SYSTEMS OF DIAGONAL EQUATIONS OVER $\mathfrak{p}$-ADIC FIELDS

MICHAEL P. KNAPP

1. *Introduction*

Let $\mathbb{K}$ be a $\mathfrak{p}$-adic field, and consider the system $\mathbf{F} = (F_1, \ldots, F_R)$ of diagonal equations

$$\begin{array}{ccc} a_{11}x_1^k + \ldots + a_{1N}x_N^k = 0 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ a_{R1}x_1^k + \ldots + a_{RN}x_N^k = 0 \end{array} \tag{1}$$

with coefficients in $\mathbb{K}$. It is an interesting problem in number theory to determine when such a system possesses a nontrivial $\mathbb{K}$-rational solution. In particular, we define $\Gamma^*(k, R, \mathbb{K})$ to be the smallest natural number such that any system of R equations of degree k in N variables with coefficients in $\mathbb{K}$ has a nontrivial $\mathbb{K}$-rational solution provided only that $N \geqslant \Gamma^*(k, R, \mathbb{K})$. For example, when $k = 1$, ordinary linear algebra tells us that $\Gamma^*(1, R, \mathbb{K}) = R+1$ for any field $\mathbb{K}$. We also define $\Gamma^*(k, R)$ to be the smallest integer N such that $\Gamma^*(k, R, \mathbb{Q}_p) \leqslant N$ for all primes p.

When $\mathbb{K} = \mathbb{Q}_p$, much is known about this problem. In the case where $R = 1$, Davenport and Lewis [**5**] showed that $\Gamma^*(k, 1) \leqslant k^2+1$ for each k, with equality holding whenever $k = p-1$ for some prime p. When $R = 2$ and k is odd, Davenport and Lewis [**6**] showed that $\Gamma^*(k, 2) \leqslant 2k^2+1$. For general R, a conjecture of Artin's suggests that one should have $\Gamma^*(k, R) \leqslant Rk^2+1$, but this is not known in any case other than the three above. Despite the inability to obtain the conjectured bound, several authors have found upper bounds for $\Gamma^*(k, R)$. Davenport and Lewis [**7**] obtained the bound $\Gamma^*(k, R) \leqslant [9R^2k\log(3Rk)]$ for all odd k, and the bound $\Gamma^*(k, R) \leqslant [48R^2k^3\log(3Rk^2)]$ for all even k larger than 2. This was improved in most cases by Low, Pitman and Wolff [**11**], who showed that the bound $\Gamma^*(k, R) \leqslant [48Rk^3\log(3Rk^2)]$ is sufficient for all k larger than 2, and that the bound $\Gamma^*(k, R) \leqslant 2R^2k\log k$ holds whenever k is odd and sufficiently large.

Recently, Brüdern and Godinho [**3**] obtained the bound $\Gamma^*(k, R) \leqslant R^3k^2$ whenever R and k are at least 3, except for the case in which $R = 3$ and k is a power of 2, when one has $\Gamma^*(k, 3) \leqslant 36k^2$. This bound is better than those of Low, Pitman and Wolff and Davenport and Lewis when k is even and suitably large compared with R. Also, this result is notable because it shows that a bound of the form $\Gamma^*(k, R) \ll_R k^2$ is possible for all values of k.

The primary purpose of this paper is to make an improvement on the bound of Brüdern and Godinho through methods involving the use of Teichmüller representatives. To this end, we will prove the following theorem.

Received 28 July 1999; revised 1 May 2000.

2000 *Mathematics Subject Classification* 11D72, 11D88, 11E76, 11G25.

Work partially supported through a fellowship from the David and Lucile Packard Foundation.

J. London Math. Soc. (2) 63 (2001) 257–267. © London Mathematical Society 2001.

THEOREM 1. *For any $R \in \mathbb{N}$ and $k \geqslant 2$, one has*

$$\Gamma^*(k, R) \leqslant 4R^2k^2.$$

We note that when $R > 4$, the conclusion recorded in this theorem plainly improves on the aforementioned bound $\Gamma^*(k, R) \leqslant R^3k^2$ of Brüdern and Godinho [**3**]. If k is even and R is suitably small in terms of k, then this is the best known bound on $\Gamma^*(k, R)$. Theorem 1 is actually a corollary of a more precise estimate which is somewhat more complicated to state.

THEOREM 2. *Suppose that $\mathbb{K} = \mathbb{Q}_p$, and that $k = p^\tau k_0$, where $(k_0, p) = 1$. Then the following statements are true.*

(i) *If $p \neq 2$, then the system* (1) *of R diagonal equations of degree k has a nontrivial solution over $\mathbb{Q}_p$ provided only that*

$$N \geqslant R^2k^2 + R^2kk_0\left(\frac{p^\tau - 1}{p-1}\right) + 2Rk - R^2k.$$

Hence, whenever $p \neq 2$, one has $\Gamma^*(k, R, \mathbb{Q}_p) \leqslant \frac{3}{2}R^2k^2$.

(ii) *If $p = 2$, then the system* (1) *has a nontrivial solution over $\mathbb{Q}_p$ provided only that*

$$N \geqslant 4R^2k^2 - R^2kk_0 + 2Rk - R^2k.$$

Therefore, one has $\Gamma^*(k, R, \mathbb{Q}_2) \leqslant 4R^2k^2$.

When $\mathbb{K}$ is a finite extension of $\mathbb{Q}_p$, much less is known. In the case where $R = 1$, Birch [**2**] has recorded the bound $\Gamma^*(k, 1, \mathbb{K}) \leqslant (2\tau + 3)^k (m^2k)^{k-1}$, where $m = (k_0, p^f - 1)$, the numbers τ and k_0 are as in the statement of Theorem 2, and p^f is the cardinality of the residue field of $\mathbb{K}$. Note that this bound implies that

$$\Gamma^*(k, 1, \mathbb{K}) \leqslant \left(\frac{2\log k}{\log 2} + 3\right)^k k^{3k-3},$$

which is independent of the field $\mathbb{K}$, and indeed this appears to be the only such estimate in the literature. If the bound on $\Gamma^*(k, 1, \mathbb{K})$ is allowed to depend on the degree n of $\mathbb{K}$ over $\mathbb{Q}_p$, then Dodson [**8**] has shown that the bound $\Gamma^*(k, 1, \mathbb{K}) \leqslant 16n^2k^2(\log k)^2$ holds. If $\mathbb{K}$ is an unramified extension of $\mathbb{Q}_p$, then Dodson notes that his method leads to the bound $\Gamma^*(k, 1, \mathbb{K}) \leqslant 36k^2(\log k)^2$, which is independent of the degree of $\mathbb{K}$ over $\mathbb{Q}_p$. Additionally, Skinner [**13**] has shown that if $\mathbb{K}$ is a finite extension of $\mathbb{Q}_p$ and $k = p^\tau$ is a power of p, then the bound $\Gamma^*(k, 1, \mathbb{K}) \leqslant k((k+1)^{2\tau+1} - 1) + 1$ holds. (Skinner claims in [**13**] to prove this bound for all exponents k, but this is incorrect. The crucial error is in the proof of his Lemma 5, in which he uses Hensel's lemma to lift kth power residues modulo the maximal ideal of $\mathbb{K}$ to kth powers in $\mathbb{K}$. Unfortunately, some of these may be the zero residue, in which case Hensel's lemma may not be applied.) From the above bounds for one equation, bounds for systems of equations can be derived from statements due to Leep and Schmidt [**10**]. While proving their second basic inequality, Leep and Schmidt show that one has

$$\Gamma^*(k, R, \mathbb{K}) \leqslant \Gamma^*(k, 1, \mathbb{K})^R,$$

which provides a bound exponential in R. Furthermore, as a consequence of [**10**, Theorem 1] one also has

$$\Gamma^*(k, R, \mathbb{K}) \ll_{k, \mathbb{K}} R^{2^{k-1}},$$

giving a bound polynomial in R. However, one feels that the correct bounds should be rather smaller than these.

The methods used to prove Theorem 2 may also be used to develop a bound for $\Gamma^*(k,R,\mathbb{K})$ for arbitrary $\mathfrak{p}$-adic fields $\mathbb{K}$.

THEOREM 3. *Let $\mathbb{K}$ be a finite extension of $\mathbb{Q}_p$ of degree n. Suppose that the ramification index of $\mathbb{K}$ is e, and that the residue field of $\mathbb{K}$ contains $q = p^f$ elements. Suppose that $k = p^\tau k_0$, where $(k_0, p) = 1$. Then the system of equations* (1) *of degree k has a nontrivial $\mathbb{K}$-rational solution provided that*

$$N \geqslant R^2 k k_0 \left(\frac{q^{2e\tau+1}-1}{q-1}\right) - R^2 k + 2Rk.$$

Note that Theorem 3 implies that one has

$$\begin{aligned}\Gamma^*(k,R,\mathbb{K}) &\leqslant R^2 k k_0 \left(\frac{q^{2e\tau+1}-1}{q-1}\right) - R^2 k + 2Rk \\ &\leqslant R^2 k^{2+2n\tau} \\ &\leqslant R^2 k^{2+2n(\log k)/(\log 2)}.\end{aligned}$$

Although this bound is still not as strong as could be desired, and in particular is not independent of the degree of $\mathbb{K}$ over $\mathbb{Q}_p$, it does at least show that one has $\Gamma^*(k,R,\mathbb{K}) \ll_{k,\mathbb{K}} R^2$.

We prove these theorems by making a small improvement on the method of Brüdern and Godinho. We begin by employing a suitable normalization process, and then using the idea due to Low, Pitman and Wolff of partitioning the matrix of coefficients of our system into disjoint submatrices which are all nonsingular modulo a generator of the maximal ideal of $\mathbb{K}$. We then attempt to find a nonsingular solution of the system modulo a suitably high power of the maximal ideal of $\mathbb{K}$. By setting variables corresponding to columns of the same submatrix equal to each other, we obtain a new system of congruences to solve. Our improvement is to now solve this system by restricting our variables to be elements of the Teichmüller set $T_{\mathbb{K}} = \{x \in \mathbb{K} \mid x^q = x\}$. Suppose first that k has the special form $k = q^t k_0$. If $x \in T_{\mathbb{K}}$, then we have $x^k = x^{k_0}$, and so we need only to solve a system of congruences of degree k_0. We solve these congruences through an extension of a theorem of Schanuel. Finally, we use Hensel's lemma to lift this solution to a $\mathbb{K}$-rational solution of (1). Should k not have the above shape, we show that when solving the system of congruences, k may be replaced by a different exponent which does have this form, and apply our argument to the resulting set of equations.

This plan is actually employed to prove both Theorem 2 and Theorem 3. The reason why Theorem 2 is not merely a special case of Theorem 3 is that in the general case one needs to use the standard version of Hensel's lemma to lift a solution of a congruence to a solution in $\mathbb{K}$. However, the theory of kth power residues of rational integers leads to a better version of Hensel's lemma when $\mathbb{K} = \mathbb{Q}_p$.

2. *Normalization and preliminaries*

In what follows, $\mathbb{K}$ will be a finite extension of degree n of the field $\mathbb{Q}_p$, with maximal ideal generated by π. The ramification index will be denoted e, and we set $f = n/e$ so that $\pi^e = p$ and the residue field of $\mathbb{K}$ modulo π has cardinality $q = p^f$. Let

$\mathfrak{O}_{\mathbb{K}}$ represent the integers of $\mathbb{K}$. Finally, we denote by $\mathbb{F}_q$ the finite field containing q elements.

Before we can prove the theorems, we must discuss the concept of the normalization of a system of equations. For $1 \leqslant j \leqslant N$, let $\mathbf{a}_j$ denote the column in the matrix of coefficients of (1) corresponding to the variable x_j, and set

$$\Theta(\mathbf{F}) = \prod_{1 \leqslant i_1 < \ldots < i_R \leqslant R} \det((\mathbf{a}_{i_j})_{1 \leqslant j \leqslant R}).$$

By a standard argument involving the compactness of the $\mathfrak{p}$-adic field $\mathbb{K}$, we may assume that $\Theta(\mathbf{F}) \neq 0$, and we make this assumption throughout this paper. (One may see [**7**, pp. 572–573] for an example of this argument. As with [**7**, Lemma 11], which will be quoted shortly, although this fact is written down only for the case $\mathbb{K} = \mathbb{Q}_p$, it is not hard to see that it extends to general $\mathfrak{p}$-adic fields by merely replacing occurrences of $\mathbb{Q}_p$ and p by $\mathbb{K}$ and π respectively.) Next, we say that two systems of additive equations with coefficients in $\mathfrak{O}_{\mathbb{K}}$ are *equivalent* if one can be obtained from the other through a combination of the following three operations:

(i) replacing a variable x_i by $\pi^{\alpha} x_i$ for some integer α;
(ii) dividing one or more equations by an integral power of π;
(iii) taking nonsingular $\mathfrak{O}_{\mathbb{K}}$-linear combinations of the equations.

A system $\mathbf{F}$ is said to be *π-normalized* if both $\Theta(\mathbf{F}) \neq 0$ and the power of π dividing $\Theta(\mathbf{F})$ is less than or equal to the power of π dividing $\Theta(\mathbf{G})$ for all systems $\mathbf{G}$ equivalent to $\mathbf{F}$. Since any system $\mathbf{F}$ with $\Theta(\mathbf{F}) \neq 0$ is equivalent to a π-normalized system, it suffices to show that Theorems 2 and 3 are true for π-normalized systems. The major benefit of working with normalized systems is that they have nice properties when considered modulo π. In particular, we have the following lemma.

LEMMA 1. *A π-normalized system of additive forms can be written (after renumbering the variables) as*

$$F_i = f_i(x_1, \ldots, x_r) + \pi g_i(x_{r+1}, \ldots, x_N)$$

for $i = 1, \ldots, R$, where $r \geqslant N/k$, and if $1 \leqslant i \leqslant r$, then the coefficient of x_i in at least one of the forms is not divisible by π.

Moreover, if we form any s linear combinations of $F_1, \ldots, F_R$ (these combinations being independent modulo π), and denote by q_s the number of variables that occur in at least one of these combinations with a coefficient not divisible by π, then for each s with $1 \leqslant s \leqslant R-1$, we have

$$q_s \geqslant \frac{sN}{Rk}.$$

Proof. After making the changes indicated above, this is [**7**, Lemma 11]. □

Following Brüdern and Godinho, let A be the matrix of coefficients of the variables $x_1, \ldots, x_r$. In [**3**], they use the idea of Low, Pitman and Wolff [**11**] of showing when $\mathbb{K} = \mathbb{Q}_p$ that A has many disjoint $R \times R$ submatrices which are nonsingular modulo p. We repeat this argument here, making the trivial modifications necessary to apply it to general $\mathfrak{p}$-adic fields $\mathbb{K}$.

Let $\mu(d)$ be the maximal number of columns of A which lie in a d-dimensional linear subspace of $\mathbb{F}_q^R$. Then for $1 \leqslant s \leqslant R$, we have

$$q_s + \mu(R-s) = r.$$

The idea of Low, Pitman and Wolff was to make use of the following lemma, which is a special case of a theorem of Aigner [**1**].

LEMMA 2. *Let A be an $R \times r$ matrix over a field $\mathbb{K}$ and let m be a positive integer. The matrix A includes m disjoint $R \times R$ submatrices which are nonsingular over $\mathbb{K}$ if and only if we have*

$$r - l \geqslant m(R - \operatorname{rank}((\mathbf{a}_{i_j})_{1 \leqslant j \leqslant l}))$$

for any $l \leqslant r$ and $1 \leqslant i_1 < i_2 < \ldots < i_l \leqslant r$.

Proof. This is [**11**, Lemma 1]. One can find in [**1**] a proof of Aigner's more general result for matroids. □

Observe that this lemma implies that A has m disjoint submatrices with determinants not divisible by π if and only if

$$r - \mu(d) \geqslant m(R - d) \quad \text{for all } 0 \leqslant d \leqslant R,$$

and that this is equivalent to the condition

$$q_s \geqslant ms$$

for $1 \leqslant s \leqslant R$. Hence, we have the following lemma, which is (after making the obvious changes) [**3**, Lemma 2].

LEMMA 3. *Suppose that* (1) *is a π-normalized system written in the form given in Lemma* 1. *Then the $R \times r$ matrix A contains at least* $[N/(Rk)]$ *disjoint submatrices which are nonsingular modulo π.*

Now that we have dealt with normalization, we need to prove a lemma about solutions of congruences modulo powers of π. This generalizes the result given by Schanuel in [**12**]. As above, the proof of this lemma differs only trivially from Schanuel's proof, but is given here for completeness.

LEMMA 4. *Suppose that $\mathbb{K}$ is a finite extension of $\mathbb{Q}_p$ of degree n, with maximal ideal generated by π. For $1 \leqslant i \leqslant R$, let F_i be a polynomial of degree k_i in N variables with coefficients in $\mathfrak{O}_{\mathbb{K}}$ and no constant term. Finally, let $T_{\mathbb{K}} = \{x \in \mathfrak{O}_{\mathbb{K}} \mid x^q = x\}$ be the set of Teichmüller representatives of $\mathfrak{O}_{\mathbb{K}}/(\pi)$. Then the system of equations*

$$F_i(x_1, \ldots, x_N) \equiv 0 \pmod{\pi^{v_i}} \quad (1 \leqslant i \leqslant R)$$

has a nontrivial solution in $T_{\mathbb{K}}^N$ provided that

$$N > \sum_{i=1}^{R} k_i \frac{q^{v_i} - 1}{q - 1}.$$

Proof. For any function $F \in \mathfrak{O}_{\mathbb{K}}[x_1, \ldots, x_N]$, define $\Delta F = \Delta^{(1)} F$ by

$$\begin{aligned}(\Delta F)(x_1, \ldots, x_N) = \pi^{-1}(&F(x_1, \ldots, x_N)^q - F^{(q)}(x_1^q, \ldots, x_N^q) \\ &+ F^{(q)}(x_1, \ldots, x_N) - F(x_1, \ldots, x_N)),\end{aligned}$$

where $F^{(q)}(x_1, \ldots, x_N)$ denotes the polynomial obtained by raising each coefficient of F to the qth power, and define $\Delta^{(j)} F = \Delta(\Delta^{(j-1)} F)$ when $j > 1$. Note that ΔF is a polynomial in $x_1, \ldots, x_N$ of degree at most $q \cdot \deg F$. Since the map sending x to x^q is the identity homomorphism on $\mathbb{F}_q$, all of the coefficients of the polynomial in parentheses are elements of $\mathfrak{O}_{\mathbb{K}}$ congruent to 0 modulo π. Hence, all of the coefficients of ΔF are in $\mathfrak{O}_{\mathbb{K}}$. Furthermore, if $F(x_1, \ldots, x_N) = c$ is a constant polynomial, then we have

$$\Delta F = \Delta c = \pi^{-1}(c^q - c).$$

We wish to show that we have $c \equiv 0 \pmod{\pi^v}$ if and only if the numbers $\Delta^{(0)}c, \Delta^{(1)}c, \ldots, \Delta^{(v-1)}c$ are all congruent to 0 modulo π, where we set $\Delta^{(0)}c = c$ for convenience. To see this, note that if $|\Delta^{(j)}c|_\pi < 1$, where we have normalized so that $|\pi|_\pi = p^{-1/e}$, then we have

$$|\Delta^{(j+1)}c|_\pi = |\pi^{-1}((\Delta^{(j)}c)^q - \Delta^{(j)}c)|_\pi = p^{1/e}|\Delta^{(j)}c|_\pi.$$

Using this equation inductively, we find that if $|c|_\pi = p^{-\tau/e}$, with $\tau > 0$, then we have $|\Delta^{(i)}c|_\pi = p^{(i-\tau)/e}$ provided that $|\Delta^{(i-1)}c|_\pi < 1$, and hence $|\Delta^{(j)}c|_\pi < 1$ for $0 \leqslant j \leqslant \tau - 1$.

Suppose that $c \equiv 0 \pmod{\pi^v}$. Then $|\Delta^{(0)}c|_\pi \leqslant p^{-v/e}$, and the above statement implies that $|\Delta^{(j)}c|_\pi < 1$ for $0 \leqslant j \leqslant v-1$. We therefore have $\Delta^{(j)}c \equiv 0 \pmod{\pi}$ whenever $0 \leqslant j \leqslant v-1$. Conversely, suppose that $\Delta^{(0)}c, \Delta^{(1)}c, \ldots, \Delta^{(v-1)}c$ are all congruent to 0 modulo π. Then the equation displayed above holds for each j with $0 \leqslant j \leqslant v-2$, and we find inductively that $|c|_\pi = p^{(1-v)/e}|\Delta^{(v-1)}c|_\pi \leqslant p^{-v/e}$, whence c is congruent to 0 modulo π^v.

Now suppose that $a_1, \ldots, a_N$ are elements of $T_{\mathbb{K}}$. Then since $a_j^q = a_j$, we have $F^{(q)}(a_1^q, \ldots, a_n^q) = F^{(q)}(a_1, \ldots, a_N)$, and hence

$$\begin{aligned}(\Delta F)(a_1, \ldots, a_N) &= \pi^{-1}[F(a_1, \ldots, a_N)^q - F(a_1, \ldots, a_N)] \\ &= \Delta(F(a_1, \ldots, a_N)).\end{aligned}$$

Therefore, for such elements, one has $F(a_1, \ldots, a_N) \equiv 0 \pmod{\pi^v}$ if and only if $(\Delta^{(j)}F)(a_1, \ldots, a_N) \equiv 0 \pmod{\pi}$ for $0 \leqslant j \leqslant v-1$.

In view of the above discussion, solving the system of equations

$$F_i(x_1, \ldots, x_N) \equiv 0 \pmod{\pi^{v_i}} \quad (1 \leqslant i \leqslant R)$$

nontrivially with variables in $T_{\mathbb{K}}$ is equivalent to nontrivially solving the system of congruences

$$(\Delta^{(j)}F_i)(a_1, \ldots, a_N) \equiv 0 \pmod{\pi} \quad (1 \leqslant i \leqslant R, 1 \leqslant j \leqslant v_i).$$

This is a system of $v_1 + \ldots + v_R$ congruences modulo π, where for each i we have v_i congruences of degrees at most $k_i, k_i q, \ldots, k_i q^{v_i - 1}$. By the Chevalley–Warning theorem (see [**14**]), we may do this provided that

$$N > \sum_{i=1}^{R} \sum_{j=0}^{v_i - 1} k_i q^j = \sum_{i=1}^{R} k_i \left(\frac{q^{v_i} - 1}{q - 1}\right),$$

which is our desired bound. □

Now that we have a lemma allowing us to solve congruences modulo powers of π, we need one more lemma which tells us that we may 'lift' such a solution to a solution of an equation over $\mathbb{K}$. Therefore, we give one form of Hensel's lemma.

LEMMA 5. *Suppose that $F(\mathbf{x})$ is a polynomial in r variables, with coefficients in $\mathbb{K}$, and $\mathbf{a} \in \mathfrak{O}_{\mathbb{K}}^r$ satisfies the equation*

$$|F(\mathbf{a})|_\pi < \left|\frac{\partial F}{\partial x_i}(\mathbf{a})\right|_\pi^2$$

for some variable x_i. Then there exists a unique $\mathbf{a}^* \in \mathfrak{O}_{\mathbb{K}}^r$ such that both $F(\mathbf{a}^*) = 0$ and

$$\max_{1 \leqslant i \leqslant r} |a_i^* - a_i|_\pi \leqslant \left|\frac{\partial f}{\partial x_i}(\mathbf{a})\right|_\pi^{-1} |F(\mathbf{a})|_\pi.$$

In particular, if $\mathbf{a}$ is nontrivial modulo π, then $\mathbf{a}^*$ is a nontrivial solution of $F(\mathbf{x}) = 0$.

While versions of Hensel's lemma exist which simultaneously lift solutions of several congruences to solutions of equations in $\mathbb{K}$, we will see later that this version is all that we need. A very thorough exposition of Hensel's lemma may be found in [**9**, Chapter 5].

3. *The proof of Theorem* 3

In order to prove Theorem 3, we follow in general the method used by Brüdern and Godinho to prove [**3**, Theorem 4]. As mentioned above, we may assume that our system (1) is reduced. Then, by Lemma 3, it is equivalent to a system

$$\begin{array}{c} b_{1,1}x_1^k + \ldots + b_{1,r}x_r^k + b_{1,r+1}x_{r+1}^k + \ldots + b_{1,N}x_N^k = 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \quad \vdots \\ b_{R,1}x_1^k + \ldots + b_{R,r}x_r^k + b_{R,r+1}x_{r+1}^k + \ldots + b_{R,N}x_N^k = 0, \end{array} \tag{2}$$

which is in the form given in Lemma 1 and has the property that the matrix of coefficients of the variables $x_1, \ldots, x_r$ contains at least $s = [N/(Rk)]$ disjoint submatrices $B_0, \ldots, B_{s-1}$ which are nonsingular modulo π. By relabeling variables if necessary, we may assume that for $0 \leqslant l \leqslant s-1$, the columns of B_l correspond to the variables $x_{lR+1}, \ldots, x_{(l+1)R}$. We now set $x_{sR+1} = \ldots = x_N = 0$, and attempt to solve the system of congruences

$$\begin{array}{c} b_{1,1}x_1^k + \ldots + b_{1,r}x_r^k \equiv 0 \pmod{\pi^{2e\tau+1}} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ b_{R,1}x_1^k + \ldots + b_{R,r}x_r^k \equiv 0 \pmod{\pi^{2e\tau+1}}. \end{array} \tag{3}$$

Next, for $0 \leqslant l \leqslant s-2$ and $1 \leqslant i \leqslant R$, we define

$$c_{i,l} = \sum_{j=lR+1}^{(l+1)R} b_{i,j},$$

and consider the system of equations

$$\sum_{l=0}^{s-2} c_{i,l} y_l^{k_0} + \sum_{j=R(s-1)+1}^{Rs} b_{i,j} y_j^{k_0} \equiv 0 \pmod{\pi^{2e\tau+1}} \quad (1 \leqslant i \leqslant R). \tag{4}$$

This is a system of R equations of degree k_0 in $s+R-1$ variables. By Lemma 4, we can solve this system with each of the variables in $T_{\mathbb{K}}$ provided that

$$s + R - 1 > Rk_0\left(\frac{q^{2e\tau+1} - 1}{q-1}\right).$$

That is,

$$s \geqslant Rk_0\left(\frac{q^{2e\tau+1}-1}{q-1}\right)-R+2. \tag{5}$$

Writing $\tau = \alpha f+\beta$ with $0 \leqslant \beta < f$, we see that since the elements x of $T_{\mathbb{K}}$ have the property that $x^q = x$, any solution of the system (4) is also a solution of the system

$$\sum_{l=0}^{s-2} c_{i,l}\, y_l^{k_0 p^{(\alpha+1)f}} + \sum_{j=R(s-1)+1}^{Rs} b_{i,j}\, y_j^{k_0 p^{(\alpha+1)f}} \equiv 0 \pmod{\pi^{2e\tau+1}} \qquad (1 \leqslant i \leqslant R). \tag{6}$$

We now set $x_{lR+1} = \ldots = x_{(l+1)R} = y_l^{p^{f-\beta}}$ for $0 \leqslant l \leqslant s-2$, and $x_j = y_j^{p^{f-\beta}}$ when $R(s-1)+1 \leqslant j \leqslant Rs$. Since we have

$$(y_l^{p^{f-\beta}})^k = y_l^{k_0 p^{(\alpha+1)f}},$$

the vector $(x_1, \ldots, x_r)$ is then a solution of the system (3).

Now, we must lift this solution to a solution of (2) over $\mathbb{K}$. Since the matrix of coefficients of the second summation in (6) is nonsingular modulo π, at least one of the y_l with $0 \leqslant l \leqslant s-2$ must be nonzero modulo π. Without loss of generality, suppose that $y_0 \not\equiv 0 \pmod{\pi}$. Consider the matrix B_0 of coefficients of the variables $x_1, \ldots, x_R$ in (2). Recall that this matrix is nonsingular modulo π, and hence nonsingular. One may therefore apply elementary row operations to the system (2) which transform B_0 into a diagonal matrix. Since we have set $x_{sR+1} = \ldots = x_N = 0$, when we look at the resulting system modulo $\pi^{2e\tau+1}$ we obtain a system

$$\begin{array}{ccccccc} b_{1,1}x_1^k & + & b_{1,R+1}x_{R+1}^k & + \ldots + & b_{1,r}x_r^k & \equiv 0 & \pmod{\pi^{2e\tau+1}} \\ \vdots & & \vdots & & \vdots & \vdots & \\ b_{R,R}x_R^k & + & b_{R,R+1}x_{R+1}^k & + \ldots + & b_{R,r}x_r^k & \equiv 0 & \pmod{\pi^{2e\tau+1}}, \end{array} \tag{7}$$

which is equivalent to (3). Since elementary row operations do not change the determinant of a matrix, the image of B_0 under this transformation is still nonsingular modulo π, and so $b_{1,1}b_{2,2}\ldots b_{R,R} \not\equiv 0 \pmod{\pi}$. Finally, because we have only taken linear combinations of equations, our solution of (3) is also a solution of (7).

We now wish to find $\zeta_1, \ldots, \zeta_R \in \mathbb{K}$ such that we have

$$\begin{array}{cccccc} b_{1,1}\zeta_1^k & + & b_{1,R+1}x_{R+1}^k & + \ldots + & b_{1,r}x_r^k & = 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ b_{R,R}\zeta_R^k & + & b_{R,R+1}x_{R+1}^k & + \ldots + & b_{R,r}x_r^k & = 0. \end{array} \tag{8}$$

If we consider $x_{R+1}, \ldots, x_r$ to be fixed, then we have a system of R equations, each in only one variable ζ_i, and the variable in each equation is different. Hence, we may use Hensel's lemma on each equation separately. If, for $1 \leqslant i \leqslant R$, we set

$$G_i(t_i) = b_{i,i}\,t_i^k + (b_{i,R+1}x_{R+1}^k + \ldots + b_{i,r}x_r^k),$$

then we have $G_i(x_i) \equiv 0 \pmod{\pi^{2e\tau+1}}$, whence

$$|G_i(x_i)|_\pi < p^{-2\tau},$$

and

$$|G_i'(x_i)|_\pi = |kb_{i,i}x_i^{k-1}|_\pi = |k|_\pi = p^{-\tau}.$$

Hence, for each value of i, we have

$$|G_i(x_i)|_\pi < |G_i'(x_i)|_\pi^2.$$

Therefore, we may apply Lemma 5 to see that the desired ζ_i all exist, and we have found a $\mathbb{K}$-rational solution of (8). When combined with our having set $x_{r+1} = \ldots = x_N = 0$, this provides us with a $\mathbb{K}$-rational solution of the system (2). Finally, since the systems (2) and (1) are equivalent, there is a nontrivial $\mathbb{K}$-rational solution of (1).

Hence, there is a nontrivial solution of (1) provided only that the lower bound given in (5) holds. However, by Lemma 3, this bound will hold if

$$\left[\frac{N}{Rk}\right] \geqslant Rk_0\left(\frac{q^{2e\tau+1}-1}{q-1}\right) - R + 2,$$

which is true provided that

$$N \geqslant R^2kk_0\left(\frac{q^{2e\tau+1}-1}{q-1}\right) - R^2k + 2Rk,$$

which is the desired bound.

4. *The proof of Theorem* 2

Throughout this section, we will assume that $\mathbb{K} = \mathbb{Q}_p$. While one could trivially obtain a bound on the value of N needed in this case by setting $e = f = 1$ and $q = p$ and applying Theorem 3, the following version of Hensel's lemma, which is a consequence of the theory of kth power residues of rational integers, makes it possible to do better.

LEMMA 6. *Suppose that $p^\tau \| k$, and define $\gamma = \gamma(k,p)$ by*

$$\gamma = \begin{cases} 1 & \text{if } \tau = 0 \\ \tau + 1 & \text{if } \tau > 0 \text{ and } p > 2 \\ \tau + 2 & \text{if } \tau > 0 \text{ and } p = 2. \end{cases}$$

Then if the congruence

$$ax^k + b \equiv 0 \pmod{p^\gamma}$$

with $ab \not\equiv 0 \pmod{p}$ is soluble, then the equation

$$ax^k + b = 0$$

has a nonzero solution in $\mathbb{Q}_p$.

A proof of this result can be found in [**4**, p. 36].

If $\gamma = 1$, then $p \nmid k$, and the following result of Brüdern and Godinho shows that the theorem is true in this case.

LEMMA 7. *Let p be a prime, $p \nmid k$, and $N \geqslant Rk(R(k,p-1) - R + 2)$. Then the system of equations* (1) *admits a nontrivial solution in $\mathbb{Q}_p$.*

Proof. This is [**3**, Theorem 3]. □

In the other cases, we proceed as in the proof of Theorem 3, except that we use Lemma 7 instead of the standard version of Hensel's lemma (Lemma 5). In particular, we need to solve the system (4), except that the congruences are now modulo p^γ. If $p \neq 2$, then $\gamma = \tau + 1$, and by Lemma 4 we can solve this system with elements of the Teichmüller set provided that

$$s \geqslant Rk_0\left(\frac{p^{\tau+1}-1}{p-1}\right) + 2 - R.$$

Therefore, as above, it is enough to have

$$\left[\frac{N}{Rk}\right] \geqslant Rk_0\left(\frac{p^{\tau+1}-1}{p-1}\right)+2-R,$$

and this is true provided that

$$\begin{aligned} N &\geqslant Rk\left(Rk_0\left(\frac{p^{\tau+1}-1}{p-1}\right)+2-R\right) \\ &= R^2k^2+R^2kk_0\left(\frac{p^{\tau}-1}{p-1}\right)+2Rk-R^2k, \end{aligned}$$

as desired. Note that if $R \geqslant 2$, then this bound is strictly less than $\frac{3}{2}R^2k^2$. Moreover, when $R = 1$, one can use the bound $N \geqslant k^2+1$ due to Davenport and Lewis [**5**] to show that the bound $N \geqslant \frac{3}{2}R^2k^2$ also suffices in this situation. Hence, this bound suffices whenever $p \neq 2$.

In the final case, we have $p = 2$ and $\gamma = \tau+2$. Then we wish to solve the sysem (4), again with the congruences now being modulo p^{γ}. By Lemma 4, this can be done provided that

$$s \geqslant Rk_0(2^{\tau+2}-1)+2-R.$$

Again, it suffices to have

$$\left[\frac{N}{Rk}\right] \geqslant Rk_0(2^{\tau+2}-1)+2-R,$$

and this is true provided that

$$\begin{aligned} N &\geqslant Rk(Rk_0(2^{\tau+2}-1)+2-R) \\ &= 4R^2k^2-R^2kk_0+2Rk-R^2k, \end{aligned}$$

as desired. It is trivial to see that this lower bound cannot be greater than $4R^2k^2$. Therefore, if $N \geqslant 4R^2k^2$, then the system (1) has a nontrivial $\mathbb{Q}_2$-rational solution.

Acknowledgements. The author would like to thank Professor Trevor Wooley for his helpful suggestions and encouragement. This work forms part of my PhD thesis for the University of Michigan.

References

1. M. Aigner, *Combinatorial theory* (Springer, New York, 1979).
2. B. J. Birch, 'Diagonal equations over 𝔭-adic fields', *Acta Arith.* 9 (1964) 291–300.
3. J. Brüdern and H. Godinho, 'On Artin's conjecture. I: Systems of diagonal forms', *Bull. London Math. Soc.* 31 (1999) 305–313.
4. H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities* (Ann Arbor Publishers, Ann Arbor, MI, 1962).
5. H. Davenport and D. J. Lewis, 'Homogeneous additive equations', *Proc. Roy. Soc. London Ser.* A 274 (1963) 443–460.
6. H. Davenport and D. J. Lewis, 'Two additive equations', *Proc. Sympos. Pure Math.* 12 (1967) 74–98.
7. H. Davenport and D. J. Lewis, 'Simultaneous equations of additive type', *Proc. Roy. Soc. London Ser.* A 264 (1969) 557–595.
8. M. Dodson, 'Some estimates for diagonal equations over 𝔭-adic fields', *Acta Arith.* 40 (1982) 117–124.
9. M. J. Greenberg, *Lectures on forms in many variables* (W. A. Benjamin, New York, 1969).

10. D. B. LEEP and W. M. SCHMIDT, 'Systems of homogeneous equations', *Invent. Math.* 71 (1983) 539–549.

11. L. LOW, J. PITMAN and A. WOLFF, 'Simultaneous diagonal congruences', *J. Number Theory* 29 (1988) 31–59.

12. S. H. SCHANUEL, 'An extension of Chevalley's theorem to congruences modulo prime powers', *J. Number Theory* 6 (1974) 284–290.

13. C. M. SKINNER, 'Solvability of $\mathfrak{p}$-adic diagonal equations', *Acta Arith.* 75 (1996) 251–258.

14. E. WARNING, 'Bemerkung zur vorstehenden Arbeit von Herrn Chevalley', *Abh. Math. Sem. Univ. Hamburg* 11 (1935) 76–83.

Department of Mathematics
University of Michigan
525 East University Avenue
Ann Arbor
MI 48109-1109
USA