ON THE COMPLEXITY OF FACTORING POLYNOMIALS WITH INTEGER COEFFICIENTS

by

Erich Kaltofen

A Thesis Submitted to the Graduate

Faculty of Rensselear Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Computer Science

Approved by the
Examining Committee

____
Bobby F. Caviness, Thesis Adviser

____
Charles M. Fiduccia, Member

____
Mukkai S. Krishnamoorthy, Member

____
Robert McNaughton, Member

____
B. David Saunders, Member

Rensselaer Polytechnic Institute
Troy, New York

December 1982

CONTENTS

To my parents, who waited so long,
and to my wife, for all her support.

ACKNOWLEDGEMENTS

I shall not begin without naming some of the scientists who have made contributions to this thesis. I would like to thank all members of the examining committee for their assistance. Prof. Paul Wang gave valuable suggestions for section 1.4. Dr. David Musser developed parts of section 2.2 besides listening so patiently during my visits at General Electric. The remark following theorem 2.2 is due to Dr. Andrew Odlyzko. Section 2.3 contains partially the findings of Prof. Gary Miller and he and Prof. Hendrik Lenstra deserve some credit for theorem 2.6. I owe the topic of chapter 3 to Prof. George Collins whose enthusiasm for my work was one of its greatest rewards. Prof. Hans Zassenhaus provided an important manuscript in connection with algorithm 3.1. Some simplifying remarks in chapter 3.5 are due to Prof. Hendrik Lenstra.

I also wish to acknowledge the support by the National Science Foundation and by the Department of Energy. This thesis could not have been printed without the generosity of the General Electric Research Laboratory in Schenectady, New York, or without the support of the Department for Mathematical Sciences at Kent State University.

I must not begin before expressing my gratitude to my advisers Prof. Bobby Caviness and Prof. B. David Saunders without whose friendship I could not have done it.

ABSTRACT

The complexity of the Berlekamp-Hensel algorithm for factoring polynomials in one or more variables with integer coefficients can become exponential in the individual variable degrees of the input polynomial due to the fact that, after factoring the projected polynomial and lifting its factors to sufficiently large coefficients, one may need to combine exponentially many lifted factors to obtain the true integer factors. In the univariate case, where the projection is taking the coefficients modulo a prime number, we can find worst case polynomials by prescribing that their Galois groups consist only of permutations with short cycles. Using the Chebotarev density theorem we then are able to construct succinct certificates for our hard-to-factor polynomials. In the multivariate case the projection is evaluation of selected variables at integral points. By computing the minimal polynomial of the approximation for a root we are able to replace the factor combination process by solving a system of linear equations. The growth analysis for the size of the rational numbers involved shows that, provided the number of variables is fixed, our algorithm reduces the problem in polynomial time in the total degree and coefficient length to the problem of factoring univariate polynomials, which has recently been solved in polynomial time as well. Therefore our algorithm can factor multivariate polynomials with a fixed number of variables in polynomial time in the total degrees and coefficient lengths, except for splitting a possible constant factor into its prime divisors. The evaluation process also leads us to the study of the Hilbert irreducibility theorem, an effective version of which provides us with an alternate polynomial time reduction from multivariate to bivariate polynomial factorization and irreducibility testing.

1. Overview

1.1. Review of Earlier Developments and Our Results

The problem of factoring polynomials with integer coefficients has a long and distinguished history. D. Knuth traces the first attempts back to Isaac Newton's Arithmetica Universalis (1707) and to the astronomer Friedrich T. v. Schubert who in 1793 presented a finite-step algorithm to compute the factors of a univariate polynomial with integer coefficients (cf. [Knuth 81, Sec. 4.6.2]). A notable criterion for determining irreducibility was given by F. G. Eisenstein in 1850 [Eisenstein 1850, p. 166]. L. Kronecker rediscovered Schubert's method in 1882 and also gave algorithms for factoring polynomials with two or more variables or with coefficients in algebraic extensions [Kronecker 1882, Sec. 4, pp. 10-13]. Exactly one hundred years have passed since then, and though early computer programs relied on Kronecker's work [Johnson 66], modern polynomial factorization algorithms and their analysis depend on major advances in mathematical research during this period of time.

When the long-known finite-step algorithms were first put on computers they turned out to be highly inefficient. The fact that almost any uni- or multivariate polynomial of degree up to 100 and with coefficients of a moderate size (up to 100 bits) can be factored by modern algorithms in a few minutes of computer time indicates how successfully this problem has been attacked during the past fifteen years. However, until very recently, some rare polynomials still took an exponential number of steps in the degree of the polynomial to be factored by these modern algorithms. It is the main theme of this thesis to investigate these hard-to-factor polynomials and, in the multivariate case, to give a new algorithm which does not suffer from this exponential worst-case complexity.

In 1967 E. Berlekamp devised an ingenious algorithm which factors univariate polynomials over Z_p, p a prime number, whose running time is of order $O(n^3+prn^2)$ where n is the degree of the polynomial and r the number of actual factors (cf. [Knuth 81, Sec. 4.6.2]). The incredible speed of this algorithm suggested factoring integer polynomials by first factoring them modulo certain small primes and then reconstructing the integer factors by some mean such as Chinese remaindering [Knuth 69, Sec.

4.6.2]. H. Zassenhaus discussed in his landmark 1969 paper [Zassenhaus 69] how to apply the "Hensel lemma" to lift in k iterations a factorization modulo p to a factorization modulo p^{2^k}, provided that the integral polynomial is squarefree and remains squarefree modulo p. Readers familiar with basic field theory will know that if a polynomial over a field of characteristic 0 has repeated roots, then the greatest common divisor (GCD) of the polynomial and its derivative is nontrivial. Hence casting out multiple factors is essentially a polynomial GCD process, but we will come back to this problem in a later section. Squarefreeness is preserved modulo all but a reasonable small number of primes. Given a bound for the size of the coefficients of any possible polynomial factor, one then lifts the modular factorization to a factorization modulo p^{2^k} such that $p^{2^k}/2$ supersedes this coefficient bound. At this point either factors with balanced residues modulo p^{2^k} are already the integral factors or one needs to multiply some factors together to obtain a true factor over the integers. The slight complication arising from a leading coefficient not equal to unity can be resolved in various easy ways.

D. Musser [Musser 71, Musser 76] and, using his ideas, P. Wang in collaboration with L. Rothschild [Wang and Rothschild 75], generalized the Hensel lemma to obtain factorization algorithms for multivariate integral polynomials. Subsequently, P. Wang has incorporated various improvements to these multivariate factorization algorithms [Wang 77, Wang 78, Wang 79b]. In 1973 J. Moses and D. Yun found the Hensel construction suitable for multivariate GCD computations (now called the EZGCD algorithm) [Moses and Yun 73], and D. Yun has used this algorithm for the squarefree decomposition process of multivariate polynomials [Yun 76b]. In 1979 G. Collins published a thorough analysis of the average time behavior for the univariate Berlekamp-Hensel algorithm [Collins 79], while in the same year improved algorithms for squarefree factorization [Wang and Trager 79] and Chinese remaindering on sparse multivariate polynomials appeared [Zippel 79].

To completely factor a univariate polynomial over the integers means, of course, to also factor the common divisor of all its coefficients. This thesis does not discuss the topic of factorization of integers and we will not consider this problem as a part of polynomial factorization. However, some comparisons are in order. Factoring large random integers is much harder than factoring integral polynomials. This was partially confirmed by a polynomial-time reduction from polynomial to integer factorization, which is, however, subject to an old number theoretic conjecture [Adleman and Odlyzko 81]. The problem of finding polynomially long irreducibility proofs ("succinct certificates") was first solved for prime numbers in 1975 [Pratt 75] and has recently also been achieved for densely encoded integral polynomials [Cantor 81]. A polynomial-time irreducibility test for prime numbers depending on the validity of the generalized Riemann hypothesis (GRH) was discovered in 1976 (cf. [Knuth 81, Sec. 4.5.4]). P. Weinberger obtained the corresponding result for densely encoded integer polynomials [Weinberger 81, Knuth 81, p. 632, Exercise 38]. In 1971 E. Berlekamp pointed out that the modular projection and lifting algorithm may take an exponential number of trial factor combinations [Berlekamp 70]. It was not until 1982 when A. Lenstra, H. Lenstra and L. Lovász overcame this problem by reconstructing the integral factors from the lifted modular factors using a new integer lattice algorithm. Their algorithm takes at worst polynomially many steps in the degree and coefficient size of the input polynomial [Lenstra et al. 82].

Little work has been done on the theoretical analysis of the multivariate versions of the Berlekamp-Hensel algorithm. Similar to the univariate case, the steps involved may require an exponential number of trial factor combinations, though this problem may be probabilistically controllable by virtue of the Hilbert Irreducibility Theorem. G. Viry has also shown how to replace the trial divisions of multivariate polynomials by a simple degree test [Viry 80]. In chapter 3 we will prove that it is only polynomially harder to factor densely encoded multivariate integer polynomials with a fixed number of variables than integer polynomials with just one variable. Together with the recent work on univariate factorization mentioned above, this implies that multivariate integer polynomials with a fixed number of variables can be factored in time polynomial in their total degree and coefficient size.

We will discover a characterization for those univariate polynomials that are hard-to-factor by the Berlekamp-Hensel algorithm as a property of their Galois groups. This property can then be most precisely expressed by the Chebotarev Density Theorem and its recently discovered effective versions. We then can obtain an alternate construction of succinct certificates for our hard to factor polynomials. Our algorithm can be also used to compute the Galois group of univariate integer polynomials though we require the computation of a resolvent for their splitting fields, i.e. the minimal polynomial for one of its primitive elements. Whether our technique is then more efficient than a simple substitution method is yet to be investigated. We will present all of our results on univariate polynomials in chapter 2.

The next section of this chapter establishes our notation. Section 1.3 contains a detailed description of the univariate Berlekamp-Hensel algorithm which we will need as a reference in chapter 2. The last section of this chapter gives an overview of conventional methods for factoring multivariate integer polynomials, in particular of the Kronecker and multivariate Hensel algorithms which we will refer to in chapter 3. The detailed organization of chapters 2 and 3 can be found in their corresponding introductions. We conclude with a list of open problems in chapter 4.

1.2. Notation

By Z we denote the integers, by Q the rationals and by C the complex numbers. Z_p denotes the prime residues modulo p. If D is an integral domain, $D[x_1, \ldots, x_v]$ denotes the polynomials in $x_1, \ldots, x_v$ over D; $\deg_{x_1}(f)$ denotes the highest degree of x_1 in $f \in D[x_1, \ldots, x_v]$, $\deg_{x_1,x_2}(f)$ the highest total degree of monomials x_1 and x_2 in f, and $\deg(f) = \deg_{x_1, \ldots, x_v}(f)$ the total degree of f. The coefficient of the highest power of x_v in f is referred to as the leading coefficient of f in x_v and will be denoted by $\mathrm{ldcf}_{x_v}(f)$. We call f monic in x_v if $\mathrm{ldcf}_{x_v}(f)$ is the unity of D. As is well-known, $D[x_1, \ldots, x_v]$ is a unique factorization domain (UFD) provided that D is a UFD. In this case the content of $f \in D[x_1, \ldots, x_v]$ in x_v, $\mathrm{cont}_{x_v}(f)$, is the greatest common divisor (GCD) of all coefficients of $f(x_v)$ as elements in $D[x_1, \ldots, x_{v-1}]$. The primitive part of f in x_v is defined as

$$\mathrm{pp}_{x_v}(f) = f / \mathrm{cont}_{x_v}(f)$$

and we call f primitive in x_v if $f = \mathrm{pp}_{x_v}(f)$. We also note that the total degree of any monomial in a factor of f is less than or equal to the total degree of that monomial in f. The infinity norm of $f \in Q[x_1, \ldots, x_v]$, the maximum of the absolute values of the rational coefficients of f, will be denoted by $| f |$. The sum of the absolute values of the coefficients will be denoted by $| f |_1$, the square norm by $| f |_2$.

By $\left[\begin{matrix} n \\ m \end{matrix} \right]$ we denote the binomial coefficient $\dfrac{n!}{m!\,(n-m)!}$.

1.3. The Berlekamp-Hensel Algorithm

Given a polynomial $\bar{h}(x) \in Z[x]$ we seek to compute its content and all its primitive irreducible polynomial factors $g_{ij}(x) \in Z[x]$, that is

$$\bar{h}(x) = \mathrm{cont}(\bar{h}) \prod_{i=1}^{r} \left[\prod_{j=1}^{s_i} g_{ij}(x) \right]^i$$

with all g_{ij} irreducible and pairwise distinct. The complete algorithm consists of three separate steps, namely

Algorithm 1.1: [Factorization of $\bar{h}(x) \in Z[x]$:]

(C) [Content computation:] The integer GCD of all coefficients of $\bar{h}$ constitutes the cont($\bar{h}$), $h \leftarrow \bar{h}/\mathrm{cont}(\bar{h})$. [$h$ is now a primitive polynomial.]

(S) [Squarefree decomposition of h:] Compute squarefree polynomials $f_i(x) \in Z[x]$, $1 \le i \le r$, GCD(f_j, f_k) = 1 for $1 \le j \ne k \le r$ such that $h(x) = \prod_{i=1}^{r} (f_i(x))^i$.

(F) [Factor the squarefree f_i:] FOR $i \leftarrow 1, \ldots, r$ DO

Compute irreducible polynomials $g_{ij}(x) \in Z[x]$, $1 \le j \le s_i$, such that $f_i(x) = \prod_{j=1}^{s_i} g_{ij}(x)$. □

Step (C) is a repeated integer GCD computation and shall not be discussed further.

The computational aspects of step (S) were first investigated by E. Horowitz following an idea of R. Tobey in 1969 (cf. [Horowitz 71]) whose algorithms were later improved by D. Musser [Musser 76], D. Yun [Yun 76b] and P. Wang and B. Trager [Wang and Trager 79]. We shall briefly present D. Yun's algorithm:

Algorithm 1.2: [Squarefree decomposition of a primitive polynomial h:]

(S1) $g(x) \leftarrow$ GCD($h(x)$, d$h(x)$/dx) where d$h(x)$/dx = $h'(x)$ is the derivative of h w.r.t. x.

$$c_1(x) \leftarrow \frac{h(x)}{g(x);} \quad d_1(x) \leftarrow \frac{1}{g(x)} \frac{\mathrm{d}h(x)}{\mathrm{d}x} - \frac{\mathrm{d}c_1(x)}{\mathrm{d}x}.$$

[Assume that $h = \prod_{i=1}^{r} f_i^i$ with the f_i squarefree and pairwise relatively prime. Then

$$g = \prod_{i=2}^{r} f_i^{i-1}, \quad c_1 = \prod_{i=1}^{r} f_i, \quad \frac{h'}{g} = \sum_{i=1}^{r} \left[i \ f_i' \prod_{j=1, j \neq i}^{r} f_j \right]$$

which is relatively prime to g since GCD$(f_i, f_i') = 1$ (The f_i are squarefree!). Thus

$$d_1 = \sum_{i=2}^{r} \left[(i-1) f_i' \prod_{j=1, j \neq i}^{r} f_j \right].]$$

(S2) FOR $k \leftarrow 1,2, \cdots$ UNTIL $c_k = 1$ DO

[At this point

$$c_k = \prod_{i=k}^{r} f_i, \quad d_k = \sum_{i=k+1}^{r} \left[(i-k) f_i' \prod_{j=k, j \neq i}^{r} f_j \right].]$$

$$f_k(x) \leftarrow \mathrm{GCD}(c_k(x), d_k(x)); \quad c_{k+1}(x) \leftarrow \frac{c_k(x)}{f_k(x)}; \quad d_{k+1}(x) \leftarrow \frac{d_k(x)}{f_k(x)} - \frac{\mathrm{d}c_{k+1}(x)}{\mathrm{d}x}. \ \square$$

The reader should be able to derive the correctness of this algorithm from the embedded comments. It is important that the cofactor of h' in the GCD computation of step (S1) and that of d_k in step (S2) are relatively prime to the computed GCDs. This enables one to use, besides the modular GCD algorithm, the EZGCD algorithm [Moses and Yun 73] whose general version needs the above algorithm if both cofactors have a common divisor with the GCD. The relation between polynomial GCDs and squarefree decompositions is even more explicit (cf. [Yun 77]).

Step (F) is the actual heart of the algorithm. As outlined in the introduction, various substeps are needed for the Berlekamp-Hensel algorithm:

Algorithm 1.3: [Factorization of a primitive, squarefree polynomial f:]

(F1) [Choice of a modulus:] Find a prime number p which neither divides ldcf$(f(x))$ nor the resultant of $f(x)$ and $\mathrm{d}f(x)/\mathrm{d}x$. The latter is equivalent to the condition that $f(x)$ modulo p is squarefree.

By trying various primes in connection with the distinct factorization procedure we may also attempt to minimize the number of modular factors in the next step.

(F2) [Modular factorization:] Factor $f(x)$ modulo p completely by the Berlekamp algorithm [Knuth 81, Chap. 4.6.2], namely compute irreducible polynomials $u_1(x), \ldots, u_r(x) \in Z_p[x]$ such that $\text{ldcf}(u_1) \equiv \text{ldcf}(f)$ (modulo p), $u_2, \ldots, u_r$ are monic and $u_1(x) \cdots u_r(x) \equiv f(x)$ (modulo p).

(F3) [Factor coefficient bound:] Compute an integer $B(f)$ such that all coefficients of any possible factor of $f(x)$ in $Z[x]$ are absolutely bounded by $B(f)$ (see lemma 3.5a).

(F4) [Lift modular factors:] $q \leftarrow p$;

```
FOR $k \leftarrow 1,2, \cdots$ UNTIL $q \geq 2\,B(f)$ DO
    $q \leftarrow q^2$; [At this point $q = p^{2^k}$.]
    Compute polynomials $u_i^{(k)}(x) \in Z_q[x]$ such that $u_1^{(k)} \cdots u_r^{(k)} \equiv f(x)$ (modulo $q$),
    $\text{ldcf}(u_1^{(k)}) \equiv \text{ldcf}(f)$ (modulo $q$) and $u_i^{(k)} \equiv u_i$ (modulo $p$) where the coefficients of $u_i^{(k)}$ are
    interpreted as $p$-adic approximations.
```

(F5) [Form trial factor combinations:]

```
$h(x) \leftarrow f(x)$; $C \leftarrow \{2, \ldots, r\}$; $s \leftarrow 0$; $j \leftarrow 1$;
REPEAT $t \leftarrow s$;
    FOR $m \leftarrow j, \ldots,$ cardinality of $C$ DO
        FORALL subsets $\{i_1, \ldots, i_m\}$ of $C$ DO
            Test whether $g(x) = \text{pp}\left[\text{ldcf}(h)\, u_{i_1}^{(k)} \cdots u_{i_m}^{(k)} \;(\text{modulo } p^{2^k})\right]$ divides $h$, where
            $k$ is the number of iterations in (F4) and the modulus is balanced before taking
            the primitive part over the integers. If so then set $s \leftarrow s+1$; $g_s(x) \leftarrow g(x)$;
            $h(x) \leftarrow h(x)/g(x)$; $j \leftarrow m$; $C \leftarrow C$ minus $\{i_1, \ldots, i_m\}$; and exit both FOR
```

```
                loops.
            END FORALL
        END FOR
    UNTIL $t = s$ [No more factors discovered in the FOR loops]

    $s \leftarrow s+1$; $g_s(x) \leftarrow h(x)$
    [All factors are computed as $f(x) = g_1(x) \cdots g_s(x)$.] □
```

We must scrutinize various steps further. By the choice of p in step (F1) $\overline{f}(x) = f(x)$ modulo p is of the same degree as $f(x)$ and the inverse of ldcf($\overline{f}$) exists in Z_p. We factor the monic polynomial ldcf$(\overline{f})^{-1}$ $\overline{f}(x)$ first into distinct degree factors and then into irreducibles in step (F2). To satisfy the condition on the ldcf(u_1) we multiply the monic u_1 by ldcf($\overline{f}$) in Z_p. Step (F4) utilizes the "Hensel-lemma" and various lifting techniques have been investigated [Zassenhaus 69, Musser 71, Wang 79a]. The following algorithm is due to P. Wang:

Algorithm 1.4: [Hensel Lifting Algorithm:]

[Given polynomials $f(x) \in Z[x]$, q relatively prime to ldcf(f), $u_1^*(x), \ldots, u_r^*(x) \in Z_q[x]$ such that ldcf(u_1) $\equiv$ ldcf(f) (modulo q), $u_2^*, \ldots, u_r^*$ monic and

$$u_1^*(x) \cdots u_r^*(x) \equiv f(x) \text{ (modulo } q). \qquad \text{(A)}$$

Furthermore given polynomials $v_1^*(x), \ldots, v_r^*(x) \in Z_q[x]$ with deg(v_i^*) < deg(u_i^*) for 1 ≤ i ≤ r, and if we set $\hat{u}_i^* = \prod_{j=1, j \neq i}^{r} u_i^*$ then

$$v_1^*(x)\hat{u}_1^*(x) + \cdots + v_r^*(x)\hat{u}_r^*(x) \equiv 1 \text{ (modulo } q).$$

The goal is to produce polynomials $u_1^{**}(x), \ldots, u_r^{**}(x), v_1^{**}(x), \ldots, v_r^{**}(x) \in Z_{q^2}[x]$ which satisfy the same conditions as the single-starred polynomials if we replace the modulus q by q^2.]

(H1) Replace ldcf(u_1^*) by ldcf(f) modulo q^2;

[Lift u_i^* by computing $\overline{u}_i^* \in Z_q[x]$ such that $u_i^{**} = u_i^* + q\overline{u}_i^*$ with deg($\overline{u}_i^*$) < deg(u_i^*) for $i \geq 1$.]

$$t(x) \leftarrow \left[f(x) - \prod_{i=1}^{r} u_i^*(x) \right] \text{ modulo } q^2;$$

[The above replacement guarantees deg(t) < deg(f). Also all coefficients of t are divisible by q because of (A).]

$t(x) \leftarrow t(x)/q$; [Integer division, hence $t(x) \in Z_q[x]$. We need to determine $\overline{u_i}^*$ with

$$\overline{u_1}^*(x)\hat{u}_1^*(x) + \cdot \cdot \cdot + \overline{u_r}^*(x)\hat{u}_r^*(x) = t(x). \qquad \text{(B)]}$$

```
FOR $i \leftarrow 1, . . . , r$ DO

    $\overline{u_i}^*(x) \leftarrow$ remainder($t(x)v_i^*(x)$, $u_i^*(x)$) in $Z_q[x]$; $u_i^{**}(x) \leftarrow u_i^*(x) + q\overline{u_i}^*(x)$.
```

[Obviously the polynomials tv_i^* solve (B) but do not satisfy the degree constraint for the $\overline{u_i}^*$. Hence the $\overline{u_i}^*$ solve (B) modulo $\prod_{i=1}^{r} u_i^*$ but since all degrees are less than deg(f) there must be equality.]

(H2) [Lift v_i^* by computing $\overline{v_i}^* \in Z_q[x]$ such that $v_i^{**} = v_i^* + q\overline{v_i}^*$ and deg($\overline{v_i}^*$) < deg(u_i^*).]

$$b(x) \leftarrow \left[\left[1 - \sum_{i=1}^{r} v_i^*(x)\hat{u}_i^*(x) \right] \text{ modulo } q^2 \right] / q;$$

[Again the division is integral and $b(x) \in Z_q[x]$ with deg(b) < deg(f).]

```
FOR $i \leftarrow 1, . . . , r$ DO

    $\overline{v_i}^*(x) \leftarrow$ remainder($b(x)v_i^*(x)$, $u_i^*(x)$) in $Z_q[x]$;
    $v_i^{**}(x) \leftarrow v_i^*(x) + q\overline{v_i}^*(x)$. □
```

In order to use the above algorithm within the loop of step (F4) we also need to initialize the $v_i(x)$ in Z_p with

$$1/\prod_{i=1}^{r} u_i(x) = \sum_{i=1}^{r} \frac{v_i(x)}{u_i(x)} \quad \text{and} \quad \deg(v_i) < \deg(u_i).$$

To do this one can use the extended Euclidian algorithm $r-1$ times or use fast partial fraction decomposition algorithms [Kung and Tong 77, Abdali et al. 77].

Step (H2) is not necessary if one only considers the first solution v_i and corrects u_i^* from modulus q to modulus pq by calculating $\hat{u}_i^*$ in $Z_p[x]$. This method is referred to as "linear lifting" whereas our algorithm has quadratic p-adic convergence. We also lift all factors in parallel while earlier versions proceeded with one factor and its cofactor at a time. It is not clear which technique is preferable (cf. [Yun 76a, Zassenhaus 78]), though the parallel quadratic approach seems superior [Wang 79a]. In order to prevent p^{2^k} from overshooting $B(f)$ by too much one may calculate the last correction polynomials $\hat{u}_i^*$ with a smaller modulus than q.

As we will show in section 2.2, in the worst case step (F5) is the dominant step in our algorithm. Therefore one is advised to test whether the constant coefficient of $g(x)$ divides that of $f(x)$.

D. Musser has carefully analysed a variation of steps (F1) - (F5), the result of which is the following [Musser 71]: Let $f = g_1 \cdots g_s$ in $Z[x]$, $\deg(g_1) \le \deg(g_2) \le \cdots \le \deg(g_s)$, and let

$$\mu = \begin{cases} \max\limits_{i=2,\ldots,s} \{\deg(g_{i-1}), \lfloor \deg(g_i)/2 \rfloor\} & \text{if } s>1 \\ \lfloor \deg(f)/2 \rfloor & \text{if } f \text{ is irreducible}. \end{cases}$$

If f factors into r polynomials modulo p then

$$\min(2^r, r^\mu)\, \mu n^2\, (n+\log(B(f)))^2$$

dominates the complexity of the factorization problem. This bound depends intrinsically on r which is one reason why one should attempt to minimize this number in step (F2). If one does not, the algorithm still performs quite well − on the average. An n-th degree polynomial in $Z_p[x]$ has an average of $\log(n)$ factors as p tends to infinity and 2^r averages $n+1$ where r is the number of modular factors (cf. [Knuth 81, Sec.4.6.2., Exercise 5]. However, almost all integer polynomials are irreducible (cf. [Knuth 81, Sec.4.6.2, Exercise 27]), and one may not expect almost all inputs to our algorithm to behave that way since a user probably tries to factor polynomials which are expected to be composite. In this matter G. Collins has shown, subject to two conjectures, that if we restrict our set to those polynomials which factor over the integers into factors of degree $d_1, d_2, \ldots, d_s$ for a given additive decomposition of $n = d_1 + \ldots + d_s$, the average number of trial combinations will be below n^2. This result only holds if one processes combinations of m factors at a time as we did in step (F5) ("cardinality procedure"), because if one chooses to test combinations of a possible total degree ("degree procedure") the average behavior may be exponential in n [Collins 79].

The worst case complexity of the Berlekamp-Hensel algorithm is unfortunately exponential in n, the degree of f. This is because, as we will prove in detail in section 2.2, there exist irreducible integer polynomials of arbitrarily large degree which factor over every prime into linear or quadratic factors. This means that we must test at least $2^{n/2-1}-1$ trial factor combinations to show that no integral factor occurs. We will also show that the number of binary digits of the coefficients of those polynomials is about their degree (cf. theorem 2.3) which makes the worst case of the Berlekamp-Hensel algorithm truly exponential in its input size. Here the following remark is appropriate. We always assume that our algorithm operates on densely encoded polynomials. If we allow sparse encoding schemes, various primitive operations on the input polynomials such as GCD computations are NP-hard (cf. [Plaisted 78]) and the factorization problem actually requires exponential space. In order to substantiate the last claim we consider the polynomial x^n-1 whose sparse encoding requires $O(\log n)$ bits. However, following earlier developments, R. Vaughan [Vaughan 75] has shown that for infinitely many n the cyclotomic polynomials Ψ_n, which constitute irreducible factors of x^n-1, have coefficients absolutely larger than $\exp(n^{\log 2/log\ \log n})$.

One question about our algorithm remains to be answered. That is how the choice of various primes in step (F1) can influence later steps, especially step (F5). It is clear that if a polynomial f factors modulo p_1 into all quadratic and modulo p_2 into all cubic factors, then the degrees of integral factors must be multiples of six. Indeed if the degree sets of factorizations modulo various primes are completely incompatible we know the input polynomial to be irreducible without the need of steps (F2) - (F5). For this situation D. Musser has developed an interesting model which, given a random irreducible polynomial $f(x) \in Z[x]$ of degree n, shows how to derive the average number $\mu(n)$ of factorizations modulo distinct primes $p_1, \ldots, p_{\mu(n)}$ needed to prove f irreducible [Musser 78]. His approach is based on the fact that the degrees $d_1, \ldots, d_r$ of a factorization $f \equiv g_1 \cdots g_r$ modulo p, $d_i = \deg(g_i)$ for $1 \le i \le r$ and p a random prime correspond to the cycle lengths of a random permutation

$$(1, \ldots, d_1)(d_1+1, \ldots, d_1+d_2) \cdots (d_1+ \cdots +d_{r-1}+1, \ldots, d_1+ \cdots +d_r)$$

of n elements. We will show in section 2.3 that this property remains valid for any given polynomial provided that its Galois group is the full symmetric group. This result is in accordance with D. Musser's observation since almost all polynomials have the symmetric group as Galois group [Gallagher 72].

1.4. Factorization of Multivariate Integer Polynomials

We shall begin this chapter with Kronecker's algorithm which, for certain coefficient domains (such as C), is still the only one known.

Algorithm 1.5: [Kronecker Factorization of $f(x_1, \ldots, x_v) \in D[x_1, \ldots, x_v]$ with D being a unique factorization domain.]

(K1) [Compute degree bound:] Obtain an integer d larger than the degree of f in any single variable.

(K2) [Reduction:] $\overline{f}(y) \leftarrow S_d(f) = f\left[y, y^d, \ldots, y^{d^{v-1}}\right]$.

(K3) [Factorization:] Factor $\overline{f}(y)$ into irreducibles, i.e., $\overline{f}(y) = \overline{g}_1(y) \cdots \overline{g}_s(y)$, $\overline{g}_i(y) \in$ D[y] for $1 \leq i \leq s$.

(K4) [Inverse reduction and trial division:] For all products $\overline{g}_{i_1}(y) \cdots \overline{g}_{i_m}(y)$ (similar to step (F5) in algorithm 1.3) perform the following test:

$$g_{i_1, \ldots, i_m}(x_1, \ldots, x_v) \leftarrow S_d^{-1}(\overline{g}_{i_1} \cdots \overline{g}_{i_m})$$

where S_d^{-1} is the inverse of S_d which is additive and

$$S_d^{-1}\left[\lambda y^{b_1 + db_2 + \cdots + d^{v-1}b_v}\right] = \lambda x_1^{b_1} \cdots x_v^{b_v}$$

with $0 \leq b_i < d$ for $1 \leq i \leq v$, $\lambda \in Z$.

Test whether $g_{i_1, \ldots, i_m}$ divides f and if so remove this irreducible factor from f and proceed with its co-factor. □

The correctness of this algorithm follows easily from the fact that no variable in any factor of f can occur with degree d or higher. The running time of the algorithm depends on of how fast the univariate polynomial $\overline{f}(y)$ can be factored, the degree of which can be substantially large. It should be clear that step (K4) can take time exponential in the degree of f, e.g., if $D = C$ and f is irreducible. Unfortunately this exponential worst case complexity remains true for $D = Z$ as we will show in section

3.2. In this case, the Hensel lemma has produced a much more efficient approach. In the following we will take a closer look at this algorithm.

The overall structure of the multivariate factorization algorithm is remarkably close to that of algorithm 1.1. First we choose a main variable x, i.e., the input polynomial $\overline{h} \in Z[y_1, \ldots, y_v, x]$. The content computation of step (C) now becomes a GCD computation in $Z[y_1, \ldots, y_v]$. The squarefree decomposition performed in step (S) can also be achieved by algorithm 1.2 if we replace the derivatives d/dx by partial derivatives $\partial/\partial x$ and the GCDs by multivariate polynomial GCDs. However, in this case P. Wang's and B. Trager's algorithm becomes more efficient [Wang and Trager 79].

The idea of their algorithm is to find an evaluation point $(b_1, \ldots, b_v)$ such that if

$$h(y_1, \ldots, y_v, x) = \prod_{i=1}^{r} f_i(y_1, \ldots, y_v, x)^i$$

is the squarefree decomposition of h, and

$$h(b_1, \ldots, b_v, x) = \overline{h}(x) = \prod_{i=1}^{\overline{r}} \overline{f}_i(x)^i$$

is that of $\overline{h}$, then $r = \overline{r}$ and $f_i(b_1, \ldots, b_v, x) = \overline{f}_i(x)$, $1 \le i \le r$. Under these conditions

$$f_r \text{ divides } g = \frac{1}{(r-1)!}\left[\frac{\partial}{\partial x}\right]^{r-1}(h),$$
$$\overline{f}_r \text{ divides } \overline{g} = \frac{1}{(r-1)!}\left[\frac{d}{dx}\right]^{r-1}(\overline{h})$$

and we can lift the equation

$$g(y_1, \ldots, y_v, x) \equiv \overline{f}_r(x)\left[\frac{\overline{g}(x)}{\overline{f}_r(x)}\right] \text{ modulo } (y_1 - b_1, \ldots, y_v - b_v)$$

to determine f_r from the univariate square decomposition of $\overline{h}$, provided $\overline{g}/\overline{f}_r \ne 1$. Evaluation points for which the above conditions do not hold are, as in the modular multivariate GCD algorithm, very rare. (Cf. lemma 3.1.)

Step (F), the complete factorization of a squarefree polynomial $f(y_1, \ldots, y_v, \text{x})$, is again a major challenge. As in the above squarefree decomposition algorithm we evaluate the minor variables y_i at integers b_i, $1 \le i \le v$, then factor the resulting univariate polynomial $f(b_1, \ldots, b_v, x)$ and finally rebuild multivariate factor candidates by a Hensel lifting algorithm with respect to the ideal P generated by $\{(y_1 - b_1), \ldots, (y_v - b_v)\}$. Instead of presenting a complete algorithm we shall work out a simple example and refer the reader to the papers by P. Wang [Wang and Rothschild 75, Wang 77, 78, 79b] and D. Musser [Musser 76] for the details.

Example 1.1: Factor

$$\begin{aligned} f(y,z,x) = x^3 &+ ((y+2)z + 2y + 1)x^2 \\ &+ ((y+2)z^2 + (y^2+2y+1)z + 2y^2 + y)x \\ &+ (y+1)z^3 + (y+1)z^2 + (y^3+y^2)z + y^3 + y^2. \end{aligned}$$

The polynomial is monic and squarefree.

Step F1: Choose an evaluation point which preserves degree and squarefreeness but contains as many zero components as possible.

$y=0$, $z=0$: $f(0,0,x) = x^3 + x^2$ is not squarefree

$y=1$, $z=0$: $f(1,0,x) = x^3 + 3x^2 + 3x + 2$ is squarefree.

Translate variables for nonzero components

$$\begin{aligned} f(w+1,z,x) = {} & x^3+3x^2+3x+2+(2x^2+5x+5)w \\ & +(2x+4)w^2+w^3+((3x^2+4x+2)+(x^2+4x+5)w \\ & +(x+4)w^2+w^3)z+((3x+2)+(x+1)w)z^2+(2+w)z^3 \end{aligned}$$

By $f_{ij}(x)$ we denote the coefficient of $w^j z^i$.

Step F2: Factor $f_{00}(x) = g_{00}(x) h_{00}(x)$ in $Z[x]$. We get

$$x^3 + 3x^2 + 3x + 2 = (x+2)\,(x^2 + x + 1).$$

Step F3: Compute highest degrees of w and z in factors of

$$f(w+1,z,x) = g(w,z,x)\; h(w,z,x)\colon \deg_w(g,h) \le 3,\ \deg_z(g,h) \le 2.$$

Step F4: Lift g_{00} and h_{00} to highest degrees in w and z. We set

$$g(w,z,x) = g_{00}(x) + g_{01}(x)w + g_{02}(x)w^2$$
$$+ \cdots + (g_{10}(x) + g_{11}(x)w + \cdots)z + \cdots$$

and

$$h(w,z,x) = h_{00}(x) + h_{01}(x)w + h_{02}(x)w^2$$
$$+ \cdots + (h_{10}(x) + h_{11}(x)w + \cdots)z$$
$$+ (h_{20}(x) + h_{21}(x)w + \cdots)z^2 + \cdots$$

and compute g_{01}, h_{01}, g_{02}, h_{02}, . . . , g_{10}, h_{10}, g_{11}, h_{11}, . . . , g_{20}, h_{20}, $\cdots$ in that sequence. Since f is monic $\deg(g_{ij}) \le 1$ and $\deg(h_{ij}) \le 2$ for $i+j \ge 1$. Multiplying g times h with undetermined g_{ij}, h_{ij} we get $g_{00}h_{01} + h_{00}g_{01} = f_{01}$ whose unique solution is

$$(x+2)(x+2) + (x^2+x+1)\cdot 1 = 2x^2+5x+5,$$

by the extended Euclidean algorithm. In the next step we get

$$g_{00}h_{02} + h_{00}g_{02} = f_{02}-g_{01}h_{01}$$

which is solved by

$$(x+2)\cdot 1 + (x^2+x+1)\cdot 0 = 2x+4 - 1\cdot(x+2).$$

Finally

$$g_{00}h_{03} + h_{00}g_{03} = f_{03} - g_{01}h_{02} - g_{02}h_{01},$$

or

$$(x+2)\cdot 0 + (x^2+x+1)\cdot 0 = 1 - 1\cdot 1 - 0\cdot(x+2).$$

This gives factor candidates for

$$f(w+1,0,x) = ((x+2)+1\cdot w+0\cdot w^2)((x^2+x+1)+(x+2)w+w^2)$$

and a trial division shows them to be true factors.

We now lift z:

$$g_{00}h_{10} + h_{00}g_{10} = f_{10},$$

or

$$(x+2)x + (x^2+x+1)\cdot 2 = 3x^2+4x+2;$$

$$g_{00}h_{11} + h_{00}g_{11} = f_{11} - g_{01}h_{10} - g_{10}h_{01},$$

or

$$(x+2)\cdot 0 + (x^2+x+1)\cdot 1 = x^2+4x+5 - 1\cdot x - 2(x+2);$$

$$g_{00}h_{20} + h_{00}g_{20} = f_{20} - g_{10}h_{10},$$

or

$$(x+2)\cdot 1 + (x^2+x+1)\cdot 0 = 3x+2-2x.$$

All other equations have 0 as their right-hand sides.

The factor candidates are

$$f(w+1,z,x) = \Big[(x+2)+w+(2+w)z\Big]\ \Big[(x^2+x+1)+(x+2)w+w^2+xz+z^2\Big]$$

which are the actual factors. Setting $w = y-1$ we obtain

$$f(y,z,x) = \Big[x+yz+y+z+1\Big]\ \Big[x^2+(y+z)x+y^2+z^2\Big].$$

In factoring the above sample polynomial we followed the algorithm by P. Wang [Wang 78]. Our construction is actually a linear lifting technique. There is also the possibility of quadratic lifting [Musser 76], but in the multivariate case, the linear algorithm seems to be more efficient [Yun 76a]. If more than two univariate factors are present, one can again lift each one iteratively or lift them in parallel as we demonstrated for the univariate case.

Various complications have been identified with the multivariate Hensel algorithm.

a) The *leading coefficient* problem: In our example we dealt with a monic polynomial in which case the leading coefficients of all factors are known. If a polynomial leading coefficient is present, one can impose it on one factor as in the univariate case, but this leads most likely to dense factor candidates. P. Wang describes an algorithm to predetermine the actual leading coefficients of the factors, which avoids this intermediate expression growth [Wang 78, Sec.3]. However, in section 3.3 we will choose yet another method which is not very efficient, in practice, but which is guaranteed to work in polynomial time.

b) The *bad zero* problem: In our example, y had to be evaluated at 1 in order to preserve squarefreeness. The change of variables $y_i = w_i + b_i$ for $b_i \neq 0$ can make the polynomial $f(w_1+b_1, \ldots, w_v+b_v, x)$ dense. P. Wang suggests to compute the coefficients $f_{i_1 \cdots i_v}(x)$ of $w_1^{i_1} \cdots w_v^{i_v}$ by Taylor's formula without performing the substitution

$$f_{i_1 \cdots i_v}(x) = \frac{1}{i_1! \cdots i_v!} \left[\frac{\partial}{\partial y_1} \right]^{i_1} \cdots \left[\frac{\partial}{\partial y_v} \right]^{i_v} f(y_1, \ldots, y_v, x) \Bigg|_{y_i = b_i}$$

See also R. Zippel's work on preserving sparseness [Zippel 79].

c) The *extraneous factors* problem: This problem is the same as in the univariate case, namely that $f(b_1, \ldots, b_v, x)$ has more factors than $f(y_1, \ldots, y_v, x)$ (in which case we call $b_1, \ldots, b_v$ "unlucky"). One immediate consequence may be that the correction coefficients $g_{i_1 \cdots i_v}(x)$, $h_{i_1 \cdots i_v}(x)$ have non-integral coefficients. In order to avoid working with denominators one can choose to work with coefficients modulo a prime which preserves the squarefreeness of $f(b_1, \ldots, b_v, x)$, and as a first step lift the coefficients. A good factor coefficient bound is given in lemma 3.3. The algorithm 3.1 of chapter 3 provides a solution for this problem if the number of variables is fixed.

Various implementation issues can be found in [Moore and Norman 81]. A good set of polynomials for benchmarking an actual implementation of the factorization algorithm can be found in [Claybrook 76].

Little is known about the average computing time of the multivariate Hensel algorithm. The worst case complexity can be exponential in the degree of the main variable depending on what evaluation points one chooses. In section 3.2 we will show how to construct irreducible polynomials for which various evaluations yield all linear factors. However, unlike in the univariate case, it cannot happen that an irreducible polynomial factors for all possible evaluations. Actually, quite the opposite is true due to the following theorem.

Theorem 1.1 *(Hilbert Irreducibility Theorem):* Let $f(y_1, \ldots, y_v, x_1, \ldots, x_t)$ be irreducible in $Z[y_1, \ldots, y_v, x_1, \ldots, x_t]$. By $U(N)$ we denote the number of v-tuples $(b_1, \ldots, b_v) \in Z^v$ such that $|b_i| \leq N$ for $1 \leq i \leq v$ and $f(b_1, \ldots, b_v, x_1, \ldots, x_t)$ is reducible in $Z[x_1, \ldots, x_t]$. Then there exist constants α and C (depending on f) such that $U(N) \leq C(2N+1)^{v-\alpha}$ and $0 < \alpha < 1$. (Cf. [Knobloch 55]). □

Unfortunately, no polynomial upper bounds on the length of C seem to be known which would make the theorem useful for "realistic" evaluations. We will formulate the open problem 2 in section 4 in this connection. In practice lucky evaluations seem quite frequent.

A special problem is to test a polynomial $f(x_1, \ldots, x_v) \in Z[x_1, \ldots, x_v]$ for absolute irreducibility, that is, to test $f(x_1, \ldots, x_v)$ for irreducibility in $C[x_1, \ldots, x_v]$. The first criterion probably goes back to E. Noether [Noether 22] which also implies that if $f(x_1, \ldots, x_r)$ is absolutely irreducible, then $f(x_1, \ldots, x_r)$ remains irreducible modulo almost all prime numbers. Unfortunately, the first such prime number may be very large. A more efficient test for absolute irreducibility can be found in [Heintz and Sieveking 81].

2. Hard-to-Factor Polynomials and Galois Groups

2.1. Introduction and Review of the Galois Theory

In section 2.2 we will generalize a class of univariate polynomials with integral coefficients attributed to H.P.F. Swinnerton-Dyer by E.R. Berlekamp [Berlekamp 70, p.733]. We use Galois theoretical methods to prove their properties of interest. Some of these results were published earlier in [Kaltofen et al. 81].

These polynomials are of particular interest for the Berlekamp-Hensel factorization algorithm 1.3, which determines factors modulo p and lifts them to find the integral factors of a polynomial. Because the polynomials in the class we will define are irreducible over the integers but have a large number of factors modulo p for every prime p, the Berlekamp-Hensel algorithm behaves badly on them. In determining their irreducibility in step (F5) algorithm 1.3 needs a number of operations that is exponential in the degree and coefficient lengths of the polynomials.

As we will see in lemma 1.3 below, the degrees of modular factors of univariate polynomials are closely related to the cycles of the permutations in their Galois groups. While we use this relation in an elementary fashion in section 2.2, we will formulate it as the mathematically deep Chebotarev Density Theorem in section 2.3. This new insight will also provide us with an alternate construction of succinct certificates for normal polynomials and those whose Galois group is small (of polynomial cardinality in their degrees). Our construction actually provides us with a deterministic algorithm for constructing the Galois groups but its efficiency compared to standard techniques needs further investigation.

We will use some well-known properties of the cyclotomic polynomials in various places later and shall mention them now: Let r be an integer with $r \geq 2$ and let ζ_r be a primitive r-th root of unity. There always exist $\phi(r)$ distinct primitive r-th roots of unity in an extension field of Q or Z_q provided that q is a prime number not dividing r. By ϕ we denote Euler's totient function. These are the powers of ζ_r whose exponents are relatively prime to r. Then

$$\Psi_r(x) = \prod_{\substack{i=1\\ \mathrm{GCD}(i,r)=1}}^{r} (x-\zeta_r^i) = \prod_{d\,|\,r}(x^d-1)^{\mu(r/d)}$$

denotes the r-th cyclotomic polynomial which has all integer coefficients (or their residues modulo q if the ground field is Z_q). By $d \mid r$ we mean that d is a divisor of r and μ denotes the Möbius function: $\mu(n) = (-1)^m$ if n is squarefree and has m distinct prime divisors, $\mu(1) = 1$, and otherwise $\mu(n) = 0$. (Cf. [van der Waerden 53, p.112].)

If $\zeta_r = \exp(2\pi i/r)$ (i.e. the ground field is Q) then Ψ_r is irreducible over Z [van der Waerden 53, p.162].

Lemma 2.1: Let q be a prime number and let m and r be positive integers such that r is relatively prime to q. Then

$$\Psi_{rq^m}(x) \equiv \Psi_r(x)^{\phi(q^m)} \pmod{q}.$$

Proof: First we notice that for any integral polynomial f and any integer $i \geq 0$, $f(x^{q^i}) \equiv f(x)^{q^i}$ (mod q). Then by using the formulas for the cyclotomic polynomials and the Möbius function given above the stated congruence can be easily shown. □

By the Galois group of a polynomial we mean the automorphism group of its splitting field over the field of its coefficients. Then the Galois group of Ψ_r over Q is isomorphic to U_r under multiplication modulo r [van der Waerden 53, p.162]. U_r denotes the set of integral residues modulo r which are relatively prime to r.

The next two lemmas will help explain why the polynomials of section 2.2 split into so many factors modulo any prime number. First we show what happens to the Galois group when an integral polynomial is projected onto a polynomial over a residue field.

Lemma 2.2: Let f be a monic separable polynomial in $Z[x]$ and let $\overline{f} \in Z_q[x]$ be its natural projection modulo a prime number q. If f is separable (over Z_q) the Galois group of f over Z_q is a subgroup (as a permutation group on the suitably arranged roots) of the Galois group of f over Q. (Cf.

[van der Waerden 53, p.190].) □

Lemma 2.3: Let $f \in Z_q[x]$ with q prime. Assume that all elements of the Galois group of f (as permutations on the distinct roots of f) are written as products of disjoint cycles. Then f does not contain an irreducible factor with degree greater than the length of the longest cycle.

Proof: The statement follows immediately from the statement made about the generating element of the Galois group of f in [van der Waerden 53, p.191]. □

We now summarize some properties of Galois fields. Let GF(q^n) be the splitting field of $x^{q^n}-x$ as a polynomial in x with coefficients in Z_q, q being a prime number. Then GF(q^n) is a finite field with q^n elements of characteristic q whose multiplicative group is cyclic. All fields with q^n elements are isomorphic to GF(q^n) and hence it is called the Galois field with q^n elements. The degree $[GF(q^n):Z_q]$ is n and GF(q^n) has exactly one subfield with q^m elements, GF(q^m), provided that m divides n. The automorphism group on GF(q^n) is isomorphic to Z_n under addition and one of its generators maps each element α of GF(q^n) into α^q (the so called *Frobenius automorphism)* [van der Waerden 53, p.115].

Let f be an irreducible polynomial of degree n with coefficients in Z_q, q being prime, and let α be a root of f. Since $Z_q(\alpha)$ is isomorphic to $Z_q[x]/(f(x))$, the residues modulo f, $Z_q(\alpha)$ contains q^n elements and thus $\alpha \in GF(q^n)$. The remaining roots of f are $\alpha^q, \ldots, \alpha^{q^{n-1}}$ because of the structure of the Galois group mentioned above.

2.2. Univariate Polynomials That Are Hard to Factor

Let n be a positive integer and let r be an integer with $r \geq 2$. By ζ_r we denote $\exp(2\pi i/r)$, the first primitive r-th root of unity. Let $p_1, \ldots, p_n$ be n distinct positive prime numbers. By $f_{r;p_1,\ldots,p_n}(x)$ we denote the monic univariate polynomial in x whose roots are

$$\zeta_r^{i_1} p_1^{1/r} + \cdots + \zeta_r^{i_n} p_n^{1/r}$$

with $1 \leq i_1, \ldots, i_n \leq r$.

All $f_{r;p_1,\ldots,p_n}$ have integral coefficients and are irreducible polynomials of degree r^n over the integers. If r is a prime number, the following will be shown: If the coefficients of $f_{r;p_1,\ldots,p_n}$ are projected into a field of residues modulo any prime number q, Z_q, the image polynomials $f_{r;p_1,\ldots,p_n}$ (mod q) factor into irreducible polynomials over Z_q which have degree at most r.

If $r = 2$ this construction gives a slightly simpler version of the Swinnerton-Dyer polynomials which treat $\sqrt{-1}$ as an additional prime number. But our Galois theoretical proofs can be easily extended to yield this special case.

The condition of r being a prime number is not crucial for the unpleasant running time behavior for the factorization of these polynomials. For composite r the degrees of the irreducible factors in the modular domain are then bounded by r^2 (we will actually prove a somewhat better bound).

A modified version of these polynomials is also presented because of its closely related properties: By $f^*_{r;p_1,\ldots,p_n}$ we denote the polynomial whose roots are

$$\zeta_r^{i_0} + \zeta_r^{i_1} p_1^{1/r} + \cdots + \zeta_r^{i_n} p_n^{1/r}$$

where $1 \leq i_{0,}\ i_1, \ldots, i_n \leq r$ and $\mathrm{GCD}(i_0, r) = 1$.

Again all $f^*_{r;p_1,\ldots,p_n}$ are integer polynomials which factor modulo any prime q into polynomials whose degrees are bounded as for $f_{r;p_1,\ldots,p_n}$. If r is 2, 4, 6 or an odd integer, $f^*_{r;p_1,\ldots,p_n}$ is also irreducible over the integers. Otherwise these polynomials may be reducible but we can guarantee that all factors over the integers are of degree at least $2r^n$.

If $n = 0$, $f^*_{r;\varnothing}$ are the cyclotomic polynomials $\Psi_r(x)$. We will show that for certain composite r the maximum degree of factors in any residue field implies a super-polynomial running time for the Berlekamp-Hensel factorization algorithm. This fact is discussed in [Musser 75, p.302]. D. Knuth [Knuth 81, p.437] uses Berlekamp's algorithm to prove the modular factorization property for Ψ_8.

We need some number theoretic facts which we shall establish now. Let r be an integer with $r \geq 2$. As above, by U_r we denote the set of residues modulo r which are relatively prime to r. This set forms a group under multiplication modulo r and there exists a minimal non-negative integer $\lambda(r)$ such that for each $s \in U_r$: $s^{\lambda(r)} \equiv 1 \pmod r$. We call $\lambda(r)$ the minimum universal exponent modulo r. It is known (cf. [Knuth 81, p.19]) that

$$\lambda(2) = 1,\ \lambda(4) = 2,\ \lambda(2^\alpha) = 2^{\alpha-2} \text{ for } \alpha \geq 3$$
$$\lambda\left[2^{\alpha_0} q_1^{\alpha_1} \cdots q_n^{\alpha_n}\right] = \mathrm{LCM}\left[\lambda(2^{\alpha_0}), \phi(q_1^{\alpha_1}), \ldots, \phi(q_n^{\alpha_n})\right]$$

where the q_i are distinct odd prime numbers, ϕ is Euler's totient function and LCM means the least common multiple. Let p_i be the i-th consecutive prime number. As a consequence of Tchebycheff's theorem $p_i < 2^i$ for all $i > 1$ [Hardy and Wright 79, Theorem 418, p.343]. This enables us to prove the following:

Lemma 2.4: Let j be an integer with $j \geq 2$. Then there are infinitely many positive integers m (namely the product of the first k odd prime numbers with k sufficiently large) such that

$$\frac{\phi(m)}{\lambda(m)} > \log_2(\phi(m))^j.$$

Proof: Let $m = p_2 \cdots p_k$. Then

$$\phi(m) = (p_2-1) \cdots (p_k-1) < 2^{k(k+1)/2-1}$$

by the above estimate for p_i. Therefore $\log_2(\phi(m))^j < k^{3j}$. Also

$$\lambda(m) = \mathrm{LCM}(p_2-1, \ldots, p_k-1) < 2(p_2-1)/2 \cdots (p_k-1)/2 = 2^{2-k}\ \phi(m).$$

Hence $\phi(m)/\lambda(m) > 2^{k-2} > k^{3j}$ for k chosen large enough. Therefore for all sufficiently large k:

$\phi(m)/\lambda(m) > \log_2(\phi(m))^j$. □

In the proof of theorem 2.2 below we will make use of the fact that for every prime number r and for all $s \in U_r - \{1\}$: $\frac{s^{r-1}-1}{s-1}$ is a multiple of r. This follows from the Fermat theorem ($a^{\phi(b)} \equiv 1 \pmod{b}$ for $(a,b) = 1$) and the fact that r is a prime number. In order to treat composite r we generalize this matter:

Lemma 2.5: Let r be a positive composite integer. By $\eta(r)$ we denote the minimum exponent such that for each $s \in U_r - \{1\}$: $\frac{s^{\eta(r)}-1}{s-1}$ is divisible by r. Then $\eta(r) \le r\lambda(r)$. In fact, $\eta(r) \le d\lambda(r)$ where $d = \mathrm{LCM}(\{(s-1,r) \mid s \in U_r - \{1\}\})$.

Proof: Since for any s, $\mathrm{GCD}(s-1,r)$ divides r so must d and therefore $d \le r$. We claim that $(s^{d\lambda(r)}-1)/(s-1)$ is a multiple of r: To prove this we first factor $s^{d\lambda(r)}-1$ as

$$(s^{\lambda(r)}-1)\ (s^{(d-1)\lambda(r)}+s^{(d-2)\lambda(r)}+ \cdots +1).$$

Now the left factor is a multiple of r. It is therefore sufficient to show that the right factor is a multiple of d since that means it can absorb any factor of r in $s-1$ (by definition of d). But

$$s^{k\lambda(r)} \equiv (s^{\lambda(r)})^k \equiv 1 \pmod{d} \text{ for } 0 \le k \le d-1$$

since d divides r and thus

$$(s^{(d-1)\lambda(r)}+ \cdots +1) \equiv d.1 \equiv 0 \pmod{d},$$

as required. Therefore $\eta(r) \le d\lambda(r) \le r\lambda(r)$. □

Let f and g be two monic polynomials whose coefficients lie in some integral domain R. Let α_i, $1 \le i \le \deg(f)$ and β_j, $1 \le j \le \deg(g)$ denote their roots respectively. Since the polynomial

$$\prod_{i=1}^{\deg(f)} \prod_{j=1}^{\deg(g)} (x-\alpha_i-\beta_j)$$

is symmetric in both the α_i and the β_j it follows from the fundamental theorem of symmetric functions [van der Waerden 53, p.79] that its coefficients also lie in R. There is a resultant method which makes it feasible to actually compute this polynomial:

Lemma 2.6: Let R be an integral domain and let f and g be monic polynomials in $R[x]$. Then the resultant

$$(-1)^{\deg(f)\deg(g)}\,\mathrm{res}_y(f(x-y),g(y))$$

is a monic polynomial in $R[x]$ of degree $\deg(f)\deg(g)$ whose roots are $\alpha_i+\beta_j$ where α_i $(1 \le i \le \deg(f))$ are the roots of f and β_j $(1 \le j \le \deg(g))$ are the roots of g. (Cf. [Loos 82].) □

Now we mention a slight generalization of Eisenstein's irreducibility criterion, which can be used to show the irreducibility of some but not all of our polynomials.

Lemma 2.7: Let $f(x) = a_0+a_1x+\cdots+a_nx^n \in Z[x]$. If there exists a prime number p and an exponent i relatively prime to n such that $p^i \mid a_0, p^i \mid a_1, \ldots, p^i \mid a_{n-1}, p \nmid a_n, p^{i+1} \nmid a_0$ then f is irreducible over Z.
(Cf. [Kőnigsberger 1895].) □

Notice that the condition $\mathrm{GCD}(i,n) = 1$ in the above lemma is also necessary, because $x^4+4x^3+8x^2+8x+4 = (x^2+2x+2)^2$ yields a counterexample if this is not the case.

The next two lemmas constitute the key for our irreducibility proofs. By $[K:F]$ we denote the degree of a field K over a subfield F and by $F(\theta_1, \ldots, \theta_n)$ we denote the field F extended by the elements $\theta_1, \ldots, \theta_n$.

Lemma 2.8: Let r be an integer with $r \ge 2$, ζ_r a primitive r-th root of unity, and let $p_1, \ldots, p_n$ be distinct positive primes:

a) $[Q(p_1^{1/r}, \ldots, p_n^{1/r}):Q] = r^n$.

b) If $r \ge 3$ then $2r^n \le [Q(\zeta_r, p_1^{1/r}, \ldots, p_n^{1/r}):Q] \le \phi(r)\, r^n$.

c) If r is odd or 2, 4, or 6 then $[Q(\zeta_r, p_1^{1/r}, \ldots, p_n^{1/r}):Q] = \phi(r)\, r^n$.

Proof: Part a) is proven in [Besicovitch 40]. Part b) follows immediately from part a) and the fact that for $r \ge 3$ every ζ_r is a non-real number of algebraic degree $\phi(r)$ over Q. Part c) is proven for odd r in

[Richards 74] which is also a special case of [Caviness 68, Theorem 10, p.50]. If $r = 2$ part c) is actually the same as part a) because $\zeta_2 = -1$. For $r = 4$ or 6 we combine part b) and the fact that both $\phi(4)$ and $\phi(6)$ are 2. □

Notice that part c) may not hold for even $r \geq 8$ depending on what primes $p_1, \ldots, p_n$ are chosen. Counterexamples may be constructed using the fact that $\sqrt{2} \in Q(\zeta_8)$ or $\sqrt{5} \in Q(\zeta_{10})$.

Lemma 2.9: Let r be an integer ≥ 2, ζ_r a primitive r-th root of unity, and let $p_1, \ldots, p_n$ be distinct prime numbers. Then $p_n^{1/r}$ is not an element of the field $Q(\zeta_r, p_1^{1/r}, \ldots, p_{n-1}^{1/r})$.

Proof: If r is 2 the fact follows from part a) of lemma 2.8. By F_k we denote the field $Q(\zeta_r, p_1^{1/r}, \ldots, p_k^{1/r})$ with $1 \leq k \leq n$. Now assume that $r \geq 3$ and $p_n^{1/r} \in F_{n-1}$ which implies that $F_n = F_{n-1}$. Applying part b) of lemma 2.8 we get $2r^n \leq [F_n:Q] = [F_{n-1}:Q] \leq \phi(r)r^{n-1}$, which is impossible. □

The following lemma will enable us to give an alternate proof of theorem 2.2:

Lemma 2.10: Let $m \in Z_q$, q being prime, and let r be an integer greater than 1.

a) A necessary and sufficient condition for the existence of an r-th root of m in Z_q (i.e. a residue b such that $b^r \equiv m \pmod{q}$) is that r is either relatively prime to $q-1$ or $m^{(q-1)/d} \equiv 1 \pmod{q}$ where $d = \mathrm{GCD}(q-1, r)$. (Cf. [van der Waerden 53, p.118, Exercise 2].)

b) The polynomial $x^r - m$ in $Z_q[x]$ has at least one root α such that $\alpha \in \mathrm{GF}(q^d)$ and d divides r.

Proof: a) Let g be a primitive root of Z_q i.e. a generating element of $Z_q - \{1\}$ with multiplication. Then $g, g^2, \ldots, g^{q-1}=1$ are distinct residues modulo q. If $\mathrm{GCD}(r, q-1)=1$ then $g^r, g^{2r}, \ldots, g^{(q-1)r}=1$ are also distinct residues and therefore exactly one element is equal to m. If b is an r-th root of m then by Lagrange's theorem

$$b^{(q-1)r/d} \equiv m^{(q-1)/d} \equiv 1 \pmod{q} \text{ where } d = \mathrm{GCD}(q-1, r).$$

We finally prove that if $m^{(q-1)/d} \equiv 1 \pmod{q}$ then $m \equiv g^{jr} \pmod{q}$ for some integer $j \geq 1$. Assume m

$\equiv g^i \pmod{q}$ and d does not divide i. Then $q-1$ does not divide $i(q-1)/d$ and therefore

$$m^{(q-1)/d} \equiv g^{i(q-1)/d} \not\equiv 1 \pmod{q}.$$

Hence $m \equiv g^{kd} \pmod{q}$ with $k \geq 1$. Since $\mathrm{GCD}(r/d, q-1) = 1$ there exists a $j \equiv k(r/d)^{-1} \pmod{q-1}$ and therefore $m \equiv g^{kd} \equiv g^{jr} \pmod{q}$. (Notice that this proof works also if we replace Z_q by $\mathrm{GF}(q^n)$.)

b) We use induction on r: For $r = 1$ the statement is trivial. Assume that $r > 1$. We now distinguish two cases:

Case 1: There is a factor $r_1 > 1$ of r such that an r_1-th root m_1 of m exists in Z_q. By part a) we already know that this is always true if $r \nmid q-1$. Let $r_2 = r/r_1$. Then

$$x^r - m \equiv x^{r_1 r_2} - m_1^{r_1} \equiv (x^{r_2} - m_1)(x^{(r_1-1)r_2} + x^{(r_1-2)r_2} m_1 + \cdots + m_1^{r_1-1}) \pmod{q}.$$

Applying the induction hypothesis to $x^{r_2} - m_1$ yields the statement for r.

Case 2: r divides $q-1$. Let α be a root of $x^r - m$ and let ζ_r be a primitive r-th root of unity both of which lie in some Galois field. Let h be the minimal polynomial of α over Z_q whose constant coefficient be denoted by h_0. Since $h(x)$ divides $x^r - m \equiv (x-\alpha)(x-\zeta_r\alpha) \cdots (x-\zeta_r^{r-1}\alpha)$, it follows that $h_0 = \zeta_r^t \alpha^s$ with $s = \deg(h)$ and t some positive integer. Therefore $h_0^r = m^s$. If $d = \mathrm{GCD}(s,r)$ we can find suitable integers u, v such that $us + vr = d$. Then $m^d = m^{us} m^{vr} = h_0^{ur} m^{vr} = (h_0^u m^v)^r$ which implies that m^d possesses an r-th root in Z_q. From part a) we conclude that $m^{d(q-1)/r} \equiv 1 \pmod{q}$ and further that there exists an r/d-th root of m. If $s < r$ then $r/d > 1$ and we can apply case 1. Otherwise $x^r - m$ is already irreducible. □

Case 2 of the above proof yields an interesting side result: Let q be a prime and r an integer dividing $q-1$. Then $x^r - m$ is irreducible over Z_q if $m^{(q-1)/d} \not\equiv 1 \pmod{q}$ for all divisors d of r. As we showed in part a) of the above lemma this is true for all m of the form g^i where g is a primitive root of Z_q and $\mathrm{GCD}(i,r)$=1. Hence by picking a random residue m the probability that $x^r - m$ does not factor in $Z_q[x]$ is $\phi(r)/r$. Choosing an r-th degree polynomial randomly only yields a probability of $1/r$ [Rabin 80]. Moreover it follows from theorem 328 in [Hardy and Wright 79, p.267] that

$\phi(r)/r > 0.56/\log\log r$ for almost all r. Therefore in searching for irreducible polynomials in $Z_q[x]$ of degree r, r a divisor of $q-1$, we will succeed considerably sooner by choosing the above polynomials than entirely random ones.

Theorem 2.1: Let r be an integer with $r \geq 2$ and let $p_1, \ldots, p_n$ be distinct prime numbers. Then $f_{r;p_1,\ldots,p_n}$ and $f^*_{r;p_1,\ldots,p_n}$ have integer coefficients and the following irreducibility conditions hold:

a) $f_{r;p_1,\ldots,p_n}$ is irreducible over the integers and each irreducible factor of $f^*_{r;p_1,\ldots,p_n}$ over the integers with $r \geq 3$ has degree at least $2r^n$.

b) If $r = 2, 4, 6$ or odd then $f^*_{r;p_1,\ldots,p_n}$ is irreducible.

Proof: Using lemma 2.6 inductively we see that the coefficients of $f_{r;p_1,\ldots,p_n}$ and $f^*_{r;p_1,\ldots,p_n}$ are integers and that their degrees are r^n and $\phi(r)r^n$ respectively. (Notice that Ψ_r has integer coefficients as mentioned before.) First we prove by induction that $p_1^{1/r} + \cdots + p_n^{1/r}$ is a primitive element of $Q(p_1^{1/r}, \ldots, p_n^{1/r})$. We make use of the construction of a primitive element given in [van der Waerden 53, p.126]: Let $\alpha_1 = p_1^{1/r} + \cdots + p_{n-1}^{1/r}$ and $\alpha_2, \ldots, \alpha_{r^{n-1}}$ be the remaining roots of $f_{r;p_1,\ldots,p_{n-1}}$. By the induction hypothesis $Q(\alpha_1) = Q(p_1^{1/r}, \ldots, p_{n-1}^{1/r})$. The minimal polynomial of α_1 is of degree $[Q(\alpha_1):Q]$ which is r^{n-1} by lemma 2.8. Therefore $f_{r;p_1,\ldots,p_{n-1}}$ is this minimal polynomial. Let $\beta_1 = p_n^{1/r}, \beta_2, \ldots, \beta_r$ be the roots of $x^r - p_n$ which is irreducible by Eisenstein's criterion (lemma 2.7). Then $\alpha_1+\beta_1$ is a primitive element of

$$Q(\alpha_1,\beta_1) = Q(p_1^{1/r}, \ldots, p_n^{1/r})$$

provided that $\alpha_1+\beta_1 \neq \alpha_i+\beta_j$ for $1 \leq i \leq r^{n-1}$ and $1 < j \leq r$. For the sake of contradiction assume that this condition cannot be achieved, namely there exist an i and a $j > 1$ such that $\alpha_1-\alpha_i = \beta_1-\beta_j$. Since $\beta_j = \zeta_r^k p_n^{1/r}$ for some $k \geq 1$ it follows that $\alpha_1-\alpha_i = p_n^{1/r}(1-\zeta_r^k)$ and therefore $p_n^{1/r} = (\alpha_1-\alpha_i)/(1-\zeta_r^k)$ which is an element of $Q(\zeta_r, p_1^{1/r}, \ldots, p_{n-1}^{1/r})$, in contradiction to lemma 2.9. Noticing that Ψ_r is irreducible we can prove in exactly the same way that $\zeta_r + p_1^{1/r} + \cdots + p_n^{1/r}$ is a primitive element of $Q(\zeta_r, p_1^{1/r}, \ldots, p_n^{1/r})$. (However, the α_i will be the roots of an irreducible factor of $f^*_{r;p_1,\ldots,p_{n-1}}$.)

We now conclude that the minimal polynomials of these primitive elements are of the same degree as the field extensions obtained by adjoining them to the rationals which we know by lemma 2.8, part a) and c). Therefore $f_{r;p_1,...,p_n}$ and, in the case that $r = 2$, 4, 6 or an odd integer, $f^*_{r;p_1,...,p_n}$ are these minimal polynomials and hence must be irreducible. All irreducible factors of $f^*_{r;p_1,...,p_n}$ have degree at least $2r^n$ because all roots are primitive elements by the argument above and the lower bound of the corresponding field extension is known from lemma 2.8b). □

Theorem 2.2: Let r be an integer with $r \geq 2$ and let $p_1, ..., p_n$ be prime numbers. For any prime number q the following factorization properties hold for the projected polynomials $f_{r;p_1,...,p_n}$ (mod q) and $f^*_{r;p_1,...,p_n}$ (mod q):

a) The maximum degree of any irreducible factor of both polynomials over the residue field modulo q is at most $r\lambda(r)$. Special case: If r is a prime number the maximum degree is r.

b) If $n = 0$ then the maximum degree of an irreducible factor of $f^*_{r;\varnothing}$ (mod q) = Ψ_r (mod q) is $\lambda(r)$.

Proof: a) We first show that the length of the longest cycle in any permutation of the Galois group of $f_{r;p_1,...,p_n}$ or $f^*_{r;p_1,...,p_n}$ is at most $\max(r,\eta(r))$, where $\eta(r)$ is as defined in lemma 2.5. Let σ be an automorphism on $Q(\zeta_r, p_1{}^{1/r}, ..., p_n{}^{1/r})$. As such it has to map the roots of the polynomials Ψ_r and $x^r - p_i$ into roots of the same polynomials. In particular $\sigma(\zeta_r) = \zeta_r^{s_\sigma}$ where ζ_r is a primitive r-th root of unity and s_σ is relatively prime to r. Also $\sigma(p_i{}^{1/r}) = \zeta_r^{m_i}\, p_i{}^{1/r}$, where the m_i depend also on σ ($1 \leq i \leq n$). (Notice that ζ_r generates all distinct r-th roots of unity.) We now distinguish two cases:

Case 1: $s_\sigma = 1$. Applying σ r times we get $\sigma^{(r)}(p_i{}^{1/r}) = p_i{}^{1/r}$ for all $1 \leq i \leq n$ and therefore $\sigma^{(r)}$ maps each root of $f_{r;p_1,...,p_n}$ and $f^*_{r;p_1,...,p_n}$ onto itself which is to say that the permutation corresponding to σ has cycles of length at most r.

Case 2: $s_\sigma > 1$. By lemma 2.5 we know that both

$$s_\sigma^{\eta(r)} \equiv 1 \pmod{r} \text{ and } \frac{s_\sigma^{\eta(r)}-1}{s_\sigma-1} \equiv 0 \pmod{r}.$$

A short computation shows that then

$$\sigma^{(\eta(r))}(\zeta_r) = \zeta_r \text{ and } \sigma^{(\eta(r))}(p_i^{1/r}) = p_i^{1/r}$$

for all $1 \le i \le n$. Therefore the cycle lengths of the permutation corresponding to σ are at most $\eta(r)$.

Cases 1 and 2 together prove the statement made initially. If the image polynomials are separable we are finished by virtue of the lemmas 2.2, 2.3 and 2.5. But we can repeat the above arguments for automorphisms on the splitting field of $f_{r;p_1,\ldots,p_n} \pmod q$ itself because as we mentioned before the properties of r-th roots of unity carry over for ground fields of characteristic q, provided that q does not divide r. Finally let q^m be the highest power of q dividing r. By using the identity introduced in the proof of lemma 2.1 and by using lemma 2.1 itself we can determine the multiplicities of the roots of $\Psi_r \pmod q$ and $x^r - p_i \pmod q$ (which lie in some Galois field). Therefore

$$f_{r;p_1,\ldots,p_n} \equiv \left[f_{r/q^m;p_1,\ldots,p_n} \right]^{q^{mn}} \pmod q$$

and

$$f^*_{r;p_1,\ldots,p_n} \equiv \left[f^*_{r/q^m;p_1,\ldots,p_n} \right]^{\phi(q^m)q^{mn}} \pmod q.$$

It follows from the formula for λ given at the beginning of this section that $\lambda(r/q^m)$ divides $\lambda(r)$. Then by lemma 2.5 and the already proven theorem for the case that q does not divide r we conclude that the maximum degree in this case is $r/q^m \ \lambda(r/q^m) < r\lambda(r)$. If r is a prime number the above proof together with the remark made above lemma 2.5 actually gives the degree bound r.

b) If $\Psi_r \pmod q$ is separable we know its Galois group to be a subgroup of U_r under multiplication modulo r. (This by lemma 2.2 but one may verify it directly.) The definition of λ and lemma 2.3 then lead to the statement. If $\Psi_r \pmod q$ is inseparable q necessarily divides r. Again putting together the above, lemma 2.1 and the fact that $\lambda(r/q^m)$ divides $\lambda(r)$ proves the theorem for this case. □

In special cases the bound $r\lambda(r)$ is actually too pessimistic: If the image polynomial is separable or more generally if q does not divide r we have actually proven that the bound is $\max(r,\eta(r))$ which may be considerably smaller than $r\lambda(r)$. One can show that this is generally true by proving that $\eta(d)$

is not larger than $\eta(r)$ for any divisor d of r. By lemma 2.10a) we also know that each p_i possesses an r-th root in Z_q if r is relatively prime to $q-1$. Then the maximum degree over Z_q can be bounded by $\lambda(r)$ instead. The second case of lemma 2.10a) applies as well.

We now present a second proof of theorem 2 expanding ideas from [Berlekamp 70, p.734] with the help of lemma 2.10b). However, this method does not introduce the function η and therefore in view of the preceeding remarks is somewhat weaker.

Alternate Proof of Theorem 2.2: If q divides r we must apply the same reduction as in the last part of the previous proof. Now assume that $q \nmid r$. We will show part b) first:

b) Let α be a root of an irreducible factor g of Ψ_r (mod q). Then g is separable and the remaining roots are α^q, $\alpha^{q^2}, \ldots,$ $\alpha^{q^{\deg(g)-1}}$. However $q^{\lambda(r)} \equiv 1$ (mod r) and also $\alpha^r = 1$ which implies $\alpha^{q^{\lambda(r)}} = \alpha$. Therefore $\deg(g) \le \lambda(r)$.

a) By lemma 2.10 b) and the observation about the subfields of a Galois field we know that at least one root of each x^r-p_i, $1 \le i \le n$ lies in GF(q^r). From part b) above we conclude that all r-th roots of unity are in a GF(q^s) with $s \le \lambda(r)$. Therefore all roots of x^r-p_i and Ψ_r lie in GF(q^{rs}) and hence any sum of them does also. If $f_{r;p_1,\ldots,p_n}$ (mod q) or $f^*_{r;p_1,\ldots,p_n}$ (mod q) had an irreducible factor g of degree greater than rs then one of its roots would generate GF$(q^{\deg(g)})$. But we know that this root lies in GF(q^{rs}). Therefore $\deg(g) \le rs \le r\lambda(r)$. □

One may use lemma 2.6 in connection with a method to compute cyclotomic polynomials [Knuth 81, Sec.4.6.2, Exercise 32] to actually generate sample polynomials.

Example 2.1: $n=0$:

$$f^*_{8;0}(x) = \Psi_8(x) = x^4+1,\ \lambda(x) = 2. \tag{1}$$

$$f^*_{12;0}(x) = \Psi_{12}(x) = x^4-x^2+1,\ \lambda(12) = 2. \tag{2}$$

$$f^*_{15;0}(x) = \Psi_{15}(x) = x^8-x^7+x^5-x^4+x^3-x+1,\ \lambda(15) = 4. \tag{3}$$

$n=1$:

$$f^{*}_{3;2}(x) = x^6+3x^5+6x^4+3x^3+9x+9. \tag{4}$$

$$\begin{aligned} f^{*}_{8;2}(x) = {} & (x^{16}+4x^{12}-16x^{11}+80x^9+2x^8+160x^7 \\ & +128x^6-160x^5+28x^4-48x^3+128x^2-16x+1) \\ & (x^{16}+4x^{12}+16x^{11}-80x^9+2x^8-160x^7 \\ & +128x^6+160x^5+28x^4+48x^3+128x^2+16x+1). \end{aligned} \tag{5}$$

$n=2$:

$$f_{2;2,3}(x) = x^4-10x^2+1. \tag{6}$$

$$f_{3;2,3}(x) = x^9-15x^6-87x^3-125. \tag{7}$$

$$f_{4;2,3}(x) = x^{16}-20x^{12}+666x^8-3860x^4+1. \tag{8}$$

$$f_{5;2,3}(x) = x^{25}-25x^{20}-3500x^{15}-57500x^{10}+21875x^5-3125. \tag{9}$$

$$\begin{aligned} f^{*}_{3;2,3}(x) = {} & x^{18}+9x^{17}+45x^{16}+126x^{15}+189x^{14}+27x^{13}-540x^{12} \\ & -1215x^{11}+1377x^{10}+15444x^9+46899x^8+90153x^7 \\ & +133893x^6+125388x^5+29160x^4-32076x^3+26244x^2 \\ & -8748x+2916. \end{aligned} \tag{10}$$

$n=3$:

$$f_{2;2,3,5} = x^8-40x^6+352x^4-960x^2+576. \tag{11}$$

$$f_{2;-1,2,3} = x^8-16x^6+88x^4+192x^2+144. \tag{12}$$

Example 2.1 illustrates very well our results: All but polynomial (5) are irreducible over the integers. Since $\sqrt{2} \in Q(\zeta_8)$ we also know that (5) must be composite. Notice here that lemma 2.7 cannot be used to show the irreducibility of (4), (9), (11) and (12). All the polynomials (1)-(12) factor in any modular field into polynomials of smaller degrees and make excellent test cases for implementations of the Berlekamp-Hensel factorization algorithm. E.g. polynomial (10) factors

$$\begin{aligned} \text{mod } 7\text{:} \quad & (x^3+x^2+4x+3)(x^3+2x^2+5x+5)(x^3+2x^2+4x+2) \\ & (x^3+x^2+3x+5)(x^3+2x^2+2x+3)(x^3+x^2+x+2) \\ \text{mod } 17\text{:} \quad & (x^2+12x+16)(x^2+16x+7)(x^2+9x+13)(x^2+9x+9) \\ & (x^2+16x+16)(x^2+12x+9)(x^2+5x+7)(x^2+16x+1)(x+8)^2 \\ \text{mod } 103\text{:} \quad & (x^3+9x^2+27x+25)(x^3+62x^2+11x+28)(x^3+73x^2+94x+28) \\ & (x^3+39x^2+95x+32)(x^3+92x^2+6x+25)(x^3+43x^2+67x+32) \\ \text{mod } 1979\text{:} \quad & (x^2+1823x+1632)(x^2+85x+6)(x^2+828x+749) \\ & (x^2+1069x+6)(x^2+1069x+749)(x^2+1069x+1632) \\ & (x^2+1069x+878)(x^2+85x+1744)(x^2+828x+1744) \end{aligned}$$

The variation of the maximum degree bound for different primes will be explained in section 2.3.

The Berlekamp-Hensel factorization algorithm contains the following "bottleneck" [Knuth 81, p.434]: If f is a polynomial of degree k and splits in a chosen residue field into j irreducible factors then one must perform at least $2^{j-1}-1$ trial divisions to prove its irreducibility over the integers in step (F5) of algorithm 1.3. In the case of $f_{r;p_1,\ldots,p_n}$, $k = r^n$ and $j \geq r^{n-1}/\lambda(r)$ and hence at least $2^{r^{n-1}/\lambda(r)-1}-1$ steps are executed. Fixing r gives an $O(2^k)$ lower timing bound for these inputs. We will show below that the lengths of the coefficients are bounded by $O(k \log\log(k))$ and thus the worst case time complexity of the Berlekamp-Hensel algorithm is indeed an exponential function of the degree and coefficient lengths of its inputs. Since the degrees of all irreducible factors of $f_{r;p_1,\ldots,p_n} \pmod{q}$ are independent of n the modifications of this algorithm suggested in [Musser 78] do not eliminate the exponential running time behavior.

The cyclotomic polynomials Ψ_m with m chosen as in lemma 2.4 are significant because even if $\Psi_m \pmod{q}$ is not squarefree then the multiplicities of its factors are prime divisors of m, which are a small numbers compared to $\deg(\Psi_m) = \phi(m)$. The number of irreducible modular factors causes a super-polynomial running time for the Berlekamp-Hensel algorithm due to lemma 2.4.

Finally we establish certain bounds for the coefficients of our polynomials when the primes p_i are as small as possible. For a polynomial f, let $| f |_1$ denote the sum of the absolute values of the coefficients of f.

Theorem 2.3: Let r be an integer ≥ 2 and let $p_1, \ldots, p_n$ be the first n primes.

a) $\log(| f |_1) = O(\deg(f) \log\log(\deg(f)))$ for $f = f_{r;p_1,\ldots,p_n}$ and for $f = f^*_{r;p_1,\ldots,p_n}$.

b) $\log_2(| \Psi_m |_1) \leq \phi(m)$ for $m \geq 1$.

Proof: Given $f(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x_k \in Z[x]$, let B denote the maximum of the absolute values of the roots of f. Then, since the coefficients are the elementary symmetric functions of the roots, it follows that $a_i \le \begin{bmatrix} k \\ i \end{bmatrix} B^{k-i}$ for $0 \le i \le k$. Therefore $| f |_1 \le (B+1)^k$.

a) For $f = f_{r;p_1, \ldots, p_n}$ the maximum absolute value of the roots is $B = p_1^{1/r} + \cdots + p_n^{1/r}$ and for $f^* = f^*_{r;p_1, \ldots, p_n}$ it is $B^* = 1+B$. Using the prime number distribution law [Hardy and Wright 79, Theorem 8, p.10] we know that p_i is of order $O(i \log(i))$, so that B and B^* are of order $O(n^2)$. Since r is fixed n is of order $O(\log \deg(f))$ and $O(\log \deg(f^*))$. Taking the logarithm of the previous inequality for the norm immediately establishes part a).

b) Every root of Ψ_m has absolute value 1 and hence $| \Psi_m |_1 \le (1+1)^{\phi(m)}$. □

2.3. Computation of Galois Groups

In this section, we will exploit the Chebotarev Density Theorem for the explicit construction of the Galois groups of univariate integer polynomials. In order to formulate our results, we need a constructive definition of the Frobenius element. This requires some fundamental results of the theory of algebraic integers which we shall present now. The reader can find the corresponding proofs in most any book on algebraic number theory, e.g. [Narkiewicz 74, Chap.2 and Chap.8].

Let f be a monic irreducible polynomial over Z, such that f is also normal, i.e., any root of f generates the complete splitting field of f. Let α be a root of f. If K is an algebraic number field then all elements, whose minimal polynomials are monic, form a subring of K, the ring of algebraic integers of K denoted by R_K. Let $n = \deg(f)$ and hence $[Q(\alpha):Q] = n$. There exist algebraic integers $b_1, \ldots, b_n \in R_{Q(\alpha)}$ such that $\{b_1, \ldots, b_n\}$ generates $R_{Q(\alpha)}$ over Z. We call $\{b_1, \ldots, b_n\}$ an integral basis of $R_{Q(\alpha)}$. We shall use two notations for the Galois group of $Q(\alpha)$. $G_{Q(\alpha)}$ denotes the group of all automorphisms of $Q(\alpha)$, G_f the group of all corresponding permutations of roots of f. The trace of an element $x \in Q(\alpha)$ is defined as

$$T_{Q(\alpha)}(x) = \sum_{\sigma \in G_{Q(\alpha)}} \sigma(x),$$

which is a rational number. The discriminant of the field $Q(\alpha)$ is defined by

$$D_{Q(\alpha)} = \det\,[T_{Q(\alpha)}(b_i b_j)]_{i,j=1,\ldots,n}$$

which is independent of the choice for the integral basis $\{b_i\}$. By $1/m\ R_{Q(\alpha)}$, $m \in Z$, we denote those algebraic numbers, which when multiplied by m become algebraic integers. Let Δ_f denote the discriminant of f, i.e., $\Delta_f = \mathrm{res}(f, f')$.

Lemma 2.11. If $f \in Z[x]$ is monic, normal and irreducible with the root α then $\Delta_f = m^2 D_{Q(\alpha)}$, where m is the index of the Z-module $Z[\alpha]$ generated by $\{1, \alpha, \ldots, \alpha^{n-1}\}$ in $R_{Q(\alpha)}$. Furthermore, $R_{Q(\alpha)} \subseteq 1/m\ Z[\alpha] \subseteq 1/\Delta_f\ Z[\alpha]$.

Proof: The equation $\Delta_f = m^2 D_{Q(\alpha)}$ is a special case of [Narkiewicz 74, Proposition 2.6]. There it is proven that for an integral basis $\{b_i\}$ of $R_{Q(\alpha)}$ the set $\{c_1, \ldots, c_n\}$ with

$$c_i = \sum_{k=1}^{i} d_{ik} b_k \quad (d_{ik} \in Z), \tag{A}$$

under the condition that $c_i \in Z[\alpha]$ and d_{ii} positive and as small as possible, generates $Z[\alpha]$ over Z. Also $m = \prod_{i=1}^{n} d_{ii}$ which, by Cramer's rule applied to (A), yields $b_i \in 1/m\ Z[\alpha]$. Since $\{b_i\}$ forms an integral basis, we obtain $R_{Q(\alpha)} \subseteq 1/m\ Z[\alpha]$. □

A prime number p which does not divide Δ_f is called unramified. By f_p we mean f taking each coefficient modulo p; $f_p \in Z_p[x]$. We shall now focus on the prime ideals in $R_{Q(\alpha)}$. The following lemma contains a classical result by E. Kummer.

Lemma 2.12: Let $f \in Z[x]$ be monic, normal and irreducible with the root α and let p be an unramified prime. Furthermore, let $f_p \equiv f_1 \cdots f_r$ modulo p with f_i monic and $(f_i)_p$ irreducible in GF$(p)[x]$.

a) The degrees of all f_i are equal, i.e.

$$\deg f_1 = \cdots = \deg f_r = s.$$

b) The only prime ideals in $R_{Q(\alpha)}$ containing pZ are

$$P_i = pR_{Q(\alpha)} + f_i(\alpha)R_{Q(\alpha)},$$

which are also maximal.

Proof: a) Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the roots of f. By lemma 2.11, there exist polynomials $g_i \in 1/m\ Z[x]$ such that $g_i(\alpha) = \alpha_i$, $2 \le i \le n$. Now let α_p be a root of $(f_i)_p$. Since p does not divide m, all other roots of f_p can be computed by $(g_i)_p(\alpha_p)$. Therefore, the splitting field of f_p is GF(p^s) and $\deg(f_i)_p = s$.

b) Assume $P \supseteq pZ$ is a prime ideal in $R_{Q(\alpha)}$. Since $f_1(\alpha) \cdots f_r(\alpha) \in pZ[\alpha] \subseteq P$, at least one $f_i(\alpha) \in P$. Therefore, $P_i \subseteq P$ for some $i \in \{1, \ldots, r\}$.

Let $y \in P$. By lemma 2.11 we can write $y = g(\alpha)$ with $g(x) \in 1/m\ Z[x]$. Let $\overline{g} = m(m^{-1} \bmod p)g$. We show first that if $(f_i)_p$ divides $\overline{g}_p$ then $y \in P_i$. Then we prove that if $(f_i)_p$ does not divide $\overline{g}_p$ then $P = R_{Q(\alpha)}$, which also establishes the maximality of P_i, and finally that P_i is prime.

First assume that $\overline{g} = \overline{q}f_i + \overline{r}$ with $\overline{q}$, $\overline{r} \in Z[x]$, $\overline{r}_p$=0. Then $(g - \overline{q}f_i)(\alpha) = pr(\alpha) \in R_{Q(\alpha)}$ with $r(x) \in 1/m\ Z[x]$. We have to show that $r(\alpha) \in R_{Q(\alpha)}$. Let $pr(\alpha_1), \ldots, pr(\alpha_k)$ be the conjugates of $pr(\alpha)$. Then $\prod_{i=1}^{k} (x - pr(\alpha_k)) \in Z[x]$ and since p does not divide m, we conclude that also $\prod_{i=1}^{k} (x - r(\alpha_k)) \in Z[x]$, i.e., $r(\alpha) \in R_{Q(\alpha)}$. Therefore

$$y = g(\alpha) = \overline{q}(\alpha)f_i(\alpha) + pr(\alpha) \in P_i.$$

Now assume that $(f_i)_p$ does not divide $\overline{g}_p$. Then, since $(f_i)_p$ is irreducible, the resultant $R = \mathrm{res}(f_i, \overline{g}) \neq 0 \bmod p$, but $R \in P$ ($\overline{g}(\alpha) \in P$) and therefore $P = R_{Q(\alpha)}$.

We proceed to show that $1 \notin P_i$. Assume it were, i.e., there exist $y_1 = g_1(\alpha)$, $y_2 = g_2(\alpha) \in R_{Q(\alpha)}$ such that $py_1 + f_i(\alpha)y_2 = 1$. Then $(f_i)_p(\overline{g}_2)_p \equiv 1 \bmod f_p$ which is impossible since $(f_i)_p$ divides f_p.

Finally, we establish the primality of P_i. Let $y_1 = g_1(\alpha)$, $y_2 = g_2(\alpha) \in R_{Q(\alpha)}$ s.t. $y_1 y_2 \in P_i$. Then $(\overline{g_1 g_2})(\alpha) \in P_i$ and therefore $(f_i)_p$ divides $(\overline{g_1 g_2})_p$. Since $(f_i)_p$ is irreducible, $(f_i)_p$ divides either one of the polynomials $(\overline{g}_1)_p$ or $(\overline{g}_2)_p$, say the first one. From the above it follows then that $y_1 \in P_i$. □

Lemma 2.12 actually shows that the fields $R_{Q(\alpha)}$ mod P_i, GF(p)[x] mod $(f_i)_p$, and GF(p^s) are isomorphic. From lemma 2.2 and the remarks on Galois fields in section 2.1 we deduce that the Galois group G_{f_p} is a cyclic subgroup of G_f with one distinguished generator, isomorphic to the Frobenius automorphism $\beta \to \beta^p$ for all $\beta \in$ GF(p^s). Let $\overline{f}$ be a monic integer polynomial of degree $\overline{n}$, not necessarily irreducible, whose splitting field is $Q(\alpha)$. Assume further that $\overline{f}_p$ is squarefree. We now identify the Frobenius automorphism under the permutations of $G_{\overline{f}}$. Let

$$\overline{f}(x) \equiv (x - g_1(\alpha))(x - g_2(\alpha)) \cdots (x - g_{\overline{n}}(\alpha)) \bmod f(\alpha) \tag{2.1}$$

with $g_k \in 1/m\ Z[\alpha]$, and let

$$h_{k,i} = (g_k)_p \bmod (f_i)_p,\ 1 \le k \le \overline{n}.$$

If α_i is a root of $(f_i)_p$, then the elements $h_{k,i}(\alpha_i)$ are all roots of $\overline{f}_p$. Let

$$\overline{f}_1 \cdots \overline{f}_{\overline{r}} \equiv \overline{f} \bmod p$$

be a factorization of $\overline{f}$ into irreducibles with $\deg(\overline{f}_i) = \overline{s}_i$. Let $h_{k_j,i}(\alpha_i)$ be a root of $(\overline{f}_j)_p$ and let $h_{k^{(t)},i}$ denote $h_{k,i}{}^t \bmod (f_i)_p$. Notice that if t is a power of p the index $k^{(t)}$ is uniquely determined because $\overline{f}_p$ does not possess multiple roots. Then the permutation

$$\left[k_1, k_1^{(p)}, \ldots, k_1^{(p^{\overline{s}_1 - 1})}\right] \ \cdots \ \left[k_{\overline{r}}, k_{\overline{r}}^{(p)}, \ldots, k_{\overline{r}}^{(p^{\overline{s}_r - 1})}\right] \tag{2.2}$$

is the image of the Frobenius automorphism in $G_{\overline{f}}$ using the root enumeration of (2.1).

At this point we shall be more precise in what we mean by root enumeration. Until now there was no need in explicitly stating what algebraic structure $Q(\alpha)$ had. However, in the previous paragraph it is important to set $Q(\alpha)$ to $Q[x]/(f(x))$ and α to the projection of x in that domain. We then can make the factorization (2.1) canonical by imposing the condition $\deg(g_k) < n = \deg(f)$. In order to find a proper embedding of $G_{\overline{f}_p}$ into $G_{\overline{f}}$ we now can use any irreducible factor f_i of $f \bmod p$. The splitting field of $\overline{f}_p$ is then $\left[Q[x]/(f(x))\right]/(p, f_i(x))$ and the natural projection maps canonical roots of $\overline{f}$ onto unique roots of $\overline{f}_p$ since the later polynomial is squarefree. Notice that our permutation (2.2) not only depends on f_i but also on the initial resolvent f.

We call the image of the permutation (2.2) in $G_{Q(\alpha)}$ the Frobenius element of P_i in $Q(\alpha)$ and denote it by $\left[\dfrac{Q(\alpha)/Q}{P_i}\right]$. The Frobenius element generates exactly those automorphisms in $Q(\alpha)$ which map elements of P_i into P_i, the so-called decomposition group of P_i in $Q(\alpha)$. Given f and the factorization (2.1) we have described an algorithm of how to identify Frobenius elements by permutations in $G_{\overline{f}}$.

The Frobenius element depends on the prime ideal P_i. For a given automorphism $\sigma \in G_{Q(\alpha)}$ we define the conjugate class of σ, $C_\sigma = \{\tau\,\sigma\,\tau^{-1} \mid \tau \in G_{Q(\alpha)}\}$. Notice that the corresponding permutations of a conjugate class possess the same cycle structure. The group $G_{Q(\alpha)}$ can be partitioned into finitely many such classes.

Lemma 2.13.: The set

$$\left\{ \left[\frac{Q(\alpha)/Q}{P_i} \right] \;\Big|\; i=1,\ldots,r \right\}$$

forms a complete conjugate class of $G_{Q(\alpha)}$, denoted by $F_{Q(\alpha)}(p)$. (Cf. [Narkiewicz 74, Proposition 7.12 and Theorem 4.2].) □

The question arises whether any conjugate class of $G_{Q(\alpha)}$ can be realized by $F_{Q(\alpha)}(p)$, letting p range over all unramified prime numbers for f. The first positive answer goes back to L. Dirichlet for the special case of f being the m-th cyclotomic polynomial Ψ_m. This case is equivalent to the problem whether there are infinitely many primes in an arithmetic progression. In 1896, G. Frobenius showed that certain unions of conjugate classes will always be realized by Frobenius elements, and in 1926 N. Chebotarev proved his classical theorem.

Theorem 2.4 *(Chebotarev Density Theorem):*

Let C be a conjugate class of G_f and let card(S) denote the cardinality of the set S. If

$$N_C(x) = \text{card}\,\{\, p \;|\; p \text{ prime}, p \le x \text{ s.t. } F_{Q(\alpha)}(p) = C \,\}$$

then

$$N_C(x) = \left[\frac{\text{card}(C)}{\text{card}(G_f)} + \varepsilon(x) \right] \frac{x}{\log(x)}.$$

where $\lim\limits_{x\to\infty} \varepsilon(x) = 0$. (Cf. [Chebotarev 26] or [Narkiewicz 74, Theorem 7.10].) □

The nature of the error term $\varepsilon(x)$ in theorem 2.4 has been quite successfully analyzed in the past decade.

Theorem 2.5. Let $N_C(x)$ be as in theorem 2.4.

a) There exist absolute constant b_1 and b_2 such that $N_C\left[(b_1 D_{Q(\alpha)})^{b_2}\right] \ge 1$. (Cf. [Lagarias et al. 79].)

b) Assume that the Generalized Riemann Hypothesis (GRH) holds. Then $N_C \left[70 \log(D_{Q(\alpha)})^2\right] \geq 1$. (Cf. [Lagarias and Odlyzko 77] and [Oesterlé 79].) □

Theorem 2.5 and lemma 2.13 gives us a handle how to compute the complete Galois group of $\overline{f}$. After factoring $\overline{f}$ according to (2.1), we calculate all Frobenius elements for the prime ideals $P_1, \ldots, P_{r_p}$ over pZ, p unramified and p less then the bounds in theorem 2.5, by the procedure discussed above. We will eventually find n distinct permutations constituting $G_{\overline{f}}$. Until now we always assumed f, a resolvent for $Q(\alpha)$, to be irreducible. We shall now show how to establish this fact in non-deterministic polynomial time in n, the degree of f, and $\log(\,|f|\,)$. Though this result has been superseded by the results in [Cantor 81] and [Lenstra et al. 82], we believe that its proof reveals valuable new insight.

Algorithm 2.1: [Succinct Certificates for normal irreducible polynomials]

[For a given monic polynomial $f[x] \in Z[x]$ of degree n this algorithm verifies f to be normal and irreducible.]

(F) Guess a factorization

$$(x-\alpha)\,(x-g_2(\alpha)) \,\cdots\, (x-g_n(\alpha)) \equiv f(x) \bmod f(\alpha)$$

with $g_i(\alpha) \in 1/m\ Z[\alpha]$, m^2 a factor of Δ_f, $\deg(g_i) < n$, and the numerators of the coefficients of g_i absolutely smaller than the respective coefficient bound [Weinberger and Rothschild 76, Lemma 8.3].

(N) For all i with $2 \leq i \leq n$ verify that $g_i^{(k)}(\alpha) \bmod f(\alpha) \in \{\alpha, g_2(\alpha), \ldots, g_n(\alpha)\}$, for $2 \leq k \leq n$, and that

$$g_i^{(n)}(\alpha) \equiv \alpha \bmod f(\alpha)$$

where $g_i^{(n)}$ denotes $g_i(g_i \cdots (g_i))$ with $n-1$ substitutions.

(P) Guess a number $c \leq n$ and integers p_λ, $1 \leq \lambda \leq c$, not larger than $b_1\Delta_f^{b_2}$, b_1 and b_2 from theorem 2.5a) such that the following conditions are verifiable.

We show that all p_λ are prime numbers. For this step we use the prime certificates by [Pratt 75]. We also prove that all $f \bmod p_\lambda$ are squarefree.

For all p_λ we perform the following computation:

We factor f into irreducibles mod p_λ, i.e.

$$f_1 \cdots f_{r_\lambda} \equiv f \bmod p_\lambda.$$

The factors f_i can be tested for irreducibility by the distinct degree factorization [Knuth 81, Sec.4.6.2]. The proof of lemma 2.12 shows that all f_i have the same degree $s_\lambda = n/r_\lambda$. For each f_i we now construct the permutation $\sigma_{\lambda,i}$ corresponding to the Frobenius element determined by f_i according to the algorithm following the proof of lemma 2.12. Notice that the irreducible factor of f belonging to f_i is not needed in this construction. In order to compute $h_{k,i}^{p^t} \bmod (f_i)_{p_\lambda}$ we use binary exponentiation [Knuth 81, Sec.4.6.3].

(C) We verify that

$$\mathrm{card}(\{\sigma_{\lambda,i} \mid 1 \le \lambda \le c,\ 1 \le i \le r_\lambda\}) = n. \quad \square$$

Theorem 2.6: Algorithm 2.1 certifies f to be normal and irreducible in non-deterministic polynomial time in $\deg(f)$ and $\log(|f|)$.

Proof: It is easily established that algorithm 2.1 works in non-deterministic polynomial time. If the input polynomial f is indeed irreducible and normal we can find the described certificate by lemma 2.13 and theorem 2.5. The condition in step (N) is then satisfied because there exists a root α which can be mapped to any $g_i(\alpha)$ by some automorphism on $Q(\alpha)$, whose order is divisible by n.

Now let us assume that f is not irreducible. Steps (F) and (N) guarantee that f is at least normal, i.e. any root generates the complete splitting field of f. To prove this we need to express every root $g_j(\alpha)$ as a polynomial of $g_i(\alpha)$, e.g.

$$g_j(\alpha) = g_j(g_i^{(n-1)})(g_i(\alpha)).$$

Therefore, $f = f_1 \cdots f_t$ with $f_1, \ldots, f_t$ irreducible and having the same splitting field. As we process different primes p_λ our algorithm computes the Frobenius elements of G_f using the distinct resolvents $f_1, \ldots, f_t$. For each resolvent f_i we get at most the complete Galois group of G_f using the root enumeration of step (F) w.r.t. that resolvent. Hence, for each f_i we get at most $n/t-1$ permutations not equal to the identity. Therefore, for all resolvents we can get at most $n-t$ distinct permutations unequal to the identity. Since $t > 1$ our algorithm can produce at most $n-1$ distinct permutations in step (P) and the test in step (C) will always fail. □

Once we can certify normal polynomials irreducible it is an easy generalization to certify a polynomial $\bar{f}$ of degree $\bar{n}$ with small Galois group irreducible. Here small shall mean of polynomial cardinality in $\bar{n}$. We shall prove $\bar{f}$ irreducible by presenting its Galois group. First we guess a normal irreducible polynomial f whose splitting field contains that of $\bar{f}$. In order to show that $\log(|f|)$ can be chosen polynomial in $\bar{n}$ and $\log(|\bar{f}|)$ we consider the construction of a primitive element for the splitting field of $\bar{f}$ [van der Waerden 53, p.126]. If $\alpha_1, \ldots, \alpha_{\bar{n}}$ are the roots of $\bar{f}$ then we can select integers $b_2, \ldots, b_{\bar{n}}$ sufficiently small such that the roots of f are of the form $\alpha_{i_1} + b_2\alpha_{i_2} + \cdots + b_{\bar{n}}\alpha_{i_{\bar{n}}}$. We now factor $\bar{f}$ according to (2.1) and compute Frobenius elements as described above. We actually obtain the permutations corresponding to the automorphisms restricted to the smaller splitting field. If a transitive set of permutations is obtained our input polynomial $\bar{f}$ is certified to be irreducible.

In view of theorem 2.5b) one may consider to use our algorithm deterministically. We require the computation of a resolvent f as well as the factorization (2.1). For small Galois groups both tasks can be achieved in polynomial time using the algorithms in [Trager 76] in connection with the polynomial time factorization algorithm for univariate integer polynomials. However, once we have a resolvent f we can obtain its Galois group by first factoring

$$f(x) \equiv (x-\alpha)(x-g_2(\alpha)) \cdots (x-g_n(\alpha)) \bmod f(\alpha)$$

and then calculating the permutation σ_i corresponding to the automorphism $\alpha \to g_i(\alpha)$ by the substitutions

$$g_k(g_i(\alpha)) \bmod f(\alpha) = g_{k_i}(\alpha),\ 2 \le k \le n$$

implying that $\sigma_i(k) = k_i$. This approach avoids the necessity for the GRH though the algorithm using the identification of Frobenius elements may be more efficient, in practice.

It is interesting to ask how quickly one can obtain a generating set for the Galois group G_f of an arbitrary integer polynomial f. One can easily show that every finite group G can be generated by at most $\lfloor \log_2(\mathrm{card}(G)) \rfloor$ elements which implies that a small generating set S for G_f exists. Once such a set is known various questions about the group such as solvability can be answered in time polynomial in card(S). The open problem 4 in chapter 4 asks to compute S in time polynomial in deg(f) and log$(|f|)$.

We conclude this section with an example for which the calculations were carried out on [Macsyma 77].

Example 2.2:

$$f(x) = x^6 + 3x^5 + 6x^4 + 3x^3 + 9x + 9.$$

Then

$$f(x) = (x-\alpha)\,(x-g_2(\alpha))\,(x-g_3(\alpha))\,(x-g_4(\alpha))\,(x-g_5(\alpha))\,(x-g_6(\alpha))$$

with

$$g_2(\alpha) = \frac{1}{3}\,(\alpha^5 + 2\alpha^4 + 4\alpha^3 + 6),$$

$$g_3(\alpha) = \frac{1}{9}\,(4\alpha^5 + 6\alpha^4 + 12\alpha^3 - 12\alpha^2 + 9\alpha + 27),$$

$$g_4(\alpha) = -\frac{1}{9}\,(\alpha^5 - 12\alpha^2 + 9),$$

$$g_5(\alpha) = -\frac{1}{9}\,(\alpha^5 + 3\alpha^4 + 6\alpha^3 + 6\alpha^2 + 9\alpha + 18),$$

$$g_6(\alpha) = -\frac{1}{9}\,(5\alpha^5 + 9\alpha^4 + 18\alpha^3 - 6\alpha^2 + 9\alpha + 45).$$

$p = 5$:

$$f_p \equiv f_1\, f_2\, f_3 \bmod 5 \text{ with}$$
$$f_1 \equiv x^2 + 2,\quad f_2 \equiv x^2 - x + 1,\ \text{ and }\ f_3 \equiv x^2 - x + 2.$$

$h_{1,1}(\alpha_1) = \alpha_1,\ h_{1^{(5)},1}(\alpha_1) = -\ \alpha_1 = h_{3,1}(\alpha_1);$

$h_{2,1}(\alpha_1) = 2\alpha_1 - 2,\ h_{2^{(5)},1}(\alpha_1) = -\ 2\alpha_1 - 2 = h_{6,1}(\alpha_1);$

$h_{4,1}(\alpha_1) = -\ \alpha_1 - 2,\ h_{4^{(5)},1}(\alpha_1) = \alpha_1 - 2 = h_{5,1}(\alpha_1).$

$$\left[\frac{Q(\alpha)/Q}{P_1}\right] = (1\ \ 3)\ (2\ \ 6)\ (3\ \ 4).$$

$h_{1,2}(\alpha_2) = \alpha_2,\ h_{1^{(5)},2}(\alpha_2) = -\ \alpha_2 + 1 = h_{2,1}(\alpha_2);$

$h_{3,2}(\alpha_2) = -\ 2\alpha_2 - 1,\ h_{3^{(5)},2}(\alpha_2) = 2\alpha_2 + 2 = h_{4,2}(\alpha_2);$

$h_{5,2}(\alpha_2) = \alpha_2 + 2,\ h_{5^{(5)},2}(\alpha_2) = -\ \alpha_2 - 2 = h_{6,2}(\alpha_2).$

$$\left[\frac{Q(\alpha)/Q}{P_2}\right] = (1\ \ 2)\ (3\ \ 4)\ (5\ \ 6).$$

$h_{1,3}(\alpha_3) = \alpha_3,\ h_{1^{(5)},3}(\alpha_3) = -\ \alpha_3 + 1 = h_{5,3}(\alpha_3);$

$h_{2,3}(\alpha_3) = -\ 2\alpha_3 + 1,\ h_{2^{(5)},3}(\alpha_3) = 2\alpha_3 - 1 = h_{4,3}(\alpha_3);$

$h_{3,3}(\alpha_3) = 2\alpha_3 + 2,\ h_{3^{(5)},3}(\alpha_3) = -\alpha_3 + 1 = h_{5,3}(\alpha_3).$

$$\left[\frac{Q(\alpha)/Q}{P_3}\right] = (1\ \ 5)\ (2\ \ 4)\ (3\ \ 6).$$

$p = 7$:

$$f_p \equiv f_1\, f_2 \bmod 7 \text{ with}$$
$$f_1 \equiv x^3 + x^2 - 2x - 3 \ \text{ and } \ f_2 \equiv x^3 + 2x^2 - x - 3.$$

$h_{1,1}(\alpha_1) = \alpha_1,\ h_{1^{(7)},1}(\alpha_1) = 2\alpha_1 - 2 = h_{4,1}(\alpha_1),$

$\qquad h_{4^{(7)},1}(\alpha_1) = -\ 3\alpha_1 + 1 = h_{6,1}(\alpha_1);$

$h_{2,1}(\alpha_1) = 2\alpha_1,\ h_{2^{(7)},1}(\alpha_1) = -\ 3\alpha_1 + 3 = h_{5,1}(\alpha_1),$

$\qquad h_{5^{(7)},1}(\alpha_1) = \alpha_1 + 2 = h_{3,1}(\alpha_1).$

$$\left[\frac{Q(\alpha)/Q}{P_4}\right] = (1\ \ 4\ \ 6)\ (2\ \ 5\ \ 3).$$

$h_{1,2}(\alpha_2) = \alpha_2,\ h_{1^{(7)},2}(\alpha_2) = 2\alpha_2 + 3 = h_{6,2}(\alpha_2),$

$\qquad h_{6^{(7)},2}(\alpha_2) = -\ 3\alpha_2 + 2 = h_{4,2}(\alpha_2);$

$h_{2,2}(\alpha_2) = -3\alpha_2,\ h_{2^{(7)},2}(\alpha 2) = \alpha_2 - 2 = h_{3,2}(\alpha_2),$

$\qquad h_{3^{(7)},2}(\alpha_2) = 2\alpha_2 + 1 = h_{5,2}(\alpha_2).$

$$\left[\frac{Q(\alpha)/Q}{P_5}\right] = (1\ \ 6\ \ 4)\ (2\ \ 3\ \ 5).$$

$p = 31$:

$$f \equiv (x-14)\,(x-12)\,(x-9)\,(x-1)\,(x+2)\,(x+6) \bmod 31.$$

$$\left[\frac{Q(\alpha)/Q}{P_6}\right] = (1)\,(2)\,(3)\,(4)\,(5)\,(6).$$

The Galois group of f has 3 conjugate classes,

$$C_1 = \{(1)\,(2)\,(3)\,(4)\,(5)\,(6)\},$$
$$C_2 = \{(1\ \ 3)\,(2\ \ 6)\,(3\ \ 4),\ (1\ \ 2)\,(3\ \ 4)\,(5\ \ 6),\ (1\ \ 5)\,(2\ \ 4)\,(3\ \ 6)\}$$
$$C_3 = \{(1\ \ 4\ \ 6)\,(2\ \ 5\ \ 3),\ (1\ \ 6\ \ 4)\,(2\ \ 3\ \ 5)\}$$

The group is the symmetric group on 3 elements.

Note added on November 17, 1997: In Step (N) the computation of $g_i^{(n)}$ may cause exponential growth in the coefficient length. However, if f is normal, each $g_i^{(k)}$ must be found among the roots of f.

3. Polynomial Time Reductions from Multivariate to Univariate Integer Polynomial Factorization

3.1. Introduction

Both the classical Kronecker algorithm (algorithm 1.5) and the modern multivariate Hensel algorithm of section 1.4 solve the problem of factoring multivariate polynomials with integer coefficients by reduction to factorization of univariate polynomials and reconstruction of the multivariate factors from the univariate ones. However, as we will see in section 3.2, the running time of both methods suffers from the fact that, in rare cases, an exponential number of factor candidates obtained from the univariate factorization may have to be tested to determine the true factors. In this chapter we will present a new algorithm which does not require exponential time in its worst case. But before we can state our result precisely, we need to clarify what we mean by input size. We will assume that our input polynomials are densely encoded, that is all coefficients including zeros are listed. Hence, the size of a polynomial with v variables, given that the absolutely largest coefficient has l digits and the highest degree of any variable is n, is of order $O(ln^v)$.

Let v, the number of variables, be a fixed integer. We will show that the problem of determining all irreducible factors of v-variate polynomials is polynomial time Turing-reducible to completely factoring univariate polynomials. Recently, A. Lenstra, H. Lenstra, and L. Lovász [Lenstra et al. 82] have shown that factoring univariate polynomials is achievable in polynomial time. Their algorithm excludes the full factorization of a possible common integer content, which all coefficients of the polynomial to be factored might have. Therefore, our result implies the following theorem. Factoring an integer polynomial with a fixed number of variables into irreducibles, except for the constant factors, can be accomplished in time polynomial in the total degree and the size of its coefficients.

In [Kaltofen 82] we have already established a polynomial time reduction from multivariate to bivariate polynomial factorization. However, our new algorithm is less complex, though still exponential in the number of variables. On the other hand, the results in [Kaltofen 82] imply a polynomial time reduction for irreducibility testing, which our new algorithm does not provide. Therefore, we will include this older irreducibility test in section 3.5.

If one does not fix the number of variables, our definition of input size may not be appropriate since the input size then grows exponentially with the number of variables. In this case, sparsity concepts are definitely needed (cf. [Zippel 79]), but little is known about even the space complexity under these conditions. In section 4 open problem 1 corresponds to this question.

3.2. Exponential Cases for the Kronecker and Hensel Algorithms

We only consider bivariate polynomials though the constructions easily generalize. First, we discuss some exponential cases for the Kronecker algorithm 1.5. This algorithm transforms the bivariate polynomial $f(z,x)$ into $\overline{f}(y) = f(y^d, y)$, $d = \max(\deg_z(f), \deg_x(f))+1$. It requires time exponential in the degree of f in the case where f is irreducible, but $\overline{f}$ splits into linear factors. It is easy to construct such f's, as we do below, by working backward from $\overline{f}(y)$.

Example 3.1:

$$\begin{aligned}\overline{f}(y) &= (y-4)(y-3)(y-2)(y-1)(y+1)(y+2)(y+3)(y+4)\\ &= y^8 - 30y^6 + 273y^4 - 820y^2 + 576.\end{aligned}$$

Set $d=3$: $f_1(z,x) = z^2x^2-30z^2+273xz-820x^2+576$ which is irreducible. In step (K4), algorithm 1.4 needs 127 trial combinations to determine irreducibility.

Set $d=5$: $f_2(z,x) = x^3z-30xz+273x^4-820x^2+576$ which is irreducible because $\deg_z(f)=1$.

This condition can always be enforced by choosing d large enough and yields exponential cases of arbitrarily high degree.

Example 3.2. Let $n = (\prod_{i=2}^{k} p_i)-2$ with p_i the i-th prime number. Let $f_3(z,x) = x^n - z^2$, which is irreducible by lemma 2.7, since n is odd. We obtain $\overline{f}_3(y) = y^n(1-y^{n+2})$ where $1-y^{n+2}$ factors into 2^{k-1} cyclotomic polynomials. Since n is of order $O(2^{k\ \log(k)})$ the number of possible factor candidates cannot be polynomial in n.

The abundance of univariate factors disappears as soon as we choose a slightly different evaluation. For example,

$$f_1(3x^6,x) = 9x^8-270x^6+819x^4-820x^2+576$$

and

$$f_2(2x^5,x) = 2x^8-60x^6+273x^4-820x^2+576$$

are both irreducible. We have used such evaluations in [Kaltofen 82] and, as we will see by theorem 3.3, it is highly probable that substituting $2x^d$ or $3x^d$ for z in $f(z,x)$ already preserves the irreducibility of f. However, to prove that a multiplier of polynomial length definitely works is a much harder task, and we have only succeeded in showing this for the multivariate to bivariate reduction.

To construct an irreducible polynomial $f(y_1, \ldots, y_v, x)$ such that $f(0, \ldots, 0, x)$ has all linear factors is quite easy. The following example demonstrates the construction of a polynomial which has all linear factors for various evaluation points.

Example 3.3. Let $f(y,x)$ have $\deg_y(f) \le 3$ and

$$\begin{aligned} f(-1,x) &= (x-2)(x-1)(x+1)(x+2) = x^4-5x^2+4,\\ f(0,x) &= (x-1)x(x+1)(x+2) = x^4+2x^3-x^2-2x,\\ f(1,x) &= (x-2)(x-1)x(x+1) = x^4-2x^3-x^2+2x,\end{aligned}$$

and $f(2,x) = x^4+2$. By interpolation $f(y,x) \in Q[y,x]$ is determined uniquely, namely

$$\begin{aligned} f(y,x) = x^4 &+ (2y^3-3y^2-3y+2)x^3 + (\frac{5}{6}y^3-2y^2+\frac{7}{6}y-1)x^2\\ &+ (-2y^3+3y^2+3y-2)x - \frac{1}{3}y^3+2y^2-\frac{5}{3}y.\end{aligned}$$

We can also remove the rational denominators, namely

$$\begin{aligned} \bar{f}(y,x) &= 6^4 f(y,\frac{x}{6})\\ &= x^4 + (12y^3-18y^2-18y+12)x^3 + (30y^3-72y^2+42y-36)x^2\\ &\quad + (-432y^3+648y^2+648y-432)x - 432y^3+2592y^2-2160y.\end{aligned}$$

Since $f(2,x)$ is irreducible, so is $\bar{f}(y,x)$, but

$$\begin{aligned} \bar{f}(-1,x) &= (x-12)(x-6)(x+6)(x+12),\\ \bar{f}(0,x) &= (x-6)x(x+6)(x+12),\\ \bar{f}(1,x) &= (x-12)(x-6)x(x+6).\end{aligned}$$

The above construction obviously generalizes for arbitrarily high degrees but the number of unlucky evaluation points (i.e. those integers b for which $f(b,x)$ splits into linear factors) is bounded by the degree in y. The Hilbert Irreducibility Theorem states that for any irreducible polynomial

$f(y,x) \in Z[y,x]$ there exists an integer b such that $f(b,x)$ remains irreducible. It can be shown that the ratio of unlucky points to the size of the interval, from which the points are taken, tends to zero as the size of the interval goes to infinity (theorem 1.1). The reader is referred to Appendix B for a bibliography on the Hilbert Irreducibility Theorem. Unfortunately, we do not understand the distribution of unlucky evaluation points of small size. In this connection, we state open problem 2 in section 4.

3.3 The Reduction Algorithm

In this section, we shall discuss the proposed algorithm in detail. This algorithm uses ideas from an algorithm for univariate factorization proposed by [Zassenhaus 81]. After the algorithm we give a correctness proof using the theory of subresultants. Its complexity analysis is deferred to the next section. We wish to emphasize that the following version can be improved significantly by performing various steps at once. However, we are most interested in the theoretical result, namely that the algorithm works in polynomial time, and we have not yet investigated the conditions under which a highly tuned version of this algorithm might out-perform the Hensel algorithm in practice. (Cf. open problem 2 in section 4.)

Algorithm 3.1:

[Given $\overline{f}(z_1, \ldots, z_v, x) \in Z[z_1, \ldots, z_v, x]$, this algorithm constructs one irreducible factor $\overline{g}(z_1, \ldots, z_v, x) \in Z[z_1, \ldots, z_v, x]$ of $\overline{f}$.]

(I) [Precondition $\overline{f}$:]

IF cont($\overline{f}$) or pp($\overline{f}$) is univariate THEN factor it by a univariate factorization algorithm and return one irreducible factor, ELSE

(I1) Determine a primitive squarefree factor $\overline{s}(z_1, \ldots, z_v, x)$ of $\overline{f}$ by a multivariate version of algorithm 1.2 or Wang and Trager's algorithm as described in section 1.4.

(I2) [Make $\overline{s}$ monic in x:] $n \leftarrow \deg_x(\overline{s})$; $c(z_1, \ldots, z_v) \leftarrow \mathrm{ldcf}_x(\overline{s})$;

$$s(z_1, \ldots, z_v, x) \leftarrow c(z_1, \ldots, z_v)^{n-1}\, \overline{s}\left[z_1, \ldots, z_v, \frac{x}{c(z_1, \ldots, z_v)}\right].$$

[Notice that s is monic in x, an irreducible factor of which can be back-transformed to an irreducible factor of $\overline{s}$ (see step (E2)).]

(I3) [Find good integral evaluation points $w_1, \ldots, w_v$ such that $s(w_1, \ldots, w_v, x)$ is squarefree.]

FORALL integers w_i with $|w_i| \le \left\lceil \frac{(2n-1)}{2} \deg_{z_i}(s) \right\rceil$, $1 \le i \le v$, DO

Test whether $s(w_1, \ldots, w_v, x)$ is squarefree. If so, exit loop.

$f(y_1, \ldots, y_v, x) \leftarrow s(y_1+w_1, \ldots, y_v+w_v, x)$.

(F) [Factor $f(0, \ldots, 0, x)$:]

Compute an irreducible factor $t(x)$ of $f(0, \ldots, 0, x)$; $m \leftarrow \deg(t)$.

[Let β be a root of t. In the following, we will perform computations in $Q(\beta)$, whose elements are represented as polynomials in $Q[\beta]$ modulo t.]

(N) [Newton iteration. For purposes of later analysis, we emulate the Newton iteration by a Hensel lifting algorithm. We adopt the following vector notation: $\underset{\sim}{k} \equiv (k_1, \ldots, k_v)$, $\underset{\sim}{0} \equiv (0, \ldots, 0)$, $\underset{\sim}{y}^{\underset{\sim}{k}} \equiv y_1^{k_1} \cdots y_v^{k_v}$, $\underset{\sim}{k} \pm \underset{\sim}{k}' \equiv (k_1 \pm k_1', \ldots, k_v \pm k_v')$, $\underset{\sim}{k} \leq \underset{\sim}{k}'$ if, for all i, $k_i \leq k_i'$, $|\underset{\sim}{k}| \equiv k_1 + \cdots + k_v$, if $\underset{\sim}{k} \geq \underset{\sim}{0}$ and $-\infty$ otherwise.

Let J be the ideal in $Q(\beta)[y_1, \ldots, y_v]$ generated by $\{y_1, \ldots, y_v\}$. The goal is to construct

$$\alpha_j(y_1, \ldots, y_v) = \sum_{i=0}^{j} \sum_{|\underset{\sim}{k}|=i} a_{\underset{\sim}{k}}(\beta)\, \underset{\sim}{y}^{\underset{\sim}{k}}$$

for $j = 1, 2, \cdots$ such that

$$f(y_1, \ldots, y_v, \alpha_j(y_1, \ldots, y_v)) \equiv 0 \bmod J^{j+1}.]$$

Rewrite $f(y_1, \ldots, y_v, x) = \sum_{\underset{\sim}{k} \geq \underset{\sim}{0}} f_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}$

[Since f is monic and $\deg_x(f) = n$, $\deg(f_k) < n$ for $|\underset{\sim}{k}| \geq 1$.]

(N1) [Initialize for Hensel lifting:]

$g_{\underset{\sim}{0}}(x) \leftarrow x - \beta$; $h_{\underset{\sim}{0}}(x) \leftarrow f_{\underset{\sim}{0}}(x) / g_{\underset{\sim}{0}}(x) \in Q(\beta)[x]$.

(N2) [Bound for approximation:]

$d \leftarrow \deg_{y_1, \ldots, y_v}(f)$; $K \leftarrow d(2n-1)$.

FORALL $\underset{\sim}{k} \geq \underset{\sim}{0}$ with $1 \leq |\underset{\sim}{k}| \leq K$ DO steps (N3) and (N4). [The $\underset{\sim}{k}$ must be generated in an order such that $|\underset{\sim}{k}|$ is non-decreasing.]

(N3) IF $|\underset{\sim}{k}| = 1$ THEN $b_{\underset{\sim}{k}}(x) \leftarrow f_{\underset{\sim}{k}}(x)$ ELSE

$$b_{\underset{\sim}{k}}(x) \leftarrow f_{\underset{\sim}{k}}(x) - \sum_{\underset{\sim}{0}\le\underset{\sim}{s}\le\underset{\sim}{k},\ 1\le|\underset{\sim}{s}|\le|\underset{\sim}{k}|-1} g_{\underset{\sim}{s}}(x)h_{\underset{\sim}{k}-\underset{\sim}{s}}(x).$$

(N4) [Solve $g_{\underset{\sim}{0}}(x)h_{\underset{\sim}{k}}(x) + h_{\underset{\sim}{0}}(x)g_{\underset{\sim}{k}}(x) = b_{\underset{\sim}{k}}(x)$ with $g_{\underset{\sim}{k}}(x)$, $h_{\underset{\sim}{k}}(x) \in Q(\beta)[x]$, $\deg(g_{\underset{\sim}{k}}) = 0$, $\deg(h_{\underset{\sim}{k}}) = n-2$.]

$$a_{\underset{\sim}{k}} \leftarrow g_{\underset{\sim}{k}}(x) \leftarrow b_{\underset{\sim}{k}}(\beta) / f_{\underset{\sim}{0}}'(\beta);$$
$$h_{\underset{\sim}{k}}(x) \leftarrow \left[b_{\underset{\sim}{k}}(x) - h_{\underset{\sim}{0}}(x) g_{\underset{\sim}{k}}(x) \right] / g_{\underset{\sim}{0}}(x).$$

(N5) $\alpha_K \leftarrow \beta + \sum_{0\le|\underset{\sim}{k}|\le K} a_{\underset{\sim}{k}} \underset{\sim}{y}^{\underset{\sim}{k}}$

FOR $i \leftarrow 0, \ldots, n-1$ DO $\alpha_K^{(i)} \leftarrow \alpha_K^{\ i} \bmod J^{K+1}$.

(L) [Find minimal polynomial for α_K:]

FOR $I \leftarrow m, \ldots, n-1$ DO

Try to solve the equation

$$\alpha_K^{(I)} + \sum_{i=0}^{I-1} u_i(y_1, \ldots, y_v)\alpha_K^{(i)} \equiv 0 \bmod J^{K+1} \tag{3.1}$$

with undetermined coefficients for $u_i(y_1, \ldots, y_v) \in Q[y_1, \ldots, y_v]$ such that $\deg_{y_1, \ldots, y_v}(u_i) \le d$. [There are $I \begin{bmatrix} v+d \\ d \end{bmatrix}$ unknowns in $m \begin{bmatrix} v+K \\ K \end{bmatrix}$ linear equations. (Cf. lemma 3.4.)]

If there exists a solution, set

$g(y_1, \ldots, y_v, x) \leftarrow x^I + \sum_{i=0}^{I-1} u_i(y_1, \ldots, y_v)x^i$ and GOTO (E).

[We will prove that g is an irreducible factor of f.]

$g \leftarrow f$. [In this case f is irreducible.]

(E) [Recover non-monic factor $\bar{g}(z_1, \ldots, z_v, x)$:]

(E1) $g(z_1, \ldots, z_v, x) \leftarrow g(z_1 - w_1, \ldots, z_v - w_v, x)$.

(E2) $\bar{g}(z_1, \ldots, z_v, x) \leftarrow \mathrm{pp}_x(g(z_1, \ldots, z_v, c(z_1, \ldots, z_v)x))$. □

We shall now prove the correctness of the above algorithm. Obviously, if $g(y_1, \ldots, y_v, x)$

divides f then $g(z_1, \ldots, z_v, x)$ divides $s(z_1, \ldots, z_v, x)$. The proof for the correctness of the transformations in the steps (I2) and (E2) is quite easy and can be found in [Knuth 81, p.438, Exercise 18]. We first must show that step (I3) will yield good evaluation points.

Lemma 3.1: Let $s(z_1, \ldots, z_v, x) \in Z[z_1, \ldots, z_v, x]$ be monic of degree n in x and squarefree. Then there exist integers w_i with $|w_i| \le \left\lceil \frac{(2n-1)}{2} \deg_{z_i}(s) \right\rceil$, $1 \le i \le v$, such that $s(w_1, \ldots, w_v, x)$ is squarefree in $Z[x]$.

Proof: Let $n = \deg_x(s)$, $d_i = \deg_{z_i}(s)$ for $1 \le i \le v$. Since s is squarefree, its discriminant

$$\Delta(z_1, \ldots, z_v) = \mathrm{res}_x(s, \partial s/\partial x) \ne 0$$

[van der Waerden 53, p.86]. Since Δ is the given resultant, it follows that $\deg_{z_i}(\Delta) \le (2n-1)d_i$ for $1 \le i \le v$. If we write $\Delta(z_1, \ldots, z_v)$ as a polynomial in $Z[z_2, \ldots, z_v]$ with coefficients in $Z[z_1]$, not all these coefficients can be zero. Let $u(z_1)$ be one particular non-vanishing coefficient. Since $\deg(u) \le (2n-1)d_1$ there exists an integer w_1 with $|w_1| \le \left\lceil \frac{(2n-1)}{2} d_1 \right\rceil$ and $u(w_1) \ne 0$. Therefore $\Delta(w_1, z_2, \ldots, z_v) \ne 0$ and the lemma now follows by induction on the number of variables. □

Next, we must demonstrate that the steps (N3) and (N4) actually produce a root $\alpha_j(y_1, \ldots, y_v)$. Step (N1) sets up the basis for the Hensel lifting of the equation

$$g_{\underset{\sim}{0}}(x)\, h_{\underset{\sim}{0}}(x) \equiv f(y_1, \ldots, y_v, x) \bmod J.$$

If we have computed the sequences of polynomials $\{g_{\underset{\sim}{s}}(x)\}$ and $\{h_{\underset{\sim}{s}}(x)\}$, $\underset{\sim}{0} \le \underset{\sim}{k}$ $1 \le |\underset{\sim}{s}| \le |\underset{\sim}{k}| - 1$, then in step (N4) we want to compute $g_{\underset{\sim}{k}}(x)$ and $h_{\underset{\sim}{k}}(x) \in Q(\beta)[x]$ such that

$$\left[\sum_{\underset{\sim}{k} \ge \underset{\sim}{0}} g_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}\right] \left[\sum_{\underset{\sim}{k} \ge \underset{\sim}{0}} h_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}\right] = \sum_{\underset{\sim}{k} \ge \underset{\sim}{0}} f_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}$$

which implies that $h_{\underset{\sim}{k}}$ and $g_{\underset{\sim}{k}}$ must satisfy

$$g_{\underset{\sim}{0}}(x) h_{\underset{\sim}{k}}(x) + h_{\underset{\sim}{0}}(x) g_{\underset{\sim}{k}}(x) = b_{\underset{\sim}{k}}(x). \qquad (3.2)$$

Note that $g_{\underset{\sim}{0}}(\beta) = 0$ and $h_{\underset{\sim}{0}}(\beta) = f'_{\underset{\sim}{0}}(\beta)$. The second equation follows from the fact that if $\beta, \beta_2, \ldots, \beta_n$ are all roots of $f_{\underset{\sim}{0}}(x)$ then $h_{\underset{\sim}{0}}(x) = \prod_{i=2}^{n}(x - \beta_i)$ and $h_{\underset{\sim}{0}}(\beta) = \prod_{i=2}^{n}(\beta - \beta_i) = f'_{\underset{\sim}{0}}(\beta)$. Therefore the unique solution of (3.2) with $\deg(g_{\underset{\sim}{k}}) = 0$ is $a_k = b_{\underset{\sim}{k}}(\beta)/f'_{\underset{\sim}{0}}(\beta)$. If we now solve (3.2) for $h_{\underset{\sim}{k}}(x)$ we get

$$h_{\underset{\sim}{k}}(x) = (b_{\underset{\sim}{k}}(x) - h_{\underset{\sim}{0}}(x) g_{\underset{\sim}{k}}(x))/g_{\underset{\sim}{0}}(x)$$

which is a polynomial in x since $b_{\underset{\sim}{k}}(\beta) - h_{\underset{\sim}{0}}(\beta) a_{\underset{\sim}{k}} = 0$, and is of degree at most $n-2$. As we will see in section 3.4, the solution for (3.2) with $\deg(g_{\underset{\sim}{k}}) < \deg(g_{\underset{\sim}{0}})$ and $\deg(h_{\underset{\sim}{k}}) < \deg(h_{\underset{\sim}{0}})$ is uniquely determined by a linear system in n unknowns, whose coefficient matrix is the resultant of $g_{\underset{\sim}{0}}(x)$ and $h_{\underset{\sim}{0}}(x)$, which in our case happens to be equal to $f'_{\underset{\sim}{0}}(\beta)$.

We now know that

$$f(y_1, \ldots, y_v, \alpha_K(y_1, \ldots, y_v)) \equiv 0 \bmod J^{K+1}$$

because

$$\left[x - \sum_{0 \le |\underset{\sim}{k}| \le K} a_{\underset{\sim}{k}} \underset{\sim}{y}^{k} \right] \left[\sum_{0 \le |\underset{\sim}{k}| \le K} h_{\underset{\sim}{k}}(x) \underset{\sim}{y}^{k} \right] \equiv f(y_1, \ldots, y_v, x) \bmod J^{K+1}.$$

The polynomial $g(y_1, \ldots, y_v, x)$ is constructed in step (L) such that

$$g(y_1, \ldots, y_v, \alpha_K(y_1, \ldots, y_v)) \equiv 0 \bmod J^{K+1}$$

We will now prove that g must divide f. Our argument will show that if g does not divide f, then (3.1) has a solution for $I < \deg(g)$. One main condition for this to be true is that our approximation is of order K, as determined in step (N2). First, we must prove a simple lemma.

Lemma 3.2: Let $g(y_1, \ldots, y_v, x)$ divide $f(y_1, \ldots, y_v, x)$ in $Z[y_1, \ldots, y_v, x]$ and assume that $g(0, \ldots, 0, \beta) = 0$ in $Q(\beta)$. Then

$$g(y_1, \ldots, y_v, \alpha_j(y_1, \ldots, y_v)) \equiv 0 \bmod J^{j+1}$$

for all $j \ge 1$ with $\alpha_j(y_1, \ldots, y_v)$ as computed in steps (N3)-(N5).

Proof: The reason is simply that since $x - \alpha_j(y_1, \ldots, y_v)$ divides $f(y_1, \ldots, y_v, x) \bmod J^{j+1}$ and β is a root of single multiplicity $x - \alpha_j(y_1, \ldots, y_v)$ must also divide $g(y_1, \ldots, y_v, x) \bmod J^{j+1}$. This argument can be made formal but we shall provide a more indirect proof. Let p be the first index such that

$$g(y_1, \ldots, y_v, \alpha_p(y_1, \ldots, y_v)) \ne 0 \bmod J^{p+1}.$$

Because p is the first index

$$g(y_1, \ldots, y_v, \alpha_p(y_1, \ldots, y_v)) \equiv \sum_{|\underset{\sim}{k}|=p} \gamma_{\underset{\sim}{k}} \underset{\sim}{y}^{\underset{\sim}{k}} \bmod J^{p+1}$$

with at least one $\gamma_{\underset{\sim}{k}} \neq 0$. Let h be the cofactor of g, i.e. $f = g\,h$. Since β is a single root, $r = h(0, \ldots, 0, \beta) \neq 0$. Therefore

$$g(y_1, \ldots, y_v, \alpha_p(y_1, \ldots, y_v))\, h(y_1, \ldots, y_v, \alpha_p(y_1, \ldots, y_v))$$
$$\equiv \sum_{|\underset{\sim}{k}|=p} \gamma_{\underset{\sim}{k}} r\, \underset{\sim}{y}^{\underset{\sim}{k}} \neq 0 \bmod J^{p+1}$$

in contradiction to $\alpha_p(y_1, \ldots, y_v)$ being the p-th approximation of a root of f. □

Theorem 3.1: If step (L) finds a solution for (3.1) then $g(y_1, \ldots, y_v, x)$ derived from it is irreducible and divides $f(y_1, \ldots, y_v, x)$. Hence, the first solution for (3.1), if any, must be integral.

Proof: Let

$$D(y_1, \ldots, y_v, x) = \mathrm{GCD}(f(y_1, \ldots, y_v, x), g(y_1, \ldots, y_v, x))$$

and let $I = \deg_x(g)$, $j = \deg_x(D)$. By $S_j(y_1, \ldots, y_v, x)$ we denote the j-th subresultant of f and g as polynomials in x with coefficients in $Z[y_1, \ldots, y_v]$ (cf. [Brown, Traub 71, Section 5]). There exist polynomials $U_j(y_1, \ldots, y_v, x)$, $V_j(y_1, \ldots, y_v, x) \in Z[y_1, \ldots, y_v, x]$ such that $U_j f + V_j g = S_j$. Therefore, D divides S_j and since D is monic $\mathrm{ldcf}_x(S_j)\, D = S_j$. Since $\mathrm{ldcf}_x(S_j)$ is a subdeterminant of the resultant of f and g w.r.t. x, its total degree in $y_1, \ldots, y_v$ can be bounded by

$$\deg_{y_1, \ldots, y_v}(\mathrm{ldcf}_x(S_j)) \leq (I+n)d \leq K.$$

However,

$$(U_j f + V_j g)(y_1, \ldots, y_v, \alpha_K(y_1, \ldots, y_v))$$
$$\equiv \mathrm{ldcf}_x(S_j)\, D(y_1, \ldots, y_v, \alpha_K(y_1, \ldots, y_v)) \equiv 0 \bmod J^{K+1}$$

and therefore $D(0, \ldots, 0, \beta) = 0$. By lemma 3.2 it follows that $D(y_1, \ldots, y_v, \alpha_K(y_1, \ldots, y_v)) \equiv 0 \bmod J^{K+1}$ and from the minimality of I in (3.1), we conclude that g is irreducible and $g = D$ which divides f. □

This concludes the correctness proof for our algorithm. In the case that $v = 1$ the bound K of step (N2) can be improved to $\lceil d(2n-1)/m \rceil$ (cf. [Lenstra et al. 82, Proposition 2.7]). However, this improvement seems not to carry over for the general case, the reason being that $Q(y_1, \ldots, y_v)$ is not a

Euclidean domain. An example executed on [Macsyma 77] can be found in Appendix A.

3.4 Complexity Analysis of the Reduction Algorithm

The goal of this section is to prove that algorithm 3.1 takes, for a fixed number of variables v, polynomially many steps in deg($\overline{f}$) log($|\,\overline{f}\,|$), provided that we can factor f_0 in time polynomial in deg(f_0) log($|\,f_0|$). We wish to emphasize again that our main interest is in a polynomial time upper bound, but that we are not concerned about the best we could do by either fine tuning our algorithm or by determining sharper upper bounds. We also do not consider the influence of the underlying data structure used to represent the multivariate polynomials on our algorithm performance. In the analysis below we formulate the asymptotic complexity as a function in the total degree rather than the maximum degree of individual variables. Since the number of variables is fixed both notions for the degree are codominant.

Step (I): To obtain a squarefree factor $\overline{s}$ of $\overline{f}$, we make use of squarefree decomposition algorithms all of which employ polynomial GCD computations. All of the GCD algorithms such as the primitive remainder, subresultant or the modular algorithm [Brown 71], or the EZGCD algorithm [Moses and Yun 76], take for a fixed number of variables polynomially many steps in the maximum degree of the input polynomials and the size of their coefficients. That this time bound extends to the squarefree factorization process is shown, e.g., in [Yun 77]. Of course, deg($\overline{s}$) ≤ deg($\overline{f}$) in step (I1), and a good bound for $|\,\overline{s}\,|$ can be determined by the following lemma.

Lemma 3.3: Let $g_1, \ldots, g_m \in C[x_1, \ldots, x_v]$, let $f = g_1 \cdots g_m$ and let $n_j = \deg_{x_j}(f)$, $n = \sum_{j=1}^{v} n_j$. Then

$$\prod_{i=1}^{m} |\,g_i\,| \le 2^n\, |\,f\,| \prod_{j=1}^{v} \left[\frac{n_j+1}{2} \right]^{1/2} \le e^n\, |\,f\,|$$

with $e < \sqrt{6} \approx 2.44949$. (Cf. [Gel'fond 60, pp.135-139].) □

Therefore $|\,\overline{s}\,| \le e^{(v+1)\deg(\overline{f})}\, |\,\overline{f}\,|$. That the steps (I2) and (I3) take polynomial time is quite easily established. As a matter of fact, some of the GCD algorithms used for the squarefree decomposition of

$\overline{f}$ in step (I1) already provide the points $w_1, \ldots, w_v$ of step (I3) as a by-product. Step (I2) substantially, but yet polynomially, increases $\deg(s)$ and $\log(|s|)$. (E.g.

$$\deg(s) \le n\ \deg(\overline{s}) \quad \text{and} \quad |s| \le (\deg(\overline{s})+1)^{v\ n}\ |\overline{s}|^{n};$$

cf. lemma 3.4 and lemma 3.7.) Step (I3) again may increase $|f|$ but Taylor's formula of section 1.4, p.21, provides a quick polynomial size estimate. (E.g.

$$|f| \le v^{\deg(s)} \deg(s)^{v} \deg(s)^{2\deg(s)}\ |s|;$$

cf. lemma 3.1 and lemma 3.4.) We will not present the explicit polynomial time bound for step (I) because the following bounds for the steps (F), (N) and (L) clearly dominate the worst case complexity of step (I).

Step (F): As A. Lenstra, H. Lenstra and L. Lovász have recently shown, $t(x)$ can be computed in at most $O\left[\deg(f_{\underset{\sim}{0}})^{12} + \deg(f_{\underset{\sim}{0}})^{9} \log(|f_{\underset{\sim}{0}}|_2)^3\right]$ steps [Lenstra et al. 82].

Step (N): We first count the number of additions, subtractions and multiplications over $Q(\beta)$ (which we shall call ASM ops) needed for this step. Then we bound the absolute value of all elements of $Q(\beta)$ which appear as intermediate results. Finally, we bound the size of all computed rational numerators and denominators, and then we count the number of rational operations. The most difficult task will be to compute size bounds.

We can ignore the time it takes to retrieve the polynomials $f_{\underset{\sim}{k}}(x)$ as well as the execution time for step (N1). In order to count the number of times steps (N3) and (N4) are performed we need a lemma.

Lemma 3.4: There exist $\begin{bmatrix} v+j-1 \\ v-1 \end{bmatrix} < (j+1)^{v-1}$ different v-dimensional integer vectors $\underset{\sim}{k}$ with $\underset{\sim}{k} \ge \underset{\sim}{0}$ and $|\underset{\sim}{k}| = j$. The number of vectors with $|\underset{\sim}{k}| \le j$ is $\begin{bmatrix} v+j \\ v \end{bmatrix} < (j+1)^{v}$.

Proof: One chooses from v components j times allowing repetition. For $|\underset{\sim}{k}| \le j$ one introduces an additional dummy component. □

Therefore, steps (N3) and (N4) are executed less than $(K+1)^{\nu}$ times. Step (N3) takes $O(K^{\nu} n)$ ASM ops in $Q(\beta)$. Clearly this bound dominates the complexity of step (N4). Hence α_K can be calculated in $O(K^{2\nu} n)$ ASM ops.

We now proceed to compute an upper bound B_1 for all absolute values of the coefficients of α_K in $Q(\beta)$. We actually use a slightly more general approach which we will also use in section 3.5.

Lemma 3.5: Let $f(x) = g(x)\, h(x)$ be a non-trivial factorization of $f(x) \in Z[x]$, monic, squarefree of degree n in $C[x]$.

a) Then both $|g|$, $|h| \le 2^n |f|_2 \le \sqrt{n+1}\, 2^n |f|$ and if β is any root of f, $|\beta| \le 2|f|$.

b) If M is any $(n-1)$ by $(n-1)$ submatrix of the Sylvester matrix of f and g, then

$$|\det(M)| \le T(f) = \left[n\ 2^n\ |f| \right]^{n-1}.$$

c) The resultant of f and g is bounded by $1/S(f) \le |\mathrm{res}(g,h)| \le 2T(f)$ with

$$S(f) = (4|f|)^{(n-1)(n-2)/2}.$$

Proof: a) The bound for $|f|$ and $|g|$ is the Landau-Mignotte bound translated to maximum norms [Mignotte 74]. Assume $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ and let $\beta \in C$ with $|\beta| \ge 2|f|$. Then

$$|a_{n-1}\beta^{n-1} + \cdots + a_0| \le |f| \frac{|\beta|^n - 1}{|\beta| - 1} < |\beta|^n$$

because $|f| \ge 1$. Therefore $f(\beta) \ne 0$.

b) By part a), we know that each entry in the Sylvester matrix of f and g is bounded by $\sqrt{n+1}\, 2^n |f|$. Hadamard's determinant inequality [Knuth 81, Sec.4.6.1, Exercise 15] then gives the bound.

c) Let $g(x) = (x-\beta_1) \cdots (x-\beta_k)$ and $h(x) = (x-\beta_{k+1}) \cdots (x-\beta_n)$. Then

$$\mathrm{res}(g,\, h) = \prod_{i=1,\ldots,k;\, j=k+1,\ldots,n} (\beta_i - \beta_j)$$

and the discriminant of f, $\Delta = \prod_{i \ne j} (\beta_i - \beta_j)$, is an integer not equal 0 [van der Waerden 53, pp.87-89].

From a) we conclude that $|\beta_i - \beta_j| \le 4|f|$ for $1 \le i < j \le n$. Therefore

$$1 \le \sqrt{|\Delta|} = \left[\prod_{1\le i<j\le k} |\beta_i - \beta_j|\right] \, |\mathrm{res}(f,g)| \, \left[\prod_{k+1\le i<j\le n} |\beta_i - \beta_j|\right]$$
$$\le |\mathrm{res}(f,g)| \, (4|f|)^{(n-1)(n-2)/2}$$

because $k(k-1) + (n-k)(n-k-1) \le (n-1)(n-2)$ for $1 \le k \le n-1$. The upper bound follows from b) and the fact that f is monic. □

The following lemma estimates the size of a general version of the Catalan numbers.

Lemma 3.6: Let $d_{\underset{\sim}{k}} = 1$ for all v-dimensional vectors $\underset{\sim}{k}$ with $|\underset{\sim}{k}| = 1$ and let

$$d_{\underset{\sim}{k}} = \sum_{\underset{\sim}{0} \le \underset{\sim}{s} \le \underset{\sim}{k},\, 1 \le |\underset{\sim}{s}| \le |\underset{\sim}{k}|-1} d_{\underset{\sim}{s}} \, d_{\underset{\sim}{k}-\underset{\sim}{s}} \text{ for } \underset{\sim}{k} \ge \underset{\sim}{0}, \ |\underset{\sim}{k}| \ge 2.$$

Then

$$d_{\underset{\sim}{k}} = \frac{1}{|\underset{\sim}{k}|} \left[\begin{matrix} 2|\underset{\sim}{k}|-2 \\ |\underset{\sim}{k}|-1 \end{matrix}\right] \frac{|\underset{\sim}{k}|!}{k_1! \cdots k_v!} < (4v)^{|\underset{\sim}{k}|}.$$

Proof: Let $G(y_1, \ldots, y_v) = \sum_{|\underset{\sim}{k}| \ge 1} d_{\underset{\sim}{k}} \underset{\sim}{y}^{\underset{\sim}{k}}$ be the generating function for $d_{\underset{\sim}{k}}$ Then

$$G(y_1, \ldots, y_v)^2 = G(y_1, \ldots, y_v) - (y_1 + \cdots + y_v)$$

and thus

$$G(y_1, \ldots, y_v) = \frac{1}{2}\left[1 - \sqrt{1 - 4(y_1 + \cdots + y_v)}\right] = \sum_{i=1}^{\infty} \frac{1}{i} \left[\begin{matrix} 2i-2 \\ i-1 \end{matrix}\right] (y_1 + \cdots + y_v)^i$$

which yields our formula. Since $\dfrac{|\underset{\sim}{k}|!}{k_1! \cdots k_v!}$ is a multinomial coefficient, it is less than $v^{|\underset{\sim}{k}|}$. Similarly the given binomial coefficient is less than $2^{2|\underset{\sim}{k}|}$. □

We are now in the position to formulate and prove the main theorem on the coefficient growth for the Hensel lifting algorithm. This theorem also resolves the growth problem left open by [Kung and Traub 78] who considered the Newton iteration for the case that $v = 1$.

Theorem 3.2: Let $f(y_1, \ldots, y_v, x) \in Z[y_1, \ldots, y_v, x]$ be monic of degree n in x, such that $f_{\underset{\sim}{0}}(x) = f(0, \ldots, 0, x)$ is squarefree. Let β be an algebraic integer generating a subfield of the splitting field for $f_{\underset{\sim}{0}}$ By $Z[\beta]$ we denote the ring generated by Z and $\{\beta\}$ whose elements are polynomials in β

with integer coefficients of degree $[Q(\beta):Q]-1$. Let $g_{\underset{\sim}{0}}(x)\, h_{\underset{\sim}{0}}(x) = f_{\underset{\sim}{0}}(x)$ be a non-trivial factorization of $f_{\underset{\sim}{0}}$ in $(Z[\beta])[x]$. Then there exist unique polynomials $g_{\underset{\sim}{k}}(x)$, $h_{\underset{\sim}{k}}(x) \in Q(\beta)[x]$ with $\underset{\sim}{k} \ge \underset{\sim}{0}$ $|\underset{\sim}{k}| \ge 1$ and $\deg(g_{\underset{\sim}{k}}) < \deg(g_{\underset{\sim}{0}})$, $\deg(h_{\underset{\sim}{k}}) < \deg(h_{\underset{\sim}{0}})$ such that

$$f(y_1, \ldots, y_v, x) = \left[\sum_{\underset{\sim}{k}\ge\underset{\sim}{0}} g_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}\right] \left[\sum_{\underset{\sim}{k}\ge\underset{\sim}{0}} h_{\underset{\sim}{k}}(x)\, \underset{\sim}{y}^{\underset{\sim}{k}}\right].$$

Furthermore, let

$$\frac{1}{\operatorname{res}(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}})} = \frac{1}{R}\, r(\beta) \text{ with } R \in Z,\ r(\beta) \in Z[\beta],$$

and let $S(f_{\underset{\sim}{0}})$ and $T(f_{\underset{\sim}{0}})$ be as defined in lemma 3.5. Finally, let $N(f) = \max(n^2, n\, |f|)$, and let $d_{\underset{\sim}{k}}$ be as defined in lemma 3.6. Then for all $\underset{\sim}{k} \ge \underset{\sim}{0}$ with $|\underset{\sim}{k}| \ge 1$

$$R^{2|\underset{\sim}{k}|-1}\, g_{\underset{\sim}{k}}(x),\ R^{2|\underset{\sim}{k}|-1}\, h_{\underset{\sim}{k}}(x) \in (Z[\beta])[x]$$

and, independently of which root β of $f_{\underset{\sim}{0}}$ we choose,

$$|g_{\underset{\sim}{k}}|,\ |h_{\underset{\sim}{k}}| \le d_{\underset{\sim}{k}} \left[N(f)\, S(f_{\underset{\sim}{0}})\, T(f_{\underset{\sim}{0}})\right]^{2|\underset{\sim}{k}|-1}.$$

Proof: The existence and uniqueness of $g_{\underset{\sim}{k}}$ and $h_{\underset{\sim}{k}}$ follows from the fact that (3.2) has a unique solution with the given degree constraints, $b_{\underset{\sim}{k}}$ being computed as in step (N3). Now let $C_{\underset{\sim}{k}} = \max(|g_{\underset{\sim}{k}}|, |h_{\underset{\sim}{k}}|, |f|)$ and let $D_{\underset{\sim}{k}} = |b_{\underset{\sim}{k}}|$. Since $\deg(g_{\underset{\sim}{s}}) < \deg(g_{\underset{\sim}{0}})$ and $\deg(h_{\underset{\sim}{k}-\underset{\sim}{s}}) < \deg(h_{\underset{\sim}{0}})$ we conclude that

$$|g_{\underset{\sim}{s}} h_{\underset{\sim}{k}-\underset{\sim}{s}}| \le (n-1)\, |g_{\underset{\sim}{s}}|\, |h_{\underset{\sim}{k}-\underset{\sim}{s}}| \le (n-1)\, C_{\underset{\sim}{s}} C_{\underset{\sim}{k}-\underset{\sim}{s}}$$

By definition $C_{\underset{\sim}{s}} \ge |f|$ and thus we obtain from (N3)

$$D_{\underset{\sim}{k}} \le n \sum_{\underset{\sim}{0} \le \underset{\sim}{s} \le \underset{\sim}{k},\ 1 \le |\underset{\sim}{s}| \le |\underset{\sim}{k}|-1} C_{\underset{\sim}{s}} C_{\underset{\sim}{k}-\underset{\sim}{s}} \tag{A}$$

If we solve (3.2) by undetermined coefficients for $g_{\underset{\sim}{k}}$ and $h_{\underset{\sim}{k}}$ we encounter the Sylvester matrix of $g_{\underset{\sim}{0}}$ and $h_{\underset{\sim}{0}}$, $A(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}})$, as the coefficient matrix, namely

$$(\vec{h}_{\underset{\sim}{k}}\ \vec{g}_{\underset{\sim}{k}})\, A(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}}) = \vec{b}_{\underset{\sim}{k}} \tag{B}$$

By $\vec{p}$ we mean the coefficient vector $(p_m, \ldots, p_0)$ of the polynomial $p(x) = p_m x^m + \cdots + p_0$. Using Cramer's rule for (B) and the fact that

$$\frac{1}{|\det(A(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}}))|} = \frac{1}{|\operatorname{res}(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}})|} \le S(f_{\underset{\sim}{0}})$$

(by lemma 3.5b), we get the estimate

$$C_{\underset{\sim}{k}} \le \max(\,|\,f\,|\,,\, nD_{\underset{\sim}{k}} S(f_{\underset{\sim}{0}})\, T(f_{\underset{\sim}{0}})). \tag{C}$$

Our claims now follow by induction on $|\underset{\sim}{k}|$.

Case $|\underset{\sim}{k}| = 1$: Since $b_{\underset{\sim}{k}} = f_{\underset{\sim}{k}} \in Z[x]$, Cramer's rule applied to (B) yields $R\ g_{\underset{\sim}{k}}\ R\ h_{\underset{\sim}{k}} \in (Z[\beta])[x]$. (Notice that β is an algebraic integer.) Also $D_{\underset{\sim}{k}} \le |\,f\,|$ and hence by (C)

$$C_{\underset{\sim}{k}} \le \max(\,|\,f\,|\,, n\,|\,f\,|\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}})) \le d_{\underset{\sim}{k}} N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}}).$$

Case $|\underset{\sim}{k}| > 1$: By hypothesis and from (N3) we obtain $R^{2\,|\underset{\sim}{k}|-2}\ b_{\underset{\sim}{k}} \in (Z[\beta])[x]$. Cramer's rule applied to (B) then yields $R^{2\,|\underset{\sim}{k}|-1}\ g_{\underset{\sim}{k}}\ R^{2\,|\underset{\sim}{k}|-1}\ h_{\underset{\sim}{k}} \in (Z[\beta])[x]$. From (A) with the hypothesis we also get

$$\begin{aligned}
b_{\underset{\sim}{k}} &\le n \sum_{\underset{\sim}{0}\le\underset{\sim}{s}\le\underset{\sim}{k},\ 1\le|\underset{\sim}{s}|\le|\underset{\sim}{k}|-1} C_{\underset{\sim}{s}}\, C_{\underset{\sim}{k}-\underset{\sim}{s}} \\
&\le n\ (N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}}))^{2\,|\underset{\sim}{k}|-2} \left[\sum_{\underset{\sim}{0}\le\underset{\sim}{s}\le\underset{\sim}{k},\ 1\le|\underset{\sim}{s}|\le|\underset{\sim}{k}|-1} d_{\underset{\sim}{s}}\, d_{\underset{\sim}{k}-\underset{\sim}{s}}\right] \\
&= n\ (N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}}))^{2\,|\underset{\sim}{k}|-2} d_{\underset{\sim}{k}}
\end{aligned}$$

By (C) we finally obtain

$$\begin{aligned}
C_{\underset{\sim}{k}} &\le \max(\,|\,f\,|\,,\, n\ D_{\underset{\sim}{k}} S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}})) \\
&\le d_{\underset{\sim}{k}} \frac{n^2}{N(f)}\ (N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}}))^{2\,|\underset{\sim}{k}|-1} \\
&\le d_{\underset{\sim}{k}}(N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}}))^{2\,|\underset{\sim}{k}|-1}. \quad \square
\end{aligned}$$

Since the polynomials $g_{\underset{\sim}{k}}$ and $h_{\underset{\sim}{k}}$ are unique we can conclude from theorem 3.2

$$|\,a_{\underset{\sim}{k}}| \le d_{\underset{\sim}{k}} \Big[N(f)\ S(f_{\underset{\sim}{0}})\ T(f_{\underset{\sim}{0}})\Big]^{2\,|\underset{\sim}{k}|-1} \text{ for } 1 \le |\underset{\sim}{k}| \le K.$$

From the lemmas 3.5 and 3.6 we obtain

$$\begin{aligned}
|\,\alpha_K\,| \le B_1(f) &= (4v)^K \Big[n^2\,|\,f\,|\ (4\,|\,f\,|\,)^{n^2\!/2}\ 2^{n^2}(n\,|\,f\,|\,)^n\Big]^{2K-1} \\
&< (4v)^K\ (2n\,|\,f\,|\,)^{2Kn^2},
\end{aligned} \tag{3.3}$$

assuming that $n \ge 4$. Obviously, $\log(B_1(f))$ is polynomial in $\deg(f)$ and $\log(\,|\,f\,|\,)$.

We now demonstrate for the polynomials $g_{\underset{\sim}{0}} = x - \beta$ and $h_{\underset{\sim}{0}}$ as computed in step (N1), that

$$\frac{R}{\mathrm{res}(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}})} \in Z[\beta], \text{ with } R = \mathrm{res}(t(x), f_{\underset{\sim}{0}}'(x)),$$

where t is the minimal polynomial of β. Let $\beta_2, \ldots, \beta_n$ be the roots of $h_{\underset{\sim}{0}}$ Then

$$\operatorname{res}(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}}) = \prod_{i=2}^{n} (\beta - \beta_i) = f'_{\underset{\sim}{0}}(\beta).$$

There exist polynomials $A(x)$ and $B(x) \in Z[x]$ such that $At + Bf'_{\underset{\sim}{0}} = R$. Thus $R/f'_{\underset{\sim}{0}}(\beta) = B(\beta) \in Z[\beta]$, which we wanted to show. Now let $m = \deg(t)$. By lemma 3.5a) $|\, t \,| \le \sqrt{n+1}\, 2^n\, |\, f_{\underset{\sim}{0}} |$, and using Hadamard's determinant inequality for the resultant $\operatorname{res}(t, f'_{\underset{\sim}{0}}(x))$ we obtain

$$\begin{aligned} |\, R \,| &\le \left[\sqrt{(m+1)(n+1)}\; 2^n\; |\, f_{\underset{\sim}{0}} | \right]^{n-1} \left[\sqrt{n}\;\; n\; |\, f_{\underset{\sim}{0}} | \right]^{m} \\ &< \left[(n+1)\; 2^n\;\; |\, f_{\underset{\sim}{0}} | \right]^{m+n} < (2n\; |\, f \,|)^{n^{3/2}}, \end{aligned} \tag{3.4}$$

for $n \ge 4$. Again, we note that $\log(|\, R \,|)$ is bounded by a polynomial in $\deg(f)$ and $\log(|\, f \,|)$.

From theorem 3.2 we can also conclude that

$$R^{2\,|\underset{\sim}{k}|-1} a_{\underset{\sim}{k}} \in Z[\beta] \text{ for } 1 \le |\underset{\sim}{k}| \le K. \tag{3.5}$$

We now extend our estimates to the powers of α_K mod J^{K+1} as well as count the ASM ops for step (N5).

Lemma 3.7: Let $a_{\underset{\sim}{0}} = \beta$ and let $\alpha_K^{(i)} = \sum_{0 \le |\underset{\sim}{k}| \le K} a_{\underset{\sim}{k}}^{(i)}\, \underset{\sim}{y}^{\underset{\sim}{k}}$ for $2 \le i \le n-1$, then

$$|\, a_{\underset{\sim}{k}}^{(i)} \,| \le (K+1)^{v(i-1)} B_1(f)^i \text{ and } R^{2\,|\underset{\sim}{k}|-1} a_{\underset{\sim}{k}}^{(i)} \in Z[\beta],$$

with R as defined above. All $\alpha_K^{(i)}$, $2 \le i \le n-1$, can be computed in $O(K^{2v} n)$ ASM ops.

Proof: It is easy to show that

$$a_{\underset{\sim}{k}}^{(i+1)} = \sum_{0 \le \underset{\sim}{s} \le \underset{\sim}{k}} a_{\underset{\sim}{s}}^{(i)}\; a_{\underset{\sim}{k}-\underset{\sim}{s}},\; 0 \le |\underset{\sim}{k}| \le K,\; i \ge 1,$$

where there are less than $(|\underset{\sim}{k}|+1)^v \le (K+1)^v$ terms under the right hand sum. The lemma now follows by induction on i. □

Therefore we get from (3.3) for all $0 \le i \le K$ and for $n \ge 4$

$$|\, \alpha_K^{(i)} \,| \le B_2(f) = ((K+1)^v B_1(f))^{n-1} < 2^{3v\; nK}\; (2n\; |\, f \,|)^{2K(n^3-n^2)}. \tag{3.6}$$

Lemma 3.7 also establishes that the common denominator of any rational coefficient computed throughout step (N) is R^{2K-1}. We are now in the position of estimating the size of any numerator of the rational coefficients of ${\alpha_K}^{(i)}$, $1 \le i \le n-1$. To do this, we shall state a well-known lemma.

Lemma 3.8: Let β be any root of $t(x) \in Z[x]$, monic, squarefree of degree m. Let A be a real upper bound for the absolute value of any conjugate of β. Assume that

$$\left| \sum_{i=0}^{m-1} c_i \beta^i \right| \le C \text{ with } c_i \in Z.$$

Furthermore, let D be the absolute value of the discriminant of t. Then

$$| c_i | \le \frac{C\ m!\ A^{m(m-1)/2}}{\sqrt{D}},\ 0 \le i < m.$$

(Cf. [Weinberger and Rothschild 76, Lemma 8.3].) □

In our case, we can choose $A = 2\,|f_{\underset{\sim}{0}}|$ by lemma 3.5a), $C = B_2(f)\, R^{2K-1}$, and $D \ge 1$. Therefore, if we bring all rationals computed in step (N) to the common denominator R^{2K-1}, we have shown that the absolute values of the numerators are bounded by

$$B_3(f,m) = R^{2K-1}\ B_2(f)\ m!\ (2\,|f_{\underset{\sim}{0}}|)^{m(m-1)/2} < 2^{3vnK}\ (2n\ |f|)^{3Kn^3}, \tag{3.7}$$

using (3.4), (3.6) and $n \ge 4$. Though this bound is quite large, it is of length polynomial in $\deg(f)$ and $\log(|f|)$. This bound also implies, that all ASM ops are computable in time polynomial in $\deg(f)$ and $\log(|f|)$. Addition and subtraction in $Q(\beta)$ means adding or subtracting the numerators of polynomials in $Q[\beta]$ of degree $m-1$, after eventually multiplying them with a power of R to produce a common denominator. Multiplication in $Q(\beta)$ is multiplication of $m-1$ degree polynomials in $Q[\beta]$ followed by a remainder computation w.r.t. $t(\beta)$. Again a common denominator can be extracted a priori. Any AMS op takes at most $O(m^2)$ integral operations.

Step (L): Let $u_i(y_1, \ldots, y_v) = \sum_{0 \le |\underset{\sim}{k}| \le d} u_{i,\underset{\sim}{k}} \underset{\sim}{y}^{\underset{\sim}{k}}$ and let

$${\alpha_K}^{(i)} = \sum_{0 \le |\underset{\sim}{k}| \le K} \left[\sum_{j=0}^{m-1} a_{\underset{\sim}{k}j}^{(i)}\ \beta^j \right] \underset{\sim}{y}^{\underset{\sim}{k}}$$

Then (3.1) can be written as

$$a_{\underset{\sim}{k}j}^{(I)} + \sum_{i=0}^{I-1} \sum_{0 \le |\underset{\sim}{s}| \le d} a_{\underset{\sim}{k}-\underset{\sim}{s}j}^{(i)}\ u_{i,\underset{\sim}{s}} = 0 \tag{3.8}$$

for $0 \le |\underset{\sim}{k}| \le K$, $j = 0, \ldots, m-1$. By lemma 3.4, it follows that (3.8) consists of

$p = m \begin{bmatrix} v+K \\ K \end{bmatrix} < m(K+1)^v$ equations in $q = I \begin{bmatrix} v+d \\ d \end{bmatrix} < (n-1)(d+1)^v$ unknowns. Applying Gaussian elimination to (3.8) takes $O(pq^2)$ rational operations. It is easy to show that this is the dominant operation count, which, expressed in input terms, is

$$O(m \; n^{v+3} \, d^{3v}). \tag{3.9}$$

From the previous analysis, we know that all $a_{\underset{\sim}{k},j}^{(i)}$ can be brought to the common denominator R^{2K-1} and their numerators, $\mathrm{num}(a_{\underset{\sim}{k},j}^{(i)})$, then satisfy $| \mathrm{num}(a_{\underset{\sim}{k},j}^{(i)}) | \le B_3(f, m)$. As can be shown with little effort, all intermediate rationals computed during the Gaussian elimination process are fractions of subdeterminants of the coefficient matrix for (3.8) extended by the vector of constants [Gantmacher 58, Chap.2]. It is not necessary to calculate the GCD of the numerator and denominator of a newly obtained rational since, as can also be shown, the denominator of the row used for the elimination in subsequent rows divides the numerators and denominators in these rows after the elimination step. Thus Hadamard's determinant inequality produces a bound for the size of any intermediately computed integer which is polynomial in $\deg(f) \log(| f |)$. E.g. one such bound is

$$B_4(f, m) = \left[\sqrt{q} \;\; B_3(f, m) \right]^q$$

whose logarithm is by (3.7) of order

$$\log(B_4(f, m)) = O(d^{v+2} \, vn^4 \log(4n \, | f |)), \tag{3.10}$$

assuming that $d \ge n$. Hence, step (L) also takes at most polynomial time in $\deg(f)$ and $\log(| f |)$. Notice that (3.9) and (3.10) give a very crude bound for the complexity of the steps (N) and (L). Since we know that any solution of (3.8) must be integral of quite a small size, due to lemma 3.3, a Chinese remaindering algorithm could be used to solve (3.8) [McClellan 73] and we believe that this approach will be much more efficient, in practice.

Step (E): Step (E1) is the counterpart of the transformation of step (I3). Step (E2) is similar to step (I2), but also involves a content computation. Both steps can obviously be performed in time polynomial in $\deg(g)$ and $\log(| g |)$.

3.5. Multivariate Irreducibility Testing

As we have seen in section 3.3, in order to establish the irreducibility of the polynomial f by algorithm 3.1 we need to factor f_0. Reducibility of f_0 does, of course, not imply reducibility of f. The following theorem partially fills this gap by constructing from a polynomial $f(y_1, \ldots, y_v, x)$, monic in x such that $f(0, \ldots, 0, x)$ is squarefree, a polynomial $g(y_1, x)$ in time polynomial in deg(f) and log($|f|$), such that g is irreducible if and only if f is irreducible. Unfortunately, our approach does not allow us to eliminate y_1. We could include this as an open problem, but in view of the polynomial time algorithm for univariate factorization a solution appears to be not so significant.

In the next theorem we will need the algebraic closure of polynomial domains and we shall introduce the theory now. Let K be a field of characteristic 0, $\overline{K}$ its algebraic closure. By $K(t)^*$ we denote the fractional power series domain in t over $\overline{K}$, any element of which is of the form $\sum_{i \geq -k} a_i \ t^{i/q}$ with $k, q \in Z$, $q \geq 1$ and $a_i \in \overline{K}$ for $i \geq -k$. It is a classical result in complex analysis that every algebraic function admits a fractional power series expansion which converges in a neighborhood of zero. We state this theorem in its algebraic version as a lemma.

Lemma 3.9 *(Puiseux's Theorem):* $\overline{K[t]}$, the algebraic closure of $K[t]$, can be embedded into $K(t)^*$. (Cf. [van der Waerden 39, pp. 50-54] or [Walker 50, pp. 97-106].) □

We write $K(t_1, \ldots, t_v)^*$ for $(\cdots(K(t_1)^*)(t_2)^* \cdots)(t_v)^*$ and notice that $K(t_1, \ldots, t_v)^*$ contains the algebraic closure of $K[t_1, \ldots, t_v]$.

Theorem 3.3: Let $f(y_1, \ldots, y_v, x) \in Z[y_1, \ldots, y_v, x]$ be monic of degree n in x such that $f_0 = f(0, \ldots, 0, x)$ is squarefree. Let $T(f_0)$ be as in lemma 3.5, and let $N(f)$ be as in theorem 3.2. Furthermore, assume that $f(y_1, \ldots, y_v, x)$ is irreducible. Let $d = \deg_{y_1, y_2}(f)$ and $M = \deg_{y_3, \ldots, y_v}(f)$. Then for any integer c with

$$|c| \geq B_5(f) = 2\,(4v)^{2d+M}(2\,N(f)\,T(f_0)^2)^{4d+2M-1}$$

$f(y_1, cy_1, y_3, \ldots, y_v, x)$ is irreducible in $Z[y_1, y_3, \ldots, y_v, x]$.

Proof: By lemma 3.9 and the subsequent remark the polynomial $g_c(y_1,y_3, \ldots ,x) = f(y_1,cy_1,y_3, \ldots ,x)$ has n roots in $Q(y_1,y_3, \ldots ,y_v)^*$ and $f(y_1,y_2, \ldots ,y_v,x)$ has n roots in $Q(y_1,y_2, \ldots ,y_v)^*$. Each of the roots of g_c corresponds to a root of f with $y_2 = cy_1$. Hence each factor of $g_c(y_1, y_3, \ldots , x) \in Q(y_1, y_3, \ldots , y_v)^*[x]$ corresponds to a factor of $f(y_1, y_2, \ldots , x) \in Q(y_1, y_2, \ldots , y_v)^*[x]$ with $y_2 = cy_1$.[#] We will show that for an integer c of the stated size no factor derived from f in such a way can be an integral polynomial dividing g_c. For simplicity we write $\utilde{y}$ for the variables y_1, y_2 and $\utilde{x}$ for the variables $y_3, \ldots , y_v$, x and again use our vector notation but now all vectors have either 2 or $v-1$ components. Our plan is the following: We first show that any candidate factor $h(y_1, y_2, \ldots , x)$ of $f(y_1, y_2, \ldots , x) \in Q(y_1, \ldots , y_v)^*[x]$ contains at least one monomial $b_{\utilde{m}\utilde{p}}\utilde{y}^{\utilde{p}}\utilde{x}^{\utilde{m}}$ with $b_{\utilde{m}\utilde{p}} \neq 0$ and $d < |\utilde{p}| \leq 2d$. From it we get a monomial $t(c)\, y_1^{|\utilde{p}|}\utilde{x}^{\utilde{m}}$ in $h(y_1,cy_1,\utilde{x})$ where $t(c)$ is a non-zero polynomial in c. Since $\deg_{y_1}(g_c) \leq d$, no polynomial factor of g_c has a degree in y_1 higher than d. By choosing c larger than the absolute value of any root of $t(c)$ we force

$$\deg_{y_1}(h(y_1,cy_1, \ldots ,x)) \geq |\utilde{p}| > d$$

and hence $h(y_1,cy_1, \ldots ,x)$ cannot be a polynomial dividing $g_c(y_1,y_3, \ldots ,x)$. Let

$$h(y_1, \ldots ,y_v,x) = \sum_{\utilde{k}\geq\utilde{0}} \sum_{\utilde{i}\geq\utilde{0}} b_{\utilde{k}\utilde{i}}\utilde{y}^{\utilde{i}}\utilde{x}^{\utilde{k}}$$

be the product of a subset of the linear factors of $f(x)$ with the constant coefficients in $Q(y_1, \ldots ,y_v)^*$ and let

$$\overline{h}(y_1, \ldots ,y_v,x) = \sum_{\utilde{k}\geq\utilde{0}} \sum_{\utilde{i}\geq\utilde{0}} \overline{b}_{\utilde{k}\utilde{i}}\utilde{y}^{\utilde{i}}\utilde{x}^{\utilde{k}}$$

be its cofactor, i.e. $f = h\overline{h}$. We first can assume that

$$h(0,0,y_3, \ldots ,x) = \sum_{\utilde{k}\geq\utilde{0}} b_{\utilde{k}\utilde{0}}\utilde{x}^{\utilde{k}} \in Z[y_3, \ldots ,x].$$

Otherwise $h(y_1,cy_1,y_3, \ldots ,x)$ could not be an integer polynomial for any choice of c. Similarly we can also assume that $b_{\utilde{k}\utilde{i}}$, $\overline{b}_{\utilde{k}\utilde{i}}$ are zero for all $\utilde{i}$ and $\utilde{k}$ with $|\utilde{k}^1| > M$, where $|\utilde{k}^1|$ is the vector derived from $\utilde{k}$ by removing its last component.[@] Otherwise, even if $h(y_1,cy_1,y_3, \ldots ,x)$ were an integer

[#] Prof. H. Lenstra points out that, by using the uniqueness of the Hensel lifting procedure (theorem 3.2), the need for Puiseux theorem can be completely avoided to establish this correspondence.

[@] This argument if false. See SIAM J. Comput., vol. 14, no. 2, p. 485 (1995) for the correct argument (note added November 17, 1997).

polynomial, it could not divide g_c because its or its cofactor's degree in $y_3, \ldots, y_v$ were to high.

Now there must exist at least one $b_{\underset{\sim}{k}\underset{\sim}{i}}$ or $\bar{b}_{\underset{\sim}{k}\underset{\sim}{i}}$ with

$$d < |\underset{\sim}{i}| \le 2d \text{ and } \left[b_{\underset{\sim}{k}\underset{\sim}{i}} \ne 0 \textit{ or } \bar{b}_{\underset{\sim}{k}\underset{\sim}{i}} \ne 0 \right].$$

To see this, assume the contrary. Then

$$\left[\sum_{0 \le |\underset{\sim}{k}| \le M, 0 \le |\underset{\sim}{i}| \le d} b_{\underset{\sim}{k}\underset{\sim}{i}} \underset{\sim}{y}^{\underset{\sim}{i}} \underset{\sim}{x}^{\underset{\sim}{k}} \right] \left[\sum_{0 \le |\underset{\sim}{k}| \le M, 0 \le |\underset{\sim}{i}| \le d} \bar{b}_{\underset{\sim}{k}\underset{\sim}{i}} \underset{\sim}{y}^{\underset{\sim}{i}} \underset{\sim}{x}^{\underset{\sim}{k}} \right] = f(y_1, \ldots, y_v, x)$$

since no term $r \underset{\sim}{y}^{\underset{\sim}{i}} \underset{\sim}{x}^{\underset{\sim}{k}}$, r a nonzero rational, with $d < |\underset{\sim}{i}| \le 2d$ in the left product could be canceled by higher terms in the product of the complete expansion of h and $\bar{h}$. (Notice that f does not contain a monomial in y_1, y_2 of degree larger than d.) But this contradicts the fact that f is irreducible. Without loss of generality we now can assume the existence of $\underset{\sim}{m}\underset{\sim}{p}$ such that

$$b_{\underset{\sim}{m}\underset{\sim}{p}} \ne 0 \text{ with } 0 \le |\bar{\underset{\sim}{m}}^1| \le M \text{ and } d < |\underset{\sim}{p}| \le 2d.$$

We collect all non-zero $b_{\underset{\sim}{m}\underset{\sim}{j}}$ with $j_1 + j_2 = p_1 + p_2$. Let $\underset{\sim}{q}$ be that of all vectors $\underset{\sim}{j}$ whose second component is largest. Then the following inequalities hold

$$q_2 \le q_1 + q_2 \le 2d \text{ and } d < j_1 + j_2 \le 2d. \tag{A}$$

We now consider the coefficient $t(c)$ of $y_1^{p_1+p_2} \underset{\sim}{x}^{\underset{\sim}{m}}$ in $g(y_1, cy_1, y_3, \ldots, x)$, namely

$$t(c) = b_{\underset{\sim}{m}\underset{\sim}{q}} c^{q_2} + \cdots + b_{\underset{\sim}{m}\underset{\sim}{j}} c^{j_2} + \cdots + b_{\underset{\sim}{m}(p_1+p_2,0)}.$$

By lemma 3.5a) the absolute value of any root of $t(c)$ is bounded by $2\,|t/\mathrm{ldcf}(t)|$. We now apply theorem 3.2 with $\beta = 1$, $g_{\underset{\sim}{0}}(x) = h(0, \ldots, 0, x)$ and $h_{\underset{\sim}{0}}(x) = \bar{h}(0, \ldots, 0, x)$. Notice, that if $R = \mathrm{res}(g_{\underset{\sim}{0}}, h_{\underset{\sim}{0}})$ then $1/|R| \le 1$ and hence we can set $S(f_{\underset{\sim}{0}}) = 1$. Therefore $R^{2(|\bar{\underset{\sim}{m}}^1| + |\underset{\sim}{q}|)-1} b_{\underset{\sim}{m}\underset{\sim}{q}} \in Z$ and $|\bar{\underset{\sim}{m}}^1| + |\underset{\sim}{q}| \le M+2d$ by (A). Since $b_{\underset{\sim}{m}\underset{\sim}{q}} \ne 0$ it follows that

$$\left| \frac{1}{b_{\underset{\sim}{m}\underset{\sim}{q}}} \right| \le R^{2M+4d-1} \le (2\,T(f_{\underset{\sim}{0}}))^{2M+4d-1},$$

the last inequality by lemma 3.5c). Also by theorem 3.2, lemma 3.6 and (A)

$$|b_{\underset{\sim}{m}\underset{\sim}{j}}| \le (4v)^{|\bar{\underset{\sim}{m}}^1| + |\underset{\sim}{j}|} (N(f)T(f_{\underset{\sim}{0}}))^{2(\underset{\sim}{m}+\underset{\sim}{j})-1} \le (4v)^{M+2d} (N(f)T(f_{\underset{\sim}{0}}))^{2M+4d-1}.$$

Therefore

$$2\ |\,t\Lambda\mathrm{dcf}(t)\,|\ \le 2\,(4v)^{2d+M}\ (2\,N(f)\,T(\underset{\sim}{f}_0)^2)^{4d+2M-1}. \qquad \text{(B)}$$

Thus, for any integer c absolutely larger than the right-hand side of (B) we know that $t(c) \neq 0$, and therefore $h(y_1, cy_1, y_3, \ldots, x)$ contains a non-zero monomial $t(c)y_1^{p_1+p_2}\underset{\sim}{x}^m$ and cannot be a polynomial factor of $g_c(y_1, y_3, \ldots, x)$, as argued above. Our given bound then obviously works for any factor candidate h. □

Our irreducibility test can now be constructed easily by induction. We compute the bounds $c_1, \ldots, c_{v-1}$ such that for the sequence of polynomials $f_1 = f$,

$$f_2(y_1, y_3, \ldots, x) = f_1(y_1, c_1y_1, y_3, \ldots, x),$$
$$f_3(y_1, y_4, \ldots, x) = f_2(y_1, c_2y_1, y_4, \ldots, x), \ldots,$$
$$f_v(y_1, x) = f_{v-1}(y_1, c_{v-2}y_1, x),\ g = f_v,$$

we have $c_i = B_5(f_i)$. Since v is assumed to be fixed and since $B_5(f_i)$ is of size polynomial in $\deg(f_i)$ and $\log(\,|\,f_i\,|\,)$, g can be constructed in time polynomial in $\deg(f)$ and $\log(\,|\,f\,|\,)$. By theorem 3.3, g is irreducible if f is irreducible. On the other hand, if $f = f_1f_2$ then

$$g(y_1, x) = f_1(y_1, c_1y_1, \ldots, c_{v-1}y_1, x)\ f_2(y_1, c_1y_1, \ldots, c_{v-1}y_1, x).$$

One can prove theorem 3.3 for the more general substitution $y_2 = c\ y_1^s$, s being an arbitrary positive integer. Since the bound $B_5(f)$ grows monotonicly in $|\,f\,|$ we can, in the case that f is reducible, find a bound for c using lemma 3.3 such that the given substitution maps all irreducible factors of f into irreducible polynomials in one less variable. Together with a Kronecker style algorithm this then leads to a different polynomial time reduction from multivariate to bivariate polynomial factorization. In the case of $v = 2$ the complete proof is given in [Kaltofen 82], which, following the lines of the proof for theorem 3.3, is readily extended to any fixed v. Instead of using Kronecker's algorithm one can also apply the multivariate Hensel lifting algorithm (see section 1.4) with the coefficients in $Q(y_1)$. Since our evaluation guarantees that no extraneous factors can occur all computed coefficients actually lie in $Z[y_1]$.

The type of substitution $y_2 = c\ y_1^s$ is derived from a version of the Hilbert Irreducibility Theorem [Franz 31] and theorem 3.3 can be regarded as its effective counterpart. For the classical Hilbert

Irreducibility Theorem, no such an effective formulation seems to be known. (See open problem 2 in section 4.)

4. Conclusion

We have discussed the phenomenon of extraneous factors during the uni- and multivariate Hensel algorithm which can cause in both cases exponential running time in the degree of the polynomial to be factored. Two classical theorems are central for the analysis of this problem. The Chebotarev Density Theorem in the univariate and the Hilbert Irreducibility Theorem in the multivariate case. An effective version of the first provided us with an algorithm for the determination of Galois groups, an effective version of the second provided a reduction algorithm from multivariate to bivariate irreducibility testing. We also have shown how to overcome the extraneous factor problem in the multivariate case by approximating a root and then determining its minimal polynomial, which lead us to solving a system of linear equations. We conclude this thesis with a list of open problems.

Problem 1: Does there exist a polynomial $p(d,v)$ and an infinite sequence of polynomials $f(x_1, \ldots, x_v) \in Z[x_1, \ldots, x_v]$ with the following property: If d is the maximum of the degrees in the individual variables then any f contains less than $p(d,v)$ monomials with non-zero coefficients; moreover, there does not exist a polynomial $q(d,v)$ such that any factor of f contains less than $q(d,v)$ monomials with non-zero coefficients? In simple words, are there sparse polynomials with dense factors?

Problem 2: Does there exist an infinite sequence of irreducible polynomials $f(y,x) \in Z[y,x]$, $n = \deg(f)$, such that for no polynomial $p(n)$ any polynomial $f(i,x)$ is irreducible for an integer i with $|i| < p(n)$? This problem asks whether there is an effective version of the Hilbert Irreducibility Theorem.

Problem 3: Devise an algorithm which, for a fixed number v, computes irreducible factors of $f(x_1, \ldots, x_v) \in Z_p[x_1, \ldots, x_v]$ in time polynomial in $p \deg(f)$. Algorithm 3.1 partially solves this problem provided that we can find good evaluation points in step (I3). Can one determine irreducibility of f in deterministic or probabilistic time polynomial in $\log(p) \deg(f)$?

Problem 4: Given a monic irreducible polynomial $f \in Z[x]$, $n = \deg(f)$, construct a generating set for the Galois group of f in time polynomial in $n \log(|f|)$. Notice that such a set contains at most $\log_2(n!) = O(n \log(n))$ elements.

Da steh' ich nun, ich armer Tor,
und bin so klug als wie zuvor.

J. W. v. Goethe - Faust

(Here I stand with all my lore,
poor fool, no wiser than before.)

REFERENCES

[Abdali et al. 77]

Abdali, S. K., Caviness, B. F., Pridor, A.: Modular Polynomial Arithmetic in Partial Fraction Decomposition. Proc. MACSYMA Users’ Conf. 77. Washington, D.C.: NASA 1977, 253-261.

[Adleman and Odlyzko 81]

Adleman, L. M., Odlyzko, A. M.: Irreducibility Testing and Factorization of Polynomials. Proc. 22nd Symp. Foundations Comp. Sci. IEEE, 409-418 (1981).

[Berlekamp 70]

Berlekamp, E. R.: Factoring Polynomials over Large Finite Fields. Math. Comp. **24,** 713-735 (1970).

[Besicovitch 40]

Besicovitch, A. S.: On the Linear Independence of Fractional Powers of the Integers. J. London Math. Soc. **15,** 1-3 (1940).

[Cantor 81]

Cantor, D. G.: Irreducible Polynomials with Integral Coefficients Have Succinct Certificates. J. of Algorithms **2,** 385-392 (1981).

[Caviness 68]

Caviness, B. F.: On Canonical Forms and Simplification. Ph.D. thesis, Carnegie-Mellon Univ., 1968.

[Chebotarev 26]

Chebotarev,N. G.: Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören. Math. Ann. **95,** 191-228 (1926).

[Claybrook 76]

Claybrook, B. G.: Factorization of Multivariate Polynomials over the Integers. ACM SIGSAM Bulletin **10,** 13 (Feb. 1976).

[Collins 79]

Collins, G. E.: Factoring Univariate Polynomials in Polynomial Average Time. Springer Lecture Notes in Comp. Sci. **72,** 317-329 (1979).

[Eisenstein 1850]

Eisenstein, F. G.: Über die Irreductibilität und einige andere Eigenschaften der Gleichung, von welcher die Theilung der ganzen Lemniscate abhängt. J. f. d. reine u. angew. Math. **39,** 160-179 (1850).

[Franz 31]

Franz, W.: Untersuchungen zum Hilbertschen Irreduzibilitätssatz. Math. Zeitschr. **33,** 275-293 (1931).

[Gallagher 72]

Gallagher, P. X.: Probabilistic Galois Theory. AMS Proc. Symp. in Pure Math., Analytic Number Theory, 91-102 (1972).

[Gantmacher 59]

Gantmacher, F. R.: Matrix Theory, vol. 1. New York: Chelsea 1959.

[Gel’fond 60]

Gel’fond, A. O.: Transcendental and Algebraic Numbers. New York: Dover Publ. 1960.

[Hardy and Wright 79]

Hardy, G. H., Wright, E. M.: An Introduction to the Theory of Numbers. Oxford: Oxford Univ. Press 1979.

[Heintz and Sieveking 81]

Heintz, J., Sieveking, M.: Absolute Primality of Polynomials is Decidable in Random Polynomial Time in the Number of Variables. Springer Lecture Notes Comp. Sci. **115,** 16-28 (1981).

[Horowitz 71]

Horowitz, E.: Algorithms for Partial Fraction Decomposition and Rational Function Integration. In Petrick, S. (ed.): Proc. 2nd Symp. on Symbolic and Algebraic Manipulation. ACM, 441-457 (1971).

[Johnson 66]

Johnson, S. C.: Tricks for Improving Kronecker’s Method. Bell Laboratories Report 1966.

[Kaltofen 82]

Kaltofen, E.: A Polynomial Reduction from Multivariate to Bivariate Integer Polynomial Factorization. ACM Proc. Symp on Theory Comp. 1982, 261-266.

[Kaltofen et al. 81]

Kaltofen, E., Musser, D. R., Saunders, B. D.: A Generalized Class of Polynomials that Are Hard to Factor. In Wang, P. (ed.): Proc. ACM Symp. on Symbolic and Algebraic Comp. 1981. ACM, 188-194.

[Knobloch 55]

Knobloch, H.-W.: Zum Hilbertschen Irreduzibilitätssatz. Abh. math. Sem. d. Univ. Hamburg **19,** 176-190 (1955).

[Knuth 69]

Knuth, D. E.: The Art of Computer Programming, vol.2, Seminumerical Algorithms. Reading, MA: Addison Wesley 1969.

[Knuth 81]

Knuth, D. E.: The Art of Computer Programming, vol.2, Seminumerical Algorithms, 2nd ed. Reading, MA: Addison Wesley 1981.

[Königsberger 1895]

Königsberger, L.: Über den Eisensteinschen Satz von der Irreductibilität algebraischer Gleichungen. J. f. reine angew. Math. **115,** 53-78 (1895).

[Kronecker 1882]

Kronecker, L.: Grundzüge einer arithmetischen Theorie der algebraischen Grössen. J. f. d. reine u. angew. Math. **92,** 1-122 (1882).

[Kung and Tong 77]

Kung, H. T., Tong, D. M.: Fast Algorithms for Partial Fraction Decomposition. SIAM J. Comp. **6,** 582-593 (1977).

[Kung and Traub 78]

Kung, H. T., Traub, J. F.: All Algebraic Functions Can Be Computed Fast. J. ACM **25,** 245-260 (1978).

[Lagarias and Odlyzko 77]

Lagarias, J. C., Odlyzko, A. M.: Effective Versions of the Chebotarev Density Theorem. In Fröhlich A. (ed.): Algebraic Number Fields (L-Functions and Galois Properties). New York: Academic Press 1977.

[Lagarias et al. 79]

Lagarias, J. C., Montgomery, H. L., Odlyzko, A. M.: A Bound for the Least Prime Ideal in the Chebotarev Density Theorem. Inventiones Math. **54,** 271-296 (1979).

[Lenstra et al. 82]

Lenstra, A. K., Lenstra, H. W., Lovász, L.: Factoring Polynomials with Rational Coefficients. Report 82-05. Amsterdam: Mathematisch Instituut 1982.

[Loos 82]

Loos, R.: Computing in Algebraic Extensions. Computing Supplement **4.** Vienna: Springer Verlag 1982.

[Macsyma 77]

MACSYMA Reference Manual. Version 9, Cambridge MA: M.I.T. 1977.

[McClellan 73]

McClellan, M. T.: The Exact Solution of Systems of Linear Equations with Polynomial Coefficients. J. ACM **20,** 563-588 (1973).

[Mignotte 74]

Mignotte, M.: An Inequality about Factors of Polynomials. Math. Comp. **28,** 1153-1157 (1974).

[Moore and Norman 81]

Moore, P. M. A., Norman, A. C.: Implementing a Polynomial Factorization Problem. In Wang, P. (ed.): Proc. ACM Symp. on Symbolic and Algebraic Comp. 1981. ACM, 109-116.

[Moses and Yun 73]

Moses, J., Yun, D. Y. Y.: The EZGCD Algorithm. Proc. 1973 ACM National Conf., 159-166.

[Musser 71]

Musser, D. R.: Algorithms for Polynomial Factorization. Ph.D. thesis and TR 134, Univ. of Wisconsin 1971.

[Musser 76]

Musser, D. R.: Multivariate Polynomial Factorization. J. ACM **22,** 291-308 (1976).

[Musser 78]

Musser, D. R.: On the Efficiency of a Polynomial Irreducibility Test. J. ACM **25,** 271-282 (1978).

[Narkiewicz 74]

Narkiewicz, W.: Elementary and Analytic Theory of Algebraic Numbers. Warsaw: Polish Science Publ. 1974.

[Noether 22]

Noether, E.: Ein algebraisches Kriterium für absolute Irreduzibilität. Math. Ann. **85,** 26-33 (1922).

[Oesterlé 79]

Oesterlé, J.: Versions Effectives du Théorème de Chebotarev sous L'Hypothèse de Riemann Généralisée. Soc. Math. France Asterisque **61,** 165-167 (1979).

[Plaisted 78]

Plaisted, D. A.: Some Polynomial and Integer Divisibility Problems are NP-Hard. SIAM J. Comp. **7,** 458-464 (1978).

[Pratt 75]

Pratt, V. R.: Every Prime Has a Succinct Certificate. SIAM J. Comp. **4,** 214-220 (1975).

[Rabin 80]

Rabin, M. O.: Probabilistic Algorithms in Finite Fields. SIAM J. Comp. **9,** 273-280 (1980).

[Richards 74]

Richards, I.: An Application of Galois Theory to Elementary Arithmetic. Adv. in Math. **13,** 268-273 (1974).

[Trager 76]

Trager, B. M.: Algebraic Factoring and Rational Function Integration. In Jenks, R. (ed.): Proc. ACM Symp. on Symbolic and Algebraic Comp. 1976. ACM, 219-226.

[van der Waerden 39]

van der Waerden, B. L.: Einführung in die algebraische Geometrie. Berlin: Springer Verlag 1939.

[van der Waerden 53]

van der Waerden, B. L.: Modern Algebra, vol.1. Engl. transl. by F. Blum. New York: Ungar Publ. 1953.

[Vaughan 75]

Vaughan, R. C.: Bounds for the Coefficients of Cyclotomic Polynomials. Michigan Math. J. **21,** 289-295 (1975).

[Viry 80]

Viry, G.: Factorisation des Polynômes a Plusieurs Variables. RAIRO Informatique Théorique **14,** 209-223 (1980).

[Walker 50]

Walker, J. R.: Algebraic Curves. New York: Dover Publ. 1950.

[Wang 77]

Wang, P. S.: Preserving Sparseness in Multivariate Polynomial Factorization. Proc. MACSYMA Users' Conf. 77. Washington, D.C.: NASA 1977, 55-61.

[Wang 78]

Wang, P. S.: An Improved Multivariate Polynomial Factoring Algorithm. Math. Comp. **32,** 1215-1231 (1978).

[Wang 79a]

Wang, P. S.: Parallel p-adic Constructions in the Univariate Polynomial Factoring Algorithm. In Lewis, V. (ed.): Proc. MACSYMA Users' Conf. 79. Cambridge, MA: MIT 1979, 310-318.

[Wang 79b]

Wang, P. S.: Analysis of the p-adic Construction of Multivariate Correction Coeficients in Polynomial Factorization: Iteration vs. Recursion. Springer Lecture Notes in Comp. Sci. **72,** 291-300 (1979).

[Wang and Rothschild 75]

Wang, P. S., Rothschild, L. P.: Factoring Multivariate Polynomials over the Integers. Math. Comp. **29,** 935-950 (1975).

[Wang and Trager 79]

Wang, P. S., Trager, B. M.: New Algorithms for Polynomial Square-free Decomposition over the Integers. SIAM J. Comp. **8,** 300-305 (1979).

[Weinberger 81]

Weinberger, P. J.: Finding the Number of Factors of a Polynomial. 1981, manuscript.

[Weinberger and Rothschild 76]

Weinberger, P. J., Rothschild, L. P.: Factoring Polynomials over Algebraic Number Fields. ACM Trans. Math. Software **2,** 335-350 (1976).

[Yun 76a]

Yun, D. Y. Y.: Hensel Meets Newton - Algebraic Construction in an Analytic setting. Analytic Computational Complexity. Traub J. ed. NY: Academic Press 1976.

[Yun 76b]

Yun, D. Y. Y.: On Squarefree Decomposition Algorithms. In Jenks, R. (ed.): Proc. ACM Symp. on Symbolic and Algebraic Comp. 1976. ACM, 26-35.

[Yun 77]

Yun, D. Y. Y.: On the Equivalence of Polynomial GCD and Squarefree Factorization Problems. Proc. MACSYMA Users’ Conf. 77. Washington, D.C.: NASA 1977, 65-70.

[Zassenhaus 69]

Zassenhaus, H.: On Hensel Factorization I. J. Number Theory **1,** 291-311 (1969).

[Zassenhaus 78]

Zassenhaus, H.: A Remark on the Hensel Factorization Method. Math. Comp. **32,** 287-292 (1978).

[Zassenhaus 81]

Zassenhaus, H.: Polynomial Time Factoring of Integral Polynomials. ACM SIGSAM Bulletin **15,** 6-7, (May 1981).

[Zippel 79]

Zippel, R. E.: Probabilistic Algorithms for Sparse Polynomials. Ph.D. thesis, MIT 1979.

APPENDIX A

```
     Vaxima 1.50
  Sat Mar  6 01:41:14 1982

(c2) /* Sample run of algorithm 3.1: */

/* Inhibit garbage collection message. */

gcprint:false$

(c3) /* The following bi-variate polynomial is squarefree and
monic in x as well as 0 is already a useable evaluation
point for y. Therefore, step (I) is not needed. */

f:x↑6+x↑5+(2*y+4)*x↑4+(y+3)*x↑3+(y↑2+3*y+5)*x↑2+(2-y)*x-y↑2+y+2;
      2   2               2    4              3                            6
(d3) x  (y  + 3 y + 5) - y  + x  (2 y + 4) + x  (y + 3) + y + x (2 - y) + x

                                                                     5
                                                                + x  + 2

(c4) /* Step (F): */

factor(subst(0,y,f));
                          2        2        2
(d4)                    (x  + 1) (x  + 2) (x  + x + 1)

(c5) /* We choose β=%i, the imaginary unit. */

/* Step (N): */

/* Step (N1): */

g[0]:x-%i;
(d5)                             x - %i

(c6) h[0]:quotient(subst(0,y,f),g[0]);
      5             4             3               2
(d6) x  + (%i + 1) x  + (%i + 3) x  + (3 %i + 2) x  + (2 %i + 2) x + 2 %i

(c7) /* Step (N2). We actually choose K smaller than described,
though this does not influence the outcome of latter steps */

K:5;
(d7)                                5

(c8) /* Precompute inverse of f’(0,%i)=h[0](%i). */
r:ratsimp(1/subst(%i,x,h[0]));
```

```
                                  1
(d8)                            - -
                                  2

(c9) /* Computation of g[j] and h[j] using b[j]: */

for j:1 thru K do (
   display(f[j]:ratcoeff(f,y,j)),
   /* Step (N2): */
   if j=0
      then display(b[j]:f[j])
      else display(b[j]:ratsimp(f[j]-sum(g[s]*h[j-s],s,1,j-1))),
   display(g[j]:ratsimp(subst(%i,x,b[j])*r)),
   display(h[j]:quotient(b[j]-h[0]*g[j],g[0]))
   );
                             4    3      2
                     f  = 2 x  + x  + 3 x  - x + 1
                      1

                             4    3      2
                     b  = 2 x  + x  + 3 x  - x + 1
                      1

                              g  = %i
                               1

                  4             3             2
       h  = - %i x  + (4 - %i) x  + (%i + 3) x  + (%i + 5) x + 3 %i
        1

                                  2
                            f  = x  - 1
                             2

               4                 3               2
       b  = - x  + (- 4 %i - 1) x  + (2 - 3 %i) x  + (1 - 5 %i) x + 2
        2

                                   5 %i
                            g  = - ----
                             2      2

             4                88888 3                  2
       5 %i x  + (5 %i - 12) x  + (- 5 %i - 12) x  + (- 8 %i - 6) x - 6 %i
  h  = -------------------------------------------------------------------
   2                               2

                              f  = 0
                               3
                     4                 3                  2
            - 10 x  + (- 32 %i - 10) x  + (10 - 27 %i) x  + (13 - 31 %i) x + 21
      b  = - -----------------------------------------------------------------------
       3                                   2

                                    25 %i
                              g  = -----
                               3     2

h  =
 3

                   4                3                  2
          25 %i x  + (25 %i - 60) x  + (- 17 %i - 60) x  + (- 37 %i - 48) x - 29 %i
        - -----------------------------------------------------------------------------
                                      2

                               f  = 0
                                4

b  =
 4

                 4                   3                  2
         - 125 x  + (- 380 %i - 125) x  + (109 - 330 %i) x  + (164 - 376 %i) x + 238
         ------------------------------------------------------------------------------
                                      4

                                   619 %i
                             g  = - ------
                              4       8

                 4                  3                     2
h  = (619 %i x  + (619 %i - 1488) x  + (- 391 %i - 1488) x
 4

                                 + (- 910 %i - 1248) x - 762 %i)/8

                               f  = 0
                                5

                     4                   3                     2
b  = - (- 1738 x  + (- 5164 %i - 1738) x  + (1430 - 4545 %i) x
 5

                                   + (2299 - 5123 %i) x + 3209)/8

                                 4291 %i
                             g  = -------
                              5      8
                 4                      3                         2
h  = - (4291 %i x  + (4291 %i - 10320) x  + (- 2611 %i - 10320) x
 5

                                  + (- 6283 %i - 8832) x - 5373 %i)/8

(d9)                          done

(c10) /* Assign K-th order approximation of root of f. */

a[K]:%i+sum(-g[j]*y↑j,j,1,K);
                    5          4          3         2
           4291 %i y    619 %i y    25 %i y    5 %i y
(d10)    - ---------- + --------- - -------- + ------- - %i y + %i
               8            8          2          2

(c11) /* This command verifies our approximation as it was proven. */

remainder(subst(a[K],x,f),y↑(K+2));
                                      6
                            63729 %i y
(d11)                       -----------
                                 8

(c12) /* Compute powers of a[K] mod y↑(K+1). */

asquare[K]:remainder(a[K]↑2,y↑(K+1));
                      5        4       3      2
(d12)           1290 y  - 186 y  + 30 y  - 6 y  + 2 y - 1

(c13) acube[K]:remainder(a[K]*asquare[K],y↑(K+1));
                5            4           3          2
      18537 %i y  - 2667 %i y  + 428 %i y  - 84 %i y  + 24 %i y - 8 %i
(d13) ---------------------------------------------------------------
                               8

(c14) afourth[K]:remainder(a[K]*acube[K],y↑(K+1));
                      5        4       3       2
(d14)         - 3684 y  + 528 y  - 84 y  + 16 y  - 4 y + 1

(c15) /* Set up undetermined polynomials of possible factor. */

u[0]:w0+v0*y+u0*y↑2;
                              2
(d15)                     u0 y  + v0 y + w0

(c16) u[1]:w1+v1*y+u1*y↑2;
                              2
(d16)                     u1 y  + v1 y + w1

(c17) u[2]:w2+v2*y+u2*y↑2;
```

```
                                2
(d17)                        u2 y  + v2 y + w2

(c18) u[3]:w3+v3*y+u3*y↑2;
                                2
(d18)                        u3 y  + v3 y + w3

(c19) /* Compute equation (3.1) for I=2. */

L2:remainder(asquare[K]+u[1]*a[K]+u[0],y↑(K+1));
                                                5
(d19) - ((4291 %i w1 - 619 %i v1 + 100 %i u1 - 10320) y

                                         4
 + (- 619 %i w1 + 100 %i v1 - 20 %i u1 + 1488) y

                                     3
 + (100 %i w1 - 20 %i v1 + 8 %i u1 - 240) y

                                          2
 + (- 20 %i w1 + 8 %i v1 - 8 %i u1 - 8 u0 + 48) y

 + (8 %i w1 - 8 %i v1 - 8 v0 - 16) y - 8 %i w1 - 8 w0 + 8)/8

(c20) t2:[]$

(c21) /* Retrieve linear equations for the coefficients. */

for i:0 thru K do
    for j:0 thru 1 do (
        display(s2[i,j]:ratcoeff(ratcoeff(L2,y,i),%i,j)),
        t2:cons(s2[i,j],t2)
        );
                             s2     = w0 - 1
                               0, 0

                               s2     = w1
                                 0, 1

                             s2     = v0 + 2
                               1, 0

                            s2     = v1 - w1
                              1, 1

                             s2     = u0 - 6
                               2, 0

                                5 w1 - 2 v1 + 2 u1
                       s2     = ------------------
                         2, 1           2
```

```
                     s2      = 30
                       3, 0

                       25 w1 - 5 v1 + 2 u1
              s2      = - -------------------
                3, 1               2

                   s2      = - 186
                     4, 0

                   619 w1 - 100 v1 + 20 u1
             s2      = -----------------------
               4, 1               8

                    s2      = 1290
                      5, 0

                   4291 w1 - 619 v1 + 100 u1
           s2      = - -------------------------
             5, 1                  8

(d21)                                    done

(c22) /* Try to solve the system. */

errcatch(linsolve(t2,
                    [w0,v0,u0,w1,v1,u1]));
Dependent equations eliminated:  (1 5 7)
Inconsistent equations:  (2 4 6)

(d22)                                     []

(c23) /* Compute equation (3.1) for I=3. */

L3:remainder(acube[K]+u[2]*asquare[K]+u[1]*a[K]+u[0],y↑(K+1));
(d23) ((10320 w2 - 4291 %i w1 - 1488 v2 + 619 %i v1 + 240 u2 - 100 %i u1

                5
 + 18537 %i) y  + (- 1488 w2 + 619 %i w1 + 240 v2 - 100 %i v1 - 48 u2

                          4
 + 20 %i u1 - 2667 %i) y  + (240 w2 - 100 %i w1 - 48 v2 + 20 %i v1 + 16 u2

                       3
 - 8 %i u1 + 428 %i) y  + (- 48 w2 + 20 %i w1 + 16 v2 - 8 %i v1 - 8 u2

                              2
 + 8 %i u1 + 8 u0 - 84 %i) y  + (16 w2 - 8 %i w1 - 8 v2 + 8 %i v1 + 8 v0

 + 24 %i) y - 8 w2 + 8 %i w1 + 8 w0 - 8 %i)/8
```

```
(c24) t3:[]$

(c25) /* Retrieve linear equations for the coefficients. */

for i:0 thru K do
   for j:0 thru 1 do (
      display(s3[i,j]:ratcoeff(ratcoeff(L3,y,i),%i,j)),
      t3:cons(s3[i,j],t3)
      );
                       s3      = w0 - w2
                         0, 0

                        s3      = w1 - 1
                          0, 1

                     s3      = 2 w2 - v2 + v0
                       1, 0

                     s3      = - w1 + v1 + 3
                       1, 1

                 s3      = - 6 w2 + 2 v2 - u2 + u0
                   2, 0

                        5 w1 - 2 v1 + 2 u1 - 21
                 s3      = -----------------------
                   2, 1                2

                   s3      = 30 w2 - 6 v2 + 2 u2
                     3, 0

                        25 w1 - 5 v1 + 2 u1 - 107
                s3      = - -------------------------
                  3, 1                 2

                 s3      = - 186 w2 + 30 v2 - 6 u2
                   4, 0

                      619 w1 - 100 v1 + 20 u1 - 2667
               s3      = ------------------------------
                 4, 1                 8

                 s3      = 1290 w2 - 186 v2 + 30 u2
                   5, 0

                     4291 w1 - 619 v1 + 100 u1 - 18537
             s3      = - --------------------------------
               5, 1                   8

(d25)                            done
```

```
(c26) /* Try to solve the system. */

errcatch(linsolve(t3,
                  [w0,v0,u0,w1,v1,u1,w2,v2,u2]));
Inconsistent equations:  (5 7 1)

(d26)                                    []

(c27) /* Compute equation (3.1) for I=4. */

L4:remainder(afourth[K]+u[3]*acube[K]+u[2]*asquare[K]+u[1]*a[K]+u[0],y↑(K+1));
(d27) ((18537 %i w3 + 10320 w2 - 4291 %i w1 - 2667 %i v3 - 1488 v2 + 619 %i v1

                                             5
 + 428 %i u3 + 240 u2 - 100 %i u1 - 29472) y

 + (- 2667 %i w3 - 1488 w2 + 619 %i w1 + 428 %i v3 + 240 v2 - 100 %i v1

                                         4
 - 84 %i u3 - 48 u2 + 20 %i u1 + 4224) y

 + (428 %i w3 + 240 w2 - 100 %i w1 - 84 %i v3 - 48 v2 + 20 %i v1 + 24 %i u3

                            3
 + 16 u2 - 8 %i u1 - 672) y  + (- 84 %i w3 - 48 w2 + 20 %i w1 + 24 %i v3

                                                               2
 + 16 v2 - 8 %i v1 - 8 %i u3 - 8 u2 + 8 %i u1 + 8 u0 + 128) y

 + (24 %i w3 + 16 w2 - 8 %i w1 - 8 %i v3 - 8 v2 + 8 %i v1 + 8 v0 - 32) y

 - 8 %i w3 - 8 w2 + 8 %i w1 + 8 w0 + 8)/8

(c28) t4:[]$

(c29) /* Retrieve linear equations for the coefficients. */

for i:0 thru K do
    for j:0 thru 1 do (
        display(s4[i,j]:ratcoeff(ratcoeff(L4,y,i),%i,j)),
        t4:cons(s4[i,j],t4)
        );
                            s4     = - w2 + w0 + 1
                              0, 0

                               s4     = w1 - w3
                                 0, 1

                          s4     = 2 w2 - v2 + v0 - 4
                            1, 0
```

```
                    s4      = 3 w3 - w1 - v3 + v1
                      1, 1

                s4      = - 6 w2 + 2 v2 - u2 + u0 + 16
                  2, 0

                    21 w3 - 5 w1 - 6 v3 + 2 v1 + 2 u3 - 2 u1
          s4      = - ----------------------------------------
            2, 1                      2

                  s4      = 30 w2 - 6 v2 + 2 u2 - 84
                    3, 0

                 107 w3 - 25 w1 - 21 v3 + 5 v1 + 6 u3 - 2 u1
         s4      = -------------------------------------------
           3, 1                      2

               s4      = - 186 w2 + 30 v2 - 6 u2 + 528
                 4, 0

               2667 w3 - 619 w1 - 428 v3 + 100 v1 + 84 u3 - 20 u1
     s4      = - --------------------------------------------------
       4, 1                          8

              s4      = 1290 w2 - 186 v2 + 30 u2 - 3684
                5, 0

           18537 w3 - 4291 w1 - 2667 v3 + 619 v1 + 428 u3 - 100 u1
   s4      = -------------------------------------------------------
     5, 1                          8

(d29)                                    done

(c30) /* Try to solve the system. */

linsolve(t4,
         [w0,v0,u0,w1,v1,u1,w2,v2,u2,w3,v3,u3]);
Solution

(e30)                                   w1 = 0

(e31)                                   v1 = 0

(e32)                                   w0 = 2

(e33)                                   u2 = 0

(e34)                                   w2 = 3

(e35)                                   v2 = 1
```

```
(e36)                          u0 = 0

(e37)                         v0 = - 1

(e38)                          u1 = 0

(e39)                          w3 = 0

(e40)                          v3 = 0

(e41)                          u3 = 0
(d41)    [e30, e31, e32, e33, e34, e35, e36, e37, e38, e39, e40, e41]

(c42) /* Substitute solution into factor. */

g:ev(x↑4+u[3]*x↑3+u[2]*x↑2+u[1]*x+u[0],%);
                        2              4
(d42)                  x  (y + 3) - y + x  + 2

(c43) /* Test whether it divides f as was proven. */

remainder(f,g);
(d43)                              0

(c44) /* Finally we demonstrate what happens if K is too small (4). */

a[4]:remainder(a[K],y↑5);
                    4           3          2
             619 %i y  - 100 %i y  + 20 %i y  - 8 %i y + 8 %i
(d44)        ------------------------------------------------
                                  8

(c45) asquare[4]:remainder(asquare[K],y↑5);
                          4       3      2
(d45)               - 186 y  + 30 y  - 6 y  + 2 y - 1

(c46) acube[4]:remainder(acube[K],y↑5);
                     4           3          2
             2667 %i y  - 428 %i y  + 84 %i y  - 24 %i y + 8 %i
(d46)      - --------------------------------------------------
                                  8

(c47) afourth[4]:remainder(afourth[K],y↑5);
                        4       3       2
(d47)               528 y  - 84 y  + 16 y  - 4 y + 1

(c48) /* Compute equation (3.1) for I=4. */

L4:remainder(afourth[4]+u[3]*acube[4]+u[2]*asquare[4]+u[1]*a[4]+u[0],y↑5);

(d48) - ((2667 %i w3 + 1488 w2 - 619 %i w1 - 428 %i v3 - 240 v2 + 100 %i v1
                                4
 + 84 %i u3 + 48 u2 - 20 %i u1 - 4224) y

 + (- 428 %i w3 - 240 w2 + 100 %i w1 + 84 %i v3 + 48 v2 - 20 %i v1 - 24 %i u3

                      3
 - 16 u2 + 8 %i u1 + 672) y  + (84 %i w3 + 48 w2 - 20 %i w1 - 24 %i v3 - 16 v2

                                          2
 + 8 %i v1 + 8 %i u3 + 8 u2 - 8 %i u1 - 8 u0 - 128) y

 + (- 24 %i w3 - 16 w2 + 8 %i w1 + 8 %i v3 + 8 v2 - 8 %i v1 - 8 v0 + 32) y

 + 8 %i w3 + 8 w2 - 8 %i w1 - 8 w0 - 8)/8

(c49) t4:[]$

(c50) /* Retrieve linear equations for the coefficients. */

for i:0 thru 4 do
   for j:0 thru 1 do (
      display(s4[i,j]:ratcoeff(ratcoeff(L4,y,i),%i,j)),
      t4:cons(s4[i,j],t4)
      );
                      s4      = - w2 + w0 + 1
                        0, 0

                        s4      = w1 - w3
                          0, 1

                     s4      = 2 w2 - v2 + v0 - 4
                       1, 0

                    s4      = 3 w3 - w1 - v3 + v1
                      1, 1

                s4      = - 6 w2 + 2 v2 - u2 + u0 + 16
                  2, 0

                     21 w3 - 5 w1 - 6 v3 + 2 v1 + 2 u3 - 2 u1
           s4      = - ----------------------------------------
             2, 1                        2

                  s4      = 30 w2 - 6 v2 + 2 u2 - 84
                    3, 0

                 107 w3 - 25 w1 - 21 v3 + 5 v1 + 6 u3 - 2 u1
          s4      = -------------------------------------------
            3, 1                         2
```

```
                  s4      = - 186 w2 + 30 v2 - 6 u2 + 528
                    4, 0

                  2667 w3 - 619 w1 - 428 v3 + 100 v1 + 84 u3 - 20 u1
         s4      = - -------------------------------------------------
           4, 1                             8

(d50)                               done

(c51) /* Try to solve the system. */

linsolve(t4,
        [w0,v0,u0,w1,v1,u1,w2,v2,u2,w3,v3,u3]);

Solution

                                     u3
(e51)                             w1 = --
                                     9

                                    2 u3
(e52)                            v1 = ----
                                     3

                                  v2 + 15
(e53)                           w0 = -------
                                     8

                                  9 v2 - 9
(e54)                           u2 = --------
                                     8

                                  v2 + 23
(e55)                           w2 = -------
                                     8

                                    v2 - 1
(e56)                           u0 = - ------
                                      8

                                  3 v2 - 7
(e57)                           v0 = --------
                                     4

                                      u3
(e58)                            u1 = - --
                                      9

                                     u3
(e59)                             w3 = --
                                     9
```

```
                                 8 u3
(e60)                       v3 = ----
                                  9

(d60)       [e51, e52, e53, e54, e55, e56, e57, e58, e59, e60]

(c61) /* We now specialize u3=0 and v2=0 in the above solution. */

g:ev(ev(x↑4+u[3]*x↑3+u[2]*x↑2+u[1]*x+u[0],%),u3=0,v2=0);
                   2              2
                  y     2  23    9 y     7 y    4   15
(d61)             -- + x  (-- - ----) - --- + x  + --
                  8        8     8       4         8

(c62) /* g does not divide f, however */

resultant(f,g,x);

/user/vaxima/rat/result being loaded.
[fasl /user/vaxima/rat/result.o]
             10      6        5        4        3        2
(d62)  5184 y   (81 y  + 324 y  - 135 y  - 806 y  + 865 y  - 98 y + 169)

(c63) /* which is divisible by y↑4 explaining the problem. */
```

APPENDIX B

A Short Bibliography of Hilbert's Irreducibility Theorem

[Hilbert 1892]

Hilbert, D.: Über die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten. Journal für die reine und angewandte Mathematik **110,** 104-29 (1892).

[Skolem 21]

Skolem, T.: Untersuchungen über die möglichen Verteilungen ganzzahliger Lösungen gewisser Gleichungen. Skrifter udgivne af Videnskaps-Selskapet i Christinia **17,** (1921).

[Dörge 26]

Dörge, K.: Zum Hilbertschen Irreduzibilitätssatz. Mathematische Annalen **95,** 84-97 (1926).

[Dörge 26]

Dörge, K.: Über die Seltenheit der reduziblen Polynome und der Normalgleichungen. Mathematische Annalen **95,** 247-56 (1926).

[Dörge 26]

Dörge, K.: Ein Beitrag zur Theorie der diophantischen Gleichungen mit zwei Unbekannten. Mathematische Zeitschrift **24,** 193-8 (1926).

[Dörge 27]

Dörge, K.: Einfacher Beweis des Hilbertschen Irreduzibilitätssatzes. Mathematische Annalen **96,** 176-82 (1927).

[Dörge 30]

Dörge, K.: Bemerkung zum Hilbertschen Irreduzibilitätssatz. Mathematische Annalen **102,** 521-30 (1930).

[Franz 31]

Franz, W.: Untersuchungen zum Hilbertschen Irreduzibilitätssatz. Math. Zeitschrift **33,** 275-293 (1931).

[Eichler 38]

Eichler, M.: Zum Hilbertschen Irreduzibilitätssatz. Mathematische Annalen **116,** 742-8 (1938).

[Inaba 44]

Inaba, E.: Über den Hilbertschen Irreduzibilitätssatz. Japanese J. of Math. **19,** 1-25 (1944).

[Knobloch 55]

Knobloch, H.-W.: Zum Hilbertschen Irreduzibilitätssatz. Abhandlungen aus dem Mathematischen Seminar and der Universität Hamburg **19,** 176-90 (1955).

[Knobloch 56]

Knobloch, H.-W.: Die Seltenheit der reduziblen Polynome. Jahresbericht des deutschen Mathematikervereins **59,** Abteilung 1, 12-9 (1956).

[Lang 62]

Lang, S.: Diophantine Geometry. New York: Interscience Publ. 1962.

[Dörge 65]

Dörge, K.: Abschätzung der reduziblen Polynome. Mathematische Annalen **160,** 59-63 (1965).

[Schinzel 65]

Schinzel, A.: On Hilbert's Irreducibility Theorem. Ann. Polon. Math. **16,** 333-340 (1965).

[Fried 74]

Fried, M.: On Hilbert's Irreducibility Theorem. J. Number Theory **6,** 211-231 (1974).