ON THE LOCAL SUM CONJECTURE IN TWO DIMENSIONS

ROBERT FRASER AND JAMES WRIGHT

Abstract. The *local sum conjecture* is a variant of some of Igusa's questions on exponential sums put forward by Denef and Sperber in [7]. In a remarkable paper [6] by Cluckers, Mustata and Nguyen, this conjecture has been established in all dimensions, using sophisticated, powerful techniques from a research area blending algebraic geometry with ideas from logic. The purpose of this paper is to give an elementary proof of this conjecture in two dimensions which follows Varčenko's treatment of euclidean oscillatory integrals based on Newton polyhedra for good coordinate choices. Another elementary proof is given by Veys [18] from an algebraic geometric perspective.

1. Introduction

In their seminal paper [7], Denef and Sperber formulated the following *local sum conjecture*. Let $\phi \in \mathbb{Z}[X_1, \dots, X_n]$ and consider the *local* exponential sum

$$S_0 = S_0(\phi, p^s) \ := \ \frac{1}{p^{sn}} \sum_{\substack{x \in [\mathbb{Z}/p^s\mathbb{Z}]^n \\ x \equiv 0 \pmod{p}}} e^{2\pi i \phi(x)/p^s},$$

the local sum being a truncation of the complete exponential sum

$$S = S(\phi, p^s) \ := \ \frac{1}{p^{sn}} \sum_{x \in [\mathbb{Z}/p^s\mathbb{Z}]^n} e^{2\pi i \phi(x)/p^s}$$

which selects the terms $x = (x_1, \dots, x_n)$ where $p \,|x_j$ for all $1 \le j \le n$. The conjecture postulates that there exists a constant C, independent of p and s, and a finite set $\mathcal{P} = \mathcal{P}_\phi$ of primes such that for all $p \notin \mathcal{P}$,

$$|S_0| \ \le \ C s^{n-1} p^{-\sigma_c s} \tag{1}$$

where $\sigma_c = \sigma_c(\phi)$ is the complex oscillation index[1] of ϕ at 0. We will recall the precise definitions for this and other notions in Section 3. The conjecture (1) is related to one of the Igusa conjectures on exponential sums which posits similar uniform bounds for S when ϕ is any homogeneous polynomial.

In [7], Denef and Sperber proved (1) when ϕ is $\mathbb{C}$-nondegenerate and when an auxilary condition on the vertices of Newton polyhedron holds (in [5], Cluckers removed this auxiliary condition). The notion of $\mathbb{C}$-nondegenerate was introduced in [11] and will be defined in Section 3. In this case the complex oscillation index

[1]In the literature, oscillation indices tend to be defined as negative numbers. We will consider their absolute values and define them as positive numbers.

$\sigma_c(\phi) = 1/d(\phi)$ is the reciprocal of the *Newton distance* $d(\phi)$ of ϕ, see [7]. In the same paper, Denef and Sperber also established the Igusa conjecture under the same hypotheses; in [4], Cluckers removed the auxiliary condition on the vertices, establishing the Igusa conjecture when ϕ is $\mathbb{C}$-nondegenerate.

As mentioned in the abstract, Cluckers, Mustata and Nguyen [6] established the local sum conjecture (1) in all dimensions and much more; they also established the Igusa conjecture for complete exponential sums $S(\phi, p^s)$ where ϕ is a general homogeneous polynomial, where the exponent is replaced by the log-canonical threshold of ϕ. In two dimensions, this will always be equal to the complex oscillation index. Earlier, the Igusa conjecture was established in two dimensions [21] by an elementary argument. Afterwards, Lichtin [12] gave an alternative proof of the two dimensional Igusa conjecture from a different perspective. Albarracín-Mantilla and León-Cardenal [2] gave a detailed description of the behaviour of the non-truncated sum S in two dimensions based on the poles of the corresponding local zeta function under an additional nondegeneracy condition.

One difficulty in higher dimensions is that it is hard to get one's hands on the oscillation index $\sigma_c(\phi)$ in a precise way for general ϕ. For our arguments, it is essential that we work with an explicit description of σ_c. Although such a description is not available in higher dimensions, we can describe σ_c explicitly in two dimensions. To see this, consider the case when ϕ is homogeneous so that it can be factored

$$\phi(x,y) \;=\; c\, x^{\alpha} y^{\beta} \prod_{j=1}^{N} (y - \xi_j x)^{n_j} \tag{2}$$

for some roots $\{\xi_j\}_{j=1}^N$ lying in $\mathbb{Q}^{\mathrm{alg}}$; see (11) in Section 4. It can be shown that $\sigma_c(\phi) = 1/\max(m_{\mathbb{Q}}, d/2)$ where d is the degree of the polynomial ϕ and $m_{\mathbb{Q}} = \max(\alpha, \beta, \{n_j : \xi_j \in \mathbb{Q}\})$; see [21] and Section 3.

This gives an explicit description of the oscillation index for general homogeneous ϕ in two variables. Below we will see that a more involved description can be made for general polynomials ϕ in two variables. Such a concrete description is not available in higher dimensions. Furthermore one easily sees that ϕ is $\mathbb{C}$-*degenerate* precisely when there is at least one root ξ_j with multiplicity n_j larger than one. This provides many examples of degenerate homogeneous polynomials.

As mentioned above, when ϕ is $\mathbb{C}$-nondegenerate, the oscillation index $\sigma_c(\phi)$ is given in terms of the Newton distance $d(\phi)$, a quantity we can easily compute. Nevertheless, in two dimensions, we can get our hands on the oscillation index since it is known[2] that $\sigma_c(\phi) = 1/h(\phi)$ where $h(\phi) := \sup_z d_z(\phi)$ is the so-called *height* of ϕ. Here the supremum is taken over all local coordinate systems $z = (x, y)$ of the origin (real-analytic coordinate systems if the phase ϕ is real-analytic and smooth coordinate systems if ϕ is smooth) and $d_z(\phi)$ denotes the Newton distance of ϕ in the coordinates z.

[2]This equation is usually stated for the real oscillation index σ_r. In two dimensions, we have that $\sigma_r = \sigma_c$; see Section 3.

In two dimemsions, the supremum $\sup_z d_z(\phi) = d_{z_0}(\phi)$ is attained in the definition of the height $h(\phi)$; any such coordinate system z_0 is called *adapted*. The height $h(\phi)$ gives us the exact decay rate for the corresponding euclidean oscillatory integral

$$I_\psi(\phi,\lambda) \;:=\; \int_{\mathbb{R}^2} e^{2\pi i\lambda\phi(x)}\psi(x)\,dx$$

where ϕ is a smooth, real-valued function, λ is a large real parameter and $\psi \in C_c^\infty$. When ϕ is of finite type, we have for some constant C_ϕ depending on ϕ but not on λ

$$|I_\psi(\phi,\lambda)| \;\le\; C_\phi\,\log^\nu(|\lambda|)|\lambda|^{-1/h(\phi)} \tag{3}$$

for large λ and all $\psi \in C^\infty$ supported in a sufficiently small neighbourhood of 0. Here $\nu(\phi) \in \{0,1\}$ is the so-called *Varčenko's exponent* (also known as the multiplicity of the oscillation index). We emphasise that in higher dimensions the Varčenko exponent $\nu(\phi)$ lies in $\{0,1,\dots,n-1\}$. Furthermore,

$$\lim_{\lambda\to\infty} \frac{\lambda^{1/h}}{\log^\nu(\lambda)}\, I_\psi(\phi,\lambda) \;=\; c\,\psi(0) \tag{4}$$

where $c = c_\phi$ is nonzero.[3] In this generality, the results in (3) and (4) were established by Ikromov and Müller in [8] and [9]. Their work was influential in our analysis establishing the following.

Theorem 1.1. *Let $\phi \in \mathbb{Z}|X,Y]$. Then there exists a finite set $\mathcal{P}$ of primes and a constant $C = C_\phi$ such that for any $p \notin \mathcal{P}$ and $s \ge 1$,*

$$|S_0(\phi,p^s)| \;\le\; C\; s^{\nu(\phi)}p^{-s/h(\phi)} \tag{5}$$

holds for all ϕ except for an exceptional class $\mathcal{E}$. For $\phi \in \mathcal{E}$, the estimate (5) *holds with $\nu = 1$; that is $|S_0(\phi,p^s)| \le Csp^{-s/h(\phi)}$ holds for $\phi \in \mathcal{E}$.*

The class $\mathcal{E}$ consists of those polynomials of the form

$$\phi(x,y) \;=\; \quad a(by^2+cxy+dx^2)^m \;+\; \text{higher order terms}^4$$

where the quadratic polynomial $by^2+cxy+dx^2$ is irreducible over the rationals $\mathbb{Q}$. For example when $\phi(x,y) = a(x^2+y^2)^m$, we have $h(\phi) = m$ and $\nu(\phi) = 0$. However when $m \ge 2$ and $p \equiv 1 \bmod 4$, then $|S_0(\phi,p^s)| \sim sp^{-s/m}$ for infinitely many $s \ge 1$. Furthermore when $p \equiv 3 \bmod 4$, then $|S_0(\phi,p^s)| \sim p^{-s/m}$ for infinitely many $s \ge 1$. These calculations are not difficult; see for example [21] where more general bounds are derived.

The estimate (5) in Theorem 1.1 is a slight strengthening in the two dimensional case only in that the exponent $\nu(\phi)$ in (5) is more precise, determining exactly when it matches the euclidean case. More importantly, we establish (5) using elementary arguments, only basic p-adic analysis is used. A key step in our argument will follow ideas from Ikromov and Müller in [8] in the euclidean setting which in turn were inspired from the arguments developed in [14] which gave an elementary

[3]The existence of this limit is proved under the additional condition that the principal face of ϕ in adapted coordinates is compact.

[4]We will describe this class precisely in section 3.

treatment of Karpushkin's work [10] on stability bounds of euclidean oscillatory integral estimates in two dimensions.

The main effort in this paper is to rework euclidean arguments in the local field setting. Basic euclidean arguments for estimating oscillatory integrals rely heavily on the order structure of the reals (in applications of the mean value and intermediate value theorems which are implicitly used in integration by parts arguments). We need to readdress these arguments, relying more on rudimentary sublevel set estimates (bounds for the number of solutions to polynomial congruences) in place of integration by parts arguments. These sublevel set bounds will be derived from a higher order Hensel lemma and so matters are kept on an elementary level.

In our argument, we will decompose the sum $S_0(\phi, p^s)$ in the same manner as Denef and Sperber [7]. This decomposition partitions the sum $S_0(\phi, p^s)$ into components corresponding to each compact face of the Newton diagram of ϕ. We will obtain a bound for each component of the sum. This bound will be sufficient to prove Theorem 1.1 provided that the coordinate system is chosen appropriately.

Assuming this coordinate change has been applied, we can describe the finite set $\mathcal{P}$ as follows. For each compact face τ of the Newton polyhedron of ϕ, there is a natural associated part of ϕ, a quasi-homogeneous polynomial ϕ_τ (see Section 3). Let $\mathcal{A}$ denote the collection of algebraic numbers consisting of the roots ξ as well as the differences of distinct roots $\xi_j - \xi_k$ of each ϕ_τ. The set $\mathcal{P}$ consists of

- prime numbers that are at most $\deg \phi$
- prime divisors of the coefficients of ϕ
- prime divisors of the coefficients of the minimal polynomials of each $\xi \in \mathcal{A}$.

Notation. All constants $C, c, c_0 > 0$ throughout this paper will depend only on the polynomial ϕ, although the values of these constants may change from line to line. Often it will be convenient to suppress explicitly mentioning the constants C or c in these inequalities and we will use the notation $A \lesssim B$ between positive quantities A and B to denote the inequality $A \leq CB$ (we will also denote this as $A = O(B)$). When we want to emphasise the dependence of the implicit constant in $A \lesssim B$ on a parameter k, we write $A \lesssim_k B$ to denote $A \leq C_k B$. Finally we use the notation $A \sim B$ to denote that both inequalities $A \lesssim B$ and $B \lesssim A$ hold.

Acknowledgements. We would like to thank Allan Greenleaf and Malabika Pramanik for informative and enlightening discussions about oscillatory integrals.

We are grateful to the anonymous referee both for their speedy review and for their very useful comments and suggestions, which greatly improved the final version of the paper.

We would like to thank Dimitrios Karslidis for providing a translation of Theorem 4.4 of [17].

This material is based upon work supported by the National Science Foundation under Award No. 1803086.

2. Outline of the paper

In the next section we will define precisely the various notions introduced above, including reviewing the Newton polyhedron, diagram and distance of a polynomial. We will also give a quick review of the required p-adic analysis that we will use and show how we can lift our exponential sum S_0 over $\mathbb{Z}/p^s\mathbb{Z}$ to an oscillatory integral over the p-adic field. This will illustrate the close analogy between these kinds of exponential sums and euclidean oscillatory integrals. In Sections 4 and 5 we will derive a basic bound for $S_0(\phi, p^s)$ which will imply (5) in Theorem 1.1 when the coordinates $z = (x, y)$ of our given polynomial $\phi(x, y)$ are *adapted*.

This basic bound will employ a useful estimate for exponential sums in one variable which depends on a generalisation of the classical Hensel lemma. We will outline the proof of this one dimensional bound in Section 7.

The main effort then will be to find a change of variables to put our polynomial ϕ into adapted coordinates. In general the change of variables that accomplishes this will be analytic. Attempting to keep our analysis on an elementary level, we will find a polynomial change of variables

$$p(x, y) \;=\; (p_1(x, y), p_2(x, y)) \in \mathbb{Q}[X, Y]$$

so that the new phase $\tilde{\phi}(x, y) = \phi(p(x, y))$ will be a polynomial with rational coefficients. The polynomial $\tilde{\phi}(x, y)$ will not quite be in adapted coordinates but nevertheless the bound established in Sections 4 and 5 will be sufficient to prove Theorem 1.1.

To produce this change of variables, we will follow an algorithm due to Ikromov and Müller [8] in the euclidean setting. They, in turn, blend ideas from two different algorithms due to Varčenko [16] and Phong, Stein and Sturm [14]. This will be carried out in Section 6. The algorithm producing this change of variables with rational coefficients is based on the clustering of the roots of ϕ which can be expressed in terms of Puiseux series.

3. Definitions and preliminaries

A good reference for the following basic results and definitions regarding oscillatory integrals can be found in [3].

Oscillation indices. Any polynomial $\phi \in \mathbb{Z}[X_1, \ldots, X_n]$ can be viewed as a real-valued phase and so oscillation indices make sense for ϕ. The *complex oscillation index* $\sigma_c(\phi)$ is defined as the supremum of β's where the bounds $|\int_\Gamma e^{2\pi i\lambda\phi(x)}dx| \leq C_\Gamma\lambda^{-\beta}$ hold for large $\lambda > 1$ and all n-dimensional chains Γ in a sufficiently small neighbourhood of 0 in $\mathbb{C}^n$, such that the imaginary part ϕ is strictly positive on

the boundary of Γ. The *real oscillation index* $\sigma_r(\phi)$ is defined as the supremum of β where the bounds $|\int_{\mathbb{R}^n} e^{2\pi i\lambda\phi(x)}\psi(x)dx| \leq C\lambda^{-\beta}$ hold for large $\lambda > 1$ and all smooth ψ supported in some neighbourhood of 0. In general these indices are difficult to compute. However when ϕ satisfies a certain nondegeneracy condition, then these numbers have a simple description.

The Newton polyhedron and diagram. To describe this nondegeneracy condition, we need to recall the definition of the Newton polyhedron of a polynomial ϕ; we will restrict ourselves to two dimensions although these notions make sense in any dimension. Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ include zero. For any polynomial $\phi(x, y) = \sum_{j,k} c_{j,k}x^j y^k$, we call the set $\mathcal{S}(\phi) := \{(j, k) \in \mathbb{N}^2 \setminus \{0\} : c_{j,k} \neq 0\}$, the *reduced* support of ϕ. The *Newton polyhedron* $\Gamma(\phi)$ of ϕ is the convex hull of the union of all quadrants $(j, k) + \mathbb{R}^2_+$ in $\mathbb{R}^2$ with $(j, k) \in \mathcal{S}(\phi)$. Let $\Delta(\phi)$ be the collection of compact faces (vertices and edges) of $\Gamma(\phi)$. The *Newton diagram* $\mathcal{N}_d(\phi)$ is the union of the faces in $\Delta(\phi)$.

For each face γ of $\Gamma(\phi)$, we set $\phi_\gamma(x, y) = \sum_{(j,k)\in\gamma} c_{j,k}x^j y^k$. We say that ϕ is $\mathbb{C}$-*nondegenerate* ($\mathbb{R}$-*nondegenerate*) if for every compact face $\tau \in \Delta(\phi)$,

$$\nabla\phi_\tau(x, y) \;=\; (\frac{\partial\phi_\tau}{\partial x}(x, y),\; \frac{\partial\phi_\tau}{\partial y}(x, y))$$

never vanishes in $(\mathbb{C} \setminus \{0\})^2$ $((\mathbb{R} \setminus \{0\})^2)$.

The Newton distance and the height function. If we use coordinates (t_1, t_2) for points in the plane containing the Newton polyhedron, consider the point (d, d) in this plane where the bisectrix $t_1 = t_2$ intersects the boundary of $\Gamma(\phi)$. The coordinate $d = d(\phi)$ is called the *Newton distance* of ϕ in the coordinates $z = (x, y)$. The *principal face* $\pi(\phi)$ is the face of minimal dimension (an edge or vertex) which contains the point (d, d). Following [8], we call $\phi_{\pi(\phi)}$ the *principal part* of ϕ and denote it by ϕ_{pr}.

When ϕ is $\mathbb{R}$-nondegenerate, then the real oscillation index is equal to the reciprocal of the Newton distance;[5] $\sigma_r(\phi) = 1/d(\phi)$. Similarly, when ϕ is $\mathbb{C}$-nondegenerate, then the complex oscillation index is also equal to $1/d(\phi)$ and so the real and complex indices agree in this case.

In two dimensions, we can still get our hands on the real oscillation index $\sigma_r(\phi)$ for general ϕ since $\sigma_r(\phi) = 1/h(\phi)$ is the reciprocal of the height $h(\phi) := \sup_z d_z$ where d_z is the Newton distance of ϕ in the coordinates $z = (x, y)$. Furthermore the supremum is attained $h(\phi) = d_{z_0}$ and we call any such coordinate system z_0 *adapted*. This is no longer the case in higher dimensions.

The notions of Newton polyhedron $\Gamma(\phi)$, Newton diagram $\mathcal{N}_d(\phi)$, Newton distance $d(\phi)$ as well as principal face $\pi(\phi)$ and principal part ϕ_{pr} easily extend from polynomials to any real-analytic function. This will be useful in Section 6.

[5]This is true in two dimensions but we need to assume in addition that $d(\phi) > 1$ in higher dimensions.

The Varčenko exponent. The *Varčenko exponent* $\nu(\phi)$ was introduced in [16] and is defined to be zero unless $h(\phi) \geq 2$ and in this case, when the principal face $\pi(\phi^z)$ of ϕ^z in an adapted coordinate system z is a vertex, we define $\nu(\phi)$ to be 1. Otherwise we set $\nu(\phi) = 0$.

Complex σ_c versus real σ_r. The complex oscillation index $\sigma_c(\phi)$ is smaller (and can be strictly smaller in dimensions three or more) than the real oscillation index $\sigma_r(\phi)$; see [3], Lemma 13.6. However in two dimensions, they agree. This follows from Theorem 4.4 in [17], which states that for polynomials in two variables, the complex oscillation index is given by $\frac{1}{h_c(f)}$, where $h_c(f)$ is the maximum of $d_z f$ with respect to any *complex* coordinate transformation z. Combining this with the observation that the proof of Theorem 3.3 of [8] can be easily modified to give a condition under which a coordinate system is not $\mathbb{C}$-adapted to f, implying that a coordinate system in two variables is $\mathbb{C}$-adapted if and only if it is $\mathbb{R}$-adapted, we observe that $\sigma_c = \sigma_r$ in two dimensions.

In general, it is the complex oscillation index $\sigma_c(\phi)$ and *not* the real oscillation index $\sigma_r(\phi)$ which governs the decay bounds for the exponential sums S_0 and S. Many simple examples show this; for example, consider the homogeneous polynomial $\phi(x, y, z) = (x^2 + y^2 + z^2)^2$ which is $\mathbb{R}$-nondegenerate (but $\mathbb{C}$-degenerate) and so $\sigma_r(\phi) = 1/d(\phi) = 3/4$. However a simple computation using Gauss sums shows that $|S(\phi, p^k)| \leq 10p^{-k/2}$ for all $k \geq 1$ and when k is even, $c_0 p^{-k/2} \leq |S(\phi, p^k)|$ for large primes p. Here $\sigma_c(\phi) = 1/2$. For lower bounds in great generality, see [6], Proposition 3.9.

The exceptional class $\mathcal{E}$. With the notions of the Newton diagram and the principal part of ϕ, we can now describe the exceptional class $\mathcal{E}$ precisely. It is the class of polynomials ϕ whose principal part $\phi_{\mathrm{pr}}(x, y) = a(bx^2 + cxy + dy^2)^m$ where the quadratic polynomial $bx^2 + cxy + dy^2$ is irreducible over the rationals $\mathbb{Q}$.

The p-adic number field. We fix a prime p and define the p-adic absolute value[6] $|\cdot| = |\cdot|_p$ on the field of rationals $\mathbb{Q}$ as follows. For integers $a \in \mathbb{Z}$, we define $|a| := p^{-k}$ where $k \geq 0$ is the largest power such that p^k divides a. This p-adic absolute value extends to all rationals a/b by $|a/b| = |a|/|b|$ and satisfies the basic conditions $|uv| = |u||v|$ and $|u + v| \leq |u| + |v|$ for all rationals $u, v \in \mathbb{Q}$, giving $\mathbb{Q}$ a metric space structure $d(u, v) = |u - v|$. The p-adic absolute value in fact satisfies a stronger version of triangle inequality called the ultrametric inequality: $|u + v| \leq \max(|u|, |v|)$. This implies $|u + v| = |u|$ if $|v| < |u|$ and so if $v \in B_r(u) := \{w \in \mathbb{Q} : |w - u| \leq r\}$, then $B_r(v) = B_r(u)$.

The p-adic field $\mathbb{Q}_p$ is the completion of the rational field $\mathbb{Q}$ with respect to the metric defined by the p-adic absolute value. The elements in the completed field

[6]We will also use the notation $|z|$ for the usual absolute value on elements $z \in \mathbb{C}$ but the context will make it clear which absolute value is being used.

$x \in \mathbb{Q}_p$ can be represented by a Laurent series

$$x \;=\; \sum_{j=-N}^{\infty} a_j p^j, \qquad a_j \in \mathbb{Z}/p\mathbb{Z} \;=\; \{0, 1, \cdots, p-1\}, \tag{6}$$

convergent with respect to $|\cdot| = |\cdot|_p$ which extends uniquely to all of $\mathbb{Q}_p$ by $|x| = p^N$ where $a_{-N} \neq 0$ is the first term of the series representation (6). We also define $|0| = 0$.

The compact unit ball $B_1(0) = \{x \in \mathbb{Q}_p : |x| \leq 1\}$ plays a special role as it is a ring due to the ultrametric inequality. We call this compact ring the *ring of p-adic integers* and denote it by $\mathbb{Z}_p$. Hence $\mathbb{Q}_p$ is a locally compact abelian group and has a unique Haar measure μ which we normalise so that $\mu(\mathbb{Z}_p) = 1$. To carry out Fourier analysis on $\mathbb{Q}_p$, we fix a non-principal additive character e defined by

$$\mathrm{e}(x) \;:=\; e^{2\pi i[\sum_{j=-N}^{-1} a_j p^j]} \;\text{ where } x \text{ is represented as in (6).}$$

All other characters χ on $\mathbb{Q}_p$ are given by $\chi(x) = \mathrm{e}(vx)$ for some $v \in \mathbb{Q}_p$. Hence the Fourier dual of $\mathbb{Q}_p$ is itself.

Hensel's lemma. The following basic lemma harks back to the origins of p-adic analysis and it, together with a generalisation described in Section 7, will be useful for us.

Lemma 3.1. *Let $g \in \mathbb{Z}[X]$ such that $g(x_0) \equiv 0$ mod p^s for some integer x_0. If $p^\delta || g'(x_0)$ (or $|g'(x_0)| = p^{-\delta}$) for some $\delta < s/2$, then there exists a unique $x \in \mathbb{Z}_p$ such that $g(x) = 0$ and $x \equiv x_0$ mod $p^{s-\delta}$.*

For a proof of Hensel's lemma, see [15], Chapter 1.6.

The sum $S_0(\phi, p^s)$ as an oscillatory integral. It is natural to analyse $S_0(\phi, p^s)$ by lifting this sum to an oscillatory integral defined over the p-adic field $\mathbb{Q}_p$.

First let us see how the complete exponential sum $S(\phi, p^s)$ can be written as the following oscillatory integral; we claim that

$$S(\phi, p^s) \;=\; \iint_{\mathbb{Z}_p \times \mathbb{Z}_p} \mathrm{e}(p^{-s}\phi(x,y))\, d\mu(x) d\mu(y) \tag{7}$$

holds. Consider a pair x_0, y_0 of integers and note that for any $x \in B_{p^{-s}}(x_0)$ and $y \in B_{p^{-s}}(y_0)$, we have $\mathrm{e}(p^{-s}\phi(x,y)) = \mathrm{e}(p^{-s}\phi(x_0,y_0))$. This simply follows from the definition of the character e. Hence the oscillatory integral in (7) can be written as

$$\sum_{(x_0,y_0) \in [\mathbb{Z}/p^s\mathbb{Z}]^2} \iint_{B_{p^{-s}}(x_0) \times B_{p^{-s}}(y_0)} \mathrm{e}(p^{-s}\phi(x,y))\, d\mu(x) d\mu(y)$$

$$= \sum_{(x_0,y_0) \in [\mathbb{Z}/p^s\mathbb{Z}]^2} \mathrm{e}(p^{-s}\phi(x_0,y_0)) \mu(B_{p^{-s}})^2 \;=\; p^{-2s} \sum_{(x_0,y_0) \in [\mathbb{Z}/p^s\mathbb{Z}]^2} e^{2\pi i \phi(x_0,y_0)/p^s}$$

and this last sum is our complete exponential sum $S(\phi, p^s)$. The last equality follows since $\mathrm{e}(p^{-s}\phi(x_0,y_0)) = e^{2\pi i \phi(x_0,y_0)/p^s}$ by the definition of e.

A similar argument shows that our local sum $S_0(\phi, p^s)$ has the following oscillatory integral representation:

$$S_0(\phi, p^s) \;=\; \iint_{|x|,|y|\le p^{-1}} \mathrm{e}(p^{-s}\phi(x,y))\, d\mu(x) d\mu(y). \tag{8}$$

To simplify notation, we will denote the Haar measure $d\mu(x)$ by dx and $\mu(E)$ by $|E|$.

4. A reduction of Theorem 1.1 to a basic bound for $S_0(\phi, p^s)$

In this section we will give a basic bound on the oscillatory integral in (8) which represents our local sum $S_0(\phi, p^s)$. This bound by itself will fall short in proving the desired bound (5) in Theorem 1.1 and so one of our main tasks will be to find a change of variables in (8) so that the bound formulated in this section, with the transformed phase under this change of variables, is sufficient to establish Theorem 1.1.

First though, we observe that we may assume that $\nabla\phi(0,0) = 0$. If $\nabla\phi(0,0) \neq 0$, say $\partial_x\phi(0,0) = c \neq 0$ and since c is a coefficient of ϕ, we have $p \nmid c$ whenever $p \notin \mathcal{P}$. Then for any integer $y \equiv 0 \bmod p$, consider the polynomial $g \in \mathbb{Z}[X]$ defined by $g(x) = \phi(x,y)$ and note that $p \nmid g'(x)$ for every $x \equiv 0 \bmod p$. Hence by Hensel's lemma, the map $x \to g(x)$ defines a bijection on $\{x \in \mathbb{Z}/p^s\mathbb{Z} : x \equiv 0 \bmod p\}$ so that

$$\sum_{\substack{x\in\mathbb{Z}/p^s\mathbb{Z}\\ x\equiv 0 \pmod p}} e^{2\pi i g(x)/p^s} \;=\; \sum_{\substack{u=0\\ p\,|\,u}}^{p^s-1} e^{2\pi i u/p^s}$$

which is equal to zero when $s \geq 2$, and equal to 1 when $s = 1$. Hence

$$S_0(\phi, p^s) \;=\; \frac{1}{p^{2s}} \sum_{\substack{(x,y)\in[\mathbb{Z}/p^s\mathbb{Z}]^2\\ x,y\equiv 0 \pmod p}} e^{2\pi i\phi(x,y)/p^s} \;=\; 0 \tag{9}$$

when $s \geq 2$ and equal to p^{-2} when $s = 1$.

A key result in this paper is the following.

Theorem 4.1. *For any $\phi \in \mathbb{Z}[X,Y]$ with $\nabla\phi(0,0) = 0$, we can find a polynomial $\psi \in \mathbb{Q}[X]$ such that if $\tilde{\phi}(x,y) = \phi(x, y+\psi(x))$, then $h(\phi) = h(\tilde{\phi}) = h(\tilde{\phi}_{\mathrm{pr}})$.*

This result was established in the euclidean setting by Ikromov and Müller, [8]. We follow their argument closely but with an extra effort to ensure that the polynomial ψ we end up with has rational coefficients. We postpone the proof until Section 6.

Theorem 4.1 is useful because there is a convenient expression for the height of the principal part of a polynomial. This characterisation will be given later in Proposition 4.4.

This change of variables $(x,y) \to (x, y+\psi(x))$ depends only on ϕ. As described in the introduction, the exceptional set of primes $\mathcal{P}$ include the prime divisors

of the coefficients of the transformed phase $\tilde{\phi}(x,y) = \phi(x, y+\psi(x)) \in \mathbb{Q}[X,Y]$. Hence for $p \notin \mathcal{P}$, $\tilde{\phi}$ will not only be a polynomial with rational coefficients, these coefficients will be units (their p-adic absolute values are equal to 1) in the ring of p-adic integers $\mathbb{Z}_p$. Hence implementing this change of variables in (8) shows that $S_0(\phi, p^s) = S_0(\tilde{\phi}, p^s)$.

Therefore in order to prove Theorem 1.1, according to Theorem 4.1, it suffices to assume that the phase ϕ in the oscillatory integral (8) representing the local sum S_0 is a polynomial with rational coefficients lying in $\mathbb{Z}_p$ and $h(\phi) = h(\phi_{\mathrm{pr}})$.

We begin with the decomposition

$$S_0(\phi, p^s) \;=\; \sum_{\tau \in \Delta(\phi)} \sum_{\substack{\vec{l} \in \mathbb{N}^2 \\ F(\vec{l}) = \tau}} \iint_{|x|=p^{-l_1}, |y|=p^{-l_2}} \mathrm{e}(p^{-s}\phi(x,y))\, dxdy \tag{10}$$

introduced in [7]. Here, for each $\vec{l} = (l_1, l_2) \in \mathbb{N}^2$, $F(\vec{l})$ is the face of $\Gamma(\phi)$ of largest dimension which is contained in the supporting line of $\Gamma(\phi)$ perpendicular to $\vec{l}$. In other words,

$$F(\vec{l}) \;=\; \{\vec{t} \in \Gamma(\phi) : \vec{t} \cdot \vec{l} = N(\vec{l})\} \;\text{ where }\; N(\vec{l}) \;:=\; \min_{\vec{t} \in \Gamma(\phi)} \vec{t} \cdot \vec{l}.$$

Note that $F(\vec{l})$ is a compact face of $\Gamma(\phi)$ if and only if $l_1 l_2 \neq 0$ which explains why only compact faces $\Delta(\phi)$ enter into the decomposition of S_0 above.

For each compact face $\tau \in \Delta(\phi)$ and each $\vec{l}$ such that $\tau = F(\vec{l})$, write $\phi(x,y) = \phi_\tau(x,y) + \eta_\tau(x,y)$. Then

$$\phi_\tau(p^{l_1}x, p^{l_2}y) \;=\; \sum_{(j,k) \in \tau} p^{(j,k)\cdot\vec{l}}\, c_{j,k} x^j y^k \;=\; p^{N(\vec{l})} \phi_\tau(x,y)$$

and

$$\eta_\tau(p^{l_1}x, p^{l_2}y) \;=\; \sum_{(j,k) \in \Gamma(\phi) \setminus \tau} p^{(j,k)\cdot\vec{l}}\, c_{j,k} x^j y^k \;=\; p^{N(\vec{l})}\, p\, g_\tau(x,y)$$

for some polynomial $g \in \mathbb{Z}_p[X,Y]$.

Changing variables to normalise the region of integration, we have

$$S_0(\phi, p^s) = \sum_{\tau \in \Delta(\phi)} \sum_{\substack{\vec{l} \in \mathbb{N}^2 \\ F(\vec{l}) = \tau}} p^{-l_1 - l_2} \iint_{|x|,|y|=1} \mathrm{e}(p^{-s+N(\vec{l})}(\phi_\tau(x,y) + p g_\tau(x,y)))\, dxdy.$$

Now let us fix a compact face $\tau \in \Delta(\phi)$. If $\tau = \{(\alpha, \beta)\}$ is a vertex, then $\phi_\tau(x,y) = cx^\alpha y^\beta$ is a monomial where c is a rational number with $|c| = 1$. If τ is a compact edge, then

$$\tau \;\subset\; \{(t_1, t_2) : qt_1 + mt_2 = n\}$$

for some integers $(m, q, n) = (m_\tau, q_\tau, n_\tau)$ with $\gcd(m,q) = 1$ and ϕ_τ is a quasi-homogeneous polynomial, $\phi_\tau(r^{\kappa_1}x, r^{\kappa_2}y) = r\phi_\tau(x,y)$ for $r > 0$ where $\kappa_1 = q/n$ and

$\kappa_2 = m/n$. The polynomial ϕ_τ consists of at least two terms and so by homogeneity, we can factor[7]

$$\phi_\tau(x,y) \;=\; c\,x^\alpha y^\beta \prod_{j=1}^{N} (y^q - \xi_j x^m)^{n_j} \tag{11}$$

for some roots $\{\xi_j\}_{j=1}^N$ lying in $\mathbb{Q}^{\rm alg}$. It will be convenient for us to think of the roots $\{\xi_j\}$ as lying in some finite field extension of $\mathbb{Q}_p$. Again c is a rational with $|c| = 1$.

Assume, without loss of generality, that $1 \le q \le m$. We define the *homogeneous distance*[8] $d_\tau = d_\tau(\phi)$ of ϕ_τ as

$$d_\tau \;:=\; \frac{1}{\kappa_1 + \kappa_2} \;=\; \frac{n}{m+q} \;=\; \frac{q\alpha + m\beta + qmM}{m+q} \tag{12}$$

where $M := \sum_{j=1}^N n_j$. The point (d_τ, d_τ) on the bisectrix lies on the line $\{(t_1,t_2) : qt_1 + mt_2 = n\}$ containing τ. Hence $d_\tau(\phi) \le d(\phi)$ for every compact edge $\tau \in \Delta(\phi)$. If τ is the principal face, then $d_\tau(\phi) = d(\phi)$ is the Newton distance of ϕ.

We will use the notation $m_{\rm pr}(\phi)$ to denote the maximal multiplicity among the roots $\{\xi_j\}_{j=1}^N$ appearing in the factorisation (11) of the principal part $\phi_{\rm pr}$ of ϕ when $\pi(\phi)$ is a compact edge.

The following simple lemma will be useful.

Lemma 4.2. *Let $\tau \in \Delta(\phi)$ be a compact edge.*

(a) If τ is not the principal face, then $M = \sum_{j=1}^N n_j \le d_\tau$. Furthermore strict inequality $M < d_\tau$ holds unless an endpoint of τ lies on the bisectrix.

(b) Suppose that $\tau = \pi(\phi)$ is the principal face, $m_{\rm pr}(\phi) > d(\phi)$, and j_* is such that $n_{j_*} := m_{\rm pr}(\phi)$. Then $n_j < d(\phi)$ for all $j \ne j_*$ and $\xi_{j_*} \in \mathbb{Q}$.

(c) Again suppose that $\tau = \pi(\phi)$ but now $n_{j_*} = m_{\rm pr}(\phi) = d(\phi)$. Then either $\phi_{\rm pr}(x,y) = c(y^2 + bxy + dx^2)^n$ is a power of a quadratic form or

$$\phi_{\rm pr}(x,y) \;=\; c\,x^\alpha y^\beta \prod_{j=1}^{N} (y - \xi_j x^m)^{n_j}, \tag{13}$$

$n_j < n_{j_*}$ for all $j \ne j_*$ and $\xi_{j_*} \in \mathbb{Q}$.

Proof. To prove (a), let us suppose that τ lies below the bisectrix so that the left endpoint $(\alpha, \beta + qM)$ of τ satisfies $\alpha \ge \beta + qM$. Hence by (12),

$$d_\tau \;\ge\; \frac{q(\beta + qM) + qmM}{q+m} \;\ge\; \frac{(\beta + qM) + mM}{q+m} \;\ge\; M$$

[7]See [8] for details.

[8]We are borrowing terminology from [8].

since $q \geq 1$. If the left endpoint $(\alpha, \beta + qM)$ does not lie on the bisectrix, then we have the strict inequality $\alpha > \beta + qM$ which in turn implies that the strict inequality $d_\tau > M$ holds.

To prove (b), suppose that $n_{j_*} > d(\phi)$ and that there exists a $1 \leq j \leq N$ with $j \neq j_*$ such that $d(\phi) \leq n_j$. Then $M > 2d(\phi)$ and so by (12),

$$d(\phi) \;\geq\; \frac{qmM}{m+q} \;>\; \frac{2qm}{m+q} d(\phi) \;\geq\; d(\phi)$$

which is a contradiction. Hence $n_j < d(\phi)$ for all $j \neq j_*$. To show that $\xi_{j_*} \in \mathbb{Q}$, we argue by contradiction once again and suppose that the degree of ξ_{j_*} over the rationals is at least 2. Since the conjugates of ξ_{j_*} all lie among the roots $\{\xi_j\}_{j=1}^N$, we would be able to find a conjugate ξ_j with $j \neq j_*$. As all conjugates must have the same multiplicity, we see that $n_j = n_{j_*}$ which we have just seen is impossible.

Finally to prove (c), suppose that $n_{j_*} = d(\phi)$ and that there is a $1 \leq j \leq N$ with $j \neq j_*$ and $n_j = n_{j_*}$. Hence $M \geq 2n_{j_*} = 2d(\phi)$ and so

$$d(\phi) \;\geq\; \frac{qmM}{m+q} \;\geq\; \frac{2qm}{m+q} d(\phi),$$

implying $2qm/(m+q) \leq 1$ and hence $q = m = 1$. Plugging this back into (12), we have

$$2d(\phi) \;= \alpha + \beta + M$$

and since $M \geq 2d(\phi)$, this gives a contradiction unless $\alpha = \beta = 0$ and $M = 2d(\phi)$. Hence $d(\phi) = n_{j_*} = n_j$ and

$$\phi_{\mathrm{pr}}(x, y) \;=\; c(y - \xi_1 x)^n (y - \xi_2 x)^n \;=\; c(y^2 + bxy + dx^2)^n$$

is a power of a quadratic form with $n = n_j$. Otherwise we have $n_j < n_{j_*} = d(\phi)$ for all $j \neq j_*$ and reasoning as part (b), we also conclude that $\xi_{j_*} \in \mathbb{Q}$.

□

The following theorem contains our basic estimate for S_0 which will imply Theorem 1.1 via Theorem 4.1.

Theorem 4.3. *Suppose that the coefficients of $\phi \in \mathbb{Q}[X, Y]$ are units in $\mathbb{Z}_p$.*

(a) If $\pi(\phi)$ is a compact edge, then

$$|S_0(\phi, p^s)| \;\lesssim_{\deg \phi}\; \begin{cases} p^{-s/d(\phi)} & \text{if } m_{\mathrm{pr}}(\phi) < d(\phi) \\ s p^{-s/d(\phi)} & \text{if } m_{\mathrm{pr}}(\phi) = d(\phi) \\ p^{-s/m_{\mathrm{pr}}} & \text{if } m_{\mathrm{pr}}(\phi) > d(\phi) \end{cases}. \tag{14}$$

(b) If $\pi(\phi)$ is a vertex, then

$$|S_0(\phi, p^s)| \;\lesssim_{\deg \phi}\; s p^{-s/d(\phi)} \tag{15}$$

and this improves to $|S_0(\phi, p^s)| \lesssim_{\deg \phi} p^{-s/d(\phi)}$ when the vertex $\pi(\phi) = (1, 1)$.

(c) If $\pi(\phi)$ is an unbounded edge, then

$$|S_0(\phi, p^s)| \ \lesssim_{\deg \phi} \ p^{-s/d(\phi)}. \tag{16}$$

To see how Theorem 4.3 implies Theorem 1.1 under the assumption $h(\phi) = h(\phi_{\rm pr})$ (which we can make by Theorem 4.1), we need the following characterisation of $h(\phi_{\rm pr})$.

Proposition 4.4. *Suppose that the principal face $\pi(\phi)$ of $\phi \in \mathbb{Q}[X, Y]$ is a compact edge. Then*

$$h(\phi_{\rm pr}) \ = \ \max(d(\phi), m_{\rm pr}(\phi)).$$

Proof. In [21] it was shown that when $\psi(x, y) = c\, x^\alpha y^\beta \prod_{j=1}^N (y^q - \xi_j x^m)^{n_j}$ is a quasi-homogeneous polynomial with rational coefficients, then $h(\psi) = \max(d_h(\psi), m_{\mathbb{Q}}(\psi))$ where

$$m_{\mathbb{Q}}(\psi) \ := \ \max(\alpha, \beta, n_j : \xi_j \in \mathbb{Q})$$

and $d_h(\psi)$ is the homogeneous distance given in (12). This result in [21] is a minor adjustment of the corresponding euclidean result found in [8].

Note that if the principal face $\pi(\phi)$ is a compact edge, then the left endpoint $(\alpha, \beta + qM)$ lies above the bisectrix (so that $\alpha < \beta + qM$) and the right endpoint $(\alpha + mM, \beta)$ lies below the bisectrix (so that $\beta < \alpha + mM$). Hence by (12) we see that $\max(\alpha, \beta) < d(\phi)$ and so by Lemma 4.2 part (b), we see that

$$h(\phi_{\rm pr}) \ = \ \max(d_h(\phi_{\rm pr}), m_{\mathbb{Q}}(\phi_{\rm pr})) \ = \ \max(d(\phi), m_{\rm pr}(\phi)). \tag{17}$$

□

Recall that the Varčenko exponent $\nu(\phi) = 1$ if and only if $h(\phi) \geq 2$ and there is an adapted coordinate system in which the principal face is a vertex. According to Ikromov and Müller in [8] (Corollaries 2.3 and 4.3), a coordinate system $z = (x, y)$ is adapted to $\phi(x, y)$ if and only if one of the following conditions is satisfied:

(a) $\pi(\phi)$ is a compact edge and $m_{\rm pr}(\phi) \leq d(\phi)$;

(b) $\pi(\phi)$ is a vertex; or

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

Hence if the principal face $\pi(\phi)$ of our polynomial $\phi(x, y)$ is a vertex (d, d) (where necessarily $d = d(\phi)$), then the coordinates $z = (x, y)$ are adapted and so $\nu(\phi) = 1$ when $d = h(\phi) \geq 2$ and $\nu(\phi) = 0$ when $d = 1$. In this case the bound in (15) implies $|S_0(\phi, p^s)| \leq C s^{\nu(\phi)} p^{-s/h(\phi)}$, establishing Theorem 1.1 in this case. When $\pi(\phi)$ is an unbounded edge, the coordinates are adapted and hence $d(\phi) = h(\phi)$. Thus (16) establishes Theorem 1.1.

Next suppose that $\pi(\phi)$ is a compact edge and $m_{\rm pr}(\phi) \neq d(\phi)$. Then by Proposition 4.4, the bound (14) implies $|S_0(\phi, p^s)| \leq C p^{-s/h(\phi)}$ since we are assuming $h(\phi) = h(\phi_{\rm pr})$. This establishes Theorem 1.1 in this case. Finally suppose that $\pi(\phi)$ is a

compact edge and $m_{\mathrm{pr}}(\phi) = d(\phi)$. Then $h(\phi) = d(\phi)$ by Proposition 4.4. If $\phi \in \mathcal{E}$, then (14) establishes Theorem 1.1 in this case.

Hence we may suppose that $\phi \notin \mathcal{E}$. In this case, we claim that there is a coordinate system in which the principal face is a vertex so that (15) can be used to show that Theorem 1.1 holds in this case as well. Lemma 4.2 part (c) implies that

$$\phi_{\mathrm{pr}}(x,y) \;=\; c\,x^{\alpha}y^{\beta}\prod_{j=1}^{N}(y-\xi_j x^m)^{n_j} \tag{18}$$

with $\xi_{j_*} \in \mathbb{Q}$. Recall that when $\phi_{\mathrm{pr}}(x,y) = c(y^2+bxy+dx^2)^n$ is a power of a quadratic, then $\phi_{\mathrm{pr}}(x,y) = \; c\,(y-\xi_1 x)^n(y-\xi_2)^n$ where $\xi_1, \xi_2 \in \mathbb{Q}$ since $\phi \notin \mathcal{E}$. This is of the form (18). It is simple matter to see that the change of variables $(x,y) \to (x, y+\xi_{j_*}^m)$ transforms our polynomial to one whose principal part is a vertex.

Therefore Theorem 1.1 follows from Theorems 4.1 and 4.3.

5. Proof of Theorem 4.3

A key step in the proof of the bounds for the local sum $S_0(\phi, p^s)$ in Theorem 4.3 will be to freeze one of the variables and estimate a sum in the other variable. Equivalently, we will reduce to bounding a one dimensional oscillatory integral and for this, we will employ the following useful bound.

Proposition 5.1. *Let $\psi \in \mathbb{Z}_p[X]$ and suppose there is an $n \geq 1$ such that $\psi^{(n)}(x)/n! \not\equiv 0$ mod p for all x lying in some subset $S \subseteq \mathbb{Z}/p\mathbb{Z}$. Then there exists a constant C, depending on the degree of ψ (and not on S or p) such that*

$$\Big|\sum_{x_0\in S}\int_{B_{p^{-1}}(x_0)} \mathrm{e}(p^{-s}\psi(x))\,dx\Big| \;\leq\; Cp^{-s/n} \tag{19}$$

holds for all $s \geq 2$. Furthermore when $n = 1$, the sum in (19) vanishes.

When $S = \mathbb{Z}/p\mathbb{Z}$, Proposition 5.1 was proved in [20] using a higher order Hensel lemma. However the proof given in [20] also gives the strengthening stated here where we consider a truncated integral (or sum) on which we know some derivative of ψ is non-degenerate. We will outline the proof of Proposition 5.1 in Section 7.

Now let us recall the decomposition (10) of the oscillatory integral representation of $S_0(\phi, p^s)$ and write $S_0(\phi, p^s) = \sum_{\tau\in\Delta(\phi)} I_\tau$ where

$$I_\tau \;:=\; \sum_{\substack{\vec{l}\in\mathbb{N}^2\\ F(\vec{l})=\tau}} p^{-l_1-l_2}\iint_{|x|,|y|=1} \mathrm{e}(p^{-s+N(\vec{l})}(\phi_\tau(x,y)+pg_\tau(x,y)))\,dxdy.$$

We will provide a bound for each I_τ with $\tau \in \Delta(\phi)$. We split into two cases: the case in which τ is a compact edge and the case in which τ is a vertex.

When τ is a compact edge. In this case, ϕ_τ is a quasi-homogeneous polynomial which can be factored

$$\phi_\tau(x,y) \;=\; c\,x^\alpha y^\beta \prod_{j=1}^{N}(y^q-\xi_j x^m)^{n_j};$$

see (11). If $\vec{l}$ is such that $F(\vec{l})=\tau$, then the vector $\vec{l}$ is perpendicular to the line $\{(t_1,t_2): qt_1+mt_2=n\}$ containing τ if and only if $\vec{l}=l(q,m)$ for some integer $l\geq 1$. Hence $N(\vec{l})=ln$ and so

$$I_\tau \;=\; \sum_{l\geq 1} p^{-l(m+q)} \iint_{|x|,|y|=1} \mathrm{e}(p^{-s+ln}(\phi_\tau(x,y)+pg_\tau(x,y)))\,dxdy. \tag{20}$$

Set $\kappa=\lceil \frac{s}{n}\rceil -1$ so that $s=\kappa n+r$ where $1\leq r\leq n$ and split $I_\tau=I_\tau^1+I_\tau^2$ into two parts where $I_\tau^1=\sum_{l\geq\kappa+1} I_{\tau,l}$ and $I_\tau^2=\sum_{l\leq\kappa} I_{\tau,l}$; here $I_{\tau,l}$ denotes the integral in (20). Note that $l\geq\kappa+1$ precisely when $s-ln\leq 0$ and hence the integrand in $I_{\tau,l}$ is identically equal to 1. Thus $I_{\tau,l}=(1-p^{-1})^2$ for such l and so

$$I_\tau^1 \;=\; (1-p^{-1})^2 \sum_{l\geq\kappa+1} p^{-l(m+q)} \;\lesssim\; p^{-(\kappa+1)(m+q)} \;\leq\; p^{-s(m+q)/n} \;=\; p^{-s/d_\tau}.$$

Since $d_\tau\leq d(\phi)$, we have

$$|I_\tau^1| \;\lesssim\; p^{-s/d(\phi)} \tag{21}$$

which is smaller than the bounds (14), (15), (16) in the statement of Theorem 4.3. Hence (21) gives an acceptable contribution for Theorem 4.3.

Let us now concentrate on bounding each integral $I_{\tau,l}$ arising in I_τ^2 when $L:=s-nl\geq 2$. In this case we will use Proposition 5.1 to bound

$$\mathcal{I}_{\tau,L} \;:=\; \iint_{|x|,|y|=1} \mathrm{e}(p^{-L}(\phi_\tau(x,y)+pg_\tau(x,y)))\,dxdy.$$

Set

$$X \;=\; \{(x_0,y_0)\in[\mathbb{Z}/p\mathbb{Z}]^2 : x_0y_0\neq 0\}$$

and note that the region of integration in the integral $\mathcal{I}_{\tau,L}$ is precisely the set of $(x,y)\in\mathbb{Z}_p^2$ such that (x,y) is congruent mod p to an element of X.

We will split $X=Z_0\cup Z_1\cup\cdots\cup Z_N$ according to roots $\{\xi_j\}$ of ϕ_τ. All the roots arising from the quasi-homogeneous polynomials ϕ_τ with $\tau\in\Delta(\phi)$ are algebraic numbers lying in our set $\mathcal{A}$ defining the exceptional primes $\mathcal{P}$ and hence lie in a finite field extension of $\mathbb{Q}_p$. Therefore each p-adic absolute value $|\cdot|=|\cdot|_p$ extends uniquely to these elements. Elementary considerations (see [21]) show that for $p\notin\mathcal{P}$, $|\xi_j|_p=1$ and $|\xi_j-\xi_k|_p=1$ whenever $j\neq k$.

We define

$$\begin{aligned} Z_0 &:= \{(x_0,y_0)\in X : \phi_\tau(x_0,y_0)\not\equiv 0 \bmod p\} \;\text{ and}\\ Z_j &:= \{(x_0,y_0)\in X : |y_0^q-\xi_j x_0^m|<1\} \qquad \text{for } 1\leq j\leq N. \end{aligned}$$

Note that Z_j may be empty if there are no ordered pairs of elements of $(x_0,y_0)\in(\mathbb{Z}/p\mathbb{Z})^2$ for which the inequality defining Z_j holds. Furthermore, the Z_j are disjoint: if $|y_0^q-\xi_jx_0^m|<1$ and $j\neq j'$ then the ultrametric inequality shows that

$|y_0^q - \xi_{j'} x_0^m| = |y_0^q - \xi_j x_0^m + \xi_j x_0^m - \xi_{j'} x_0^m| = 1$, since $|\xi_j x_0^m - \xi_{j'} x_0^m| = 1$ by the separation of the roots.

This gives us a disjoint decomposition of X. Accordingly, we split $\mathcal{I}_{\tau,L} = \sum_{j=0}^{N} \mathcal{I}_j$ where

$$\mathcal{I}_j \ := \ \sum_{(x_0,y_0)\in Z_j} \iint_{B_{p^{-1}}(x_0,y_0)} \mathrm{e}(p^{-L}(\phi_\tau(x,y) + p g_\tau(x,y)))\, dxdy.$$

Here $B_{p^{-1}}(x_0,y_0) = \{(x,y)\in\mathbb{Z}_p^2 : \max(|x-x_0|,|y-y_0|) \le p^{-1}\}$ consists of those elements of $\mathbb{Z}_p^2$ that are congruent to (x_0,y_0) modulo p.

First we claim that $\mathcal{I}_0 = 0$. In fact, each integral appearing in the sum defining $\mathcal{I}_0$ vanishes. Fix $(x_0,y_0)\in Z_0$ and let I_{x_0,y_0} denote the corresponding integral in $\mathcal{I}_0$. By a simple extension of Euler's homogeneous function theorem to the quasi-homogeneous case, we have $(qx,my)\cdot\nabla\phi_\tau(x,y) = n\phi_\tau(x,y)$ and so $\nabla\phi_\tau(x_0,y_0) \not\equiv 0 \mod p$. Set $\varphi(x,y) = \phi_\tau(x,y) + p g_\tau(x,y)$ and note that $\nabla\varphi(x_0,y_0) \not\equiv 0 \mod p$. Hence the argument establishing (9) shows $I(x_0,y_0) = 0$ and thus $\mathcal{I}_0 = 0$.

Let us now examine the other terms $\mathcal{I}_j$, $1\le j\le N$. We have

$$\begin{aligned}\mathcal{I}_j \ &= \ \sum_{(x_0,y_0)\in Z_j} \iint_{B_{p^{-1}}(x_0,y_0)} \mathrm{e}(p^{-L}\varphi(x,y))\, dxdy \\ &= \ \sum_{x_0\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\}} \sum_{y_0\in Z_{j,x_0}} \iint_{B_{p^{-1}}(x_0,y_0)} \mathrm{e}(p^{-L}\varphi(x,y))\, dxdy\end{aligned}$$

where $Z_{j,x_0} = \{y_0\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\} : (x_0,y_0)\in Z_j\}$.

Interchanging the sum in y_0 and the x integration, we have

$$\mathcal{I}_j \ = \ \sum_{x_0\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\}} \int_{B_{p^{-1}}(x_0)} \Big(\sum_{y_0\in Z_{j,x_0}} \int_{B_{p^{-1}}(y_0)} \mathrm{e}(p^{-L}\varphi(x,y))\, dy\Big)\, dx.$$

Denoting $\mathrm{Inner}_{x_0}(x)$ as the sum in y_0, we have

$$\mathcal{I}_j \ = \ \sum_{x_0\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\}} \int_{B_{p^{-1}}(x_0)} \mathrm{Inner}_{x_0}(x)\, dx.$$

For any fixed $x_0\in\mathbb{Z}/p\mathbb{Z}\setminus\{0\}$ and x such that $|x-x_0|\le p^{-1}$, define $\psi_x(y)$ to be the function $\varphi(x,y)$. Thus we have

$$\mathrm{Inner}_{x_0}(x) \ = \ \sum_{y_0\in Z_{j,x_0}} \int_{B_{p^{-1}(y_0)}} \mathrm{e}(p^{-L}\psi_x(y))\, dy,$$

putting us in a position to employ our bound (19) since it is straightforward to check that $\psi_x^{(n_j)}(y_0)/n_j! \not\equiv 0 \mod p$ for every $y_0\in Z_{j,x_0}$. Hence the uniform bound

$$|\mathrm{Inner}_{x_0}(x)| \ \le \ C_{\deg\phi}\, p^{-L/n_j}$$

holds and when $n_j = 1$, we have in fact $\mathrm{Inner}_{x_0}(x) = 0$. This implies that

$$|\mathcal{I}_j| \ \le \ C_{\deg\phi}\, p^{-L/n_j} \ \text{ but when } n_j = 1, \ \ \mathcal{I}_j = 0. \tag{22}$$

Therefore $|\mathcal{I}_{\tau,L}| \leq C_{\deg\phi}\, p^{-L/m_\tau}$, where m_τ is the maximal multiplicity of the roots $\{\xi_j\}$ of ϕ_τ. This gives us a bound on the sum of those terms in I_τ^2 where $L = s - ln \geq 2$; write $I_\tau^2 = I_\tau^{2,1} + I_\tau^{2,2}$ where $I_\tau^{2,1} := \sum_{\substack{1\leq l\leq\kappa \\ s-ln\geq 2}} I_{\tau,l}$ so that

$$|I_\tau^{2,1}| \;\leq\; C\Big(\sum_{1\leq l\leq\kappa} p^{-l(m+q)} p^{ln/m_\tau}\Big) p^{-s/m_\tau} \;=\; \Big(\sum_{1\leq l\leq\kappa} p^{-ln[\frac{1}{d_\tau}-\frac{1}{m_\tau}]}\Big) p^{-s/m_\tau}.$$

Hence

$$|I_\tau^{2,1}| \;\lesssim_{\deg\phi}\; \begin{cases} p^{-s/d_\tau} & \text{if } m_\tau < d_\tau \\ sp^{-s/d_\tau} & \text{if } m_\tau = d_\tau \\ p^{-s/m_\tau} & \text{if } m_\tau > d_\tau \end{cases}. \tag{23}$$

If $\tau = \pi(\phi)$ is the principal face, then $d_\tau = d(\phi)$ and $m_\tau = m_{\mathrm{pr}}(\phi)$ so that (23) gives an acceptable contribution to the bound (14) in Theorem 4.3. Now suppose that the compact edge τ is not the principal face. By Lemma 4.2 part (a), we conclude that $m_\tau \leq d_\tau$. Furthermore, if the endpoint of τ does not lie on the bisectrix, then in fact $m_\tau < d_\tau$ and so (23) implies $|I_\tau^{2,1}| \lesssim p^{-s/d_\tau} \lesssim p^{-s/d(\phi)}$ and this is an acceptable bound as before.

Finally suppose that an endpoint of τ lies on the bisectrix. Then the principal face $\pi(\phi)$ is a vertex and $d_\tau = d(\phi)$. The bound (23) gives an acceptable contribution to the bound (15) in Theorem 4.3 **unless** the vertex $\pi(\phi)$ is $(1,1)$. In this case $m_\tau = d_\tau = d(\phi) = 1$ and the formula (12) for d_τ shows two possible outcomes: (1) either $q = M = \alpha = 1$ and $\beta = 0$ in which case $\phi_\tau = ax(y - \xi x^m)$ for some $\xi \in \mathbb{Q}$ or (2) $q = M = m = \beta = 1$ and $\alpha = 0$ in which case $\phi_\tau = ay(y - \xi x)$ for some $\xi \in \mathbb{Q}$. In either case, $\mathcal{I}_{\tau,L} = \mathcal{I}_0 + \mathcal{I}_j$ where $n_j = 1$. Hence by (22) we see that $\mathcal{I}_{\tau,L} = 0$ implying in turn $I_\tau^{2,1} = 0$ in this case.

It remains to treat $I_\tau^{2,2}$ where we are summing the integrals $I_{\tau,l}$ for $1 \leq l \leq \kappa$ and $s - ln = 1$. The condition $s - ln = 1$ can only occur if $l = \kappa$ and $s \equiv 1 \bmod n$. Hence $I_\tau^{2,2} = I_{\tau,\kappa}$ and $s - \kappa n = 1$ so that

$$I_\tau^{2,2} \;=\; p^{-\kappa(m+q)} \iint_{|x|,|y|=1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\, dxdy$$

which is an exponential sum over a finite field. We claim that the bound

$$|I_\tau^{2,2}| \;\lesssim_\phi\; p^{-s/d(\phi)} \tag{24}$$

holds and as we have seen before, this is an acceptable bound.

First we can apply the Weil bound [19] for finite field sums (say to the y integral) to see that

$$|I_\tau^{2,2}| \;\lesssim_\phi\; p^{-(\frac{1}{2}-\frac{1}{d_\tau})} p^{-s/d_\tau};$$

here we used the identity $\kappa(m+q) = (s-1)(m+q)/n = (s-1)/d_\tau$. Therefore if $d_\tau \geq 2$, we obtain the bound (24). We now treat the case when $d_\tau < 2$.

First suppose that τ is not the principal face. Then $m_\tau \leq d_\tau$ by Lemma 4.2 which implies $m_\tau = 1$ (and hence $d_\tau \geq 1$) so that $\phi_\tau(x,y) = ax^\alpha y^\beta(y^q - \xi x^m)$ for some $\xi \in \mathbb{Q}$. If τ lies below the bisectrix, then the second coordinate of the left endpoint must be equal to 1. Hence $\beta + q = 1$ implying $q = 1$, $\beta = 0$ and so

$\phi_\tau(x,y) = ax^\alpha(y - \xi x^m)$. Similarly if τ lies above the bisectrix, then $\alpha = 0$, $m = 1$ and so $\phi_\tau(x,y) = ay^\beta(y - \xi x)$. In either case ϕ_τ is either linear in y or linear in x which implies that one of the integrals

$$\int_{|y|\le 1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dy, \qquad \int_{|x|\le 1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dx$$

is equal to 0. Hence

$$\Big|\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dxdy\Big| \;\lesssim\; p^{-1}$$

which in turn implies

$$|I_\tau^{2,2}| \;\le\; C_\phi\; p^{-(1-\frac{1}{d_\tau})} p^{-s/d_\tau} \;\le\; C_\phi\; p^{-s/d(\phi)},$$

establishing (24) in this case.

Now suppose that τ is the principal face. Then $d_\tau = d(\phi)$. If $d_\tau < 1$, then τ cannot contain any lattice points away from the coordinate axes. Hence $\alpha = \beta = 0$, $M = 1$ and so $\phi_\tau(x,y) = a(y^q - \xi x^m)$ for some $\xi \in \mathbb{Q}$. Using the formula (12), $d_\tau = qm/(m+q)$ and the restriction $d_\tau < 1$ shows $q = 1$. Therefore

$$\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dxdy \;=\; -p^{-1}\int_{|x|=1} \mathrm{e}(p^{-1}ax^m)\,dx \tag{25}$$

and if $m = 1$, the above integral is $O(p^{-2})$ leading to the bound

$$|I_\tau^{2,2}| \;\le\; C_\phi\; p^{-(2-\frac{1}{d_\tau})} p^{-s/d_\tau} \;\le\; C_\phi\; p^{-s/d(\phi)}$$

which proves (24). When $m \ge 2$, we are stuck with the bound $O(p^{-3/2})$ arising from a character sum estimate for the integral in (25) but in this case, we have $d_\tau = m/(m+1) \ge 2/3$ and so

$$|I_\tau^{2,2}| \;\le\; C_\phi\; p^{-(\frac{3}{2}-\frac{1}{d_\tau})} p^{-s/d_\tau} \;\le\; C_\phi\; p^{-s/d(\phi)}$$

which once again proves (24).

Finally suppose that τ is the principal face but $d_\tau \ge 1$. Since

$$\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dxdy \;=\; \iint_{\mathbb{Z}_p^2} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dxdy \;+\; O(p^{-1}),$$

we can use Cluckers's bound [5, Theorem 3.2.1] to conclude that

$$\Big|\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}(\phi_\tau(x,y)))\,dxdy\Big| \;\lesssim\; p^{-1/d_\tau}$$

which implies

$$|I_\tau^{2,2}| \;\le\; C_\phi\; p^{-s/d_\tau} \;\le\; C_\phi\; p^{-s/d(\phi)},$$

establishing (24) in all cases.

When τ is a vertex. We will now consider the case where $\tau = (\alpha, \beta)$ is a vertex. This means that $\phi_\tau(x,y) = cx^\alpha y^\beta$ a monomial. However, the sum over $\vec{l}$ will consist of more than just integer multiples of a fixed vector.

Assume that τ is the endpoint of two edges e_1 and e_2, where e_2 lies below (to the right of) τ and e_1 lies above (to the left of) τ. Hence if the edges are compact,

$$e_1 \subset \{(t_1,t_2) : q_1 t_1 + m_1 t_2 = n_1\} \text{ and } e_2 \subset \{(t_1,t_2) : q_2 t_1 + m_2 t_2 = n_2\}$$

for some positive integers $(q_j, m_j), j = 1,2$ with $\gcd(q_j, m_j) = 1$. If the e_2 is unbounded (that is, it is a horizontal line), then $e_2 \subset \{(t_1,t_2) : t_2 = \beta\}$. Likewise if e_1 is unbounded (vertical), then $e_1 \subset \{(t_1,t_2) : t_1 = \alpha\}$.

If both edges e_1 and e_2 are compact, then $F(\vec{l}) = \tau$ if and only if $\vec{l} = (l_1, l_2)$ satisfies

$$\frac{m_1}{q_1} < \frac{l_2}{l_1} < \frac{m_2}{q_2}.$$

If one of the edges is unbounded, the corresponding upper or lower restriction of the ratio l_2/l_1 is removed; for example, if e_2 is an infinite horizontal edge and e_1 is compact, then $F(\vec{l}) = \tau$ if and only if $m_1/q_1 < l_2/l_1$. We will, without loss of generality, assume that $\alpha \le \beta$.

Then $N(\vec{l}) = l_1\alpha + l_2\beta$ and our integral I_τ to bound is

$$\begin{aligned} I_\tau &= \sum_{\substack{\vec{l}\in\mathbb{N}^2 \\ F(\vec{l})=\tau}} p^{-l_1-l_2} \iint_{|x|,|y|=1} \mathrm{e}(p^{-s+N(\vec{l})}(cx^\alpha y^\beta + pg_\tau(x,y)))\,dxdy \\ &= \sum_{\substack{l_1,l_2\ge 1 \\ l_1\frac{r_1}{t_1} < l_2 < l_1\frac{r_2}{q_2}}} p^{-l_1-l_2} \iint_{|x|,|y|=1} \mathrm{e}(p^{-s+N(\vec{l})}(cx^\alpha y^\beta + pg_\tau(x,y)))\,dxdy \end{aligned}$$

with the understanding that if one of edges e_1 and/or e_2 is unbounded, then the corresponding restriction on the ratio l_2/l_1 does not appear.

We decompose $I_\tau = I_\tau^1 + I_\tau^2 + I_\tau^3$ into three pieces according to whether $N(\vec{l}) \ge s$, $N(\vec{l}) = s-1$ and $N(\vec{l}) \le s-2$, respectively. When $N(\vec{l}) \ge s$, the integrand is identically equal to 1 and so

$$I_\tau^1 \;\le\; (1-p^{-1})^2 \sum_{\substack{l_1,l_2\ge 1 \\ N(\vec{l})\ge s \\ l_2/l_1 < r_2/q_2}} p^{-l_1-l_2}$$

if the edge e_2 is compact. When e_2 is unbounded, the only restriction on the sum over $\vec{l} = (l_1,l_2)$ with $l_1, l_2 \ge 1$ is $N(\vec{l}) = l_1\alpha + l_2\beta \ge s$. This is a geometric series, and a straightforward argument shows

$$|I_\tau^1| \;\lesssim\; \begin{cases} p^{-s/d_{e_2}} & \text{if } \alpha \ne \beta \\ sp^{-s/d_{e_2}} & \text{if } \alpha = \beta \end{cases}. \tag{26}$$

Next let us turn our attention to

$$I_\tau^3 \;=\; \sum_{\substack{l_1,l_2\geq 1\\ s-N(\vec{l})\geq 2\\ m_1/q_1<l_2/l_1<m_2/q_2}} p^{-l_1-l_2}\iint_{|x|,|y|=1} \mathrm{e}(p^{-s+N(\vec{l})}\varphi_\tau(x,y))\,dxdy$$

where $\varphi_\tau(x,y)=cx^\alpha y^\beta+pg_\tau(x,y)$. Since $\nabla\varphi_\tau(x,y)\not\equiv 0 \bmod p$ for any $(x,y)\in\mathbb{Z}_p^2$ satisfying $|x|=|y|=1$, the same argument above showing that $\mathcal{I}_0=0$ shows that $I_\tau^3=0$.

Finally, the treatment of

$$I_\tau^2 \;=\; \sum_{\substack{l_1,l_2\geq 1\\ N(\vec{l})=s-1\\ r_1/t_1<l_2/l_1<r_2/q_2}} p^{-l_1-l_2}\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}cx^\alpha y^\beta)\,dxdy$$

follows along the same lines for I_τ^1, showing that (26) holds for I_τ^2 as well. Hence we have established Theorem 4.3 except in the solitary case that the principal face of ϕ is $(1,1)$ where we need to improve the bound for I_τ to $|I_\tau|\lesssim p^{-s}$ in order to finish the proof of Theorem 4.3.

The last step. When the vertex $\tau=(1,1)$, then $\tau=\pi(\phi)$ and $d(\phi)=1$. Here we will show the improved bound $|I_\tau|\lesssim p^{-s}$ which will conclude the proof of Theorem 4.3.

Recall the decomposition $I_\tau=I_\tau^1+I_\tau^2+I_\tau^3$ above where $I_\tau^3=0$ and in this case,

$$I_\tau^1 \;=\; (1-p^{-1})^2 \sum_{\substack{\vec{l}:\,N(\vec{l})\geq s\\ m_1/q_1<l_2/l_1<m_2/q_2}} p^{-l_1-l_2} \tag{27}$$

with the understanding that $l_1,l_2\geq 1$ and if one of edges e_1 and/or e_2 is unbounded, then the corresponding restriction on the ratio l_2/l_1 does not appear. Also

$$I_\tau^2 \;=\; \sum_{\substack{l_1,l_2\geq 1\\ N(\vec{l})=s-1\\ r_1/t_1<l_2/l_1<r_2/q_2}} p^{-l_1-l_2}\iint_{|x|,|y|=1} \mathrm{e}(p^{-1}cxy)\,dxdy$$

so that we can write

$$I_\tau^2 \;=\; -(1-p^{-1})p^{-1} \sum_{\substack{\vec{l}:\,N(\vec{l})=s-1\\ r_1/t_1<l_2/l_1<r_2/q_2}} p^{-l_1-l_2}.$$

Thus we see that in this case (when $\tau=(1,1)$), I_τ is a difference of two explicit sums of positive terms. A careful examination of this difference will exhibit the additional cancellation we seek.

We will show this when the edges e_1 and $_2$ are both infinite so the restrictions on $\vec{l}=(l_1,l_2)$ are $l_1,l_2\geq 1$ and either $N(\vec{l})\geq s$ or $N(\vec{l})=s-1$. The case when one

edge (or both) is compact is similar. In this case, $N(\vec{l}) = l_1 + l_2$ and so

$$I_\tau^1 \ = (1-p^{-1})^2 \sum_{\substack{l_1,l_2\geq 1\\ l_1+l_2\geq s}} p^{-l_1-l_2} \ = \ (1-p^{-1})^2 \sum_{N\geq s} (N-1)p^{-N}$$

and by the geometric series formula,

$$I_\tau^1 \ = \ (s-1)p^{-s} \ + \ -(s-2)p^{-s-1}.$$

In a similar but easy manner,

$$I_\tau^2 \ = \ -(1-p^{-1})p^{-1} \sum_{\substack{l_1,l_2\geq 1\\ l_1+l_2=s-1}} p^{-l_1-l_2} \ = \ -(1-p^{-1})(s-2)p^{-s}$$

and so $I_\tau = I_\tau^1 + I_\tau^2 = p^{-s}$ which shows the desired cancellation between the two terms I_τ^1 and I_τ^2.

This completes the proof of Theorem 4.3.

6. Proof of Theorem 4.1

Here we give the proof of Theorem 4.1 by developing an appropriate variant of an algorithm due Ikromov and Müller in [8] which produces an adapted coordinate system for any real-analytic function f. This algorithm constructs a series of changes of variables, and except for the final one, all are given by a simple polynomial map. The goal will be to show that the polynomial change of variables reached by the penultimate stage satisfies the conclusion of Theorem 4.1.

6.1. Conditions for Adapted Coordinate Systems.

For this section we will work entirely with *real-analytic* functions f. We will observe what happens when we apply the algorithm from [8] to a polynomial with rational coefficients.

The key observation is the one made in [21]: Corollary 2.3 from [8] is valid in any perfect field K; in particular, it is valid over $\mathbb{Q}$. The content of this corollary is to relate the roots of a quasi-homogeneous polynomial f to its *homogeneous distance* $d(f)$. A polynomial $f \in K[X,Y]$ being quasi-homogeneous makes sense in any field K and can be factored as

$$f(x,y) \ = \ c\, x^\alpha y^\beta \prod_{j=1}^{N} (y^q - \xi_j x^m)^{n_j}$$

where $c \in K$ and the roots $\{\xi_j\}_{j=1}^N$ lie in some finite field extension of K. Here $\gcd(m,q) = 1$ and $\kappa_1 := q/n, \kappa_2 := m/n$ are the dilation parameters so that $f(r^{\kappa_1}x, r^{\kappa_2}y) = rf(x,y)$ for $r > 0$. Recall that the homogeneous distance of f is defined as

$$d(f) \ = \ \frac{1}{\kappa_1+\kappa_2} \ = \ \frac{q\alpha + m\beta + qmM}{q+m}$$

where $M := \sum_{j=1}^N n_j$. Finally set $n_0 = \alpha$ and $n_{N+1} = \beta$.

We now reproduce the version Corollary 2.3 from [8] as it appeared in [21].

Lemma 6.2 ([8], [21]). *Let K be a perfect field and $f \in K[X,Y]$ be a quasi-homogeneous polynomial as above. Without loss of generality suppose that $\kappa_2 \geq \kappa_1$ or $1 \leq q \leq m$.*

1. *If there is a multiplicity $n_{j_*} > d(f)$ for some $0 \leq j_* \leq N+1$, then all the other multiplicities must be strictly less than $d(f)$; that is, $n_j < d(f)$ for all $0 \leq j \neq j_* \leq N+1$. In particular, there is at most one multiplicity n_j, $0 \leq j \leq N+1$ with $n_j > d(f)$.*
2. *If $\kappa_2/\kappa_1 \notin \mathbb{N}$, then $M = \sum_{j=1}^{N} n_j < d(f)$.*
3. *If $\kappa_2/\kappa_1 \in \mathbb{N}$, then $n_j \leq d(f)$ for any $1 \leq j \leq N$ such that $\xi_j \notin K$.*

The corollary says that the multiplicity of every root $\xi_j, 1 \leq j \leq N$, is bounded by $d(f)$ unless $\kappa_2/\kappa_1 \in \mathbb{N}$, in which case there is at most one root $\xi_j, 1 \leq j \leq N$ with multiplicity exceeding $d(f)$. If such a root exists, it necessarily lies in K and we shall call it the *principal root* of f.

We will need the following theorem in [8].

Theorem 6.3 (Ikromov-Müller). *Let f be a real-analytic function near the origin with $f(0,0) = 0$ and $\nabla f(0,0) = 0$. Then the given coordinates are not adapted to f if and only if the following hold true:*

1. *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 $\kappa_1, \kappa_2 > 0$. Swapping coordinates if necessary, we may assume $\kappa_2 \geq \kappa_1$.*
2. *$\frac{\kappa_2}{\kappa_1} \in \mathbb{N}$. Note that this implies that $q = 1$ in (11).*
3. *The inequality $m_{pr}(f) > d(f)$ holds.*

Moreover, in this case, an adapted coordinate system for f_{pr} is given by $y_1 := x_1$, $y_2 := x_2 - a x_1^m$, where a is the root of f_{pr} in the sense of (11) with the maximum multiplicity. The height of f_{pr} is then given by $h(f_{pr}) = m_{pr}(f)$.

We will apply Theorem 6.3 in the case when f has rational coefficients. In this case, when the principal face is a compact edge, $m_{\mathrm{pr}}(f) = \max_{1 \leq j \leq N} n_j$ where the $\{n_j\}$ are the multiplicities of the roots of the principal part $f_{\mathrm{pr}}(x,y) = cx^\alpha y^\beta \prod_{j=1}^{N} (y^q - \xi_j x^m)^{n_j}$ of f. We have $f_{\mathrm{pr}} \in \mathbb{Q}[X,Y]$ is a quasi-homogeneous polynomial with rational coefficients and we apply Lemma 6.2 with $K = \mathbb{Q}$ to conclude that if $n_{j_*} = m_{\mathrm{pr}}(f) > d(f)$, then the principal root $\xi_{j_*} \in \mathbb{Q}$ of f is a rational number.

We will adopt the following terminology from [8]. If a pair of dilation parameters $\kappa = (\kappa_1, \kappa_2)$ is chosen so that $L_\kappa = \{(t_1, t_2) : \kappa_1 t_1 + \kappa_2 t_2 = 1\}$ is a supporting line of the Newton polygon (that is, it contains a face $\tau = \tau_\kappa$ of the Newton diagram $\mathcal{N}_d(f)$), then we call $f_\tau(x_1, x_2) = \sum_{(j,k) \in \tau} c_{j,k} x_1^j x_2^k$ the κ-principal part of f. Abusing notation, we will sometimes denote this by f_κ. Note that $f_\kappa(x_1, x_2)$ is a quasi-homogeneous polynomial such that $f_\kappa(r^{\kappa_1} x_1, r^{\kappa_2} x_2) = r f_\kappa(x_1, x_2)$.

6.4. **Prerequisites to the Algorithm.** The Weierstrass preparation theorem holds for the ring $\mathbb{Q}\{x_1, x_2\}$ of convergent power series with rational coefficients. This can be seen by either modifying the proof of the Weierstrass preparation theorem for real coefficients given in [13] or observing that the Weierstrass preparation theorem holds for both $\mathbb{R}\{x_1, x_2\}$ (see [13]) and for the rings $\mathbb{Q}[[x_1, x_2]]$ of formal power series over $\mathbb{Q}$ and $\mathbb{R}[[x_1, x_2]]$ of formal power series over $\mathbb{R}$ (see [1]), and observing that the uniqueness of the factorisation in $\mathbb{R}[[x_1, x_2]]$ given by the Weierstrass preparation theorem implies that the factorisation over $\mathbb{Q}[[x_1, x_2]]$ and $\mathbb{R}\{x_1, x_2\}$ are the same.

This means that given an analytic function $f \in \mathbb{Q}\{x, y\}$, convergent in a neighbourhood of the origin (with the real topology on $\mathbb{Q}$), where $f(0, x_2) = x_1^{\nu_1} x_2^{\nu_2} f'(x_1, x_2)$, and where $f'(0, x_2) = x_2^m g(x_2)$, $g(0) \neq 0$, we can write f in the form

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

where

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

where $U(0, 0) \neq 0$ and $g_j(0) = 0$ for all j. Furthermore $g_1, \ldots, g_m \in \mathbb{Q}\{x_1\}$ are uniquely determined, not just in $\mathbb{Q}\{x_1\}$, but also as formal power series in the larger rings $\mathbb{Q}[[x_1]]$ and $\mathbb{R}[[x_1]]$. The unit U is also uniquely defined as a power series in $\mathbb{R}[[x_1, x_2]]$.

We may assume that g_m is not zero. Then the roots $r(x_1)$ of $F(x_1, x_2)$ have a Puiseux series expansion

$$r(x_1) \;=\; c_1 x_1^a + c_2 x_1^b \;+\; \cdots$$

where, importantly for us, the nonzero *coefficients* c_l's lie in $\mathbb{Q}^{\text{alg}}$, the algebraic closure of $\mathbb{Q}$ and the *exponents* $0 < a < b < \cdots$ are a strictly increasing sequence of rational numbers. A reference showing the existence of a *formal* Puiseux expansion of this form is Abhyankar's book [1]. Combining this with the usual Puiseux theorem for real power series as we did for the Weierstrass preparation theorem shows that the series describing each root is convergent.

The Puiseux expansion of two or more distinct roots r of F may agree for the first few terms and it will be important for us to quantify this.

We introduce the following notation from [8]. Let $a_1 < \cdots < a_n$ be the distinct leading exponents of the roots of F so that each root $r(x_1) = c x_1^{a_l} + O(x_1^A)$ for some $c \neq 0$, $1 \leq l \leq n$ and for some $A > a_l$. For each $l \in \{1, 2, \ldots, n\}$, we denote by $\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]$ the collection of roots with leading exponent a_l. Next, for every $1 \leq l \leq n$, let $\{c_l^{(\alpha)}\}$ denote the collection of distinct, leading nonzero coefficients appearing in the expansion of a root with leading exponent a_l and let $\left[\begin{smallmatrix} \alpha \\ l \end{smallmatrix}\right]$ denote the collection of roots with leading exponent a_l and leading coefficient $c_l^{(\alpha)}$.

We continue to the second exponent in the expansion; for every l_1 and α_1, we let $\{a_{l_1,l}^{(\alpha_1)} : l \geq 1\}$ denote the collection of distinct exponents appearing in the second term of the Puiseux expansion of the roots in $\left[\begin{smallmatrix} \alpha_1 \\ l_1 \end{smallmatrix}\right]$. Proceeding in this way, we can

express each root r as

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

where the nonzero coefficients c_l lie in $\mathbb{Q}^{\text{alg}}$ and

$$c_{l_1,\ldots,l_p}^{(\alpha_1,\ldots,\alpha_{p-1},\beta)} \neq c_{l_1,\ldots,l_p}^{(\alpha_1,\ldots,\alpha_{p-1},\gamma)}$$

whenever $\beta \neq \gamma$. Also

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

so that the terms in r have increasing exponents. Furthermore the exponents are positive rational numbers.

The root cluster $\left[\begin{smallmatrix} \alpha_1 & \cdots & \alpha_p \\ l_1 & \cdots & l_p \end{smallmatrix}\right]$ denotes the collection of roots whose first p leading terms are indexed by $l_1, \alpha_1, l_2, \alpha_2, \ldots, l_p, \alpha_p$. We will also introduce clusters where the last *exponent* has been picked but not the last *coefficient*. These are denoted $\left[\begin{smallmatrix} \alpha_1 & \cdots & \alpha_{p-1} & \cdot \\ l_1 & \cdots & l_{p-1} & l_p \end{smallmatrix}\right]$ and equal the union over α_p of the clusters $\left[\begin{smallmatrix} \alpha_1 & \cdots & \alpha_{p-1} & \alpha_p \\ l_1 & \cdots & l_{p-1} & l_p \end{smallmatrix}\right]$. The notation N[cluster] will denote the number of roots in a cluster.

Since each a_l corresponds to the cluster $[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}]$, the collection of roots of F can be expressed as the union over all l of these clusters. Then we can write

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

where

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

The advantage of this decomposition is that it allows us to read off the vertices of the Newton polygon.

Lemma 6.5. *The points* (A_l, B_l) *where*

$$A_l = \nu_1 + \sum_{\mu \leq l} a_\mu N[\begin{smallmatrix} \cdot \\ \mu \end{smallmatrix}] \quad \text{and} \quad B_l = \nu_2 + \sum_{\mu \geq l+1} N[\begin{smallmatrix} \cdot \\ \mu \end{smallmatrix}]$$

are the vertices of the Newton polygon of f.

Here l ranges between 0 and n. When $l = 0$, we set $a_0 = 0$ so that $A_0 = \nu_1$ and $B_0 = \nu_2 + m$ where m is the degree of $F(x_1, x_2)$ as a polynomial in x_2; that is, the sum of the multiplicities of the roots of F. When $l = n$, $B_n = \nu_2$.

Proof. The Newton polygon of f is the same as the Newton polygon of $x_1^{\nu_1} x_2^{\nu_2} F(x_1, x_2)$.

Consider any $\kappa > 0$ not among the exponents $\{a_1, a_2, \ldots, a_n\}$ and choose $0 \leq l_\kappa \leq n$ so that $a_{l_\kappa} < \kappa < a_{l_\kappa + 1}$ (if $a_n < \kappa$, choose l_n). Let $L_\kappa = \{(t_1, t_2) : t_1 + \kappa t_2 = c_\kappa\}$ be a supporting line of the Newton polygon of f. It either intersects the Newton diagram in a vertex or a compact edge as κ is a positive, finite number. In fact we will see that L_κ intersects the Newton diagram in a vertex.

We say that a monomial $x_1^c x_2^d$ in the Puiseux expansion of F has degree $c + \kappa d$ with respect to the weight $(1, \kappa)$. A necessary and sufficient condition for a point (c_0, d_0) to lie on L_κ is that it has minimal $(1, \kappa)$-degree among all the pairs (c, d) arising as a monomial $x_1^c x_2^d$ in the Puiseux expansion of F.

For each factor $x_2 - r(x_1)$ arising in F with the root $r(x_1)$ belonging to $[\,_l^i]$, the term x_2 has $(1, \kappa)$-degree equal to κ and the minimal $(1, \kappa)$-degree among the terms in the Puiseux expansion of $r(x_1)$ is a_l. Hence the lowest-degree $(1, \kappa)$-monomial appearing in F is $x_1^{A_{l_\kappa}} x_2^{B_{l_\kappa}}$ since we take the $x_1^{a_l}$ term for $l \le l_\kappa$ and the x_2 term for $l > l_\kappa$. This shows that L_κ intersects the Newton diagram at the vertex $(A_{l_\kappa}, B_{l_\kappa})$.

Note that each A_l must be an integer (it is obvious that B_l is an integer) since the vertices of the Newton diagram are lattice points. □

Now, notice that $A_l - A_{l-1} = -a_l(B_l - B_{l-1})$, since $N[\,_l^i]$ is equal to $B_l - B_{l-1}$. From this it immediately follows that the slope of the line connecting (A_{l-1}, B_{l-1}) to (A_l, B_l) is $-1/a_l$. Therefore the line connecting (A_{l-1}, B_{l-1}) to (A_l, B_l) is given by $y = -(1/a_l)(x - A_l) + B_l$.

This line intersects the bisectrix at (d_l, d_l) where $d_l = -(1/a_l)(d_l - A_l) + B_l$, so $d_l = \frac{A_l + a_l B_l}{1 + a_l}$. If we index this line $L_{\kappa^l} = \{(t_1, t_2) : \kappa_1^l t_1 + \kappa_2^l t_2 = 1\}$ by the dilation parameters $\kappa^l = (\kappa_1^l, \kappa_2^l)$, then

$$\kappa_1^l \;=\; \frac{1}{A_l + a_l B_l}, \quad \text{and} \quad \kappa_2^l \;=\; \frac{a_l}{A_l + a_l B_l} \quad \text{so that } a_l \;=\; \frac{\kappa_2^l}{\kappa_1^l}.$$

The vertical edge, which passes through $(\nu_1, \nu_2 + m)$ (here m is the sum of the multiplicities of all the roots $r(x_1)$ in F), intersects the bisectrix at (ν_1, ν_1), and the horizontal edge, passing through (A_n, ν_2), is contained in a line intersecting the bisectrix at (ν_2, ν_2). So the distance $d(f)$ is given by $\max(\nu_1, \nu_2, \max_l d_l)$.

Finally, we observe that the κ^l-principal part of f is the same as the κ^l-principal part of

$$c\, x_1^{\nu_1} x_2^{\nu_2} \prod_{j,\alpha} (x_2 - c_j^{(\alpha)} x_1^{a_j})^{N[\,_j^\alpha]}$$

where $c = U(0, 0)$. Since the κ^l-principal part of $x_2 - c_j^{(\alpha)} x_1^{a_j}$ equals $c_j^{(\alpha)} x_1^{a_j}$ if $j < l$ and equals x_2 if $l < j$, we have

$$f_{\kappa^l}(x_1, x_2) = c_l x_1^{A_{l-1}} x_2^{B_l} \prod_{\alpha} (x_2 - c_l^{(\alpha)} x_1^{a_l})^{N[\,_l^\alpha]}. \tag{28}$$

In view of (28) we say that the edge $[(A_{l-1}, B_{l-1}), (A_l, B_l)]$ is associated to the cluster of roots $[\,_l^i]$.

6.6. **The Algorithm.** We are now ready to describe the algorithm. Suppose that $f(x_1, x_2)$ is a real-analytic function near $(0, 0)$ with rational coefficients. Furthermore suppose that the coordinates (x_1, x_2) are not adapted (otherwise there is nothing to do).

We apply Theorem 6.3 part (a) to conclude that the principal face $\pi(f)$ is a compact edge which lies on a uniquely determined line $L_\kappa = \{(t_1, t_2) : \kappa_1 t_1 + \kappa_2 t_2 = 1\}$ with $\kappa_1, \kappa_2 > 0$. The principal part f_{pr} is just the κ-principal part of f. By Lemma 6.5 the compact edges of $\mathcal{N}_d(f)$ are given by $[(A_{l-1}, B_{l-1}), (A_l, B_l)]$ with $1 \le l \le n$. Choose λ so that the principal face $\pi(f)$ of f is $\tau_\lambda := [(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$. Therefore by (28), we have

$$f_{\mathrm{pr}}(x_1, x_2) \;=\; f_{\kappa^\lambda}(x_1, x_2) \;=\; c\, x_1^{A_{\lambda-1}} x_2^{B_\lambda} \prod_{\alpha} (x_2 - c_\lambda^{(\alpha)} x_1^{a_\lambda})^{N\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]}. \tag{29}$$

The slope of τ_λ is $-1/a_\lambda$ so that $a_\lambda = \kappa_2/\kappa_1$. By Theorem 6.3 part (b), $a_\lambda \in \mathbb{N}$. Furthermore by part (c), there exists an index β such that

$$m_{\mathrm{pr}}(f) \;=\; N\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right] \;>\; d(f) \;=\; \frac{A_\lambda + a_\lambda B_\lambda}{1 + a_\lambda}, \;\text{ and }\; c_\lambda^{(\beta)} \in \mathbb{Q}. \tag{30}$$

The root $c_\lambda^{(\beta)}$ is the principal root of f_{pr}.

The first step is to apply $x = \sigma(y)$ where $y_1 := x_1$ and $y_2 := x_2 - c_\lambda^{(\beta)} x_1^{a_\lambda}$ and put $\tilde{f} = f \circ \sigma$. Then $\tilde{f}(y_1, y_2)$ is equal to $f(y_1, y_2 + c_\lambda^{(\beta)} y_1^{a_\lambda})$. Since a_λ is an integer, this is a polynomial change of variables.

We want to see what happens to the Newton diagram from this change of variables. We will use the $\widetilde{\ }$ notation to denote quantities in the variables (y_1, y_2); for example $\tilde{U}(y_1, y_2) = U(y_1, y_2 + c_\lambda^{(\beta)} y_1^{a_\lambda})$. Hence

$$\tilde{f}(y_1, y_2) \;=\; \tilde{U}(y_1, y_2) y_1^{\nu_1} (y_2 + c_\lambda^{(\beta)} y_1^{a_\lambda})^{\nu_2} \prod_{l,\alpha} \big(y_2 - (r(y_1) - c_\lambda^{(\beta)} y_1^{a_\lambda})\big)^{N\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]}$$

and so each root $\tilde{r}(y_1)$ of $\tilde{f}$ has the form

$$\tilde{r}(y_1) \;=\; c_l^{(\alpha_1)} y_1^{a_l} - c_\lambda^{(\beta)} y_1^{a_\lambda} \;+\; \text{higher order terms.}$$

For $l < \lambda$, the lowest degree term in the root is left unchanged, so we have $\tilde{a}_l = a_l$. Furthermore the multiplicities $\widetilde{N\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}$ are the same as the corresponding multiplicities for f.

For $l > \lambda$, any root r in any cluster $\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]$ (including the $x_2^{\nu_2}$ term) is transformed into a root with leading exponent a_λ. The same happens for roots in $\left[\begin{smallmatrix}\cdot\\ \lambda\end{smallmatrix}\right]$ that are not in $\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]$. Finally if $r \in \left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]$, then the leading exponent of $\tilde{r}$ is of the form $a_{\lambda, l_2}^{(\beta)} > a_\lambda$.

Following Ikromov and Müller [8], we separately consider two cases depending on whether or not there is a root that maps to a root with leading exponent a_λ.

Case 1: This is the case where there is at least one root that maps to a root with leading exponent a_λ. This implies that $\tilde{a}_\lambda = a_\lambda$. We have $\tilde{B}_\lambda = N\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]$ since the roots $\tilde{r}$ with leading exponent greater than a_λ are precisely those roots corresponding to $r \in \left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]$.

We then see that $\tilde{A}_\lambda = \tilde{A}_{\lambda-1} + a_\lambda B_\lambda - a_\lambda N\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]$ so that

$$(\tilde{A}_\lambda, \tilde{B}_\lambda) = \left(A_{\lambda-1} + a_\lambda B_\lambda - a_\lambda N\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right], N\left[\begin{smallmatrix}\beta\\ \lambda\end{smallmatrix}\right]\right).$$

The inequality in (30) is equivalent to the statement that $\tilde{A}_\lambda < \tilde{B}_\lambda$. Therefore the edge $[(\tilde{A}_{\lambda-1}, \tilde{B}_{\lambda-1}), (\tilde{A}_\lambda \tilde{B}_\lambda)]$ lies entirely above the bisectrix and is thus not the principal face. Hence the principal face is associated to some subcluster $\left[\begin{smallmatrix}\beta & \cdot\\ \lambda & \lambda_2\end{smallmatrix}\right]$ in the original coordinates (or is a horizontal edge in which case the new coordinates are adapted).

Case 2: This is the other case. Now there is no root with leading exponent a_λ in the new coordinates and again the principal face corresponds to a subcluster of the same form (or is an unbounded edge in which case we are done).

If $\tilde{f}$ is not yet expressed in an adapted coordinate system (so that the conditions (1)-(3) in Theorems 6.3 still hold), we continue the procedure. Now, the later steps are similar. If the conditions (1)-(3) are satisfied, we again take the principal root, which is known to exist and is a rational number. In terms of the original coordinates, we now have a change of coordinates $x = \sigma_{(2)}(y)$ of the form

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

where the coefficients are, once again, rational, and the exponents are integers, and now the new principal face will be a compact edge associated to a further subcluster of the original root cluster, or it will be an unbounded edge, in which case the new coordinates are adapted.

We iterate this procedure. If this procedure terminates after finitely many steps, then we have arrived at a polynomial shear transformation that converts the coordinates into adapted coordinates. The conclusion of Theorem 4.1 therefore follows.

On the other hand, it is possible that this procedure does not terminate after finitely many steps. In this case, the multiplicities

$$N_k := N\left[\begin{smallmatrix}\beta & \beta_2 & \cdots & \cdot\\ \lambda & \lambda_2 & \cdots & \lambda_{k+1}\end{smallmatrix}\right]$$

are a nonincreasing sequence of positive integers and hence eventually constant. We can therefore find a polynomial $\psi_0 \in \mathbb{Q}[X]$ such that the function $f_0(x_1, x_2) := f(x_1, x_2 + \psi_0(x_2))$ has an analytic root

$$\rho(x_1) := c_\lambda^{(\beta)} x_1^{a_\lambda} + \cdots$$

where each coefficient of this root is rational and where ρ is not a polynomial. Furthermore, ψ_0 can be chosen so that $\rho(x_1)$ is the only root with leading exponent λ, but the root ρ may have higher multiplicity.

Now if we take $\tilde{f}(y_1, y_2) := f_0(y_1, y_2 + c_\lambda^{(\beta)} y_1^{a_\lambda})$, the previous arguments imply that the principal face of $\tilde{f}$ must be the final non-horizontal edge in the Newton diagram. Furthermore $\tilde{f}$ does not have a vanishing root because this would imply that f_0 has a root $c_\lambda^{(\beta)} x_1^{a_\lambda}$, which cannot exist because that would contradict the multiplicity assumption on f and the particular choice of ψ_0.

We claim that $\tilde{f}$ satisfies the conclusion of Theorem 4.1. We will do this by making a further, non-polynomial change of variables that yields an adapted coordinate system.

By the construction of $\tilde{f}$, the vertices of the Newton polyhedron of $\tilde{f}$ are given by $(A_0, B_0), \ldots, (A_\lambda B_\lambda)$ where $B_\lambda = 0$ and the principal edge is $[(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$, where $A_{\lambda-1} < B_{\lambda-1}$. From (29), we see that the principal part of $\tilde{f}$ is

$$\tilde{f}_{\mathrm{pr}}(x_1, x_2) = c x_1^{A_{\lambda-1}} (x_2 - c_\lambda^{(\beta)} x_1^{a_\lambda})^N$$

where $N = B_{\lambda-1} > \nu_1$.

We will now apply Proposition 4.4 to show that the height of $\tilde{f}_{\mathrm{pr}}$ is equal to N. Since the principal face of $\tilde{f}$ is the compact edge $[(A_{\lambda-1}, B_{\lambda-1}), (A_\lambda, B_\lambda)]$ and $A_{\lambda-1} < B_{\lambda-1} = N$, we see that $d(f_{\mathrm{pr}}) < N$. But the root $c_\lambda^{(\beta)}$ has multiplicity N as a root in the sense of the factorisation (11), so this must be the principal root of $\tilde{f}_{\mathrm{pr}}$ and thus the height of $\tilde{f_{\mathrm{pr}}}$ is N by Proposition 4.4.

We now consider the function $f^*(y_1, y_2)$ given by $\tilde{f}(y_1, y_2 + \rho(y_1))$. The nonzero roots $\tilde{r}$ are given by $r - \rho$ with $r \in [\,]_l$ for some $l < \lambda$ and they have the same multiplicities and leading exponents as r. This change of variables deletes the last vertex of the Newton polygon since the last factor changes into y_2^N and the principal face is now an unbounded horizontal edge. Therefore the Newton distance is N, the multiplicity of the vanishing root and so the height of f, the height of f^*, the height of $\tilde{f}$, and the height of $\tilde{f}_{\mathrm{pr}}$ are all equal to N.

The completes the proof of Proposition 4.1.

7. Hensel's lemma and the proof of Proposition 5.1

A weaker version of Proposition 5.1 was established in [20] but the argument given in [20] readily extends to give a proof of Proposition 5.1. Here we give an outline of the proof which relies on a generalisation of the classical Hensel lemma. The following result was established in [20].

Lemma 7.1. *Let $g \in \mathbb{Z}_p[X]$ with $p > \deg(g)$. Suppose there exists an integer $L \geq 1$ such that for any $x_0 \in \mathbb{Z}_p$,*

1. $|g^{(k+1)}(x_0)g(x_0)| \ < \ |g^{(k)}(x_0)g'(x_0)|$, for all $1 \leq k \leq L-1$, *and*

2. $|g(x_0)| \ < \ |g^{(L)}(x_0)g'(x_0)|$.

Then there exists a unique $x \in \mathbb{Z}_p$ such that $g(x) = 0$ and $|x - x_0| \leq |g(x_0)g'(x_0)^{-1}|$.

Remarks:

1. The lemma is valid for all primes p but then the derivatives $g^{(k)}(x)$ appearing in the statement of the lemma need to be replaced by $g^{(k)}(x)/k!$.

2. The $L = 1$ case is the classical statement of Hensel's lemma. In this case, condition 1 is vacuous and 2 reduces to the usual hypothesis $|g(x_0)| < |g'(x_0)|^2$. In particular if $g(x_0) \equiv 0 \bmod p^s$ and $p^\delta || g'(x_0)$ where $\delta < s/2$, then $|g(x_0)| < |g'(x_0)|^2$. The conclusion implies that there exists a unique $x \in \mathbb{Z}_p$ with $x \equiv x_0 \bmod p^{s-\delta}$ and $g(x) = 0$.

3. The lemma holds in any field K, complete with respect to any nontrivial nonarchimedean absolute value $|\cdot|$ and $g \in \mathfrak{o}[X]$ where $\mathfrak{o} = \{x \in K : |x| \le 1\}$.

4. The proof is a small variant of the usual proof of Hensel's lemma using the Newton formula to produce an approximating sequence to a solution of a polynomial equation.

We now turn to the proof of Proposition 5.1 where we seek to prove the following: suppose $\psi \in \mathbb{Z}_p[X]$ and that for some $n \ge 1$, $\psi^{(n)}(x_0)/n! \not\equiv 0 \bmod p$ for all $x_0 \in S$ in some set $S \subseteq \mathbb{Z}/p\mathbb{Z}$. Then for

$$I \;:=\; \sum_{x_0 \in S} \int_{B_{p^{-1}}(x_0)} \mathrm{e}(p^{-s}\psi(x))\, dx,$$

we have $|I| \le C p^{-s/n}$ for all $s \ge 2$ with a constant C depending only on n and the degree of ψ. This is the bound (19).

When $n = 1$ then each integral in the above sum over S vanishes. This follows in the same way we showed $\mathcal{I}_0 = 0$ in the proof of Theorem 4.3.

Suppose now $n \ge 2$, and, to simplify matters, we will assume that $s \equiv 0 \bmod n$. The other cases are slightly more involved, especially the case $s \equiv 1 \bmod n$ but here we just want to give a general outline how to prove (19). When $s \equiv 0 \bmod n$, then $s = tn$ for some $t \ge 1$. We write

$$I \;=\; \sum_{x_0 \in S} \sum_{\substack{u_0 \in \mathbb{Z}/p^t\mathbb{Z} \\ u_0 \equiv x_0 \bmod p}} \int_{B_{p^{-t}}(u_0)} \mathrm{e}(p^{-nt}\psi(x))\, dx \;=$$

$$\sum_{x_0 \in S} \sum_{\substack{u_0 \in \mathbb{Z}/p^t\mathbb{Z} \\ u_0 \equiv x_0 \bmod p}} p^{-t} \int_{|u| \le 1} \mathrm{e}(p^{-nt}\psi(u_0 + p^t u))\, du = \sum_{x_0 \in S} \sum_{\substack{u_0 \in \mathbb{Z}/p^t\mathbb{Z} \\ u_0 \equiv x_0 \bmod p}} p^{-t} \mathrm{e}(p^{-nt}\psi(u_0)) T_{x_0,u_0}$$

where

$$T_{x_0,u_0} \;:=\; \int_{|u| \le 1} \mathrm{e}\big(p^{-(n-1)t} \sum_{r=1}^{n-1} \frac{1}{r!} \psi^{(r)}(u_0) p^{t(r-1)} u^r\big) du.$$

We break up the sum over

$$R \;:=\; \{(x_0, u_0) \in S \times \mathbb{Z}/p^t\mathbb{Z} : x_0 \equiv u_0 \bmod p\} \;=\; R_1 \cup \cdots \cup R_n$$

into n disjoint sets where

$$R_1 \;=\; \{(x_0, u_0) \in R : |\psi^{(n-1)}(u_0)| \le p^{-t}\},$$

$$R_2 \;=\; \{(x_0, u_0) \in R : |\psi^{(n-1)}(u_0)| > p^{-t} \text{ and } |\psi^{(n-2)}(u_0)| \le p^{-t} |\psi^{(n-1)}(u_0)|\}$$

$$\vdots$$

$$R_{n-1} \;=\; \Big\{(x_0,u_0)\in R : |\psi''(u_0)| > p^{-t}|\psi'''(u_0)| > \cdots > p^{-(n-2)t} \\ \text{and}\;\; |\psi'(u_0)| \le p^{-t}|\psi''(u_0)|\Big\},$$

and

$$R_n = \Big\{(x_0,u_0)\in R : |\psi'(u_0)| > p^{-t}|\psi''(u_0)| > \cdots > p^{-(n-2)t}|\psi^{(n-1)}(u_0)| > p^{-(n-1)t}\Big\}.$$

We make the following claim:

- $\#R_j \le \deg(\psi),\ 1\le j\le n-1$; and
- $T_{x_0,u_0}=0$ for every $(x_0,u_0)\in R_n$.

For $j=1$, we apply the classical Hensel lemma (the $L=1$ case in Lemma 7.1) to $g(x)=\psi^{(n-1)}(x)$ to deduce that for every $(x_0,u_0)\in R_1$, there exists a unique $x\in\mathbb{Z}_p$ such that $\psi^{(n-1)}(x)=0$ and $x\equiv u_0 \bmod p^t$. Hence $\#R_1\le\deg(g)\le\deg(\psi)$.

Next for $(x_0,u_0)\in R_2$, consider $g(x)=\psi^{(n-2)}(x)$ so that $|g(u_0)|\le p^{-t}|g'(u_0)|$ and $|g'(u_0)|>p^{-t}$. Once again the classical version of Hensel implies that there exists a unique $x\in\mathbb{Z}_p$ such that $g(x)=0$ and $x\equiv u_0 \bmod p^t$. Hence $\#R_2\le\deg(g)\le\deg(\psi)$.

Now for $(x_0,u_0)\in R_j$ with $3\le j\le n-1$, consider $g(x)=\psi^{(n-j)}(x)$ so that $|g(u_0)|\le p^{-t}|g'(u_0)|$ and $|g'(u_0)|>p^{-t}|g''(u_0)|>\cdots>p^{-(j-1)t}$. Applying Lemma 7.1 with $L=j-1$ shows that there exists a unique $x\in\mathbb{Z}_p$ with $g(x)=0$ and $x\equiv u_0 \bmod p^t$. Hence $\#R_j\le\deg(g)\le\deg(\psi)$.

Finally for $(x_0,u_0)\in R_n$, we define $\sigma=t(n-1)-t-\nu$ where $p^{-\nu}:=|\psi''(u_0)|$. Note that $(x_0,u_0)\in R_n$ implies that $|\psi''(u_0)|>p^{-(n-1)t+t}$ and so $t+\nu<(n-1)t$, implying $\sigma\ge 1$. Hence, setting

$$\Psi(u) \;:=\; \sum_{r=1}^{n-1}\frac{1}{r!}\psi^{(r)}(u_0)p^{t(r-1)}u^r,$$

we have

$$T_{x_0,u_0} \;=\; \int_{|u|\le 1}\mathrm{e}(p^{-(n-1)t}\Psi(u))du \;=\; \sum_{w\in\mathbb{Z}/p^\sigma\mathbb{Z}}\int_{B_{p^{-\sigma}}(w)}\mathrm{e}(p^{-(n-1)t}\Psi(u))du$$

$$= \sum_{w\in\mathbb{Z}/p^\sigma\mathbb{Z}} p^{-\sigma}\int_{|y|\le 1}\mathrm{e}(p^{-(n-1)t}\Psi(w+p^\sigma y))\,dy.$$

Now observe that $\Psi(w+p^\sigma y)=\Psi(w)+p^\sigma\psi'(u_0)y+$

$$\frac{1}{2}\psi''(u_0)p^t\big((w+p^\sigma y)^2-w^2\big)+\cdots+\frac{1}{(n-1)!}\psi^{(n-1)}(u_0)p^{t(n-2)}\big((w+p^\sigma y)^{n-1}-w^{n-1}\big).$$

However since $(x_0,u_0)\in R_n$,

$$\left|\frac{1}{2}\psi''(u_0)p^t\big((w+p^\sigma y)^2-w^2\big)\right| = p^{-t-\sigma-\nu} = \; p^{-t(n-1)}$$

and, by comparing $\psi^{(j)}(u_0)$ to $\psi''(u_0)$ and using the fact that $\sigma > 1$, we have for $3 \leq j \leq n-1$:

$$\begin{aligned}
&\left|\frac{1}{j!}\psi^{(j)}(u_0)p^{t(j-1)}\big((w+p^\sigma y)^j - w^j\big)\right| \\
&\leq p^{(j-2)t}\left|\psi''(u_0)p^{t(j-1)}\big((w+p^\sigma y)^j - w^j\big)\right| \\
&\leq \left|\psi''(u_0)p^t p^\sigma\right| \\
&= p^{-t(n-1)}.
\end{aligned}$$

This means that the $j \geq 2$ terms in the sum defining Ψ are divisible by $p^{t(n-1)}$. Hence

$$T_{x_0,u_0} \;=\; \sum_{w\in\mathbb{Z}/p^\sigma\mathbb{Z}} p^{-\sigma}\mathrm{e}(p^{-(n-1)t}\Psi(w)) \int_{|y|\leq 1} \mathrm{e}(p^{-(n-1)t+\sigma}\psi'(u_0)y)\,dy$$

and this last integral is equal to zero since $(x_0,u_0) \in R_n$ implies

$$|\psi'(u_0)| > p^{-t}|\psi''(u_0)| \;=\; p^{-t-\nu} \;=\; p^{\sigma-t(n-1)}$$

and so $p^{t(n-1)-\sigma} \nmid \psi'(u_0)$.

This establishes the claim which implies

$$|I| \;\leq\; \left|p^{-t}\sum_{j=1}^{n-1}\sum_{(x_0,u_0)\in R_j} \mathrm{e}(p^{-nt}\psi(u_0))T_{x_0,u_0}\right| \;\leq\; (n-1)\deg(\psi)p^{-t} \;=\; C\,p^{-s/n},$$

giving us (19).

References

[1] S. S. Abhyankar. *Algebraic geometry for scientists and engineers*, volume 35 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1990.

[2] A. A. Albarracín-Mantilla and E. León-Cardenal. Igusa's local zeta functions and exponential sums of arithmetically non degenerate polynomials. *J. Théor. Nombres Bordeaux*, 30(1):331–354, 2018.

[3] V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko. *Singularities of differentiable maps. Volume 2*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2012. Monodromy and asymptotics of integrals, Translated from the Russian by Hugh Porteous and revised by the authors and James Montaldi, Reprint of the 1988 translation.

[4] R. Cluckers. Igusa and Denef-Sperber conjectures on nondegenerate p-adic exponential sums. *Duke Math. J.*, 141(1):205–216, 2008.

[5] R. Cluckers. Exponential sums: questions by Denef, Sperber, and Igusa. *Trans. Amer. Math. Soc.*, 362(7):3745–3756, 2010.

[6] R. Cluckers, M. Mustaţă, and K. H. Nguyen. Igusa's conjecture for exponential sums: optimal estimates for nonrational singularities. *Forum Math. Pi*, 7:e3, 2019.

[7] J. Denef and S. Sperber. Exponential sums mod p^n and Newton polyhedra. Number suppl., pages 55–63. 2001. A tribute to Maurice Boffa.

[8] I. A. Ikromov and D. Müller. On adapted coordinate systems. *Trans. Amer. Math. Soc.*, 363(6):2821–2848, 2011.

[9] I. A. Ikromov and D. Müller. Uniform estimates for the Fourier transform of surface carried measures in $\mathbb{R}^3$ and an application to Fourier restriction. *J. Fourier Anal. Appl.*, 17(6):1292–1332, 2011.

[10] V. N. Karpushkin. A theorem on uniform estimates for oscillatory integrals with a phase depending on two variables. *Trudy Sem. Petrovsk.*, (10):150–169, 238, 1984.

[11] A. G. Kušnirenko. The Newton polygon, and Milnor numbers. *Funkcional. Anal. i Priložen.*, 9(1):74–75, 1975.
[12] B. Lichtin. On a conjecture of Igusa. *Mathematika*, 59(2):399–425, 2013.
[13] R. Narasimhan. *Introduction to the theory of analytic spaces*. Lecture Notes in Mathematics, No. 25. Springer-Verlag, Berlin-New York, 1966.
[14] D. H. Phong, E. M. Stein, and J. A. Sturm. On the growth and stability of real-analytic functions. *Amer. J. Math.*, 121(3):519–554, 1999.
[15] A. M. Robert. *A course in p-adic analysis*, volume 198 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.
[16] A. N. Varčenko. Newton polyhedra and estimates of oscillatory integrals. *Funkcional. Anal. i Priložen.*, 10(3):13–38, 1976.
[17] A. N. Varčenko. Asymptotic Hodge structure on vanishing cohomology. *Izv. Akad. Nauk SSSR Ser. Mat.*, 45(3):540–591, 688, 1981.
[18] W. Veys. On the log canonical threshold and numerical data of a resolution in dimension 2. *Manuscripta Mathematica*, to appear.
[19] A. Weil. On some exponential sums. *Proc. Nat. Acad. Sci. U. S. A.*, 34:204–207, 1948.
[20] J. Wright. From oscillatory integrals and sublevel sets to polynomial congruences and character sums. *J. Geom. Anal.*, 21(1):224–240, 2011.
[21] J. Wright. Exponential sums and polynomial congruences in two variables: the quasi-homogeneous case. preprint, http://arxiv.org/abs/1202.2686.

Maxwell Institute of Mathematical Sciences and the School of Mathematics, University of Edinburgh, JCMB, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, Scotland

Email address: robert.fraser@@ed.ac.uk

Maxwell Institute of Mathematical Sciences and the School of Mathematics, University of Edinburgh, JCMB, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, Scotland

Email address: J.R.Wright@@ed.ac.uk