

ON ADAPTED COORDINATE SYSTEMS

ISROIL A. IKROMOV AND DETLEF MÜLLER

ABSTRACT. The notion of an adapted coordinate system for a given real-analytic function, introduced by V. I. Arnol'd, plays an important role, for instance, in the study of asymptotic expansions of oscillatory integrals. In two dimensions, A. N. Varchenko gave sufficient conditions for the adaptiveness of a given coordinate system and proved the existence of an adapted coordinate system for analytic functions without multiple components. Varchenko's proof is based on a two-dimensional resolution of singularities result.

In this article, we present a more elementary approach to these results, which is based on the Puiseux series expansion of roots of the given function. This approach is inspired by the work of D. H. Phong and E. M. Stein on the Newton polyhedron and oscillatory integral operators. It applies to arbitrary real-analytic functions, and even to arbitrary smooth functions of finite type. In particular, we show that Varchenko's conditions are in fact necessary and sufficient for the adaptiveness of a given coordinate system and that adapted coordinates always exist in two dimensions, even in the smooth, finite type setting. For analytic functions, a construction of adapted coordinates by means of Puiseux series expansions of roots has already been carried out in work by D. H. Phong, E. M. Stein and J. A. Sturm on the growth and stability of real-analytic function, as we learned after the completion of this paper. In contrast to their work, however, our proof more closely follows Varchenko's algorithm for the construction of an adapted coordinate system, which turns out to be useful for the extension to the smooth setting.

CONTENTS

1. Introduction	2822
2. Preliminaries	2823
2.1. The principal part of f associated to a supporting line of the Newton polyhedron as a mixed homogeneous polynomial	2824
3. Conditions for adaptiveness of a given coordinate system	2828
3.1. On the effect of a change of coordinates on the Newton diagram	2828
4. The real-analytic case	2835
4.1. Description of the Newton polyhedron in terms of the roots	2835
4.2. Existence of adapted coordinates in the analytic setting	2838
5. The smooth case	2845
Acknowledgements	2848
References	2848

Received by the editors April 8, 2008 and, in revised form, September 23, 2008.

2000 *Mathematics Subject Classification*. Primary 35D05, 35D10, 35G05.

Key words and phrases. Oscillatory integral, Newton diagram.

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

1. INTRODUCTION

It is an obvious fact that the asymptotic behavior of an oscillatory integral of the form

$$I(\lambda) := \int_{\mathbb{R}^n} e^{i\lambda f(x)} a(x) dx$$

does not change under a smooth change of variables $x = \varphi(y)$.

This observation is already employed in the proof of van der Corput's lemma (see, e.g., [7]), according to which the asymptotic behavior of a one-dimensional oscillatory integral is determined by the maximal order of vanishing at the critical points of the phase function f .

In higher dimensions, the problem of determining the exact asymptotic behavior of an oscillatory integral is substantially more difficult. V. I. Arnol'd conjectured in [1] that the asymptotic behavior of the oscillatory integral $I(\lambda)$ is determined by the Newton polyhedron associated to the phase function f in a so-called "adapted" coordinate system. For some special cases this conjecture was then indeed verified by means of Arnol'd's classification of singularities (see [2]). Later, however, A. N. Varchenko [8] disproved Arnol'd's conjecture in dimensions three and higher.

Moreover, in the same paper he was able to verify Arnol'd's conjecture for two-dimensional oscillatory integrals. In particular, in two dimensions, Varchenko gave sufficient conditions for the adaptiveness of a given coordinate system, proved the existence of an adapted coordinate system for analytic functions (without multiple components), and showed that the leading term of the asymptotic expansion of $I(\lambda)$ can be constructed from the Newton polyhedron associated to the phase function f in such an adapted coordinate system. These two-dimensional results were based on a strong result on the resolution of singularities for this case. Note also that in dimensions higher than two adapted coordinate systems may not exist, as has also been shown by Varchenko [8].

The purpose of this article is to present a more elementary and concrete approach to the latter results in two dimensions, which is based on the Puiseux series expansion of roots of the given function f but does not require any stronger two-dimensional resolution result. This approach was inspired by the work of D. H. Phong and E. M. Stein in their seminal paper [5]. Our methods apply to arbitrary real-analytic functions and even extend to arbitrary smooth functions of finite type. Moreover, we show that Varchenko's conditions are in fact necessary and sufficient for the adaptiveness of a given coordinate system, even in the smooth, finite type setting.

After completion of this paper, we have learned that a construction of adapted coordinates by means of Puiseux series expansions of roots has already been carried out in the work by D. H. Phong, E. M. Stein and J. A. Sturm [6] (compare Theorems 4 and 5, and the remark at the bottom of page 539). In contrast to [6], however, our proof more closely follows Varchenko's algorithm for the construction of an adapted coordinate system, which turns out to be useful for the extension to the smooth setting.

The results and techniques in this paper are basic to the subsequent article [4] in which we study estimates for maximal functions, oscillatory integrals and restriction theorems associated to hypersurfaces in \mathbb{R}^3 .

2. PRELIMINARIES

Let f be a smooth real-valued function defined on a neighborhood of the origin in \mathbb{R}^2 with $f(0,0) = 0$, $\nabla f(0,0) = 0$, and consider the associated Taylor series

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

of f centered at the origin. The set

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

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

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

i.e., that the function f is of finite type at the origin. If f is real-analytic, so that the Taylor series converges to f near the origin, this just means that $f \neq 0$. The *Newton polyhedron* $\mathcal{N}(f)$ of f 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 $(j, k) \in \mathcal{T}(f)$. The associated *Newton diagram* $\mathcal{N}_d(f)$ in the sense of Varchenko [8] 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*, or shorter *distance* $d = d(f)$ 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* $\pi(f)$ of the Newton polyhedron of f is the face of minimal dimension containing the point (d, d) . Deviating from the notation in [8], we shall call the series

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

the *principal part* of f . In the case where $\pi(f)$ is compact, f_p is a mixed homogeneous polynomial; otherwise, we shall consider f_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 f 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 f , we shall always consider smooth coordinate systems, and if f 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 f is defined by

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

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 f if $h(f) = d_x$.

A. N. Varchenko [8] proved that if f is a real-analytic function (without multiple components) near the origin in \mathbb{R}^2 , then there exists a local analytic coordinate

system which is adapted to f , and he mentioned that the condition about multiple components can be dropped.

It may be interesting at this point to remark that, of course, the notions introduced above extend to smooth functions in more than two real variables. However, as shown by Varchenko [8], in dimensions higher than two, adapted coordinate systems may not exist, even in the analytic setting.

2.1. The principal part of f 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 an 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 f on \mathbb{R}^2 is said to be κ -homogeneous of degree a if $f(\delta_r x) = r^a f(x)$ 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 κ -degree of f . For instance, the monomial $x_1^j x_2^k$ has κ -degree $\kappa_1 j + \kappa_2 k$.

If f 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 $\mathcal{N}(f)$ of f . Then the non-trivial polynomial

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

is κ -homogeneous of degree a ; it will be called the κ -principal part of f . By definition, we then have

$$(2.1) \quad f(x_1, x_2) = f_\kappa(x_1, x_2) + \text{terms of higher } \kappa\text{-degree.}$$

More precisely, we mean by this that every point (j, k) in the Taylor support of the remainder term $f - f_\kappa$ lies on a line $\kappa_1 t_1 + \kappa_2 t_2 = d$ with $d > a$ parallel to, but above the line L_κ ; i.e., we have $\kappa_1 j + \kappa_2 k > a$. Moreover, clearly

$$\mathcal{N}_d(f_\kappa) \subset \mathcal{N}_d(f).$$

The following lemma gives an equivalent description of the notion “terms of higher κ -degree”, which is quite useful in applications, since it may be used to essentially reduce many considerations to the case of polynomial functions. It will mostly be applied without further mentioning.

Lemma 2.1. *Assume that f is a smooth function defined near the origin, and let $c \geq 0$. Assume that $\kappa_1 \leq \kappa_2$, and choose $m \geq 1$ in \mathbb{N} such that $\kappa_1 m > c$. Then f consists of terms of κ -degree greater than or equal to c in the above sense (i.e., $\kappa_1 j + \kappa_2 k \geq c$ for every $(j, k) \in \mathcal{T}(f)$) if and only if there exists a polynomial function F with $\mathcal{T}(F) \subset \{(j, k) \in \mathbb{N} : \kappa_1 j + \kappa_2 k \geq c\}$ and smooth functions a_{jk} , for $j + k = m$, such that*

$$(2.2) \quad f(x_1, x_2) = F(x_1, x_2) + \sum_{j+k=m} x_1^j x_2^k a_{jk}(x_1, x_2).$$

Notice that $\kappa_1 j + \kappa_2 k > c$ whenever $j + k = m$.

Proof. Assume that $\kappa_1 j + \kappa_2 k \geq c$ for every $(j, k) \in \mathcal{T}(f)$. If we then choose for F the Taylor polynomial of degree $m - 1$ of f at the origin, the representation (2.2) follows from Taylor’s formula.

Conversely, it is obvious that the representation (2.2) implies that $\kappa_1 j + \kappa_2 k \geq c$ for every $(j, k) \in \mathcal{T}(f)$. □

Sometimes it will be convenient to extend these definitions to the case where $\kappa_1 = 0$ or $\kappa_2 = 0$. In that case, the κ -principal part $\sum_{(j,k) \in L_\kappa} c_{jk} x_1^j x_2^k$ of f will just be considered as a formal power series, unless f is real-analytic, where it is real-analytic, too.

Let $P \in \mathbb{R}[x_1, x_2]$ be a mixed homogeneous polynomial, and assume that $\nabla P(0, 0) = 0$. Following [4], 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.

The following proposition will be a useful tool.

If m_1, \dots, m_n are positive integers, then we denote by (m_1, \dots, m_n) their greatest common divisor.

Proposition 2.2. *Let P be a (κ_1, κ_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 κ_1 and κ_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

$$\kappa_1 = \frac{q}{m}, \kappa_2 = \frac{p}{m}, \quad (p, q, m) = 1,$$

so that in particular $p \geq q$. Then $(p, q) = 1$, and there exist non-negative integers α_1, α_2 and a $(1, 1)$ -homogeneous polynomial Q such that the polynomial P can be written as

$$(2.3) \quad P(x_1, x_2) = x_1^{\alpha_1} x_2^{\alpha_2} Q(x_1^p, x_2^q).$$

More precisely, P can be written in the form

$$(2.4) \quad P(x_1, x_2) = cx_1^{\nu_1} x_2^{\nu_2} \prod_{l=1}^M (x_2^q - \lambda_l x_1^p)^{n_l},$$

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 α_1, α_2 in (2.3)).

Let us put $n := \sum_{l=1}^M n_l$. The distance $d(P)$ of P can then be read from (2.4) as follows:

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

$$(2.5) \quad d(P) = \frac{1}{\kappa_1 + \kappa_2} = \frac{\nu_1 q + \nu_2 p + pqn}{q + p}.$$

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

Proof. The proof of Proposition 2.2 is based on elementary number theoretic arguments. Denote by A the set of all solutions $(\alpha, \beta) \in \mathbb{N}^2$ of the linear equation

$$(2.6) \quad \alpha \kappa_1 + \beta \kappa_2 = 1.$$

Then

$$(2.7) \quad P(x_1, x_2) = \sum_{(\alpha, \beta) \in A} c_{\alpha, \beta} x_1^\alpha x_2^\beta$$

for suitable coefficients $c_{\alpha, \beta}$.

If P is not of the form $P(x_1, x_2) = cx_1^{\nu_1} x_2^{\nu_2}$, then the equation (2.6) has at least two different solutions $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{N}^2$ for which the coefficients $c_{\alpha, \beta}$

in (2.7) are non-vanishing. The corresponding equations $\alpha_1\kappa_1 + \beta_1\kappa_2 = 1$ and $\alpha_2\kappa_1 + \beta_2\kappa_2 = 1$ determine the numbers κ_1, κ_2 uniquely and show that they are rational.

Assume next that $\kappa_1 \leq \kappa_2$, and write

$$\kappa_1 = \frac{q}{m}, \kappa_2 = \frac{p}{m}, \quad (p, q, m) = 1.$$

Then choose $(\alpha_0, \beta_0) \in A$ with α_0 maximal. Since $p\mathbb{Z} + q\mathbb{Z} + m\mathbb{Z} = \mathbb{Z}$, the identity (2.6), which is equivalent to

$$\alpha q + \beta p = m,$$

implies that $p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}$ (since $A \neq \emptyset$), so that

$$(2.8) \quad (p, q) = 1.$$

Notice that $(\alpha, \beta) \in A$ is equivalent to $(\alpha - \alpha_0)q + (\beta - \beta_0)p = 0$. Then (2.8) implies that $\alpha - \alpha_0 = -sp$, $\beta - \beta_0 = sq$ for some $s \in \mathbb{Z}$, or, equivalently,

$$\alpha = \alpha_0 - sp, \beta = \beta_0 + sq.$$

Since α_0 is maximal for A , we have $s \geq 0$. Choose $s_1 \in \mathbb{Z}$ maximal with $\alpha_1 := \alpha_0 - s_1p \geq 0$. Then for any $(\alpha, \beta) \in A$ we have the relation

$$\alpha = \alpha_1 + (s_1 - s)p, \quad \beta = \beta_0 + sq.$$

Notice that $s, s_1 - s \in \mathbb{N}$. So, every monomial $x_1^\alpha x_2^\beta$ with $(\alpha, \beta) \in A$ can be written as

$$x_1^\alpha x_2^\beta = x_1^{\alpha_1} x_2^{\beta_0} (x_1^p)^{s_1 - s} (x_2^q)^s.$$

This in combination with (2.7) yields (2.3).

In order to prove (2.4), write

$$Q(y_1, y_2) = cy_2^n + c_1y_2^{n-1}y_1 + \dots + c_ny_1^n,$$

where n is the degree of Q . We may then assume that $c \neq 0$, for otherwise we can pull out some power of $y_1 = x_1^p$ from Q in (2.3). Assuming without loss of generality that $c = 1$, we can write

$$Q(y_1, y_2) = y_1^n Q\left(1, \frac{y_2}{y_1}\right) = y_1^n \prod_{j=1}^n \left(\frac{y_2}{y_1} - \lambda_j\right) = \prod_{j=1}^n (y_2 - \lambda_j y_1),$$

where $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are the roots of the polynomial $Q(1, y_2)$, listed with their multiplicities. This yields (2.4).

To compute the distance $d(P)$, observe that the vertices of the Newton polyhedron of P are given by $X_1 := (\nu_1, \nu_2 + nq)$ and $X_2 := (\nu_1 + np, \nu_2)$, which, by our assumptions, are different points on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. One then easily computes that

$$(2.9) \quad \kappa_1 = \frac{q}{\nu_1 q + \nu_2 p + pqn}, \quad \kappa_2 = \frac{p}{\nu_1 q + \nu_2 p + pqn}.$$

Thus, if the principal face of $\mathcal{N}(P)$ is compact, then it is the interval connecting these two points and hence it lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$, and we immediately obtain (2.5). Otherwise, if $\nu_1 \leq \nu_2$, then the principal face is the horizontal half-line with left endpoint X_1 , so that $d(P) = \nu_2$, and similarly $d(P) = \nu_1$, if $\nu_1 \geq \nu_2$.

From the geometry of $\mathcal{N}(P)$, it is then clear that we always have the identity $d(P) = \max\{\nu_1, \nu_2, \frac{1}{\kappa_1 + \kappa_2}\}$. □

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 $x_2 = \lambda_l^{1/q} x_1^{p/q}$. 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,

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

and (2.9) shows that

$$(2.10) \quad d_h(P) = \frac{\nu_1 q + \nu_2 p + pqn}{q + p}.$$

Notice also that

$$(2.11) \quad m(P) = \max\{\nu_1, \nu_2, \max_{l \in R} n_l\},$$

where the index set $R := \{l = 1, \dots, M : \lambda_l \in \mathbb{R}\}$ corresponds to the set of real roots of P which do not lie on a coordinate axis.

Corollary 2.3. *Let P be a (κ_1, κ_2) -homogeneous polynomial of degree one as in Proposition 2.2, and consider the representation (2.4) of P . We again put $n := \sum_{l=1}^M n_l$.*

- (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 $x_2 = \lambda_l^{1/q} x_1^{p/q}$ of P has multiplicity $n_l < d_h(P)$.*
- (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 $n_0 := \nu_1, n_{M+1} := \nu_2$, choose $l_0 \in \{0, \dots, M+1\}$ so that $n_{l_0} = \max_{l=0, \dots, M+1} n_l$, and assume that $n_{l_0} > d_h(P)$, then $n_l < d_h(P)$ for every $l \neq l_0$.*
- (c) *If $1 < \kappa_2/\kappa_1 \in \mathbb{N}$, then there exists at most one real root (x_1^0, x_2^0) of P on the unit circle S^1 with $x_1^0 \neq 0$, which has multiplicity greater than or equal to $d_h(P)$.*

Proof. If $q \geq 2$, then $p > q \geq 2$, so that $\frac{1}{p} + \frac{1}{q} < 1$. Hence by (2.10)

$$d_h(P) \geq \frac{pqn}{q + p} > n.$$

This proves (a).

To prove (b), suppose next that $q = 1$, and assume that $n_{l_1}, n_{l_2} \geq d_h(P)$ and $n_{l_1} + n_{l_2} > 2d_h(P)$, where $l_1 < l_2$. If $l_1 > 0$, then we arrive at the contradiction

$$d_h(P) \geq \frac{p(n_{l_1} + n_{l_2})}{p + 1} > \frac{2pd_h(P)}{p + 1} \geq d_h(P).$$

Similarly, if $l_1 = 0$, then we obtain the contradiction

$$d_h(P) \geq \frac{n_0 + pn_{l_2}}{p + 1} > \frac{d_h(P) + pd_h(P)}{p + 1} = d_h(P).$$

To prove (c), we argue similarly as in case (b), noticing that here we have $0 < l_1 < l_2$, and $p \geq 2$. Then

$$d_h(P) \geq \frac{p(n_{l_1} + n_{l_2})}{p + 1} \geq \frac{2pd_h(P)}{p + 1} > d_h(P).$$

□

The corollary shows in particular 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 $x_2 = \lambda_{l_0} x_1^p$ with multiplicity exceeding $d(P)$. If such a root exists, we shall call it the *principal root of P* .

3. CONDITIONS FOR ADAPTEDNESS OF A GIVEN COORDINATE SYSTEM

In [8], Varchenko has provided various sufficient conditions for the adaptedness of a given coordinate system (at least under certain non-degeneracy conditions), which prove to be useful. The goal of this section is to provide a new, elementary approach to these results in the general case. Our approach has the additional advantage of extending to the category of smooth functions.

3.1. On the effect of a change of coordinates on the Newton diagram. We have to understand what effect a change of coordinates has on the Newton diagram. To this end, we shall make use of the following auxiliary result.

Lemma 3.1. *Let $m \in \mathbb{N}$, $m \geq 1$, be given, and denote by μ the weight $\mu := (1, m)$. Moreover, let P be a μ -homogeneous polynomial. Then its Newton diagram $\mathcal{N}_d(P)$ is a compact interval $[(A_0, B_0), (A_1, B_1)]$ joining two vertices $(A_0, B_0), (A_1, B_1)$ (which may coincide), where we shall assume that $A_0 \leq A_1$.*

Moreover, let $x = \varphi(y)$ be a change of coordinates of the form $x_1 = y_1$ and $x_2 = y_2 + a_2 y_1^m$ (with $a_2 \neq 0$) if $m \geq 1$, or $x_1 = y_1 + a_1 y_2$ and $x_2 = y_2 + a_2 y_1$ if $m = 1$ (with $a_j \neq 0$). Denote by \tilde{P} the polynomial $P \circ \varphi$. Then \tilde{P} is μ -homogeneous of the same degree as P , and its Newton diagram $\mathcal{N}_d(\tilde{P})$ is an interval of the form $[(\tilde{A}_0, \tilde{B}_0), (\tilde{A}_1, \tilde{B}_1)]$, with $\tilde{A}_0 \leq \tilde{A}_1$.

Assume that either

- (i) *the interval $\mathcal{N}_d(P)$ lies in the closed half-space above the bisectrix, i.e., $j \leq k$ for every $(j, k) \in \mathcal{N}_d(P)$ (Figure 1)*
- (ii) *or the point (A_0, B_0) lies above or on the bisectrix, i.e., $A_0 \leq B_0$, and $m(P) \leq d(P)$, where $m(P)$ again denotes the maximal order of vanishing of P along the unit circle S^1 and $d(P)$ denotes the distance.*

Then the point $(\tilde{A}_0, \tilde{B}_0)$ lies in the closed half-space above the bisectrix and the point $(\tilde{A}_1, \tilde{B}_1)$ lies in the closed half-space below the bisectrix. In particular, the Newton diagram $\mathcal{N}_d(\tilde{P})$ intersects the bisectrix (Figure 2).

Proof. We have to show that conditions (i) or (ii) imply

$$(3.1) \quad \tilde{A}_0 \leq \tilde{B}_0 \text{ and } \tilde{A}_1 \geq \tilde{B}_1.$$

Now, by Proposition 2.2, we can write P in the form

$$(3.2) \quad P(x_1, x_2) = x_1^\alpha x_2^\beta \prod_l (x_2 - c_l x_1^m)^{n_l},$$

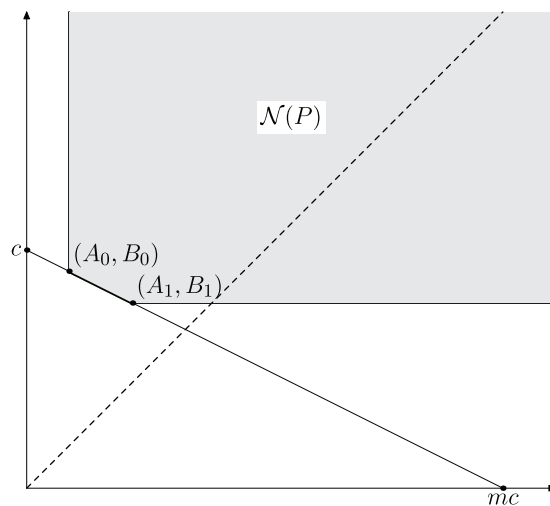


FIGURE 1

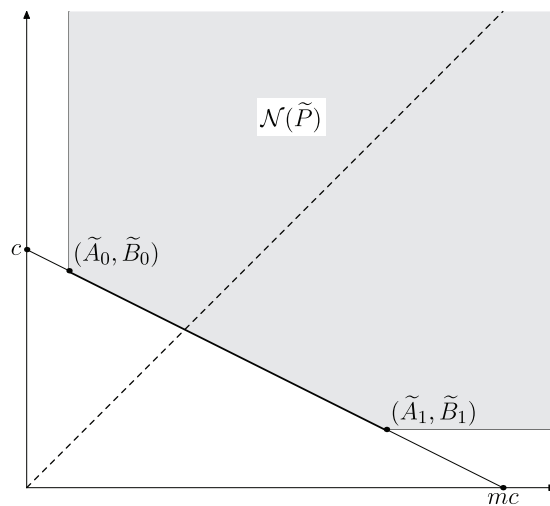


FIGURE 2

where the c_l 's are the non-trivial distinct complex roots of the polynomial $P(1, x_2)$ and the n_l 's are their multiplicities.

It is easy to read from (3.2) that the vertices of the Newton polyhedron of P are given by (A_0, B_0) and (A_1, B_1) , where

$$(3.3) \quad A_0 := \alpha, \quad B_0 := \beta + N, \quad A_1 := \alpha + mN, \quad B_1 := \beta,$$

and that

$$d_h(P) = \frac{A_1 + mB_1}{1 + m} = \frac{\alpha + m(\beta + N)}{1 + m}$$

(compare also [5]); here, we have put

$$N := \sum_l n_l.$$

Assume first that $m \geq 1$ and $x_1 = y_1, x_2 = y_2 + a_2 y_1^m$. Then, by (3.2), we have

$$(3.4) \quad \tilde{P}(y_1, y_2) = y_1^\alpha (y_2 + a_2 y_1^m)^\beta \prod_l (y_2 - (c_l - a_2) y_1^m)^{n_l}.$$

By looking at the terms of lowest power of y_1 in (3.4), we see that $\tilde{A}_0 = \alpha, \tilde{B}_0 = \beta + N$, i.e.,

$$(\tilde{A}_0, \tilde{B}_0) = (A_0, B_0),$$

so that $\tilde{A}_0 \leq \tilde{B}_0$. We next identify the term of highest power of y_1 in (3.4) in order to determine $(\tilde{A}_1, \tilde{B}_1)$.

If $a_2 \neq c_l$ for every l , then $\tilde{A}_1 = \alpha + m\beta + mN$ and $\tilde{B}_1 = 0$; hence $\tilde{A}_1 > \tilde{B}_1$.

Assume next that $a_2 = c_{l_0}$ coincides with a non-trivial root (which is then necessarily real). Then we obtain

$$(3.5) \quad \tilde{A}_1 = \alpha + m\beta + m(N - n_{l_0}), \quad \tilde{B}_1 = n_{l_0}.$$

Now, if condition (i) holds, then by (3.3) we have

$$(3.6) \quad \alpha + mN \leq \beta,$$

which by Corollary 2.3 easily implies $n_{l_0} \leq \alpha + m(\beta + N - n_{l_0})$. Also, if (ii) holds, then

$$n_{l_0} \leq m(P) \leq d_h(P) = \frac{\alpha + m(\beta + N)}{1 + m},$$

which is equivalent to $n_{l_0} \leq \alpha + m(\beta + N - n_{l_0})$. Thus both conditions imply that

$$(3.7) \quad \tilde{A}_1 \geq \tilde{B}_1,$$

so that (3.1) is satisfied.

There remains the case $m = 1$ and $x_1 = y_1 + a_1 y_2, x_2 = y_2 + a_2 y_1^m$, with $a_1, a_2 \neq 0$. In this case,

$$(3.8) \quad \tilde{P}(y_1, y_2) = (y_1 + a_1 y_2)^\alpha (y_2 + a_2 y_1)^\beta \prod_l \left((1 - c_l a_1) y_2 + (a_2 - c_l) y_1 \right)^{n_l},$$

where $a_1, a_2 \neq 0$. Our assumptions imply that either (3.6) holds or that for every l such that c_l is real we have

$$(3.9) \quad n_l \leq d_h(P) = \frac{\alpha + \beta + N}{2}.$$

Notice also that it may happen that $c_{l_1} a_1 = 1$ or $c_{l_0} = a_2$, but at most for one l_0 and one l_1 , and that these indices must be different, since φ has a non-degenerate Jacobian at 0.

Now, if $c_l a_1 \neq 1$ for every l , then

$$\tilde{A}_0 = 0, \quad \tilde{B}_0 = \alpha + \beta + N,$$

and if $c_{l_1} a_1 = 1$, then

$$\tilde{A}_0 = n_{l_1}, \quad \tilde{B}_0 = \alpha + \beta + N - n_{l_1}.$$

Moreover, if $c_l \neq a_2$ for every l , then

$$\tilde{A}_1 = \alpha + \beta + N, \quad \tilde{B}_1 = 0,$$

and if $c_{l_0} = a_2$, then

$$\tilde{A}_1 = \alpha + \beta + N - n_{l_0}, \quad \tilde{B}_1 = n_{l_0}.$$

Using (3.6), respectively (3.9), it is easy to check that $\tilde{A}_0 \leq \tilde{B}_0$ and $\tilde{A}_1 \geq \tilde{B}_1$ in all cases. \square

Now fix a smooth function f defined near the origin satisfying $f(0, 0) = 0$ and $\nabla f(0, 0) = 0$.

Choose $\kappa := (\kappa_1, \kappa_2)$ with $\kappa_j \geq 0$ such that the principal face $\pi(f)$ of the Newton polyhedron lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. Notice that κ is uniquely determined unless $\pi(f)$ is a vertex. In the latter case, we shall assume that $\kappa_1, \kappa_2 > 0$. Flipping coordinates x_1 and x_2 , if necessary, we shall assume without loss of generality that $\kappa_1 \leq \kappa_2$. Notice that then $\kappa_1 = 0$ if and only if the principal face is non-compact. In this case, it is a half-line which lies on the horizontal line L_κ given by $t_2 = 1/\kappa_2$. Also, if $\kappa_1 > 0$, then the principal face is compact and the principal part $f_p = f_\kappa$ of f is a κ -homogeneous polynomial of degree one.

Next consider a smooth local coordinate system y at the origin given by $x = \varphi(y)$, and put

$$\tilde{f}(y_1, y_2) := f(\varphi(y_1, y_2)).$$

Flipping coordinates y_1 and y_2 , if necessary, we may assume without loss of generality that $(x_1, x_2) = (\varphi_1(y_1, y_2), \varphi_2(y_1, y_2))$ satisfies $\frac{\partial \varphi_j(0,0)}{\partial y_j} \neq 0$ for $j = 1, 2$.

Therefore, we can write the functions φ_1, φ_2 in the form

$$(3.10) \quad \varphi_1(y_1, y_2) = y_1 \psi_1(y_1, y_2) + \eta_1(y_2), \quad \varphi_2(y_1, y_2) = y_2 \psi_2(y_1, y_2) + \eta_2(y_1),$$

where $\psi_1, \psi_2, \eta_1, \eta_2$ are smooth functions satisfying

$$\psi_1(0, 0) \neq 0, \quad \psi_2(0, 0) \neq 0, \quad \eta_1(0) = \eta_2(0) = 0.$$

Since separate scaling of the coordinates x_1 and x_2 does not change the Newton polyhedron, we may further assume that $\psi_1(0, 0) = \psi_2(0, 0) = 1$.

Denote by k_j the order of vanishing of η_j at 0, $j = 1, 2$. Then clearly $k_j \geq 1$. We shall see that if $k_2 = \infty$, i.e., if the function η_2 is flat at 0, then the distance will not change under the change of coordinates; i.e., $d_x = d_y$.

On the other hand, if k_2 is finite, then according to (3.10), the main term of φ_2 is of the form $y_2 + a_2 y_1^{k_2}$, with $a_2 \neq 0$. Therefore, we then introduce a second weight $\mu := (1, k_2)$. Then the μ -principal part of $\varphi_2(y_1, y_2)$ is given by $y_2 + a_2 y_1^{k_2}$, which is μ -homogeneous of degree k_2 . Recall that $L_\mu := \{(t_1, t_2) \in \mathbb{R}^2 : t_1 + k_2 t_2 = d\}$ denotes the supporting line to the Newton polyhedron $\mathcal{N}(f)$. Notice that the line L_κ has slope $\frac{\kappa_1}{\kappa_2} \leq 1$ and that L_μ has slope $\frac{1}{k_2} \leq 1$. The effect of the change of coordinates on the Newton diagram is then related to the interplay between these two homogeneities κ and μ , in particular the relation between the slopes of the corresponding supporting lines.

We put $\frac{\kappa_2}{\kappa_1} := \infty$, if $\kappa_1 = 0$.

Lemma 3.2. *In the situation described above, the following hold true:*

- (a) *Assume either that*
 - (i) $\kappa_2 > \kappa_1$, and either $k_2 = \infty$ or $\infty > k_2 > \frac{\kappa_2}{\kappa_1}$ (so that the line L_μ is less steep than the line L_κ), or that
 - (ii) $\kappa_1 = \kappa_2$ and $k_1, k_2 > 1$.

Then $d_y = d_x$.

(b) If $k_2 < \frac{\kappa_2}{\kappa_1}$ (so that the line L_μ is steeper than the line L_κ), then $d_y \leq d_x$.

In particular, if $\kappa_1 = 0$, i.e., if the principal face is non-compact, then the coordinates (x_1, x_2) are adapted to f .

Proof. (a) **The case where $\kappa_1 = 0$ and $k_2 = \infty$.**

In this case $n := \frac{1}{\kappa_2} \in \mathbb{N}$, $n \geq 1$, so that the principal face is a half-line lying on the horizontal line L_κ given by $t_2 = n$, and $d_x = n$. Moreover, η_2 is flat at the origin; i.e.,

$$(3.11) \quad \partial_{y_1}^j \eta_2(0) = 0$$

for every $j \in \mathbb{N}$. In view of the structure of the Newton polyhedron that we assume, it is easy to see that f can then be written in the form

$$f(x_1, x_2) = x_2^n x_1^{n_1} g(x_1) + x_2^{n+1} h(x_1, x_2) + \sum_{k=0}^{n-1} a_k(x_1) x_2^k,$$

where g and h are smooth functions and $g(0) \neq 0$, and where a_1, \dots, a_{n-1} are flat functions at the origin. Moreover, $n_1 < n$. This implies that

$$\tilde{f}(y_1, y_2) = \varphi_2(y_1, y_2)^n \varphi_1(y_1, y_2)^{n_1} G(y_1, y_2) + \varphi_2(y_1, y_2)^{n+1} H(y_1, y_2) + R(y_1, y_2)$$

for some smooth functions G, H , and R , where $G(0, 0) \neq 0$, and where R is flat at the origin. Since $\frac{\partial \varphi_k}{\partial y_k}(0, 0) = 1$, by the product rule and (3.11) this easily implies that $\partial_{y_1}^{n_1} \partial_{y_2}^n \tilde{f}(0, 0) \neq 0$ and $\partial_{y_1}^l \partial_{y_2}^j \tilde{f}(0, 0) = 0$ if $j < n$, or if $j = n$ and $l < n_1$, so that $(n_1, n) \in \mathcal{N}(\tilde{f})$, but no lattice point below or to the left of this point belongs to $\mathcal{N}(\tilde{f})$. This shows that $d_y = d_x = n$.

(b) **The case where $k_2 > \frac{\kappa_2}{\kappa_1}$ and $\kappa_1 > 0$, and either $\kappa_2 > \kappa_1$, or $\kappa_1 = \kappa_2$ and $k_1 > 1$.**

Let us first state some general observations. If φ_κ denotes the κ -principal part of φ , then it is easily seen, e.g., by Lemma 2.1, that

$$\tilde{f}(y_1, y_2) = f_\kappa \circ \varphi_\kappa(y_1, y_2) + \text{terms of higher } \kappa\text{-degree},$$

so that

$$\tilde{f}_\kappa = f_\kappa \circ \varphi_\kappa.$$

Moreover, $f_\kappa \circ \varphi_\kappa$ is a κ -homogeneous polynomial, so that its Newton diagram $\mathcal{N}_d(\tilde{f}_\kappa)$ is again a compact interval (possibly a single point). In case this interval intersects the bisectrix, too, then the principal face of $\mathcal{N}(\tilde{f})$ is contained in it. Moreover, if φ_κ has κ -degree one, then $\mathcal{N}_d(f_\kappa)$ again lies on the line L_κ , and consequently we have $d_y = d_x$.

Notice that the same conclusion holds true if x_1, x_2 are κ -homogeneous of any degree δ and if φ_κ has the same κ -degree δ .

Now, if $k_2 > \frac{\kappa_2}{\kappa_1}$, then $k_2 \kappa_1 > \kappa_2$, and thus, by (3.10), $x_2 = y_2$, up to terms of higher κ -degree. Moreover, if $\kappa_2 > \kappa_1$ or $k_1 > 1$, then also $x_1 = y_1$, up to terms of higher κ -degree, so that $\varphi_\kappa(y_1, y_2) = (y_1, y_2)$. Thus the reasoning above applies, and we see that $d_y = d_x$.

c) **The case where $k_2 < \frac{\kappa_2}{\kappa_1}$ and $\kappa_1 \geq 0$.**

In this case, the μ -principal part of $\varphi_2(y_1, y_2)$ is given by $y_2 + a_2 y_1^{k_2}$, and the μ -principal part of $\varphi_1(y_1, y_2)$ is given by y_1 if $k_2 > 1$ and by $y_1 + a_1 y_2$ if $k_2 = 1$, with $a_1 \neq 0$ if and only if $k_1 = 1$.

Choose $d > 0$ such that the line $L_\mu := \{(t_1, t_2) \in \mathbb{R}^2 : t_1 + k_2 t_2 = d\}$ is the supporting line to the Newton polyhedron $\mathcal{N}(f)$, and let f_μ be the μ -principal part of f . Then obviously $d \in \mathbb{N}$. Since the line L_μ is steeper than the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$, it is clear from the geometry of the Newton polyhedron that $L_\mu \cap \mathcal{N}(f)$ is a compact interval I_μ (possibly a single point) lying in the closed half-space above the bisectrix; i.e.,

$$f_\mu(x) = \sum_{(j,k) \in I_\mu} c_{jk} x_1^j x_2^k, \quad \text{where } j \leq k \text{ for every } (j, k) \in I_\mu.$$

Applying Lemma 3.1 to the μ -homogeneous polynomial f_μ , we see that the Newton diagram of $\tilde{f}_\mu = f_\mu \circ \varphi_\mu$ intersects the bisectrix. The principal face of $\mathcal{N}(\tilde{f})$ therefore lies on the line L_μ , and since this line is steeper than the line L_κ containing the principal face of $\mathcal{N}(f)$, it is clear from the geometry that $d_y \leq d_x$.

Finally, assume that $\kappa_1 = 0$. Then either $k_2 = \infty$ or $k_2 < \frac{\kappa_2}{\kappa_1}$, and so we always have $d_y \leq d_x$. Therefore, the coordinates (x_1, x_2) are adapted to f . \square

Theorem 3.3. *Let f be a real-analytic (respectively smooth) function near the origin, with $f(0, 0) = 0$ and $\nabla f(0, 0) = 0$. Assume that the given coordinates are not adapted to f . Then all of the following conditions hold true:*

- (a) *The principal face $\pi(f)$ of the Newton polyhedron is a compact edge. It thus lies on a uniquely determined line $\kappa_1 t_1 + \kappa_2 t_2 = 1$, with $\kappa_1, \kappa_2 > 0$. Flipping coordinates x_1, x_2 , if necessary, we assume without loss of generality that $\kappa_1 \leq \kappa_2$.*
- (b) *$\frac{\kappa_2}{\kappa_1} \in \mathbb{N}$.*
- (c) *Any root of f_p of maximal order $m(f_p)$ on the unit circle S^1 lies away from the coordinate axes, and we have $m(f_p) > d(f)$. In particular, there exists a unique non-trivial real root a of the polynomial $f_p(1, x_2)$ with multiplicity $n_a = m(f_p) > d(f)$.*

Moreover, in this case, if we put $m := \frac{\kappa_2}{\kappa_1} \in \mathbb{N}$, then an adapted coordinate system for the principal part f_p of f is defined by $y_1 := x_1, y_2 := x_2 - ax_1^m$. The height of f_p is then given by $h(f_p) = m(f_p)$.

Conversely, if the conditions (a)-(c) are satisfied, then the given coordinates are not adapted to the principal part f_p of f .

Notice that for mixed homogeneous polynomials $f = f_p$ the theorem gives a necessary and sufficient condition for the adaptedness of the coordinates. We shall see later that the same is indeed true for general functions f .

Also observe that the root in (c) with multiplicity $n_a > d(f)$ is the principal root of f_p .

We finally remark that condition (b) is in fact redundant, since it is a consequence of (a) and (c), in view of Corollary 2.3. We nevertheless included it in order to stress that (b) is a necessary condition for non-adaptedness.

Proof of Theorem 3.3. Assume that the coordinates (x_1, x_2) are not adapted to f . Then there exists a smooth change of coordinates $x = \varphi(y)$ at the origin such that

$$(3.12) \quad d_y > d_x.$$

Arguing similarly as in the proof of Lemma 3.2, we may assume that φ satisfies (3.10), with

$$\psi_1(0, 0) = \psi_2(0, 0) = 1, \quad \eta_1(0) = \eta_2(0) = 0.$$

Denote again by k_j the order of vanishing of η_j at 0, $j = 1, 2$, and choose the weight $\kappa = (\kappa_1, \kappa_2)$ as before in such a way that the principal face of $\mathcal{N}(f)$ lies on the line L_κ . Again, assume that $\kappa_1 \leq \kappa_2$.

From (3.12) and Lemma 3.2 it then follows that if $\kappa_2 > \kappa_1$, then

$$\frac{\kappa_2}{\kappa_1} = k_2 \in \mathbb{N},$$

and if $\kappa_1 = \kappa_2$, then $k_1 = 1$ or $k_2 = 1$. In the latter case, by symmetry in the variables x_1 and x_2 , let us assume without loss of generality that $k_2 = 1$. In particular, the principal face $\pi(f)$ is compact, hence either a compact edge or a vertex. It is therefore an interval $[(A_0, B_0), (A_1, B_1)]$ joining two vertices $(A_0, B_0), (A_1, B_1)$ (which may coincide), i.e.,

$$\pi(f) = [(A_0, B_0), (A_1, B_1)].$$

We shall assume that $A_0 \leq A_1$. Since this interval intersects the bisectrix $t_1 = t_2$, we then have

$$(3.13) \quad A_0 \leq B_0, \quad A_1 \geq B_1.$$

In the sequel, let us write $m := k_2$ and consider the κ -homogeneous polynomial $f_p = f_\kappa$. Its Newton diagram is given by the interval $[(A_0, B_0), (A_1, B_1)]$.

We next show that $m(f_\kappa) > d_x$. To this end, assume to the contrary that $m(f_\kappa) \leq d_x$, so that f_κ satisfies condition (ii) in Lemma 3.1. Observe that the κ -principal part of $\varphi_2(y_1, y_2)$ is given by $y_2 + a_2 y_1^m$, where $a_2 \neq 0$, and the κ -principal part of $\varphi_1(y_1, y_2)$ is given by y_1 , unless $\kappa_1 = \kappa_2$ (hence $m = 1$) and $k_1 = 1$, when it is given by $y_1 + a_1 y_2$, with $a_1 \neq 0$. Therefore, the κ -principal part \tilde{f}_κ of \tilde{f} is given by

$$\tilde{f}_\kappa(y_1, y_2) = f_\kappa(y_1 + a_1 y_2, y_2 + a_2 y_1^m),$$

where $a_2 \neq 0$ and $a_1 \neq 0$ if and only if $m = k_1 = 1$. Lemma 3.1 then shows that the Newton diagram of \tilde{f}_κ intersects the bisectrix so that the principal face of $\mathcal{N}(\tilde{f})$ lies on the line L_κ ; hence $d_y = d_x$. This contradicts (3.12).

We have thus seen that our assumptions on f imply that $\frac{\kappa_2}{\kappa_1} \in \mathbb{N}$ and $m(f_p) > d(f)$. Assuming these properties, write the principal part $f_p = f_\kappa$ of f according to Proposition 2.2 in the form

$$(3.14) \quad f_\kappa(x_1, x_2) = x_1^\alpha x_2^\beta \prod_l (x_2 - c_l x_1^m)^{n_l},$$

where the c_l 's are the non-trivial distinct complex roots of the polynomial $f_\kappa(1, x_2)$ and the n_l 's are their multiplicities.

Observe that clearly $\mathcal{N}(f_\kappa)$ is contained in the half-plane $\{t_1 \geq \alpha\}$, so that $d_x = d(f_\kappa) \geq \alpha$. Similarly, we see that $d_x \geq \beta$. Consequently, there must be an l_0 with real root c_{l_0} so that $n_{l_0} = m(f_p)$. Notice that this root is unique, by Corollary 2.3. Let us remark at this point that this excludes the possibility that the principal face of f is a vertex, for then $f_p(x)$ would be of the form $c x_1^\alpha x_2^\beta$, so that $f_p(1, x_2)$ has no non-trivial root. Consequently, $\pi(f)$ is a compact edge.

We show that in this case the change of coordinates $x_1 := y_1, x_2 := y_2 + a_{l_0} y_1^m$ leads to adapted coordinates. To this end, we shall refer to the notation and results in the proof of Lemma 3.1, applied to $P := f_p$.

We have seen that the function \tilde{f}_p representing f_p in the new coordinates (y_1, y_2) is again a κ -homogeneous polynomial, so that its Newton diagram is a compact

interval $[(\tilde{A}_0, \tilde{B}_0), (\tilde{A}_1, \tilde{B}_1)]$ whose endpoints are given by $(\tilde{A}_0, \tilde{B}_0) = (A_0, B_0)$ and (3.5). Since now $n_{l_0} > d_x$, the order signs \geq in (3.7) have to be replaced by $<$, i.e.,

$$\tilde{A}_1 < \tilde{B}_1.$$

This implies that the interval $[(\tilde{A}_0, \tilde{B}_0), (\tilde{A}_1, \tilde{B}_1)]$ lies in the half-plane $t_2 > t_1$. Consequently, the principal face of $\mathcal{N}(\tilde{f}_p)$ lies on the horizontal line $t_2 = \tilde{B}_1$; hence $d_y = \tilde{B}_1 = m(\tilde{f}_p)$. Since this face is non-compact, the coordinates (y_1, y_2) are adapted, due to Lemma 3.2.

Notice, however, that the same change of coordinates will in general not lead to an adapted coordinate system for f – this may require further, higher order terms in y_1 , in addition to $a_{l_0}y_1^m$.

Finally, assume conversely that all of the conditions (a)-(c) are satisfied. Then the coordinates (x_1, x_2) are not adapted to f_p , since, as we have seen, we can change coordinates for f_p in such a way that, in the new coordinates (y_1, y_2) , the distance is given by $d_y = m(\tilde{f}_p) > d_x$. This concludes the proof of Theorem 3.3. \square

As a corollary, we obtain the following characterization of the height in the case of a κ -homogeneous polynomial.

Corollary 3.4. *Let P be a κ -homogeneous polynomial as in Proposition 2.2. Then*

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

Proof. We adopt the notation from Proposition 2.2. If $\frac{\kappa_2}{\kappa_1} \notin \mathbb{N}$, then the coordinates are adapted to P , and by Corollary 2.3 (a) and (2.11) we have $\max\{m(P), d_h(P)\} = \max\{\nu_1, \nu_2, d_h(P)\} = d(P) = h(P)$.

So, assume next that $\frac{\kappa_2}{\kappa_1} \in \mathbb{N}$. Then, according to Corollary 2.3 (b), there is at most one real root of P on the unit circle with multiplicity $n_{l_0} > d_h(P)$.

If there is no such root, then $m(P) \leq d_h(P)$, and so the coordinates are adapted. This shows that $h(P) = d(P) = d_h(P) = \max\{m(P), d_h(P)\}$.

If there is such a root, and if it lies on one of the coordinate axes, then $n_{l_0} = \nu_1$ or $n_{l_0} = \nu_2$, hence $m(P) = \max\{\nu_1, \nu_2\}$, and the claim follows. We may thus assume that the root does not lie on a coordinate axis, so that $1 \leq l_0 \leq M$. Then $\nu_1, \nu_2 \leq d_h(P)$, which implies that the principal face $\pi(P)$ of the Newton polyhedron of P must be a compact edge. Then, by Theorem 3.3, the coordinates are not adapted to P , and $h(P) = m(P) > d(P)$. Since $d(P) \geq d_h(P)$, the conclusion also follows in this case. \square

4. THE REAL-ANALYTIC CASE

In this section we shall provide yet another proof of Varchenko’s result that every real-analytic function f admits an analytic coordinate system which is adapted to f . Our proof combines Varchenko’s algorithm for the construction of an adapted coordinate system with the use of the Puiseux series expansion of roots of f in the work by Phong, Stein and Sturm [6]. This approach will turn out to be useful for the extension to the smooth setting.

4.1. Description of the Newton polyhedron in terms of the roots. We begin by recalling some notions and results from the seminal article [5] by Phong and Stein.

Again assume f to be real-analytic and real-valued. By the Weierstraß preparation theorem we can then write

$$f(x_1, x_2) = U(x_1, x_2)x_1^{\nu_1}x_2^{\nu_2}F(x_1, x_2)$$

near the origin, where $F(x_1, x_2)$ is a pseudopolynomial of the form

$$F(x_1, x_2) = x_2^m + g_1(x_1)x_2^{m-1} + \dots + g_m(x_1)$$

and U, g_1, \dots, g_m are real-analytic functions satisfying $U(0, 0) \neq 0, g_j(0) = 0$. Observe that the Newton polyhedron of f is the same as that of $x_1^{\nu_1}x_2^{\nu_2}F(x_1, x_2)$. We shall also assume without loss of generality that g_m is a non-trivial function so that the roots $r(x_1)$ of the equation $F(x_1, x_2) = 0$, considered as a polynomial in x_2 , are all non-trivial.

It is well known that these roots can be expressed in a small neighborhood of 0 as Puiseux series

$$r(x_1) = c_{l_1}^{\alpha_1}x_1^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2}x_1^{a_{l_1 l_2}^{\alpha_1}} + \dots + c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_p}x_1^{a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}}} + \dots,$$

where

$$\begin{aligned} c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1} \beta} &\neq c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1} \gamma} \quad \text{for } \beta \neq \gamma, \\ a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}} &> a_{l_1 \dots l_{p-1}}^{\alpha_1 \dots \alpha_{p-2}}, \end{aligned}$$

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

By the cluster $\begin{bmatrix} \alpha_1 & \dots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix}$, we shall designate all the roots $r(x_1)$, counted with their multiplicities, which satisfy

$$(4.1) \quad r(x_1) - c_{l_1}^{\alpha_1}x_1^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2}x_1^{a_{l_1 l_2}^{\alpha_1}} + \dots + c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_p}x_1^{a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}}} = O(x_1^b)$$

for some exponent $b > a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}}$. We also introduce the clusters $\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix}$ by

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

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

$$\begin{aligned} N \begin{bmatrix} \alpha_1 & \dots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix} &:= \text{number of roots in } \begin{bmatrix} \alpha_1 & \dots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix}, \\ N \begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix} &:= \text{number of roots in } \begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix}. \end{aligned}$$

Let $a_1 < \dots < a_l < \dots < a_n$ be the distinct leading exponents of all the roots of F . Each exponent a_l corresponds to the cluster $\begin{bmatrix} \cdot \\ l \end{bmatrix}$ so that the set of all roots of

F can be divided as $\bigcup_{l=1}^n \begin{bmatrix} \cdot \\ l \end{bmatrix}$. Then we may write

$$f(x_1, x_2) = U(x_1, x_2)x_1^{\nu_1}x_2^{\nu_2} \prod_{l=1}^n \Phi \begin{bmatrix} \cdot \\ l \end{bmatrix} (x_1, x_2),$$

where

$$\Phi \begin{bmatrix} \cdot \\ l \end{bmatrix} (x_1, x_2) := \prod_{r \in [l]} (x_2 - r(x_1)).$$

We introduce the following quantities:

$$(4.2) \quad A_l = A \begin{bmatrix} \cdot \\ l \end{bmatrix} := \nu_1 + \sum_{\mu \leq l} a_\mu N \begin{bmatrix} \cdot \\ \mu \end{bmatrix}, \quad B_l = B \begin{bmatrix} \cdot \\ l \end{bmatrix} := \nu_2 + \sum_{\mu \geq l+1} N \begin{bmatrix} \cdot \\ \mu \end{bmatrix}.$$

Notice that B_l is just the number of all roots with leading exponent strictly greater than a_l (where here we interpret the trivial roots corresponding to the factor $(x_2 - 0)^{\nu_2}$ in our representation of $f(x_1, x_2)$ as roots with exponent $+\infty$).

If N_{all} denotes the total number of all roots of f away from the axis $x_1 = 0$, including the trivial ones and counted with their multiplicities, then we can also write

$$(4.3) \quad B_l = N_{\text{all}} - \sum_{\mu \leq l} N \begin{bmatrix} \cdot \\ \mu \end{bmatrix}.$$

Then, similarly as for the reduced Newton diagram in [5], the vertices of the Newton diagram $\mathcal{N}_d(f)$ of f are the points (A_l, B_l) , $l = 0, \dots, n$, and the Newton polyhedron $\mathcal{N}(f)$ is the convex hull of the set $\bigcup_l ((A_l, B_l) + \mathbb{R}_+^2)$.

Notice also that

$$A_l + a_l B_l = A_{l-1} + a_l B_{l-1}.$$

Let $L_l := \{(t_1, t_2) \in \mathbb{N}^2 : \kappa_1^l t_1 + \kappa_2^l t_2 = 1\}$ denote the line passing through the points (A_{l-1}, B_{l-1}) and (A_l, B_l) . It is easy to see that it is given by

$$(4.4) \quad \begin{aligned} \kappa_1^l &= \frac{\Delta B_l}{A_l \Delta B_l - B_l \Delta A_l} = \frac{1}{A_l + a_l B_l}, \\ \kappa_2^l &= \frac{\Delta A_l}{A_l \Delta B_l - B_l \Delta A_l} = \frac{a_l}{A_l + a_l B_l}, \end{aligned}$$

where $\Delta A_l := A_l - A_{l-1}$, $\Delta B_l := B_l - B_{l-1}$. This implies that

$$(4.5) \quad \frac{\kappa_2^l}{\kappa_1^l} = a_l,$$

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

$$d_l := \frac{A_l + a_l B_l}{1 + a_l}.$$

Moreover, the vertical edge of the Newton polyhedron, which passes through the point $(A_0, B_0) = (\nu_1, \nu_2 + m)$, intersects the bisectrix at (ν_1, ν_1) , and the horizontal edge, which passes through the point $(A_n, B_n) = (A_n, \nu_2)$, intersects the bisectrix at (ν_2, ν_2) . We therefore conclude that the distance $d(f)$ is given by $d(f) = \max\{\nu_1, \nu_2, \max_{l=1, \dots, n} d_l\}$.

Finally, fix l , and let us determine the κ^l -principal part f_{κ^l} of f corresponding to the supporting line L_l . To this end, observe that f has the same κ^l -principal part as the function

$$U(0, 0) x_1^{\nu_1} x_2^{\nu_2} \prod_{\alpha, \mu} (x_2 - c_\mu^\alpha x_1^{a_\mu})^{N \begin{bmatrix} \alpha \\ \mu \end{bmatrix}}.$$

Moreover, the κ^l -principal part of $x_2 - c_\mu^\alpha x_1^{a_\mu}$ is given by $c_\mu^\alpha x_1^{a_\mu}$ if $\mu < l$, and by x_2 if $\mu > l$. This implies that

$$(4.6) \quad f_{\kappa^l}(x_1, x_2) = c_l x_1^{A_{l-1}} x_2^{B_l} \prod_{\alpha} \left(x_2 - c_l^\alpha x_1^{a_l} \right)^N \begin{bmatrix} \alpha \\ l \end{bmatrix}.$$

In view of this identity, we shall say that the edge $[(A_{l-1}, B_{l-1}), (A_l, B_l)]$ is *associated* to the cluster of roots $\begin{bmatrix} \cdot \\ l \end{bmatrix}$. We collect these results in the following lemma.

Lemma 4.1. *The vertices of the Newton diagram $\mathcal{N}_d(f)$ of f are the points (A_l, B_l) , $l = 0, \dots, n$, with A_j, B_j given by (4.2), and its edges are the intervals $[(A_{l-1}, B_{l-1}), (A_l, B_l)]$, $l = 1, \dots, n$. Moreover, the distance between the Newton polyhedron and the origin is given by*

$$(4.7) \quad d(f) = \max \left\{ A_0, B_n, \max_{l=1, \dots, n} \frac{A_l + a_l B_l}{1 + a_l} \right\},$$

and the κ^l -principal part of f corresponding to the supporting line L_l through the edge $[(A_{l-1}, B_{l-1}), (A_l, B_l)]$ is given by (4.6).

4.2. Existence of adapted coordinates in the analytic setting.

Theorem 4.2. *Let f be a non-trivial real valued real-analytic function defined on a neighborhood of the origin in \mathbb{R}^2 with $f(0, 0) = 0, \nabla f(0, 0) = 0$. Choose $\kappa_1, \kappa_2 \geq 0$ such that the principal face $\pi(f)$ of the Newton polyhedron of f lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. Without loss of generality, we may assume that $\kappa_2 \geq \kappa_1$. Then there exists a real-analytic function $\psi(x_1)$ of x_1 near the origin with $\psi(0) = 0$ such that an adapted coordinate system (y_1, y_2) for f near 0 is given by $y_1 := x_1, y_2 := x_2 - \psi(x_1)$.*

The function ψ can in fact be chosen as one of the roots, respectively a leading partial sum of the Puiseux series expansion of one of the roots, of the equation $f(x_1, x_2) = 0$, considered as an equation in x_2 .

Notice that Theorem 4.2 was proved by A. N. Varchenko in [8] for analytic functions (without multiple components) by means of a two-dimensional resolution of singularities result, and by D. H. Phong, E. M. Stein and J. A. Sturm [6] by means of the Puiseux series expansion of roots of f . Our proof will be based on the Puiseux series expansion of roots too, and will in fact give an explicit description of an adapted coordinate system in terms of these roots. However, in contrast to [6], we shall follow more closely Varchenko’s algorithm for constructing adapted coordinates, which will become useful for the extension to the smooth setting.

Proof. We may and shall assume without loss of generality that $U \equiv 1$. If the coordinate system is adapted, then we can choose $\psi = 0$. According to Theorem 3.3, this applies in any of the following three cases:

- (a) The principal face is unbounded. Since we are assuming that $\kappa_2 \geq \kappa_1$, by Lemma 4.1 this means that $A_n < B_n$, i.e., that

$$(4.8) \quad \nu_1 + \sum_{l=1}^n a_l N \begin{bmatrix} \cdot \\ l \end{bmatrix} < \nu_2.$$

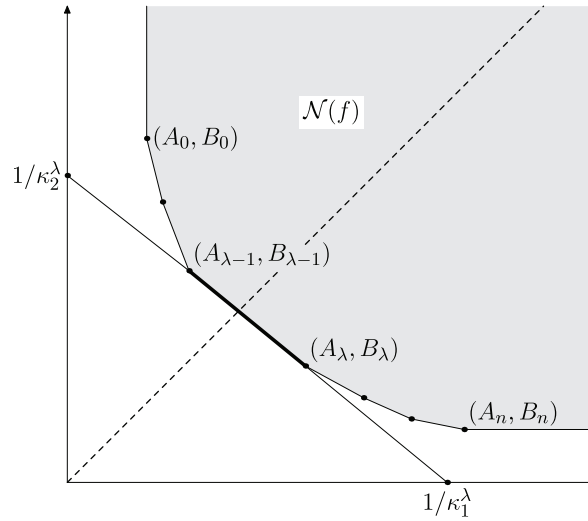


FIGURE 3

(b) $\pi(f)$ consists of a vertex. Then choose $0 \leq \lambda \leq n$ so that $\pi(f) = \{(A_\lambda, A_\lambda)\}$. This happens iff

$$(4.9) \quad \nu_1 + \sum_{l \leq \lambda} a_l N \begin{bmatrix} \cdot \\ l \end{bmatrix} = \nu_2 + \sum_{l \geq \lambda+1} N \begin{bmatrix} \cdot \\ l \end{bmatrix}.$$

(c) $\pi(f)$ is a compact edge $[(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$, but

$$(4.10) \quad a_\lambda = \frac{\kappa_2^\lambda}{\kappa_1^\lambda} \notin \mathbb{N}$$

or

$$(4.11) \quad a_\lambda \in \mathbb{N} \text{ and } N \begin{bmatrix} \alpha \\ \lambda \end{bmatrix} \leq d(f) \text{ whenever } c_\lambda^\alpha \in \mathbb{R}.$$

There remains the case where the principal face $\pi(f)$ is a compact edge $\pi(f) = [(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$ ($1 \leq \lambda \leq n$), $a_\lambda \in \mathbb{N}$ and there is an index β such that

$$(4.12) \quad m(f_p) = N \begin{bmatrix} \beta \\ \lambda \end{bmatrix} > d(f) = \frac{A_\lambda + a_\lambda B_\lambda}{1 + a_\lambda} \text{ and } c_\lambda^\beta \in \mathbb{R}$$

(Figure 3). Notice that in view of Corollary 2.3 the index β is unique, and $c_\lambda^\beta x_1^{a_\lambda}$ is the principal root of f_p .

We then apply an algorithm due to Varchenko [8].

Step 1 (Figure 4). We apply the real change of variables $x = \varphi(y)$ given by $y_1 := x_1$, $y_2 := x_2 - c_\lambda^\beta x_1^{a_\lambda}$ and again put $\tilde{f} := f \circ \varphi$. In order to describe the effect of this change of variables to the Newton polyhedron, let us denote all quantities associated to \tilde{f} with a superscript $\tilde{\cdot}$.

Observe first that the roots \tilde{r} of \tilde{f} are of the form

$$(4.13) \quad \tilde{r}(y_1) = c_l^{\alpha_1} y_1^{a_l} - c_\lambda^{\beta} y_1^{a_\lambda} + c_{l_2}^{\alpha_1 \alpha_2} y_1^{a_{l_2}^{\alpha_1}} + \dots,$$

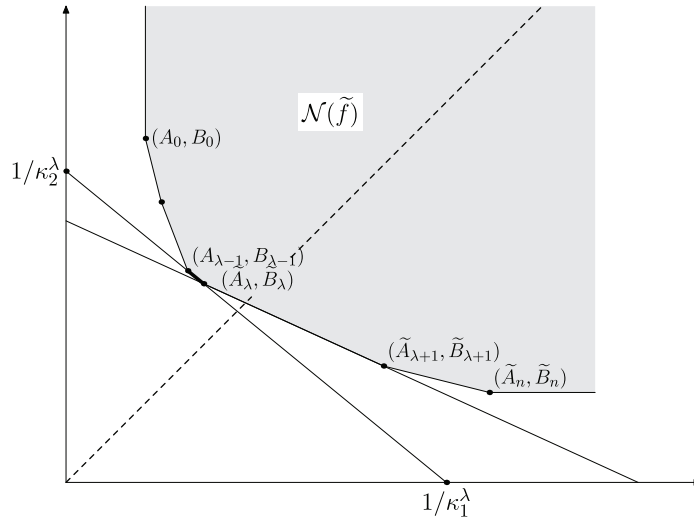


FIGURE 4

respectively $\tilde{r}(y_1) = -c_\lambda^\beta y_1^{a_\lambda}$ (namely those corresponding to the trivial roots of f). This shows that the leading exponents

$$\tilde{a}_1 < \tilde{a}_2 < \dots < \tilde{a}_l < \dots$$

of these roots are given by $\tilde{a}_l = a_l$ for $l < \lambda$, with the same multiplicities $N \left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix} \right] = N \left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix} \right]$ (where $\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix} \right]$ denotes the cluster of roots \tilde{r} associated to the index l).

For the vertices $(\tilde{A}_l, \tilde{B}_l)$ of $\mathcal{N}(\tilde{f})$ we thus obtain from (4.2), (4.3) that

$$(4.14) \quad (\tilde{A}_l, \tilde{B}_l) = (A_l, B_l) \quad \text{for } l < \lambda.$$

Moreover, any root r belonging to a cluster $\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix} \right]$ with $l > \lambda$ is transformed into a root \tilde{r} with the leading exponent a_λ . Finally, if r belongs to $\left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right]$, then either the leading exponent of \tilde{r} is a_λ (namely if $\alpha_1 \neq \beta$) or it is of the form $a_{\lambda l_2}^\beta > a_\lambda$. We therefore distinguish two cases.

Case 1. $\nu_2 + B_\lambda + \left(N \left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right] - N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right] \right) > 0.$

Then there exists at least one root \tilde{r} with the leading exponent a_λ , so that $\tilde{a}_\lambda = a_\lambda$. Moreover, since \tilde{B}_λ is the number of roots \tilde{r} with leading exponent strictly greater than a_λ , we see that $\tilde{B}_\lambda = N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right]$. Similarly, the number $N \left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right]$ is the same as the number of roots r with leading exponent strictly greater than a_λ or equal to a_λ , but then with index $\alpha_1 \neq \beta$; hence $N \left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right] = B_{\lambda-1} - N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right]$. This implies

$\tilde{A}_\lambda = \tilde{A}_{\lambda-1} + \tilde{a}_\lambda N \left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right] = A_{\lambda-1} + a_\lambda (B_{\lambda-1} - N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right]) = A_\lambda + a_\lambda B_\lambda - a_\lambda N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right]$. In combination, we thus have

$$(4.15) \quad (\tilde{A}_\lambda, \tilde{B}_\lambda) = \left(A_\lambda + a_\lambda B_\lambda - a_\lambda N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right], N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right] \right).$$

But, estimate (4.12) is equivalent to

$$\tilde{A}_\lambda < \tilde{B}_\lambda,$$

so that the edge $[(\tilde{A}_{\lambda-1}, \tilde{B}_{\lambda-1}), (\tilde{A}_\lambda, \tilde{B}_\lambda)]$, which lies on the same line as the edge $\pi(f) = [(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$ and has the same left vertex, is not the principal face of the Newton polyhedron of \tilde{f} . Finally, it is evident from (4.13) that in this case $\tilde{a}_{\lambda+k} = a_{\lambda k}^\beta$ if $k > 0$ (unless there is no cluster $\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix} \right]$ with $l > \lambda$).

This shows that in this case the principal face of $\mathcal{N}(\tilde{f})$ is either associated to a cluster of roots \tilde{r} which corresponds to a cluster of roots $\left[\begin{smallmatrix} \beta & \cdot \\ \lambda & \lambda_2 \end{smallmatrix} \right]$ in the original coordinates or is a horizontal, unbounded edge (so that the new coordinates are adapted).

Case 2. $\nu_2 + B_\lambda + \left(N \left[\begin{smallmatrix} \cdot \\ \lambda \end{smallmatrix} \right] - N \left[\begin{smallmatrix} \beta \\ \lambda \end{smallmatrix} \right] \right) = 0.$

Then there is no root \tilde{r} with the leading exponent a_λ , and so again the conclusion stated at the end of the previous case applies.

In both cases we see that the principal face of the new Newton polyhedron $\mathcal{N}(\tilde{f})$ will be less steep than the one of $\mathcal{N}(f)$, so that $d(\tilde{f}) > d(f)$.

Subsequent steps. Now, either the new coordinates y are adapted, in which case we are finished, or we can apply the same procedure to \tilde{f} . Composing the change of coordinates from the first step with the one from the second step, we see that we then can find a change of coordinates $x = \varphi_{(2)}(y)$ of the form

$$x_1 := y_1, \quad x_2 := y_2 - (c_\lambda^\beta x_1^{a_\lambda} + c_{\lambda\lambda_2}^{\beta\beta_2} y_1^{a_{\lambda\lambda_2}^\beta}),$$

with $a_{\lambda\lambda_2}^\beta \in \mathbb{N}$ and $c_{\lambda\lambda_2}^{\beta\beta_2} \in \mathbb{R}$, such that the following holds: If the function $f_{(2)} := f \circ \varphi_{(2)}$ expresses the function f in the new coordinates, then the principal face of the Newton polyhedron of $f_{(2)}$ is either associated to a cluster of roots which corresponds to a cluster of roots $\left[\begin{smallmatrix} \beta & \beta_2 & \cdot \\ \lambda & \lambda_2 & \lambda_3 \end{smallmatrix} \right]$ in the original coordinates or is a horizontal, unbounded edge (so that the new coordinates are adapted).

Now, if we iterate this procedure, then either this procedure will stop after finitely many steps or it will continue infinitely. If it stops, it is clear that we will have arrived at a new, adapted coordinate system of the form

$$x_1 := y_1, \quad x_2 := y_2 - (c_\lambda^\beta y_1^{a_\lambda} + \dots + c_{\lambda\lambda_2 \dots \lambda_p}^{\beta\beta_2 \dots \beta_p} y_1^{a_{\lambda\lambda_2 \dots \lambda_p}^{\beta\beta_1 \dots \beta_{p-1}}}),$$

and Theorem 4.2 holds with a polynomial function $\psi(y_1)$.

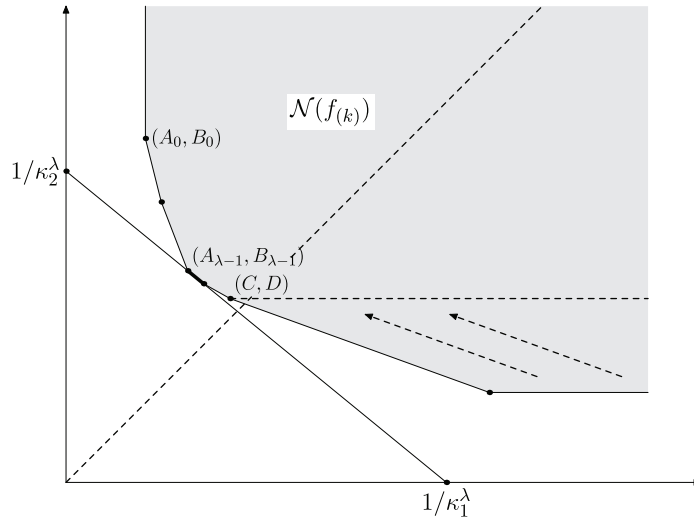


FIGURE 5

Final step (Figures 5, 6). Assume that the procedure does not terminate. Then, in each further step we pass to a new Newton diagram with a principal face associated to a cluster of roots (in the original coordinates) $\begin{bmatrix} \beta & \beta_2 & \cdots & \cdot \\ \lambda & \lambda_2 & \cdots & \lambda_{k+1} \end{bmatrix}$, which is a sub-cluster of the previous cluster $\begin{bmatrix} \beta & \beta_2 & \cdots & \cdot \\ \lambda & \lambda_2 & \cdots & \lambda_k \end{bmatrix}$. In particular, the corresponding multiplicities $N_k := N \begin{bmatrix} \beta & \cdots & \cdot \\ \lambda & \cdots & \lambda_{k+1} \end{bmatrix}$ form a decreasing sequence, which eventually must become constant. Choose $N, k_0 \in \mathbb{N}$ such that $N_k = N$ for every $k \geq k_0$. Replacing the original coordinate system by the one obtained in the k_0 -th step, we may assume without loss of generality that $k_0 = 0$, i.e.,

$$(4.16) \quad N = N \begin{bmatrix} \cdot \\ \lambda \end{bmatrix} = N \begin{bmatrix} \beta & \cdot \\ \lambda & \lambda_2 \end{bmatrix} = \cdots = N \begin{bmatrix} \beta & \beta_2 & \cdots & \cdot \\ \lambda & \lambda_2 & \cdots & \lambda_{k+1} \end{bmatrix} = \cdots$$

This clearly implies that $N \begin{bmatrix} \beta & \beta_2 & \cdots & \cdot \\ \lambda & \lambda_2 & \cdots & \lambda_{k+1} \end{bmatrix} = N \begin{bmatrix} \beta & \beta_2 & \cdots & \beta_{k+1} \\ \lambda & \lambda_2 & \cdots & \lambda_{k+1} \end{bmatrix}$ for every k and that each of the clusters $\begin{bmatrix} \beta & \beta_2 & \cdots & \cdot \\ \lambda & \lambda_2 & \cdots & \lambda_{k+1} \end{bmatrix}$ contains exactly one and the same root (of multiplicity N), namely

$$(4.17) \quad \rho(x_1) := c_\lambda^\beta x_1^{a_\lambda} + \cdots + c_{\lambda\lambda_2 \cdots \lambda_{k+1}}^{\beta\beta_2 \cdots \beta_{k+1}} x_1^{a_{\lambda\lambda_2 \cdots \lambda_{k+1}}^{\beta\beta_1 \cdots \beta_k}} + \cdots$$

(so that in fact $\lambda_k = \beta_k = 1$ for every $k \geq 2$). Moreover, our procedure shows that all coefficients in this series must be real and all exponents positive integers, so that $\rho(x_1)$ is a real valued, real-analytic function of x_1 .

If we apply our first change of coordinates $y_1 := x_1, y_2 := x_2 - c_\lambda^\beta x_1^{a_\lambda}$ in this situation, then we see that the leading exponents of the new roots \tilde{r} are given by

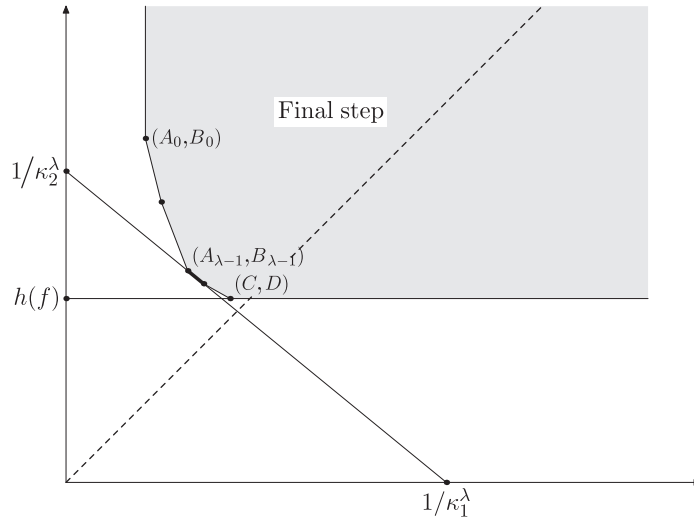


FIGURE 6

$\tilde{a}_1 = a_1, \dots, \tilde{a}_\lambda = a_\lambda$ and $\tilde{a}_{\lambda+1} = a_{\lambda\lambda_2}^\beta$ in Case 1, and by $\tilde{a}_1 = a_1, \dots, \tilde{a}_{\lambda-1} = a_{\lambda-1}$ and $\tilde{a}_\lambda = a_{\lambda\lambda_2}^\beta$ in Case 2. Moreover, it is clear from our discussion that the last edge, associated to the cluster of roots with the biggest leading exponent $a_{\lambda\lambda_2}^\beta$, must be the principal edge. Moreover, in the new coordinates y , the function \tilde{f} can have no vanishing root, since such a root would have corresponded to a root $c_\lambda^\beta x_1^{a_\lambda}$, which, by (4.17), cannot exist in the cluster $\begin{bmatrix} \cdot \\ \lambda \end{bmatrix}$.

By passing to these new coordinates, we may therefore assume in addition to (4.16) that the cluster $\begin{bmatrix} \cdot \\ \lambda \end{bmatrix}$ is the cluster associated to the principal face of the Newton polyhedron of f and that $\nu_2 = 0$, i.e., that

$$(4.18) \quad f_p(x_1, x_2) = cx_1^{\nu_1} (x_2 - c_\lambda^\beta x_1^{a_\lambda})^N.$$

The Newton polyhedron of f thus has vertices $(A_0, B_0), \dots, (A_\lambda, B_\lambda)$, and the principal edge is given by $[(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$, where, according to (4.2), (4.16),

$$(4.19) \quad A_{\lambda-1} < B_{\lambda-1} = N \quad \text{and} \quad B_\lambda = 0.$$

Let us finally apply the change of coordinates $y_1 := x_1, y_2 := x_2 - \rho(x_1)$. The non-zero roots \tilde{r} of \tilde{f} are then given by $\tilde{r} = r - \rho$, with $r \in \begin{bmatrix} \cdot \\ l \end{bmatrix}$ for some $l < \lambda$, and they have the same multiplicities and leading exponents as the corresponding roots r . In view of (4.2), the vertices of the Newton diagram $\mathcal{N}(\tilde{f})$ are thus given by the points $(A_0, B_0), \dots, (A_{\lambda-1}, B_{\lambda-1})$, i.e., the effect of the change of coordinates on the Newton diagram is the removal of the last vertex (A_λ, B_λ) . Since, by (4.19), the vertex $(A_{\lambda-1}, B_{\lambda-1})$ lies above the bisectrix, we see that the principal face of $\mathcal{N}(\tilde{f})$ is the horizontal half-line emerging from the point $(A_{\lambda-1}, B_{\lambda-1})$ along the

line $t_2 = N$, a non-compact set. According to Theorem 3.3, the coordinates y are thus adapted to f and the height $h(f)$ is given by

$$(4.20) \quad h(f) = N.$$

□

The proof suggests the following definitions. If

$$r(x_1) = c_{l_1}^{\alpha_1} x_1^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2} x_1^{a_{l_1 l_2}^{\alpha_1}} + \dots + c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_p} x_1^{a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}}} + \dots$$

is any root of f (more precisely, of F), then any leading part

$$\sum_{p=1}^K c_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_p} x_1^{a_{l_1 \dots l_p}^{\alpha_1 \dots \alpha_{p-1}}},$$

with $1 \leq K \leq \infty$, which is real analytic, i.e, where all exponents appearing in this series are positive integers, will be called an *analytic root jet*.

We have seen that the function ψ constructed by Varchenko’s algorithm in a unique way is indeed an analytic root jet, which we call the *principal root jet*.

Our proof even reveals that the conditions (a)-(c) in Theorem 3.3 are necessary and sufficient for the adaptedness of the given coordinate system for arbitrary analytic functions f . In the statements of the following corollaries we shall always make the following

General Assumptions. *The function $f(x_1, x_2)$ is a real valued real-analytic function defined on a neighborhood of the origin in \mathbb{R}^2 with $f(0, 0) = 0, \nabla f(0, 0) = 0$. Choose $\kappa_1, \kappa_2 \geq 0$ such that the principal face $\pi(f)$ of the Newton polyhedron of f lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$, and assume that $\kappa_2 \geq \kappa_1$.*

Corollary 4.3. *The given coordinates (x_1, x_2) are not adapted to f if and only if the conditions (a)-(c) in Theorem 3.3 are satisfied. In particular, the given coordinates are adapted to f if and only if they are adapted to the principal part f_p of f .*

Moreover, we always have $h(f) \leq h(f_p)$.

Proof. The necessity of these conditions for non-adaptedness has been proved in Theorem 3.3. Assume conversely that (a)-(c) hold true. In that case, we have seen in the proof of Theorem 4.2 that there exists a change of coordinates which strictly increases the height, so that the original coordinates are not adapted.

Since the conditions (a)-(c) depend in fact only on f_p , we see in particular that the given coordinates are adapted to f if and only if they are adapted to the principal part f_p of f .

Finally, if the coordinates (x_1, x_2) are adapted to f (hence also to f_p), then we clearly have $h(f) = h(f_p)$. Otherwise, the proof of Theorem 4.2 shows that the first change of coordinates $y_1 := x_1, y_2 := x_2 - c_\lambda^\beta x_1^\lambda$ that we considered reduces the principal edge $\pi(f) = [(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$ to the shorter interval (possibly of length zero) $[(A_{\lambda-1}, B_{\lambda-1}), (\tilde{A}_\lambda, \tilde{B}_\lambda)]$ on the same line, but lying above the bisectrix. Notice that the coordinates (y_1, y_2) are already adapted to the principal part f_p and that, by (4.15), (4.12) and Theorem 3.3, we have $\tilde{B}_\lambda = N \begin{bmatrix} \beta \\ \lambda \end{bmatrix} = m(f_p) = h(f_p)$.

But the point $(\tilde{A}_\lambda, \tilde{B}_\lambda)$ will be contained in the Newton diagrams of f associated to all subsequent systems of coordinates that we constructed by our algorithm (compare (4.14), applied to the coordinates y), which means that the principal face

of f in our final, adapted coordinate system must lie in the half-space $t_2 \leq h(f_p)$, so that $h(f) \leq h(f_p)$. \square

Corollary 4.4. (a) *We can always find a change of coordinates $x = \varphi(y)$ at 0 of the form $y_1 := x_1, y_2 := x_2 - \psi(x_1)$, such that the coordinates (y_1, y_2) are adapted to $\tilde{f} := f \circ \varphi$ and the following hold true:*

If the principal face $\pi(\tilde{f})$ is compact and lies on the line $\tilde{\kappa}_1 t_1 + \tilde{\kappa}_2 t_2 = 1$, with $\tilde{\kappa}_1 \leq \tilde{\kappa}_2$, then ψ is a polynomial of degree strictly less than $\tilde{\kappa}_2/\tilde{\kappa}_1$.

This applies, in particular, if the height $h(f)$ of f is a non-integer rational number.

(b) *There always exists a change of coordinates $x = \varphi(y)$ of the form $y_1 := x_1, y_2 := x_2 - \eta(x_1)$ at the origin, with a polynomial function $\eta(y_1)$, such that in the new coordinates y we have $h(f) = h(\tilde{f}) = h(\tilde{f}_p)$. Here, we have again put $\tilde{f} := f \circ \varphi$.*

Proof. Indeed, the algorithm that we devised in the proof of Theorem 4.2 in order to construct an adapted coordinate system shows that we can arrive at an adapted coordinate system with a polynomial function ψ , unless we have to choose for ψ one of the roots r with infinitely many non-trivial terms in its Puiseux series expansion. In the latter case, the principal face in the adapted coordinate system that we constructed is non-compact and the height is an integer, as we have seen. Thus, if the principal face in the adapted coordinate system is compact, the algorithm must terminate after a finite number of steps. Also, in Step 1, the degree of the polynomial used in the change of coordinates is given by a_λ , where by (4.10) $a_\lambda = \kappa_2/\kappa_1$ is just the inverse of the slope of the principal edge of the Newton polyhedron of f . However, the proof shows that the slope of the principal face strictly decreases by the change of coordinates in Step 1, and the same applies to all subsequent steps. If we apply this to the last change of coordinates before achieving adapted coordinates, we see that the function ψ is a polynomial of degree $m < \tilde{\kappa}_2/\tilde{\kappa}_1$. This proves (a).

Moreover, in the case where our algorithm does not terminate after finitely steps in our algorithm (which all consist of polynomial changes of coordinates), we may assume that the polynomial \tilde{f}_p corresponding to the principal face of the Newton diagram has a unique root of multiplicity $N = h(f)$ (compare (4.16), (4.17) and (4.20)). If we choose the coordinates y which we obtain at this stage, we then have $h(\tilde{f}_p) = m(\tilde{f}_p) = N = h(f) = h(\tilde{f})$, so that (b) is also proven. \square

5. THE SMOOTH CASE

We shall finally extend Theorem 4.2 to the smooth setting.

Theorem 5.1. *Let f be a real valued smooth function of finite type defined on a neighborhood of the origin in \mathbb{R}^2 with $f(0, 0) = 0, \nabla f(0, 0) = 0$. Choose $\kappa_1, \kappa_2 \geq 0$ such that the principal face $\pi(f)$ of the Newton polyhedron of f lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$. Without loss of generality, we may assume that $\kappa_2 \geq \kappa_1$. Then there exists a smooth function $\psi(x_1)$ of x_1 near the origin with $\psi(0) = 0$ such that an adapted coordinate system (y_1, y_2) for f near 0 is given by $y_1 := x_1, y_2 := x_2 - \psi(x_1)$.*

Proof. We shall proceed in a very similar way as in the analytic setting, again following the idea of Varchenko’s algorithm [8] and adopting the same notation as before.

If the coordinates are adapted to f , then we may choose $\psi := 0$ and we are finished.

Otherwise, again by Theorem 3.3, the principal face $\pi(f)$ is a compact edge. Moreover, we have $\kappa_2/\kappa_1 =: m_1 \in \mathbb{N}$ and $m(f_p) = m(f_\kappa) > d(f)$. Let us then choose a real root $x \mapsto b_1x_1^{m_1}$ of the principal part f_p of f of maximal multiplicity $N_0 := m(f_p)$, i.e., the principal root. Then $b_1 \neq 0$, again by Theorem 3.3,

Step 1. We apply the real change of variables $x = \varphi(y)$ given by $y_1 := x_1, y_2 := x_2 - b_1x_1^{m_1}$, and put $\tilde{f} := f \circ \varphi$. Let us again endow all quantities associated to \tilde{f} with a superscript $\tilde{\cdot}$. Now, if the coordinates y are adapted to f , we choose $\psi(x_1) := b_1x_1^{m_1}$ and we are finished.

Otherwise, the principal face $\pi(\tilde{f})$ is a compact edge, and we have $\tilde{\kappa}_2/\tilde{\kappa}_1 =: m_2 \in \mathbb{N}$ and $N_1 := m(\tilde{f}_p) > d(\tilde{f})$. Recall that the principal part \tilde{f}_p of \tilde{f} is $\tilde{\kappa}$ -homogeneous of degree one. We claim that

$$(5.1) \quad m_2 > m_1 \quad \text{and} \quad N_1 \leq N_0.$$

Indeed, recall that the effect of our change of coordinates φ on the Newton polyhedron is such that it preserves all lines $\kappa_1t_1 + \kappa_2t_2 = c$. Let us therefore choose $m \in \mathbb{N}$ so big that we have $j + k \leq m$ for every point (j, k) lying on any such line passing through any point in the Newton diagram $\mathcal{N}_d(f)$ of f , and denote by F the Taylor polynomial of order m of f . Then it is clear that f and F have the same principal faces and parts, and the same applies to \tilde{f} and $\tilde{F} := F \circ \varphi$. I.e., we have $f_p = F_p$ and $\tilde{f}_p = \tilde{F}_p$. We can therefore apply our results for the analytic case to the polynomial function F and obtain (5.1).

This argument also shows that the change of coordinates increases the distance, i.e., $d(\tilde{f}) > d(f)$.

Subsequent steps. Now, either the new coordinates y are adapted, in which case we are finished, or we can apply the same procedure to \tilde{f} , etc. In this way, we obtain a sequence of functions $f_{(k)} = f \circ \varphi_{(k)}$ with $f_{(0)} := f$ and $f_{(k+1)} := \tilde{f}_{(k)}$, which can be obtained from the original coordinates x by means of a change of coordinates $x = \varphi_{(k)}(y)$ of the form

$$y_1 := x_1, y_2 := x_2 - \sum_{l=1}^k b_l x_1^{m_l},$$

with positive integers $m_1 < m_2 < \dots < m_k < m_{k+1} < \dots$ and real coefficients $b_l \neq 0$. Moreover, if $N_k := m((f_{(k)})_p)$ denotes the maximal order of vanishing of the principal part of $f_{(k)}$ along the unit circle S^1 , then we have

$$N_0 \geq N_1 \geq \dots \geq N_k \geq N_{k+1} \geq \dots.$$

Either this procedure will stop after finitely many steps or it will continue infinitely. If it stops, say, at the k -th step, it is clear that we will have arrived at an adapted coordinate system $x = \varphi_{(k)}(y)$, with a polynomial function $\psi(x_1) = \sum_{l=1}^k b_l x_1^{m_l}$.

Final step. Assume that the procedure does not terminate. Since the maximal multiplicities N_k of the roots of the principal part of $f_{(k)}$ form a decreasing sequence, we again find some $k_0, N \in \mathbb{N}$ such that $N_k = N$ for every $k \geq k_0$. By comparing the effect of the change of coordinates in each step of order $k \geq k_0$ with the effect

on the Taylor polynomial of sufficiently high degree, we see from the corresponding result (4.18) in the analytic case that the principal part of $f_{(k)}$ is of the form

$$(f_{(k)})_p(x) = c_k x_1^{\nu_1} (x_2 - b_{k+1} x_1^{m_{k+1}})^N,$$

where $\nu_1 < N$. It is mixed homogeneous of degree one with respect to the weight $\kappa^{(k)}$, given by

$$\kappa_1^{(k)} := \frac{1}{\nu_1 + N m_{k+1}}, \quad \kappa_2^{(k)} := \frac{m_{k+1}}{\nu_1 + N m_{k+1}},$$

so that, for $k \geq k_0$,

$$(5.2) \quad f_{(k)}(x) = c_k x_1^{\nu_1} (x_2 - b_{k+1} x_1^{m_{k+1}})^N + \text{terms of higher } \kappa^{(k)\text{-degree.}}$$

Now, according to a classical lemma of Borel (cf. [3], Theorem 1.2.6), we can find a smooth function $\rho(x_1)$ near the origin whose Taylor series is the formal series $\sum_{l=1}^{\infty} b_l x_1^{m_l}$. Consider the smooth change of coordinates $x := \varphi(y)$ given by $y_1 := x_1, y_2 := x_2 - \rho(x_1)$, and put $\tilde{f} := f \circ \varphi$. We claim that the coordinates y are adapted to \tilde{f} .

Indeed, we have

$$\tilde{f}(y) = f_{(k)} \circ (\varphi_{(k)}^{-1} \circ \varphi)(y) = f_{(k)} \left(y_1, y_2 + (\rho(y_1) - \sum_{l=1}^k b_l y_1^{m_l}) \right),$$

where $\rho(y_1) - \sum_{l=1}^k b_l y_1^{m_l}$ has the Taylor series $\sum_{l=k+1}^{\infty} b_l y_1^{m_l}$. In view of (5.2), this shows that

$$(5.3) \quad \tilde{f}(y) = c_k y_1^{\nu_1} y_2^N + R(y),$$

where R is a smooth function consisting of terms of $\kappa^{(k)}$ -degree strictly larger than one, for every $k \geq k_0$; i.e.,

$$(5.4) \quad \frac{1}{\nu_1 + N m_{k+1}} j_1 + \frac{m_{k+1}}{\nu_1 + N m_{k+1}} j_2 > 1 \quad \text{for every } (j_1, j_2) \in \mathcal{T}(R).$$

Since $m_k \rightarrow \infty$ as $k \rightarrow \infty$, this implies that $j_2 \geq N$. Moreover, if $j_2 = N$, then the left-hand side of (5.4) is bounded by 1 if $j_1 \leq \nu_1$, so that we must have $j_1 > \nu_1$. In combination, (5.3) and (5.4) show that the Newton polyhedron $\mathcal{N}(\tilde{f})$ of \tilde{f} contains the point (ν_1, N) , but no further point on the left to this point on the line $t_2 = N$, and that all other points of $\mathcal{N}(\tilde{f})$ are contained in the open half-plane above this line. Since $\nu_1 < N$, this shows that the principal face of $\mathcal{N}(\tilde{f})$ is the unbounded horizontal half-line given by $t_1 \geq \nu_1, t_2 = N$; hence the coordinates are adapted to \tilde{f} . \square

The following corollary is immediate from the proof of Theorem 5.1 and the proofs of Corollaries 4.3 and 4.4, which carry over to the smooth setting.

Corollary 5.2. *Let f be a real valued smooth function of finite type defined on a neighborhood of the origin in \mathbb{R}^2 with $f(0,0) = 0, \nabla f(0,0) = 0$. Then the height $h(f)$ is a rational number.*

Moreover, if we choose $\kappa_1, \kappa_2 \geq 0$ such that the principal face $\pi(f)$ of the Newton polyhedron of f lies on the line $\kappa_1 t_1 + \kappa_2 t_2 = 1$ and if we assume without loss of generality that $\kappa_2 \geq \kappa_1$, then Corollaries 4.3 and 4.4 remain true in this smooth setting.

ACKNOWLEDGEMENTS

The authors would like to thank D. H. Phong for informing them about the results on adapted coordinates in [6], Michael Kempe for his support in creating the graphics included in this article, and the referee for several helpful comments.

REFERENCES

- [1] V. I. Arnol'd, *Remarks on the method of stationary phase and on the Coxeter numbers*, Uspekhi Mat. Nauk. **28 (5)** (1973), 17-44. English transl. Russ. Math. Surv. **28 (5)** (1973), 19-58. MR0397777 (53:1635)
- [2] V. I. Arnol'd, C. M. Guseyn-zade, and A. N. Varchenko, *Singularities of Differentiable Maps*, Vol. II. Monodromy and asymptotic of Integrals. Nauka, Moscow (1984), English transl. Birhauser, Boston, Basel, Berlin, 1988. MR966191 (89g:58024)
- [3] L. 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. MR1996773
- [4] I. A. Ikromov, M. Kempe, and D. Müller, *Estimates for maximal functions associated with hypersurfaces in \mathbb{R}^3 and related problems of harmonic analysis*, Acta Math. **204** (2010), 151-271.
- [5] D. H. Phong and E. M. Stein, *The Newton polyhedron and oscillatory integral operators*, Acta Math. **179(1)** (1997), 105-152. MR1484770 (98j:42009)
- [6] D. H. Phong, E. M. Stein, and J. A. Sturm, *On the growth and stability of real-analytic functions*, Amer. J. Math. **121 (3)** (1999), 519-554. MR1738409 (2002a:58025)
- [7] E. M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. MR1232192 (95c:42002)
- [8] A. N. Varchenko, *Newton polyhedra and estimates of oscillating integrals*, Funct. Anal. and Appl. **10(3)** (1976), 175-196.

DEPARTMENT OF MATHEMATICS, SAMARKAND STATE UNIVERSITY, UNIVERSITY BOULEVARD 15,
703004, SAMARKAND, UZBEKISTAN
E-mail address: ikromov1@rambler.ru

MATHEMATISCHES SEMINAR, C.A.-UNIVERSITÄT KIEL, LUDEWIG-MEYN-STRASSE 4, D-24098
KIEL, GERMANY
E-mail address: mueller@math.uni-kiel.de
URL: <http://analysis.math.uni-kiel.de/mueller/>