

EXPONENTIAL SUMS: QUESTIONS BY DENEUF, SPERBER, AND IGUSA

RAF CLUCKERS

ABSTRACT. We prove the remaining part of the conjecture by Denef and Sperber [Denef, J. and Sperber, S., *Exponential sums mod p^n and Newton polyhedra*, Bull. Belg. Math. Soc., **suppl.** (2001) 55-63] on nondegenerate local exponential sums modulo p^m . We generalize Igusa’s conjecture in the introduction of [Igusa, J., *Lectures on forms of higher degree*, Lect. Math. Phys., Springer-Verlag, **59** (1978)] from the homogeneous to the quasi-homogeneous case and prove the nondegenerate case as well as the modulo p case. We generalize some results by Katz in [Katz, N. M., *Estimates for “singular” exponential sums*, Internat. Math. Res. Notices (1999) no. 16, 875-899] on finite field exponential sums to the quasi-homogeneous case.

1. INTRODUCTION

1.1. Global sums: From homogeneous to quasi-homogeneous. Let f be a polynomial over \mathbb{Z} in n variables. As in [1], [2], we look at the “global” exponential sums

$$S_f\left(\frac{1}{N}\right) := \frac{1}{N^n} \sum_{x \in \{0, \dots, N-1\}^n} \exp(2\pi i \frac{f(x)}{N}),$$

where N varies over the positive integers. In order to bound $|S_f(1/N)|$, it suffices to find bounds when $N = p^m$ for integers $m > 0$ and prime numbers p .

When f is homogeneous and nondegenerate w.r.t. its Newton polyhedron at zero, Igusa’s conjecture for the toric resolution of f conjectures that there exists $c > 0$ such that

$$\left|S_f\left(\frac{1}{p^m}\right)\right| < cp^{-\sigma m} m^{n-1},$$

for all primes p and integers $m > 0$, where σ is the largest rational number such that $(1/\sigma, \dots, 1/\sigma)$ lies on the Newton polyhedron at zero $\Delta_0(f)$ of f . In a later paper, Denef and Sperber conjectured for the same f that the stronger bound

$$\left|S_f\left(\frac{1}{p^m}\right)\right| < cp^{-\sigma m} m^{\kappa-1}$$

holds for some c , uniformly in large enough primes p and all integers $m > 0$, with κ the codimension in \mathbb{R}^n of the smallest face of $\Delta_0(f)$ which contains $(1/\sigma, \dots, 1/\sigma)$.

Received by the editors September 4, 2008.

2010 *Mathematics Subject Classification*. Primary 11L07, 11S40; Secondary 11L05.

Key words and phrases. Exponential sums, nondegenerate polynomials, Igusa’s conjecture on exponential sums, Igusa’s local zeta functions, motivic integration.

The author was a postdoctoral fellow of the Fund for Scientific Research - Flanders (Belgium) (F.W.O.).

©2009 American Mathematical Society
Reverts to public domain 28 years from publication

Both these conjectures are proved by Denef and Sperber [3] under the extra condition that no vertex of $\Delta_0(f)$ belongs to $\{0, 1\}^n$ and by the author [1] in general.

In this paper, we show that both these conjectures also hold when f is quasi-homogeneous and nondegenerate w.r.t. $\Delta_0(f)$. Quasi-homogeneous means that there exist integers $a_i > 0$ such that $f(x_1^{a_1}, \dots, x_n^{a_n})$ is homogeneous. This gives evidence to our conjecture of [2] that the analogue of Igusa's conjecture for exponential sums of [4] holds for all quasi-homogeneous polynomials, and not only for homogeneous ones.

We give some more evidence for this conjecture by proving the analogue of Igusa's conjecture (with the motivic oscillation index instead of the above σ , as in [2]), for all quasi-homogeneous polynomials (also degenerate ones) for $m = 1$, thereby generalizing [2] and some results by Katz of [5] to the quasi-homogeneous case.

This work generalizes most of the known evidence for Igusa's conjecture to evidence for its generalization to quasi-homogeneous polynomials. (To our knowledge, the case of isolated singularities of [4] is only done for homogeneous polynomials and remains open for quasi-homogeneous ones.)

1.2. Local sums. Since for more general f than quasihomogeneous f , Igusa's conjecture is not even conjecturally understood, Denef and Sperber designed a local variant of Igusa's conjecture which they conjecture to hold for all polynomials; see [3]. We treat only the nondegenerate case. Let g be a polynomial in n variables which is nondegenerate w.r.t. its Newton polyhedron at zero $\Delta_0(g)$. Denef and Sperber studied the "local" exponential sum

$$T_g\left(\frac{1}{p^m}\right) := \frac{1}{p^{mn}} \sum_{x \in \{jp \mid j=1, \dots, p^{m-1}\}} \exp(2\pi i \frac{g(x)}{p^m}),$$

for p a prime and $m > 0$ an integer, and conjectured that there exists c such that

$$|T_g\left(\frac{1}{p^m}\right)| < cp^{-\sigma(g)m} m^{\kappa(g)-1}$$

for large enough prime numbers p and all integers $m > 0$, with $\sigma(g)$ and $\kappa(g)$ as σ and κ but for g instead of for f . Denef and Sperber [3] proved this under the condition that no vertex of $\Delta_0(g)$ belongs to $\{0, 1\}^n$. In this paper we prove this conjecture of [3] for all polynomials g which are nondegenerate w.r.t. $\Delta_0(g)$ (thereby removing the condition of [3] that no vertex of $\Delta_0(g)$ belongs to $\{0, 1\}^n$).

1.3. Global sums: Igusa's question mod p . Theorem 3.2.1 below is the modulo p case for quasi-homogeneous polynomials of Igusa's conjecture for exponential sums in [4]. In this theorem, the polynomial h is not necessarily nondegenerate. It generalizes the main result of [2] from the homogeneous to the quasi-homogeneous case. The modulo p^2 case is already proven in [2] for all polynomials. In order to prove Theorem 3.2.1 we generalize some results by Katz [5]; see Theorem 7.4 below.

2. A DICTIONARY

2.1. From finite sums to p -adic integrals. For \mathbb{Q}_p the field of p -adic numbers, $x = (x_1, \dots, x_n)$ variables running over \mathbb{Q}_p^n , let $|dx|$ be the unique (real-valued) Haar measure on \mathbb{Q}_p^n so normalized that \mathbb{Z}_p^n has measure one. For $a \in \mathbb{Q}_p$, the complex number

$$\exp(2\pi i a \bmod \mathbb{Z}_p) := \exp(2\pi i a')$$

does not depend on the choice of representant a' in $\mathbb{Z}[1/p]$ of $a \bmod \mathbb{Z}_p$, and will be denoted by $\exp(2\pi ia)$.

Then, for f a polynomial in n variables over \mathbb{Q} , one has the equality

$$S_f\left(\frac{1}{p^m}\right) = \int_{\mathbb{Z}_p^n} \exp(2\pi i \frac{f(x)}{p^m}) |dx|,$$

and we will more generally consider for all $y \in \mathbb{Q}_p$ the integral

$$S_{f, \mathbb{Q}_p}(y) := \int_{\mathbb{Z}_p^n} \exp(2\pi i y f(x)) |dx|.$$

2.2. From \mathbb{Q}_p to finite field extensions of \mathbb{Q}_p and to $\mathbb{F}_q((t))$. If f is a polynomial over $\mathcal{O}[1/N]$ for some ring of integers \mathcal{O} and $N > 0$ an integer, when K is a non-Archimedean local field which is an algebra over $\mathcal{O}[1/N]$ (thus p -adic or of the form $\mathbb{F}_q((t))$), and $\psi_K : K \rightarrow \mathbb{C}^\times$ is a nontrivial additive character which is 1 on \mathcal{O}_K and different from 1 on some element of order -1 , then we write for $y \in K$,

$$S_{f, K, \psi_K}(y) := \int_{\mathcal{O}_K^n} \psi_K(yf(x)) |dx|,$$

with \mathcal{O}_K the valuation ring of K and $|dx|$ the Haar measure, normalized so that \mathcal{O}_K^n has measure one.

Similarly, if g is a polynomial over $\mathcal{O}[1/N]$ and with K and ψ_K as above in this section, we write

$$T_{g, K, \psi_K}(y) := \int_{\mathcal{M}_K^n} \psi_K(yg(x)) |dx|$$

with \mathcal{M}_K the maximal ideal of \mathcal{O}_K and \mathcal{M}_K^n the n -fold Cartesian product of \mathcal{M}_K with itself.

For $K = \mathbb{Q}_p$ one has

$$T_g\left(\frac{1}{p^m}\right) = T_{g, \mathbb{Q}_p, \exp(2\pi i \cdot)}\left(\frac{1}{p^m}\right)$$

and

$$S_f\left(\frac{1}{p^m}\right) = S_{f, \mathbb{Q}_p, \exp(2\pi i \cdot)}\left(\frac{1}{p^m}\right).$$

Write $|\cdot|_K$ for the standard norm on K . So, the norm of a uniformizer of \mathcal{O}_K equals $\frac{1}{q_K}$ with q_K the number of elements in the residue field of \mathcal{O}_K . Let $|\cdot|$ denote the norm on \mathbb{C} .

2.3. Nondegenerate polynomials. Let f be a nonconstant polynomial over \mathbb{C} in n variables with $f(0) = 0$.¹ Write $f(x) = \sum_{i \in \mathbb{N}^n} a_i x^i$ with $a_i \in \mathbb{C}$. The *global Newton polyhedron* $\Delta_{\text{global}}(f)$ of f is the convex hull of the support $\text{Supp}(f)$ of f , with

$$\text{Supp}(f) := \{i \mid i \in \mathbb{N}^n, a_i \neq 0\}.$$

The *Newton polyhedron* $\Delta_0(f)$ of f at the origin is

$$\Delta_0(f) := \Delta_{\text{global}}(f) + \mathbb{R}_+^n$$

¹When $f(0) \neq 0$, then there is no harm in replacing f by $f - f(0)$: all corresponding changes in the paper are easily made.

with $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ and $A + B = \{a + b \mid a \in A, b \in B\}$ for $A, B \subset \mathbb{R}^n$. For a subset I of \mathbb{R}^n define

$$f_I(x) := \sum_{i \in I \cap \mathbb{N}^n} a_i x^i.$$

By the *faces* of I we mean I itself and each nonempty convex set of the form

$$\{x \in I \mid L(x) = 0\},$$

where $L(x) = b_0 + \sum_{i=1}^n b_i x_i$ with $b_i \in \mathbb{R}$ is such that $L(x) \geq 0$ for each $x \in I$. For \mathcal{I} a collection of subsets of \mathbb{R}^n , call f *nondegenerate with respect to \mathcal{I}* when f_I has no critical points² on $(\mathbb{C}^\times)^n$ for each I in \mathcal{I} , where $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$. As is common terminology, call f nondegenerate w.r.t. $\Delta_0(f)$ when f is nondegenerate w.r.t. the compact faces of $\Delta_0(f)$.

For $k \in \mathbb{R}_+^n$ put

$$\begin{aligned} \nu(k) &= k_1 + k_2 + \dots + k_n, \\ N(f)(k) &= \min_{i \in \Delta_0(f)} k \cdot i, \\ F(f)(k) &= \{i \in \Delta_0(f) \mid k \cdot i = N(f)(k)\}, \end{aligned}$$

where $k \cdot i$ is the standard inproduct on \mathbb{R}^n , and where i also runs over \mathbb{R}_+^n . Denote by $F_0(f)$ the smallest face of $\Delta_0(f)$ which has nonempty intersection with the diagonal $\{(t, \dots, t) \mid t \in \mathbb{R}\}$ and let $(1/\sigma(f), \dots, 1/\sigma(f))$ be the intersection point of the diagonal with $F_0(f)$. Let $\kappa(f)$ be the codimension of $F_0(f)$ in \mathbb{R}^n . If there is no possible confusion, we write σ instead of $\sigma(f)$, $N(k)$ instead of $N(f)(k)$, $F(k)$ for $F(f)(k)$, and κ for $\kappa(f)$.

2.4. Notation. Often in this paper, f is a quasi-homogeneous polynomial and g a polynomial, both over $\mathcal{O}[1/N]$ and in n variables, with \mathcal{O} a ring of integers and $N > 0$ an integer, such that f and g are nonzero, $f(0) = g(0) = 0$ and such that f is nondegenerate w.r.t. $\Delta_0(f)$ and g is nondegenerate w.r.t. $\Delta_0(g)$. If f and g are such we will say that they are as in section 2.4.

By K is usually meant a non-Archimedean local field that is an algebra over $\mathcal{O}[1/N]$ and by q_K the number of elements in the residue field of \mathcal{O}_K . If f and g , K and q_K are such, we will say that they are as in section 2.4.

3. THE MAIN RESULTS

Let f , g , and K be as in section 2.4 and use the notation of sections 2.2 and 2.3.

3.1. Theorem. *There exists c , only depending on $\Delta_0(f)$, resp. on $\Delta_0(g)$, such that for all K with large enough residual characteristic, all ψ_K as in section 2.2, and all y in K with $\text{ord}_K(y) < 0$,*

$$(3.1.1) \quad |S_{f,K,\psi_K}(y)| < c|y|_K^{-\sigma(f)} |\text{ord}_K(y)|^{\kappa(f)-1},$$

resp.

$$(3.1.2) \quad |T_{g,K,\psi_K}(y)| < c|y|_K^{-\sigma(g)} |\text{ord}_K(y)|^{\kappa(g)-1}.$$

²Sometimes one requires that f_I has no *singular* points on $(\mathbb{C}^\times)^n$, instead of no critical points on $(\mathbb{C}^\times)^n$. Recall that x_0 is a critical point of f_I if and only if $\text{grad } f_I(x_0)$ is the zero vector.

In the case that f is moreover homogeneous, a slightly weaker form than (3.1.1) has been conjectured by Igusa [4] and (3.1.1), (3.1.2) have been conjectured by Denef and Sperber [3]. The similar conjecture by Igusa mentioned in the beginning of the introduction (namely for all primes p) then follows by standard arguments; cf. [2]. Denef and Sperber prove (3.1.2) when no vertex of $\Delta_0(g)$ belongs to $\{0, 1\}^n$. As briefly mentioned in the introduction, (3.1.1) for homogeneous f is proven in [3] under the condition that no vertex of $\Delta_0(f)$ belongs to $\{0, 1\}^n$ and in [1] in full generality. Recall that in Theorem 3.1, f is allowed to be quasi-homogeneous and g is general.

3.2. Let h be any quasi-homogeneous polynomial over $\mathcal{O}[1/N]$ (so, h is not necessarily nondegenerate). Let α_h be the motivic oscillation index of h as defined in [2] according to a suggestion by Jan Denef. As in [2], we conjecture that there exists c such that for all K with large enough residual characteristic (where K runs over non-Archimedean local fields that are algebras over $\mathcal{O}[1/N]$), all ψ_K as in section 2.2, and all y in K with $\text{ord}_K(y) < 0$,

$$(3.2.1) \quad |S_{h,K,\psi_K}(y)| < c|y|_K^{\alpha_h} |\text{ord}_K(y)|^{n-1}.$$

Generalizing the main theorem of [2] to the quasi-homogeneous case, we prove the following evidence for this conjecture.

3.2.1. Theorem. *Statement (3.2.1) holds for all $y \in K$ of orders -1 and -2 . That is, there exists c such that for all K of large enough residual characteristic with K an algebra over $\mathcal{O}[1/N]$, all ψ_K as in section 2.2, and all $y \in K$ of order -1 or order -2 , statement (3.2.1) holds.*

Note that (3.1.1) of Theorem 3.1 also constitutes evidence for this conjecture.

3.3. A third main result is Theorem 7.4 on finite field exponential sums and generalizes some results by Katz [5] to the quasi-homogeneous case (not necessarily nondegenerate). Proposition 6.2 represents a new kind (a similar bound in the homogeneous case was introduced in [1]) of very simple bounds for nondegenerate finite field exponential sums for quasi-homogeneous polynomials, based on the combinatorics of the Newton Polyhedron.

3.4. The underlying idea of the proofs of these questions, apart from the ideas and results of [3], [5], [1], and [2], is that one can make a homogeneous polynomial out of a quasi-homogeneous polynomial $f(x)$ in finitely many steps, by replacing one of the variables x_i by yx_i with y a new variable, that is, by replacing $f(x)$ by

$$f_1(x, y) := f(x_1, x_2, \dots, x_{i-1}, x_i y, x_{i+1}, \dots, x_n)$$

and so on, and then one compares the considered objects for f and for f_1 .

4. A DENEFF - SPERBER FORMULA FOR S_{f,K,ψ_K} AND T_{g,K,ψ_K}

The following proposition has essentially the same proof as Proposition (2.1) of [3], but is slightly more general. We give the proof for the convenience of the reader.

4.1. Proposition. *Let f, g, K , and q_K be as in section 2.4. There exists $M > 0$ such that if the residual characteristic of K is greater than M , then for all y in K with $\text{ord}_K(y) < 0$,*

$$(4.1.1) \quad S_{f,K,\psi_K}(y) = d_K \cdot \sum_{\tau \text{ face of } \Delta_0(f)} (A_f(K, y, \tau) + E(K, y, f_\tau, \psi_K) B_f(K, y, \tau))$$

and

$$(4.1.2) \quad T_{g,K,\psi_K}(y) = d_K \cdot \sum_{\substack{\tau \text{ compact} \\ \text{face of } \Delta_0(\mathfrak{g})}} (A_g(K, y, \tau) + E(K, y, g_\tau, \psi_K)B_g(K, y, \tau)),$$

with

$$\begin{aligned} d_K &:= (1 - q_K^{-1})^n, \\ A_f(K, y, \tau) &:= \sum_{\substack{k \in \mathbb{N}^n \\ F(f)(k) = \tau \\ N(f)(k) \geq -\text{ord}(y)}} q_K^{-\nu(k)}, \\ B_f(K, y, \tau) &:= \sum_{\substack{k \in \mathbb{N}^n \\ F(f)(k) = \tau \\ N(f)(k) = -\text{ord}(y) - 1}} q_K^{-\nu(k)}, \end{aligned}$$

and similarly for $A_g(K, y, \tau)$ and $B_g(K, y, \tau)$, and, for $h(x)$ either f_τ or g_τ ,

$$(4.1.3) \quad E(K, y, h, \psi_K) = \frac{1}{(q_K - 1)^n} \sum_{x \in (\mathbb{G}_m(\mathbb{F}_{q_K}))^n} \varphi_y(h(x)),$$

with $\varphi_y : \mathbb{F}_{q_K} \rightarrow \mathbb{C}^\times$ a nontrivial additive character on \mathbb{F}_{q_K} sending b to $\psi_K(y\pi_K^{-\text{ord}_y-1}b')$, with b' a representant in \mathcal{O}_K of b and π_K a uniformizer of \mathcal{O}_K .

Proof. Rewrite the definition

$$S_{f,K,\psi_K}(y) = \int_{\mathcal{O}_K^n} \psi_K(yf(x))|dx|$$

as

$$S_{f,K,\psi_K}(y) = \sum_{\tau \text{ face of } \Delta_0(f)} \sum_{\substack{k \in \mathbb{N}^n \\ F(f)(k) = \tau}} \int_{\text{ord}_K x=k} \psi_K(yf(x))|dx|.$$

Put $x_j = \pi_K^{k_j} u_j$ for $k \in \mathbb{N}^n$, with π_K a uniformizer of \mathcal{O}_K . Then $|dx| = q_K^{-\nu(k)}|du|$ and

$$f(x) = \pi_K^{N(f)(k)}(f_{F(f)(k)}(u) + \pi_K(\dots)),$$

for x with $\text{ord}_K x = k$, where the dots take values in \mathcal{O}_K . Hence, $S_{f,K,\psi_K}(y)$ equals

$$(4.1.4) \quad \sum_{\tau \text{ face of } \Delta_0(f)} \sum_{\substack{k \in \mathbb{N}^n \\ F(f)(k) = \tau}} q_K^{-\nu(k)} \int_{u \in (\mathcal{O}_K^\times)^n} \psi_K(y\pi_K^{N(f)(k)}(f_\tau(u) + \pi_K(\dots)))|du|,$$

with \mathcal{O}_K^\times the group of units in \mathcal{O}_K . Because of the nondegenerateness assumptions, for τ any face of $\Delta_0(f)$ and when the residue field characteristic of K is large enough, the reduction $f_\tau \text{ mod } \mathcal{M}_K$ has no critical points on $(\mathbb{F}_{q_K}^\times)^n$ (this holds indeed for all faces and not only for the compact faces by the quasi-homogeneity of f). Hence, the integral in (4.1.4) is zero whenever $\text{ord}_K(y) + N(f)(k) \leq -2$. When $\text{ord}_K(y) + N(f)(k) \geq 0$, the integral over $(\mathcal{O}_K^\times)^n$ in (4.1.4) is just the measure of $(\mathcal{O}_K^\times)^n$ and thus equals $(1 - q_K^{-1})^n$. When $\text{ord}_K(y) + N(f)(k) = -1$, the integral

over $(\mathcal{O}_K^\times)^n$ in (4.1.4) equals $q_K^{-n}(q_K - 1)^n E(K, y, h, \psi_K)$. Equation (4.1.1) now follows from the fact that $S_{f,K,\psi_K}(y)$ equals (4.1.4).

Equation (4.1.2) follows in a similar way. Namely, one writes

$$T_{g,K,\psi_K}(y) = \sum_{\tau \text{ face of } \Delta_0(g)} \sum_{\substack{k \in (\mathbb{N} \setminus \{0\})^n \\ F(g)(k) = \tau}} \int_{\text{ord}_K x=k} \psi_K(yg(x))|dx|,$$

which equals

$$T_{g,K,\psi_K}(y) = \sum_{\tau \text{ compact face of } \Delta_0(g)} \sum_{\substack{k \in \mathbb{N}^n \\ F(g)(k) = \tau}} \int_{\text{ord}_K x=k} \psi_K(yg(x))|dx|,$$

and one proceeds as for f . □

5. BOUNDS FOR $\nu(k)$, A_f , B_f , A_g , AND B_g

We recall two results from [1], about lower bounds for $\nu(k)$ in terms of $N(f)(k)$ and $N(g)(k)$, and upper bounds for A_f , B_f , A_g , and B_g .

5.1. Proposition ([1], Theorem 4.1). *Let h be any nonzero polynomial in n variables with $h(0) = 0$. Let τ be a face of $\Delta_0(h)$. Then one has for all k in \mathbb{N}^n with $F(h)(k) = \tau$ that*

$$(5.1.1) \quad \nu(k) \geq \sigma(h)(N(h)(k) + 1) - \sigma(h_\tau).$$

The main point in Proposition 5.1, and one of the main differences with the approach from [3], is that one subtracts $\sigma(h_\tau)$ on the right hand side of (5.1.1). Subtracting $\sigma(h)$ would yield trivial bounds since one has $\nu(k) \geq \sigma(h)N(h)(k)$ for all $k \in \mathbb{N}_+^n$.

The bounds for A_f and A_g in Proposition 5.2 are essentially proven in [3] and follow from Lemma (3.3) of [3], which is recalled in Lemma 5.3 below (in [1] the proof of these bounds is repeated from [3]). The bounds for B_f and B_g are essentially proven in [1] and follow from Lemma (3.3) of [3] and Proposition 5.1.

5.2. Proposition ([1], [3]). *Let f , g , A , and B be as in Proposition 4.1. Then there exists a real number $c > 0$ such that for all K as in section 2.4, all faces τ of $\Delta_0(f)$, resp. of $\Delta_0(g)$, and all y in K with $\text{ord}_K(y) < 0$,*

$$(5.2.1) \quad A_f(K, y, \tau) \leq c|y|_K^{-\sigma(f)}|\text{ord}_K(y)|^{\kappa(f)-1}$$

and

$$(5.2.2) \quad B_f(K, y, \tau) \leq c|y|_K^{-\sigma(f)}q_K^{\sigma(f_\tau)}|\text{ord}_K(y)|^{\kappa(f)-1},$$

resp.,

$$(5.2.3) \quad A_g(K, y, \tau) \leq c|y|_K^{-\sigma(g)}|\text{ord}_K(y)|^{\kappa(g)-1}$$

and

$$(5.2.4) \quad B_g(K, y, \tau) \leq c|y|_K^{-\sigma(g)}q_K^{\sigma(g_\tau)}|\text{ord}_K(y)|^{\kappa(g)-1}.$$

Moreover, one can choose c depending on $\Delta_0(f)$ and $\Delta_0(g)$ only.

Proof. We give the proof for B_g for the convenience of the reader. The proof for B_f is similar, and the proofs for A_f and A_g can be found in [3]. To derive (5.2.4) from Lemma 5.3 and Proposition 5.1, use for C the topological closure of the convex hull of $\{0\} \cup \{k \in \mathbb{N}^n \mid F(g)(k) = \tau\}$, and note that $C^{\text{int}} \cap \mathbb{N}^n = \{k \in \mathbb{N}^n \mid F(g)(k) = \tau\}$. Clearly $\kappa(g) \geq 1$ and $\kappa(g) \geq \dim\{k \in C \mid \nu(k) = N(g)(k)\sigma\}$. By (5.1.1), $\nu(k) \geq N(g)(k)\sigma(g) + \sigma(g) - \sigma(g_\tau)$ for all $k \in C^{\text{int}} \cap \mathbb{N}^n$. So for γ one can take $\sigma(g) - \sigma(g_\tau)$, which is nonnegative. \square

5.3. Lemma ([3], Lemma (3.3)). *Let C be a convex polyhedral cone in \mathbb{R}_+^n generated by vectors in \mathbb{N}^n , and let L be a linear form in n variables with coefficients in \mathbb{N} . We denote by C^{int} the interior of C in the sense of Newton polyhedra. Let $\sigma > 0$ and $\gamma \geq 0$ be real numbers satisfying*

$$(5.3.1) \quad \nu(k) \geq L(k)\sigma + \gamma, \quad \text{for all } k \in C^{\text{int}} \cap \mathbb{N}^n.$$

Put

$$e = \dim\{k \in C \mid \nu(k) = L(k)\sigma\}.$$

Then there exists a real number $c > 0$ such that for all $m \in \mathbb{N}$ and for all $q \in \mathbb{R}$ with $q \geq 2$,

$$(5.3.2) \quad \sum_{\substack{k \in C^{\text{int}} \cap \mathbb{N}^n \\ L(k) = m}} q^{-\nu(k)} \leq cq^{-m\sigma - \gamma} (m + 1)^{\max(0, e - 1)}.$$

6. ESTIMATES FOR FINITE FIELD EXPONENTIAL SUMS

For each prime number p , let ψ_p be a nontrivial additive character from \mathbb{F}_p to \mathbb{C}^\times , and likewise, for each power q of p , let ψ_q be a nontrivial additive character from \mathbb{F}_q to \mathbb{C}^\times .

6.1. Lemma. *Let f be a polynomial over some number field in the n variables x_1, \dots, x_n . (Thus f is not necessarily quasi-homogeneous.) Let $g(x_1, \dots, x_n, y)$ be the polynomial $f(x_1y, x_2, \dots, x_n)$ in the $n + 1$ variables (x, y) . Suppose that $f(0) = g(0) = 0$. Then $\sigma(f) = \sigma(g)$. Moreover, if f is nondegenerate w.r.t. $\Delta_0(f)$, then g is nondegenerate w.r.t. $\Delta_0(g)$.*

Proof. Clearly, for any point $P = (p_1, \dots, p_n)$ on $\Delta_0(f)$, the point $P' = (p_1, \dots, p_n, p_1)$ lies on $\Delta_0(g)$. Indeed, this holds for points P in the support of f , thus also for P in the convex hull of the support of f , and thus for general P in $\Delta_0(f)$. Vice versa, for any point $Q = (q_1, \dots, q_{n+1})$ on $\Delta_0(g)$, the point $Q' = (q_1, \dots, q_n)$ lies in $\Delta_0(f)$. From this it follows that $\sigma(f) = \sigma(g)$. The statement about the nondegenerateness is immediate since we performed a coordinate transformation on the torus \mathbb{G}_m^{n+1} induced by a \mathbb{Z} -module automorphism of \mathbb{Z}^{n+1} . \square

The following generalizes Corollary 6.4 of [1] from the homogeneous case to the quasi-homogeneous case. Corollary 6.4 of [1] is proven in [1] using results by Katz [5], by Segers [6], and by the author [2].

6.2. Proposition. *Let f be as in section 2.4. (In particular, f is quasi-homogeneous.) Then there exists a $a > 0$ such that for each large enough prime p , and each power q of p such that \mathbb{F}_q is an algebra over $\mathcal{O}[1/N]$, one has*

$$(6.2.1) \quad \left| \frac{1}{(q-1)^n} \sum_{x \in \mathbb{G}_m^n(\mathbb{F}_q)} \psi_q(f(x)) \right| < aq^{-\sigma(f)}$$

and

$$(6.2.2) \quad \left| \frac{1}{q^n} \sum_{x \in \mathbb{A}^n(\mathbb{F}_q)} \psi_q(f(x)) \right| < a q^{-\sigma(f)},$$

for all choices of ψ_q . Moreover, a can be chosen depending on $\Delta_0(f)$ only.

Proof. We first prove (6.2.1). First suppose that f is homogeneous. If moreover $f(x)$ is of the form $\sum_{i=1}^n a_i x_i$, (6.2.1) follows from the obvious inequality $\sigma(f) \leq n$. If f is nonlinear (that is, $f(x)$ is not of the form $\sum_{i=1}^n a_i x_i$) and homogeneous, then the proposition holds by Corollary 6.4 of [1]; see also section 9 of [1]. So, the case of homogeneous f is done. Now let f be quasi-homogeneous. After finitely many steps as in going from f to g in Lemma 6.1 and renumbering the variables x_i , one can go from a quasi-homogeneous polynomial to a homogeneous polynomial h in possibly more variables. Going from f to g as in Lemma 6.1 consists of two steps: first, one defines a polynomial $f_0(x, y)$ which equals $f(x)$, that is, one considers f as a polynomial in one more variable y . Then one performs the transformation on the torus, (x, y) to $(x_1 y, x_2, \dots, x_n, y)$, to obtain g . Under both these steps, the sum (6.2.1) remains unaltered, since one divides by $(q - 1)^n$ with n the number of variables in the left hand side of (6.2.1), and since in the second step we just perform a transformation on the torus. Lemma 6.1 together with the homogeneous case now proves (6.2.1).

The inequality (6.2.2) is not used in this paper and follows from (3.1.1) of Theorem 3.1 with argument y of order -1 . □

7. A COROLLARY OF RESULTS BY KATZ

The results in this section will only serve to prove Theorem 3.2.1, not to prove Theorem 3.1. In this section as well as in Theorem 3.2.1 we focus on quasi-homogeneous polynomials in general (thus not necessarily nondegenerate ones).

Call a collection of polynomials f_i in n variables *quasi-homogeneous with the same weights* if there are integers $a_j > 0$, not depending on i , such that $f_i(x_1^{a_1}, \dots, x_n^{a_n})$ is homogeneous for each i .

7.1. Lemma. *Let f_i be quasi-homogeneous polynomials with the same weights, for $i = 1, \dots, d$. Suppose that f_d is nonconstant. Let X be the locus of the f_i for $i = 1, \dots, d - 1$ in \mathbb{A}^n and let Y be the locus of the f_i for $i = 1, \dots, d$ in \mathbb{A}^n . Then*

$$\dim X \leq \dim Y + 1,$$

where the dimension of the empty scheme is -1 .

Proof. Clearly this statement holds when the f_i are homogeneous polynomials, by intersection theory. Since the f_i have the same weights, let $a_j, j = 1, \dots, n$ be positive integers such that each of the $f_i(x_1^{a_1}, \dots, x_n^{a_n})$ is homogeneous. Consider the application

$$G : \mathbb{A}^n \rightarrow \mathbb{A}^n : x \mapsto (x_1^{a_1}, \dots, x_n^{a_n}).$$

The application G is finite to one; hence $G^{-1}(X)$ has the same dimension as X , and $G^{-1}(Y)$ has the same dimension as Y . By the above statement for homogeneous polynomials, the dimension of $G^{-1}(X)$ is less than or equal to the dimension of $G^{-1}(Y)$ plus 1. The lemma is proven. □

7.2. Corollary. *Let f be a quasi-homogeneous polynomial over a field of characteristic zero in the n variables x_1, \dots, x_n . Let $g(x_1, \dots, x_n, y)$ be the polynomial $f(x_1y, x_2, \dots, x_n)$ in $n+1$ variables. Let C_f be the closed subscheme of \mathbb{A}^n given by $\text{grad } f = 0$, and let C_g be the closed subscheme of \mathbb{A}^{n+1} given by $\text{grad } g = 0$. Suppose that C_f contains the point 0 . Then*

$$\dim C_g \leq \dim C_f + 1,$$

where the dimension of the empty scheme is said to be -1 .

Proof. For $i = 1, \dots, n$ let $f_i(x)$ be the polynomial $\frac{\partial f}{\partial x_i}(x)$, let $g_i(x, y)$ be the polynomial $\frac{\partial g}{\partial x_i}(x, y)$, and let $g_0(x, y)$ be $\frac{\partial g}{\partial y}(x, y)$. Note that the polynomials f_i are quasi-homogeneous with the same weights. Similarly, the polynomials g_i are quasi-homogeneous with the same weights.

By the chain rule for differentiation one has

$$\begin{aligned} g_0(x, y) &= x_1 f_1(x_1y, x_2, \dots, x_n), \\ g_1(x, y) &= y f_1(x_1y, x_2, \dots, x_n), \end{aligned}$$

and, for $i > 1$,

$$g_i(x, y) = f_i(x_1y, x_2, \dots, x_n).$$

On the part U_1 of \mathbb{A}^{n+1} where $x_1y \neq 0$ the bound is clear. Indeed, $U_1 \cap C_g$ is the subscheme of \mathbb{A}^{n+1} given by $x_1y \neq 0$ and $f_i(x_1y, x_2, \dots, x_n) = 0$ for $i = 1, \dots, n$, which has dimension at most $\dim C_f + 1$.

Now work on the part U_2 where $x_1 = 0$ and $y \neq 0$. Then $U_2 \cap C_g$ is the subscheme of \mathbb{A}^{n+1} given by $f_i(x) = 0$ for $i = 1, \dots, n$, $y \neq 0$, and $x_1 = 0$, which has dimension at most $\dim C_f + 1$.

Finally work on the part U_3 , where $x_1 = 0$ and $y = 0$. Then $U_3 \cap C_g$ is the subscheme of \mathbb{A}^{n+1} given by $x_1 = y = 0$ and $0 = f_i(x)$ for $i = 2, \dots, n$. We know that f_1 is nonconstant since C_f contains 0 . So, the subscheme of \mathbb{A}^n given by $0 = f_i(x)$ for $i = 2, \dots, n$ has at most dimension $\dim C_f + 1$ by Lemma 7.1. Hence, the dimension of $C_g \cap U_3$ has at most dimension $\dim C_f + 1$ and the corollary is proven. \square

The following is a trivial lemma.

7.3. Lemma. *Let f be a nonconstant, quasi-homogeneous polynomial in n variables x_1, \dots, x_n over a field k of characteristic zero. Then exactly one of the following two statements holds: either 0 is a critical point of f , or f contains a term of the form $a_i x_i$ with $a_i \neq 0$ in k for some $i \in \{1, \dots, n\}$.*

From Lemma 7.1 and Corollary 7.2, Katz' results of [5] can be generalized to quasi-homogeneous polynomials; see Theorem 7.4. Namely, Theorem 7.4 generalizes Katz' results [5] and some of their corollaries in [1] from homogeneous to quasi-homogeneous polynomials, on \mathbb{G}_m^n and on \mathbb{A}^n .

7.4. Theorem. *Let f be a quasi-homogeneous polynomial in n variables over $\mathcal{O}[1/N]$ for some ring of integers \mathcal{O} and integer $N > 0$. Suppose that 0 is a critical point of f . Let d be the dimension of the locus of $\text{grad } f = 0$, with the dimension of the empty scheme equal to -1 . Then there exists a $\delta > 0$ such that for each large enough prime p , each power q of p such that \mathbb{F}_q is an algebra over $\mathcal{O}[1/N]$, one has*

$$(7.4.1) \quad \left| \frac{1}{(q-1)^n} \sum_{x \in \mathbb{G}_m^n(\mathbb{F}_q)} \psi_q(f(x)) \right| < a q^{-\frac{n+d}{2}}$$

and

$$(7.4.2) \quad \left| \frac{1}{q^n} \sum_{x \in \mathbb{A}^n(\mathbb{F}_q)} \psi_q(f(x)) \right| < aq^{\frac{-n+d}{2}},$$

for all choices of ψ_q .

Proof. We first prove (7.4.1). By [1], Corollary 6.1 and section 9, (7.4.1) holds for nonlinear homogeneous f . Now let f be quasi-homogeneous and suppose that 0 is a critical point. After finitely many steps as from going from f to g in the proof of Corollary 7.2 and after renumbering the coordinates, one can go from a quasi-homogeneous polynomial to a nonlinear homogeneous polynomial h in possibly more variables. Going from f to g consists of two steps: first, one defines a polynomial $f_0(x, y)$ which equals $f(x)$; that is, one considers one more variable y . Then one performs the transformation on the torus (x, y) to $(x_1y, x_2, \dots, x_n, y)$ to obtain g . Under both these steps, the sum which is the left hand side of (7.4.1) remains unaltered, since one divides by $(q - 1)^n$ with n the number of variables in the left hand side of (7.4.1), and since one just performs a transformation on the torus. Corollary 7.2 now proves (7.4.1).

Now let us prove (7.4.2). Let $f_0(\hat{x})$ be the polynomial $f(0, \hat{x})$ in the $n - 1$ variables $\hat{x} = (x_2, \dots, x_n)$. Clearly f_0 is quasi-homogeneous in $n - 1$ variables. Write C_f for the locus of $\text{grad } f = 0$ in \mathbb{A}^n and C_{f_0} for the locus of $\text{grad } f_0 = 0$ in \mathbb{A}^{n-1} . By (7.4.1) it is enough to show that $n - 1 + \dim C_{f_0} \leq n + d$. Thus we only have to show that

$$(7.4.3) \quad \dim C_{f_0} \leq \dim C_f + 1,$$

where we recall that $\dim C_f = d$. This inequality (7.4.3) follows from writing

$$f(x) = x_1 \tilde{f}(x) + f_0(\hat{x})$$

with \tilde{f} a polynomial in x , and from Lemma 7.1, as follows. By Lemma 7.3, \tilde{f} is nonconstant. For $i = 1, \dots, n$, let $f_i(x)$ be the polynomial $\frac{\partial f}{\partial x_i}(x)$ and for $j = 2, \dots, n$, let $f_{0j}(\hat{x})$ be the polynomial $\frac{\partial f_0}{\partial x_j}(\hat{x})$. Let Y be the locus in \mathbb{A}^n of the polynomials f_{0j} and the equation $x_1 = 0$. Let Z be the intersection of C_f with $x_1 = 0$. Then clearly

$$(7.4.4) \quad \dim Y = \dim C_{f_0}.$$

Since $f_1 = \tilde{f} + x_1 \partial \tilde{f} / \partial x_1$, it follows that Z equals the intersection of Y with $\tilde{f}(x) = 0$. The polynomials f_i and \tilde{f} are quasi-homogeneous with the same weights. By Lemma 7.1 and since \tilde{f} is nonconstant, $\dim Y \leq \dim Z + 1$. Clearly also $\dim Z \leq d$. Now (7.4.3) follows by (7.4.4), and thus the proposition is proved. \square

8. PROOFS OF THE MAIN THEOREMS

Proof of Theorem 3.1. By (6.2.1) of Proposition 6.2, one finds immediately that there is a c such that for all K and q_K as in section 2.4 of large enough residue characteristic, and all faces τ of $\Delta_0(f)$, resp. all compact faces τ of $\Delta_0(g)$,

$$(8.0.5) \quad |E(K, y, f_\tau, \psi_K)| < cq_K^{-\sigma(f_\tau)},$$

resp.

$$(8.0.6) \quad |E(K, y, g_\tau, \psi_K)| < cq_K^{-\sigma(g_\tau)},$$

since automatically each f_τ , resp. each g_τ , is quasi-homogeneous for such τ . Moreover, c only depends on $\Delta_0(f)$, resp. on $\Delta_0(g)$, by Proposition 6.2. Now use Proposition 4.1, Proposition 5.2, and (8.0.5), (8.0.6). \square

The above proof of Theorem 3.1 is similar to the one of the main theorem of [1]. The key new ingredient in its proof is Proposition 6.2 for quasi-homogeneous polynomials instead of for homogeneous polynomials. Note that Proposition 6.2 was used both for (3.1.1) and for (3.1.2).

Proof of Theorem 3.2.1. By Lemma 7.3 we may suppose that 0 is a critical point of f . Otherwise, that is, when 0 is not a critical point, one has $S_{h,K,\psi_K}(y) = 0$ for the K and y under consideration. The theorem then follows from (7.4.2) of Theorem 7.4 and from the fact that the motivic oscillation index α_h of h satisfies

$$\frac{-n+d}{2} \leq \alpha_h,$$

with d the dimension of the locus of $\text{grad } h = 0$, by Theorem 5.1 of [2]. \square

ACKNOWLEDGMENT

The author would like to thank J. Denef, E. Hrushovski, F. Loeser, and especially J. Nicaise for inspiring discussions during the preparation of this paper.

REFERENCES

1. R. Cluckers, *Igusa and Denef-Sperber conjectures on nondegenerate p -adic exponential sums*, Duke Math. J. **141** (2008), no. 1, 205–216. MR2372152 (2009b:11138)
2. ———, *Igusa’s conjecture on exponential sums modulo p and p^2 and the motivic oscillation index*, Internat. Math. Res. Not. IMRN **2008** (2008), no. 4, article ID rnm118, 20 pages. MR2424173
3. J. Denef and S. Sperber, *Exponential sums mod p^n and Newton polyhedra*, Bull. Belg. Math. Soc. Simon Stevin **suppl.** (2001), 55–63. MR1900398 (2003b:11080)
4. J. Igusa, *Lectures on forms of higher degree (notes by S. Raghavan)*, Lectures on Mathematics and Physics, Tata Institute of Fundamental Research, vol. 59, Springer-Verlag, 1978. MR546292 (80m:10020)
5. N. Katz, *Estimates for “singular” exponential sums*, Internat. Math. Res. Not. IMRN (1999), no. 16, 875–899. MR1715519 (2001d:11084)
6. D. Segers, *Lower bound for the poles of Igusa’s p -adic zeta functions*, Math. Ann. **336** (2006), no. 3, 659–669. MR2249763 (2007g:11154)

DEPARTEMENT WISKUNDE, KATHOLIEKE UNIVERSITEIT LEUVEN, CELESTIJNENLAAN 200B, B-3001 LEUVEN, BELGIUM

Current address: Laboratoire Painlevé, Université Lille 1, Cité Scientifique, 59655 Villeneuve d’Ascq Cedex France

E-mail address: raf.cluckers@wis.kuleuven.be

URL: <http://www.wis.kuleuven.be/algebra/Raf/>