ACTA ARITHMETICA
XCII.3 (2000)

Estimates for complete multiple exponential sums

by

J. H. Loxton (Sydney)

1. Introduction. Let f be a polynomial in $\mathbb{Z}[\mathbf{x}]$ in the n variables $\mathbf{x} = (x_1, \ldots, x_n)$, with integer coefficients and of total degree d, say, greater than 1. For a positive integer q and such a polynomial f, we define the *complete multiple exponential sum*

$$S(f; q) = \sum_{\mathbf{x} \bmod q} e_q(f(\mathbf{x})),$$

where the sum is taken over a complete set of residues for $\mathbf{x}$ modulo q and $e_q(t) = e^{2\pi it/q}$.

The study of these sums is readily motivated by applications in analytic number theory and elsewhere. The first important estimates for sums in one variable appear in the work of Weyl (1916) on uniform distribution. This led to van der Corput's method with applications to the zeta function, the divisor problem and other problems in multiplicative number theory. Multiple exponential sums first appeared in work on the Epstein zeta function by Titchmarsh (1934). (Graham and Kolesnik (1991) discuss the history and recent results.) On the other hand, and of more immediate relevance to what follows, Hardy and Littlewood (1919) found a new method for tackling problems in additive number theory such as the problems of Waring and Goldbach. The treatment of the major arcs by this method involves complete exponential sums. (See, for example, Vaughan (1981).)

As a consequence of his proof of the Weil conjectures, Deligne (1974) showed that, for a prime p,

$$|S(f; p)| \le (d-1)^n p^{n/2},$$

provided that the homogeneous part of f of highest degree is non-singular modulo p. The applications to the Hardy–Littlewood method require non-trivial estimates of $S(f; q)$ for any q. As we will see, such estimates can be obtained from the case for a prime modulus by relatively elementary means.

2000 *Mathematics Subject Classification*: 11L07, 11G25.

If, in particular, $f(x)$ is a polynomial in one variable x, then very precise estimates for $S(f;q)$ are known in terms of invariants associated with the polynomial f. (See Loxton and Vaughan (1985) and Loh (1997).)

The results for polynomials in several variables are much less precise. Chubarikov (1976) proved the general estimate

$$|S(f;q)| \le e^{7d'n}3^{n\nu(q)}\tau(q)^{n-1}q^{n-1/d'},$$

provided that the content of f is prime to q, where d' is the maximum degree of f in any variable, $\nu(q)$ is the number of distinct prime divisors of q and $\tau(q)$ is the number of divisors of q. (See also Arkhipov, Karatsuba and Chubarikov (1987).) The example $f(x_1,\dots,x_n) = ax_1^{d'}\dots x_n^{d'}$ shows that the dependence on q is best possible. The experience with sums in one variable suggests that it is the high order singularity at the origin which leads to such an extremely large sum. At the opposite extreme, Loxton and Smith (1982) obtain a much smaller bound, namely $S(f;q) \ll q^{n/2}$, which applies when the projective variety defined by the equations $\operatorname{grad} f = 0$ is non-singular. The aim of this paper is to obtain bounds which are sensitive to the geometric properties of f and improve on the general bounds of Arkhipov, Karatsuba and Chubarikov (1987).

It is easy to see that $S(f;q)$ has a multiplicative property with respect to q. That is, if q_1 and q_2 are relatively prime integers and the integers m_1 and m_2 are such that $m_1q_1 + m_2q_2 = 1$, then

$$S(f;q_1q_2) = S(m_2f;q_1)S(m_1f;q_2).$$

Hence it suffices to examine the exponential sums $S(f;p^\alpha)$ with prime power modulus. In this paper, we essay an attack on this problem based on the use of the Newton polyhedron of a polynomial in several variables and illustrate the accuracy of the bound by an analysis of sums formed from polynomials of degree 2, where, of course, there are classical evaluations. A future paper will continue the analysis with explicit and precise estimates for polynomials of degree 3.

Our results imply estimates of the shape

$$S(f;p^\alpha) \ll p^{n\alpha(1-1/(2e))+\dim(\operatorname{grad} f)\alpha/(2e)},$$

where e is the maximum order of a singularity of the variety defined by the equation $\operatorname{grad} f = 0$, $\dim(\operatorname{grad} f)$ is the dimension of this variety, α is sufficiently large and the implied constant is independent of α. Extreme cases as discussed above can be expressed in terms of elementary quantities.

(a) In general,

$$S(f;p^\alpha) \ll p^{n\alpha(1-1/(2d))+\dim(\operatorname{grad} f)\alpha/(2d)},$$

where d is the total degree of f, $\dim(\operatorname{grad} f)$ is the dimension of the variety defined by the equations $\operatorname{grad} f = 0$, α is sufficiently large and the implied constant is independent of α. (Theorem 1, Corollary 2.) This estimate has the same quality as the general bounds of Arkhipov, Karatsuba and Chubarikov (1987), but may be better when nd' is relatively large compared to $2d$.

(b) In case the variety defined by the equations $\operatorname{grad} f = 0$ is nonsingular,

$$S(f;p^\alpha) \ll p^{n([(\alpha+1)/2]+\delta)},$$

where δ is the p-adic order of a certain discriminant, α is sufficiently large and the implied constant is independent of α and p. (Theorem 2, Corollary.) This is a more precise version of the type of estimate obtained in Loxton and Smith (1982).

We use standard p-adic notation. Thus, throughout, p denotes a rational prime and, for x in $\mathbb{Q}$, $\operatorname{ord}_p x$ denotes the highest power of p dividing x. (By convention, $\operatorname{ord}_p 0 = \infty$.) For a vector $\mathbf{x} = (x_1, \ldots, x_n)$, we write $\operatorname{ord}_p \mathbf{x} = \min_{1\le j\le n} \operatorname{ord}_p x_j$. We can embed the p-adic rationals $\mathbb{Q}_p$ in a complete algebraically closed field Ω_p and we continue to write ord_p for the extension of the valuation to Ω_p.

2. Simultaneous congruences. Let p be a prime and let $f(\mathbf{x})$ be a polynomial in $\mathbb{Z}[\mathbf{x}]$ in the n variables $\mathbf{x} = (x_1, \ldots, x_n)$. In this section, we give an upper bound for the sum

$$S(f;p^\alpha) = \sum_{\mathbf{x} \bmod p^\alpha} e_{p^\alpha}(f(\mathbf{x}))$$

in terms of the quantity

$$N(\operatorname{grad} f;p^\alpha) = |\{\mathbf{x} \bmod p^\alpha : \operatorname{grad} f(\mathbf{x}) \equiv \mathbf{0} \bmod p^\alpha\}|,$$

which counts the number of solutions of the simultaneous congruences $\partial f/\partial x_j \equiv 0 \bmod p^\alpha$. The results of this section are adapted from Loxton and Smith (1982).

PROPOSITION 1. *Suppose $\alpha > 1$ and set $\theta = [\alpha/2]$. Then*

$$|S(f;p^\alpha)| \le p^{n(\alpha-\theta)} N(\operatorname{grad} f;p^\theta).$$

Proof. Set $\gamma = \alpha - \theta$, so that $2\gamma \ge \alpha$ and $\gamma \ge \theta \ge 1$. We rewrite the sum $S(f;p^\alpha)$ by setting

$$\mathbf{x} = \mathbf{u} + p^\gamma \mathbf{v},$$

so that $\mathbf{x}$ runs through the residue classes modulo p^α as $\mathbf{u}$ and $\mathbf{v}$ respectively run through the residue classes modulo p^γ and p^θ. By a Taylor expansion

$$f(\mathbf{x}) = f(\mathbf{u}) + p^\gamma \operatorname{grad} f(\mathbf{u}) \cdot \mathbf{v} \bmod p^\alpha$$

and so

$$S(f;p^\alpha) = \sum_{\mathbf{u} \bmod p^\gamma} e_{p^\alpha}(f(\mathbf{u})) \sum_{\mathbf{v} \bmod p^\theta} e_{p^\alpha}(p^\gamma \operatorname{grad} f(\mathbf{u}) \cdot \mathbf{v}).$$

The inner sum vanishes unless all the components of grad $f(\mathbf{u})$ are congruent to 0 modulo p^θ. If this condition is satisfied, then the inner sum is equal to $p^{n\theta}$ because each term is equal to 1. Therefore,

$$S(f;p^\alpha) = p^{n\theta} \sum e_{p^\alpha}(f(\mathbf{u})),$$

where the sum is taken over all $\mathbf{u}$ modulo p^γ such that grad $f(\mathbf{u}) \equiv \mathbf{0} \bmod p^\theta$. Since there are $p^{n(\gamma-\theta)}$ points $\mathbf{u}$ modulo p^γ corresponding to each solution of the above congruences modulo p^θ, we have

$$|S(f;p^\alpha)| \le p^{n\theta+n(\gamma-\theta)} N(\operatorname{grad} f;p^\theta),$$

as required.

If α is odd, we can obtain a slightly sharper estimate than the one given by Proposition 1. To this end, let $H_f(\mathbf{u})$ denote the *Hessian matrix* $H_f(\mathbf{u}) = (\partial^2 f/\partial x_i \partial x_j(\mathbf{u}))$ and define

$$K_f(\mathbf{u}) = \{\mathbf{v} \bmod p : \mathbf{v}H_f(\mathbf{u}) \equiv \mathbf{0} \bmod p\}.$$

PROPOSITION 2. *Suppose $\alpha = 2\theta + 1$ with $\theta \ge 1$. Then*

$$|S(f;p^\alpha)| \le p^{n\alpha/2} \sum |K_f(\mathbf{u})|^{1/2},$$

where the sum is taken over all $\mathbf{u} \bmod p^\theta$ *such that* grad $f(\mathbf{u}) \equiv \mathbf{0} \bmod p^\theta$ *and in addition, in case p is odd,* grad $f(\mathbf{u}) \cdot \mathbf{v} \equiv 0 \bmod p^{\theta+1}$ *for all* $\mathbf{v}$ *in* $K_f(\mathbf{u})$.

Proof. From the proof of Proposition 1,

$$S(f;p^\alpha) = p^{n\theta} \sum e_{p^\alpha}(f(\mathbf{x})),$$

where the sum is taken over all $\mathbf{x}$ modulo p^γ such that grad $f(\mathbf{x}) \equiv \mathbf{0} \bmod p^\theta$ and $\gamma = \theta+1$. Here, we write $\mathbf{x} = \mathbf{u}+p^\theta\mathbf{v}$, so that $\mathbf{x}$ runs through the residue classes modulo p^γ as $\mathbf{u}$ and $\mathbf{v}$ respectively run through the residue classes modulo p^θ and p. By a Taylor expansion,

$$f(\mathbf{x}) = f(\mathbf{u}) + p^\theta \operatorname{grad} f(\mathbf{u}) \cdot \mathbf{v} + \tfrac{1}{2}p^{2\theta}\mathbf{v}H_f(\mathbf{u})\mathbf{v}^t \bmod p^\alpha.$$

Hence,

$$S(f;p^\alpha) = p^{n\theta} \sum e_{p^\alpha}(f(\mathbf{u}))G_f(\mathbf{u}),$$

where the sum is taken over all $\mathbf{u}$ modulo p^θ such that grad $f(\mathbf{u}) \equiv \mathbf{0} \bmod p^\theta$ and $G_f(\mathbf{u})$ denotes the *Gaussian sum*

$$G_f(\mathbf{u}) = \sum_{\mathbf{v} \bmod p} e_p\big(\tfrac{1}{2}\mathbf{v}H_f(\mathbf{u})\mathbf{v}^t + p^{-\theta} \operatorname{grad} f(\mathbf{u}) \cdot \mathbf{v}\big).$$

To estimate $G_f(\mathbf{u})$, consider

$$|G_f(\mathbf{u})|^2 = \sum_{\mathbf{v},\mathbf{w}} e_p\big(\tfrac{1}{2}\mathbf{v}H_f(\mathbf{u})\mathbf{v}^t - \tfrac{1}{2}\mathbf{w}H_f(\mathbf{u})\mathbf{w}^t + p^{-\theta}\operatorname{grad} f(\mathbf{u})\cdot(\mathbf{v}-\mathbf{w})\big).$$

Write $\mathbf{v} = \mathbf{w} + \mathbf{z}$ and carry out the summation over $\mathbf{w}$. This gives

$$|G_f(\mathbf{u})|^2 = p^n \sum_{\mathbf{z}H_f(\mathbf{u})\equiv 0 \bmod p} e_p\big(\tfrac{1}{2}\mathbf{z}H_f(\mathbf{u})\mathbf{z}^t + p^{-\theta}\operatorname{grad} f(\mathbf{u})\cdot\mathbf{z}\big).$$

We can replace $\mathbf{z}$ here by $\mathbf{z}+\mathbf{v}$ where $\mathbf{v}$ is any point in $K_f(\mathbf{u})$, so we have

$$|G_f(\mathbf{u})|^2 = e_p\big(\tfrac{1}{2}\mathbf{v}H_f(\mathbf{u})\mathbf{v}^t + p^{-\theta}\operatorname{grad} f(\mathbf{u})\cdot\mathbf{v}\big)|G_f(\mathbf{u})|^2.$$

Hence, $G_f(\mathbf{u})$ is 0 unless the argument of the exponential function is 0 mod p for all $\mathbf{v}$ in $K_f(\mathbf{u})$. If p is odd, this condition is equivalent to $p^{-\theta}\operatorname{grad} f(\mathbf{u})\cdot\mathbf{v} \equiv 0 \bmod p$ for all $\mathbf{v}$ in $K_f(\mathbf{u})$ and we have $|G_f(\mathbf{u})|^2 = p^n|K_f(\mathbf{u})|$ which gives the required estimate. If $p = 2$, the condition for $G_f(\mathbf{u})$ to be non-zero does not simplify, but we still have $|G_f(\mathbf{u})|^2 \le p^n|K_f(\mathbf{u})|$.

3. Basins of attraction of zeros. Let $\mathbf{f} = (f_1,\ldots,f_m)$ be an m-tuple of polynomials in $\mathbb{Z}_p[\mathbf{x}]$ and let $V(\mathbf{f})$ be the variety defined by the vector equation $\mathbf{f}(\mathbf{x}) = 0$ over Ω_p. As usual, the *degree* of the variety, $\deg\mathbf{f}$, is the product of the total degrees of the polynomials f_i and its *dimension*, $\dim\mathbf{f}$, is the maximum dimension of an irreducible component of $V(\mathbf{f})$. Note that the variety $V(\mathbf{f})$ contains a point $\xi = (\xi_1,\ldots,\xi_n)$ in $\mathbb{Z}_p^n$ if and only if the congruences $\mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \bmod p^\alpha$ are soluble for each α which, in the present context, is the case of most interest. We write

$$V_0(\mathbf{f}) = V(\mathbf{f}) \cap \mathbb{Z}_p^n.$$

We now turn to estimates for the quantity

$$N(\mathbf{f};p^\alpha) = |\{\mathbf{x} \bmod p^\alpha : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \bmod p^\alpha\}|,$$

which counts the number of solutions of the simultaneous congruences $f_i(\mathbf{x}) \equiv 0 \bmod p^\alpha$ for $1 \le i \le m$. A solution of these congruences is an approximate zero of each of the polynomials f_i and might be expected to fall near a point of $V_0(\mathbf{f})$. Consequently, for each point ξ in $V_0(\mathbf{f})$ and $\alpha \ge 0$, we define

$$\Gamma_\xi(\alpha) = \{\mathbf{x} \bmod p^\alpha : \operatorname{ord}_p\mathbf{f}(\mathbf{x}) \ge \alpha,\ \operatorname{ord}_p(\mathbf{x} \bmod p^\alpha - \xi) = \max_{\eta \text{ in } V_0(\mathbf{f})} \operatorname{ord}_p(\mathbf{x} \bmod p^\alpha - \eta)\}.$$

(Here, $\operatorname{ord}_p(\mathbf{x} \bmod p^\alpha - \xi)$ stands for the minimum of $\operatorname{ord}_p(\mathbf{y}-\xi)$ taken over all $\mathbf{y} \equiv \mathbf{x} \bmod p^\alpha$, so that $\Gamma_\xi(\alpha)$ only depends on $\xi \bmod p^\alpha$ and contains the complete residue class of $\mathbf{x}$ whenever $\mathbf{x}$ is "captured" by ξ.) We measure the size of $\Gamma_\xi(\alpha)$ by means of

$$\gamma_\xi(\alpha) = \min\{\operatorname{ord}_p(\mathbf{x} \bmod p^\alpha - \xi) : \mathbf{x} \text{ in } \Gamma_\xi(\alpha)\}.$$

If $\Gamma_\xi(\alpha)$ is non-empty, it follows that $0 \le \gamma_\xi(\alpha) \le \alpha$.

In order to get the most effective results from the geometry of $V(\mathbf{f})$, we need to lift the polynomials $\mathbf{f} \bmod p^\alpha$ in the most appropriate way. To this end, let d be the maximum of the total degrees of the polynomials f_i and let $\mathbf{g}$ run through all m-tuples of polynomials in $\mathbb{Z}_p[\mathbf{x}]$ with $\mathbf{g} \equiv \mathbf{f} \bmod p^\alpha$ and $\deg g_i \le d$. Define the *dimension of* $V(\mathbf{f})$ *at level* α by

$$\dim_\alpha \mathbf{f} = \max\{\dim \mathbf{g} : \mathbf{g} \equiv \mathbf{f} \bmod p^\alpha \text{ and } \deg g_i \le d\}$$

and call an m-tuple, $\mathbf{f}^{(\alpha)}$, say, at which the maximum is attained a *canonical representative for* $\mathbf{f}$ *at level* α. If α is sufficiently large, these adjustments are not needed and we have $\dim_\alpha \mathbf{f} = \dim \mathbf{f}$.

PROPOSITION 3. *Let* $\mathbf{f}$ *be an* m*-tuple of polynomials in* $\mathbb{Z}_p[\mathbf{x}]$ *and let* $\mathbf{f}^{(\alpha)}$ *be a canonical representative at level* α. *Suppose that* $V_0(\mathbf{f}^{(\alpha)})$ *is non-empty and let*

$$\varrho(\alpha) = \min\{\gamma_\xi(\alpha) : \xi \text{ in } V_0(\mathbf{f}^{(\alpha)})\}.$$

Then

$$N(\mathbf{f}; p^\alpha) \le (\deg \mathbf{f}^{(\alpha)}) p^{n(\alpha - \varrho(\alpha)) + (\dim_\alpha \mathbf{f})\varrho(\alpha)}.$$

Proof. To simplify the notation in the proof, we assume that $\mathbf{f}^{(\alpha)} = \mathbf{f}$. Let $V_0(\mathbf{f}; p^\alpha)$ denote the set of integral points $\mathbf{x} \bmod p^\alpha$ satisfying the simultaneous congruences $f_i(\mathbf{x}) \equiv 0 \bmod p^\alpha$. Note first that each point of $V_0(\mathbf{f}; p^\alpha)$ lies in some $\Gamma_\xi(\alpha)$ with ξ on $V_0(\mathbf{f})$ and

$$\Gamma_\xi(\alpha) \subseteq B_\xi(\alpha) = \{\mathbf{x} \bmod p^\alpha : \operatorname{ord}_p(\mathbf{x} - \xi) \ge \varrho(\alpha)\}$$

so that the number of distinct $\mathbf{x} \bmod p^\alpha$ in $\Gamma_\xi(\alpha)$ is at most $p^{n(\alpha - \varrho(\alpha))}$. If $\operatorname{ord}_p(\xi - \eta) \ge \varrho(\alpha)$, then both $\Gamma_\xi(\alpha)$ and $\Gamma_\eta(\alpha)$ lie inside $B_\xi(\alpha)$, so we need only count one such ξ in each residue class modulo $p^{\varrho(\alpha)}$. To finish the proof, we need an estimate for the number of residue classes modulo $p^{\varrho(\alpha)}$ represented by the points of $V_0(\mathbf{f})$.

We can make an integral unimodular linear change of coordinates so that no coordinate function x_j is constant on any component of $V(\mathbf{f})$ of positive dimension. Since this does not change the parameters in the statement of the proposition, nor the number of residue classes we seek to count, we can suppose that $V(\mathbf{f})$ has this property. Pick $\xi_1, \ldots, \xi_n$ in $\mathbb{Z}_p$ and consider the points of $V(\mathbf{f})$ with $x_j = \xi_j$ for $1 \le j \le \dim \mathbf{f}$. These form a variety of dimension 0 and so, by Bezout's theorem, they comprise a finite set of cardinality at most $\deg \mathbf{f}$. By allowing the ξ_j to run through the residue classes modulo $p^{\varrho(\alpha)}$, we pick up all the possible residue classes modulo $p^{\varrho(\alpha)}$ on $V(\mathbf{f})$, so their number is bounded by $(\deg \mathbf{f}) p^{(\dim \mathbf{f})\varrho(\alpha)}$.

The required estimate for $N(\mathbf{f}; p^\alpha)$ follows on combining the results of the two preceding paragraphs.

After the last proposition, we are interested in locating zeros of polynomial equations by successive approximation. The traditional tool for this purpose is Hensel's lemma, one version of which runs as follows.

PROPOSITION 4. *Let* $\mathbf{f} = (f_1, \ldots, f_n)$ *be an* n*-tuple of polynomials in* $\mathbb{Z}_p[\mathbf{x}]$ *and let* $J(\mathbf{x})$ *denote the* $n \times n$ *Jacobian matrix with entries* $\partial f_i/\partial x_j$. *Set* $\delta(\mathbf{x}) = \mathrm{ord}_p \det J(\mathbf{x})$. *Suppose* $\mathbf{x}_0 = (x_{01}, \ldots, x_{0n})$ *is a point of* Ω_p^n *with* $\mathrm{ord}_p x_{0j} \ge 0$ *for each* j *and* $\mathrm{ord}_p \mathbf{f}(\mathbf{x}_0) > 2\delta(\mathbf{x}_0)$. *Then there is a unique point* ξ *in* Ω_p^n *with* $\mathbf{f}(\xi) = 0$ *and* $\mathrm{ord}_p(\xi - \mathbf{x}_0) \ge \mathrm{ord}_p \mathbf{f}(\mathbf{x}_0) - \delta(\mathbf{x}_0)$.

Proof. Expanding the polynomials $\mathbf{f}$ about $\mathbf{x}_0$ gives

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + J(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \mathbf{h}(\mathbf{x} - \mathbf{x}_0),$$

where the entries of $\mathbf{h}$ are polynomials which have p-adic integer coefficients and vanish together with all their first order derivatives at the origin. By hypothesis, $J(\mathbf{x}_0)$ is non-singular, so we can find $\mathbf{y}_0$ satisfying $\mathbf{f}(\mathbf{x}_0) + J(\mathbf{x}_0)\mathbf{y}_0 = 0$. Set $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}_0$. Then

$$\mathrm{ord}_p \mathbf{f}(\mathbf{x}_1) \ge 2\mathrm{ord}_p \mathbf{y}_0 \ge 2(\mathrm{ord}_p \mathbf{f}(\mathbf{x}_0) - \delta(\mathbf{x}_0)) > \mathrm{ord}_p \mathbf{f}(\mathbf{x}_0)$$

and

$$\mathrm{ord}_p(J(\mathbf{x}_1) - J(\mathbf{x}_0)) \ge \mathrm{ord}_p \mathbf{y}_0 \ge \mathrm{ord}_p \mathbf{f}(\mathbf{x}_0) - \delta(\mathbf{x}_0) > \delta(\mathbf{x}_0).$$

Thus, $\mathbf{x}_1$ has the properties

$$\mathrm{ord}_p \mathbf{f}(\mathbf{x}_1) \ge 2(\mathrm{ord}_p f(\mathbf{x}_0) - \delta(\mathbf{x}_0)), \quad \delta(\mathbf{x}_1) = \delta(\mathbf{x}_0),$$
$$\mathrm{ord}_p(\mathbf{x}_1 - \mathbf{x}_0) \ge \mathrm{ord}_p \mathbf{f}(\mathbf{x}_0) - \delta(\mathbf{x}_0).$$

By repeating this construction with $\mathbf{x}_1$ in place of $\mathbf{x}_0$ and so on, we generate a Cauchy sequence $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ in Ω_p^n and $\xi = \lim \mathbf{x}_k$ is the desired solution.

Loxton and Vaughan demonstrated that the p-adic Newton polygon leads to sharper estimates for exponential sums in one variable than Hensel's lemma. Atan and Loxton (1986) explored extensions of the technique to multiple exponential sums. We sketch the idea in order to derive an alternative approximation theorem to replace Hensel's lemma in this case.

The *Newton polyhedron* of the polynomial $f(\mathbf{x}) = \sum a_{s_1 \ldots s_n} x_1^{s_1} \ldots x_n^{s_n}$ in $\mathbb{Z}_p[\mathbf{x}]$ is the lower convex hull of the points $(s_1, \ldots, s_n, \mathrm{ord}_p a_{s_1 \ldots s_n})$. The Newton polyhedron allows us to predict the size of the zeros of f in the following way.

PROPOSITION 5. *Let* f *be a polynomial in* $\mathbb{Z}_p[\mathbf{x}]$. *There is a zero* ξ *in* Ω_p *with* $f(\xi) = 0$ *and* $\mathrm{ord}_p \xi_j = \lambda_j$ *for each* j *if and only if the vector* $(\lambda_1, \ldots, \lambda_n, 1)$ *is normal to a face of the Newton polyhedron of* f, *or normal to an edge and between the normals to the adjacent faces.*

Proof. Suppose that $f(\xi) = 0$ and write $T_{\mathbf{s}} = a_{\mathbf{s}} \xi^{\mathbf{s}} = a_{s_1 \ldots s_n} \xi_1^{s_1} \ldots \xi_n^{s_n}$ and $P_{\mathbf{s}} = (s_1, \ldots, s_n, \mathrm{ord}_p a_{\mathbf{s}})$. Since $f(\xi) = 0$, the minimum, M, of the numbers $\mathrm{ord}_p T_{\mathbf{s}}$ is attained for at least 2 choices of $\mathbf{s}$, say $\mathbf{s} = \mathbf{u}$ and $\mathbf{s} = \mathbf{v}$. The corresponding points $P_{\mathbf{u}}$ and $P_{\mathbf{v}}$ lie on the hyperplane

$\Pi : x_1 \mathrm{ord}_p \xi_1 + \ldots + x_n \mathrm{ord}_p \xi_n + x_{n+1} = M$ which is a support hyperplane for the Newton polyhedron and its normal $(\xi_1, \ldots, \xi_n, 1)$ is normal to a face of the polyhedron, or normal to an edge and between the normals to the adjacent faces.

Conversely, suppose the rational vector $\nu = (\nu_1, \ldots, \nu_n, 1)$ is normal to the edge determined by the endpoints $P_{\mathbf{u}}$ and $P_{\mathbf{v}}$ on the Newton polyhedron and lies between the normals to the adjacent faces and suppose $\mathrm{ord}_p \xi_j = \nu_j$ for each j. By considering the projection on ν of the vector joining the point $P_{\mathbf{u}}$ to $P_{\mathbf{s}}$, we see that $\mathrm{ord}_p a_{\mathbf{u}} \xi^{\mathbf{u}} \le \mathrm{ord}_p a_{\mathbf{s}} \xi^{\mathbf{s}}$, with equality for $\mathbf{s} = \mathbf{v}$. Again, suppose $\xi_2, \ldots, \xi_n$ have $\mathrm{ord}_p \xi_j = \nu_j$ and write $g(x) = f(x, \xi_2, \ldots, \xi_n) = \sum c_i x^i$, with $c_i = \sum a_{i,s_2,\ldots,s_n} \xi_2^{s_2} \ldots \xi_n^{s_n}$. Since the terms $a_{\mathbf{u}} \xi_2^{u_2} \ldots \xi_n^{u_n}$ and $a_{\mathbf{v}} \xi_2^{v_2} \ldots \xi_n^{v_n}$ dominate c_{u_1} and c_{v_1} respectively, we can choose $\xi_2, \ldots, \xi_n$ so that $\mathrm{ord}_p c_{u_1} = \mathrm{ord}_p a_{\mathbf{u}} \xi_2^{u_2} \ldots \xi_n^{u_n}$ and $\mathrm{ord}_p c_{v_1} = \mathrm{ord}_p a_{\mathbf{v}} \xi_2^{v_2} \ldots \xi_n^{v_n}$. It may be necessary here to make an extension of the residue field in order to find such $\xi_2, \ldots, \xi_n$ and guarantee that the leading terms do not vanish. The Newton polygon of the polynomial $g(x)$ has a segment of slope $-\nu_1$ joining the points $(u_1, \mathrm{ord}_p c_{u_1})$ and $(v_1, \mathrm{ord}_p c_{v_1})$, so we can find ξ_1 with $g(\xi_1) = 0$ and $\mathrm{ord}_p \xi_1 = \nu_1$. This completes the construction of a p-adic zero ξ with $f(\xi) = 0$ and $\mathrm{ord}_p \xi_j = \nu_j$ for each j.

The next proposition uses the construction to estimate the distance from an approximate zero of a polynomial to its nearest zero. The statement uses the usual conventions for monomials $\mathbf{s} = (s_1, \ldots, s_n)$ with $|\mathbf{s}| = s_1 + \ldots + s_n$.

Proposition 6. *Let $f(\mathbf{x})$ be a polynomial in $\mathbb{Z}_p[\mathbf{x}]$ with degree at most d and let $\mathbf{x}_0$ be a point in $\mathbb{Z}_p^n$. Set*

$$\delta = \max_{\mathbf{s} \ne \mathbf{0}} \frac{1}{|\mathbf{s}|} \left(\mathrm{ord}_p f(\mathbf{x}_0) - \mathrm{ord}_p \frac{1}{\mathbf{s}!} \frac{\partial^{\mathbf{s}} f}{\partial \mathbf{x}^{\mathbf{s}}}(\mathbf{x}_0) \right).$$

Then

$$\max\{\mathrm{ord}_p(\xi - \mathbf{x}_0) : f(\xi) = 0 \text{ and } \xi \text{ is in } \Omega_p^n\} = \delta.$$

Proof. We can take $\mathbf{x}_0 = \mathbf{0}$. Write $f(\mathbf{x}) = \sum a_{\mathbf{s}} x^{\mathbf{s}}$ and suppose

$$\delta = \max_{\mathbf{s} \ne \mathbf{0}} \frac{1}{|\mathbf{s}|} \mathrm{ord}_p \frac{a_{\mathbf{0}}}{a_{\mathbf{s}}} = \frac{1}{|\mathbf{u}|} \mathrm{ord}_p \frac{a_{\mathbf{0}}}{a_{\mathbf{u}}}.$$

The choice of $\mathbf{u}$ means that the plane

$$\frac{\mathrm{ord}_p a_{\mathbf{0}} - \mathrm{ord}_p a_{\mathbf{u}}}{u_1 + \ldots + u_n}(x_1 + \ldots + x_n) + x_{n+1} = \mathrm{ord}_p a_{\mathbf{0}}$$

containing the points $(0, \ldots, 0, \mathrm{ord}_p a_{\mathbf{0}})$ and $(u_1, \ldots, u_n, \mathrm{ord}_p a_{\mathbf{u}})$ is a support hyperplane for the Newton polyhedron. Consequently, there is a p-adic zero ξ with $f(\xi) = 0$ and $\mathrm{ord}_p \xi = \delta$ and there are no zeros η with $\mathrm{ord}_p \eta > \delta$.

The last proposition improves Hensel's lemma in some respects. In contrast to Hensel's lemma, it yields useful information even when all the first order derivatives $\partial f/\partial x_j(\mathbf{x}_0)$ are zero. A more precise result for polynomials in 2 variables, together with a fuller proof, is given in Atan and Loxton (1986). Unfortunately, the Newton polyhedron technique does not lead directly to satisfactory estimates for the common zeros of several polynomials. The reasons for this will appear below.

4. A general estimate. Let $\mathbf{f} = (f_1, \ldots, f_m)$ be an m-tuple of polynomials in $\mathbb{Z}_p[\mathbf{x}]$. We define the p-adic *content*, $c(f_i)$, of the polynomial f_i to be the largest power of p dividing all the coefficients of f_i and we set $c(\mathbf{f}) = \min_{1\le i\le m} c(f_i)$.

We also define the *slope* $\delta_{f_i}(\mathbf{x}_0)$ of the Newton polyhedron of f_i at $\mathbf{x}_0$ by

$$\delta_{f_i}(\mathbf{x}_0) = \max_{\mathbf{s}\ne\mathbf{0}} \frac{1}{|\mathbf{s}|}\left(\mathrm{ord}_p f_i(\mathbf{x}_0) - \mathrm{ord}_p \frac{1}{\mathbf{s}!}\frac{\partial^{\mathbf{s}} f_i}{\partial \mathbf{x}^{\mathbf{s}}}(\mathbf{x}_0)\right)$$

and we call the vectors $\mathbf{s}$ at which the maximum is attained and $|\mathbf{s}|$ is minimal the *critical orders* of f_i at $\mathbf{x}_0$. Set

$$\delta_{\mathbf{f}}(\alpha) = \min\{\delta_{f_i}(\mathbf{x}_0) : f_i(\mathbf{x}_0) \equiv 0 \bmod p^\alpha,\ 1 \le i \le m\}.$$

This p-adic *slope* serves as a convenient substitute for the p-adic discriminant of $\mathbf{f}$ (see below).

Clearly,

$$\delta_{\mathbf{f}}(\alpha) \ge \frac{\alpha - c(\mathbf{f})}{d},$$

but local information about the singularities of $V(\mathbf{f})$ leads to stronger results. For example, if the highest order of a singular point on $V(\mathbf{f})$ is e, we have

$$\delta_{\mathbf{f}}(\alpha) \ge \frac{\alpha - \gamma(\mathbf{f})}{e},$$

where the constant $\gamma(\mathbf{f})$ can be obtained from the p-adic orders of the derivatives of the f_i of order up to e. The non-singular case, $e = 1$, is done explicitly in Section 5 below.

THEOREM 1. *Let $\mathbf{f} = (f_1, \ldots, f_m)$ be an m-tuple of polynomials in $\mathbb{Z}_p[\mathbf{x}]$ with degrees at most d and slope $\delta_{\mathbf{f}}(\alpha)$ and let $\mathbf{f}^{(\alpha)}$ be a canonical representative at level α. Let $\dim_\alpha \mathbf{f}$ denote the dimension of the variety defined by the equations $\mathbf{f}^{(\alpha)} = \mathbf{0}$ over Ω_p and suppose this variety contains points defined over $\mathbb{Z}_p$. Then*

$$N(\mathbf{f}; p^\alpha) \le \begin{cases} p^{n\alpha} & \text{if } \alpha \le \delta_{\mathbf{f}}(\alpha),\\ d^m p^{n(\alpha-\delta_{\mathbf{f}}(\alpha)) + \delta_{\mathbf{f}}(\alpha)\dim_\alpha \mathbf{f}} & \text{if } \alpha > \delta_{\mathbf{f}}(\alpha). \end{cases}$$

Proof. The estimate for $\alpha \le \delta_{\mathbf{f}}(\alpha)$ is trivial, so consider the second case.

Again, to simplify the notation in the proof, we suppose $\mathbf{f}^{(\alpha)} = \mathbf{f}$ and write $\delta_{\mathbf{f}}(\alpha) = \delta$. We require a parameter t which is a p-adic unit and algebraic of degree m over $\mathbb{Q}_p$ and which generates the full residue field extension, that is $[\mathbb{F}_p[t] : \mathbb{F}_p] = m$. Set

$$F(\mathbf{x}) = \mathrm{Norm}_{\mathbb{Q}_p[t]/\mathbb{Q}_p}(f_1(\mathbf{x}) + tf_2(\mathbf{x}) + \ldots + t^{m-1}f_m(\mathbf{x})).$$

Suppose $\mathbf{x}_0$ is in $\mathbb{Z}_p^n$. Since t has degree m over $\mathbb{F}_p$, $\mathrm{ord}_p\mathbf{f}(\mathbf{x}_0) \geq \alpha$ if and only if $\mathrm{ord}_p F(\mathbf{x}_0) \geq m\alpha$.

Write $f_i(\mathbf{x}) = p^{c_i}g_i(\mathbf{x})$, where g is a polynomial in $\mathbb{Z}_p[\mathbf{x}]$ whose coefficients are not all zero modulo p. Let $c = \min c_i$. Then $F(\mathbf{x}) = p^{mc}G(\mathbf{x})$, where $G(\mathbf{x})$ is in $\mathbb{Z}_p[\mathbf{x}]$ and

$$G(\mathbf{x}) \equiv \mathrm{Norm}_{\mathbb{Q}_p[t]/\mathbb{Q}_p} \sum\nolimits' g_i(\mathbf{x})t^{i-1} \bmod p,$$

where $\sum'$ denotes the sum taken over those indices i for which $c_i = c$. Again, because t has degree m over $\mathbb{F}_p$, the polynomial G cannot vanish identically modulo p. We can apply the same argument to the polynomials $f_i(\mathbf{x}_0 + \mathbf{x})$ for any $\mathbf{x}_0$ in $\mathbb{Z}_p^n$.

In the same way, we can estimate the slope δ_F of F at $\mathbf{x}_0$. Let $\mathbf{s}_i$ be a critical order of f_i for each i. Consider only those i for which the corresponding δ_{f_i} is equal to δ and, where there is a choice of critical $\mathbf{s}_i$, take the first such $\mathbf{s}_i$ in the lexicographical ordering. Suppose $\mathbf{s}_1$, say, is the first of these selected critical orders in the lexicographical ordering. Then the derivative of order $\mathbf{s} = m\mathbf{s}_1$ of F potentially gives rise to a critical order of F and the coefficient does not vanish because of the choice of t and the direction of differentiation. In fact, the leading term is

$$\mathrm{Norm}_{\mathbb{Q}_p[t]/\mathbb{Q}_p} \frac{\partial^{\mathbf{s}_1}}{\partial \mathbf{x}^{\mathbf{s}_1}}(f_1(\mathbf{x}) + tf_2(\mathbf{x}) + \ldots + t^{m-1}f_m(\mathbf{x}))$$

and so $\delta_F(\mathbf{x}_0) = m\delta$.

We can therefore apply Proposition 3 to the system of polynomials $\mathbf{f}$ and use Proposition 6 for the polynomial F to estimate $\varrho(\alpha) \geq \delta$ and so obtain the assertion of the theorem.

Corollary 1. *Let $\mathbf{f} = (f_1, \ldots, f_m)$ be an m-tuple of polynomials in $\mathbb{Z}_p[\mathbf{x}]$ with degrees at most d. Suppose the variety defined by the equations $\mathbf{f} = \mathbf{0}$ over Ω_p contains points defined over $\mathbb{Z}_p$ and the highest order of a singular point on $V(\mathbf{f})$ is e. Then, if α is sufficiently large*,

$$N(\mathbf{f}; p^{\alpha}) \ll p^{n\alpha(1-1/e)+(\dim \mathbf{f})\alpha/e},$$

where the implied constant is independent of α.

In fact, since $\delta_{\mathbf{f}}(\alpha) \geq (\alpha - \gamma(\mathbf{f}))/e$ with a constant $\gamma(\mathbf{f})$ independent of α, a more precise estimate valid for sufficiently large α is

$$N(\mathbf{f}; p^{\alpha}) \leq d^m p^{n\alpha(1-1/e)+n\gamma(\mathbf{f})/e+(\dim_\alpha \mathbf{f})(\alpha-\gamma(\mathbf{f}))/e}.$$

Corollary 2. *Suppose $\alpha > 1$ and set $\theta = [\alpha/2]$. Let f be a polynomial in $\mathbb{Z}[\mathbf{x}]$ with degree at most d and p-adic content $c(f)$. Let $\mathbf{g}^{(\theta)}$ be a canonical representation at level θ for the polynomials* grad f *and suppose the variety defined by the equations $\mathbf{g}^{(\theta)} = \mathbf{0}$ over Ω_p has points defined over $\mathbb{Z}_p$. Then*

$$|S(f;p^\alpha)| \le \begin{cases} p^{n\alpha} & \text{if } \theta \le \delta_{\operatorname{grad} f}(\theta),\\ d^n p^{n(\alpha-\delta_{\operatorname{grad} f}(\theta))+\delta_{\operatorname{grad} f}(\theta)\dim_\theta(\operatorname{grad} f)} & \text{if } \theta > \delta_{\operatorname{grad} f}(\theta)\end{cases}$$

and, if the highest order of a singular point on the variety $V(\operatorname{grad} f)$ is e, then

$$|S(f;p^\alpha)| \ll p^{n(\alpha-\theta/e)+\dim_\theta(\operatorname{grad} f)\theta/e}$$

when α is sufficiently large, with an implied constant independent of α.

The required inequalities follow from Proposition 1. If $\theta > \delta_{\operatorname{grad} f}(\theta)$, we get a completely specified inequality by using the estimates $\delta_{\operatorname{grad} f}(\theta) \ge (\theta - c(\operatorname{grad} f))/d$ and $c(\operatorname{grad} \mathbf{f}) \le \log \deg f/\log p$, namely

$$|S(f;p^\alpha)| \le d^{n+1} p^{n(\alpha-\theta/d)+nc(f)/d+\dim_\theta(\operatorname{grad} f)(\theta-c(f))/d}.$$

5. Estimation with linear forms. Let $\mathbf{f} = (f_1, \ldots, f_m)$ be an m-tuple of polynomials in $\mathbb{Z}_p[\mathbf{x}]$. We aim to construct solutions of the equations $\mathbf{f} = \mathbf{0}$ in $\mathbb{Z}_p$ by successively refining solutions of the congruences $\mathbf{f} \equiv \mathbf{0} \bmod p^\alpha$. In the process, we can make integral unimodular transformations of the variables and of the polynomials, without changing the solutions of the system. It will often be convenient to transform the system to an equivalent system with $\sum_i c(f_i)$ as large as possible and we refer to this as the *normal form* of the system.

If the polynomials $\mathbf{f}$ are linear, the bounds can be obtained directly. The result illustrates the source of the term involving $\dim \mathbf{f}$ in the estimate of Theorem 1 and of the discriminant which appears in the work of Loxton and Smith (1982) and in the estimates below.

Proposition 7. *Let $\mathbf{f} = (f_1, \ldots, f_m)$ be a vector of linear functions in $\mathbb{Z}[\mathbf{x}]$ represented by $\mathbf{f} = A\mathbf{x} + \mathbf{b}$, where A is an $m \times n$ matrix. Suppose A has rank r and let δ denote the minimum of the p-adic orders of the $r \times r$ non-singular submatrices of A. Then*

$$N(\mathbf{f};p^\alpha) \le p^{\min(n\alpha,(n-r)\alpha+\delta)}.$$

Proof. The first bound of $p^{n\alpha}$ is trivial and the second arises as follows. The matrix A is equivalent through integral unimodular transformations to a matrix A_S in the Smith normal form,

$$A_S = \begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix},$$

in which $A' = \operatorname{diag}(a_1, \ldots, a_r)$ is an $r \times r$ non-singular diagonal matrix. In the new coordinates, we write $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')^t$, where $\mathbf{x}'$ comprises the first r

components of $\mathbf{x}$ and $\mathbf{x}''$ the remainder. The congruences $A\mathbf{x} \equiv \mathbf{b} \bmod p^\alpha$ are equivalent to $A'\mathbf{x}' \equiv \mathbf{b}' \bmod p^\alpha$. Here, each congruence $a_j x'_j \equiv b'_j \bmod p^\alpha$ determines $x'_j \bmod p^{\alpha-\delta_j}$, where $\delta_j = \mathrm{ord}_p a_j$. Since $\delta = \mathrm{ord}_p \det A' = \delta_1 + \ldots + \delta_r$, the number of solutions for $\mathbf{x}'$ modulo p^α is at most p^δ. For each of these, there are $p^{(n-r)\alpha}$ choices for $\mathbf{x}''$ modulo p^α, so the total number of solutions for $\mathbf{x}$ modulo p^α is at most $p^{(n-r)\alpha+\delta}$.

COROLLARY. *Let p be an odd prime and $f(x,y)$ be a polynomial in $\mathbb{Z}[x,y]$ of degree* 2. *Let r be the rank of the matrix*

$$H_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$$

and suppose $r > 0$. Define the quantity δ by

$$\delta = \begin{cases} \mathrm{ord}_p(f_{xx}f_{yy} - f_{xy}^2) & \text{if } r = 2, \\ \mathrm{ord}_p(f_{xx}, f_{xy}, f_{yy}) & \text{if } r = 1. \end{cases}$$

Then

$$|S(f;p^\alpha)| \le p^{\min(2\alpha,(2-r/2)\alpha+\delta)}.$$

Proof. In case $\alpha = 1$, we can evaluate the Gauss sum $S(f;p)$ directly: $|S(f;p)| = p$ if $r = 2$ and $\delta = 0$, $|S(f;p)| = p^{3/2}$ if $r = 1$ and $\delta = 0$, and we have at worst the trivial bound $|S(f;p)| \le p^2$ otherwise.

If α is even, the result follows from Proposition 1. Suppose then $\alpha = 2\theta + 1 > 1$ is odd. If $\dim K_f(\mathbf{u}) = 0$ in Proposition 2, then

$$|S(f;p^\alpha)| \le p^\alpha N(f_x, f_y; p^\theta) \le p^{\alpha+\min(2\theta,(2-r)\theta+\delta)}$$

and this is less than the required bound. If $\dim K_f(\mathbf{u}) = 1$, then

$$|S(f;p^\alpha)| \le p^{\alpha+1/2} U,$$

say, where U counts the number of points $\mathbf{u}$ modulo p^θ which satisfy $\mathrm{grad}\, f(\mathbf{u}) \equiv tp^\theta \mathbf{z} \bmod p^{\theta+1}$, $\mathbf{z}$ is a fixed vector orthogonal to $K_f(\mathbf{u})$ and t is taken modulo p. If $\mathbf{u}$ is one solution of this congruence, then $\mathbf{u} + p^\theta \mathbf{v}$ is another whenever $\mathbf{v}H_f(\mathbf{u}) \equiv s\mathbf{z} \bmod p$ for some s taken modulo p. Consequently, $U = p^{-2}V$, where V is the number of points $\mathbf{u}$ taken modulo $p^{\theta+1}$ and satisfying the same congruence as before. We can estimate the number of these points by Proposition 7. Hence,

$$|S(f;p^\alpha)| \le p^{\alpha+1/2-2} V \le p^{\alpha-1/2+\min(2(\theta+1),(2-r)(\theta+1)+\delta)}$$

and this gives the result. Finally, if $\dim K_f(\mathbf{u}) = 2$, then

$$|S(f;p^\alpha)| \le p^{\alpha+1-2} N(f_x, f_y; p^{\theta+1}),$$

because each solution modulo p^θ of the congruences involved in computing $N(f_x, f_y; p^{\theta+1})$ corresponds to p^2 solutions taken modulo $p^{\theta+1}$. Again, the result follows.

In general, we define a p-adic *discriminant* as follows. For a polynomial $g(\mathbf{x})$, $\widetilde{g}$ denotes the associated homogeneous form in $x_0, x_1, \ldots, x_n$. We assume that the polynomials $\widetilde{f}_i$ and $\partial \widetilde{f}_i/\partial x_j$ for $1 \le i \le m$ and $1 \le j \le n$ have no common zeros. Let $\mathcal{I}_k$ denote the ideal generated by the polynomials f_i and the $k \times k$ subdeterminants of the Jacobian matrix $J(\mathbf{f}) = (\partial f_i/\partial x_j)$ and let $\widetilde{\mathcal{I}}_k$ be the associated homogeneous ideal. The ideals $\widetilde{\mathcal{I}}_j$ form a decreasing chain $\widetilde{\mathcal{I}}_1 \supset \widetilde{\mathcal{I}}_2 \supset \ldots$ By the Hilbert Nullstellensatz, $\sqrt{\widetilde{\mathcal{I}}_1} = (x_0, \ldots, x_n)$ and, in particular, $\mathcal{I}_1$ contains a non-zero element of $\mathbb{Z}_p$. Let δ_k denote the minimal p-adic order of the elements of $\mathcal{I}_k \cap \mathbb{Z}_p$, provided the intersection is non-empty, and set $\delta_k = 0$ otherwise. The *rank of the system*, denoted by $r(\mathbf{f})$, is the largest value of k such that the generators of $\widetilde{\mathcal{I}}_k$ have no common zeros. We call the corresponding value of δ_k the p-adic discriminant and denote it by $\delta(\mathbf{f})$. The variety defined by the equations $\mathbf{f} = \mathbf{0}$ is *non-singular* exactly when $r(\mathbf{f}) = n$. In that case, the discriminant $\delta(\mathbf{f})$ follows Krull's definition used in Loxton and Smith (1982).

THEOREM 2. *Let* $\mathbf{f} = (f_1, \ldots, f_n)$ *be an* n*-tuple of polynomials in* $\mathbb{Z}_p[\mathbf{x}]$ *with rank* $r(\mathbf{f}) = n$ *and* p*-adic discriminant* $\delta(\mathbf{f})$. *Then*

$$N(\mathbf{f}; p^\alpha) \le \begin{cases} p^{n\alpha} & \text{if } 0 < \alpha \le \delta(\mathbf{f}), \\ p^{n(\delta(\mathbf{f})+1)} & \text{if } \delta(\mathbf{f}) + 1 \le \alpha \le 2\delta(\mathbf{f}), \\ (\deg \mathbf{f}) p^{n\delta(\mathbf{f})} & \text{if } \alpha > 2\delta(\mathbf{f}). \end{cases}$$

P r o o f. For $\alpha \le \delta(\mathbf{f}) + 1$, trivially $N(\mathbf{f}; p^\alpha) \le p^{n\alpha}$.

Suppose $\delta(\mathbf{f})+1 < \alpha \le 2\delta(\mathbf{f})$ and abbreviate $\delta = \delta(\mathbf{f})$. Let $\mathbf{x}_0$ be any solution of the congruences $\mathbf{f}(\mathbf{x}_0) \equiv 0 \bmod p^{\delta+1}$ and write $\varepsilon = \operatorname{ord}_p \det J(\mathbf{f})(\mathbf{x}_0)$. We can find the required solutions modulo p^α by solving the congruence

$$\mathbf{f}(\mathbf{x}_0 + p^{\delta+1-\varepsilon}\mathbf{x}) \equiv \mathbf{f}(\mathbf{x}_0) + p^{\delta+1-\varepsilon} J(\mathbf{f})(\mathbf{x}_0)\mathbf{x} \equiv \mathbf{0} \bmod p^\alpha$$

for $\mathbf{x} \bmod p^{\alpha-\delta-1+\varepsilon}$. Here, $J(\mathbf{f})(\mathbf{x}_0)$ must have rank n and $\varepsilon \le \delta$, since otherwise the polynomials f_i and $\det J(\mathbf{f})$ evaluated at $\mathbf{x}_0$ all have p-adic order exceeding δ, contrary to the definition of the discriminant. By Proposition 7, the number of solutions for $\mathbf{x} \bmod p^{\alpha-\delta-1+\varepsilon}$ is at most p^ε and, summing over all possible choices for $\mathbf{x}_0 \bmod p^{\delta+1-\varepsilon}$ gives $N(\mathbf{f}; p^\alpha) \le p^{n(\delta+1-\varepsilon)+\varepsilon} \le p^{n(\delta+1)}$.

Finally, suppose $\alpha > 2\delta(\mathbf{f})$ and let $V(\mathbf{f})$ be the variety defined by the polynomials f_i. If $N(\mathbf{f}; p^\alpha)$ is non-zero, then Hensel's lemma (Proposition 4) yields a point of $V(\mathbf{f}) \cap \mathbb{Z}_p^n$ and, in the notation of Proposition 3, $\varrho(\alpha) = \alpha - \delta(\mathbf{f})$. The required result then follows from Proposition 3 because $\dim \mathbf{f} = 0$.

COROLLARY. *Let* f *be a polynomial in* $\mathbb{Z}[\mathbf{x}]$. *Suppose* $\operatorname{grad} f$ *has rank* n *and set* $\delta = \delta(\operatorname{grad} \mathbf{f})$ *and* $\theta = [\alpha/2]$. *Then*

$$|S(f;p^\alpha)| \le \begin{cases} p^{n\alpha} & \text{if } 1 < \alpha \le 2\delta + 1, \\ p^{n(\alpha-\theta+\delta+1)} & \text{if } 2\delta + 2 \le \alpha \le 4\delta + 1, \\ (\deg f - 1)^n p^{n(\alpha-\theta+\delta)} & \text{if } \alpha > 4\delta + 1. \end{cases}$$

Theorem 2 and its corollary improve the result of Loxton and Smith (1982). The parameters in that paper are essentially the same since the discriminant should properly have been defined as it is here.

References

G. I. Arkhipov, A. A. Karatsuba and V. N. Chubarikov, *Theory of Multiple Exponential Sums*, Moscow, 1987 (in Russian).

K. A. Atan and J. H. Loxton, *Newton polyhedra and solutions of congruences*, in: Diophantine Analysis, J. H. Loxton and A. J. van der Poorten (eds.), London Math. Soc. Lecture Note Ser. 109, Cambridge, 1986, 67–82.

V. N. Chubarikov, *Multiple rational trigonometric sums and multiple integrals*, Mat. Zametki 20 (1976), 61–68 (in Russian); English transl.: Math. Notes 20 (1976).

P. Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. 43 (1974), 273–307.

G. H. Hardy and J. E. Littlewood, *A new solution of Waring's problem*, Quart. J. Math. 48 (1919), 272–293.

S. W. Graham and G. Kolesnik, *Van der Corput's Method of Exponential Sums*, London Math. Soc. Lecture Note Ser. 126, Cambridge, 1991.

W. K. A. Loh, *Exponential sums on reduced residue systems*, Canad. Math. Bull. 41 (1997), 187–195.

J. H. Loxton and R. A. Smith, *Estimates for multiple exponential sums*, J. Austral. Math. Soc. 33 (1982), 125–134.

J. H. Loxton and R. C. Vaughan, *The estimation of complete exponential sums*, Canad. Math. Bull. 28 (1985), 440–454.

E. C. Titchmarsh, *On Epstein's zeta function*, Proc. London Math. Soc. (2) 36 (1934), 485–500.

R. C. Vaughan, *The Hardy–Littlewood Method*, Cambridge Tracts in Math. 80, Cambridge, 1981.

H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. 77 (1916), 313–352.

Macquarie University
Sydney, NSW 2109, Australia
E-mail: John.Loxton@mq.edu.au

Received on 17.10.1995
and in revised form on 25.5.1999 (3611)