On The Complexity of Polynomial Factorization Over P-adic Fields

Olga Erzsébet Veres

A Thesis

In the Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy (Mathematics) at

Concordia University

Montreal, Quebec, Canada

March, 2009

Your file Votre référence
ISBN: 978-0-494-63368-7
Our file Notre référence
ISBN: 978-0-494-63368-7

ABSTRACT

On The Complexity of Polynomial Factorization Over P-adic Fields

Olga Erzsébet Veres, Ph.D.
Concordia University, 2009

Let p be a rational prime and $\Phi(x)$ be a monic irreducible polynomial in $\mathbf{Z}_p[x]$. Based on the work of Ore on Newton polygons (Ore, 1928) and MacLane's characterization of polynomial valuations (MacLane, 1936), Montes described an algorithm for the decomposition of the ideal $p\mathcal{O}_K$ over an algebraic number field (Montes, 1999).

We give a simplified version of the Montes algorithm with a full MAPLE implementation which tests the irreducibility of $\Phi(x)$ over $\mathbf{Q}_p$. We derive an estimate of the complexity of this simplified algorithm in the worst case, when $\Phi(x)$ is irreducible over $\mathbf{Q}_p$. We show that in this case the algorithm terminates in at most

$$O\big((\deg \Phi)^{3+\epsilon} v_p(\operatorname{disc} \Phi)^{2+\epsilon}\big)$$

bit operations.

Lastly, we compare the "one-element" and "two-element" variations of the Zassenhaus "Round Four" algorithm with the Montes algorithm.

ACKNOWLEDGMENTS

I would like to acknowledge with great gratitude the excellent guidance and consistent help of my supervisor, Dr. David Ford. Throughout the time I have been working on my research and thesis he did not cease to encourage and support me. Without his mathematical expertise, his valuable advice, and without his time consuming and meticulous corrections of the text this research could not have been achieved.

I would also like to thank the program director Dr. Galia Dafni and the Department of Mathematics and Statistics of Concordia University for their counsel and financial assistance during my program.

It is worth mentioning my parents who with their steadfast love sustained me from the beginning to the end of my studies. Finally I am grateful to my husband Paul for his patience, love, and enduring kindness, but most of all to God, since it is by His grace that this work was accomplished.

Contents

List of Symbols

$\alpha_{r,\nu}$	x-coordinate of the left endpoint of some segment
$\widetilde{\alpha}_{r,K}$	x-coordinate of the left endpoint of $\mathcal{S}_{r,K}$
$\beta_{r,\nu}$	x-coordinate of the right endpoint of some segment
$\widetilde{\beta}_{r,K}$	x-coordinate of the right endpoint of $\mathcal{S}_{r,K}$
$\Delta(c_1, \ldots, c_n)$	discriminant of $c_1, \ldots, c_n$
$\Delta(K)$	discriminant of the field K
$\operatorname{disc} F$	discriminant of F
$\underset{v_p^*}{\sim}$	equivalence relation
$\underset{v_t}{\sim}$	equivalence relation
η_k	the coefficient of the associated polynomial
η	root of F
$\mathbb{F}_{q_r}$	finite field with q_r elements
$\operatorname{Gal}(\mathbb{F}_{q_r}/\mathbb{F}_{q_0})$	Galois group of $\mathbb{F}_{q_r}/\mathbb{F}_{q_0}$
$H_{t,\nu,\delta}$	polynomial constructed in Algorithm 1
$\mathcal{L}_{r,\nu}$	line with slope $-d_r/e_r$ and $(0, \nu/e_r)$ y-intercept
$N(a)$	norm of an element a
$N(\mathfrak{a})$	norm of an ideal $\mathfrak{a}$
$\mathcal{N}_r(K)$	Newton polygon of order r of a polynomial K
$\#A$	size of set A

$O(\cdot)$	big-O
$O^{\sim}(\cdot)$	soft-O
$\mathcal{O}_K$	ring of integers of the field K
$\omega_r(F)$	pseudo-valuation of F with respect to v_r
$\mathfrak{p}_{b,a}$	simple prime ideal
$\varphi_r(x)$	irreducible, monic polynomial in $\mathbf{Z}_p[x]$ with degree n_r
$\psi_r(x)$	irreducible, monic polynomial in $\mathbf{F}_{q_r}[Y]$ with non-zero constant term and of degree f_r
$\psi_r^*(Y)$	polynomial in $\mathbf{F}_p[Y]$ of degree f_r^*
$\widehat{\psi}_r(Z,Y)$	polynomial in $\mathbf{F}_p[Z][Y]$
$\Psi_{S,K}^{(r)}(Y)$	level-r associated polynomial of K with respect S
$\widetilde{\Psi}_K^{(r)}(Y)$	natural level-r associated polynomial of K
$\widehat{\Psi}_K^{(r)}(Y)$	extended natural level-r associated polynomial of K
$Q(a)$	algebraic extension of Q
$\mathbf{Q}_p$	field of p-adic numbers
$\mathbf{Q}_p(\alpha)$	algebraic extension of $\mathbf{Q}_p$
$\mathcal{S}_{r,K}$	segment of $\mathcal{N}_r(K)$ with slope $-d_r/e_r$
$\mathcal{T}_{r,\nu}$	longest segment of $\mathcal{L}_{r,\nu}$ with endpoints having nonnegative integer coordinates
v_r	valuation on $\mathbf{Q}_p(x)$
v_ψ	valuation associated with ψ
$v^*(\alpha)$	valuation
$\mathbf{Z}_p$	ring of p-adic integers
ξ_r	root of ψ_r

Introduction

Factorization According to Zassenhaus.

In an algebraic number field K with ring of integers $\mathcal{O}_K$, factorization of the ideal $p\mathcal{O}_K$, for p prime, can be determined via polynomial factorization over the field of p-adic numbers $\mathbf{Q}_p$ (Hensel, 1908).

If $K = \mathbf{Q}(\alpha)$ for a given $\alpha \in \mathcal{O}_K$ such that the index $\left[\mathcal{O}_K : \mathbf{Z}[\alpha]\right]$ is not divisible by p then the factorization of the ideal $p\mathcal{O}_K$ can be determined by polynomial factorization modulo p (Dedekind, 1871, 1876, 1878). In practice, efficient techniques for polynomial factorization modulo p (Berlekamp, 1967, 1970; Cantor and Zassenhaus, 1970) combined with Hensel lifting (Hensel, 1908; Zassenhaus, 1975) solve the problem of factoring $p\mathcal{O}_K$ in a straightforward and effective manner when p does not divide the index.

The complications arising when p divides the index $\left[\mathcal{O}_K : \mathbf{Z}[\alpha]\right]$ have been the subject of considerable study. Current ideas are derived from the "Round Four" algorithm of Zassenhaus (Zassenhaus, 1975), which has evolved into two main variations, the "one-element" method (Ford, Pauli, and Roblot, 2002) and the "two-element" method (Pauli, 2001). Variations of the one-element method are used by MAPLE and PARI. The two-element method is used, *e.g.*, by Magma.

The Approach of Montes.

The algorithm of Montes (Montes, 1999) is in a separate category.

Given a prime p and a monic polynomial $F(X)$ in $\mathbf{Z}_p[X]$, the Montes algorithm determines the number of irreducible factors of $F(X)$ in $\mathbf{Z}_p[X]$ and their respective degrees. (It is understood that, in practice, the algorithm works with a sufficiently precise approximation of $F(X)$ in $\mathbf{Z}[X]$.)

The algorithm exploits classical results of Ore on Newton polygons and provides an alternative to the methods based on ideas of Zassenhaus.

A familiar application of Newton polygons gives the p-adic valuations of roots of a polynomial in $\mathbf{Z}_p[X]$. If $F(X) \in \mathbf{Z}_p[X]$ has two roots with different p-adic values then Hensel-lifting techniques can be applied to construct a non-trivial p-adic factorization of F to any desired degree of precision.

This process constitutes "level 0" of the Montes algorithm.

For each factor of F revealed at level 0, the algorithm proceeds to higher levels, either to discover a refined factorization or to establish irreducibility.

At levels $r \geq 1$ the algorithm constructs the following:

- $\mathcal{N}_r(F)$, the Newton polygon of F with respect to the valuation v_r;
- a valuation v_{r+1} on $\mathbf{Q}_p[X]$;
- an irreducible monic polynomial $\varphi_{r+1}(X) \in \mathbf{Z}_p[X]$;
- the "associated polynomial" $\Psi^r_{S,F}(X) \in \mathbf{F}_{q_r}[X]$ for each segment S of the Newton polygon $\mathcal{N}_r(F)$.

The number of edges of $\mathcal{N}_r(F)$ and the number of distinct irreducible factors of $\Psi^r_{S,F}(X)$ give information for the factorization of F: if either is more than one then F is reducible.

Chapter Summary.

In Chapter 1 we present a short history of the development of factorization algorithms previous to the Montes algorithm.

In Chapter 2 we give definitions and theorems which are used in the Montes algorithm, together with the construction of φ_r, the so-called "key polynomial" (MacLane, 1936), for $r > 1$.

Our goal being to give an estimate of the complexity of the worst case of the Montes algorithm, we have simplified the algorithm so that it merely decides the question of irreducibility of a given polynomial. It is apparent that irreducibility is the most costly case for the original algorithm (*i.e.*, the case that reaches the most levels). In this case the Newton polygon at each level is a single segment (a necessary condition for irreducibility), and so our modified algorithm operates under the assumption that this is always the case; the failure of this condition terminates the modified algorithm.

An important gain from this approach is a substantial simplification of the notation, with a corresponding clarification of the operation of the algorithm. It should be noted that in the interest of simplicity we have abandoned most of the original notation in (Montes, 1999) and invented our own.

In Chapter 3 we give a complete MAPLE implementation of the modified Montes algorithm. In parallel with the presentation of this implementation we give complexity estimates for the various steps, ultimately arriving at an estimate of

$$O\big(n^{3+\epsilon} v_p(\operatorname{disc} F)^{2+\epsilon}\big)$$

bit operations, with $n = \deg F$, for the entire (modified) algorithm.

In Chapter 4 we compare the one-element and two-element methods with the Montes algorithm.

Chapter 1

History

1.1 Kummer

Our discussion of ideal factorization necessarily begins with Kummer.

In 1844, Kummer pointed out that unique factorization into primes is impossible for certain (algebraic) numbers. He was the first to discover the possibility of "ideal" factorization in cyclotomic fields (although he used different terminology).

At the time Fermat's "Last" theorem and the higher reciprocity laws were topics of wide interest. Kummer himself was concerned with computations of cyclotomic integers. He denoted a *cyclotomic integer* (or *complex number*, to use Kummer's terminology) by

$$f(\alpha) = a_0 + a_1\alpha + \cdots + a_{\lambda-1}\alpha^{\lambda-1},$$

with $a_0, \ldots, a_{\lambda-1}$ rational integers and λ an odd prime, with α being a λ-th primitive root of unity:

$$\alpha^{\lambda} = 1.$$

Computing with cyclotomic integers, (Kummer, 1847) gives prime factorizations of prime numbers $p < 100$ such that $p \equiv 1 \pmod{\lambda}$ and $\lambda \leq 19$.

Also in (Kummer, 1847) are given cyclotomic integers $f(\alpha)$ such that

$$N(f(\alpha)) = p$$

for $p < 1000$ and $5 \leq \lambda \leq 19$.

Thus in each of these cases there is a unique factorization of the prime p.

In the case $\lambda = 23$ Kummer found cyclotomic integers $f(\alpha)$ with the property that $N(f(\alpha)) = p$, namely

$$599 = N(1 + \alpha^{15} - \alpha^{16}), \quad 691 = N(1 + \alpha + \alpha^5), \quad 829 = N(1 + \alpha^{11} + \alpha^{20}).$$

However for $p = 47$ and $p = 139$ he found different representations. Specifically

$$47 = f(\alpha)f(\alpha^4)f(\alpha^{-7})f(\alpha^{-5})f(\alpha^3)f(\alpha^{-11})f(\alpha^2)f(\alpha^8)f(\alpha^9)f(\alpha^{-10})f(\alpha^6),$$
$$139 = g(\alpha)g(\alpha^4)g(\alpha^{-7})g(\alpha^{-5})g(\alpha^3)g(\alpha^{-11})g(\alpha^2)g(\alpha^8)g(\alpha^9)g(\alpha^{-10})g(\alpha^6),$$

with

$$f(\alpha) = \alpha^{10} + \alpha^{13} + \alpha^8 + \alpha^{15} + \alpha^7 + \alpha^{16},$$
$$g(\alpha) = \alpha^{10} + \alpha^{13} + \alpha^8 + \alpha^{15} + \alpha^4 + \alpha^{19},$$

and, since $f(\alpha) = f(\alpha^{-1})$ and $g(\alpha) = g(\alpha^{-1})$,

$$47^2 = N(f(\alpha)), \qquad 139^2 = N(g(\alpha)),$$

and hence the conditions

$$47 = N(h(\alpha)), \qquad 139 = N(k(\alpha)),$$

cannot be satisfied by any cyclotomic integers $h(\alpha)$, $k(\alpha)$. However, since

$$47 \cdot 139 = N(1 - \alpha + \alpha^{21}),$$

it follows that $1 - \alpha + \alpha^{21}$ has no nontrivial factor, and yet is not a prime. A detailed explanation appears in (Edwards, 1977) and (Edwards, 1980).

Since the factorization of some rational primes into algebraic primes was impossible. Kummer introduced ideal prime factors.

For given primes q and λ with $\lambda \neq q$ he considered f the smallest positive integer such that $q^f \equiv 1 \pmod{\lambda}$ and $e = (\lambda - 1)/f$, which is an integer since $q^{\lambda-1} \equiv 1 \pmod{\lambda}$.

Kummer did not give a definition of ideal prime factors, he described them, gave some properties of them, and laws of divisibility by them; nor did he give a definition of ideal numbers. In (Kummer, 1851) he wrote:

> "Nous remarquons aussi que la notion du nombre ou facteur complexe idéal sera employée aussi bien dans le sens plus large où les nombres complexes *existants*, comme cas particuliers, sont compris parmi les nombres complexes *idéaux*, que dans le sens plus étroit où les nombres *idéaux* signifient le contraire des nombres complexes *existants*, de méme que, dans l'Algèbre, le mot *imaginaire* est employé dans ce double sens."

In (Kummer, 1851) and (Kummer, 1846) some properties of ideal prime factors are proved. (In what follows the terms *cyclotomic integer* and *complex number* have the same meaning.)

> "The product of two or more complex numbers has precisely the same ideal prime factors as the factors taking (taken) together."

> "Each complex number, represented as a product of some factors, is divisible by q if and only if it contains all e ideal prime factors of q."

> "A complex number, containing all the ideal prime factors of q, containing each at least n times, is divisible by q^n."

> "If the complex number $f(\alpha)$ contains n ideal prime factors of the number q (belonging to the exponent f), in other words all this factors are different of not, the norm $N(f(\alpha))$ contains all the time the factor q^{nf}, but it never contains a higher power of q."

"Each given complex number contains only a finite number of ideal prime factors, perfectly determined."

"Two complex numbers, containing (having) the same ideal prime factors, differ only by a complex unit, by which they can be multiplied."

"In order that a complex number $f(\alpha)$ be divisible by $\varphi(\alpha)$ it is necessary and sufficient that all the ideal prime factors of the divisor $\varphi(\alpha)$ be contained in the dividend $f(\alpha)$."

Considering these properties one can see that ideal numbers (complex numbers) have the unique factorization property.

In the case λ not prime, Kummer generalized the theory of ideal numbers. He extended the theory to cyclotomic numbers that are roots of the equation

$$w^\lambda = D(\alpha),$$

where $D(\alpha)$ is a λ-th root of unity. Kummer also tried to extend his theory for factorization of some algebraic numbers.

An example appears in (Edwards, 1980) with $\alpha = \sqrt{-3}$ for which Kummer's generalization fails. The reason for this failure is that, if $\mu = 2$, $\nu = 1 + \sqrt{-3}$, and $\rho = 2^3$, then μ does not divide ν, although μ^k does divide $\rho\nu^k$ for every positive integer k.

In the following section we will describe work of Dedekind that led to a generalization of Kummer's theory.

1.2 Dedekind

In generalizing Kummer's theory Dedekind gave the definition of algebraic numbers and determined many of their properties.

Definition. A (real or complex) number α is called an *algebraic integer* (or simply an *integer*) if it is a root of an equation $P(x) = x^n + p_1 x^{n-1} + \cdots + p_n$, where $p_1, \ldots, p_n$ are rational integers.

Some basic properties are the following.

1. The sum, difference, and product of two algebraic integers are algebraic integers.
2. Each root of a monic polynomial with algebraic integer coefficients is an algebraic integer.
3. A rational integer is an algebraic integer.
4. All conjugates of an algebraic integer are algebraic integers.

The ideas of *ring of (algebraic) integers* and *integral basis* of an algebraic number field are due to Dedekind.

Other basic definitions are the following.

- The *norm* $N(b)$ of a number b is the product of the n conjugate numbers $b^{(1)}, b^{(2)}, \ldots, b^{(n)}$, *i.e.*,
$$N(b) = b^{(1)}b^{(2)}\cdots b^{(n)}.$$
- The *discriminant* $\Delta(c^{(1)}, c^{(2)}, \ldots, c^{(n)})$ of the numbers $c_1, c_2, \ldots, c_n$ is
$$\Delta(c_1, c_2, \ldots, c_n) = \det\bigl(\textstyle\sum \pm c^{(1)}c^{(2)}\cdots c^{(n)}\bigr)^2.$$
- A main invariant of an algebraic number field K is its *discriminant* $\Delta(K)$ which is a nonzero rational integer defined by
$$\Delta(K) = \Delta(\omega_1, \omega_2, \ldots, \omega_n)$$
where $(\omega_1, \omega_2, \ldots, \omega_n)$ an arbitrary integral basis for the ring of integers $\mathcal{O}_K$. The value of $\Delta(K)$ does not depend upon the choice of integral basis.
- Let $\theta \in \mathcal{O}_K$ with $K = \mathrm{Q}(\theta)$ and let $n = [K : \mathrm{Q}]$. The *index* of θ is defined as the positive integer $k = [\mathcal{O}_K : \mathrm{Z}[\theta]]$, where
$$\mathrm{Z}[\theta] = \{\, a_0 + a_1\theta + \cdots a_{n-1}\theta \mid a_i \in \mathrm{Z},\ i = 0, \ldots, n-1 \,\}.$$

The following relation holds:

$$\Delta(1, \theta, \ldots, \theta^{n-1}) = k^2 \Delta(K).$$

- The integer a is *divisible* by the integer b if $a = bc$ for some integer c.

Having defined the ring of integers $\mathcal{O}_K$, Dedekind generalized Kummer's theory in the following important theorem (Dedekind, 1871).

Theorem 1. *If x, a, b are nonzero integers in $\mathcal{O}_K$ such that xa^r is divisible by b^r for $r = 0, 1, 2, \ldots$, then a is divisible by b.*

Example. Taking $\alpha = \sqrt{-3}$ it can be shown that $\mathcal{O}_K = \mathbf{Z}\left[\frac{1}{2}(1 + \sqrt{-3})\right]$, and in this ring 2 divides $1 + \sqrt{-3}$.

Dedekind considered the set of all numbers $\alpha \in \mathcal{O}_K$ which are divisible by a given ideal number and called this set an *ideal*. In this way he found a correspondence between a given ideal number and a given ideal. Based on the properties of algebraic integers and on the elementary theory of divisibility, Dedekind gave the definition of an ideal.

A subset $\mathfrak{a} \subseteq \mathcal{O}_K$ is called an *ideal* if it satisfies the following two properties.

I. If any two elements $a, b \in \mathfrak{a}$, then their sum and difference $a \pm b \in \mathfrak{a}$.

II. If $a \in \mathfrak{a}$ and $x \in \mathcal{O}_K$, then their product $ax \in \mathfrak{a}$.

Dedekind also defined the divisibility of ideals and the notion of prime ideal.

An ideal $\mathfrak{a}$ is *divisible* by the ideal $\mathfrak{d}$ if $\mathfrak{a} = \mathfrak{d}\mathfrak{b}$ for some ideal $\mathfrak{b}$.

An ideal $\mathfrak{p}$ not equal to $\mathcal{O}_K$ or the zero ideal is said to be a *prime ideal* if, whenever a product of integers ab is in $\mathfrak{p}$, so is either a or b.

The lemma below, from (Dedekind, 1871), is used in the definition of a simple prime ideal.

> **Lemma 1.** *If an integer c is not an element of an ideal $\mathfrak{a}$, then there is an integer a divisible by c such that the roots $x \in \mathcal{O}_K$ of the congruence*
>
> $$ax \equiv 0 \pmod{\mathfrak{a}}$$
>
> *form a prime ideal.*

Definition. Given an integer $b \in \mathcal{O}_K$, which is not a unit, and any integer $a \in \mathcal{O}_K$, it follows from Lemma 1 that there is a prime ideal $\mathfrak{p}$ such that

$$\mathfrak{p} = \mathfrak{p}_{b,a} = \{\, x \in \mathcal{O}_K \mid xa \equiv 0 \pmod{b} \,\}.$$

Such a prime ideal is called *simple* prime ideal.

If $x \in \mathfrak{p}$ then we can say the simple prime ideal $\mathfrak{p}$ divides x.

Given b, a as above Dedekind considered the the r-th power of $\mathfrak{p}$,

$$\mathfrak{p}^r = \{\, y \in \mathcal{O}_K \mid ya^r \equiv 0 \pmod{b^r} \,\}$$

where r is a nonnegative rational integer. (Note that $\mathfrak{p}^0 = \mathcal{O}_K$.)

Some important results about simple ideals are given in (Dedekind, 1871, §163, 5).

> **Theorem 2.** *Let b and c be two integers. If, for every simple prime ideal $\mathfrak{p}$, every power of $\mathfrak{p}$ dividing b also divides c, then b divides c.*

One can see the similarities between the properties of prime ideal factors and simple prime ideals. An interesting presentation concerning these similarities can be found in (Edwards. 1980).

An immediate result of Theorem 2 is the following.

Corollary 1. *Any principal ideal (b) is the least common multiple of all powers of simple prime ideals dividing b.*

Dedekind gave this important consequence of Theorem 2.

Corollary 2. *Each prime ideal is a simple prime ideal.*

From now on, Dedekind could use prime ideals instead of simple prime ideals.

Dedekind showed that an ideal $\mathfrak{a}$ divides an ideal $\mathfrak{b}$ if and only if $\mathfrak{a} \subseteq \mathfrak{b}$. This important theorem may be summarized briefly by saying that *to divide is to contain* and leads immediately to the key result of this theory.

Theorem 3. *Each ideal $\mathfrak{a}$ different from $\mathcal{O}_K$ is a prime ideal or it can be written uniquely as a product of prime ideals.*

Dedekind defined the norm $N(\mathfrak{a})$ of an ideal $\mathfrak{a}$ as the number of mutually incongruent integers in $\mathcal{O}_K$ modulo $\mathfrak{a}$ and this number is finite assuming $\mathfrak{a} \neq 0$.

An important result is the following (Dedekind, 1876, 1877).

Proposition 1. *The norm of a product of ideals is equal to the product of the norms of the factors: $N(\mathfrak{ab}) = N(\mathfrak{a})N(\mathfrak{b})$.*

Each prime ideal in $\mathcal{O}_K$ occurs in the factorization of exactly one rational prime p. If the prime ideal $\mathfrak{p}$ divides $p\mathcal{O}_K$, then $N(\mathfrak{p}) = p^f$ and f is called the *inertial degree* of the prime ideal $\mathfrak{p}$. (Note that $\mathcal{O}_K/\mathfrak{p}$ is a finite field of degree f over $\mathbb{F}_p$.) In general, the factorization of $p\mathcal{O}_K$ has the form

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g},$$

where the prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$ are assumed to be mutually distinct from each other. The exponent $e_i \geq 1$ is known as the *ramification index* of the prime ideal $\mathfrak{p}_i$. If f_i is the inertial degree of $\mathfrak{p}_i$, the following general relation holds

$$n = \sum_{i=1}^{g} e_i f_i.$$

where $n = [K : \mathbf{Q}]$. If any ramification index e_i is greater than one then the rational prime p is said to *ramify in the field* K. Only a finite number of rational primes ramify in a given algebraic number field K and Dedekind proved the following crucial theorem regarding ramification.

Theorem 4. *A rational prime p ramifies in a number field K if and only if p divides the discriminant $\Delta(K)$.*

An important result later proved by Minkowski states that $|\Delta(K)| > 1$ whenever $[K : \mathbf{Q}] > 1$, and coupled with Theorem 4 we see that at least one rational prime ramifies in any number field $K \neq \mathbf{Q}$. The key result of this section with regard to the factorization methods developed later in this thesis is the following theorem; see (Dedekind, 1878, §2) and also (Dedekind, 1876, 1877).

Theorem 5. *Let $K = \mathbf{Q}(\theta)$, where $\theta \in \mathcal{O}_K$, and let k be the index of θ whose minimal polynomial is denoted by $F(x)$. Assume the prime p does not divide the index k and that*

$$F \equiv F_1^{e_1} \cdots F_g^{e_g} \pmod{p},$$

where $F_1, \ldots, F_g$ are distinct irreducible polynomials (incongruent prime functions) of degrees $f_1, \ldots, f_g$, respectively. Then the ideal $p\mathcal{O}_K$ has g distinct prime ideal factors and each prime function F_i corresponds to a specific prime ideal $\mathfrak{p}_i$ whose inertial degree is f_i and whose ramification index is e_i. Furthermore, $\mathfrak{p}_i = \gcd(p\mathcal{O}_K, F_i(\theta)\mathcal{O}_K)$.

Given the result in Theorem 5, it is natural to wonder if it is always possible to find an integer in any given number field whose index is one. Dedekind saw that there are some number fields which contain no integer of index one. The cubic field $K = \mathbf{Q}(\theta)$, where $F(x) = x^3 - x^2 - 2x - 8$ is the minimal polynomial of θ, is a classic example of such a field.

1.3 Hensel

Hensel, like Dedekind, was occupied with unique factorization of algebraic numbers. He believed there had to be an analogy between algebraic function theory and algebraic number theory. He first wrote about this in (Hensel, 1897), showing that the decomposition of algebraic numbers into prime factors can be replaced by a simpler approach using the expansion of algebraic functions around an arbitrary point, and he introduced the theory on which this statement is based. The first basic result is the following.

Proposition 2. *If K is a field and $K(\alpha)$ is an extension of K of degree n then any element of $K(\alpha)$ satisfies some polynomial of degree n*

$$(*) \qquad F(X) = 0$$

Considering $(*)$ as a congruence modulo p^M, with p is a prime number and M arbitrary large, he showed the following.

Proposition 3. *The congruence*

$$F(X) \equiv 0 \pmod{p^M}$$

possesses exactly the same number of roots as its degree. The n roots $X_1, \ldots, X_n$ always can be expanded in power series which progress by increasing powers of p and have at most a finite number of initial member with negative exponents. Thus

$$X_i = A_{-h}\, p^{-h} + \cdots + A_{-1}\, p^{-1} + A_0 + A_1\, p + \cdots,$$

for $i = 1, \ldots, n$, where h is a nonnegative integer.

Hensel was aware of the work of Kummer, and later Dedekind and Kronecker, in extending an algebraic number field while preserving unique factorization. Hensel

was inspired by Weierstrass's theory of the representation of algebraic functions as infinite power series

$$f(x) = \sum_{i=n}^{\infty} a_i x^i$$

where n is any integer.

It is an important fact that this representation of f as an infinite power series is not unique, and this led Hensel to give different representations, the *p-adic representations* for any prime p, of algebraic numbers. Hensel studied the properties of p-adic numbers and developed the theory of their use, including the well-known Hensel's Lemma.

Hensel considered the *field of p-adic numbers* $\mathbf{Q}_p$. He called a p-adic number a *p-adic integer* if in its there are only positive exponents of p. (The ring of p-adic integers is denoted $\mathbf{Z}_p$.)

Assume

$$f(x) = A_0 x^n + A_1 x^{n-1} + \cdots + A_n$$

is a polynomial with p-adic coefficients, and let $A_i^{(k)}$ be the k^{th} (rational) approximation of A_i, for $i = 0, \ldots, n$. Then

$$f^{(k)}(x) = A_0^{(k)} x^n + A_1^{(k)} x^{n-1} + \cdots + A_n^{(k)}$$

is called the *k^{th} approximation value of $f(x)$*.

Hensel was preoccupied with factorization of such polynomials into irreducible factors. The problem of decomposability was determined by the following.

> **Proposition 4.** *Let $F(x)$ be a p-adic function with discriminant $D(F) = p^\delta E$ where $\delta \geq 1$ is an integer. Then $F(x)$ decomposes in lower degree polynomials if and only if the δ^{th} approximation value $F^{(\delta)}$ modulo $p^{\delta+1}$ decomposes, namely each decomposition*
>
> $$F^{(\delta)}(x) \equiv \overline{f}(x)\overline{g}(x) \pmod{p^{\delta+1}}$$
>
> *specifies a unique decomposition*
>
> $$F(x) = f(x)g(x)$$

in $\mathbf{Z}_p[x]$, *with* $\overline{f}(x)$ *and* $\overline{g}(x)$ *being approximation values of* $f(x)$ *and* $g(x)$ *respectively.*

The next proposition is a consequence of the preceding.

Proposition 5. *Consider the modular factorisation*

$$F(x) \equiv f_0(x)\, g_0(x) \pmod{p^{r+1}}$$

with $r+1 > 2\rho$ *and* ρ *the p-adic valuation of* $R_x(f_0(x), g_0(x))$, *the resultant of* $f_0(x)$ *and* $g_0(x)$. *Then there is a factorisation*

$$F(x) = f(x)\, g(x)$$

in $\mathbf{Z}_p[x]$ *such that the* $(r-\rho)^{th}$ *approximation value of* $f(x)$ *and* $g(x)$ *are* $f_0(x)$ *and* $g_0(x)$ *respectively.*

From the propositions above Hensel derived a theorem of great importance.

Theorem 6 (Hensel's Lemma). *Let* $F(X)$ *be a polynomial in* $\mathbf{Z}_p[x]$ *and let* $f_0(x)$ *and* $g_0(x)$ *be polynomials in* $\mathbf{Z}[x]$ *such that*

$$F(x) \equiv f_0(x)\, g_0(x) \pmod{p}.$$

Assume further that the resultant $R_x(f_0(x), g_0(x))$ *is not divisible by* p. *Then there exists a factorisation*

$$F(x) = f(x)\, g(x)$$

in $\mathbf{Z}_p[x]$ *such that the* 0^{th} *approximation values of* $f(x)$ *and* $g(x)$ *are* $f_0(x)$ *and* $g_0(x)$ *respectively.*

In (Hensel, 1918) there appears an explicit procedure (now known as *Hensel lifting*) to construct arbitrarily precise p-adic approximations to $f(x)$ and $g(x)$.

Hensel lifting was the starting point for the subsequent work of Zassenhaus.

Lastly we present the theorem of Hensel giving the relation between polynomial factorization and ideal factorization.

Theorem 7. *Let $K = \mathbf{Q}(\alpha)$ be a algebraic extension of $\mathbf{Q}$ and let p be a rational prime not dividing the index of α. Let $F(X)$ be a polynomial in $\mathbf{Z}_p[x]$ such that $F(\alpha) = 0$. Suppose $F(X)$ has the factorization*

$$F(x) = F_1(x) \cdots F_h(x)$$

into distinct irreducible factors $F_1(x), \ldots, F_h(x)$ in $\mathbf{Z}_p[x]$, with

$$F_i(x) \equiv \mathcal{F}_i(x)^{e_i} \pmod{p},$$

$\mathcal{F}_i(x)$ irreducible modulo p, and $\deg \mathcal{F}_i = f_i$, for $i = 1, \ldots, h$. Then

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_h^{e_h}$$

with

$$\mathfrak{p}_i = p\,\mathcal{O}_K + \mathcal{F}_i(\alpha)\,\mathcal{O}_K$$

and $\mathfrak{p}_i$ having inertial degree f_i and ramification index e_i, for $i = 1, \ldots, h$.

1.4 Zassenhaus

In (Zassenhaus, 1969) Zassenhaus developed a method based on Hensel's lifting procedure to factorize a polynomial with rational integer coefficients.

Let

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n = \prod_1^n (x - \xi_i)$$

be a monic polynomial with rational integer coefficients. Defining

$$\Phi f = \max_{1 \le i \le n} \left(|a_i| / \binom{n}{i} \right)^{1/i}$$

we have

$$\min_{1 \le i \le n} |\xi_i| \;\le\; \Phi f \;\le\; \max_{1 \le i \le n} |\xi_i| \;\le\; \frac{\Phi f}{\sqrt[n]{2} - 1}$$

and it follows that

$$\Phi g \;\le\; \frac{\Phi f}{\sqrt[n]{2} - 1}$$

for any factor $g(x)$ of $f(x)$. Assuming

$$g(x) = x^m + b_1 x^{m-1} + \cdots + b_m$$

with $m \leq \lfloor n/2 \rfloor$, it is easily seen that

$$\begin{aligned} \max_{1 \leq i \leq m} |b_i| &\leq \max_{1 \leq i \leq m} \binom{m}{i} \left(\frac{\Phi f}{\sqrt[n]{2} - 1} \right)^i \\ &\leq \max_{1 \leq i \leq \lfloor n/2 \rfloor} \binom{\lfloor n/2 \rfloor}{i} \left(\frac{\Phi f}{\sqrt[n]{2} - 1} \right)^i. \end{aligned}$$

The Zassenhaus Hensel factorization algorithm combines Hensel lifting with Berlekamp's algorithm for factorization modulo p, where p is any prime not dividing the discriminant of f. Taking e such that

$$p^e > 2 \max_{1 \leq i \leq \lfloor n/2 \rfloor} \binom{\lfloor n/2 \rfloor}{i} \left(\frac{\Phi f}{\sqrt[n]{2} - 1} \right)^i$$

and g and m as given above, it follows that

$$\max_{1 \leq i \leq m} |b_i| < p^e/2.$$

From the Berlekamp algorithm we have

$$f(x) \equiv g_{1,1}(x) \cdots g_{r,1}(x) \pmod{p}$$

with $g_{1,1}(x), \ldots, g_{r,1}(x)$ irreducible modulo p. Applying $e - 1$ iterations of Hensel lifting yields

$$f(x) \equiv g_{1,e}(x) \cdots g_{r,e}(x) \pmod{p^e}.$$

We know that any factor of f of degree at most $\lfloor n/2 \rfloor$ will have coefficients in the range $[-M, +M]$, where $M = \lfloor p^e/2 \rfloor$. Products $g_{i_1}(x) \cdots g_{i_k}(x)$ of degree at most $\lfloor n/2 \rfloor$, with coefficients reduced to the range $[-M, +M]$, are tested as factors of $f(x)$. If f is reducible then such a factor will be found; otherwise f is irreducible.

It is clear that this algorithm is exponential in its worst case: if f is irreducible over $\mathbf{Z}$ but splits into linear factors modulo p then 2^{n-1} tests will be required to establish the irreducibility of f. (In practice this exponential behavior is rarely encountered.)

The Hensel-Zassenhaus algorithm was subsequently improved.

- Based on the inequality

$$\prod_{i=1}^{n} |\xi_i| \leq \left(1 + \sum_{i=1}^{n} |a_i|^2\right)^{1/2}$$

from (Specht, 1949), Mignotte, in (Mignotte, 1974), sharpened Zassenhaus's bound on the coefficients of g. For $k = 1, \ldots, m$ we have

$$|b_k| \leq \binom{m}{k}\left(1 + \sum_{i=1}^{n}|a_i|^2\right)^{1/2} \leq \binom{\lfloor n/2 \rfloor}{\lfloor n/4 \rfloor}\left(1 + \sum_{i=1}^{n}|a_i|^2\right)^{1/2}.$$

- In (Lenstra, Lenstra, Lovász, 1982) lattice basis reduction is applied to the testing phase to give a polynomial-time algorithm.

In (Zassenhaus, 1975) Zassenhaus was again occupied with factorization of polynomials with coefficients in $\mathbf{Z}$, this time considering the case when the prime p divides the polynomial discriminant. This work gave rise to the original version of the "Round Four" algorithm (Ford, 1978, 1987) for the computation of integral bases and factoriazion of polynomials over the field of p-adic numbers.

1.5 Ore and MacLane

The central technique of the Round Four algorithm is the attempted construction of a root of a polynomial in $\mathbf{Z}_p[x]$ as a power series with respect to a uniformizing element of an algebraic extension of $\mathbf{Q}_p$.

In contrast, Montes and Nart, building on work of Ore and MacLane, developed the idea of generalized Newton polygons, derived from valuations of the ring $\mathbf{Q}_p[x]$.

The Contributions of Ore

Considering the factorization in $\mathbf{Z}_p[x]$ of a monic irreducible polynomial with integer coefficients

$$F(x) = F_1(x) \cdots F_g(x)$$

with

$$F_i \equiv \psi_i^{e_i} \pmod{p}$$

for $i = 1, \ldots, g$, Ore constructed the *Newton polygon* $\mathcal{N}_1(F_i)$ of each factor F_i (Ore, 1928). For this construction let $\varphi_i(x) \in \mathbf{Z}[x]$ be monic and irreducible such that

$$\varphi_i \equiv \psi_i \pmod{p}.$$

For $i = 1, \ldots, g$, Ore considered the following representation of F:

$$F(x) = \sum_{j=0}^{m} p^{\alpha_j} Q_j(x) \varphi_i(x)^j$$

with $m = \lfloor \deg F / \deg \varphi_i \rfloor$, $\deg Q_j < \deg \varphi_i$, and $p \nmid Q_j$, for $j = 0, \ldots, m$. The Newton polygon $\mathcal{N}_1(F_i)$ consists of the lower convex hull of the set of points

$$\{ (0, \alpha_0), (1, \alpha_1), \ldots, (m, 0) \}.$$

Remark. The complete definition of Newton polygons appears as Definition 4 in Section 2.1 below.

The edges of $\mathcal{N}_1(F_i)$ provide information about factorizations of F_i and ramification indices.

For each edge S_i, Ore defined an *associated polynomial* $\Psi^{(1)}_{S_i, F_i}$ with coefficients in a finite field $\mathbf{F}_{q_i}$, where $q_i = p^{\deg \psi_i}$. A factorization of this associated polynomial gives a factorization of F_i. In particular, if an irreducible factor of the associated polynomial appears with multiplicity one then the corresponding factor of F_i is irreducible.

Remark. The complete definition of the associated polynomial appears as Definition 10 in Section 2.3 below.

Theorem 8 (Ore: Theorem of the Product). *Let $F(x) \in \mathbf{Z}_p[x]$ be a product*

$$F(x) = F_1(x) \cdots F_\delta(x)$$

of monic polynomials in $\mathbf{Z}_p[x]$. Then the edges with negative slope of $\mathcal{N}_1(F)$ are constructed by joining the edges of $\mathcal{N}_1(F_1), \ldots, \mathcal{N}_1(F_\delta)$ with positive length and negative

slope. Moreover, if $\mathcal{S}$ is the segment with slope $-d/e$ of $\mathcal{N}_1(F)$ then

$$\Psi^{(1)}_{S,F}(Y) = \prod_{|\mathcal{S}_i|>0} \Psi^{(1)}_{S_i,F_i}(Y)$$

where $\mathcal{S}_i$ is the segment of $\mathcal{N}_1(F_i)$ of slope $-d/e$, for $i = 1, \ldots, \delta$.

Theorem 9 (Ore: Theorem of the Polygon). *Let $F(x)$ be a monic polynomial in $\mathbf{Z}_p[x]$. Let $\mathcal{S}_1, \ldots, \mathcal{S}_\gamma$ be the edges of $\mathcal{N}_1(F)$ of negative slope and let*

$$-d_1/e_1, \ldots, -d_\gamma/e_\gamma$$

be their respective slopes. Then $F(x)$ admits a factorization

$$F(x) = F_1(x) \cdots F_\gamma(x).$$

Each factor $F_i(x)$ is a monic polynomial in $\mathbf{Z}_p[x]$ for $i = 1, \ldots, \gamma$, with $N_1(F_i)$ consisting of the single edge $\mathcal{T}_i$ having slope $-d_i/e_i$ and

$$\Psi^{(1)}_{\mathcal{T}_i,F_i}(Y) = \Psi^{(1)}_{\mathcal{S}_i,F}(Y).$$

Moreover, if θ_i is a root of $F_i(x)$ then

$$v(\varphi_1(\theta_i)) = d_i/e_i.$$

Theorem 10 (Ore: Theorem of the Associated Polynomial). *Let $F(x) \in \mathbf{Z}_p[x]$ be a monic polynomial such that $\mathcal{N}_1(F)$ consists of a single segment $\mathcal{S}$, with $\mathcal{S}$ having slope $-d_1/e_1$.*

Assume that the associated polynomial of F with respect to $\mathcal{S}$ has the factorization

$$\Psi^{(1)}_{S,F}(Y) = \psi_1(Y)^{a_1} \cdots \psi_\delta(Y)^{a_\delta}$$

with $\psi_i(Y), \ldots, \psi_i(Y)$ distinct irreducible polynomials in $\mathbf{F}_{q_1}[Y]$.

Then $F(x)$ admits a factorization

$$F(x) = G_1(x) \cdots G_\delta(x)$$

where

- $G_i(x)$ *is a monic polynomial in* $\mathbf{Z}_p[x]$,
- $\mathcal{N}_1(G_i)$ *consists of a single segment* $\mathcal{S}_i$,
- $\mathcal{S}_i$ *has slope* $-d_1/e_1$, *and*
- $\Psi^{(1)}_{\mathcal{S}_i,G_i}(Y) = \psi_i(Y)^{a_i}$,

for $i = 1, \ldots, \delta$. *Moreover, if* $a_i = 1$ *then* G_i *is irreducible.*

The Contributions of MacLane

Montes's work is built on the valuation theory of MacLane, who characterized all valuations of the polynomial ring $\mathbf{Z}[x]$ in terms of "inductive values" of $\mathbf{Z}[x]$.

Definition 1. If K is a ring then the map $v : K[x] \to \mathbf{Z} \cup \{\infty\}$ is called a *valuation* of $K[x]$ if

i) $v(F) = \infty$ if and only if $F = 0$,

ii) $v(FG) = v(F) + v(G)$ for all $F(x)$, $G(x) \in K[x]^*$,

iii) $v(F + G) \geq \min\{ v(F), v(G) \}$ for all $F(x)$, $G(x) \in K[x]$.

Definition 2. Let W be a valuation of $\mathbf{Z}[x]$ and let $F(x)$, $G(x) \in \mathbf{Z}[x]$.

We write

$$F \mid_W G$$

to express the condition

$$W(G - QF) > W(G)$$

for some $Q(x) \in \mathbf{Z}[x]$.

MacLane defined inductively the values $V_1, V_2, \ldots, V_k$ such that each value V_k is obtained from the value V_{k-1} using a suitable key polynomial ϕ_k.

A *key polynomial* with respect to the valuation W of $\mathbf{Z}[x]$ is a monic polynomial $\phi(x)$ in $\mathbf{Z}[x]$ such that

i) if $\phi \mid_w F\,G$ then $\phi \mid_w F$ or $\phi \mid_w G$, and

ii) if $\phi \mid_w F$ and $F \neq 0$ then $\deg F \geq \deg \phi$.

First Stage Valuation V_1

Let a rational prime p be given, let v_p denote the standard p-adic valuation of $\mathbf{Q}$, and let

$$F(x) = A_n\,x^n + \cdots A_1\,x + A_0$$

be a polynomial in $\mathbf{Z}[x]$.

Define the key polynomial for the first stage to be $\phi_1(x) = x$ and define

$$V_1 : \mathbf{Z}[x] \rightarrow \mathbf{Z}^{\geq 0}$$

by $V_1(\phi_1) = \mu_1$, with μ_1 an arbitrary nonnegative integer, and in general

$$V_1(F) = \min\{\, v_p(A_i) + i\mu_1 \mid 0 \leq i \leq n \,\}.$$

We denote this definition compactly by

$$V_1 = [\, v_p,\, \phi_1,\, \mu_1 \,].$$

We note that if $\mu_1 = 0$ then $V_1(F) = v(\text{content}(F))$.

k^{th} Stage Valuation V_k, $k \geq 2$

Definition 3. For $k \geq 2$ we choose ϕ_k to be a key polynomial with respect to V_{k-1} such that

i) $\deg \phi_k \geq \deg \phi_{k-1}$, and

ii) $V_{k-1}(\phi_k - \phi_{k-1}) = \min\{\, \mu_{k-1},\, V_{k-1}(\phi_k) \,\}$.

Considering the expansion

$$F(x) = A_{k,0}(x) + A_{k,1}(x)\,\phi_k(x) + \cdots + A_{k,m}(x)\,\phi_k(x)^m$$

with $A_{k,i}(x) \in \mathbf{Z}[x]$ and $\deg A_{k,i} < \deg \phi_k$ for $i = 0, \ldots, m$, we define

$$V_k : \mathbf{Z}[x] \to \mathbf{Z}^{\geq 0}$$

by $V_k(\phi_k) = \mu_k > V_{k-1}(\phi_k)$, with μ_k a positive integer, and in general

$$V_k(F) = \min \{ V_{k-1}(A_{k,i}) + i\mu_k \mid 0 \leq i \leq m \}.$$

We denote this definition compactly by

$$V_k = [\, V_{k-1},\ \phi_k,\ \mu_k \,].$$

MacLane proved that V_k is a valuation on $\mathbf{Z}[x]$.

Let $V_1, V_2, \ldots, V_k, \ldots$ be an infinite sequence of values defined as above. MacLane defined a *limit-valuation* as

$$V_\infty(f(x)) = \lim_{k \to \infty} V_k(f(x)),$$

and proved that it is a valuation on $\mathbf{Z}[x]$. He gave the following theorem.

Theorem (Maclane). *If every value of the field K is discrete, then every non-archimedean value W of the ring $K[x]$ can be represented either as an inductive or as a limit-valuation.*

Remark. As we have seen, Ore worked with first stage valuations. MacLane notes that "similar 'second-stage' values V_2 appear implicitly in the irreducibility investigations of Ore, Kürschák, and Rella."

Chapter 2

The Montes Algorithm

2.1 Newton Polygons

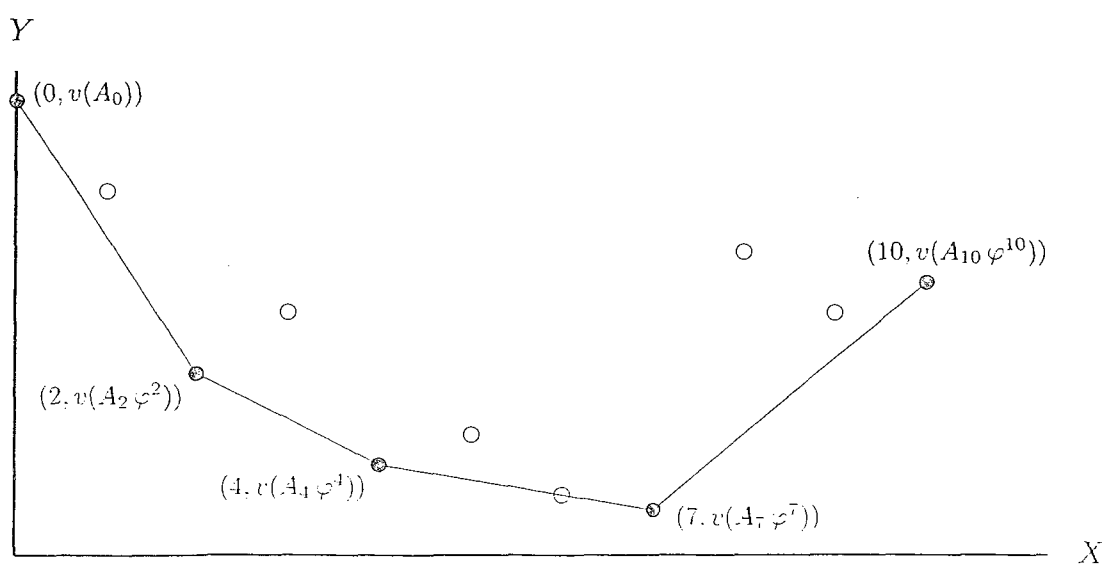
Definition 4. Let $v : \mathbf{Z}[x] \to \mathbf{Z}^{\geq 0}$ be a valuation of $\mathbf{Z}[x]$ and let $\varphi(x)$ be an irreducible monic polynomial in $\mathbf{Z}[x]$. Suppose $F(x) \in \mathbf{Z}[x]$ has the φ-adic expansion

$$F(x) = A_0(x) + A_1(x)\,\varphi(x) + \cdots + A_m(x)\,\varphi(x)^m$$

with $\deg A_i < \deg \varphi$ for $i = 0, \ldots, m$. Then *the Newton polygon of F with respect to v and φ* is the lower convex hull of the set of points

$$\left\{ \left(i,\, v(A_i) + iv(\varphi)\right) \;\middle|\; 0 \leq i \leq m \right\}$$

and is denoted by $\mathcal{N}_{v,\varphi}(F)$.



As we will show below, the Montes algorithm constructs

- a sequence $v_0, v_1, \ldots,$ of valuations of $\mathbf{Q}_p[x]$,
- a sequence $\varphi_0, \varphi_1, \ldots,$ of irreducible monic polynomials in $\mathbf{Z}_p[x]$,
- a sequence $-d_0/e_0, -d_1/e_1, \ldots,$ of rational slopes,

with d_r and e_r relatively prime, $d_0/e_0 = 0$, and $-d_r/e_r < 0$ if $r > 0$.

Definition 5. For $F(x) \in \mathbf{Z}[x]$ and $r \geq 0$ we define

$$\mathcal{N}_r(F) = \mathcal{N}_{v_r,\varphi_r}(F).$$

2.2 Valuations in the Montes Algorithm

Let $v : \mathbf{Q}_p \to \mathbf{Z} \cup \{\infty\}$ denote the standard p-adic valuation on $\mathbf{Q}_p$.

Definition 6. For $r \geq 0$ we define the valuation

$$v_r : \mathbf{Q}_p[x] \to \mathbf{Z} \cup \{\infty\}$$

as follows. If $F(x) \in \mathbf{Q}_p[x]^*$ then

$$v_r(F) = \begin{cases} v(\text{content}_p(F)) & \text{if } r = 0, \\ x_{r-1}\, d_{r-1} + y_{r-1}\, e_{r-1} & \text{if } r \geq 1, \end{cases}$$

where (x_{r-1}, y_{r-1}) is any point on the edge of $\mathcal{N}_{r-1}(F)$ with slope $-d_{r-1}/e_{r-1}$.

By definition $\varphi_0(x) = x$, and it follows that $v_0(\varphi_0) = 0$.

For $\theta \in \mathbf{Q}_p$ we have

$$v_r(\theta) = e_{r-1} v_{r-1}(\theta) = e_{r-1} e_{r-2} v_{r-2}(\theta) = (e_{r-1} e_{r-2} \cdots e_1) v(\theta).$$

Since $\mathcal{N}_{r-1}(\varphi_{r-1})$ consists of the single point $(1, v_{r-1}(\varphi_{r-1}))$ it follows that

$$v_r(\varphi_{r-1}) = e_{r-1} v_{r-1}(\varphi_{r-1}) + d_{r-1}$$

and from the construction of φ_r it follows that

$$v_r(\varphi_r) = e_{r-1} f_{r-1} v_r(\varphi_{r-1}).$$

Since $v_1(\varphi_1) = e_0 f_0 v_1(\varphi_0) = e_0 f_0 v_1(x) = 0$ we have $v_1 = v_0$.

If the φ_{r-1}-adic expansion of F is

$$F(x) = B_0(x) + B_1(x)\,\varphi_{r-1}(x) + \cdots + B_k(x)\,\varphi_{r-1}(x)^k$$

it can be shown that

$$v_r(F) = \min\left\{\, e_{r-1} v_{r-1}\big(B_j\big) + j v_r\big(\varphi_{r-1}\big) \;\middle|\; 0 \le j \le k \,\right\}.$$

In an obvious generalization of MacLane's notation we have

$$v_r = \begin{cases} [\,v,\, x,\, 0\,] & \text{if } r = 0, \\ [\,e_{r-1} v_{r-1},\, \varphi_{r-1},\, \overline{\mu}_r\,], \text{ with } \overline{\mu}_r = v_r\big(\varphi_{r-1}\big), & \text{if } r \ge 1. \end{cases}$$

2.3 Miscellaneous Definitions

At any given time the algorithm operates at some "level", say level r, with $r \ge 0$. At level r the algorithm is concerned with the "φ_r-adic" expansion of a given polynomial, from which is determined a "slope" $-d_r/e_r$, with d_r and e_r coprime. $d_0 = 0$, $e_1 = 1$, and $d_r > 0$ and $e_r > 0$ for $r \ge 1$.

Definition 7. For $r \ge 0$ we define

$$m_r = (1/d_r) \bmod e_r.$$

For positive integers r and ν we define

$$\alpha_{r,\nu} = \nu d_r^{-1} \bmod e_r,$$

$$\beta_{r,\nu} = (\nu - \alpha_{r,\nu} d_r)/e_r,$$

$$\mathcal{T}_{r,\nu} = \left\{\, (\alpha_{r,\nu} + \lambda e_r,\; \beta_{r,\nu} - \lambda d_r) \mid 0 \le \lambda \le \lfloor \beta_{r,\nu}/d_r \rfloor \,\right\},$$

$$\mathcal{L}_{r,\nu} = \left\{\, (x, y) \mid d_r x + e_r y = \nu \,\right\}.$$

For $r \ge 1$ and $K(X)$ a nonzero polynomial in $\mathbf{Z}_p[X]$ we define

- $\mathcal{S}_{r,K}$ to be the segment of $\mathcal{N}_r(K)$ having slope $-d_r/e_r$,
- $(\widetilde{\alpha}_{r,K},\, \widetilde{\beta}_{r,K})$ to be the left endpoint of $\mathcal{S}_{r,K}$,
- $(\widetilde{\alpha}_{r,K} + \widetilde{\gamma}_{r,K} e_r,\, \widetilde{\beta}_{r,K} - \widetilde{\gamma}_{r,K} d_r)$ to be the right endpoint of $\mathcal{S}_{r,K}$,
- $\widetilde{\nu}_{r,K} = d_r \widetilde{\alpha}_{r,K} + e_r \widetilde{\beta}_{r,K}$,

For $r \geq 0$ we define

$$\begin{aligned} &\overline{\mu}_r = 0, && \overline{\nu}_r = 0, && \text{if } r = 0, \\ &\overline{\mu}_r = d_{r-1} + e_{r-1}\overline{\nu}_{r-1}, && \overline{\nu}_r = e_{r-1} f_{r-1} \overline{\mu}_r, && \text{if } r \geq 1. \end{aligned}$$

Remark 1. It is readily seen that

- if $\beta_{r,\nu} \geq 0$ then $\mathcal{T}_{r,\nu}$ is the longest segment of $\mathcal{L}_{r,\nu}$ with endpoints having nonnegative integer coordinates, and
- if $\beta_{r,\nu} < 0$ then $\mathcal{T}_{r,\nu}$ is empty.

For example, if $d_2 = 3$ and $e_2 = 4$ then the segment $\mathcal{T}_{2,41}$ has endpoints $(3, 8)$ and $(11, 2)$ and the segment $\mathcal{T}_{2,5}$, with $\beta_{2,5} = -1$, is empty.

For $r \geq 1$ we have $(\alpha_{r,\overline{\nu}_{r+1}},\, \beta_{r,\overline{\nu}_{r+1}}) = (0,\, f_r \overline{\mu}_{r+1})$.

For $r \geq 1$ it is easily shown that $\overline{\mu}_r = v_r(\varphi_{r-1})$ and $\overline{\nu}_r = v_r(\varphi_r)$.

It is always the case that $\mathcal{S}_{r,K} \subseteq \mathcal{T}_{r,\widetilde{\nu}_{r,K}} \subseteq \mathcal{L}_{r,\widetilde{\nu}_{r,K}}$.

The line $\mathcal{L}_{r,\widetilde{\nu}_{r,K}}$ is the tangent line to $\mathcal{N}_r(K)$ of slope $-d_r/e_r$.

The set of integer points on $\mathcal{S}_{r,K}$ is given by

$$\{\, (\widetilde{\alpha}_{r,K} + j e_r,\, \widetilde{\beta}_{r,K} - j d_r) \mid j = 0,\, \ldots,\, \widetilde{\gamma}_{r,K} \,\}.$$

If $\deg K < \deg \varphi_r$ then $\mathcal{N}_r(K) = \{(0, v_r(K))\}$ so $v_{r+1}(K) = e_r v_r(K)$.

If $K(X)$ is a nonzero polynomial in $\mathbf{Z}_p[X]$ with

$$K(X) = A_{r,0}(X) + A_{r,1}(X)\, \varphi_r(X) + \cdots + A_{r,n}(X)\, \varphi_r(X)^n$$

its φ_r-adic expansion and

$$J_{r,K} = \{\, k \mid 0 \le k \le n,\ A_{r,k}(X) \ne 0 \,\}$$

then $\mathcal{N}_r(K)$ is the lower convex hull of the set

$$\{\, (k, v_r(A_{r,k}\varphi_r^k)) \mid k \in J_{r,K} \,\} = \{\, (k, y_k) \mid k \in J_{r,K} \,\}.$$

Since $K(X) \in \mathbf{Z}_p[X]$ we have $y_k \ge 0$ for $k \in J_{r,K}$. It follows that

$$\widetilde{\nu}_{r,K} = \min\{\, d_r k + e_r y_k \mid k \in J_{r,K} \,\},$$

$$\widetilde{\alpha}_{r,K} = \min\{\, k \mid k \in J_{r,K},\ d_r k + e_r y_k = \widetilde{\nu}_{r,K} \,\},$$

$$\widetilde{\beta}_{r,K} = y_{\widetilde{\alpha}_{r,K}} = v_r(A_{r,\widetilde{\alpha}_{r,K}}) + \widetilde{\alpha}_{r,K}\overline{\nu}_r.$$

Furthermore

$$\begin{aligned} v_{r+1}(K) = \widetilde{\nu}_{r,K} &= d_r\widetilde{\alpha}_{r,K} + e_r\widetilde{\beta}_{r,K} \\ &= d_r\widetilde{\alpha}_{r,K} + e_r v_r(A_{r,\widetilde{\alpha}_{r,K}}) + e_r\widetilde{\alpha}_{r,K}\overline{\nu}_r \\ &= \widetilde{\alpha}_{r,K}(d_r + e_r\overline{\nu}_r) + e_r v_r(A_{r,\widetilde{\alpha}_{r,K}}) \\ &= \widetilde{\alpha}_{r,K}\overline{\mu}_{r+1} + v_{r+1}(A_{r,\widetilde{\alpha}_{r,K}}). \end{aligned}$$

Definition 8 (*Sum of Segments*). If

$$S_1 = [(\alpha_1, \beta_1), (\alpha_1', \beta_1')], \quad S_2 = [(\alpha_2, \beta_2), (\alpha_2', \beta_2')]$$

are two segments with the same slope $-d_r/e_r$ then their sum is the segment of slope $-d_r/e_r$ with endpoints

$$(\alpha, \beta) = (\alpha_1 + \alpha_2,\ \beta_1 + \beta_2), \quad (\alpha', \beta') = (\alpha_1' + \alpha_2',\ \beta_1' + \beta_2').$$

Definition 9. For $r \ge 1$ and $k \ge 0$ and $\mathcal{S}$ an arbitrary segment of slope $-d_r/e_r$ with left endpoint (α, β) we define

$$\Theta_1(r, s) = m_s f_s \overline{\mu}_{s+1} n_r / n_{s+1} \qquad \text{for } s = 0, \ldots, r-1,$$

$$\Omega_r = \xi_0^{\Theta_1(r,0)} \cdots \xi_{r-1}^{\Theta_1(r,r-1)},$$

$$\Theta_2(\mathcal{S}, r, k) = \left\lfloor m_{r-1} \frac{(\beta - kd_r) - (\alpha + ke_r)\overline{\nu}_r}{e_{r-1}} \right\rfloor,$$

$$\Gamma_{\mathcal{S},r,k} = \Omega_r^{\alpha + ke_r} \xi_{r-1}^{\Theta_2(\mathcal{S},r,k)} \in \mathbf{F}_{q_r}.$$

Definition 10 (*Associated Polynomial*). Let $r \geq 0$, let α and β be nonnegative integers, and let $\mathcal{S}$ be an arbitrary segment of slope $-d_r/e_r$ with left endpoint (α, β). Assume $K(X) \in \mathbf{Z}_p[X]$ has φ_r-adic expansion

$$K(X) = A_0(X) + A_1(X)\,\varphi_r(X) + \cdots + A_n(X)\,\varphi_r(X)^n$$

with $d_r j + e_r v_r(A_j \varphi_r^j) \geq d_r\alpha + e_r\beta$ for $j = 0, \ldots, n$ and let

$$J = \left\{\, k \;\middle|\; 0 \leq k \leq \lfloor (n-\alpha)/e_r \rfloor,\ (\alpha + ke_r, v_r(A_{\alpha+ke_r}\varphi_r^{\alpha+ke_r})) \in \mathcal{S} \,\right\}.$$

We define the *level-r associated polynomial of K with respect to $\mathcal{S}$* to be

$$\Psi_{\mathcal{S},K}^{(r)}(Y) = \textstyle\sum_{k\in J} \eta_k Y^k$$

with $\eta_k \in \mathbf{F}_{q_r}$ constructed as follows.

- If $r = 0$ we let
$$\eta_k = \overline{A}_{\alpha+ke_0}.$$
- If $r = 1$ we let $B_k(X) = A_{\alpha+ke_1}(X) \,/\, p^{\beta-kd_1}$ and we let
$$\eta_k = \overline{B}_k(\xi_0).$$
- If $r \geq 2$ we let $\nu_k = v_r(A_{\alpha+ke_r})$ and we set
$$\eta_k = \Gamma_{\mathcal{S},r,k}^{-1}\,\Psi_{\mathcal{T}_{r-1,\nu_k},A_{\alpha+ke_r}}^{(r-1)}(\xi_{r-1}).$$

Example 1. Let $r \geq 1$, let $K(X) = 1$, and let $\mathcal{S} = \mathcal{T}_{r,0} = \{\,(0,0)\,\}$. Then

$$K = A_0\,\varphi_r^0, \quad A_0 = 1, \quad \mathcal{N}_r(K) = \{\,(0,0)\,\}, \quad J = \{\,0\,\}, \quad \nu_0 = 0,$$

and we have

$$\Psi_{\mathcal{S},K}^{(r)}(Y) = \Psi_{\mathcal{T}_{r,0},1}^{(r)}(Y) = \eta_0 Y^0 = \eta_0.$$

If $r = 1$ then

$$\Psi_{\mathcal{S},K}^{(r)} = \eta_0 = \overline{B}_0(\xi_0) = 1$$

and if $r \geq 2$ then by definition we have

$$\Gamma_{\mathcal{T}_{r,0},r,0} = \Omega_r^0 \, \xi_{r-1}^{\Theta_2(\mathcal{T}_{r,0},r,0)} = \xi_{r-1}^{\Theta_2(\mathcal{T}_{r,0},r,0)} = \xi_{r-1}^0 = 1$$

and therefore

$$\begin{aligned}
\Psi_{S,K}^{(r)}(Y) = \eta_0 &= \Gamma_{\mathcal{T}_{r,0},r,0}^{-1} \Psi_{\mathcal{T}_{r-1,\nu_0},A_0}^{(r-1)}(\xi_{r-1}) \\
&= \Gamma_{\mathcal{T}_{r,0},r,0}^{-1} \Psi_{\mathcal{T}_{r-1,0},1}^{(r-1)}(\xi_{r-1}) \\
&= \Gamma_{\mathcal{T}_{r,0},r,0}^{-1} \Gamma_{\mathcal{T}_{r-1,0},r-1,0}^{-1} \Psi_{\mathcal{T}_{r-2,0},1}^{(r-2)}(\xi_{r-2}) \\
&= \Gamma_{\mathcal{T}_{r,0},r,0}^{-1} \Gamma_{\mathcal{T}_{r-1,0},r-1,0}^{-1} \cdots \Gamma_{\mathcal{T}_{2,0},2,0}^{-1} \Psi_{\mathcal{T}_{1,0},1}^{(1)}(\xi_1) = 1.
\end{aligned}$$

Example 2. Let $r \geq 1$, let $K = \varphi_r^{e_r f_r}$, and let $\mathcal{S} = \mathcal{T}_{r,\overline{\nu}_{r+1}}$. Then

$$K = A_{e_r f_r} \varphi_r^{e_r f_r}, \quad A_{e_r f_r} = 1, \quad \mathcal{N}_r(K) = \{ (e_r f_r, \, e_r f_r \overline{\nu}_r) \},$$

$$(\alpha, \beta) = (0, f_r \overline{\mu}_{r+1}), \quad J = \{ f_r \}, \quad \nu_{f_r} = v_r(A_{e_r f_r}) = 0.$$

If $r = 1$ we have

$$B_{f_1}(X) = A_{\alpha + f_1 e_1}(X)/p^{\beta - f_1 d_1} = A_{e_1 f_1}(X)/p^{f_1 \overline{\mu}_2 - f_1 d_1} = 1$$

so that $\eta_{f_1} = 1 = \Omega_1^{-e_1 f_1}$ and therefore

$$\Psi_{\mathcal{T}_{1,\overline{\nu}_2}, \varphi_1^{e_1 f_1}}^{(1)}(Y) = \eta_{f_1} Y^{f_1} = \Omega_1^{-e_1 f_1} Y^{f_1}.$$

If $r \geq 2$ then

$$\Psi_{\mathcal{T}_{r-1,\nu_{f_r}}, A_{e_r f_r}}^{(r-1)}(Y) = \Psi_{\mathcal{T}_{r-1,0},1}^{(r-1)}(Y) = 1$$

and we have

$$\Psi_{\mathcal{T}_{r,\overline{\nu}_{r+1}}, \varphi_r^{e_r f_r}}^{(r)}(Y) = \eta_{f_r} Y^{f_r}$$

with

$$\eta_{f_r} = \Gamma_{\mathcal{S},r,f_r}^{-1} \Psi_{\mathcal{T}_{r-1,\nu_{f_r}}, A_{e_r f_r}}^{(r-1)}(\xi_{r-1}) = \Gamma_{\mathcal{S},r,f_r}^{-1}.$$

By definition

$$\begin{aligned}
\Theta_2(\mathcal{S}, r, f_r) &= \left\lfloor m_{r-1} \frac{(\beta - f_r d_r) - (\alpha + f_r e_r) \overline{\nu}_r}{e_{r-1}} \right\rfloor \\
&= \left\lfloor m_{r-1} \frac{f_r \overline{\mu}_{r+1} - f_r d_r - f_r e_r \overline{\nu}_r}{e_{r-1}} \right\rfloor = 0
\end{aligned}$$

so that

$$\Gamma_{S,r,f_r} = \Omega_r^{\alpha+f_r e_r} \xi_{r-1}^{\Theta_2(S,r,f_r)} = \Omega_r^{e_r f_r}$$

and therefore

$$\Psi^{(r)}_{\mathcal{T}_{r,\bar{\nu}_{r+1}},\varphi_r^{e_r f_r}}(Y) = \Omega_r^{-e_r f_r} Y^{f_r}.$$

Definition 11. We define the *natural level-r associated polynomial of* K to be

$$\widetilde{\Psi}_K^{(r)} = \Psi^{(r)}_{\mathcal{S}_{r,K},K}$$

and the *extended natural level-r associated polynomial of* K to be

$$\widehat{\Psi}_K^{(r)} = \Psi^{(r)}_{\mathcal{T}_{r,\bar{\nu}_{r,K}},K}.$$

Remark 2. For $r \geq 0$ and K and $\mathcal{S}$ as above we have the following.

- If we let $\Omega_0 = 1$ then for $r \geq 1$ we have

$$\Omega_r = \Omega_{r-1}^{e_{r-1}f_{r-1}} \xi_{r-1}^{m_{r-1}f_{r-1}\bar{\mu}_r}.$$

- $\Psi^{(r)}_{\mathcal{S},K}(Y) \in \mathbb{F}_{q_r}[Y]$.
- $\widetilde{\Psi}_K^{(r)}$ has nonzero constant term.
- $\widehat{\Psi}_K^{(r)} = \Psi^{(r)}_{\mathcal{T}_{r,\nu_{r+1}(K)},K}$.
- $\widehat{\Psi}_K^{(r)}(Y) = Y^{(\widetilde{\alpha}_{r,K}-\alpha_{r,\bar{\nu}_{r,K}})/e_r}\, \widetilde{\Psi}_K^{(r)}(Y)$.

Lemma 2. *If* $r \geq 1$ *and* $F(X)$, $G(X) \in \mathbf{Z}_p[X]^*$ *with* $\widetilde{\nu}_{r,F} = \widetilde{\nu}_{r,G} = \widetilde{\nu}$ *then*

$$\Psi^{(r)}_{\mathcal{S}_r,F+G} = \Psi^{(r)}_{\mathcal{S}_r,F} + \Psi^{(r)}_{\mathcal{S}_r,G}$$

for any segment $\mathcal{S}_r \subseteq \mathcal{L}_{r,\widetilde{\nu}}$.

Proof. Let $H = F + G$. It is clear that either $\mathcal{S}_{r,H} \subseteq \mathcal{L}_{r,\widetilde{\nu}}$ or else $\mathcal{S}_{r,H}$ lies entirely above $\mathcal{L}_{r,\widetilde{\nu}}$.

Let (α_r, β_r) be the left endpoint of $\mathcal{S}_r$. We consider the φ_r-adic expansions

$$F = \sum_{j\geq 0} A_j \varphi_r^j, \quad G = \sum_{j\geq 0} B_j \varphi_r^j, \quad H = \sum_{j\geq 0} C_j \varphi_r^j,$$

with $C_j = A_j + B_j$ for $j \geq 0$, and we let

$$\mathcal{J}_A = \{\, k \mid (\alpha_r + ke_r,\, v_r(A_{\alpha_r+ke_r}\varphi_r^{\alpha_r+ke_r})) \in \mathcal{S}_r \,\},$$

$$\mathcal{J}_B = \{\, k \mid (\alpha_r + ke_r,\, v_r(B_{\alpha_r+ke_r}\varphi_r^{\alpha_r+ke_r})) \in \mathcal{S}_r \,\},$$

$$\mathcal{J}_C = \{\, k \mid (\alpha_r + ke_r,\, v_r(C_{\alpha_r+ke_r}\varphi_r^{\alpha_r+ke_r})) \in \mathcal{S}_r \,\},$$

$$\mathcal{J}_U = \mathcal{J}_A \cup \mathcal{J}_B \cup \mathcal{J}_C.$$

It is evident that

$$(\mathcal{J}_A - \mathcal{J}_B) \cup (\mathcal{J}_B - \mathcal{J}_A) \subseteq \mathcal{J}_C \subseteq \mathcal{J}_A \cup \mathcal{J}_B = \mathcal{J}_U.$$

For $k \geq 0$ there exist a_k, b_k, $c_k \in \mathbb{F}_{q_r}$ such that

$$\Psi^{(r)}_{\mathcal{S}_r,F}(Y) = \sum_{k\in\mathcal{J}_A} a_k Y^k = \sum_{k\in\mathcal{J}_U} a_k Y^k,$$

$$\Psi^{(r)}_{\mathcal{S}_r,G}(Y) = \sum_{k\in\mathcal{J}_B} b_k Y^k = \sum_{k\in\mathcal{J}_U} b_k Y^k,$$

$$\Psi^{(r)}_{\mathcal{S}_r,H}(Y) = \sum_{k\in\mathcal{J}_C} c_k Y^k = \sum_{k\in\mathcal{J}_U} c_k Y^k.$$

It is obvious that

$$\sum_{k\in(\mathcal{J}_U-\mathcal{J}_A)} a_k Y^k = \sum_{k\in(\mathcal{J}_U-\mathcal{J}_B)} b_k Y^k = \sum_{k\in(\mathcal{J}_U-\mathcal{J}_C)} c_k Y^k = 0.$$

We will proceed by induction on r.

Let $r = 1$. Then for $k \geq 0$ we have

$$a_k = \overline{A}_k(\xi_0), \quad b_k = \overline{B}_k(\xi_0), \quad c_k = \overline{A}_k(\xi_0) + \overline{B}_k(\xi_0) = a_k + b_k$$

and therefore

$$\begin{aligned}\Psi^{(1)}_{\mathcal{S}_1,F}(Y) + \Psi^{(1)}_{\mathcal{S}_1,G}(Y) &= \sum_{k\in\mathcal{J}_U} a_k Y^k + \sum_{k\in\mathcal{J}_U} b_k Y^k \\ &= \sum_{k\in\mathcal{J}_U} (a_k + b_k) Y^k = \sum_{k\in\mathcal{J}_U} c_k Y^k = \Psi^{(1)}_{\mathcal{S}_1,H}(Y).\end{aligned}$$

Now assume $r \geq 2$. Let $A(X)$, $B(X) \in \mathbf{Z}_p[X]^*$ and let $C = A + B$.

- If $\widetilde{\nu}_{r-1,A} < \widetilde{\nu}_{r-1,B}$ and $\mathcal{S}_{r-1} = \mathcal{T}_{r,\widetilde{\nu}_{r-1,A}}$ it is clear that $\Psi^{(r-1)}_{S_{r-1},B} = 0$ and

$$\Psi^{(r-1)}_{S_{r-1},C} = \Psi^{(r-1)}_{S_{r-1},A} = \Psi^{(r-1)}_{S_{r-1},A} + \Psi^{(r-1)}_{S_{r-1},B}.$$

- If $\widetilde{\nu}_{r-1,A} > \widetilde{\nu}_{r-1,B}$ and $\mathcal{S}_{r-1} = \mathcal{T}_{r,\widetilde{\nu}_{r-1,B}}$ it is clear that $\Psi^{(r-1)}_{S_{r-1},A} = 0$ and

$$\Psi^{(r-1)}_{S_{r-1},C} = \Psi^{(r-1)}_{S_{r-1},B} = \Psi^{(r-1)}_{S_{r-1},A} + \Psi^{(r-1)}_{S_{r-1},B}.$$

- If $\widetilde{\nu}_{r-1,A} = \widetilde{\nu}_{r-1,B}$ and $\mathcal{S}_{r-1} = \mathcal{T}_{r,\widetilde{\nu}_{r-1,A}} = \mathcal{T}_{r,\widetilde{\nu}_{r-1,B}}$ then

$$\Psi^{(r-1)}_{S_{r-1},C} = \Psi^{(r-1)}_{S_{r-1},A} + \Psi^{(r-1)}_{S_{r-1},B}$$

  by induction.

It now follows that

$$\widehat{\Psi}^{(r-1)}_{C_{\alpha+ke_r}}(\xi_{r-1}) = \widehat{\Psi}^{(r-1)}_{A_{\alpha+ke_r}}(\xi_{r-1}) + \widehat{\Psi}^{(r-1)}_{B_{\alpha+ke_r}}(\xi_{r-1})$$

for all $k \in \mathcal{J}_U$. By definition we have

$$a_k = \begin{cases} \Gamma^{-1}_{S,r,k}\, \widehat{\Psi}^{(r-1)}_{A_{\alpha+ke_r}}(\xi_{r-1}) & \text{if } k \in \mathcal{J}_A, \\ 0 & \text{if } k \notin \mathcal{J}_A, \end{cases}$$

$$b_k = \begin{cases} \Gamma^{-1}_{S,r,k}\, \widehat{\Psi}^{(r-1)}_{B_{\alpha+ke_r}}(\xi_{r-1}) & \text{if } k \in \mathcal{J}_B, \\ 0 & \text{if } k \notin \mathcal{J}_B, \end{cases}$$

$$c_k = \begin{cases} \Gamma^{-1}_{S,r,k}\, \widehat{\Psi}^{(r-1)}_{C_{\alpha+ke_r}}(\xi_{r-1}) & \text{if } k \in \mathcal{J}_C, \\ 0 & \text{if } k \notin \mathcal{J}_C. \end{cases}$$

Hence $c_k = a_k + b_k$ for all $k \in \mathcal{J}_U$. □

Lemma 3. *If r, α, and β are nonnegative integers and $K(X)$ is a nonzero polynomial in $\mathbf{Z}_p[X]$ then the point*

$$\left(\alpha + ke_r,\, v_r(K\,\varphi_r^{\alpha+ke_r})\right)$$

lies on the line $\mathcal{L}_{r,\,d_r\alpha+e_r\beta}$ if and only if

$$v_r(K) = (\beta - kd_r) - (\alpha + ke_r)\,\overline{\nu}_r\,.$$

Proof. We have

$$\begin{aligned}
& (\alpha + ke_r, v_r(K\varphi_r^{\alpha+ke_r})) \in \mathcal{L}_{r,\, d_r\alpha+e_r\beta} \\
\iff \quad & d_r(\alpha + ke_r) + e_r v_r(K\varphi_r^{\alpha+ke_r}) = d_r\alpha + e_r\beta \\
\iff \quad & kd_re_r + e_r v_r(K\varphi_r^{\alpha+ke_r}) = e_r\beta \\
\iff \quad & kd_r + v_r(K\varphi_r^{\alpha+ke_r}) = \beta \\
\iff \quad & kd_r + v_r(K) + (\alpha + ke_r)\,\overline{\nu}_r = \beta \\
\iff \quad & v_r(K) = (\beta - kd_r) - (\alpha + ke_r)\,\overline{\nu}_r\,.
\end{aligned}$$

□

Lemma 4. *If $r \geq 1$ and $\nu \geq \overline{\nu}_{r+1}$ and $0 \leq i \leq f_r - 1$ then*

$$(\beta_{r,\nu} - id_r) - (\alpha_{r,\nu} + ie_r)\,\overline{\nu}_r = \frac{\nu - (\alpha_{r,\nu} + ie_r)\,\overline{\mu}_{r+1}}{e_r} > \overline{\nu}_r.$$

Proof. Since $\nu \geq \overline{\nu}_{r+1} = e_r f_r \overline{\mu}_{r+1}$ we have

$$\begin{aligned}
e_r\big((\beta_{r,\nu} - id_r) - (\alpha_{r,\nu} + ie_r)\,\overline{\nu}_r\big) &= e_r(\beta_{r,\nu} - id_r) - e_r(\alpha_{r,\nu} + ie_r)\,\overline{\nu}_r \\
&= \beta_{r,\nu}e_r - ie_rd_r - e_r(\alpha_{r,\nu} + ie_r)\,\overline{\nu}_r \\
&= \nu - \alpha_{r,\nu}d_r - ie_rd_r - e_r(\alpha_{r,\nu} + ie_r)\,\overline{\nu}_r \\
&= \nu - (\alpha_{r,\nu} + ie_r)(e_r\overline{\nu}_r + d_r) \\
&= \nu - (\alpha_{r,\nu} + ie_r)\,\overline{\mu}_{r+1}
\end{aligned}$$

and

$$\begin{aligned}
\nu - (\alpha_{r,\nu} + ie_r)\,\overline{\mu}_{r+1} &\geq \nu - ((e_r - 1) + (f_r - 1)e_r)\,\overline{\mu}_{r+1} \\
&= \nu - (e_rf_r - 1)\,\overline{\mu}_{r+1} \\
&\geq \overline{\mu}_{r+1} \\
&= e_r\overline{\nu}_r + d_r \\
&> e_r\overline{\nu}_r.
\end{aligned}$$

□

2.4 Pseudo-valuations

Let $\psi_r(Y)$ be a monic irreducible polynomial with nonzero constant term of degree f_r in $\mathbf{F}_{q_r}[Y]$. Taking v_{ψ_r} to be the valuation associated with ψ_r we give the definition of the pseudo-valuation ω_r on $\mathbf{Z}_p[x]$ and enumerate some of its properties.

Definition 12. (*Pseudo-valuation*) Let v be a valuation of $\mathbf{Q}_p(X)$.

A *pseudo-valuation of* $\mathbf{Q}_p(X)$ *with respect to* v is a mapping

$$\omega : \mathbf{Q}_p(X)^* \to \mathbf{Z}$$

satisfying three conditions for all $F(X)$ and $G(X)$ in $\mathbf{Z}_p[X]^*$:

(1) $\omega(FG) = \omega(F) + \omega(G)$;

(2) $\omega(F/G) = \omega(F) - \omega(G)$;

(3) If $v(F) = v(G)$ and $\omega(F) \neq \omega(G)$ then

$$v(F+G) = v(F) = v(G),$$

$$\omega(F+G) = \min\{\,\omega(F), \omega(G)\,\}.$$

Proposition 6. *The mapping* $\omega_r : \mathbf{Q}_p(X)^* \to \mathbf{Z}$ *defined by*

$$\omega_r(K) = \begin{cases} v_{\psi_0}(\overline{K/p^{v_1(K)}}) & \text{if } r = 1, \\ v_{\psi_{r-1}}(\widetilde{\Psi}_K^{(r-1)}) & \text{if } r \geq 2. \end{cases}$$

is a pseudo-valuation of $\mathbf{Q}_p(X)$ *with respect to the valuation* v_r.

Proof. If $r = 1$ it is clear that ω_1 satisfies properties (1) and (2) of the definition. For property (3), assume $v_1(F) = v_1(G) = \lambda$ and $\omega_1(F) \neq \omega_1(G)$. Then

$$v_{\psi_0}\big(\overline{F/p^\lambda}\big) = \omega_1(F) \neq \omega_1(G) = v_{\psi_0}\big(\overline{G/p^\lambda}\big)$$

by assumption, and it follows from the definition of the valuation of a polynomial that $v_1(F+G) = \lambda$. It is also clear that

$$\omega_1(F+G) = v_{\psi_0}\big(\overline{(F+G)/p^\lambda}\big) = v_{\psi_0}\big(\overline{F/p^\lambda} + \overline{G/p^\lambda}\big)$$

and it follows that

$$\omega_1(F+G) = \min\{\,\omega_1(F),\,\omega_1(G)\,\}.$$

The case $r \geq 2$ is an immediate consequence of Lemma 5 below. □

Lemma 5. *If $r \geq 2$ and $F(X), G(X) \in \mathbf{Z}_p[X]^*$, with $v_r(F) = v_r(G)$ and $\omega_r(F) < \omega_r(G)$, then $v_r(F+G) = v_r(F)$ and $\omega_r(F+G) = \omega_r(F)$.*

Proof. We adopt the following notation.

- We let $\widetilde{\nu} = v_r(F) = v_r(G)$.
- We let $F = \sum_j A_j\,\varphi_{r-1}^j$ be the φ_{r-1}-adic expansion of F.
- We let $G = \sum_j B_j\,\varphi_{r-1}^j$ be the φ_{r-1}-adic expansion of G.
- We let $\mathcal{S}$ be the shortest segment containing both $\mathcal{S}_{r-1,F}$ and $\mathcal{S}_{r-1,G}$.
- For $j \geq 0$ we let $\lambda_j = (\widetilde{\nu} - \overline{\mu}_r j)/e_{r-1}$.

For $j \geq 0$ and $C_j(X) \in \mathbf{Z}_p[X]$ we observe that

$$\big(j, v_{r-1}(C_j\,\varphi_{r-1}^j)\big) \text{ lies } \left\{\begin{array}{c}\text{above}\\ \text{on}\\ \text{below}\end{array}\right\} \mathcal{L}_{r-1,\widetilde{\nu}} \text{ if and only if } \left\{\begin{array}{l} v_{r-1}(C_j) > \lambda_j,\\ v_{r-1}(C_j) = \lambda_j,\\ v_{r-1}(C_j) < \lambda_j.\end{array}\right.$$

We know $v_r(F+G) \geq \widetilde{\nu}$. If we assume $v_r(F+G) > \widetilde{\nu}$ then we have

$$\forall j : v_{r-1}(A_j) \geq \lambda_j, \quad \forall j : v_{r-1}(B_j) \geq \lambda_j, \quad \forall j : v_{r-1}(A_j + B_j) > \lambda_j,$$
$$\exists j : v_{r-1}(A_j) = \lambda_j, \quad \exists j : v_{r-1}(B_j) = \lambda_j.$$

It follows that $v_{r-1}(A_j) = \lambda_j$ if and only if $v_{r-1}(-B_j) = \lambda_j$ and therefore

$$\mathcal{S}_{r-1,F} = \mathcal{S}_{r-1,-G} = \mathcal{S}.$$

Applying Lemma 2 we have

$$\widetilde{\Psi}_F^{(r-1)} = \Psi_{\mathcal{S},F}^{(r-1)} = \Psi_{\mathcal{S},-G}^{(r-1)} + \Psi_{\mathcal{S},F+G}^{(r-1)} = \Psi_{\mathcal{S},-G}^{(r-1)} = \widetilde{\Psi}_{-G}^{(r-1)}$$

which implies $\omega_r(F) = \omega_r(-G)$, a contradiction. It follows that $v_r(F+G) = \widetilde{\nu}$.

We have $\mathcal{S}_{r-1,F+G} \subseteq \mathcal{L}_{r-1,\widetilde{\nu}}$. By definition the respective left and right endpoints of $\mathcal{S}$ are (α_L, β_L) and (α_R, β_R), where

$$\alpha_L = \min\{\,\widetilde{\alpha}_{r-1,F},\, \widetilde{\alpha}_{r-1,G}\,\}, \quad \beta_L = (\widetilde{\nu} - d_{r-1}\alpha_L)/e_{r-1},$$

$$\alpha_R = \max\{\,\widetilde{\alpha}_{r-1,F},\, \widetilde{\alpha}_{r-1,G}\,\}, \quad \beta_R = (\widetilde{\nu} - d_{r-1}\alpha_R)/e_{r-1}.$$

It is clear that if $j < \alpha_L$ or $j > \alpha_R$ then $v_{r-1}((A_j + B_j)\,\varphi_{r-1}^j) > \lambda_j$ and it follows immediately that $\mathcal{S}_{r-1,F+G} \subseteq \mathcal{S}$.

Let $H(X) = X$. Since $\mathcal{S}_{r-1,F} \subseteq \mathcal{S}_{r-1,F+G} \subseteq \mathcal{S}$ and $\mathcal{S}_{r-1,G} \subseteq \mathcal{S}_{r-1,F+G} \subseteq \mathcal{S}$ there exist nonnegative integers s, s', s'' such that

$$H^s\,\widetilde{\Psi}_{F+G}^{(r-1)} = \Psi_{S,F+G}^{(r-1)} = \Psi_{S,F}^{(r-1)} + \Psi_{S,G}^{(r-1)} = H^{s'}\widetilde{\Psi}_F^{(r-1)} + H^{s''}\widetilde{\Psi}_G^{(r-1)}.$$

Since $\gcd(\psi_{r-1}, H) = 1$ and $\omega_r(F) < \omega_r(G)$ it follows that

$$\begin{aligned}\omega_r(F+G) &= v_{\psi_{r-1}}(H^s\,\widetilde{\Psi}_{F+G}^{(r-1)}) \\ &= v_{\psi_{r-1}}(H^{s'}\widetilde{\Psi}_F^{(r-1)} + H^{s''}\widetilde{\Psi}_G^{(r-1)}) \\ &= \omega_r(F).\end{aligned}$$

□

Remark 3. We can observe easily that

- ω_1 is not a valuation; if $F(X) = -\psi_0(X)$ and $G(X) = \psi_0(X) + p$ then $\omega_1(F) = \omega_1(G) = 1$ but $\omega_1(F+G) = \omega_1(p) = 0 < \min\{\,\omega(F), \omega(G)\,\}$.

- $\omega_1(F) = 0$ if and only if $F'(\xi_0) \neq 0$, where $F'(x) = \overline{F(x)/p^{v_1(F)}}$.

- If F has φ_1-adic expansion

$$F(X) = \textstyle\sum_{i=0}^{\lfloor n/n_1 \rfloor} A_i(X)\,\varphi_1(X)^i,$$

with $n = \deg F$, $n_1 = \deg \varphi_1$, and $\lambda_i = v_1(A_i)$, then

$$\omega_1(F) = \min\{\,i : \lambda_i = v_1(F)\,\}.$$

Indeed, if $\lambda = v_1(F)$ then

$$\begin{aligned}\omega_1(F) &= v_{\psi_0}\big(\overline{F/p^\lambda}\big)\\ &= v_{\psi_0}\Big(\textstyle\sum_{j=0}^{\lfloor n/n_1\rfloor}\overline{A_j(X)\,\varphi_1(X)^j/p^\lambda}\Big)\\ &= v_{\psi_0}\Big(\textstyle\sum_{j=0}^{\lfloor n/n_1\rfloor}\big(\overline{A_j(X)/p^\lambda}\big)\,\overline{\varphi_1(X)}^j\Big)\\ &= v_{\psi_0}\Big(\textstyle\sum_{(i:\lambda=\lambda_i)}\big(\overline{A_i(X)/p^{\lambda_i}}\big)\,\psi_0(X)^i\Big)\\ &= \min\{\,i : \lambda_i = v_1(F)\,\}\end{aligned}$$

since $\sum_{(i:\lambda_i>\lambda)}\overline{A_i(X)/p^{\lambda_i}} = 0$.

Remark 4. If F is either a monic polynomial or a polynomial such that $v_1(F) = 0$, and if ψ_0 is an irreducible factor of $\overline{F}$ with degree of multiplicity a_0, then

$$\omega_1(F) = v_{\psi_0}(\overline{F}) = v_{\psi_0}(\Psi^{(0)}_{S_0,F}) = a_0.$$

Remark 5. Assume $r \geq 2$.

- $\omega_r(F) = 0$ if and only if $\psi_{r-1} \nmid \widetilde{\Psi}_F^{(r-1)}$.
- If $F(X) \in \mathbf{Z}_p[X]^*$ is a polynomial, then

$$\omega_{r-1}(F) \geq v_{\varphi_{r-1}}(F) + e_{r-1}f_{r-1}\omega_r(F).$$

Proof. By the definition of $\mathcal{N}_{r-1}(F)$ the difference $\omega_{r-1}(F) - v_{\varphi_{r-1}}(F)$ is greater than or equal to the length of the projection of $\mathcal{S}_{r-1,F}$ onto the x-axis, but by definition this is $e_{r-1}\widetilde{\gamma}_{r-1,F}$, and $\widetilde{\gamma}_{r-1,F} \geq f_{r-1}\omega_r(F)$. □

- If ξ_{r-1} is a root of ψ_{r-1} and $F(X) \in \mathbf{Z}_p[X]^*$ then

$$\omega_r(F) = 0 \iff \widetilde{\Psi}_F^{(r-1)}(\xi_{r-1}) \neq 0$$

and $\omega_r(F) = 0$ if $\deg F < n_r$.

Proof. The first affirmation is evident from the definition of ω_r. Now, if $\deg F < n_r = n_{r-1}e_{r-1}f_{r-1}$, then from the φ_{r-1}-adic expansion of F,

$$F(X) = \sum_{i<e_{r-1}f_{r-1}} A_i(X)\,\varphi_{r-1}(X)^i,$$

we get that the length of the projection of $\mathcal{S}_{r-1,F}$ on the x-axis is less than $e_{r-1}f_{r-1}$; it follows that $\deg \widetilde{\Psi}_F^{(r-1)} < f_{r-1} = \deg \psi_{r-1}$, *i.e.*, $\omega_r(F) = 0$. □

- If $1 \le k \le r$ then $\omega_k(\varphi_r) = \prod_{i=k}^{r-1} e_i f_i$.

Proof. We will use induction on k.

Let $k = 1$. By the construction of φ_r we have $\overline{\varphi}_r = \overline{\varphi}_1^{e_{r-1}f_{r-1}\cdots e_1 f_1}$ and hence

$$\omega_1(\varphi_r) = \prod_{i=1}^{r-1} e_i f_i.$$

Now we assume

$$\omega_k(\varphi_r) = \prod_{i=k}^{r-1} e_i f_i$$

for some k in the range $1 \le k \le r-1$.

By the construction of φ_r we have

$$\omega_k(\varphi_r) = e_k g_k = e_k f_k a_k = e_k f_k \omega_{k+1}(\varphi_r).$$

Therefore, applying the induction hypothesis, we have

$$\omega_{k+1}(\varphi_r) = \frac{\omega_k(\varphi_r)}{e_k f_k} = \frac{1}{e_k f_k} \prod_{i=k}^{r-1} e_i f_i = \prod_{i=k+1}^{r-1} e_i f_i.$$ □

- If $F(X)$ has φ_r-adic expansion

$$F(X) = \sum_{i=0}^{\lfloor n/n_r \rfloor} A_i(X)\,\varphi_r(X)^i$$

with $n = \deg F$, $n_r = \deg \varphi_r$, and $\lambda_i = v_r(A_i\,\varphi^i)$, then

$$\omega_r(F) = \min\{\, i : \lambda_i = v_r(F)\,\}.$$

Proof. Let $\lambda = v_r(F)$. Then

$$\begin{aligned} F(X) &= \textstyle\sum_{i=0}^{\lfloor n/n_r \rfloor} A_i(X)\,\varphi_r(X)^i \\ &= \textstyle\sum_{(i:\lambda_i=\lambda)} A_i(X)\,\varphi_r(X)^i + \sum_{(i:\lambda_i>\lambda)} A_i(X)\,\varphi_r(X)^i \\ &= G(X) + H(X). \end{aligned}$$

Thus $G(X) = \sum_{(i:\lambda_i=\lambda)} A_i(X)\,\varphi_r(X)^i$, with $v_r(A_i\,\varphi_r^i) = \lambda$ for each i in this summation. It is clear that $v_r(G) = \lambda$. By the above properties of ω_r we get

$$\omega_r(A_i\,\varphi_r^i) = 0 + i \cdot 1 = i$$

and since ω_r is a pseudo-valuation we have

$$\omega_r(G) = \min\{\,\omega_r(A_i\,\varphi_r^i) \mid \lambda_i = \lambda\,\} = \min\{\,i \mid \lambda_i = \lambda\,\}.$$

Now we can see that $v_r(H) = v_r(F - G) > \lambda$. Therefore the points of $\mathcal{N}_{r-1}(H)$ are above the line $\mathcal{L}_{r-1,G}$ so we have

$$\mathcal{S}_{r-1,F} = \mathcal{S}_{r-1,G} = \mathcal{S}$$

which implies

$$\widetilde{\Psi}_F^{(r-1)} = \Psi_{\mathcal{S},F}^{(r-1)} = \Psi_{\mathcal{S},G}^{(r-1)} + \Psi_{\mathcal{S},H}^{(r-1)} = \Psi_{\mathcal{S},G}^{(r-1)} = \widetilde{\Psi}_G^{(r-1)}.$$

It follows from the definition of ω_r that $\omega_r(F) = \min\{\,i : \lambda_i = \lambda\,\}$. □

2.5 A Fundamental Property of Valuations

Proposition 7. *If $r \geq 1$ and $F(X) \in \mathbf{Z}_p[X]$ has φ_r-adic expansion*

$$F(X) = A_{r,0}(X) + A_{r,1}(X)\,\varphi_r(X) + \cdots + A_{r,n}(X)\,\varphi_r(X)^n$$

then $v_r(F) = \min_{0 \leq j \leq n}\{\, v_r(A_{r,j}\,\varphi_r^j)\,\}$.

Proof. Let $y_{\min} = \min_{0 \leq j \leq n}\{\, v_r(A_{r,j}\,\varphi_r^j)\,\}$ and let

$$\mathcal{J} = \{\, j \mid 0 \leq j \leq n,\ v_r(A_{r,j}\,\varphi_r^j) = y_{\min}\,\}.$$

The φ_r-adic expansion of F can be written

$$F = H + K$$

with $H = \sum_{j \in \mathcal{J}} A_{r,j}\,\varphi_r^j$ and $K = \sum_{j \notin \mathcal{J}} A_{r,j}\,\varphi_r^j$. We have

$$v_r(F - H) = v_r(K) = v_r\big(\textstyle\sum_{j \notin \mathcal{J}} A_{r,j}\,\varphi_r^j\big) \geq \min_{j \notin \mathcal{J}}\{\, v_r(A_{r,j}\,\varphi_r^j)\,\} > y_{\min}.$$

From the definition of ω_r and Proposition 9 we have

$$\omega_r(\varphi_r) = v_{\psi_{r-1}}(\widetilde{\Psi}_{\varphi_r}^{(r-1)}) = v_{\psi_{r-1}}(c_{r-1}\psi_{r-1}) = 1.$$

For $j \in \mathcal{J}$ we know that $0 \leq \deg A_{r,j} < \deg \varphi_r = e_{r-1}f_{r-1}\deg \varphi_{r-1}$ and therefore the φ_{r-1}-adic expansion of $A_{r,j}$ is of the form

$$A_{r,j}(X) = \textstyle\sum_{k=0}^{e_{r-1}f_{r-1}-1} B_{r,j,k}(X)\,\varphi_{r-1}(X)^k.$$

Hence the integer points on $S_{r-1,A_{r,j}}$ must all belong to the set

$$\{\, (\widetilde{\alpha}_{r-1,A_{r,j}} + k\,e_{r-1},\ \widetilde{\beta}_{r-1,A_{r,j}} - k\,d_{r-1}) \mid 0 \leq k \leq f_{r-1} - 1\,\}$$

which implies $\deg \widetilde{\Psi}_{A_{r,j}}^{(r-1)} < f_{r-1} = \deg \psi_{r-1}$ and it follows that

$$\omega_r(A_{r,j}) = v_{\psi_{r-1}}(\widetilde{\Psi}_{A_{r,j}}^{(r-1)}) = 0.$$

For $j,\ k \in \mathcal{J}$ with $j < k$ we have

$$v_r(A_{r,j}\,\varphi_r^j) = v_r(A_{r,k}\,\varphi_r^k) = y_{\min},$$

$$\omega_r(A_{r,j}\,\varphi_r^j) = j < k = \omega_r(A_{r,k}\,\varphi_r^k).$$

It follows from Lemma 5 that $v_r(H) = y_{\min} < v_r(K)$ and therefore

$$v_r(F) = v_r(H + K) = v_r(H) = y_{\min} = \min_j\{\, v_r(A_{r,j}\,\varphi_r^j)\,\}. \qquad \square$$

2.6 Algorithm 1 and the Construction of φ_r

Algorithm 1. *Given d_s, e_s, f_s, etc., for $1 \le s \le r$ and given*

- *an integer t in the range $1 \le t \le r$,*
- *an integer $\nu \ge \overline{\nu}_{t+1}$,*
- *a nonzero polynomial $\delta(Y) \in \mathbf{F}_{q_t}[Y]$ of degree less than f_t,*

to construct a polynomial $H_{t,\nu,\delta}(X) \in \mathbf{Z}_p[X]$ such that

- $\deg H_{t,\nu,\delta} < n_{t+1}$,
- $v_{t+1}(H_{t,\nu,\delta}) = \nu$,
- $\Psi^{(t)}_{\mathcal{T}_{t,\nu},\, H_{t,\nu,\delta}}(Y) = \delta(Y)$.

Construction. Let $\zeta_0, \ldots, \zeta_{f_t-1}$ in $\mathbf{F}_{q_t}$ be such that

$$\delta(Y) = \textstyle\sum_{i=0}^{f_t-1} \zeta_i Y^i.$$

Since $\delta(Y) \neq 0$ the set $J_\delta = \{\, i \mid 0 \le i \le f_t - 1,\ \zeta_i \neq 0 \,\}$ is not empty.

For $i \in J_\delta$ we construct $K_i(X)$ as follows.

- We take $\delta_i(Y)$ to be the unique polynomial in $\mathbf{F}_{q_{t-1}}[Y]$ of degree less than f_{t-1} such that
$$\delta_i(\xi_{t-1}) = \Gamma_{\mathcal{T}_{t,\nu},t,i}\, \zeta_i.$$
- If $t = 1$ we take $P_i(X)$ to be a polynomial in $\mathbf{Z}_p[X]$ of degree less than f_0 such that $\overline{P}_i(Y) = \delta_i(Y)$ and we set
$$K_i(X) = p^{\beta_{1,\nu} - id_1} P_i(X).$$
- If $t \ge 2$ we let $\nu_i = (\beta_{t,\nu} - id_t) - (\alpha_{t,\nu} + ie_t)\overline{\nu}_t$ and we set
$$K_i(X) = H_{t-1,\,\nu_i,\,\delta_i}(X).$$

Having constructed $K_i(X)$ for $i \in J_\delta$, we set

$$H_{t,\nu,\delta}(X) = \sum_{i \in J_\delta} K_i(X)\, \varphi_t(X)^{\alpha_{t,\nu}+ie_t}. \qquad \square$$

Proposition 8 (Montes, Proposition 3.2). *Algorithm 1 is correct.*

Proof. We will proceed by induction on t.

If $t = 1$ then we have $H_{1,\nu,\delta}(X) = \sum_{i \in J_\delta} p^{\beta_{1,\nu}-id_1} P_i(X)\, \varphi_1(X)^{\alpha_{1,\nu}+ie_1}$.

- By assumption $\nu \geq \overline{\nu}_2 = e_1 f_1 d_1$. Hence for $i \in J_\delta$ we have

$$\begin{aligned}(\beta_{1,\nu} - id_1)e_1 &= \beta_{1,\nu}\, e_1 - id_1 e_1 \\ &= \nu - \alpha_{1,\nu}\, d_1 - id_1 e_1 \\ &\geq \nu - (e_1 - 1)d_1 - (f_1 - 1)d_1 e_1 \\ &= \nu - e_1 f_1 d_1 + d_1 \\ &> 0.\end{aligned}$$

  It follows that $\beta_{1,\nu} - id_1 > 0$ for all $i \in J_\delta$ and therefore $H_{1,\nu,\delta}(X) \in \mathbf{Z}_p[X]$.

- We have

$$\begin{aligned}\deg H_{1,\nu,\delta} &\leq \max_{i \in J_\delta} \{ \deg P_i + (\alpha_{1,\nu} + ie_1) \deg \varphi_1 \} \\ &< n_1 + (e_1 - 1 + (f_1 - 1)e_1)n_1 = e_1 f_1 n_1 = n_2.\end{aligned}$$

- For $i \in J_\delta$ we have $\overline{P}_i(\xi_0) = \zeta_i \neq 0$ so that $v_2(P_i) = 0$. It follows that

$$\begin{aligned}v_2(H_{1,\nu,\delta}) &= \min_{i \in J_\delta} \{ v_2\big(p^{\beta_{1,\nu}-id_1} P_i\, \varphi_1^{\alpha_{1,\nu}+ie_1}\big) \} \\ &= \min_{i \in J_\delta} \{ (\beta_{1,\nu} - id_1)v_2(p) + v_2(P_i) + (\alpha_{1,\nu} + ie_1)v_2(\varphi_1) \} \\ &= \min_{i \in J_\delta} \{ (\beta_{1,\nu} - id_1)e_1 + (\alpha_{1,\nu} + ie_1)d_1 \} \\ &= \beta_{1,\nu} e_1 + \alpha_{1,\nu} d_1 \\ &= \nu.\end{aligned}$$

- We have

$$H_{1,\nu,\delta}(X) = \sum_{i\in J_\delta} A_{\alpha_{1,\nu}+ie_1}(X)\,\varphi_1(X)^{\alpha_{1,\nu}+ie_1}$$

with $A_{\alpha_{1,\nu}+ie_1}(X) = p^{\beta_{1,\nu}-id_1}P_i(X)$ for $i \in J_\delta$. It follows that

$$\Psi^{(1)}_{T_{1,\nu},\,H_{1,\nu,\delta}}(Y) = \sum_{i\in J_\delta} \overline{B}_i(\xi_0)\,Y^i$$

with

$$B_i(X) = \frac{A_{\alpha_{1,\nu}+ie_1}(X)}{p^{\beta_{1,\nu}-(\alpha_{1,\nu}+ie_1-\alpha_{1,\nu})d_1/e_1}} = \frac{p^{\beta_{1,\nu}-id_1}P_i(X)}{p^{\beta_{1,\nu}-id_1}} = P_i(X)$$

for $i \in J_\delta$, and therefore

$$\Psi^{(1)}_{T_{1,\nu},\,H_{1,\nu,\delta}}(Y) = \delta(Y).$$

Now we assume Algorithm 1 is correct for $t = s-1$ for some $s \geq 2$. By this assumption, if $\mu \geq \overline{\nu}_s$ and $\eta(Y)$ is any nonzero polynomial in $\mathbf{F}_{q_{s-1}}[Y]$ of degree less than f_{s-1} then Algorithm 1 returns a polynomial $H_{s-1,\mu,\eta}(X) \in \mathbf{Z}_p[X]$ with

$$\deg H_{s-1,\mu,\eta} < n_s, \quad v_s(H_{s-1,\mu,\eta}) = \mu, \quad \Psi^{(s-1)}_{T_{s-1,\mu},\,H_{s-1,\mu,\eta}}(Y) = \delta(Y).$$

We will prove the algorithm is correct for $t = s$.

From Lemma 4 we have $\nu_i > \overline{\nu}_s$ and thus the construction gives

$$H_{s,\nu,\delta}(X) = \sum_{i\in J_\delta} H_{s-1,\nu_i,\delta_i}(X)\,\varphi_s(X)^{\alpha_{s,\nu}+ie_s}.$$

- Since $H_{s-1,\nu_i,\delta_i}(X) \in \mathbf{Z}_p[X]$ for $i \in J_\delta$, it is clear that $H_{s,\nu,\delta}(X) \in \mathbf{Z}_p[X]$.

- Since $\alpha_{s,\nu} \leq e_s - 1$ we have

$$\begin{aligned}\deg H_{s,\nu,\delta} &\leq \max_{i\in J_\delta}\{\deg H_{s-1,\nu_i,\delta_i} + (\alpha_{s,\nu}+ie_s)\deg\varphi_s\}\\ &< n_s + ((e_s-1)+(f_s-1)e_s)n_s = e_sf_sn_s = n_{s+1}.\end{aligned}$$

- Since $\deg H_{s-1,\nu_i,\delta_i} < n_s \leq n_{s+1}$ we have

$$v_{s+1}(H_{s-1,\nu_i,\delta_i}) = e_sv_s(H_{s-1,\nu_i,\delta_i}) = e_s\nu_i$$

and from Lemma 4 we have $e_s\nu_i + (\alpha_{s,\nu} + ie_s)\overline{\mu}_{s+1} = \nu$. It follows that

$$\begin{aligned} v_{s+1}(H_{s,\nu,\delta}) &= \min{}_{i\in J_\delta}\left\{ v_{s+1}(H_{s-1,\nu_i,\delta_i}\,\varphi_s^{\alpha_{s,\nu}+ie_s}) \right\} \\ &= \min{}_{i\in J_\delta}\left\{ v_{s+1}(H_{s-1,\nu_i,\delta_i}) + (\alpha_{s,\nu} + ie_s)v_{s+1}(\varphi_s) \right\} \\ &= \min{}_{i\in J_\delta}\left\{ e_s\nu_i + (\alpha_{s,\nu} + ie_s)\overline{\mu}_{s+1} \right\} \\ &= \nu. \end{aligned}$$

- For $i \in J_\delta$ the polynomial H_{s-1,ν_i,δ_i} was constructed so that
  - $v_s(H_{s-1,\nu_i,\delta_i}) = \nu_i$,
  - $S_{s-1,H_{s-1,\nu_i,\delta_i}} \subseteq \mathcal{T}_{s-1,\nu_i}$,
  - $\Psi^{(s-1)}_{T_{s-1,\nu_i},H_{s-1,\nu_i,\delta_i}}(Y) = \delta_i(Y)$.

  Writing $\widehat{\alpha}_s, \nu, \delta$ for $(\widetilde{\alpha}_{s,H_{s,\nu,\delta}} - \alpha_{s,\nu})/e_r$ it follows that

$$\begin{aligned} \Psi^{(s)}_{T_{s,\nu},H_{s,\nu,\delta}}(Y) &= Y^{\widehat{\alpha}_{s,\nu,\delta}}\,\Psi^{(s)}_{S_{s,H_{s,\nu,\delta}},H_{s,\nu,\delta}}(Y) \\ &= Y^{\widehat{\alpha}_{s,\nu,\delta}}\,\textstyle\sum_{i\in J_\delta}\,\Gamma^{-1}_{T_{s,\nu},s,i}\,\Psi^{(s-1)}_{T_{s-1,\nu_i},H_{s-1,\nu_i,\delta_i}}(\xi_{s-1})\,Y^{i-\widehat{\alpha}_{s,\nu,\delta}} \\ &= \textstyle\sum_{i\in J_\delta}\,\Gamma^{-1}_{T_{s,\nu},s,i}\,\Psi^{(s-1)}_{T_{s-1,\nu_i},H_{s-1,\nu_i,\delta_i}}(\xi_{s-1})\,Y^{i} \\ &= \textstyle\sum_{i\in J_\delta}\,\Gamma^{-1}_{T_{s,\nu},s,i}\,\delta_i(\xi_{s-1})\,Y^{i} \\ &= \textstyle\sum_{i=0}^{f_s-1}\,\zeta_i\,Y^{i} \\ &= \delta(Y). \end{aligned}$$

□

Proposition 9. *Let d_s, e_s, f_s, φ_s, ψ_s, etc., be given for $1 \le s \le r$, let*

$$\gamma_r(Y) = c_r(\psi_r(Y) - Y^{f_r})$$

where $c_r = \Omega_r^{-e_r f_r} \in \mathbf{F}^*_{q_r}$ and let

$$\varphi_{r+1}(X) = \varphi_r(X)^{e_r f_r} + H_{r,\overline{\nu}_{r+1},\gamma_r}(X).$$

Then $\varphi_{r+1}(X)$ is a monic polynomial in $\mathbf{Z}_p[X]$ with the following properties.

- $\deg \varphi_{r+1} = n_{r+1}$.

- $\mathcal{N}_r(\varphi_{r+1})$ *consists of the single segment* $\mathcal{S}_{r,\varphi_{r+1}}$.
- $v_{r+1}(\varphi_{r+1}) = \overline{\nu}_{r+1}$.
- $\widetilde{\Psi}^{(r)}_{\varphi_{r+1}}(Y) = c_r\psi_r(Y)$.
- φ_{r+1} *is irreducible over* $\mathbf{Z}_p$.

Proof. Let $\zeta_0, \ldots, \zeta_{f_r-1} \in \mathbf{F}_{q_r}$ be such that $\gamma_r(Y) = \sum_{i=0}^{f_r-1} \zeta_i Y^i$ and define

$$J_{\gamma_r} = \{\, i \mid 0 \le i \le f_r - 1,\ \zeta_i \ne 0 \,\}.$$

By the definition of the associated polynomial ψ_r has nonzero constant term, so $\zeta_0 = c_r\psi_r(0) \ne 0$ and hence $0 \in J_{\gamma_r}$.

- Since $\varphi_r(X) \in \mathbf{Z}_p[X]$ and $H_{r,\overline{\nu}_{r+1},\gamma_r}(X) \in \mathbf{Z}_p[X]$ it is clear that

  $$\varphi_{r+1}(X) = \varphi_r(X)^{e_r f_r} + H_{r,\overline{\nu}_{r+1},\gamma_r}(X) \in \mathbf{Z}_p[X]$$

  and since $\deg H_{r,\overline{\nu}_{r+1},\gamma_r} < n_{r+1}$ it is clear that φ_{r+1} is monic with

  $$\deg \varphi_{r+1} = e_r f_r \deg \varphi_r = e_r f_r n_r = n_{r+1}.$$

- The φ_r-adic expansion of $\varphi_{r+1}(X)$ is

  $$\varphi_{r+1}(X) = \textstyle\sum_{i \in J_{\gamma_r}} K_i(X)\, \varphi_r(X)^{\alpha_{r,\overline{\nu}_{r+1}} + ie_r} + \varphi_r(X)^{e_r f_r}$$

  and therefore $\mathcal{N}_r(\varphi_{r+1})$ is the lower convex hull of the set

  $$S = \{\, (\alpha_{r,\overline{\nu}_{r+1}} + ie_r,\ v_r(K_i) + (\alpha_{r,\overline{\nu}_{r+1}} + ie_r)\overline{\nu}_r) \mid i \in J_{\gamma_r} \,\} \cup \{\, (e_r f_r,\ e_r f_r \overline{\nu}_r) \,\}.$$

  If $r = 1$ then for $i \in J_{\gamma_1}$ we have $K_i(X) = p^{\beta_{1,\overline{\nu}_2} - id_1} P_i(X)$ with $v_1(P_i) = 0$ and, since $\overline{\nu}_1 = 0$ and $v_1(p) = 1$, it follows that

  $$\begin{aligned} v_1(K_i) + (\alpha_{1,\overline{\nu}_2} + ie_1)\overline{\nu}_1 &= v_1(K_i) \\ &= (\beta_{1,\overline{\nu}_2} - id_1)v_1(p) + v_1(P_i) \\ &= \beta_{1,\overline{\nu}_2} - id_1. \end{aligned}$$

If $r \geq 2$ then for $i \in J_{\gamma_r}$ we have $K_i(X) = H_{r-1,\nu_i,\delta_i}(X)$, so that

$$v_r(K_i) = \nu_i = (\beta_{r,\overline{\nu}_{r+1}} - id_r) - (\alpha_{r,\overline{\nu}_{r+1}} + ie_r)\overline{\nu}_r$$

and thus

$$v_r(K_i) + (\alpha_{r,\overline{\nu}_{r+1}} + ie_r)\overline{\nu}_r = \beta_{r,\overline{\nu}_{r+1}} - id_r.$$

In either case we have

$$S = \{\, (\alpha_{r,\overline{\nu}_{r+1}} + ie_r,\ \beta_{r,\overline{\nu}_{r+1}} - id_r) \mid i \in J_{\gamma_r} \,\} \cup \{\, (e_r f_r,\ e_r f_r \overline{\nu}_r) \,\}.$$

It is clear that $d_r\alpha + e_r\beta = \overline{\nu}_{r+1}$ for every point $(\alpha, \beta) \in S$. It follows that $\mathcal{N}_r(\varphi_{r+1})$ consists of a single segment and that this segment has endpoints $(\alpha_{r,\overline{\nu}_{r+1}}, \beta_{r,\overline{\nu}_{r+1}})$ and $(e_r f_r,\ e_r f_r \overline{\nu}_r)$. Hence $\mathcal{N}_r(\varphi_{r+1}) = \mathcal{S}_{r,\varphi_{r+1}}$.

- It follows from the definition of v_{r+1} that

$$v_{r+1}(\varphi_{r+1}) = \alpha_{r,\overline{\nu}_{r+1}} d_r + \beta_{r,\overline{\nu}_{r+1}} e_r = \overline{\nu}_{r+1}.$$

- It was shown in Example 2 that

$$\Psi^{(r)}_{\mathcal{T}_{r,\overline{\nu}_{r+1}},\varphi_r^{e_r f_r}}(Y) = \Omega_r^{-e_r f_r} Y^{f_r} = c_r Y^{f_r}.$$

It is clear that $\mathcal{S}_{r,\varphi_{r+1}} \subseteq \mathcal{T}_{r,\overline{\nu}_{r+1}}$, and since $\mathcal{S}_{r,\varphi_{r+1}}$ and $\mathcal{T}_{r,\overline{\nu}_{r+1}}$ have the same left endpoint it follows that

$$\begin{aligned}
\widetilde{\Psi}^{(r)}_{\varphi_{r+1}}(Y) &= \Psi^{(r)}_{\mathcal{S}_{r,\varphi_{r+1}},\varphi_{r+1}}(Y) \\
&= \Psi^{(r)}_{\mathcal{T}_{r,\overline{\nu}_{r+1}},\varphi_{r+1}}(Y) \\
&= \Psi^{(r)}_{\mathcal{T}_{r,\overline{\nu}_{r+1}},\varphi_r^{e_r f_r}}(Y) + \Psi^{(r)}_{\mathcal{T}_{r,\overline{\nu}_{r+1}},H_{r,\overline{\nu}_{r+1},\gamma_r}}(Y) \\
&= c_r Y^{f_r} + \gamma_r(Y) \\
&= c_r \psi_r(Y).
\end{aligned}$$

- Since the associated polynomial

$$\widetilde{\Psi}^{(r)}_{\varphi_{r+1}}(Y) = c_r \psi_r(Y)$$

is irreducible over $\mathbf{F}_{q_r}$, it follows from the theorem of the associated polynomial that φ_{r+1} is itself irreducible over $\mathbf{Z}_p$. □

By construction, the degree of any irreducible factor of F is greater than or equal to n_r. From this fact and Proposition 4.1 of (Pauli, 2001) we have the following.

Proposition 10 (Pauli). *Let $\eta_1, \ldots, \eta_N$ be the roots of monic square-free polynomial $F(X)$ in an algebraic closure of $\mathbf{Q}_p$. Assume that each irreducible factor of F has degree greater or equal to $\deg \varphi_r$ and that*

$$v(\varphi_r(\eta_j)) > \frac{2v(\operatorname{disc} F)}{N}$$

for $j = 1, \ldots, N$. Then $\deg F = n_r$ and hence $F(X)$ is irreducible over $\mathbf{Q}_p$.

See also Proposition 15 below.

Remark 6. From theorem of the polygon we have

$$v(\varphi_r(\eta_1)) = \cdots = v(\varphi_r(\eta_N)) = \frac{1}{e_1 \cdots e_{r-1}}\left(v_r(\varphi_r) + \frac{d_r}{e_r}\right) = \frac{\overline{\mu}_{r+1}}{e_1 \cdots e_r}.$$

Thus Proposition 10 gives a termination condition for the Montes algorithm:

$$\frac{\overline{\mu}_{r+1}}{e_1 \cdots e_r} > \frac{2v(\operatorname{disc} F)}{N}.$$

Corollary 3. *If $e_r f_r = e_{r+1} f_{r+1} = \cdots = e_{r+m} f_{r+m} = 1$ and*

$$\frac{1}{e_1 \cdots e_{r-1}}\left(v_r(\varphi_r) + d_r + \cdots + d_{r+m}\right) > \frac{2v(\operatorname{disc} F)}{N}$$

then $\deg F = n_r$ and hence $F(X)$ is irreducible over $\mathbf{Q}_p$.

Proof. Let $\eta_1, \ldots, \eta_N$ be as in Proposition 10.

By the properties of φ_{k+1} we know

$$v_{k+1}(\varphi_{k+1}) = e_k f_k (e_k v_k(\varphi_k) + d_k) = v_k(\varphi_k) + d_k$$

for $k = r, \ldots, r + m - 1$, and by the theorem of the polygon we have

$$v(\varphi_{k+1}(\eta_j)) = \frac{1}{e_1 \cdots e_{k-1} e_k}\left(v_{k+1}(\varphi_{k+1}) + \frac{d_{k+1}}{e_{k+1}}\right)$$

for $j = 1, \ldots, N$. Thus we can write

$$v_{r+n}(\varphi_{r+n}) = v_r(\varphi_r) + d_r + \cdots + d_{r+n-1}$$

for $n = 1, \ldots, m$. Hence for $j = 1, \ldots, N$ we have

$$\begin{aligned} v(\varphi_{r+m}(\eta_j)) &= \frac{1}{e_1 \cdots e_{r-1} e_r \cdots e_{r+m-1}} \Big(v_{r+m}(\varphi_{r+m}) + \frac{d_{r+m}}{e_{r+m}} \Big) \\ &= \frac{1}{e_1 \cdots e_{r-1}} \Big(v_r(\varphi_r) + d_r + \cdots + d_{r+m} \Big) \\ &> \frac{2v(\operatorname{disc} F)}{N}. \end{aligned}$$

By Proposition 10 we conclude that F is irreducible over $\mathbf{Q}_p$. □

Proposition 11. *The Montes algorithm terminates.*

Proof. Let $N = \deg F$. The algorithm constructs the sequence $\varphi_1, \varphi_2, \ldots$ with $\deg \varphi_r = n_r$, $1 \leq n_1 \leq n_2 \leq \ldots \leq N$, and $n_{r+1} = n_r$ if and only if $e_r f_r = 1$.

It is clear that the case $e_r f_r \geq 2$ can occur at most $\log_2 N$ times, since $n_{r+1} = e_r f_r n_r$ and N is an upper bound on n_r.

Furthermore, from Corollary 3 it is clear that the case $e_r f_r = 1$ cannot occur infinitely often, since for each root η of F the terms of the sequence

$$v(\varphi_{r+m}(\eta)) = \frac{1}{e_1 \cdots e_{r-1}} \Big(v_r(\varphi_r) + d_r + \cdots + d_{r+m} \Big)$$

are bounded above by

$$\frac{2v(\operatorname{disc} F)}{N}$$

and increase by

$$\frac{d_{r+m}}{e_1 \cdots e_{r-1}} \geq \frac{1}{e_1 \cdots e_{r-1}} \geq \frac{1}{N}$$

at each increment of m.

It follows that the Montes algorithm constructs only finitely many levels and therefore must terminate. □

2.7 Three Important Theorems

2.7.1 The Theorem of the Product

The following two theorems appear, in a more general form, as Theorem 6.1 in (Montes, 1999).

Theorem 11 (*Theorem of the Product: Segments*). *Let $r \geq 1$, let $F_1(x)$ and $F_2(x)$ be nonzero polynomials in $\mathbf{Z}_p[x]$, and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be the segments of slope $-d_r/e_r$ of the Newton polygons $\mathcal{N}_r(F_1)$ and $\mathcal{N}_r(F_2)$ respectively. Then $\mathcal{S}_1 + \mathcal{S}_2$ is the segment with slope $-d_r/e_r$ of $\mathcal{N}_r(F_1F_2)$.*

Proof. We denote

$$F = F_1F_2, \quad \mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2, \quad \mathcal{L} = \mathcal{L}_{r,\tilde{\nu}_{r,F}} = \mathcal{L}_{r,v_{r+1}(F)}.$$

We can assume that $\varphi_r \nmid F_1$ and $\varphi_r \nmid F_2$.

Let F_1 and F_2 have φ_r-adic expansions

$$F_1 = \sum_{j=0}^{s} B_j\varphi_r^j, \quad F_2 = \sum_{k=0}^{t} C_k\varphi_r^k.$$

Then

$$F = \sum_{i=0}^{s+t}\Big(\sum_{j+k=i} B_jC_k\Big)\varphi_r^i.$$

Since $\deg B_jC_k \leq 2n_r - 2$, each term $\sum_{j+k=i} B_jC_k$ has φ_r-adic expansion

$$\sum_{j+k=i} B_jC_k = D_i = D_{i,1}\varphi_r + D_{i,0}$$

with $\deg D_{i,0} \leq n_r - 1$ and $\deg D_{i,1} \leq n_r - 2$. Taking

$$D_{-1} = D_{-1,0} = D_{-1,1} = 0$$

the φ_r-adic expansion of F is

$$F = \sum_{i=0}^{s+t+1} A_i\varphi_r^i$$

with $A_i = D_{i-1,1} + D_{i,0}$ for $i = 0, \ldots, s+t+1$.

Let us denote $\Delta_i = v_r(D_i\varphi_r^i)$, with $\Delta_{-1} = \infty$, and $u_i = v_r(A_i\varphi_r^i)$.

We have to show that $\mathcal{S}$ is an edge of $\mathcal{N}_r(F)$.

Let (α_1, β_1), (α_1', β_1'), (α_2, β_2), (α_2', β_2') be the left and right endpoints of $\mathcal{S}_1$ and the left and right endpoints of $\mathcal{S}_2$ respectively. By definition

$$(\alpha, \beta) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2), \quad (\alpha', \beta') = (\alpha_1' + \alpha_2', \beta_1' + \beta_2')$$

are respectively the left and right endpoints of $\mathcal{S}$.

For $j \geq 0$ and $k \geq 0$ we have

$$\begin{aligned}
&d_r j + e_r v_j > \widetilde{\nu}_{r,F_1} \text{ if } j < \alpha_1, && d_r k + e_r w_k > \widetilde{\nu}_{r,F_2} \text{ if } k < \alpha_2,\\
&d_r j + e_r v_j = \widetilde{\nu}_{r,F_1} \text{ if } j = \alpha_1, && d_r k + e_r w_k = \widetilde{\nu}_{r,F_2} \text{ if } k = \alpha_2,\\
&d_r j + e_r v_j \geq \widetilde{\nu}_{r,F_1} \text{ if } \alpha_1 < j < \alpha_1', && d_r k + e_r w_k \geq \widetilde{\nu}_{r,F_2} \text{ if } \alpha_2 < k < \alpha_2',\\
&d_r j + e_r v_j = \widetilde{\nu}_{r,F_1} \text{ if } j = \alpha_1', && d_r k + e_r w_k = \widetilde{\nu}_{r,F_2} \text{ if } k = \alpha_2',\\
&d_r j + e_r v_j > \widetilde{\nu}_{r,F_1} \text{ if } j > \alpha_1', && d_r k + e_r w_k > \widetilde{\nu}_{r,F_2} \text{ if } k > \alpha_2'.
\end{aligned}$$

Lemma 6. *If $\gamma < \alpha$ or $\gamma > \alpha'$ then (γ, Δ_γ) lies above $\mathcal{L}$.*

Proof. For all $\gamma \geq 0$ we have

$$\begin{aligned}
\Delta_\gamma &= v_r\left(\textstyle\sum_{j+k=\gamma} B_j\, C_k\, \varphi_r^\gamma\right)\\
&\geq \min\{\, v_r(B_j\varphi_r^j) + v_r(C_k\varphi_r^k) \mid j + k = \gamma \,\}\\
&= \min\{\, v_j + w_k \mid j + k = \gamma \,\}.
\end{aligned}$$

If $\gamma < \alpha$ and $j + k = \gamma$ then either $j < \alpha_1$ or $k < \alpha_2$ and hence

$$\widetilde{\nu}_{r,F} = \widetilde{\nu}_{r,F_1} + \widetilde{\nu}_{r,F_2} < d_r j + e_r v_j + d_r k + e_r w_k = d_r\gamma + e_r(v_j + w_k).$$

Taking $j' + k' = \gamma$ such that $v_{j'} + w_{k'} = \min\{\, v_j + w_k \mid j + k = \gamma \,\}$ we have

$$\widetilde{\nu}_{r,F} < d_r\gamma + e_r(v_{j'} + w_{k'}) \leq d_r\gamma + e_r\Delta_\gamma$$

and thus (γ, Δ_γ) lies above $\mathcal{L}$.

A similar argument shows that (γ, Δ_γ) lies above $\mathcal{L}$ if $\gamma > \alpha'$. □

Lemma 7. *If* $(j,k) = (\alpha_1, \alpha_2)$ *then*

$$v_j + w_k = \beta$$

and if $j + k \leq \alpha$ *and* $(j,k) \neq (\alpha_1, \alpha_2)$ *then*

$$v_j + w_k > \beta.$$

Proof. It is clear that $v_{\alpha_1} + w_{\alpha_2} = \beta_1 + \beta_2 = \beta$.

If $j + k \leq \alpha$ and $(j,k) \neq (\alpha_1, \alpha_2)$ then $j < \alpha_1$ or $k < \alpha_2$. If $j < \alpha_1$ then

$$d_r j + e_r v_j > \widetilde{\nu}_{r,F_1}, \qquad d_r k + e_r w_k \geq \widetilde{\nu}_{r,F_2};$$

$$d_r \alpha_1 + e_r \beta_1 = \widetilde{\nu}_{r,F_1}, \qquad d_r \alpha_2 + e_r \beta_2 = \widetilde{\nu}_{r,F_2};$$

$$v_j > \frac{\widetilde{\nu}_{r,F_1} - d_r j}{e_r} = \beta_1 + \frac{d_r}{e_r}(\alpha_1 - j), \qquad w_k \geq \frac{\widetilde{\nu}_{r,F_2} - d_r k}{e_r} = \beta_2 - \frac{d_r}{e_r}(k - \alpha_2);$$

$$\begin{aligned} v_j + w_k &> \beta_1 + \frac{d_r}{e_r}(\alpha_1 - j) + \beta_2 - \frac{d_r}{e_r}(k - \alpha_2) \\ &= \beta_1 + \beta_2 + \frac{d_r}{e_r}(\alpha_1 - j - k + \alpha_2) = \beta. \end{aligned}$$

A similar argument applies if $k < \alpha_2$, and the result follows. □

Remark 7. It follows from Lemma 7 that $\Delta_\alpha = \beta$, hence (α, Δ_α) lies on $\mathcal{L}$.

Lemma 8. *If* $(j,k) = (\alpha'_1, \alpha'_2)$ *then*

$$v_j + w_k = \beta'$$

and if $j + k \geq \alpha'$ *and* $(j,k) \neq (\alpha'_1, \alpha'_2)$ *then*

$$v_j + w_k > \beta'.$$

Proof. The proof is closely similar to that of Lemma 7. □

Remark 8. It follows from Lemma 8 that $\Delta_{\alpha'} = \beta'$, hence $(\alpha', \Delta_{\alpha'})$ lies on $\mathcal{L}$.

Lemma 9. *If* $\gamma_i \geq 0$ *then* $u_\gamma \geq \min\{\Delta_{\gamma-1}, \Delta_\gamma\}$.

Proof. For $\gamma \geq 0$ the φ_r-adic expansion of D_γ is $D_{\gamma,0} + D_{\gamma,1}\varphi_r$.

It follows from Proposition 7 that

$$v_r(D_\gamma) = \min\{\, v_r(D_{\gamma,0}),\, v_r(D_{\gamma,1}\varphi_r)\,\}$$

and thus

$$\Delta_\gamma = v_r(D_\gamma\varphi_r^\gamma) = \min\{\, v_r(D_{\gamma,0}\varphi_r^\gamma),\, v_r(D_{\gamma,1}\varphi_r^{\gamma+1})\,\}.$$

Since

$$v_r(A_\gamma\varphi_r^\gamma) = v_r(D_{\gamma-1,1}\varphi_r^\gamma + D_{\gamma,0}\varphi_r^\gamma) \geq \min\{\, v_r(D_{\gamma-1,1}\varphi_r^\gamma),\, v_r(D_{\gamma,0}\varphi_r^\gamma)\,\}$$

it follows that

$$u_\gamma = v_r(A_\gamma\varphi_r^\gamma) \geq \min\{\, \Delta_{\gamma-1}, \Delta_\gamma\,\}. \qquad \square$$

Lemma 10. *If $\gamma < \alpha$ or $\gamma > \alpha'$ then (γ, u_γ) lies above $\mathcal{L}$.*

Proof. If $\gamma < \alpha$ or $\gamma > \alpha'$ then the point (γ, Δ_γ) lies above $\mathcal{L}$ and the point $(\gamma - 1, \Delta_{\gamma-1})$ lies on or above $\mathcal{L}$, *i.e.*,

$$\Delta_\gamma > \frac{\widetilde{\nu}_{r,F} - d_r\gamma}{e_r}, \qquad \Delta_{\gamma-1} \geq \frac{\widetilde{\nu}_{r,F} - d_r(\gamma-1)}{e_r} > \frac{\widetilde{\nu}_{r,F} - d_r\gamma}{e_r},$$

so that

$$u_\gamma \geq \min\{\, \Delta_{\gamma-1}, \Delta_\gamma\,\} > \frac{\widetilde{\nu}_{r,F} - d_r\gamma}{e_r}.$$

Hence (γ, u_γ) lies above $\mathcal{L}$. $\square$

Lemma 11. *It is the case that $u_\alpha = \beta$ and $u_{\alpha'} = \beta'$.*

Proof. Since $D_\alpha = \sum_{j+k=\alpha} B_jC_k$ we have

$$v_r(D_\alpha\varphi_r^\alpha - B_{\alpha_1}C_{\alpha_2}\varphi_r^\alpha) \geq \min\{\, v_j + w_k \mid j+k = \alpha,\ (j,k) \neq (\alpha_1, \alpha_2)\,\}$$
$$> v_{\alpha_1} + w_{\alpha_2} = v_r(B_{\alpha_1}C_{\alpha_2}\varphi_r^\alpha)$$

and it follows that

- $\widetilde{\nu}_{r-1,D_\alpha} = v_r(D_\alpha) = v_r(B_{\alpha_1}C_{\alpha_2}) = \widetilde{\nu}_{r-1,B_{\alpha_1}C_{\alpha_2}}$,
- $\mathcal{N}_{r-1}(D_\alpha - B_{\alpha_1}C_{\alpha_2})$ lies entirely above the line $\mathcal{L}_{r-1,\widetilde{\nu}_{r-1,D_\alpha}}$,
- $\mathcal{S}_{r-1,D_\alpha} = S_{r-1,B_{\alpha_1}C_{\alpha_2}}$.

We apply Lemma 2 to obtain

$$\Psi^{(r-1)}_{T,D_\alpha} = \Psi^{(r-1)}_{T,B_{\alpha_1}C_{\alpha_2}} + \Psi^{(r-1)}_{T,D_\alpha - B_{\alpha_1}C_{\alpha_2}} = \Psi^{(r-1)}_{T,B_{\alpha_1}C_{\alpha_2}}$$

with $\mathcal{T} = \mathcal{S}_{r-1,D_\alpha} = S_{r-1,B_{\alpha_1}C_{\alpha_2}}$. From the definition of ω_r we obtain

$$\omega_r(D_\alpha) = \omega_r(B_{\alpha_1}C_{\alpha_2}) = \omega_r(B_{\alpha_1}) + \omega_r(C_{\alpha_2}).$$

Since $\deg B_j < n_r$ and $\deg C_k < n_r$ we have $\omega_r(B_{\alpha_1}) = \omega_r(C_{\alpha_2}) = 0$ and thus

$$\omega_r(D_\alpha) = 0.$$

The φ_r-adic expansion of D_α is $D_{\alpha,0} + D_{\alpha,1}\varphi_r$, hence $v_r(D_\alpha) = v_r(D_{\alpha,0})$, and therefore

$$v_r(D_{\alpha,0}\varphi_r^\alpha) = v_r(D_\alpha\varphi_r^\alpha) = \Delta_\alpha.$$

It is clear that $v_r(D_{\alpha-1,1}\varphi_r^\alpha) \geq \Delta_{\alpha-1}$, and since $(\alpha-1, \Delta_{\alpha-1})$ lies above $\mathcal{L}$ and (α, Δ_α) lies on $\mathcal{L}$ it follows that $\Delta_{\alpha-1} > \Delta_\alpha$. Since

$$A_\alpha\varphi_r^\alpha = D_{\alpha-1,1}\varphi_r^\alpha + D_{\alpha,0}\varphi_r^\alpha$$

we conclude that $u_\alpha = v_r(A_\alpha\varphi_r^\alpha) = \Delta_\alpha = \beta$.

In the same way it can be shown that $u_{\alpha'} = \beta'$. □

Lastly, since no vertices of $\mathcal{N}_r(F)$ can lie below $\mathcal{L}$ it follows that

$$d_r\gamma + e_r u_\gamma \geq \widetilde{\nu}_{r,F}$$

if $\alpha < \gamma < \alpha'$. Hence $\mathcal{S}$ must be a segment of $\mathcal{N}_r(F)$. □

Theorem 12 (*Theorem of the Product: Associated Polynomials*). *Let F_1, F_2, $\mathcal{S}_1$, $\mathcal{S}_2$, and r be as in Theorem 11. Then*

$$\Psi^{(r)}_{\mathcal{S}_1+\mathcal{S}_2,F_1F_2}(Y) = \Psi^{(r)}_{\mathcal{S}_1,F_1}(Y)\,\Psi^{(r)}_{\mathcal{S}_2,F_2}(Y).$$

Proof. Throughout this proof we will use the notation from Theorem 11.

Let us define

$$J = \{\, i \mid (\alpha + ie_r, u_{\alpha+ie_r}) \in \mathcal{S} \,\},$$

$$J_1 = \{\, j \mid (\alpha_1 + je_r, v_{\alpha_1+je_r}) \in \mathcal{S}_1 \,\},$$

$$J_2 = \{\, k \mid (\alpha_2 + ke_r, w_{\alpha_2+ke_r}) \in \mathcal{S}_2 \,\}.$$

Using a similar argument as in Lemma 6 we have the following.

Lemma 12. *If $\alpha < \gamma < \alpha'$ then (γ, Δ_γ) lies above or on $\mathcal{L}$.*

Lemma 13. *If $i \in J$ then $u_{\alpha+ie_r} = \Delta_{\alpha+ie_r}$.*

Proof. Let $\gamma = \alpha + ie_r$. Then the case $\gamma = \alpha$ and $\gamma = \alpha'$ is clear from Lemma 7 and Lemma 8.

Since $(\gamma - 1, \Delta_{\gamma-1})$ and (γ, Δ_γ) lie above or on $\mathcal{L}$ by Lemma 12 and Lemma 6 and $u_\gamma \geq \min\{\, \Delta_{\gamma-1}, \Delta_\gamma \,\}$ by Lemma 9 we then get $u_\gamma = \Delta_\gamma$. □

Lemma 14. *Let $\alpha < \gamma < \alpha'$. Consider*

$$A_\gamma(x) = D_{\gamma-1,1}(x) + D_{\gamma,0}(x).$$

Then $v_r(D_{\gamma-1,1}) > v_r(D_{\gamma,0})$.

Proof. Let us review the following notation:

$$D_\gamma(x) = D_{\gamma,1}(x)\varphi_r(x) + D_{\gamma,0}(x)$$

$$D_{\gamma-1}(x) = D_{\gamma-1,1}(x)\varphi_r(x) + D_{\gamma-1,0}(x).$$

Then by the property of v_r we have

$$\Delta_{\gamma-1} \leq v_r(D_{\gamma-1,1}\varphi_r^\gamma).$$

On the other hand

$$\begin{aligned} u_\gamma &\geq \min\{\, v_r(D_{\gamma-1,1}\varphi_r^\gamma), v_r(D_{\gamma,0}\varphi_r^\gamma\,\}) \\ &\geq \min\{\, \Delta_{\gamma-1}, v_r(D_{\gamma,0}\varphi_r^\gamma\,\}). \end{aligned}$$

Since $u_\gamma = \Delta_\gamma < \Delta_{\gamma-1}$, we have

$$v_r(D_{\gamma-1,1}\varphi_r^\gamma) \geq \Delta_{\gamma-1} > v_r(D_{\gamma,0}\varphi_r^\gamma) \qquad \square$$

To prove the theorem we will proceed by induction on r.

Let $r = 1$. We need the following notation

$$\begin{aligned} B'_j(x) &= B_{\alpha_1+je_r}(x)/p^{\beta_1-jd_r} & C'_k(x) &= C_{\alpha_2+ke_r}(x)/p^{\beta_2-kd_r} \\ A'_i(x) &= A_{\alpha+ie_r}(x)/p^{\beta-id_r} & D'_i(x) &= D_{\alpha+ie_r}(x)/p^{\beta-id_r}. \end{aligned}$$

Then by the definition of the associated polynomial we have

$$\begin{aligned} \Psi^{(r)}_{S_1+S_2,F_1F_2}(Y) &= \textstyle\sum_{i\in J} \eta_i Y^i = \sum_{i\in J} \overline{A}'_i(\xi_0)Y^i \\ &= \textstyle\sum_{i\in J}\big(\overline{D}'_{i-1,1}(\xi_0) + \overline{D}'_{i,0}(\xi_0)\big)Y^i = \sum_{i\in J} \overline{D}'_{i,0}(\xi_0)Y^i. \end{aligned}$$

We obtain the last equality by Lemma 14.

Now we will calculate the right hand side of the required equality.

$$\begin{aligned} \Psi^{(r)}_{S_1,F_1}(Y)\,\Psi^{(r)}_{S_2,F_2}(Y) &= \big(\textstyle\sum_{j\in J_1} \eta_j Y^j\big)\big(\sum_{k\in J_2} \eta_k Y^k\big) \\ &= \textstyle\sum_{i\in J}\big(\sum_{j+k=i} \eta_j\eta_k\big)Y^i \\ &= \textstyle\sum_{i\in J}\big(\sum_{j+k=i} \overline{B}'_j(\xi_0)\overline{C}'_k(\xi_0)\big)Y^i \\ &= \textstyle\sum_{i\in J} \overline{D}'_i(\xi_0)Y^i. \end{aligned}$$

Since

$$D_i(x) = D_{i,1}(x)\varphi_r(x) + D_{i,0}(x)$$

and $u_i = \Delta_i$, then

$$D_i'(x) = D_{i,0}'(x).$$

Now we assume the theorem is correct for $r = t-1$ for some $t \geq 2$. By this assumption we have

$$\Psi^{(t-1)}_{S_1+S_2,F_1F_2}(Y) = \Psi^{(t-1)}_{S_1,F_1}(Y)\,\Psi^{(t-1)}_{S_2,F_2}(Y).$$

We will prove the theorem is correct for $r = t$.

Remark 9. For $t \geq 0$ and K and S are as above we have the following. If $\nu_k \equiv \beta - kd_t \pmod{e_{t-1}}$ then $\alpha_{t-1,\nu_k} = (\beta - kd_t)\,m_{t-1}$ **mod** e_{t-1} and therefore

$$(\alpha + ke_t)\,\Theta_1(t, t-1) + \Theta_2(S, t, k) = \big(m_{t-1}(\beta - kd_t) - \alpha_{t-1,\nu_k}\big)/e_{t-1}\,.$$

Then we have the following

$$\Gamma^{-1}_{S,t,i} = \Omega^{\alpha+ie_t}_{t-1}\xi^{\theta(t-1,i)}_{t-1},$$

with $\theta(t-1,i) = (\alpha_{t-1,\nu_i} - m_{t-1}(\beta - id_t))$ and $\nu_i = v_t(A_{\alpha+ie_t})$.

It is sufficient to show that

$$\begin{aligned}\eta_i &= \Gamma^{-1}_{S,t,i}\Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i},A_{\alpha+ie_t}}(\xi_{t-1}) \\ &= \textstyle\sum_{j+k=i}\eta_j\eta_k.\end{aligned}$$

By definition

$$\begin{aligned}\eta_i &= \Gamma^{-1}_{S,t,i}\Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i},A_{\alpha+ie_t}}(\xi_{t-1}) \\ &= \Omega^{\alpha+ie_t}_{t-1}\xi^{\theta(t-1,i)}_{t-1}\Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i},A_{\alpha+ie_t}}(\xi_{t-1}) \\ &= \Omega^{\alpha+ie_t}_{t-1}\xi^{\theta(t-1,i)}_{t-1}\Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i},D_{\alpha+ie_t,0}}(\xi_{t-1})\end{aligned}$$

The last equality is obtaind by the same observation as in the particular case $t = 1$.

To calculate the right hand side of the required equality we observe the following for $j \in J_1$, $k \in J_2$ and $j + k = i \in J$:

$$\mathcal{T}_{t-1,\nu_j}, B_{\alpha+je_t} + \mathcal{T}_{t-1,\nu_k}, C_{\alpha+ke_t} \subseteq \mathcal{T}_{t-1,\nu_i}, B_{\alpha+je_t} C_{\alpha+ke_t}$$

$$\alpha_{t-1,\nu_j} + \alpha_{t-1,\nu_k} = \alpha_{t-1,i} + \delta_{j,k} e_{t-1}$$

$$\theta(t-1,j) + \theta(t-1,k) = \theta(t-1,i) + \delta_{j,k}$$

where

$$\delta_{j,k} = \begin{cases} 1 & \text{if } \alpha_{t-1,\nu_j} + \alpha_{t-1,\nu_k} \geq e_{t-1}, \\ 0 & \text{if } \alpha_{t-1,\nu_j} + \alpha_{t-1,\nu_k} < e_{t-1}. \end{cases}$$

Then we have

$$\begin{aligned}
\textstyle\sum_{j+k=i} \eta_j \eta_k &= \textstyle\sum_{j+k=i} \big(\Omega_{t-1}^{\alpha_1+je_t} \xi_{t-1}^{\theta(j,t-1)} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_j}, B_{\alpha_1+je_t}(\xi_{t-1})} \big) \\
&\qquad \big(\Omega_{t-1}^{\alpha_2+ke_t} \xi_{t-1}^{\theta(k,t-1)} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_k}, C_{\alpha_2+ke_t}(\xi_{t-1})} \big) \\
&= \Omega_{t-1}^{\alpha+ie_t} \big(\textstyle\sum_{j+k=i} \xi_{t-1}^{\theta(j,t-1)+\theta(k,t-1)+\delta_{j,k}} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_j}+\mathcal{T}_{t-1,\nu_k}, B_{\alpha_1+je_t} C_{\alpha_1+ke_t}}(\xi_{t-1}) \big) \\
&= \Omega_{t-1}^{\alpha+ie_t} \big(\textstyle\sum_{j+k=i} \xi_{t-1}^{\theta(t-1,i)} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i}, B_{\alpha_1+je_t} C_{\alpha_1+ke_t}}(\xi_{t-1}) \big) \\
&= \Omega_{t-1}^{\alpha+ie_t} \xi_{t-1}^{\theta(t-1,i)} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i} D_{\alpha+ie_t}}(\xi_{t-1}) \\
&= \Omega_{t-1}^{\alpha+ie_t} \xi_{t-1}^{\theta(t-1,i)} \Psi^{(t-1)}_{\mathcal{T}_{t-1,\nu_i} D_{\alpha+ie_t,0}}(\xi_{t-1}).
\end{aligned}$$

The second-to-last equality follows from Lemma 2 and the last equality by the same observation as in the particular case $t = 1$. □

2.7.2 The Theorem of the Polygon

Theorem 13 (*Theorem of the Polygon*). *Let $r \geq 1$. Assume $\mathcal{N}_r(F)$ consists of a single segment and $\varphi_j \nmid F$ for $j = 0, 1, \ldots, r$. Then*

i) $\deg F = e_r g_r n_r$,

ii) *the endpoints of $\mathcal{N}_r(F)$ are*

$$\big(0,\ d_r g_r + e_r g_r v_r(\varphi_r)\big), \quad \big(e_r g_r,\ e_r g_r v_r(\varphi_r)\big).$$

iii) *if η is a root of F then*

$$v\big(\varphi_r(\eta)\big) = \frac{1}{e_0 e_1 \cdots e_{r-1}}\big(v_r(\varphi_r) + d_r/e_r\big),$$

iv) *if $g_r = 1$ then F is irreducible.*

Proof. The first two items are obvious.

To prove (iii), let $G(X) = \sum_{i=0}^{n} a_i X^i$ be the minimal polynomial of $\varphi_r(\eta)$. Using the Viete relations between the coefficients of a polynomial and its roots we get

$$v(a_0) = nv(\varphi_r(\eta)),$$

i.e.,

$$-\frac{v(a_0)}{n} = -v(\varphi_r(\eta))$$

which is the slope of $\mathcal{N}_0(G)$. The polynomial G being irreducible, $\mathcal{N}_0(G)$ is a single segment, and each point $(i, v(a_i))$ is either on that segment or above it. We can write

$$v(a_i) \geq (\beta - id) \geq v(a_0) - i\frac{v(a_0)}{n} = (n-i)v(\varphi_r(\eta)),$$

for any i, $0 \leq i \leq n$.

Taking the polynomial

$$H(X) = G(\varphi_r(X)) = \textstyle\sum_{i=0}^{n} a_i \varphi_r(X)^i \in \mathcal{O}[X]$$

we get

$$\begin{aligned} v_r(a_i\varphi_r^i) &= v_r(a_i) + iv_r(\varphi_r) \\ &= e_0 e_1 \cdots e_{r-1} v(a_i) + iv_r(\varphi_r) \\ &\geq e_0 e_1 \cdots e_{r-1}(n-i)v(\varphi_r(\eta)) + iv_r(\varphi_r) \\ &= nv_r(\varphi_r) + (n-i)(e_0 e_1 \cdots e_{r-1} v(\varphi_r(\eta)) - v_r(\varphi_r)). \end{aligned}$$

The conditions $i = 0$ or $i = n$ equals the above inequality. Hence $\mathcal{N}_r(H)$ is a single segment with slope

$$\frac{v_r(a_0) - nv_r(\varphi_r)}{-n} = -e_0 e_1 \cdots e_{r-1} v(\varphi_r(\eta)) - v_r(\varphi_r).$$

We have $F \mid H$ or $H \mid F$ since η is a root of H, thus by the Theorem of the Product $\mathcal{N}_r(F)$ and $\mathcal{N}_r(H)$ have the same slope, *i.e.*,

$$-\frac{d_r}{e_r} = -e_0 e_1 \cdots e_{r-1} v(\varphi_r(\eta)) - v_r(\varphi_r),$$

which gives the desired formula.

Statement (iv) follows from the Theorem of the Product.

Assume $F = F_1 F_2$ and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be the segments of $\mathcal{N}_r(F_1)$ and $\mathcal{N}_r(F_2)$ of slope $-d_r/e_r$, respectively, and let $\mathcal{S} = \mathcal{N}_r(F)$. By the Theorem of the Product we have $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$.

If $|\mathcal{S}_1| = |\mathcal{S}_2| = 0$ then $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ would consist of a single point, implying $g_r = 0$, which is assumed not to be the case.

If, say, $|\mathcal{S}_1| > 0$ then the x-coordinate of the right endpoint of $\mathcal{S}_1$ would be at least e_r, so that $\deg F_1 \geq e_r n_r = \deg F$, implying that $\deg F_2 = 0$. Hence F must be irreducible. □

2.7.3 The Theorem of the Associated Polynomial

Definition 13. For the statement of the crucial Theorem 14 below we need to establish some notation.

- We say the monic polynomial $G(X) \in \mathbf{Z}_p[X]$ has r-type ψ if
  - $\varphi_s(X) \nmid G(X)$,
  - $\mathcal{N}_s(G)$ consists of the single segment $\mathcal{S}_{s,G}$,
  - there exist $c_s > 0$ and $\lambda_s \in \mathbf{F}_{q_s}$ such that

$$\widetilde{\Psi}_G^{(s)}(Y) = \lambda_s \widetilde{\psi}_s(Y)^{c_s}$$

with $\lambda_0 = 1$, $\widetilde{\psi}_s = \psi_s$ if $s < r$, and $\widetilde{\psi}_r = \psi$.

for $s = 0, \ldots, r$.

- With $m_0 = 0$ and $m_i = (1/d_i) \bmod e_i$ for $1 \le i \le r$ we define

$$m_i' = \frac{m_i d_i - 1}{e_i}$$

so that $m_i d_i - m_i' e_i = 1$.

- We let $\pi_0(x) = p$ and for $1 \le i \le r$ we define

$$\Phi_i(x) = \frac{\varphi_i(x)}{\pi_{i-1}(x)^{f_{i-1} v_i(\varphi_{i-1})}}, \quad \mu_i(x) = \frac{\Phi_i(x)^{e_i}}{\pi_i(x)^{d_i}}, \quad \pi_{i+1}(x) = \frac{\Phi_i(x)^{m_i}}{\pi_i(x)^{m_i'}}.$$

Remark 10. We have $m_0 = 0$ and $m_0' = -1$, so that

$$\begin{array}{lll}
\pi_0 = p & \Phi_0 = x & \mu_0 = x \\
\pi_1 = p & \Phi_1 = \varphi_1 & \mu_1 = \dfrac{\varphi_1^{e_1}}{p^{d_1}} \\
\pi_2 = \dfrac{\varphi_1^{m_1}}{p^{m_1'}} & \Phi_2 = \dfrac{\varphi_2}{p^{d_1 f_1}} & \mu_2 = \dfrac{\varphi_2^{e_2}\, p^{m_1' d_2}}{\varphi_1^{m_1 d_2}\, p^{d_1 f_1 e_2}} \\
\vdots & \vdots & \vdots \\
\pi_k = \dfrac{\Phi_{k-1}^{m_{k-1}}}{\pi_{k-1}^{m_{k-1}'}} & \Phi_k = \dfrac{\varphi_k}{\pi_{k-1}^{f_{k-1} v_k(\varphi_{k-1})}} & \mu_k = \dfrac{\Phi_k^{e_k}}{\pi_k^{d_k}} \\
\vdots & \vdots & \vdots
\end{array}$$

Theorem 14 (*Theorem of the Associated Polynomial*). *Let $r \ge 1$. Assume $F(X)$ is monic with $(r-1)$-type ψ_{r-1} and that the factorization*

$$\widetilde{\Psi}_F^{(r)}(Y) = \lambda\, \psi_{r,1}(Y)^{c_{r,1}} \cdots \psi_{r,\gamma}(Y)^{c_{r,\gamma}}$$

is given, where $\lambda \in \mathbf{F}_{q_r}$ and $\psi_{r,1}(Y), \ldots, \psi_{r,\gamma}(Y)$ are distinct irreducible monic polynomials in $\mathbf{F}_{q_r}[Y]$ with respective degrees $f_{r,1}, \ldots, f_{r,\gamma}$. Then $F(X)$ has the factorization

$$F(X) = G_{r,1}(X) \cdots G_{r,\gamma}(X)$$

with $G_{r,1}, \ldots, G_{r,\gamma}$ satisfying the following.

- *$G_{r,i}(X)$ is a monic polynomial in $\mathcal{O}_K[X]$ of degree $e_r f_{r,i} c_{r,i} n_r$.*
- *$G_{r,i}(X)$ has r-type $\psi_{r,i}$.*

- *Let η_i be a root of $G_{r,i}$, let $\mu_{i,r} = \mu_r(\eta_i)$, let $\sigma_i \in \mathrm{Gal}(\mathbf{F}_{q_r}/\mathbf{F}_{q_0})$ be such that $\sigma_i(\xi_j) = \overline{\mu}_{i,j}$ for $1 \leq j \leq r-1$, and let $\tau_i = \sigma_i^{-1}$. Then $v_1(\mu_{i,r}) = 0$ and $\psi_{r,i}$ is the minimal polynomial of $\overline{\mu}_{i,r}^{\tau_i}$ over $\mathbf{F}_{q_r}$.*

Furthermore, if $c_{r,1} = \cdots = c_{r,\gamma} = 1$ then each of $G_{r,1}, \ldots, G_{r,\gamma}$ is irreducible.

Proof. See (Ore, 1928; Montes, Nart, 1992; Montes, 1999). □

Chapter 3

The Modified Montes Algorithm

3.1 Simplification of the Original Algorithm

Given a prime p and an irreducible monic polynomial $\Phi(x)$ in $\mathbf{Z}[x]$, the Montes algorithm finds the inertial degrees and ramification indices of the prime factors of the ideal $p\mathcal{O}_K$ in $\mathcal{O}_K$, with K the extension of $\mathbf{Q}$ generated by a root of Φ and $\mathcal{O}_K$ its ring of integers.

As is well known, this data can be derived from the factorization of $\Phi(x)$ into irreducible factors in $\mathbf{Z}_p[x]$.

In terms of execution time, the worst case for the Montes algorithm is the case with Φ irreducible over $\mathbf{Q}_p$. Each of the tests that would reveal reducibility must fail, and hence the maximum number of such tests must be performed.

Our intention is to analyze the complexity of this worst case. Hence there will be no need to consider the less time-consuming cases, and the algorithm can be simplified considerably.

Our simplified version of the algorithm, of which a full MAPLE implementation is given below, is a test for irreducibility only. Its output is `true` if Φ is irreducible and `false` otherwise, with no further information being given.

A considerable advantage to this approach is a substantial simplification in the nota-

tion required. We have in fact mostly abandoned the original notation and invented our own.

A complete MAPLE implementation of the modified algorithm appears in section 3.3 below. An extend example is given in Appendix B.

3.2 Complexity of Fundamental Operations

Notation.

We let $\langle \alpha \rangle$ denote the number of operations required to compute α.

We use the notation

$$f(n) \in O(n^{k+\epsilon})$$

as an alternative to the perhaps more familiar "soft-O" notation

$$f(n) \in O^{\sim}(n^k) \equiv f(n) \in O(n^k (\ln n)^c)$$

for some positive constant c (von zur Gathen and Gerhard, 1999).

For $n \geq 3$ and q a prime power we define the following.

$$\mathsf{L}(n) = \ln n \ln\ln n \qquad \mathsf{F}(n,q) = n\,\mathsf{M}(n)\ln(qn)$$

$$\mathsf{M}(n) = n\,\mathsf{L}(n) \qquad \mathsf{R}(n,q) = \mathsf{M}(n)\ln\ln(qn)$$

$$\mathsf{K}(q) = \mathsf{M}(\ln q)\ln\ln q$$

Arithmetic in $\mathbf{Z}_p$.

We are concerned with the reducibility of a monic polynomial $F_0(x) \in \mathbf{Z}_p[x]$ for some prime p.

Let p^{δ_Φ} denote the p-adic reduced discriminant of this polynomial (Ford, Pauli, and Roblot, 2002, Appendix A). If $F_1(x) \in \mathbf{Z}[x]$ with

$$F_1(x) \equiv F_0(x) \pmod{p^{2\delta_\Phi+1}\mathbf{Z}_p[x]}$$

then $F_0(x)$ is reducible in $\mathbf{Z}_p[x]$ if and only if $F_1(x)$ is reducible in $\mathbf{Z}_p[x]$. Thus in our computations p-adic integers can be represented as rational approximations with $2\delta_\Phi + 1$ p-adic digits of precision, *i.e.*, as rational integers reduced modulo $p^{2\delta_\Phi+1}$.

Shönhage and Strassen have shown that the time required to perform an arithmetic operation on two rational integers of length m is $O(\mathsf{M}(m))$; see (von zur Gathen and Gerhard, 1999, Ch.8, §8.3).

It follows that if we represent p-adic integers in this fashion then the cost of an arithmetic operation is

$$O(\mathsf{M}(\delta_\Phi \ln p)).$$

For clarity we will omit this factor from our subsequent complexity estimates; these estimates can therefore be interpreted as the cost in arithmetic operations in $\mathbf{Z}_p$.

Arithmetic in $\mathbf{F}_q$.

By (von zur Gathen and Gerhard, 1999, Ch.14, §14.7), a single operation in $\mathbf{F}_q$ can be performed in $O(\mathsf{K}(q))$ word operations.

Under the simplifying assumption that $\ln p \in O(1)$ we have

$$\ln q_r = f_{r-1}^* \ln p \in O(f_{r-1}^*)$$

and thus the cost of an operation in $\mathbf{F}_{q_r}$ is

$$\begin{aligned} O(\mathsf{K}(q_r)) &= O\big(\mathsf{M}(\ln q_r) \ln\ln q_r\big) \\ &\subseteq O\big(f_{r-1}^* (\ln f_{r-1}^*)^2 \ln\ln f_{r-1}^*\big) \\ &\subseteq O\big(f_{r-1}^{*(1+\epsilon)}\big). \end{aligned}$$

For $\alpha \in \mathbf{F}_{q_r}$ and any integer n the cost of computing α^n is

$$O(\ln q_r\, \mathsf{K}(q_r)) \subseteq O(f_{r-1}^* f_{r-1}^{*(1+\epsilon)}) = O(f_{r-1}^{*(2+\epsilon)})$$

since we may assume $0 \le n \le q_r - 1$.

By (Shoup, 1994, Theorem 10), the cost for constructing an irreducible polynomial of degree n over the finite field $\mathbf{F}_q$ is

$$O\big((n^2 \log n + n \log q)\, \mathsf{L}(n)\big).$$

Polynomial Arithmetic.

The number of operations required to evaluate a polynomial of degree n at a given point using Horner's rule is $O(n)$.

By (Schönhage and Strassen, 1971) and (Cantor and Kaltofen, 1991), the number of operations needed to multiply two polynomials of degree at most n is

$$O(\mathsf{M}(n)).$$

It follows that the number of operations needed to compute the m^{th} power of a polynomial of degree n is

$$O\big(nm \ln^2(nm)\big) \subseteq O\big((nm)^{1+\epsilon}\big).$$

Let q be a prime power and let $K = \mathbf{F}_q$. Then by (von zur Gathen and Gerhard, 1999, Ch 14, §14.4 and §14.5), the number of operations in K needed to factorize a polynomial of degree n over K is

$$O(\mathsf{F}(n, q))$$

and the number of operations in K needed to find all roots in K of polynomial in $K[x]$ of degree n is

$$O(\mathsf{R}(n, q)).$$

Let $\varphi(x)$ be a monic polynomial in $\mathbf{Z}_p[x]$ of degree n_φ, let $f(x)$ be a polynomial in $\mathbf{Z}_p[x]$ of degree n, and let $k_\varphi = \lfloor n/n_\varphi \rfloor$. Let $E(f, k_\varphi)$ denote the number of operations in $\mathbf{Z}_p$ needed to compute the φ-adic expansion

$$f(x) = \textstyle\sum_{i=1}^{k_\varphi} a_i(x)\, \varphi^i(x).$$

From (von zur Gathen and Gerhard, 1999, Ch 5, §5.11), we have

$$E(f, k_\varphi) \in O(k_\varphi(k_\varphi + 1)n_\varphi^2) = O(n_\varphi^2 k_\varphi^2) = O(n^2).$$

To evaluate a polynomial $f(x)$ at n points, $\theta_1, \ldots, \theta_n$ we have to compute $f(x)$ **mod** $(x - \theta_i)$, which, by (Aho, Hopcroft and Ullman, 1974, Ch.8, §8.5), requires $O(n \ln^2 n)$ arithmetic operations in $\mathbf{Z}_p$.

Matrix Arithmetic.

By (Strassen, 1969), finding the inverse of a $n \times n$ matrix over a field K requires

$$O(n^{\log_2 7}) \subseteq O(n^{2.81})$$

operations in K.

3.3 Complexity of the Modified Algorithm

We give a complete MAPLE implementation of the modified Montes algorithm, with proofs and explanatory comments interspersed.

We begin with an outline, showing the three major phases of the algorithm.

The algorithm begins in phase L_0 ("level 0"), then alternates between phase L_1 and phase L_2 ("level r", for $r = 1, 2, \ldots$).

- input: $\Phi(x) \in \mathbf{Z}[x]$ monic and irreducible, $p \in \mathbf{Z}$ prime
- output: $\begin{cases} \text{TRUE} & \text{if } \Phi(x) \text{ is irreducible over } \mathbf{Q}_p[x], \\ \text{FALSE} & \text{if } \Phi(x) \text{ is reducible over } \mathbf{Q}_p[x]. \end{cases}$

$\mathbf{L_0}$. ○ Factorize Φ modulo p:

$$\Phi \equiv \psi_{0,1}^{a_{0,1}} \cdots \psi_{0,\rho_0}^{a_{0,\rho_0}} \pmod{p}.$$

○ If $\rho_0 > 1$ then **return** FALSE.

If $\rho_0 = 1$ and $a_{0,1} = 1$ then **return** TRUE.

○ Set $r \leftarrow 0$ and define

$$\varphi_r(x) = x, \quad n_r = 1, \quad d_r = 0, \quad e_r = 1,$$

$$\psi_r = \psi_{r,1}, \quad f_r = \deg \psi_r, \quad \xi_r \text{ a root of } \psi_r.$$

○ Replace $r \leftarrow r + 1$.

$\mathbf{L_1}$. ○ If $r = 1$ let $\varphi_1(x)$ be a monic polynomial in $\mathbf{Z}[x]$ such that

$$\overline{\varphi}_1 = \psi_0.$$

○ If $r > 1$ construct H_{r-1} according to Algorithm 1 and let

$$\varphi_r = \varphi_{r-1}^{e_{r-1}f_{r-1}} + H_{r-1}.$$

○ Define

$$n_r = e_{r-1}f_{r-1}n_{r-1} = \deg \varphi_r.$$

○ If $e_{r-1}f_{r-1} = 1$ then replace $\varphi_{r-1} \leftarrow \varphi_r$ and $r \leftarrow r - 1$.

○ Go to L_2.

$\mathbf{L_2}$. ∘ If $\varphi_r = \Phi$ then **return** TRUE.

If $\varphi_r \mid \Phi$ and $\varphi_r \neq \Phi$ then **return** FALSE.

∘ Let $\mathcal{S}_{r,1}, \ldots, \mathcal{S}_{r,\delta_r}$ be the segments of $\mathcal{N}_r(\Phi)$ and let $\gamma_{r,k}+1$ be the number of points on $\mathcal{S}_{r,k}$ with integer coordinates, for $k = 1, \ldots, \delta_r$.

∘ If $\delta_r > 1$ then **return** FALSE.

If $\delta_r = 1$ and $\gamma_{r,1} = 1$ then **return** TRUE.

∘ Let $-d_r/e_r$ be the slope of $\mathcal{S}_{r,1}$ and construct

$$\widetilde{\Psi}_{\Phi}^{(r)}(x) = c_r \psi_{r,1}^{a_{r,1}} \cdots \psi_{r,\rho_r}^{a_{r,\rho_r}} \in \mathbf{F}_{q_r}[x]$$

with $c_r \in \mathbf{F}_{q_r}$ a nonzero constant.

∘ If $\rho_r > 1$ then **return** FALSE.

If $\rho_r = 1$ and $a_{r,1} = 1$ then **return** TRUE.

∘ Define

$$\psi_r = \psi_{r,1}, \quad f_r = \deg \psi_r, \quad \xi_r \text{ a root of } \psi_r.$$

∘ Replace $r \leftarrow r + 1$.

∘ Go to L_1.

```
#############################################################################
unprotect(norm,trace): unassign(norm,trace): with(linalg): with(padic,ordp):
#############################################################################

montes := proc (F0, p0)

local F1, vr, xp, u, v: global p, x, y, z:

p := p0: x := 'x': y := 'y': z := 'z':                    ##### global ######

vr := ordp(rres(F0,diff(F0,x)),p): xp := 1 + 2*vr: F1 := mods(F0,p^xp):

u := montL0(F1):

if u then v := "irreducible" else v := "reducible" fi:

printf("\n   F is %s over Q_%d.\n\n",v,p):

end:

#############################################################################
```

Given the polynomial $\Phi = F_0$ and a prime $p = p_0$, the coefficients of Φ are reduced modulo $p^{2\delta_\Phi+1}$ to yield the approximation F_1.

The procedure `montL0`(F_1) performs the algorithm proper, returning `true` if $F_1(x)$ is irreducible in $\mathbf{Z}_p[x]$ and `false` if $F_1(x)$ is reducible in $\mathbf{Z}_p[x]$.

In what follows it implicit that the results of arithmetic operations in $\mathbf{Z}$ are reduced reduced modulo $p^{2\delta_\Phi+1}$; for clarity we have suppressed these operations.

```
#############################################################################

rres := proc (f, g)                    #####  reduced resultant of f, g  #####

local A, n:  global x:

A := ihermite(sylvester(f,g,x)):  n := rowdim(A):  return(A[n,n]):

end:

#############################################################################
```

```
#############################################################################

aub := proc (A)  return(op(2,op(2,eval(A))))  end:   ###### array upper bound

#############################################################################

ival := proc (v)

###   input:  v = integer
###
###  output:  p-adic valuation of v

global p:

if v = 0 then return(+infinity) else return(ordp(v,p)) fi

end:

#############################################################################

phexp := proc (r, F)

###   input:  r = level
###           F = polynomial
###
###  output:  coefficients of phi_r-adic expansion of F (array)

global p, d, e, f, m, n, mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z:

local j, q, A, B:

if F = 0 then q := 0 else q := floor(degree(F,x)/n[r]) fi:

A := F:  B := array(0..q):

for j from 0 to q do B[j] := sort(rem(A,phi[r],x,'A')) od:

return(eval(B)):

end:

#############################################################################
```

It is clear that

$$\langle \mathrm{ival}(v) \rangle \in O(\ln(1+|v|))$$

and that for all r we have

$$\langle \mathrm{phexp}(r,F) \rangle \in O((\deg F)^2).$$

```
#############################################################################

valf := proc (r, F)

###   input:  r = level
###           F = polynomial in x
###
###  output:  v = v_r(F)

global p, d, e, f, m, n, mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z:

local j, k, v, w, A:

if r = 0 then
    v := ival(content(F)):                                    #####  Step 1
else
    A := phexp(r-1,F):                                        #####  Step 2
    k := aub(A):
    v := +infinity:
    for j from 0 to k do                                      #####  Step 3
        w := d[r-1]*j + e[r-1]*valf(r-1,A[j]):                #####  Step 4
        if w < v then v := w fi:
    od:
fi:

return(v)

end:

#############################################################################
```

Theorem 15. *For $r \geq 1$ let us define*

$$\omega(r,d) = \max\{\, \langle \mathtt{valf}(r,g) \rangle \mid g(x) \in \mathbf{Z}_p[x],\ \deg g \leq d \,\}.$$

Then $\omega(r,d) \in O(d^{2+\epsilon})$.

Proof. Assuming $\deg g \leq d$ with $d \geq n_{r-1}$, we observe the following.

1. time: $\langle \mathtt{valf}(0,g) \rangle \in O(d)$ (since $|\mathrm{content}(F)| \leq p^{2\delta_\Phi+1}$).

2. time: $\langle \mathtt{phexp}(r-1,g) \rangle \in O(d^2)$.

3. The number of iterations in the for-loop is $1 + k = 1 + \lfloor d/n_{r-1} \rfloor$.

4. time: at most $C_0 + \omega(r-1, n_{r-1}-1)$. with $C_0 \in O(1)$.

Step 4 requires at most

$$(1 + \lfloor d/n_{r-1} \rfloor)(C_0 + \omega(r-1, n_{r-1} - 1))$$

operations. It follows that

$$\omega(r, d) \leq h_\omega(r) + (1 + \lfloor d/n_{r-1} \rfloor)\omega(r-1, n_{r-1} - 1)$$

with $h_\omega(r) \in O(d^2)$.

Proposition 12. *For $r \geq 1$ it is the case that*

$$\omega(r, n_r - 1) \leq h_\omega(r) + e_{r-1} f_{r-1} \omega(r-1, n_{r-1} - 1))$$

with $h_\omega(r) \in O(n_r^2)$.

Proof. We have

$$1 + \left\lfloor \frac{n_r - 1}{n_{r-1}} \right\rfloor = \left\lceil \frac{n_r}{n_{r-1}} \right\rceil = e_{r-1} f_{r-1}$$

and therefore

$$\begin{aligned}
\omega(r, n_r) &\leq \textstyle\sum_{i=1}^{r} h_\omega(i) \prod_{j=i}^{r-1} e_j f_j + \omega(0, n_0) \prod_{j=0}^{r-1} e_j f_j \\
&\in O\big(r n_r^2 + n_r\, \omega(0, n_0)\big) \\
&\subseteq O\big(n_r^2 \log_2 n_r + n_r n_0\big) \\
&\subseteq O(n_r^2 \ln n_r) \\
&\subseteq O(n_r^{2+\epsilon}).
\end{aligned}$$

□

In the general case the for-loop at Step 3 makes at most $1 + \lfloor d/n_{r-1} \rfloor$ iterations. Hence the time-complexity of the for-loop is

$$\Big(1 + \Big\lfloor \frac{d}{n_{r-1}} \Big\rfloor\Big)\, \omega(r-1, n_{r-1} - 1) \in O\Big(\frac{d}{n_{r-1}} n_{r-1}^{2+\epsilon}\Big) = O\big(d n_{r-1}^{1+\epsilon}\big)$$

and we have

$$\omega(r, d) \in O\big(d^2 + d n_{r-1}^{1+\epsilon}\big) \subseteq O\big(d^{2+\epsilon}\big).$$

Lastly, we note that the case $d < n_{r-1}$ is simpler and yields a similar result. We omit the details. □

```
#############################################################################

valc := proc (r, A)

###   input:  r = level
###           A = array (0..k) of polynomials
###
###  output:  P = array ( [0,v_r(A_0 phi_r^0)], ..., [k,v_r(A_k phi_r^k)] )

global p, d, e, f, m, n, mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z:

local j, k, P:

k := aub(A):  P := array(0..k):

for j from 0 to k do P[j] := [ j, valf(r,A[j]) + j*nub[r] ] od:

return(eval(P)):

end:

#############################################################################
```

Let $\#A$ denote the number of entries in the array A. We have

$$\langle \texttt{P[j]} \rangle \in O\big(\omega(r, n_r)\big)$$

and therefore

$$\textstyle\sum_{j=0}^{k} \langle \texttt{P[j]} \rangle \in O\big(\sum_{j=0}^{k} \omega(r, n_r)\big) = O\big(\#A \cdot \omega(r, n_r)\big)$$

which implies

$$\big\langle \texttt{valc}(r, A) \big\rangle \in O\big(\#A \cdot \omega(r, n_r)\big) \subseteq O\big(\#A \cdot n_r^2 \ln n_r\big).$$

```
#############################################################################
zrho := proc (r, h)  global p, rho, z:                    #####  z <-- rho_r
return(subs(z=rho[r],eval(h)) mod p)
end:
#############################################################################
rhoz := proc (r, h)  global p, psl, rho, z:               #####  rho_r <-- z
if psl[r] then return(eval(h) mod p)
         else return(subs(rho[r]=z,eval(h)) mod p) fi:
end:
#############################################################################
rhoy := proc (r, h)  global p, psl, rho, y:               #####  rho_r <-- y
if psl[r] then return(eval(h) mod p)
         else return(subs(rho[r]=y,eval(h)) mod p) fi:
end:
#############################################################################
rhox := proc (r, h)  global p, psl, rho, x:               #####  rho_r <-- x
if psl[r] then return(eval(h) mod p)
         else return(subs(rho[r]=x,eval(h)) mod p) fi:
end:
#############################################################################
rhhrs := proc (r, s, h)  global p, psl, rho, rhh:  #####  rho_s <-- rhh_{r,s}
if psl[r] or r = s then return(eval(h) mod p)
                   else return(subs(rho[s]=rhh[r,s],eval(h)) mod p) fi:
end:
#############################################################################
```

The procedures zrho, rhoz, rhoy, rhox, and rhhrs are of a technical nature and were necessitated by our decision not to use the MAPLE GF package. As a consequence we were obliged to provide several different representations of elements in finite fields: as polynomials in x, y, or z, or in RootOf notation. These five procedures convert between the various representations. The details of these representations are given in Appendix A.1.

The execution time of each procedure is no worse than the time required to evaluate a polynomial, of degree d, say, at an element of $\mathbf{F}_{q_r}$. Using Horner's method, the time required for this evaluation is

$$O(df_{r-1}^{*(1+\epsilon)}).$$

```
#############################################################################

FFFacts := proc (h, r)                    #####  finite field factorization #####

###    input:  h(z,y) = polynomial in F_p[z,y], z <-- rho[r]
###                r = level
###
###   output:  factors of h mod p ([ coeff, list ]) in F_p[z,y]

global p, d, e, f, m, n, mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z:

local j, k, w:

if psl[r] then
    w := Factors(h) mod p:                                          ##### Step 1
else
    w := Factors(zrho(r,h),rho[r]) mod p:                           ##### Step 2
    w[1] := rhoz(r,w[1]):                                           ##### Step 3
    for j from 1 to nops(w[2]) do
       w[2][j][1] := rhoz(r,w[2][j][1])                             ##### Step 4
    od:
fi:

return(w)

end:

#############################################################################
```

With d_h denoting the degree of h we observe the following.

1. time: $O(\mathsf{F}(d_h, q_r)) = O(d_h\,\mathsf{M}(d_h)\ln(q_r d_h))$.

2. time: $O(d_h f_{r-1}^{*(1+\epsilon)} + \mathsf{F}(d_h, q_r))$.

3. time: $O\big(f_{r-1}^{*(1+\epsilon)}\big)$. [Note that $h(\rho_r, y) = w_1 \prod_j w_{2,j,1}(y)^{w_{2,j,2}}$.]

4. time: $O\big(\sum_j \deg w_{2,j,1}\, f_{r-1}^{*(1+\epsilon)}\big) \subseteq O\big(d_h f_{r-1}^{*(1+\epsilon)}\big)$.

Consequently

$$\big\langle \mathtt{FFFacts}(h,r) \big\rangle \in O\big(d_h\,\mathsf{M}(d_h)\ln(q_r d_h) + d_h f_{r-1}^{*(1+\epsilon)}\big).$$

Since

$$\begin{aligned} d_h\,\mathsf{M}(d_h)\ln(q_r d_h) &= d_h^2 \ln d_h \ln\ln d_h (f_{r-1}^* \ln p + \ln d_h) \\ &= d_h^2 \ln^2 d_h \ln\ln d_h + f_{r-1}^* d_h^2 \ln p \ln d_h \ln\ln d_h \end{aligned}$$

it follows that

$$\big\langle \mathtt{FFFacts}(h,r) \big\rangle \in O\big(d_h^{2+\epsilon} f_{r-1}^* + d_h f_{r-1}^{*(1+\epsilon)}\big) \subseteq O\big(d_h^{2+\epsilon} f_{r-1}^{*(1+\epsilon)}\big).$$

```
###############################################################################

randrt := proc (h, r)                    #####  finite field factorization #####

###    input:  h(z,y) = polynomial in F_p[z,y], z <-- rho[r]
###                r = level
###
###   output:  root of h (in F_p[z])

global p, d, e, f, m, n, mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z:

local s, w:

w := FFFacts(h,r):

return(simplify(y - zrho(r,w[2][1][1]) mod p))

end:

###############################################################################
```

We note in passing that the MAPLE finite-field factorization procedure returns factors in random order; hence the name `randrt`.

It is clear that

$$\big\langle \mathtt{randrt}(h,r) \big\rangle \in O\big(d_h^{2+\epsilon} f_{r-1}^{*(1+\epsilon)}\big).$$

```
#############################################################################

lchull := proc (P)  local S, j1, j2, j, k, w, w0:

###   input:  P = sequence of points  (array 0..k)
###
###  output:  S = vertices of lower convex hull (list)

k := aub(P):

for j from k by -1 to 0 do
    if P[j][2] < infinity then j1 := j fi:
od:

S := [ P[j1] ]:

while j1 < k do
    j2 := j1:  w0 := +infinity:
    for j from j1+1 to k do
        w := (P[j][2] - P[j1][2])/(P[j][1] - P[j1][1]):
        if w <= w0 then w0 := w: j2 := j fi:
    od:
    S := [ op(S), P[j2] ]:  j1 := j2:
od:

return(eval(S))

end:

#############################################################################

lcsgmt := proc (r, P)  local j, k, v, w, P1, P2:  global d, e:

###   input:  r = level
###           P = sequence of points  (array 0..k)
###
###  output:  S = segment of lower convex hull of slope -d_r/e_r

k := aub(P):  v := +infinity:

for j from 0 to k do
    w := d[r]*P[j][1] + e[r]*P[j][2]:
    if w < v then  P1 := P[j]:  v := w:  fi:
    if w = v then  P2 := P[j]:  fi:
od:

return([P1,P2]):

end:

#############################################################################
```

```
#############################################################################

fsegs := proc (S)  local Dx, Dy, g:

###   input:  S = segment (two points)
###
###  output:  g = number of "fundamental" segments in S

Dx := S[2,1] - S[1,1]:
Dy := S[2,2] - S[1,2]:

g := igcd(Dx,Dy):

return(g)

end:

#############################################################################

slopes := proc (S)  local t, k, j, w:

###   input:  S = vertices of Newton polygon (list)
###
###  output:  slopes of edges (list)

t := [ ]:
k := nops(S):
for j from 2 to k do
    w := (S[j][2] - S[j-1][2])/(S[j][1] - S[j-1][1]):
    t := [ op(t), w ]:
od:

return(eval(t))

end:

#############################################################################
```

These simple procedures provide basic operations on polygons and segments.

It is clear that

$$\langle \texttt{lchull}(P) \rangle \in O(\#P^2), \qquad \langle \texttt{fsegs}(S) \rangle \in O(1),$$

$$\langle \texttt{lcsgmt}(r, P) \rangle \in O(\#P), \qquad \langle \texttt{slopes}(S) \rangle \in O(\#S),$$

where $\#P$ and $\#S$ denote the number of points in P and S respectively.

```
##############################################################################

Upsilon := proc (r)                              #####  construct Upsilon_r  #####

###   input:  r = level (r > 0)

local Ups, fr, fs, fq, h, j, k, w:  global p, f, pss, rhh, xi, y, z:

fr := f[r]:  fs := degree(pss[r],y):  fq := fs / fr:

Ups := matrix(fs,fs):

for j from 0 to fr - 1 do                                          #####  Step 1
for k from 0 to fq - 1 do                                          #####  Step 2
    w := simplify(rhh[r,r-1]^k*xi[r]^j):                           #####  Step 3
    w := sort(rhoz(r,w),z):
    for h from 0 to fs - 1 do                                      #####  Step 4
        Ups[1+h,1+j+k*fr] := coeff(w,z,h) mod p:
    od:
od:
od:

return(evalm(Ups)):

end:

##############################################################################
```

We observe the following.

1. The number of iterations occurring in the for-loop is f_r.

2. The number of iterations occurring in the for-loop is f^*_{r-1}.

3. Since $\widehat{\rho}_{r,r-1} \in \mathbf{F}_p[\rho_r] = \mathbf{F}_{q_{r+1}}$ and $\xi_r \in \mathbf{F}_p[\rho_r] = \mathbf{F}_{q_{r+1}}$ the cost is

$$O(f_r^{*(2+\epsilon)} + f_r^{*(2+\epsilon)}) = O(f_r^{*(2+\epsilon)}).$$

4. The number of iterations in the for-loop is f^*_r.

Together these imply

$$\big\langle \texttt{Upsilon}(r) \big\rangle \in O\big(f_r f^*_{r-1}(f_r^{*(2+\epsilon)} + f^*_r)\big) = O\big(f_r^{*(3+\epsilon)}\big).$$

```
##############################################################################

Gamma := proc (S, r, k)                    #####  Gamma in F_p[rho_{r-1}]  #####

local alpha, beta, ake, bkd, T2, G:  global p, d, e, m, nub, xi, Omg:

alpha := S[1][1]:  ake := alpha + k*e[r]:
beta  := S[1][2]:  bkd := beta  - k*d[r]:

T2 := floor(m[r-1]*(bkd - ake*nub[r])/e[r-1]):

G := simplify(Omg[r]^ake * xi[r-1]^T2) mod p:

return(G):

end:

##############################################################################
```

Since $\Omega_r \in \mathbf{F}_{q_r}$ and $\xi_{r-1} \in \mathbf{F}_{q_r}$ we have

$$\big\langle \mathrm{Gamma}(S,r,k) \big\rangle \in O\big(f_{r-1}^{*(2+\epsilon)} + f_{r-1}^{*(2+\epsilon)}\big) = O\big(f_{r-1}^{*(2+\epsilon)}\big).$$

```
##############################################################################

AP := proc (r, S, A, P)

###  input:   r = level
###           S = segment of slope -d_r/e_r
###           A = phi-adic expansion
###           P = [ [ j, v(A_j phi_r^j) ] ]
###
### output:  FS = associated polynomial in F_p[rho_{r-1}][y]

global p, d, e, f, m, n,
       mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z, Omg:

local g, j, k, s, alpha, beta, gamma, ae, bd,
      nuk, Tnuk, eta, vA, J, AA, PA, FA, GA, FS:

alpha := S[1][1]:  beta := S[1][2]:  gamma := min(S[2][1],aub(P)):

g := floor( (gamma - alpha) / e[r] ):      #####  to stay within both S and P

ae := array(0..g):
bd := array(0..g):

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

J := { }:
for k from 0 to g do
    ae[k] := alpha + k*e[r]:  bd[k] := beta - k*d[r]:
    if P[ae[k]][2] = bd[k] then J := J union { k } fi:
od:

if r = 0 then  ###################################

    FS := add((A[ae[k]] mod p) * y^k, k = J):                    #####  Step 1

elif r = 1 then  #################################

    FS := add((subs(x=rho[r-1],A[ae[k]]) / p^bd[k]) * y^k, k = J):

    FS := simplify(FS) mod p:                #####  FS in Fp[y] #####  Step 2

else  ############################################           ###########  r > 1

    eta := array(0..g):

    for k in J do                                                #####  Step 3

        nuk  := bd[k] - ae[k]*nub[r]:
        Tnuk := cT(r-1,nuk):                                     #####  Step 4

        AA   := phexp(r-1,A[ae[k]]):  ###  expansion of A_ae[k] #####  Step 5

        PA   := valc(r-1,AA):        ###  v_{r-1} points (array) #####  Step 6

        FA   := AP(r-1,Tnuk,AA,PA):                              #####  Step 7
        FA   := simplify(rhhrs(r-1,r-2,FA)):

        GA   := Gamma(S,r,k):                                    #####  Step 8

        eta[k] := simplify((1/GA)*subs(y=xi[r-1],FA)) mod p:     #####  Step 9

    od:

    FS := add(eta[k] * y^k, k = J) mod p:                        ####  Step 10

fi:  ##############################################

FS := sort(FS,y):

return(eval(FS)):

end:

##############################################################################
```

Theorem 16. *For $r \geq 2$ let us define*

$$\varrho(r,d) = \max\{\, \langle \mathtt{AP}(r, S, A_{g,r}, P_{g,r}) \rangle \mid g(x) \in \mathbf{Z}_p[x],\ \deg g \leq d \,\}$$

where

- *S is a segment having slope $-d_r/e_r$ and contained in the first quadrant,*
- *$A_{g,r}$ is the array of coefficients of the φ_r-adic expansion of g,*
- *$P_{g,r}$ is the corresponding sequence of "valuation points".*

Then $\varrho(r,d) \in O(dn_r^{1+\epsilon})$.

Proof. We note that if $d = \deg g$ then the array $A_{g,r}$ has $1 + \lfloor d/n_r \rfloor$ entries,

$$A_{g,r} = \left[\, A_{g,r,0},\ A_{g,r,1},\ \ldots,\ A_{g,r,\lfloor d/n_r \rfloor} \,\right],$$

and if $d_1 \leq d_2$ then $\varrho(r,d_1) \leq \varrho(r,d_2)$.

We will first estimate $\varrho(r,d)$ for the special case $d < n_{r+1}$. In this case we have

$$\#A_{g,r} = \#P_{g,r} = 1 + \left\lfloor \frac{d}{n_r} \right\rfloor \leq 1 + \left\lfloor \frac{n_{r+1}-1}{n_r} \right\rfloor = \left\lceil \frac{n_{r+1}}{n_r} \right\rceil = e_r f_r,$$

$$\#J \leq 1 + \left\lfloor \frac{d}{n_r e_r} \right\rfloor \leq 1 + \left\lfloor \frac{n_{r+1}-1}{n_r e_r} \right\rfloor = \left\lfloor f_r + \frac{n_r e_r - 1}{n_r e_r} \right\rfloor = f_r.$$

For this special case we note the following.

1. time: $\varrho(0,d)$.
2. time: $\varrho(1,d)$.
3. The for-loop makes at most f_r iterations.
   4. time: $O(1)$.
   5. time: $\langle \mathtt{phexp}(r-1, A_{\alpha+k\epsilon_r}) \rangle \leq \pi(r-1, n_r) \in O(n_r^2)$.
   6. time: $\langle \mathtt{valc}(r-1, \mathtt{AA}) \rangle \in O(e_{r-1} f_{r-1} n_{r-1}^2 \ln n_{r-1})$
      $= O(n_r n_{r-1} \ln n_{r-1})$.

7. time: at most $\varrho(r-1, n_r - 1)$.
8. time: $\langle \texttt{Gamma}(S, r, k) \rangle \in O(f_{r-1}^{*(2+\epsilon)})$
9. time: $O(\deg \texttt{FA} \cdot f_{r-1}^{*(1+\epsilon)}) \subseteq O(f_{r-1} f_{r-1}^{*(1+\epsilon)}) \subseteq O(f_{r-1}^{*(2+\epsilon)})$.
10. time: $O(f_r)$.

Thus the time-complexity of the for-loop at Step 3 is

$$f_r\big(1 + n_r^2 + n_r n_{r-1} \ln n_{r-1} + \varrho(r-1, n_r - 1) + f_{r-1}^{*(2+\epsilon)}\big)$$
$$= f_r + f_r n_r (n_r + n_{r-1} \ln n_{r-1}) + f_r f_{r-1}^{*(2+\epsilon)} + f_r \varrho(r-1, n_r - 1)$$

and this yields the following.

Proposition 13. *For $r \geq 2$ we have*

$$\varrho(r, n_{r+1} - 1) \leq h_\rho(r) + f_r \varrho(r-1, n_r - 1)$$

with $h_\rho(r) \in O\big(f_r n_r (n_r + n_{r-1} \ln n_{r-1}) + f_r^{*(2+\epsilon)}\big)$.

Writing ϱ_r for $\varrho(r, n_{r+1} - 1)$ and h_r for $h_\rho(r)$ it follows that

$$\begin{aligned}
\varrho_r &\leq h_r + f_r \varrho_{r-1} \\
&\leq h_r + f_r(h_{r-1} + f_{r-1}\varrho_{r-2}) \\
&\leq h_r + f_r(h_{r-1} + f_{r-1}(h_{r-2} + f_{r-2}\varrho_{r-3})) \\
&\leq h_r + f_r(h_{r-1} + f_{r-1}(h_{r-2} + f_{r-2}(h_{r-3} + f_{r-3}\varrho_{r-4}))) \\
&= h_r + \tfrac{f_r^*}{f_{r-1}^*}(h_{r-1} + \tfrac{f_{r-1}^*}{f_{r-2}^*}(h_{r-2} + \tfrac{f_{r-2}^*}{f_{r-3}^*}(h_{r-3} + \tfrac{f_{r-3}^*}{f_{r-4}^*}(\varrho_{r-4})))) \\
&= \tfrac{f_r^*}{f_r^*} h_r + \tfrac{f_r^*}{f_{r-1}^*} h_{r-1} + \tfrac{f_r^*}{f_{r-2}^*} h_{r-2} + \tfrac{f_r^*}{f_{r-3}^*} h_{r-3} + \tfrac{f_r^*}{f_{r-4}^*} \varrho_{r-4} \\
&\leq \textstyle\sum_{j=0}^{r-2} \tfrac{f_r^*}{f_{r-j}^*} h_{r-j} + \tfrac{f_r^*}{f_1^*} \varrho_1 \\
&= f_r^* \varrho_1 / f_1^* + \textstyle\sum_{k=2}^{r} f_r^* h_k / f_k^*
\end{aligned}$$

with

$$\textstyle\sum_{k=2}^{r} f_r^* h_k / f_k^* \in O\big(r f_r n_r (n_r + n_{r-1} \ln n_{r-1}) + r f_r^{*(2+\epsilon)}\big)$$

(see Remark 11 below). We also have $\varrho_1 \in O(n_1 f_1^{*(1+\epsilon)})$ so that

$$f_r^* \varrho_1 / f_1^* \in O(f_r^* n_1 f_1^*) = O(f_r^* f_0^* f_1^*) \subseteq O(f_r^{*(2+\epsilon)}).$$

In the general case we know that the for-loop makes at most $1 + g_r$ iterations, with $g_r = \lfloor d/(n_r e_r) \rfloor$. In this case the time-complexity of the for-loop is

$$\begin{aligned}
&(1+g_r)(1+n_r^2+n_r n_{r-1}\ln n_{r-1}+\varrho_{r-1}+f_{r-1}^{*(2+\epsilon)}) \\
&\quad \in O\Big(\tfrac{d}{n_r e_r}\big[n_r^2+n_r n_{r-1}\ln n_{r-1}+f_{r-1}^{*(2+\epsilon)} \\
&\qquad\qquad +(r-1)f_{r-1}n_{r-1}(n_{r-1}+n_{r-2}\ln n_{r-2})+(r-1)f_{r-1}^{*(2+\epsilon)}\big]\Big) \\
&\quad \subseteq O\Big(\tfrac{d}{n_r e_r}\big[r n_r^2+r n_r n_{r-1}\ln n_{r-1}+r f_{r-1}^{*(2+\epsilon)}\big]\Big) \\
&\quad = O\Big(\tfrac{1}{e_r} r d n_r+\tfrac{1}{e_r} r d n_{r-1}\ln n_{r-1}+\tfrac{1}{e_r^*} r d f_{r-1}^{*(1+\epsilon)}\Big) \subseteq O(d n_r^{1+\epsilon}).
\end{aligned}$$

Since Step 12 requires time $O(d/(n_r e_r)) \subseteq O(d)$, the theorem follows. □

Remark 11. There exists a positive constant C such that

$$\begin{aligned}
f_r^* h_k / f_k^* &= f_r f_{r-1}\cdots f_{k+1} h_k \\
&\le C f_r f_{r-1}\cdots f_{k+1}\big(f_k n_k(n_k+n_{k-1}\ln n_{k-1})+f_k^{*(2+\epsilon)}\big) \\
&\le C\big(f_r f_{r-1}\cdots f_{k+1} f_k n_k(n_k+n_{k-1}\ln n_{k-1})+f_r^{*(2+\epsilon)}\big) \\
&\le C\big(f_r n_r(n_r+n_{r-1}\ln n_{r-1})+f_r^{*(2+\epsilon)}\big)
\end{aligned}$$

for $k = 2, \ldots, r$.

```
################################################################################

cT := proc (r, nu)  local g, a_0, b_0, a_g, b_g:  global d, e:

###  input:   r = level with d_r > 0
###          nu = positive integer
###
### output:  longest integer-endpoint segment of L_{r,nu} in first quadrant

a_0 := (nu/d[r]) mod e[r]:
b_0 := (nu - a_0 * d[r])/e[r]:

  g := floor(b_0/d[r]):

a_g := a_0 + g * e[r]:
b_g := b_0 - g * d[r]:

return([ [a_0,b_0], [a_g,b_g] ])

end:

################################################################################
```

It is clear that

$$\langle \mathrm{cT}(r,\nu) \rangle = O(1).$$

```
##############################################################################

Hr := proc (t, nu, delta)                                    #####  Algorithm 1

###  input:      t = integer in { 1, ..., r }
###             nu = integer (at least bar{nu}_{t+1})
###          delta = nonzero polynomial in F_q_t[Y] of degree < f_t
###
### output:  H_{t,nu,delta}

global p, d, e, f, m, n, nub, phi, pss, rho, x, y, z, Upsinv:

local h, Gz, Vk, Vm, fr, fs, fq,
      i, zeta, ae, bd, J, K, Ttnu, atnu, btnu, del, dnu, H:

Ttnu := cT(t,nu):  atnu := Ttnu[1][1]:  btnu := Ttnu[1][2]:

zeta := array(0..f[t]-1):

J := { }:

for i from 0 to f[t] - 1 do                                  #####  Step 1

    zeta[i] := coeff(delta,y,i) mod p:          ###  in F_q = F_p[rho_{t-1}]

    zeta[i] := simplify(zeta[i]) mod p:

    if zeta[i] <> 0 then J := J union { i } fi:

od:

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

for i in J do                                                 #####  Step 2

    ae[i] := atnu + i * e[t]:
    bd[i] := btnu - i * d[t]:

    if t = 1 then        #####  xi_0 = rho_0, Gamma_{T_{1,nu},1,i} = 1  #####

        K[i] := p^bd[i] * rhox(t-1,zeta[i]):                  #####  Step 3

    else

        fr := f[t-1]:  fs := degree(pss[t-1],y):  fq := fs / fr:

        Gz := simplify( Gamma(Ttnu,t,i) * zeta[i] ):
        Gz := rhoz(t-1,Gz):                                   #####  Step 4

        Vk := vector(fs):
        for h from 0 to fs-1 do                               #####  Step 5
            Vk[1+h] := coeff(Gz,z,h):
        od:
        Vm := evalm(Upsinv[t-1] &* Vk):                       #####  Step 6

        dnu := bd[i] - ae[i]*nub[t]:
        del := add( add( Vm[1+j+k*fr] * rho[t-1]^k,
                         k=0..fq-1 ) * y^j, j=0..fr-1 ) mod p:  #####  Step 7
        del := simplify(del) mod p:

        K[i] := Hr(t-1,dnu,del):                              #####  Step 8

    fi:

od:

H := add(K[i]*phi[t]^ae[i],i=J):                              #####  Step 9
H := sort(collect(H,x),x):

return(H):

end:

###############################################################################
```

Theorem 17. *For $t \geq 2$ let us define*

$$\kappa(t,\nu) = \max\left\{ \left\langle \mathrm{Hr}(t,\nu,\delta) \right\rangle \mid \delta(x) \in \mathbf{F}_{q_t}[x],\ \deg\delta \leq f_t \right\}.$$

Then $\kappa(t,\nu) \in O(tn_{t+1}^{3+\epsilon})$.

Proof. We observe the following.

1. The for-loop at Step 1 makes f_t iterations.
2. The for-loop at Step 2 makes at most f_t iterations.
   3. time: $\kappa(0,\nu) \in O(f_0 f_0^{*(1+\epsilon)}) = O(f_0^{*(2+\epsilon)}) = O(n_1^{2+\epsilon})$
   4. time: $O(f_{t-1}^{*(2+\epsilon)} + \mathsf{K}(q_t) + f_{t-1}^*) \subseteq O(f_{t-1}^{*(2+\epsilon)})$
   5. time: $O(f_{t-1}^*)$
   6. time: $O(f_{t-1}^{*(3+\epsilon)}) + f_{t-1}^{*\alpha} \subseteq O(f_{t-1}^{*(3+\epsilon)}$, with $2 < \alpha \leq 3$
   7. time: $O(f_{t-1}^* \mathsf{K}(q_t)) \subseteq O(f_{t-1}^* f_{t-1}^{*(1+\epsilon)}) = O(f_{t-1}^{*(2+\epsilon)})$
   8. time: at most $\kappa(t-1,\nu)$
9. time: $O(n_t \ln\ln n_t \ln e_t f_t) \subseteq O\big((e_t f_t n_t)^{(1+\epsilon)}\big) = O\big(n_{t+1}^{(1+\epsilon)}\big)$.

The time spent in the for-loop at Step 2, excluding Step 8, is

$$O\big(f_t(n_1 + f_{t-1}^{*(3+\epsilon)} + f_{t-1}^* + f_{t-1}^{*(2+\epsilon)} + f_{t-1}^{*(2+\epsilon)})\big) \subseteq O\big(f_t f_{t-1}^{*(3+\epsilon)}\big)$$

and the time spent executing Step 8 is at most

$$f_t \kappa(t-1,\nu).$$

Taking account of Step 9, it now follows that

$$\kappa(t,\nu) \leq h_\kappa(t) + f_t \kappa(t-1,\nu)$$

with

$$h_\kappa(t) \in O\big(f_t f_{t-1}^{*(3+\epsilon)} + n_{t+1}^{1+\epsilon}\big).$$

As in the proof of Proposition 13 we have

$$\kappa(t,\nu) \leq (f_t^*/f_1^*)\,\kappa(1.\nu) + \textstyle\sum_{i=2}^t (f_t^*/f_i^*)\, h_\kappa(i)$$

with

$$\textstyle\sum_{i=2}^t (f_t^*/f_i^*)\, h_\kappa(i) \in O\big(t f_t f_{t-1}^{*(3+\epsilon)} + t n_{t+1}^{1+\epsilon}\big).$$

since

$$\max\{\,(f_t^*/f_i^*)h_\kappa(i) \mid 2 \le i \le t\,\} = h_\kappa(t).$$

We also have

$$\kappa(1,\nu) \in O(n_1^{2+\epsilon} + n_2^{1+\epsilon}).$$

Since $\kappa(1,\nu) \le \kappa(t,\nu)$ we obtain

$$\kappa(t,\nu) \in O\big(tf_t^{*(3+\epsilon)} + tn_{t+1}^{1+\epsilon}\big) \subseteq O(tn_{t+1}^{3+\epsilon}). \qquad \square$$

Remark 12. The computation of $\delta_i(Y)$ in Steps 4 through 7 above is explained in Appendix A.2.

```
##############################################################################

montL0 := proc (F)                              #####  Level 0 -- Initialization

###   input:  F = monic polynomial (in x)
###
###  output:  true   if F is irreducible over Q_p
###           false  if F is reducible over Q_p

global PHI, p, d, e, f, m, n,
       mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z, Omg:

local r, FL0, vF, wF, psi:

wF := Factors(subs(x=y,F)) mod p:

if nops(wF[2]) > 1 then    ####################################################

              printf("   F has multiple irreducible factors mod p.  <<\n\n"):

    FL0 := false:                ###  reducible (==> Hensel lifting)

elif wF[2][1][2] = 1 then  ####################################################

              printf("   F is irreducible mod p.  <<\n\n"):

    FL0 := true:                 ###  irreducible mod p  ###

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

else                        ###################################################

    if subs(x=0,F) mod p = 0 then

        PHI := sort(collect(subs(x=x+1,F),x),x):
        psi := sort(collect(subs(y=y+1,wF[2][1][1]),y),y):

    else

        PHI := sort(collect(F,x),x):
        psi := sort(collect(wF[2][1][1],y),y):

    fi:

    r := 0:  d[r] := 0:  mub[r] := 0:    m[r] := 0:
             e[r] := 1:  nub[r] := 0:  Omg[r] := 1:

    phi[r] := x:      n[r] := degree(phi[r],x):

    psh[r] := psi:    f[r] := degree(psh[r],y):

    pss[r] := psi:  psl[r] := evalb(degree(pss[r],y) = 1):

    rho[r] := RootOf(psi) mod p:  xi[r] := rho[r]:  rhh[r,r] := rho[r]:

    FL0 := montL1(r+1):

fi:                         ###################################################

return(FL0):

end:

###############################################################################
```

Taking $n_\Phi = \deg F$ we have

$$\langle \mathtt{montL0}(F) \rangle \leq \mu_0(n_\Phi) + \langle \mathtt{montL1}(1) \rangle$$

with

$$\mu_0(n_\Phi) \in O(\mathsf{F}(n_\Phi, p)) = O(n_\Phi \, \mathsf{M}(n_\Phi) \ln(p n_\Phi)) \subseteq O(n_\Phi n_\Phi^{1+\epsilon}) = O(n_\Phi^{2+\epsilon}).$$

```
#############################################################################

montL1 := proc (r)

###   input:  r = level (r > 0)
###
###  output:  true   if PHI is irreducible over Q_p
###           false  if PHI is reducible over Q_p

global PHI, p, d, e, f, m, n,
       mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z, Omg:

local FL1, A, P, S, T, h, j, k, s, fs, nfs, wfs, gamma, Phi:

#####  0. Initialize  ##########################################################

mub[r] := d[r-1] + e[r-1]*nub[r-1]:
nub[r] := e[r-1]*f[r-1]*mub[r]:
  n[r] := n[r-1]*e[r-1]*f[r-1]:                                  #####  Step 1

Omg[r] := rhhrs(r-1,r-2,Omg[r-1])^(e[r-1]*f[r-1])
                        * xi[r-1]^(m[r-1]*f[r-1]*mub[r]):

Omg[r] := simplify(Omg[r]) mod p:       ###  in F_p[rho_{r-1}]  #####  Step 2

if r = 1 then  ##############################################################

    phi[r] := subs(y=x,psh[r-1]):                                #####  Step 3

else

    gamma  := Omg[r-1]^(-e[r-1]*f[r-1])*(zrho(r-2,psh[r-1]) - y^f[r-1]):

    gamma  := sort(collect(simplify(gamma),y),y) mod p:          #####  Step 4

    phi[r] := phi[r-1]^(e[r-1]*f[r-1]) + Hr(r-1,nub[r],gamma):   #####  Step 5

    phi[r] := sort(collect(phi[r],x),x):

fi:  ########################################################################

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

if (r > 1) and (e[r-1]*f[r-1] = 1) then  ####################################

    phi[r-1] := phi[r]:

    FL1 := montL2(r-1):

else

    FL1 := montL2(r):

fi:  ##########################################################################

return(FL1)

end:

##############################################################################
```

We observe the following.

1. time: $O(1)$.
2. time: $O(f_{r-1}^{*(2+\epsilon)})$.
3. time: $O(n_1)$.
4. time: $O\big(f_{r-2}^{*(1+\epsilon)} + f_{r-2}^{*(2+\epsilon)} + f_{r-1}f_{r-2}^{*(1+\epsilon)}\big) \subseteq O(f_{r-1}^{*(2+\epsilon)})$.
5. time: $O(rn_r^{3+\epsilon})$, since the time required to calculate $\varphi_{r-1}^{e_{r-1}f_{r-1}}$ is

$$O((n_{r-1}e_{r-1}f_{r-1})^{1+\epsilon}) = O(n_r^{1+\epsilon})$$

and $\big\langle \mathtt{Hr}(r-1, \overline{\nu}_r, \gamma)\big\rangle \in O(rn_r^{3+\epsilon})$.

Remark 13. If $e_{r-1}f_{r-1} > 1$ then the total time is at most

$$\mu_1(r) + \big\langle \mathtt{montL2}(r)\big\rangle$$

with $\mu_1(r) \in O(rn_r^{3+\epsilon})$.

Remark 14. The case $e_{r-1}f_{r-1} = 1$ can recur at most

$$2\,\frac{e_{r-2}^*}{n_\Phi}\,v_p(\operatorname{disc}\Phi) \;\le\; 2\,v_p(\operatorname{disc}\Phi)$$

times, where $n_\Phi = \deg\Phi$ (see Proposition 10 above).

```
##############################################################################

montL2 := proc (r)                                    #####  Continuation of montL1

###   input:  r = level (r > 0)
###
###  output:  true   if PHI is irreducible over Q_p
###           false  if PHI is reducible over Q_p

global PHI, p, d, e, f, m, n,
       mub, nub, phi, pss, psl, rho, psh, xi, rhh, x, y, z, Upsinv:

local FL2, A, P, S, T, h, j, k, s, fs, nfs, wfs,
      gamma, chig, Ng, Dg, Ag, Bg, Gg, Phi:

#####  1. Create N_r(PHI), d_r, e_r  ########################################

A := phexp(r,PHI):          ###  phi_r-adic expansion of PHI    #####  Step 1

P := valc(r,A):             ###  v_r points           (array)   #####  Step 2

S := lchull(P):             ###  vertices of N_r(PHI)  (list)   #####  Step 3

T := slopes(S):             ###  slopes of edges       (list)   #####  Step 4

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

Let $n_\Phi = \deg \Phi$. Then $\#A = \#P = 1 + \lfloor n_\Phi / n_r \rfloor$.

The execution times of Steps 1 through 4 are as follows.

1. time: $\langle \mathtt{phexp}(r, \Phi) \rangle \in O(n_\Phi^2)$.

2. time: $\langle \mathtt{valc}(r, A) \rangle \in O(n_\Phi n_r \ln n_r) \subseteq O(n_\Phi n_r^{1+\epsilon})$.

3. time: $\langle \mathtt{lchull}(1 + \lfloor n_\Phi / n_r \rfloor) \rangle \in O((n_\Phi / n_r)^2)$.

4. time: $\langle \mathtt{slopes}(S) \rangle \in O(\#S)$.

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

if A[0] = 0 then              #####  BREAK if phi_r | PHI  ####################

    if n[r] = degree(PHI,x)
    then FL2 := true:         printf("   phi_%d = PHI  <<",r):
    else FL2 := false:        printf("   A_%d,0 = 0, phi_%d =/= PHI  <<",r,r):
    fi:                       printf("\n\n"):

elif nops(T) > 1 then         #####  BREAK if > 1 segment (Newton)  ############

    FL2 := false:             printf("   #E = %a  <<\n\n",nops(T)):

elif fsegs(S) = 1 then        #####  BREAK if irreducible (Eisenstein)  #######

    FL2 := true:              printf("   -d/e = %a",op(T)):
                              printf("   g = %a  <<\n\n",fsegs(S)):

else                          #####  CONTINUE  ################################

    d[r] := numer(-T[1]):
    e[r] := denom(-T[1]):
    m[r] := (1/d[r]) mod e[r]:

    #####  2. Create tAP^r_PHI  ##############################################

    fs  := AP(r,S,A,P):       #####  associated polynomial    #####  Step 5
    wfs := FFFacts(fs,r-1):                                   #####  Step 6
    nfs := nops(wfs[2]):

    if nfs > 1 then           #####  BREAK if > 1 irreducible factor  #####

            printf("   AP_%d has %d distinct irreducible factors.  <<\n\n",
                                                                  r,nfs):
        FL2 := false:

    elif wfs[2][1][2] = 1 then  #####  BREAK if AP irreducible          #####

            printf("   AP_%d is irreducible.  <<\n\n",r):

        FL2 := true:

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

The execution times of Steps 5 and 6 are as follows.

5. time: $\varrho(r, n_\Phi) \in O(n_\Phi n_r^{1+\epsilon})$.

6. time: $\langle \mathrm{FFFacts}(r-1) \rangle \in O(d_\Psi^{2+\epsilon} f_{r-2}^{*(1+\epsilon)})$.
   with $d_\Psi = \deg \widetilde{\Psi}_\Phi^{(r)} \leq \lfloor n_\Phi/(n_r e_r) \rfloor$.

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

    else                          #####  CONTINUE  ############################

        #####  3. Create psh_r, etc.  ########################################

        psh[r] := rhoz(r-1,wfs[2][1][1]):                          #####  Step 7

        f[r] := degree(psh[r],y):

        k := mul(f[j],j=0..r):
        h := y^k:
        while not Irreduc(h) mod p do                              #####  Step 8
            h := y^k + randpoly(y, degree=k-1, coeffs=rand(p)):
        od:
        pss[r] := sort(h,y):                                  ########  in F_p[y]

        rho[r] := RootOf(pss[r]) mod p:
        psl[r] := evalb(degree(pss[r],y) = 1):

        for s from r by -1 to 0 do
            if s = r then
                rhh[r,s] := rho[r]                                 #####  Step 9
            elif s = r-1 then
                rhh[r,s] := randrt(pss[s],r)                       ####  Step 10
            else
                rhh[r,s] := simplify(rhhrs(r,s+1,rhh[s+1,s]))      ####  Step 11
            fi:
        od:
        xi[r] := randrt(subs(z=rhh[r,r-1],psh[r]),r):              ####  Step 12

        Upsinv[r] := Inverse(Upsilon(r)) mod p:                    ####  Step 13

        #####  4. Next Level  ###############################################

        FL2 := montL1(r+1):                                        ####  Step 14

    fi:

fi:                          ###################################################

return(FL2)

end:

################################################################################
```

The execution times of Steps 7 through 14 are as follows.

7. time: $O(f_r f_{r-1}^{*(1+\epsilon)}) \subseteq O(f_r^{*(1+\epsilon)})$

8. expected time: $O(f_r^{*2}\,\mathsf{L}(f_r^*)\log f_r^* + f_r^*\,\mathsf{L}(f_r^*\log q_r)) \subseteq O(f_r^{*(2+\epsilon)})$.

9. time: $O(1)$

10. time: $O(f_{r-1}^{*(2+\epsilon)} f_{r-1}^*) = O(f_{r-1}^{*(3+\epsilon)})$

11. $\deg \widehat{\rho}_{s+1,s} = f_s$, $0 \le s \le r-2$;
    total time: $O\big(\sum_{s=0}^{r-2} f_s f_{r-1}^{*(1+\epsilon)}\big) \subseteq O\big(r f_{r-1}^{*(2+\epsilon)}\big)$

12. time: $O(f_r f_{r-1}^{*(1+\epsilon)} + f_r^{*(2+\epsilon)} f_{r-1}^*) = O(f_r^{*(2+\epsilon)} f_{r-1}^*) \subseteq O(f_r^{*(3+\epsilon)})$

13. time: $O(f_{r-1}^{*(3+\epsilon)} + f_{r-1}^{*\alpha}) \subseteq O(f_{r-1}^{*(3+\epsilon)})$, with $2 < \alpha \le 3$

14. time: $\big\langle \mathtt{montL1}(r+1) \big\rangle$

The total time (excluding step 14) is given by

$$\begin{aligned}
\mu_2(r) \in O\big(& n_\Phi^2 + (n_\Phi/n_r) n_r^2 \ln n_r + (n_\Phi/n_r)^2 \\
&\qquad + n_\Phi n_r^{1+\epsilon} + d_\Psi^{2+\epsilon} f_{r-2}^{*(1+\epsilon)} + r f_{r-1}^{*(2+\epsilon)} + f_r^{*(3+\epsilon)}\big) \\
\subseteq O\big(& n_\Phi^2 + n_\Phi n_r \ln n_r + n_\Phi n_r^{1+\epsilon} + d_\Psi^{2+\epsilon} f_{r-2}^{*(1+\epsilon)} + r f_{r-1}^{*(2+\epsilon)} + f_r^{*(3+\epsilon)}\big) \\
\subseteq O\big(& n_\Phi^2 + n_\Phi n_r^{1+\epsilon} + d_\Psi^{2+\epsilon} f_{r-2}^{*(1+\epsilon)} + r f_{r-1}^{*(2+\epsilon)} + f_r^{*(3+\epsilon)}\big) \\
\subseteq O\big(& n_\Phi^{2+\epsilon} + f_r^{*(3+\epsilon)}\big) \\
\subseteq O\big(& n_\Phi^{3+\epsilon}\big).
\end{aligned}$$

Proposition 14. *We have*

$$\big\langle \mathtt{montL2}(r) \big\rangle \le \mu_2(r) + \big\langle \mathtt{montL1}(r+1) \big\rangle$$

with $\mu_2(r) \in O(n_\Phi^{3+\epsilon})$.

Remark 15. If $e_r f_r \neq 1$ then

$$\big\langle \mathtt{montL1}(r+1) \big\rangle \le \mu_1(r+1) + \big\langle \mathtt{montL2}(r+1) \big\rangle$$

with $\mu_1(r+1) \in O\big((r+1)\, n_{r+1}^{3+\epsilon}\big)$.

Conclusion.

It will be convenient to let

$$\begin{array}{lll} L_0(F) & \text{denote} & \texttt{montL0}(F), \\ L_1(r) & \text{denote} & \texttt{montL1}(r) \quad \text{for } r \geq 1, \quad \text{and} \\ L_2(r) & \text{denote} & \texttt{montL2}(r) \quad \text{for } r \geq 1. \end{array}$$

We will estimate the time required for the chain of computations

$$L_0(F) \to L_1(1) \to L_2(1) \to L_1(2) \to L_2(2) \to \cdots \to L_1(m) \to L_2(m)$$

where $m \leq \lfloor \log_2 \deg F \rfloor$.

Substituing the time required for each term in the chain and assuming that the algorithm terminates at level m we get the following.

$$\begin{aligned} \big\langle L_0(F) \big\rangle &\leq \mu_0(n) + \big\langle L_1(1) \big\rangle \\ &\leq \mu_0(n) + \mu_1(1) + \big\langle L_2(1) \big\rangle \\ &\leq \mu_0(n) + \mu_1(1) + \mu_2(1) + \big\langle L_1(2) \big\rangle \\ &\leq \mu_0(n) + \mu_1(1) + \mu_2(1) + \mu_1(2) + \big\langle L_2(2) \big\rangle \\ &\leq \mu_0(n) + \mu_1(1) + \mu_2(1) + \mu_1(2) + \mu_2(2) + \big\langle L_1(3) \big\rangle \\ &\leq \mu_0(n) + \mu_1(1) + \mu_2(1) + \mu_1(2) + \mu_2(2) + \mu_1(3) + \big\langle L_2(3) \big\rangle \\ &\vdots \\ &\leq \mu_0(n) + \textstyle\sum_{r=1}^{m} \mu_1(r) + \sum_{r=1}^{m-1} \mu_2(r) + \big\langle L_2(m) \big\rangle \\ &\in O\big(n_\Phi^{2+\epsilon} + m^2 n_\Phi^{3+\epsilon} + (m-1) n_\Phi^{3+\epsilon} + n_\Phi^{2+\epsilon}\big) \\ &\subseteq O\big(n_\Phi^{3+\epsilon}\big). \end{aligned}$$

Note that, since the algorithm terminates at level m, we have

$$\big\langle L_2(m) \big\rangle \in O(n_\Phi^{2+\epsilon}).$$

We will give a bound $B_2(r)$ for the time taken by the sequence

$$L_1(r) \to L_2(r-1) \to L_1(r)$$

which happens when $e_{r-1}f_{r-1} = 1$ for some $r > 1$. We have

$$\big\langle L_1(1)\big\rangle \leq \mu_1(r) + \big\langle L_2(r-1)\big\rangle$$

when $e_{r-1}f_{r-1} = 1$, and on the other hand

$$\big\langle L_2(r-1)\big\rangle \leq \mu_2(r-1) + \big\langle L_1(r)\big\rangle.$$

Thus we can take

$$B_2(r) = \mu_1(r) + \mu_2(r-1) \in O(rn_r^{2+\epsilon} + n_\Phi^{3+\epsilon}) \subseteq O(n_\Phi^{3+\epsilon}).$$

From Proposition 10 it follows that the sequence

$$L_1(r) \rightarrow L_2(r-1) \rightarrow L_1(r)$$

can occur at most $2v_p(\operatorname{disc}\Phi)$ times in the course of the computation.

It now follows that the Montes algorithm terminates in time

$$O\big(n_\Phi^{3+\epsilon} + 2n_\Phi^{3+\epsilon}v_p(\operatorname{disc}\Phi)\big) \subseteq O\big(n_\Phi^{3+\epsilon}v_p(\operatorname{disc}\Phi)\big).$$

Since this estimate counts operations in $\mathbf{Z}_p$ and $\mathbf{F}_p$ together, it follows that the bit-complexity of the Montes algorithm is

$$O\big(n_\Phi^{3+\epsilon}v_p(\operatorname{disc}\Phi)^{2+\epsilon}\big).$$

Chapter 4

Comparisons

4.1 The One-Element Algorithm

The *One-Element Method* is in essence the original Round Four algorithm of Zassenhaus (Ford, 1987), with various improvements in detail. In what follows we will refer to the version given in (Ford, Pauli, and Roblot, 2002).

Let $f(x)$ be a monic polynomial in $\mathbf{Z}_p[x]$ with nonzero discriminant, let K be the extension of $\mathbf{Q}_p$ generated by a root α of f, and let $\mathcal{O}_K$ be the ring of integers of K. We will let v denote the extension of the standard p-adic valuation of $\mathbf{Q}_p$ to K.

A prime element π of $\mathcal{O}_K$ has minimal positive valuation:

$$0 < v(\pi) = 1/e \leq v(p) = 1$$

and there is no $\theta \in \mathcal{O}_K$ such that $0 < v(\theta) < v(\pi)$.

The one-element method exploits the fact that if the polynomial g is irreducible over $\mathbf{Q}_p$ then all roots of g are algebraic conjugates over $\mathbf{Q}_p$. Otherwise stated, if f is reducible over $\mathbf{Q}_p$ then there must exists two roots of f that are not conjugate.

Let R_π be a complete set of representatives of $\pi\mathcal{O}_K$ in $\mathcal{O}_K$ and let

$$\omega = \pi^e/p = \lambda_0(\alpha) + \lambda_1(\alpha)\pi + \lambda_2(\alpha)\pi^2 + \cdots$$

with $\lambda_0(\alpha), \lambda_1(\alpha), \lambda_2(\alpha), \ldots \in R_\pi$.

The (not necessarily distinct) roots $\omega^{\sigma_1}, \ldots, \omega^{\sigma_r}$ of μ_ω are given by

$$(*) \qquad \omega^{\sigma_j} = \lambda_0(\alpha^{\sigma_j}) + \lambda_1(\alpha^{\sigma_j})\pi^{\sigma_j} + \lambda_2(\alpha^{\sigma_j})(\pi^{\sigma_j})^2 + \cdots$$

where $\{\sigma_1, \ldots, \sigma_r\} = \mathrm{Gal}(K/\mathbf{Q}_p)$. Note that although $\lambda_k(\alpha^\sigma)$ changes with σ, the *polynomial* $\lambda_k(x)$ does not. This fact is employed in the construction of the minimal polynomial $\mu_\omega(x)$ of ω.

As the algorithm progresses an element

$$\omega_{k-1} = \lambda_0(\alpha) + \lambda_1(\alpha)\pi + \cdots + \lambda_{k-1}(\alpha)\pi^{k-1}$$

will have been constructed and it will be necessary to find $\lambda_k(x)$ such that

$$v\Big(\frac{\omega - \omega_{k-1}}{\pi^k} - \lambda_k(\alpha)\Big) > 0.$$

If it happens that there is more than one choice for $\lambda_k(\alpha) \in R_\pi$ then there is more than one choice for $\mu_\omega(x)$ and this leads to a factorization of $f(x)$.

On the other hand, if the expansion $(*)$ can be extended sufficiently far it would follow that $\deg \mu_\omega = \deg f$ and thus f would necessarily be irreducible (Ford, Pauli, and Roblot, 2002, Proposition 4.5).

We will now give a brief sketch the algorithm.

Assuming $f(x) = (x - \xi_1) \cdots (x - \xi_n)$ and $\theta(x) \in \mathbf{Q}_p(x)$, we define

$$\chi_\theta(t) = \big(t - \theta(\xi_1)\big) \cdots \big(t - \theta(\xi_1)\big) = \mathrm{Res}_x\big(f(x, t - \theta(x)\big).$$

The algorithm constructs a polynomial $\alpha(x) \in \mathbf{Q}(x)$ and seeks to determine if the elements $\alpha(\xi_1), \ldots, \alpha(\xi_n)$ are conjugate over $\mathbf{Q}_p$. If so then f is irreducible over $\mathbf{Q}_p$; if not then a proper factorization of f is constructed.

Reducibility is established either by $\overline{\chi}_\alpha(x) \in \mathbf{F}_p[x]$ having more than one distinct irreducible factor or by $\mathcal{N}_0(\chi_\alpha)$ having more than one edge.

We define $-D_\alpha/E_\alpha$ to be the slope of the (unique) edge of $\mathcal{N}_0(\chi_\alpha)$, with $E_\alpha > 0$ and $\gcd(D_\alpha/E_\alpha) = 1$.

The irreducible polynomial $\overline{\nu}_\alpha(x) \in \mathbf{F}_p[x]$ is given by $\overline{\chi}_\alpha(x) = \overline{\nu}_\alpha(x)^e$ for some $e > 0$, and we define $F_\alpha = \deg \overline{\nu}_\alpha$.

Irreducibility of f is established if $E_\alpha F_\alpha = n$.

Initially $\alpha(x) \leftarrow x$. The algorithm iteratively constructs $\alpha'(x)$ with either $E_{\alpha'} > E_\alpha$ and $F_{\alpha'} = F_\alpha$ or $E_{\alpha'} \geq E_\alpha$ and $F_{\alpha'} > F_\alpha$, replacing $\alpha(x) \leftarrow \alpha'(x)$, until a terminating condition is achieved.

It is a consequence of Proposition 10 that the algorithm will terminate before $v(\mu_\omega(\alpha))$ exceeds $2v(\operatorname{disc} \chi_\omega)/n$.

4.2 The Two-Element Algorithm

In (Pauli, 2001) Pauli presented the *Two Element Method* for factorization of polynomials over local fields, together with a complexity analysis.

In the following we recall this algorithm.

Let $f(x)$ be a squarefree monic polynomial in $\mathbf{Z}_p[x]$ with

$$f(x) = (x - \xi_1) \cdots (x - \xi_n).$$

Initially: $E \leftarrow 1, \quad \varphi \leftarrow x$

The algorithm constructs the following:

- a sequence $\varphi_1, \varphi_2, \ldots, \varphi_h \in \mathbf{Z}_p[x]$, with $E \leftarrow \operatorname{lcm}(E_{\varphi_1}, E_{\varphi_2}, \ldots, E_{\varphi_h})$;
- a sequence $\eta_1, \eta_2, \ldots, \eta_k \in \mathcal{O}$ such that $\mathbf{Q}_p(\eta_1, \eta_2, \ldots, \eta_k)$ is an unramified extension of $\mathbf{Q}_p$, with $F \leftarrow \left[\mathbf{Q}_p(\eta_1, \eta_2, \ldots, \eta_k) : \mathbf{Q}_p\right]$.

Termination:

- If $EF = n$ then f is irreducible over $\mathbf{Q}_p$.

  In this case the algorithm returns elements $\gamma, \pi \in \mathcal{O}$ such that

- $\mathbf{Q}_p(\gamma)$ is an unramified extension of $\mathbf{Q}_p$ of degree F, and
- $\mathbf{Q}_p(\gamma, \pi)$ is a totally ramified extension of $\mathbf{Q}_p(\gamma)$ of degree E.

- If, for some $\alpha(x)$ with $\chi_\alpha(x) \in \mathbf{Z}_p[x]$, either of the conditions
  - $\overline{\chi}_\alpha(x) = \overline{\nu}_\alpha(x)^e$ with $e > 0$ for some irreducible $\overline{\nu}_\alpha(x)$,
  - $v\big(\alpha(\xi_1)\big) = \cdots = v\big(\alpha(\xi_n)\big)$,

  is violated, then f is reducible over $\mathbf{Q}_p$.

  In this case the algorithm returns a proper factorization of f.

We now cite two results from (Pauli, 2001).

The first, Proposition 15 below, ensures the termination of the Two Element algorithm. The termination of the One Element Method is also a consequence of this proposition, and we used it to establish the termination for the Montes algorithm as well.

Proposition 15 (Pauli). *Let*

$$f(x) = (x - \xi_1) \cdots (x - \xi_n),$$
$$\varphi(x) = (x - \alpha_1) \cdots (x - \alpha_m)$$

be two polynomials in $\mathbf{Q}_p[x]$ *such that* $f(x)$ *is squarefree and the degree of any irreducible factor of* $f(x)$ *is greater than or equal to* m. *Assume further that*

$$v(\mathrm{disc}(f)) > \frac{n}{2} \max_{1 \le i \le n} v(\varphi(\xi_i)).$$

Then $f(x)$ *is irreducible over* $\mathbf{Q}_p$.

The second result is an estimate of the complexity of the two-element method for the general case of polynomial factorization over a finite extension of $\mathbf{Q}_p$.

For the special case of factorization over $\mathbf{Q}_p$, Pauli's estimate simplifies to

$$O\big(N^{3+\epsilon} v_p(\mathrm{disc}\,\Phi)^{1+\epsilon} + N^{2+\epsilon} v_p(\mathrm{disc}\,\Phi)^{2+\epsilon}\big)$$

bit operations, where $N = \deg \Phi$.

We now consider an equivalence relation given in (MacLane, 1936).

Definition 14. Let $\alpha(x)$, $\beta(x)$ two nonzero polynomials in $\mathbf{Q}_p[x]$ and v be a valuation on $\mathbf{Q}_p(x)$. Then the equivalence relation $\underset{v}{\sim}$ on $\mathbf{Q}_p[x]$ is given by

$$\alpha \underset{v}{\sim} \beta \iff v(\alpha - \beta) > v(\alpha).$$

We use the following notation from (Ford, Pauli, and Roblot, 2002) and (Pauli, 2001).

For $\alpha(x)$ a nonzero polynomial in $\mathbf{Q}_p[x]$ we denote

$$v_p^*(\alpha) = \min_{1 \le i \le n} \{ v_p(\alpha(\xi_i)) \}.$$

Note that if $\alpha \in \mathbf{Q}_p$ then $v_p^*(\alpha) = v_p(\alpha)$.

The polynomials $\varphi_1, \varphi_2, \ldots,$ constructed in (Pauli, 2001) satisfy

$$\varphi_{t+1} = \varphi_t - \delta_t \psi_t$$

with $v_p^*(\varphi_t) = v_p^*(\psi_t)$ for $t \ge 1$.

Remark 16. We have

$$\varphi_{t+1} \underset{v_p^*}{\sim} \varphi_t.$$

By the construction of δ_t in Algorithm 5.1 in (Pauli, 2001), we have $v_p^*(\delta_t) \neq 0$, so that

$$v_p^*(\varphi_{t+1} - \varphi_t) = v_p^*(\delta_t \psi_t) > v_p^*(\psi_t) = v_p^*(\varphi_t).$$

Now we consider the construction of φ_{t+1} from the Montes algorithm. By Proposition 9 we have $v_t(\varphi_t) = \overline{\nu}_t$ and $v_t(\varphi_{t+1}) = e_t f_t \overline{\nu}_t$, since $(e_t f_t, e_t f_t \overline{\nu}_t)$ is the right endpoint of $\mathbf{N}_r(\varphi_{t+1})$. Then we have the following.

Remark 17. The equivalence

$$\varphi_{t+1} \underset{v_t}{\sim} \varphi_t$$

is false.

The "key" polynomial φ_{t+1} constructed in (MacLane, 1936) also does not satisfy the above equivalence relation.

4.3 Further Development

Conspicuously absent from recent research is any attempt at a complexity analysis of the one-element method.

Experimental results suggest that the one-element algorithm is comparatively fast, but it remains to be discovered whether this is due merely to the relative simplicity of the algorithm (reducing the overhead costs) or if it is the result of an intrinsically superior complexity.

Bibliography

Aho, A.V., Hopcroft, J.E. and Ullman, J.D., *The design and analysis of computer algorithms*, Addison-Wesley Publishing Company, 1974.

Berlekamp, E., *Factoring Polynomials over Finite Fields*, Bell Systems Technical Journal **46** (1967).

Berlekamp, E., *Factoring polynomials over large finite fields*, Math. Comp. **24** (1970).

Cannon, J., *et al.* The Magma Computational Algebra System, University of Sydney, 2000.

Cantor, D. G. and Kaltofen, E., *On fast multiplication of polynomials over arbitrary algebras*,

Cantor, D. G.,and Zassenhaus, H., *A new algorithm for factoring polynomials over finite fields*, Math. Comp. **36** (1981).

Dedekind, R., *Theory of Algebraic Integers*, 1877. (Translated by John Stillwell, Cambridge University Press, 1996.)

Dedekind, R., *Supplement X to Vorlesungen über Zahlentheorie von P.G. Lejeune Dirichlet (2nd ed.)*, Vieweg, Braunschweig, 1871; also (in part) Werke, vol 3, 223–261.

Dedekind, R., *Sur la théorie des nombres entiers algébriques*, Gauthier-Villars, 1877; also Bull. des Sci. math. Astron., (1), 11 (1876) 278–288; (2), 1 (1877) 17–41, 69–92, 144–164, 207–248 and (in part) Werke, vol 3, 263–296.

Dedekind, R., *Über den Zusammenhang zwischen der Theorie der Ideale und der Theorie der höheren Kongruenzen*, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen 23 (1878), 1–23.

Edwards, H. M., *Fermat's Last Theorem. A genetic introduction to number theory, ed.*, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

Edwards, H. M., *The genesis of Ideal Theory*, Arch. History Ex. Sci., 23, 1980, 321–378.

Ford, D., *On the Computation of the maximal order in a Dedekind domain*, PhD Dissertation, Ohio State University, 1978.

Ford, D., *The construction of maximal orders over a Dedekind domain*, J. Symb. Comp. **4** (1987) 69–75.

Ford D. and Letard P., *Implementing the Round Four maximal order algorithm*, J. Théor. Nombres de Bordeaux **6** (1994) 39–80,

Ford, D., Pauli, S. and Roblot, X.-F., *A Fast Algorithm for Polynomial Factorization over* $\mathbf{Q}_p$, Journal de Théorie des Nombres de Bordeaux **14** (2002) 151–169.

Gathen, J. and Gerhard J., J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, 1999.

Hensel, K., *Über eine neue Begründlung der Theorie der algebraischen Zahlen*, Jahresbericht der Deutschen Mathematiker-Vereinigung 6 (1897), 83–88.

Hensel, K., *Theorie der algebraischen Zahlen*, Teubner, Leipzig, Berlin, 1908.

Hensel, K., *Eine neue Theorie der algebraischen Zahlen*, Mathematische Zeitschrift 2 (1918), 433–452. [H7]

Kummer, E. E., *De numeris complexis, qui radicibus unitatis et numeris integris realibus constant*, Gratulationschrift der Univ. Breslau zur Jubelfeier der Univ.

Königsberg; Reprint, Jour. de Math. 12 (1847) 185–212 and Collected Papers, vol 1, 165–192.

Kummer, E. E., *Zur Theorie der Complexen Zahlen*, Monatsber. Akad. Wiss. Berlin, 1846, 87–96; also Jour. für Math. (Crelle) 35 (1847) 319–326 and Collected Papers, vol 1, 203–210.

Kummer, E. E., *Mémoire sur la théorie des nombres composés de racines de l'unité et de nombres entier*, Jour. de Math. 16 (1851) 377–498 and Collected Papers, vol 1, 363–484.

Lenstra, A. K., Lenstra, H. W., Jr., and Lovász, L. *Factoring polynomials with rational coefficients*, Mathematische Annalen 261 (4), (1982), 515–534.

MacLane, S., *A Construction for absolute values in polynomial rings*, Trans. Amer. Math. Soc. 40 (1936), 363–395.

Mignotte, M., *An inequality about factors of polynomials*, Math Comp, vol 28, (1974), 1153–1157.

Montes, J., Nart, E., *On a theorem of Ore*, Journal of Algebra 146, (1992), 318–334

Montes, J., *Polígonos de Newton de orden superior y aplicaciones aritméticas*, PhD thesis, Universitat de Barcelona, 1999.

Ore, Ö. *Newtonsche Polygone in der Theorie der algebraischen Körper*, Math. Ann **99** (1928).

Pauli, S., *Factoring Polynomials over Local Fields*, Journal of Symbolic Computation 32 (5), (2001), 533–547.

Pohst, M. E., and Zassenhaus, H. *Algorithmic algebraic number theory*, Cambridge University Press. 1989.

Shoup, V., *Fast Construction of Irreducible Polynomials over Finite Fields*, Journal of Symbolic Computations. 17. (1994), 371–394.

Schönhage, A. and Strassen, V., *Schnelle Multiplikation großer Zahlen*, Computing **7** (1971), 281–292.

Specht, W., *Abschätzungen der Wurzeln albebraischer Gleichungen*, Math Zeit, vol 52, (1949), 310–321.

Strassen, V., *Gaussian Elimination is not Optimal*, Numerische Mathematik, vol 13, (1969), 354–356.

Zassenhaus, H., *On Hensel factorization,I*, Journal of Number Theory, vol. 1., (1969), 291–311.

Zassenhaus, H., *On Hensel factorization, II*, Symposia Mathematica XV, Instituto Di Alta Matematica, Academic Press, New York-London, (1975), 499–513.

Appendix A

Finite Field Computations

A.1 Implementing Finite Fields

We decided not to use the MAPLE GF package, choosing instead to represent a finite field directly as

$$\mathbf{F}_p[x] \,/\, \psi^*(x)\mathbf{F}_p[x]$$

with $\psi^*(x)$ an irreducible monic polynomial in $\mathbf{F}_p[x]$.

By definition $q_0 = p$.

For each $r \geq 0$ the algorithm finds a monic polynomial $\psi_r(x)$ in $\mathbf{F}_{q_r}[x]$, irreducible over $\mathbf{F}_{q_r}$, having degree f_r and a root ξ_r.

The field $\mathbf{F}_{q_{r+1}}$ is given by

$$\mathbf{F}_{q_{r+1}} = \mathbf{F}_{q_r}[\xi_r] = \mathbf{F}_p[\xi_0, \ldots, \xi_r] = \mathbf{F}_p[\rho_r]$$

with ρ_r an arbitrary root of an arbitrary irreducible monic polynomial $\psi_r^*(x) \in \mathbf{F}_p[x]$, with $\deg \psi_r^* = f_r^* = f_0 \cdots f_r$.

In general an element of $\mathbf{F}_{q_r}$ is represented as a polynomial in $\mathbf{F}_p[\rho_{r-1}]$, and this (rather inconveniently) necessitates expressing ξ_s and ρ_s as polynomials in ρ_r, for $0 \leq s \leq r$.

To summarize, for $r \geq 0$ we have the following:

$$\mathbf{F}_{q_r} = \begin{cases} \mathbf{F}_p & \text{if } r = 0, \\ \mathbf{F}_{q_{r-1}}[\xi_{r-1}] = \mathbf{F}_p[\xi_0, \ldots, \xi_{r-1}] = \mathbf{F}_p[\rho_{r-1}] & \text{if } r > 0; \end{cases}$$

$$[\,\mathbf{F}_{q_{r+1}} : \mathbf{F}_{q_r}\,] = f_r \text{ and } q_{r+1} = q_r^{f_r} = p^{f_r^*}.$$

For $0 \leq s \leq r$ we adopt the following notation.

$$\begin{array}{lll} \psi_r(Y) \in \mathbf{F}_{q_r}[Y] & 0 = \psi_r(\xi_r) & \deg \psi_r = f_r \\ \psi_r^*(Y) \in \mathbf{F}_p[Y] & 0 = \psi_r^*(\rho_r) & \deg \psi_r^* = f_r^* = f_0 \cdots f_r \\ \widehat{\psi}_r(Z, Y) \in \mathbf{F}_p[Z][Y] & \psi_r(Y) = \widehat{\psi}_r(\rho_{r-1}, Y) & \\ \widehat{\rho}_{r,s} \in \mathbf{F}_p[\rho_r] & 0 = \psi_s^*(\widehat{\rho}_{r,s}) & \\ \xi_r \in \mathbf{F}_p[\rho_r] & 0 = \psi_r(\xi_r) = \widehat{\psi}_r(\rho_{r-1}, \xi_r) & \end{array}$$

We construct $\widehat{\psi}_r$, ψ_r^*, ρ_r, $\widehat{\rho}_{r,r-1}$, ξ_r as follows.

1. $\lambda_r\, \psi_r(Y)^{a_r} \leftarrow \texttt{Factors}(\widetilde{\Psi}_F^{(r)}(Y), \rho_{r-1})$, with $\lambda_r \in \mathbb{F}_{q_r}$, $a_r > 1$.
2. $f_r = \deg \psi_r$ and $\widehat{\psi}_r(Z, Y)$ such that $\widehat{\psi}_r(\rho_{r-1}, Y) = \psi_r(Y)$ are given.
3. Choose $\psi_r^*(Y)$ random in $\mathbf{F}_p[Y]$, monic of degree f_r^*.
4. Set $\rho_r = \texttt{RootOf}(\psi_r^*)$. *(formal)*
5. $\{\, Y - \mu(\rho_r)^{p^k} \mid k = 0, \ldots, f_{r-1}^* - 1 \,\} \leftarrow \texttt{Factors}(\psi_{r-1}^*(Y), \rho_r)$.
6. Choose $\widehat{\rho}_{r,r-1} \leftarrow \mu(\rho_r)$. *(arbitrary)* $[\,\psi_{r-1}^*(\widehat{\rho}_{r,r-1}) = 0\,]$
7. For $s = r-1, r-2, \ldots, 1$: $\widehat{\rho}_{r,s-1} \leftarrow \texttt{subs}(\rho_s = \widehat{\rho}_{r,s}, \widehat{\rho}_{s,s-1})$.
8. $\{\, Y - \widehat{\eta}_k(\rho_r) \mid k = 0, \ldots, f_r - 1 \,\} \leftarrow \texttt{Factors}(\widehat{\psi}_r(\widehat{\rho}_{r,r-1}, Y), \rho_r)$.
9. Choose $\xi_r \leftarrow \widehat{\eta}_k(\rho_r)$. *(arbitrary)* $[\,\widehat{\psi}_r(\rho_{r-1}, \xi_r) = 0\,]$

A.2 Computing $\delta_i(Y)$

The polynomial $H_{t,\nu,\delta}(Y)$ in Algorithm 1 is computed as $\texttt{Hr}(t,\nu,\delta)$. Here we give the (somewhat complicated) details of the construction of $\delta_i(Y)$ for a given $i \in J_\delta$.

Computing Υ_r.

If $r > 0$ we construct $\Upsilon_r \in \mathbf{F}_p^{f_r^* \times f_r \times f_{r-1}^*}$ such that

$$\rho_{r-1}^k \, \xi_r^j \;=\; \textstyle\sum_{h=0}^{f_r^*-1} (\Upsilon_r)_{h,j,k} \, \rho_r^h$$

for $j = 0, \ldots, f_r - 1$, $k = 0, \ldots, f_{r-1}^* - 1$.

In practice we construct $\widetilde{\Upsilon}_r \in \mathbf{F}_p^{f_r^* \times f_r^*}$ and $\widetilde{M} \in \mathbf{F}_p^{f_r^*}$ such that

$$(\widetilde{\Upsilon}_r)_{1+h,1+j+kf_r} = (\Upsilon_r)_{h,j,k}, \quad \widetilde{M}_{1+j+kf_r} = M_{j,k},$$

for $h = 0, \ldots, f_r^* - 1$, $j = 0, \ldots, f_r - 1$, $k = 0, \ldots, f_{r-1}^* - 1$.

Deriving δ_i from Υ_{t-1}.

Given $i \in J$ and $t \geq 2$, let

$$\Gamma_{\mathcal{T}_{t,\nu},t,i} \, \zeta_i \;=\; \kappa_{i,0} + \kappa_{i,1} \, \rho_{t-1} + \cdots + \kappa_{i,f_{t-1}^*-1} \, \rho_{t-1}^{f_{t-1}^*-1} \;\in\; \mathbb{F}_p[\rho_{t-1}] \;=\; \mathbb{F}_{q_t}.$$

For $j = 0, \ldots, f_{t-1} - 1$, $k = 0, \ldots, f_{t-2}^* - 1$, let $M_{j,k} \in \mathbf{F}_p$ satisfy

$$\textstyle\sum_{j=0}^{f_{t-1}-1} \sum_{k=0}^{f_{t-2}^*-1} (\Upsilon_{t-1})_{h,j,k} \, M_{j,k} \;=\; \kappa_{i,h}$$

for $h = 0, \ldots, f_{t-1}^* - 1$, and let

$$\delta_i(Y) \;=\; \textstyle\sum_{j=0}^{f_{t-1}-1} \big(\sum_{k=0}^{f_{t-2}^*-1} M_{j,k} \, \rho_{t-2}^k \big) \, Y^j.$$

Then $\delta_i(Y) \in \mathbf{F}_p[\rho_{t-2}][Y] = \mathbf{F}_{q_{t-1}}[Y]$ and

$$\begin{aligned}
\delta_i(\xi_{t-1}) &= \sum_{j=0}^{f_{t-1}-1} \sum_{k=0}^{f^*_{t-2}-1} M_{j,k}\, \rho_{t-2}^k\, \xi_{t-1}^j \\
&= \sum_{j=0}^{f_{t-1}-1} \sum_{k=0}^{f^*_{t-2}-1} M_{j,k} \sum_{h=0}^{f^*_{t-1}-1} (\Upsilon_{t-1})_{h,j,k}\, \rho_{t-1}^h \\
&= \sum_{h=0}^{f^*_{t-1}-1} \sum_{j=0}^{f_{t-1}-1} \sum_{k=0}^{f^*_{t-2}-1} (\Upsilon_{t-1})_{h,j,k}\, M_{j,k}\, \rho_{t-1}^h \\
&= \sum_{h=0}^{f^*_{t-1}-1} \kappa_{i,h}\, \rho_{t-1}^h \\
&= \Gamma_{\mathcal{T}_{t,\nu},t,i}\, \zeta_i.
\end{aligned}$$

Appendix B

An Extended Example

Let $p = 2$ and let

$$F(x) = x^{16} - 12x^{14} - 84x^{13} - 196x^{12} + 2856x^{11} + 6328x^{10} - 42336x^9 - 64820x^8 - 171824x^7 - 225360x^6 - 203232x^5 + 261872x^4 + 215776x^3 + 221280x^2 + 127328x + 2256.$$

The *reduced resultant* of F is $512 = p^9$ and hence it would suffice if all computations in $\mathbf{Z}_p$ were performed modulo $p^{19} = 524288$. (We have omitted this reduction in this example.)

$\mathbf{L_0}$. Since $F(0) \equiv 0 \pmod{p}$ we set

$$\Phi(x) \leftarrow F(x+1)$$

and now

$$\Phi(x) = x^{16} + 16x^{15} + 108x^{15} + 308x^{13} - 560x^{12} - 6048x^{11} - 3220x^{10} + 62260x^9 + 81862x^8 - 841760x^7 - 4504236x^6 - 11496820x^5 - 17916176x^4 - 17316592x^3 - 9498860x^2 - 2114868x + 129833$$

with

$$\Phi(x) \equiv (x+1)^{16} \pmod{p}.$$

At level 0 we have

$$\begin{array}{lll} d_0 = 0 & \overline{\mu}_0 = 0 & m_0 = 0 \\ e_0 = 1 & \overline{\nu}_0 = 0 & \Omega_0 = 1 \\ \varphi_0(x) = x & n_0 = 1 & \\ \psi_0(y) = y + 1 & f_0 = 1 & \psi_0^*(y) = y + 1 \\ \xi_0 = \rho_0 & \widehat{\rho}_{0,0} = \rho_0 & \end{array}$$

with $\rho_0 = 1$.

L_1. Ascending to level 1 we have

$$\begin{array}{ll} \overline{\mu}_1 = d_0 + e_0\overline{\nu}_0 = 0 & n_1 = n_0 e_0 f_0 = 1 \\ \overline{\nu}_1 = e_0 f_0 \overline{\mu}_1 \quad = 0 & \Omega_1 = \Omega_0^{e_0 f_0} \xi_0^{m_0 f_0 \overline{\mu}_1} = 1 \end{array}$$

and we set

$$\varphi_1(x) \leftarrow x + 1 \equiv \psi_0(x) \pmod{p}.$$

L_2. Below are the φ_1-adic coefficients and valuation points of Φ.

$$\begin{array}{ll} A_0 = 2256 & P_0 = (0, 4) \\ A_1 = 127328 & P_1 = (1, 5) \\ A_2 = 221280 & P_2 = (2, 5) \\ A_3 = 215776 & P_3 = (3, 5) \\ A_4 = 261872 & P_4 = (4, 4) \\ A_5 = -203232 & P_5 = (5, 5) \\ A_6 = -225360 & P_6 = (6, 4) \\ A_7 = -171824 & P_7 = (7, 4) \\ A_8 = -64820 & P_8 = (8, 2) \\ A_9 = -42336 & P_9 = (9, 5) \\ A_{10} = 6328 & P_{10} = (10, 3) \\ A_{11} = 2856 & P_{11} = (11, 3) \\ A_{12} = -196 & P_{12} = (12, 2) \\ A_{13} = -84 & P_{13} = (13, 2) \\ A_{14} = -12 & P_{14} = (14, 2) \\ A_{15} = 0 & P_{15} = (15, \infty) \\ A_{16} = 1 & P_{16} = (16.0) \end{array}$$

It follows that $\mathcal{N}_1(\Phi)$ is the single segment $\mathcal{S}_1$ with endpoints

$$\{\,(0,4),\,(16,0)\,\}$$

so that

$$g_1 = 4 > 1, \qquad d_1 = 1, \qquad e_1 = 4, \qquad m_1 = 1.$$

In computing $\Psi^{(1)}_{\mathcal{S}_1,\Phi}(y) = \mathtt{AP}(1,\mathcal{S}_1,A,P)$ we find $J = \{\,0,\,2,\,4\,\}$ so that

$$\Psi^{(1)}_{\mathcal{S}_1,\Phi}(y) = \Big(\frac{\overline{A_{16}}}{p^0}\Big)y^4 + \Big(\frac{\overline{A_8}}{p^2}\Big)y^2 + \Big(\frac{\overline{A_0}}{p^4}\Big) = y^4 + y^2 + 1 = (y^2+y+1)^2.$$

We now have

$$\psi_1(y) = y^2 + y + 1 \qquad f_1 = 2 \qquad \psi_1^*(y) = y^2 + y + 1$$
$$\xi_1 = \rho_1 \qquad\qquad\qquad \widehat{\rho}_{1,1} = \rho_1 \qquad \widehat{\rho}_{1,0} = 1$$

with ρ_1 an arbitrary root of ψ_1^*.

L$_1$. At level 2 we have

$$\overline{\mu}_2 = d_1 + e_1\overline{\nu}_1 = 1 \qquad\qquad n_2 = n_1 e_1 f_1 = 8$$
$$\overline{\nu}_2 = e_1 f_1 \overline{\mu}_2 \quad = 8 \qquad\qquad \Omega_2 = \Omega_1^{e_1 f_1}\xi_1^{m_1 f_1 \overline{\mu}_2} = \rho_1 + 1.$$

We set

$$\gamma_2(y) \leftarrow \Omega_1^{-e_1 f_1}(\psi_1(y) - y^{f_1}) = y + 1$$

and call $\mathtt{Hr}(t,\nu,\delta)$ with

$$t = r - 1 = 1, \qquad \nu = \overline{\nu}_2 = 8, \qquad \delta(y) = \gamma_2(y) = y + 1.$$

Then

$$J_\delta = \{\,0,1\,\} \qquad \zeta_0 = 1 \qquad \mathtt{ae}_0 = 0 \qquad \mathtt{bd}_0 = 2 \qquad K_0 = 4$$
$$\zeta_1 = 1 \qquad \mathtt{ae}_1 = 4 \qquad \mathtt{bd}_1 = 1 \qquad K_1 = 2$$

giving

$$H_{1,\overline{\nu}_2,\gamma_2}(x) = 4\,\varphi_1(x)^0 + 2\,\varphi_1(x)^4 = 2x^4 + 8x^3 + 12x^2 + 8x + 6$$

and we set

$$\varphi_2(x) \leftarrow \varphi_1(x)^{e_1 f_1} + H_{1,\overline{\nu}_2,\gamma_2}(x)$$
$$= x^8 + 8x^7 + 28x^6 + 56x^5 + 72x^4 + 64x^3 + 40x^2 + 16x + 7.$$

L$_2$. We have $e_1 f_1 = 8 > 1$ and

$$\varphi_2(x) \equiv (x+1)^8 \pmod{p}.$$

Below are the φ_2-adic coefficients and valuation points of Φ.

$$\begin{aligned} A_0 &= -177536x^7 - 1480768x^6 - 5274912x^5 - 9985280x^4 && P_0 = (0,22) \\ &\quad - 10389248x^3 - 5326048x^2 - 396768x + 813600 && P_1 = (1,21) \\ A_1 &= -12x^6 - 156x^5 - 800x^4 + 976x^3 && P_2 = (2,16) \\ &\quad + 12700x^2 - 22188x - 97688 \\ A_2 &= 1 \end{aligned}$$

It follows that $\mathcal{N}_2(\Phi)$ is the single segment $\mathcal{S}_2$ with endpoints

$$\{\,(0,22),\,(2,16)\,\}$$

so that

$$g_2 = 2 > 1, \qquad d_2 = 3, \qquad e_2 = 1, \qquad m_2 = 0.$$

In computing $\Psi^{(2)}_{\mathcal{S}_2,\Phi}(y) = \mathtt{AP}(2,\mathcal{S}_2,A,P)$ we find $J = \{\,0,\,2\,\}$ so that

$$\Psi^{(2)}_{\mathcal{S}_2,\Phi}(y) = \eta_2\, y^2 + \eta_0 = (\rho_1+1)y^2 + \rho_1 = (\rho_1+1)(y+\rho_1)^2$$

We now have

$$\psi_2(y) = y + \rho_1 \qquad f_2 = 1 \qquad \psi_2^*(y) = y^2 + y + 1$$
$$\xi_2 = \rho_2 \qquad \widehat{\rho}_{2,2} = \rho_2 \qquad \widehat{\rho}_{2,1} = \rho_2 \qquad \widehat{\rho}_{2,0} = 1$$

with $\rho_2 = \rho_1$.

L$_1$. At level 3 we have

$$\overline{\mu}_3 = d_2 + e_2\overline{\nu}_2 = 11, \qquad \overline{\nu}_3 = e_2 f_2 \overline{\mu}_3 = 11, \qquad n_3 = n_2 e_2 f_2 = n_2.$$

We set

$$\gamma_3(y) \leftarrow \Omega_2^{-e_2 f_2}(\psi_2(y) - y^{f_2}) = (\rho_1+1)^{-1}\rho_1 = \rho_1 + 1$$

and call $\mathtt{Hr}(t,\nu,\delta)$ with

$$t = r - 1 = 2. \qquad \nu = \overline{\nu}_3 = 11. \qquad \delta(y) = \gamma_3(y) = \rho_1 + 1.$$

Then

$$J_\delta = \{\,0\,\}, \quad \zeta_0 = \rho_1 + 1, \quad \mathtt{ae}_0 = \alpha_{t,\nu} = 0, \quad \mathtt{bd}_0 = \beta_{t,\nu} = 11.$$

Calling $\mathtt{Hr}(t-1, \nu_0, \delta_0)$ with $\nu_0 = 11$, $\delta_0(y) = y$ yields

$$K_0(x) = H_{t-1,\nu_0,\delta_0}(x) = 2x^7 + 14x^6 + 42x^5 + 70x^4 + 70x^3 + 42x^2 + 14x + 2.$$

Hence

$$H_{t,\nu,\delta}(x) = K_0(x)\,\varphi_2(x)^{\alpha_{t,\nu}} = K_0(x)$$

and, since $e_2 f_2 = 1$, we set

$$\begin{aligned}\varphi_2(x) &\leftarrow \varphi_2(x)^{e_2 f_2} + H_{t,\nu,\delta}(x) \\ &= x^8 + 10x^7 + 42x^6 + 98x^5 + 142x^4 + 134x^3 + 82x^2 + 30x + 9.\end{aligned}$$

L$_2$. We have

$$\varphi_2(x) \equiv (x+1)^8 \pmod{p}.$$

Below are the φ_2-adic coefficients and valuation points of Φ.

$$\begin{aligned} A_0 &= -220192x^7 - 1767424x^6 - 6097568x^5 - 11463488x^4 && P_0 = (0, 26) \\ &\quad - 12316192x^3 - 7036224x^2 - 1254048x + 296576 && P_1 = (1, 21) \\ A_1 &= -4x^7 - 36x^6 - 200x^5 - 664x^4 + 1748x^3 && P_2 = (2, 16) \\ &\quad + 8060x^2 - 33920x - 18536 \\ A_2 &= 1 \end{aligned}$$

It follows that $\mathcal{N}_2(\Phi)$ is the single segment $\mathcal{S}_2$ with endpoints

$$\{\,(0, 26),\ (2, 16)\,\}$$

so that

$$g_2 = 2 > 1, \qquad d_2 = 5, \qquad e_2 = 1, \qquad m_2 = 0.$$

In computing $\Psi^{(2)}_{\mathcal{S}_2,\Phi}(y) = \mathtt{AP}(2, \mathcal{S}_2, A, P)$ we find $J = \{\,0,\ 1,\ 2\,\}$ so that

$$\Psi^{(2)}_{\mathcal{S}_2,\Phi}(y) = \eta_2\, y^2 + \eta_1\, y + \eta_0$$

with

$$\begin{aligned}
\eta_0 &= \Gamma^{-1}_{S_2,2,0}\Psi^{(1)}_{T_{1,\nu_0},A_\alpha}(\xi_1) &&= \rho_1,\\
\eta_1 &= \Gamma^{-1}_{S_2,2,1}\Psi^{(1)}_{T_{1,\nu_1},A_{\alpha+e_2}}(\xi_1) &&= 1,\\
\eta_2 &= \Gamma^{-1}_{S_2,2,2}\Psi^{(1)}_{T_{1,\nu_2},A_{\alpha+2e_2}}(\xi_1) &&= \rho_1+1,
\end{aligned}$$

and therefore

$$\Psi^{(2)}_{S_2,\Phi}(y) = (\rho_1+1)\,y^2 + y + \rho_1 = (\rho_1+1)(y+\rho_1+1)(y+1).$$

Since $\Psi^{(2)}_{S_2,\Phi}$ has two distinct irreducible factors, it follows from the Theorem of the Associated Polynomial that $\Phi(x)$, and hence $F(x)$, is reducible in $\mathbf{Q}_2[x]$.