On the solvability of $\mathfrak{p}$-adic diagonal equations

Antoni Rangachev

under the direction of
Mr. Mohsen Bahramgiri
Massachusetts Institute of Technology

Research Science Institute
July 31, 2007

Abstract

This paper deals with some fundamental questions in the study of the diagonal diophantine equation $a_1x_1^k+\cdots+a_sx_s^k=0$ over a finite extension K of the field $\mathbb{Q}_p$ of p-adic numbers, namely some new upper bounds on the number of variables that ensure their solvability.

1 Introduction

A basic problem in the study of diophantine equations is that of determining sufficient conditions for ensuring their solvability. Let p be a prime, let $\mathbb{Q}_p$ denote the field of p-adic numbers, and let K be a finite extension of $\mathbb{Q}_p$. One of the fundamental questions in the theory of diophantine equations is, when does a diagonal equation, i.e. an equation of the form

$$a_1x_1^k + \cdots + a_sx_s^k = 0, \tag{1}$$

where the coefficients a_i are in the ring $\mathfrak{O}_K$ of integers of K, 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 [5] that it suffices to have $s \geqslant k^2+1$. Generally, suppose $k = p^t m$, with $(m, p) = 1$. Let f be the residue class degree of K, and $d = (m, p^f - 1)$. The best known result in the case of arbitrary fields was established by Birch [2] who showed that for any K, it suffices to have $s \geqslant (2t + 3)^k(d^2k)^{k-1}$. In 1996, Skinner [1] proved that when $k = p^t$ it suffices to have $s \geqslant k((k + 1)^{\max(2t,1)} - 1) + 1$. The original formulation of Skinner's result is that the inequality above holds for every k. Unfortunately, it later turned out that there was an error in Skinner's proof. In an attempt to reconstruct this result we have obtained the following three results:

Theorem 1 If $s \geq k((nk + 1)^{\max(2t,1)} - 1) + 1$, then any equation of the form (1) has a non-trivial solution over K, where $n = [K/\mathbb{Q}_p]$.

This statement is in the same range of Skinner's claim. The result is sharp, when $K = \mathbb{Q}_p$ and $k = p - 1$

Theorem 2 If $s \geq k(p^{n \max(2t,1)}-1)+1$, then any equation of the form (1) has a non-trivial solution over K, where $n = [K/\mathbb{Q}_p]$.

This statement is an improvement of Skinner's general claim for k sufficiently large.

Theorem 3 If $s \geq k^3 + 1$, then any equation of the form (1), satisfying the additional

restrictions $(k,p)=1$ and $(a_i,p)=1$, has a non-trivial solution.

This statement gives us essential information for the case $(k,p)=1$, which was not treated by Skinner's method.

2 Notation and preliminaries

In order to describe our new results we need some notation, which we adopt from [1]. We denote by $\mathfrak{O}$ 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 of K, and $\mathfrak{o}$ is the ring of integers of L. We recall that $\{1,\pi,\ldots,\pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$. For more detailed information see Appendix A.

Let $\Gamma(k)$ be the least positive integer such that if $s\geq\Gamma(k)$, then any equation of the form (1) is solvable non-trivially over K. By $\Gamma_1(k)$ we denote the least positive integer such that any equation of the form (1) has a solution satisfying $a_i\not\equiv 0 \pmod{\pi}$ for all i.

We say that $\mathbf{x}$ is a "non-trivial solution modulo π^n" if $\mathbf{x}=(x_1,\ldots,x_s)\in\mathfrak{O}^s$ is a solution of (1) modulo π^n and if additionally $x_j\not\equiv 0 \pmod{\pi}$ for some j. By $\Phi(k,n)$ we 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 modulo π^n. Throughout the paper, we denote by N any of the integers $kn+1$, p^n or k^3+1.

2.1 The reduction lemma

The following lemma reduces the proof of our three main results to showing that $\Phi(k,e)\leq N$. Recall that e is the ramification index of K over $\mathbb{Q}_p$. Then

Lemma 2.1.1 (*Skinner* [1])

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

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

3. $\Phi(k,(v+1)e) \leq \Phi(k,e)\Phi(k,ve) \leq \Phi(k,e)^{v+1}$ for an arbitrary positive integer v.

4. If $\Phi(k,e) \leq N$, then $\Gamma(k) \leq k(N^{\max(2t,1)} - 1) + 1$.

Proof. 1. Using the fact that every element of K can be written in the form $x = u\pi^{v_p(x)}$, where u is a unit, we can write all the coefficients a_i in the form $a_i = \pi^{r_i k + c_i} b_i$, where $v_p(a_i) = r_i k + c_i$, with $r_i \geq 0, 0 \leq c_i < k$ and $(b_i, \pi) = 1$. If $s > k(c-1)$, then by the Pigeonhole Principle at least c of the c_i are the same. We may assume the corresponding indices to be $i = 1, \ldots, c$. Thus it suffices to find a non-trivial solutions to the equation

$$b_1 x_1^k + \cdots + b_c x_c^k = 0, \quad (b_i, \pi) = 1. \tag{2}$$

The existence of a solution is guaranteed as $c \geq \Gamma_1(k)$.

2. We may assume that $a_i \not\equiv 0 \pmod{\pi}$ 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) $\pmod{\pi^r}$. Let $\mathbf{x} = (x_1, \ldots, x_s)$ be such a solution. We may assume that $x_1 \not\equiv 0 \pmod{\pi}$. Choose $y_2, \ldots, y_s \in \mathfrak{o}$ such that $y_i \equiv x_i \pmod{\pi^r}$. Let $d = \sum_{i=2}^{s} a_i y_i^k$. Since $a_1 x_1^k + d \equiv 0 \pmod{\pi^r}$, it follows from Hensel's Lemma that we can find $y_1 \in \mathfrak{o}$ such that $y_1 \equiv x_1 \pmod{\pi^r}$ and $a_1 y_1^k + d = 0$. Thus $\mathbf{y} = (y_1, \ldots, y_s)$ is a non-trivial solution of (1).

3. Let $h = \Phi(k, ve)$ and $l = \Phi(k,e)$ and for $j = 0, \ldots, l-1$ write

$$F_j(\mathbf{x_j}) = a_{jh+1} x_{jh+1}^k + \cdots + a_{(j+1)h} x_{(j+1)h}^k,$$

where $\mathbf{x}_j = (x_{jh+1}, \ldots, x_{j+1}h)$. Then (1) becomes

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

Thus, it suffices to find a non-trivial solution to the congruence

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

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

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

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

$$f_0 t_0^k + \cdots + f_{l-1} t_{l-1}^k \equiv 0 \pmod{\pi^{(v+1)e}}, \tag{4}$$

where $f_j \equiv 0 \pmod{\pi^{ve}}$ for $0 \leq j < l$. 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 $\pi^{(v+1)e}$.

4. First we consider the case $\max(2et, 1) = 2et$. Substituting $r = 2t - 1$ in the inequality in 1., we obtain

$$\Gamma(k) \leq k(\Gamma_1(k) - 1) + 1 \leq k(\Phi(k, 2et) - 1) + 1 \leq k(N^{\max(2t,1)} - 1) + 1.$$

Now let $\max(2et, 1) = 1$. Since $\Phi(k, r)$ is an increasing function in r, we have $\Phi(k, 1) \leq \Phi(k, e) \leq N$, which proves the desired inequality.

2.2 Chevalley-Warning Theorem

In this section we discuss some classical results concerning the solvability of equations over finite fields.

Let q be a power of a prime number p, and let $\mathbb{F}_q$ be a field with q elements. Let also, $\mathbb{F}_q[x_1, \ldots, x_n]$ be the ring of the polynomials in n variables over $\mathbb{F}_q$.

In 1935 Artin conjectured that if $f(x) \in \mathbb{F}_q[x_1, \ldots, x_n]$ satisfying $f(0, \ldots, 0) = 0$ and $n > \deg f$, then f has at least one non-trivial zero in K^n. Here we shall prove the more general

Theorem 2.2.1 (*Chevalley-Warning*, cf.,eg., [8]) Let $f_i \in K[x_1, \ldots, x_n]$ be polynomials in n variables such that $\sum \deg f_i < n$, and let V be the set of their common zeros in $\mathbb{F}_q^n$. Then

$$\mathrm{Card}(V) \equiv 0 \pmod{p}$$

Proof. We use the following

Lemma 2.2.2 Let u be a non-negative integer. The sum $S(X^u) = \sum_{x \in \mathbb{F}_q} x^u$ is equal to -1 if $u \geq 1$ and u is divisible by $q-1$; it is equal to 0 otherwise.

Proof. If $u = 0$, all the terms of the sum are equal to 1; then $S(X^u) = q \cdot 1 = 0$ because $\mathbb{F}_q$ is of characteristic p. If $u \geq 1$ and u is divisible by $q-1$, we have $0^u = 0$ and $x^u = 1$ if $x \neq 0$. Hence $S(X^u) = (q-1) \cdot 1 = -1$.

Finally, if $u \geq 1$ and not divisible by $q-1$, the fact that $\mathbb{F}_q^*$ (The multiplicative group of non zero elements of $\mathbb{F}_q$ is cyclic of order $q-1$ shows that there exists $y \in \mathbb{F}_q^*$ such that $y^u \neq 1$. Then:

$$S(X^u) = \sum_{x \in \mathbb{F}_q^*} x^u = \sum_{x \in \mathbb{F}_q^*} y^u x^u = y^u S(X^u),$$

whence $(1 - y^u)S(X^u) = 0$, or $S(X^u) = 0$.

Now put $P = \prod_\alpha (1 - f_\alpha^{q-1})$ and let $x \in \mathbb{F}_q^n$. If $x \in V$, all the $f_\alpha(x)$ are zero and $P(x) = 1$; if $x \notin V$, one of the $f_\alpha(x)$ is nonzero and $f_\alpha^{q-1} = 1$, hence $P(x) = 0$. If, for every polynomial f, we put $S(f) = \sum_{x \in \mathbb{F}_q} f(x)$, we have

$$\mathrm{Card}(V) \equiv S(P) \pmod{p}$$

and we reduce the claim to showing that $S(P) = 0$.Now the assumption $\deg f_\alpha < n$ implies that $\deg P < n(q-1)$; thus P is a linear combination of monomials $X^u = X_1^{u_1} \cdots X_n^{u_n}$ with

$\sum u_i < n(q-1)$. It suffices to prove that $S(X^u) = 0$ for such monomial X^u, and this follows from the lemma since at least one u_i is smaller than $q-1$.

Corollary 2.2.3 If $\sum \deg f_\alpha < N$ and if the f_α have no constant term, then the f_α have a common non-trivial zero.

2.3 Algebraic number theory lemmas

Before we are able to prove our results, we need some standart algebraic number theory properties of the ring $\mathfrak{O}$ and its maximal ideal $\mathfrak{p}$.

Lemma 2.3.1 There exists an isomorphism such that $\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong \mathfrak{O}/\mathfrak{p}$.

Proof. Let $\mathfrak{a} \neq 0$ be an arbitrary ideal of $\mathfrak{O}$ and $x \neq 0$ an element in $\mathfrak{a}$ with the smallest possible value $v(x) = n$. Then one can write $x = u\pi^n$, where u is a unit. The last equality gives the inclusion $\pi^n\mathfrak{O} \subseteqq \mathfrak{a}$. Now let y be an element of $\mathfrak{O}$. We can write $y = \varepsilon\pi^m$, where ε is a unit. By our assumption $m = v(y) \geq n$, hence $y = (\varepsilon\pi^{m-n})\pi^n \in \pi^n\mathfrak{O}$, so that $\mathfrak{a} = \pi^n\mathfrak{O}$. The isomorphism results from the correspondence $a\pi^n \mapsto a \pmod{p}$.

Lemma 2.3.2 Denoting by $|\mathfrak{O}/(\pi^e)|$ the number of elements of the quotient $\mathfrak{O}/(\pi^e)$, we have

$$|\mathfrak{O}/(\pi^e)| = p^{ef} = p^n.$$

Proof. This follows upon combining Lemma 2.3.1 with the fact that $|\mathfrak{O}/(\pi)| = p^f$ and $|\mathfrak{O}/(\pi^e)| = |\mathfrak{O}/(\pi)| \dots |(\pi)^{e-1}/(\pi)^e|$.

Lemma 2.3.3 If $x \in \mathfrak{O}$ and $x \equiv 0 \pmod{\pi}$, then $x \equiv 0 \pmod{\pi^e}$.

Proof. The congruence $x \equiv 0 \pmod{\pi}$ is equivalent to $x = \pi x_1$, where x_1 is an element of $\mathfrak{O} \setminus \mathfrak{o}$. Then we can write

$$x_1 = x_{1,0} + x_{1,1}\pi + \cdots + x_{1,e-1}\pi^{e-1},$$

where $x_{1,i} \in \mathfrak{o}$. Substituting x_1 in $x = \pi x_1$ we obtain the identity

$$px_{1,e-1} - x + x_{1,0}\pi + \cdots + x_{1,e-2}\pi^{e-1} = 0,$$

which implies $px_{1,e-1} = x$, i.e. $x \equiv 0 \pmod{\pi^e}$.

3 Proof of Theorem 1

By the Reduction Lemma 2.1.1 we only have to show that when $s \geq nk+1$, any congruence of the form

$$a_1 x_1^k + \cdots + a_s x_s^k \equiv 0 \pmod{\pi^e}, \quad a_i \in \mathfrak{O}, \tag{5}$$

has a non-trivial solution modulo π^e. Since $\{1, \ldots, \pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$ we can write

$$a_i = a_{i,0} + a_{i,1}\pi + \cdots + a_{i,e-1}\pi^{e-1}.$$

Now let $k = p^t$. We look for solutions x_i only from the ring $\mathfrak{o}$. Solving (3) is equivalent to solving the system

$$\begin{aligned} \sum_i^s a_{i,0}(x_1)^{p^t} &\equiv 0 \pmod{p} \\ &\vdots \\ \sum_i^s a_{i,e-1}(x_s)^{p^t} &\equiv 0 \pmod{p} \end{aligned}$$

over $\mathfrak{o}$. Since the system consists only congruences modulo p, it suffices to solve it over the field $L(p) = \mathfrak{o}/(p)$, which is a field of characteristic p. Because the correspondence $x \mapsto x^{p^t}$ is an automorphism in $L(p)$, system (6) is equivalent to

$$\begin{pmatrix} a_{1,0} & a_{2,0} & \dots & a_{s,0} \\ \vdots & \ddots & \vdots & \\ a_{1,e-1} & a_{2,e-1} & \dots & a_{s,e-1} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} \equiv 0 \pmod{p}, \tag{6}$$

where $x_i^{p^t}$ is replaced with y_i. Now we wish to find solutions such that y_i is an m-th power for $i = 1, \dots, s$. By the Chevalley-Warning theorem, if $s > em$ then (6) has a non-trivial solution over $L(p)$, say $(y_1, \dots, y_s) = (z_1^m, \dots, z_s^m)$. To summarize, we found $x_1, \dots, x_s$ not all zero mod π, such that

$$a_1 x_1^{p^t} + \dots + a_s x_s^{p^t} \equiv 0 \pmod{\pi^e}$$

and

$$x_i \equiv z_i^m \pmod{\pi}.$$

We can divide the set of all x_i into two sets. Let these two sets be A, the set of all solutions x_i that are coprime with π, and B, the set of all solutions x_j that are divisible by π. Without loss generality we may assume that $A = \{x_1, \dots, x_r\}$ and $B = \{x_{r+1}, \dots, x_s\}$, where $r \in \mathbb{N}$.

Now consider the set A and the polynomial $f(X) = X^m - x_i$, where $i = 1, \dots, r$. For every element x_i of this set we have

$$x_i \equiv 0 \pmod{\pi}$$

Since $(m, p) = 1$ and $(z_i, \pi) = 1$, we have $f'(z_i) \not\equiv 0 \pmod{\pi}$. Now it follows from Hensel's lemma that there exists a solution $\hat{\mathbf{z}} = (\hat{z_1}, \dots, \hat{z_r})$ of the equation $f(X) = 0$, which as a consequence implies the congruence $\hat{z_i}^m \equiv 0 \pmod{\pi^e}$ for $i = 1, \dots, r$.

By definition, all elements of the set B are divisible by π. By virtue of Lemma 2.3.3,

we have the congruence $x_i \equiv 0 \pmod{\pi^e}$ for $i = r+1, \ldots, s$. The last is equivalent to $x_i \equiv x_i^m \pmod{\pi^e}$. Finally, the solution of (5) is given by $\mathbf{x} = (\hat{z_1}, \ldots, \hat{z_r}, x_{r+1} \ldots, x_s)$. The inequalities $e \leq n$ and $m \leq k$ imply the desired one, namely $s \geq k((nk+1)^{\max(2t,1)} - 1) + 1$.

4 Proof of Theorem 2

By the same considerations as in the previous section it suffices to show that when $s \geq p^n$, the congruence (5) has a non-trivial solution modulo π^e.

Consider the sequence $0, a_1, a_1 + a_2, \ldots, a_1 + a_2 + \cdots + a_s$. We look for solutions x_i in the set $\{0, 1\}$. By virtue of $s \geq p^n$ and Lemma 2.3.2, at least two terms of the sequence above are congruent modulo π^e. Let their difference be $a_u + \cdots + a_v$. Then the solution of the congruence is given by $(x_u, x_2, \ldots, x_v) = (1, 1, \ldots, 1)$ and $x_i = 0$ for $i < n$ or $i > v$.

5 Proof of Theorem 3

Using the Reduction Lemma again, we have to show that when $s \geq mk+1$ the congruence (5) has a non-trivial solution mod π^e, with the additional restrictions $(a_i, p) = 1$ and $(k, m) = 1$. We look for solutions x_i of the form

$$x_i = x_{i,0} + x_{i,m}\pi + \ldots + x_{i,cm}\pi^{cm}, \text{ with } x_{i,j} \in \mathfrak{o} \text{ and } c = \lfloor \tfrac{e}{m} \rfloor.$$

Put $l = \lfloor \frac{e}{p^t m} \rfloor = \lfloor \frac{e}{k} \rfloor$. Then

$$(x_i)^{p^t} \equiv (x_{i,0})^{p^t} + (x_{i,m}\pi^m)^{p^t} + \ldots + (x_{i,lm}\pi^{lm})^{p^t} \pmod{\pi^e}$$

Now let $k = p^t$. Again, a necessary and sufficient condition to solve (5) is to solve the system.

$$\begin{aligned}
\sum_{i=1}^{s} a_{i,0}(x_{i,0})^{p^t} &\equiv 0 \pmod{p} \\
\sum_{i=1}^{s} a_{i,1}(x_{i,0})^{p^t} &\equiv 0 \pmod{p} \\
&\vdots \\
\sum_{i=1}^{s} a_{i,mp^t-1}(x_{i,0})^{p^t} &\equiv 0 \pmod{p} \\
\sum_{i=1}^{s} a_{i,mp^t}(x_{i,0})^{p^t} + \sum_{i=1}^{s} a_{i,0}(x_{i,1})^{p^t} &\equiv 0 \pmod{p} \\
&\vdots \\
\sum_{i=1}^{s} a_{i,2mp^t-1}(x_{i,0}^{p^t}) + \sum_{i=1}^{s} a_{i,mp^t-1}(x_{i,1})^{p^t} &\equiv 0 \pmod{p} \\
&\vdots \\
\sum_{i=1}^{s} a_{i,(l+1)mp^t-1}(x_{i,0})^{p^t} + \sum_{i=1}^{s} a_{i,(l-1)mp^t-1}(x_{i,1})^{p^t} + \cdots + \sum_{i=1}^{s} a_{i,mp^t-1}(x_{i,lm})^{p^t} &\equiv 0 \pmod{p}
\end{aligned}$$

By the same arguments as in the proof of Theorem 1, it suffices to solve the system over $L(p)$, where again $(x_{i,j})^{pt}$ are replaced with the elements $y_{i,j}$ from the ring $\mathfrak{o}$, with the additional condition that each of them should be an m-th power. By the Chevalley-Warning theorem, if $(l+1)s > (l+1)m^2p^t$, then the system above has a non-trivial solution over $L(p)$, say $(y_{1,0}, \ldots, y_{s,lm}) = (z_{1,0}^m, \ldots, z_{s,lm}^m)$. So we have found $(x_1, \ldots, x_s)$, not all zero modulo π^e, such that

$$a_1 x_1^{p^t} + \cdots + a_s x_s^{p^t} \equiv 0 \pmod{\pi^e}$$

and

$$x_i \equiv x_{i,v_i m} \pi^{v_i m}, \tag{7}$$

where x_{i,v_i} is the first coefficient which is not divisible by π and each is an m-th power modulo π. Therefore, the congruence (7) is equivalent to

$$x_i/\pi^{v_i m} \equiv y_{i,v_i}^m \pmod{p}.$$

Consider the polynomial $f(x) = X^m - a$, where we write a for $x_i/\pi^{v_i m}$. Since $(y_{i,v_i}, \pi) = 1$, then by Hensel's lemma there exists a solution to the equation $f(X) = 0$, which implies the congruence

$$x_i/\pi^{v_i m} \equiv \hat{z_i}^m \pmod{\pi^e}.$$

Since $v_i m \leq e$ (Not all solutions are divisible by π^e), we have

$$x_i \equiv (z_i \pi^{v_i})^m \pmod{\pi^e}.$$

Finally, when $s \geq k^3 + 1$ there exists a solution $\mathbf{x} = (x_1, \ldots, x_s)$ satisfying the congruence (5). We now distinguish two general cases: either at least one of the x_i is not divisible by π, or else each x_i is divisible by π. The first case guarantees that we have obtained a non-trivial solution modulo π. Thus the desired solution is given by $\mathbf{x} = ((z_1\pi^{v_1})^m, \ldots, (z_s\pi^{v_s})^m)$
The second case, i.e., each x_i is divisible by π, is of special interest. Here, we can not claim that the obtained solution is the desired one, because by the definition of $\Phi(k, e)$ at least one of the x_i must be divisible by π. What we do here is the same as what we did in the proof of the Reduction Lemma. By virue of the fact that at least one x_i is not divisible by π^e, after a reduction modulo π^r, where r is the greatest integer such that π^r divides $a_1x_1^k + \cdots + a_sx_s^k$, we obtain a congruence of the form

$$a_j(\hat{x_j})^k + d \equiv 0 \pmod{\pi^{e-r}},$$

where $r = \operatorname{ord}_\pi \hat{x_j}$, and $d = a_1(\hat{x_1}^k) + \cdots + a_{j-1}(\hat{x}_{j-1})^k + a_{j+1}(\hat{x}_{j+1})^k + \cdots + a_s(\hat{x}_s)^k$. Since $(a_j, \pi) = (k, \pi) = 1$, then there exists a solution X to the equation $f(X) = a_j(\hat{x_j})^k + d$. Finally, the solution of (5) is given by

$$\mathbf{x} = (\hat{x_1}, \ldots, X, \ldots, \hat{x_s}).$$

6 Conclusion

When studying Skinner's work [1], we managed to develop a refinement of his method, obtaining some new upper bounds for the number of variables of a diagonal form over an extension of $\mathbb{Q}_p$. Our further goal will be to prove the general case of his claim. Let us note that the results we have obtained could be generalized to systems of diagonal equations.

7 Acknowledgments

I would like to to express my deep gratitude to my mentor Mr. Mohsen Bahramgiri of the Massachusetts Institute of Technology for his ongoing support and inspiring discussions. I am grateful to Prof. Hartley Rogers for providing me with such a great mentor. I am extremely thankful to Prof. Christopher Skinner for posing the problem and encouraging me to work in this field. Many thanks are due to my tutor Chris Mihelich for his constant support and professional editing of this paper. I greatly appreciate the support of the CEE, especially of the RSI staff and its director Mathew Paschke. Finally, I would like to acknowledge with thanks the help of my Bulgarian friends and colleagues Prof. Ivan Chipchakov (for enriching my knowledge about the history of the problem), Kaloyan Slavov (for sharing with me his fascination for RSI), Vesselin Dimitrov(for acting as virtual nobody and sending number of valuable comments till the deadline of the submission), and Dr. Jenny Sendova (for helping

me with some ping-pong lessons when I was stuck with the proofs).

A Introduction to p-adic numbers

In this section we provide some preliminaries from the theory of p-adic numbers following [3] and [4].

A.1 Absolute Values on a Field

Definition A.1.1 Let K be a field. An **absolute value** on K is a real-valued function $|\ |: K \longrightarrow \mathbb{R}_+$ on K satisfying the following three properties:

1. We have $|x|_v \geq 0$ for all $x \in K$, and $|x| = 0$ if and only if $x = 0$.

2. For all $x, y \in K$, we have $|xy| = |x||y|$.

3. For all $x, y \in K$, we have $|x + y| \leq |x| + |y|$.

We say an absolute value on K is non-archimedian if it satisfies the additional condition:

4. $|x + y| \leq max(|x|, |y|)$;otherwise, we will say that the absolute value is archimedian.

The absolute value which is such that $|x| = 1$ for all $x \neq 0$ is called **trivial**.

Definition A.1.2 Fix a prime number $p \in \mathbb{Z}$. The p-adic valuation on $\mathbb{Z}$ is the function

$$v_p : \mathbb{Z} - \{0\} \longrightarrow \mathbb{R}$$

defined as follows: for each integer $n \in \mathbb{Z}$, $n \neq 0$, let $v_p(n)$ be the unique positive integer satisfying

$$n = p^{v_p(n)}n' \qquad \text{with } (p, n') = 1.$$

We extend $v_p(n)$ to the field of rational numbers as follows: if $x = a/b \in \mathbb{Q}$, then

$$v_p(x) = v_p(a) - v_p(b).$$

Lemma A.1.3 For all x and $y \in \mathbb{Q}$, we have

i) $v_p(xy) = v_p(x) + v_p(y)$

ii) $v_p(x + y) \geq \min(v_p(x), v_p(y))$,

with the obvious conventions with respect to $v_p(0) = +\infty$.

Definition A.1.4 For any $x \in \mathbb{Q}$, we define the p-adic absolute value of x by

$$|x|_p = p^{-v_p(x)}$$

if $x \neq 0$, and we set $|0|_p = 0$.

Corollary A.1.4 The function $|\ |_p$ is non-archimedian absolute value on $\mathbb{Q}$.

A.2 Basic Properties

An absolute value of K defines a metric. The distance between two elements x, y of K in this metric is $|x - y|$. Thus an absolute value defines a topology on K. Two absolute values are called **equivalent** if they define the same topology. If they do not, they are called not equivalent. We observe that $|1| = |1^2| = |(-1)^2| = |1|^2$ whence $|1| = |-1| = 1$. Also, $|-x| = |x|$ for all $x \in K$, and $|x^{-1}| = |x|^{-1}$ for all $x \neq 0$.

Theorem A.2.1 Let $|\ |_1$ and $|\ |_2$ be non-trivial absolute values on a field K. They are equivalent if and only if the relation

$$|x|_1 < 1$$

implies $|x|_2 < 1$. If they are equivalent, then there exists a number $\lambda > 0$ such that $|x|_1 = |x|_2^{\lambda}$ for all $x \in K$.

Now we come to the main theorem in this section. It says that we have already found all the absolute values on $\mathbb{Q}$.

Theorem A.2.2 (Ostrowski) Every non-trivial absolute value on $\mathbb{Q}$ is equivalent to one of the absolute values $|\ |_p$, where either p is a prime number or $p = \infty$.

A.3 Completions

Definition A.3.1 Let K be a field and let $|\ |$ be an absolute value on K.

i) A sequence of elements $x_n \in K$ is called a Cauchy sequence if for every $\varepsilon > 0$ one can find a bound M such that we have $|x_n - x_m| < \varepsilon$ wnenever $m, n \geq M$.

ii) The field K is called complete with respect to $|\ |$ if every Cauchy sequence of elements of K has a limit.

Lemma A.3.2 A sequence (x_n) of rational numbers is a Cauchy sequence with respect to a non-archimedian absolute value $|\ |$ if and only if we have

$$\lim_{n\to\infty} |x_{n+1} - x_n| = 0.$$

Lemma A.3.3 The field $\mathbb{Q}$ is not complete with respect to any of its nontrivial absolute values.

Since $\mathbb{Q}$ is not complete, we need to construct completion. As we shall see the completion is the field obtained by adding to $\mathbb{Q}$ the limits of all Cauchy sequences.

Definition A.3.4 Let $|\ | = |\ |_p$ be a non-archimedian absolute value on $\mathbb{Q}$. We denote by $\mathfrak{C}$, or $\mathfrak{C}_{\mathfrak{p}}(\mathbb{Q})$ if we want to emphasize p and $\mathbb{Q}$, the set of all Cauchy sequences of elements

of $\mathbb{Q}$:

$$\mathfrak{C} = \mathfrak{C}_p(\mathbb{Q}) = \{(x_n) : (x_n) \text{ is a Cauchy sequence with respect to } ||_p\}.$$

Theorem A.3.5 Defining

$$(x_n) + (y_n) = (x_n + y_n)$$

$$(x_n).(y_n) = (x_n y_n)$$

makes C a commutative ring with unity.

Definition A.3.6 We define $\mathfrak{N} \subset \mathfrak{C}$ to be the ideal

$$\mathfrak{N} = \{(x_n) : \lim_{n\to\infty} |x_n|_p = 0$$

Lemma A.3.7 $\mathfrak{N}$ is a maximal ideal of $\mathfrak{C}$.

We want to identify sequences that differ by elements of $\mathfrak{N}$, on the grounds that they ought to have the same limit. This is done in the standard way, by taking the quotient of the ring $\mathfrak{C}$ by the ideal $\mathfrak{N}$.

Definition A.3.8 We define the field of p-adic numbers to be the quotient of the ring $\mathfrak{C}$ by its maximal ideal $\mathfrak{N}$:

$$\mathbb{Q}_p = \mathfrak{C}/\mathfrak{N}.$$

A.4 Exploring $\mathbb{Q}_p$

The aim of this section is to explore the field $\mathbb{Q}_p$ which we have just constructed.

Lemma A.4.1 For each $x \in \mathbb{Q}_p \;\; x \neq 0$, there exists an integer $n \in \mathbb{Z}$ such that $|x|_p = p^{-n}$.

Another way of saying this is in terms of the p-adic valuation v_p. What the lemma says is:

Lemma A.4.2 For each $x \in \mathbb{Q}_p \;\; x \neq 0$, there exists an integer $v_p(x)$ such that $|x|_p = p^{-1v_p(x)}$.

In other words, the p-adic valuation v_p extends to $\mathbb{Q}_p$.

Now we begin to explore the structure of $\mathbb{Q}_p$.

Definition A.4.3 The ring of p-adic integers is

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}.$$

The ring $\mathbb{Z}_p$ is called the **valuation ring** of $|\ |$.

Rings that contain a unique maximal ideal whose complement consists of invertible elements are called **local rings**.

Theorem A.4.4 The ring $\mathbb{Z}_p$ of p-adic integers is a local ring whose maximal ideal is the principal ideal $p\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p < 1\}$.

The p-adic units are invertible elements of $\mathbb{Z}_p$. We will denote the set of all such elements by $\mathbb{Z}_p^\times$. Since $x \in \mathbb{Z}_p$ means $|x| \leq 1$ and $x^{-1} \in \mathbb{Z}_p$ means $|-1| = |x|^{-1} \leq 1$, we see that

$$\mathbb{Z}_p^\times = \{x \in \mathbb{Q}_p : |x| = 1\}.$$

It is also easy to see that

$$\mathbb{Z}_p^\times \cap \mathbb{Q} = \{\tfrac{a}{b} \in \mathbb{Q} : (p, ab) = 1\}.$$

As in every ring, the p-adic units form a group.

A.5 Hensel's Lemma

The theorem known as "Hensel's Lemma" is probably the most important algebraic property of the p-adic numbers (and of other fields like $\mathbb{Q}_p$, which are complete with respect to a non-arcimedian valuation). Basically, it says that one can decide quite easily whether a polynomial has roots in $\mathbb{Z}_p$. The rest involves finding an "approximate" root of the polynomial, and then verifying a condition on the derivative of the polynomial.

Theorem A.5.1 (Hensel's Lemma) Let $F(x) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ be a

polynomial whose coefficients are in $\mathbb{Z}_p$. Suppose that there exists a p-adic integer $\alpha_1 \in \mathbb{Z}_p$ such that

$$F(\alpha_1) \equiv 0 \pmod{p^{2r-1}\mathbb{Z}_p}$$

and

$$F'(\alpha_1) \not\equiv 0 \pmod{p^r\mathbb{Z}_p},$$

where $F'(x)$ is the (formal) derivative of $F(x)$. Then there exists a p-adic integer $\alpha \in \mathbb{Z}_p$ such that $\alpha \equiv \alpha_1 \pmod{p^r\mathbb{Z}p}$ and $F(\alpha) = 0$.

A.6 Properties of Finite extensions

This section deals with introducing some notions and basic facts about finite extensions of $\mathbb{Q}_p$. On one level, what we want to say is that the structure we have found in $\mathbb{Q}_p$ extends without effort. Our main interest, however, is to see what information this gives us about finite extensions of $\mathbb{Q}_p$. In all of this section, K will be a finite extension of degree n of $\mathbb{Q}_p$, and we will write $|\ \ | = |\ \ |_p$ for p-adic absolute value (extended to K). We define the p-adic absolute value on K by the formula

$$|x| = \sqrt[n]{|\mathbb{N}_{K/\mathbb{Q}_p}(x)|_p},$$

where $\mathbb{N}_{K/\mathbb{Q}_p}(x)$ is the norm of the element x.

Definition A.6.1 Let K be a finite extension of $\mathbb{Q}_p$, and let $|\ \ |$ be the p-adic absolute value on K. For any $x \in K$, $x \neq 0$, we define the p-adic valuation $v_p(x)$ to be the unique rational number satisfying

$$|x| = p^{-v_p(x)}.$$

We extend the definition formally by setting $v_p(0) = +\infty$. It is useful to notice that since we know exactly how to compute the p-adic absolute value of an element of K, we also know

how to compute v_p. Here is the formula: for any $x \in K^\times$,

$$v_p(x) = \frac{1}{n} v_p(\mathbb{N}_{K/\mathbb{Q}_p}(x)).$$

This reduces computing v_p to computing norms.

Theorem A.6.2 The p-adic valuation v_p is a homomorphism from the multiplicative group $K^\times$ to the additive group $\mathbb{Q}$. Its image is of the form $\frac{1}{e}\mathbb{Z}$, where e is a divisor of $n = [K : \mathbb{Q}_p]$.

Definition A.6.3 Let $K/\mathbb{Q}_p$ be a finite extension, and let $e = e(K/\mathbb{Q}_p)$ be unique positive integer (dividing $n = [K : \mathbb{Q}_p]$) defined by

$$v_p(K^\times) = \frac{1}{e}\mathbb{Z}.$$

Wer call e the ramification index of K over $\mathbb{Q}_p$. We say the extension $K/\mathbb{Q}_p$ is unramified if $e = 1$. We say the extension $K/\mathbb{Q}_p$ is ramified if $e > 1$, and totally ramified if $e = n$. Finally, we write $f = f(K/\mathbb{Q}_p) = n/e$.

Definition A.6.4 Let $K/\mathbb{Q}_p$ be a finite extension, and let $e = e(K/\mathbb{Q}_p)$. We say an element $\pi \in K$ is a **uniformizer** if $v_p(\pi) = 1/e$.

We should remark that in the unramified case, we have $e = 1$, and we can (and usually will) simply take $\pi = p$.

Now we can describe the algebraic structure of K. First of all, recall that we defined the valuation ring

$$\mathfrak{O} = \mathfrak{O}_K = \{\in K : |x| \leq 1\} = \{x \in K : v_p(x) \geq 0\}$$

and its maximal ideal

$$\mathfrak{p} = \mathfrak{p}_k = \{x \in K : |x| \leq 1\} = \{x \in K : V_p(x) > 0\}.$$

As we saw above $\mathfrak{O}$ is a local ring, and the residue field is the quotient

$$\mathbf{k} = \mathfrak{O}_K/\mathfrak{p}_k$$

Theorem A.6.5 Let notations be as above, and fix a uniformizer π in K. Then

1. The ideal $\mathfrak{p}_K \subset \mathfrak{O}_k$ is principal, and π is a generator.

2. The residue field $\mathbf{k}$ is a finite extension of the field with p-elements $\mathbb{F}_p$ whose degree is f, i.e., $f = [\mathbf{k} : \mathbf{F}_p]$, so that $\mathbf{k} = \mathbb{F}_{p^f}$ is the finite field with p^f elements.

3. All the nonzero ideals of $\mathfrak{O}$ are given by $\mathfrak{p}^n = \pi^n\mathfrak{O} = \{x \in K : v(x) \geq n\}$, where $n \in \mathbb{N}$.

Definition A.6.6 Let $K/\mathbb{Q}_p$ be a finite extension. Then, the composite of all unramified subextensions is called the **maximal unramified subfield** of $K/\mathbb{Q}_p$.

Theorem A.6.7 The set $\{1, \pi, \ldots, \pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$.

At the end we recall the end we recall the Hensel's lemma for extensions of the of p-adic numbers.

Theorem A.6.8 (Hensel's Lemma) Let K be a finite extension of $\mathbb{Q}_p$, and let π be a uniformizer. Let $F(X) = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$ be a polynomial whose coefficients are in $\mathfrak{O}_K$. Suppose that there exists an integer $\alpha_1 \in \mathfrak{O}_K$ such that

$$F(\alpha_1) \equiv 0 \pmod{\mathfrak{p}_K^{2r-1}}$$

and

$$F'(\alpha_1) \not\equiv 0 \pmod{\mathfrak{p}_K^{r}},$$

where $F'(X)$ is the (formal) derivative of $F(X)$. Then there exists an integer $\alpha \in \mathfrak{O}_K$ such that $\alpha \equiv \alpha_1 \pmod{\mathfrak{p}_K^r}$ and $F(\alpha) = 0$.

References

[1] C.M. Skinner. Solvability of $\mathfrak{p}$-adic diagonal equations. Acta Arithmetica 75 (1996), 251–258.

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

[3] F.Q. Gouvea. *P-adic Numbers*. Springer-Verlag, Berlin (1997).

[4] S. Lang. *Algebra*. Addison-Wesley (1997).

[5] H. Davenport, D.J. Lewis. Homogeneous additive equations. Proc. Royal Soc. London Ser. 274 (1964), 443–460.

[6] S. Lang. *Algebraic Number Theory*. Springer-Verlag, New York (1994).

[7] T.D. Wooley. On the local solubility of diophantine systems. Compositio Mathematica 111 (1998), 149–165.

[8] D.B. Leep and W.M. Schmidt. Systems of homogeneous equations. Inventiones Mathematicae 71 (1983), 539–549.

[9] J.P. Serre. *A Course in Arithmetic*. Springer, New York (1973).

[10] J.W. Cassels and A. Frohlich. *Algebraic Number Theory*. Academic Press, London (1967). c

[11] J. Neukirch. *Algebraic Number Theory*. Springer, New York (1999).

[12] C. Chevalley. Démonstration d'une hypothèse de M. Artin. Abhand. Math. Sem. Hamburg (1936).

[13] L.J. Mordell. *Diophantine Equations*. Academic Press, New York (1969).