Factoring Polynomials over p-Adic Fields

David G. Cantor and Daniel M. Gordon

Center for Communications Research
4320 Westerra Court, San Diego, CA 92121
{dgc,gordon}@ccrwest.org

Abstract. We give an efficient algorithm for factoring polynomials over finite algebraic extensions of the p-adic numbers. This algorithm uses ideas of Chistov's random polynomial-time algorithm, and is suitable for practical implementation.

1 Introduction

Factoring polynomials over the p-adic numbers $\mathbb{Q}_p$ is an important problem in computational number theory. One application is determining the prime ideals of a number field $\mathbb{Q}(\alpha)$, and how a given rational prime p factors into prime ideals in that field. See Cohen [10] and the references cited therein for some methods currently in use.

These algorithms, while generally good in practice, will take exponential time for some polynomials. A. L. Chistov ([7], [8], and [9]) has given an algorithm which runs in random polynomial time for all polynomials, but would be very difficult to implement efficiently. In this paper we give a random polynomial-time algorithm which works well in practice. The algorithm is non-deterministic only because all known efficient algorithms for factoring polynomials over finite fields $\mathbb{F}_{p^n}$ ([3], [5]) are non-deterministic. Note that any polynomial-time p-adic factoring algorithm can factor polynomials over $\mathbb{F}_{p^n}$ in polynomial time. It has been implemented in PARI, and is available on the second author's web site [13].

We will factor polynomials over a finite algebraic extension K of $\mathbb{Q}_p$. See Chapter 5 of [14] for properties of these extensions. Let π be a uniformizer of K. In the case when K is an unramified extension of $\mathbb{Q}_p$, we choose $\pi = p$. For x in the ring of integers O_K of K, $\bar{x}$ will denote the image of x in the residue class field $\overline{K}$. We will fix a set of representatives $\mathcal{A} = \{0, a_1, \ldots a_{p^f-1}\} \subset O_K$ for the elements of $\overline{K}$. This set may be lifted to representatives for unramified extensions of K in a straightforward manner.

The valuation of an element $x \in K$ will be denoted $|x|$, and its order by $\operatorname{ord} x$. We assume that $|\ |$ has been normalized so that $|p| = 1/p$. There is a unique extension of the valuation $|\ |$ on K to its algebraic closure $\widetilde{K}$; we assume that $|\ |$ has been so extended.

Just as for real numbers, one cannot, in general, explicitly represent a p-adic number exactly, but only an approximation which is a rational number. Thus our algorithm will find approximations to factors of $F(X)$. Elements x of K may

W. Bosma (Ed.): ANTS-IV, LNCS 1838, pp. 185–208, 2000.

be written $x = \sum_{i=-m}^{\infty} a_i \pi^i$, with $a_i \in \mathcal{A}$. In Section 8 we discuss where this series can be truncated to guarantee a correct answer.

Let $F(X)$ be a monic polynomial with coefficients in O_K which has no repeated factors. See Zippel [29, pp 294–295] for a simple method of removing repeated factors. Unlike Chistov's algorithm, our method does not require computing in, or even constructing, ramified extensions of K. The algorithm is applied recursively, at each step either finding a new factor or terminating with an irreducible factor and certificate of its irreducibility. The certificate of irreducibility will be a generalized Eisenstein polynomial with coefficients in the maximal unramified (over K) subfield of $K(x)$, where x is a root of the irreducible factor.

The **p-adic Factor** algorithm works by looking for a polynomial $A(X)$ for which we can determine the factorization of

$$R(Y) = \mathrm{Res}_X(F(X), Y - A(X)). \tag{1.1}$$

In Section 2 we show that a factor of $R(Y)$ lets us find a factor of $F(X)$, and a certificate of irreducibility for $R(Y)$ also applies to $F(X)$. Once such an $A(X)$ is found, we apply the information to $F(X)$ and, if necessary, recurse on remaining factors of the original polynomial.

The standard "easy" method for factoring a polynomial over the p-adics, the Newton diagram method, is given in Section 3. If the Newton diagram of the polynomial is not a straight line, then Hensel's Lemma may be used to find a factor. If the Newton diagram is a straight line with slope k/n, where n is the degree of $F(X)$ and k is relatively prime to n, then $F(X)$ is irreducible.

Otherwise the Newton diagram method fails, and we use an extension of Hensel's Lemma given in Section 4.1. We proceed by looking at the factorization of $F(X)$ in $\overline{K}$. If the reduction $F^*(X)$ (defined in Section 3) has two relatively prime factors, then using Hensel's Lemma we may lift these to factors over K. If $F^*(X)$ is the power of an irreducible polynomial of degree $d \geq 2$, then we may factor $F(X)$ over an unramified extension of degree d of K, leading to a factorization of $F(X)$ over K. These methods form the basis of the **Hensel Factor** routine given in Section 4.2. The only case **Hensel Factor** cannot handle is when

$$R(Y) = a_n \left(Y^r - b\pi^s\right)^m + \text{ [terms above the Newton diagram]}. \tag{1.2}$$

In this case we have $\operatorname{ord} A(x) = s/r$ for each root x of $F(X)$ in $\widetilde{K}$, the closure of K. The **p-adic Factor** algorithm then finds a new polynomial $A(X)$ such that either **Hensel Factor** successfully factors $R(Y)$, or (1.2) still holds with either $\operatorname{ord} A(x)$ or $\deg A(X)$ increased. Since $\deg A(X) < n$, and $\operatorname{ord} A(x)$ is bounded by Corollary 5.8, this will terminate after a bounded number of steps.

In Section 7 we illustrate how the algorithm works on two examples. Section 8 gives a worst-case bound for the bit complexity of the algorithm

$$O\left(n^{8+\epsilon} \log^3 |\, \Delta_F \,| \log^2 p^k\right), \tag{1.3}$$

where n is the degree of $F(X)$, Δ_F is the discriminant of $F(X)$, and k is the degree of K over $\mathbb{Q}_p$.

Our algorithm may be extended to any local field complete with respect to a discrete rank-1 valuation, under the assumptions that the residue class field is perfect and that an algorithm for factoring polynomials defined over the residue-class field is given. For example, applying it to the field $\mathbb{F}_q((X))$ of Laurent series, it can be used to resolve singularities of plane curves. A future paper will extend the algorithm to other local fields, and include some proofs which have been omitted here due to space constraints.

We thank Stephen DiPippo and Robert Segal for many helpful discussions. John Cannon told us of developments with MAGMA's local rings and fields package, and informed us that the MAGMA group has developed a similar algorithm for factoring polynomials over $\mathbb{Q}_p$, which is currently being implemented.

2 Some Criteria for Factorization

In this section we give simple criteria for polynomial factorization and polynomial irreducibility. Let $\mathrm{Res}_X(A(X), B(X))$ denote the resultant of two polynomials $A(X)$ and $B(X)$. See Lang [19] or Cassels [6] for details. Due to space constraints we omit proofs of the lemmas in this section. They follow in a straightforward way from the properties of the resultant.

Lemma 2.1. *Suppose that $F(X)$ and $A(X)$ are polynomials in the field $K[X]$ with $F(X)$ monic of degree n. Put*

$$R(Y) = \mathrm{Res}_X(F(X), Y - A(X)). \tag{2.2}$$

Then

1. *$R(Y)$ is a monic polynomial of degree n in Y and*
2. *the polynomial $F(X)$ divides the polynomial $R(A(X))$.*

The following lemma provides a way of factoring a polynomial.

Lemma 2.3. *Suppose that $F(X)$ and $A(X)$ are polynomials in $K[X]$, with $F(X)$ monic. Put*

$$R(Y) = \mathrm{Res}_X(F(X), Y - A(X)). \tag{2.4}$$

Suppose further that $R(Y) = R_1(Y)R_2(Y)$ is a factorization of $R(Y)$ into relatively prime, non-constant factors. Then

$$F(X) = F_1(X)F_2(X), \tag{2.5}$$

where

$$F_1(X) = \gcd(F(X), R_1(A(X)) \text{ and } F_2(X) = \gcd(F(X), R_2(A(X)), \tag{2.6}$$

is a factorization of $F(X)$ into relatively prime, non-constant factors. Furthermore,

$$\deg F_1(X) = \deg R_1(Y) \quad \text{and} \quad \deg F_2(X) = \deg R_2(Y). \tag{2.7}$$

The following Lemma provides a partial converse to Lemma 2.3.

Lemma 2.8. *Suppose that $F(X)$ is a monic polynomial of degree n, that $A(X)$ is a polynomial, and that both have coefficients in the field K. If the polynomial $R(Y) = \mathrm{Res}_X(F(X), Y - A(X))$ is irreducible over K, then $F(X)$ is also irreducible over K.*

If neither Lemma 2.3 nor Lemma 2.8 applies, we may need to go to an unramified extension field of K. The following lemma shows how irreducible factors of $F(X)$ over an extension field L of K lead to irreducible factors over K.

Lemma 2.9. *Suppose that $F(X)$ is a monic polynomial in $K[X]$ with no repeated factors of degree ≥ 1, that L is a finite algebraic extension of K, and that $G(X)$ is a monic, irreducible, polynomial in $L[X]$ of degree ≥ 1 which divides $F(X)$. Put $H(X) = \mathrm{Norm}_{L/K} G(X)$. Then,*

1. *$\gcd(F(X), H(X))$ is an irreducible factor of degree ≥ 1 of $F(X)$ in $K[X]$; and*
2. *if the field extension L/K is generated by the coefficients of $G(X)$, then $H(X)$ is already an irreducible factor of $F(X)$ in $K[X]$.*

3 Newton Diagrams

In this section we give our notation for Newton diagrams and some related items. For details see Artin [1], Cassels [6], or Gouvea [14, Section 6.4].

Suppose that

$$R(Y) = \sum_{i=0}^{n} a_i Y^i \tag{3.1}$$

is a polynomial in $K[Y]$ of (exact) degree $n \geq 1$. As usual, we associate to $R(Y)$ a finite, non-empty point set $\mathfrak{S} \subset \mathbb{R}^2$ consisting of points $(i, \operatorname{ord} a_i) \in \mathbb{R}^2$ corresponding to each nonzero term $a_i Y^i$ of $R(Y)$.

Definition 3.2. We define, as is customary, the *Newton diagram* of $R(Y)$ to be the lower boundary of the convex hull of $\mathfrak{S}$.

Following Cassels [6], we use the following definition:

Definition 3.3. Suppose that $R(Y)$ is a given by (3.1). We shall call $R(Y)$ *pure* if $a_0 \neq 0$, $n \geq 1$, and the Newton diagram of $R(Y)$ is a straight line.

If the Newton diagram is not pure, we may immediately factor $R(Y)$. The following is is well known (see Cassels [6]), and is also a corollary of our Theorem 4.21.

Lemma 3.4. *Suppose that $R(Y) = \sum_{i=0}^{k} a_i Y^i$ is a polynomial of degree $k \geq 1$ and that a_0 is not zero. If the polynomial $R(Y)$ is not pure (so that its Newton diagram consists of two or more straight line-segments necessarily of different slopes), then $R(Y)$ factors into two non-constant polynomials in $K[Y]$.*

If the Newton diagram is pure, we may sometimes use its slope to show that $R(Y)$ is irreducible.

Lemma 3.5. (Generalized Eisenstein criterion) *Suppose $R(Y)$ is pure, and its Newton diagram has slope k/n, where k is an integer relatively prime to n. Then $R(Y)$ is irreducible.*

Proof. If y is a root of $R(y)$ in $\widetilde{K}$, then $\operatorname{ord} y = k/n$. Hence $K(y)/K$ is a totally ramified extension and has degree n, so $R(Y)$ is irreducible. □

Remark 3.6. The customary form of Eisenstein's criterion is the special case when $k = -1$ (see, for example, [29]).

Now suppose that $R(Y)$ is pure and has slope $-s/r$. Because the points $(0, \operatorname{ord} a_0)$ and $(n, \operatorname{ord} a_n)$ are the end-points of the Newton diagram, n must be an integral multiple of r, say, $n = mr$. Put

$$\alpha_i = a_{ri} / \left(a_n \pi^{s(m-i)}\right) \tag{3.7}$$

so that $\alpha_i \in O_K$. We can then write

$$R(Y) = a_n \sum_{i=0}^{m} \alpha_i \pi^{s(m-i)} Y^{ri} + \text{ [terms above the Newton diagram]}. \tag{3.8}$$

Here "terms above the Newton diagram" refers to those non-zero terms of $R(Y)$ whose corresponding points in the Newton set $\mathfrak{S}$ lie strictly above the Newton diagram. These are the non-zero terms of the form $a_i Y^i$ for which $\operatorname{ord} a_i > s(m-i)/r + \operatorname{ord} a_n$.

Definition 3.9. Suppose $R(Y)$ as given by (3.1) is pure and suppose that the α_i are given by (3.7). Define

$$R^*(Y) = \sum_{i=0}^{m} \bar{\alpha}_i Y^i. \tag{3.10}$$

The polynomial $R^*(Y)$ is monic and has coefficients in $\overline{K}$. In the next section we will show how to factor $F(X)$ using Hensel's Lemma if we can write $R^*(Y)$ as the product of two relatively prime factors, perhaps over an extension field of K. Otherwise, we will use a reduction method extending the one used by Chistov [8].

4 Factoring with Hensel's Lemma

4.1 Hensel's Lemma

Hensel's Lemma refers to an algorithm, due to Hensel [17], which shows how to find a factorization of a polynomial $R(Y) \in K[Y]$ from an "approximate

factorization". Here we describe an extension of this algorithm. The extension is related to that of Artin [1]. The main novelty is Corollary 4.30. In the special case when the slope of the Newton diagram of $R(Y)$ is zero, it is well known. Dealing with general slopes avoids the need to go to ramified extension fields as in [8], making the algorithm much more practical.

Definition 4.1. Suppose that λ is a positive real number. If

$$A(Y) = \sum_{i=0}^{k} a_i Y^i \in K[Y]. \tag{4.2}$$

define its λ-*norm* $\|A(Y)\|_\lambda$ to be $\max_i |a_i|\lambda^i$. If λ is understood we shall write simply $\|A(Y)\|$ instead of $\|A(Y)\|_\lambda$.

When $A(Y)$ is the constant polynomial a_0, that is, when $n = 0$, then $\|A(Y)\|_\lambda = |a_0|$, independent of λ. Suppose $\lambda = |\pi|^{s/r}$, then, $\|aX^r\|_\lambda = |a\pi^s|$. If $A(Y) = \sum_{i=0}^{n} a_i Y^i$ is pure (see definition 3.3) with slope $-s/r$ then $\|A(Y)\|_\lambda = |a_0|$.

Lemma 4.3. *Suppose that*

1. $A(Y) = \sum_{i=0}^{k} a_i Y^i$ *is a polynomial in* $K[Y]$ *of degree* k*;*
2. $B(Y) = \sum_{i=0}^{l} b_i Y^i$ *is a non-zero polynomial in* $K[Y]$ *of degree* $l \le k$*;*
3. $\|B(Y)\| = \|b_l Y^l\|$*; equivalently,* $|b_l|\lambda^l = \max_i |b_i|\lambda^i$.

Define $C(Y) = A(Y) - Y^{k-l}(a_k/b_l)B(Y)$. *In other words,* $C(Y)$ *is the first remainder and* $(a_k/b_l)Y^{k-l}$ *is the first quotient obtained when dividing* $A(Y)$ *by* $B(Y)$ *using the classical division algorithm. Then*

1. $\|C(Y)\| \le \|A(Y)\|$, *and*
2. $\|(a_k/b_l)Y^{k-l}\| \le \|A(Y)\|/\|B(Y)\|$.

Proof. Define $b_i = 0$ when $i < 0$. Then

$$C(Y) = \sum_{i=1}^{k} \Big(a_{k-i} - \frac{a_k b_{l-i}}{b_l}\Big) Y^{k-i}. \tag{4.4}$$

Hence,

$$\begin{aligned} \|C(Y\| &= \max_{1\le i\le k} \lambda^{k-i} \Big| a_{k-i} - \frac{a_k b_{l-i}}{b_l} \Big| \\ &\le \max_{1\le i\le k} \max \Big(\lambda^{k-i}|a_{k-i}|,\ \frac{\lambda^k |a_k| \lambda^{l-i} |b_{l-i}|}{\lambda^l |b_l|} \Big) \\ &\le \max_{0\le i\le k} \max \left(\lambda^{k-i}|a_{k-i}|,\ \lambda^k |a_k| \right) \\ &= \|A(Y)\|. \end{aligned} \tag{4.5}$$

The remainder of the proof is clear. □

Lemma 4.6. *Suppose that $A(Y)$ and $B(Y)$ are polynomials satisfying hypothesis* 1, 2, *and* 3 *of Lemma 4.3. Suppose that $Q(Y)$ and $V(Y)$ are the quotient and remainder, respectively, when $A(Y)$ is divided by $B(Y)$; that is,*

$$A(Y) = B(Y)Q(Y) + V(Y), \tag{4.7}$$

where $A(Y)$, $B(Y)$, $Q(Y)$, and $V(Y)$ are polynomials in $K[Y]$ such that $\deg V(Y) < \deg B(Y)$. Then

$$\|V(Y)\| \le \|A(Y)\| \quad \text{and} \quad \|Q(Y)\| \le \|A(Y)\|/\|B(Y)\|. \tag{4.8}$$

Proof. Apply Lemma 4.3 repeatedly. □

Lemma 4.9. *Suppose that we are given a 7-tuple*

$$\big(k, \mu, B(Y), C(Y), u(Y), v(Y), \epsilon(Y)\big) \tag{4.10}$$

where k is a positive integer, where μ is real number ≥ 1, and where the remaining five entries are polynomials in $K[Y]$. Suppose that the following conditions are satisfied:

1. *$B(Y) = \sum_{i=0}^{l} b_i Y^i$ and $C(Y) = \sum_{i=0}^{m} c_i Y^i$ are non-zero polynomials in $K[Y]$ of degrees, respectively, l and m, such that*

$$\|B(Y)\| = \|b_l Y^l\| = \|C(Y)\| = 1; \tag{4.11}$$

2. *$\|u(Y)\| \le \mu$ and $\|v(Y)\| \le \mu$;*
3. *$\|u(Y)B(Y) + v(Y)C(Y) - 1\| < 1$;*
4. *$\deg \epsilon(Y) \le k$ and $l + m \le k$.*

Then there exist a pair of polynomials $(U(Y), V(Y))$, each in $K[Y]$, such that:

1. *$\|U(Y)\| \le \mu\|\epsilon(Y)\|$ and $\deg U(Y) \le k - l$;*
2. *$\|V(Y)\| \le \mu\|\epsilon(Y)\|$ and $\deg V(Y) \le l - 1$;*
3. *$\|U(Y)B(Y) + V(Y)C(Y) - \epsilon(Y)\| < \|\epsilon(Y)\|$.*

Proof. From hypothesis 3 we obtain

$$\|\epsilon(Y)u(Y)B(Y) + \epsilon(Y)v(Y)C(Y) - \epsilon(Y)\| < \|\epsilon(Y)\| \tag{4.12}$$

Let $Q(Y)$ be the quotient and $V(Y)$ be the remainder when $\epsilon(Y)v(Y)$ is divided by $B(Y)$; that is, $\epsilon(Y)v(Y) = Q(Y)B(Y) + V(Y)$, where $Q(Y)$ and $V(Y)$ are polynomials in $K[Y]$ with $\deg V(Y) \le l - 1$. By Lemma 4.6,

$$\|V(Y)\| \le \|\epsilon(Y)v(Y)\| \le \mu\|\epsilon(Y)\| \tag{4.13}$$

and

$$\|Q(Y)\| \leq \|\epsilon(Y)v(Y)\|/\|B(Y)\| \leq \mu\|\epsilon(Y)\| \tag{4.14}$$

Next,

$$\begin{aligned}
&\epsilon(Y)u(Y)B(Y) + \epsilon(Y)v(Y)C(Y) - \epsilon(Y) \\
&\quad = \epsilon(Y)u(Y)B(Y) + \big(Q(Y)B(Y) + V(Y)\big)C(Y) - \epsilon(Y) \\
&\quad = \big(\epsilon(Y)u(Y) + Q(Y)C(Y)\big)B(Y) + V(Y)C(Y) - \epsilon(Y) \\
&\quad = U'(Y)B(Y) + V(Y)C(Y) - \epsilon(Y),
\end{aligned} \tag{4.15}$$

where

$$U'(Y) = \epsilon(Y)u(Y) + Q(Y)C(Y). \tag{4.16}$$

Then,

$$\begin{aligned}
\|U'(Y)\| &\leq \max(\|\epsilon(Y)u(Y)\|, \|Q(Y)C(Y)\|) \\
&\leq \mu\|\epsilon(Y)\|
\end{aligned} \tag{4.17}$$

and

$$\|U'(Y)B(Y) + V(Y)C(Y) - \epsilon(Y)\| < \|\epsilon(Y)\|. \tag{4.18}$$

The polynomial $V(Y)$ already meets the requirements of the Lemma. We show that we can modify $U'(Y)$ to obtain the required polynomial $U(Y)$. Write

$$U'(Y) = \sum_i u_i Y^i. \tag{4.19}$$

If any monomial $u_i Y^i$ satisfies $\|u_i Y^i\| < \|\epsilon(Y)\|$, then we may replace u_i by 0; this will not affect the validity of (4.18). Define $U(Y)$ to be the polynomial obtained from $U'(Y)$ by replacing all such monomials $u_i Y^i$ by 0. Then,

$$\|U(Y)B(Y) + V(Y)C(Y) - \epsilon(Y)\| < \|\epsilon(Y)\|. \tag{4.20}$$

Put $j = \deg U(Y)$. If $j \leq k - l$, we are done. If not, then, the term of highest degree in the product $U(Y)B(Y)$ has degree $j+l > k$. Since $\deg V(Y)C(Y) \leq l-1+m < k$ and $\deg \epsilon(Y) \leq k$, the term of highest degree in the product $U(Y)B(Y)$ must also be the term of highest degree in the left-hand side of $U(Y)B(Y) + V(Y)C(Y) - \epsilon(Y)$. The norm of this term is $\|u_j Y^j\|\,\|b_l Y^l\| \geq \|\epsilon(Y)\|$. This contradicts (4.20) and shows that $j + l \leq k$, equivalently $\deg U(Y) \leq k - l$. □

For the remainder of this section we assume that λ is a rational power of $|\pi|$. Specifically, $\lambda = |\pi|^{s/r}$, where r and s are relatively prime integers with $r \geq 1$. In particular, we require that if $s = 0$, then $r = 1$. Under this assumption, the norm $\|A(Y)\|$ of any non-zero polynomial $A(Y) \in K[Y]$ will be an integral power of $|\pi|^{1/r}$.

We can now state the form of Hensel's Lemma that we use.

Theorem 4.21. (Hensel's Lemma) *Suppose that h is a non-negative integer and that we are given a 5-tuple of polynomials*

$$\big(R(Y), B_0(Y), C_0(Y), u(Y), v(Y)\big) \tag{4.22}$$

each with coefficients in K such that

1. *$R(Y)$ has degree k and satisfies $\|R(Y)\| = 1$;*
2. *$B_0(Y) = \sum_{i=0}^{l} b_i Y^i$ has degree l and satisfies $\|B_0(Y)\| = \|b_l Y^l\| = 1$;*
3. *$C_0(Y) = \sum_{i=0}^{m} c_i Y^i$ has degree m and satisfies $\|C_0(Y)\| = 1$;*
4. *$\|R(Y) - B_0(Y)C_0(Y)\| \le |\pi|^{(2h+1)/r}$;*
5. *$\|u(Y)\| \le |\pi|^{-h/r}, \quad \|v(Y)\| \le |\pi|^{-h/r}$;*
6. *$\|u(Y)B_0(Y) + v(Y)C_0(Y) - 1\| < 1$.*

Then there exist polynomials $B(Y)$ and $C(Y)$ in $K[Y]$ such that

1. *$R(Y) = B(Y)C(Y)$;*
2. *$\|B(Y) - B_0(Y)\| < |\pi|^{h/r}$;*
3. *$\|C(Y) - C_0(Y)\| < |\pi|^{h/r}$;*
4. *$\deg B(Y) = \deg B_0(Y)$.*

Proof. We first show that we may assume that $k \ge m + l$. If $k < l + m$, then the term of highest degree of $R(Y) - B_0(Y)C_0(Y)$ is $-b_l c_m Y^{m+1}$ whose norm, by hypotheses (2) and (4), satisfies

$$\|b_l Y^l\| \, \|c_m Y^m\| = \|b_l c_m Y^{l+m}\| \le |\pi|^{(2h+1)/r}, \tag{4.23}$$

so that $\|c_m Y^m\| \le |\pi|^{(2h+1)/r}$. It follows that if we replace $C_0(Y)$ by by the lower degree polynomial $C_0(Y) - c_m Y^m$ and replace m by the degree of this new $C_0(Y)$, then the hypotheses remain satisfied. For the remainder of this proof we assume that $k \ge l + m$.

We shall construct sequences of polynomials $\{B_i(Y)\}$ and $\{C_i(Y)\}$ for $i = 1, 2, \ldots$ such that

1. $\|B_i(Y) - B_{i-1}(Y)\| \le |\pi|^{(h+i)/r}$ and $\deg B_i(Y) = l$;
2. $\|C_i(Y) - C_{i-1}(Y)\| \le |\pi|^{(h+i)/r}$ and $\deg C_i(Y) \le m - l$;
3. $\|R(Y) - B_i(Y)C_i(Y)\| \le |\pi|^{(2h+i+1)/r}$.

Putting $B(Y) = \lim_{i\to\infty} B_i(Y)$ and $C(Y) = \lim_{i\to\infty} C_i(Y)$ will complete the proof.

We proceed by induction on the variable i, starting with $i = 1$. Put $\epsilon_i(Y) = R(Y) - B_{i-1}(Y)C_{i-1}(Y)$ so that, by hypothesis (when $i = 1$) or induction (when $i > 1$), $\|\epsilon_i(Y)\| \le |\pi|^{(2h+i)/r}$. Apply Lemma 4.9 to the 7-tuple

$$(k, |\pi|^{-h}, B_i(Y), C_i(Y), u(Y), v(Y), \epsilon_i(Y)). \tag{4.24}$$

Lemma 4.9 returns a pair of polynomials which we denote $(U_i(Y), V_i(U))$. These polynomials satisfy

1. $\|U_i(Y)\| \le |\pi|^{(h+i)/r}$ and $\deg U_i(Y) \le m-1$;
2. $\|V_i(Y)\| \le |\pi|^{(h+i)/r}$ and $\deg V_i(Y) \le l-1$;
3. $\|U_i(Y)B_0(Y) + V_i(Y)C_0(Y) - \epsilon_i(Y)\| \le |\pi|^{(2h+i+1)/r}$.

Define

$$B_i(Y) = B_{i-1}(Y) + V_i(Y), \qquad C_i(Y) = C_{i-1}(Y) + U_i(Y) \tag{4.25}$$

Then

$$\begin{aligned}
&\|R(Y) - B_i(Y)C_i(Y)\| \\
&\quad = \|R(Y) - \big(B_{i-1}(Y) + V_i(Y)\big)\big(C_{i-1}(Y) + U_i(Y)\big)\| \\
&\quad = \Big\|\big(R(Y) - B_{i-1}(Y)C_{i-1}(Y)\big) \\
&\qquad\quad - \big(U_i(Y)B_{i-1}(Y) + V_i(Y)C_{i-1}(Y)\big) - U_i(Y)V_i(Y)\Big\| \\
&\quad = \Big\|\big(\epsilon_i(Y) - \big(U_i(Y)B_{i-1}(Y) + V_i(Y)C_{i-1}(Y)\big)\big) \\
&\qquad\quad - U_i(Y)V_i(Y)\Big\| \\
&\quad \le \max\big(|\pi|^{2h+1}, |\pi|^{2h+2i}\big) \\
&\quad = |\pi|^{(2h+i+1)/r}
\end{aligned} \tag{4.26}$$

□

The proof of Hensel's Lemma consists of an algorithm. If only approximations to the factors $R(Y)$ and $B(Y)$ are needed, then the algorithm is finite. We shall call the algorithm *Hensel's Lemma*, also.

Now suppose that we are given a polynomial $R(Y)$ which is pure and whose Newton diagram has slope $-s/r$, where r and s are relatively prime integers with $r > 0$. The degree of $R(Y)$ must be a multiple of r, say kr. Both of the points $(0, \operatorname{ord} a_0)$ and $(kr, \operatorname{ord} a_{kr})$ must lie on this segment. We can write

$$R(Y) = \sum_{i=0}^{k} a_i \pi^{-is} Y^{ir} + \text{ [terms above the Newton diagram]} \tag{4.27}$$

where $|a_i| \le 1$ for $0 \le i \le k$, and where, In the $\lambda = |\pi|^{s/r}$ norm,

$$\|R(Y)\| = |a_0| = \|a_{kr}Y^{kr}\| = |a_k|. \tag{4.28}$$

Equation (4.27) can be restated as

$$\|R(Y) - \sum_{i=0}^{k} a_i \pi^{-is} Y^{ir}\| < \|R(Y)\|. \tag{4.29}$$

When this is the case we have

Corollary 4.30. *Suppose that $R(Y)$ is a pure polynomial of degree kr, of the form* (4.27) *which satisfies* (4.28) *and suppose further that the polynomial*

$R^*(Y) = \sum_{i=0}^{k} \bar{a}_i Y^i$ satisfies $R^*(Y) = \beta(Y)\gamma(Y)$ where $\beta(Y)$ and $\gamma(Y)$ are monic, relatively prime polynomials in $\overline{K}[Y]$. Then $R(Y) = B(Y)C(Y)$ where $B(Y)$ and $C(Y)$ are relatively prime polynomials in $K[Y]$ satisfying $B^*(Y) = \beta(Y)$ and $C^*(Y) = \gamma(Y)$.

Proof. By multiplying $R(Y)$ by an appropriate power of π, we may assume that $\|R(Y)\| = 1$. Suppose that $\deg \beta(Y) = l$ and $\deg \gamma(Y) = m$. There exist polynomials $\mu(Y)$ and $\nu(Y)$ in $\overline{K}[Y]$ such that $\mu(Y)\beta(Y) + \nu(Y)\gamma(Y) = 1$ and such that $\deg \mu(Y) < m$ and $\deg \nu(Y) < l$. Choose elements b_i, c_i, u_i, and v_i in K such that

$$\begin{aligned} \beta(Y) &= \sum_{i=0}^{l} \bar{b}_i Y^i, \qquad \gamma(Y) = \sum_{i=0}^{m} \bar{c}_i Y^i, \\ \mu(Y) &= \sum_{i=0}^{m-1} \bar{u}_i Y^i, \qquad \nu(Y) = \sum_{i=0}^{l-1} \bar{v}_i Y^i. \end{aligned} \tag{4.31}$$

Define

$$\begin{aligned} B_0(Y) &= \sum_{i=0}^{l} b_i \pi^{-is} Y^{ir}, \qquad C_0(Y) = \sum_{i=0}^{m} c_i \pi^{-is} Y^{ir}, \\ u(Y) &= \sum_{i=0}^{m-1} u_i \pi^{-is} Y^{ir}, \qquad v(Y) = \sum_{i=0}^{l-1} v_i \pi^{-is} Y^{ir}. \end{aligned} \tag{4.32}$$

Then $B_0(Y)^* = \beta(Y)$, $C_0(Y)^* = \gamma(Y)$, $u(Y)^* = \mu(Y)$ and $v(Y)^* = \mu(Y)$. Apply Theorem 4.21 with $h = 0$ to the 5-tuple

$$(R(Y), B_0(Y), C_0(Y), u(Y), v(Y)). \tag{4.33}$$

The result will be two polynomials $B(Y)$ and $C(Y)$ which meet the requirements of this corollary. □

The special case of this Corollary when $C(Y)$ is pure with horizontal Newton diagram appears as Lemma 4.1 in [6].

4.2 Hensel Factor

We may now define **Hensel Factor**, an important subroutine of our algorithm. It takes as input a triple $(K, F(X), A(X))$, where K is a field, $F(X)$ is a polynomial of degree ≥ 2 to be factored, and $A(X)$ is a non-zero polynomial of degree $< \deg F(X)$. We will say the algorithm *succeeds* if one of Lemmas 3.4, 3.5 or Corollary 4.30 apply. If Lemma 3.5 holds, then $(K, F(X), A(X))$ forms a certificate for the irreducibility of $F(X)$, and we are done. If Lemma 3.4 or Corollary 4.30 hold, then we have found a factor $G(X)$ of $F(X)$ over a field L, and we recursively call **p-adic Factor** with input $(L, G(X))$. If none of the lemmas apply, we say it *fails*.

Hensel Factor. Input $(K, F(X), A(X))$.

1. Compute $R(Y) = \mathrm{Res}_X(F(X), Y - A(X))$.

   Comment. Each of the elements $A(x)$, where x is a root of $F(X)$, is a root of $R(Y)$. If the resultant $R(Y)$ were a monomial, then the n distinct roots x of $F(X)$ would satisfy the polynomial $A(X)$, of degree $< n$. Thus $R(Y)$ is not a monomial.

2. There are now four sub-cases, at most one of which can hold:

   (a) The polynomial $R(Y)$ is not pure.

   Factor $R(Y)$ using Lemma 3.4. Then factor $F(X)$ using Lemma 2.3. Let $G(X)$ be a factor of least degree. Restart **p-adic Factor** with the pair $(K, G(X))$.

   (b) The polynomial $R(Y)$ is pure and $R^*(Y)$ can be written as a product of two relatively prime factors, each of degree ≥ 1 in $\overline{K}[X]$.

   Factor $R(Y)$ using Corollary 4.30 of Hensel's Lemma. Then factor $F(X)$ using Lemma 2.3. Let $G(X)$ be a factor of least degree. Restart **p-adic Factor** with the pair $(K, G(X))$.

   (c) The polynomial $R(Y)$ is pure and $R^*(Y)$ is the e^{th} power of an irreducible monic polynomial $\alpha(Y)$ of degree ≥ 2 in $\overline{K}[Y]$.

   Choose a polynomial $u(Y) \in K[Y]$ such that $\bar{u}(Y) = \alpha(Y)$. Denote by L the unramified extension field of K obtained by adjoining a root y of $u(Y)$ to K. Put $\beta(Y) = (Y - \bar{y})^e$ and put $\gamma(Y) = R^*(Y)/\beta(Y)$. Then $R^*(Y) = \beta(Y)\gamma(Y)$ where $(\beta(Y), \gamma(Y)) = 1$. By Corollary 4.30 we can factor $R(Y)$ as $R(Y) = B(Y)C(Y)$ where $B^*(Y) = \beta(Y)$. Factor $F(X)$ over L using Lemma 2.3 with $R_1(Y) = B(Y)$ and $R_2(Y) = C(Y)$. Let $F_1(X)$ be the factor of $F(X)$ corresponding to $R_1(Y)$. Restart **p-adic Factor** with the pair $(L, F_1(X))$.

   Comment. Note that the field L is determined uniquely by K and $\alpha(Y)$; it is independent of the specific choice of $u(Y)$ (see Artin [1, page 69, Theorem 2A]). Moreover, if x is a root of $F_1(X)$ in $\widetilde{K}$, then $\bar{y} = \overline{F_1(x)}$. Hence the field L is contained in the field $K(x)$.

   (d) The polynomial $R(Y)$ is pure and the slope of its Newton diagram is k/n where $(k, n) = 1$.

   By Lemma 3.5, $F(X)$ is irreducible and the algorithm terminates with the triple $(K, F(X), A(X))$.

3. None of the four cases (2a), (2b), (2c), or (2d) applies, so that $R^*(Y)$ is a power of a linear factor in $\overline{K}[Y]$.

   Return **failure**

5 Some Technical Lemmas

We state here some simple results which will be used in the next section. We first have a lemma from elementary number theory. Its proof is constructive.

Lemma 5.1. *Suppose that h is a positive integer and that for $1 \le j \le h$ we are given fractions s_j/r_j where r_j and s_j are relatively prime positive integers. Define $t_0 = 1$ and for $1 \le j \le h$, define $t_j = \mathrm{lcm}(r_1, r_2, \ldots, r_j)$. Then, for any integer u, there exist integers e_j, for $1 \le j \le h$, satisfying $0 \le e_j < t_j/t_{j-1}$ and such that*

$$\sum_{j=1}^{h} e_j s_j/r_j - u/t_h \tag{5.2}$$

is an integer.

Proof. The proof proceeds by induction on h. When $h = 1$, then $t_1 = r_1$, and the unique choice for e_1 is the least non-negative, integral solution to $e_1 s_1 \equiv u \pmod{r_1}$.

Suppose that $h > 1$. We will show that there exist integers v and e_h such that $0 \le e_h < t_h/t_{h-1}$ and such that

$$e_h s_h/r_h + v/t_{h-1} - u/t_h \tag{5.3}$$

is an integer. This will reduce the problem to the $h-1$ case with u replaced by v. Multiplying (5.3) by t_h shows that we must choose e_h and v to satisfy

$$e_h s_h t_h/r_h + v t_h/t_{h-1} \equiv u \pmod{t_h} \tag{5.4}$$

Now suppose that p is a prime dividing t_h, that $p^\alpha \| r_h$ (this means that p^α is the exact power of p dividing r_h), and that $p^\beta \| t_{h-1}$. Put $\gamma = \max(\alpha, \beta)$. Since $t_h = \mathrm{lcm}(t_{h-1}, r_h)$, we see that $p^\gamma \| t_h$. Then $p^{\gamma-\alpha} \| (t_h/r_h)$ and $p^{\gamma-\beta} \| (t_h/t_{h-1})$. If $\alpha = \gamma$, then p divides r_h, hence does not divide s_h, so that p does not divide $s_h t_h/r_h$. If $\beta = \gamma$, then p does not divide t_h/t_{h-1}. Thus p divides at most one of $s_h t_h/r_h$ and t_h/t_{h-1}. It follows that $s_h t_h/r_h$ and t_h/t_{h-1} are relatively prime. Hence there exists a solution e_h and v to (5.4) (even with equality replacing congruence). For any integer k the pair $(e_h + k t_h/t_{h-1}, v - k s_h t_h/r_h)$ is also a solution of (5.4). Replacing e_h by $e_h + k t_h/t_{h-1}$ for an appropriate integer k allows us to choose e_h to satisfy $0 \le e_h < t_h/t_{h-1}$. □

This immediately gives the following corollary, which will be used in the algorithm to construct a polynomial $E(X)$ with specified values of $E(x)$ for the roots x of $F(X)$.

Corollary 5.5. *Suppose that h, the fractions s_j/r_j and the integers t_j satisfy the hypotheses of Lemma 5.1. Suppose that $A_1, A_2, \ldots, A_h$ are elements of $\widetilde{\mathbb{Q}}_p$ such that* $\mathrm{ord}\, A_j = s_j/r_j$. *Then for any integer u there exist integers $e_1, e_2, \ldots, e_h$ satisfying $0 \le e_j < t_j/t_{j-1}$ and an integer e_0 such that* $\mathrm{ord}\, \pi^{e_0} \prod_{j=1}^{h} A_j^{e_j} = u/t$.

The next lemma shows that if a monic polynomial of degree m is "small" at $n > m$ distinct points, then at least two of these points must be "close" to each other. If the points are given in advance, then there is a limit to how "small" the polynomial can be at all n points.

Lemma 5.6. *Suppose that $x_1, x_2, \ldots, x_n$ are elements of $\widetilde{K}$ and that $A(X)$ is a monic polynomial in $\widetilde{K}[X]$ of degree $m < n$. Then* $\min_{j \neq j'} |x_j - x_{j'}|^m \leq \max_i |A(x_i)|$.

Proof. Put $\epsilon = \max_j |A(x_j)|$. We can write $A(X) = \prod_{i=1}^{m}(X - \theta_i)$ where the $\theta_i \in \widetilde{\mathbb{Q}}_p$ are the roots of $A(X)$. Then for each j,

$$\epsilon \geq |A(x_j)| = \prod_{i=1}^{m} |x_j - \theta_i|. \tag{5.7}$$

Not all of the factors $|x_j - \theta_i|$ on the right-hand side of (5.7) can be $> \epsilon^{1/m}$. Hence there must exist a value of i, call it $\sigma(j)$, such that $|x_j - \theta_{\sigma(j)}| \leq \epsilon^{1/m}$. By doing this for all j, we obtain a map σ from the set $\{1, 2, \ldots, n\}$ to the set $\{1, 2, \ldots, m\}$. Since $n > m$, there must be two values, $j \neq j'$ such that $\sigma(j) = \sigma(j')$. Call this common value k. Then both $|x_j - \sigma_k| \leq \epsilon^{1/m}$ and $|x_{j'} - \sigma_k| \leq \epsilon^{1/m}$. Hence $|x_j - x_{j'}| \leq \epsilon^{1/m}$ □

Corollary 5.8. *Suppose that $F(X)$ is a monic polynomial in $O_K[X]$ of degree n with distinct roots $x_1, x_2, \ldots, x_n$. If $A(X)$ is a monic polynomial in $K[X]$ of degree $m < n$, then, for at least one i, we have $|A(x_i)| \geq |\Delta_F|^m$.*

Proof. Because all $|x_i| \leq 1$, we have

$$\Delta_F = \prod_{i \neq j} |x_i - x_j| \leq \min_{i \neq j} |x_i - x_j| \tag{5.9}$$

Now apply Lemma 5.6. □

6 The *p*-Adic Factor Algorithm

In this section, we describe the main algorithm. It will find an irreducible factor $H(X)$ of $F(X)$ along with a certificate that $H(X)$ is irreducible. To completely factor $F(X)$, the algorithm may have to be repeated, perhaps several times, with $F(X)/H(X)$ replacing $F(X)$ until this quotient is 1.

The algorithm will attempt to factor $F(X)$ using **Hensel Factor** with $A(X) = X$. This will fail only when $F^*(X)$ has the form $(X - \alpha)^m$. When this occurs, the algorithm will systematically look for a polynomial $A(X) \in K[X]$ for which **Hensel Factor** succeeds.

Because the algorithm is recursive and both the polynomial to be factored and the local field may change during the course of the algorithm we will, for the

remainder of this paper, denote by $F_0(X)$ the original polynomial to be factored over the original field K_0.

The input to the algorithm is a pair $(K, F(X))$, where K is either K_0 or a finite, unramified extension of K_0, and $F(X)$ is a monic polynomial of degree $n \geq 2$ with coefficients in O_K dividing $F_0(X)$. We assume $F(X)$ has no multiple factors and $F(0) \neq 0$. Since we compute approximations to the factors, $F(X)$ will not in general be known exactly. In Section 8 we determine how much precision is needed to avoid errors in the factorization.

The **p-adic Factor** algorithm will return a field L which is an unramified extension of K of degree $\leq n$, a polynomial $G(X)$ in $L[X]$ dividing $F(X)$, and a polynomial $B(X) \in L[X]$ of degree $< \deg G(X)$. By Lemma 3.5, the triple $(L, G(X), B(X))$ provides the proof that $G(X)$ is irreducible.

By Lemma 2.9,

$$H(X) = \mathrm{Norm}_{L/K}\, G(X). \tag{6.1}$$

is an irreducible factor of $F(X)$. As noted above, the algorithm may then be called recursively on the pair $(K, F(X)/H(X))$ to complete the factorization of $F(X)$.

Section 6.1 presents the algorithm, after which Section 6.2 describes in more detail what certain steps are doing, and why they work.

6.1 The Algorithm

p-adic Factor. Input: $(K, F(X))$.

Step 1. Apply **Hensel Factor** to $(K, F(X), X)$ (in this case $\mathrm{Res}_X(F(X), Y - X) = F(Y)$).

Step 2. We reach this step only if **Hensel Factor** did not succeed in Step 1, so $F^*(X)$ is a power of a linear polynomial. Choose $\alpha \in \mathcal{A}$ such that

$$F(X) = (X^r - \alpha\pi^s)^m + \text{ [terms above the Newton diagram]} \tag{6.2}$$

where
(a) $\bar{\alpha}$ is the unique root of $F^*(X)$ in $\overline{K}$ and $\mathrm{ord}\,\alpha = 0$;
(b) $r < n$ and $m > 1$;
(c) $mr = n$; $\gcd(r, s) = 1$;
(d) the Newton diagram of $F(X)$ has slope $-s/r$.

Step 3. We initiate the outer loop by putting $A_1(X) = X$, $R_1(Y) = F(Y)$, $r_1 = r$, $s_1 = s$, $t_0 = 1$, and $t_1 = r_1$.

Step 4. (Outer loop) For $h = 1, 2, \ldots$, perform Steps 5 through 11.

Step 5. To begin the inner loop, put

$$\begin{aligned} B_0(X) &= A_h(X)^{t_h/t_{h-1}}, \\ S_0(Y) &= \mathrm{Res}_X(F(X), Y - B_0(X)), \\ u_0 &= s_h t_h^2/(r_h t_{h-1}). \end{aligned}$$

Step 6. (Inner Loop) For $i = 0, 1, \ldots$, perform Steps 7 through 10.

Step 7. Use Corollary 5.5 to choose integers e_j, for $0 \le j \le h$, such that
(a) $0 \le e_j \le t_j/t_{j-1} - 1$ when $1 \le j \le h$,
(b) $e_0 + \sum_{j=1}^{h} e_j s_j/r_j = u_i/t_h$ (in the notation of Corollary 5.5, $e_0 = u/t_h - \sum_{j=1}^{h} e_j s_j/r_j$).

Define a polynomial $E(X)$ by

$$E(X) = \pi^{e_0} A_1(X)^{e_1} A_2(x)^{e_2} \cdots A_h(X)^{e_h}. \tag{6.3}$$

Step 8. Define

$$C(X) = B_i(X)\, E(X)^{-1} \pmod{F(X)} \tag{6.4}$$

and

$$T(Y) = \mathrm{Res}_X(F(X), Y - C(X)). \tag{6.5}$$

Apply **Hensel Factor** to the triple $(K, F(X), C(X))$.

Step 9. Put $B(X) = B_i(X) - \alpha E(X)$ and $S(Y) = \mathrm{Res}_X(F(X), Y - B(X)$. Apply **Hensel Factor** to the triple $(K, F(X), B(X))$.

Step 10. If the common value $\mathrm{ord}\, B(x)$ can be written in the form u/t_h, where u is an integer, then put $B_{i+1}(X) = B(X)$, $S_{i+1}(Y) = S(Y)$, $u_{i+1} = u$. and continue the "inner loop" by returning to Step 6.

Step 11. Denote the common value of $\mathrm{ord}\, B(x)$ by s_{h+1}/r_{h+1}, where r_{h+1} and s_{h+1} are relatively prime, non-negative integers as before. Put $A_{h+1}(X) = B(X)$, $R_{h+1}(Y) = S(Y)$, and $t_{h+1} = \mathrm{lcm}(t_h, r_{h+1})$.
(a) If $t_{h+1} < n$ continue the "outer loop" by returning to Step 4, with h increased by 1.
(b) Otherwise use Corollary 5.5 to choose integers e_j for $0 \le j \le h+1$ such that
i. $0 \le e_j \le t_j/t_{j-1} - 1$ when $1 \le j \le h$ and
ii. $\sum_{j=1}^{h+1} e_j s_j/r_j - 1/t_{h+1} = e_0$;
Define $E(X)$ by

$$E(X) = \pi^{e_0} A_1(X)^{e_1} A_2(x)^{e_2} \cdots A_h(X)^{e_h} \tag{6.6}$$

and apply **Hensel Factor** to the triple $(K, F(X), E(X))$.

6.2 Discussion of the Algorithm

In Step 2, each (unknown) root x of $F(X)$ has $\mathrm{ord}\, x = s/r$ by (6.2). This shows that the ramification index of each of the n field extensions of the form $K(x)/K$ is divisible by r.

Starting with $A_1(X) = X$ at Step 3, the outer loop defines a finite sequence of polynomials $A_1(X), A_2(X), \ldots$ and a corresponding sequence of pairs of non-negative integers, $(r_1, s_1), (r_2, s_2), \ldots$, where each of the pairs (r_i, s_i) are relatively prime. We have $R_h(Y) = \mathrm{Res}_X(F(X), Y - A_h(X))$, $t_0 = 1$, and for $h \ge 0$, define $t_h = \mathrm{lcm}(r_1, r_2, \ldots, r_h)$. The the following properties are easily checked:

1. Each of the r_h and each of the t_h divides n.
2. The polynomial $A_h(X)$ is monic of degree t_{h-1}.
3. There exists an element $\alpha \in \mathcal{A}$ such that $\operatorname{ord} \alpha = 0$ and

$$R_h(Y) = (Y^{r_h} - \alpha\pi^{s_h})^{n/r_h} + \text{ [terms above the Newton diagram].}$$

It follows that for each root x of $F(X)$, we have

$$\operatorname{ord} A_h(x) = s_h/r_h. \tag{6.7}$$

Thus the multiplicative group generated by $|\pi|, |A_1(x)|, |A_2(x)|, \cdots, |A_h(x)|$ is independent of the choice of x and contains the value group of K. Hence, for each root x of $F(x)$, the ramification index of the field extension $K(x)/K$ is divisible by r_h.

4. The integer r_h does not divide t_{h-1} and for each root x of $F(X)$, the ramification index of of the field extension $K(x)/K$ is divisible by t_h. It follows that $t_1 < t_2 < \cdots < t_h \leq n$.

Since t_i is a proper divisor of t_{i+1}, we must have $h \leq \log_2 n$. This limits the number of steps of the outer loop.

To determine $A_{h+1}(X)$, we attempt in the inner loop to find a monic polynomial $B(X)$ of degree t_h satisfied by all roots x of $F(X)$. Since $F(X)$ has $n > t_h$ distinct roots, this attempt must fail. Its failure either leads to a situation where we can factor $F(X)$ using Hensel's lemma or leads to the determination of $A_{h+1}(X)$. The inner loop finds $A_{h+1}(X)$ by defining a sequence of polynomials

$$B_0(X), B_1(X), B_2(X), \ldots \tag{6.8}$$

and a corresponding, strictly increasing, finite sequence of non-negative integers $u_0 < u_1 < u_2, \ldots$.

Each polynomial $B_i(X)$ is monic of degree t_h. Each root x of $F(X)$ will satisfy $\operatorname{ord} B_i(x) = u_i/t_h$. Corollary 5.8 provides an upper bound for u_i and hence the sequence $B_0(X), B_1(X), \ldots$ will be finite.

In Step 7, we have constructed $E(X)$ so that $\operatorname{ord} E(x) = u_i/t_h$ for every root x of $F(X)$. Since $\deg A_j(X) \leq t_{j-1}$, we obtain, from Step 7a, have

$$\begin{aligned} \deg E(X) &\leq \sum_{j=1}^{h} (t_j/t_{j-1} - 1)t_{j-1} \\ &= \sum_{j=1}^{h} (t_j - t_{j-1}) \\ &= t_h - 1. \end{aligned} \tag{6.9}$$

In Step 8, (6.4) is valid because $E(X)$ and $F(X)$ have no common zeros. The polynomial $T(Y)$ is monic of degree n and, for each root x of $F(X)$, we have $|C(x)| = |B_i(x)/E(x)| = 1$. Consequently, the Newton diagram of $T(Y)$ is the horizontal line-segment connecting the points $(0, 0)$ and $(n, 0)$. It follows that

the polynomial $T^*(Y)$ is monic of degree n and its constant term is not zero. If **Hensel Factor** fails, then we can write

$$T(Y) = (Y - \alpha)^n + \text{ [terms above the Newton diagram]} \tag{6.10}$$

where $\alpha \in \mathcal{A}$ and $\operatorname{ord} \alpha = 0$.

After Step 10, since $B_i(X)$ is monic of degree t_h and $\deg E(X) < t_h$, $B(X)$ is monic of degree t_h. By the definition of α, we have $\operatorname{ord} B_i(x) - \alpha E(x) > 0$ for each root x of $F(X)$. It follows that $\operatorname{ord} B(x) > \operatorname{ord} B_i(X)$ for each such x. If **Hensel Factor** fails, then $\operatorname{ord} B(x)$ is the same for all roots x of $F(X)$ and is $> u_i/t_h$.

Put $\delta = \operatorname{ord} |\, \Delta_F \,|$. Step 6 will increase i by 1. Since t_h divides n and the the u_i are non-negative integers and strictly increasing we have $u_i/t_h \geq i/n$. By Corollary 5.8, we see that $u_i/t_h \leq \delta n$. Thus $i \leq \delta n^2$. This means that for each value of h, the inner-loop is performed at most δn^2 times.

In Step 11a, r_{h+1} does not divide t_h, so that $t_{h+1} > t_h$. In Step 11b, we have $\operatorname{ord} E(x) = 1/n$ for every root x of $F(X)$, so case 2d of **Hensel Factor** will succeed, and this will lead to finding an irreducible factor of $F_0(X)$.

7 Two Examples

We decided to implement the algorithm, both to verify its correctness and practicality, and to allow experimentation. The first decision was to choose a mathematical package in which to implement it. MAGMA [4] was the original choice, but a package to perform local field operations was delayed several times, so the implementation was done in GP instead. GP is a part of the PARI system developed by Henri Cohen [2]. It does support p-adic fields, and is flexible enough to support unramified extension fields of the p-adics relatively easily. A new version of MAGMA with local fields has recently appeared, so a port of the algorithm to MAGMA is planned.

The resulting code is available at the second author's web site [13]. Because of the overhead of GP, it is slower than the PARI routine `factorpadic` for most polynomials. An implementation in C using the PARI library would run in about the same time as `factorpadic` for most polynomials.

For an example of how the algorithm functions, we will factor the polynomial

$$F(X) = (X - 4)^2(X^2 - 2) + 2^{100} \tag{7.1}$$

over $\mathbb{Q}_2$.

If we apply **p-adic Factor** to this polynomial, it starts by attempting to apply **Hensel Factor**. The Newton diagram of $R(Y) = F(Y)$ is not pure, so using Hensel's Lemma we find factors

$$G_1(X) = (X^2 - 2) + (2^{101} + 2^{105} + \cdots)\, X + (2^{99} + 2^{102} + \cdots) \tag{7.2}$$

and

$$G_2(X) = (X - 4)^2 + (2^{101} + 2^{102} + \cdots)X + (2^{99} + 2^{100} + \cdots) \tag{7.3}$$

Attempting to factor $G_1(X)$, we call **Hensel Factor** again. This time, the Newton diagram is pure, and we are in subcase (2d). Thus $G_1(X)$ is irreducible.

$G_2(X)$ is also pure, but its slope and degree are both even, so **Hensel Factor** does not apply. We have $G_2^*(X) = (X-1)^2$.

In Step 2 of **p-adic Factor**, we have $\alpha = 1$, $r = 1$, $s = 2$, and $n = m = 2$. We arrive in Step 7 with $E(X) = 4$, $C(X) = X/4$, and

$$T(Y) = Y^2 - 2Y + (1 + 2^{95} + \cdots). \tag{7.4}$$

The Newton diagram of $T(Y)$ is now horizontal, but $T^*(Y) = (Y-1)^2$ is still a power of a linear polynomial, so the call to **Hensel Factor** in Step 8 fails.

In Step 9, we have $\alpha = 1$ and $B(X) = X - 4$. This gives

$$S(Y) = Y^2 + (2^{101} + \cdots)\, Y + (2^{99} + \cdots). \tag{7.5}$$

The call to **Hensel Factor** in Step 9 now goes to subcase (2d), and we have proved that $G_2(X)$ is irreducible, completing the factorization of $F(X)$.

Very few polynomials make it all the way through the inner loop more than once. One that does is

$$F(X) = (X^2 - 2 - 2^{20})(X^2 - 2 + 2^{20}) \tag{7.6}$$

over $\mathbb{Q}_2$.

We have $F^*(X) = (X-1)^2$, so **Hensel Factor** fails. In Step 2 we choose $\alpha = 1$, $r = 2$, $s = 1$, $m = 2$, and $n = 1$. Entering the inner loop, we find $E(X) = 2$, $C(X) = X^2/2$, and

$$T(Y) = Y^4 - 4Y^3 + (6 - 2^{39})\, Y^2 + (-4 + 2^{40})\, Y + (1 - 2^{39} + 2^{76}). \tag{7.7}$$

Again, **Hensel Factor** fails. In Step 9 we set $B(X) = X^2 - 2$, and have

$$S(Y) = Y^4 - 2^{41}Y^2 + 2^{80}. \tag{7.8}$$

Hensel Factor fails on $S(Y)$, and $\operatorname{ord} B(x) = 20$ for each root x of $F(X)$, so we continue the inner loop. Returning to Step 7, we have $E(X) = 2^{20}$, $C(X) = 2^{-20}X^2 - 2^{-19}$, and $T(Y) = Y^4 - 2Y^2 + 1$. Once again, **Hensel Factor** fails.

Finally, we succeed in Step 9. This time we have $B(X) = X^2 - 2 - 2^{20}$, and $S(Y) = Y^4 + 2^{22}Y^3 + 2^{42}Y^2$. The factor of Y^2 in $S(Y)$ yields the factor

$$G_1(X) = X^2 - 2 - 2^{20}. \tag{7.9}$$

Both this factor and the other one immediately are shown to be irreducible by subcase (2d) of **Hensel Factor**.

8 Bounds on Required Precision and Complexity

From the discussion in Section 6.2, it is clear that the loops of **p-adic Factor** will be executed a polynomial number of times in n and $\log |\, \Delta_F \,|$. Therefore, to

show that **p-adic Factor** is a random polynomial-time algorithm, we only need to bound the precision needed in the computations.

In general, we can only approximately compute the factors of the p-adic polynomial $F(X)$. This causes two problems. First, in the gcd computation in Lemma 2.3:

$$F_i(X) = \gcd(F(X), R_i(A(X)), \tag{8.1}$$

we do not know the R_i exactly, and so terms in the computation that appear to be zero may not be. In this situation it is difficult to give a reasonable a priori estimate of the accuracy of $R_i(Y)$ that is needed to compute the gcd to the desired accuracy.

To circumvent this difficulty, we give an alternative method of computing $F_i(X)$, which involves solving a system of linear equations.

Lemma 8.2. *Suppose that $F(X) \in K[X]$ is a monic polynomial of degree n with distinct roots $x_1, x_2, \ldots, x_n$ in the algebraic closure $\overline{K}$ of K. Suppose that $A(X) \in K[X]$. Put $y_i = A(x_i)$, and suppose that the y_i are distinct. Put $R(Y) = \mathrm{Res}_X(F(X), Y - A(X))$. Then there exists a polynomial $B(X) \in K[Y]$ of degree $\leq n-1$ such that $B(A(X)) \equiv X \pmod{F(X)}$. Furthermore, if $R(Y) = R_1(Y)R_2(Y)$ is a nontrivial factorization of $R(Y)$, then $F(X) = F_1(X)F_2(X)$ where $F_i(X) = \mathrm{Res}_Y(R_i(Y), X - B(Y))$. Finally, $\deg F_i(X) = \deg R_i(Y)$.*

Proof. We first show that the n polynomials $A(X)^k \pmod{F}(X)$ for $0 \leq k \leq n-1$ are linearly independent over K. Suppose that we have a relation

$$\sum_{i=0}^{n-1} b_i A(X)^i \equiv 0 \pmod{F(X)}. \tag{8.3}$$

Substituting the values $x = x_k$ into (8.3) yields the system of linear equations

$$\sum_{i=0}^{n-1} b_i y_k^i = 0 \quad \text{for} \quad 1 \leq i \leq n. \tag{8.4}$$

The matrix of the equations (8.4) is a Vandermonde. Since the y_k are distinct it is nonsingular. This shows that all of the b_i are zero. It follows that the equation

$$\sum_{i=0}^{n-1} b_i A(X)^i \equiv X \pmod{F(X)} \tag{8.5}$$

has a unique solution $b_0, b_1, \ldots b_{n-1}$. Put $B(Y) = \sum_{i=0}^{n-1} b_i Y^i$. Then

$$B(A(X)) = \sum_{i=0}^{n-1} b_i A(X)^i \equiv X \pmod{F(X)}. \tag{8.6}$$

Suppose that $\deg R_1(Y) = r$. By renumbering we may suppose that the roots of $R_1(Y)$ are $y_1, y_2, \ldots, y_r$ where $r < n$. The roots x of $F_1(X)$ are those x for which there exists y such that $R(y) = 0$ and $x - B(y) = 0$. Thus the roots of $F_1(X)$ are $x_1, x_2, \ldots, x_r$ where $x_i = B(y_i)$. This shows that $F_1(X)$ is a factor of $F(X)$ of degree r. Similarly, $F_2(X)$ is a factor of $F(X)$ of degree $n - r$. It is immediate from the definition of resultant that $\deg F_i(X) = \deg R_i(X)$. □

The other potential problem of using approximations to $R_i(Y)$ is that, if we do not use sufficient accuracy, the factorization might be changed. Corollaries 8.7 and 8.19 give bounds on the accuracy needed to preserve the correct factorization.

Corollary 8.7. *Suppose that $R(Y)$, $B_0(Y)$, and $C_0(Y)$ are polynomials in Y of degrees k, l, and m, respectively, and*

$$\|R(Y) - B_0(Y)C_0(Y)\| < |\operatorname{Res}_Y(B_0(Y), C_0(Y))|^2. \tag{8.8}$$

Then if the polynomials $R(Y)$, $B_0(Y)$, and $C_0(Y)$ satisfy hypotheses 1, 2, *and* 3 *of Hensel's Lemma, there exist an integer h and polynomials $u(Y)$ and $v(Y)$ such that h and the 5-tuple $\big(R(Y), B_0(Y), C_0(Y), u(Y), v(Y)\big)$ satisfy the hypotheses and hence the conclusions of Hensel's Lemma.*

Proof. Put

$$h = r \cdot \operatorname{ord} \operatorname{Res}(B_0(Y), C_0(Y)). \tag{8.9}$$

Then, using this value of h, hypothesis 4 of Hensel's Lemma is satisfied.

We will choose polynomials $u(Y)$ and $v(Y)$ in $K[Y]$ of degrees $\le m-1$ and $\le l-1$, repectively, to satisfy

$$u(Y)B_0(Y) + v(Y)C_0(Y) = 1. \tag{8.10}$$

Suppose that $u(Y) = \sum_{i=0}^{m-1} u_i Y^i$ and $v(Y) = \sum_{i=0}^{l-1} v_i Y^i$. Equation (8.10) amounts to a system of $l + m$ linear equations in the $l + m$ unknowns, u_0, $u_1, \ldots, u_{m-1}$ and $v_0, v_1, \ldots, v_{l-1}$. The matrix of this system of linear equations is, up to sign, the Sylvester (resultant) matrix of $B_0(Y)$ and $C_0(Y)$ (see, for example, [10], Section 3.3.2). Since the determinant of this matrix is non-zero, the coefficients of $u(Y)$ and $v(Y)$ are uniquely determined elements of K, not all 0. We may estimate them as elements of the field $\widetilde{K}$. Choose $\tau \in \widetilde{K}$ to satisfy $\tau^r = \pi^{-s}$ so that $|\tau| = \|\pi^{-r/s} = 1/\lambda$. Put

$$\begin{aligned} u^\tau(Y) &= u(Y/\tau), & B_0^\tau(Y) &= B_0(Y/\tau), \\ v^\tau(Y) &= v(Y/\tau), & C_0^\tau(Y) &= C_0(Y/\tau). \end{aligned} \tag{8.11}$$

Then,

$$\begin{aligned} \|u^\tau(Y)\|_1 &= \|u(Y)\|, & \|B_0^\tau(Y)\|_1 &= \|B_0(Y)\|, \\ \|v^\tau(Y)\|_1 &= \|v(Y)\|, & \|C_0^\tau(Y)\|_1 &= \|C_0(Y)\|. \end{aligned} \tag{8.12}$$

Substituting Y/τ for Y, equation (8.10) becomes

$$u^\tau(Y)B_0^\tau(Y) + v^\tau(Y)C_0^\tau(Y) = 1 \tag{8.13}$$

As above, equation (8.13) may be considered as a system of linear equations in the coefficients of $u^\tau(z)$ and $v^\tau(z)$, which may be obtained from the matrix of equation (8.10) by elementary row operations, giving

$$|u_i/\tau^i| \le 1/|\operatorname{Res}_Y(B_0(Y), C_0(Y))| \ . \tag{8.14}$$

It follows that

$$\|u(Y)\| \le |\operatorname{Res}_Y(B_0(Y), C_0(Y))|^{-1}, \tag{8.15}$$

and similarly

$$\|v(Y)\| \le |\operatorname{Res}_Y(B_0(Y), C_0(Y))|^{-1}. \tag{8.16}$$

Thus the remaining hypotheses of Hensel's Lemma hold. □

This corollary shows that if $R(Y)$ is computed to accuracy given by (8.9), then any factorization found will be correct. To show that a proof of irreducibility is also not changed by small pertubations of $R(Y)$, we first need two easy lemmas.

Lemma 8.17. *Suppose that $B(Y)$ and $C(Y)$ are polynomials in $K[Y]$ whose product $A(Y)$ is pure. Then both $B(Y)$ and $C(Y)$ are pure. Furthermore, the Newton diagrams of the three polynomials $A(Y)$, $B(Y)$, and $C(Y)$ have the same slope.*

Proof. This follows by repeated applications of Theorem 3.1 and Lemma 3.2 of Chapter 6 of [6].

Lemma 8.18. *Suppose that $A(Y)$ and $B(Y)$ are polynomials in $K[Y]$ of the same degree k. Suppose further that $\|A(Y) - B(Y)\| < \|A(Y)\|$. Then, if $A(Y)$ is pure, so is $B(Y)$ and their Newton diagrams have the same slope.*

Proof. Put $\alpha = \|A(Y))\|$. Suppose that $A(Y) = \sum_{i=0}^k a_i Y^i$ and that $B(Y) = \sum_{i=0}^k b_i Y^i$. Then $|a_i| \le \alpha\lambda^{-i}$ and $|a_i - b_i| < \alpha\lambda^{-i}$. It follows that $|b_i| \le \alpha\lambda^{-i}$. Since $|a_0| = |\alpha|$ and $|a_k| = |\alpha|\lambda^{-k}$, we see that $|b_0| = |\alpha|$ and that $|b_k| = |\alpha|\lambda^{-k}$.

Corollary 8.19. *Suppose that $R(Y)$ is an irreducible polynomial of degree n satisfying $\|R(Y)\| = 1$, so that, in particular, $R(Y)$ is pure. Suppose that the Newton diagram of $R(Y)$ has slope $-s/r \le 0$. If $R_0(Y)$ is a polynomial of degree n satisfying $\|R_0(Y)\| = 1$ and $\|R_0(Y) - R(Y)\| < \min(1, |\Delta_{R_0}|)^2$, then $R_0(Y)$ is irreducible.*

Proof. Suppose that $R_0(Y)$ factors as $R_0(Y) = B_0(Y)C_0(Y)$. By Lemma 8.18, $R_0(Y)$ is pure, and by Lemmas 8.17 both $B_0(Y)$ and $C_0(Y)$ are pure, and their Newton diagrams have slope $-s/r$. We may assume that $\|B_0(Y)\| = \|C_0(Y)\| = 1$. Using the definitions and standard properties of the resultant and discriminant (see Lang [19, pp 200–204]), we find that $|\,\Delta_{R_0}\,| = |\,\Delta_R\,|$, and that $|\operatorname{Res}(B_0(Y), C_0(Y))| \geq |\,\Delta_R\,|$. Hence

$$\|R(Y) - R_0(Y)\| < |\operatorname{Res}(B_0(Y), C_0(Y))|^2. \tag{8.20}$$

By Corollary 8.7, $R(Y)$ factors, contradicting the hypotheses. □

Theorem 8.21. *Let K be an extension of degree k of $\mathbb{Q}_p$, and $F(X) \in K[X]$ have degree n. Algorithm* **p-adic Factor** *will find an irreducible factor of $F(X)$ in random time*

$$O\left(n^{8+\epsilon} \log^3 |\,\Delta_F\,| \log^2 p^k\right). \tag{8.22}$$

Proof. By Corollaries 8.7 and 8.19, we will find the correct factorization if we compute terms to $O(|\,\Delta_F\,|^2)$ precision. Note that, although we are starting in an extension of degree k of $\mathbb{Q}_p$, we may need to go to an extension of degree n of that field.

The dominant computation is the resultant, which in worst case takes time $O(n^4 \log^2(|\,\Delta_F\,|^2 np^{nk}))$ (see [10], Section 3.3). From the discussion in Section 6.2, the outer loop of the algorithm will be executed at most $O(\log n)$ times, and the inner loop at most $O(n^2 \log |\,\Delta_F\,|)$ times. When **Hensel Factor** succeeds, we may have to call **p-adic Factor** on a factor of degree at most $n/2$, so that no more than $O(\log n)$ recursive calls will be needed. Combining these bounds, we have (8.22). □

The implied constant in (8.22) depends upon the choice of uniformizer π and representatives $\mathcal{A}$. Note that this is a pessimistic worst-case bound. Most polynomials factor on the first call to **Hensel Factor**, and it takes an effort to construct a polynomial which goes through the inner and outer loops more than once. Since we have not used fast arithmetic algorithms, and it is unlikely that all the worst cases can occur simultaneously, with a more detailed analysis the $n^{8+\epsilon}$ in (8.22) can be improved.

References

1. Emil Artin. *Algebraic numbers and algebraic functions.* Gordon and Breach, 1967.
2. C. Batut, K. Belabas, D. Bernardi, H. Cohen, and M. Olivier. User's guide to PARI-GP, for version 2.0.10, July 1998.
3. Elwyn R. Berlekamp. Factoring polynomials over large finite fields. *Mathematics of Computation*, 24:713–735, 1970.
4. W. Bosma, J. Cannon, and C. Playoust. The magma algebra system I: The user language. *J. Symb. Comp.*, 24:235–269, 1997.
5. David G. Cantor and Hans Zassenhaus. A new algorithm for factoring polynomials over finite fields. *Mathematics of Computation*, 36:587–592, 1981.

6. J. W. S. Cassels. *Local Fields*. Cambridge University Press, 1986. ISBN 0-521-30484-9 (hard cover) or 0-521-31525-5 (paper back).
7. A. L. Chistov. Efficient factorization of polynomials over local fields. *Soviet Math. Doklady*, 35:430–433, 1987. Translated from Russian original.
8. A. L. Chistov. Efficient factoring polynomials over local fields and its applicatons. In *Proceedings of the international congress of mathematicians, Kyoto, Japan, August 21-29, 1990*, pages 1509–1519, Vol. 2. Springer Verlag, 1991. ISBN 0-387-70047-1.
9. A. L. Chistov. Algorithm of polynomial complexity for factoring polynomials over local fields. *Journal of mathematical sciences*, 70:1912–1933, 1994. Translated from Russian original.
10. Henri Cohen. *A course in computational algebraic number theory*. Springer Verlag, 1994. ISBN 0-387-55640-0 or 3-546-55640-0.
11. David Ford and Pascal Letard. Implementing the round four maximal order algorithm. *Journal de Théorie des Nombres des Bordeaux*, 6:33–80, 1994.
12. Patrizia Gianni, Victor Miller, and Barry Trager. Decomposiition of algebras. Unpublished.
13. Daniel M. Gordon. `http://sdcc12.ucsd.edu/~xm3dg`. Web Site.
14. Fernando Gouvea. *p-adic Numbers*. Springer-Verlag, 1993. ISBN 0-387-56844-1.
15. W. B. Gragg. The Padé table and its relation to certain algorithms of numerical analysis. *SIAM Review*, 14:–62, 1972.
16. Helmut Hasse. *Number Theory*. Springer Verlag, 1980. ISBN 0-387-08275-1.
17. Kurt Hensel. *Theorie der Algebraischen Zahlen*. B. G. Teubner, 1908.
18. Dexter Kozen. Efficient resolution of singularities of plane curves. In *Proceedings 14th conference on foundations of software technology and theoretical computer science*, 1994.
19. Serge Lang. *Algebra*. Addison-Wesley, 1984.
20. R. Loos. Generalized polynomial remainder sequences. In B. Buchberger, G. E. Collins, and R. Loos, editors, *Computer Algebra Symbolic and Algebraic Computation, second edition*, pages 115–136. Springer Verlag, 1983. ISBN 0-387-81776-X.
21. Daniel A. Marcus. *Number Fields*. Springer-Verlag, 1977. ISBN 0-387-90279-1 or 3-540-90279-1.
22. Władysław Narkiewicz. *Elementary and analytic theory of algebraic numbers, second edition*. Springer Verlag, 1989. ISBN 0-387-51250-9.
23. Michael E. Pohst. *Computational algebraic number theory*. Birkhäuser Verlag, 1993. ISBN 0-8176-2913-0 or 3-7643-2913-0.
24. Michael E. Pohst and Hans Zassenhaus. *Algorithmic algebraic number theory*. Cambridge University Press, 1989. ISBN 0-521-33060-2.
25. Paulo Ribenboim. *The theory of classical valuations*. Springer Verlag, 1999. ISBN 0-387-98525-5.
26. Ian Stewart and David Tall. *Algebraic Number Theory*. Chapman and Hall, 1987. ISBN 0-412-29870-8 or 0-412-29690-X.
27. André Weil. *Basic Number Theory*. Springer Verlag, 1970.
28. Edwin Weiss. *Algebraic Number Theory*. McGraw-Hill, 1963.
29. Richard Zippel. *Effective Polynomial Computation*. Kluwer Academic Press, 1993. ISBN 0-7923-9375-9.