ACTA ARITHMETICA
LXXV.3 (1996)

Solvability of $\mathfrak{p}$-adic diagonal equations

by

Christopher M. Skinner (Princeton, N.J.)

1. Introduction. Let p be a prime, let $\mathbb{Q}_p$ denote the p-adic numbers, and let K be a finite extension of $\mathbb{Q}_p$. One of the fundamental questions in the study of diophantine equations asks: when does an equation of the form

$$a_1x_1^k + \ldots + a_sx_s^k = 0, \quad a_i \in K,\ k \geq 2, \tag{1}$$

have a non-trivial solution over K? (By "non-trivial solution" we mean a non-zero vector $\mathbf{x} = (x_1, \ldots, x_s) \in K^s$ satisfying (1).) When $K = \mathbb{Q}_p$, it is well known that it suffices to have $s \geq k^2 + 1$. More generally, suppose $k = p^t m$, $(m,p) = 1$, f is the residue class degree of K, and $d = (m, p^f - 1)$. Birch [B] has shown that for any K, it suffices to have $s \geq (2t+3)^k(d^2k)^{k-1}$. It is the purpose of this note to improve the result of Birch, by essentially reducing the exponent k to $\log k$. Specifically, we prove the following theorem.

Theorem. *If $s \geq k((k+1)^{\max(2t,1)} - 1) + 1$, then any equation of the form* (1) *has a non-trivial solution over K. In particular*, *if $(k,p) = 1$, then it suffices to have $s \geq k^2 + 1$.*

If K is unramified over $\mathbb{Q}_p$, then it is possible to replace the $2t$ of the Theorem with a constant. A proof of such a result is indicated in [D]. It is also possible to generalize the results of Schmidt [S] for simultaneous additive equations, at least in the case $(k,p) = 1$. However, in order to keep our exposition as elementary as possible, we do not treat either of these problems in this paper.

2. Notation and preliminaries. In what follows, $\mathfrak{O}$ is the ring of integers of K, $\mathfrak{p} = (\pi)$ is the maximal ideal of $\mathfrak{O}$, f is the residue class degree of K, e is the ramification index of p, and t and m are integers such that $k = p^t m$, with $(m,p) = 1$. Also, L is the maximal unramified subfield

The author was supported by an N.S.F. Graduate Fellowship and, during a visit to the University of Michigan, by the David and Lucile Packard Foundation.

of K, and $\mathfrak{o}$ is the ring of integers of L. Recall that $\{1, \pi, \ldots, \pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$.

Clearly, we lose no generality by assuming that $a_i \in \mathfrak{O}$ for all i, so henceforth we shall do so.

We write $\Gamma(k)$ for the least positive integer such that if $s \geq \Gamma(k)$, then any equation of the form (1) is solvable non-trivially over K. We use $\Gamma_1(k)$ to denote the similar function for those equations of the form (1) with the additional restriction that $a_i \not\equiv 0 \bmod \pi$ for all i.

We write that $\mathbf{x}$ is a "non-trivial solution mod π^n" if $\mathbf{x} = (x_1, \ldots, x_s) \in \mathfrak{O}^s$ is a solution of (1) modulo π^n and if $x_j \not\equiv 0 \bmod \pi$ for some j. We let $\Phi(k, n)$ denote the least positive integer such that if $s \geq \Phi(k, n)$, then any equation of the form (1) has a non-trivial solution mod π^n.

Our first lemma reduces the proof of the Theorem to showing that $\Phi(k, e) \leq k + 1$.

Lemma 1. (i) $\Gamma(k) \leq k(\Gamma_1(k) - 1) + 1$.

(ii) $\Gamma_1(k) \leq \Phi(k, \max(2et, 1))$.

(iii) $\Phi(k, (r + 1)e) \leq \Phi(k, e)\Phi(k, re) \leq \Phi(k, e)^{r+1}$.

(iv) *If* $\Phi(k, e) \leq (k + 1)$, *then*

$$\Gamma(k) \leq k((k + 1)^{\max(2t,1)} - 1) + 1.$$

Proof. (i) Write $a_i = \pi^{r_i k + c_i} b_i$ with $r_i \geq 0$, $0 \leq c_i < k$ and $(b_i, \pi) = 1$. If $s > k(c - 1)$, then by the Box Principle at least c of the c_i's are the same. We may assume the corresponding i's to be $i = 1, \ldots, c$. Thus it suffices to find a non-trivial solution of the equation

$$b_1 x_1^k + b_2 x_2^k + \ldots + b_c x_c^k = 0, \qquad (b_i, \pi) = 1. \tag{2}$$

That such a solution exists if $c \geq \Gamma_1(k)$ is a consequence of the definition of $\Gamma_1(k)$.

(ii) Assume $a_1 \not\equiv 0 \bmod \pi$ for all i. Put $r = \max(1, 2te)$. If $s \geq \Phi(k, r)$, then by the definition of $\Phi(k, r)$, there exists a non-trivial solution of (1) mod π^r. Let $\mathbf{x} = (x_1, \ldots, x_s)$ be such a solution. We may assume that $x_1 \not\equiv 0$ mod π. Choose $y_2, \ldots, y_s \in \mathfrak{o}$ such that $y_i \equiv x_i \bmod \pi^r$. Let $d = \sum_{i=2}^{s} a_i y_i^k$. Since $a_1 x_1^k + d \equiv 0 \bmod \pi^r$, it follows from Hensel's Lemma [La, II, Prop. 2] that we can find $y_1 \in \mathfrak{o}$ such that $y_1 \equiv x_1 \bmod \pi^r$ and $a_1 y_1^k + d = 0$. Thus $\mathbf{y} = (y_1, \ldots, y_c)$ is a non-trivial solution of (1).

(iii) Let $h = \Phi(k, re)$, $l = \Phi(k, e)$ and let

$$F_j(\mathbf{x}_j) = a_{jh+1} x_{jh+1}^k + \ldots + a_{(j+1)h} x_{(j+1)h}^k, \qquad j = 0, \ldots, l - 1.$$

Then (1) becomes

$$F_0(\mathbf{x}_0) + F_1(\mathbf{x}_1) + \ldots + F_{l-1}(\mathbf{x}_{l-1}) + \sum_{i=lh+1}^{s} a_i x_i^k = 0.$$

Thus, it suffices to find a non-trivial solution of

$$F_0(\mathbf{x}_0) + \ldots + F_{l-1}(\mathbf{x}_{l-1}) \equiv 0 \bmod \pi^{(r+1)e}. \tag{3}$$

By definition of $\Phi(k, re)$ there exist non-trivial solutions $\mathbf{y}_j$ of the l equations

$$F_j(\mathbf{x}_j) \equiv 0 \bmod \pi^{re}, \quad j = 0, \ldots, l-1.$$

Let $f_j = F_j(\mathbf{y}_j)$. Substituting $\mathbf{x}_j = t_j \mathbf{y}_j$ in (3) we get the new equation

$$f_0 t_0^k + \ldots + f_{l-1} t_{l-1}^k \equiv 0 \bmod \pi^{(r+1)e}, \quad f_j \equiv 0 \bmod \pi^{re}. \tag{4}$$

From the definition of $\Phi(k, e) = l$, (4) has a non-trivial solution $\mathbf{t} = (t_0, \ldots, t_{\Phi(k,e)-1})$. Thus, $\mathbf{y} = (t_0 \mathbf{y}_0, \ldots, t_{\Phi(k,e)-1} \mathbf{y}_{\Phi(k,e)-1}, 0, \ldots, 0) \in \mathfrak{o}^s$ is a non-trivial solution of (1) modulo $\pi^{(r+1)e}$.

(iv) This follows upon combining parts (i)–(iii). ∎

3. Some results about linear systems. Before we can prove that $\Phi(k, e) \le k + 1$, we need some facts about linear systems of a particular type.

In this section, F is an arbitrary field, and for any non-negative integers a and b, $\mathbf{M}_{a,b}(F)$ is the ring of matrices over F of size $a \times b$.

Let c, r, and n be positive integers, and let

$$A_{ij} \in \mathbf{M}_{r_i,n}(F), \quad i = 1, \ldots, c, \ j = 1, \ldots, i, \ r_i \le r, \tag{5a}$$

be arbitrary matrices. We allow "empty" matrices (i.e. $r_i = 0$). Consider the block matrix

$$A = \begin{pmatrix} A_{11} & 0 & & \ldots & 0 \\ A_{21} & A_{22} & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ A_{c1} & & \ldots & & A_{cc} \end{pmatrix}. \tag{5b}$$

Definition. We say that any matrix A of the form (5a,b) is (c, r, n)-*good* if

1. for each i, the non-zero row vectors of A_{ii} are linearly independent over F, and

2. for each q, the qth row of $(A_{i1}\ A_{i2}\ \ldots\ A_{ii})$ is non-zero iff the qth row of A_{ii} is non-zero.

Note that both conditions are trivially satisfied by matrices with $r_i = 0$. The following lemma partially motivates our use of the adjective "good."

Lemma 2. *Suppose A is (c,r,n)-good with $n > r$, and suppose $\mathbf{X} = (x_1, \ldots, x_n)$ is a non-zero solution of the linear system*

$$A_{11}\mathbf{X} = \mathbf{0}.$$

(*For* A_{11} *empty*, *any* $\mathbf{X}$ *is a solution.*) *Then the linear system*

$$A\mathbf{Y} = \mathbf{0} \tag{6}$$

has a solution $\mathbf{Y} = (y_1, \ldots, y_{cn})$ *such that* $y_i = x_i$ *for* $i = 1, \ldots, n$.

Proof. We will proceed by induction on c. The claim is trivially true for $c = 1$. Suppose $c > 1$. Write

$$A = \begin{pmatrix} B_1 & 0 \\ B_2 & A_{cc} \end{pmatrix}.$$

B_1 is $(c-1, r, n)$-good, so by hypothesis there exists a solution $\mathbf{Y}_1 = (y_1, \ldots, y_{(c-1)n})$ of the linear system

$$B_1\mathbf{Y}_1 = \mathbf{0}$$

such that $y_1 = x_1$ for $i = 1, \ldots, n$. Let $\mathbf{D} = B_2\mathbf{Y}_1$. It follows from Part 2 of the definition of a good matrix that the qth entry of $\mathbf{D}$ is zero if the qth row of A_{cc} is zero. By Part 1 of the definition of a good matrix, the non-zero rows of A_{cc} are linearly independent. Thus, since $n > r \geq \operatorname{rank}(A_{cc})$ the linear system

$$A_{cc}\mathbf{Y}_2 = -\mathbf{D}$$

has a solution in F. It follows that $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2)$ is the desired solution to (6). ∎

Next, we consider a slightly more general system, though still of a very special type. Again, let c, r, n be positive integers. Let

$$M_{i,j} \in \mathbf{M}_{r_j,n}(F), \qquad i = 1, \ldots, c,\ j = 1, \ldots, c - i + 1,\ \sum_{j=1}^{c} r_j \leq r. \tag{7a}$$

We allow empty matrices (i.e. $r_j = 0$). Consider the block matrix

$$M = \begin{pmatrix} M_{1,1} & 0 & & \ldots & 0 \\ M_{1,2} & 0 & & \ldots & 0 \\ \vdots & \vdots & & & \vdots \\ M_{1,c} & 0 & & \ldots & 0 \\ M_{2,1} & M_{1,1} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ M_{2,c-1} & M_{1,c-1} & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ M_{c,1} & M_{c-1,1} & M_{c-2,1} & \ldots & M_{11} \end{pmatrix}. \tag{7b}$$

LEMMA 3. *If M is any matrix of the form* (7a,b), *then there exists an invertible matrix P such that $M' = PM$ is (c, r, n)-good.*

Proof. We will proceed again by induction on c. There is an invertible Q such that $QM_{1,1} = \binom{N_{1,1}}{0}$, where the rows of $N_{1,1}$ are non-zero and linearly independent. Suppose that $N_{1,1}$ has ν rows, so that $QM_{1,1}$ has $r_1 - \nu$ zero rows. For every $k = 1, \ldots, c$,

$$Q(M_{k,1}\ M_{k-1,1}\ \ldots\ M_{1,1}) = \begin{pmatrix} N_{k,1} & \ldots & N_{2,1} & N_{1,1} \\ N^*_{k,1} & \ldots & N^*_{2,1} & 0 \end{pmatrix}. \tag{8}$$

Thus, there exists an invertible matrix P_1 such that

$$P_1 M = \begin{pmatrix} N_{1,1} & 0 & & \ldots & 0 \\ N^*_{2,1} & & & & \\ M_{1,2} & & & & \\ \vdots & \vdots & & & \vdots \\ M_{1,c} & 0 & & \ldots & 0 \\ N_{2,1} & N_{1,1} & 0 & \ldots & 0 \\ N^*_{3,1} & N^*_{2,1} & & & \\ M_{2,2} & M_{1,2} & & & \\ \vdots & \vdots & \vdots & & \vdots \\ M_{2,c-1} & M_{1,c-1} & 0 & & 0 \\ \vdots & \vdots & & & \vdots \\ N_{c,1} & N_{c-1,1} & N_{c-2,1} & \ldots & N_{1,1} \\ 0 & 0 & \ldots & \ldots & 0 \end{pmatrix}, \tag{9}$$

where there are $r_1 - \nu$ rows of zeros at the bottom. Put

$$R_{i,1} = \begin{pmatrix} N_{i,1} \\ N^*_{i+1,1} \\ M_{i,2} \end{pmatrix}, \qquad i = 1, \ldots, c-1,$$

$$R_{i,j} = M_{i,j+1}, \quad i = 1, \ldots, c-1, \quad j = 2, \ldots, c-i.$$

Let $v_j =$ (number of rows of $R_{i,j}$). Then by (8) and the definition of M, we see that

$$\sum_{j=1}^{c-1} v_j = \sum_{j=1}^{c} r_j \le r. \tag{10}$$

Put

$$R = \begin{pmatrix} R_{1,1} & 0 & & \dots & 0 \\ R_{1,2} & 0 & & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ R_{1,c-1} & 0 & & \dots & 0 \\ R_{2,1} & R_{1,1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ R_{2,c-2} & R_{1,c-2} & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ R_{c-1,1} & R_{c-2,1} & R_{c-3,1} & \dots & R_{11} \end{pmatrix}.$$

Then

$$P_1 M = \begin{pmatrix} R & 0 \\ * & N_{1,1} \\ 0 & 0 \end{pmatrix}.$$

From (10) it follows that R is of the form (7a,b) with c replaced by $c-1$. By the induction hypothesis, there exists an invertible P_2 such that $P_2 R$ is $(c-1, r, n)$-good. Then

$$\begin{pmatrix} P_2 & 0 \\ 0 & I \end{pmatrix} P_1 M = \begin{pmatrix} P_2 R & 0 \\ * & N_{1,1} \\ 0 & 0 \end{pmatrix}.$$

This is clearly (c, r, n)-good, and we have found the desired P. ∎

4. Proof of the Theorem. By Lemma 1, we need only show that any equation of the form

$$a_1 x_1^k + \ldots + a_s x_s^k \equiv 0 \bmod \pi^e, \qquad a_i \in \mathfrak{O}, \tag{11}$$

has a non-trivial solution mod π^e, provided $s \geq k+1$.

For any $x \in \mathfrak{O}$ we have

$$x = x_0 + x_1 \pi + \ldots + x_{e-1} \pi^{e-1}, \qquad x_i \in \mathfrak{o}.$$

Put $c = [e/p^t]$. Then

$$x^{p^t} \equiv x_0^{p^t} + x_1^{p^t} \pi^{p^t} + \ldots + x_c^{p^t} \pi^{cp^t} \bmod \pi^e.$$

Write

$$a_i = \sum_{j=0}^{e-1} a_{i,j} \pi^j, \qquad x_i = \sum_{j=0}^{e-1} x_{i,j} \pi^j.$$

By the above comments, to solve (11) for $k = p^t$ it is sufficient to solve the system

$$(12)\quad \begin{aligned}
&\sum_{i=1}^{s} a_{i,0}x_{i,0}^{p^t} \equiv 0 \bmod p,\\
&\quad\vdots\\
&\sum_{i=1}^{s} a_{i,p^t-1}x_{i,0}^{p^t} \equiv 0 \bmod p,\\
&\sum_{i=1}^{s} a_{i,p^t}x_{i,0}^{p^t} + \sum_{i=1}^{s} a_{i,0}x_{i,1}^{p^t} \equiv 0 \bmod p,\\
&\quad\vdots\\
&\sum_{i=1}^{s} a_{i,2p^t-1}x_{i,0}^{p^t} + \sum_{i=1}^{s} a_{i,p^t-1}x_{i,1}^{p^t} \equiv 0 \bmod p,\\
&\quad\vdots\\
&\sum_{i=1}^{s} a_{i,(c+1)p^t-1}x_{i,0}^{p^t} + \sum_{i=1}^{s} a_{i,cp^t-1}x_{i,1}^{p^t} + \ldots + \sum_{i=1}^{s} a_{i,p^t-1}x_{i,c}^{p^t} \equiv 0 \bmod p,
\end{aligned}$$

over $\mathfrak{o}$. Here $a_{i,j} = 0$ if $j \geq e$.

LEMMA 4. *If $s \geq k+1$, then any system of the form* (12) *has a non-trivial solution such that*

(i) $x_{j,0} \not\equiv 0 \bmod p$ *for some* j.
(ii) $x_{j,0}$ *is an m-th power* $\bmod\, p$ *for all* j.

Proof. Since p is unramified in L, $L(p) = \mathfrak{o}/(p)$ is a finite field of characteristic p. Thus, $x \mapsto x^{p^t}$ is an automorphism of $L(p)$. Therefore, to solve (12) it suffices to solve the associated linear system (i.e. replace $x_{i,j}^{p^t}$ with $y_{i,j}$) over the field $L(p)$. We wish to find a solution such that $y_{i,0}$ is an mth power for $i = 1, \ldots, s$.

Observe that the matrix of coefficients of (12) is in the form of (7a,b), with c replaced by $c+1$, $r = p^t$, and $n = s$. By Lemma 3, (12) is equivalent via elementary row operations to a system whose coefficient matrix is $(c+1, p^t, s)$-good. Suppose this new matrix is given by

$$\begin{pmatrix} B_{11} & 0 & & \ldots & 0 \\ B_{21} & B_{22} & 0 & \ldots & 0 \\ * & * & * & * & * \end{pmatrix}, \qquad B_{ij} \in \mathbf{M}_{r_i,s}(L(p)),\ r_i \leq p^t.$$

By the Thereom of Chevalley–Warning [Se, I, Thm. 3], if $s > p^t m = k$, then the system $B_{11}\mathbf{Y}_1 = \mathbf{0}$ has a non-trivial solution over $L(p)$, say $\mathbf{Y}_1 = (y_1, \ldots, y_s)$, such that each y_i is an mth power. By Lemma 2 this can be extended to a solution $\mathbf{Y}$ of the linear system associated with (12). By the remarks in the first paragraph of this proof, $\mathbf{Y}$ corresponds to a solution of (12). ∎

The proof of the Theorem now follows upon combining Lemma 1 with the following lemma.

LEMMA 5. *For any k, an equation of the form* (11) *has a non-trivial solution* $\bmod \pi^e$ *provided $s \geq k+1$. Therefore, $\Phi(k,e) \leq k+1$.*

Proof. By the previous lemma and the comments preceding it, we can find $x_1, \ldots, x_s$, not all zero modulo π, such that

$$a_1 x_1^{p^t} + \ldots + a_s x_s^{p^t} \equiv 0 \bmod \pi^e,$$

and

$$x_i \equiv y_i^m \bmod \pi, \quad i = 1, \ldots, s.$$

Since $(m,p) = 1$, it follows from Hensel's Lemma that for each i we can find $z_i \in \mathfrak{O}$ such that $z_i^m \equiv x_i \bmod \pi^e$. Thus $\mathbf{z} = (z_1, \ldots, z_s)$ is the desired solution of (11). ∎

References

[B] B. J. Birch, *Diagonal equations over $\mathfrak{p}$-adic fields*, Acta Arith. 9 (1964), 291–300.

[D] M. M. Dodson, *Some estimates for diagonal equations over $\mathfrak{p}$-adic fields*, ibid. 40 (1981), 117–124.

[La] S. Lang, *Algebraic Number Theory*, Springer, 1986.

[S] W. Schmidt, *The solubility of certain p-adic equations*, J. Number Theory 19 (1984), 63–80.

[Se] J.-P. Serre, *A Course in Arithmetic*, Springer, 1973.

Department of Mathematics
Princeton University
Princeton, New Jersey 08544
U.S.A.
E-mail: cmcls@math.princeton.edu

Received on 9.6.1995
and in revised form on 2.10.1995 (2803)