NOTE ON SIMULTANEOUS CONGRUENCES

Trevor D. Wooley

Abstract. We prove, by an essentially elementary argument, that the number of non-singular solutions of a system of d simultaneous congruences, to a prime power modulus, in d variables is at most the product of the degrees of the polynomials defining the congruences.

1. Introduction

In this note we establish that the number of non-singular solutions of a system of d simultaneous congruences, to a prime power modulus, in d variables is at most the product of the degrees of the polynomials defining the congruences. While the treatment of the case $d = 1$ is entirely elementary, and occurs in any introduction to the theory of numbers, the natural generalisation to systems of congruences appears to be inaccessible in the literature. Very recently, the author received an enquiry from Professor K. Ford concerning bounds in this problem, which he required in work [1] on the asymptotic formula in Waring's problem. To be precise, let $\widetilde{G}(k)$ denote the least number t such that for all $s \geq t$, and all large natural numbers n, we have the expected asymptotic formula in Waring's problem, that is

$$\text{card}\left\{\mathbf{x} \in \mathbb{N}^s : n = x_1^k + \cdots + x_s^k\right\} = (\mathfrak{S}_{s,k}(n) + o(1))\,\frac{(\Gamma(1+1/k))^s}{\Gamma(s/k)}n^{s/k-1},$$

where $\mathfrak{S}_{s,k}(n)$ denotes the usual singular series in Waring's problem (see Vaughan [5, §2.6]). Then it transpires that the results of this note are required to obtain Ford's new bound $\widetilde{G}(k) \leq (1+o(1))k^2 \log k$, which improves the previous bound $\widetilde{G}(k) \leq (2 + o(1))k^2 \log k$ due to Wooley [6, Corollary 1.2]. Since the above result on congruences had also previously arisen in unpublished work [7] of the author on exponential sums over smooth numbers, we feel justified in publishing this note.

The precise form of our result on simultaneous congruences is contained in the following theorem, which we establish by an elementary argument in §2 below.

Theorem 1. *Let $f_1, \ldots, f_d$ be polynomials in $\mathbb{Z}[x_1, \ldots, x_d]$ with respective degrees $k_1, \ldots, k_d$, and write*

$$J(\mathbf{f}; \mathbf{x}) = \det\left(\frac{\partial f_j}{\partial x_i}(\mathbf{x})\right)_{1 \leq i,j \leq d}. \tag{1}$$

When p is a prime number, and s is a natural number, let $\mathcal{N}(\mathbf{f}; p^s)$ denote the number of solutions of the simultaneous congruences

$$f_j(x_1, \ldots, x_d) \equiv 0 \pmod{p^s} \quad (1 \leq j \leq d), \tag{2}$$

with $1 \leq x_i \leq p^s$ $(1 \leq i \leq d)$ and $(J(\mathbf{f}; \mathbf{x}), p) = 1$. Then $\mathcal{N}(\mathbf{f}; p^s) \leq k_1 \ldots k_d$.

The upper bound provided by Theorem 1 is plainly best possible, in view of the example $f_j(\mathbf{x}) = x_j^{k_j} - 1$ $(1 \leq j \leq d)$. For when p is a prime number with $p \equiv 1 \pmod{k_1 \ldots k_d}$, each congruence $y^{k_j} \equiv 1 \pmod{p}$ has precisely k_j solutions distinct modulo p, whence $\mathcal{N}(\mathbf{f}; p) = k_1 \ldots k_d$.

Since Hensel's Lemma permits one to lift solutions modulo p to solutions modulo p^s, it suffices to establish Theorem 1 when $s = 1$. While in most cases a classical version of Bézout's Theorem suffices for this purpose, the varieties corresponding to the polynomials f_i may not be "in general position",

Research supported in part by NSF grant DMS-9303505, an Alfred P. Sloan Research Fellowship and a Fellowship from the David and Lucile Packard Foundation.

and in such circumstances one must employ the formidable modern theory. For example, when a and b are fixed, such systems as

$$\sum_{i=1}^{2}(x_i^{k_j}-y_i^{k_j})\equiv a^{k_j}-b^{k_j} \pmod p \quad (1\le j\le 4)$$

are not in general position, since the solution set contains the singular line $(\mathbf{x},\mathbf{y})=(a,t,b,t)$. Versions of Bézout's Theorem adequate for such general situations involve concepts from commutative algebra unfamiliar to many number theorists (see [4, §I.7]). Thus, although we sketch in §3 how such a treatment may be used to prove Theorem 1, it seems desirable to devise a simple proof, more in keeping with the elementary statement of Theorem 1. Professor E. Bombieri has suggested that one might deform the underlying varieties, so that the non-singular solutions of the system of congruences correspond to solutions of an associated non-singular system. In general this deformation procedure is difficult to control, so in §2 we develop an elimination procedure which allows us to perform deformations restricted to one variable only. Such deformations are easily handled, and consequently the most difficult aspect of our whole argument is a standard version of Hensel's Lemma.

The author is grateful to Professors E. Bombieri, J. S. Milne and N. Dummigan for their valuable suggestions. He is also very grateful to Professor A. Granville for a productive discussion concerning effective elimination procedures. Finally, the author gladly acknowledges the perceptive comments of the referee that have improved our exposition.

We adopt the notation of using bold-face symbols for vectors, leaving the number of components implicit. Thus, for example, we may write $\mathbf{x}$ to denote $(x_1,\dots,x_d)$ and $\mathbf{y}$ to denote $(y_1,\dots,y_{d+1})$. We denote the p-adic numbers by $\mathbb{Q}_p$, with the usual p-adic valuation $|\cdot|_p$ normalised so that for $n\in\mathbb{Z}$ one has $|n|_p=p^{-r}$ when $p^r|n$ and $p^{r+1}\nmid n$. Finally, when $a,b\in\mathbb{Q}_p$, we write $a\equiv b \pmod{p^s}$ when $|a-b|_p\le p^{-s}$, and we write $p^t\|a$ when $|a|_p=p^{-t}$.

2. An elementary proof of Theorem 1

Our main tool in proving Theorem 1 is an effective elimination procedure provided by Lemma 2 below. In order to establish that lemma we require an intuitively obvious result concerning the non-vanishing of polynomials.

Lemma 1. *Suppose that $F(\mathbf{y})$ is a non-trivial polynomial in $\mathbb{Z}[y_1,\dots,y_t]$. Then for every non-zero integer n there exists $\mathbf{a}\in\mathbb{Z}^t$ such that $F(n\mathbf{a})$ is non-zero.*

Proof. The number of zeros of a non-trivial polynomial in a single variable is finite, and so the lemma is immediate when $t=1$. When $t>1$, write $F(\mathbf{y})$ as $\sum_i G_i(y_1,\dots,y_{t-1})y_t^i$, for suitable polynomials G_i. From the case $t=1$, it follows that $F(\mathbf{y})$ is non-zero for some $\mathbf{y}\in n\mathbb{Z}^t$ if there is an i for which there exists $\mathbf{z}\in n\mathbb{Z}^{t-1}$ with $G_i(\mathbf{z})\ne 0$. The lemma therefore follows by induction.

We next establish an effective elimination procedure which lies at the heart of our proof of Theorem 1.

Lemma 2. *Let $f_1,\dots,f_d$ be polynomials in $\mathbb{Z}[x_1,\dots,x_d]$ with respective degrees $k_1,\dots,k_d$. Let p be a prime number, let s be a positive integer, and suppose that there exists $\mathbf{c}\in\mathbb{Z}^d$ such that $p\nmid J(\mathbf{f};\mathbf{c})$. Then for each j with $1\le j\le d$, there is a non-trivial polynomial $\Psi_j\in\mathbb{Z}[y_1,\dots,y_{d+1}]$, and polynomials $g_1,\dots,g_d\in\mathbb{Z}[x_1,\dots,x_d]$, with the following properties:*

(i) $g_i\equiv f_i \pmod{p^s}$ for $1\le i\le d$;

(ii) $\Psi_j(g_1,\dots,g_d,x_j)$ is identically zero;

(iii) the degree of $\Psi_j(y_1,\dots,y_{d+1})$ with respect to y_{d+1} is at most $k_1\dots k_d$;

(iv) $\Psi_j\notin\mathbb{Z}[y_1,\dots,y_d]$.

Proof. We begin by establishing the existence of a generic eliminant polynomial Φ_j which avoids explicit reference to the coefficients of the f_j. When k is a non-negative integer, define the set $\mathcal{I}(k)$ by

$$\mathcal{I}(k)=\big\{(i_1,\dots,i_d)\in\mathbb{Z}^d : i_1\ge 0,\dots,i_d\ge 0, \text{ and } i_1+\cdots+i_d\le k\big\}. \tag{3}$$

On abbreviating the monomial $x_1^{i_1}x_2^{i_2}\dots x_d^{i_d}$ to x^σ, with $\sigma=(i_1,\dots,i_d)$, we may write $f_j(\mathbf{x})$ in the form $f_j(\mathbf{x})=\sum_{\sigma\in\mathcal{I}(k_j)}a_{j\sigma}x^\sigma$ for suitable integers $a_{j\sigma}$. Let $t_{j\sigma}$, for $\sigma\in\mathcal{I}(k_j)$ and $1\le j\le d$, be mutually

independent transcendental elements independent of $x_1, \dots, x_d$. We define the polynomial $F_j(\mathbf{x}, \mathbf{t}_j)$ associated to $f_j(\mathbf{x})$ by $F_j(\mathbf{x}, \mathbf{t}_j) = \sum_{\sigma \in \mathcal{I}(k_j)} t_{j\sigma} x^\sigma$. Write $\delta = \sum_{j=1}^d \operatorname{card}(\mathcal{I}(k_j))$, and relabel the $t_{j\sigma}$, in any convenient manner, simply as t_l $(1 \le l \le \delta)$. Let T denote the field extension $\mathbb{Q}(t_1, \dots, t_\delta)$, and consider the polynomials in $T[x_1, \dots, x_d]$ of degree at most D as a vector space, $\mathcal{V}_D$, over T. Since the monomials x^σ with $\sigma \in \mathcal{I}(D)$ form a basis for $\mathcal{V}_D$, one has $\dim \mathcal{V}_D = \operatorname{card}(\mathcal{I}(D))$. Then by considering the constant term in the Maclaurin expansion of

$$(z^{-D} + z^{1-D} + \dots)(1-z)^{-d} = z^{-D}(1-z)^{-d-1},$$

one deduces from (3) that

$$\dim \mathcal{V}_D = \binom{d+D}{d}. \tag{4}$$

Consider next the set $\mathcal{S}_B$ of polynomials $F_1^{a_1} \dots F_d^{a_d} x_j^b$ with $a_1 \ge 0, \dots, a_d \ge 0$, $k_1 a_1 + \dots + k_d a_d \le D - b$ and $0 \le b \le B$. Write $S(b; \mathbf{n})$ for the number of d-tuples $\mathbf{a}$ with

$$a_1 \ge 0, \dots, a_d \ge 0 \quad \text{and} \quad \sum_{i=1}^d (k_i a_i + n_i) \le D - b.$$

Then whenever $0 \le n_i \le k_i - 1$ $(1 \le i \le d)$, one has $S(b; \mathbf{n}) \le S(b; \mathbf{0})$, whence

$$k_1 \dots k_d S(b; \mathbf{0}) \ge \sum_{n_1=0}^{k_1-1} \cdots \sum_{n_d=0}^{k_d-1} S(b; \mathbf{n}) = \operatorname{card}(\mathcal{I}(D-b)) = \binom{d+D-b}{d}.$$

Consequently,

$$\begin{aligned} \operatorname{card}(\mathcal{S}_B) = \sum_{b=0}^B S(b; \mathbf{0}) &\ge (k_1 \dots k_d)^{-1} \sum_{b=0}^B \binom{d+D-b}{d} \\ &= \frac{B+1}{k_1 \dots k_d} \binom{d+D}{d} \left(1 + O(D^{-1})\right), \end{aligned} \tag{5}$$

where in the last expression, the implicit constant in the O-notation depends at most on B, $\mathbf{k}$ and d. On taking D sufficiently large in terms of $\mathbf{k}$ and d, it follows from (4) and (5) that whenever $B \ge k_1 \dots k_d$, one has $\operatorname{card}(\mathcal{S}_B) > \dim \mathcal{V}_D$, whence the polynomials in $\mathcal{S}_B$ are linearly dependent over T. Thus there exists a non-trivial polynomial Φ_j in $\mathbb{Z}[y_1, \dots, y_{d+1}, t_1, \dots, t_\delta]$, which by construction has degree at most $k_1 \dots k_d$ with respect to y_{d+1}, such that $\Phi_j(F_1, \dots, F_d, x_j, t_1, \dots, t_\delta)$ is identically zero.

If we now substitute the integer coefficients $a_{j\sigma}$ for their surrogate indeterminates $t_{j\sigma}$, in pursuit of our desired eliminant polynomial, then we leave open the possibility that $\Phi_j(\mathbf{F}, x_j, \mathbf{a})$ is trivial. We avoid such difficulties by making a deformation. Consider Φ_j as a polynomial in $y_1, \dots, y_{d+1}$ with coefficients which are polynomials in $t_1, \dots, t_\delta$, and let $h(\mathbf{t})$ be any non-trivial such coefficient. It follows from Lemma 1 that there exist integers $b_1, \dots, b_\delta$ with $h(\mathbf{a} + p^s \mathbf{b}) \ne 0$, whence $\Phi_j(\mathbf{y}, \mathbf{a} + p^s \mathbf{b})$ is non-trivial as a polynomial in $\mathbb{Z}[\mathbf{y}]$. We define $\Psi_j(\mathbf{y})$ to be $\Phi_j(\mathbf{y}, \mathbf{a} + p^s \mathbf{b})$, and write $g_j(\mathbf{x})$ for $F_j(\mathbf{x}, \mathbf{a} + p^s \mathbf{b})$ $(1 \le j \le d)$. Then since $g_j(\mathbf{x}) \equiv F_j(\mathbf{x}, \mathbf{a}) \pmod{p^s}$, we have $g_j(\mathbf{x}) \equiv f_j(\mathbf{x}) \pmod{p^s}$, which establishes the property (i). Moreover properties (ii) and (iii) follow from the construction of Ψ_j.

It remains only to prove assertion (iv). We consider any non-trivial polynomial Ψ_j of least degree for which there exist polynomials $g_1, \dots, g_d$ in $\mathbb{Z}[x_1, \dots, x_d]$ satisfying conditions (i), (ii) and (iii) of the statement of the lemma. If we suppose that $\Psi_j \in \mathbb{Z}[y_1, \dots, y_d]$, then $\partial \Psi_j / \partial y_{d+1} = 0$, and on differentiating the relation $\Psi_j(g_1, \dots, g_d, x_j) = 0$ with respect to x_k by using the chain rule, we obtain

$$\sum_{i=1}^d \frac{\partial \Psi_j}{\partial y_i}(g_1, \dots, g_d, x_j) \frac{\partial g_i}{\partial x_k}(\mathbf{x}) = 0 \quad (1 \le k \le d). \tag{6}$$

By property (i) we have $J(\mathbf{f}; \mathbf{x}) \equiv J(\mathbf{g}; \mathbf{x}) \pmod{p^s}$, so that by hypothesis there exists $\mathbf{c} \in \mathbb{Z}^d$ such that $p \nmid J(\mathbf{g}; \mathbf{c})$, whence the matrix with entries $\partial g_i / \partial x_k$ $(1 \le i, k \le d)$ has non-trivial determinant. It therefore follows from (6) that

$$\frac{\partial \Psi_j}{\partial y_i}(g_1, \dots, g_d, x_j) = 0 \quad (1 \le i \le d). \tag{7}$$

Now observe that for each i, the polynomial $\partial\Psi_j/\partial y_i$ has lower degree than Ψ_j. Then since we have supposed that Ψ_j has minimal degree, the equation (7) implies that the polynomials $\partial\Psi_j/\partial y_i$ $(1 \le i \le d)$ are identically zero. Thus Ψ_j is a non-trivial polynomial in $\mathbb{Z}[y_{d+1}]$, which contradicts our initial assumption that $\Psi_j \in \mathbb{Z}[y_1, \dots, y_d]$. This establishes the property (iv), and completes the proof of the lemma.

We require the following version of Hensel's lemma, which is only a slight refinement of Greenberg [2, Proposition (5.20)].

Lemma 3. *Let $f_1, \dots, f_d$ be polynomials in $\mathbb{Q}_p[x_1, \dots, x_d]$. Suppose that $\mathbf{a} \in \mathbb{Q}_p^d$ satisfies the system of congruences $f_j(\mathbf{a}) \equiv 0 \pmod{p^s}$ $(1 \le j \le d)$, and that $p^\delta \| J(\mathbf{f}; \mathbf{a})$ with $2\delta < s$. Then there exists a unique $\mathbf{b} \in \mathbb{Q}_p^d$ with $f_j(\mathbf{b}) = 0$ $(1 \le j \le d)$, and $\mathbf{b} \equiv \mathbf{a} \pmod{p^{s-\delta}}$.*

Proof. When $s > 2\delta$, and $\mathbf{a} \in \mathbb{Q}_p^d$ satisfies the hypotheses of the lemma, it follows from Greenberg [2, Proposition (5.20)] that there exists a unique $\mathbf{b} \in \mathbb{Q}_p^d$ with $f_j(\mathbf{b}) = 0$ $(1 \le j \le d)$, and $\mathbf{b} \equiv \mathbf{a} \pmod{p^{\delta+1}}$. Let t be the largest integer such that for each i with $1 \le i \le d$, one has $p^t | (b_i - a_i)$. Then $t \ge \delta + 1$. We suppose that $t < s - \delta$, and derive a contradiction. For each i write $h_i = (b_i - a_i)p^{-t}$, so that the h_i are p-adic integers, and for some i we have $p \nmid h_i$. On using the binomial theorem to provide a series expansion for the f_j, one deduces that

$$f_j(\mathbf{a} + \mathbf{h}p^t) \equiv f_j(\mathbf{a}) + p^t \sum_{i=1}^{d} h_i \frac{\partial f_j}{\partial x_i}(\mathbf{a}) \pmod{p^{2t}} \quad (1 \le i \le d).$$

Then since $f_j(\mathbf{b}) = 0$ and $f_j(\mathbf{a}) \equiv 0 \pmod{p^s}$ $(1 \le j \le d)$, we have

$$\sum_{i=1}^{d} h_i \frac{\partial f_j}{\partial x_i}(\mathbf{a}) \equiv 0 \pmod{p^{\min\{t, s-t\}}}.$$

But $\delta < \min\{t, s-t\}$ and $p^\delta \| J(\mathbf{f}; \mathbf{a})$, so that necessarily $\mathbf{h} \equiv \mathbf{0} \pmod{p}$. But $p \nmid h_i$ for some i, so we have derived a contradiction. Thus $t \ge s - \delta$, whence $\mathbf{b} \equiv \mathbf{a} \pmod{p^{s-\delta}}$, which completes the proof of the lemma.

The proof of Theorem 1. We adopt the notation of the statement of Theorem 1, and note that without loss of generality, we may suppose that $\mathcal{N}(\mathbf{f}; p^s)$ is non-zero, whence there exists a $\mathbf{c} \in \mathbb{Z}^d$ with $p \nmid J(\mathbf{f}; \mathbf{c})$. Our first objective is to make a change of variables which allows us to resolve the solutions counted by $\mathcal{N}(\mathbf{f}; p^s)$ by examining a single coordinate. We first note that a trivial estimate yields $\mathcal{N}(\mathbf{f}; p^s) \le p^{ds}$, so that the number, N, of solutions of the system (2) counted by $\mathcal{N}(\mathbf{f}; p^s)$ is certainly finite. Moreover by Lemma 3, these solutions lift uniquely to solutions $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(N)}$ in $\mathbb{Q}_p^d$ of the system $f_j(\mathbf{y}) = 0$ $(1 \le j \le d)$ subject to $p \nmid J(\mathbf{f}; \mathbf{y})$. When $y_i^{(l)} \ne y_i^{(m)}$, define u_{ilm} to be the largest integer such that $p^{u_{ilm}} | (y_i^{(l)} - y_i^{(m)})$, and otherwise define u_{ilm} to be zero. Put

$$u = \max_{1 \le i \le d} \max_{1 \le l < m \le N} u_{ilm},$$

and write $v_j = p^{2(j-1)u}$ $(1 \le j \le d)$. We make the substitution $y_1 = \sum_{i=1}^d v_i z_i$ and $y_j = z_j$ $(2 \le j \le d)$, and write $h_j(\mathbf{z}) = f_j(\mathbf{y})$. Notice that the solutions of the system $f_j(\mathbf{y}) = 0$ $(1 \le j \le d)$, subject to $p \nmid J(\mathbf{f}; \mathbf{y})$, are in one-to-one correspondence with those of the system $h_j(\mathbf{z}) = 0$ $(1 \le j \le d)$ subject to $p \nmid J(\mathbf{h}; \mathbf{z})$. Then by the uniqueness of the lifting action implied by Lemma 3, we deduce that $\mathcal{N}(\mathbf{f}; p^s) = \mathcal{N}(\mathbf{h}; p^{2du})$.

We assert that the N solutions $\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(N)}$ counted by $\mathcal{N}(\mathbf{h}; p^{2du})$ have distinct coordinates $z_1^{(j)}$ modulo p^{2du}. For if $z_1^{(i)} \equiv z_1^{(j)} \pmod{p^{2du}}$, then

$$y_1^{(i)} - \sum_{l=2}^{d} v_l y_l^{(i)} \equiv y_1^{(j)} - \sum_{l=2}^{d} v_l y_l^{(j)} \pmod{p^{2du}}, \tag{8}$$

whence $y_1^{(i)} \equiv y_1^{(j)} \pmod{p^{2u}}$. But $p^{u+1} \nmid (y_1^{(i)} - y_1^{(j)})$ when $y_1^{(i)} \ne y_1^{(j)}$, by the definition of u, and hence we obtain $y_1^{(i)} = y_1^{(j)}$. On substituting the latter relation into (8), we obtain $\sum_{l=2}^d v_l y_l^{(i)} \equiv \sum_{l=2}^d v_l y_l^{(j)}$

(mod p^{2du}), so that a similar argument yields $y_2^{(i)} = y_2^{(j)}$. By repeating this argument, we ultimately deduce that $\mathbf{y}^{(i)} = \mathbf{y}^{(j)}$. Thus, on writing $\mathcal{N}_1(\mathbf{h}; p^{2du})$ for the number of distinct z_1 for which $\mathbf{z}$ is a solution counted by $\mathcal{N}(\mathbf{h}; p^{2du})$, it follows from the conclusion of the previous paragraph that $\mathcal{N}(\mathbf{f}; p^s) = \mathcal{N}_1(\mathbf{h}; p^{2du})$.

We now apply Lemma 2 to deduce that there is a non-trivial polynomial $\Psi_1 \in \mathbb{Z}[y_1, \dots, y_{d+1}]$, and polynomials $g_1, \dots, g_d$ in $\mathbb{Z}[x_1, \dots, x_d]$, such that $g_i \equiv h_i \pmod{p^{2du}}$ $(1 \le i \le d)$, and properties (ii), (iii) and (iv) of the statement of Lemma 2 hold with $j = 1$. Write $\Psi_1(\mathbf{y})$ as a polynomial in y_{d+1} with coefficients in $\mathbb{Z}[y_1, \dots, y_d]$, and let $\psi_1(y_1, \dots, y_d)$ be the leading such coefficient (which is non-trivial, by property (iv) of Lemma 2). It follows from Lemma 1 that there exist integers $a_1, \dots, a_d$ such that $\psi_1(p^{2du}a_1, \dots, p^{2du}a_d)$ is non-zero, whence $\Psi_1(p^{2du}a_1, \dots, p^{2du}a_d, x_1)$ is a non-trivial polynomial in x_1.

But if z_1 is counted by $\mathcal{N}_1(\mathbf{h}; p^{2du})$, then it is also counted by $\mathcal{N}_1(\mathbf{g} - p^{2du}\mathbf{a}; p^{2du})$. Moreover each solution $\mathbf{z}$ counted by $\mathcal{N}(\mathbf{g} - p^{2du}\mathbf{a}; p^{2du})$ can, by Lemma 3, be lifted uniquely to a solution $\mathbf{w} \in \mathbb{Q}_p^d$ of the system $g_j(\mathbf{w}) = p^{2du}a_j$ $(1 \le j \le d)$ with $\mathbf{w} \equiv \mathbf{z} \pmod{p^{2du}}$. Thus the number of z_1 counted by $\mathcal{N}_1(\mathbf{g} - p^{2du}\mathbf{a}; p^{2du})$ is bounded above by the number of solutions $w_1 \in \mathbb{Q}_p$ of the equation $\Psi_1(p^{2du}a_1, \dots, p^{2du}a_d, w_1) = 0$, since by property (ii) of Lemma 2, $\Psi_1(g_1, \dots, g_d, w_1)$ is identically zero. Moreover, the polynomial $\Psi_1(p^{2du}\mathbf{a}, w_1)$ is non-trivial in w_1, by the conclusion of the previous paragraph. Thus the number of z_1 counted by $\mathcal{N}_1(\mathbf{g} - p^{2du}\mathbf{a}; p^{2du})$ is bounded above by the degree of $\Psi_1(\mathbf{y})$ with respect to y_{d+1}, which by property (iii) of Lemma 2 is at most $k_1 \dots k_d$. Theorem 1 follows immediately.

3. A sketch proof of Theorem 1 using Bézout's theorem

We now give a sketch of the proof of Theorem 1 using a suitable version of Bézout's Theorem. We note that several of the details require not insubstantial verification, and that the proof of the version of Bézout's Theorem which we require depends heavily on a knowledge of commutative algebra uncommon amongst workers in the field of application mentioned in the introduction. We shall make use of the notation of Hartshorne [4, Chapter I.7], so that, in particular, $i(Y, H; Z_j)$ denotes the intersection multiplicity of the varieties Y and H along Z_j, and $\deg V$ denotes the degree of a variety V. The following lemma embodies the version of Bézout's Theorem which we require.

Lemma 4. *Let Y be a variety in $\mathbb{P}^d$, and let H be a hypersurface not containing Y. Let $Z_1, \dots, Z_s$ be the irreducible components of $Y \cap H$. Then*

$$\sum_{j=1}^{s} i(Y, H; Z_j) \deg Z_j = (\deg Y)(\deg H).$$

Proof. This is Hartshorne [4, Theorem 7.7 of Chapter 1].

We start our proof of Theorem 1 by using Lemma 3 to lift solutions counted by $\mathcal{N}(\mathbf{f}; p^s)$ uniquely to solutions $\mathbf{y} \in \mathbb{Q}_p^d$ of the system $f_j(\mathbf{y}) = 0$ $(1 \le j \le d)$ subject to $p \nmid J(\mathbf{f}; \mathbf{y})$. We may then work in the completion, Ω_p, of the algebraic closure, $\overline{\mathbb{Q}}_p$, of $\mathbb{Q}_p$. We homogenise the polynomials $f_j(y_1, \dots, y_d)$ to obtain polynomials $\widetilde{f}_j(y_0, \dots, y_d)$, and view the solutions of the equations $\widetilde{f}_j(\mathbf{y}) = 0$ as defining hypersurfaces H_j in $(\Omega_p\mathbb{P})^d$ of degree k_j. Let $\mathcal{N}^*(\mathbf{f}; p)$ denote the number of points, $\mathbf{y}$, in $(\Omega_p\mathbb{P})^d$ lying on the intersection of the hypersurfaces H_j with $1 \le j \le d$, subject to the conditions $y_0 \ne 0$ and $J(\mathbf{f}; \mathbf{y}) \ne 0$. Notice that in our definition of $\mathcal{N}^*$, we do not count solutions at infinity, although we do permit singular solutions at infinity. Plainly, $\mathcal{N}(\mathbf{f}; p^s) \le \mathcal{N}^*(\mathbf{f}; p)$.

When $1 \le j \le d$, we define the set of varieties $\mathcal{W}_j$ as follows. We set $\mathcal{W}_1 = \{H_1\}$, and when $j \ge 2$ we define $\mathcal{W}_j$ to be the union of the dimension $d - j$ irreducible components of $Y \cap H_j$, for $Y \in \mathcal{W}_{j-1}$. We observe that the irreducible components of $Y \cap H_j$, for $Y \in \mathcal{W}_{j-1}$, of dimension exceeding $d - j$, do not contain points counted by $\mathcal{N}^*(\mathbf{f}; p)$. For if Z is any such component, then $Z \cap H_{j+1} \cap \dots \cap H_d$ is made up of components of dimension at least 1. Moreover any points $\mathbf{x} \in (\Omega_p\mathbb{P})^d$ which are not at infinity, and lie on such a component, must be singular with $J(\mathbf{f}; \mathbf{x}) = 0$, by using a suitable analogue of the Implicit Function Theorem. Thus such points are indeed not counted by $\mathcal{N}^*(\mathbf{f}; p)$.

We now provide an estimate for $\mathrm{card}(\mathcal{W}_d)$. When $1 \le j \le d$, let $s(j)$ denote the cardinality of $\mathcal{W}_j$, and write the elements of $\mathcal{W}_j$ as $Z_1^{(j)}, \dots, Z_{s(j)}^{(j)}$. We assert that for each j one has the inequality

$$\sum_{i=1}^{s(j)} \deg Z_i^{(j)} \le \prod_{i=1}^{j} k_i. \tag{9}$$

Our assertion certainly holds when $j = 1$. Suppose that it holds also for $j = J$. When $Z \in \mathcal{W}_J$, consider the intersection $Z \cap H_{J+1}$. If Z is contained in H_{J+1}, then $\dim(Z \cap H_{J+1}) = \dim Z > d - J - 1$, so that $Z \cap H_{J+1}$ is not contained in $\mathcal{W}_{J+1}$. Meanwhile, if Z is not contained in H_{J+1}, then the components of $Z \cap H_{J+1}$, which we write as $X_1, \dots, X_s$, must, by Lemma 4, satisfy the inequality

$$\sum_{i=1}^{s} \deg X_i \le \sum_{i=1}^{s} i(Z, H_{J+1}; X_i) \deg X_i = (\deg Z)(\deg H_{J+1}).$$

Consequently,

$$\sum_{i=1}^{s(J+1)} \deg Z_i^{(J+1)} \le \sum_{i=1}^{s(J)} (\deg Z_i^{(J)})(\deg H_{J+1}) = k_{J+1} \sum_{i=1}^{s(J)} \deg Z_i^{(J)},$$

and hence, by the inequality (9) with $j = J$,

$$\sum_{i=1}^{s(J+1)} \deg Z_i^{(J+1)} \le \prod_{i=1}^{J+1} k_i,$$

which establishes our assertion for $J + 1$. Thus our assertion follows for all j by induction.

From the case $j = d$ of (9), it follows that the number of components of $\cap_{i=1}^{d} H_i$ of dimension zero is at most $k_1 \dots k_d$. Consequently, $\mathcal{N}^*(\mathbf{f}; p) \le k_1 \dots k_d$, and Theorem 1 follows immediately.

References

1. K. B. Ford, *New estimates for mean values of Weyl sums* (preprint).
2. M. J. Greenberg, *Lectures on forms in many variables*, W. A. Bejamin, Inc., New York, Amsterdam, 1969.
3. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers, 5th ed.*, Oxford University Press, Oxford, 1979.
4. R. Hartshorne, *Algebraic geometry*, Springer-Verlag, Berlin, 1977.
5. R. C. Vaughan, *The Hardy-Littlewood Method*, Cambridge University Press, 1981.
6. T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.
7. T. D. Wooley, *On exponential sums over smooth numbers* (preprint).

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1003
E-mail address: wooley@math.lsa.umich.edu