Factoring multivariate polynomials over algebraic number fields

Arjen K. Lenstra
Centrum voor wiskunde en informatica
Kruislaan 413
1098 SJ Amsterdam
The Netherlands

Abstract

We present an algorithm to factor multivariate polynomials over algebraic number fields that is polynomial-time in the degrees of the polynomial to be factored. The algorithm is an immediate generalization of the polynomial-time algorithm to factor univariate polynomials with rational coefficients.

1. Introduction

We show that the algorithm from [7] to factor univariate polynomials with rational coefficients can be generalized to multivariate polynomials with coefficients in an algebraic number field. As a result we get an algorithm that is polynomial-time in the degrees and the coefficient-size of the polynomial to be factored.

An outline of the algorithm is as follows. First the polynomial $f \in \mathbf{Q}(\alpha)[X_1, X_2, ..., X_t]$ is evaluated in a suitably chosen integer point $(X_2 = s_2, X_3 = s_3, ..., X_t = s_t)$. Next, for some prime number p, a p-adic irreducible factor $\tilde{h}$ of the resulting polynomial $\tilde{f} \in \mathbf{Q}(\alpha)[X_1]$ is determined up to a certain precision. We then show that the irreducible factor h_0 of f for which $\tilde{h}$ is a p-adic factor of $\tilde{h}_0$, belongs to a certain integral lattice, and that h_0 is relatively short in this lattice. This enables us to compute this factor h_0 by means of the so-called *basis reduction algorithm* (cf. [7: Section 1]).

As [7] is easily available, we do not consider it to be necessary to recall the basis reduction algorithm here; we will assume the reader to be familiar with this algorithm and its properties.

Although the algorithm presented in this paper is polynomial-time, we do not think it is a useful method for practical purposes. Like the other generalizations of the algorithm from [7], which can be found in [8; 9; 10; 11], the algorithm will be slow, because the basis reduction algorithm has to be applied to huge dimensional lattices with large entries. In practice, a combination of the methods from [6], [14], and [15] can be recommended (cf. [6]).

2. Preliminaries

In this section we introduce some notation, and we derive an upper bound for the coefficients of factors of multivariate polynomials over algebraic number fields.

Let the algebraic number field $\mathbf{Q}(\alpha)$ be given as the field of rational numbers $\mathbf{Q}$ extended by a root α of a prescribed *minimal polynomial* $F \in \mathbf{Z}[T]$ with leading coefficient equal to one; i.e. $\mathbf{Q}(\alpha) \simeq \mathbf{Q}[T]/(F)$. Similarly, we define $\mathbf{Z}[\alpha] = \mathbf{Z}[T]/(F)$ as a ring of polynomials in α over $\mathbf{Z}$ of degree $< l$, where l denotes the degree δF of F.

Let $f \in \mathbf{Q}(\alpha)[X_1, X_2, ..., X_t]$ be the polynomial to be factored, with the number of variables $t \geqslant 2$. By $\delta_i f = n_i$ we denote the degree of f in X_i, for $1 \leqslant i \leqslant t$. We often use n instead of n_1. We put $N_i = \prod_{k=i}^{t}(n_k + 1)$, and $N = N_1$. Let $lc_0(f) = f$. For $1 \leqslant i \leqslant t$ we define $lc_i(f) \in \mathbf{Q}(\alpha)[X_{i+1}, X_{i+2}, ..., X_t]$ as the leading coefficient with respect to X_i of $lc_{i-1}(f)$, and we put $lc(f) = lc_t(f)$. Finally, we define the *content* $\text{cont}(f) \in \mathbf{Q}(\alpha)[X_2, X_3, ..., X_t]$ of f as the greatest common divisor of the coefficients of f with respect to X_1. Without loss of generality we may assume that $2 \leqslant n_i \leqslant n_{i+1}$ for $1 \leqslant i < t$, that f is *monic* (i.e. $lc(f) = 1$), and that $\delta_i \operatorname{cont}(f) = 0$ for $2 \leqslant i \leqslant t$.

Let $d \in \mathbb{Z}_{>0}$ be such that $f \in \frac{1}{d}\mathbb{Z}[\alpha][X_1, X_2, ..., X_t]$, and let discr$(F)$ denote the discriminant of F. It is well-known (cf. [15]) that if we take $D = d\,|\mathrm{discr}(F)|$, then all monic factors of f are in $\frac{1}{D}\mathbb{Z}[\alpha][X_1, X_2, ..., X_t]$ (in fact it is sufficient to take $D = d{\cdot}s$, where s is the largest integer such that s^2 divides discr(F), but this integer s might be too difficult to compute).

We now introduce some notation, similar to [8: Section 1]. Suppose that we are given a prime number p such that

(2.1) $\quad p$ does not divide D.

For $G = \sum_i a_i T^i \in \mathbb{Z}[T]$ we denote by G_l or $G \bmod p^l$ the polynomial $\sum_i (a_i \bmod p^l) T^i \in (\mathbb{Z}/p^l\mathbb{Z})[T]$, for any positive integer l. Suppose furthermore that we are given some positive integer k, and that p is chosen in such a way that a polynomial $H \in \mathbb{Z}[T]$ exists such that

(2.2) $\quad H$ has leading coefficient equal to one,

(2.3) $\quad H_k$ divides F_k in $(\mathbb{Z}/p^k\mathbb{Z})[T]$,

(2.4) $\quad H_1$ is irreducible in $(\mathbb{Z}/p\mathbb{Z})[T]$,

(2.5) $\quad (H_1)^2$ does not divide F_1 in $(\mathbb{Z}/p\mathbb{Z})[T]$.

Clearly H_1 divides F_1 in $(\mathbb{Z}/p\mathbb{Z})[T]$, and $0 < \delta H \leqslant I$. In the sequel we will assume that conditions (2.1), (2.2), (2.3), (2.4), and (2.5) are satisfied.

By $\mathbb{F}_q$ we denote the finite field containing $q = p^{\delta H}$ elements. From (2.4) we have $\mathbb{F}_q \simeq (\mathbb{Z}/p\mathbb{Z})[T]/(H_1) \simeq \{\sum_{i=0}^{\delta H-1} a_i \alpha_1^i \colon a_i \in \mathbb{Z}/p\mathbb{Z}\}$, where $\alpha_1 = T \bmod (H_1)$ is a zero of H_1. Furthermore we put $W_k(\mathbb{F}_q) = (\mathbb{Z}/p^k\mathbb{Z})[T]/(H_k) = \{\sum_{i-1}^{\delta H-1} a_i \alpha_k^i \colon a_i \in \mathbb{Z}/p^k\mathbb{Z}\}$, where $\alpha_k = T \bmod (H_k)$ is a zero of H_k. Notice that $W_k(\mathbb{F}_q)$ is a ring containing q^k elements, and that $W_1(\mathbb{F}_q) \simeq \mathbb{F}_q$. For $a \in \mathbb{Z}[\alpha]$ we denote by $a \bmod (p^l, H_l) \in W_l(\mathbb{F}_q)$ the result of the canonical mapping from $\mathbb{Z}[\alpha] = \mathbb{Z}[T]/(F)$ to $W_l(\mathbb{F}_q) = (\mathbb{Z}/p^l\mathbb{Z})[T]/(H_l)$ applied to a, for $l = 1, k$. For $\tilde{g} = \sum_i \frac{a_i}{D} X_1^i \in \frac{1}{D}\mathbb{Z}[\alpha][X_1]$ we denote by $\tilde{g} \bmod (p^l, H_l)$ the polynomial $\sum_i (((D^{-1} \bmod p^l) a_i) \bmod (p^l, H_l)) X_1^i \in W_l(\mathbb{F}_q)[X_1]$ (notice that $D^{-1} \bmod p^l$ exists due to (2.1)).

We derive an upper bound for the height of a monic factor g of f. As usual, for $g = \sum_{i_1}\sum_{i_2} \cdots \sum_{i_t} \sum_j a_{i_1 i_2 \ldots i_t j} \alpha^j X_i^{i_1} X_2^{i_2} \ldots X_t^{i_t} \in \mathbb{Q}(\alpha)[X_1, X_2, ..., X_t]$, the *height* $g_{\max}$ is defined as $\max |a_{i_1 i_2 \ldots i_t j}|$, and the *length* $|g|$ as $(\sum a_{i_1 i_2 \ldots i_t j}^2)^{½}$. Similarly, for a polynomial h with complex coefficients, we define its height $h_{\max}$ as the maximum of the absolute values of its complex coefficients.

For any choice of $\alpha \in \{\alpha_1, \alpha_2, ..., \alpha_I\}$, where $\alpha_1, \alpha_2, ..., \alpha_I$ are the conjugates of α, we can regard g as a polynomial g_α with complex coefficients. We define $\|g\|$ as $\max_{1 \leqslant i \leqslant I} (g_{\alpha_i})_{\max}$. From [3] we have

$$\|g\| \leqslant e^{\sum_{i=1}^{t} n_i} \|f\|.$$

In [8: Section 4] we have shown that this leads to

$$g_{\max} \leqslant e^{\sum_{i=1}^{t} n_i} \|f\| I (I-1)^{(I-1)/2} |F|^{I-1} |\mathrm{discr}(F)|^{-½}. \tag{2.6}$$

From [13] we know that the length $|F|$ of F is an upper bound for the absolute value of the conjugates of α, so that

$$\|f\| \leqslant f_{\max} \sum_{i=0}^{I-1} |F|^i,$$

which yields, combined with (2.6),

$$g_{\max} \leqslant e^{\sum_{i=1}^{t} n_i} f_{\max} I(I-1)^{(I-1)/2} |F|^{I-1} |\mathrm{discr}(F)|^{-1/2} \sum_{i=0}^{I-1} |F|^i. \tag{2.7}$$

The upper bound for the height of monic factors of f, as given by the right hand side of (2.7), will be denoted by B_f. Because $|\mathrm{discr}(F)| \geqslant 1$, we find

$$\log B_f = O\left(\sum_{i=1}^{t} n_i + \log f_{\max} + I \log(I\,|F|)\right). \tag{2.8}$$

3. Factoring multivariate polynomials over algebraic number fields

We describe an algorithm to compute the irreducible factorization of f in $\mathbb{Q}(\alpha)[X_1, X_2, ..., X_t]$.

Let $s_2, s_3, ..., s_t \in \mathbb{Z}_{>0}$ be a $(t-1)$-tuple of integers. For $g \in \mathbb{Q}(\alpha)[X_1, X_2, ..., X_t]$ we denote by $\tilde g_j$ the polynomial $g \bmod ((X_2 - s_2), (X_3 - s_3), ..., (X_j - s_j)) \in \mathbb{Q}(\alpha)[X_1, X_{j+1}, X_{j+2}, ..., X_t]$; i.e. $\tilde g_j$ is g with s_i substituted for X_i, for $2 \leqslant i \leqslant j$. Notice that $\tilde g_1 = g$ and that $\tilde g_j = \tilde g_{j-1} \bmod (X_j - s_j)$. We put $\tilde g = \tilde g_t$.

Suppose that a polynomial $\tilde h \in \mathbb{Z}[\alpha][X_1]$ is given such that

(3.1) $\tilde h$ is monic,

(3.2) $\tilde h \bmod (p^k, H_k)$ divides $\tilde f \bmod (p^k, H_k)$ in $W_k(\mathbb{F}_q)[X_1]$,

(3.3) $\tilde h \bmod (p, H_1)$ is irreducible in $\mathbb{F}_q[X_1]$,

(3.4) $(\tilde h \bmod (p, H_1))^2$ does not divide $\tilde f \bmod (p, H_1)$ in $\mathbb{F}_q[X_1]$.

We put $l = \delta_1 \tilde h$, so $0 < l \leqslant n$. By $h_0 \in \frac{1}{D}\mathbb{Z}[\alpha][X_1, X_2, ..., X_t]$ we denote the unique, monic, irreducible factor of f such that $\tilde h \bmod (p^k, H_k)$ divides $\tilde h_0 \bmod (p^k, H_k)$ in $W_k(\mathbb{F}_q)[X_1]$ (cf. (3.2), (3.3), (3.4)).

(3.5) Let $m = m_1, m_2, m_3, ..., m_t$ be a t-tuple of integers satisfying $l \leqslant m < n$ and $0 \leqslant m_i \leqslant \delta_i lc_{i-1}(f)$ for $2 \leqslant i \leqslant t$, and let $M = 1 + I \sum_{i=1}^{t} m_i N_{i+1}$ (where of course $N_{t+1} = 1$). We define $L \subset (\frac{\mathbb{Z}}{D})^M$ as the lattice of rank M, consisting of the polynomials $g \in \frac{1}{D}\mathbb{Z}[\alpha][X_1, X_2, ..., X_t]$ for which

(i) $\delta_1 g \leqslant m$ and $\delta_i g \leqslant n_i$ for $2 \leqslant i \leqslant t$;

(ii) If $\delta_j lc_{j-1}(g) = m_j$ for $1 \leqslant j \leqslant i$, then $\delta_{i+1} lc_i(g) \leqslant m_{i+1}$ for $1 \leqslant i < t$;

(iii) If $\delta_i lc_{i-1}(g) = m_i$ for $1 \leqslant i \leqslant t$, then $lc(g) \in \mathbb{Z}$;

(iv) $\tilde h \bmod (p^k, H_k)$ divides $\tilde g \bmod (p^k, H_k)$ in $W_k(\mathbb{F}_q)[X_1]$.

Here M-dimensional vectors and polynomials satisfying conditions (i), (ii), and (iii), are identified in the usual way (cf. [8: (2.6); 11: (2.2)]). For notational convenience we only give a basis for L in the case that $m_i = n_i$ for $2 \leqslant i \leqslant t$; the general case can easily be derived from this:

$$\begin{aligned} &\{\tfrac{1}{D} p^k \alpha^j X_1^i : 0 \leqslant j < \delta H, 0 \leqslant i < l\} \\ \cup\ &\{\tfrac{1}{D} \alpha^{j-\delta H} H(\alpha) X_1^i : \delta H \leqslant j < I, 0 \leqslant i < l\} \\ \cup\ &\{\tfrac{1}{D} \alpha^j \tilde h X_1^{i-l} : 0 \leqslant j < I, l \leqslant i \leqslant m\} \\ \cup\ &\{\tfrac{1}{D} \alpha^j X_1^{i_1} \prod_{r=2}^{t} (X_r - s_r)^{i_r} : 0 \leqslant j < I, 0 \leqslant i_1 \leqslant m, 0 \leqslant i_r \leqslant n_r \\ &\qquad\qquad \text{for } 2 \leqslant r \leqslant t, (i_2, i_3, ..., i_t) \neq (0, 0, ..., 0),\end{aligned}$$

$$\text{and } (i_1, i_2, i_3, \ldots, i_t) \neq (m, n_2, n_3, \ldots, n_t)\}$$

$$\cup \; \{X_1^m \prod_{r=2}^{t} (X_r - s_r)^{n_r}\}$$

(cf. [8: (2.6); 11: (2.19)], (2.2), and (3.1)).

(3.6) **Proposition.** *Let b be a non-zero element of L and let*

$$\tilde{B}_j = f_{\max}^m b_{\max}^n (n+m)! \left[DN_2 (1+F_{\max})^{I-1} \prod_{i=2}^{j} s_i^{n_i} \right]^{n+m}, \tag{3.7}$$

for $1 \leqslant j \leqslant t$, where $f_{\max}^m$ denotes $(f_{\max})^m$.

Suppose that

$$s_j \geqslant ((n+m)n_j + 1)^{1/2} \tilde{B}_{j-1} \tag{3.8}$$

for $2 \leqslant j \leqslant t$, and

$$p^{k\delta H} \geqslant |F|^{I-1} (I^{1/2} \tilde{B}_t)^I. \tag{3.9}$$

Then $gcd(f, b) \neq 1$ in $\mathbf{Q}(\alpha)[X_1, X_2, \ldots, X_t]$.

Proof. Denote by $R = R(Df, Db) \in \mathbf{Z}[\alpha][X_2, X_3, \ldots, X_t]$ the resultant of Df and Db (with respect to the variable X_1). An outline of the proof is as follows. First we prove that an upper bound for $(\tilde{R}_j)_{\max}$ is given by $\tilde{B}_j$. Combining this with (3.8), we then see that $X_j = s_j$ cannot be a zero of $\tilde{R}_{j-1}$ if $\tilde{R}_{j-1} \neq 0$, for $2 \leqslant j \leqslant t$. This implies that the assumption that $R \neq 0$ (i.e. $gcd(f, b) = 1$) leads to $\tilde{R} \neq 0$. We then apply a result from [6], and we find with (3.9) that $\tilde{R} \bmod (p^k, H_k) \neq 0$. But this is a contradiction, because $\tilde{h} \bmod (p^k, H_k)$ divides both $\tilde{f} \bmod (p^k, H_k)$ and $\tilde{b} \bmod (p^k, H_k)$ in $W_k(\mathbf{F}_q)[X_1]$. We conclude that $R = 0$, so that $gcd(f, b) \neq 1$ in $\mathbf{Q}(\alpha)[X_1, X_2, \ldots, X_t]$.

If a and b are two polynomials in any number of variables over $\mathbf{Q}(\alpha)$, having l_a and l_b terms respectively, then

$$(a \cdot b)_{\max} \leqslant a_{\max} b_{\max} \min(l_a, l_b)(1+F_{\max})^{I-1}. \tag{3.10}$$

From (3.10) we easily derive an upper bound for $(\tilde{R}_j)_{\max}$, because $\tilde{R}_j \in \mathbf{Z}[\alpha][X_{j+1}, X_{j+2}, \ldots, X_t]$ is the resultant of $D\tilde{f}_j$ and $D\tilde{b}_j$:

$$(\tilde{R}_j)_{\max} \leqslant (D\tilde{f}_j)_{\max}^m (D\tilde{b}_j)_{\max}^n (n+m)! N_{j+1}^{n+m-1} (1+F_{\max})^{(I-1)(n+m-1)}. \tag{3.11}$$

It follows from $\tilde{f}_j = \tilde{f}_{j-1} \bmod (X_j - s_j)$, that $(\tilde{f}_j)_{\max} \leqslant (\tilde{f}_{j-1})_{\max}(n_j + 1)s_j^{n_j}$, so that

$$(\tilde{f}_j)_{\max} \leqslant f_{\max} \prod_{i=2}^{j} (n_i + 1) s_i^{n_i}. \tag{3.12}$$

Combining (3.11), (3.12), and a similar bound for $(\tilde{b}_j)_{\max}$, we obtain

$$(\tilde{R}_j)_{\max} < f_{\max}^m b_{\max}^n (n+m)! (DN_2 \prod_{i=2}^{j} s_i^{n_i})^{n+m} (1+F_{\max})^{(I-1)(n+m-1)}, \tag{3.13}$$

for $1 \leqslant j < t$. (Remark that (3.13) with "$<$" replaced by "$\leqslant$" holds for $j = t$.)

Now assume, for some j with $2 \leqslant j \leqslant t$, that $\tilde{R}_{j-1}$ is unequal to zero. We prove that $\tilde{R}_j \neq 0$. Because $\tilde{R}_j = \tilde{R}_{j-1} \bmod (X_j - s_j)$, the condition $\tilde{R}_j = 0$ would imply that all polynomials in $\mathbf{Z}[X_j]$ that result from $\tilde{R}_{j-1}$ by grouping together all terms with identical exponents in α and X_{j+1} up to X_t, have $(X_j - s_j)$ as a factor. These polynomials have degree (in X_j) at most $(n+m)n_j$, so that we get, with the result from [12], that

$$|s_j| \leqslant ((n+m)n_j + 1)^{1/2} (\tilde{R}_{j-1})_{\max}.$$

Combined with (3.13) and (3.7) this is a contradiction with (3.8). We conclude that $\tilde{R}_j \neq 0$ if $\tilde{R}_{j-1} \neq 0$ for any j with $2 \leqslant j \leqslant t$, so that the assumption $gcd(f, b) = 1$ (i.e. $R \neq 0$) leads to $\tilde{R} \neq 0$.

Assume that $H_k(\mathrm{T})$ divides $\tilde{R}(\mathrm{T}) \in \mathbb{Z}[\mathrm{T}]$ in $(\mathbb{Z}/p^k\mathbb{Z})[\mathrm{T}]$, i.e. $\tilde{R} \bmod (p^k, H_k) = 0$. The polynomial $H_k(\mathrm{T})$ is also a divisor of $F(\mathrm{T})$ in $(\mathbb{Z}/p^k\mathbb{Z})[\mathrm{T}]$, so that $gcd(F(\mathrm{T}), \tilde{R}(\mathrm{T})) = 1$ and [6: Theorem 2] lead to

$$p^{k\delta H} \leqslant |F|^{I-1}(I^{1/2}\tilde{R}_{\max})^I .$$

With the remark after (3.13) and (3.7) this is a contradiction with (3.9), so that $\tilde{R} \bmod (p^k, H_k) \neq 0$. This concludes the proof of (3.6). □

(3.14) **Proposition.** *Let $b_1, b_2, \ldots, b_M$ be a reduced basis for L (cf. [7: Section 1]), where L and M are as in (3.5), and let*

(3.15) $$B_j = (n+m)!(M2^{M-1})^{n/2} \left[B_f DN_2(1+F_{\max})^{I-1} \prod_{i=2}^{j} s_i^{n_i} \right]^{n+m} ,$$

for $2 \leqslant j \leqslant t$, where B_f is as in Section 2. Suppose that

(3.16) $$s_j \geqslant ((n+m)n_j + 1)^{1/2} B_{j-1}$$

for $2 \leqslant j \leqslant t$, that

(3.17) $$p^{k\delta H} \geqslant |F|^{I-1}(I^{1/2}B_t)^I ,$$

and that f does not contain multiple factors. Then

(3.18) $$(b_1)_{\max} \leqslant (M2^{M-1})^{1/2} B_f$$

and h_0 divides b_1, if and only if $h_0 \in L$.

Proof. If h_0 divides b_1, then $h_0 \in L$, because $b_1 \in L$; this proves the "if"-part.

To prove the "only if"-part, suppose that $h_0 \in L$. Because h_0 is a monic factor of f, we have from (2.7) that $(h_0)_{\max} \leqslant B_f$. With [7: (1.11)] and $h_0 \in L$ this gives $|b_1| \leqslant (M2^{M-1})^{1/2} B_f$, so that (3.18) holds, because $(b_1)_{\max} \leqslant |b_1|$. Because of (3.18), (3.16), (3.17), (3.15), and the definition of B_f, we can apply (3.6), which yields $gcd(f, b_1) \neq 1$.

Now suppose that h_0 does not divide b_1. This implies that h_0 also does not divide $r = gcd(f, b_1)$, where r can be assumed to be monic. But then $\tilde{h} \bmod (p^k, H_k)$ divides $(\tilde{f}/\tilde{r}) \bmod (p^k, H_k)$, so that Proposition (3.6) can be applied with f replaced by f/r. Conditions (3.8) and (3.9) are satisfied because $(f/r)_{\max} \leqslant B_f$ (cf. (2.7)) and because of (3.16), (3.17), and (3.15). It follows that $gcd(f/r, b_1) \neq 1$, which contradicts $r = gcd(f, b_1)$ because f does not contain multiple factors. □

(3.19) We describe how to compute the irreducible factor h_0 of f. Suppose that f does not contain multiple factors, and that the polynomial $\tilde{h}$, the $(t-1)$-tuple $s_2, s_3, \ldots, s_t$, and the prime power p^k are chosen such that (3.1), (3.2), (3.3), (3.4), (3.16), and (3.17) are satisfied with, for (3.16) and (3.17), m replaced by $n-1$. Remember that we also have to take care that conditions (2.1), (2.2), (2.3), (2.4), and (2.5) on p and H are satisfied.

We apply the basis reduction algorithm (cf. [7: Section 1]) to a sequence of M_j-dimensional lattices as in (3.5), where the $M_j = 1 + I\sum_{i=1}^{t} m_i N_{i+1}$ run through the range of admissible values for $m_1, m_2, \ldots, m_t$ (cf. (3.5)), in such a way that $M_j < M_{j+1}$. (So, for $m = l, l+1, \ldots, n-1$, and $m_i = 0, 1, \ldots, \delta_i lc_{i-1}(f)$ for $i = t, t-1, \ldots, 2$ in succession.) According to (3.14), the first vector b_1 that we find that satisfies (3.18) equals $\pm h_0$ (remember that b_1 belongs to a basis for the lattice), so that we can stop if such a vector is found. If for none of the lattices a vector satisfying (3.18) is found, then h_0 is not contained in any of these lattices according to (3.14), so that $h_0 = f$.

(3.20) **Proposition.** *Assume that the conditions in (3.19) are satisfied. The polynomial h_0 can be computed in $O((\delta_1 h_0 IN_2)^4 k \log p)$ arithmetic operations on integers having binary length $O(INk \log p)$.*

Proof. Observing that $\log(INp^{2k}) = O(k \log p)$ (cf (3.17), (3.15), and (2.8)), the proof immediately follows from (3.19), (3.5), and [7: (1.26), (1.37)]. □

(3.21) We now show how $s_2, s_3, \ldots, s_t$ and p can be chosen in such a way that the conditions in (3.19) can be satisfied. The algorithm to factor f then easily follows by repeated application of (3.19).

We assume that f does not contain multiple factors, so that the resultant $R = R(df, df')$ of df and its derivative df' with respect to X_1 is unequal to zero. First we choose $s_2, s_3, \ldots, s_t \in \mathbb{Z}_{>0}$ minimal such that (3.16) is satisfied with m replaced by $n-1$. It follows from (3.16), (3.15), (2.8) and $\log D = O(\log d + I \log(I\,|F|))$ (because $D = d\,|\mathrm{discr}(F\,|)$, that

$$\begin{aligned}\log s_j &= O(\log((n+m)n_j) + \log B_{j-1}) \\ &= O(InN + n(\log B_f + \log D + I \log(1+F_{\max}) + \sum_{i=1}^{j-1} n_i \log s_i)) \\ &= O(n(IN + \log(df_{\max}) + I \log(I\,|F\,|) + \sum_{i=1}^{j-1} n_i \log s_i))\end{aligned}$$

for $2 \leqslant j \leqslant t$, so that

$$\log s_j = O(n(IN + \log(df_{\max}) + I \log(I\,|F\,|)) \prod_{i=2}^{j-1} (1 + nn_i))$$

and

$$\sum_{i=2}^{t} n_i \log s_i = O(n^{t-2} N (IN + \log(df_{\max}) + I \log(I\,|F\,|))). \tag{3.22}$$

From the proof of (3.6) it follows that, for this choice of $s_2, s_3, \ldots, s_t$ the resultant $R \in \mathbb{Z}[\alpha]$ of $d\tilde{f}$ and $d\tilde{f}'$ is unequal to zero.

Next we choose p minimal such that p does not divide D or $\mathrm{discr}(F)$, and such that $\tilde{R} \not\equiv 0 \bmod p$. Clearly

$$\prod_{q \text{ prime},\, q<p} q \leqslant d \,\mathrm{discr}(F)\, \tilde{R}_{\max}$$

which yields, together with

$$\prod_{q \text{ prime},\, q<p} q > e^{Ap}$$

for all $p > 2$ and some constant $A > 0$ (cf. [4: Section 22.2]), that

$$p = O(\log d + I \log(I\,|F\,|) + \log \tilde{R}_{\max}). \tag{3.23}$$

Similar to (3.13) we obtain

$$\tilde{R}_{\max} \leqslant f_{\max}^{2n-1} n^n (2n-1)! \left[dN_2 \prod_{i=2}^{t} s_i^{n_i} \right]^{2n-1} (1+F_{\max})^{(I-1)(2n-2)},$$

so that we get, using (3.22)

$$\log \tilde{R}_{\max} = O(n^{t-1} N (IN + \log(df_{\max}) + I \log(I\,|F\,|))).$$

Combining this with (3.23) we conclude that

$$p = O(n^{t-1} N (IN + \log(df_{\max}) + I \log(I\,|F\,|))). \tag{3.24}$$

Notice that (2.1) is now satisfied. In order to compute a polynomial $H \in \mathbb{Z}[T]$ satisfying (2.2), (2.4), (2.5), and (2.3) with k replaced by 1, we factor $F \bmod p$ by means of Berlekamp's algorithm [5: Section 4.6.2] and we choose H as an irreducible factor of $F \bmod p$ for which $\tilde{R} \bmod (p, H_1) \neq 0$; such a polynomial H exists because $\tilde{R} \bmod p \neq 0$. Conditions (2.4) and (2.3) with k replaced by 1 are clear from the construction of H, and because we may assume that H

has leading coefficient equal to one, (2.2) also holds. The condition that $\mathrm{discr}(F) \bmod p \neq 0$, finally, guarantees that $F \bmod p$ does not contain multiple factors, so that (2.5) is satisfied.

We choose k minimal such that (3.17) holds, so that

$$k \log p = O(I(InN + n \log(df_{\max}) + In \log(\mathrm{I}\,|F|) + n \sum_{i=2}^{t} n_i \log s_i) + \log p)$$

(cf. (3.15) and (2.8)), which gives, with (3.22) and (3.24)

$$k \log p = O(In^{t-1}N(IN + \log(df_{\max}) + I \log(I\,|F|))). \tag{3.25}$$

Now we apply Hensel's lemma [5: Exercise 4.6.22] to modify H in such a way that (2.3) holds for this value of k (this is possible because (2.3) already holds for $k = 1$), and finally we apply Berlekamp's algorithm as described in [1: Section 5] and Hensel's lemma as in [14] to compute the irreducible factorization of $\tilde{f} \bmod (p^k, H_k)$ in $W_k(\mathbb{F}_q)[\mathrm{X}_1]$. Condition (3.4) is satisfied for each irreducible factor $\tilde{h} \bmod (p^k, H_k)$ of $\tilde{f} \bmod (p^k, H_k)$ because $\tilde{R} \bmod (p, H_1) \neq 0$, and (3.1), (3.2), and (3.3) are clear from the construction of $\tilde{h}$.

We have shown how to choose $s_2, s_3, \ldots, s_t$ and p, and how to satisfy the conditions in (3.19). We are now ready for our theorem.

(3.26) **Theorem.** *Let f be a monic polynomial in $\frac{1}{d}\mathbb{Z}[\alpha][\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_t]$ with $t \geq 2$, of degree n_i in X_i, and $2 \leq n = n_1 \leq n_2 \leq \ldots \leq n_t$. The irreducible factorization of f can be found in $O(n^{t-1}(IN)^5(IN + \log(df_{\max}) + I \log(I\,|F|)))$ arithmetic operations on integers having binary length $O(n^{t-1}(IN)^2(IN + \log(df_{\max}) + I \log(I\,|F|)))$, where $N = \prod_{i=1}^{t}(n_i + 1)$.*

Proof. If f does not contain multiple factors, then f can be factored by repeated application of (3.19). In that case (3.26) follows from (3.21), (3.20), (3.25), and the well-known estimates for the application of Berlekamp's algorithm and Hensel's lemma (cf. [5; 1] and [16]).

If f contains multiple factors, then we first have to compute the monic *gcd* g of f and its derivative with respect to X_1, and the factoring algorithm is then applied to f/g. The cost of factoring f/g satisfies the same estimates as above, because $(f/g)_{\max} \leq B_f$ (cf. (2.7)), and this dominates the costs of the computation of g, which can be done by means of the subresultant algorithm (cf. [2]). □

References

1. E.R. Berlekamp, Factoring polynomials over large finite fields, *Math. Comp.* **24** (1970), 713-735.
2. W.S. Brown, The subresultant PRS algorithm, *ACM Transactions on mathematical software* **4** (1978), 237-249.
3. A.O. Gel'fond, Transcendental and algebraic numbers, Dover Publ., New York 1960.
4. G.H. Hardy, E.M.Wright, An introduction to the theory of numbers, Oxford University Press 1979.
5. D.E. Knuth, The art of computer programming, vol. 2, Seminumerical algorithms, Addison-Wesley, Reading, second edition 1981.
6. A.K. Lenstra, Lattices and factorization of polynomials over algebraic number fields, Proceedings Eurocam 82, LNCS 144, 32-39.
7. A.K. Lenstra, H.W. Lenstra, Jr., L. Lovász, Factoring polynomials with rational coefficients, *Math. Ann.* **261** (1982), 515-534.

8. A.K. Lenstra, Factoring polynomials over algebraic number fields, Report IW 213/82, Mathematisch Centrum, Amsterdam 1982 (also Proceedings Eurocal 83, LNCS 162, 245-254).

9. A.K. Lenstra, Factoring multivariate polynomials over finite fields, Report IW 221/83, Mathematisch Centrum, Amsterdam 1983 (also Proceedings 15th STOC, 189-192).

10. A.K. Lenstra, Factoring multivariate integral polynomials, Report IW 229/83, Mathematisch Centrum, Amsterdam 1983 (also Proceedings 10th ICALP, LNCS 154, 458-465).

11. A.K. Lenstra, Factoring multivariate integral polynomials, II, Report IW 230/83, Mathematisch Centrum, Amsterdam 1983.

12. M. Mignotte, An inequality about factors of polynomials, *Math. Comp.* **28** (1974), 1153-1157.

13. J. Stoer, Einführung in die numerische Mathematik I, Springer, Berlin 1972.

14. P.S. Wang, Factoring multivariate polynomials over algebraic number fields, *Math. Comp.* **30** (1976), 324-336.

15. P.J. Weinberger, L.P. Rothschild, Factoring polynomials over algebraic number fields, *ACM Transactions on mathematical software* **2** (1976), 335-350.

16. D.Y.Y. Yun, The Hensel lemma in algebraic manipulation, MIT, Cambridge 1974; reprint: Garland Publ. Co., New York 1980.