Newton Polyhedra and p-Adic Estimates of Zeros of Polynomials in $\Omega_p[x, y]$

KAMEL ARIFFIN MOHD ATAN
Mathematics Department,
Faculty of Science and Environmental Studies,
Universiti Pertanian Malaysia,
43400 Serdang, Selangor, Malaysia.

Key words: p-adic estimates, zeros of polynomials, indicator diagrams.

ABSTRACT

Newton polyhedron associated with a polynomial in $\Omega_p[x, y]$ is introduced. Existence of a relationship between a Newton polyhedron and zeros of its associated polynomial is proved. This relationship is used to arrive at the $p$-adic estimates of the zeros. An upper bound to the $p$-adic orders of these zeros is found using the Newton polyhedron method.

1. INTRODUCTION

The role of the Newton polygon in obtaining properties of zeros of polynomials in one variable is quite well-known. For example, the Newton polygon can be usefully applied in proving Puiseux's theorem (Walker, 1962; Lefschetz, 1953). A Sathaye (1983) also considered generalised Newton-Puiseux expansion.

In this paper we consider extending the Newton polygon idea in the $p$-adic case to polynomials in two variables and call it the Newton polyhedron method. We will derive $p$-adic properties of zeros of such polynomials from their associated Newton polyhedrons, as asserted in Theorem 2.1 and Theorem 2.2.

In the $p$-adic case Koblitz (1977) discusses the Newton polygon method for polynomials and power series in $\Omega_p[x]$ where $\Omega_p$ denotes the completion of the algebraic closed field of the field of $p$-adic numbers $\mathbb{Q}_p$. Here estimates concerning zeros of polynomials are derived from the properties of the associated Newton polygon. Loxton and Smith (1982) investigated the application of the Newton polygon technique although a different method was eventually used to arrive at their result.

With $p$ denoting a prime, we define the valuation $|x|_p$ on $\mathbb{Q}_p$ as usual. That is

$$|x|_p = \begin{cases} p^{-\text{ord } x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

where $\text{ord } x$ denotes the highest power of $p$ dividing $x$ and $\text{ord } x = \infty$ if $x = 0$. This valuation extends uniquely from $\mathbb{Q}_p$ to $\overline{\mathbb{Q}}_p$, the algebraic closure of $\mathbb{Q}_p$, and to $\Omega_p$ and $\Omega_p$ is complete and algebraically closed.
Birch and McCann (1967) showed that if \( f(x_1, \ldots, x_n) \) is a polynomial with integral coefficients and \( \text{ord}_p(f(a_1, \ldots, a_n)) > \text{ord}_p(f) \) where \( D(f) \) is the discriminant of \( f \) for some integers \( a_1, \ldots, a_n \), then there are \( p \)-adic integers \( a'_1, \ldots, a'_n \) such that \( f(a'_1, \ldots, a'_n) = 0 \) and \( \text{ord}_p(a'_i - a_i) > 0 \) for \( i = 1, \ldots, n \).

From the properties of \( p \)-adic zeros established for polynomials in two variables from the associated Newton polyhedron, we give an estimate for the \( p \)-adic order of zeros of polynomials in two variables with coefficients in \( \mathbb{Q}_p \).

Our assertion is as follows.

**Theorem 1.1**

Let \( f \) be a polynomial in \( \mathbb{Q}_p[x, y] \). Let

\[
\delta = \max \left\{ \frac{1}{r, s} \left\{ \text{ord}_p f(\alpha, \beta) - \text{ord}_p \left( \frac{f^{(r+s)}}{r!s!}(\alpha, \beta) \right) \right\} \right\}
\]

for some \( \alpha, \beta \) in \( \mathbb{Q}_p \) where the maximum is taken over all pairs of non-negative integers \( (r, s) \). Then \( f \) has a zero \( (\xi, \eta) \) in \( \mathbb{Q}_p \), such that \( \text{ord}_p f(\xi, \eta) = \delta \) and every zero \( (\xi, \eta) \) of \( f \) satisfies \( \text{ord}_p f(\xi - \alpha, \eta - \beta) \leq \delta \).

**Theorem 2.1**

Let \( p \) be a prime and \( f \) be a polynomial in \( \Omega_p[x, y] \). If \( (\xi, \eta) \) is a zero of \( f \) then \( \text{ord}_p \xi, \text{ord}_p \eta, 1 \) is normal to an edge in \( N \) and falls between the upward-pointing normals to the faces of \( N \) adjacent to this edge.

**Proof:** Let \( f(x, y) = \sum a_{ij} x^i y^j \) where the \( a_{ij} \) are in \( \mathbb{Q}_p \). If \( (\xi, \eta) \) is a zero of \( f \) it follows that there are at least two terms \( T_i, T_j \) say of \( f(\xi, \eta) \) which attain the minimum \( \text{ord}_p \) that is \( \text{ord}_p T_i = \text{ord}_p T_j = \text{ord}_p \eta \) and \( \text{ord}_p \xi \). Thus, the corresponding points \( P_{ij} \) and \( P_{mn} \) to \( T_i \) and \( T_m \) respectively satisfy the equation:

\[
x \text{ord}_p \xi + y \text{ord}_p \eta + z = M
\]

where \( M = \min \text{ord}_p T_{ij} \). That is \( P_{ij} \) and \( P_{mn} \) lie on the plane \( Z \) say whose equation is given in (1).

Since \( M \leq \text{ord}_p T_i \) for every \( 0 \leq i, j \leq n \), it follows that every point \( P_{ij} \) in \( S \) lies on or above the plane \( Z \), and the line segment \( E \) joining \( P_{ij} \) to \( P_{mn} \) is either an edge or lies on a face \( F \) of \( N \). In either case \( E \) lies in the plane \( Z \). The normal \( \eta = (\text{ord}_p \xi, \text{ord}_p \eta, 1) \) to \( Z \) is normal to \( F \) containing \( E \). Clearly \( \eta \) is normal to at least one of the edges of \( F \). Further, since \( N \) lies above \( Z \), \( \eta \) must lie...
between the upward-pointing normals to the faces of $N_i$ adjacent to the edges of $F$.

We will prove the converse to the above theorem. First we have the following lemma.

**Lemma 2.1:** Let $L$ be a finite extension of $\mathbb{Q}_p$.

Let $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{i=0}^{n} b_i x^i$ be polynomials in $L[x]$. Let $\lambda$, $\mu$ be in $Q$ and let $\lambda_0 = \min(\text{ord } a_i + i \lambda)$, $\mu_1 = \min(\text{ord } b_i + i \mu)$, $\lambda_1 = \min(\text{ord } a_i + i \lambda)$.

Then there is a $\pi$ in $\mathbb{Q}_p$ with $\text{ord } \pi = \lambda_0$ and $\text{ord } f(\pi) = \lambda_0$ and $\text{ord } g(\pi) = \mu_1$.

**Proof:** Let $K$ be a finite unramified extension of $L$ with prime element $\pi$ chosen so that the residue field is sufficiently large as required below. Let $\Sigma$ be a set of representatives in $K$ for the residue field. Write

$$a_i = \pi^{\lambda_0} \sum_{i=0}^{n} a_i \pi^i \quad (0 \leq i \leq m)$$

where $\lambda_0 \geq 0$ and the $a_i$ are in $\Sigma$ and $a_i^{(0)} + 0$. Consider

$$\xi = \pi^{\lambda_1} \sum_{i=0}^{n} a_i^{(0)} \pi^i \quad \xi \geq 0$$

where the $a_i$ are in $\Sigma$ and $a_i^{(0)} + 0$. Then

$$f(\xi) = \pi^{\lambda_0} \sum_{i=0}^{n} a_i^{(0)} \pi^i \equiv 0 (\text{mod } \pi^{\mu_0 + 1}).$$

If the residue field is sufficiently large, we can find $C_i$ in $\Sigma$ so that

$$\Sigma a_i^{(0)} C_i \equiv 0 \quad (\text{mod } \pi)$$

and this gives the required $\xi$ in $\mathbb{Q}_p$ for the polynomial $f$.

Similarly, by letting

$$b_i = \pi^{\mu_1 - i} + \epsilon_i \sum_{i=0}^{n} b_i \pi^i \quad (0 \leq i \leq n)$$

where $\epsilon_i \geq 0$, the $b_i$ are in $\Sigma$ and $b_i^{(0)} + 0$ and considering

$$\xi = \pi^{\lambda_1} \sum_{i=0}^{n} C_i \pi^i \quad \xi \geq 0$$

as above, we see that $\xi$ can also be chosen to be the required element in $\mathbb{Q}_p$ for the polynomial $g$.

**Theorem 2.2**

Let $f$ be a polynomial in $\mathbb{Q}_p[x, y]$. Let $E$ be a non-vertical edge of $N_i$ common to two adjacent faces $F_1$ and $F_2$. Suppose $\eta = (\lambda, \mu, 1)$ is normal to $E$ and lies between the upward-pointing normals to $F_1$ and $F_2$. Then there are $\xi$ and $\eta$ in $\mathbb{Q}_p$ such that $\text{ord } \xi = \lambda$, $\text{ord } \eta = \mu$ and $f(\xi, \eta) = 0$.

**Proof:** Let $f(x, y) = \sum_{i,j} a_{i,j} x^i y^j$ and let $V$ and $W$ be the end-points of the edge $E$ on $N_i$. Then $e_{mn} = (m - r, n - s, \text{ord } a_{mn} - \text{ord } a_{nn})$ is a vector from $V$ to $W$.

Choose $\xi$, $\eta$ in $\mathbb{Q}_p$ with $\text{ord } \xi = \lambda$, $\text{ord } \eta = \mu$. We show first that the terms corresponding to $V$ and $W$ dominate in $f(\xi, \eta)$. Since $\eta$ is orthogonal to $e_{mn}$ we have

$$(m - r) \text{ord } \xi + (n - s) \text{ord } \pi + \text{ord } a_{mn} - \text{ord } a_{nn} = 0,$$

that is

$$\text{ord } a_{mn} \xi + \eta = \text{ord } a_{mn} \xi = \eta.$$

Let $\pi_1$ and $\pi_2$ be the upward-pointing normals to the faces $F_1$ and $F_2$, respectively, normalised to have third component equal to 1.
Since \( \sim \) is in the plane of \( n_1 \) and \( n_2 \) and lies between them, we can write

\[
\sim = \gamma n_1 + (1 - \gamma) n_2
\]

with \( 0 \leq \gamma \leq 1 \).

Let \( V \) be any point in \( N_f \) and let \( \sim \) be the vector from \( V \) to \( V' \). From the definition of \( N_f \), the line segment from \( V \) to \( V' \) lies on or above the planes determined by \( F_i \) and \( F_j \). So, we have

\[
\sim \cdot n = \gamma \sim \cdot n_1 + (1 - \gamma) \sim \cdot n_2 \geq 0,
\]

that is

\[
\text{ord}_p a_{ij} \xi \eta^i \geq \text{ord}_p a_{rs} \xi \eta^s.
\]

Hence, as asserted

\[
\text{ord}_p a_{rs} \xi \eta^s = \min_{ij} \text{ord}_p a_{ij} \xi \eta^i.
\]

(2) We can suppose \( r \neq m \). Otherwise the same argument applies after interchanging \( x \) and \( y \). Choose \( \eta \) in \( \mathbb{Q}_p \) with \( \text{ord}_p \eta = \mu \) as specified below, and write

\[
g(x) = f(x, \eta) = \sum_{k} c_k (\eta) x^k
\]

where

\[
c_k (\eta) = \sum a_{kj} \eta^j
\]

By part (1), \( \text{ord}_p a_{nj} \eta^j = \min_{ij} \text{ord}_p a_{ij} \eta^j \) and \( \text{ord}_p a_{mn} \eta^n = \min_{ij} \text{ord}_p a_{mj} \eta^j \). By Lemma (2.1) we can choose \( \eta \) in \( \mathbb{Q}_p \) with \( \text{ord}_p \eta = \mu \) so that \( c_\eta (\eta) = \text{ord}_p a_{\eta} \eta^s \) and \( \text{ord}_p a_{\eta} \eta^s \). Using (1) again, we have

\[
\text{ord}_p c_\eta (\eta) = \text{ord}_p a_{\eta} \eta^s + \lambda m \leq \text{ord}_p c_k (\eta) + \lambda k,
\]

for each \( k \). Thus the line segment joining the points \( (r, \text{ord}_p c_\eta (\eta)) \) and \( (m, \text{ord}_p c_k (\eta)) \) having slope \( -\lambda \), is part of the Newton polygon of \( g(x) \). By a standard theorem (see Koblitz (1977) lemma 4, page 90), therefore, there is an \( \xi \) in \( \mathbb{Q}_p \) with \( \text{ord}_p \xi = \lambda \) and \( g(\xi) = 0 \). This choice of \( \xi \) and \( \eta \) satisfies the requirements of the theorem.

3. \textit{p-adic ESTIMATE OF ZEROS OF A POLYNOMIAL IN } \( \Omega_p[x, y] \)

\textbf{Definition 3.1}

Let \( (\mu, \lambda, 1) \) be the normalised upward-pointing normals to the faces \( F_i \) of \( N_f \) of a polynomial \( f(x, y) \) in \( \Omega_p[x, y] \). We map \( (\mu_1, \lambda_1, 1) \) to the point \( (\mu, \lambda, 1) \) in the \( x-y \) plane. If \( F_i \) and \( F_j \) are adjacent faces in \( N_f \), sharing a common edge, we construct the straight line joining \( (u, \lambda) \) and \( (v, \lambda) \). If \( F \) shares a common edge with a vertical face \( F \) say \( \alpha x + \beta y = \gamma \) in \( N_f \), we construct the straight line segment joining \( (\mu, \lambda) \) and the appropriate point at infinity that corresponds to the normal of \( F \), that is the segment along a line with slope \( -\alpha / \beta \). We call the set of lines so obtained the Indicator diagram associated with \( N_f \).

Hence by the above definition if \( (\mu, \lambda) \) is a point lying on the straight line segment joining \( (\mu_1, \lambda_1) \) and \( (\mu_2, \lambda_2) \) say in the Indicator diagram of an \( N_f \), then \( (\mu, \lambda, 1) \) is normal to the common edge of the faces to which \( (\mu_1, \lambda_1, 1) \) and \( (\mu_2, \lambda_2, 1) \) are normal. It follows by Theorem 2.2 that the point \( (\mu, \lambda) \) gives the \( p \)-adic order of a zero \( (\xi, \eta) \) of the associated polynomial \( f \).

\textbf{Definition 3.2}

We call the segment in an Indicator diagram of an \( N_f \) that corresponds to the initial edges passing through the vertical axis in \( \mathbb{R}^3 \), the initial segments of the Indicator diagram.

Let \( f(x, y) = \sum a_{ij} x^i y^j \) be a polynomial in \( \Omega_p[x, y] \) of degree \( n \), and let \( (\alpha, \beta, 1) \) be a point of degree \( n \), and let \( (\alpha, \beta, 1) \) be a point of degree \( n \). Then the equation of an initial segment of the Indicator diagram associated with \( N_f \) is of the form \( rx + sy = \alpha \), obtained by considering the relationships of normals \( (x, y, 1) \) to the edges \( E_{rs} \) of \( N_f \), which join the points \( V_{rs}: (0, 0, \text{ord}_p a_{rs}) \) and \( V_{rs}: (r, s, \text{ord}_p a_{rs}) \).
By making use of the Newton polyhedron method we give the following theorem.

**Theorem 3.1**

Let \( f(x, y) = \sum a_{ij} x^i y^j \) be a polynomial in \( \mathbb{Q}_p[x, y] \) and let

\[
\delta = \max_{r, s} \frac{1}{(r + s)} \left( \text{ord}_p a_{ro} - \text{ord}_p a_{rs} \right)
\]

where the maximum is taken over all pairs \((r, s)\).

Then \( f \) has a zero \((\xi, \eta)\) in \( \mathbb{Q}_p^2 \) with \( \text{ord}_p (\xi, \eta) = \delta \) and every zero \((\xi, \eta)\) satisfies \( \text{ord}_p (\xi, \eta) \leq \delta \).

**Proof:** We note first that the maximum defining \( \delta \) occurs for an initial edge in \( \mathbb{N}_i \). Let \( r \) and \( s \) denote a pair of initial edges common to an initial face \( F \) in \( \mathbb{N}_i \) such that by the convexity of \( \mathbb{N}_i \) and the consecutive ordering of the initial edges,

\[
\text{ord}_p (a) \quad \text{for every} \quad j, \quad j \leq j \leq k
\]

Let \( \xi_j \) denote the segments in the Indicator diagram corresponding to the edges \( r, s \) in \( \mathbb{N}_i \).

Thus \( \xi_j \) and \( \xi_{j+1} \) in the Indicator diagram associated with \( \mathbb{N}_i \) are adjacent segments sharing a common vertex. Now the equation of \( \xi_j \) is given by

\[
r x + s y = \alpha
\]

where \( \alpha = \text{ord}_p a_{ro} - \text{ord}_p a_{rs} \). Since \( \frac{r_j}{s_j} > \frac{r_{j+1}}{s_{j+1}} \), \( \xi_j \) is steeper than \( \xi_{j+1} \). As this is true for every \( j, 1 \leq j \leq k \), the set of initial segments in the Indicator diagram associated with \( \mathbb{N}_i \) has a convex shape. The line \( y = x \) intersects some initial segment \( \xi_j \) at a point \( (\mu_j, \mu_m) \). Also, for every \( j, \mu_j + m \) the line \( y = x \) intersects lines \( \xi_j \) produced at some points \( (\mu_j, \mu_m) \).

Since \( \frac{r_{j+t}}{s_{j+t}} < \frac{r_j}{s_j} < \frac{r_{j-t}}{s_{j-t}} \), it follows that \( \mu_m \geq \mu_j \) for every \( t, 0 < t < m, m + t \leq k \). Hence

\[
\mu_m \geq \mu_j
\]

for every \( j, 1 \leq j \leq k \).

By considering the equation of \( \xi_j \), there are \( r_j, s_j \) such that

\[
\mu_j = \frac{1}{r_j + s_j} \text{ord}_p a_{ro}
\]

Then, clearly by (1),

\[
\mu_m = \max_{j} \frac{1}{r_j + s_j} \text{ord}_p a_{ro}
\]

By the convexity of the set of initial segments in the Indicator diagram associated with \( \mathbb{N}_i \), for every point \((\mu_j, \lambda_j)\) in the initial segments \( \xi_j \), we have

\[
\mu_m \geq \min_{j} \mu_j
\]

for every \( j, \mu_j + m \). By Theorem 2.2 and Definition 3.1 there exist \( \xi_j, \eta_j \) in \( \mathbb{Q}_p \) such that \( \text{ord}_p (\xi_j, \eta_j) = \mu_m \), \( \text{ord}_p (\xi_j, \eta_j) = \mu_m \) and \( f(\xi_j, \eta_j) = 0 \), and there are \( \xi_j, \eta_j \) in \( \mathbb{Q}_p \) such that \( \text{ord}_p (\xi_j, \eta_j) = \mu_m \) and \( f(\xi_j, \eta_j) = 0 \). Our assertion then follows from (2), (3) and letting \( \delta = \mu_m \).

The above theorem is a special case of Theorem 1.1 whose proof is as follows.

**Proof of Theorem 1.1.** Let \( g(X, Y) \) be the resulting polynomial on expanding \( f(X + \alpha, Y + \beta) \) using Taylor’s theorem. That is,

\[
g(X, Y) = f(\alpha, \beta) + \sum_{r, s \geq 0} \frac{f(r+s)(\alpha, \beta)}{r!s!}(X-\alpha)^r(Y-\beta)^s
\]

By Theorem 3.1 there are \( \gamma_1, \gamma_2 \) in \( \mathbb{Q}_p \) such that \( f(\gamma_1, \gamma_2) = 0 \) and \( \text{ord}_p (\gamma_1, \gamma_2) = \delta \) and every zero \((X', Y')\) of \( g \) satisfies \( \text{ord}_p (X', Y') \leq \delta \).

Set \( \xi = \gamma_1 + \alpha, \eta = \gamma_2 + \beta \) and \( \xi = X' + \alpha, \eta = Y' + \beta \). Then \((\xi, \eta)\) are zeroes of \( f \) satisfying the requirements of the theorem.
CONCLUSION

Theorems 2.1 and 2.2 assert the existence of relationships between zeros of polynomials in $\Omega, [x, y]$ and their associated Newton polyhedra. This relationship is already well-known for one-variable polynomials with coefficients in $\Omega$. Newton polyhedra associated with polynomials in two variables with coefficients in $\Omega$ is treated in more detail in Mohd Atan (1984). The result of Theorem 1.1 gives an improvement to a result by Birch and McCann (1967) for polynomials in two variables.

ACKNOWLEDGEMENTS

The author would like to express his gratitude to Assoc. Prof. J. Loxton of the University of New South Wales for suggesting the problem and for discussions leading to the result.

REFERENCES


(Received 16 July, 1985)