On Systems of Diagonal Forms II

Michael P. Knapp

1. Introduction

In a recent paper [8], we considered the system $\mathbf{F}$ of homogeneous additive forms

$$\begin{aligned} F_1(\mathbf{x}) &= a_{11}x_1^{k_1} + \cdots + a_{1s}x_s^{k_1} \\ \vdots &\quad\quad \vdots \quad\quad\quad \vdots \\ F_R(\mathbf{x}) &= a_{R1}x_1^{k_R} + \cdots + a_{Rs}x_s^{k_R} \end{aligned}$$

with coefficients in a p-adic field $\mathbb{Q}_p$. If we had $R \geq 2$, $k_1 > k_2 > \cdots > k_R$ (in particular, the degrees of the forms were all different) and $p > k_1 - k_R + 1$, then we gave a bound on s, polynomial in terms of the degrees, which guarantees that the system $\mathbf{F} = \mathbf{0}$ has a non-trivial solution in p-adic integers. One could view this as a "hybrid" between the results of Ax & Kochen [2] and Lewis & Montgomery [9].

2000 *Mathematics Subject Classification.* 11D72 (primary), 11E76, 11E95 (secondary).

Work supported by NSF grant DMS-0344082.

The Ax-Kochen theorem states that the (smaller) bound

$$s > k_1^2 + \cdots + k_R^2$$

suffices provided that p is sufficiently large, but gives no bound on how large p needs to be. The Lewis-Montgomery bound, on the other hand, implies that any bound on s which works for all primes must exhibit exponential growth. Thus one may think of the results in [8] as bounding how small the prime p can be before exponential growth on s is required.

The purpose of the present paper is to eliminate the restriction in our previous result that the degrees of the forms be distinct. In this situation, we are now able to prove that a polynomial bound on s suffices for any system, provided that p is mildly large. In particular, we prove the following theorem.

Theorem 1. *Let p be a prime number. Consider the system $\mathbf{F} = \mathbf{0}$ of diagonal forms with coefficients in $\mathbb{Z}_p$ given by*

$$\begin{array}{rccccc} & F_1(\mathbf{x}) & = & a_{11}x_1^{k_1} + \cdots + a_{1s}x_s^{k_1} & = & 0 \\ (1) & \vdots & & \vdots & \vdots & \vdots \\ & F_R(\mathbf{x}) & = & a_{R1}x_1^{k_v} + \cdots + a_{Rs}x_s^{k_v} & = & 0. \end{array}$$

Suppose that $k_1 > k_2 > \cdots > k_v$ are the distinct degrees of the forms, and that for $1 \le i \le v$ there are R_i forms of degree k_i. Assume that

$p > \max\{k_1 - k_v + 1, 2\}$. *If we have*

$$s \geq \frac{3}{2}R\left(\sum_{i=1}^{v} R_i k_i\right)^2 + R\sum_{i=1}^{v} R_i k_i - \sum_{i=1}^{v} k_i + v$$

then this system has a solution with each variable in $\mathbb{Z}_p$ *and at least one variable not equal to 0.*

Note that in this theorem, $\sum_{i=1}^{v} k_i$ represents the sum of the distinct degrees in the system, and $\sum_{i=1}^{v} R_i k_i$ represents the sum of the degrees with multiplicities taken into account.

If all of the coefficients are in $\mathbb{Z}$, then Theorem 1 says that the system (1) has a nontrivial integral solution in every p-adic field $\mathbb{Q}_p$ for which p is at least the given bound. Note that if $v \geq 2$, then Theorem 1 implies that the bound

$$s \geq \frac{3}{2}R^3\left(k_1 + k_2 + \cdots + k_v\right)^2$$

suffices for these primes. One can compare this with the result given in [4] for systems of forms which all have the same degree (although this result is improved in [6]).

Theorem 1 is a trivial consequence of the following theorem, which is unfortunately somewhat more complicated to state.

Theorem 2. *Let* $\mathbf{F}(x_1, \ldots, x_s)$ *be a system as in Theorem 1, and assume again that* $p > \max\{k_1 - k_v + 1, 2\}$. *For* $1 \leq i \leq v$, *define the numbers* τ_i *and* $\tilde{k}_i$ *so that* $k_i = p^{\tau_i}\tilde{k}_i$ *for each* i, *with* $(p, \tilde{k}_i) = 1$. *Further, define*

$$t = 1 + \sum_{j=1}^{v} R_j \tilde{k}_j \frac{p^{\tau_j+1} - 1}{p - 1}.$$

If we have

$$s \geq t \sum_{i=1}^{v} (1 + R_1 + \cdots + R_{i-1}) R_i k_i - \sum_{i=1}^{v} k_i + v$$

then the system $\mathbf{F} = \mathbf{0}$ *has a solution with each variable in* $\mathbb{Z}_p$ *and at least one variable not equal to zero.*

We note that Theorem 1 follows from Theorem 2 since each of the R_i is at least 1 and $p/(p-1)$ is at most $3/2$. We also mention that we only require $k_1 > \cdots > k_v$ as a convenience in the proof of Lemma 8 below. If we replace the expression $k_1 - k_v$ in these theorems by

$$\text{(largest degree)} - \text{(smallest degree)},$$

then we may relabel the degrees to possibly obtain a smaller bound on s.

Our proof of Theorem 2 is essentially a three-step process. First, we apply a normalization procedure which allows us to assume that the

system **F** has certain desirable properties. Next we reduce the forms in **F** modulo various powers of p, obtaining a system of congruences, which we show must have a nonsingular solution. Finally, we are able to lift this nonsingular solution to a solution of (1) through an application of Hensel's lemma.

To this end, Section 2 of this paper gives the notation that we will use throughout the proof and some preliminary results. Section 3 describes our normalization procedure and proves that normalized systems have the properties we will need. In Section 4, we prove an auxiliary result stating that under certain conditions, a particular matrix can be guaranteed to be nonsingular modulo a prime p. This lemma will eventually guarantee that our solution of congruences is a nonsingular solution. Finally, in Section 5 we use the results of the preceding sections to complete the proof of Theorem 2.

2. Notation and Preliminaries

Although much of the notation to be used throughout this paper is given in the statements of Theorems 1 and 2, we repeat it here to have it all available in one location. Let p be a prime number. We will

consider systems of the shape

$$\begin{aligned} F_1(\mathbf{x}) &= a_{11}x_1^{k_1} + \cdots + a_{1s}x_s^{k_1} &= 0 \\ \vdots & \quad \vdots & \quad \vdots \quad \vdots \\ F_R(\mathbf{x}) &= a_{R1}x_1^{k_v} + \cdots + a_{Rs}x_s^{k_v} &= 0, \end{aligned}$$

where the first R_1 forms are of degree k_1, the next R_2 forms are of degree k_2, and so on. Note that we have $R_1 + R_2 + \cdots + R_v = R$. Here, $k_1 > k_2 > \cdots > k_v$ are the distinct degrees of the forms, and all of the coefficients are elements of $\mathbb{Z}_p$. Given the degrees and the prime p, we define numbers τ_i and $\tilde{k}_i$ so that $k_i = p^{\tau_i}\tilde{k}_i$ for each i and $(p, \tilde{k}_i) = 1$. Other notational conveniences that we will use require more context to define, and so we will defer those definitions until a more appropriate point.

Next, we state a few lemmata which will be needed as we complete the proof of Theorem 2. The first of these is a theorem of Alon [1, Theorem 1.2], which guarantees that a polynomial of a particular shape is not identically zero. This will be useful in Section 4 when we show that we can guarantee that certain matrices have a nonzero determinant modulo p.

Lemma 1. *Let $\mathbb{F}$ be an arbitrary field, and let $f = f(x_1, \ldots, x_s)$ be a polynomial in $\mathbb{F}[x_1, \ldots, x_s]$. Suppose the degree of f is $\sum_{i=1}^{s} t_i$, where*

each t_i is a nonnegative integer, and suppose the coefficient of $\prod_{i=1}^{s} x_i^{t_i}$ in f is nonzero in $\mathbb{F}$. Then, if $Y_1, \dots, Y_s$ are subsets of $\mathbb{F}$ with $|Y_i| > t_i$, then there exist $y_1 \in Y_1, y_2 \in Y_2, \dots, y_s \in Y_s$ so that $f(y_1, \dots, y_s) \neq 0$.

Our next lemma appears to have been independently discovered by Browkin [3] and Schanuel [11]. This is a version of Chevalley's Theorem for congruences modulo powers of a prime. That is, this lemma provides a number of variables sufficient to guarantee that a system of congruences modulo powers of a fixed prime has a nontrivial solution.

Lemma 2. *Suppose that $f_1, \dots, f_R$ are (not necessarily homogeneous) polynomials without constant terms, and for each i, let d_i be the (total) degree of f_i. Here we make no restrictions on the degrees, and in particular they need not be distinct. Consider the system of congruences*

$$f_i(x_1, \dots, x_s) \equiv 0 \pmod{p^{\tau_i}}, \qquad (1 \leq i \leq R).$$

If we have

$$s > \sum_{i=1}^{R} \frac{d_i(p^{\tau_i} - 1)}{p - 1},$$

then this system has a nontrivial solution in which each variable lies in the Teichmüller set

$$T_p = \{x \in \mathbb{Z}_p : x^p = x\}.$$

The final lemma in this section is a version of Hensel's Lemma. This is Lemma 4 of [7], which shows that a nonsingular solution of a system of homogeneous additive congruences can be lifted to a p-adic solution of the system.

Lemma 3. *Consider a system of additive equations*

$$
\begin{array}{rcccccc}
 & F_1(\mathbf{x}) & = & a_{11}x_1^{d_1} + \cdots + a_{1s}x_s^{d_1} & = & 0 \\
(2) & \vdots & & \vdots & & \vdots & \vdots \\
 & F_R(\mathbf{x}) & = & a_{R1}x_1^{d_R} + \cdots + a_{Rs}x_s^{d_R} & = & 0,
\end{array}
$$

where we again make no assumptions about the degrees of the forms. Let p be a prime number, and for $1 \le i \le R$ we define numbers τ_i and $\tilde{d}_i$ such that $d_i = p^{\tau_i}\tilde{d}_i$ with $(p, \tilde{d}_i) = 1$. Further, for $1 \le i \le R$, we define

$$
\gamma_i = \begin{cases} \tau_i & \text{if } p \text{ is odd} \\ \tau_i + 1 & \text{if } p = 2. \end{cases}
$$

Let h be a positive integer and suppose that $\mathbf{z}$ is a nontrivial solution of the system of congruences

$$
F_i(\mathbf{x}) \equiv 0 \pmod{p^{2h+\gamma_i-1}}, \qquad (1 \le i \le R)
$$

such that the matrix

$$\begin{bmatrix} a_{11}z_1^{d_1-1} & \dots & a_{1s}z_s^{d_1-1} \\ \vdots & & \vdots \\ a_{R1}z_1^{d_R-1} & \dots & a_{Rs}z_s^{d_R-1} \end{bmatrix}$$

has an $R \times R$ submatrix M such that

$$\det M \not\equiv 0 \pmod{p^h}.$$

Then the system (2) has a solution $\mathbf{y} \in \mathbb{Z}_p^s$ such that $\mathbf{y} \equiv \mathbf{z} \pmod{p^h}$.

If a solution of (2) modulo various powers of p satisfies the determinant condition above, we will say that this solution is nonsingular modulo p^h.

3. Normalization

In this section we define our normalization procedure and show that if the system (1) is normalized, then it is explicit in a large number of variables when considered modulo p. That is, when a normalized system is reduced modulo p, a large number of variables will appear in at least one form with a nonzero coefficient. Since our normalization originates with the one given by Davenport & Lewis in [5] for systems of additive forms of equal degrees, we begin by recalling the function

they used to normalize such systems. Suppose that we have a system $\mathbf{F}$ of R homogeneous additive forms

$$F_i(\mathbf{x}) = a_{i1}x_1^k + \cdots + a_{is}x_s^k, \qquad (1 \le i \le R)$$

all of degree k. Let $[\mathbf{a}_j]_{1\le j\le s}$ be the matrix of coefficients of the system, where $\mathbf{a}_j$ is the vector of coefficients of the variable x_j. Define the function

$$\Theta(\mathbf{F}) = \prod_{1\le j_1<j_2<\cdots<j_R\le s} \det([\mathbf{a}_{j_n}]_{1\le n\le R}).$$

In [5, Lemma 16], Davenport & Lewis showed that Θ satisfies the following properties.

Lemma 4. *1) If a system $\mathbf{F}'$ is defined by*

$$\mathbf{F}'(\mathbf{x}) = \mathbf{F}(p^{w_1}x_1, p^{w_2}x_2, \ldots, p^{w_s}x_s),$$

then we have

$$\Theta(\mathbf{F}') = p^{kRMw/s}\Theta(\mathbf{F}),$$

where $M = s(s-1)\cdots(s-R+1)$ and $w = w_1 + \cdots + w_s$.

2) If a system $\mathbf{F}''$ is defined by taking linear combinations of the forms in $\mathbf{F}$, so that

$$F_i'' = \sum_{j=1}^{R} d_{ij}F_i \qquad (1 \le i \le R),$$

then we have

$$\Theta(\mathbf{F}'') = D^M \Theta(\mathbf{F}),$$

where $D = \det([d_{ij}]_{1 \le i,j \le R})$.

We now turn to our situation, where the forms may have different degrees. Given a system $\mathbf{F}$, we first define two fundamental operations which may be performed on the forms, similar to the ones mentioned in Lemma 4. First, we may multiply each variable by a power of p, yielding a new system

$$\mathbf{F}'(\mathbf{x}) = \mathbf{F}(p^{w_1} x_1, \ldots, p^{w_s} x_s).$$

Second, we may take linear combinations of the forms, provided that for each linear combination, the forms involved all have the same degree and the total number of forms of each degree does not change. That is, we may form a new system $\mathbf{F}''$ defined by

$$F_i''(\mathbf{x}) = \sum_{\substack{n \\ \deg F_n = \deg F_i}} d_{in} F_n(\mathbf{x}).$$

Note that these operations commute. If a system $\mathbf{G}$ with coefficients in $\mathbb{Z}_p$ can be obtained from $\mathbf{F}$ through a combination of these operations, we will say that $\mathbf{G}$ is *equivalent* to $\mathbf{F}$. Note that when $\mathbf{F}$ and $\mathbf{G}$ are equivalent, $\mathbf{F}$ possesses a nontrivial zero if and only if $\mathbf{G}$ does.

Next, for positive integers j, a, b with $1 \leq j \leq v$ and $b-a+1 \geq R_j$, let $(\mathbf{F}_j, a, b)$ be the subsystem of $\mathbf{F}$ consisting only of the forms of degree k_j and only the variables x_i with $a \leq i \leq b$. For example, if $\mathbf{F}$ is the system of forms

$$x_1^{k_1} + 2x_2^{k_1} + 3x_3^{k_1}$$

$$4x_1^{k_2} + 5x_2^{k_2} + 6x_3^{k_2}$$

$$7x_1^{k_2} + 8x_2^{k_2} + 9x_3^{k_2},$$

then $(\mathbf{F}_2, 2, 3)$ is the system

$$5x_2^{k_2} + 6x_3^{k_2}$$

$$8x_2^{k_2} + 9x_3^{k_2}.$$

We define a function $\partial_j(\mathbf{F}, a, b)$ by

$$\partial_j(\mathbf{F}, a, b) = \Theta(\mathbf{F}_j, a, b).$$

That is, we define ∂_j by applying the Davenport-Lewis function to the subsystem $(\mathbf{F}_j, a, b)$ of $\mathbf{F}$. Note that performing a combination of the fundamental operations on $\mathbf{F}$ yields a combination of fundamental operations on $(\mathbf{F}_j, a, b)$. This is because if we form a linear combination of the forms of $\mathbf{F}$, we must use forms of the same degree, and because we do not change the number of forms of degree k_j. Hence Lemma 4 applies to ∂_j, and so we immediately get the following lemma.

Lemma 5. *1) If the system* $\mathbf{F}'$ *is defined by* $\mathbf{F}' = \mathbf{F}(p^{w_1}x_1, \ldots, p^{w_s}x_s)$, *then we have*

$$\partial_j(\mathbf{F}', a, b) = p^{k_j R_j M_j w/(b-a+1)}\partial_j(\mathbf{F}),$$

where $M_j = (b-a+1)(b-a)\cdots(b-a-R_j+2)$ *and* $w = w_a + \cdots + w_b$.

2) If the system $\mathbf{F}''$ *is defined by*

$$F_i'' = \sum_{\substack{n \\ \deg F_n = \deg F_i}} d_{in}F_n(\mathbf{x}),$$

then we have

$$\partial_j(\mathbf{F}'', a, b) = D_j^{M_j}\partial_j(\mathbf{F}, a, b),$$

where $D_j = \det([d_{in}]_{\deg F_n = \deg F_i = k_j})$.

Now we can begin to define the function we will use to normalize $\mathbf{F}$. Pick integers $1 < n_0 < n_1 < \cdots < n_v = s$ such that $n_0 \geq R$ and $n_i - n_{i-1} \geq R_i$ for each i. Define numbers M_i and L_i by

$$M_i = n_0(n_0 - 1)\cdots(n_0 - R_i + 1)$$

and

$$L_i = (n_i - n_{i-1})(n_i - n_{i-1} - 1)\cdots(n_i - n_{i-1} - R_i + 1).$$

Next, define the numbers K and M by

$$K = \prod_{i=1}^{v} k_i \qquad \text{and} \qquad M = \prod_{i=1}^{v} M_i,$$

and for $1 \leq i \leq v$, define $k_i' = K/k_i$, and $M_i' = M/M_i$. Finally, for $1 \leq i \leq v$, we define

$$N_i = \frac{(n_i - n_{i-1})k_i' RM}{n_0 R_i L_i}.$$

At last, for any system $\mathbf{F}$, we can define our normalization function $\partial(\mathbf{F})$ by

$$\partial(\mathbf{F}) = \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}, 1, n_0)^{k_i' M_i'}\right] \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}, n_{i-1}+1, n_i)^{N_i}\right].$$

We will say that the system $\mathbf{F}$ with coefficients in $\mathbb{Z}_p$ is *p-normalized* if $\partial(\mathbf{F}) \neq 0$ and the power of p dividing $\partial(\mathbf{F})$ is less than or equal to the power of p dividing $\partial(\mathbf{G})$ for any system $\mathbf{G}$ with coefficients in $\mathbb{Z}_p$ which is equivalent to $\mathbf{F}$.

We now prove a lemma showing that the function ∂ behaves nicely under the fundamental operations. This is the analogue of Lemma 4 for our normalization function.

Lemma 6. *Suppose that* $\mathbf{F}$ *is a system of additive forms.*

1) If the system $\mathbf{F}'$ *is defined by* $\mathbf{F}' = \mathbf{F}(p^{w_1}x_1, \ldots, p^{w_s}x_s)$*, then we have*

$$\partial(\mathbf{F}') = p^{KRMw/n_0}\partial(\mathbf{F}),$$

where $w = w_1 + \cdots + w_s$.

2) If the system $\mathbf{F}''$ *is defined by*

$$F_j'' = \sum_{\substack{n \\ \deg F_n = \deg F_j}} d_{jn} F_n,$$

then we have

$$\partial(\mathbf{F}'') = \left(\prod_{i=1}^{v} D_i^{A_i} \right) \partial(\mathbf{F}),$$

where

$$D_i = \det([d_{jn}]_{\deg F_n = \deg F_j = k_i})$$

and

$$A_i = M k_i' + \frac{(n_i - n_{i-1}) k_i' R M}{n_0 R_i}.$$

Proof. For part 1) of the lemma, we have

$$\begin{aligned} \partial(\mathbf{F}') &= \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}', 1, n_0)^{k_i' M_i'} \right] \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}', n_{i-1}+1, n_i)^{N_i} \right] \\ &= \left[\prod_{i=1}^{v} \left(p^{B_i} \partial_i(\mathbf{F}, 1, n_0) \right)^{k_i' M_i'} \right] \left[\prod_{i=1}^{v} \left(p^{C_i} \partial_i(\mathbf{F}, n_{i-1}+1, n_i) \right)^{N_i} \right] \end{aligned}$$

where

$$B_i = \frac{k_i R_i M_i (w_1 + \cdots + w_{n_0})}{n_0}$$

and

$$C_i = \frac{k_i R_i L_i (w_{n_{i-1}+1} + \cdots + w_{n_i})}{(n_i - n_{i-1})}.$$

Combining the powers of p and simplifying, we find that $\partial(\mathbf{F}') = p^A\partial(\mathbf{F})$, where

$$\begin{aligned} A &= \sum_{i=1}^{v} B_i k_i' M_i' + \sum_{i=1}^{v} C_i N_i \\ &= \sum_{i=1}^{v} \frac{R_i(w_1 + \cdots + w_{n_0})KM}{n_0} + \sum_{i=1}^{v} \frac{(w_{n_{i-1}+1} + \cdots + w_{n_i})KRM}{n_0} \\ &= \frac{(w_1 + \cdots + w_{n_0})KM}{n_0} \sum_{i=1}^{v} R_i + \frac{KRM}{n_0} \sum_{i=1}^{v} (w_{n_{i-1}+1} + \cdots + w_{n_i}) \\ &= \frac{KRM}{n_0} \left(w_1 + w_2 + \cdots + w_s\right). \end{aligned}$$

To prove part 2) of the lemma, note that we have

$$\begin{aligned} \partial(\mathbf{F}'') &= \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}'', 1, n_0)^{k_i' M_i'}\right] \left[\prod_{i=1}^{v} \partial_i(\mathbf{F}'', n_{i-1}+1, n_i)^{N_i}\right] \\ &= \left[\prod_{i=1}^{v} \left(D_i^{M_i} \partial_i(\mathbf{F}, 1, n_0)\right)^{k_i' M_i'}\right] \left[\prod_{i=1}^{v} \left(D_i^{L_i} \partial_i(\mathbf{F}, n_{i-1}+1, n_i)\right)^{N_i}\right] \\ &= \partial(\mathbf{F}) \cdot \prod_{i=1}^{v} D_i^{M_i k_i' M_i' + L_i N_i} \\ &= \partial(\mathbf{F}) \cdot \prod_{i=1}^{v} D_i^{M k_i' + (n_i - n_{i-1}) k_i' RM/n_0 R_i}, \end{aligned}$$

as desired.□

Now we can prove that p-normalized forms have certain desirable properties. These properties essentially state that if the forms in $\mathbf{F}$, or linear combinations of these forms, are reduced modulo p, then many variables will appear in the reduced forms with a nonzero coefficient.

While we develop here only a few results along these lines, we note that it is possible to go further and develop results analogous to Lemmata 2.2 and 2.3 of [12].

Lemma 7. *Let the system* $\mathbf{F}$ *of additive forms be p-normalized.*

1) Suppose that when this system is reduced modulo p, there are m variables which have a nonzero coefficient in at least one of the reduced forms. Then we have

$$m \geq \sum_{i=1}^{v} \frac{1}{k_i} \left(\frac{n_0 R_i}{R} + n_i - n_{i-1} \right).$$

2) Suppose that F is one of the forms in $\mathbf{F}$*, having degree* k_l*. If* q_l *denotes the number of variables which appear with a nonzero coefficient when F is reduced modulo p, then we have*

$$q_l \geq \frac{1}{k_l} \left(\frac{n_0}{R} + \frac{n_l - n_{l-1}}{R_l} \right).$$

3) Consider the forms in $\mathbf{F}$ *of degree* k_l*, and suppose that we make any* $H \leq R_l$ *linear combinations of these forms which are linearly independent over* $\mathbb{Z}/p\mathbb{Z}$*. Suppose that when these linear combinations are reduced modulo p, there are* $Q = Q(k_l, H)$ *variables which appear in at least one of these reduced forms with a nonzero coefficient. Then we*

have

$$Q \geq \frac{H}{k_l}\left(\frac{n_0}{R} + \frac{n_l - n_{l-1}}{R_l}\right).$$

Proof. Since part 2 is a special case of part 3 of this lemma, we will only give explicit proofs of part 1 and part 3. To prove part 1, suppose that the variables which are explicit modulo p are $x_i, i \in I$. Define a system $\mathbf{F}'$ by $\mathbf{F}' = p^{-1}\mathbf{F}(\mathbf{y})$, where

$$y_i = \begin{cases} px_i & \text{if } i \in I \\ x_i & \text{if } i \notin I, \end{cases}$$

and note that all the coefficients of this system are in $\mathbb{Z}_p$. By Lemma 6, we have $\partial(\mathbf{F}') = p^A\partial(\mathbf{F})$, where

$$A = \frac{KRMm}{n_0} - \sum_{i=1}^{v}\left(R_iMk_i' + \frac{(n_i - n_{i-1})k_i'RM}{n_0}\right).$$

Now, since the system $\mathbf{F}$ is p-normalized, we must have $A \geq 0$, whence we obtain

$$\begin{aligned} m &\geq \frac{n_0}{KRM}\sum_{i=1}^{v}\left(R_iMk_i' + \frac{(n_i - n_{i-1})k_i'RM}{n_0}\right) \\ &= \sum_{i=1}^{v}\frac{1}{k_i}\left(\frac{n_0R_i}{R} + n_i - n_{i-1}\right), \end{aligned}$$

as desired.

In order to prove part 3, we first complete the original H linear combinations to a set of R_l linear combinations of the forms of degree k_l which are linearly independent over $\mathbb{Z}/p\mathbb{Z}$. Let $\mathbf{F}'$ be the system consisting of these R_l forms and the $R - R_l$ forms in $\mathbf{F}$ of degrees other than k_l. Let $F'_j, j \in J$, be the forms in $\mathbf{F}'$ produced by our original H linear combinations, and suppose that $x_i, i \in I$ are the variables which have a nonzero coefficient when at least one of these H forms are reduced modulo p. Finally, consider the system $\mathbf{F}''$ defined by

$$F''_j = \begin{cases} p^{-1}F'_j(\mathbf{y}) & \text{if } j \in J \\ F'_j(\mathbf{y}) & \text{if } j \notin J, \end{cases}$$

where

$$y_i = \begin{cases} px_i & \text{if } i \in I \\ x_i & \text{if } i \notin I. \end{cases}$$

Note that all of the coefficients in both $\mathbf{F}'$ and $\mathbf{F}''$ are elements of $\mathbb{Z}_p$. By Lemma 6, we have

$$\partial(\mathbf{F}'') = p^{MRKQ/n_0} p^{-H(Mk'_l + (n_l - n_{l-1})k'_l RM/n_0 R_l)} \partial(\mathbf{F}')$$

and

$$\partial(\mathbf{F}') = D^{Mk'_l + (n_l - n_{l-1})k'_l RM/n_0 R_l} \partial(\mathbf{F}),$$

where D is the determinant of the matrix corresponding to the linear combinations taken to produce the R_l forms of degree k_l in $\mathbf{F}'$. Hence

we may write $\partial(\mathbf{F}'') = p^A D^B \partial(\mathbf{F})$, where

$$A = \frac{MRKQ}{n_0} - H\left(Mk_l' + \frac{(n_l - n_{l-1})k_l' RM}{n_0 R_l}\right)$$

and B is the exponent on D above. Since these combinations are linearly independent over $\mathbb{Z}/p\mathbb{Z}$, we have $p \not| D$. Now, since the system $\mathbf{F}$ is p-normalized, we must have $A \geq 0$, or in other words

$$\frac{MRKQ}{n_0} - H\left(Mk_l' + \frac{(n_l - n_{l-1})k_l' RM}{n_0 R_l}\right) \geq 0.$$

Solving this for Q yields

$$Q \geq \frac{H}{k_l}\left(\frac{n_0}{R} + \frac{n_l - n_{l-1}}{R_l}\right),$$

as desired. This completes the proof of the lemma. □

4. Nonsingular Matrices

This section is devoted to the proof of the following lemma, which shows that we can guarantee that certain matrices are nonsingular modulo a prime p.

Lemma 8. *Suppose that R is a positive integer, that $R_1, \ldots, R_v$ are positive integers such that $R_1 + \cdots + R_v = R$, that $k_1, \ldots, k_v$ are positive*

integers with $k_1 > k_2 > \cdots > k_v$, *and that* p *is a prime satisfying* $p > k_1 - k_v + 1$. *Suppose that* A *is the matrix*

$$A = \begin{bmatrix} a_{11}x_1^{k_1-1} & \cdots & a_{1R}x_R^{k_1-1} \\ \vdots & & \vdots \\ a_{R1}x_1^{k_v-1} & \cdots & a_{RR}x_R^{k_v-1} \end{bmatrix},$$

where the exponent on the x_i *is* $k_1 - 1$ *in the first* R_1 *rows, the exponent is* $k_2 - 1$ *in the next* R_2 *rows, and so on. Suppose further that for* $1 \le n \le v$, *the* $R_n \times R_n$ *submatrix*

$$A_n = [a_{ij}]_{R_1+\cdots+R_{n-1}+1 \le i,j \le R_1+\cdots+R_n}$$

is nonsingular modulo p *(i.e.* $\det A_n \not\equiv 0 \pmod{p}$*). Then there exist numbers* $b_2, \ldots, b_R$, *all nonzero modulo* p, *such that if we set*

$$x_2 = b_2x_1, \ldots, x_R = b_Rx_1$$

and let x_1 *be any number which is nonzero modulo* p, *then the matrix* A *will be nonsingular modulo* p.

To begin the proof, notice that if we make the substitution

$$x_j = b_jx_1, \qquad 2 \le j \le R,$$

and then pull as many factors of x_1 and b_j as possible out of the determinant, we obtain

$$\det A = x_1^{R_1k_1+\cdots+R_vk_v-R}(b_2\cdots b_R)^{k_v-1}\det B,$$

where B is the matrix

$$B = \begin{bmatrix} a_{11} & a_{12}b_2^{k_1-k_v} & \cdots & a_{1R}b_R^{k_1-k_v} \\ \vdots & \vdots & & \vdots \\ a_{R1} & a_{R2} & \cdots & a_{RR} \end{bmatrix}.$$

Therefore it suffices to show that we can choose $b_2,\ldots,b_R$ all nonzero modulo p such that B is nonsingular modulo p. We prove this by induction on v, the number of distinct degrees. If $v = 1$, then the matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1R} \\ \vdots & & \vdots \\ a_{R1} & \cdots & a_{RR} \end{bmatrix}$$

is nonsingular modulo p by hypothesis, and so it suffices to set $b_2 = \cdots = b_R = 1$.

Suppose now that the lemma is true when there are $v-1$ distinct degrees. Set $m = R_1 + \cdots + R_{v-1}$, and choose values for $b_2,\ldots,b_m$, all nonzero modulo p, which make the upper left-hand $m \times m$ submatrix C of B nonsingular modulo p. Now we just need to show that we can choose values for $b_{m+1},\ldots,b_R$ which work. Since the submatrix C of

B is nonsingular modulo p, we can perform elementary row operations on B to obtain a matrix

$$\left[\begin{array}{c|c} C & * \\ \hline 0_p & M \end{array}\right],$$

where all the entries of $\mathbf{0_p}$ are divisible by p, the asterisk could represent any matrix, and M has the form

$$M = [a_{ij} - p_{ij}(b_j)]_{m+1 \le i,j \le R},$$

where each $p_{ij}(b_j)$ is a polynomial of degree at most $k_1 - k_v$. Since C is nonsingular modulo p, it suffices to show that we can choose $b_{m+1}, \ldots, b_R$ to make M nonsingular modulo p. Note that if it happens that each $p_{ij}(b_j)$ is identically zero, then setting $b_{m+1} = \cdots = b_R = 1$ yields the desired conclusion.

Suppose then that at least one of the $p_{ij}(b_j)$ is not identically zero. Then expanding $\det M$ yields a polynomial $\det M = q(b_{m+1}, \ldots, b_R)$ in which $\deg_{b_j} q \le k_1 - k_v$ for each variable b_j. Select a term in this polynomial with highest total degree d which has a nonzero coefficient modulo p (if several such terms exist, any one of them can be selected). Without loss of generality, suppose that the variables involved in this term are $b_{m+1}, b_{m+2}, \ldots, b_r$. Set $b_{r+1} = \cdots = b_R = 1$. Then $\det M$

becomes a polynomial of total degree d in $b_{m+1}, \ldots, b_r$ and contains a term

$$b_{m+1}^{d_{m+1}} b_{m+2}^{d_{m+2}} \cdots b_r^{d_r}$$

whose coefficient is nonzero modulo p, and

$$d_{m+1} + \cdots + d_r = d.$$

Also, for each j, we have $d_j \leq k_1 - k_v$. Since $p > k_1 - k_v + 1$, we have $p - 1 > d_j$ for each j. Then by Lemma 1, there exist values of $b_{m+1}, \ldots, b_r \in \{1, 2, \ldots, p-1\}$ such that

$$q(b_{m+1}, \ldots, b_r, 1, \ldots, 1) \not\equiv 0 \pmod{p}.$$

The lemma follows upon noting that the value of each of $b_1, \ldots, b_R$ in our construction is nonzero modulo p.

5. Proof of Theorem 2

In this section, we will prove Theorem 2 using the tools developed in Sections 2, 3, and 4. Suppose that $\mathbf{F}$ is a system as in (1) defined over $\mathbb{Z}_p$ and that $p > \max\{k_1 - k_v + 1, 2\}$. Let s_0 be the bound on the number of variables given in Theorem 2. Assume briefly that we can prove that the theorem is true for systems containing exactly s_0 variables. If $\mathbf{F}$ contains more than s_0 variables, then setting $s - s_0$ of

them equal to zero yields a new system $\mathbf{F}'$ of diagonal forms in exactly s_0 variables. Since the theorem applies to $\mathbf{F}'$, we know that $\mathbf{F}'$ has a nontrivial zero, and hence so does $\mathbf{F}$. Thus, in order to prove Theorem 2, it suffices to prove it for systems containing exactly s_0 variables.

Suppose then that $\mathbf{F}$ has exactly s_0 variables. Note that since our normalization function ∂ is a polynomial in the coefficients of $\mathbf{F}$, a standard argument (see for example pages 572-573 of [5]) shows that if the theorem is true for all systems with $\partial(\mathbf{F}) \neq 0$, then it is true for all systems with $\partial(\mathbf{F}) = 0$ as well. Hence we may assume that $\partial(\mathbf{F}) \neq 0$. Additionally, since every system with $\partial(\mathbf{F}) \neq 0$ is equivalent to a p-normalized system, we may assume without loss of generality that $\mathbf{F}$ is p-normalized. When we normalize, we must specify the values of $n_0, n_1, \ldots, n_v$ used in the normalization procedure. To this end, we set $n_0 = R$ and for $1 \leq i \leq v$, we define n_i by the recursion

$$n_i = n_{i-1} + tR_ik_i - k_i - R_i + t(R_1 + \cdots + R_{i-1})R_ik_i + 1,$$

noting that we have $n_v = s_0$.

Let A be the matrix of coefficients of $\mathbf{F}$. We wish to find a collection of submatrices of A which satisfy the following properties.

(1) Let t be as defined in the theorem. Then for $1 \leq j \leq v$, there are t submatrices of size $R_j \times R_j$ in the collection.

(2) Each submatrix of size $R_j \times R_j$ contains only coefficients of the forms of degree k_j.

(3) Each submatrix in the collection is nonsingular modulo p (i.e. if B is one of the submatrices, then $\det B \not\equiv 0 \pmod{p}$).

(4) Each column of A contributes elements to at most one submatrix in the collection.

To accomplish this, we follow the method used by Brüdern & Godinho in [4], which adapts the work of Low, Pitman & Wolff in [10]. In the rest of this section, when we use the term "nonsingular," we will really mean "nonsingular modulo p."

First, we will find t disjoint $R_1 \times R_1$ nonsingular submatrices of A. Let A_1 be the submatrix of A containing the coefficients of the forms of degree k_1. Then A_1 is an $R_1 \times s_0$ matrix. Let $s_{1,m}$ be the minimal number of nonzero columns in any m linearly independent linear combinations of the rows of A_1. Then Brüdern & Godinho show in [4] that A_1 contains t disjoint nonsingular $R_1 \times R_1$ submatrices if and only if

$$s_{1,m} \geq tm \quad \text{for all} \quad 0 \leq m \leq R_1.$$

Clearly this condition holds for $m = 0$. Now, from Lemma 7, we have

$$\begin{aligned} s_{1,m} &= Q(k_1, m) \\ &\geq \frac{m}{k_1}\left(1 + \frac{tk_1R_1 - k_1 - R_1 + 1}{R_1}\right) \\ &= tm - \frac{m}{R_1}\left(1 - \frac{1}{k_1}\right). \end{aligned}$$

Since $\frac{m}{R_1} \leq 1$ and $1 - \frac{1}{k_1} < 1$, we see that $s_{1,m} > tm - 1$, and since $s_{1,m}$ is an integer we have $s_{1,m} \geq tm$ as desired. So we can form t disjoint nonsingular submatrices of A_1 (and hence also of A) of size $R_1 \times R_1$.

These submatrices involve a total of tR_1 columns of A. When we find our submatrices corresponding to the degree k_2 forms, we do not want these columns involved. Hence we let A_2 be the submatrix of coefficients of the degree k_2 forms, excluding the coefficients from any columns of A which were involved in our previous nonsingular submatrices. Then A_2 has R_2 rows and $s_0 - tR_1$ columns. Again, we find from [4] that if $s_{2,m}$ is the minimal number of nonzero columns in any m linearly independent linear combinations of the rows of A_2, then the submatrices we want exist provided that

$$s_{2,m} \geq tm \quad \text{for all} \quad 0 \leq m \leq R_2.$$

Again, this is clearly true if $m = 0$. From the normalization process, we know that any m linearly independent linear combinations of the forms of degree k_2 are explicit in at least

$$Q(k_2, m) \geq \frac{m}{k_2}\left(\frac{n_0}{R} + \frac{n_2 - n_1}{R_2}\right)$$

variables. However, up to tR_1 of these variables might not contribute to the columns of A_2. Hence, we know that the number of nonzero columns in any m linearly independent linear combinations of the rows of A_2 is at least

$$\begin{aligned} s_{2,m} &\geq \frac{m}{k_2}\left(1 + \frac{tR_2k_2 - k_2 - R_2 + tR_1R_2k_2 + 1}{R_2}\right) - tR_1 \\ &= tm + (m-1)tR_1 - \frac{m}{R_2}\left(1 - \frac{1}{k_2}\right) \end{aligned}$$

whenever $m \geq 1$. Again, we can see that this bound is strictly larger than $tm - 1$, and since $s_{2,m}$ is an integer, we obtain $s_{2,m} \geq tm$ when $m \geq 1$. Hence we have the bound that we need, and can get the required $R_2 \times R_2$ nonsingular submatrices of A.

We can now proceed inductively to show that we can find all of the necessary nonsingular submatrices. Assume that we can find the submatrices associated with the forms of degrees $k_1, k_2, \ldots, k_{l-1}$. These submatrices involve $t(R_1 + \cdots + R_{l-1})$ columns of A which must not

contribute to the submatrices for the forms of degree k_l. Let A_l be the matrix containing the coefficients of the forms of degree k_l, but excluding any coefficients coming from a column of A which was involved in any of the submatrices we have found previously. Let $s_{l,m}$ be the minimal number of nonzero columns in any m linearly independent linear combinations of the rows of A_l. As before, we can find t disjoint nonsingular $R_l \times R_l$ submatrices of A_l if and only if we have

$$s_{l,m} \geq tm \quad \text{for all} \quad 0 \leq m \leq R_l. \tag{3}$$

Once again, this is trivially true for $m = 0$. If $m \geq 1$, then we have

$$\begin{aligned} s_{l,m} &\geq Q(k_l, m) - t(R_1 + R_2 + \cdots + R_{l-1}) \\ &\geq \frac{m}{k_l}\left(1 + \frac{tR_lk_l - k_l - R_l + t(R_1 + \cdots + R_{l-1})R_lk_l + 1}{R_l}\right) \\ &\qquad - t(R_1 + \cdots + R_{l-1}) \\ &= mt + (m-1)t(R_1 + \cdots + R_{l-1}) - \frac{m}{R_l}\left(1 - \frac{1}{k_l}\right) \\ &> mt - 1. \end{aligned}$$

Again, since $s_{l,m}$ is an integer, this implies that the inequality in (3) holds for $1 \leq m \leq R_l$. Hence this inequality holds for all the necessary values of m, and so the matrices we are looking for do indeed exist.

Next, by relabeling the variables if necessary, we may assume that the $t(R_1 + \cdots + R_v) = tR$ columns of A involved in our nonsingular submatrices are the columns corresponding to the first tR variables. We may also assume that the coefficients of $x_1, \ldots, x_R$ correspond (in order) first to the columns of one of the $R_1 \times R_1$ nonsingular submatrices, then the columns of one of the nonsingular $R_2 \times R_2$ submatrices, and so on, ending with the columns of one of the $R_v \times R_v$ nonsingular submatrices. Further, we may assume that this is true for each set of R variables

$$x_{jR+1}, \ldots, x_{(j+1)R}, \qquad (0 \le j \le t-1).$$

Now, for each degree k_j, define the numbers τ_j and $\tilde{k}_j$ as in the statement of Theorem 2. Consider the system

$$\begin{array}{ccccc} a_{11}x_1^{k_1} & +\cdots+ & a_{1s}x_s^{k_1} & \equiv 0 & (\mathrm{mod}\ p^{\tau_1+1}) \\ \vdots & & \vdots & \vdots & \\ a_{R1}x_1^{k_v} & +\cdots+ & a_{Rs}x_s^{k_v} & \equiv 0 & (\mathrm{mod}\ p^{\tau_v+1}). \end{array} \tag{4}$$

Note that we have arranged the variables so that the coefficients of each set of R variables

$$x_{jR+1}, \ldots, x_{(j+1)R}, \qquad (0 \le j \le t-1)$$

satisfy the nonsingularity hypothesis of Lemma 8. Thus for each j we can find numbers

$$b_{jR+2}, \ldots, b_{(j+1)R},$$

all nonzero modulo p, such that if we set

$$x_{jR+2} = b_{jR+2}x_{jR+1}, \ \ldots, \ x_{(j+1)R} = b_{(j+1)R}x_{jR+1},$$

then the matrix

$$\begin{bmatrix} a_{1,jR+1}x_{jR+1}^{k_1-1} & \cdots & a_{1,(j+1)R}x_{(j+1)R}^{k_1-1} \\ \vdots & & \vdots \\ a_{R,jR+1}x_{jR+1}^{k_v-1} & \cdots & a_{R,(j+1)R}x_{(j+1)R}^{k_v-1} \end{bmatrix}$$

is nonsingular modulo p whenever $x_{jR+1} \not\equiv 0 \pmod{p}$. If we make these identifications and set $x_i = 0$ whenever $i > tR$, we obtain a new system of congruences

$$\begin{array}{ll} & c_{1,1}x_1^{k_1} + c_{1,2}x_{R+1}^{k_1} + \cdots + c_{1,t}x_{(t-1)R+1}^{k_1} \equiv 0 \pmod{p^{\tau_1+1}} \\ (5) \qquad & \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ & c_{R,1}x_1^{k_v} + c_{R,2}x_{R+1}^{k_v} + \cdots + c_{R,t}x_{(t-1)R+1}^{k_v} \equiv 0 \pmod{p^{\tau_v+1}}. \end{array}$$

Note that any nontrivial solution of (5) yields a nontrivial solution of (4). In fact, because of the way we chose the multipliers b_j, this solution of (4) will be nonsingular modulo p.

We wish to solve (5) with all the variables lying in the Teichmüller set

$$T_p = \{x \in \mathbb{Z}_p \ : \ x^p = x\}.$$

Note that if we write $k = p^\tau \tilde{k}$ with $(p, \tilde{k}) = 1$, then we have

$$x^k = (x^{p^\tau})^{\tilde{k}} = x^{\tilde{k}}$$

whenever $x \in T_p$. Hence any nontrivial solution of the system

$$\begin{array}{cccc} & c_{1,1}x_1^{\tilde{k}_1} + c_{1,2}x_{R+1}^{\tilde{k}_1} + \cdots + c_{1,t}x_{(t-1)R+1}^{\tilde{k}_1} \equiv 0 & \pmod{p^{\tau_1+1}} \\ (6) & \vdots & \vdots & \vdots \\ & c_{R,1}x_1^{\tilde{k}_v} + c_{R,2}x_{R+1}^{\tilde{k}_v} + \cdots + c_{R,t}x_{(t-1)R+1}^{\tilde{k}_v} \equiv 0 & \pmod{p^{\tau_v+1}} \end{array}$$

with all variables in T_p will also be a solution of (5).

Now, since we have

$$t > \sum_{j=1}^{v} R_j \tilde{k}_j \frac{p^{\tau_j+1} - 1}{p - 1},$$

Lemma 2 tells us that we can solve the system (6) nontrivially with each variable in T_p. As noted above, this leads to a solution of (4) which is nonsingular modulo p. Then Lemma 3 shows that we can lift this solution of (4) to a solution of (1). Finally, since the solution of (4) contains at least one variable which is not divisible by p and our solutions of (1) and (4) are congruent modulo p, the solution of (1) is a nontrivial solution. This completes the proof of the theorem.

References

1. N. Alon, *"Combinatorial Nullstellensatz"*, Combin. Probab. Comput. **8** (1999), 7–29.
2. J. Ax and S. Kochen, *"Diophantine problems over local fields I"*, Amer. J. Math. **87** (1965), 605–630.
3. J. Browkin, *"On zeros of forms"*, Bull. Acad. Polon. Sci. S'er. Sci. Math. Astronom. Phys. **17** (1969), 611–616.
4. J. Brüdern and H. Godinho, *"On Artin's conjecture, I: systems of diagonal forms"*, Bull. London Math. Soc. **31** (1999), 305–313.
5. H. Davenport and D. J. Lewis, *"Simultaneous equations of additive type"*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **264** (1969), 557–595.
6. M. Knapp, *"Systems of diagonal equations over p-adic fields"*, J. London Math. Soc. (2) **63** (2001), 257–267.
7. ———, *"Diagonal equations of different degrees over p-adic fields"*, Acta Arith. **126.2** (2007), 139–154.
8. ———, *"On systems of diagonal forms"*, J. Aust. Math. Soc. **82** (2007), 221–236.
9. D. J. Lewis and H. L. Montgomery, *"On zeros of p-adic forms"*, Michigan Math J. **30** (1983), 83–87.
10. L. Low, J. Pitman, and A. Wolff, *"Simultaneous diagonal congruences"*, J. Number Theory **29** (1988), 31–59.
11. S. H. Schanuel, *"An extension of Chevalley's theorem to congruences modulo prime powers"*, J. Number Theory **6** (1974), 284–290.

12. T. D. Wooley, *"On simultaneous additive equations I"*, Proc. London Math. Soc. (3) **63** (1991), 1–34.

Michael P. Knapp

Mathematical Sciences Department

Loyola College

4501 North Charles Street

Baltimore, MD 21210-2699

Email: mpknapp@loyola.edu