P-ADIC NUMBERS AND DIOPHANTINE EQUATIONS

By

KAY WILLIAM DUNDAS

Bachelor of Science
Fort Hays Kansas State College
Hays, Kansas
1960

Master of Arts
Fort Hays Kansas State College
Hays, Kansas
1961

Submitted to the Faculty of the Graduate College
of Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF EDUCATION
July, 1972

P-ADIC NUMBERS AND DIOPHANTINE EQUATIONS

Thesis Approved:

Jeanne Agnew
Thesis Adviser

John Jewett

Robert T. Alciatore

Vernon Troxel

D Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to acknowledge the following people for the roles they played in the preparation of this thesis. To my thesis adviser, Dr. Jeanne Agnew, goes a special thank you for her inspirational teaching as well as her understanding guidance. I would also like to thank Dr. John Jewett, Dr. Vernon Troxel, and Dr. Robert Alciatore for serving on my advisory committee. For particular contributions to my mathematics training, I thank Wilmont Toalson, Dr. Hiroshi Uehara, Dr. Forrest Whitfield, and Dr. John Jobe.

The thesis was typed by Velda Davis. I appreciate the excellent work she has done, especially with the difficult notation.

Finally, I thank my family. The constant encouragement provided by my parents, Mr. and Mrs. W. L. Dundas, and my wife's parents, Mr. and Mrs. Ralph Peterson, was a great help. The most important factor to the success of this endeavor was the cooperation and support of my wife, Mildred, and sons, Barry and Rodger.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	POWERS OF P-ADIC NUMBERS	5
III.	WARING'S PROBLEM	18
IV.	ARTIN'S CONJECTURE	28
SELECTED BIBLIOGRAPHY		72

CHAPTER I

INTRODUCTION

The field of number theory is known for having problems that are easy to state but difficult to solve. Problems that have traditionally been referred to as diophantine problems are good examples of this phenomenon. The essential ingredient of a diophantine problem is proving the existence of integral solutions of a set of equations or inequalities. A beginning number theory student can understand the statement of such problems and soon discovers why they have intrigued mathematicians for centuries. As the student acquires the basic techniques in number theory, he discovers that a primary factor in solving diophantine problems is his own ingenuity and the ingenuity of those who have preceded him. Methods that have been devised, while frequently elementary in nature, display a creativity that appears to be unending.

Attempts to devise methods for solving certain types of diophantine problems contributed to the development of the field known as the p-adic numbers. The basic idea behind these methods is the following. If $f(x_1,x_2, \dots ,x_s) = 0$ is to have a solution in integers, the congruence $f(x_1,x_2, \dots ,x_s) \equiv 0 \bmod p^{n+1}$ must have a solution in integers for every $n \geq 0$. It would be very convenient if the converse of this statement were true. However, as the following example demonstrates, this is not the case when $p = 3$. Similar examples can be cited for any p.

Example 1.1 The equation $x^2 = 7$ obviously has no solution in integers. However, consider the congruence $x^2 - 7 \equiv 0 \bmod 3^{n+1}$ where $n \geq 0$. The following congruences can easily be verified.
$2^2 - 7 \equiv 0 \bmod 3$, $5^2 - 7 \equiv 0 \bmod 3^2$, $14^2 - 7 \equiv 0 \bmod 3^3$, $68^2 - 7 \equiv 0 \bmod 3^4$, $68^2 - 7 \equiv 0 \bmod 3^5$. The values $\{2,5,14,68,68\}$ can be considered as the first five elements of a sequence $\{A_n\}$ where $A_n^2 - 7 \equiv 0 \bmod 3^{n+1}$. To show that the remainder of the sequence can be constructed, suppose $A_i^2 - 7 \equiv 0 \bmod 3^{i+1}$ for some $i \geq 0$. Then $A_i^2 - 7 = 3^{i+1}t$ for some integer t. In order to construct A_{i+1}, consider the following:

$$\begin{aligned}(A_i + 3^{i+1}x)^2 - 7 &= A_i^2 + 2A_i 3^{i+1}x + 3^{2i+2}x^2 - 7 \\ &= 3^{i+1}t + 2A_i 3^{i+1}x + 3^{2i+2}x^2 \\ &= 3^{i+1}(t + 2A_i x) + 3^{2i+2}x^2 .\end{aligned}$$

Since $(2A_i,3) = 1$, there exists an integer a_{i+1} such that $t + 2A_i a_{i+1} \equiv 0 \bmod 3$. This implies that

$$(A_i + 3^{i+1}a_{i+1})^2 - 7 \equiv 0 \bmod 3^{i+2} .$$

Therefore, A_{i+1} can be defined as $A_i + 3^{i+1}a_{i+1}$.

From this example, it is apparent that a sequence $\{A_n\}$ can be constructed where $A_n = \sum_{i=0}^{n} 3^i a_i$ and $A_n^2 - 7 = 0 \bmod 3^{n+1}$ for every $n \geq 0$. Each a_i is obtained by solving a congruence mod 3 so the condition $0 \leq a_i \leq 2$ can be imposed. In the example, the values of the first five a_i are $a_0 = a_3 = 2$, $a_1 = a_2 = 1$, and $a_4 = 0$. Note that a solution to the congruence $x^2 \equiv 7 \bmod 3$ is essential to the construction. That is, 7 must be a quadratic residue mod 3.

The following definitions are a more formal presentation of the ideas suggested by the example. No further attempt is made to justify the definitions. For a complete development of the p-adic numbers, see Agnew (1).

Definition 1.1 Let p be a prime and let A_i and a_i represent non-negative integers:

(1) A sequence $\{A_n\}$ is a p-adic sequence if $A_n \equiv A_{n-1} \bmod p^n$ for every $n \geq 1$.

(2) A p-adic sequence is in canonical form if $A_n = \sum_{i=0}^{n} a_i p^i$ where $0 \leq a_i \leq p-1$.

(3) A p-adic number α is defined by $\alpha = p^m \sum_{i=0}^{\infty} a_i p^i$ where $0 \leq a_i \leq p-1$. The field of p-adic numbers is denoted by Q_p.

(4) A p-adic number α is a p-adic integer if $m = 0$. The ring of p-adic integers is denoted by O_p.

(5) A p-adic integer α is a unit in O_p if $a_0 \neq 0$.

With these definitions, the equation $x^2 = 7$ from example 1.1 has a solution in O_3. Since the congruence $x^2 - 7 \equiv 0 \bmod 5$ has no solution in integers, the equation $x^2 = 7$ cannot be solved in O_5. One might conjecture that a solution in integers for the equation $f(x_1, x_2, \ldots, x_s) = 0$ exists whenever solutions exist in O_p for every prime p. This conjecture is much more difficult to disprove, but the equation $3x^3 + 4y^3 + 5z^3 = 0$ can be used to show that it is indeed false.

As suggested by the previous discussion, the main value of the p-adic integers as a tool in diophantine problems is for showing when solutions in integers do not exist. That is, when an equation has no solution in O_p for some p it has no solution in integers. In this thesis, a more positive approach is taken. The problems that are considered are posed in a p-adic setting and the solutions are p-adic. No attempt is made to relate the solutions to problems involving integers. One value of working with the p-adic numbers in this way is that they provide an unfamiliar system that is simple enough for a developing mathematician to make discoveries on his own.

The necessary background for reading the thesis is provided by a basic number theory course plus a course in which the p-adic numbers have been developed. Much of the material is a generalization of results found in the first chapter of Borevich and Shafarevich (3) so familiarity with this book would be most helpful.

The characterization of squares that appears in (3) was the motivation for Chapter II. This chapter, which is basic to the other two, is devoted to a development of a characterization of nth powers in O_p. Chapter III is an investigation of Waring's problem in a p-adic setting. This problem is easier to solve in the p-adic setting and the investigation produces some rather surprising results. Chapter IV is a study of Artin's conjecture for homogeneous forms in Q_p. The original conjecture is shown to be false and a weakened conjecture for diagonal forms is substituted. The remainder of the chapter provides a complete proof of the weakened conjecture.

CHAPTER II

POWERS OF P-ADIC NUMBERS

Several interesting diophantine problems in the field of number theory involve integral powers. One such problem is Waring's problem which is to be investigated in a p-adic setting in Chapter III. Basic to such an investigation is a usable characterization of the integral powers in O_p. The main objective of this chapter is to develop such a characterization. The first three theorems are essential to the development and suggestive of the primary ingredients of the characterization.

<u>Theorem 2.1</u> Let p be a prime and $n = mp^k$. Then if a and α are p-adic integers, $(a + \alpha p)^n \equiv a^n \bmod p^{k+1}$.

Proof: Using the binomial expansion

$$(a + \alpha p)^n = a^n + na^{n-1}(\alpha p) + \frac{n(n-1)a^{n-2}(\alpha p)^2}{2} + \cdots + (\alpha p)^n .$$

Since $n = mp^k$ every term in the expansion except a^n contains the factor p^{k+1}. It follows that $(a + \alpha p)^n \equiv a^n \bmod p^{k+1}$. ▲

<u>Theorem 2.2</u> Let $n = 2^k m$ where $k > 0$. Then for any 2-adic integer α, $(1 + 2\alpha)^n \equiv 1 \bmod 2^{k+2}$.

Proof: Consider first $(1 + 2\alpha)^2 = 1 + 4\alpha + 4\alpha^2 = 1 + 4\alpha(\alpha + 1)$. Since either α or $\alpha + 1$ is divisible by 2, it follows that $(1 + 2\alpha)^2 = 1 + 8\beta$ for some β in O_p. Now let $n = 2t$ and consider

$$(1+2\alpha)^n = [(1+2\alpha)^2]^t = (1+8\beta)^t$$

$$= 1 + t(8\beta) + \frac{t(t-1)(8\beta)^2}{2} + \cdots + (8\beta)^t .$$

Since $t = 2^{k-1}m$, each term in the expansion except 1 contains the factor 2^{k+2}. It follows that $(1+2\alpha)^n \equiv 1 \bmod 2^{k+2}$. ▲

Theorem 2.3 Let a be an integer and p be a prime. Then $a^{p^k} \equiv a \bmod p$.

Proof: The proof is by induction on k. If $k=1$, then $a^p \equiv a \bmod p$ by Fermat's theorem. Assume then that $a^{p^{k-1}} \equiv a \bmod p$. Therefore, $a^{p^k} = (a^{p^{k-1}})^p \equiv a^{p^{k-1}} \equiv a \bmod p$. ▲

Suppose that $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ is a unit in O_p, $p \neq 2$. Then $\varepsilon = a_0 + \alpha p$ where $\alpha = \sum_{i=1}^{\infty} a_i p^{i-1}$. According to theorem 2.1,

$$\varepsilon^n = (a_0 + \alpha p)^n \equiv a_0^n \bmod p^{k+1}$$

where $n = mp^k$ and theorem 2.3 implies

$$a_0^n = (a_0^m)^{p^k} \equiv a_0^m \bmod p.$$

From these two observations, we can conclude that $\varepsilon^n = \sum_{i=0}^{\infty} b_i p^i$ implies $b_0 \equiv a_0^m \bmod p$. That is, b_0 must be an mth power residue mod p. This in turn implies that $a_0^m = b_0 + \beta p$ for some β in O_p. Therefore,

$$\varepsilon^n \equiv a_0^n = (a_0^m)^{p^k} = (b_0 + \beta p)^{p^k} \equiv b_0^{p^k} \bmod p^{k+1} .$$

This development shows that when $p \neq 2$ necessary conditions for a unit $\varepsilon = \sum_{i=0}^{\infty} b_i p^i$ in O_p to be an nth power are the following. If $n = mp^k$ where $(m,p) = 1$, b_0 must be an mth power residue mod p and the congruence $\varepsilon \equiv b_0^{p^k} \bmod p^{k+1}$ must hold. The following characterization states that these conditions are also sufficient and provides similar conditions for units in O_2.

Characterization of nth powers in O_p. Let $n = mp^k$ where $(m,p) = 1$. Then the following conditions are necessary and sufficient for a unit $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ to be an nth power in O_p.

(1) The integer a_0 is an mth power residue mod p.

(2) If $p = 2$ and $k > 0$ then $\varepsilon \equiv 1 \bmod 2^{k+2}$.

(2′) If $p \neq 2$ then $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$.

This characterization will be verified in several steps which are undertaken in the following theorems. The overall plan is to consider mth powers where $(m,p) = 1$, then p^kth powers, and finally mp^kth powers.

Theorem 2.4 Let $f(x_1,x_2, \ldots ,x_s)$ be a polynomial whose coefficients are p-adic integers. Suppose $\alpha_1,\alpha_2, \ldots ,\alpha_s$ are p-adic integers such that for some i, $1 \leq i \leq s$,

$$f(\alpha_1,\alpha_2, \ldots ,\alpha_s) \equiv 0 \bmod p^{2e+1}$$

$$\frac{\partial f}{\partial x_i} (\alpha_1,\alpha_2, \ldots ,\alpha_s) \equiv 0 \bmod p^e$$

$$\frac{\partial f}{\partial x_i} (\alpha_1,\alpha_2, \ldots ,\alpha_s) \not\equiv 0 \bmod p^{e+1}$$

where e is a nonnegative integer. Then there exist p-adic integers $\theta_1, \theta_2, \ldots, \theta_s$ such that $f(\theta_1, \theta_2, \ldots, \theta_s) = 0$ and $\theta_i \equiv \alpha_i \bmod p^{e+1}$ for every i, $1 \leq i \leq s$.

Proof: See Borevich and Shafarevich (3, p. 42). ▲

<u>Theorem 2.5</u> If p is a prime and $(a,p) = 1$, then the congruence $x^n \equiv a \bmod p$ has $(n, p-1)$ solutions or no solutions according as $a^{(p-1)/(n,p-1)} \equiv 1 \bmod p$ or $a^{(p-1)/(n,p-1)} \not\equiv 1 \bmod p$.

Proof: See Niven and Zuckerman (11, p. 54). ▲

<u>Theorem 2.6</u> Let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ be a unit in O_p and $(m,p) = 1$. Then ε is an mth power in O_p if and only if a_0 is an mth power residue mod p.

Proof: Suppose first that ε is an mth power in O_p. Then there exists a $\delta = \sum_{i=0}^{\infty} b_i p^i$ in O_p for which $\delta^m = \varepsilon$. Now $\delta = b_0 + \beta p$ for some β in O_p and by theorem 2.1, $\delta^m = (b_0 + \beta p)^m \equiv b_0^m \bmod p$. Since $\varepsilon \equiv a_0 \bmod p$ and $\varepsilon = \delta^m$, it follows that $a_0 \equiv b_0^m \bmod p$. Hence, a_0 is an mth power residue mod p.

Now, suppose that a_0 is an mth power residue mod p. The proof that ε is an mth power will be accomplished when the equation $x^m - \varepsilon = 0$ is shown to have a solution in O_p. If $d = (m, p-1)$, there exist integers r and s, $s < 0$, such that $d = rm + s(p-1)$. Since d divides m, a_0 is a dth power residue so there exists an integer c such that $c^d \equiv a_0 \bmod p$. Now in order to apply theorem 2.4, let $f(x) = x^m - \varepsilon$. Then

$$f(c^r) = c^{rm} - \varepsilon = c^{d-s(p-1)} - \varepsilon = (c^d)^{1-s(p-1)/d} - \varepsilon$$

or

$$f(c^r) \equiv (a_0)^{1-s(p-1)/d} - \varepsilon \bmod p.$$

Since a_0 is an mth power residue mod p, theorem 2.5 implies $a_0^{(p-1)/d} \equiv 1 \bmod p$. It follows that $f(c^r) \equiv a_0 - \varepsilon \bmod p$ or $f(c^r) \equiv 0 \bmod p$. On the other hand, $f'(x) = mx^{m-1}$ so $f'(c^r) = m(c^r)^{m-1} \not\equiv 0 \bmod p$. Therefore, theorem 2.4 implies the existence of a δ in O_p for which $f(\delta) = 0$. It follows that $\varepsilon = \delta^m$ and the proof is complete. ▲

<u>Theorem 2.7</u> Let ε be a unit in O_p, $p \neq 2$, such that $\varepsilon \equiv 1 \bmod p^{k+1}$. Then ε is a p^kth power in O_p.

Proof: The method of proof will be to construct a p-adic sequence $\{B_n\}$, $B_n = \sum_{i=0}^{n} b_i p^i$, with the property that $B_n^{p^k} \equiv \varepsilon \bmod p^{n+1}$ for every n. Then, if $\delta = \sum_{i=0}^{\infty} b_i p^i$, $\delta^{p^k} \equiv \varepsilon \bmod p^{n+1}$ for every n and hence $\delta^{p^k} = \varepsilon$. Actually, we will prove the slightly stronger result that $B_n^{p^k} \equiv \varepsilon \bmod p^{n+k+1}$ for every n. The construction is by induction on n. If $n = 0$, then $1^{p^k} \equiv 1 \bmod p^{k+1}$ so $B_0 = b_0 = 1$. Now, suppose $B_{n-1} = \sum_{i=0}^{n-1} b_i p^i$ has been determined so that $(B_{n-1})^{p^k} \equiv \varepsilon \bmod p^{n+k}$. This implies that $(B_{n-1})^{p^k} = \varepsilon + \alpha p^{n+k}$ for some α in O_p. Now, $(B_{n-1}, p) = 1$ so there exists an integer b_n where $0 \leq b_n < p$ and $\alpha + b_n(B_{n-1})^{p^k - 1} \equiv 0 \bmod p$. With this choice for b_n, $B_n = B_{n-1} + b_n p^n = \sum_{i=0}^{n} b_i p^i$. When $(B_{n-1} + b_n p^n)^{p^k}$ is expanded the first two terms are

$$(B_{n-1})^{p^k} + p^k(B_{n-1})^{p^k-1}(b_n p^n) .$$

The third term $p^k(p^k - 1)(B_{n-1})^{p^k-2}(b_n p^n)^2/2$ and all remaining terms contain the factor p^{n+k+1}. This implies that

$$B_n^{p^k} \equiv (B_{n-1})^{p^k} - [b_n(B_{n-1})^{p^k-1}]p^{n+k} \bmod p^{n+k+1}$$

$$= \varepsilon + \alpha p^{n+k} + [b_n(B_{n-1})^{p^k-1}]p^{n+k}$$

$$= \varepsilon + [\alpha + b_n(B_{n-1})^{p^k-1}]p^{n+k}$$

$$\equiv \varepsilon \bmod p^{n+k+1} .$$

Therefore, $\{B_n\}$ is defined by induction and $\delta^{p^k} = \varepsilon$ where $\delta = \lim_{n\to\infty} B_n$.

▲

<u>Theorem 2.8</u> Let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ be a unit in O_p where $p \neq 2$. Then, ε is a p^kth power in O_p if and only if $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$.

Proof: Suppose first that ε is a p^kth power in O_p and let $\delta^{p^k} = \varepsilon$ where $\delta = \sum_{i=0}^{\infty} b_i p^i$. By theorem 2.3, $b_0^{p^k} \equiv b_0 \bmod p$. If $\delta = b_0 + \beta p$, by theorem 2.1,

$$\varepsilon = \delta^{p^k} = (b_0 + \beta p)^{p^k} \equiv b_0^{p^k} \bmod p^{k+1} .$$

It follows that $b_0 \equiv \varepsilon \equiv a_0 \bmod p$ which implies that $a_0 = b_0$. Therefore, $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$ completing the proof.

Now suppose $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ is a unit in O_p and $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$. In order to apply theorem 2.4, consider the function $f(x,y) = p^k(x^{p^k} - \varepsilon y)$. Observe first that $f(a_0,1) = p^k(a_0^{p^k} - \varepsilon) \equiv 0 \bmod p^{2k+1}$. Also, $\frac{\partial f}{\partial y} = -p^k\varepsilon$ so $\frac{\partial f}{\partial y}(a_0,1) = -p^k\varepsilon \equiv 0 \bmod p^k$ while $\frac{\partial f}{\partial y}(a_0,1) \not\equiv 0 \bmod p^{k+1}$. Therefore, by theorem 2.4, there exist μ and δ in O_p such that $f(\mu,\delta) = 0$ where $\mu \equiv a_0 \bmod p^{k+1}$ and

$\delta \equiv 1 \bmod p^{k+1}$. This implies that $f(\mu,\delta) = p^k(\mu^{p^k} - \varepsilon\delta) = 0$ or $\mu^{p^k} = \varepsilon\delta$. Now, by theorem 2.7, $\delta \equiv 1 \bmod p^{k+1}$ implies that δ is a p^kth power in O_p. It follows, since $\varepsilon = \delta^{-1}\mu^{p^k}$, that ε is a p^kth power in O_p also. ▲

Theorem 2.9 Let ε be a unit in O_2 and $k > 0$. Then ε is a 2^kth power in O_2 if and only if $\varepsilon \equiv 1 \bmod 2^{k+2}$.

Proof: If ε is a 2^k power in O_2, there exists a δ in O_2 for which $\delta^{2^k} = \varepsilon$. Since ε is a unit in O_2, δ must be a unit also, so let $\delta = 1 + 2\alpha$. Theorem 2.2 implies that $\delta^{p^k} = (1+2\alpha)^{2^k} \equiv 1 \bmod 2^{k+2}$. Therefore, $\varepsilon \equiv 1 \bmod 2^{k+2}$.

Now, suppose $\varepsilon \equiv 1 \bmod 2^{k+2}$. The proof that ε is a 2^kth power is the same as the proof in theorem 2.7 with two alterations. In the first place, $p = 2$. The other difference is that when the sequence $\{B_n\}$ is constructed so that $B_n^{2^k} \equiv \varepsilon \bmod 2^{n+k+1}$ there are two choices for B_1. Since $1^{2^k} \equiv 1 \bmod 2^{k+1}$, $B_0 = b_0 = 1$. It is also true that $1^{2^k} \equiv (1+2)^{2^k} \equiv 1 \bmod 2^{k+2}$ so B_1 can be chosen as either 1 or 3. Once B_1 is chosen, the construction proceeds exactly as in theorem 2.7. The result is that $\{B_n\}$ is constructed by induction so that $\delta = \lim_{n\to\infty} B_n$ and $\delta^{p^k} = \varepsilon$. ▲

The fact that two values of δ can be constructed so that $\delta^{2^k} = \varepsilon$ is natural since 2^k is an even number. Suppose $\delta = \sum_{i=0}^{\infty} a_i 2^i$ is one 2-adic unit for which $\delta^{2^k} = \varepsilon$. Then,

$$\delta = 1 + \sum_{i=1}^{\infty} a_i 2^i \quad \text{and} \quad -\delta = 1 + \sum_{i=1}^{\infty} (1 - a_i) 2^i .$$

Now, $(-\delta)^{2^k} = \varepsilon$ so $\{1 + \sum_{i=1}^{n} (1 - a_i) 2^i\}$ must be the other 2-adic sequence that can be constructed in theorem 2.9.

The characterization has now been established for mth powers with $(m,p) = 1$ and for p^kth powers. When ε is an mp^kth power in O_p, it is obvious that ε is both an mth power and a p^kth power in O_p. In terms of the characterization, this statement reads, when ε is an nth power, conditions (1) and (2 or $2'$) are satisfied. It is now necessary to show that when conditions (1) and (2 or $2'$) are satisfied ε is an nth power. That is, when ε is both an mth and a p^kth power, it must also be an mp^kth power in O_p. This is shown to be the case by theorem 2.10 due to the fact that $(m,p^k) = 1$. With this theorem, the proof of the characterization of nth powers of units in O_p is complete.

<u>Theorem 2.10</u> Let ε be a unit in O_p such that ε is both an mth power and an nth power in O_p where $(m,n) = 1$. Then ε is an mnth power in O_p.

Proof: Let $\varepsilon = \delta^m$ and $\varepsilon = \mu^n$ where δ and μ are units in O_p. Because $(m,n) = 1$, there exist integers r and s such that $1 = rm + sn$. Therefore,

$$\varepsilon = \varepsilon^{rm+sn} = (\varepsilon^r)^m(\varepsilon^s)^n = (\mu^{nr})^m(\delta^{ms})^n = (\mu^r\delta^s)^{mn}$$

Since μ and δ are units in O_p, $\mu^r\delta^s$ is an element in O_p for any integers r and s. It follows that ε is an mnth power in O_p. ▲

Given a specific unit in O_p, one could determine whether or not it is an nth power by checking the two conditions of the characterization that has just been established. However, for application purposes, the criterion in the following theorem is much more practical.

Theorem 2.11 Let $n = mp^k$ where $(m,p) = 1$ and let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ be a unit in O_p. Define t as $k+1$ if $p \neq 2$ and $k+2$ if $p = 2$. Then ε is an nth power in O_p if and only if $\varepsilon \equiv \delta^n \bmod p^t$ for some δ in O_p.

Proof: If ε is an nth power in O_p, then $\varepsilon = \delta^n$ for some δ in O_p. Consequently, $\varepsilon \equiv \delta^n \bmod p^{k+2}$ which verifies the only if statement for all p.

Now suppose $\varepsilon \equiv \delta^n \bmod p^{k+1}$ where $p \neq 2$. Then if $\delta^n = \sum_{i=0}^{\infty} c_i p^i$, $c_i = a_i$ for $0 \leq i \leq k$. Since δ^n is an nth power, the characterization states that c_0 is an mth power residue mod p and $\delta^n \equiv c_0^{p^k} \bmod p^{k+1}$. Now $c_0 = a_0$ and $\varepsilon \equiv \delta^n \bmod p^{k+1}$ so a_0 is an mth power residue mod p and $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$. It follows that ε is an nth power in O_p. The same argument holds for $p = 2$ except $\varepsilon \equiv \delta^n \bmod 2^{k+2}$ so $c_i = a_i$ for $0 \leq i \leq k+1$. ▲

The next theorem is included here because its proof makes use of the first condition in the characterization of nth powers and its conclusions are important to the developments in Chapters III and IV.

Theorem 2.12 Let $n = mp^k$, $(m,p) = 1$, $d = (m, p-1)$, and let α, β, and ε be units in O_p. Then

(1) $\alpha x^n \equiv \beta \bmod p$ has a solution n integers if and only if $\alpha y^d \equiv \beta \bmod p$ has a solution in integers.

(2) ε is an mth power in O_p if and only if ε is a dth power in O_p.

Proof: To prove (1), suppose first that θ is a solution for $\alpha x^n \equiv \beta \bmod p$. By definition, $d = (m, p-1) = (n, p-1)$ so $n = ds$ for some integer s. This implies that $\alpha x^n = \alpha(x^s)^d \equiv \beta \bmod p$. It follows that θ^s is a solution of $\alpha y^d \equiv \beta \bmod p$. Now suppose θ is a solution of $\alpha y^d \equiv \beta \bmod p$. Since $d = (n, p-1)$, there exist integers r and s for which $d = nr + (p-1)s$. This gives

$$\alpha y^d = \alpha y^{nr+(p-1)s} = \alpha(y^r)^n(y^{p-1})^s \equiv \beta \bmod p .$$

Since both α and β are units and θ is a solution of $\alpha y^d \equiv \beta \bmod p$, θ must be a unit also. Therefore, $\theta^{p-1} \equiv 1 \bmod p$ and it follows that θ^r is a solution of $\alpha x^n \equiv \beta \bmod p$.

To prove (2), let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$. Since $(m,p) = 1$, ε is an mth power in O_p if and only if a_0 is an mth power residue mod p. Likewise, $(d,p) = 1$ so ε is a dth power in O_p if and only if a_0 is a dth power residue mod p. Therefore, to prove (2), it suffices to show that a_0 is an mth power residue mod p if and only if a_0 is a dth power residue mod p. This is true since $x^m \equiv a_0 \bmod p$ has a solution in integers if and only if $y^d \equiv a_0 \bmod p$ has a solution in integers, which is a special case of (1). ▲

Having established a characterization for nth powers of units in O_p, one might consider nth powers of all p-adic integers or even all p-adic numbers. Actually, the extension to include all p-adic numbers is a very small one. All non-zero p-adic numbers can be represented uniquely as εp^t where ε is a unit in O_p and t is an integer. A p-adic number in this form is an nth power if and only if ε is an nth power in O_p and t is a multiple of n.

This chapter is concluded with an interesting result that developed from considering the significance of the second condition of the nth power criteria.

The condition for a unit ε to be an mth power where $(m,p) = 1$ is very direct. If $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$, then a_0 is either an mth power residue or nonresidue mod p and ε is classified immediately. The condition for p^kth powers is less direct. For example, in O_5, if $a_0 = 2$ what values of ε, if any, are 5th, 5^2th, 5^3th, ... powers? A little arithmetic shows that $2^5 \equiv 7 \bmod 5^2$, $2^{25} \equiv 57 \bmod 5^3$, and $2^{125} \equiv 182 \bmod 5^4$. So 7, 57, and 182 are respectively 5th, 5^2th, and 5^3th powers in O_5. Also, $\varepsilon_1 = 2 + 1{\cdot}5 + 5^2\alpha$, $\varepsilon_2 = 2 + 1{\cdot}5 + 2{\cdot}5^2 + 5^3\alpha$, and $\varepsilon_3 = 2 + 1{\cdot}5 + 2{\cdot}5^2 + 1{\cdot}5^3 + 5^4\alpha$ are respectively 5th, 5^2th, and 5^3th powers for any α in O_5. When the ε_i are written in this way, the coefficients suggest two possible conjectures regarding the further coefficients. One is that the first $i+1$ coefficients of ε_i and ε_{i+1} agree. This conjecture is verified when the sequence $\{2^{5^n}\}$ is shown, in the next theorem, to be a 5-adic sequence. The other possible conjecture is that the pattern 2,1,2,1 of the first coefficients is repeated. That this is incorrect is seen by direct computation since the next coefficient is a 3 instead of a 2. An indirect argument which shows that no such pattern could continue is the following. The 5-adic integer $2 + 1{\cdot}5 + 2{\cdot}5^2 + 1{\cdot}5^3 + \cdots$ corresponds to the rational number $-7/24$. However, the 5-adic integer that corresponds to the sequence $\{2^{5^n}\}$ is an element of the set $\mathcal{S}_5$ in the next theorem. As the theorem develops, it will be obvious since $(-7/24)^5 \neq -7/24$ that $-7/24$ is not an element of $\mathcal{S}_5$.

<u>Theorem 2.13</u> Let p be any prime. Then there exists a set $\mathcal{S}_p$ in O_p with exactly p elements ($p-1$ units and zero) with the property that for every α in $\mathcal{S}_p$, $\alpha^{p^k} = \alpha$ for every integer k.

Proof: To prove the existence of the $p-1$ units in $\mathcal{S}_p$, it is sufficient to show that for each a_0, $1 \leq a_0 \leq p-1$, exactly one unit ε exists with the property that $\varepsilon \equiv a_0 \bmod p$ and $\varepsilon^{p^k} = \varepsilon$ for every k. If $p = 2$, one is the only unit with this property, so $\mathcal{S}_2 = \{0,1\}$. If $p \neq 2$ for a given a_0, $1 \leq a_0 \leq p-1$, ε can be constructed as follows. For each $i \geq 0$, define A_i by $A_i \equiv a_0^{p^i} \bmod p^{i+1}$ and $1 \leq A_i \leq p^{i+1} - 1$. As defined, $A_i \equiv a_0 \bmod p$ for every i and

$$A_i - A_{i-1} \equiv a_0^{p^i} - a_0^{p^{i-1}} = a_0^{p^{i-1}}(a_0^{p^i - p^{i-1}} - 1) \equiv 0 \bmod p^i .$$

The final congruence is true because $p^i - p^{i-1} = \varphi(p^i)$ where φ is Euler's function. Therefore, $A_i \equiv A_{i-1} \bmod p^i$ for $i \geq 1$ which implies that $\{A_i\}$ is a p-adic sequence. The condition that $1 \leq A_i \leq p^{i+1} - 1$ is not needed to obtain this result, but it gives the sequence $\{A_i\}$ canonical form. Now, if ε is defined by $\varepsilon = \lim_{n\to\infty} A_n$, $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ where $A_n = \sum_{i=0}^{n} a_i p^i$. By definition $\varepsilon \equiv a_0^{p^k} \bmod p^{k+1}$ for every k so ε is a p^kth power in O_p for every k. To show that for a given a_0 only one such ε exists, suppose ε and ε' exist such that $\varepsilon \equiv a_0 \equiv \varepsilon' \bmod p$ and both ε and ε' are p^kth powers for every k. Then $\varepsilon \equiv a_0^{p^k} \equiv \varepsilon' \bmod p^{k+1}$ for every k and $\varepsilon \equiv \varepsilon' \bmod p^{k+1}$ for every k implies that $\varepsilon = \varepsilon'$. To show that $\varepsilon^{p^k} = \varepsilon$ for every k let i be a fixed positive integer and consider ε^{p^i}. Since ε is a p^kth power for every k, ε^{p^i} is a p^kth power for every k. Also since $a_0^{p^i} \equiv a_0 \bmod p$, $\varepsilon^{p^i} \equiv a_0 \bmod p$. We have just shown that ε is the only element in O_p that satisfies these two conditions, therefore $\varepsilon^{p^i} = \varepsilon$. This argument is valid for any i so $\varepsilon^{p^k} = \varepsilon$ for every k.

To show that zero is the only non-unit in $\mathcal{S}_p$, consider any non-unit β where $\beta \neq 0$. Represent β as εp^r where ε is a unit in O_p and r is an integer greater than one. This representation shows that β cannot be a power higher than the rth power. Hence, β cannot be a p^kth power for every k, so β cannot be an element of $\mathcal{S}_p$. This completes the proof of the theorem. ▲

The set $\mathcal{S}_p$ has several interesting properties. If g is a primitive root mod p and ε is the element of $\mathcal{S}_p$ that corresponds to the sequence $\{g^{p^n}\}$, then each element of the set $\{\varepsilon, \varepsilon^2, \ldots, \varepsilon^{p-1}\}$ is a distinct element in $\mathcal{S}_p$. Therefore, $\mathcal{S}_p - \{0\} = \{\varepsilon, \varepsilon^2, \ldots, \varepsilon^{p-1}\}$ which is a cyclic group under multiplication. Also, since 1 is in $\mathcal{S}_p$ and $\varepsilon^{p-1} \equiv 1 \bmod p$, it follows that $\alpha^{p-1} = 1$ for every α in $\mathcal{S}_p - \{0\}$. This implies one more property, that $\mathcal{S}_p$ contains the p distinct p-adic roots of the equation $x^p - x = 0$.

CHAPTER III

WARING'S PROBLEM

The problem referred to as Waring's problem is the following. Given a positive integer n, find a positive integer $g(n)$ such that each positive integer is the sum of $g(n)$ nth powers of nonnegative integers. Since criteria have been established in Chapter II for determining nth powers in Q_p the groundwork has been done for considering Waring's problem in a p-adic setting. The objective of this chapter is to investigate the number of nth powers needed to represent any p-adic integer.

Suppose α is a non-unit in O_p. Then $\alpha = 1 + (\alpha - 1)$ and $\alpha - 1$ is a unit in O_p. Since 1 is an nth power in O_p, it follows that if any unit in O_p can be represented as the sum of $g(n)$ nth powers, any p-adic integer can be represented as the sum of $g(n) + 1$ nth powers. This observation indicates that an investigation of the units in O_p will supply information about all p-adic integers. The theorems of this chapter are, therefore, designed to investigate the following problem.

> <u>Waring's problem for p-adic integers</u>. Given a positive integer n and a prime p, determine the smallest positive integer $g(n)$ such that every unit in O_p can be represented as the sum of $g(n)$ nth powers in O_p.

<u>Theorem 3.1</u> Let $n = mp^k$ where $(m,p) = 1$. Then any unit in O_p can be represented as the sum of fewer than p^{k+1} nth powers in O_p if $p \neq 2$ and fewer than 2^{k+2} nth powers in O_2 if $p = 2$.

Proof: Let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ be a unit in O_p where $p \neq 2$. Then

$$\varepsilon = \sum_{i=0}^{k} a_i p^i + \sum_{i=k+1}^{\infty} a_i p^i .$$

Now, let $N = \sum_{i=0}^{k} a_i p^i$ and $\alpha p^{k+1} = \sum_{i=k+1}^{\infty} a_i p^i$. Then

$$\varepsilon = N + \alpha p^{k+1} = \sum_{j=1}^{N} 1 + \alpha p^{k+1} = \sum_{j=1}^{N-1} 1 + (1 + \alpha p^{k+1}) .$$

By theorem 2.11, $1 + \alpha p^{k+1}$ is an nth power in O_p so ε is expressed as the sum of N nth powers. Since $0 \leq a_i \leq p-1$ for every i, by definition $N < p^{k+1}$. Similarly, when $p = 2$,

$$\varepsilon = \sum_{i=0}^{\infty} a_i 2^i = \sum_{i=0}^{k+1} a_i 2^i + \sum_{i=k+2}^{\infty} a_i 2^i = \sum_{j=1}^{N-1} 1 + (1 + \alpha 2^{k+2})$$

where $N = \sum_{i=0}^{k+1} a_i 2^i < 2^{k+2}$. ▲

This theorem shows that an upper bound for $g(n)$ is available for any n. The interesting part of the problem is to investigate special cases to determine if and when this upper bound can be lowered. Theorem 3.2 shows that for every p there exist values of n for which the upper bound of theorem 3.1 cannot be lowered.

<u>Theorem 3.2</u> Let p be a prime and $n = (p-1)p^k$, $k \geq 0$. Then there

exists a unit in O_p that cannot be written as the sum of fewer than $p^{k+1}-1$ nth powers if $p \neq 2$ or $2^{k+2}-1$ nth powers if $p=2$.

Proof: Let $\varepsilon = \sum_{i=0}^{k} (p-1)p^i$ where $p \neq 2$. Note that ε is a unit and $\varepsilon = p^{k+1}-1$. Suppose $\alpha_1^n + \alpha_2^n + \cdots + \alpha_s^n$ is any sum of nth powers of p-adic integers where $s < p^{k+1}-1$. Since $n = (p-1)p^k = \varphi(p^{k+1})$, $\alpha_i^n \equiv 1 \bmod p^{k+1}$ when α_i is a unit in O_p. On the other hand, if α_i is a non-unit, $\alpha_i^n \equiv 0 \bmod p^{k+1}$. Therefore, each α_i^n is congruent to 1 or 0 mod p^{k+1}. It follows that there exists an integer a such that $0 \leq a \leq s < p^{k+1}-1$ and $\alpha_1^n + \alpha_2^n + \cdots + \alpha_s^n \equiv a \bmod p^{k+1}$. Now, $p^{k+1}-1 \not\equiv a \bmod p^{k+1}$ since $0 \leq a < p^{k+1}-1$. Therefore, $\varepsilon \neq \alpha_1^n + \alpha^n + \cdots + \alpha_s^n$ when $s < p^{k+1}-1$.

When $p=2$ the argument is identical except $n = 2^k$ and the conclusion is that $\varepsilon = \sum_{i=0}^{k+1} 2^i = 2^{k+2}-1$ cannot be represented as the sum of fewer than $2^{k+2}-1$ nth powers in O_2. ▲

Now, return to the case $n = mp^k$ where $(m,p)=1$. It has just been shown that when $m = p-1$, $g(n)$ attains the maximum value of $p^{k+1}-1$ if $p \neq 2$ and $2^{k+2}-1$ if $p=2$. The next theorem shows that $g(n)$ attains the minimum value when $n=m$ and $(m,p-1)=1$. In this case, every unit is an nth power in O_p; that is, $g(n)=1$.

<u>Theorem 3.3</u> Let $(n,p)=1$ and ε be a unit in O_p. Then ε can be represented as the sum of $d = (n,p-1)$ nth powers in O_p.

Proof: If $\varepsilon = \sum_{i=0}^{\infty} a_i 2^i$ is a unit in O_2 then $a_0 = 1$. Therefore, a_0 is an nth power residue mod 2 and by theorem 2.6 ε is an nth power in O_2. It follows that ε can be written as the sum of $(n,2-1)=1$ nth power in O_2.

When $p \neq 2$ theorem 2.12 states that when $(n,p) = 1$ a unit in O_p is an nth power if and only if it is a dth power in O_p. Therefore, to prove the theorem in this case, it suffices to show that ε can be represented as the sum of d dth powers of units in O_p. First let $\varepsilon = \sum_{i=0}^{\infty} a_i p^i$ and show that the congruence $x_1^d + x_2^d + \cdots + x_d^d \equiv a_0 \bmod p$ has a solution in integers. This problem can be stated as follows. Let Z_p be the finite field of integers mod p, and let G be the multiplicative group $Z_p - \{0\}$. Then, given any element a_0 of G show that $x_1^d + x_2^d + \cdots + x_d^d = a_0$ has a solution in Z_p. Define the subgroup H by $H = \{x^d \mid x \in G\}$ and K_r as the set of elements in G that can be represented as a sum of r dth powers of elements in Z_p. As a set, G consists of the elements $\{1,2, \ldots ,p-1\}$. Therefore, every element g in G can be written as $\sum_{i=1}^{g} 1$ which implies that $K_{p-1} = G$. Let $t = \min\{r \mid K_r = G\}$ and consider the set difference $K_t - K_{t-1}$. By the definition of t this difference is not empty, so let $x \in K_t - K_{t-1}$. Since $x \in K_t$, there exist $x_1, x_2, \ldots ,x_t$ in Z_p for which $x = \sum_{i=1}^{t} x_i^d$. In order to show that $K_{t-1} - K_{t-2}$ is not empty, define x' in K_{t-1} by $x' = \sum_{i=1}^{t-1} x_i^d$. Suppose that $x' \in K_{t-2}$. Then,

$$x' = \sum_{i=1}^{t-2} y_i^d \quad \text{for some} \quad y_1, y_2, \ldots ,y_{t-2} \quad \text{in} \quad Z_p \,.$$

It follows that

$$x = x' + x_t^d = \sum_{i=1}^{t-2} y_i^d + x_t^d$$

which implies that $x \in K_{t-1}$. This is a contradiction since $x \in K_t - K_{t-1}$. Therefore, $x' \notin K_{t-2}$ and it follows that $x' \in K_{t-1} - K_{t-2}$. By a similar argument, if $x'' = \sum_{i=1}^{t-2} x_i^d$, then

$x'' \in K_{t-2} - K_{t-3}$. In general

$$\sum_{i=1}^{s} x_i^d \in K_s - K_{s-1} \quad \text{for} \quad 2 \leq s \leq t \; .$$

Thus, in the sequence of inclusions $K_1 \subset K_2 \subset \cdots \subset K_t$ each inclusion has been shown to be proper.

The next objective is to show that not only is $K_s - K_{s-1}$ non-empty for every s, $2 \leq s < t$, but that each such set contains a coset of H. This follows if $x \in K_s - K_{s-1}$ implies $xH \subset K_s - K_{s-1}$. Suppose $y \in xH$; that is, $y = xz^d$ for some $z^d \in H$. Since $x \in K_s$, x can be represented as $\sum_{i=1}^{s} x_i^d$ and $y = xz^d = \sum_{i=1}^{s} (x_i z)^d$. This implies that $y \in K_s$. To show that $y \notin K_{s-1}$, assume the contrary that $y = \sum_{i=1}^{s-1} y_i^d$. Thus, $y = xz^d = \sum_{i=1}^{s-1} y_i^d$ which implies that $x = \sum_{i=1}^{s-1} (y_i/z)^d$. This is a contradiction since $x \notin K_{s-1}$. This contradiction shows that $y \notin K_{s-1}$ completing the proof that $y \in K_s - K_{s-1}$. Since y is an arbitrary element of xH, it follows that $xH \subset K_s - K_{s-1}$.

Recalling the sequence $K_1 \subset K_2 \subset \cdots \subset K_t$, by definition $K_1 = H$ and $K_t = G$. It has been shown that each $K_s - K_{s-1}$ contains at least one coset of H. Therefore, G must contain at least t different cosets of H. If $o(X)$ denotes the number of elements in X, this conclusion is written $o(H)\cdot t \leq o(G)$. Therefore, $t \leq o(G)/o(H)$ or t is less than or equal to the index of H in G. Since $H = \{x^d \mid x \in G\}$ is a subgroup of G, the index of H in G can be computed as the number of distinct values of x in G for which $x^d = 1$. By Lagrange's theorem, $x^d = 1$ has at most d incongruent solutions mod p. Therefore, the index of H is not more than d. It follows that $t \leq o(G)/o(H) \leq d$. Originally, t was defined as the smallest number such that every element of G can be represented as a sum of t dth power of elements in Z_p. Thus, since $a_0 \in G$ and $t \leq d$, there exists a solution in Z_p of the equation $x_1^d + x_2^d + \cdots + x_d^d = a_0$. Furthermore, since $a_0 \neq 0$, at least one x_i must be non-zero. Let $(b_1, b_2, \ldots, b_d)$ be a solution to the above equation and assume, without loss of generality, that $b_d \neq 0$. Therefore, $b_1^d + b_2^d + \cdots + b_d^d \equiv a_0 \equiv \varepsilon \bmod p$ or $\varepsilon = b_1^d + b_2^d + \cdots + b_d^d + \beta p$ for some β in O_p. By definition of Z_p, $b_d \neq 0$ implies that b_d is a unit in O_p, so theorem 2.11 implies that $b_d^d + \beta p$ is a dth power in O_p. Therefore, ε is represented as the sum of d dth powers in O_p which completes the proof of the theorem. ▲

After a proof of this length, one would hope for a significant improvement over previous results. With the condition $(n,p) = 1$, theorem 3.1 implies that $g(n) \leq p - 1$. Theorem 3.2, under the same condition, provides the specific case $n = p - 1$ for which $g(n) = p - 1$. In this case, since $(n, p-1) = p - 1$, theorem 3.3 offers no improvement. However, since $(n, p-1)$ is a divisor of $p - 1$, the conclusion that $g(n) \leq (n, p-1)$ is an improvement whenever $(n, p-1) \neq p - 1$. For example, in O_{71}, suppose $(n,71) = 1$. The possible values of $(n,70)$ are 1,2,5,7,10,14,35, and 70. The value of $g(n)$ will be 70 only if n is a multiple of 70. If n does not contain one of the factors 2,5, or 7, then $g(n) = 1$. That is, every unit in O_{71} is an nth power.

After the case where $(n,p) = 1$, it seems natural to investigate the opposite situation when $n = p^k$ with $k \geq 1$. The investigation of this case begins with the following lemma.

<u>Lemma 3.1</u> Let p be an odd prime and $k \geq 1$. Then the congruence $x^{p^k} + y^{p^k} + z^{p^k} \equiv 0 \bmod p$ has a solution (a,b,c) in integers such that $a^{p^k} + b^{p^k} + c^{p^k} \not\equiv 0 \bmod p^2$.

Proof: The binomial expansion of $(p-1)^{p^k}$ shows that $(p-1)^{p^k} \equiv -1 \bmod p^2$ which implies that $(p-1)^{p^k} \not\equiv p-1 \bmod p^2$. This along with the fact that $1^{p^k} \equiv 1 \bmod p^2$ makes the following definition possible. Let $t = \min\{r \mid r$ is a positive integer and $r^{p^k} \not\equiv r \bmod p^2\}$. With this definition, since $t-1 < t$, $(t-1)^{p^k} \equiv t-1 \bmod p^2$. Therefore, $1^{p^k} + (-t)^{p^k} + (t-1)^{p^k} \equiv 1 - t^{p^k} + t - 1 \bmod p^2$. Because of the definition of t, the right side of this congruence, $t - t^{p^k}$, cannot be congruent to zero mod p^2. On the other hand, since $a^{p^k} \equiv a \bmod p$ for every integer a, it follows that

$$1^{p^k} + (-t)^{p^k} + (t-1)^{p^k} \equiv 1 - t + t - 1 \equiv 0 \bmod p .$$

Therefore, $(1,-t,t-1)$ is a suitable solution, completing the proof of the lemma. ▲

In terms of p-adic integers, this lemma can be stated as follows. For any odd prime p and any positive integer k there exist p-adic integers (x,y,z) such that $x^{p^k} + y^{p^k} + z^{p^k} = \varepsilon p$ where ε is a unit in O_p.

<u>Theorem 3.4</u> Let α be a unit in O_p where $p \neq 2$. Then α can be represented as the sum of $(3^{k+1}-1)/2$ p^kth powers in O_p.

Proof: Note first that for every p-adic integer β, there exist p-adic integers x and μ such that $\beta = x^{p^k} + \mu p$. To see this, let $\beta = \sum_{i=0}^{\infty} b_i p^i$. Then $\beta \equiv b_0 \equiv b_0^{p^k} \bmod p$ and it follows that

$\beta = b_0^{p^k} + \mu p$ for some μ in O_p. Now, let (x,y,z) be p-adic integers provided by lemma 3.1 such that $x^{p^k} + y^{p^k} + z^{p^k} = \varepsilon p$ where ε is a unit in O_p. The following construction shows how to represent α as a sum of $(3^{k+1} - 1)/2$ p^kth powers. First, determine p-adic integers x_0 and μ_1 so that $\alpha = x_0^{p^k} + \mu_1 p$. Note that x_0 must be a unit in O_p since α is a unit in O_p. Now, determine x_1 and μ_2 so that $\mu_1\varepsilon^{-1} = x_1^{p^k} + \mu_2 p$. This gives

$$\alpha = x_0^{p^k} + (x_1^{p^k} + \mu_2 p)\,\varepsilon p = x_0^{p^k} + x_1^{p^k}\varepsilon p + \mu_2\varepsilon p^2 .$$

Next, determine x_2 and μ_3 so that $\mu_2\varepsilon^{-1} = x_2^{p^k} + \mu_3 p$. This gives

$$\begin{aligned}\alpha &= x_0^{p^k} + x_1^{p^k}\,\varepsilon p + \mu_2\varepsilon^{-1}(\varepsilon p)^2 \\ &= x_0^{p^k} + x_1^{p^k}\,\varepsilon p + (x_2^{p^k} + \mu_3 p)(\varepsilon p)^2 \\ &= x_0^{p^k} + x_1^{p^k}\,\varepsilon p + x_2^{p^k}(\varepsilon p)^2 + \mu_3\varepsilon^2 p^3 .\end{aligned}$$

Repeating this process $k+1$ times produces the result

$$\alpha = x_0^{p^k} + x_1^{p^k}(\varepsilon p) + x_2^{p^k}(\varepsilon p)^2 + \cdots + x_k^{p^k}(\varepsilon p)^k + \mu_{k+1}\varepsilon^k p^{k+1} .$$

Thus,
$$\begin{aligned}\alpha &= (x_0^{p^k} + \mu_{k+1}\varepsilon^k p^{k+1}) + \sum_{i=1}^{k} x_i^{p^k}(\varepsilon p)^i \\ &= (x_0^{p^k} + \mu_{k+1}\varepsilon^k p^{k+1}) + \sum_{i=1}^{k} x_i^{p^k}(x^{p^k} + y^{p^k} + z^{p^k})^i .\end{aligned}$$

Since x_0 is a unit in O_p, theorem 2.11 implies that $x_0^{p^k} + \mu_{k+1}\varepsilon^k p^{k+1}$ is a p^kth power in O_p. When expanded

$x_i^{p^k}(x^{p^k} + y^{p^k} + z^{p^k})^i$ produces 3^i terms, each of which is a p^kth power in O_p. For example, when $x_2^{p^k}(x^{p^k} + y^{p^k} + z^{p^k})^2$ is expanded, the result is $(x_2x^2)^{p^k} + (x_2y^2)^{p^k} + (x_2z^2)^{p^k} + (x_2xy)^{p^k} + (x_2xy)^{p^k} + (x_2xz)^{p^k} + (x_2xz)^{p^k} + (x_2yz)^{p^k} + (x_2yz)^{p^k}$. The net result is that α is expressed as the sum of $\sum_{i=0}^{k} 3^i = (3^{k+1} - 1)/2$ p^kth powers in O_p. ▲

The number of terms predicted in this theorem is considerably smaller than in theorem 3.1. The value of $(3^{k+1} - 1)/2$ is half as large as the $p^{k+1} - 1$ value in theorem 3.1 even when $p = 3$. As p becomes large, the improvement is considerably better. The next theorem is an example showing that, in general, the value $(3^{k+1} - 1)/2$ in theorem 3.3 cannot be lowered.

<u>Theorem 3.4</u> For every k, there exists a unit in O_3 that cannot be represented as the sum of fewer than $(3^{k+1} - 1)/2$, 3^kth powers in O_3.

Proof: Let $\varepsilon = \sum_{i=0}^{k} 3^i$. Note that ε is a unit in O_3 and $\varepsilon = (3^{k+1} - 1)/2$. Suppose $\alpha_1^{3^k} + \alpha_2^{3^k} + \cdots + \alpha_s^{3^k}$ is any sum of 3^kth powers in O_3 where $s < (3^{k+1} - 1)/2$. If α_i is a unit in O_3, then $\alpha_i = \pm 1 + 3\beta$ for some β in O_3. The binomial expansion shows that $\alpha_i^{3^k} = (\pm 1 + 3\beta)^{3^k} \equiv \pm 1 \bmod 3^{k+1}$. However, if α_i is a non-unit in O_3, $\alpha_i = 3\beta$ for some β in O_3 and $a_i^{3^k} = (3\beta)^{3^k} \equiv 0 \bmod 3^{k+1}$. Therefore, the value of each element in the sum $\alpha_1^{3^k} + \alpha_2^{3^k} + \cdots + \alpha_s^{3^k}$ is either 1, -1, or 0 mod 3^{k+1}. It follows that there exists an integer a such that $|a| \leq s$ and $\alpha_1^{3^k} + \alpha_2^{3^k} + \cdots + \alpha_s^{3^k} \equiv a \bmod 3^{k+1}$. Now $|a| \leq s$ and $s < (3^{k+1} - 1)/2$ imply that $-(3^{k+1} - 1)/2 < a < (3^{k+1} - 1)/2$. Since the set of integers $\{r \mid -(3^{k+1} - 1)/2 \leq r \leq (3^{k+1} - 1)/2\}$ constitutes a complete residue system mod 3^{k+1}, it follows that $a \not\equiv (3^{k+1} - 1)/2 \bmod 3^{k+1}$.

That is, $\varepsilon \not\equiv a \bmod 3^{k+1}$ which implies $\varepsilon \neq \alpha_1^{3^k} + \alpha_2^{3^k} + \cdots + \alpha_s^{3^k}$. Therefore, the unit $\varepsilon = (3^{k+1} - 1)/2$ cannot be represented as the sum of fewer than $(3^{k+1} - 1)/2$ 3^kth powers completing the proof. ▲

In terms of the function $g(n)$, theorem 3.4 shows that for odd primes, $g(n) \leq (3^{k+1} - 1)/2$ when $n = p^k$. Equality holds when $p = 3$ according to theorem 3.5.

As previously noted, the results of this chapter can be extended to include all p-adic integers due to the fact that any non-unit α can be represented as the nth power 1 plus the p-adic unit $\alpha - 1$. The results can also be extended to include all p-adic numbers as follows. Let εp^t be any non-zero p-adic number. Determine integers r and s so that $t = rn + s$ where $s > 0$. Then $\varepsilon p^t = \varepsilon p^s (p^r)^n$ and εp^s is a p-adic integer. As indicated above, εp^s can be represented as $x_1^n + x_2^n + \cdots + x_{h(n)}^n$ where $h(n) = g(n) + 1$. It follows that $\varepsilon p^t = y_1^n + y_2^n + \cdots + y_{h(n)}^n$ where $y_i = x_i p^r$ for every i, $1 \leq i \leq h(n)$.

CHAPTER IV

ARTIN'S CONJECTURE

The original conjecture made by Artin, as it pertains to p-adic numbers, was the following. If a homogeneous form of degree n with coefficients in Q_p contains more than n^2 variables, it must have a non-trivial zero in O_p. The definition of a homogeneous form requires only that each term be of the same degree and, in general, such a form is difficult to work with. In this respect, it is fortunate that Artin's conjecture in its original form has been proven false. The most famous counterexample was given by Terjanian (12).

Terjanian observed that the function $g(x) = g(x_1,x_2,x_3)$ defined by

$$g(x) = x_1^4 + x_2^4 + x_3^4 - x_1^2x_2^2 - x_2^2x_3^2 - x_1^2x_3^2 - (x_1 + x_2 + x_3)(x_1x_2x_3)$$

has the following properties: $g(x) \equiv 1 \bmod 4$ if some x_i is odd and $g(x) \equiv 0 \bmod 16$ if every x_i is even. He then constructed the form

$$f = g(x) + g(y) + g(z) + 4g(u) + 4g(v) + 4g(w) \ .$$

This form is homogeneous of degree 4 with 18 variables. According to the conjecture, it should have a non-trivial zero in O_2. The fact that $f \equiv 0 \bmod 16$ only if each of the variables is even means that for any zero $(\theta_1,\theta_2, \ldots ,\theta_{18})$ each θ_i must be even. Suppose $(\theta_1,\theta_2, \ldots ,\theta_{18})$ is a non-trivial zero of f in O_2. Each of the non-zero θ_i has the form $\varepsilon_i 2^{k_i}$ where ε_i is a unit in O_2. If k be the minimum k_i, then $(2^{-k}\theta_1, 2^{-k}\theta_2, \ldots, 2^{-k}\theta_{18})$ is another zero of f. However, at least one $2^{-k}\theta_i$ is not even so this cannot be a zero of f. We must conclude that f has no non-trivial zeros in O_2.

A paper by Browkin (4) gives an even more dramatic counterexample. By using a tremendous construction, he demonstrates that for any prime the number of variables needed to insure non-trivial zeros for forms of degree n is not less than n^3.

In view of these counterexamples, the conjecture must be weakened in order to present an interesting problem. The objective of this chapter will be to investigate and eventually prove such a weakened conjecture.

Definition 4.1 A diagonal form is an expression of the form

$$\alpha_1 x_1^n + \alpha_2 x_2^n + \cdots + \alpha_s x_s^n .$$

This is also referred to in the literature as an additive form or simply as a linear combination of nth powers. When the α_i are p-adic numbers, this expression is called a diagonal form in Q_p. If the α_i are all units in O_p, the form is referred to as a unit diagonal form in O_p.

Artin's conjecture for diagonal forms. If a diagonal form in Q_p of degree n contains more than n^2 variables, it must have a non-trivial zero in O_p.

Since a diagonal form is homogeneous, this conjecture is a special case of Artin's original conjecture.

A particularly interesting aspect of this conjecture is that it can be shown to be the best possible. That is, for a given prime p, there exists a diagonal form in Q_p of some degree n which contains n^2 variables, but has only the trivial zero in O_p. The proper degree for such a form is not difficult to guess. As noted in theorem 3.2, since $\varphi(p^{k+1}) = (p-1)p^k$, $\varepsilon^n \equiv 1 \bmod p^{k+1}$ when $n = (p-1)p^k$ and ε is a unit in O_p. For $p=2$ the slightly stronger result, $\varepsilon^n \equiv 1 \bmod 2^{k+2}$ can be stated. This does not improve the result of the next theorem so $p=2$ will not be considered as a special case here. For any non-unit α, it is trivially true that $\alpha^n \equiv 0 \bmod p^{k+1}$. This along with $\varepsilon^n \equiv 1 \bmod p^{k+1}$ for any unit ε implies that the form

$$g_0 = x_1^n + x_2^n + \cdots + x_s^n, \quad s < p^{k+1}$$

has the property that $g_0 \equiv 0 \bmod p^{k+1}$ only when every x_i is a non-unit. Extend this idea to consider $g_0 + p^{k+1}g_1$ where

$$g_1 = y_1^n + y_2^n + \cdots + y_t^n, \quad t < p^{k+1} .$$

It follows that $g_0 + p^{k+1}g_1 \equiv 0 \bmod p^{2(k+1)}$ only if each x_i and each y_i is a non-unit in O_p. This construction suggests that the form

$$g = g_0 + p^{k+1}g_1 + p^{2(k+1)}g_2 + \cdots + p^{q(k+1)}g_q$$

would have a relatively large number of variables and still have only the trivial zero.

The value of q depends on k and must be chosen correctly in order to allow g to contain the maximum number of variables. The following lemma determines the correct choice for q. To get an idea of the relationship between the lemma and g note that when any non-zero p-adic number εp^t is substituted in a form g_r, the contribution to the sum g is a term of the form $\varepsilon^n p^{tn+r(k+1)}$.

Lemma 4.1 Let $q = [n/(k+1)] - 1$ and $s_i = t_i n + r_i(k+1)$ where r_i, t_i, k are integers, $0 \leq r_i \leq q$ and $k \geq 0$; $[x]$ denotes the greatest integer less than or equal to x. Then

(1) $s_i = s_j$ implies $t_i = t_j$ and $r_i = r_j$ and

(2) $s_i \neq s_j$ implies $|s_i - s_j| \geq k + 1$.

Proof: To prove (1), suppose $s_i = s_j$ or $t_i n + r_i(k+1) = t_j n + r_j(k+1)$. This gives $(t_i - t_j)n = (r_j - r_i)(k+1)$ which implies that either $t_i = t_j$ and $r_i = r_j$ or n divides $|r_j - r_i|(k+1)$. However, by definition $|r_j - r_i| \leq q < n/(k+1)$ so $|r_j - r_i|(k+1) < n$. Therefore, n can divide $|r_j - r_i|(k+1)$ only if $|r_j - r_i| = 0$. This gives the desired result $r_i = r_j$, and $t_i = t_j$ follows immediately.

To prove (2) suppose, without loss of generality, $s_i > s_j$ or $t_i n + r_i(k+1) > t_j n + r_j(k+1)$. There are two cases to consider: $t_i = t_j$ and $t_i \neq t_j$. If $t_i = t_j$, then $r_i > r_j$ so that

$$s_i - s_j = (r_i - r_j)(k+1) \geq k+1 .$$

If $t_i \neq t_j$, let $t_i > t_j$. It follows that

$$(t_i - t_j)n \geq n \geq (q+1)(k+1) \geq (r_j - r_i + 1)(k+1)$$

which gives

$$(t_i - t_j)n + (r_i - r_j)(k+1) \geq k+1 .$$

This is the same as $s_i - s_j > k + 1$. A similar argument holds for

$t_i < t_j$ giving $s_j - s_i \geq k+1$. The net result is $|s_i - s_j| \geq k+1$. ▲

The following definitions will be useful in the next theorem. Let

$$G_k = g_0 + p^{k+1}g_1 + p^{2(k+1)}g_2 + \cdots + p^{q(k+1)}g_q \ .$$

Each g_r has the form $x_1^n + x_2^n + \cdots + x_s^n$ where $n = (p-1)p^k$ and $s = p^{k+1} - 1$. The sets of variables contained in the g_r are pairwise disjoint and $q = [n/(k+1)] - 1$. If we denote the number of variables in G_k as N_k, then

$$N_k = (q+1)s = [n/(k+1)](p^{k+1} - 1) \ .$$

<u>Theorem 4.1</u> The form G_k described in the previous paragraph has the following properties:

(1) For any $k \geq 0$, G_k has only the trivial zero.

(2) Given any $\epsilon > 0$, there exist infinitely many k for which $N_k > n^{2-\epsilon}$.

(3) When $p \neq 2$, $N_0 = n^2$.

Proof: To prove (1), suppose the contrary; that is, G_k has a non-trivial zero $(\theta_1, \theta_2, \ldots, \theta_{N_k})$. Then each non-zero θ_i can be written as $\varepsilon_i p^{t_i}$ and

$$G_k(\theta_1, \theta_2, \ldots, \theta_{N_k}) = \sum_{\theta_i \neq 0} \varepsilon_i^n p^{s_i}$$

where $s_i = t_i n + r_i(k+1)$ and $0 \leq r_i \leq q$. Then let s be defined as $\min \{s_i \mid 0 \leq i \leq N_k, \ \theta_i \neq 0\}$ and write

$$G_k = p^s \sum_{\theta_i \neq 0} \varepsilon_i^n p^{s_i - s} .$$

This implies that

$$\sum_{\theta_i \neq 0} \varepsilon_i^n p^{s_i - s} = 0 .$$

In order to analyze this result, let f be the sum of the terms where $s_i = s$ and h be the sum of the terms where $s_i > s$. Thus,

$$\sum_{\theta_i \neq 0} \varepsilon_i^n p^{s_i - s} = f + h = 0 .$$

Now, as a result of lemma 4.1, $s_i = s_j$ implies $r_i = r_j$ so all terms in f are from the same g_r. The second statement in lemma 4.1 shows that $s_i - s \geq k + 1$ when $s_i \neq s$. Therefore, h has a factor of p^{k+1}. Now, since $f + h = 0$, it follows that $f \equiv 0 \bmod p^{k+1}$. However, this is impossible since f contains at most $p^{k+1} - 1$ terms of the form ε_i^n each of which is congruent to 1 mod p^{k+1}. This contradiction completes the proof of (1).

To prove (2) it will suffice to show $\lim_{k \to \infty} \log_n N_k = 2$. This will be accomplished by showing that $\log_n N_k$ is bounded on one side by 2 and on the other side by a function whose limit is 2. The following inequalities are used without proof. Each can be shown to be true when $k > 4$ using elementary methods.

$$\frac{p^k}{k^2} < \frac{p^k}{k+1} - 1 \quad , \quad p^k < p^{k+1} - 1 \quad , \quad p^{\sqrt{k}} > k .$$

First consider the following:

$$N_k = \left[\frac{(p-1)p^k}{k+1}\right](p^{k+1}-1) < \frac{(p-1)p^k}{k+1}(p^{k+1}) = \frac{(p-1)^2 p^{2k} p}{(p-1)(k+1)} \ .$$

So $\quad N_k < \frac{n^2 p}{(p-1)(k+1)} < n^2$. This can be written

$$\log_n N_k < 2 \ . \tag{4.1}$$

To find a function that bounds $\log_n N_k$ on the other side, observe first that

$$\frac{p^k}{k^2} < \frac{p^k}{k+1} - 1 \le \left[\frac{(p-1)p^k}{k+1}\right] .$$

Combining this with $p^k < p^{k+1} - 1$ gives

$$\frac{p^k}{k^2}(p^k) < \left[\frac{(p-1)p^k}{k+1}\right](p^{k+1}-1) = N_k \ .$$

The resulting inequality $p^{2k}/k^2 < N_k$ can be written

$$2k \log_n p - 2 \log_n k < \log_n N_k \ . \tag{4.2}$$

The functions on the left side of (4.2) can be replaced by more familiar functions. The fact that $n = (p-1)p^k$ gives us the inequality $p^k \le n < p^{k+1}$. The right portion $n < p^{k+1}$ can be written

$$\log_n p < \frac{1}{k+1} \ . \tag{4.3}$$

The left portion $p^k \le n$ implies $n^{1/\sqrt{k}} \ge p^{\sqrt{k}} > k$ which implies

$$\frac{1}{\sqrt{k}} > \log_n k \ . \tag{4.4}$$

Now, the four inequalities (4.1), (4.2), (4.3), and (4.4), give

$$\frac{2k}{k+1} - \frac{2}{\sqrt{k}} < 2k \log_n p - 2 \log_n k < \log_n N_k < 2 .$$

From this it follows immediately that $\lim_{k\to\infty} \log_n N_k = 2$.

Statement (3) follows from the definition of N_k since by direct substitution $N_0 = (p-1)^2 = n^2$. ▲

The results of this theorem have some interesting aspects. Conclusion (3) shows that when $p \neq 2$, the n^2 in the conjecture cannot be reduced. Conclusion (2) shows that if n^2 were replaced by n^s where $s < 2$, then for any prime p infinitely many forms can be constructed for which the conjecture is false. It is interesting to note that the number of variables in the construction exceeds $n^{2-\epsilon}$ as k becomes large while the power where G_k actually contains n^2 variables occurs when $k = 0$.

The prime $p = 2$ is conspicious by its absence in conclusion (3). Including 2 here produces the uninteresting conclusion that a form consisting of one first power has only the trivial zero. In order to fill this gap, consider the form g defined by

$$g = x_1^2 + x_2^2 + x_3^2 + x_4^2 .$$

This form has only the trivial zero in O_2. To see this first note that $x_i^2 \equiv 1 \bmod 8$ when x_i is odd and $x_i^2 \equiv 0 \bmod 4$ when x_i is even. The argument is then similar to the Terjanian counterexample, $g \equiv 0 \bmod 8$ only if every x_i is even so any non-trivial zero must contain all even values. However, any such zero would produce another non-trivial zero whose values are not all even which is a contradiction.

Having established that the conjecture is in some sense the best possible, we turn our attention to proving that it is true. Consider what must be accomplished. Given a diagonal form

$$f = \alpha_1 x_1^n + \alpha_2 x_2^n + \cdots + \alpha_s x_s^n,$$

we must show that $s > n^2$ implies that non-trivial zeros of f exist. In general, the α_i are p-adic numbers which have the form εp^t where t could be any integer either positive or negative. If the values of the t's could be limited to relatively small positive integers, f would be easier to work with. The following example demonstrates how this can be done.

Example 4.1 Let f be the following form with coefficients in Q_2.

$$x_1^6 + 2^{-5}x_2^6 + 3\cdot 2^6 x_3^6 + 2^{14}x_4^6 + 5\cdot 2^{-18}x_5^6 + 3\cdot 2^{-11}x_6^6 + 7\cdot 2^2 x_7^6$$

$$+ 2^9 x_8^6 + 17\cdot 2^{-1}x_9^6 .$$

This form can be written as

$$(x_1)^6 + 2(2^{-1}x_2)^6 + 3(2x_3)^6 + 2^2(2^2x_4)^6 + 5(2^{-3}x_5)^6 + 3\cdot 2(2^{-2}x_6)^6$$

$$+ 7\cdot 2^2(x_7)^6 + 2^3(2x_8)^6 + 17\cdot 2^5(2^{-1}x_9)^6 .$$

When y_i is substituted for each of the corresponding expressions in parenthesis f is written

$$y_1^6 + 2y_2^6 + 3y_3^6 + 2^2y_4^6 + 5y_5^6 + 3\cdot 2y_6^6 + 7\cdot 2^2y_7^6 + 2^3y_8^6 + 17\cdot 2^5y_9^6 \ .$$

After this substitution, the powers of 2 in the coefficients are limited to the integers from 0 to 5. Any zero that is found in terms of the y_i will produce a zero in terms of the x_i by simply reversing the substitution. Another transformation that will prove helpful is the following grouping of f.

$$(y_1^6 + 3y_3^6 + 5y_5^6) + 2(y_2^6 + 3y_6^6) + 2^2(y_4^6 + 7y_7^6) + 2^3(y_8^6) + 2^5(17y_9^6) \ .$$

This puts f in the form $f_0 + 2f_1 + 2^2f_2 + 2^3f_3 + 2^4f_4 + 2^5f_5$ where $f_0 = y_1^6 + 3y_3^6 + 5y_5^6$, $f_1 = y_2^6 + 3y_6^6$, $f_2 = y_4^6 + 7y_7^6$, $f_3 = y_8^6$, $f_4 = 0$, and $f_5 = 17y_9^6$.

The significant feature of this grouping is that each f_i has coefficients which are units in O_2; that is, each f_i is a unit diagonal form in O_2. The following theorem formalizes this transformation and shows that it can always be accomplished.

Theorem 4.2 Let f be a diagonal form of degree n with coefficients in Q_p. For the purpose of determining zeros, f can be assumed to be of the form $f_0 + pf_1 + p^2f_2 + \cdots + p^{n-1}f_{n-1}$ where each f_i is either zero or a unit diagonal form in O_p of degree n.

Proof: Let $f = \alpha_1x_1^n + \alpha_2x_2^n + \cdots + \alpha_sx_s^n$ where each α_i is a non-zero p-adic number. Each α_i can be uniquely represented as $\varepsilon_ip^{t_i}$ where ε_i is a unit in O_p and t_i is an integer. Now, for each t_i there exist integers a_i and b_i so that $t_i = a_in + b_i$ and $0 \leq b_i < n$.

These values can be used as follows:

$$\alpha_i x_i^n = \varepsilon_i p^{t_i} x_i^n = \varepsilon_i p^{a_i n + b_i} x_i^n = \varepsilon_i p^{b_i} (p^{a_i} x_i)^n .$$

This indicates that the substitution $y_i = p^{a_i} x_i$ gives

$$f = \varepsilon_1 p^{b_1} y_1^n + \varepsilon_2 p^{b_2} y_2^n + \cdots + \varepsilon_s p^{b_s} y_s^n$$

where $0 \leq b_i \leq n-1$. To obtain the desired representation of f, group the terms by ascending powers of p and factor out the p^{b_i}. In this way f is written as a function of the y_i and if

$$(y_1, y_2, \dots, y_s) = (\theta_1, \theta_2, \dots, \theta_s)$$

is a zero of f, then

$$(x_1, x_2, \dots, x_s) = (\theta_1 p^{-a_1}, \theta_2 p^{-a_2}, \dots, \theta_s p^{-a_s})$$

also is a zero of f. ▲

In theorem 4.2, the zeros of f were not stated to be in O_p. This does not lessen the value of the theorem because any zero of f can be used to produce a zero in O_p by multiplying each component of the zero by one sufficiently large power of p. Having established that this transformation is always possible, it can be assumed, when convenient, that diagonal forms in Q_p have this representation.

To appreciate the advantage of this representation for f, consider example 4.1 again. In this example $f_0 = y_1^6 + 3y_3^6 + 5y_5^6$. Now take $y_1 = 0$, $y_3 = 1$, and $y_5 = 1$; then $f_0(0,1,1) = 8$. In view of the result of theorem 2.11, the fact $6 = 3 \cdot 2$ implies that a 2-adic unit ε

is a 6th power if $\varepsilon \equiv 1 \bmod 8$. Also 3 is a unit in O_2 which means that $1/3$ is a unit in O_2. These facts imply that $1-8/3$ is a 6th power in O_2. Now let $\delta^6 = 1-8/3$ so that $f_0(0,\delta,1) = 0$. The net result is that a non-trivial zero of f_0 has been constructed. If all other variables in f_1, f_2, f_3, f_4, and f_5 are assigned the value zero, a non-trivial zero of f is produced. This zero is in terms of the y_i, but can be written in terms of the x_i by letting $x_3 = \delta/2$, $x_5 = 2^3(1)$ and all other x_i by zero.

The advantages of this representation are further demonstrated by the following theorems.

Theorem 4.3 Let $n = mp^k$, $(m,p) = 1$ and $g = \varepsilon_1 x_1^n + \varepsilon_2 x_2^n + \cdots + \varepsilon_s x_s^n$ where g is a unit diagonal form in O_p. Suppose the congruence $g \equiv 0 \bmod p^{k+1}$ (p^{k+2} when $p = 2$) has a solution $(\theta_1, \theta_2, \ldots, \theta_s)$ in O_p where $\theta_i \not\equiv 0 \bmod p$ for some i. Then g has a non-trivial zero in O_p.

Proof: When $p \neq 2$, $g(\theta_1, \theta_2, \ldots, \theta_s) \equiv 0 \bmod p^{k+1}$ implies that for some α in O_p

$$\varepsilon_1 \theta_1^n + \varepsilon_2 \theta_2^n + \cdots + \varepsilon_s \theta_s^n = \alpha p^{k+1} .$$

Now assume, with no loss of generality, that $\theta_1 \not\equiv 0 \bmod p$ and consider the fact that

$$\varepsilon_1(\theta_1^n - \varepsilon_1^{-1}\alpha p^{k+1}) + \varepsilon_2 \theta_2^n + \cdots + \varepsilon_s \theta_s^n = 0 .$$

Using the criteria established in theorem 2.11, $\theta_1^n - \varepsilon_1^{-1}\alpha p^{k+1}$ is an nth power in O_p because it is congruent to the nth power $\theta_1^n \bmod p^{k+1}$.

Therefore, there is a δ in O_p such that $\delta^n = \theta_1^n - \varepsilon_1^{-1}\alpha p^{k+1}$. It follows that $(\delta, \theta_2, \theta_3, \ldots, \theta_s)$ is a non-trivial zero of g. The argument for $p = 2$ is exactly the same except p^{k+1} is replaced by 2^{k+2} in the appropriate places. ▲

Theorem 4.4 Suppose every unit diagonal form in O_p of degree n with more than s variables has a non-trivial zero in O_p. Then every diagonal form in Q_p of degree n with more than ns variables has a non-trivial zero in O_p.

Proof: Let $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_p of degree n. If f has more than ns variables, some f_i must have more than s variables. Let f_r denote an f_i with more than s variables. By hypothesis, f_r must have a non-trivial zero in O_p. Assigning the values from this non-trivial zero of f_r and zero to each of the variables in $f_0, f_1, \ldots, f_{r-1}, f_{r+1}, \ldots, f_s$ produces a non-trivial zero of f. ▲

These three theorems provide a method for finding non-trivial zeros for diagonal forms in Q_p. Given a diagonal form f of degree n first write f in the form $f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$. Next, find an i for which $f_i \equiv 0 \bmod p^{k+1}$ (p^{k+2} if $p = 2$) has a solution in O_p. The k is determined by $n = mp^k$, $(m,p) = 1$ and at least one value in the solution must be a unit in O_p. Using theorem 4.3, this solution produces a non-trivial zero of f_i. As indicated in theorem 4.4, a non-trivial zero of any f_i will produce a non-trivial zero of f.

The theorems also provide a method for proving the conjecture in some important special cases. The first cases investigated are diagonal forms in Q_p of degree m where $(m,p) = 1$ and diagonal forms of degree p^k, $p \neq 2$. In view of the nature of nth powers as established in Chapter II, these seem like natural cases to consider. Solving the form of degree m involves a congruence mod p while the form of degree p^k requires a congruence mod p^{k+1}. In view of this, it seems surprising that the conjecture is easier to prove in the latter case. However, in the course of the investigation of forms of degree $m, (m,p) = 1$, several results are demonstrated which are necessary in the later work. The next few theorems provide a method for proving the conjecture for diagonal forms in Q_p of degree $m, (m,p) = 1$.

<u>Theorem 4.5</u> (Lagrange's theorem) Let p be a prime and $f(x)$ be a polynomial of degree n whose coefficients are integers. The congruence $f(x) \equiv 0 \bmod p$ has at most n incongruent solutions mod p unless each coefficient of $f(x)$ is congruent to zero mod p.

Proof: See Niven and Zuckerman (11, p. 44). ▲

<u>Theorem 4.6</u> Let $f(x_1, x_2, \ldots, x_s)$ be a polynomial of degree less than p in each x_i. Suppose $f(\theta_1, \theta_2, \ldots, \theta_s) \equiv 0 \bmod p$ for every $(\theta_1, \theta_2, \ldots, \theta_s)$ where $\theta_i = 0, 1, 2, \ldots, p-1$ for each i. Then the coefficients of f must all be congruent to zero mod p.

Proof: Consider f as a polynomial f_1 in x_1 having coefficients which are polynomials in $x_2, x_3, \ldots, x_s$. This polynomial $f_1(x_1)$ has degree less than p and $f_1(x_1) \equiv 0 \bmod p$ has p incongruent solutions mod p. Therefore, Lagrange's theorem implies that each of its coefficients, the polynomials in $x_2, x_3, \ldots, x_s$, must be congruent to zero mod p for every $(\theta_2, \theta_3, \ldots, \theta_s)$. Now each of these coefficient polynomials can be considered as a polynomial in x_2 whose coefficients are polynomials in $x_3,x_4, \dots ,x_s$. Again, from Lagrange's theorem, the coefficient polynomials in $x_3,x_4, \dots ,x_s$ must all be congruent to zero mod p for every $(\theta_3,\theta_4, \dots ,\theta_s)$. This process can be repeated until the step where polynomials in x_s are obtained whose coefficients are the original integer coefficients of f. These polynomials are of degree less than p in x_s and are all congruent to zero mod p for $x_s = 0,1, \dots ,p-1$. Therefore, their coefficients and, hence, all coefficients of f must be congruent to zero mod p. ▲

An example will help to clarify the argument in theorem 4.6.

<u>Example 4.2</u> Let $f(x,y) = a_1x^4y^3 + a_2x^4y^2 + a_3x^2y^2 + a_4x^2y + a_5y^4 + a_6$. Suppose $f(\theta_1,\theta_2) \equiv 0 \bmod 5$ for every (θ_1,θ_2). Write

$$f(x,y) = (a_1y^3 + a_2y^2)x^4 + (a_3y^2 + a_4y)x^2 + (a_5y^4 + a_6) .$$

Then, f can be considered as a polynomial in x whose coefficients are polynomials in y. Now consider a fixed value of θ_2 for y so $f(\theta_1,\theta_2) \equiv 0 \bmod 5$ for $\theta_1 = 0,1,2,3,4$. This implies that each coefficient polynomial must be congruent to zero mod 5 for the fixed θ_2. We can use this argument for 5 different values of θ_2, so we have $a_1y^3 + a_2y^2$, $a_3y^2 + a_4y$, and $a_5y^4 + a_6$ all congruent to zero mod 5 for 5 incongruent values of y. The conclusion follows that each a_i, the original coefficients of f, must be congruent to zero mod 5.

An important consequence of this theorem is the following. Let $f(x_1,x_2, \dots ,x_s) \equiv g(x_1,x_2, \dots ,x_s) \bmod p$ where f and g are both of degree less than p for each x_i. Then if the congruence holds for every $(\theta_1,\theta_2, \ldots ,\theta_s)$, the polynomials f and g must be identical mod p. That is, all corresponding coefficients must be congruent mod p.

Theorem 4.7 (Chevalley's theorem) Let $f(x_1,x_2, \ldots ,x_s)$ be a polynomial of degree less than s with integral coefficients and whose constant term is zero. Then $f \equiv 0 \bmod p$ has a solution $(\theta_1,\theta_2, \ldots ,\theta_s)$ where $\theta_i \not\equiv 0 \bmod p$ for some i.

Proof: By Fermat's theorem $x^p \equiv x \bmod p$ so each exponent in each term of f can be reduced to one of the values $0,1, \ldots ,p-1$ without affecting the solution set of the congruence. Therefore, f can be assumed to have degree less than p in each x_i. In order to prove the theorem, assume the contrary, for every $(\theta_1,\theta_2, \ldots ,\theta_2)$ where $\theta_i \not\equiv 0 \bmod p$ for some i, $f(\theta_1,\theta_2, \ldots ,\theta_s) \not\equiv 0 \bmod p$. With this assumption in mind, consider the following congruence:

$$1-[f(x_1,x_2, \ldots ,x_s)]^{p-1} \equiv (1-x_1^{p-1})(1-x_2^{p-1}) \ldots (1-x_s^{p-1}) \bmod p .$$

This congruence holds for all $(\theta_1,\theta_2, \ldots ,\theta_s)$. To see this, first let every θ_i be congruent to zero mod p. By hypothesis, f has no constant term so $f(\theta_1,\theta_2, \ldots ,\theta_s) \equiv 0 \bmod p$ and it follows that both sides are congruent to one mod p. Now consider the other possibility that some θ_i is not congruent to zero mod p. By assumption $f(\theta_1,\theta_2, \ldots ,\theta_s) \not\equiv 0 \bmod p$ so the left side is congruent to zero mod p. The right side is also congruent to zero mod p since one of its factors is congruent to zero mod p. So the congruence holds for all $(\theta_1,\theta_2, \ldots ,\theta_s)$. Applying the previous theorem, the polynomials on the right and left sides must be identical mod p. However, when the right side is expanded, it contains the term

$$(-1)^s x_1^{p-1} x_2^{p-1} \cdots x_s^{p-1}$$

which is of degree $s(p-1)$. On the left side, the fact that f has degree less than s means that no term can have degree as great as $s(p-1)$. This contradiction completes the proof of the theorem. ▲

This theorem and the previous one are also true if the coefficients of the given polynomials are p-adic integers. To see this for Chevalley's theorem, let α be any p-adic integer that is a coefficient of f. If $\alpha = \sum_{i=0}^{\infty} a_i p^i$, then $\alpha \equiv a_0 \bmod p$. Now, if each p-adic coefficient is replaced by its corresponding a_0 a new polynomial, say $f_1(x_1, x_2, \ldots, x_s)$, with integer coefficients is produced. Chevalley's theorem gives a solution $(\theta_1, \theta_2, \ldots, \theta_s)$ in integers for $f_1(x_1, x_2, \ldots, x_s) \equiv 0 \bmod p$ where for some i, $\theta_i \not\equiv 0 \bmod p$. If each a_0 coefficient in f_1 is now replaced by the corresponding p-adic value the congruence still holds since only a multiple of p is added. Therefore, $(\theta_1, \theta_2, \ldots, \theta_s)$ is a solution for $f \equiv 0 \bmod p$. The θ_i that is not congruent to zero $\bmod p$ can be considered as a unit in O_p.

Chevalley's theorem is strong enough to use in proving the conjecture for mth powers where $(m,p) = 1$. However, a stronger result can be established for diagonal forms and since this result will be needed later, the following theorem is included here.

<u>Theorem 4.8</u> Let $g = \varepsilon_1 x_1^n + \varepsilon_2 x_2^n + \cdots + \varepsilon_{d+1} x_{d+1}^n$ be a unit diagonal form in O_p where $n = mp^k$, $(m,p) = 1$, and $d = (m, p-1)$. Then, the congruence $g \equiv 0 \bmod p$ has a solution in integers where $x_1 = 1$.

Proof: Recall from theorem 2.12 that the congruence $\varepsilon x^n \equiv a \bmod p$ has a solution in integers if and only if the congruence $\varepsilon x^d \equiv a \bmod p$ has a solution in integers. Therefore, the congruence

$$\varepsilon_1 x_1^d + \varepsilon_2 x_2^d + \cdots + \varepsilon_{d+1} x_{d+1}^d \equiv 0 \bmod p$$

has a solution in integers with $x_1 = 1$ if and only if $g \equiv 0 \bmod p$ has a solution in integers with $x_1 = 1$. In order to prove the theorem, assume the contrary, for every choice of integers $x_2, x_3, \ldots, x_{d+1}$, it is true that

$$\varepsilon_1 + \varepsilon_2 x_2^d + \cdots + \varepsilon_{d+1} x_{d+1}^d \not\equiv 0 \bmod p \ .$$

It then follows from Fermat's theorem that the congruence

$$(\varepsilon_1 + \varepsilon_2 x_2^d + \cdots + \varepsilon_{d+1} x_{d+1}^d)^{p-1} - 1 \equiv 0 \bmod p$$

holds for every choice of $x_2, x_3, \ldots, x_{d+1}$. When the left side is expanded and each exponent is reduced to a value less than p, we can apply theorem 4.6 and conclude that each coefficient must be congruent to zero mod p. This expansion can be accomplished using the multinomial formula with the following result:

$$\sum \frac{(p-1)!}{a_1! a_2! \ldots a_{d+1}!} (\varepsilon_1)^{a_1} (\varepsilon_2 x_2^d)^{a_2} \cdots (\varepsilon_{d+1} x_{d+1}^d)^{a_{d+1}} - 1 \ .$$

The sum is taken over all combinations for which $0 \leq a_i \leq p-1$ and $a_1 + a_2 + \cdots + a_{d+1} = p-1$. Each of these combinations occurs in exactly one of the following cases:

(1) $a_1 = 0$ and $a_i = (p-1)/d$ for all i, $2 \leq i < d+1$.

(2) $a_1 = 0$ and $0 \leq a_i < (p-1)/d$ for some i, $2 \leq i \leq d+1$.

(3) $a_1 > 0$.

In case (1), the resulting term is

$$\left[\left(\frac{p-1}{d}\right)!\right]^{-d} (p-1)!(\varepsilon_2\varepsilon_3 \cdots \varepsilon_{d+1})^{\frac{p-1}{d}} x_2^{p-1}x_3^{p-1} \cdots x_{d+1}^{p-1}. \quad (4.5)$$

Note that the coefficient of this term is not congruent to zero mod p. Each term in case (2) has at least one exponent that is less than $p-1$ since $0 \leq da_i < p-1$ for some i and da_i is the exponent of x_i. This means that none of the terms from case (2) can be combined with the term from case (1). In case (3), the fact that $a_1 > 0$ means that $a_2 + a_3 + \cdots + a_{d+1} = p-1-a_1 < p-1$. The sum has d terms so for some i, $0 \leq a_i < (p-1)/d$. Therefore, as in case (2), each term in case (3) has at least one exponent that is less than $p-1$. Since none of the terms from cases (2) or (3) combine with the term (4.5) above, this term must occur in the sum exactly as written. However, its coefficient is not congruent to zero mod p which is a contradiction to theorem 4.6. Hence, the assumption that

$$\varepsilon_1 + \varepsilon_2 x_2^d + \cdots + \varepsilon_{d+1} x_{d+1}^d \equiv 0 \bmod p$$

had no solution must be false. It follows that a solution exists for $g \equiv 0 \bmod p$ with $x_1 = 1$. ▲

We now return to the task of proving the conjecture that diagonal forms in Q_p of degree n containing more than n^2 variables have non-trivial zeros in O_p. As a result of theorem 4.4, we need only show that any unit diagonal form in O_p of degree n with more than n variables has a non-trivial zero in O_p.

Theorem 4.9 Let $g = \varepsilon_1 x_1^m + \varepsilon_2 x_2^m + \cdots + \varepsilon_s x_s^m$ be a unit diagonal form in O_p where $s > m$ and $(m,p) = 1$. Then g has a non-trivial zero in O_p.

Proof: Chevalley's theorem implies that the congruence $g \equiv 0 \bmod p$ has an integral solution $(\theta_1, \theta_2, \cdots, \theta_s)$ where $\theta_i \not\equiv 0 \bmod p$ for some i. Without loss of generality, suppose $\theta_1 \not\equiv 0 \bmod p$. Then

$$\varepsilon_1 \theta_1^m + \varepsilon_2 \theta_2^m + \cdots + \varepsilon_s \theta_s^m = \alpha p$$

for some α in O_p. This can be written as

$$\varepsilon_1 (\theta_1^m - \varepsilon_1^{-1} \alpha p) + \varepsilon_2 \theta_2^m + \cdots + \varepsilon_s \theta_s^m = 0 \ .$$

Now from theorem 2.11, $\theta_1^m - \varepsilon_1^{-1} \alpha p$ is an mth power in O_p so there is a δ in O_p for which $\delta^m = \theta_1^m - \varepsilon_1^{-1} \alpha p$. It follows that $(\delta, \theta_2, \theta_3, \cdots, \theta_s)$ is a non-trivial zero of g. ▲

Note that this theorem can be strengthened using theorem 4.8. We can use $s > d$ where $d = (m, p-1)$ instead of $s > m$. The two conditions are the same only when $m = p-1$. This is similar to the result of theorem 4.1 in which $p-1$ was the degree of the form containing n^2 variables and having only the trivial zero.

Theorem 4.10 Let $g = \varepsilon_1 x_1^{p^k} + \varepsilon_2 x_2^{p^k} + \cdots + \varepsilon_s x_s^{p^k}$ be a unit diagonal form in O_p where $p \neq 2$ and $s > p^k$. Then g has a non-trivial zero in O_p.

Proof: Consider the set $\mathcal{S}$ of integers defined by

$$\mathcal{S} = \{a \mid a \equiv \varepsilon_i^{p-1} \bmod p^{k+1},\ 0 < a < p^{k+1},\ 1 \leq i \leq s\} \ .$$

Since each ε_i is a unit in O_p, we have $\varepsilon_i^{p-1} \equiv 1 \bmod p$ for every i. This implies that $\mathcal{S}$ contains at most p^k elements since there are exactly p^k integers between 1 and p^{k+1} which are congruent to 1 mod p. By hypothesis, $s > p^k$ so for some $i \neq j$ we must have

$$\varepsilon_i^{p-1} \equiv \varepsilon_j^{p-1} \bmod p^{k+1} .$$

Without loss of generality, let $\varepsilon_i = \varepsilon_1$ and $\varepsilon_j = \varepsilon_2$. Raising each side of the congruence to the power $(p^{k-1} + p^{k-2} + \cdots + p + 1)$ and using the identity $(p-1)(p^{k-1} + p^{k-2} + \cdots + 1) = p^k - 1$ gives

$$\varepsilon_1^{p^k-1} \equiv \varepsilon_2^{p^k-1} \bmod p^{k+1} \quad \text{or} \quad \varepsilon_1\varepsilon_2^{p^k} \equiv \varepsilon_2\varepsilon_1^{p^k} \bmod p^{k+1} .$$

This implies that for some α in O_p

$$\varepsilon_1\varepsilon_2^{p^k} - \varepsilon_2\varepsilon_1^{p^k} = \alpha p^{k+1} \quad \text{or} \quad \varepsilon_1(\varepsilon_2^{p^k} - \varepsilon_1^{-1}\alpha p^{k+1}) + \varepsilon_2(-\varepsilon_1)^{p^k} = 0 .$$

Now $\varepsilon_2^{p^k} - \varepsilon_1^{-1}\alpha p^{k+1}$ is a p^kth power in O_p so there exists a δ in O_p for which

$$\delta^{p^k} = \varepsilon_2^{p^k} - \varepsilon_1^{-1}\alpha p^{k+1} .$$

It follows that $(\delta, -\varepsilon_1, 0, \ldots, 0)$ is a non-trivial zero of g. ▲

<u>Corollary 4.1</u> Let f be a diagonal form in Q_p of degree n and containing more than n^2 variables. If $n = m$, $(m,p) = 1$ or if $n = p^k$, $p \neq 2$, then f has a non-trivial zero in O_p.

Proof: The proof follows directly from theorems 4.4, 4.9, and 4.10. ▲

In theorem 4.10, the result depended strongly on the fact that when

$p \neq 2$, p^k is odd. The next theorem shows that the restriction to odd powers produces a result that is often considerably better than the conjecture suggests. Note that $(p-1)p^k$ is excluded by this restriction when p is odd.

<u>Theorem 4.11</u> Let $g = \varepsilon_1 x_1^n + \varepsilon_2 x_2^n + \cdots + \varepsilon_s x_s^n$ be a unit diagonal form in O_p of odd degree $n = mp^k, (m,p) = 1$, $p \neq 2$. Then if $s > (k+1)\log_2 p$, g must have a non-trivial zero in O_p.

Proof: Consider the set $\mathcal{S}$ defined by

$$\mathcal{S} = \{a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots a_s\varepsilon_s \mid a_i = 1 \text{ or } 0,\ 1 \leq i \leq s\} .$$

This set $\mathcal{S}$ contains at most 2^s elements. Now, since $s > (k+1)\log_2 p$ implies $2^s > p^{k+1}$, there must be at least two elements of $\mathcal{S}$ which are congruent mod p^{k+1}. This gives

$$a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_s\varepsilon_s \equiv a_1'\varepsilon_1 + a_2'\varepsilon_2 + \cdots + a_s'\varepsilon_s \bmod p^{k+1}$$

where each a_i and a_i' is either 1 or 0 and for at least one i, $a_i \neq a_i'$. This result can be written as

$$(a_1 - a_1')\varepsilon_1 + (a_2 - a_2')\varepsilon_2 + \cdots + (a_s + a_s')\varepsilon_s = \alpha p^{k+1}$$

for some α in O_p where each $a_i - a_i'$ is either 1, -1, or 0 and at least one $a_i - a_i'$ is not zero. Since n is odd and each $a_i - a_i'$ has one of the values 1, -1, and 0, we have $(a_i - a_i')^n = a_i - a_i'$ for every i. Assume, without loss of generality, that $a_1 - a_1'$ is not zero. Then $(a_1 - a_1')^n - \varepsilon_1^{-1}\alpha p^{k+1}$ is an nth power in O_p. Therefore, for some δ in O_p, $\delta^n = (a - a')^n - \varepsilon_1^{-1}\alpha p^{k+1}$ and it follows that

$(\delta,\ a_2 - a_2',\ \dots\ ,a_s - a_s')$ is a non-trivial zero of g. ▲

It is interesting to see the conditions under which this result improves on the conjecture and how much improvement is made. Combining the result of theorem 4.4 with the result of this theorem we have that any diagonal form in Q_p of odd degree $n = mp^k$ and more than $n(k+1)\log_2 p$ variables must have a non-trivial zero in O_p. This value $n(k+1)\log_2 p$ will be less than the n^2 in the conjecture whenever $(k+1)\log_2 p < n$. If $k = 0$, this compares m and $\log_2 p$ and $m > \log_2 p$ provided $p < 2^m$. If $k > 0$ it is not difficult to show that $n > (k+1)\log_2 p$ except when $m = 1$, $p = 3$, and $k = 1$. In general, when $k > 0$ the comparison of n and $(k+1)\log_2 p$ is similar to the comparison of n and $\log_2 n$ which are significantly different, especially for large n.

In theorem 4.10 diagonal forms in O_2 of degree 2^k were excluded because the method of proof required that p^k be odd for every k. This case is more difficult because 2^k is even. However, as we will show later, the proof of the conjecture for diagonal forms in Q_2 of degree 2^k provides a proof for all diagonal forms in Q_2. Also, the methods that are devised for proving the conjecture in this case suggest methods for proving the conjecture for odd primes. We begin devising these methods by considering the following example.

<u>Example 4.3</u> Let $f = f_0 + 2f_1 + 4f_2 + 8f_3$ where

$$f_0 = x_1^4 + 9x_2^4 + 17x_3^4 + 25x_4^4 + 33x_5^4 \quad , \quad f_1 = x_6^4 + 5x_7^4$$

$$f_2 = x_8^4 + 9x_9^4, \quad \text{and} \quad f_3 = x_{10}^4 \ .$$

Note that f_0 has more than 4 terms, but $f_0 \equiv 0 \bmod 16$ only if $x_i \equiv 0 \bmod 2$ for each i, $1 \leq i \leq 5$. Therefore, f_0 cannot be used to construct a zero of f as has been the case in previous examples and theorems. However, if $x_i = 1$ for $i = 1,2,7,8,10$ and $x_i = 0$ for $i = 3,4,5,6,9$ the result is $f_0 = 10$, $2f_1 = 10$, $4f_2 = 4$, and $8f_3 = 8$. This gives $f = 32$. Now, since $1 - 32 \equiv 1 \bmod 2^4$, -31 is a 2^2th power in O_2. Therefore, for some δ in O_2, $\delta^4 = -31$ and when $x_1 = \delta$, $x_i = 1$ for $i = 2,7,8,10$ and $x_i = 0$ for $i = 3,4,5,6,9$, this is a non-trivial zero of f. Another non-trivial zero could be constructed using x_2 instead of x_1. Since 9 is a unit in O_2, $1/9$ also is a unit in O_2 and $1 - 32/9 \equiv 1 \bmod 2^4$. Therefore, $1 - 32/9 = -23/9$ is a 4th power in O_2 and if $\delta^4 = -23/9$, $x_2 = \delta$, $x_i = 1$ for $i = 1,7,8,10$ and $x_i = 0$ for $i = 3,4,5,6,9$ is a non-trivial zero of f. It is important to note that the x_i from f_1, f_2, and f_3 cannot be used as the x_1 and x_2 were used to construct a non-trivial zero because as variables in f their coefficients are not units in O_2.

This example demonstrates a method that will be used in constructing all non-trivial zeros in the remainder of this chapter. First a solution in integers to an appropriate congruence will be constructed. Then, this solution will be used to construct a zero of f. This method can succeed only if the integral solution assigns a value to some variable in f_0 which is not congruent to zero mod p. If f_0 does not contain a sufficient number of variables the method fails. The following example shows how such a problem can be overcome.

Example 4.4 Let f be the form

$$(x_1^4 + 9x_2^4) + 2(x_3^4) + 4(x_4^4 + 9x_5^4 + 17x_6^4 + 25x_7^4 + 33x_8^4) + 8(x_9^4 + 5x_{10}^4) .$$

Suppose the x_i are replaced by $2y_i$ for $i = 1,2,3$. The first three terms of f then become $16(y_1^4 + 9y_2^4) + 32(y_3^4)$. After rearranging the terms, f can be written as

$$4[(x_4^4 + 9x_5^4 + 17x_6^4 + 25x_7^4 + 33x_8^4) + 2(x_9^4 + 5x_{10}^4) + 4(y_1^4 + 9y_2^4) + 8(y_3^4)] .$$

This form is essentially 4 times the form in example 4.3. The non-trivial zero of that form, adapted to the proper variables here, gives $x_4 = \delta$, $x_5 = x_{10} = y_1 = y_3 = 1$ and $x_6 = x_7 = x_8 = x_9 = y_2 = 0$. This produces a non-trivial zero of f in terms of x_i when $x_1 = 2y_1 = 2$, $x_2 = 2y_2 = 0$, and $x_3 = 2$.

The following theorem uses this type of substitution to effect a cyclic permutation of the f_i so that any f_i can be placed in the first position.

<u>Theorem 4.12</u> Let $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_p of degree n. Then f has a non-trivial zero in O_p if and only if the form

$$g_r = f_r + pf_{r+1} + \cdots + p^{n-1-r}f_{n-1} + p^{n-r}f_0 + \cdots + p^{n-1}f_{r-1}$$

$0 \le r \le n-1$, has a non-trivial zero in O_p.

Proof: Let $h_1(x_1,x_2, \ldots ,x_t) = f_0 + pf_1 + \cdots + p^{r-1}f_{r-1}$ and $h_2 = p^r f_r + p^{r+1}f_{r+1} + \cdots + p^{n-1}f_{n-1}$. It follows that $f = h_1 + h_2$ and $g_r = p^{n-r}h_1 + p^{-r}h_2$. If $\alpha_1x_1^n + \alpha_2x_2^n + \cdots + \alpha_tx_t^n$ represents h_1, then by direct substitution

$$h_1(px_1,px_2, \ldots ,px_t) = p^n h_1(x_2,x_2, \ldots ,x_t) .$$

A similar statement can be made about h_2. Now suppose $(\theta_1,\theta_2, \ldots ,\theta_s)$ is a non-trivial zero of f. It follows that

$$g_r(\theta_1,\theta_2, \ldots ,\theta_t,p\theta_{t+1}, \ldots ,p\theta_s)$$

$$= p^{n-r}h_1(\theta_1,\theta_2, \ldots ,\theta_t) + p^{-r}h_2(p\theta_{t+1},p\theta_{t+2}, \ldots ,p\theta_s)$$

$$= p^{n-r}h_1(\theta_1,\theta_2, \ldots ,\theta_t) + p^{n-r}h_2(\theta_{t+1},\theta_{t+2}, \ldots ,\theta_s)$$

$$= p^{n-r}[h_1(\theta_1,\theta_2, \ldots ,\theta_t) + h_2(\theta_{t+1},\theta_{t+2}, \ldots ,\theta_s)]$$

$$= p^{n-r}f(\theta_1,\theta_2, \ldots ,\theta_t) = 0 .$$

Therefore, $(\theta_1,\theta_2, \ldots ,\theta_t,p\theta_{t+1}, \ldots ,p\theta_s)$ is a non-trivial zero of g_r. Now suppose $(\theta_1,\theta_2, \ldots ,\theta_s)$ is a non-trivial zero of g_r. In this case $(p\theta_1,p\theta_2, \ldots ,p\theta_t,\theta_{t+1}, \ldots ,\theta_s)$ is shown to be a non-trivial zero of f by the following:

$$f(p\theta_1,p\theta_2, \ldots ,p\theta_t,\theta_{t+1}, \ldots ,\theta_s)$$

$$= h_1(p\theta_1,p\theta_2, \ldots ,p\theta_t) + h_2(\theta_{t+1},\theta_{t+2}, \ldots ,\theta_s)$$

$$= p^n h_1(\theta_1,\theta_2, \ldots ,\theta_t) + h_2(\theta_{t+1},\theta_{t+2}, \ldots ,\theta_s)$$

$$= p^r[p^{n-r}h_1(\theta_1,\theta_2, \ldots ,\theta_t) + p^{-r}h_2(\theta_{t+1},\theta_{t+2}, \ldots ,\theta_s)]$$

$$= p^r g_r(\theta_1,\theta_2, \ldots ,\theta_s) = 0 .$$ ▲

In previous problems, we have used the fact that when f contains more than s variables, the average number of variables in the f_i is more than s/n. It was natural to observe that this implies some f_i must contain more than s/n variables. As a result of theorem 4.12, we may assume without loss of generality that f_0 itself contains more than s/n variables. The observation about the f_i can be extended to consider all consecutive pairs $\{f_i, f_{i+1}\}$. Since there are n pairs and each variable is included in exactly 2 pairs the average number of variables contained in each pair is greater than $2s/n$. Therefore, some pair must contain more than $2s/n$ variables. Continuing this concept we can consider all sets $\{f_i, f_{i+1}, \ldots, f_{i+j-1}\}$ of j consecutive forms. Some such set should contain more than js/n variables. Theorem 4.13 from a paper by Lewis and Davenport (8) shows that such a result is not only possible but that an even stronger result can be obtained.

Theorem 4.13 Let $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ be a diagonal form with s variables where each f_i is a unit diagonal form in O_p of degree n. Then there exists a diagonal form $g = g_0 + pg_1 + \cdots + p^{n-1}g_{n-1}$ with the following properties:

(1) g has a non-trivial zero in O_p if and only if f has a non-trivial zero in O_p.

(2) If M_i denotes the number of variables in g_i then

$$M_0 \geq \frac{s}{n},\ M_0 + M_1 \geq \frac{2s}{n},\ \ldots,\ M_0 + M_1 + \cdots + M_{n-2} \geq \frac{(n-1)s}{n},$$

and $M_0 + M_1 + \cdots + M_{n-1} = s.$

Proof: To prove this theorem, it will be shown that there exists an r for which g_r in theorem 4.12 has property (2). Denote the number of variables in each f_i as N_i and consider the infinite periodic sequence $\{N_i\}$ where $N_i = N_{n+i}$. In this sequence, any segment $\{N_t, N_{t+1}, \ldots, N_{t+n-1}\}$ has the property that the N_i denote exactly the number of variables in the unit diagonal forms in g_r where

$r \equiv t \bmod n$ and $0 \leq r \leq n-1$. The proof of the theorem will be complete when the existence of an r is demonstrated for which

$$N_r + N_{r+1} + \cdots + N_{r+t-1} \geq \frac{ts}{n}$$

for all t, $1 \leq t \leq n$. To prove that such an r exists, assume the contrary that for every r there exists a t, $1 \leq t \leq n$, for which

$$N_r + N_{r+1} + \cdots + N_{r+t-1} < \frac{ts}{n} .$$

First define $u_i = N_i - s/n$. By the definitions of N_i and u_i, $u_0 + u_1 + \cdots + u_{n-1} = 0$. By assumption, for every integer a there exists another integer $b > a$ for which

$$N_a + N_{a+1} + \cdots + N_b < \frac{(b-a+1)s}{n} .$$

This can be written as

$$(N_a - \frac{s}{n}) + (N_{a+1} - \frac{s}{n}) + \cdots + (N_b - \frac{s}{n}) < 0$$

and in terms of the u_i as $u_a + u_{a+1} + \cdots + u_b < 0$. Now consider the following sequence of ordered pairs. Let a_1 be any integer and determine b_1 so that

$$u_{a_1} + u_{a_1+1} + \cdots + u_{b_1} < 0 .$$

Now let $a_2 = b_1 + 1$ and determine b_2 so that

$$u_{a_2} + u_{a_2+1} + \cdots + u_{b_2} < 0 .$$

Continuing this process, there must eventually be some $i < j$ such that $a_i \equiv a_j \bmod n$. This result allows us to establish a contradiction. Consider the sum

$$u_{a_i} + u_{a_i+1} + \cdots + u_{a_j-1} .$$

This sum can be considered as segments of the form

$$u_{a_s} + u_{a_s} + \cdots + u_{b_s}, \quad i \leq s \leq j-1 .$$

The sum of each segment is less than zero so the entire sum must be less than zero. This sum can also be considered as segments, each of length n, and each containing a complete set of the original u_i. Each of these segments has the sum zero so the entire sum must be zero. This contradiction proves that the assumption is false and completes the proof ot the theorem. ▲

The two previous theorems have applications for odd primes as well as for $p = 2$. The next section concentrates on proving the conjecture in the 2-adic case. Recall from example 4.3 that the zero of f was obtained by first assigning values of one or zero to each x_i. The result of this assignment was $f_0 = 10$, $2f_1 = 10$, $4f_2 = 4$, and $8f_3 = 8$. Now consider the progressive sums $f_0 = 10 = 5 \cdot 2$, $f_0 + 2f_1 = 20 = 5 \cdot 2^2$, $f_0 + 2f_1 + 4f_2 = 24 = 3 \cdot 2^3$, and $f_0 + 2f_1 + 4f_2 + 8f_3 = 32 = 2^5$. An important observation is that the value of each sum contains at least one more factor of 2 than the value of the previous sum. This is necessary in order to obtain the result of $f \equiv 0 \bmod 2^4$ using variables from f_0. A second observation that can be made from this example is that when an x_i is assigned the value one this has the effect of picking the coefficient of the x_i to be retained in the sum while assigning the value zero to an x_i has the effect of deleting its coefficient from the sum. In constructing a zero for f we will be concerned with obtaining a sum whose value contains a certain power of 2. This construction will be accomplished by retaining or deleting the coefficients of f. We begin with the following definition. The use of β and ε in the definition is consistent with the usual convention of letting β denote a p-adic integer and ε denote a unit in O_2.

Definition 4.2 A 2-adic integer that is divisible by 2^i will be said to be of the $2^i\beta$ type. A 2-adic integer of the $2^i\beta$ type that is not of the $2^{i+1}\beta$ type will be said to be of the $2^i\varepsilon$ type.

Lemma 4.2 Given 2^i terms of the $2^j\beta$ type one can construct 2^{i-n} terms of the $2^{j+n}\beta$ type where $0 \leq n \leq i$.

Proof: First partition the 2^i terms of the $2^j\beta$ type into 2^{i-1} pairs. If $\{2^j\beta_1, 2^j\beta_2\}$ is one such pair, then β_1, β_2, or $\beta_1 + \beta_2$ is even. Therefore, $2^j\beta_1, 2^j\beta_2$, or $2^j(\beta_1 + \beta_2)$ is of the $2^{j+1}\beta$ type. It follows that from the 2^{i-1} pairs 2^{i-1} terms of the $2^{j+1}\beta$ type can be constructed. Similarly from 2^{i-2} pairs of these $2^{j+1}\beta$ terms 2^{i-2} terms of the $2^{j+2}\beta$ type can be constructed. Proceeding in this fashion in general 2^{i-n} terms of the $2^{j+n}\beta$ type can be constructed. ▲

Lemma 4.3 Let $2^j\varepsilon_1$ and $2^j\varepsilon_2$ be terms of the $2^j\varepsilon$ type with $\varepsilon_1 \equiv \varepsilon_2 \bmod 4$. Then, $2^j(\varepsilon_1 + \varepsilon_2)$ is of the $2^{j+1}\varepsilon$ type.

Proof: Since ε_1 and ε_2 are both congruent to one mod 2 either $\varepsilon_1 \equiv \varepsilon_2 \equiv 1 \bmod 4$ or $\varepsilon_1 \equiv \varepsilon_2 \equiv 3 \bmod 4$. In either case,

$\varepsilon_1 + \varepsilon_2 \equiv 2 \bmod 4$. Therefore, $\varepsilon_1 + \varepsilon_2$ is of the 2β type, but not of the $2^2\beta$ type. It follows that $\varepsilon_1 + \varepsilon_2$ is of the 2ε type and that $2^j(\varepsilon_1 + \varepsilon_2)$ is of the $2^{j+1}\varepsilon$ type. ▲

Lemma 4.4 Given 2^i terms of the $2^j\varepsilon$ type, one can construct $2^{i-n} - 1$ terms of the $2^{j+n}\varepsilon$ type where $0 \leq n \leq i$.

Proof: First partition the 2^i terms into two sets, one set containing the $2^j\varepsilon$ terms where $\varepsilon \equiv 1 \bmod 4$ and the other set containing those where $\varepsilon \equiv 3 \bmod 4$. Now each set can be partitioned into pairs with at most one term in each set left over. Using lemma 4.3, each pair can be used to construct a term of the $2^{j+1}\varepsilon$ type. In all, there will be at least $(2^i - 2)/2 = 2^{i-1} - 1$ terms of the $2^{j+1}\varepsilon$ type constructed. For the next step, and all succeeding steps, the number of terms at the beginning is odd so exactly one term will not be used in pairing the terms whose ε values are congruent mod 4. Therefore, $[(2^{i-1} - 1) - 1]/2 = 2^{j-2} - 1$ terms of the $2^{j+2}\varepsilon$ type can be constructed. Continuing this process, $2^{j-n} - 1$ terms of the $2^{j+n}\varepsilon$ type can be constructed where $0 \leq n \leq i$. ▲

Lemma 4.5 Given one term of the $2^j\beta$ type and $2^i - 1$ terms of the $2^j\varepsilon$ type, one can construct one term of the $2^{i+j}\beta$ type. Furthermore, the original $2^j\beta$ term can be retained as one of the terms used in constructing the $2^{i+j}\beta$ term.

Proof: Using the technique of lemma 4.4, the $2^i - 1$ terms of the $2^j\varepsilon$ type can be used to construct $2^{i-1} - 1$ terms of the $2^{j+1}\varepsilon$ type where exactly one term of the $2^j\varepsilon$ type is not used. This remaining term can be paired with the term of the $2^j\beta$ type. For this final pair, if

$\beta \equiv 1 \bmod 2$, then $2^j(\beta+\varepsilon)$ is of the $2^{j+1}\beta$ type. If $\beta \equiv 0 \bmod 2$, then the $2^j\beta$ term is of the $2^{j+1}\beta$ type. The net result is that the final pair produces a term of the $2^{j+1}\beta$ type which is constructed using the original $2^j\beta$ term. By exactly the same technique, the $2^{i-1}-1$ terms of the $2^{j+1}\varepsilon$ type and the one term of the $2^{j+1}\beta$ type can be used to construct $2^{j-2}-1$ terms of the $2^{j+2}\varepsilon$ type and one term of the $2^{j+2}\beta$ type. As in the first step, the $2^{j+1}\beta$ term and, hence, the original $2^j\beta$ term is retained in constructing the $2^{j+2}\beta$ term. In general, this process will produce $2^{i-n}-1$ terms of the $2^{j+n}\varepsilon$ type plus one term of the $2^{j+n}\beta$ type. When the step where $n=i$ is reached, there are 2^0-1 or zero terms of the $2^{j+i}\varepsilon$ type and one term of the $2^{j+i}\beta$ type. The original $2^j\beta$ term is retained in constructing the $2^{j+i}\beta$ term. ▲

Theorem 4.14 Let f be a diagonal form in Q_2 of degree $n=2^k m$ where $(m,2)=1$ and $k \geq 4$. Then if f contains at least n^2+1 variables f has a non-trivial zero in O_2.

Proof: As a result of theorems 4.2, 4.12, and 4.13, the following assumptions can be made:

(1) $f = f_0 + 2f_1 + \cdots + 2^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_2.

(2) If N_i denotes the number of variables in each f_i, then

$$N_0 \geq \frac{n^2+1}{n} > n \;,\; N_0 + N_1 > 2n \;,\; \ldots,\; N_0 + N_1 + \cdots + N_{n-1} > n^2.$$

The proof of the theorem will follow if it is possible to construct a term of the $2^{k+2}\beta$ type using a sum of coefficients from f provided some of the coefficients come from f_0. To construct this $2^{k+2}\beta$ term, consider two cases, one when $N_0 \geq 2^{k+2}$ and the other when $N_0 < 2^{k+2}$. If $N_0 \geq 2^{k+2}$ using lemma 4.2, a term of the $2^{k+2}\beta$ type can be constructed using only coefficients from f_0. If $N_0 < 2^{k+2}$, the result is not so immediate. Note first that since $n = 2^k m$ and $k \geq 4$ by assumption (2), $N_0 > n \geq 2^k \geq 2^4$. Also by assumption (2),

$$N_0 + N_1 + N_2 + N_3 + N_4 > 5n \geq 5 \cdot 2^k = 2^{k+2} + 2^k .$$

This inequality and $N_0 < 2^{k+2}$ imply that $N_1 + N_2 + N_3 + N_4 > 2^k$. Using 2^4 of the coefficients from f_0, by lemma 4.2, one term of the $2^4\beta$ type can be constructed. Now, in order to apply lemma 4.5, $2^{k+2} - 1$ terms of the $2^4\varepsilon$ type need to be constructed using the remaining terms from f_0, f_1, f_2, f_3, and f_4. To do this, begin with the $N_0 - 16$ terms that remain in f_0. As demonstrated in lemma 4.4, these $N_0 - 16$ terms can be used to construct at least $(N_0 - 16)/2 - 1$ terms of the 2ε type. The terms from f_1 are already of the 2ε type so at least $(N_0 - 16)/2 - 1 + N_1$, 2ε terms are available using f_0 and f_1. Applying the same argument to this set of 2ε terms at least $[(N_0 - 16)/2 - 1 + N_1]/2 - 1$ terms of the $2^2\varepsilon$ type can be constructed. The terms from f_2 will add N_2 terms of the $2^2\varepsilon$ type. The following expression represents the number of terms of the $2^4\varepsilon$ type that result from applying this process four times.

$$\cfrac{\cfrac{\cfrac{\cfrac{N_0 - 16}{2} - 1 + N_1}{2} - 1 + N_2}{2} - 1 + N_3}{2} - 1 + N_4$$

This expression is equal to $N_0/16 + N_1/8 + N_2/4 + N_3/2 + N_4 - 23/8$ which is greater than

$$[(N_0 + N_1 + N_2 + N_3 + N_4) + (N_1 + N_2 + N_3 + N_4)]/16 - 3$$

$$> [(2^{k+2} + 2^k) + (2^k)]/16 - 3 = 2^{k-2} + 2^{k-3} - 3 \geq 2^{k-2} - 1 .$$

Thus, when $N_0 < 2^{k+2}$, the construction of one term of the $2^4\beta$ type using coefficients from f_0 and $2^{k-2} - 1$ terms of the $2^4\varepsilon$ type using the remaining coefficients from f_0, f_1, f_2, f_3, and f_4 can be accomplished. Applying lemma 4.5, one term of the $2^{k+2}\beta$ type can be constructed which retains the $2^4\beta$ term and, hence, uses coefficients from f_0.

In the following, the two cases no longer need to be considered separately. In each case, the construction of a $2^{k+2}\beta$ term means that by assigning values of one and zero to the variables in the f_i result

$$f_0 + 2f_1 + 2^2f_2 + 2^3f_3 + 2^4f_4 = 2^{k+2}\beta$$

is obtained. In this construction, at least one variable in f_0 has been assigned the value one. Denote εx^n as a term in f_0 for which $x = 1$ and note as in previous similar situations that $1 - \varepsilon^{-1}2^{k+2}\beta$ is an nth power in O_2. Therefore, for some δ in O_2, $\delta^n = 1 - \varepsilon^{-1}2^{k+2}\beta$. Now, let $x = \delta$, assign values to the remaining variables in $f_0, f_1, \ldots, f_4$ from the previous construction and assign all variables in f_i, $i > 4$, the value zero. This produces a non-trivial zero of f. ▲

This theorem proves the conjecture for diagonal forms in Q_2 of degree $n = 2^k m$ when $k \geq 4$. When $k = 0$, $n = m$, and since $(m,2) = 1$, the conjecture is proved in corollary 4.1. This leaves the cases where

$k = 1,2$, and 3. Techniques similar to those used in theorem 4.14 also work in these cases. One approach to constructing the different solutions of $f \equiv 0 \bmod 2^{k+2}$ is the following. When $k = 1$ consider two cases, $m = 1$ and $m \geq 3$. When $m = 1$ the proof can be found in Borevich and Shafarevich (3, p. 51). When $k = 2$ consider four cases, $N_0 \geq 16$, $12 \leq N_0 \leq 15$, $8 \leq N_0 \leq 11$, and $5 \leq N_0 \leq 7$. When $k = 3$ consider three cases, $N_0 \geq 32$, $16 \leq N_0 \leq 31$, and $8 < N_0 \leq 15$. The following example is representative of these cases.

Example 4.5 Consider the case where $k = 3$ and $8 < N_0 \leq 15$. Under assumption (1) of theorem 4.14, f can be represented as $f_0 + 2f + \cdots + 2^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_2 of degree $8m$, $(m,2) = 1$. With assumption (2) of theorem 4.14, $N_0 + N_1 + N_2 + N_3 > 4n = 4(8m) \geq 32$. This and the condition $N_0 \leq 15$ implies that $N_1 + N_2 + N_3 \geq 18$. Now lemma 4.2 implies that one term of the $2^3\beta$ type can be constructed using 8 of the coefficients of f_0. From the remaining coefficients of f_0, f_1, f_2, and f_3, at least

$$\frac{\dfrac{\dfrac{N_0 - 8}{2} - 1 + N_1}{2} - 1 + N_2}{2} - 1 + N_3$$

terms of the $2^3\varepsilon$ type can be constructed. The value of this expression is not less than

$$(N_0 + N_1 + N_2 + N_3)/8 + (N_1 + N_2 + N_3)/8 - 11/4 \geq 33/8 + 18/8 - 11/4 > 3.$$

Therefore, three terms of the $2^3\varepsilon$ type can be constructed to combine with the one term of the $2^3\beta$ type that came from f_0. Using lemma 4.5 these four terms will produce one term of the $2^5\beta$ type that retains the $2^3\beta$ term. Hence, a suitable solution for $f \equiv 0 \bmod 2^5$ has been constructed and the non-trivial zero follows.

The objective of the final section of this chapter is to prove Artin's conjecture for diagonal forms in Q_p when $p \neq 2$. This has been accomplished in corollary 4.1 for forms of degree n where $(n,p) = 1$ or $n = p^k$. This section deals with forms of degree $n = mp^k$ where $(m,p) = 1$, $k > 0$, and $m > 1$. The methods are similar to those used previously, however a few additional factors need to be considered.

Given a diagonal form f of degree $n = mp^k$, f is represented as $f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ and a solution for $f \equiv 0 \bmod p^{k+1}$ is constructed. As before, it is essential that this solution assign to some variable in f_0 a value that is not congruent to zero mod p. The following development describes a procedure for constructing this solution.

Suppose a form $g = \varepsilon_1 x_1^n + \varepsilon_2 x_2^n + \cdots + \varepsilon_{d+1} x_{d+1}^n$ where $d = (m,p-1)$ is a portion of some unit diagonal form f_i of f. As shown in theorem 4.8, there exists a solution $(1,\theta_2,\theta_3, \ldots ,\theta_{d+1})$ for the congruence $g \equiv 0 \bmod p$. Therefore,

$$\varepsilon_1 + \varepsilon_2\theta_2^n + \cdots + \varepsilon_{d+1}\theta_{d+1}^n = p^t\varepsilon$$

for some $t > 0$ and some unit ε. Now substitute $\theta_i y$ for each x_i:

$$g = \varepsilon_1 y^n + \varepsilon_2\theta_2^n y^n + \cdots + \varepsilon_{d+1}\theta_{d+1}^n y^n$$

$$= (\varepsilon_1 + \varepsilon_2\theta_2^n + \cdots + \varepsilon_{d+1}\theta_{d+1}^n)y^n = p^t\varepsilon y^n .$$

With reference to the entire form f, the term $p^t\varepsilon y^n$ produced from terms of f_i can be considered as the term εy^n in f_{i+t}. This operation of constructing one term from $d+1$ terms is called a <u>contraction</u>. In each partial form g the first variable is always assigned the value one and is called the <u>distinguished</u> variable. Variables that are produced by the contraction operation are called <u>derived</u> variables. The contraction operation will be applied repeatedly to f as described by the following steps.

<u>Step 1</u> Divide f_0 into $[N_0/(d+1)]$ partial forms each containing $d+1$ terms and assign any remaining variables of f_0 the value zero. Apply the contraction operation to each of the partial forms. With the resulting derived variables, f can now be represented as

$$pf_1^1 + p_2 f_2^1 + \cdots + p^i f_i^1 + \cdots$$

where f_i^1 denotes the original f_i combined with any derived terms of the form $p^i \varepsilon y^n$.

<u>Step 2</u> This step is similar to step 1 applied to f_1^1 instead of f_0. The only difference is that f_1^1 is divided into partial forms whose first, or distinguished, variable is a derived variable. Any variables in f_1^1 that cannot be used in this way are assigned the value zero. After applying the contraction operation to all such partial forms, f can be represented as $p^2 f_2^2 + p^3 f_3^2 + \cdots + p^i f_i^2 + \cdots$ with f_i^2 denoting the original f_i combined with any qualified derived variables produced in steps 1 and 2.

<u>Step t, $t > 1$</u> This step is exactly the same as step 2 applied to f_{t-1}^{t-1} and after t steps, f can be represented as

$$p^t f_t^t + p^{t+1} f_{t+1}^t + \cdots + p^i f_i^t + \cdots$$

where f_i^t contains the original f_i and any qualified derived variables produced in steps 1 through t.

Any f_i^t where $i \geq n$ contains only derived variables since originally $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$. A more important observation is that if $i \geq k+1$, $p^i f_i^t \equiv 0 \bmod p^{k+1}$ must have a non-trivial solution. Furthermore, if f_i^t contains a derived variable some non-trivial zero of $p^i f_i^t \equiv 0 \bmod p^{k+1}$ yields a non-trivial zero of f. This zero is obtained by setting the derived variable in f_i^t equal to one and assigning corresponding values to the ancestors of this derived variable in $f_0, f_1, \ldots, f_k$. All other variables are assigned the value zero. It is important to see that when a derived variable in f_i^t is assigned the value one all distinguished variables that are ancestors of this derived variable also have the value one. Therefore, when the ancestry is traced back to an original distinguished variable in f_0 this variable will have been assigned the value one. This value in f_0 allows us to construct a non-trivial zero of f from the solution of $f \equiv 0 \bmod p^{k+1}$. The main task of the next theorem is to show that some f_i^t where $i > k+1$ must contain a derived variable whenever f contains more than n^2 variables.

The following example will help to clarify these ideas and make the following theorem more understandable.

<u>Example 4.5</u> Consider the following diagonal form in O_5:

$$f = 2x_1^{50} + x_2^{50} + 3x_3^{50} + 20x_4^{50} + 65x_5^{50} + 150x_6^{50} + 175x_7^{50}$$

$$= (2x_1^{50} + x_2^{50} + 3x_3^{50}) + 5(4x_4^{50} + 13x_5^{50}) + 25(6x_6^{50} + 7x_7^{50}) .$$

Therefore,

$$f_0 = 2x_1^{50} + x_2^{50} + 3x_3^{50} \ , \ f_1 = 4x_4^{50} + 13x_5^{50} \ , \text{ and } f_2 = 6x_6^{50} + 7x_7^{50} \ .$$

Since $50 = 2\cdot 5^2$ and $(2,5-1) = 2$, according to theorem 4.8 the contraction operation can be applied to any three terms from any f_i. In f_0 if x_1 is considered as the distinguished variable, $x_1 = 1$, $x_2 = 0$, and $x_3 = 1$ yields $f_0 = 5$. Therefore, the substitution $x_1 = 1\cdot y_1$, $x_2 = 0\cdot y_1$, and $x_3 = 1\cdot y_1$ yields $2x_1^{50} + x_2^{50} + 3x_3^{50} = 5y_1^{50}$. The resulting term, $5y_1^{50}$, can be considered as the derived term y_1^{50} in f_1^1. That is, f can be represented as

$$5(y_1^{50} + 4x_4^{50} + 13x_5^{50}) + 25(6x_6^{50} + 7x_7^{50}) \ .$$

Now the derived variable y_1 must be considered as the distinguished variable when a solution for $y_1^{50} + 4x_4^{50} + 13x_5^{50} \equiv 0 \bmod 5$ is sought. Using only ones and zeros will not produce a solution to this congruence. However, a little arithmetic shows that when $x \equiv \pm 1 \bmod 5$, $x^{50} \equiv 1 \bmod 125$ and when $x \equiv \pm 2 \bmod 5$, $x^{50} \equiv -1 \bmod 125$. Since we are attempting to solve a congruence mod 125, knowing the value of x^{50} mod 125 is as useful as knowing the actual value of x^{50}. Considering the size of 2^{50}, the value mod 125 is more useful. Therefore, $y_1 = 1$, $x_4 = 2$, and $x_5 = 1$ yields $f_1^1 = y_1^{50} + 4x_4^{50} + 13x_5^{50} \equiv 10 \bmod 125$ and the substitution $y_1 = 1\cdot y_2$, $x_4 = 2y_2$, $x_5 = 1\cdot y_2$ yields $y_1^{50} + 4x_4^{50} + 13x_5^{50} \equiv 5\cdot 2y_2^{50} \bmod 125$. Thus, the second derived variable y_2 produces the term $2y_2^{50}$ in f_2^2 and the representation of f is now $f = 25f_2^2 = 25(2y_2^{50} + 6x_6^{50} + 7x_7^{50})$. The derived variable y_2 must now be the distinguished variable. Now set $y_2 = y_3$, $x_6 = 0$, and $x_7 = 2y_3$

to obtain $f_2^2 \equiv 2y_3^{50} + 0 - 7y_3^{50} \equiv -5y_3^{50} \bmod 125$. It follows that $f_3^3 = -y_3^{50}$ and $f \equiv 125f_3^3 \equiv 0 \bmod 125$ for any value of y_3. If $y_3 = 1$, then $y_2 = y_1 = 1$ and the resulting assignment for the x_i is $x_1 = x_3 = x_5 = 1$, $x_2 = x_6 = 0$, and $x_4 = x_7 = 2$. Therefore, $f(1,0,1,2,1,0,2) = 125\,a$ for some integer a. As in previous examples, theorem 2.11 implies that $1 - 2^{-1}125a$ is a 50th power in O_5 and $(\delta,0,1,2,1,0,2)$ is a non-trivial zero of f where $\delta^{50} = 1 - 2^{-1}125a$.

This example and the preceding discussion indicate the importance of having a derived variable in some f_i^t where $i \geq k+1$. This situation will have to exist if f contains at least one derived variable after $k+1$ of the previously described steps. Lemma 4.6 establishes a usable lower bound for the number of derived variables in f after t steps.

Lemma 4.6 Let $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_p of degree $n = mp^k$, $p \neq 2$. As usual, $(m,p) = 1$, $d = (m, p-1)$, and N_i denotes the number of variables in f_i. Define S_t to be the number of derived variables in

$$p^t f_t^t + p^{t+1} f_{t+1}^t + \cdots + p^i f_i^t + \cdots$$

after t steps as outlined in the discussion prior to example 4.5. Then $S_t \geq \min\{C_t, D_t\}$ where

$$C_t = \frac{N_0}{(d+1)^{t-1}} - 1 \quad\text{and}\quad D_t = \frac{N_0}{(d+1)^t} + \frac{N_1}{(d+1)^{t-1}} + \cdots + \frac{N_{t-1}}{d+1} - 1 .$$

Proof: The proof is by induction on t. Step one produces

$[N_0/(d+1)]$ derived variables so $S_1 = [N_0/(d+1)] \geq N_0/(d+1) - 1 = D_1$. Therefore, $S_1 \geq \min\{C_1, D_1\}$. Now assume that $S_r \geq \min\{C_r, D_r\}$ and consider S_{r+1}. After r steps, f has the form

$$p^r f_r^r + p^{r+1} f_{r+1}^r + \cdots + p^i f_i^r + \cdots .$$

Let w denote the number of derived variables in f_r^r and $S_r - w$ the number of derived variables in $f_{r+1}^r, f_{r+2}^r, \ldots$. For the $(r+1)$st step, the $N_r + w$ terms of f_r^r are partitioned into partial forms each with $d+1$ terms subject to the condition that the distinguished variable in each is one of the w derived variables. If $N_r \geq dw$, w partial forms can be constructed each containing d terms from f_r plus one derived variable. If $N_r < dw$, $[(N_r + w)/(d+1)]$ partial forms can be constructed. Thus, two cases need to be considered:

(1) $N_r \geq dw$ with $S_{r+1} = (S_r - w) + w = S_r$ and

(2) $N_r < dw$ with $S_{r+1} = (S_r - w) + [(N_r + w)/(d+1)]$.

Case (1). Consider the $\min\{C_r, D_r\}$. If $C_r \leq D_r$, the minimum is attained at C_r and

$$C_r = \frac{N_0}{(d+1)^{r-1}} - 1 > \frac{N_0}{(d+1)^r} - 1 = C_{r+1} .$$

If $C_r > D_r$, the minimum is attained at D_r and

$$D_r = \frac{N_0}{(d+1)^r} + \frac{N_1}{(d+1)^{r-1}} + \cdots + \frac{N_{r-1}}{d+1} - 1 \geq \frac{N_0}{(d+1)^r} - 1 = C_{r+1} .$$

Therefore $\min\{C_r, D_r\} \geq C_{r+1}$. Since $S_{r+1} = S_r$ and $S_r \geq \min\{C_r, D_r\}$, it follows that $S_{r+1} \geq \min\{C_r, D_r\} \geq C_{r+1} \geq \min\{C_{r+1}, D_{r+1}\}$.

Case (2) First observe that

$$S_{r+1} = S_r - w + \left[\frac{N_r + w}{d+1}\right] \geq S_r - w + \frac{N_r + w - d}{d+1}$$

$$= \frac{S_r + d(S_r - w) + N_r - d}{d+1} \geq \frac{S_r}{d+1} + \frac{N_r - d}{d+1} .$$

Now, if $C_r \leq D_r$, then $S_r \geq C_r$ and

$$S_{r+1} \geq \frac{C_r}{d+1} + \frac{N_r - d}{d+1} = \left(\frac{N_0}{(d+1)^{r-1}} - 1\right)\left(\frac{1}{d+1}\right) + \frac{N_r - d}{d+1}$$

$$= \frac{N_0}{(d+1)^r} + \frac{N_r}{d+1} - 1 \geq \frac{N_0}{(d+1)^r} - 1 = C_{r+1} .$$

If $C_r > D_r$, then $S_r > D_r$ and

$$S_{r+1} \geq \frac{D_r}{d+1} + \frac{N_r - d}{d+1} = \left(\frac{N_0}{(d+1)^r} + \cdots + \frac{N_{r-1}}{d+1} - 1\right)\left(\frac{1}{d+1}\right) + \frac{N_r - d}{d+1}$$

$$= \frac{N_0}{(d+1)^{r+1}} + \frac{N_1}{(d+1)^r} + \cdots + \frac{N_r}{d+1} - 1 = D_{r+1} .$$

Again, the conclusion $S_{r+1} \geq \min\{C_{r+1}, D_{r+1}\}$ follows completing the proof of the lemma. ▲

Theorem 4.15 Let f be a diagonal form in Q_p, $p \neq 2$, of degree $n = mp^k$, $(m,p) = 1$, $k \geq 1$. Then, if f contains at least $n^2 + 1$ variables it must have a non-trivial zero in O_p.

Proof: As in theorem 4.14 the following assumptions can be made:

(1) $f = f_0 + pf_1 + \cdots + p^{n-1}f_{n-1}$ where each f_i is a unit diagonal form in O_p.

(2) If N_i denotes the number of variables in f_i, then

$$N_0 + N_1 + \cdots + N_j > (j+1)n \quad \text{for} \quad 0 \leq j \leq n-1.$$

The proof of the theorem will follow when the existence of a derived variable in some f_t^{k+1}, $t \geq k+1$, is demonstrated. That is, in the notation of lemma 4.6, it is necessary to prove that $S_{k+1} > 0$. Since $S_{k+1} \geq \min\{C_{k+1}, D_{k+1}\}$, we need only establish that $C_{k+1} > 0$ and $D_{k+1} > 0$. By assumption (2), $N_0 > n = mp^k \geq p^k \geq (d+1)^k$. Therefore, $N_0/(d+1)^k > 1$ or

$$C_{k+1} = N_0/(d+1)^k - 1 > 0 .$$

Also, by assumption (2), $N_0 + N_1 + \cdots + N_k > (k+1)n$ while

$$(k+1)n \geq 2n = 2mp^k \geq (m+1)p^k \geq (d+1)^{k+1} .$$

The final inequality is true since $m+1 \geq d+1$ and $p \geq d+1$ are implied by $d = (m, p-1)$. Now

$$D_{k+1} = \frac{N_0}{(d+1)^{k+1}} + \frac{N_1}{(d+1)^k} + \cdots + \frac{N_k}{d+1} - 1$$

$$\geq \frac{N_0 + N_2 + \cdots + N_k}{(d+1)^{k+1}} - 1 > 0$$

so it follows that $D_{k+1} > 0$.

In terms of f the result $S_{k+1} > 0$ means that after $k+1$ steps f can be represented as

$$p^{k+1} f_{k+1}^{k+1} + p^{k+2} f_{k+2}^{k+1} + \cdots + p^i f_i^{k+1} + \cdots$$

and that some f_i^{k+1} contains a derived variable. Let z denote such a derived variable. Now, assign z the value one thereby assigning a corresponding value to each ancestor of z. When each variable that is not an ancestor of z is assigned the value zero the result is that $f = \alpha p^{k+1}$ for some α in O_p. In this process, each distinguished variable that is an ancestor of z will be assigned the value one. In particular, some distinguished variable in f_0 will be assigned the value one. Let εx^n be a term in f_0 in which x has been assigned the value one. Since $1 - \varepsilon^{-1}\alpha p^{k+1}$ is an nth power in O_p, $\delta^n = 1 - \varepsilon^{-1}\alpha p^{k+1}$ for some δ in O_p. When x is assigned the value δ and the other variables in f are assigned the values indicated above, the result is a non-trivial zero of f. ▲

SELECTED BIBLIOGRAPHY

1. Agnew, Jeanne. Explorations in Number Theory. Monterey, Calif.: Brooks Cole, 1972.

2. Bhaskaran, M. "Sums of p-th Powers in a P-adic Ring." Acta Arithmetica, 15 (1969), 217-219.

3. Borevich, Z. I., and I. R. Shafarevich. Number Theory. trans. Newcomb Greenleaf. New York: Academic Press, 1966.

4. Browkin, J. "On Forms Over p-adic Fields." Bulletin de l'Academic Polonaise des Sciences, Series des Sciences Math. Astro., et Phys., 14 (1966), 489-492.

5. Chevalley, C. "Demonstration d'une Hypothesis de M. Artin." Hamburg Mathematisches Seminar, Abhandlungen, 11 (1935), 73-75.

6. Chowla, S., and G. Shimura. "On the Representation of Zero by a Liniar Combination of k-th Powers." Det Kongelige Norske Videnskabers Selokabs Forhandlinger, 36 (1963), 169-176.

7. Davenport, H. The Higher Arithmetic. 3rd ed. London: Hutchinson and Company, 1968.

8. Davenport, H., and D. J. Lewis. "Homogeneous Additive Equations." Proceedings of the Royal Society, Series A, 274 (1963), 443-460.

9. LeVeque, William J., ed. Studies in Number Theory, MAA Studies in Mathematics, The Mathematics Association of America, 6 (1969).

10. Maxwell, George. "A Note on Artin's Diophantine Conjecture." Candadian Mathematical Bulletin, 13 (1970), 119-120.

11. Niven, Ivan, and Herbert S. Zuckerman. An Introduction to the Theory of Numbers. New York: John Wiley and Sons, Inc., 1966.

12. Terjanian, Guy. "Un Contre-example a une Conjecture d'Artin." Comptes Rendus de l'Academie des Sciences, Series A, 262 (1966), 612.

13. Tornheim, Leonard. "Sums of n-th Powers in Fields of Prime Characteristic." Duke Mathematical Journal, 4 (1938), 359-362.

VITA

Kay William Dundas

Candidate for the Degree of

Doctor of Education

Thesis: P-ADIC NUMBERS AND DIOPHANTINE EQUATIONS

Major Field: Higher Education

Biographical:

Personal Data: Born in Windsor, Missouri, October 19, 1938, the son of William L. and Ruth O. Dundas.

Education: Graduated from Arnold High School, Arnold, Kansas in 1956; received the Bachelor of Science degree from Fort Hays Kansas State College, Hays, Kansas, in May, 1960; received the Master of Arts degree from Fort Hays Kansas State College, Hays, Kansas, in August, 1961; attended Rutgers, The State University, New Brunswick, New Jersey, summers of 1963 and 1964; attended the University of Arkansas, summer of 1967; completed requirements for the Doctor of Education degree at Oklahoma State University in July, 1972.

Professional Experience: Instructor of Mathematics, Fort Hays Kansas State College, Hays, Kansas, 1961-62; Instructor of Mathematics, Hutchinson Community Junior College, Hutchinson, Kansas, 1962-69; graduate teaching assistant, Department of Mathematics and Statistics, Oklahoma State University, 1969-72.