon_0 x_1^a,
\]

where \( \psi \) is given as before, with leading term \( b_1 x_1^{m_1} \) and where \( \varepsilon_0 > 0 \) can still be chosen as small as we like. Moreover, we always assume that \( x_1 > 0 \). More precisely, in view of Proposition 6.1, there only remains to prove that the maximal operator \( M^\rho_0 \) associated to this domain is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > h(\phi) \).

Now, combining what we have proved so far, the following result is easy:

**Corollary 7.1.** Let \( \phi \) and its associated functions \( \phi^a \) and \( \tilde{\phi} \) be as in the previous section, and let \( \pi(\phi^a) \) be the principal face of the Newton polyhedron of \( \phi^a \) (i.e., of \( \phi \) when expressed in adapted coordinates). If any of the following conditions is satisfied, and if the neighborhood \( \Omega \) of the point \( (0,0) \) is chosen sufficiently small, then the maximal operator \( M \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > h(\phi) \):

(a) \( \pi(\phi^a) \) is a compact edge, and there is some \( j \) with \( 2 \leq j \leq h(\phi) \) such that \( \partial_2^j \tilde{\phi}_p(1,0) \neq 0 \).

(b) \( \pi(\phi^a) \) is a vertex.

(c) \( \pi(\phi^a) \) is a unbounded.

**Proof.** Since \( \varepsilon_0 > 0 \) can still be chosen as small as we wish, if there exists some \( j \) with \( 2 \leq j \leq h(\phi) \) such that \( \partial_2^j \tilde{\phi}_p(1,0) \neq 0 \), it is clear that we can argue near \( y_2 = 0 \) exactly as in the discussion of the maximal operator \( M^{\rho_\lambda} \) and obtain for the dyadic constituents \( M^k \) of the operator \( M^{\rho_\lambda} \) the estimate

\[
||M^k f||_p \leq C 2^{-|k| \frac{\lambda}{\rho} + \frac{k}{p}} ||f||_p,
\]

provided \( p > h(\phi) \). Moreover, by (6.5), these estimates sum in \( k \) as before. This proves in particular (a).

And, we have seen before in Subsection 6.4, that in the cases (b) and (c) such an integer \( j \) exists automatically - we can indeed choose \( j = h(\phi) \). This completes the proof of the corollary. Q.E.D.

In view of this corollary, we are left with the proof of

**Proposition 7.2.** Assume that \( \pi(\phi^a) \) is a compact edge, and that

\[
\partial_2^j \tilde{\phi}_p(1,0) = 0 \text{ for every } 2 \leq j \leq h(\phi),
\]

If the neighborhood \( \Omega \) of the point \( (0,0) \) is chosen sufficiently small, then the maximal operator \( M^{\rho_0} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > h(\phi) \).
Under the assumption (7.2), it will no longer be possible to estimate the maximal operator $M_{\rho_0}$ by means of oscillatory integral estimates in the variable $x_2$ alone, but we will have to take into account the oscillations in $x_1$ too.

We shall therefore consider the Fourier transforms of the convolution kernels of the averaging operators $A_t^{\rho_0}$, i.e.,

$$\hat{A}_t^{\rho_0} \hat{f}(\xi) = e^{it\xi_3} J^{\rho_0}(t\xi) \hat{f}(\xi),$$

where

$$J^{\rho_0}(t\xi) := \int e^{it(\xi_1 x_1 + \xi_2 x_2 + \xi_3 \phi(x_1, x_2))} \rho \left( \frac{x_2 - \psi(x_1)}{\varepsilon_0 x_1^a} \right) \eta(x) \, dx, \quad \xi \in \mathbb{R}^3.$$ 

Our goal will be to derive suitable estimates of the oscillatory integrals $J^{\rho_0}(\xi)$ (compare the method in [12]). If we change to the coordinates $y_1 := x_1$, $y_2 := x_2 - \psi(x_1)$ in the integral (notice that these are adapted, by our construction of $\psi$ in Section 6, since we assume that we are in case (a)), and assume again that $y_1 > 0$, we obtain

$$J^{\rho_0}(\xi) := \int_{\mathbb{R}^2_+} e^{i(y_1 \xi_1 + \xi_2 \psi(y_1) + \xi_2 y_2 + \xi_3 \tilde{\phi}(y))} \rho \left( \frac{y_2}{\varepsilon_0 y_1^a} \right) \tilde{\eta}(y) \, dy,$$

where $\tilde{\eta}$ is again a smooth cut-off function supported in a sufficiently small neighborhood of the origin and where $\mathbb{R}^2_+$ denotes the half-plane $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0\}$.

At this point, it will be convenient to defray our notation by writing $\phi$ in place of $\tilde{\phi}$ and $\eta$ in place of $\tilde{\eta}$. This means that from now on we shall consider Fourier multipliers of the form $e^{i\xi_3} J(\xi)$, with

$$J(\xi) := \int_{\mathbb{R}^2_+} e^{i(\xi_1 x_1 + \xi_2 \psi(x_1) + \xi_2 x_2 + \xi_3 \phi(x))} \rho \left( \frac{x_2}{\varepsilon_0 x_1^a} \right) \eta(x) \, dx,$$

such that the following general assumptions are fulfilled:

**Assumptions 7.3.** The functions $\phi, \psi$ and $\eta$ are smooth functions such that

(i) $\psi$ is given by $\psi(x_1) = \sum_{l=1}^K b_l x_1^{m_l} + c_p x_1^a$, where $b_l \neq 0$ for $l = 1, \ldots, K$;

(ii) $\phi$ is of finite type, and $\phi(0) = 0, \nabla \phi(0) = 0$;

(iii) the coordinates $x$ are adapted to $\phi$, i.e., $h(\phi) = d(\phi) \geq 2$;

(iv) the principal face $\pi(\phi)$ is a compact edge, and the associated principal part $\phi_p$ of $\phi$ is $\kappa$-homogeneous of degree one, with $a = \frac{\kappa}{\kappa_1} > m_K \geq m_1 \geq 2$ (in particular, $h(\phi) = \frac{1}{|\phi|}$);

(v) $\eta$ is a smooth bump function supported in a sufficiently small neighborhood $\Omega$ of the origin.

Moreover, we may and shall assume that

$$\partial_x^j \phi_p(1, 0) = 0 \text{ for every } 2 \leq j \leq h(\phi).$$


In order to estimate the maximal operator $M^{\rho_0}$ associated to the Fourier multiplier $e^{i\xi_3 J(\xi)}$, we shall further decompose it and estimate the corresponding constituents. If $\chi$ is a bounded measurable function, we shall use the notation

$$J^\chi(\xi) := \int_{\mathbb{R}^3_+} e^{i\left(\xi_1 x_1 + \xi_2 \psi(x_1) + \xi_3 \phi(x)\right)} \rho\left(\frac{x_2}{\varepsilon_0 x_1^a}\right) \eta(x) \chi(x) \, dx.$$ 

The corresponding re-scaled Fourier multiplier operators are the averaging operators $\hat{A}_t^\chi$ given by

$$\hat{A}_t^\chi f(\xi) = e^{it\xi_3 J^\chi(t\xi)} \hat{f}(\xi), \quad t > 0,$$

with associated maximal operator $M^\chi$. Then we shall make use the following essentially well-known result in order to estimate $M^\chi$.

**Lemma 7.4.** Assume that, for some $n \in \mathbb{N}$ and $\varepsilon > 0$, the following estimate

$$(7.5) \quad |J^\chi(\xi)| \leq A_\chi \|\eta\|_{C^\infty(\mathbb{R}^3)} (1 + |\xi|)^{-\left(1/2 + \varepsilon\right)}, \quad \xi \in \mathbb{R}^3,$$

holds, where the constant $A_\chi$ is independent of $\eta$. Moreover, put

$$B_\chi := \int |\rho\left(\frac{x_2}{\varepsilon_0 x_1^a}\right) \eta(x) \chi(x)| \, dx.$$

Then, for $2 \leq p \leq \infty$,

$$\|M^\chi f\|_p \leq C(A_\chi)^{\frac{p}{2}} (B_\chi)^{1-\frac{p}{2}} \|f\|_p,$$

where the constant $C$ depends only on the $C^n$-norms of $\phi$ and $\psi$ and the $C^n$-norm of $\eta$, but not on $\chi$.

**Proof.** Observe that

$$|\frac{\partial}{\partial t}[e^{it\xi_3 J^\chi(t\xi)}]| \leq |\xi| (|J^\chi(t\xi)| + |(\nabla J^\chi)(t\xi)|),$$

where, because of (7.5),

$$|J^\chi(\xi)| + |(\nabla J^\chi)(\xi)| \leq C A_\chi (1 + |\xi|)^{-\left(1/2 + \varepsilon\right)}.$$

The desired estimate of the maximal operator for $p = 2$ follows then essentially from Littlewood-Paley theory and Sobolev’s embedding theorem (for details, compare, e.g., [28], ch.XI.1, or our discussion in Subsection 3.1). Moreover, since $B_\chi$ is just the $L^1$-norm of the convolution kernel of $A_1^\chi$, the estimate for $p = \infty$ is trivial. The general case $2 \leq p \leq \infty$ then follows by interpolation.

Q.E.D.
7.1. The case where $\partial_2 \phi_p(1, 0) \neq 0$. Let us write

$$\Phi(x, \xi) := \xi_1 x_1 + \xi_2 \psi(x_1) + \xi_2 x_2 + \xi_3 \phi(x)$$

for the complete phase function of $J$, and decompose

$$\phi = \phi_p + \phi_r.$$

As in Subsection 6.4, we perform a dyadic decomposition

$$J = \sum_{k=k_0}^{\infty} J_k,$$

where

$$J_k(\xi) := J^{\chi_k}(\xi) = \int e^{i\Phi(x, \xi)} \rho \left( \frac{x_2}{\varepsilon_0 x_1} \right) \eta(x) \chi_k(x) \, dx,$$

with $\chi_k := \chi(\delta_k x)$. Here, the $\delta_k$ denote the dilations with respect to $\kappa$, and $\chi$ is supported in an annulus $1 \leq |x| \leq R$. Moreover, by choosing the neighborhood $\Omega$ of the origin sufficiently small, we may choose $k_0$ as large as we need. Notice that then

$$|M^k f| \leq \sum_{k=k_0}^{\infty} |M^{\chi_k} f|.$$

By a change of coordinates, we obtain

$$J_k(\xi) = 2^{-k|\kappa|} \int_{\mathbb{R}^2} e^{i2^{-k} \lambda \Phi_k(x, s)} \rho \left( \frac{x_2}{\varepsilon_0 x_1} \right) \eta(\delta_{2-k} x) \chi(x) \, dx,$$

where we have put $\lambda := \xi_3$, $s = (s_1, s_2)$ and

$$\Phi_k(x, s) := s_1 x_1 + S_2 \psi_k(x_1) + s_2 x_2 + \phi_p(x_1, x_2) + \phi_{r,k}(x),$$

with

$$\psi_k(x_1) := 2^{\kappa_1 m_1 k} \psi(2^{-\kappa_1 k} x_1) = b_1 x_1^m + O(2^{-\delta_1 k}) \text{ in } C^\infty,$$

$$\phi_{r,k}(x) := 2^k \phi_r(\delta_{2-k} x) = O(2^{-\delta_2 k}) \text{ in } C^\infty,$$

(assuming without loss of generality that $\xi_3 \neq 0$), where $\delta_1, \delta_2 > 0$.

We remark that indeed $\psi_k(x)$ and $\phi_{r,k}(x)$ can be viewed as smooth functions $\tilde{\psi}(x_1, \delta)$ respectively $\tilde{\phi}_r(x, \delta)$ depending on the small parameter $\delta = 2^{-k/r}$ for some suitable positive integer $r \geq 1$ such that

$$\tilde{\psi}(x_1, 0) = b_1 x_1^m, \quad \tilde{\phi}_r(x, 0) \equiv 0.$$

Observe also that $1 - \kappa_1 m_1 > \kappa_2 - \kappa_1 m_1 > 0$, $1 - \kappa_j > 0$, so that in particular

$$|S_2| >> |s_2| \text{ and } |\lambda s_j| >> |\xi_j|.$$

Recall that in our domain of integration, we have $x_1 \sim 1$, $|x_2| \lesssim \varepsilon_0$. 

$$S_1 := 2^{(1-\kappa_1) k} \xi_1, \quad S_2 := 2^{(1-\kappa_2) k} \xi_2,$$

$$S_3 := 2^{(1-\kappa_1 m_1) k} \xi_2 = 2^{(\kappa_2 - \kappa_1 m_1) k} S_2.$$
The following proposition will be useful not only in in the present situation. Its proof will make use of estimates for oscillatory integrals given in the later Section 9.

**Proposition 7.5.** Assume that \( \phi \) and \( \psi \) satisfy the Assumptions 7.3 (but not necessarily (7.4)), and that \( \partial_2 \phi_p(1,0) \neq 0 \). If \( \varepsilon_0 \) above is chosen sufficiently small, then the following estimate

\[
|J_k(\xi)| \leq C||\eta||_{C^3(\mathbb{R}^2)} \frac{2^{-k|\kappa|}}{(1 + |2^{-k}\xi|)^{1/2 + \varepsilon}}
\]

holds true for some \( \varepsilon > 0 \), where the constant \( C \) does not depend on \( k \) and \( \xi \).

Consequently, the maximal operator \( M^{\rho_0} \) associated to the averaging operators \( A_{t_0}^{\rho_0} \), \( t > 0 \), defined by \( \hat{A}_{t_0}^{\rho_0} f(\xi) = e^{it_0\xi} J_{t_0}(t\xi) \hat{f}(\xi) \), is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > 1/|\kappa| \).

**Proof.** We shall distinguish several cases, assuming for simplicity that \( \lambda > 0 \).

1. **Case.** \( |s_1| + |S_2| \leq C \) for some large constant \( C >> 1 \). In this case, if \( k \) is sufficiently large, then we have \( |s_2| << 1 \), and since \( \partial_2 \phi_p(1,0) \neq 0 \), we can integrate by parts in \( x_2 \) and obtain

\[
|J_k(\xi)| \leq C 2^{-k|\kappa|}(1 + 2^{-k}\lambda)^{-1},
\]

hence (7.10), since, by (7.9), \( |\xi| \sim \lambda \) in this case.

2. **Case.** \( |s_1| + |S_2| \geq C \), with \( C \) as above, and either \( |s_1| << |S_2| \) or \( |s_1| >> |S_2| \).

In this case we can integrate by parts in \( x_1 \) and obtain

\[
|J_k(\xi)| \leq C 2^{-k|\kappa|}(1 + 2^{-k}\lambda(|s_1| + |S_2|))^{-1}),
\]

which again implies (7.10), since here, by (7.9), \( |\xi| \lesssim \lambda(|s_1| + |S_2|) \).

3. **Case.** \( |s_1| + |S_2| \geq C \), with \( C \) as above, and \( |s_1| \sim |S_2| \). Observe first that \( |s_1| \sim |S_2| \) implies \( |\xi_2| \sim 2^{|s_1| (m_1 - 1)k}|\xi_1| \), so that

\[
|\xi_2| >> |\xi_1|.
\]

We then write

\[
2^{-k}\lambda \Phi_k(x, s) = 2^{-k}\lambda S_2 F(x, \sigma, \delta),
\]

where

\[
F(x, \sigma, \delta) := \frac{s_1}{S_2} x_1 + \tilde{\psi}(x_1, \delta) + \sigma \left( \phi_p(x_1, x_2) + \tilde{\phi}_r(x_1, x_2, \delta) + s_2 x_2 \right)
\]

and \( \delta := 2^{-k/r} << 1 \), \( \sigma := \frac{1}{S_2} \), so that

\[
|\frac{s_1}{S_2}| \sim 1, \ |\sigma| << 1.
\]

Observe that

\[
\left| \partial_2^2 \left( \frac{s_1}{S_2} x_1 + \tilde{\psi}(x_1, 0) \right) \right| \sim 1
\]
for $x_1 \sim 1$. We also claim that the polynomial $P(x_2) := \phi_p(x_1^0, x_2)$ has degree
\[
(7.11) \quad m := \deg P \geq 2.
\]

For, otherwise, by the homogeneity of $\phi_p$, the polynomial $\phi_p$ was of the form $\phi_p(x) = c_1 x_1^n + c_2 x_1^2 x_2$, where the point $(l, 1)$ had to lie in the closed half-space above the bisectrix, since $\phi_p$ is the principal part of $\phi$. Thus $l \leq 1$, so that $d(\phi) \leq 1$, in contradiction to our assumption $d(\phi) = h(\phi) \geq 2$.

From (7.11) we conclude that there is some integer $m \geq 2$ so that
\[
|\partial_{x_2}^m(\phi_p(x_1, x_2) + s_2 x_2)| \sim 1.
\]

If we now fix $x_1^0 \sim 1$ and translate the $x_1$-coordinate by $x_1^0$, we see that we can apply Proposition 9.1 if we localize our oscillatory integral $J_k$ to a small neighborhood of $(x_1^0, 0)$ by introducing a suitable cut-off function into the amplitude, and obtain an estimate of order
\[
O(2^{-k|\kappa|}(1 + 2^{-k}(|S_2|))^{-1/2}(1 + 2^{-k}\lambda)^{-1/m})
\]
for the corresponding localized integral, uniformly in $s_1$ and $s_2$, since Proposition 9.1 also gives uniform estimates for small perturbations of such parameters. Since we can decompose $J_k(\xi)$ by means of a suitable partition of unity into such localized oscillatory integrals, we see that
\[
|J_k(\xi)| \leq C 2^{-k|\kappa|}(1 + 2^{-k}(|S_2|))^{-1/2}(1 + 2^{-k}\lambda)^{-1/m},
\]
where $m \geq 2$.

a) If we assume that $|s_2| \leq C$ for some fixed, large constant $C$, then we have $|\xi_1| << |\xi_2| << |s_2\lambda| \leq C|\lambda|$, hence $|\xi| \sim \lambda$, so that this estimate implies (7.10).

b) If $|s_2| >> 1$, then we proceed in a slightly different way. We first perform one integration by parts in $x_2$, and then apply the method of stationary phase in $x_1$. This leads to the estimate
\[
|J_k(\xi)| \leq C 2^{-k|\kappa|}(1 + 2^{-k}(|S_2|))^{-1/2}(1 + 2^{-k}|s_2|\lambda)^{-1},
\]
If now $|\xi_2| \leq \lambda$, then $|\xi| \sim \lambda$, and if $|\xi_2| \geq \lambda$, then $|\xi| \sim |\xi_2| << |s_2|\lambda$, so that again (7.10) follows.

In order to estimate the maximal operator $M^{\alpha_0}$, we observe that (7.10) implies that
\[
|J_k(\xi)| \leq C_\varepsilon 2^{-k|\kappa|}2^{k(\frac{1}{2} + \varepsilon)}(1 + |\xi|)^{-\frac{1}{2} - \varepsilon}
\]
for every sufficiently small $\varepsilon > 0$. We may therefore choose $A_{\chi_k} := C_\varepsilon 2^{-k|\kappa|}2^{k(\frac{1}{2} + \varepsilon)}$ for $\chi = \chi_k$ in Lemma 7.4. Moreover, clearly we can choose $B_{\chi_k} := C2^{-k|\kappa|}$, so that we have
\[
||M^{\chi_k}f||_p \leq C_\varepsilon 2^{-k|\kappa| - \frac{1}{p} - \varepsilon},
\]
with a constant $C_\varepsilon$ which is independent of $k$. If $p > 1/|\kappa|$, and if $\varepsilon$ is chosen small enough, these estimates sum in $k$, so that the maximal operator $M^{\alpha_0}$ is bounded on $L^p$. 


We shall indeed need a slight extension of this result to the following situation. As before, we shall always assume that $x_1 > 0$.

**Definitions.** Let $q \in \mathbb{N}^\times$ be a fixed positive integer. Assume that $\phi$ is a smooth function of the variables $x_1^{1/q}$ and $x_2$ near the origin, i.e., that there exists a smooth function $\phi^{[q]}$ near the origin such that $\phi(x) = \phi^{[q]}(x_1^{1/q}, x_2)$. If the Taylor series of $\phi^{[q]}$ is given by

$$\phi^{[q]}(x_1, x_2) \sim \sum_{j,k=0}^{\infty} c_{j,k} x_1^{j/q} x_2^k,$$

then $\phi$ has the formal Puiseux series expansion

$$\phi(x_1, x_2) \sim \sum_{j,k=0}^{\infty} c_{j,k} x_1^{j/q} x_2^k.$$

We therefore define the *Taylor-Puiseux support* of $\phi$ by

$$T(\phi) := \{(\frac{j}{q}, k) \in \mathbb{N}_q^2 : c_{jk} \neq 0\},$$

where

$$\mathbb{N}_q^2 := \left(\frac{1}{q}\mathbb{N}\right) \times \mathbb{N}.$$

The *Newton-Puiseux polyhedron* $\mathcal{N}(\phi)$ of $\phi$ at the origin is then defined to be the convex hull of the union of all the quadrants $(\frac{j}{q}, k) + \mathbb{R}_+^2$ in $\mathbb{R}^2$, with $(\frac{j}{q}, k) \in T(\phi)$. The associated *Newton-Puiseux diagram* $\mathcal{N}_d(\phi)$ is the union of all compact faces of the Newton-Puiseux polyhedron, and the notions of principal face, distance and homogeneous distance are defined as in the case of Newton diagrams. The principal part $\pi(\phi)$ is analogously defined by

$$\phi_p(x) := \sum_{(\frac{j}{q}, k) \in \pi(\phi)} c_{j,k} x_1^{j/q} x_2^k.$$

We shall then again decompose $\phi = \phi_p + \phi_r$.

**Corollary 7.6.** Proposition 7.5 remains true even under the following weaker assumptions on $\psi$ and $\phi$ in place of Assumptions 7.3, provided again that $\partial_2 \phi_p(1,0) \neq 0$:

(i) $\psi$ is given by $\psi(x_1) = \sum_{l=1}^L b_l x_1^{m_l}$, where $b_l \neq 0$ for $l = 1, \ldots, K$, and where $2 \leq m_1 < \cdots < m_L$ are positive real numbers.

(ii) $\phi$ is a smooth function of the variables $x_1^{1/q}$ and $x_2$ as above, the principal face $\pi(\phi)$ is a compact edge, and the associated principal part $\phi_p$ of $\phi$ is $\kappa$-homogeneous of degree one, where $0 < \kappa_1 < \kappa_2 < 1$ and $a := \frac{\kappa_2}{\kappa_1} > m_1$.

(iii) for the distance $d(\phi) = \frac{1}{\kappa_1}$ we have $d(\phi) \geq 2$.

(iv) $\eta$ is a smooth bump function supported in a sufficiently small neighborhood $\Omega$ of the origin.
Proof. All of our arguments extend in a straightforward manner to this setting, except perhaps for the proof of (7.11) and the straightforward application of Lemma 7.4. However, if (7.11) was false in the present situation, then we could write
\[ \phi_p(x) = c_1 x_1^{n/q} + c_2 x_1^{l/q} x_2. \]
The point \((l/q, 1)\) had to lie above the bisectrix, since \(\phi_p\) is the principal part of \(\phi\). Thus \(l < q\). Moreover, we would have \(\kappa_1 = \frac{q}{n}\), \(\kappa_2 = 1 - \frac{l}{n}\), so that
\[ |\kappa| = 1 + \frac{q - l}{n} > 1, \]
hence \(d(\phi) < 1\), in contradiction to our assumption in (iii).

As for Lemma 7.4, notice that when applying the gradient to \(J_k(\xi)\), the function \(\eta\) will be multiplied with terms like \(\phi\) or \(\psi\), which may not be smooth at \(x_1 = 0\), so that the argument in the proof of the lemma fails to hold. However, if we look at the formula for \(J_k(\xi)\) after scaling the coordinates \(x\), we find that the factor \(\eta(\delta - k x)\) will have to be replaced, for instance, by \(\phi(\delta - k x) \eta(\delta - k x)\), where we now are in the domain where \(x_1 \sim 1\), \(|x_2| \lesssim \epsilon_0\). But, in this domain, the \(C^m\)-norms of such expressions are still uniformly bounded in \(k\), so that we obtain the same type of estimate as for \(J_k(\xi)\).

Q.E.D.

As a consequence of Proposition 7.5, we see in particular that Proposition 7.2 holds true in the case where \(\partial_2 \phi_p(1, 0) \neq 0\), since here \(h(\phi) = 1/|\kappa|\).

7.2. The case where \(\phi_p\) is one of the exceptional polynomials (4.4) in Proposition 4.3. Corollary 7.6 will be useful also in order to deal with the situation where \(\phi_p\) is one of the exceptional polynomials \(P\) in (4.4), i.e.,
\[ P(y) = c(y_2^2 - \lambda_1 y_1^5)(y_2^2 - \lambda_2 y_1^5), \]
in Subsection 6.4. We were left with the domain (6.17) (in the original coordinates). Now, if we here put
\[ \psi(x_1) := b_1 x_1^2 \pm \sqrt{\frac{\lambda_1 + \lambda_2}{6}} x_1^{5/2}, \]
change coordinates as in (6.3) and call the new coordinates again \(x\), then the domain (6.17) will correspond to the domain
\[ |x_2| \leq \epsilon_0 x_1^{5/2} \]
in the present context, and we have
\[ \phi_p(x) = P(x_1, x_2 \mp \sqrt{\frac{\lambda_1 + \lambda_2}{6}} x_1^{5/2}). \]
Recall also that \(\lambda_1 + \lambda_2 > 0\), and that, in our present notation, the function \(\phi_p\) is \(\kappa\)-homogeneous of degree one, with \(\kappa_1 := \frac{1}{10}\) and \(\kappa_2 := \frac{1}{4}\). Moreover, \(h(\phi) = d(\phi_p) = 1/|\kappa| = 20/7\), \(K = 1\), \(m_1 = 2\) and \(a = m_2 = 5/2\).

So, we are again just left with the estimation of the maximal operator \(M^{\alpha_0}\) of the previous subsection. But, notice that
\[ \partial_2 P(y) = 4 y_2(y_2^2 - \frac{\lambda_1 + \lambda_2}{2} y_1^5). \]
This shows that \( \partial_2 \phi_p(1,0) \neq 0 \), and clearly the assumptions in Corollary 7.6 are satisfied, so that \( \mathcal{M}^\alpha \) is indeed bounded on \( L^p \) for \( p > h(\phi) = 20/7 \).

### 7.3. Further domain decompositions in the case where \( \partial_2 \phi_p(1,0) = 0 \).

We first observe that the Assumptions 7.3 imply in this case that \( \phi_p(1,0) \neq 0 \).

For otherwise the root \( x_2 = 0 \) had multiplicity \( N \) at least 2. On the other hand, since the coordinates \( x \) are adapted to \( \phi \), we must have \( N \leq h(\phi) \), so that (7.4) would fail to be true for \( j = N \).

We can thus write

\[
\phi_p(x_1, x_2) = x_2^B Q(x_1, x_2) + c x_1^n, \quad \text{with} \quad c \neq 0,
\]

where \( B \geq 1 \), and where \( Q \) is a \( \kappa \)-homogeneous polynomial such that \( Q(x_1,0) = bx_1^q \), \( b \neq 0 \), so that \( Q(x_1,0) \neq 0 \) for \( x_1 > 0 \). Without loss of generality, we shall assume that \( c = 1 \). Recall also that we are in the domain

\[
|x_2| \leq \varepsilon_0 x_1^a.
\]

Notice \( B \geq 2 \), since \( \partial_2 \phi_p(1,0) = 0 \), and then our assumption (7.4) implies that in fact

\[
B > h(\phi) \geq 2.
\]

In order to understand the behavior of \( \phi \) as a function of \( x_2 \), for \( x_1 \) fixed, we shall decompose

\[
\phi(x_1, x_2) = \phi(x_1,0) + \theta(x_1, x_2),
\]

and write the complete phase \( \Phi \) in the form

\[
\Phi(x, \xi) = (\xi_3 \phi(x_1,0) + \xi_1 x_1 + \xi_2 \psi(x_1)) + (\xi_3 \theta(x_1, x_2) + \xi_2 x_2),
\]

Notice that

\[
\phi(x_1,0) = x_1^n(1 + O(x_1)), \quad \psi(x_1) = b_1 x_1^m(1 + O(x_1)), \quad \theta_\kappa(x_1, x_2) = x_2^B Q(x_1, x_2),
\]

where \( \theta_\kappa \) denotes the \( \kappa \)-principal part of \( \theta \).

Now, by means of the \( \kappa \)-dilations we would like to reduce our considerations as before to the domain where \( x_1 \sim 1 \). In this domain, \( |x_2| \ll 1 \), so that \( \theta_\kappa(x) \sim x_2^B Q(x_1,0) \).

What leads to problems is that the "error term" \( \theta_{\kappa,r} := \theta - \theta_\kappa \), which consists of terms of higher \( \kappa \)-degree than \( \theta_\kappa \), may nevertheless contain terms \( c_j x_2^{l_j} x_1^{n_j} \) of lower \( x_2 \)-degree \( l_j < B \). After scaling by \( \delta_{2-k} \), so that then \( x_1 \sim 1 \) and \( |x_2| \lesssim \varepsilon_0 \), these terms will have small coefficients compared to \( x_2^B Q(x_1, x_2) \), but for \( |x_2| \) very small they may nevertheless become dominant and have to be taken into account.

Consider now the Newton polyhedron \( \mathcal{N}(\theta) \). Since the Taylor support \( T(\theta) \) arises from \( T(\phi) \) by removing all points on the \( t_1 \)-axis, we have

\[
\mathcal{N}(\partial_2 \phi) = (0,-1) + \mathcal{N}(\theta).
\]

Moreover, if we put

\[
\kappa^1 := \kappa, \quad a_1 := a = \kappa_2^1 / \kappa_1^1,
\]
then the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$ contains the point $(q, B)$ of $\mathcal{N}(\theta)$. This point is contained in the face

$$\gamma_1 = [(A_0, B_0), (A_1, B_1)], \text{ with } (A_1, B_1) := (q, B),$$

of the Newton diagram $\mathcal{N}_q(\theta)$ lying on this line. Note that possibly $(A_0, B_0) = (A_1, B_1)$.

It is also clear from the construction of $\theta$ from $\phi$ that

$$\mathcal{N}(\theta) \cap \{t_2 \geq B_1\} = \mathcal{N}(\phi) \cap \{t_2 \geq B_1\}.$$ 

(7.17)

We next describe a stopping time algorithm oriented at the level sets of $\partial_2 \theta$ which will decompose our domain (7.12) in a finite number of steps into subdomains, whose contributions to our maximal operator will be treated in different ways in the subsequent subsections. This algorithm will follow a similar line of thought as Varchenko’s algorithm (compare [11]), and it will stop at latest when we have reached a domain containing only one root of $\partial_2 \theta$ (with multiplicity).
**Case A.** $\mathcal{N}(\theta) \subset \{t_2 \geq B_1\}$

Then no term in $\theta$ has higher $x_2$-exponent than $B_1 = B$, and we stop at this point.

**Case B.** $\mathcal{N}(\theta)$ contains points below the line $t_2 = B_1$.

Then the Newton diagram $\mathcal{N}_0(\theta)$ will contain a further edge

$$\gamma'_2 = [(A_1, B_1), (A'_2, B'_2)]$$

below the line $t_2 = B_1$, lying, say, on the line $\kappa^2_1 t_1 + \kappa^2_2 t_2 = 1$ (compare figure 1). We then put

$$\kappa^2 := (\kappa^2_1, \kappa^2_2), \quad a_2 := a = \kappa^2_2 / \kappa^2_1,$$

where clearly $a_2 > a_1$.

Notice that $a_2 \in \mathbb{Q}$. We then decompose the domain (7.12) into the domains

$$E_1 := \{ N_2 x_1^{a_2} < |x_2| \leq \varepsilon_1 x_1^{a_1} \}$$

and

$$H_2 := \{ |x_2| \leq N_2 x_1^{a_2} \},$$

where $N_2$ will be any sufficiently large constant and $\varepsilon_1 := \varepsilon_0$.

In the domain $E_1$, which is again domain of transition between two different homogeneities, we stop our algorithm. It will later be treated by means of bi-dyadic decompositions.

The $\kappa^2$-homogeneous domain $H_2$ will be further decomposed as follows:

We first notice that the $\kappa^2$-homogeneous part $(\partial_2 \theta)_{\kappa^2}$ will be associated to the edge $(0, -1) + \gamma'_2 = [(A_1, B_1 - 1), (A'_2, B'_2 - 1)]$ of the Newton diagram of $\partial_2 \theta$ and is $\kappa^2$-homogeneous of degree $1 - \kappa^2_2$. Observe also that in view of (7.16) we have $(\partial_2 \theta)_{\kappa^2} = \partial_2 (\theta_{\kappa^2})$, which is a polynomial in the fractional power $x_1^{a_2}$ of $x_1$ and $x_2$.

Decomposing the polynomial $t \mapsto (\partial_2 \theta)_{\kappa^2}(1, t)$ into linear factors and making use of the $\kappa^2$-homogeneity of $(\partial_2 \theta)_{\kappa^2}$, we see that we can write it in the form

$$(\partial_2 \theta)_{\kappa^2}(x) = c_2 x_1^{A_1} x_2^{B'_2 - 1} \left( x_2 - c_2^o x_1^{a_2} \right)^{n_2^o};$$

where

$$B_1 = B'_2 + \sum_\alpha n_2^\alpha, \quad A'_2 = A_1 + a_2 \sum_\alpha n_2^\alpha,$$

with roots $c_2^\alpha \in \mathbb{C} \setminus \{0\}$ and multiplicities $n_2^\alpha \geq 1$. Let us assume in the sequel that $N_2 >> \max_\alpha |c_2^\alpha|$.

By $R_2$ we shall denote the set of all real roots $c_2^\alpha \in \mathbb{R}$, where we include also the root $d = 0$ in case that $B'_2 - 1 > 0$.

We need to understand the behavior of the complete phase function $\Phi(x, \xi)$ in display (7.14) on the domain $H_2$. Now, after dyadic decomposition with respect to the $\kappa^2$-dilation and re-scaling, we have to look at $\Phi(2^{-\kappa^2_2} x_1, 2^{-\kappa^2_2} x_2, \xi)$ in the domain where $x_1 \sim 1$ and $|x_2| \leq N_2$. We write

$$\Phi(2^{-\kappa^2_2} x_1, 2^{-\kappa^2_2} x_2, \xi) = 2^{-\kappa^2 nk} \lambda \Phi_k(x, s),$$
where

\[ \Phi_k(x, s) := x_1^n(1 + v_k(x_1)) + s_1x_1 + S_2b_1x_1^{m_1}(1 + w_k(x_1)) + 2(\kappa_1^{2n} - 1)k (\theta_{\kappa_2} + \theta_{r,k}) + s_2x_2 \]

and again \( \lambda := \xi_3 \) (assumed to be positive) and

\[ s_1 := 2\kappa_1^{2(n-1)}k \xi_1^2, \quad s_2 := 2(1-\kappa_1^2)k \xi_1^2, \quad S_2 := 2\kappa_1^{2(n-m_1)}k \xi_1^2 = 2(\kappa_1^{2(n-m_1)} + \kappa_2^2 - 1)k s_2. \]

The functions \( v_k, w_k \) and \( \theta_{r,k} \) are of order \( O(2^{-\delta k}) \) in \( C^\infty \) for some \( \delta > 0 \).

In the estimation of the corresponding oscillatory integral, the worst possible case arises when \( |s_1| \sim |S_2| \sim 1 \), so that

\[ |s_2| \sim 2^{-(\kappa_1^{2(n-m_1)} + \kappa_2^2 - 1)k}. \]

Fix now an arbitrary \( \varepsilon_2 > 0 \). For any point \( d \) in the interval \([-N_2, N_2]\) denote by \( D_2(d) \) the \( \kappa_2^2 \)-homogeneous domain (inside the half-plane \( x_1 > 0 \))

\[ D_2(d) := \{ |x_2 - dx_1^{a_2}| \leq \varepsilon_2 x_1^{a_2} \}. \]

Since we can cover the domain \( H_2 \) be a finite number of domains \( D_2(d) \), it will be sufficient to examine the contribution of the domains \( D_2(d) \).

**Case B (a).** If \( \kappa_1^2(n - m_1) + \kappa_2^2 \leq 1 \), then we have \( |s_2| \geq c > 0 \) in (7.18). In this case, it will be possible to control the corresponding oscillatory integrals if \( \varepsilon_2 \) is chosen sufficiently small, and we shall stop our algorithm with the domains \( D_2(d) \).

Indeed, if \( \kappa_1^2(n - m_1) + \kappa_2^2 < 1 \), then \( |s_2| >> 1 \), which will allow for an integration by parts with respect to \( x_2 \) as in the first case of the proof of Proposition 7.5. The worst possible case will actually arise when \( \kappa_1^2(n - m_1) + \kappa_2^2 = 1 \) and when in addition \( d \notin R_2 \), i.e., \( (\partial_\theta)_{\kappa_2}(1, d) \neq 0 \), which will indeed lead to kind of "degenerate Airy-type" integrals.

**Case B (b).** If \( \kappa_1^2(n - m_1) + \kappa_2^2 > 1 \), then \( |s_2| << 1 \) in (7.18).

(i) If \( d \notin R_2 \), then \( (\partial_\theta)_{\kappa_2}(1, d) \neq 0 \) and \( |s_2| << 1 \), so that again one can integrate by parts with respect to \( x_2 \), and again the algorithm will stop.

(ii) Assume finally that \( d \in R_2 \), so that \( (\partial_\theta)_{\kappa_2}(1, d) = 0 \) and \( |s_2| << 1 \). In this case, we introduce new coordinates

\[ y_1 := x_1, \quad y_2 := x_2 - dx_1^{a_2}, \]

and denote our original functions, when expressed in the new coordinates \( y \), by a subscript \( _{(2)} \), e.g.,

\[ \phi_{(2)}(y) := \phi(y_1, y_2 + dy_1^{a_2}). \]

\( \theta_{(2)} \) is defined by

\[ \phi_{(2)}(y_1, y_2) = \phi_{(2)}(y_1, 0) + \theta_{(2)}(y), \]

etc. Notice that in general we don’t have \( \theta_{(2)}(y) = \theta(y_1, y_2 + dy_1^{a_2}) \), but

\[ \partial_2 \theta_{(2)}(y) = \partial_2 \theta(y_1, y_2 + dy_1^{a_2}). \]
Notice that this $\kappa^2$-homogeneous change of coordinates will have the effect on the Newton-polyhedron that the edge $\gamma'_2 = [(A_1, B_1), (A'_2, B'_2)]$ of $\mathcal{N}(\theta)$ on the line $\kappa_1^2 t_1 + \kappa_2^2 t_2 = 1$ will be turned into a face

$$\gamma_2 = [(A_1, B_1), (A_2, B_2)]$$

of $\mathcal{N}(\theta_{(2)})$ on the same line, with same left end-point $(A_1, B_1)$ but possibly different right endpoint $(A_2, B_2)$ (which may even agree with the left endpoint), where still $B_2 \geq 1$.

Notice that $B_1 \geq B_2$, and that the domain $D_2(d)$ corresponds to the domain where $|y_2| \leq \varepsilon_2 y_1^{a_2}$ in the new coordinates $y$.

In the Case B(b)(ii), which is the only one where our algorithm did not stop, we see that by passing from $\phi =: \phi_{(1)}$ to $\phi_{(2)}$ and denoting the new coordinates $y$ again by $x$, we have thus reduced ourselves to the smaller, $\kappa^2$-homogeneous domain $|x_2| \leq \varepsilon_2 x_1^{a_2}$ in place of (7.12).

We observe also that since the $\kappa = \kappa^1$-homogenous part of our change of coordinates $y_1 = x_1$, $y_2 = x_2 - dx_1^{a_2}$ is given by $x_1, x_2$, i.e., by the identity mapping, the Newton polyhedra of $\theta_{(1)}$ and $\theta_{(2)}$ will have the same $\kappa^1$-principal faces and corresponding principal parts. This implies in particular that still

$$\phi_{(2)}(x_1, 0) = x_1^n (1 + O(x_1^{1/r}))$$

for some rational exponent $r > 0$. Moreover, since $a_2 > a_1 > m_1$, also the new function $\psi_{(2)}(x_1) := \psi(x_1) + dx_1^{a_2}$, which corresponds to $\psi$ in the new coordinates, will still satisfy

$$\psi_{(2)}(x_1) = b_1 x_1^{m_1} (1 + O(x_1^{1/r})).$$

Replacing $\phi$ by $\phi_{(2)}$, we can now iterate this procedure. Notice that already the function $\phi_{(2)}$ will in general be only a smooth function of $x_2$ and some fractional power of $x_1$, so that from here on we shall have to work with Newton-Puiseux polyhedra in place of Newton-polyhedra, etc..

**Example.** Let $\phi(x_1, x_2) := x_1^n + x_2^l + x_2 x_1^{n-m_1}$ and $\psi(x_1) := x_1^{m_1}$, where $n/l > m_1 \geq 2$. The coordinates $(x_1, x_2)$ are adapted to $\phi$. Notice also that in the original coordinates, say $(y_1, y_2)$, $\phi$ was given by $(y_2 - y_1^{m_1})^l + y_2 y_1^{n-m_1}$. Here

$$\phi(x_1, 0) = x_1^n, \quad \theta(x) = x_2^l + x_2 x_1^{n-m_1}, \quad \theta_{\kappa}(x) = x_2^l,$$

whereas

$$\theta_{\kappa^2}(x) = x_2^l + x_2 x_1^{n-m_1}.$$ 

Thus, if $d := 0$, we arrive at the "degenerate Airy type" situation describes in Case B (a).
Details on and modification of the algorithm. Suppose we have constructed in this way recursively a sequence

$$\phi = \phi(1), \phi(2), \ldots, \phi(L)$$

of functions, where $\phi(l)$ is obtained from $\phi(l-1)$ for $l \geq 2$ by means of a change of coordinates $y_1 := x_1$, $y_2 := x_2 - d_l x_1^a_l$ (figure 2).

Then $\phi(l)$ arises from $\phi$ by the total change of coordinates $x = \varphi(l)(y)$, where

$$y_1 := x_1, \quad y_2 := x_2 - \sum_{j=2}^{l} d_j x_1^{a_j},$$

i.e., $\phi(l) = \phi \circ \varphi(l)$, and correspondingly $\theta(l)$ is defined by

$$\phi(l)(y_1, y_2) = \phi(l)(y_1, 0) + \theta(l)(y),$$
etc. Notice that in general we don’t have $\theta(t) = \theta \circ \varphi(t)$, but

$$\partial_2 \theta(t) = \partial_2 \varphi(t) = \partial_2 \varphi.$$  

For the functions $\phi(t)(x_1, 0)$ and $\psi(t)(x_1) = \psi(x_1) + \sum_{j=2}^{l} d_j x_1^{\alpha_j}$ we then still have

$$\phi(t)(x_1, 0) = x_1^\kappa_1(1 + O(x_1^{1/r})), \quad \psi(t)(x_1) = b_1 x_1^{m_1}(1 + O(x_1^{1/r})), \quad \text{for some rational exponent } r > 0.$$  

In each step, we produce a new face $\gamma_t = [(A_{t-1}, B_{t-1}), (A_t, B_t)]$ (possibly a single point) of $\mathcal{N}(\theta(t))$, so that the Newton diagram $\mathcal{N}_d(\theta(t))$ will in particular posses the faces

$$\gamma_1 = [(A_0, B_0), (A_1, B_1)], \ldots, \gamma_l = [(A_{l-1}, B_{l-1}), (A_l, B_l)],$$  

where $B_t \geq 1$. The Newton diagram $\mathcal{N}_d(\theta(t-1))$ will have in addition a compact edge $\gamma'_l = [(A_{l-1}, B_{l-1}), (A'_l, B'_l)]$, lying on a unique line

$$\kappa_1' t_1 + \kappa_2' t_2 = 1,$$  

which contains also $\gamma_t$, such that $a_l = \frac{\kappa_1'}{\kappa_1}$. Moreover, $x_2 = d_l x_1^{\alpha_l}$ will be a real root of the $\kappa'$-homogeneous principal part $\partial_2(\theta(t-1))_\kappa'$ of $\partial_2(\theta(t-1))$ corresponding to the edge $\gamma'_l$, i.e., $\partial_2(\theta(t-1))_\kappa'(1, d_l) = 0$, where $\partial_2(\theta(t-1))_\kappa'$ is a polynomial in a fractional power of $x_1$ and in $x_2$. Moreover,

$$B_1 \geq B_2 \geq \cdots \geq B_l \geq 1 \quad \text{and} \quad m_1 < a = a_1 < a_2 < \cdots < a_l.$$  

In particular, the descending sequence $\{B_l\}_l$ must eventually become constant (unless our algorithm stops already earlier).

Our algorithm will always stop after a finite number of steps, since eventually we will have $\kappa_1'(n - m_1) + \kappa_2' \leq 1$, because $a_l \to \infty$. This is evident from the geometry of the Newton-Puiseux polyhedra $\mathcal{N}(\theta(t))$.

More precisely, in case that our algorithm did not terminate, then we could find some minimal $L \geq 1$ such that $B_l = B_L$ for every $l \geq L$. Then $B_L \geq 2$, since for $B_L = 1$ we had $\mathcal{N}_d(\theta(L)) \in \{t_2 \geq B_L\}$, and we would stop. Moreover, from $1 = \kappa_1 A_l + \kappa_2 B_l \geq 2\kappa_2'$ we conclude that $\kappa_2' \leq 1/2$.

Next, we must have that $a_l \to \infty$. For analytic $\phi$, this follows easily from the Puiseux-series expansions of roots of $\partial_2 \theta$. However, for sufficiently large $N$, the points $(t_1, t_2) \in \mathcal{N}(\theta)$ with $\kappa_1 t_1 + \kappa_2 t_2 > N$ will have no influence on the Newton-Puiseux diagrams of the functions $\theta(t)$ (compare the discussion in [11]), so that we can reduce the statement to the case of polynomials. This shows that

$$\kappa_1'(n - m_1) + \kappa_2' = \kappa_2'(1 + \frac{n - m_1}{a_l}) \leq \frac{1}{2}(1 + \frac{n - m_1}{a_l}) \leq 1$$  

for $l$ sufficiently large, and so our algorithm would stop at this step.

Let us therefore assume from now on that our algorithm terminates at step $l = L$.

Next, in case that $B_l = B_{l+1} = \cdots = B_{l+j}$ for some $j \geq 1$, then we will modify our stopping time argument as follows:
We shall skip the intermediate steps and pass from $\phi(l)$ to $\phi(l+1)$ directly, decomposing in the passage from $\phi(l)$ to $\phi(l+1)$ the domain $\{|x_2| \leq \varepsilon_l x_1^{a_l}\}$ into the bigger transition domain

$$E_l := \{N_{l+1} x_1^{a_{l+1}} < |x_2| \leq \varepsilon_l x_1^{a_l}\}$$

and the $\kappa^{l+1}$-homogeneous domain

$$H_{l+1} := \{|x_2| \leq N_{l+1} x_1^{a_{l+1}}\},$$

where $N_{l+1}$ will be any sufficiently large constant.

We may and shall therefore assume that the sequence $\{B_l\}$ is strictly decreasing, until our algorithm stops at step $L$. In particular, we have $L < B_1$, so that the number of all domains on which our algorithm will stop is finite (notice, however, that the domains arising in the course of the algorithm will depend on the choices of roots $d_j$ at every step). The corresponding domains will cover $\Omega$, so that it will suffice to study the contributions to our maximal operator of these domains.

Now, when expressed in our original coordinates $x$, then a domain on which we stop our algorithm is either a transition domain

$$E_l := \{N_{l+1} x_1^{a_{l+1}} < |x_2| - \sum_{j=2}^l d_j x_1^{a_j} \leq \varepsilon_l x_1^{a_l}\}, \quad 1 \leq l \leq L,$$

where the case $l = L$ arises only if $N(\theta(L))$ is not contained in $\{t_2 \geq B_L\}$ - otherwise, when $N(\theta(L)) \subset \{t_2 \geq B_L\}$, then we have to replace $E_L$ by the "generalized" transition domain (which is at the same time $\kappa^L$-homogeneous)

$$E_L' := \{|x_2| - \sum_{j=2}^L d_j x_1^{a_j} \leq \varepsilon_L x_1^{a_L}\},$$

where formally $a_{L+1} = \infty$ (compare Case A); or it is a domain

$$D_{l+1}(d) := \{|x_2| - \sum_{j=2}^l d_j x_1^{a_j} - dx_1^{a_{l+1}} \leq \varepsilon_{l+1} x_1^{a_{l+1}}\}, \quad 1 \leq l \leq L,$$

which is $\kappa^{l+1}$-homogeneous after applying the change of coordinates $x = \varphi(l)(y)$, where $|d| \leq N_{l+1}$, and where $\kappa^{l+1}(n-m_1)+\kappa^{l+1} \leq 1$, in case that $d = d_{l+1}$ is a real root of $\partial_2(\theta(l))_{\kappa^{l+1}(1, \cdot)}$. The case $l = L$ can here only arise if $N_{l}(\phi(L))$ is not contained in $\{t_2 \geq B_L\}$, and if $\kappa^{L+1}(n-m_1)+\kappa^{L+1} > 1$, then there is no real root of $\partial_2(\theta(L))_{\kappa^{L+1}(1, \cdot)}$.

The contribution to the oscillatory integral $J_{\rho_0}$ of a domain $E_l$, after changing to the coordinates $y$ given by $\varphi(l)$ in the integral, can be put into the form

$$J_{\lambda_l}(\xi) := \int_{\mathbb{R}^2} e^{i\Phi(l)(y, \xi)} \tilde{n}(y) \tau_l(y) dy.$$
where we put
\[ \tau_l(y) := \rho \left( \frac{y_2}{\varepsilon_l y_1^q} \right) \left( 1 - \rho \left( \frac{y_2}{N_{l+1} y_1^{q+1}} \right) \right), \]
if \( \mathcal{N}(\theta_{(l)}) \) is not contained in \( \{t_2 \geq B_l\} \), respectively
\[ \tau_l(y) := \rho \left( \frac{y_2}{\varepsilon_l y_1^q} \right), \]
if \( \mathcal{N}(\theta_{(l)}) \subset \{t_2 \geq B_l\} \); of course, this will here only be possible for \( l = L \) and will then correspond to the domain \( E_L' \). Here,
\[ \Phi_{(l)}(y, \xi) := (\xi_3 \phi_{(l)}(y_1, 0) + \xi_1 y_1 + \xi_2 \psi_{(l)}(y_1)) + (\xi_3 \theta_{(l)}(y_1, y_2) + \xi_2 y_2). \]
Similarly, the contribution of a domain \( D_{l+1}(d) \) is of the form
\[ J^{\rho_{l+1}}(\xi) := \int_{\mathbb{R}_+^2} e^{\Phi_{(l)}(y, \xi)} \eta(y) \rho_{l+1}(y_1, y_2 - dy_1^{q+1}) dy, \]
where
\[ \rho_{l+1}(y) := \rho \left( \frac{y_2}{\varepsilon_{l+1} y_1^{q+1}} \right). \]

At this point, it will again be helpful to defray the notation by writing \( \phi \) in place of \( \phi_{(l)} \), \( \psi \) in place of \( \psi_{(l)} \) etc., and assuming that \( \phi, \psi \) and \( \theta \) satisfy the following assumptions on \( \mathbb{R}_+^2 \):

**Assumptions 7.7.** The functions \( \phi \) and \( \eta \) are smooth functions of \( x_1^{1/r} \) and \( x_2 \), and \( \psi \) is a smooth function of \( x_1^{1/r} \), where \( r \) is a positive integer. If we write \( \phi(x_1, x_2) = \phi(x_1, 0) + \theta(x_1, x_2) \), then the following hold true:

(i) The Newton diagram \( \mathcal{N}_q(\theta) \) contains at least the faces
\[ \gamma_1 = [(A_0, B_0), (A_1, B_1)], \ldots, \gamma_l = [(A_{l-1}, B_{l-1}), (A_l, B_l)], \]
where \( B_1 > B_2 > \cdots > B_l \), so that \( \gamma_j \) is an edge, if \( j > 1 \), and \( B_1 > h(\phi) \geq 2 \), and in case that \( \mathcal{N}(\theta) \) is not contained in \( \{t_2 \geq B_l\} \), it contains the additional edge \( \gamma'_{l+1} = [(A_l, B_l), (A'_{l+1}, B'_{l+1})] \). The face \( \gamma_j \) lies on the line \( \kappa_1 t_1 + \kappa_2 t_2 = 1 \), where \( \kappa_1 = \kappa_2 \). Putting \( a_j := \frac{\kappa_1}{\kappa_1} \), we have
\[ a = a_1 < \cdots < a_j < a_{j+1} < \cdots. \]

(ii) We have
\[ \phi(x_1, 0) = x_1^n (1 + O(x_1^{1/r})), \quad \psi(x_1) = b_1 x_1^{m_1} (1 + O(x_1^{1/r})), \]
where \( n = 1/\kappa_1 > \kappa_2/\kappa_1 = a > m_1 \geq 2 \).

With these data, we define the phase function
\[ \Phi(x, \xi) := (\xi_3 \phi(x_1, 0) + \xi_1 x_1 + \xi_2 \psi(x_1)) + (\xi_3 \theta(x_1, x_2) + \xi_2 x_2). \]
and the oscillatory integrals

\[ J^{\tau}(\xi) := \int_{\mathbb{R}^2} e^{i\Phi(x,\xi)} \eta(x) \tau_l(x_1, x_2) \, dx, \]

and

\[ J^{\rho_{l+1}}(\xi) := \int_{\mathbb{R}^2} e^{i\Phi(x,\xi)} \eta(x) \rho_{l+1}(x_1, x_2 - dx_1^{a_{l+1}}) \, dx, \]

where again \( \eta \) denotes a smooth bump function supported in a sufficiently small neighborhood \( \Omega \) of the origin and \( \tau_l \) and \( \rho_{l+1} \) are defined as before, only with \( \theta_l \) replaced by \( \theta \).

The maximal operators corresponding to the Fourier multipliers \( e^{i\xi_3} J^{\tau} \) and \( e^{i\xi_3} J^{\rho_{l+1}} \) will again be denoted by \( M^{\tau} \) and \( M^{\rho_{l+1}} \), respectively.

In view of our previous discussion, and since we had \( h(\phi_l) = \frac{1}{|\kappa|} \), what remains to be proven is the following

**Proposition 7.8.** Assume that the neighborhood \( \Omega \) of the point \((0,0)\) is chosen sufficiently small. Then the following hold true:

(a) The maximal operator \( M^{\tau} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > \frac{1}{|\kappa|} \).

(b) The maximal operator \( M^{\rho_{l+1}} \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > \frac{1}{|\kappa|} \) provided that \( \kappa_{l+1}(n - m_1) + \kappa_{l+2} \leq 1 \) in case that \( d = d_{l+1} \) is a real root of \( \partial_2 \theta_{\kappa_{l+1}}(1, \cdot) \).

The proof will make use of estimates for oscillatory integrals with small parameters which will be given in the next section.

8. Proof of Proposition 7.8

8.1. **Estimation of** \( J^{\tau} \). Let us first assume that \( \mathcal{N}(\theta) \) is not contained in \( \{ t_2 \geq B_l \} \), so that \( \tau_l(x) := \rho \left( \frac{x_{1}\tau_l}{\varepsilon_1 x_1} \right) (1 - \rho) \left( \frac{x_2}{N_{l+1}x_2^{a_{l+1}}} \right) \). Arguing in a similar way as in Subsection 6.5, we then consider a dyadic partition of unity \( \sum_{k=0}^{\infty} \chi_k(s) = 1, \ (0 < s < 1) \) on \( \mathbb{R} \), 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 again

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

Then

\[ J^{\tau} = \sum_{j,k} J_{j,k}, \]
where

\[ J_{j,k}(\xi) := \int_{\mathbb{R}^n_+} e^{i\Phi(x,\xi)} \eta(x) \gamma_{j,k}(x) \, dx \]

\[ = 2^{-j-k} \int_{\mathbb{R}^n_+} e^{i\Phi_{j,k}(x,\xi)} \eta_{j,k}(x) \chi \otimes \chi(x) \, dx, \]

with \( \Phi_{j,k}(x, \xi) := \Phi(2^{-j}x_1, 2^{-k}x_2, \xi) \), and where the functions \( \eta_{j,k} \) are uniformly bounded in \( C^\infty \). The summation in (8.1) takes place over pairs \((j, k)\) satisfying

\[ (8.2) \quad a_l j + M \leq k \leq a_{l+1} j - M, \]

where \( M \) can still be chosen sufficiently large, because we had the freedom to choose \( \varepsilon_1 \) sufficiently small and \( N_{l+1} \) sufficiently large. In particular, we have \( j \sim k \).

Moreover, our Assumptions 7.7 on the Newton diagram of \( \theta \) imply exactly as in Subsection 6.5 that

\[ \theta_{j,k}(x) = 2^{-(A_j+B_k)} \left( c_l x_1^{A_l} x_2^{B_l} + O(2^{CM}) \right), \]

for some constants \( c_l \neq 0 \) and \( C > 0 \). Notice also that \( B_l > B_{l+1} \geq 1 \) here, so that \( B_l \geq 2 \), and that we are here only interested in the domain where

\[ x_1 \sim 1 \sim x_2. \]

In combination with our further assumptions in 7.7, we thus obtain

\[ \Phi_{j,k}(x, \xi) = 2^{-jn} \xi_3 x_1^n (1 + v_{j,k}(x_1)) + 2^{-jm} b_1 x_1^{m_1} (1 + w_{j,k}(x_1)) + 2^{-j} x_1\]

\[ + 2^{-(A_j+B_k)} \xi_3 \left( c_l x_1^{A_l} x_2^{B_l} + u_{j,k}(x_1, x_2) \right) + 2^{-k} \xi_2 x_2, \]

where the functions \( v_{j,k}, w_{j,k} \) and \( u_{j,k} \) are of order \( O(2^{-\delta(j+k)}) \) respectively \( O(2^{-\delta M}) \) in \( C^\infty \) for some \( \delta > 0 \).

**Remark 8.1.** More precisely, the functions \( v_{j,k}, w_{j,k} \) and \( u_{j,k} \) depend smoothly on the small parameters \( \delta_1 := 2^{-j/r} \) and \( \delta_2 := 2^{-k} \) respectively \( \delta_3 := 2^{-M} \) and vanish identically for \( \delta_1 = \delta_2 = 0 \) respectively \( \delta_3 = 0 \).

Assuming again without loss of generality that \( \lambda := \xi_3 > 0 \), we may thus write

\[ \Phi_{j,k}(x, \xi) = 2^{-jn} \lambda F_{j,k}(x, s, \sigma), \]

with

\[ F_{j,k}(x, s, \sigma) := x_1^n (1 + v_{j,k}(x_1)) + S_2 x_1^{m_1} (1 + w_{j,k}(x_1)) + s_1 x_1 \]

\[ + \sigma \left( c_l x_1^{A_l} x_2^{B_l} + u_{j,k}(x_1, x_2) + s_2 x_2 \right), \]

and

\[ (8.3) \quad s_1 := 2^{(n-1)j} \frac{\xi_1}{\lambda}, \quad s_2 := 2^{A_l+B_l-1)k} \xi_2 \frac{\xi_2}{\lambda}, \quad S_2 := 2^{n-m_1} b_1 \xi_2 \frac{\xi_2}{\lambda}, \quad \sigma = \sigma_{j,k} := 2^{nj-A_l-B_l k}. \]
Lemma 8.2. Under Assumptions (7.7), the following hold true:

(a) The sequence \( \{ \frac{1}{\kappa_{ij}} \} \) is increasing and the sequence \( \{ \frac{1}{\kappa_{ij}^2} \} \) is decreasing.

(b) For \( j, k \) satisfying (8.2) we have

\[
\frac{j}{\kappa_1} \ll A_t j + B_t k \ll \frac{k}{\kappa_2}.
\]

In particular,

\[
\frac{j}{\kappa_1} = nj \ll A_t j + B_t k \ll \frac{k}{\kappa_2}.
\]

Proof. (a) is evident from the geometry of the Newton diagram of \( \theta \). It follows also from the identity (4.4) in [11], according to which

\[
\frac{1}{\kappa_{ij}^2} = \frac{A_m}{a_m} + B_m = \frac{A_{m-1}}{a_m} + B_{m-1},
\]

\[
\frac{1}{\kappa_{ij}} = A_m + a_m B_m = A_{m-1} + a_m B_{m-1},
\]

since the sequence \( \{ a_m \} \) is increasing.

(b) is a consequence of (a) and the identities above. Q.E.D.

Since \( B_t \geq 1 \) and \( n > m_1 \geq 2 \), in combination with Lemma 8.2 we see that

\[
\sigma \ll 1, \ |\xi_1| \ll \lambda |s_1|, \ |\xi_2| \ll \lambda |s_2|, \text{ and also } \ |\xi_2| \ll \lambda |S_2|.
\]

Proposition 8.3. If \( M \) in (8.2) is chosen sufficiently large, then the following estimate

\[
|J_{j,k}(\xi)| \leq C||\eta||_{C^\alpha(\mathbb{R}^3)} 2^{-j-k} (1 + 2^{-nj} |\xi|)^{-1/3} (1 + 2^{-nj} \sigma_{j,k} |\xi|)^{-1/2}
\]

holds true, where the constant \( C \) does not depend on \( j, k \) and \( \xi \).

Consequently, the maximal operator \( M^n \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > 1/|\kappa| \).

Proof. We first notice that \( B_t \geq 2 \), so that \( \partial^2_{x_1} (x_1^A x_2^B) \sim 1 \).

As in the proof of Proposition 7.5 we shall distinguish several cases.

1. Case. \( |s_1| + |S_2| \ll 1 \), or \( |s_1| + |S_2| \gg 1 \) and \( |s_1| \ll |S_2| \) or \( |s_1| \gg |S_2| \).

Here, an integration by parts in \( x_1 \) yields

\[
J_{j,k}(\xi) = O \left( 2^{-j-k} (1 + 2^{-nj} \lambda (1 + |s_1| + |S_2|) \right)^{-1}),
\]

which implies (8.5) because of (8.4).

2. Case. \( |s_1| + |S_2| \gg 1 \) and \( |s_1| \sim |S_2| \).

Since \( m_1 \geq 2 \), we have \( \partial^2_{x_1} (x_1^{m_1}) \sim 1 \). Therefore, if \( s_2 \) with \( |s_2| \lesssim 1 \) is fixed, in view of Remark 8.1 we can apply Proposition 9.1 in a similar way as in the proof of Proposition 7.5, with \( \lambda \) replaced by \( 2^{-nj} \lambda (|s_1| + |S_2|) \), and obtain

\[
|J_{j,k}(\xi)| \leq C 2^{-j-k} (1 + 2^{-nj} \lambda (1 + |s_1| + |S_2|) \right)^{-1/2} (1 + 2^{-nj} \sigma \lambda (1 + |S_2|) \right)^{-1/2}.
\]
In fact, the proposition even shows that this estimate remains valid under small perturbations of $s_2$, so that we can choose the constant $C$ uniformly for $s_2$ in a fixed, compact interval.

On the other hand, if $|s_2| \gg 1$, we can obtain the even stronger estimate where the second exponent $-1/2$ is replaced by $-1$ by first integrating by parts in $x_2$ and then applying the method of stationary phase in $x_1$.

Observe at this point that if $|\xi_1| + |\xi_2| \geq \lambda$, so that $|\xi| \sim |\xi_1| + |\xi_2|$, then by (8.4)

$$|s_1| + |S_2| \gg 1.$$  

Notice also that $|s_1| \sim |S_2|$ implies, by (8.3), that $1 \sim 2^{-(m_1-1)}|\xi_2|/|\xi_1|$, hence

$$|\xi_1| \ll |\xi_2|.$$  

Thus, if $|\xi_1| + |\xi_2| \geq \lambda$ and $|s_1| \sim |S_2|$, then $|\xi| \sim |\xi_2|$, and since $|s_2|\lambda \gg |\xi_2|$, we see that (8.6) implies (8.5) in this case, as well as of course in the case where $|\xi_1| + |\xi_2| \leq \lambda$. We are thus left with the case

3. Case. $|s_1| + |S_2| \sim 1$ and $|\xi_1| + |\xi_2| \leq \lambda$, hence $|\xi| \sim \lambda$.

Since $n > m_1$, it is easy to see that in this case the polynomial $p(x_1) := x_1^n + S_2b_1x_1^{m_1} + s_1x_1$ satisfies $|p''(x_1)| + |p'''(x_1)| \neq 0$ for every $x_1 \sim 1$. Therefore, if we fix some point $x_1^0 \sim 1$, then we can either apply Proposition 9.1 or Proposition 9.2 if we localize the oscillatory integral $J_{j,k}(\xi)$ by means of a suitable cut-off function to a small neighborhood of $x_1^0$ and translate coordinates, and finally obtain by means of a suitable partition of unity in a similar way as in the previous case that

$$|J_{j,k}(\xi)| \leq C(2)^{-j-k}(1 + 2^{-nj}\lambda)^{-1/3}(1 + 2^{-nj}\lambda(1 + |s_2|))^{-1/2},$$

hence (8.5). Note again that this argument first applies for fixed $s_1, s_2, S_2$, but since Propositions 9.1 and 9.2 allow for small perturbations of parameters, the estimate above will hold uniformly in $s_1, s_2, S_2$.

Next, observe that we may replace the factor $(1 + 2^{-nj}\sigma_{j,k}|\xi|)^{-1/2}$ in (8.5) by $(1 + 2^{-nj}\sigma_{j,k}|\xi|)^{-1/6-\varepsilon}$, for any sufficiently small $\varepsilon > 0$, which leads to

$$|J_{j,k}(\xi)| \leq C ||\eta|| |_{C^3(\mathbb{R}^2)} 2^{-j-k} 2^{\frac{n}{j}2} 2^{A_{ij} + B_{ij}j}(\frac{1}{\lambda} + |\xi|)^{-\frac{1}{2} - \varepsilon} \leq C ||\eta|| |_{C^3(\mathbb{R}^2)} 2^{-j-k} 2^{\frac{n}{j}2} 2^{\frac{k}{\sigma_{j,k}^2} + \frac{1}{\lambda} + |\xi|)^{-\frac{1}{2} - \varepsilon},$$

since Lemma 8.1 shows that $A_{ij} + B_{ij}k < \frac{k}{\sigma_{j,k}^2} \leq \frac{k}{\kappa_2}$.

Lemma 7.4 then implies that the maximal operators $\mathcal{M}^{j,k}$ associated to the multipliers $J_{j,k}$ can be estimated by

$$||\mathcal{M}^{j,k}f||_p \leq C 2^{-j-k} 2^{\frac{n}{j}2} 2^{\frac{k}{\sigma_{j,k}^2} + \frac{1}{\lambda} + |\xi|)^{-\frac{1}{2} - \varepsilon}} ||f||_p$$

for every sufficiently small $\varepsilon > 0$ and $p \geq 2$. 

MAXIMAL OPERATORS ASSOCIATED TO HYPERSURFACES 73
Observe that for \( p = \frac{1}{|\kappa|} \), we have
\[
\frac{2}{3\kappa_1p} = \frac{2}{3} \frac{\kappa_1 + \kappa_2}{\kappa_1} = \frac{2}{3} (1 + a) > 1,
\]
so that for \( p > \frac{1}{|\kappa|} \) sufficiently close to \( \frac{1}{|\kappa|} \), we have
\[
\sum_{a_{ij} + M \leq k} 2^{-j-k} \frac{2^j}{2^{a_{11}} 2^{\frac{k}{2}p} (1 + \epsilon)} \leq \sum_{j \leq \frac{k}{a}, k \geq M} 2^{-j-k} \frac{2^j}{2^{a_{11}} 2^{\frac{k}{2}p} (1 + \epsilon)}
\]
\[(8.7) \leq \sum_{k \geq M} 2^{\frac{2}{3} (1 + a - \delta - 1) \frac{k}{a} - k + 2 \epsilon} \leq \sum_{k \geq M} 2^{2(\epsilon - \frac{a}{2})k}, \]
where \( \delta > 0 \) depends on \( p \). Choosing \( \epsilon \) sufficiently small, this series converges, so that \( \mathcal{M}_n \) is bounded on \( L^p \). For \( p = \infty \), the series converges as well. By real interpolation, we thus find that \( \mathcal{M}_n \) is \( L^p \)-bounded for every \( p > \frac{1}{|\kappa|} \).

Q.E.D.

The case where \( N(\theta) \subset \{ t_2 \geq B_l \} \) can be treated in a very similar way, if we formally replace \( a_{l+1} \) by \( +\infty \). Indeed, in this case we have \( \tau_l(x) := \rho \left( \frac{x_2}{c_{l+1} x_1^d} \right) \), so that condition \( (8.2) \) has to be replaced by
\[(8.8) \quad a_{ij} + M \leq k.\]
Moreover, in this case we obviously have
\[ \theta_{j,k}(x) = 2^{-(A_{ij} + B_k)} b \left( c_{l+1} x_1^d + O(2^{-\delta(j+k)}) \right) \]
for some \( \delta > 0 \). Therefore, if \( B_l \geq 2 \), we can argue exactly as before and see that Proposition 8.3 remains valid (notice that in \( (8.7) \) we only made use of \( (8.8) \)).

What remains open at this stage is the case where \( B_l = 1 \). It turns out that here the oscillatory integrals \( J_{j,k}(\xi) \) may possibly be of degenerate Airy type. We shall then need more detailed information, which we shall obtain by regarding \( \mathcal{M}_n \) rather as a maximal operator of type \( \mathcal{M}^{\rho_l} \), which will be treated in the next subsection.

8.2. Estimation of \( J^{\rho_{l+1}} \). We now consider the maximal operators \( \mathcal{M}^{\rho_{l+1}} \) in Proposition 7.8 (b). It will here be convenient to change to the \( \kappa^{l+1} \)-homogeneous coordinates
\[ y_1 := x_1, y_2 := x_2 - d x_1^{\kappa_{l+1}}. \]
This change of coordinates has the effect that we can assume that \( d = 0 \). The Newton diagram of \( \theta \) in the new coordinates will still contain the faces \( \gamma_1, \ldots, \gamma_l \), but the edge \( \gamma_{l+1} = [(A_l, B_l), (A_{l+1}, B_{l+1})] \) may change to an interval \( [(A_l, B_l), (A_{l+1}, B_{l+1})] \) on the same line \( \kappa_1^{l+1} t_1 + \kappa_2^{l+1} t_2 = 1 \), but possibly with a different right endpoint \( (A_{l+1}, B_{l+1}) \), which may even coincide with the left endpoint \( (A_l, B_l) \).
Simplifying the notation by writing \( \kappa' := \kappa^{l+1} \) and \( a' := \frac{\kappa'}{\kappa_1} = a_{l+1} \), we shall then have to estimate the oscillatory integral \( J(\xi) = J^{\rho_{l+1}}(\xi) \), with

\[
J(\xi) := \int_{\mathbb{R}^2_+} e^{i\Phi(x,\xi)} \eta(x) \rho\left(\frac{x_2}{\varepsilon_1 x_1^{\alpha}}\right) dx,
\]

where \( \varepsilon' = \varepsilon_{l+1} > 0 \) can still be chosen as small as we like, under one of the following assumptions:

(i) \( \partial_2 \theta_{\kappa'}(1, 0) = 0 \), i.e., \( B_{l+1} = 2 \), and \( \kappa'_1(n - m_1) + \kappa'_2 \leq 1 \).

(ii) \( \partial_2 \theta_{\kappa'}(1, 0) \neq 0 \), i.e., \( B_{l+1} = 1 \), and \( \kappa'(n - m_1) + \kappa'_2 \neq 1 \).

(iii) \( \partial_2 \theta_{\kappa'}(1, 0) \neq 0 \), i.e., \( B_{l+1} = 1 \), and \( \kappa'(n - m_1) + \kappa'_2 = 1 \).

The most delicate case is case (iii), which will lead to degenerate Airy-type integrals. Notice that the second condition in (iii) just means that the point \((n - m_1, 1) = (A_{l+1}, B_{l+1})\) belongs to \( N_a(\theta) \).

We shall denote the maximal operator associated to the Fourier multiplier \( e^{i\xi_3 J(\xi)} \) by \( \mathcal{M}' \).

Observe at this point that the oscillatory integral \( J^n \) for the still open case where \( B_l = 1 \) can be written in the form (8.9) too, with \( \kappa' := \kappa^l \), hence \( a' = a_l \) and \( (A_l, B_l) = (A_{l+1}, B_{l+1}) \), and since \( B_l = 1 \), it will satisfy the assumption (ii) or (iii). Notice that here necessarily \( l > 1 \).

We shall therefore in the sequel relax the condition \( a' > a_l \) and assume only that \( a' \geq a_l \) in case that \((A_l, B_l) = (A_{l+1}, B_{l+1})\) and \( B_l = 1 \). Then, as in the proof of Proposition 7.5, we can decompose

\[
J = \sum_{k=k_0}^{\infty} J_k
\]

by means of a dyadic decomposition based on the \( \kappa' \)-dilations \( \delta'_\kappa(x_1, x_2) := (r^{\kappa'_1} x_1, r^{\kappa'_2} x_2) \), where the dyadic constituent \( J_k \) of \( J \) is given, after re-scaling, by

\[
J_k(\xi) = 2^{-k|\kappa'|} \int_{\mathbb{R}^2} e^{i2^{-k} x_1 n \lambda \Phi_k(x,s)} \rho\left(\frac{x_2}{\varepsilon_1 x_1^{\alpha}}\right) \eta(\delta'_{2^{-k}} \chi) dx,
\]

where again \( \lambda := \xi_3 \) is assumed to be positive, and where

\[
\Phi_k(x, s, \sigma) := x_1^n (1 + v_k(x_1)) + s_1 x_1 + S_2 b_1 x_1^{m_1} (1 + w_k(x_1)) + \sigma \left( \theta_{\kappa'}(x_1, x_2) + \theta_{r,k}(x_1, x_2) + s_2 x_2 \right),
\]

with

\[
s_1 := 2^{\kappa'_1(n-1)k} \frac{\xi_1}{\lambda}, \quad s_2 := 2^{(1-\kappa'_2)k} \frac{\xi_2}{\lambda}, \quad S_2 := 2^{\kappa'_1(n-m_1)k} \frac{\xi_2}{\lambda}, \quad \sigma = \sigma_k := 2^{(\kappa'_1 n-1)k}.
\]
In particular, we have

\begin{equation}
S'_2 = 2^{(\kappa'_1(n-1)+\kappa'_2-1)k} S_2.
\end{equation}

Moreover, since \( \kappa'_1 < \kappa_1 = 1/n \), we have \( \kappa'_1(n - 1) > 0 \) and \( \kappa'_1 n - 1 < 0 \), and since \( 1 = \kappa'_1 A_{l+1} + \kappa'_2 B_{l+1} \geq \kappa'_2 \), we have \( 1 - \kappa'_2 > 0 \), we see that if \( \Omega \) is chosen sufficiently small so that \( k_0 \gg 1 \) in (8.10), then

\begin{equation}
(8.13) \quad |\sigma| << 1, \ |\xi_1| << \lambda |s_1|, \ |\xi_2| << \lambda |s_2|, \text{ and also } |\xi_2| << \lambda |S_2|.
\end{equation}

Recall that \( \theta_{\kappa'} \) denotes the \( \kappa' \)-homogeneous part of \( \theta \). The functions \( v_k, w_k \) and \( \theta_{r,k} \) are of order \( O(2^{-\epsilon k}) \) in \( C^\infty \) for some \( \epsilon > 0 \), and can in fact be viewed as smooth functions \( v(x_1, \delta), w(x_1, \delta) \) respectively \( \theta_r(x, \delta) \) depending also on the small parameter \( \delta = 2^{-k/r} \) for some positive integer \( r > 0 \), which vanish identically when \( \delta = 0 \).

Notice again that in our domain of integration for \( J_k(\xi) \), we have

\[
x_1 \sim 1, \ |x_2| \lesssim \epsilon',
\]

and clearly \( |\mathcal{M}'f| \leq \sum_{k=0}^{\infty} |\mathcal{M}^k f| \), if \( \mathcal{M}^k \) denotes the maximal operator associated to the Fourier multiplier \( e^{i\xi_{\delta} J_k(\xi)} \).

The following proposition will then cover Proposition 7.8(b) as well as the remaining case of Proposition 7.8(a). The constants \( l_m \) and \( c_m \) will be as in Theorem 9.3. We remark at this point that clearly

\begin{equation}
1/6 \leq l_m < 1/4.
\end{equation}

**Proposition 8.4.** If \( k_0 \) in (8.10) is chosen sufficiently large and \( \epsilon' \) sufficiently small, then

\begin{equation}
|J_k(\xi)| \leq C||\eta||_{C^\infty(\mathbb{R}^2)} 2^{-|\kappa'|k} \sigma_k^{-\left(l_m + c\right)} (2^{-\kappa'_1 nk}|\xi|)^{-1/2 - \epsilon}
\end{equation}

for some \( m \in \mathbb{N} \) with \( 2 \leq m \leq B_1 \), some constant \( c > 0 \) and every sufficiently small \( \epsilon > 0 \), where the constant \( C \) does not depend on \( k \) and \( \xi \).

Consequently, the maximal operator \( \mathcal{M}' \) is bounded on \( L^p(\mathbb{R}^3) \) for every \( p > 1/|\kappa| \).

**Proof.** We proceed in a similar way as in the proof of Proposition 8.3.

1. **Case.** \(|s_1| + |S_2| << 1\), or \(|s_1| + |S_2| \gg 1\) and \(|s_1| << |S_2|\) or \(|s_1| \gg |S_2|\).

Here, an integration by parts in \( x_1 \) yields

\[
|J_k(\xi)| \leq C2^{-|\kappa'|k}(1 + 2^{-\kappa'_1 nk}\lambda(1 + |s_1| + |S_2|))^{-1},
\]

which implies (8.15) because of (8.13).

2. **Case.** \(|s_1| + |S_2| \gg 1\) and \(|s_1| \sim |S_2|\).

Observe first that for any \( x_2^0 \sim 1 \), the polynomial \( P(x_2) := \theta_{\kappa'}(x_2^0, x_2) \) has degree \( \deg P \geq 2 \). Indeed, this is clear under assumption (i), since \( B_{l+1} \geq 2 \), and under the assumptions (ii) and (iii) it follows from \( B_l \geq 2 \), respectively \( B_{l-1} \geq 2 \) in case that \( B_l = B_{l+1} = 1 \). Clearly also \( \deg P \leq B_l \).
Therefore, if $|s_2| \lesssim 1$, we can argue in a similar way as in Case 3 of the proof of Proposition 7.5, and obtain by means of Proposition 9.1 that
\begin{equation}
|J_k(\xi)| \leq C 2^{-|\kappa'|k}(1 + 2^{-\kappa'_1 nk}(1 + |s_1| + |s_2|))^{-1/m} (1 + 2^{-\kappa'_1 nk} \sigma \lambda(1 + |s_2|))^{-1/m}
\end{equation}
for some $m$ with $2 \leq m \leq B_1$, provided $\varepsilon'$ is chosen sufficiently small.

On the other hand, if $|s_2| >> 1$, we can obtain the even stronger estimate where the second exponent $-1/m$ is replaced by $-1$ by first integrating by parts in $x_2$ and then integrating in $x_1$.

Now from (8.13) we deduce as in the proof of Proposition 8.3 that if $|\xi_1| + |\xi_2| \geq \lambda$, so that $|\xi| \sim |\xi_1| + |\xi_2|$, then we have $|s_1| + |s_2| >> 1$, and $|s_1| \sim |s_2|$ implies that $|\xi_1| << |\xi_2|$.

Thus, if $|\xi_1| + |\xi_2| \geq \lambda$ and $|s_1| \sim |s_2|$, then $|\xi| \sim |\xi_2|$, and since $|s_2| \lambda >> |\xi_2|$, we see that (8.16) implies (8.15) in this case, as well as of course in the case where $|\xi_1| + |\xi_2| \leq \lambda$, provided $\varepsilon$ is chosen small enough. We are thus left with the case

3. Case. $|s_1| + |s_2| \sim 1$ and $|\xi_1| + |\xi_2| \leq \lambda$, hence $|\xi| \sim \lambda$.

Since $n > m_1$, the polynomial $p(x_1) := x_1^n + S_2 b_1 x_1^{m_1} + s_1 x_1$ satisfies $|p''(x_1)| + |p'''(x_1)| \neq 0$ for every $x_1 \sim 1$. But, if either $|s_1| << |s_2|$ or $|s_1| >> |s_2|$, then all critical points of the polynomial $x_1^n + S_2 b_1 x_1^{m_1} + s_1 x_1$ will be non-degenerate, so that we can argue exactly as in Case 2. We shall therefore assume that

$|s_1| \sim |s_2| \sim 1$.

Now, under assumption (i), we have $\partial_2 \theta_{\kappa'}(x_1^0, 0) = 0$ whenever $x_1^0 \sim 1$, whereas $|s_2| \gtrsim 1$, by (8.12), so that
\begin{equation}
\partial_2 (\theta_{\kappa'} + s_2 x_2)(x_1^0, 0) = \partial_2 \theta_{\kappa'}(x_1^0, 0) + s_2 \neq 0.
\end{equation}
The same is true also under assumption (ii), for then either $|s_2| >> 1$ or $|s_2| << 1$ (by (8.12)), whereas $\partial_2 \theta_{\kappa'}(x_1^0, 0) \neq 0$, and it also applies in case (iii), provided $|s_2| >> 1$ or $|s_2| << 1$.

In these cases, we shall first integrate by parts in $x_2$ and then apply a Björk type version of van der Corput’s lemma in $x_1$, which results in the estimate

$|J_k(\xi)| \leq C 2^{-|\kappa'|k}(1 + 2^{-\kappa'_1 nk} \lambda)^{-1/3} (1 + 2^{-\kappa'_1 nk} \sigma \lambda(1 + |s_2|))^{-1/3}$.

By replacing the second exponent $-1$ by $-1/6 - \varepsilon$, we see in view of (8.14) that this implies (8.15).

We are thus left with the case where assumption (iii) holds true, and where $|s_2| \sim 1$. Fix $x_1^0 \sim 1$. Then $\partial_1 \partial_2 (\theta_{\kappa'} + s_2 x_2)(x_1^0, 0) \neq 0$, since $\theta_{\kappa'}(x_1, x_2) = c_0 x_1^{n-m_1} x_2 + O(x_2^2)$, where $c_0 \neq 0$.

Assume first that (8.17) holds true. Then we can again argue as before, provided we introduce in our formula for $J_k(\xi)$ an additional smooth cut-off function $a(x_1)$ supported in a sufficiently small neighborhood of $x_1^0$. 
So, assume next that $\partial^2_2(\theta_\kappa + s_2x_2)(x_1^0, 0) = 0$. Since the degree of the polynomial $P(x_2) := \theta_\kappa(x_1^0, x_2)$ satisfies $B_t \geq \deg P \geq 2$, after shifting the $x_1$ coordinates by $x_1^0$, we can apply Theorem 9.3 and obtain estimate (8.15) for some $m$ with $2 \leq m \leq B_t$, if we again introduce a cut-off function $a(x_1)$ in a sufficiently small neighborhood of $x_1^0$ into $J_k(\xi)$. Recall here that the functions $v_k, w_k$ and $\theta_{r,k}$ are smooth functions $v(x_1, \delta), w(x_1, \delta)$ respectively $\theta_r(x, \delta)$ depending also on the small parameter $\delta = 2^{-k/r}$ for some positive integer $r > 0$, which vanish identically when $\delta = 0$.

The estimate (8.15) then follows by decomposing $J_k(\xi)$ into a finite number of such "localized" integrals by means of a partition of unity. Next, in order to estimate the maximal operator $M'$, observe that (8.15) implies that for any sufficiently small $\varepsilon > 0$ we have

$$|J_k(\xi)| \leq C||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} 2^{-\frac{2}{k}nk(1/2+\varepsilon)} 2^{(1-\kappa'1n)k(l_m+c\varepsilon)} (1 + |\xi|)^{-1/2-\varepsilon}.$$  

Recalling that $1 - \kappa'_1n > 0$ and $l_m < 1/4$ by (8.14), we thus see that there is some $\delta > 0$ such that

$$|J_k(\xi)| \leq C||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} 2^{-\frac{2}{k}nk(1+|\kappa'|n)k(l_m+c\varepsilon)} (1 + |\xi|)^{-1/2-\varepsilon},$$

provided $\varepsilon$ is sufficiently small. Lemma 7.4 then implies

$$||M^k f||_p \leq C2^{-\frac{2}{k}nk} 2^{-\frac{2}{k}nk(1+|\kappa'|n)k(l_m+c\varepsilon)} ||f||_p$$

for every $p \geq 2$. Notice that

$$\frac{1 + \kappa'_1n}{2|\kappa'|} \leq 1.$$  

Indeed, we have $t := \kappa'_1n = \frac{\kappa'}{\kappa_1} \leq 1$ and $\kappa'_2 \geq \kappa_2$ by Lemma 8.2, so that

$$\frac{1 + \kappa'_1n}{2|\kappa'|} = \frac{1 + t}{2(\kappa_1t + \kappa'_2)} \leq \frac{1 + t}{2(\kappa_1t + \kappa_2)}.$$  

The latter function is increasing in $t$, so that we may replace $t$ by 1 and obtain (8.18).

The estimate (8.18) shows that the norms of the maximal operators $M^k$ sum in $k$ when $p \geq \frac{1}{|\kappa'|}$, which concludes the proof of Proposition 8.4, hence also the proof of our main result, Theorem 1.2.

Q.E.D.

9. ESTIMATES FOR OSCILLATORY INTEGRALS WITH SMALL PARAMETERS

In this section, we shall provide the estimates for oscillatory integrals that were needed in the previous sections. More precisely, we shall study oscillatory integrals

$$J(\lambda, \sigma, \delta) := \int_{\mathbb{R}^2} e^{i\lambda F(x, \sigma, \delta)} \psi(x, \delta) \, dx, \quad (\lambda > 0),$$

with a phase function $F$ of the form

$$F(x, \sigma, \delta) = \sum_{j=1}^N a_j(z_j(x, \sigma, \delta)),$$
\[ F(x_1, x_2, \sigma, \delta) := f_1(x_1, \delta) + \sigma f_2(x_1, x_2, \delta), \]
and an amplitude \( \psi \) defined for \( x \) in some open neighborhood of the origin in \( \mathbb{R}^2 \) with compact support in \( x \). The functions \( f_1, f_2 \) are assumed to be real-valued and will depend, like the function \( \psi \), smoothly on \( x \) and on small real parameters \( \delta_1, \ldots, \delta_\nu \), which form the vector \( \delta := (\delta_1, \ldots, \delta_\nu) \in \mathbb{R}^\nu \). \( \sigma \) denotes a small real parameter.

With a slight abuse of language we shall say that \( \psi \) is compactly supported in some open set \( U \subset \mathbb{R}^2 \) if there is a compact subset \( K \subset U \) such that \( \text{supp} \psi(\cdot, \delta) \subset K \) for every \( \delta \).

### 9.1. Oscillatory integrals with non-degenerate critical points in \( x_1 \).

**Proposition 9.1.** Assume that
\[
|\partial_1 f_1(0, 0)| + |\partial^2_1 f_1(0, 0)| \neq 0,
\]
and that there is some \( m \geq 2 \) such that
\[
\partial^m_2 f_2(0, 0, 0) \neq 0.
\]
Then there exists a neighborhood \( U \subset \mathbb{R}^2 \) of the origin and some \( \varepsilon > 0 \) such that for any \( \psi \) which is compactly supported in \( U \) the following estimate
\[
|J(\lambda, \sigma, \delta)| \leq \frac{C\|\psi(\cdot, \delta)\|_{C^3}}{(1 + \lambda)^{1/2}(1 + |\lambda\sigma|)^{1/m}}
\]
holds true uniformly for \( |\sigma| + |\delta| < \varepsilon \).

**Proof.** If \( \partial_1 f_1(0, 0) \neq 0 \), then we can integrate by parts in \( x_1 \) if \( \lambda > 1 \) and obtain the stronger estimate
\[
|J(\lambda, \sigma, \delta)| \leq \frac{C\|\psi(\cdot, \delta)\|_{C^3}}{1 + \lambda}.
\]
Assume therefore that \( \partial_1 f_1(0, 0) = 0 \), so that the mapping \( x_1 \mapsto f_1(x_1, 0) \) has a non-degenerate critical point at \( x_1 = 0 \). Then, by the implicit function theorem, for \( |\delta| \) sufficiently small there exists a unique critical point \( x_1 = x^0_1(\delta) \) depending smoothly on \( \delta \) of the mapping \( \xi \mapsto f_1(x_1, \delta) = 0 \), i.e., \( \partial_1 f_1(x^0_1(\delta), \delta) \equiv 0 \), where \( x^0_1(0) = 0 \).

In a similar way, we see that there is a unique, smooth function \( x^*_1(x_2, \sigma, \delta) \) for \( |x_2| + |\sigma| + |\delta| \) sufficiently small such that
\[
\partial_1 F(x^*_1(x_2, \sigma, \delta), x_2, \sigma, \delta) \equiv 0,
\]
where \( x^*_1(0, 0, 0) = 0 \). By comparison, we see that \( x^*_1(x_2, 0, \delta) = x^0_1(\delta) \), so that
\[
x^*_1(x_2, \sigma, \delta) = x^0_1(\delta) + \sigma \gamma(x_2, \sigma, \delta)
\]
for some smooth function \( \gamma \). Applying the stationary phase formula with parameters to the integration in \( x_1 \), we thus obtain
\[
J(\lambda, \sigma, \delta) = \int_{\mathbb{R}} e^{i\lambda \phi(x_2, \sigma, \delta)} a(\lambda, x_2, \sigma, \delta) \, dx_2,
\]
where
\[ \phi(x_2, \sigma, \delta) := F(x_1^0(\delta) + \sigma \gamma(x_2, \delta, \sigma), x_2, \sigma, \delta), \]
and where \( a(\lambda, x_2, \sigma, \delta) \) is a symbol of order \(-1/2\) in \( \lambda \), so that in particular
\[ |\partial_x^l a(\lambda, x_2, \sigma, \delta)| \leq C_l(1 + |\lambda|)^{-1/2}, \]
with constants \( C_l \) which are independent of \( x_2, \sigma \) and \( \delta \) (see, e.g., Sogge [25] or Hörmander [10]).

Moreover, a Taylor series expansion of \( \phi \) with respect to \( \sigma \) near \( \sigma = 0 \) shows that
\[ \phi(x_2, \sigma, \delta) = f_1(x_1^0(\delta), \delta) + \sigma f_2(x_1^0(\delta), x_2, 0, \delta) + O(\sigma) \]
in \( C^\infty \). Since \( \partial_{x_2}^3 f_1(0, 0) \neq 0 \), for \( |\sigma| \) sufficiently small we can thus apply van der Corput’s lemma (cf.[28]) to the integral \(9.2\) in \( x_2 \) and obtain the estimate \(9.1\).

Q.E.D.

9.2. Oscillatory integrals of non-degenerate Airy type.

Proposition 9.2. Assume that
\[ \partial^3 f_1(0, 0) \neq 0 \text{ and } \partial^2_x f_2(0, 0, 0) \neq 0. \]

Then there exists a neighborhood \( U \subset \mathbb{R}^2 \) of the origin and some \( \varepsilon > 0 \) such that for any \( \psi \) which is compactly supported in \( U \) the following estimate
\[ |J(\lambda, \sigma, \delta)| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^3}}{(1 + |\lambda\sigma|)^{1/2}(1 + |\lambda\sigma|)^{1/2}} \]
holds true uniformly for \( |\sigma| + |\delta| < \varepsilon \).

Proof. Consider first the case where \( \partial^2_x f_2(0, 0, 0) \neq 0 \). Then, if \( |\lambda\sigma| \gg 1 \), we first perform an integration by parts in \( x_2 \). Subsequently, we can apply van der Corput’s lemma to the integration in \( x_1 \), provided \( U \) and \( \varepsilon \) are chosen sufficiently small, and obtain the stronger estimate
\[ |J(\lambda, \sigma, \delta)| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^2}}{(1 + |\lambda\sigma|)^{1/2}(1 + |\lambda\sigma|)^{1/2}}. \]

Now, assume that \( \partial^2_x f_2(0, 0, 0) = 0 \) but \( \partial^3_x f_2(0, 0, 0) \neq 0 \). Then for \( U \) and \( \varepsilon \) chosen sufficiently small, by the implicit function theorem there exists a unique critical point \( x_2^c(x_1, \delta) \) of the function \( x_2 \mapsto f_2(x_1, x_2, \delta) \). Then, by applying the stationary phase method with small parameters to the \( x_2 \)-integration, we see that
\[ J(\lambda, \sigma, \delta) = \int_{\mathbb{R}} e^{i\lambda\phi(x_1, \sigma, \delta)} a(\lambda\sigma, x_1, \delta) \, dx_1, \]
where
\[ \phi(x_1, \sigma, \delta) := f_1(x_1, \delta) + \sigma f_2(x_1, x_2^c(x_1, \delta), \delta), \]
and where \( a(\lambda, x_1, \delta) \) is a symbol of order \(-1/2\) in \( \lambda \), so that in particular
\[
|\partial_{x_1}^l a(\lambda \sigma, x_1, \delta)| \leq C_l (1 + |\lambda \sigma|)^{-1/2},
\]
with constants \( C_l \) which are independent of \( x_1 \) and \( \delta \).

We can now apply van der Corput’s lemma to the integral (9.5) and obtain in view of (9.6) the desired estimate (9.4).

Q.E.D.

9.3. Oscillatory integrals of degenerate Airy type.

**Theorem 9.3. Assume that**
\[
|\partial_1 f_1(0, 0)| + |\partial_1^2 f_1(0, 0)| + |\partial_1^3 f_1(0, 0)| \neq 0 \quad \text{and} \quad \partial_1 \partial_2 f_2(0, 0, 0) \neq 0,
\]
and that there is some \( m \geq 2 \) such that
\[
\partial_2^l f_2(0, 0, 0) = 0 \quad \text{for} \quad l = 1, \ldots, m-1 \quad \text{and} \quad \partial_2^m f_2(0, 0, 0) \neq 0.
\]

Then there exists a neighborhood \( U \subset \mathbb{R}^2 \) of the origin and constants \( \varepsilon, \varepsilon' > 0 \) such that for any \( \psi \) which is compactly supported in \( U \) the following estimate
\[
|J(\lambda, \sigma, \delta)| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^3}}{\lambda^{\varepsilon + \varepsilon'} |\sigma|^{l_m + c_m \varepsilon}}
\]
holds true uniformly for \(|\sigma| + |\delta| < \varepsilon'\), where \( l_m := \frac{1}{6} \) and \( c_m := 1 \) for \( m < 6 \), and \( l_m := \frac{m-3}{2(2m-3)} \) and \( c_m := 2 \) for \( m \geq 6 \).

**Remark 9.4.** If \(|\partial_1 f_1(0, 0)| + |\partial_1^2 f_1(0, 0)| \neq 0\), then a stronger estimate than (9.9) follows from Proposition 9.1, since \( 1/6 \leq l_m < 1/4 \). The full thrust of Theorem 9.3 therefore lies in the case where \( \partial_1 f_1(0, 0) = \partial_1^2 f_1(0, 0) = 0 \) and \( \partial_1^3 f_1(0, 0) \neq 0 \), on which we shall concentrate in the sequel.

The proof of Theorem 9.3 will be an immediate consequence of the following two lemmas. Our first lemma allows to reduce the phase function \( F \) to some normal form and is based on Martinet’s theorem.

**Lemma 9.5. Assume that the function \( F \) satisfies the conditions of Theorem 9.3, and in addition that \( \partial_1 f_1(0, 0) = \partial_1^2 f_1(0, 0) = 0 \). Then there exist smooth functions \( X_1 = X_1(x_1, \sigma, \delta) \) and \( X_2 = X_2(x_1, x_2, \delta) \) defined in a sufficiently small neighborhood \( U \times V \subset \mathbb{R}^2 \times \mathbb{R}^{\nu+1} \) of the origin such that the following hold true:

(i) \( X_1(0, 0, 0) = X_2(0, 0, 0) = 0 \), \( \partial_1 X_1(0, 0, 0) \neq 0 \), \( \partial_2 X_2(0, 0, 0) \neq 0 \), so that we can change coordinates from \((x_1, x_2, \sigma, \delta)\) to \((X_1, X_2, \sigma, \delta)\) near the origin.

(ii) In the new coordinates \( X_1, X_2 \) for \( \mathbb{R}^2 \) near the origin, we can write \( F(x_1, x_2, \sigma, \delta) = g_1(X_1, \sigma, \delta) + \sigma g_2(X_1, X_2, \sigma, \delta) \), with \( g_1(X_1, \sigma, \delta) = X_1^3 + a_1(\sigma, \delta) X_1 + a_m(\sigma, \delta) \).
and

\[ g_2(X_1, X_2, \sigma, \delta) = X_2^m + \sum_{j=2}^{m-2} a_j(\delta) X_2^{m-j} + \left(X_1 - a_{m-1}(\sigma, \delta)\right)X_2 b(X_1, X_2, \sigma, \delta), \]

if \( m \geq 3 \), and \( g_2(X_1, X_2, \sigma, \delta) = X_2^2 \), if \( m = 2 \), where \( a_1, \ldots, a_m \) are smooth functions of the variables \( \sigma, \delta \) such that \( a_l(0, 0) = 0 \), and where \( b \) is a smooth function such that \( b(0, 0, 0, 0) \neq 0 \).

**Proof.** In a first step, we apply Martinet’s theorem (more precisely, the special case proved in [10], Theorem 7.5.13) to the function \( f_2(x_1, x_2, \delta) \). Due to our assumption (9.8), there exists a smooth function \( X_2 = X_2(x_1, x_2, \delta) \) defined in a sufficiently small neighborhood of the origin with

\[ X_2(0, 0, 0) = 0, \quad \partial_2 X_2(0, 0, 0) \neq 0, \]

so that in the new coordinate \( X_2 \) for \( \mathbb{R} \) near the origin \( f_2 \) assumes the form

\[ f_2(x_1, x_2, \delta) = X_2^m + \tilde{a}_2(x_1, \delta) X_2^{m-2} + \cdots + \tilde{a}_{m-1}(x_1, \delta) X_2 + \tilde{a}_m(x_1, \delta), \]

where \( \tilde{a}_2, \ldots, \tilde{a}_m \) are smooth functions satisfying \( \tilde{a}_l(0, 0) = 0, \ l = 2, \ldots, m - 1 \).

Notice that the case \( m = 2 \) is special, since in this case

\[ f_2(x_1, x_2, \delta) = X_2^2 + \tilde{a}_2(x_1, \delta) \]

contains no linear term in \( X_2 \).

If \( m \geq 3 \), then by assumption (9.7), we have

\[ (9.10) \quad \frac{\partial \tilde{a}_{m-1}}{\partial x_1}(0, 0) \neq 0, \]

since

\[ 0 \neq \frac{\partial^2 f_2}{\partial x_1 \partial x_2}(0, 0, 0) = \frac{\partial \tilde{a}_{m-1}}{\partial x_1}(0, 0) \frac{\partial X_2}{\partial x_2}(0, 0, 0). \]

Consequently, any smooth function \( \varphi = \varphi(x_1, \delta) \) defined in a sufficiently small neighborhood of the origin can be written in the form

\[ \varphi(x_1, \delta) = \eta(x_1, \delta) \tilde{a}_{m-1}(x_1, \delta) + \tilde{\varphi}(\delta), \]

with smooth functions \( \eta = \eta(x_1, \delta) \) and \( \tilde{\varphi}(\delta) \). Applying this observation to the functions \( \tilde{a}_l \), we can write

\[ \tilde{a}_l(x_1, \delta) = \tilde{a}_{m-1}(x_1, \delta) b_l(x_1, \delta) + a_l(\delta), \quad l = 2, \ldots, m - 2, \]

with smooth functions \( b_l(x_1, \delta) \) and \( a_l(\delta) \), where \( a_l(0) = 0 \).
We can accordingly re-write the function \( F = f_1 + \sigma f_2 \) in the form \( F = \tilde{f}_1 + \sigma \tilde{f}_2 \), where

\[
\begin{align*}
\tilde{f}_1(x_1, \sigma, \delta) &= f_1(x_1, \delta) + \sigma \tilde{a}_m(x_1, \delta), \\
\tilde{f}_2(x_1, x_2, \sigma, \delta) &= X_2^m + \tilde{a}_{m-1}(x_1, \delta) X_2 \tilde{b}(x_1, X_2, \delta) \\
&\quad + a_2(\delta) X_2^{m-2} + \ldots + a_{m-2}(\delta) X_2^2,
\end{align*}
\] (9.11)

with

\[
\tilde{b}(x_1, X_2, \delta) := 1 + b_{m-2}(x_1, \delta) X_2 + \ldots + b_2(x_1, \delta) X_2^{m-2}.
\]

In particular, \( \tilde{b}(0, 0, 0) \neq 0 \).

In a second step, we apply Martinet’s theorem to the function \( \tilde{f}_1(x_1, \sigma, \delta) \). Since \( \partial_{x_1} f_1(0, 0, 0) = \partial_{x_2}^2 f_1(0, 0, 0) = 0 \) and \( \partial_{x_2}^{3} f_1(0, 0, 0) \neq 0 \), we then see that there exists a smooth function \( X_1 = X_1(x_1, \sigma, \delta) \) defined in a sufficiently small neighborhood of the origin with

\[
X_1(0, 0, 0) = 0, \quad \partial_1 X_1(0, 0, 0) \neq 0,
\]

so that in the new coordinate \( X_1 \) for \( \mathbb{R} \) near the origin \( \tilde{f}_1 \) assumes the form

\[
\tilde{f}_1(x_1, \sigma, \delta) = X_1^3 + a_1(\sigma, \delta) X_1 + a_m(\sigma, \delta),
\]

where \( a_1, a_m \) are smooth functions such that \( a_1(0, 0) = a_m(0, 0) = 0 \).

Let us write \( \tilde{a}_{m-1}(x_1, \delta) = \alpha(X_1(x_1, \sigma, \delta), \sigma, \delta) \), so that \( \alpha \) expresses \( \tilde{a}_{m-1} \) in the new coordinates \( X_1 \). By (9.10) and the chain rule, we have

\[
\alpha(0, 0, 0) = 0, \quad \text{and} \quad \frac{\partial \alpha}{\partial X_1}(0, 0, 0) \neq 0.
\]

This implies that there exists a unique, smooth function \( a_{m-1}(\sigma, \delta) \) with \( a_{m-1}(0, 0) = 0 \), such that \( \alpha(a_{m-1}(\sigma, \delta), \sigma, \delta) \equiv 0 \). Taylor’s formula then implies that \( \alpha(X_1, \sigma, \delta) \) can be written in the form

\[
\alpha(X_1, \sigma, \delta) = (X_1 - a_{m-1}(\sigma, \delta)) \tilde{g}(X_1, \sigma, \delta),
\]

where \( \tilde{g}(X_1, \sigma, \delta) \) is a smooth function with \( \tilde{g}(0, 0, 0) \neq 0 \). This shows that

\[
\tilde{a}_{m-1}(x_1, \delta) X_2 \tilde{b}(x_1, X_2, \delta) = (X_1 - a_{m-1}(\sigma, \delta)) X_2 \tilde{g}(X_1, \sigma, \delta) \tilde{b}(x_1, X_2, \delta).
\]

When expressed in the new variables \((X_1, X_2)\), we see that in combination with (9.11) we obtain the form of \( F \) as described in (ii).

Q.E.D.

After changing coordinates, the previous lemma allows to reduce Theorem 9.3 to the estimation of two-dimensional oscillatory integrals with phase functions of the form...
\( F(x_1, x_2, \delta, \delta) = f_1(x_1, \delta) + \sigma f_2(x_1, x_2, \delta), \) where
\[
f_1(x_1, \delta) = x_1^3 + \delta_1 x_1,
\]
(9.12)
\[
f_2(x_1, x_2, \delta) = x_2^m + \sum_{j=2}^{m-2} \delta_j x_2^{m-j} + (x_1 - \delta_{m-1}) x_2 b(x_1, x_2, \sigma, \delta),
\]
if \( m \geq 3, \) and \( f_2(x_1, x_2, \delta) = x_2^2, \) if \( m = 2. \) Here, \( \sigma \) and \( \delta_1, \ldots, \delta_\nu \) are small real parameters (where \( \nu \geq m - 1 \)), the latter forming the vector \( \delta := (\delta_1, \ldots, \delta_\nu) \in \mathbb{R}^\nu, \) and \( b = b(x_1, x_2, \sigma, \delta) \) is a smooth function defined on a neighborhood of the origin with \( b(0, 0, 0, 0) \neq 0. \)

**Lemma 9.6.** Assume that the phase function \( F \) is given by (9.12). Then there exists a neighborhood \( U \subset \mathbb{R}^2 \) of the origin and constants \( \varepsilon, \varepsilon' > 0 \) such that for any \( \psi \) which is compactly supported in \( U \) the following estimate
\[
|J(\lambda, \sigma, \delta)| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^3}}{\lambda^{3+\varepsilon}|\sigma|^{l_m + c_m \varepsilon}}
\]
holds true uniformly for \( |\sigma| + |\sigma| < \varepsilon', \) where \( l_m \) and \( c_m \) are defined as in Theorem 9.3.

**Proof.** We shall prove Lemma 9.6 and Theorem 9.3 at the same time by induction over \( m. \)

If \( m = 2 \) then the phase function (9.12) is reduced to the form
\[
F(x_1, x_2) = x_1^3 + \delta_1 x_1 + \sigma x_2^2,
\]
and by applying the method of stationary phase in \( x_2 \) and van der Corput’s lemma in \( x_2 \) we easily obtain estimate (9.13), with \( l_2 = 1/6. \) This proves also Theorem 9.3 for \( m = 2. \)

Assume that \( m \geq 3, \) and that the statement of Theorem 9.3 holds for every strictly smaller value of \( m. \) We shall apply again a Duistermaat type argument, in a similar way as in Section 3, in order to prove the statement of Lemma 9.6, hence also that of Theorem 9.3, for \( m. \) To this end, we introduce the mixed-homogeneous scalings \( \Delta_\rho(x_1, x_2) := (\rho^{1/2} x_1, \rho^{1/(m-1)} x_2), \) \( \rho > 0. \) Notice that these are such that the principal part of \( f_2 \) with respect to these dilations is given by \( x_2^m + x_1 x_2 b(0, 0, \sigma, \delta). \) Then
\[
F(\Delta_\rho(x), \sigma, \delta) = \rho^{3/2} F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta),
\]
where \( F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) = f_1(x_1, \tilde{\delta}) + \tilde{\sigma} f_2(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) \) is given by
\[
f_1(x_1, \tilde{\delta}) := x_1^3 + \tilde{\delta}_1 x_1
\]
(9.14)
\[
f_2(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) := x_2^m + \sum_{j=2}^{m-2} \tilde{\delta}_j x_2^{m-j} + (x_1 - \tilde{\delta}_{m-1}) x_2 b(\Delta_\rho(x), \sigma, \delta),
\]
with $\tilde{\sigma}, \tilde{\delta}$ defined by

$$
\tilde{\sigma} := \frac{\sigma}{\rho^{2/(m-1)}}, \quad \tilde{\delta}_1 := \frac{\delta_1}{\rho}, \quad \tilde{\delta}_j := \frac{\delta_j}{\rho^{2/(m-1)}} \quad (j = 2, \ldots, m - 1),
$$

so that in particular $\tilde{\delta}_{m-1} = \frac{\delta_{m-1}}{\rho^{2}}$. Thus, if we define "dual scalings" by

$$
\Delta^*_\rho(\sigma, \delta) := (\tilde{\sigma}, \tilde{\delta}),
$$

we see that if $b$ is constant, then $F(\Delta^*_\rho(x), \sigma, \delta) = \rho^{\frac{3}{2}} F(x, \Delta^*_\rho(\sigma, \delta))$. It is then natural to introduce the quasi-norm

$$
N(\sigma, \delta) := \left| \sigma^{\frac{2(m-1)}{m}} \right| + |\delta_1| + \left| \delta_2 \right|^{m-1} + \cdots + \left| \delta_{m-2} \right|^{\frac{2(m-1)}{m}} + |\delta_{m-1}|^2,
$$

which is $\Delta^*_\rho$-homogeneous of degree $-1$, i.e., $N(\Delta^*_\rho(\sigma, \delta)) = \rho^{-1} N(\sigma, \delta)$. Given $\sigma, \delta$, we now choose $\rho$ so that $N(\tilde{\sigma}, \tilde{\delta}) = 1$, i.e.,

$$
\rho := N(\sigma, \delta).
$$

Notice that $\rho \ll 1$, and that $(\tilde{\sigma}, \tilde{\delta})$ lies in the "unit sphere"

$$
\Sigma := \{(\sigma', \delta') \in \mathbb{R}^m : N(\sigma', \delta') = 1\}.
$$

Then, after scaling, we may re-write

$$
J(\lambda, \sigma, \delta) = J(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) := \rho^{\frac{m}{2}} \int_{\mathbb{R}^2} e^{i \lambda \rho^{\frac{3}{2}} F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta)} \psi(\Delta^*_\rho(x), \delta) \chi_k(x) \, dx,
$$

where here $\rho, \sigma$ and the $\delta_j$ are small parameters. For a while, it will be convenient to consider $\tilde{\sigma}$ and the $\tilde{\delta}_j$ as additional, independent real parameters, which may not be small, but bounded.

We shall apply a dyadic decomposition to this integral. To this end, we choose $\chi_0, \chi \in C_0^\infty(\mathbb{R}^2)$ with $\text{supp}\, \chi \subset \{ \frac{B}{2} < |x| < 2B \}$ (where $B$ is a sufficiently large positive number to be fixed later) such that

$$
\chi_0(x) + \sum_{k=1}^{\infty} \chi(\Delta_{2^{-k}}(x)) = 1, \quad \text{for every } x \in \mathbb{R}^2.
$$

Accordingly we decompose the oscillatory integral

$$
J(\lambda, \sigma, \delta) = \sum_{k=0}^{\infty} J_k(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta),
$$

where

$$
J_k(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) := \rho^{\frac{m}{2}} \int_{\mathbb{R}^2} e^{i \lambda \rho^{\frac{3}{2}} F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta)} \psi(\Delta_{2^{-k}}(x), \delta) \chi_k(x) \, dx,
$$

and $\chi_k(x) := \chi(\Delta_{2^{-k}}(x))$ for $k \geq 1$. 

Assume first that $k \geq 1$. Then, by using the scaling $\Delta_{2^m}$, we get
\[ J_k(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) = (2^k \rho)^{\frac{2(m-1)}{2m}} \int_{\mathbb{R}^2} e^{i\lambda(2^k \rho) \frac{3}{2} F_k(x)} \psi(\Delta_{2^k \rho}(x), \sigma) \chi(x) \, dx, \]
where $F_k(x) := g_1(x, \tilde{\sigma}_k) + \tilde{\sigma}_k g_2(x, \tilde{\sigma}_k, \tilde{\delta}_k, 2^k \rho, \sigma, \delta)$ is given by
\[ g_1(x_1, \tilde{\delta}_k) := x_1^3 + \tilde{\delta}_{1,k} x_1, \]
\[ g_2(x, \tilde{\sigma}_k, \tilde{\delta}_k, 2^k \rho, \sigma, \delta) := x_2^m + \sum_{j=2}^{m-2} \tilde{\delta}_{j,k} x_2^{m-j} + (x_1 - \tilde{\delta}_{m-1,k}) x_2 b(\Delta_{2^k \rho}(x), \sigma, \delta), \]
with
\[ (\tilde{\sigma}_k, \tilde{\delta}_k) := (\tilde{\sigma}_k, \tilde{\delta}_{1,k}, \ldots, \tilde{\delta}_{m-1,k}) := \Delta_{2^k}^*(\tilde{\sigma}, \tilde{\delta}) = \Delta_{2^k \rho}^*(\sigma, \delta). \]

Observe that we may restrict ourselves to those $k$ for which $2^k \rho \lesssim 1/B$, since otherwise $J_k \equiv 0$. Consequently, if we choose $B$ in the definition of $\chi$ sufficiently large, then $2^k \rho << 1$, and also $|\tilde{\sigma}_k| + |\tilde{\delta}_k| << 1$. We thus see that there is some positive constant $c > 0$ such that if $x \in \text{supp} \chi$, then either $|\partial_1 g_1(x, \tilde{\delta}_k)| \geq c$, or $|\partial_2 g_2(x_1, x_2, \tilde{\sigma}_k, \tilde{\delta}_k, 2^k \rho, \sigma, \delta)| \geq c$.

Fix a point $x^0 = (x^0_1, x^0_2) \in \text{supp} \chi$, let $\eta$ be a smooth cut-off function supported in a sufficiently small neighborhood of $x^0$, and consider the oscillatory integral $J^\eta_k$ defined by
\[ J^\eta_k(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) = (2^k \rho)^{\frac{2(m-1)}{2m}} \int_{\mathbb{R}^2} e^{i\lambda(2^k \rho) \frac{3}{2} F_k(x)} \psi(\Delta_{2^k \rho}(x), \sigma) \chi(x) \eta(x) \, dx. \]

By using an integration by parts in $x_1$ in case that $|\partial_1 g_1(x^0_1, \tilde{\delta}_k)| \geq c$, respectively in $x_2$ if $|\partial_2 g_2(x^0_1, x^0_2, \tilde{\sigma}_k, \tilde{\delta}_k, 2^k \rho, \sigma, \delta)| \geq c$, and subsequently applying van der Corput’s lemma to the $x_1$-integration in the latter case, we then obtain
\[ |J^\eta_k| \leq \frac{C(2^k \rho)^{\frac{2(m-1)}{2m}} \|\psi(\cdot, \delta)\|_{C^3}}{(1 + \lambda(2^k \rho) \frac{3}{2})^{\frac{1}{2}} (1 + \lambda(2^k \rho) \frac{3}{2} |\tilde{\sigma}_k|) \frac{3}{2}} \leq \frac{C(2^k \rho)^{\frac{2(m-1)}{2m}} \|\psi(\cdot, \delta)\|_{C^3}}{\lambda(2^k \rho) \frac{3}{2}^{\frac{1}{2} + \varepsilon} |\tilde{\sigma}_k| \frac{3}{2}^{\frac{1}{2} + \varepsilon}}. \]

By means of a partition of unity argument this implies the same type of estimate
\[ (9.15) \quad |J_k| \leq \frac{C(2^k \rho)^{\frac{2(m-1)}{2m}} \|\psi(\cdot, \delta)\|_{C^3}}{|\lambda(2^k \rho) |^{\frac{3}{2} + \varepsilon} |\tilde{\sigma}_k|^{\frac{3}{2} + \varepsilon}} = C(2^k \rho)^{\frac{6-m}{2(m-1)}} - \epsilon \frac{\|\psi(\cdot, \delta)\|_{C^3}}{\lambda^{\frac{3}{2} + \varepsilon} |\sigma|^{\frac{3}{2} + \varepsilon}} \]
for $J_k$.

Consider first the case where $m < 6$. Then clearly
\[ \sum_{k \geq 1} |J_k| = \sum_{2^k \rho \leq 1} |J_k(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta)| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^3}}{\lambda^{\frac{3}{2} + \varepsilon} |\sigma|^{\frac{3}{2} + \varepsilon}}. \]

Assume next that $m \geq 6$. Then the infinite series $\sum_{k=1}^{\infty} (2^k)^{\frac{6-m}{2(m-1)} - \epsilon}$ converges. Note also that $\rho \geq |\sigma|^{\frac{2(m-1)}{2m}}$. Summing therefore over all $k \geq 1$, we obtain from (9.15)
that
\[ \sum_{k \geq 1} |J_k| \leq \frac{c \|\psi(\cdot, \delta)\|_{C^3}}{|\lambda|^{\frac{1}{2} + \varepsilon} |\sigma|^{l_m + c_m \varepsilon}}. \]

We are thus left with the integral
\[ J_0(\lambda, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) := \rho^{\frac{m}{2(m-1)}} \int_{\mathbb{R}^2} e^{i\lambda x_0 \Delta} F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) \psi(\Delta \rho(x), \delta) \chi_0(x) \, dx, \]
where \( F(x, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) \) is given by (9.14).

Let us fix a point \((\tilde{\sigma}^0, \tilde{\delta}^0) \in \Sigma\), and a point \(x^0 = (x_0^0, x_2^0) \in \text{supp} \chi_0\), and let again \(\eta\) be a smooth cut-off function supported near \(x^0\). \(J_0^0\) will be defined by introducing \(\eta\) into the amplitude of \(J_0\) in the same way as before. We shall prove that the oscillatory integral \(J_0^0\) satisfies the estimate
\[ |J_0^0| \leq \frac{C \|\psi(\cdot, \delta)\|_{C^3}}{|\lambda|^{\frac{1}{2} + \varepsilon} |\sigma|^{l_m + c_m \varepsilon}}, \]
provided \(\eta\) is supported in a sufficiently small neighborhood \(U\) of \(x^0\) and \((\tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) \in V\), where \(V\) is a sufficiently small neighborhood of the point \((\tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0)\). By means of a partion of unity argument this will then imply the same type of estimate for \(J_0\), hence for \(J\), which will conclude the proof of Lemma 9.6, hence also of Theorem 9.3.

Now, if either \(\partial_1 F(x_0^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0\) or \(\partial_2 f_2(x_0^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0\), then we can estimate \(J_0^0\) exactly like the \(J_k^0\) and get the required estimate (9.16) for \(J_0^0\).

Assume therefore next that
\[ \partial_1 F(x_0^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) = 0 \quad \text{and also} \quad \partial_2 f_2(x_0^0, x_2^0, \sigma^0, \tilde{\delta}^0, 0, 0, 0) = 0. \]

We then distinguish the following four cases:

**Case 1.** \(\tilde{\sigma}^0 \neq 0\) and \(x_0^0 \neq 0\).

Then, since \(x_0^0 \neq 0\), it is easy to see from (9.14) that \(\partial_1^2 F(x_0^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0\) as well. Note here that if we write \(b(x, \rho, \sigma, \delta) := b(\Delta \rho(x), \sigma, \delta)\), then
\[ b(x, 0, 0, 0) \equiv b(0, 0, 0, 0) \neq 0. \]

We can then argue here in a similar way as in the proof of Proposition 9.1, so let us only briefly sketch the argument. Suppose that \(x_1^c(x_2, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta)\) is a critical point of \(F\) with respect to \(x_1\). Then it is a smooth function of its variables, and if \(\rho = \sigma = \delta = 0\), then by (9.18)
\[ x_1^c = x_1^c(x_2, \tilde{\sigma}, \tilde{\delta}, 0, 0, 0) = \left( - (\tilde{\sigma} + \tilde{\delta} x_2 b(0, 0, 0, 0)) \right)^{1/2} \]

...
and
\[ F(x_1^0(x_2, \tilde{\sigma}, \tilde{\delta}, 0, 0, 0), x_2, \tilde{\sigma}, \tilde{\delta}, 0, 0, 0) \]
\[ = (x_1^0)^3 + \tilde{\delta}_1 x_1^0 + \tilde{\sigma} \left( x_2^m + \sum_{j=2}^{m-2} \tilde{\delta}_j x_2^{m-j} + (x_1^0 - \tilde{\delta}_{m-1})x_2 b(0, 0, 0, 0) \right). \]

If \( \phi \) denotes the phase function
\[ \phi(x_2, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta) := F(x_1^0(x_2, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta), x_2, \tilde{\sigma}, \tilde{\delta}, \rho, \sigma, \delta), \]
which arises after applying the method of stationary phase to the \( x_1 \)-integration, then since \( \tilde{\delta}^0 \neq 0 \), this easily shows that there exists a natural number \( N \) such that
\[ \partial^N_2 \phi(x_2, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0. \]

Consequently, we can in a second step apply van der Corput’s lemma to the \( x_2 \)-integration and obtain the estimate
\[ (9.19) \quad |J_0^0| \leq \frac{C \rho^{\frac{m-1}{2}} \| \psi(\cdot, \delta) \|_{C^3}}{|\lambda \rho^{\frac{m}{2}+\varepsilon} |\bar{\delta}|^{\frac{m}{2}+\varepsilon}}, \]
which implies (9.16) as before (just put \( k = 0 \) in our previous argument).

**Case 2.** \( \tilde{\delta}^0 \neq 0 \) and \( x_1^0 = 0 \).

Then, by (9.18), we have \( \partial^2_1 F(x_1^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) = 0 \) as well. But, again by (9.18), we also have \( \partial_1 \partial_2 F(x_1^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0 \), so that \( F \) has a non-degenerate critical point at \( x^0 \) as a function of two variables. If the neighborhoods \( U \) and \( V \) are chosen sufficiently small, we can therefore apply the stationary phase method in two variables, which leads to an even stronger estimate than the estimate (9.19), since here \( |\tilde{\sigma}| \sim 1 \).

**Case 3.** \( \tilde{\delta}^0 = 0 \) and \( \tilde{\sigma}^0 \neq 0 \).

In this case we have \( x_1^0 \neq 0 \), because of (9.17), and thus \( \partial^2_1 F(x_1^0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0 \). Moreover, in this situation we consider \( \tilde{\sigma} \) such that \( |\tilde{\sigma}| << 1 \). Since we can regard \( \tilde{\sigma} - \tilde{\sigma}^0, \tilde{\delta} - \tilde{\delta}^0 \) as small perturbation parameters if the neighborhoods \( U \) and \( V \) are chosen sufficiently small, we can therefore apply Proposition 9.1, with \( \sigma \) in this proposition replaced by \( \tilde{\sigma} \), and obtain (9.19).

**Case 4.** \( \tilde{\delta}^0 = 0 \) and \( \tilde{\sigma}^0 = 0 \).

Then, by (9.17), \( x_1^0 = 0 \) as well. In this case we make use of our induction hypothesis. Indeed, let us consider the function
\[ x_2 \mapsto f_2(0, x_2, \sigma^0, \tilde{\delta}^0, 0, 0, 0) = x_2^m + \sum_{j=2}^{m-2} \tilde{\delta}_j x_2^{m-j} - \tilde{\delta}_{m-1} x_2 b(0, 0, 0, 0). \]

Now \( x_2 = x_2^0 \) is a critical point, say of multiplicity \( \mu - 1 \), of this function, i.e., \( \partial^l_2 f_2(0, x_2^0, \sigma^0, \tilde{\delta}^0, 0, 0, 0) = 0 \) for \( l = 1, \ldots, \mu - 1 \) and \( \partial^\mu_2 f_2(0, x_2^0, \sigma^0, \tilde{\delta}^0, 0, 0, 0) \neq 0 \).
Then $\mu < m$, because at least one of the coefficients $\tilde{\delta}_j$, $j = 2, \ldots, m - 1$, does not vanish and $b(0, 0, 0, 0) \neq 0$. Moreover, at this critical point also the condition $\partial_1 \partial_2 f_2 (0, x_2^0, \tilde{\sigma}^0, \tilde{\delta}^0, 0, 0, 0) \neq 0$ is satisfied. Therefore, after translating coordinates $x_2$ by $x_2^0$, by our hypothesis we may apply the conclusion of Theorem 9.3 for $\mu$ in place of $m$ and obtain the estimate

$$|J^0_m| \leq \frac{C \rho^{2(m-1)}}{|\lambda \rho^{\frac{3}{2}}|^{\frac{2}{3}}} |\tilde{\sigma}|^{l_{\mu} + c_\mu \varepsilon},$$

provided again that $U$ and $V$ are small enough. Now, if $\mu < 6$, then this estimate agrees with (9.19), and we are done.

So, assume finally that $\mu \geq 6$. Since $l_m$ is increasing in $m$, we may replace $l_\mu$ by $l_{m-1}$ in this estimate, and clearly we have $c_\mu = c_m = 2$. Recall also that here $\tilde{\sigma} = \sigma \rho^{2(m-1)}$ and $\rho \geq |\sigma|^{\frac{2(m-1)}{2m-3}}$. Then the total exponent of $\rho$ in this estimate, except for the terms containing $\varepsilon$, is $\frac{-3}{4(m-1)(2m-5)}$, and $\rho^{\frac{3}{2}m-3} \leq |\sigma|^{\frac{2m-5}{2m-3}}$. Moreover, one computes that $|\sigma|^{\frac{3}{2}m-3} \leq |\sigma|^{-l_m-1} = |\sigma|^{-l_m}$. In a similar way, if we replace $\rho$ by $|\sigma|^{\frac{3}{2}m-3}$ in the term $|\rho^{\frac{3}{2}}\varepsilon| |\tilde{\sigma}|^{-c_{m-1}\varepsilon}$, we obtain the additional factor $|\sigma|^{-\frac{3(m-1)}{2m-3} \varepsilon} \leq |\sigma|^{-2\varepsilon}$ in the estimate for $J^0_m$. In combination, we obtain again the estimate (9.16).

This concludes the proof of the lemma as well as of Theorem 9.3.

Q.E.D.

10. Uniform estimates for oscillatory integrals with finite type phase functions of two variables

In this section we shall provide a proof of Theorem 1.10. We shall closely follow the proof of Theorem 1.2, which did already provide uniform estimates for the Fourier transforms of surface carried measures $\rho d\sigma(\xi)$ for the contribution by the region near the principal root jet. Notice that the assumption $\rho \geq 0$ that we had made for the estimation of the maximal operator $\mathcal{M}$ had only been introduced for convenience and was not needed for the estimations of oscillatory integrals. Without further mentioning, we shall use the same notation as in the various parts of the proof of Theorem 1.2.

We may assume that $S$ is the graph $S = \{ (x_1, x_2, \phi(x_1, x_2)) : (x_1, x_2) \in \Omega \}$ of a smooth real valued function of finite type $\phi \in C^\infty (\Omega)$ defined on an open neighborhood $\Omega$ of the origin in $\mathbb{R}^2$ and satisfying

$$\phi(0, 0) = 0, \nabla \phi(0, 0) = 0,$$

where $x^0 = (0, 0)$. We then have to prove

**Theorem 10.1.** There exists a neighborhood $\Omega \subset \mathbb{R}^2$ of the origin such that for every $\eta \in C^\infty_0 (\Omega)$ the following estimate holds true for every $\xi \in \mathbb{R}^3$:

$$\left| \int_{\mathbb{R}^2} e^{i \xi \phi(x_1, x_2) + \xi_1 x_1 + \xi_2 x_2} \eta(x_1, x_2) \, dx \right| \leq C \| \eta \|_{C^3(\mathbb{R}^2)} \log (2 + |\xi|)(1 + |\xi|)^{-1/h(\phi)}.$$
By decomposing $\mathbb{R}^2$ into its four quadrants, we may reduce ourselves to the estimation of oscillatory integrals of the form

$$J(\xi) := \int_{(\mathbb{R}_+)^2} e^{i(\xi_3\phi(x_1,x_2)+\xi_1x_1+\xi_2x_2)} \eta(x_1,x_2) \, dx.$$ 

Notice also that we may assume in the sequel that

$$|\xi_1| + |\xi_2| \leq \delta |\xi_3|,$$

where $0 < \delta << 1$ is a sufficiently small constant, since for $|\xi_1| + |\xi_2| > \delta |\xi_3|$ the estimate (10.1) follows by an integration by parts, if $\Omega$ is chosen small enough. Of course, we may in addition always assume that $|\xi| \geq 2$.

If $\chi$ is any integrable function defined on $\Omega$, we shall put

$$J^\chi(\xi) := \int_{(\mathbb{R}_+)^2} e^{i(\xi_3\phi(x_1,x_2)+\xi_1x_1+\xi_2x_2)} \eta(x_1,x_2) \chi(x) \, dx.$$ 

The case $h(\phi) < 2$ is contained in [8] (here, estimate (10.1) holds true even without the logarithmic term $\log(2 + |\xi|)$), so let us assume from now on that

$$h(\phi) \geq 2.$$ 

The following van der Corput type lemma, due to J. E. Björk (see [7]) and also G. I. Arhipov [1], will be useful.

**Lemma 10.2.** Assume that $f$ is a smooth real valued function defined on an interval $I \subset \mathbb{R}$ which is of polynomial type $m \geq 2$ ($m \in \mathbb{N}$), i.e., there are positive constants $c_1, c_2 > 0$ such that

$$c_1 \leq \sum_{j=2}^{m} |f^{(j)}(s)| \leq c_2 \quad \text{for every } s \in I.$$

Then for $\lambda \in \mathbb{R}$,

$$\left| \int_I e^{i\lambda f(s)} g(s) \, ds \right| \leq C ||g||c_{2(I)} (1 + |\lambda|)^{-1/m},$$

where the constant $C$ depends only on the constants $c_1$ and $c_2$.

Following Section 5 we shall begin with the easiest case where the coordinates $x$ are adapted to $\phi$. In analogy with the proof of Proposition 5.1 we then decompose $J(\xi) = \sum_{k=k_0}^{\infty} J_k(\xi)$, where

$$J_k(\xi) := \int_{(\mathbb{R}_+)^2} e^{i(\xi_3\phi(x)+\xi_1x_1+\xi_2x_2)} \eta(x) \chi_k(x) \, dx$$

$$= 2^{-k|\xi|} \int_{(\mathbb{R}_+)^2} e^{i \left( -k\xi_3\phi(x)+2^{-k\xi_1x_1+2^{-k\xi_2x_2}} \right)} \eta(\delta_{2^{-k}}(x)) \chi(x) \, dx,$$
and where \( \chi \) is supported in an annulus \( D \). Moreover, according to the proof of Corollary 5.2, we can choose the weight \( \kappa \) such that \( 0 < \kappa_1 \leq \kappa_2 < 1 \) and

\[
\frac{1}{|\kappa|} = d_h(\phi_n) \leq h(\phi_n) = h(\phi).
\]

Then, as in the proof of Proposition 5.1, given any point \( x^0 \in D \), we can find a unit vector \( e \in \mathbb{R}^2 \) and some \( m \in \mathbb{N} \) with \( 2 \leq m \leq h(\phi_n) = h(\phi) \) such that \( \partial_{e^m} \phi_n(x^0) \neq 0 \). For \( k \geq k_0 \) sufficiently large we can thus apply Lemma 10.2 to the \( x_2 \)-integration in \( J_k(\xi) \) near the point \( x^0 \). By means of a partition of unity argument, we then get

\[
|J_k(\xi)| \leq C \|\eta\|_{C^3(\mathbb{R}^2)} 2^{-k|\kappa|} (1 + 2^{-k} |\xi_3|)^{-1/m}
\]

\[
\leq C \|\eta\|_{C^3(\mathbb{R}^2)} 2^{-\frac{k}{\kappa_1\sigma}} (1 + 2^{-k} |\xi|)^{-1/h(\phi)}.
\]

The estimate (10.1) then follows by summation in \( k \).

Assume next that the coordinates \( x \) are not adapted to \( \phi \). In a first step, we then decompose \( J(\xi) = J^{1-\rho_1}(\xi) + J^{\rho_1}(\xi) \), where \( \rho_1 \) is the cut-off function introduced in Subsection 6.1 which localizes to a narrow \( \kappa \)-homogeneous neighborhood

\[
|x_2 - b_1 x_1^{m_1}| \leq \varepsilon_1 x_1^{m_1}
\]

of the curve \( x_2 = b_1 x_1^{m_1} \).

The oscillatory integral \( J^{1-\rho_1}(\xi) \) can be estimated in a similar way as in the case of adapted coordinates by means of Lemma 10.2 (compare also the proof of Lemma 6.2), so that there remains \( J^{\rho_1}(\xi) \) to be considered. To this end, we decompose the domain above as in Subsection 6.3 into the domains \( D_l \), which become \( \kappa_l \)-homogeneous in the coordinates \( y \) defined by (6.3), and the transition domains \( E_l \). Accordingly, we decompose

\[
J^{\rho_1}(\xi) = \sum_{l=l_0}^\lambda J^{\rho_1}(\xi) + \sum_{l=l_0}^{\lambda-1} J^{\rho_1}(\xi),
\]

where \( \rho_l \) and \( \tau_l \) are the cut-off functions defined in that subsection.

**Estimation of** \( J^{\rho_1}(\xi) \). In analogy with the proof of Lemma 6.4, after applying the change of coordinates (6.3) and performing a dyadic decomposition as before, only with the weight \( \kappa \) replaced by the weight \( \kappa_l \), we find that \( J^{\rho_1}(\xi) = \sum_{k=k_0}^\infty J_k(\xi) \), where

\[
J_k(\xi) = \int_{(\mathbb{R}^2)^2} e^{i \left( 2^{-k} \xi_3 \tilde{\phi}(y) + 2^{-k} \xi_1 \xi_1 y_1 + 2^{-k} \xi_2 y_2 + 2^{-k} \xi_3 y_3 \right)} \tilde{\rho}_1(y) \tilde{\eta} \tilde{\delta}_{2^{-k}}(y) \tilde{\chi}(y) dy,
\]

with \( \psi_k(y) \) etc. defined as in Subsection 6.4. In view of (6.16) for the case \( l = \lambda \) and (6.18) for the case \( l \leq \lambda - 1 \) we can then again estimate \( J_k(\xi) \) by means of Lemma 10.2 applied to the \( y_2 \)-integration and obtain that

\[
|J_k(\xi)| \leq C \|\eta\|_{C^3(\mathbb{R}^2)} 2^{-k|\kappa|} (1 + 2^{-k} |\xi_3|)^{-1/d_h(\tilde{\phi})}
\]

\[
\leq C \|\eta\|_{C^3(\mathbb{R}^2)} 2^{-\frac{k}{\kappa_1\sigma}} (1 + 2^{-k} |\xi|)^{-1/h(\phi)}.
\]
since $d_h(\tilde{\phi}_l) \leq h(\phi)$, except for the case where $l = \lambda$ and where $\tilde{\phi}_l = \tilde{\phi}_p$ is one of the exceptional polynomials $P$ given by (4.4), with $\lambda_1 + \lambda_2 > 0$. However, the contribution of the "exceptional domain" (6.17) to $J(\xi)$ can be estimated by Proposition 7.5 (compare the corresponding discussion in Subsection 7.2). Since $h(\phi) \geq 2$, the estimate (7.10) in that proposition is stronger than the desired estimate (10.1).

By summing over all $k$, we see that $J^n(\xi)$ satisfies estimate (10.1).

**Estimation of** $J^n(\xi)$. Following Subsection 6.5, we decompose

$$J^n_t(\xi) = \sum_{j,k} J_{j,k}(\xi),$$

where summation takes place over all pairs $j, k$ satisfying (6.19), i.e.,

$$a_t j + M \leq k \leq a_{t+1} j - M,$$

with $J_{j,k}(\xi)$ given by

$$J_{j,k}(\xi) := \int_{\mathbb{R}^2} e^{\left(\xi_1 \tilde{\phi}(y) + \xi_2 y_1 + \xi_3 y_2 + \xi_4 \psi(y_1)\right)} \tilde{\tau}_1(y) \tilde{\eta}(y) \chi_{j,k}(y) \, dy$$

$$= 2^{-j-k} \int_{\mathbb{R}^2} e^{\left(\xi_1 \tilde{\phi}(y) + 2^{-j} \xi_2 y_1 + 2^{-k} \xi_3 y_2 + \xi_4 \psi(2^{-j} y_1)\right)} \tilde{\tau}_{j,k}(y) \tilde{\eta}_{j,k}(y) \chi(y_1) \chi(y_2) \, dy.$$

Here, we have kept the notations from Subsection 6.5. Assume first that $\phi$ is analytic. Then, by (6.21),

$$\tilde{\phi}_{j,k}(y) = 2^{-(A_t j + B_t k)} \left( c_t y_1^{A_t} y_2^{B_t} + O(2^{-CM}) \right)$$

for some constant $C > 0$, where $A_t$ and $B_t$ are given by (6.12) and $M$ can still be chosen as large as we wish, and where

$$\tilde{\phi}_2^2(y_1^{A_t} y_2^{B_t}) \sim 1.$$

We can thus again apply Lemma 10.2, with $m = 2$, to the $y_2$-integration in $J_{j,k}(\xi)$ and obtain

$$|J_{j,k}(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-j-k} \left( 1 + 2^{-(A_t j + B_t k)} |\xi_1| \right)^{-1/2}$$

$$\sim C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-j-k} \left( 1 + 2^{-(A_t j + B_t k)} |\xi_1| \right)^{-1/2}.$$
where $I_0$ and $I_\infty$ denote the index sets

$$I_0 := \{(j, k) \in \mathbb{N}^2 : A_l j + B_l k \leq \log |\xi| \text{ and } a_l j \leq k \leq a_{l+1} j\}$$

and

$$I_\infty := \{(j, k) \in \mathbb{N}^2 : A_l j + B_l k > \log |\xi|\}.$$

These estimates can easily be summed in $j$ and $k$ by means of the following auxiliary result.

**Lemma 10.3.** Let $0 < a_1 < a_2$ and $b_1, b_2 \geq 0$ with $b_1 + b_2 > 0$ be given. For $\gamma > 0$, consider the triangle $A_\gamma := \{(t_1, t_2) \in (\mathbb{R}_+)^2 : a_1 t_1 \leq t_2 \leq a_2 t_1 \text{ and } b_1 t_1 + b_2 t_2 \geq \gamma\}$, and denote by $(0, 0)$ and $\gamma X_1$ and $\gamma X_2$, with

$$X_1 := \frac{1}{b_1 + a_1 b_2}(1, a_1) \text{ and } X_2 := \frac{1}{b_1 + a_2 b_2}(1, a_2),$$

the three vertices of $A_\gamma$. Assume that $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$ is such that

$$\mu \cdot X_1 < \mu \cdot X_2. \quad (10.4)$$

(a) If $\mu \cdot X_2 > 0$, then

$$\int_{A_\gamma} e^{\mu \cdot t} dt \leq C e^{\gamma \mu \cdot X_2}. \quad (a)$$

(b) If $\mu \cdot X_2 = 0$, then

$$\int_{A_\gamma} e^{\mu \cdot t} dt \leq C \gamma. \quad (b)$$

(c) If $\mu \cdot X_2 < 0$, then

$$\int_{A_\gamma} e^{\mu \cdot t} dt \leq C, \quad (c)$$

where the constant $C$ in these estimates depends only on the $a_j, b_j$ and $\mu$.

Similarly, if we put $B_\gamma := \{(t_1, t_2) \in (\mathbb{R}_+)^2 : a_1 t_1 \leq t_2 \leq a_2 t_1 \text{ and } b_1 t_1 + b_2 t_2 \geq \gamma\}$, then the following holds true:

(d) If $\mu \cdot X_2 < 0$, then

$$\int_{B_\gamma} e^{\mu \cdot t} dt \leq C e^{\gamma \mu \cdot X_1}. \quad (d)$$

**Proof.** Let us change to the coordinates $(x_1, x_2)$ given by

$$(t_1, t_2) = (x_1 + x_2, a_1 x_1 + a_2 x_2).$$

In these coordinates, $A_\gamma$ and $X_1, X_2$ correspond to

$$\tilde{A}_\gamma := \{(x_1, x_2) \in (\mathbb{R}_+)^2 : (b_1 + a_1 b_2)x_1 + (b_1 + a_2 b_2)x_2 \leq \gamma\}$$

and

$$\tilde{X}_1 := \left(\frac{1}{b_1 + a_1 b_2}, 0\right), \quad \tilde{X}_2 := \left(0, \frac{1}{b_1 + a_2 b_2}\right),$$

respectively.
respectively. Moreover, \( \mu \cdot t = \bar{\mu} \cdot x \), where \( \bar{\mu} \cdot \bar{X}_1 < \bar{\mu} \cdot \bar{X}_2 \), i.e.,

\[
(10.5) \quad \frac{\bar{\mu}_1}{b_1 + a_1 b_2} < \frac{\bar{\mu}_2}{b_1 + a_2 b_2}.
\]

Now, in case (a) we have \( \bar{\mu}_2 > 0 \), so that because of (10.5)

\[
\int_{A_\gamma} e^{\mu - d} dt = C \int_{A_\gamma} e^{\bar{\mu} x} dx = C \frac{1}{\bar{\mu}_2} \int_0^1 e^{\bar{\mu}_2 (\bar{X}_1 + a_1 b_2)} e^{\left( \bar{\mu}_1 - \frac{\bar{\mu}_2 (b_1 + a_1 b_2)}{b_1 + a_2 b_2} \right) x_1} dx_1
\]

\[
\leq C \frac{1}{\bar{\mu}_2} e^{\bar{\mu}_2 (a_1 b_1)} = C e^{\gamma \bar{\mu}_2} \bar{X}_2,
\]

where \( C \) depends only on \( a_1 \) and \( a_2 \). In case (b), we have \( \bar{\mu}_2 = 0 \) and \( \bar{\mu}_1 < 0 \), so that a similar estimation as before leads to

\[
\int_{A_\gamma} e^{\mu - d} \leq C \frac{\gamma \bar{\mu}_2}{b_1 + a_2 b_2},
\]

and the case (c) is obvious, since here \( \bar{\mu}_1, \bar{\mu}_2 < 0 \).

The estimate in (d) is obtained in an analogous way as the one in (a).

Q.E.D.

To estimate \( J^\gamma_0 (\xi) \), we put \( \mu := (\frac{d}{2} - 1, \frac{B}{2} - 1) \) and \( a_1 := a_t, a_2 := a_{t+1}, b_1 := A_t, b_t := B_t, \gamma := \log |\xi| \) in Lemma 6.3. Then

\[
X_1 = \frac{1}{A_t + a_t B_t} (1, a_t), \quad X_2 = \frac{1}{A_t + a_{t+1} B_t} (1, a_{t+1}),
\]

and (compare also the discussion in Subsection 6.2)

\[
\mu \cdot X_1 = \frac{1}{2} - \frac{1 + a_t}{A_t + a_t B_t} = \frac{1}{2} - \frac{1}{d_h (\bar{\phi}_{n_t})}, \quad \mu \cdot X_2 = \frac{1}{2} - \frac{1 + a_{t+1}}{A_t + a_{t+1} B_t} = \frac{1}{2} - \frac{1}{d_h (\bar{\phi}_{n_{t+1}})}.
\]

Since \( d_h (\bar{\phi}_{n_t}) < d_h (\bar{\phi}_{n_{t+1}}) \), we see that condition (10.4) is satisfied. Comparing the sum in \( J^\gamma_0 (\xi) \) with a corresponding integral and applying Lemma 6.3 we thus find that

\[
|J^\gamma_0 (\xi)| \leq C \log |\xi| \|\xi\|^{-1/2} \leq C \log (2 + |\xi|) \|\xi\|^{-1/h(\phi)},
\]

if \( \mu \cdot X_2 \leq 0 \), and

\[
|J^\gamma_0 (\xi)| \leq C |\xi|^{-1/2} \exp \left( \log |\xi| (1/2 - 1/d_h (\bar{\phi}_{n_{t+1}})) \right) \leq C \|\xi\|^{-1/d_h (\bar{\phi}_{n_{t+1}})},
\]

if \( \mu \cdot X_2 > 0 \). Since \( d_h (\bar{\phi}_{n_{t+1}}) \leq h(\phi) \), this shows that \( J^\gamma_0 (\xi) \) satisfies the estimate (10.1).

Similarly, in order to estimate \( J^\gamma_\infty (\xi) \), we put \( \mu := (-1, -1) \) in Lemma 6.3 (d). Then \( \mu \cdot X_2 = -1/d_h (\bar{\phi}_{n_{t+1}}) < 0 \), \( \mu \cdot X_1 = -1/d_h (\bar{\phi}_{n_t}) < \mu \cdot X_2 \), so that we obtain

\[
|J^\gamma_\infty (\xi)| \leq C \exp \left( \log |\xi| (-1/d_h (\bar{\phi}_{n_t})) \right) \leq C \|\xi\|^{-1/h(\phi)}.
\]

In combination, we have seen that all \( J^\gamma (\xi) \) satisfy the estimate (10.1), at least when \( \phi \) is analytic. However, the case of a general finite type function \( \phi \) can again be reduced
to the analytic case along the lines of Subsection 6.6. Notice here that we have only made use of the van der Corput type Lemma 10.2 in the preceding estimates, and this lemma allows for small perturbations of the phase function.

What remains to be estimated is the contribution of a small domain of the form (7.1) to $J(\xi)$, i.e., we are left with the oscillatory integral $J^{\rho_0}(\xi)$ which, after a change of coordinates, is given by (7.3). With a slight abuse of notation, we shall therefore adapt the notation from Section 7 and write

$$J(\xi) := J^{\rho_0}(\xi) = \int_{(\mathbb{R}^+)^2} e^{i \left( \xi_1 x_1 + \xi_2 \psi(x_1) + \xi_3 \phi(x) \right)} \rho \left( \frac{x_2}{\varepsilon_0 x_1^2} \right) \eta(x) dx,$$

where here $\phi$ and $\psi$ satisfy the Assumptions 7.3. We may also in this context assume that condition (7.4) is satisfied, since otherwise we can again obtain the desired estimate for $J(\xi)$ by means of Lemma 10.2 applied to the $x_2$-integration in $J(\xi)$. However, under these assumptions we had derived estimates for $J(\xi)$ in Sections 7 and 8, and what remains to be shown is that these estimates are sufficient also in order to establish (10.1).

If $\partial_2 \phi_p(1,0) \neq 0$, then Proposition 7.5 immediately implies the desired estimate, since $h(\phi) \geq 2$.

If $\partial_2 \phi_p(1,0) = 0$, we apply the domain decomposition algorithm of Section 7 and are left with the estimation of the oscillatory integrals $J^n_l$ and $J^{n+1}_l$ defined in that section.

We begin with $J^n(\xi) = \sum_{j,k} J_{j,k}(\xi)$, where $J_{j,k}$ is as defined in Subsection 8.1 and where summation takes place again over the set of indices $j, k$ satisfying (10.3). Observe that according to our discussion in Subsection 7.3 we have here $\kappa_1 = 1/n$, $\kappa_2/\kappa_1 \geq 2$ and $\kappa_1 A_1 + \kappa_2 B_1 = 1$, where $B_1 = B \geq 3$ (compare (7.13)). This implies that $\kappa_2 \leq 1/3$ and hence

$$\kappa_1 \leq 1/6, \quad \kappa_2 \leq 1/3.$$  

From Proposition (8.3) we then conclude that

$$|J_{j,k}(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-j-k}(1 + 2^{-nj} |\xi|)^{-\kappa_1} (1 + 2^{-nj}\sigma_{j,k}(\xi))^{-\kappa_2},$$

hence

$$|J_{j,k}(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-j-k} \left( 1 + 2^{-j} 2^{-(A_{ij}+B_{ik})\kappa_2} |\xi|^{|\kappa_1+\kappa_2|} \right)^{-1}.$$

Then

$$|J^n(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} \left( J^n_0(\xi) + J^n_\infty(\xi) \right),$$

where here

$$J^n_0(\xi) := \sum_{(j,k) \in I_0} 2^{-k+(A_{ij}+B_{ik})\kappa_2} |\xi|^{-|\kappa|},$$

$$J^n_\infty(\xi) := \sum_{(j,k) \in I_\infty} 2^{-j-k},$$
with index sets
\[ I_0 := \{ (j, k) \in \mathbb{N}^2 : j + (A_l j + B_l k) \kappa_2 \leq \log(|\xi|^{|\kappa|}) \text{ and } a_l j \leq k \leq a_{l+1} j \} \]
and
\[ I_{\infty} := \{ (j, k) \in \mathbb{N}^2 : j + (A_l j + B_l k) \kappa_2 > \log(|\xi|^{|\kappa|}) \}. \]

Since \( j \leq k/a_l \) and \( k \leq c \log |\xi| \) in \( I_0 \), summing first in \( j \) and then in \( k \) we obtain
\[ |J_0^n(\xi)| \leq C \sum_{k \leq c \log |\xi|} 2 \left( \frac{\kappa_2}{a_l} \right) (\kappa_2 - 1) k |\xi|^{-|\kappa|} C = \sum_{k \leq c \log |\xi|} 2^{(\kappa_2 - 1)k} |\xi|^{-|\kappa|}. \]

But, \( \kappa_2/\kappa_2' \leq 1 \) by Lemma 8.2, so that \( |J_0^n(\xi)| \leq C(\log |\xi|) |\xi|^{-|\kappa|}. \)

Similarly, since \((A_l j + B_l k) \kappa_2 \leq k\) (compare Lemma 8.2), we have \( j + k > \log(|\xi|^{|\kappa|}) \).

Putting \( r := j + k \), we thus see that
\[ |J_{\infty}^n(\xi)| \leq C \sum_{r \geq \log(|\xi|^{|\kappa|})} r 2^{-r} \leq C' (\log |\xi|) |\xi|^{-|\kappa|}. \]

Since \( |\kappa| = 1/h(\phi) \), we thus see that \( J_{\infty}^n(\xi) \) satisfies estimate (10.1).

What remains are the \( J_{\kappa_{l+1}}(\xi) \), respectively the oscillatory integrals \( J(\xi) \) given by (8.9), which we decompose according to (8.10) into \( J(\xi) = \sum_{k=k_0}^{\infty} J_k(\xi) \). By Proposition 8.4, we have
\[ |J_k(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} \sigma_k^{-\left(\log c + c\right)} (2^{-\|\kappa|^k |\xi|})^{-1/2 - \varepsilon/2} \]
for every sufficiently small \( \varepsilon > 0 \), where \( l_m < 1/4 \), and by the definition of \( J_k(\xi) \) in Subsection 8.2 we also have \( |J_k(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} \).

Putting in the definition of \( \sigma_k \), this implies
\[ |J_k(\xi)| \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} \left( 1 + \sigma_k^{1/2} 2^{-\frac{\kappa'|k}{2}} |\xi|^{1/2} \right)^{-1} \]
\[ \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} \left( 1 + 2^{-\frac{(1+\kappa'')k}{2}} |\xi| \right)^{-1/2} \]
\[ \leq C ||\eta||_{C^3(\mathbb{R}^2)} 2^{-|\kappa'|k} \left( 1 + 2^{-\frac{(1+\kappa'')k}{2}} |\xi| \right)^{-|\kappa|}, \]
because \( \frac{1}{|\kappa|} = h(\phi) \geq 2 \). Moreover, by (8.18), we have \( \frac{(1+\kappa'')k}{2} \leq |\kappa'| \), so that
\[ \sum_{k \leq \log |\xi|} 2^{-|\kappa'|k} 2^{-\frac{(1+\kappa'')k}{2}} |\xi|^{-|\kappa|} \leq C (\log |\xi|) |\xi|^{-|\kappa|}, \]
and
\[ \sum_{k: 2^{-\frac{(1+\kappa'')k}{2}} > |\xi|} 2^{-|\kappa'|k} \leq C |\xi|^{-\frac{2(1+\kappa'')k}{1+\kappa''}} \leq |\xi|^{-|\kappa|}. \]

This shows that also \( J(\xi) \) given by (8.9) satisfies estimate (10.1), which completes the proofs of Theorem 10.1 and Theorem 1.10.
11. Proof of the remaining statements in the Introduction and refined results

In this section, we shall prove the remaining results and claims that have been stated in the Introduction.

11.1. Invariance of the notion of height \( h(x^0, S) \) under affine transformations.

We assume that \( x^0 = (0, 0, 1) =: e_3 \) and \( T_{x^0} = \{ x_3 = 0 \} =: V \), and that our hypersurface \( S \) is the graph

\[
S = \{(x_1, x_2, 1 + \phi(x_1, x_2)) : (x_1, x_2) \in \Omega \}
\]

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

\[
\phi(0, 0) = 0, \quad \nabla \phi(0, 0) = 0.
\]

Consider an affine linear change of coordinates \( F : u \mapsto w + Au \) of \( \mathbb{R}^3 \) which fixes the point \( x^0 \), i.e., \( F(e_3) = e_3 \), and so that the derivative \( DF(x^0) \) leaves the tangent space \( T_{x^0}S \) invariant, i.e., \( A(V) = V \). Here, \( A \in GL(3, \mathbb{R}) \) and \( w \in \mathbb{R}^3 \) is a fixed translation vector. We then denote by \( B := A|_V \) the induced linear isomorphism of \( V \). If we decompose \( w = v + \mu e_3 \), with \( v \in V \) and \( \mu \in \mathbb{R} \), and write elements of \( \mathbb{R}^3 \) as \( (x, x_3) \), with \( x \in \mathbb{R}^2 \), then from \( w + Ae_3 = e_3 \) one computes that

\[
F(x, x_3) = (Bx + (1 - x_3)v, \mu + (1 - \mu)x_3).
\]

Then

\[
F(S) = \{(Bx - \phi(x)v, 1 + (1 - \mu)\phi(x)) : (x_1, x_2) \in \Omega \}.
\]

Notice that \( 1 - \mu \neq 0 \), since \( F \) is assumed to be bijective. By our assumptions on \( \phi \), the mapping \( \varphi : x \mapsto y = Bx - \phi(x)v \) is a local diffeomorphism near the origin with \( \varphi(0) = 0 \), and we can write \( F(S) \) locally as the graph of the smooth function

\[
1 + \tilde{\phi}(y) := 1 + (1 - \mu)\phi(\varphi^{-1}(y)).
\]

Since \( h(\phi) = h(\tilde{\phi}) \), we see that \( h(x^0, S) = h(x^0, F(S)) \), which proves the invariance of our notion of height \( h(x^0, S) \) under affine linear changes of coordinates.

11.2. Proof of Proposition 1.8 and remarks on the critical exponent \( p = h(x^0, S) \). We are first going to prove Proposition 1.8. As outlined in the Introduction, we may assume without loss of generality that the hypersurface \( S \) is given as the graph

\[
S = \{(x_1, x_2, 1 + \phi(x_1, x_2)) : (x_1, x_2) \in \Omega \}
\]

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

\[
\phi(0, 0) = 0, \quad \nabla \phi(0, 0) = 0,
\]

and that \( x^0 = (0, 0, 1) \), so that the affine tangent plane \( x^0 + T_{x^0}S \) is the plane \( \{ x_3 = 1 \} \). Then \( d_{T_{x^0}}(x) = |\phi(x_1, x_2)| \), so that we have to show that for every neighborhood \( \Omega \) of the origin

\[
(11.1) \quad \int_{\Omega} |\phi(x)|^{-1/p} \, dx = \infty
\]
whenever \( p < h(\phi) \). Moreover, if \( \phi \) is analytic, then we need to show that (11.1) holds also for the critical exponent \( p = h(\phi) \).

To this end, observe first that we may reduce ourselves to the case where the coordinates \( x \) are adapted to \( \phi \) by applying the change of coordinates (6.1) (compare [30] and [11]) to the integral in (11.1). Recall that then one of the following three cases applies:

1. \( \pi(\phi) \) is a compact edge, and either \( \kappa \in \mathbb{N} \) or \( \kappa \in \mathbb{N} \) and \( m(\phi_\rho) \leq d(\phi) \).
2. \( \pi(\phi) \) consists of a vertex.
3. \( \pi(\phi) \) is unbounded.

Moreover, in this case we have \( h(\phi) = d(\phi_\rho) \).

First, we consider the cases (a) and (b), where the principal face of the Newton polyhedron of \( \phi \) is a compact set.

**Proposition 11.1.** If the principal face \( \pi(\phi) \) of the Newton polyhedron of the function \( \phi \), when expressed in adapted coordinates, is compact, then (11.1) holds for every \( p \leq h(\phi) \).

**Proof.** As in the proof of Corollary 5.2, we can in this situation choose a weight \( \kappa = (\kappa_1, \kappa_2) \) such that \( h(\phi) = \frac{1}{|\kappa|} = \frac{1}{\kappa_1 + \kappa_2} \), where \( 0 < \kappa_1 \leq \kappa_2 \) without loss of generality. Then the \( \kappa \)-principal part \( \phi_\kappa \) of the function \( \phi \) is a weighted \( \kappa \)-homogeneous polynomial of degree 1.

We may also assume that \( \kappa_1 \) and \( \kappa_2 \) are rational numbers. Then we can find even positive integers \( q_1, q_2 \) and a positive integer \( r \) such that \( \kappa_1 = \frac{r}{q_1}, \kappa_2 = \frac{r}{q_2} \).

The quasi-norm \( N(x) := (x_1^{q_1} + x_2^{q_2})^{1/r} \) is then \( \kappa \)-homogeneous of degree 1 and smooth away from the origin. Denote by \( \Sigma := \{(y_1, y_2) : \rho(y_1, y_2) = 1\} \) the associated "unit circle," and let \( (y_1(\theta), y_2(\theta)) \), \( 0 \leq \theta < 1 \), be a smooth parametrization of \( \Sigma \). We can then introduce generalized polar coordinates \( (\rho, \theta) \) for \( \mathbb{R}^2 \setminus \{0\} \) by writing

\[
x_1 := \rho^{\kappa_1} y_1(\theta), \quad x_2 := \rho^{\kappa_2} y_2(\theta), \quad \rho > 0.
\]

It is well-known and easy to see that the Lebesgue measure on \( \mathbb{R}^2 \) then decomposes as

\[
dx_1 dx_2 = \rho^{k-1} d\gamma(\theta),
\]

where \( d\gamma(\theta) \) is a positive Radon measure such that \( \int_\Sigma d\gamma(\theta) > 0 \). Let us also assume without loss of generality that \( \Omega = \{(x_1, x_2) : \rho(x_1, x_2) < \varepsilon\} \), where \( \varepsilon > 0 \).

If we now decompose \( \phi = \phi_\kappa + \phi_r \) as before into the \( \kappa \)-principal part \( \phi_\kappa \) and the remainder term \( \phi_r \), and express \( \phi \) in polar coordinates \( \tilde{\phi}(\rho, \theta) := \phi(\rho^{\kappa_1} y_1(\theta), \rho^{\kappa_2} y_2(\theta)) \), then

\[
\tilde{\phi}(\rho, \theta) = \rho \left( \tilde{\phi}(1, \theta) + \tilde{\phi}_r(\rho, \theta) \right),
\]

where \( \tilde{\phi}_r(\rho, \theta) = O(\rho^\delta) \) for some \( \delta > 0 \) as \( \rho \to 0 \). In particular, also \( \tilde{\phi}_r(\rho, \theta) \) is bounded, which is all that we need. By passing to these polar coordinates, we obtain

\[
\int_\Omega |\phi(x)|^{-1/h(\phi)} dx = \int_0^\varepsilon \frac{d\rho}{\rho} \int_\Sigma |\tilde{\phi}(1, \theta) + \tilde{\phi}_r(\rho, \theta)|^{-1/h(\phi)} d\gamma(\theta) \geq c \int_0^\varepsilon \frac{d\rho}{\rho}.
\]
In the last inequality $c$ is a positive constant and therefore the integral diverges. This proves the proposition. \qquad Q.E.D.

There remains the case (c) where the principal face is unbounded.

**Proposition 11.2.** Assume that the principal face $\pi(\phi)$ of the Newton polyhedron of the function $\phi$, when expressed in adapted coordinates, is unbounded.

(i) Then (11.1) holds for every $p < h(\phi)$.

(ii) If $\phi$ is assumed to be analytic, then (11.1) holds also for $p = h(\phi)$.

**Proof.** We first prove (i), so assume that $p < h(\phi)$. Here, we can apply a similar reasoning as in the proof of case (c) in Corollary 5.2. The principal face $\pi(\phi)$ is a horizontal half-line, with left endpoint $(\nu_1, N)$, where $\nu_1 < N = h(\phi)$. Notice that $N \geq 2$, since for $N = 1$ we had $\nu_1 = 0$, which is not possible given our assumption $\nabla \phi(0,0) = 0$. We can then choose $\kappa$ with $0 < \kappa_1 < \kappa_2$ so that the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$ is a supporting line to the Newton polyhedron of $\phi$ and that the point $(\nu_1, N)$ is the only point of $\mathcal{N}(\phi)$ on this line. Moreover, we can choose $\kappa_2/\kappa_1$ as large as we wish, so that we may assume that

$$p < |\kappa|^{-1} < h(\phi).$$

Then the $\kappa$-principal part $\phi_\kappa$ of $\phi$ is of the form $\phi_\kappa(x) = cx_1^{\nu_1}x_2^N$, with $c \neq 0$, and it is $\kappa$-homogeneous of degree 1.

By passing to generalized polar coordinates as in the proof of Proposition 11.1 we then see that

$$\int_{\Omega} |\phi(x)|^{-1/p} dx = \int_0^{\varepsilon} \frac{d\rho}{\rho^{2}|\kappa|+1} \int_\Sigma |\tilde{\phi}(1, \theta) + \tilde{\phi}_r(\rho, \theta)|^{-1/p} d\gamma(\theta),$$

where again $\tilde{\phi}_r(\rho, \theta)$ is bounded. Since $\frac{1}{p} - |\kappa| > 0$, we conclude that the last integral diverges.

In order to prove (ii), observe that if $\phi$ is analytic, then there exists a non-trivial analytic function $f$ near the origin so that $\phi(x_1, x_2) = x_2^N f(x_1, x_2)$, where again $N = h(\phi)$. Then, for sufficiently small $\varepsilon > 0$, we have

$$\int_{\Omega} \frac{dx_1 dx_2}{|\phi(x_1, x_2)|^{1/h(\phi)}} \geq \int_{-\varepsilon}^{\varepsilon} \frac{dx_2}{|x_2|} \int_{-\varepsilon}^{\varepsilon} \frac{dx_1}{|f(x_1, x_2)|^{1/2}}.$$

Obviously the last integral diverges. \qquad Q.E.D.

**Remark 11.3.** If $\phi$ is a finite type smooth function and the principal face is a non-compact set then the integral $\int_{\Omega} |\phi(x)|^{-1/h(\phi)} dx$ may be convergent.

An example is given by the function $\phi(x_1, x_2) = x_2^2 + e^{-x_1^{-\alpha}}$ considered by A. Iosevich and E. Sawyer in [15]. Here we have $h(\phi) = 2$, and the associated integral converges whenever $0 < \alpha < 1$. Correspondingly, it has been shown in [15] that the maximal operator associated to the hypersurface $x_3 = 1 + x_2^2 + e^{-x_1^{-\alpha}}$ is $L^2$ bounded whenever...
0 < \alpha < 1 and unbounded for p < 2 (the latter statement follows of course also from Proposition 11.2). However, if \alpha \geq 1, then it is unbounded whenever p \leq 2.

We have thus obtained a confirmation of Iosevich-Sawyer’s conjecture for analytic hypersurfaces [15], and for smooth finite type hypersurfaces we have a partial confirmation of the conjecture. The conjecture remains open when p = h(\phi) in the case where the principal face of \phi is unbounded in an adapted coordinate system.

11.3. Proof of Theorem 1.11. By means of a smooth partition of unity consisting of non-negative functions, we may reduce ourselves to the situation where \rho is supported in a sufficiently small neighborhood of some given point z \in S. Without loss of generality we may then assume that z = 0, and that our hypersurface S is the graph

\[ S = \{(x, \phi(x)) : x \in \Omega\} \]

of a smooth function \phi defined on an open neighborhood \Omega of 0 \in \mathbb{R}^{n-1} and satisfying the conditions

\[ \phi(0) = 0, \nabla \phi(0) = 0. \]

Then the Fourier transform \( \hat{\rho d\sigma}(0, \ldots, 0, \lambda) \) of the superficial measure \( \rho d\sigma \) in direction of the unit normal to S at z = 0 is an oscillatory integral of the form

\[ J(\lambda) = \int_{\mathbb{R}^{n-1}} e^{-i\lambda \phi(x)} \eta(x) \, dx, \]

where 0 \leq \eta \in C_0^\infty(\Omega). By (1.6), we have in particular that

(11.2) \quad |J(\lambda)| \leq C_\beta (1 + |\lambda|)^{-\beta} \text{ for every } \lambda \in \mathbb{R},

where \beta > 0.

Lemma 11.4. If (11.2) holds true, then

(11.3) \quad \int_{\mathbb{R}^{n-1}} |\phi(x)|^{-\gamma} \eta(x) \, dx < \infty

for every \gamma < 1 such that \gamma < \beta.

Proof. We choose a sequence of smooth even functions \( \chi^\nu \in C_0^\infty(\mathbb{R}), \nu \geq 1 \), with compact support such that 0 \leq \chi^\nu(\lambda) \leq 1, \text{ supp} \chi^\nu \subset [-\nu - 1, -1] \cup [1, \nu + 1] and \chi^\nu(\lambda) = 1 for any \lambda \in [-\nu, -2] \cup [2, \nu]. We may clearly choose the \chi^\nu so that for any k \in \mathbb{N} their \( C^k(\mathbb{R}) \) norms are uniformly bounded with respect to \nu.

If \gamma < 1 is given such that \gamma < \beta, then we define the Schwartz functions \( \varphi_\nu \in S(\mathbb{R}) \) by their Fourier transforms

\[ \hat{\varphi_\nu}(s) := \frac{\chi^\nu(\nu s)}{|s|^\gamma}. \]

Then, by standard scaling and integration by parts arguments, one easily finds that

(11.4) \quad |\varphi_\nu(\lambda)| \leq C(1 + |\lambda|)^{\gamma-1}. \]
Consider next the integral
\[ I_\nu := \int_{\mathbb{R}} \varphi_\nu(\lambda) J(\lambda) d\lambda = \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} \varphi_\nu(\lambda) e^{-i\lambda \phi(x)} \eta(x) dx d\lambda. \]

Due to our assumptions on \( \gamma \) and the estimates (11.2) and (11.4) these integrals are uniformly bounded with respect to \( \nu \).

On the other hand since \( \varphi_\nu \) and \( \eta \) both belong to the Schwartz class, we can apply Fubini’s theorem and obtain
\[ I_\nu = \int_{\mathbb{R}^{n-1}} \eta(x) \hat{\varphi}_\nu(\phi(x)) dx. \]

Since the integrand is non-negative, this implies the lower estimate
\[ C \geq |I_\nu| \geq \int_{\frac{1}{\beta} \leq |\phi(x)| \leq 1} |\phi(x)|^{-\gamma} \eta(x) dx \]
for every \( \nu \), where \( C \) is a fixed constant. The estimate (11.3) now follows if we let \( \nu \) tend to infinity.

Q.E.D.

Theorem 1.11 is now an easy consequence of Lemma 11.4. Indeed, by Remark 1.7 it suffices to prove the estimate (1.7) only for affine tangent planes \( H = z + T_z S \) to \( S \), where \( z \in S \) is sufficiently close to the support of \( \rho \). For these, the previous reasoning applies, and since then \( d_H(x) = |\phi(x)| \), we see that (1.7) is an immediate consequence of Lemma 11.4.

**Remark 11.5.** By the same reasoning, Lemma 11.4 also shows that if \( z \in S \) and if \( 0 < \beta \in \mathfrak{B}(z, S) \) and \( \gamma < \min\{1, \beta\} \), then \( \gamma \in \mathcal{C}(z, S) \).

11.4. **Proof of Corollary 1.12.** Note first that always \( h(x^0, S) \geq 1 \). We first assume that \( h(x^0, S) > 1 \). If we had \( \beta > 1/h(x^0, S) \), then we could choose some \( p > 1 \) in this case such that \( \beta > 1/p > 1/h(x^0, S) \). Then Theorem 1.11 in combination with Proposition 1.8 would imply that \( p \leq 1/\beta \), a contradiction.

There remains the case where \( h(x^0, S) = 1 \). We may again assume that \( S \) is given as the graph of a smooth function \( \phi \), with \( \phi \) satisfying (1.2) and \( x^0 = (0,0,0) \). Assuming without loss of generality that the coordinates are adapted to \( \phi \), it is then easy to see that the Hessian matrix \( D^2 \phi(0,0) \) is non-degenerate. The asymptotic form of the method of stationary phase then shows that \( \gamma \leq 1 = 1/h(\phi) = 1/h(x^0, S) \).

Q.E.D.
11.5. **Proof of Theorem 1.13.** Let $S$ be a smooth, finite type hypersurface in $\mathbb{R}^3$, and let $x^0 \in S$ be given. Notice first that Theorem 1.10 implies that
\[
\beta_u(x^0, S) \geq 1/h(x^0, S).
\]
Moreover, by Corollary 1.12 we have $\beta_u(x^0, S) \leq 1/h(x^0, S)$. Indeed, since its proof was based on Proposition 1.8, which made only use of the affine tangent hyperplane at the point $x^0$, with the same arguments restricted to these tangent hyperplane we even obtain
\[
\beta(x^0, S) \leq 1/h(x^0, S).
\]
In combination with (1.11) these estimates imply
\[
\beta_u(x^0, S) = \beta(x^0, S) = 1/h(x^0, S) \leq 1.
\]
Observe next that if $\beta \in \mathcal{B}_u(x^0, S)$, then by Theorem 1.11 and (11.5) we have $\beta \leq 1$ and then $\beta - \varepsilon \in \mathcal{C}_u(x^0, S)$ for every sufficiently small $\varepsilon > 0$. This implies
\[
\beta_u(x^0, S) \leq \gamma_u(x^0, S),
\]
hence by (11.5) and (11.1)
\[
1/h(x^0, S) \leq \gamma_u(x^0, S) \leq \gamma(x^0, S).
\]
Finally, if $\gamma \in \mathcal{C}(x^0, S)$, then putting $p := 1/\gamma$ in Proposition 1.8 we see that $1/\gamma \geq h(x^0, S)$, hence $\gamma \leq 1/h(x^0, S)$. This implies $\gamma(x^0, S) \leq 1/h(x^0, S)$, and in combination with (11.6) we also get
\[
\gamma(x^0, S) = \gamma_u(x^0, S) = 1/h(x^0, S).
\]
This concludes the proof of Theorem 1.13
\[\text{Q.E.D.}\]

**References**

[1] G. I. Arhipov, A. A. Karacuba, and V. N. Čubarikov. Trigonometric integrals. *Izv. Akad. Nauk SSSR Ser. Mat.*, 43(5):971–1003, 1197, 1979.

[2] V. I. Arnol'd. Remarks on the method of stationary phase and on the Coxeter numbers. *Uspehi Mat. Nauk*, 28(5(173)):17–44, 1973.

[3] J. Bourgain. Averages in the plane over convex curves and maximal operators. *J. Analyse Math.*, 47:69–85, 1986.

[4] J. Bruna, A. Nagel, and S. Wainger. Convex hypersurfaces and Fourier transforms. *Ann. of Math. (2)*, 127(2):333–365, 1988.

[5] M. Cowling and G. Mauceri. Inequalities for some maximal functions. II. *Trans. Amer. Math. Soc.*, 296(1):341–365, 1986.

[6] M. Cowling and G. Mauceri. Oscillatory integrals and Fourier transforms of surface carried measures. *Trans. Amer. Math. Soc.*, 304(1):53–68, 1987.

[7] Yngve Domar. On the Banach algebra $A(G)$ for smooth sets $\Gamma \subset \mathbb{R}^n$. *Comment. Math. Helv.*, 52(3):357–371, 1977.

[8] J. J. Duistermaat. Oscillatory integrals, Lagrange immersions and unfolding of singularities. *Comm. Pure Appl. Math.*, 27:207–281, 1974.
[9] A. Greenleaf. Principal curvature and harmonic analysis. *Indiana Univ. Math. J.*, 30(4):519–537, 1981.

[10] Lars Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1990. Distribution theory and Fourier analysis.

[11] I. A. Ikromov and D. Müller. On adapted coordinate systems. *preprint*, 2007.

[12] Isroil A. Ikromov, Michael Kempe, and Detlef Müller. Damped oscillatory integrals and boundedness of maximal operators associated to mixed homogeneous hypersurfaces. *Duke Math. J.*, 126(3):471–490, 2005.

[13] A. Iosevich. Maximal operators associated to families of flat curves in the plane. *Duke Math. J.*, 76(2):633–644, 1994.

[14] A. Iosevich and E. Sawyer. Oscillatory integrals and maximal averages over homogeneous surfaces. *Duke Math. J.*, 82(1):103–141, 1996.

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

[16] V. N. Karapushkin. A theorem on uniform estimates for oscillatory integrals with a phase depending on two variables. *Trudy Sem. Petrovsk.*, (10):150–169, 238, 1984.

[18] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. *J. Amer. Math. Soc.*, 6(1):65–130, 1993.

[19] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Wave front sets, local smoothing and Bourgain’s circular maximal theorem. *Ann. of Math. (2)*, 136(1):207–218, 1992.

[20] A. Nagel, A. Seeger, and S. Wainger. Averages over convex hypersurfaces. *Amer. J. Math.*, 115(4):903–927, 1993.

[21] D. H. Phong and E. M. Stein. The Newton polyhedron and oscillatory integral operators. *Acta Math.*, 179(1):105–152, 1997.

[22] B. Randol. On the asymptotic behavior of the Fourier transform of the indicator function of a convex set. *Trans. Amer. Math. Soc.*, 139:279–285, 1969.

[23] R. Schulz. Convex hypersurfaces of finite type and the asymptotics of their Fourier transforms. *Indiana Univ. Math. J.*, 40(4):1267–1275, 1991.

[24] C. D. Sogge and E. M. Stein. Averages of functions over hypersurfaces in $\mathbb{R}^n$. *Invent. Math.*, 82(3):543–556, 1985.

[25] Christopher D. Sogge. *Fourier integrals in classical analysis*, volume 105 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1993.

[26] Christopher D. Sogge. Maximal operators associated to hypersurfaces with one nonvanishing principal curvature. In *Fourier analysis and partial differential equations (Mirafl ores de la Sierra, 1992)*, Stud. Adv. Math., pages 317–323. CRC, Boca Raton, FL, 1995.

[27] E. M. Stein. Maximal functions. I. Spherical means. *Proc. Nat. Acad. Sci. U.S.A.*, 73(7):2174–2175, 1976.

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

[29] Ingvar Svensson. Estimates for the Fourier transform of the characteristic function of a convex set. *Ark. Mat.*, 9:11–22, 1971.

[30] A.N. Varchenko. Newton polyhedra and estimates of oscillating integrals. *Funkcional. Anal. and Appl.*, 10(3):175–196, 1976.

[31] B. A. Vas’ev. The asymptotic behavior of exponential integrals, the Newton diagram and the classification of minima. *Funkcional. Anal. i Priložen.*, 11(3):1–11, 96, 1977.
