Chapter One

Introduction

Let S be a smooth hypersurface in $\mathbb{R}^3$ with Riemannian surface measure $d\sigma$. We shall assume that S is of *finite type*, that is, that every tangent plane has finite order of contact with S. Consider the compactly supported measure $d\mu := \rho d\sigma$ on S, where $0 \leq \rho \in C_0^\infty(S)$. The central problem that we shall investigate in this monograph is the determination of the range of exponents p for which a Fourier restriction estimate

$$\Big(\int_S |\widehat{f}|^2 \, d\mu\Big)^{1/2} \leq C_p \|f\|_{L^p(\mathbb{R}^3)}, \qquad f \in \mathcal{S}(\mathbb{R}^3), \tag{1.1}$$

holds true.

This problem is a special case of the more general Fourier restriction problem, which asks for the exact range of exponents p and q for which an L^p-L^q Fourier restriction estimate

$$\Big(\int_S |\widehat{f}|^q \, d\mu\Big)^{1/q} \leq C_p \|f\|_{L^p(\mathbb{R}^n)}, \qquad f \in \mathcal{S}(\mathbb{R}^n), \tag{1.2}$$

holds true and which can be formulated for much wider classes of subvarieties S in arbitrary dimension n and suitable measures $d\mu$ supported on S. In fact, as observed by G. Mockenhaupt [M00] (see also the more recent work by I. Łaba and M. Pramanik [LB09]), it makes sense in much wider settings, even for measures $d\mu$ supported on "thin" subsets S of $\mathbb{R}^n$, such as Salem subsets of the real line.

The Fourier restriction problem presents one important instance of a wide circle of related problems, such as the boundedness properties of Bochner Riesz means, dimensional properties of Kakeya type sets, smoothing effects of averaging over time intervals for solutions to the wave equation (or more general dispersive equations), or the study of maximal averages along hypersurfaces. The common question underlying all these problems asks for the understanding of the interplay between the Fourier transform and properties of thin sets in Euclidean space, for instance geometric properties of subvarieties. Some of these aspects have been outlined in the survey article [M14], from which parts of this introduction have been taken.

The idea of Fourier restriction goes back to E. M. Stein, and a first instance of this concept is the determination of the sharp range of L^p-L^q Fourier restriction estimates for the circle in the plane through work by C. Fefferman and E. M. Stein [F70] and A. Zygmund [Z74], who obtained the endpoint estimates (see also L. Hörmander [H73] and L. Carleson and P. Sjölin [CS72] for estimates on more general related oscillatory integral operators). For subvarieties of higher dimension, the first fundamental result was obtained (in various steps) for Euclidean spheres

S^{n-1} by E. M. Stein and P. A. Tomas [To75], who proved that an L^p-L^2 Fourier restriction estimate holds true for S^{n-1}, $n \geq 3$, if and only if $p' \geq 2(2/(n-1)+1)$, where p' denotes the exponent conjugate to p, that is, $1/p + 1/p' = 1$ (cf. [S93] for the history of this result). A crucial property of Euclidean spheres which is essential for this result is the non-vanishing of the Gaussian curvature on these spheres, and indeed an analogous result holds true for every smooth hypersurface S with nonvanishing Gaussian curvature (see [Gl81]).

Fourier restriction estimates have turned out to have numerous applications to other fields. For instance, their great importance to the study of dispersive partial differential equations became evident through R. Strichartz' article [Str77], and in the PDE-literature dual versions which invoke also Plancherel's theorem are often called Strichartz estimates.

The question as to which L^p-L^q Fourier restriction estimates hold true for Euclidean spheres is still widely open. It is conjectured that estimate (1.2) holds true for $S = S^{n-1}$ if and only if $p' > 2n/(n-1)$ and $p' \geq q(2/(n-1)+1)$, and there has been a lot of deep work on this and related problems by numerous mathematicians, including J. Bourgain, T. Wolff, A. Moyua, A. Vargas, L. Vega, and T. Tao (see, e.g., [Bou91], [Bou95], [W95], [MVV96], [TVV98], [W00], [TV00], [T03], and [T04] for a few of the relevant articles, but this list is far from being complete). There has been a lot of work also on conic hypersurfaces and some on even more general classes of hypersurfaces with vanishing Gaussian curvature, for instance in Barcelo [Ba85], [Ba86], in Tao, Vargas, and Vega [TVV98], in Wolff [W01], and in Tao and Vargas [TV00], and more recently by A. Vargas and S. Lee [LV10] and S. Buschenhenke [Bu12]. Again, these citations give only a sample of what has been published on this subject.

Recent work by J. Bourgain and L. Guth [BG11], making use also of multilinear estimates from work by J. Bennett, A. Carbery, and T. Tao [BCT06], has led to further important progress. Nevertheless, this and related problems continue to represent one of the major challenges in Euclidean harmonic analysis, bearing various deep connections with other important open problems, such as the Bochner-Riesz conjecture, the Kakeya conjecture and C. Sogge's local smoothing conjecture for solutions to the wave equation. We refer to Stein's book [S93] for more information on and additional references to these topics and their history until 1993, and to more recent related essays by Tao, for instance in [T04].

As explained before, we shall restrict ourselves to the study of the Stein-Tomas-type estimates (1.1). For convex hypersurfaces of finite line type, a good understanding of this type of restriction estimates is available, even in arbitrary dimension (we refer to the article [I99] by A. Iosevich, which is based on work by J. Bruna, A. Nagel and S. Wainger [BNW88], providing sharp estimates for the Fourier transform of the surface measure on convex hypersurfaces). However, our emphasis will be to allow for very general classes of hypersurfaces $S \subset \mathbb{R}^3$, not necessarily convex, whose Gaussian curvature may vanish on small, or even large subsets.

Given such a hypersurface S, one may ask in terms of which quantities one should describe the range of ps for which (1.1) holds true. It turns out that an extremely useful concept to answer this question is the notion of

Newton polyhedron. The importance of this concept to various problems in analysis and related fields has been revealed by V.I. Arnol'd and his school, in particular through groundbreaking work by A. N. Varchenko [V76] and subsequent work by V. N. Karpushkin [K84] on estimates for oscillatory integrals, and came up again in the seminal article [PS97] by D. H. Phong and E. M. Stein on oscillatory integral operators.

Indeed, there is a close connection between estimates for oscillatory integrals and L^p-L^2 Fourier restriction estimates, which had become evident already through the aforementioned work by Stein and Tomas. The underlying principles had been formalized in a subsequent article by A. Greenleaf [Gl81]. For the case of hypersurfaces, Greenleaf's classical restriction estimate reads as follows:

THEOREM 1.1 (Greenleaf). *Assume that* $|\widehat{d\mu}(\xi)| \lesssim |\xi|^{-1/h}$. *Then the restriction estimate* (1.1) *holds true for every* $p \geq 1$ *such that* $p' \geq 2(h+1)$.

Observe next that in order to establish the restriction estimate (1.1), we may localize the estimate to a sufficiently small neighborhood of a given point x^0 on S. Notice also that if estimate (1.1) holds for the hypersurface S, then it is valid also for every affine-linear image of S, possibly with a different constant if the Jacobian of this map is not one. By applying a suitable Euclidean motion of $\mathbb{R}^3$ we may and shall therefore assume in the sequel that $x^0 = (0, 0, 0)$ and that S is the graph

$$S = S_\phi = \{(x_1, x_2, \phi(x_1, x_2)) : (x_1, x_2) \in \Omega\}$$

of a smooth function ϕ defined on a sufficiently small neighborhood Ω of the origin, such that $\phi(0, 0) = 0$ and $\nabla\phi(0, 0) = 0$.

Then we may write $\widehat{d\mu}(\xi)$ as an oscillatory integral,

$$\widehat{d\mu}(\xi) = J(\xi) := \int_\Omega e^{-i(\xi_3\phi(x_1,x_2)+\xi_1x_1+\xi_2x_2)}\eta(x)\, dx_1 dx_2, \quad \xi \in \mathbb{R}^3,$$

where $\eta \in C_0^\infty(\Omega)$. Since $\nabla\phi(0, 0) = 0$, the complete phase in this oscillatory integral will have no critical point on the support of η unless $|\xi_1| + |\xi_2| \ll |\xi_3|$, provided Ω is chosen sufficiently small. Integrations by parts then show that $\widehat{\mu}(\xi) = O(|\xi|^{-N})$ as $|\xi| \to \infty$, for every $N \in \mathbb{N}$, unless $|\xi_1| + |\xi_2| \ll |\xi_3|$.

We may thus focus on the latter case. In this case, by writing $\lambda = -\xi_3$ and $\xi_j = -s_j\lambda,\ j = 1, 2$, we are reduced to estimating two-dimensional oscillatory integrals of the form

$$I(\lambda; s) := \int e^{i\lambda(\phi(x_1,x_2)+s_1x_1+s_2x_2)}\eta(x_1, x_2)\, dx_1 dx_2,$$

where we may assume without loss of generality that $\lambda \gg 1$ and that $s = (s_1, s_2) \in \mathbb{R}^2$ is sufficiently small, provided that η is supported in a sufficiently small neighborhood of the origin. The complete phase function is thus a small, linear perturbation of the function ϕ.

If $s = 0$, then the function $I(\lambda; 0)$ is given by an oscillatory integral of the form $\int e^{i\lambda\phi(x)}\eta(x)\, dx$, and it is well known ([BG69], [At70]) that for any analytic phase function ϕ defined on a neighborhood of the origin in $\mathbb{R}^n$ satisfying $\phi(0) = 0$, such an integral admits an asymptotic expansion as $\lambda \to \infty$ of the form

$$\sum_{k=0}^{\infty} \sum_{j=0}^{n-1} a_{j,k}(\eta) \lambda^{-r_k} \log(\lambda)^j, \tag{1.3}$$

provided the support of η is sufficiently small. Here, the r_k form an increasing sequence of rational numbers consisting of a finite number of arithmetic progressions, which depends only on the zero set of ϕ, and the $a_{j,k}(\eta)$ are distributions with respect to the cutoff function η. The proof is based on Hironaka's theorem on the resolution of singularities.

We are interested in the case $n = 2$. If we denote the leading exponent r_0 in (1.3) by $r_0 = 1/h$, then we find that the following estimate holds true:

$$|I(\lambda; 0)| \leq C\lambda^{-1/h} \log(\lambda)^{\nu}, \quad \lambda \gg 1, \tag{1.4}$$

where ν may be 0, or 1. Assuming that this estimate is stable under sufficiently small analytic perturbations of ϕ, then we find in particular that $I(\lambda; s)$ satisfies the same estimate (1.4) for $|s|$ sufficiently small, so that we obtain the following *uniform estimate* for $\widehat{d\mu}$,

$$|\widehat{d\mu}(\xi)| \leq C(1 + |\xi|)^{-1/h} \log(2 + |\xi|)^{\nu}, \quad \xi \in \mathbb{R}^3, \tag{1.5}$$

provided the support of ρ is sufficiently small. Greenleaf's theorem then shows that the Fourier restriction estimate (1.1) holds true for $p' \geq 2(h + 1)$, if $\nu = 0$, and at least for $p' > 2(h + 1)$, if $\nu = 1$, where $1/h$ denotes the decay rate of the oscillatory integral $I(\lambda; 0)$, hence ultimately of $\widehat{d\mu}$. Moreover, for instance for hypersurfaces with nonvanishing Gaussian curvature, this yields the sharp restriction result mentioned before.

However, as we shall see, for large classes of hypersurfaces, the relation between the decay rate of the Fourier transform of $d\mu$ and the range of $p's$ for which (1.1) holds true will not be so close anymore.

Nevertheless, uniform decay estimates of the form (1.5) will still play an important role.

The first major question that arises is thus the following one: given a smooth phase function ϕ, how can one determine the sharp decay rate $1/h$ and the exponent ν in the estimate (1.4) for the oscillatory integral $I(\lambda; 0) = \int e^{i\lambda\phi(x)}\eta(x)\,dx$? This question has been answered by Varchenko for analytic ϕ in [V76], where he identified h as the so-called height of the Newton polyhedron associated to ϕ in "adapted" coordinates and also gave a corresponding interpretation of the exponent ν. Subsequently, Karpushkin [K84] showed that the estimates given by Varchenko are stable under small analytic perturbations of the phase function ϕ, which in particular leads to uniform estimates of the form (1.5). More recently, in [IM11b], we proved, by a quite different method, that Karpushkin's result remains valid even for smooth, finite-type functions ϕ, at least for linear perturbations, which is sufficient in order to establish uniform estimates of the form (1.5).

In order to present these results in more detail, let us review some basic notations and results concerning Newton polyhedra (see [V76], [IM11a]).

1.1 NEWTON POLYHEDRA ASSOCIATED WITH ϕ, ADAPTED COORDINATES, AND UNIFORM ESTIMATES FOR OSCILLATORY INTEGRALS WITH PHASE ϕ

We shall build on the results and technics developed in [IM11a] and [IKM10], which will be our main sources, also for references to earlier and related work. Let us first recall some basic notions from [IM11a], which essentially go back to Arnol'd (cf. [Arn73], [AGV88]) and his school, most notably Varchenko [V76].

If ϕ is given as before, consider the associated Taylor series

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

of ϕ centered at the origin. The set

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

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

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

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

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

The *Newton distance* in the sense of Varchenko, or shorter *distance*, $d = d(\phi)$ between the Newton polyhedron and the origin is given by the coordinate d of the point (d, d) at which the bisectrix $t_1 = t_2$ intersects the boundary of the Newton polyhedron. (See Figure 1.1.)

The principal face $\pi(\phi)$ of the Newton polyhedron of ϕ is the face of minimal dimension containing the point (d, d). Deviating from the notation in [V76], we shall call the series

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

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

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

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

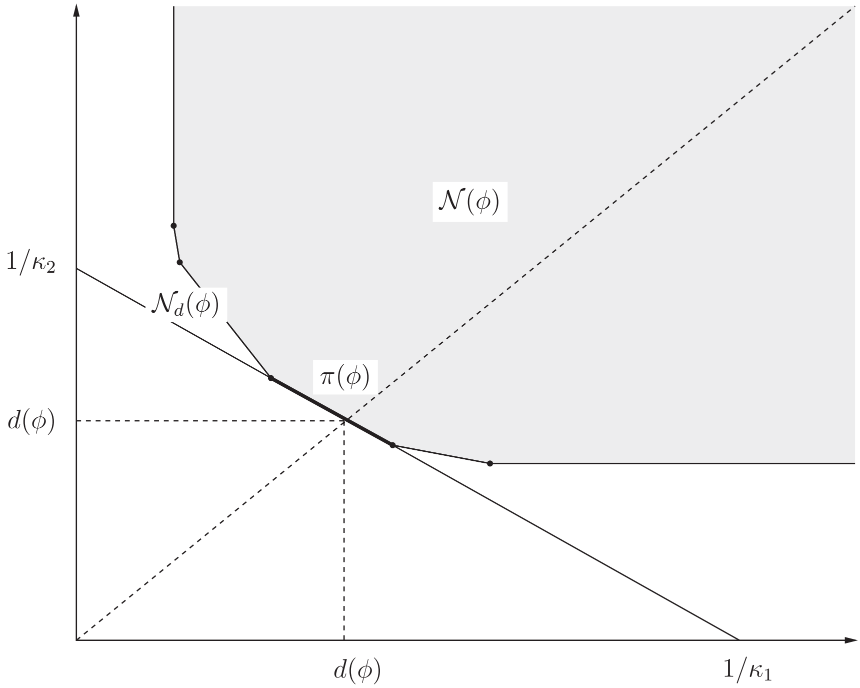

Figure 1.1 Newton polyhedron

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

A given coordinate system x is said to be *adapted* to ϕ if $h(\phi) = d_x$. In [IM11a] we proved that one can always find an adapted local coordinate system in two dimensions, thus generalizing the fundamental work by Varchenko [V76] who worked in the setting of real-analytic functions ϕ (see also [PSS99]).

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

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

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

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

It can then easily be shown (cf. Proposition 2.2 in [IM11a]) that ϕ_{pr} can be factored as

$$\phi_{\mathrm{pr}}(x_1, x_2) = c x_1^{\nu_1} x_2^{\nu_2} \prod_{l=1}^{M} (x_2^q - \lambda_l x_1^p)^{n_l}, \tag{1.6}$$

with $M \geq 1$, distinct nontrivial "roots" $\lambda_l \in \mathbb{C} \setminus \{0\}$ of multiplicities $n_l \in \mathbb{N} \setminus \{0\}$, and trivial roots of multiplicities $\nu_1, \nu_2 \in \mathbb{N}$ at the coordinate axes. Here, p and q are positive integers without common divisor, and $\kappa_2/\kappa_1 = p/q$.

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

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

is a nontrivial polynomial which is κ-homogeneous of degree 1 with respect to the dilations associated to this weight as before, which can be factored in a similar way as in (1.6). By definition, we then have

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

Adaptedness of a given coordinate system can be verified by means of the following proposition (see [IM11a]):

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

$$n(P) := \operatorname{ord}_{S^1} P \tag{1.7}$$

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

$$h(P) = \max\{n(P), d_h(P)\} \tag{1.8}$$

(see Corollary 3.4 in [IM11a]). Moreover, in [IM11a] (Corollary 4.3 and Corollary 5.3), we also proved the following characterization of adaptedness of a given coordinate system.

Proposition 1.2. *The coordinates x are adapted to ϕ if and only if one of the following conditions is satisfied:*

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

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

These conditions had already been introduced by Varchenko, who has shown that they are sufficient for adaptedness when ϕ is analytic.

We also note that in case (a) we have $h(\phi) = h(\phi_{\rm pr}) = d_h(\phi_{\rm pr})$. Moreover, it can be shown that we are in case (a) whenever $\pi(\phi)$ is a compact edge and $\kappa_2/\kappa_1 \notin \mathbb{N}$; in this case we even have $n(\phi_{\rm pr}) < d(\phi)$ (cf. [IM11a], Corollary 2.3).

1.1.1 Construction of adapted coordinates

In the case where the coordinates (x_1, x_2) are not adapted to ϕ, the previous results show that the principal face $\pi(\phi)$ must be a compact edge, that $m := \kappa_2/\kappa_1 \in \mathbb{N}$, and that $n(\phi_{\rm pr}) > d(\phi)$. One easily verifies that this implies that $p = m$ and $q = 1$ in (1.6), and that there is at least one nontrivial, real root $x_2 = \lambda_l x_1^m$ of $\phi_{\rm pr}$ of multiplicity $n_l = n(\phi_{\rm pr}) > d(\phi)$. Indeed, one can show that this root is unique. Putting $b_1 := \lambda_l$, we shall denote the corresponding root $x_2 = b_1 x_1^m$ of $\phi_{\rm pr}$ as its *principal root.*

Changing coordinates

$$y_1 := x_1, \;\; y_2 := x_2 - b_1 x_1^m,$$

we arrive at a "better" coordinate system $y = (y_1, y_2)$. Indeed, this change of coordinates will transform $\phi_{\rm pr}$ into a function $\widetilde{\phi_{\rm pr}}$, where the principal face of $\widetilde{\phi_{\rm pr}}$ will be a horizontal half line at level $t_2 = n(\phi_{\rm pr})$, so that $d(\widetilde{\phi_{\rm pr}}) > d(\phi)$, and correspondingly one finds that $d(\tilde{\phi}) > d(\phi)$ if $\tilde{\phi}$ expresses ϕ in the coordinates y (cf. [IM11a]).

In particular, if the new coordinates y are still not adapted, then the principal face of $\mathcal{N}(\tilde{\phi})$ will again be a compact edge, associated to a weight $\tilde{\kappa} = (\tilde{\kappa}_1, \tilde{\kappa}_2)$ such that $\tilde{m} := \tilde{\kappa}_2/\tilde{\kappa}_1$ is again an integer and $\tilde{m} > m \geq 1$.

Somewhat oversimplifying, by iterating this procedure, we essentially arrive at Varchenko's algorithm for the construction of an adapted coordinate system (cf. [IM11a] for details).

In conclusion, one can show (compare Theorem 5.1 in [IM11a]) that there exists a smooth real-valued function ψ (which we may choose as the *principal root jet* of ϕ) of the form

$$\psi(x_1) = b_1 x_1^m + O(x_1^{m+1}), \tag{1.9}$$

with $b_1 \in \mathbb{R} \setminus \{0\}$, defined on a neighborhood of the origin such that an adapted coordinate system (y_1, y_2) for ϕ is given locally near the origin by means of the (in general nonlinear) shear

$$y_1 := x_1, \;\; y_2 := x_2 - \psi(x_1). \tag{1.10}$$

In these adapted coordinates, ϕ is given by

$$\phi^a(y) := \phi(y_1, y_2 + \psi(y_1)). \tag{1.11}$$

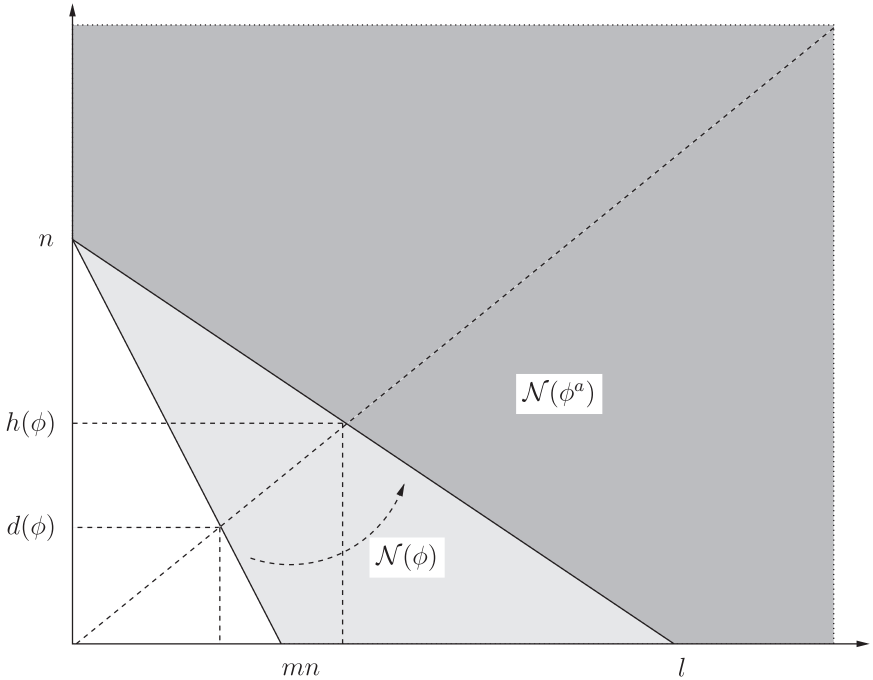

Figure 1.2 $\phi(x_1, x_2) := (x_2 - x_1^m)^n + x_1^\ell \quad (\ell > mn)$

Example 1.3.

$$\phi(x_1, x_2) := (x_2 - x_1^m)^n + x_1^\ell.$$

Assume that $\ell > mn$. Then the coordinates are not adapted. Indeed, $\phi_{\rm pr}(x_1, x_2)$ $=(x_2 - x_1^m)^n$, $d(\phi) = 1/(1/n + 1/(mn)) = mn/(m+1)$ and $n(\phi_{\rm pr}) = n > d(\phi)$. Adapted coordinates are given by $y_1 := x_1$, $y_2 := x_2 - x_1^m$, in which ϕ is expressed by $\phi^a(y) = y_2^n + y_1^\ell$. (See Figure 1.2.)

Remark 1.4. *An alternative proof of Varchenko's theorem on the existence of adapted coordinates for analytic functions ϕ of two variables has been given by Phong, Sturm, and Stein in* [PSS99], *by means of Puiseux series expansions of the roots of ϕ.*

We are now in the position to identify the exponents h and ν in (1.4) and (1.5) in terms of Newton polyhedra associated to ϕ:

If there exists an adapted local coordinate system y near the origin such that the principal face $\pi(\phi^a)$ of ϕ, when expressed by the function ϕ^a in the new coordinates, is a vertex, and if $h(\phi) \geq 2$, then we put $\nu(\phi) := 1$; otherwise, we put $\nu(\phi) := 0$. We remark [IM11b] that the first condition is equivalent to the following one: *If y is any adapted local coordinate system at the origin, then either $\pi(\phi^a)$ is a vertex or a compact edge, and $n(\phi^a_{\rm pr}) = d(\phi^a)$.*

Varchenko [V76] has shown for analytic ϕ that the leading exponent in (1.3) is given by $r_0 = 1/h(\phi)$, and $\nu(\phi)$ is the maximal j for which $a_{j,0}(\eta) \neq 0$. Correspondingly, in [IM11b] we prove, by means of a quite different method, that estimate (1.5) holds true with $h = h(\phi)$ and $\nu = \nu(\phi)$, that is, that the following estimate holds true for ϕ smooth and of finite type:

$$|\widehat{d\mu}(\xi)| \leq C(1+|\xi|)^{-1/h(\phi)} \log(2+|\xi|)^{\nu(\phi)}, \quad \xi \in \mathbb{R}^3. \tag{1.12}$$

The special case where $\xi = (0, 0, \xi_3)$ is normal to S at the origin is due to Greenblatt [Gb09].

One can also show that this estimate is sharp in the exponents even when ϕ is not analytic, except for the case where the principal face $\pi(\phi^a)$ is an unbounded edge (see [IM11b], [M14]).

1.2 FOURIER RESTRICTION IN THE PRESENCE OF A LINEAR COORDINATE SYSTEM THAT IS ADAPTED TO ϕ

Coming back to the restriction estimate (1.1) for our hypersurface S in $\mathbb{R}^3$, we begin with the case where there exists a linear coordinate system that is adapted to the function ϕ. For this case, the following complete answer was given in [IM11b].

Theorem 1.5. *Let $S \subset \mathbb{R}^3$ be a smooth hypersurface of finite type, and fix a point $x^0 \in S$. After applying a suitable Euclidean motion of $\mathbb{R}^3$, let us assume that $x^0 = 0$ and that near x^0 we may view S as the graph S_ϕ of a smooth function ϕ of finite type satisfying $\phi(0,0) = 0$ and $\nabla\phi(0,0) = 0$.*

Assume that, after applying a suitable linear change of coordinates, the coordinates (x_1, x_2) are adapted to ϕ. We then define the critical exponent p_c by

$$p_c' := 2h(\phi) + 2,$$

where p' denotes the exponent conjugate to p, that is, $1/p + 1/p' = 1$.

Then there exists a neighborhood $U \subset S$ of the point x^0 such that for every non-negative density $\rho \in C_0^\infty(U)$, the Fourier restriction estimate (1.1) *holds true for every p such that*

$$1 \leq p \leq p_c. \tag{1.13}$$

Moreover, if $\rho(x^0) \neq 0$, then the condition (1.13) *on p is also necessary for the validity of* (1.1).

Earlier results for particular classes of hypersurfaces in $\mathbb{R}^3$ (which can be seen to satisfy the assumptions of this theorem) are, for instance, in the work by E. Ferreyra and M. Urciuolo [FU04], [FU08] and [FU09], who studied certain classes of quasi-homogeneous hypersurfaces, for which they were able to prove L^p-L^q-restriction estimates when $p < \frac{4}{3}$. For further progress in the study of these classes of hypersurfaces, we refer to the work by S. Buschenhenke, A. Vargas and the second named author [BMV15]. We also like to mention work by A. Magyar [M09] on L^p-L^2 Fourier restriction estimates for some classes of analytic hypersurfaces, which preceeded [IM11b] .

As shown in [IM11b], the necessity of condition (1.13) follows easily by means of Knapp type examples (a related discussion of Knapp type arguments is given in Section 1.4). It is here where we need to assume that there is a linear coordinate system which is adapted to ϕ. The sufficiency of condition (1.13) is immediate from Greenleaf's Theorem 1.1 in combination with (1.12), in the case where $\nu(\phi) = 0$. Notice that this is true, no matter whether or not there exists a linear coordinate system that is adapted to ϕ.

If $\nu(\phi) = 1$, then we just miss the endpoint $p = p_c$, which ultimately can be dealt with by means of Littlewood-Paley theory (for more details, we refer the reader to [IM11b] and also the survey article [M14]). An analogous argument based on Littlewood-Paley theory will appear in Chapter 3.

In view of Theorem 1.5, from now on we shall always make the following assumption, unless stated explicitly otherwise.

Assumption 1.6 (NLA). *There is no linear coordinate system that is adapted to ϕ.*

Our main goal will be to understand which Fourier restriction estimates of the form (1.1) will hold under this assumption.

1.3 FOURIER RESTRICTION WHEN NO LINEAR COORDINATE SYSTEM IS ADAPTED TO ϕ—THE ANALYTIC CASE

Under the preceding assumption, but not yet assuming that ϕ is analytic, let us have another look at the first step of Varchenko's algorithm. If here $m = \kappa_2/\kappa_1 = 1$, then this leads to a linear change of coordinates of the form $y_1 = x_1$, $y_2 = x_2 - b_1 x_1$, which will transform ϕ into a function $\tilde{\phi}$ for which, by our assumption, the coordinates (y_1, y_2) are still not adapted. Replacing ϕ by $\tilde{\phi}$, it is also immediate that estimate (1.1) will hold for the graph of ϕ if and only if it holds for the graph of $\tilde{\phi}$. Replacing ϕ by $\tilde{\phi}$, we may and shall therefore always assume that our original coordinate system (x_1, x_2) is chosen so that

$$m = \frac{\kappa_2}{\kappa_1} \in \mathbb{N} \qquad \text{and } m \geq 2. \tag{1.14}$$

The next proposition will show that such a linear coordinate system is linearly adapted to ϕ in the following sense.

In analogy with Varchenko's notion of height, we can introduce the notion of *linear height* of ϕ, which measures the upper limit of all Newton distances of ϕ in linear coordinate systems:

$$h_{\text{lin}}(\phi) := \sup\{d(\phi \circ T) : T \in GL(2, \mathbb{R})\}.$$

Note that

$$d(\phi) \leq h_{\text{lin}}(\phi) \leq h(\phi).$$

We also say that a linear coordinate system $y = (y_1, y_2)$ is *linearly adapted* to ϕ if $d_y = h_{\text{lin}}(\phi)$. Clearly, if there is a linear coordinate system that is adapted to ϕ, it is in particular linearly adapted to ϕ. The following proposition, whose proof will be given in Chapter 9 (Appendix A), gives a characterization of linearly adapted coordinates under Assumption 1.6 (NLA).

Proposition 1.7. *If ϕ satisfies Assumption (NLA) and if $\phi = \phi(x)$, then the following are equivalent:*

(a) *The coordinates x are linearly adapted to ϕ.*
(b) *If the principal face $\pi(\phi)$ is contained in the line*

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

then either $\kappa_2/\kappa_1 \geq 2$ or $\kappa_1/\kappa_2 \geq 2$.

Moreover, in all linearly adapted coordinates x for which $\kappa_2/\kappa_1 > 1$, the principal face of the Newton polyhedron is the same, so that in particular the numbers $m := \kappa_2/\kappa_1$ and d_x do not depend on the choice of the linearly adapted coordinate system $x = (x_1, x_2)$.

This result shows in particular that linearly adapted coordinates always do exist under Assumption (NLA), since either the original coordinates for ϕ are already linearly adapted or we arrive at such coordinates after applying the first step in Varchenko's algorithm (when $\kappa_2/\kappa_1 = 1$ in the original coordinates).

Let us then look at the Newton polyhedron $\mathcal{N}(\phi^a)$ of ϕ^a, which expresses ϕ in the adapted coordinates (y_1, y_2) of (1.11), and denote the vertices of the Newton polyhedron $\mathcal{N}(\phi^a)$ by (A_l, B_l), $l = 0, \dots, n$, where we assume that they are ordered so that $A_{l-1} < A_l$, $l = 1, \dots, n$, with associated compact edges given by the intervals $\gamma_l := [(A_{l-1}, B_{l-1}), (A_l, B_l)]$, $l = 1, \dots, n$. The unbounded horizontal edge with left endpoint (A_n, B_n) will be denoted by γ_{n+1}. To each of these edges γ_l, we associate the weight $\kappa^l = (\kappa_1^l, \kappa_2^l)$ so that γ_l is contained in the line

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

For $l = n + 1$, we have $\kappa_1^{n+1} := 0$, $\kappa_2^{n+1} = 1/B_n$. We denote by

$$a_l := \frac{\kappa_2^l}{\kappa_1^l}, \qquad l = 1, \dots, n,$$

the reciprocal of the (modulus of the) slope of the line L_l. For $l = n + 1$, we formally set $a_{n+1} := \infty$. (See Figure 1.3.)

If $l \leq n$, then the κ^l-principal part ϕ_{κ^l} of ϕ corresponding to the supporting line L_l can easily be shown to be of the form

$$\phi_{\kappa^l}(y) = c_l\, y_1^{A_{l-1}} y_2^{B_l} \prod_\alpha \left(y_2 - c_l^\alpha\, y_1^{a_l}\right)^{N_\alpha} \tag{1.15}$$

(cf. Proposition 2.2 in [IM11a]; also [IKM10]).

Remark 1.8. *When ϕ is analytic, then this expression is linked to the Puiseux series expansion of roots of ϕ^a as follows* [PS97] (*compare also* [IM11a]): *We may then factor*

$$\phi^a(y_1, y_2) = U(y_1, y_2)\, y_1^{\nu_1} y_2^{\nu_2} \prod_r (y_2 - r(y_1)),$$

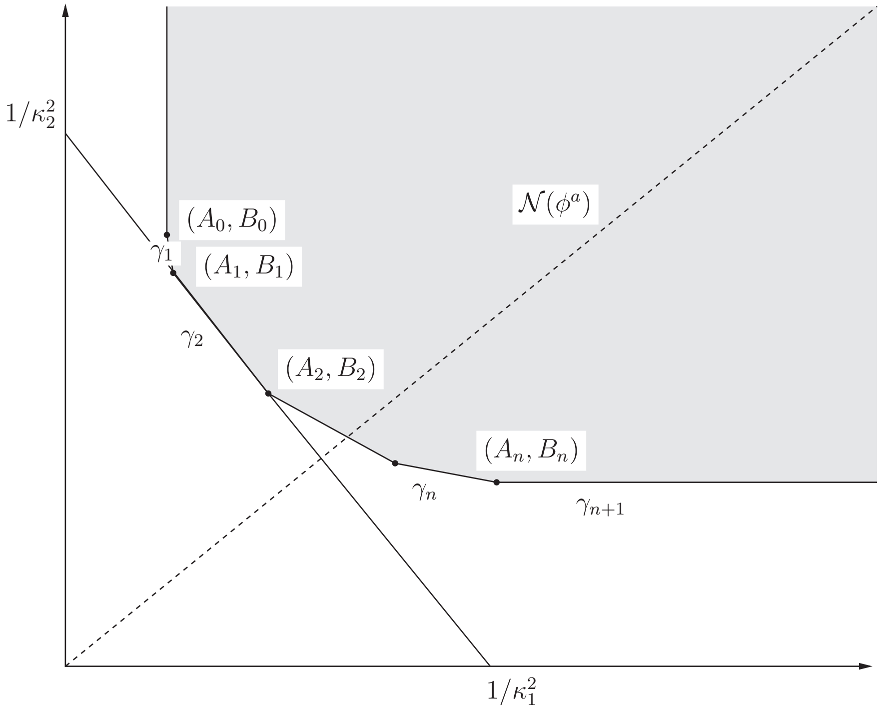

Figure 1.3 Edges and weights

where the product is indexed by the set of all nontrivial roots $r = r(y_1)$ *of* ϕ^a *(which may also be empty) and where* U *is analytic, with* $U(0,0) \neq 0$*. Moreover, these roots can be expressed in a small neighborhood of* 0 *as Puiseux series*

$$r(y_1) = c_{l_1}^{\alpha_1} y_1^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2} y_1^{a_{l_1 l_2}^{\alpha_1}} + \cdots + c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_p} y_1^{a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}}} + \cdots ,$$

where

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

$$a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}} > a_{l_1 \cdots l_{p-1}}^{\alpha_1 \cdots \alpha_{p-2}},$$

with strictly positive exponents $a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}} > 0$ *(which are all multiples of a fixed positive rational number) and nonzero 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 nonidentical roots of* ϕ^a*. The leading exponents in these series are the numbers*

$$a_1 < a_2 < \cdots < a_n.$$

One can therefore group the roots into the clusters of roots $[l], l = 1, \ldots, n$*, where the* l*th cluster* $[l]$ *consists of all roots with leading exponent* a_l*.*

Correspondingly, we can decompose

$$\phi^a(y_1, y_2) = U(y_1, y_2) y_1^{\nu_1} y_2^{\nu_2} \prod_{l=1}^{n} \Phi_{[l]}(y_1, y_2),$$

where

$$\Phi_{[l]}(y_1, y_2) := \prod_{r \in [l]} (y_2 - r(y_1)).$$

More generally, by the cluster $\begin{bmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix}$*, we shall designate all the roots* $r(y_1)$*, counted with their multiplicities, that satisfy*

$$r(y_1) - \left(c_{l_1}^{\alpha_1} y_1^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2} y_1^{a_{l_1 l_2}^{\alpha_1}} + \cdots + c_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_p} y_1^{a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}}} \right) = O(y_1^b)$$

for some exponent $b > a_{l_1 \cdots l_p}^{\alpha_1 \cdots \alpha_{p-1}}$*. Then the cluster* $[l] = [l_1]$ *will split into the clusters* $\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}$*, these cluster into the finer "subclusters"* $\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}$*, and so on.*

Observe also the following: If $\delta_s^l(y_1, y_2) = (s^{\kappa_1^l} y_1, s^{\kappa_2^l} y_2),\ s > 0$*, denote the dilations associated to the weight* κ^l*, and if* $r \in [l_1]$ *is a root in the cluster* $[l_1]$*, then one easily checks that for* $y = (y_1, y_2)$ *in a bounded set we have* $\delta_s^l y_2 = s^{\kappa_2^l} y_2$ *and* $r(\delta_s^l y_1) = s^{a_{l_1} \kappa_1^l} c_{l_1}^{\alpha_1} y_1^{a_{l_1}} + O(s^{a_{l_1} \kappa_1^l + \varepsilon})$ *as* $s \to 0$*, for some* $\varepsilon > 0$*. Consequently,*

$$\delta_s^l y_2 - r(\delta_s^l y_1) = \begin{cases} -s^{a_{l_1} \kappa_1^l} c_{l_1}^{\alpha_1} y_1^{a_{l_1}} + O(s^{a_{l_1} \kappa_1^l + \varepsilon}), & \text{if } l_1 < l, \\ s^{\kappa_2^l} (y_2 - c_l^{\alpha_l} y_1^{a_l}) + O(s^{\kappa_2^l + \varepsilon}), & \text{if } l_1 = l, \\ s^{\kappa_2^l} y_2 + O(s^{\kappa_2^l + \varepsilon}), & \text{if } l_1 > l. \end{cases}$$

This shows that the κ^l*-principal part of* ϕ^a *is given by*

$$\phi_{\kappa^l}^a = C_l y_1^{\nu_1 + \sum_{l_1 < l} |[l_1]| a_{l_1}} y_2^{\nu_2 + \sum_{l_1 > l} |[l_1]|} \prod_{\alpha_1} (y_2 - c_l^{\alpha_1} y_1^{a_l})^{N_{l,\alpha_1}}, \tag{1.16}$$

where N_{l,α_1} *denotes the number of roots in the cluster* $[l]$ *with leading term* $c_l^{\alpha_1} y_1^{a_l}$*, and where by* $|M|$ *we denote the cardinality of a set* M*.*

A look at the Newton polyhedron reveals that the exponents of y_1 *and* y_2 *in* (1.16) *can be expressed in terms of the vertices* (A_j, B_j) *of the Newton polyhedron:*

$$\nu_1 + \sum_{l_1 < l} |[l_1]| a_{l_1} = A_{l-1}, \qquad \nu_2 + \sum_{l_1 > l} |[l_1]| = B_l.$$

Notice also that

$$\prod_{\alpha_1} (y_2 - c_l^{\alpha_1} y_1^{a_l})^{N_{l,\alpha_1}} = (\Phi_{[l]})_{\kappa^l}.$$

Comparing this with (1.15)*, the close relation between the Newton polyhedron of* ϕ^a *and the Puiseux series expansion of roots becomes evident, and accordingly we say that the edge* $\gamma_l := [(A_{l-1}, B_{l-1}), (A_l, B_l)]$ *is associated to the cluster of roots* $[l]$.

Consider next the line parallel to the bisectrix

$$\Delta^{(m)} := \{(t, t+m+1) : t \in \mathbb{R}\}.$$

For any edge $\gamma_l \subset L_l := \{(t_1, t_2) \in \mathbb{R}^2 : \kappa_1^l t_1 + \kappa_2^l t_2 = 1\}$, define h_l by

$$\Delta^{(m)} \cap L_l = \{(h_l - m, h_l + 1)\},$$

that is,

$$h_l = \frac{1 + m\kappa_1^l - \kappa_2^l}{\kappa_1^l + \kappa_2^l}, \tag{1.17}$$

and define the *restriction height*, or for short, *r-height,* of ϕ by

$$h^r(\phi) := \max(d, \max_{\{l=1,\dots,n+1:\, a_l > m\}} h_l).$$

Here, $d = d_x$ denotes again the Newton distance $d(\phi)$ of ϕ with respect to our original, linearly adapted coordinates $x = (x_1, x_2)$. Recall also that by Proposition 1.7 the numbers m and d_x are well defined, that is, they do not depend on the chosen linearly adapted coordinate system x. For a more invariant description of $h^r(\phi)$, we refer the reader to Proposition 1.17.

Remarks 1.9. (a) *For L in place of L_l and κ in place of κ^l, one has $m = \kappa_2/\kappa_1$ and $d = 1/(\kappa_1 + \kappa_2)$, so that one gets d in place of h_l in* (1.17).

(b) *Since $m < a_l$, we have $h_l < 1/(\kappa_1^l + \kappa_2^l)$, hence $h^r(\phi) < h(\phi)$. On the other hand, since the line $\Delta^{(m)}$ lies above the bisectrix, it is obvious that $h^r(\phi) + 1 \geq h(\phi)$, so that*

$$h(\phi) - 1 \leq h^r(\phi) < h(\phi). \tag{1.18}$$

(See Figure 1.4.)

It is easy to see from Remark 1.9(a) that the r-height admits the following *geometric interpretation.*

By following Varchenko's algorithm (cf. Subsection 8.2 of [IKM10]), one realizes that the Newton polyhedron of ϕ^a intersects the line L of the Newton polyhedron of ϕ in a compact face, either in a single vertex or a compact edge. That is, the intersection contains at least one and at most two vertices of ϕ^a, and we choose (A_{l_0-1}, B_{l_0-1}) as the one with smallest second coordinate. Then l_0 is the smallest index l such that γ_l has a slope smaller than the slope of L, that is, $a_{l_0-1} \leq m < a_{l_0}$.

We may thus consider the *augmented Newton polyhedron* $\mathcal{N}^r(\phi^a)$ of ϕ^a, which is the convex hull of the union of $\mathcal{N}(\phi^a)$ with the half line $L^+ \subset L$ with right endpoint (A_{l_0-1}, B_{l_0-1}). Then $h^r(\phi) + 1$ is the second coordinate of the point at which the line $\Delta^{(m)}$ intersects the boundary of $\mathcal{N}^r(\phi^a)$.

We remind the reader that all notions that we have introduced so far (with the exception of those discussed in Remark 1.8) make perfect sense for arbitrary smooth functions ϕ of finite type, in particular, for analytic ϕ. For real analytic hypersurfaces, it turns out that we now have all necessary notions at hand in order to formulate the central result of this monograph. The extension to more general classes of smooth, finite-type hypersurfaces will require further notions and will be discussed in the next section (compare Theorem 1.14).

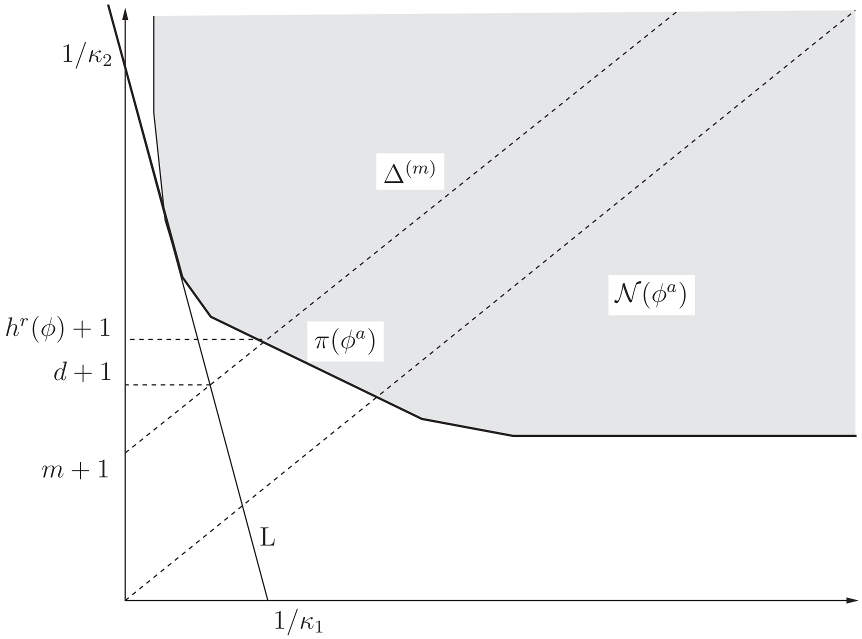

Figure 1.4 r-height

THEOREM 1.10. *Let $S \subset \mathbb{R}^3$ be a real analytic hypersurface of finite type, and fix a point $x^0 \in S$. After applying a suitable Euclidean motion of $\mathbb{R}^3$, let us assume that $x^0 = 0$ and that near x^0 we may view S as the graph S_ϕ of a real analytic function ϕ satisfying $\phi(0,0) = 0$ and $\nabla\phi(0,0) = 0$.*

Assume that there is no linear coordinate system adapted to ϕ. Then there exists a neighborhood $U \subset S$ of x^0 such that for every nonnegative density $\rho \in C_0^\infty(U)$, the Fourier restriction estimate (1.1) *holds true for every $p \geq 1$ such that $p' \geq p'_c := 2h^r(\phi) + 2$.*

REMARKS 1.11. (a) *An application of Greenleaf's result would imply, at best, that the condition $p' \geq 2h(\phi) + 2$ is sufficient for* (1.1) *to hold, which is a strictly stronger condition than $p' \geq p'_c$.*

(b) *In a preprint, which regretfully has remained unpublished and which has been brought to our attention by A. Seeger after we had found our results, H. Schulz* [Sc90] *had already observed this kind of phenomenon for particular examples of surfaces of revolution.*

EXAMPLE 1.12.

$$\phi(x_1, x_2) := (x_2 - x_1^m)^n, \qquad n, m \geq 2.$$

The coordinates (x_1, x_2) are not adapted. Adapted coordinates are $y_1 := x_1$, $y_2 := x_2 - x_1^m$, in which ϕ is given by

$$\phi^a(y_1, y_2) = y_2^n.$$

Here

$$\kappa_1 = \frac{1}{mn}, \quad \kappa_2 = \frac{1}{n},$$
$$d := d(\phi) = \frac{1}{\kappa_1 + \kappa_2} = \frac{nm}{m+1},$$

and

$$p'_c = \begin{cases} 2d+2, & \text{if } n \le m+1, \\ 2n, & \text{if } n > m+1. \end{cases}$$

On the other hand, $h := h(\phi) = n$, so that $2h + 2 = 2n + 2 > p'_c$.

1.4 SMOOTH HYPERSURFACES OF FINITE TYPE, CONDITION (R), AND THE GENERAL RESTRICTION THEOREM

Theorem 1.10 can be extended to smooth, finite-type functions ϕ under an additional Condition (R), which, however, is always satisfied when ϕ is real analytic. To state this more general result and in order to prepare a more invariant description of the notion of r-height, we need to introduce more notation. Again, we shall assume that the coordinates (x_1, x_2) are linearly adapted to ϕ.

1.4.1 Fractional shears and condition (R)

We need a few more definitions. Consider the half lines $\mathbb{R}_\pm := \{x_1 \in \mathbb{R} : \pm x_1 > 0\}$, and denote by $H^\pm := \mathbb{R}_\pm \times \mathbb{R}$ the corresponding right (respectively, left) half plane.

We say that a function $f = f(x_1)$ defined in $U \cap \mathbb{R}_+$ (respectively, $U \cap \mathbb{R}_-$), where U is an open neighborhood of the origin, is *fractionally smooth* if there exist a smooth function g on U and a positive integer q such that $f(x_1) = g(|x_1|^{1/q})$ for $x_1 \in U \cap \mathbb{R}_+$ (respectively, $x_1 \in U \cap \mathbb{R}_-$). Moreover, we shall say that a fractionally smooth function f is *flat* if $f(x_1) = O(|x_1|^N)$ as $x_1 \to 0$, for every $N \in \mathbb{N}$. Notice that this notion of flatness describes only the behavior at the origin. Observe also that a fractionally smooth function that is flat is even smooth.

Two fractionally smooth functions f and g defined on a neighborhood of the origin will be called *equivalent*, and we shall write $f \sim g$, if $f - g$ is flat. Finally, a *fractional shear* in $H^\pm$ will be a change of coordinates of the form

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

where f is real valued and fractionally smooth but not flat. If we express the smooth function ϕ on, say, the half plane H^+ as a function of $y = (y_1, y_2)$, the resulting function

$$\phi^f(y) := \phi(y_1, y_2 + f(y_1))$$

will in general no longer be smooth at the origin, but fractionally smooth in the sense described next.

For such functions, there are straightforward generalizations of the notions of Newton polyhedron, and so forth. Namely, following [IKM10] and assuming without loss of generality that we are in H^+ where $x_1 > 0$, let Φ be any smooth function of the variables $x_1^{1/q}$ and x_2 near the origin; that is, there exists a smooth function $\Phi^{[q]}$ near the origin such that $\Phi(x) = \Phi^{[q]}(x_1^{1/q}, x_2)$ (more generally, one could assume that Φ is a smooth function of the variables $x_1^{1/q}$ and $x_2^{1/p}$, where p and q are positive integers, but we won't need this generality here and shall, therefore, always assume that $p = 1$). Such functions Φ will also be called *fractionally smooth.* If the formal Taylor series of $\Phi^{[q]}$ is given by

$$\Phi^{[q]}(x_1, x_2) \sim \sum_{\alpha_1,\alpha_2=0}^{\infty} c_{\alpha_1,\alpha_2} x_1^{\alpha_1} x_2^{\alpha_2},$$

then Φ has the formal Puiseux series expansion

$$\Phi(x_1, x_2) \sim \sum_{\alpha_1,\alpha_2=0}^{\infty} c_{\alpha_1,\alpha_2} x_1^{\alpha_1/q} x_2^{\alpha_2}.$$

We therefore define the *Taylor-Puiseux support*, or *Taylor support*, of Φ by

$$\mathcal{T}(\Phi) := \{(\tfrac{\alpha_1}{q}, \alpha_2) \in \mathbb{N}_q^2 : c_{\alpha_1,\alpha_2} \neq 0\},$$

where $\mathbb{N}_q^2 := (\frac{1}{q}\mathbb{N}) \times \mathbb{N}$. The *Newton-Puiseux polyhedron* (or *Newton polyhedron*) $\mathcal{N}(\Phi)$ of Φ at the origin is then defined to be the convex hull of the union of all the quadrants $(\alpha_1/q, \alpha_2) + \mathbb{R}_+^2$ in $\mathbb{R}^2$, with $(\alpha_1/q, \alpha_2) \in \mathcal{T}(\Phi)$, and other notions, such as the notion of principal face, Newton distance, or homogenous distance, are defined in analogy with our previous definitions for smooth functions ϕ.

Coming back to our fractional shear f, assume that $f(x_1)$ has the formal Puiseux series expansion (say for $x_1 > 0$)

$$f(x_1) \sim \sum_{j \geq 0} c_j x_1^{m_j}, \tag{1.19}$$

with nonzero coefficients c_j and exponents m_j, which are growing with j and are all multiples of $1/q$. We then isolate the leading exponent m_0 and choose the weight κ^f so that $\kappa_2^f/\kappa_1^f = m_0$ and such that the line

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

is a supporting line to $\mathcal{N}(\phi^f)$.

In analogy with $h^r(\phi)$, by replacing the exponent m by m_0 and the line L by L^f, we can then define the *r-height* $h^f(\phi)$ *associated with* f by putting

$$h^f(\phi) = \max(d^f, \max_{\{l : a_l > m_0\}} h_l^f), \tag{1.20}$$

where (d^f, d^f) is the point of intersection of the line L^f with the bisectrix and where h_l^f is associated to the edge γ_l of $\mathcal{N}(\phi^f)$ by the analogue of formula (1.17), that is,

$$h_l^f = \frac{1 + m_0\kappa_1^l - \kappa_2^l}{\kappa_1^l + \kappa_2^l}, \tag{1.21}$$

if γ_l is again contained in the line L_l defined by the weight κ^l.

In a similar way as for the notion of r-height, we can reinterpret $h^f(\phi)$ *geometrically* as follows. We define the *augmented Newton polyhedron* $\mathcal{N}^f(\phi^f)$ as the convex hull of the union of $\mathcal{N}(\phi^f)$ with the half line $(L^f)^+ \subset L^f$, whose right endpoint is the vertex of $\mathcal{N}(\phi^f) \cap L^f$ with the smallest second coordinate. Then $h^f(\phi) + 1$ is the second coordinate of the point at which the line $\Delta^{(m_0)}$ intersects the boundary of $\mathcal{N}^f(\phi^f)$.

Finally, let us say that a fractionally smooth function $f(x_1)$ *agrees with the principal root jet* $\psi(x_1)$ *up to terms of higher order* if the following holds: if ψ is not a polynomial, then we require that $f \sim \psi$, and if ψ is a polynomial of degree D, then we require that the leading exponent in the formal Puiseux series expansion of $f - \psi$ is strictly bigger than D.

Such functions f will indeed arise in Section 6.5 in the course of an algorithm that will allow us to analyze the fine splitting of certain roots near the principal root jet. It is this fine splitting that will lead to terms of higher order that have to be added to ψ.

We can now formulate the extra "root" condition that we need when ϕ is nonanalytic.

Condition (R). For every fractionally smooth, real function $f(x_1)$ that agrees with the principal root jet $\psi(x_1)$ up to terms of higher order, the following holds true:

If $B \in \mathbb{N}$ is maximal such that $\mathcal{N}(\phi^f) \subset \{(t_1, t_2) : t_2 \geq B\}$, and if $B \geq 1$, then ϕ factors as $\phi(x_1, x_2) = (x_2 - \tilde{f}(x_1))^B \tilde{\phi}(x_1, x_2)$, where $\tilde{f} \sim f$ and where $\tilde{\phi}$ is fractionally smooth.

Examples 1.13. (a) If $\varphi(x_1)$ is flat and nontrivial and if $m \geq 2$ and $B \geq 2$, then the function $\phi_1(x_1, x_2) := (x_2 - x_1^m - \varphi(x_1))^B$ does satisfy Condition (R), whereas (R) fails for $\phi_2(x_1, x_2) := (x_2 - x_1^m)^B + \varphi(x_1)$. In these examples, we have $\psi(x_1) = x_1^m$.

(b) $\phi_3(x_1, x_2) := (x_2 - x_1^2 - \varphi_1(x_1))^5\big((x_2 - 3x_1^2)^2 + \varphi_2(x_1)\big)$ does satisfy Condition (R) for arbitrary flat functions $\varphi_1(x_1)$ and $\varphi_2(x_1)$. Indeed, in this example the principal face $\pi(\phi_3)$ is the interval $[(0, 7), (14, 0)]$, and so one easily finds that $d = 1/(\frac{1}{7} + \frac{1}{14}) = \frac{14}{3}$ and $\psi(x_1) = x_1^2$. Notice that $\psi(x_1)$ is a root of multiplicity $5 > \frac{14}{3}$ of $(\phi_3)_{\mathrm{pr}}$, and hence the principal root.

Real analytic functions ϕ are easily seen to satisfy Condition (R). Indeed, the definition of B implies that

$$\phi^f(y_1, y_2) = y_2^B h(y_1, y_2) + \varphi(y_1, y_2),$$

where $\varphi(y_1, y_2)$ is flat in y_1 and where the mapping $y_1 \mapsto h(y_1, 0)$ is of finite type. In particular, $g(x_1) := \phi(x_1, f(x_1)) = \phi^f(y_1, 0) = \varphi(y_1, 0)$ is flat. On the other hand, if the function ϕ is analytic, then we may factor it as

$$\phi(x_1, x_2) = U(x_1, x_2)x_1^{\nu_1}x_2^{\nu_2}\prod_{j=1}^{J}(x_2 - r_j(x_1))^{n_j},$$

where the $r_j = r_j(x_1)$ denote the distinct nontrivial roots of ϕ and n_j their multiplicities, and where U is analytic near the origin such that $U(0, 0) \neq 0$. Moreover, the roots r_j admit Puiseux series expansions near the origin (compare Remark 1.8). But then

$$g(x_1) = U(x_1, f(x_1))x_1^{\nu_1} f(x_1)^{\nu_2}\prod_{j=1}^{J}(f(x_1) - r_j(x_1))^{n_j},$$

and since g is flat, this shows that necessarily $f - r_k$ is flat for some $k \in \{1, \ldots, J\}$, that is, $f \sim r_k$. Then r_k must be a real root of ϕ, and the identity

$$\phi^f(y_1, y_2) = U(y_1, y_2 + f(y_1))y_1^{\nu_1}\big(y_2 + f(y_1)\big)^{\nu_2}\prod_{j=1}^{J}\big(y_2 + f(y_1) - r_j(y_1)\big)^{n_j}$$

shows that $B = n_k$. By choosing $\tilde{f} := r_k$, we thus find that indeed $\phi(x_1, x_2) = (x_2 - \tilde{f}(x_1))^B\tilde{\phi}(x_1, x_2)$, with $\tilde{\phi}(x_1, x_2) := U(x_1, x_2)x_1^{\nu_1}x_2^{\nu_2}\prod_{j\neq k}(x_2 - r_j(x_1))^{n_j}$.

1.4.2 The general restriction theorem and sharpness of its conditions

We can now state our main result.

Theorem 1.14. *Let $S \subset \mathbb{R}^3$ be a smooth hypersurface of finite type, and fix a point $x^0 \in S$. After applying a suitable Euclidean motion of $\mathbb{R}^3$, let us assume that $x^0 = 0$ and that near x^0 we may view S as the graph S_ϕ of a smooth function ϕ of finite type satisfying $\phi(0, 0) = 0$ and $\nabla\phi(0, 0) = 0$.*

Assume that the coordinates (x_1, x_2) are linearly adapted to ϕ but not adapted and that Condition (R) is satisfied.

Then there exists a neighborhood $U \subset S$ of x^0 such that for every nonnegative density $\rho \in C_0^\infty(U)$, the Fourier restriction estimate (1.1) *holds true for every $p \geq 1$ such that $p' \geq p'_c := 2h^r(\phi) + 2$.*

The main body of this monograph will be devoted to the proof of this theorem. A question that remains open at this stage is whether Condition (R) is really needed in this theorem, or whether it can be removed completely, respectively replaced by a weaker condition.

Before returning to the proof, we shall show by means of Knapp type examples that the conditions in this theorem are sharp in the following sense:

Theorem 1.15. *Let ϕ be smooth of finite type, and assume that the Fourier restriction estimate* (1.1) *holds true in a neighborhood of x^0. Then, if $\rho(x^0) \neq 0$, necessarily $p' \geq p'_c$.*

Since the proof will also help to illuminate the notion of r-height, we shall give it right away. In fact, we shall prove the following more general result (notice that we are making no assumption on adaptedness of ϕ here).

PROPOSITION 1.16. *Assume that the coordinates* $x = (x_1, x_2)$ *are linearly adapted to* ϕ *and that the restriction estimate* (1.1) *holds true in a neighborhood of* $x^0 = 0$, *where* $\rho(x^0) \neq 0$. *Consider any fractional shear, say on* H^+, *given by*

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

where f *is real valued and fractionally smooth but not flat. Let* $\phi^f(y) = \phi(y_1, y_2 + f(y_1))$ *be the function expressing* ϕ *in the coordinates* $y = (y_1, y_2)$. *Then, necessarily,*

$$p' \geq 2h^f(\phi) + 2.$$

Theorem 1.15 will follow by choosing for f the principal root jet ψ.

Proof. The proof will be based on suitable Knapp-type arguments.

Let us use the same notation for the Newton polyhedron of ϕ^f as we did for ϕ^a in Section 1.3, that is, the vertices of the Newton polyhedron $\mathcal{N}(\phi^f)$ will be denoted by $(A_l, B_l),\ l = 0, \ldots, n$, where we assume that they are ordered so that $A_{l-1} < A_l,\ l = 1, \ldots, n$, with associated compact edges $\gamma_l := [(A_{l-1}, B_{l-1}), (A_l, B_l)]$, $l = 1, \ldots, n$, contained in the supporting lines L_l to $\mathcal{N}(\phi^f)$ and associated with the weights κ^l. The unbounded horizontal edge with left endpoint (A_n, B_n) will be denoted by γ_{n+1}. For $l = n+1$, we have $\kappa_1^{n+1} := 0$, $\kappa_2^{n+1} = 1/B_n$. Again, we put $a_l := \kappa_2^l/\kappa_1^l$, and $a_{n+1} := \infty$.

Because of (1.20), we have to prove the following estimates:

$$p' \geq 2d^f + 2; \tag{1.22}$$

$$p' \geq 2h_l^f + 2 \quad \text{for every } l \text{ such that } a_l > m_0, \tag{1.23}$$

where, according to (1.21),

$$h_l^f = \frac{1 + m_0\kappa_1^l - \kappa_2^l}{\kappa_1^l + \kappa_2^l}.$$

Consider first any nonhorizontal edge γ_l of $\mathcal{N}(\phi^f)$ with $a_l > m_0$, and denote by D_ε^f the region

$$D_\varepsilon^f := \{y \in \mathbb{R}^2 : |y_1| \leq \varepsilon^{\kappa_1^l}, |y_2| \leq \varepsilon^{\kappa_2^l}\}, \quad \varepsilon > 0,$$

in the coordinates y. In the original coordinates x, it corresponds to

$$D_\varepsilon := \{x \in \mathbb{R}^2 : |x_1| \leq \varepsilon^{\kappa_1^l}, |x_2 - f(x_1)| \leq \varepsilon^{\kappa_2^l}\}.$$

Assume that ε is sufficiently small. Since

$$\phi^f(\varepsilon^{\kappa_1^l} y_1, \varepsilon^{\kappa_2^l} y_2) = \varepsilon(\phi_{\kappa^l}^f(y_1, y_2) + O(\varepsilon^\delta)),$$

for some $\delta > 0$, where $\phi_{\kappa^l}^f$ denotes the κ^l-principal part of ϕ^f, we have that $|\phi^f(y)| \leq C\varepsilon$ for every $y \in D_\varepsilon^f$, that is,

$$|\phi(x)| \leq C\varepsilon \quad \text{for every } x \in D_\varepsilon. \tag{1.24}$$

Moreover, for $x \in D_\varepsilon$, because $|f(x_1)| \lesssim |x_1|^{m_0}$ and $m_0 \leq a_l = \kappa_2^l/\kappa_1^l$, we have

$$|x_2| \leq \varepsilon^{\kappa_2^l} + |f(x_1)| \lesssim \varepsilon^{\kappa_2^l} + \varepsilon^{m_0\kappa_1^l} \lesssim \varepsilon^{m_0\kappa_1^l}.$$

We may thus assume that D_ε is contained in the box where $|x_1| \le 2\varepsilon^{\kappa_1^l}$, $|x_2| \le 2\varepsilon^{m_0\kappa_1^l}$. Choose a Schwartz function φ_ε such that

$$\widehat{\varphi_\varepsilon}(x_1, x_2, x_3) = \chi_0\Big(\frac{x_1}{\varepsilon^{\kappa_1^l}}\Big)\chi_0\Big(\frac{x_2}{\varepsilon^{m_0\kappa_1^l}}\Big)\chi_0\Big(\frac{x_3}{C\varepsilon}\Big),$$

where χ_0 is a smooth cutoff function supported in $[-2, 2]$ and identically 1 on $[-1, 1]$.

Then by (1.24) we see that $\widehat{\varphi_\varepsilon}(x_1, x_2, \phi(x_1, x_2)) \ge 1$ on D_ε; hence, if $\rho(0) \neq 0$, then

$$\Big(\int_S |\widehat{\varphi_\varepsilon}|^2\, \rho\, d\sigma\Big)^{1/2} \ge C_1 |D_\varepsilon|^{1/2} = C_1 \varepsilon^{(\kappa_1^l + \kappa_2^l)/2},$$

where $C_1 > 0$ is a positive constant. Since $\|\varphi_\varepsilon\|_p \simeq \varepsilon^{((1+m_0)\kappa_1^l+1)/p'}$, we find that the restriction estimate (1.1) can hold only if

$$p' \ge 2\frac{(1+m_0)\kappa_1^l + 1}{\kappa_1^l + \kappa_2^l} = 2h_l^f + 2.$$

The case $l = n+1$, where γ_l is the horizontal edge for which $h_l^f = B_n - 1$ (with $B_n = 1/\kappa_l^2$), requires a minor modification of this argument. Observe that, by Taylor expansion, in this case ϕ^f can be written as

$$\phi^f(y_1, y_2) = y_2^{B_n} h(y_1, y_2) + \sum_{j=0}^{B_n - 1} y_2^j g_j(y_1), \tag{1.25}$$

where the functions g_j are flat and h is fractionally smooth and continuous at the origin. Choose a small $\delta > 0$, and define

$$D_\varepsilon^f := \{y \in \mathbb{R}^2 : |y_1| \le \varepsilon^\delta, |y_2| \le \varepsilon^{\kappa_2^l}\}, \qquad \varepsilon > 0.$$

Then (1.25) shows that again $|\phi^f(y)| \le C\varepsilon$ for every $y \in D_\varepsilon^f$, so that (1.24) holds true again. Moreover, for $x \in D_\varepsilon$, we now find that

$$|x_2| \le \varepsilon^{\kappa_2^l} + |f(x_1)| \lesssim \varepsilon^{\kappa_2^l} + \varepsilon^{m_0\delta} \lesssim \varepsilon^{m_0\delta}$$

for δ sufficiently small. Choosing

$$\widehat{\varphi_\varepsilon}(x_1, x_2, x_3) = \chi_0\Big(\frac{x_1}{\varepsilon^{\delta}}\Big)\chi_0\Big(\frac{x_2}{\varepsilon^{m_0\delta}}\Big)\chi_0\Big(\frac{x_3}{C\varepsilon}\Big),$$

and arguing as before, we find that here (1.1) implies that

$$p' \ge 2\frac{(1+m_0)\delta + 1}{\delta + \kappa_2^l} \quad \text{for every } \delta > 0;$$

hence, $p' \ge 2B_n = 2h_{l_{n+1}}^f + 2$. This finishes the proof of (1.23).

Notice finally that the argument for the nonhorizontal edges still works if we replace the line L_l by the line L^f and the weight κ^l by the weight κ^f associated with that line. Since here $m_0\kappa_1^f = \kappa_2^f$, this leads to condition (1.22). Q.E.D.

1.5 AN INVARIANT DESCRIPTION OF THE NOTION OF *r*-HEIGHT

Finally, we can also give a more invariant description of the notion of r-height, which conceptually resembles more closely Varchenko's definition of the notion of height, only that we restrict the admissible changes of coordinates to the class of fractional shears in the half planes H^+ and H^-.

In this section, we do not work under Assumption (NLA), but we may and shall again assume that our initial coordinates (x_1, x_2) are linearly adapted to ϕ. Then we set

$$\tilde{h}^r(\phi) := \sup_f h^f(\phi),$$

where the supremum is taken over all nonflat fractionally smooth, real functions $f(x_1)$ of $x_1 > 0$ (corresponding to a fractional shear in H^+) or of $x_1 < 0$ (corresponding to a fractional shear in H^-). Then, obviously,

$$h^r(\phi) \leq \tilde{h}^r(\phi),$$

but in fact there is equality.

Proposition 1.17. *Assume that the coordinates (x_1, x_2) are linearly adapted to ϕ, where ϕ is smooth and of finite type and satisfies $\phi(0,0) = 0$, $\nabla\phi(0,0) = 0$.*

(a) *If the coordinates (x_1, x_2) are not adapted to ϕ, then for every nonflat fractionally smooth, real function $f(x_1)$ and the corresponding fractional shear in H^+ (respectively, H^-), we have $h^f(\phi) \leq h^r(\phi)$. Consequently, $h^r(\phi) = \tilde{h}^r(\phi)$.*

(b) *If the coordinates (x_1, x_2) are adapted to ϕ, then $\tilde{h}^r(\phi) = d(\phi) = h(\phi)$.*

In particular, the critical exponent for the restriction estimate (1.1) *is in all cases given by $p'_c := 2\tilde{h}^r(\phi) + 2$.*

Let us content ourselves at this stage with a short, but admittedly indirect, proof of part (b) and of part (a) under the assumption that ϕ is analytic. Our arguments will again be based on Proposition 1.16. Since these arguments will rely on the validity of Theorems 1.5 and 1.10, which is somewhat unsatisfactory, we shall give a direct, but lengthier, proof in Chapter 9, which will in addition not require analyticity of ϕ.

On the proof of Proposition 1.17. Recall that we assume that the original coordinates (x_1, x_2) are linearly adapted to ϕ.

In order to prove (a) for analytic ϕ, assume furthermore that the coordinates (x_1, x_2) are not adapted to ϕ, and let $f(x_1)$ be any nonflat fractionally smooth, real-valued function of x_1, with corresponding fractional shear, say in H^+. We have to show that

$$h^f(\phi) \leq h^r(\phi). \tag{1.26}$$

According to Theorem 1.10 the restriction estimate (1.1) holds true for $p = p_c$, where $p'_c = 2h^r(\phi) + 2$. Moreover, choosing ρ so that $\rho(x^0) \neq 0$, then

Proposition 1.16 implies that $p' \geq 2h^f(\phi) + 2$. Combining these estimates, we obtain (1.26).

In order to prove (b), we assume that the coordinates (x_1, x_2) are adapted to ϕ, so that $d(\phi) = h(\phi)$. We have to prove that

$$\tilde{h}^r(\phi) = d(\phi). \tag{1.27}$$

Let us first observe that Theorem 1.5 and Proposition 1.16 imply, in a similar way as in the proof of (a), that $2h(\phi) + 2 \geq 2h^f(\phi) + 2$; hence $d(\phi) \geq h^f(\phi)$. We thus see that

$$\tilde{h}^r(\phi) \leq d(\phi).$$

On the other hand, when the principal face $\pi(\phi)$ is compact, then we can choose a supporting line

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

to the Newton polyhedron of ϕ containing $\pi(\phi)$ and such that $0 < \kappa_1 \leq \kappa_2$. We then put $f(x_1) := x_1^{m_0}$, where $m_0 := \kappa_2/\kappa_1$. Then $d(\phi) = 1/(\kappa_1 + \kappa_2) = d^f \leq h^f(\phi) \leq \tilde{h}^r(\phi)$, and we obtain (1.27).

Assume finally that $\pi(\phi)$ is an unbounded horizontal half line, with left endpoint (A, B), where $A < B$. We then choose $f_n(x_1) := x_1^n$, $n \in \mathbb{N}$. Then it is easy to see that for n sufficiently large, the line L^{f_n} will pass through the point (A, B), and thus $\lim_{n\to\infty} h^{f_n}(\phi) = B = d(\phi)$. Therefore, $\tilde{h}^r(\phi) \geq d(\phi)$, which shows that (1.27) is also valid in this case. Q.E.D.

1.6 ORGANIZATION OF THE MONOGRAPH AND STRATEGY OF PROOF

In Chapter 2 we shall begin to prepare the proof of Theorem 1.14 by compiling various auxiliary results. This will include variants of van der Corput type estimates for one-dimensional oscillatory integrals and related sublevel estimates through "integrals of sublevel type," which will be used all over the place. In some situations, we shall also need more specific information, in particular on Airy-type oscillatory integrals, as well as on some special classes of integrals of sublevel type, which will also be provided. We shall also derive a straightforward variant of a beautiful real interpolation method that has been devised by Bak and Seeger in [BS11] and that will allow us in some cases to replace the more classical complex interpolation methods in the proof of Stein-Tomas-type Fourier restriction estimates by substantially shorter arguments. Nevertheless, complex interpolation methods will play a major role in many other situations, and a crucial tool in our application of Stein's interpolation theorem for analytic families of operators will be provided by certain uniform estimates for oscillatory sums, respectively, double sums (cf. Lemmas 2.7 and 2.9). Last, we shall derive normal forms for phase functions ϕ of linear height < 2 for which no linear coordinate system adapted to ϕ does exist. These normal forms will provide the basis for our discussion of the case $h_{\text{lin}}(\phi) < 2$ in Chapter 4.

As a first step in our proof of Theorem 1.14, in which we shall always assume that the coordinates x are linearly adapted, but not adapted to ϕ, we shall show in Chapter 3 that one may reduce the desired Fourier restriction estimate to a piece S_ψ of the surface S lying above a small, "horn-shaped" neighborhood D_ψ of the principal root jet ψ, on which $|x_2 - \psi(x_1)| \le \varepsilon x_1^m$. Here, $\varepsilon > 0$ can be chosen as small as we wish. This step works in all cases, no matter which value $h_{\text{lin}}(\phi)$ takes. The proof will give us the opportunity to introduce some of the basic tools which will be applied frequently, such as dyadic domain decompositions, rescaling arguments based on the dilations associated to a given edge of the Newton polyhedron (in this chapter given by the principal face $\pi(\phi)$ of the Newton polyhedron of ϕ), in combination with Greenleaf's restriction Theorem 1.1 and Littlewood-Paley theory, which will allow us to sum the estimates that we obtain for the dyadic pieces.

From here on, following our approach in [IKM10] and [IM11b], it will be natural to distinguish between the cases where $h_{\text{lin}}(\phi) < 2$ and where $h_{\text{lin}}(\phi) \ge 2$, since, in contrast to the first case, in the latter case a reduction to estimates for one-dimensional oscillatory integrals will be possible in many situations, which in return can be performed by means of the van der Corput–type Lemma 2.1.

Chapters 4 and 5 will be devoted to the case where $h_{\text{lin}}(\phi) < 2$. The starting point for our discussion will be the normal forms for ϕ provided by Proposition 2.11. Some of the main tools will again consist of various kinds of dyadic domain decompositions in combination with Littlewood-Paley theory and rescaling arguments. In addition, we shall have to localize frequencies to dyadic intervals in each component, which then also leads us to distinguish a variety of different cases, depending on the relative sizes of these dyadic intervals as well as of another parameter related to the Littlewood-Paley decomposition. It turns out that the particular case where $m = 2$ in (1.9) and (1.14) will require, in some situations (these are listed in Proposition 4.2), a substantially more refined analysis than the case $m \ge 3$. Indeed, in some cases, namely, those described in Proposition 4.2(a) and (b), our arguments from Chapter 4 will almost give the complete answer, except that we miss the endpoint $p = p_c$. In order to capture also the corresponding endpoint estimates, we shall devise rather intricate complex interpolation arguments in Section 5.3. These will be prototypical for many more arguments of this type that we shall devise in later chapters.

Even more of a problem will be presented by the cases described in Proposition 4.2(c). In these situations, we not only miss the endpoint estimate in our discussion in Chapter 4, but it turns out that we even have to close a large gap in the L^p range that we need to cover. In order to overcome this problem, we shall perform a further dyadic decomposition in frequency space with respect to the distance to a certain "Airy cone." This refined Airy-type analysis will be developed in Sections 5.1 and 5.2. Again, in order to capture also the endpoint $p = p_c$, we need to apply a complex interpolation argument. Useful tools in these complex interpolation arguments will be Lemmas 2.7 and 2.9 on oscillatory sums and double sums.

In Chapter 6 we shall turn to the case where $h_{\text{lin}}(\phi) \ge 2$. In a first step, following some ideas from the article [PS97] by Phong and Stein (compare also [IKM10] and [IM11b]), we shall perform a decomposition of the remaining piece S_ψ of the surface S, which will be adapted in some sense to the "root structure" of the function ϕ within the domain D_ψ. When speaking of roots, we will, in fact, always have the case of analytic ϕ in mind; for nonanalytic ϕ, these statements may no longer make strict sense, but the ideas from the analytic case may still serve as a very useful guideline.

In order to understand the root structure within our narrow horn-shaped neighborhood D_ψ of the principal root jet ψ, it is natural to look at the Newton polyhedron $\mathcal{N}(\phi^a)$ of ϕ when expressed in the adapted coordinates (y_1, y_2) given by (1.10), (1.11). More precisely, we shall associate to every edge γ_l of $\mathcal{N}(\phi^a)$ lying above the bisectrix a domain D_l, which will be homogeneous in the adapted coordinates (y_1, y_2) under the natural dilations $(y_1, y_2) \mapsto (r^{\kappa_1^l} y_1, r^{\kappa_2^l} y_2)$, $r > 0$, defined by the weight κ^l that we had associated to the edge γ_l. We shall then partition the domain D_ψ into these domains D_l, intermediate domains E_l, and a residue domain $D_{\rm pr}$ and consider the corresponding decomposition of the surface S. The remaining domain $D_{\rm pr}$, which contains the principal root jet $x_2 = \psi(x_1)$, will in some sense be associated with the principal face $\pi(\phi^a)$ of the Newton polyhedron of ϕ^a and, hence, homogeneous in the coordinates (y_1, y_2). Each domain E_l can be viewed as a "transition" domain between two different types of homogeneity (in adapted coordinates).

In the domains D_l we can again apply our dyadic decomposition techniques in combination with rescaling arguments, making use of the dilations associated with the weight κ^l, but serious new problems do arise, caused by the nonlinear change from the coordinates (x_1, x_2) to the adapted coordinates (y_1, y_2). Following again [PS97], the discussion of the transition domains E_l will be based on bidyadic domain decompositions in the coordinates (y_1, y_2).

In our discussion of the domains D_l, we shall have to distinguish three cases, Cases 1–3, depending on the behavior of the κ^l-principal part $\phi^a_{\kappa^l}$ of ϕ^a near a given point $v \neq 0$. The first case will be easy to handle, and the same is true even for the residue domain $D_{\rm pr}$. However, in the other two cases, a large difference between the domains D_l and the domain $D_{\rm pr}$ will appear. The reason for this is that for the edges γ_l lying above the bisectrix, we shall be able to prove a favorable control on the multiplicities of the roots of $\phi^a_{\kappa^l}$ (more precisely of $\partial_2\phi^a_{\kappa^l}$), but this breaks down on $D_{\rm pr}$.

We shall, therefore, start to have a closer look at the domain $D_{\rm pr}$ in Section 6.4. In order to handle Case 3, which is the case where $\phi^a_{\kappa^{l_{\rm pr}}}$ has a critical point at v, in Section 6.5 we shall devise a further decomposition of the domain $D_{\rm pr}$ into various subdomains of "type" $D_{(l)}$ and $E_{(l)}$, where each domain $D_{(l)}$ will be homogeneous in suitable "modified adapted" coordinates, and the domains $E_{(l)}$ can again be viewed as "transition domains." This domain decomposition algorithm, roughly speaking, reflects the "fine splitting" of roots of $\partial_2\phi^a$. The new transition domains $E_{(l)}$ can be treated in a similar way as the domains E_l before, and in the end we shall be left with domains of type $D_{(l)}$. Now, under the assumption that $h_{\rm lin}(\phi) \geq 5$, it turns out that these remaining domains can be handled by means of a fibration of the given piece of surface into a family of curves, in combination with Drury's Fourier restriction theorem for curves with nonvanishing torsion [Dru85]. However, that method breaks down when $h_{\rm lin}(\phi) < 5$, so that in the subsequent two chapters we shall devise an alternative approach for dealing with these remaining domains $D_{(l)}$. That approach will work equally well whenever $h_{\text{lin}}(\phi) \geq 2$.

In Chapter 7 we shall mostly consider the domains of type $D_{(1)}$, which are in some sense "closest" to the principal root jet, since it will turn out that the other domains $D_{(l)}$ with $l \geq 2$ are easier to handle (compare Section 7.10). Within the domains of type $D_{(1)}$, we shall have to deal with functions ϕ which, in suitable "modified" adapted coordinates, look like $\tilde{\phi}^a(y_1, y_2) = y_2^B\, b_B(y_1, y_2) + y_1^n \alpha(y_1) + \sum_{j=1}^{B-1} y_2^j\, b_j(y_1)$, where $B \geq 2$ and b_B is nonvanishing. In a first step, by means of some lower bounds on the r-height, we shall be able to establish favorable restriction estimates in most situations, with the exception of certain cases where $m = 2$ and $B = 3$ or $B = 4$. Along the way, in some cases we shall have to apply interpolation arguments in order to capture the endpoint estimates for $p = p_c$. Sometimes this can be achieved by means of the aforementioned variant of the Fourier restriction theorem by Bak and Seeger, whose assumptions, when satisfied, are easily checked. However, in most of these cases we shall have to apply complex interpolation, in a similar way as we did before in Section 5.3.

Eventually we can thus reduce considerations to certain cases where $m = 2$ and $B = 3$ or $B = 4$. The most difficult situations will occur when frequencies $\xi = (\xi_1, \xi_2, \xi_3)$ are localized to domains on which all components ξ_i of ξ are comparable in size. These cases, which will be discussed in Chapter 8, turn out to be the most challenging ones among all, the case $B = 3$ being the worst. Again, we shall have to apply a refined Airy-type analysis in combination with complex interpolation arguments as in Chapter 5, but further methods are needed— for instance, rescaling arguments going back to Duistermaat [Dui74]— in order to control the dependence of certain classes of oscillatory integrals on some small parameters, and the technical complexity will become quite demanding.

The monograph will conclude with the proof of Proposition 1.7 on the characterization of linearly adapted coordinates in Appendix A and a direct proof of Proposition 1.17 on an invariant description of the notion of r-height in Appendix B of Chapter 9.

Conventions: Throughout this monograph, we shall use the "variable constant" notation, that is, many constants appearing in the course of our arguments, often denoted by C, will typically have different values on different lines. Moreover, we shall use symbols such as $\sim$, $\lesssim$, or $\ll$ in order to avoid writing down constants. By $A \sim B$ we mean that there are constants $0 < C_1 \leq C_2$ such that $C_1 A \leq B \leq C_2 A$, and these constants will not depend on the relevant parameters arising in the context in which the quantities A and B appear. Similarly, by $A \lesssim B$ we mean that there is a (possibly large) constant $C_1 > 0$ such that $A \leq C_1 B$, and by $A \ll B$ we mean that there is a sufficiently small constant $c_1 > 0$ such that $A \leq c_1 B$, and again these constants do not depend on the relevant parameters.

By χ_0 and χ_1 we shall always denote smooth cutoff functions with compact support on $\mathbb{R}^n$, where χ_0 will be supported in a neighborhood of the origin and usually be identically 1 near the origin, whereas $\chi_1 = \chi_1(x)$ will be support away from the origin, sometimes in each of its coordinates x_j, that is, $|x_j| \sim 1$ for $j = 1, \dots, n$, for every x in the support of χ_1. These cutoff functions may also vary from line to line and may, in some instances, where several of such functions of different variables appear within the same formula, even designate different functions.

Also, if we speak of the *slope* of a line such as a supporting line to a Newton polyhedron, then we shall actually mean the modulus of the slope. Finally, when speaking of domain decompositions, we shall not always keep to the mathematical convention of a domain being on open connected set but shall occasionally use the word *domain* in a more colloquial way.

Finally, by $\mathbb{N}^\times$, $\mathbb{Q}^\times$, $\mathbb{R}^\times$, and so on, we shall denote the set of nonzero elements in $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, and so on.

ACKNOWLEDGMENTS

We thank the referees for numerous suggestions which greatly helped to improve the exposition. Moreover, we are indepted to Eugen Zimmermann for providing some of the graphics for this monograph. Finally, we gratefully acknowledge the support for this work by the Deutsche Forschungsgemeinschaft (DFG). We also wish to acknowledge permission to reprint excerpts from the following previously published materials:

"Uniform Estimates for the Fourier Transform of Surface Carried Measures in $\mathbb{R}^3$ and an Application to Fourier Restriction" by Isroil A. Ikromov and Detlef Müller from *Journal of Fourier Analysis and Applications*, December 2011, Volume 17, Issue 6, pp. 1292–1332 (16 July 2011) © Springer Science+ Business Media, LLC 2011. Published with permission of Springer Science+ Business Media.

"On Adapted Coordinate Systems" by Isroil A. Ikromov and Detlef Müller. First published in *Transactions of the American Mathematical Society* in 363 (2011), 2821–2848, published by the American Mathematical Society. Reprinted by permission of the American Mathematical Society.

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

Advances in Analysis: The Legacy of Elias M. Stein edited by Charles Fefferman, Alexandru Ionescu, D.H. Phong, and Stephen Wainger, Princeton University Press, 2014. Reprinted by permission of Princeton University Press and Stephen Wainger, Charles Fefferman, D.H. Phong, and Alexandru Ionescu.