MATHEMATICS OF COMPUTATION, VOLUME 30, NUMBER 134
APRIL 1976, PAGES 324–336

Factoring Multivariate Polynomials over Algebraic Number Fields

By Paul S. Wang*

Abstract. The algorithm for factoring polynomials over the integers by Wang and Rothschild is generalized to an algorithm for the irreducible factorization of multivariate polynomials over any given algebraic number field. The extended method makes use of recent ideas in factoring univariate polynomials over large finite fields due to Berlekamp and Zassenhaus. The procedure described has been implemented in the algebraic manipulation system MACSYMA.** Some machine examples with timing are included.

1. Introduction. An algorithm for the irreducible factorization of multivariate polynomials over any given algebraic number field is presented. The algebraic number field is given as an extension field of the rational numbers by specifying a minimal polynomial over the integers. In other words, we describe an algorithm for finding all the irreducible factors of a given multivariate polynomial over the field of the rationals adjoined by a root of a prescribed minimal polynomial. This algorithm is a generalization of the factoring algorithm for multivariate polynomials over the integers by Wang and Rothschild [11].

The multivariate polynomial to be factored is first reduced to a polynomial in just one variable by substituting properly selected integers for all but one variable. The resulting univariate polynomial is then factored over the given algebraic field.

There are two different approaches for the univariate factorization depending on whether a suitable small rational prime p exists such that the given minimal polynomial is irreducible modulo p. If p can be found, then the factoring is carried out via a finite field approach using methods suggested by Berlekamp [1] and Zassenhaus [13]. Otherwise, if the minimal polynomial is reducible modulo every prime, a classical method is used which transforms the factorization into factoring a multivariate polynomial of much higher degree over the rationals. The univariate factors will then be used to construct the desired multivariate factors by a "p-adic" interpolation described by Wang and Rothschild [11].

Our interest in factoring over algebraic number fields, as that in factoring over the integers, originated in the problem of indefinite integration of elementary functions in finite terms [6], [9]. However, the algorithm is also useful in computations related to groups and algebraic number fields. The entire algorithm has been implemented in

---

Received June 5, 1975; revised August 11, 1975.

AMS (MOS) subject classifications (1970). Primary 10M05, 12–04, 12C05; Secondary 10D05, 12E05.

*Work reported herein was supported by Laboratory for Computer Science (formerly Project MAC), an MIT Interdepartmental Laboratory, sponsored by the Advanced Research Projects Agency (ARPA), Department of Defense, monitored by Office of Naval Research Contract N8814-75-C-8661.

**A similar facility has been implemented for the IBM SCRATCHPAD system by Wang and Yun.

MACLISP [5] for the algebraic manipulation system MACSYMA [14] at Laboratory for Computer Science at M.I.T. The routines for arithmetic and greatest common divisor of polynomials over a given algebraic field are implemented in MACSYMA by Barry Trager. A number of machine examples with timing are included in the appendix.

Peter Weinberger, at the University of Michigan, has also been working on certain aspects of factoring over algebraic number fields. The author wishes to thank him for discussions and communications on this subject. He also wishes to thank Barry Trager and Joel Moses for their comments and suggestions.

2. **Preliminaries and Notation.** The field of rational numbers is denoted by $\mathbf{Q}$ and the rational integers by $\mathbf{Z}$. Any finite extension $\mathbf{K}$ of $\mathbf{Q}$ can be obtained by the adjunction of an algebraic number θ which satisfies $f(\theta) = 0$ where $f(x)$ is an irreducible polynomial in $\mathbf{Z}[x]$. This extension $\mathbf{K}$ is denoted as $\mathbf{Q}(\theta)$. The polynomial $f(x)$ is called the minimal polynomial of θ and $[\mathbf{K} : \mathbf{Q}] = m = \deg(f)$. We present an algorithm for the irreducible factorization of any multivariate polynomial $U(x, x_2, \ldots, x_t) \in \mathbf{K}[x, x_2, \ldots, x_t]$ over $\mathbf{Q}(\theta)$ for any given minimal polynomial $f(x)$.

An element in $\mathbf{K}$ satisfying a monic polynomial is called an algebraic integer in $\mathbf{K}$. The algebraic integers in $\mathbf{K}$ form a ring $\mathbf{R}$. It is obvious that $\mathbf{Z} \subset \mathbf{R}$. Let $f(x) = x^2 + 1$, for example; then $\mathbf{R}$ is the ring of Gaussian integers. It can be shown that if $\alpha \in \mathbf{K}$, there exist $z \in \mathbf{Z}$ such that $z\alpha \in \mathbf{R}$. Thus, we may assume, without loss of generality, that $f(x)$ is monic. Also, we may assume that U and all its factors have coefficients in $\mathbf{R}$.

An element in $\mathbf{K}$ can be written in the form $\sum_{i=0}^{m-1} c_i\theta^i/\delta$, δ and $c_i \in \mathbf{Z}$. For elements in R, there exist positive integers D such that any $\alpha \in \mathbf{R}$ can be written uniquely in the form $\alpha = \sum_{i=0}^{m-1} c_i\theta^i/D$, $c_i \in \mathbf{Z}$. The set $\{1/D, \theta/D, \ldots, \theta^{m-1}/D\}$ is known as an integral basis of $\mathbf{R}$. One such integer D is the largest integer Δ such that Δ^2 divides the discriminant of $f(x)$. We denote the discriminant by $\mathrm{Discr}(f)$ which is equal to the resultant of $f(x)$ and $df(x)/dx$ denoted by $\mathrm{Res}(f(x), f'(x))$.

By choosing a main variable, say x, we can write $U(x, x_2, \ldots, x_t) \in \mathbf{R}[x, x_2, \ldots, x_t]$ in the form

$$U(x, x_2, \ldots, x_t) = V_n x^n + \cdots + V_0$$

with $V_i \in \mathbf{R}[x_2, x_3, \ldots, x_t]$ for $i = 0, 1, \ldots, n$. $V_n \neq 0$ is the leading coefficient of U, denoted as $\mathrm{lc}(U)$. The *content* of U with respect to the main variable x, $\mathrm{CONT}(U)$, is $\mathrm{GCD}(V_0, V_1, \ldots, V_n)$; and the *principal part* of U, $\mathrm{pp}(U)$ is $U/\mathrm{CONT}(U)$. U is *primitive* if $\mathrm{CONT}(U) = 1$, and U is *squarefree* if U has no repeated factors. Any content of U, or repeated factors of U can be removed by relatively simple greatest common divisor (GCD) computations (see [3]). Thus, U may be assumed primitive and squarefree. As in factoring over $\mathbf{Z}$, the leading coefficient plays an important role in the factoring process [7], [11]. Factorization is easier if the leading coefficient is 1, for if U is monic, then any factor of U is monic. But if U is not monic, then additional computation is required to determine the leading coefficient of each factor. Therefore, we choose the main variable of U to make $\mathrm{lc}(U)$ 1 or small, in order to avoid or simplify later computations related to the leading coefficient. If several vari-

ables have a monic leading coefficient, it is best to choose the variable giving the smallest n, thus limiting the number of possible factors.

Let $p \in \mathbf{Z}$ be a prime and (p) be the ideal generated by p. We denote by $\mathbf{Z}_p$ the quotient field $\mathbf{Z}/(p)$. If the minimal polynomial $f(x)$ is irreducible mod(p), then $\mathbf{R}_p = \mathbf{R}/(P)$ is isomorphic to the Galois field GF(p^m).

For any set $F = \{f_1, f_2, \ldots, f_r\} \subset \mathbf{Z}[x_2, x_3, \ldots, x_t]$, the *ideal generated by* F, $(f_1, f_2, \ldots, f_r)$, is defined as the set

$$\{g_1 f_1 + g_2 f_2 + \cdots + g_r f_r : g_i \in \mathbf{Z}[x_2, \ldots, x_t] \quad \forall i\}.$$

The set F need not be finite. For any integer $k > 0$ and any ideal $\mathfrak{s}$, $\mathfrak{s}^k$ denotes the ideal generated by all products of the form $h_1 h_2 \cdots h_k$, $h_i \in \mathfrak{s}$, $i = 1, 2, \ldots, k$.

If A and B are polynomials and $\mathfrak{s}$ is an ideal in $\mathbf{Z}[x, x_2, \ldots, x_t]$, we define $A \equiv B \mod \mathfrak{s}$ if $A - B \in \mathfrak{s}$, i.e., if $A - B$ is divisible by an element of $\mathfrak{s}$. For example, if $\mathfrak{s} = (x_2 - a_2, x_3 - a_3, \ldots, x_t - a_t)$, $a_i \in \mathbf{Z}$, then $A(x, x_2, \ldots, x_t) \equiv A(x, a_2, \ldots, a_t) \mod \mathfrak{s}$ for $A(x, a_2, \ldots, a_t)$ is the remainder of dividing A by every $x_i - a_i$, $i = 2, \ldots, t$. $\mathfrak{s}^k$ is the ideal generated by all polynomials of the form

$$\prod_{i=2}^{t} (x_i - a_i)^{c_i} \quad \text{with} \quad \sum_{i=2}^{t} c_i = k, \qquad c_i \geqslant 0.$$

For this ideal $\mathfrak{s}$ we define, for any positive integer k,

$$A = B \mod \mathfrak{s}^k \quad \text{if } A \equiv B \mod \mathfrak{s}^k \quad \text{and} \quad \deg(A) \text{ in } x_2, \ldots, x_t < k.$$

Similarly, $A = B \mod(q)$ for any prime power $q > 2$ if $A \equiv B \mod(q)$ and the coefficients of A are between $-q/2$ and $q/2$.

3. An Outline of the Factoring Algorithm. An overall view of the algorithm is presented in the form of a brief description of each key step. Details and examples are included in later sections. To begin with, we have a primitive and squarefree polynomial $U(x, x_2, \ldots, x_t) \in \mathbf{R}[x, x_2, \ldots, x_t]$ and a monic minimal polynomial $f(x) \in \mathbf{Z}[x]$ with $\deg(f) = m$. We can assume that the constant coefficients of U are of the form $\Sigma_{i=0}^{m-1} c_i \theta^i$, $c_i \in \mathbf{Z}$. The algorithm takes the following steps in obtaining the irreducible factorization of U.

I. *Obtaining an integral basis.* Compute Discr(f) = Res$(f(x), f'(x))$. Set Δ to the largest integer such that Δ^2 divides Discr(f). Then the set $\{1/\Delta, \theta/\Delta, \ldots, \theta^{m-1}/\Delta\}$ forms an integral basis of $\mathbf{R}$. For certain forms of $f(x)$ smaller values of Δ are known (see Section 6). In such cases these smaller values are used.

II. *Substitution.*

(i) Selecting integers. Find a set of integers $\{a_2, a_3, \ldots, a_t\}$ (not necessarily distinct) such that $U(x, a_2, \ldots, a_t)$ remains squarefree and has the same degree as $U(x, x_2, \ldots, x_t)$ in the main variable x. The a_i should be small in absolute value. Best values for the a_i are $0, \pm 1$, in that order [11].

(ii) Normalizing the leading coefficient. Compute the inverse of $\alpha = \text{lc}(U(x, a_2, \ldots, a_t)) \in \mathbf{R}$. Let

$$\alpha^{-1} = \sum_{i=1}^{m-1} c_i \theta^i / \delta, \qquad c_i \in \mathbf{Z}, \quad \delta \in \mathbf{Z}.$$

Set $\tilde{U}(x) = \delta\alpha^{-1}U(x, a_2, \ldots, a_t)$ so that $\mathrm{lc}(\tilde{U})$ is an integer.

III. *Choosing a prime.* Find a small prime $p \in \mathbf{Z}$ which satisfies the following three conditions: (1) p does not divide $\mathrm{lc}(\tilde{U})$, (2) $\tilde{U}(x)$ is squarefree modulo p, and (3) $f(x)$ is irreducible over $\mathbf{Z}_p$. If such a prime p is not found within a given number of trials, the algorithm then uses a different factoring procedure as described in Section 8. Otherwise, the algorithm continues to the next step.

IV. *Factoring over* $\mathrm{GF}(p^m)$. Compute $u(x) = \tilde{U}(x) \bmod(p)$. Factor $u(x)$ over $\mathrm{GF}(p^m)$ into irreducible factors (see Section 5):

$$u(x) = u_1(x)u_2(x) \cdots u_r(x). \tag{1}$$

If $r = 1$, then $u(x)$ is irreducible over $\mathbf{Z}_p(\theta)$ which implies that $\tilde{U}(x)$, and therefore, $U(x, x_2, \ldots, x_t)$ are irreducible over $\mathbf{K}$. The algorithm ends in this case. If there are several small primes that satisfy the requirements in step III, it is usually advantageous to try more than one prime in this step. The smallest prime that produces the minimum r will be used.

V. *Construction of factors of* $\tilde{U}(x)$.

(i) Coefficient bound. Find a number $\tilde{B}$ such that for any rational number β in any coefficient of any divisor of $\tilde{U}(x)$, $\tilde{B} > \Delta\beta$. Let d be the least integer such that $p^{2^d} > 2\,\mathrm{lc}(\tilde{U})\tilde{B}$. Let $\tilde{b} = p^{2^d}$.

(ii) Constructing factors. From (1) we have

$$\tilde{U}(x) \equiv u_1(x)u_2(x) \cdots u_r(x) \mod(p). \tag{2}$$

A "p-adic" algorithm by Zassenhaus is used to construct from (2) factors $\hat{u}_1(x)$, $\hat{u}_2(x), \ldots, \hat{u}_r(x)$ such that $\hat{u}_i(x) \equiv \hat{u}_i \mod(p)$ and $\tilde{U}(x) \equiv \hat{u}_1\hat{u}_2 \cdots \hat{u}_r \mod(\tilde{b})$.

VI. *Actual factors of* $\tilde{U}(x)$ *over* $\mathbf{R}$. The algorithm TRUEFACTORS in Section 7 is applied with respect to the ideal $(\tilde{b})$ to obtain from the $\hat{u}_i$ a factorization over $\mathbf{R}$:

$$\tilde{U}(x) = \tilde{U}_1(x)\tilde{U}_2(x) \cdots \tilde{U}_s(x), \qquad 1 \leqslant s \leqslant r. \tag{3}$$

If $s = 1$, U is irreducible and the algorithm terminates. The $\tilde{U}_i$ are distinct and relatively prime and they may have rational numbers in their coefficients.

VII. *Construction of factors of* U.

(i) Coefficient bound. Let $y_i = x_i - a_i$, $i = 2, \ldots, t$, and

$$V = \delta\alpha^{-1}U(x, y_2 + a_2, \ldots, y_t + a_t).$$

Find a number B such that for any rational number β in the coefficients of any factor of $\mathrm{lc}(V)V$, $B > 2\Delta\beta$. Let d be the smallest integer such that $p^{2^d} > B$. Let $b = \max(\tilde{b}, p^{2^d})$. The prime power b is used as a modulus in part (ii).

(ii) Constructing factors. First the coefficients in the $\tilde{U}_i$ are reduced modulo b. It follows from (3) that

$$V \equiv \tilde{U}_1(x) \cdots \tilde{U}_s(x) \mod(b, \mathfrak{s}),$$

where $\mathfrak{s}$ is the ideal $(y_2, y_3, \ldots, y_t)$. A Hensel type construction by Wang and Rothschild [11] is used to compute, from the above congruence, polynomials $V_i(x, y_2, \ldots, y_t)$, $i = 1, \ldots, s$, such that $V_i \equiv \tilde{U}_i(x) \mod(b, \mathfrak{s})$ and

$$V \equiv V_1(x, y_2, \ldots, y_t) \cdots V_s(x, y_2, \ldots, y_t) \mod(b, \mathfrak{s}^h),$$

where $h = 1 +$ degree of U in $x_2, x_3, \ldots, x_t$.

VIII. *Actual factors of U.* The V_i give rise to possible factors of V. Irreducible factors U_i of U are obtained from the V_i by using the algorithm TRUEFACTORS described in Section 7. All factors will be found and we obtain

$$U(x, x_2, \ldots, x_t) = \alpha\delta^{-1} U_1(x, x_2, \ldots, x_t) \cdots U_j(x, x_2, \ldots, x_t), \quad 1 \leqslant j \leqslant s.$$

4. An Example. In this section the factoring algorithm is applied to a specific polynomial in three variables. The computation follows the steps outlined in the previous section. Let θ be a root of $f(x) = x^4 + x^3 + x^2 + x + 1 = 0$ and let $\mathbf{K} = \mathbf{Q}(\theta)$. The polynomial to be factored over $\mathbf{K}$ is

$$\begin{aligned} U(x, y, z) = {} & x^8 + 2x^7 + (-y - z^2 - 8)x^6 + (-4y + 6z^2 - 40)x^5 \\ & + (y^2 + (2z^2 - 48)y + z^4 + 32z^2 + 256)x^4 \\ & + (-4y^2 + (2z^2 + 32)y - 4z^4 + 32z^2 + 960)x^3 \\ & + (-y^3 + (-3z^2 + 28)y^2 + (2z^4 - 4z^2 + 384)y - z^6 - 32z^4 \\ & \qquad + 144z^2 - 1152)x^2 \\ & + (2y^3 + (-4z^2 + 72)y^2 + (6z^4 + 24z^2 - 576)y \\ & \qquad + 2z^6 - 48z^4 - 576z^2 + 3456)x \\ & + y^4 + (-z^2 - 12)y^3 \\ & + (z^4 + 24z^2 + 144)y^2 + (-z^6 + 24z^4 - 432z^2 - 1728)y \\ & + z^8 - 12z^6 + 144z^4 - 1728z^2 + 20736, \end{aligned}$$

which is primitive and squarefree. If x is chosen the main variable, then $t = 3$, $m = 4$, $n = 8$.

I. Since $f(x)$ is a cyclotomic polynomial, we know $\Delta = 1$ and $\{1, \theta, \theta^2, \theta^3\}$ is an integral basis of $\mathbf{R}$.

II. The values $a_2 = a_3 = 0$ are selected and

$$\begin{aligned} \tilde{U}(x) = U(x, 0, 0) = {} & x^8 + 2x^7 - 8x^6 - 40x^5 + 256x^4 \\ & + 960x^3 - 1152x^2 + 3456x + 20736. \end{aligned}$$

III. The primes 7, 13 and 17 are found to satisfy the three conditions in Step III. Both 7 and 13 give eight factors in Step IV while 17 gives only four factors. Hence, $p = 17$.

IV. $U(x) = x^8 + 2x^7 - 8x^6 - 6x^5 + x^4 + 8x^3 + 4x^2 + 5x - 4 \equiv \tilde{U}(x) \mod(17)$. And it is found that

$$\begin{aligned} U(x) \equiv {} & (x^2 - 2\theta x - 5\theta^3)(x^2 - 2\theta^2 x - 5\theta) \\ & \cdot (x^2 - 2\theta^3 x + 5\theta^3 + 5\theta^2 + 5\theta + 5) \\ & \cdot (x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x - 5\theta^2) \mod(17). \end{aligned}$$

See the example in Section 5 for details of this step.

V. The computer program computed $17^{2^3} = 6975757441$ as a coefficient bound at this point. For simplicity let us use $17^4 = 83521$ as a bound in both the univariate and multivariate stages. The Hensel construction gives

$$\begin{aligned}\tilde{U}(x) \equiv\ & (x^2 - 2\theta x + 12\theta^3)(x^2 - 2\theta^2 x + 12\theta) \\ & \cdot (x^2 - 2\theta^3 x - 12(\theta^3 + \theta^2 + \theta + 1)) \\ & \cdot (x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x + 12\theta^2) \mod(17^2).\end{aligned}$$

It turns out that the same congruence holds $\mod(17^4)$.

VI. Division tests in algorithm TRUEFACTORS show that the above congruence is actually an equality in **K**. Thus, $U(x, y, z)$ has no more than four irreducible factors.

VII. Since $a_2 = a_3 = 0$, the ideal $\mathfrak{s} = (y, z)$. We have

$$\begin{aligned}U(x, y, z) \equiv\ & (x^2 - 2\theta x + 12\theta^3)(x^2 - 2\theta^2 x + 12\theta) \\ & \cdot (x^2 - 2\theta^3 x - 12(\theta^3 + \theta^2 + \theta + 1)) \\ & \cdot (x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x + 12\theta^2) \mod(17^4, \mathfrak{s}),\end{aligned}$$

$$\begin{aligned}U(x, y, z) \equiv\ & (x^2 - 2\theta x + \theta^2 y + 12\theta^3)(x^2 - 2\theta^2 x - (\theta^3 + \theta^2 + \theta + 1)y + 12\theta) \\ & \cdot (x^2 - 2\theta^3 + \theta y - 12(\theta^3 + \theta^2 + \theta + 1)) \\ & \cdot (x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x + \theta^3 y + 12\theta^2) \mod(17^4, \mathfrak{s}^2),\end{aligned}$$

$$\begin{aligned}U(x, y, z) \equiv\ & (x^2 - 2\theta x + \theta^2 y - (\theta^3 + \theta^2 + \theta + 1)z^2 + 12\theta^3) \\ & \cdot (x^2 - 2\theta^2 x - (\theta^3 + \theta^2 + \theta + 1)y + \theta^3 z^2 + 12\theta) \\ & \cdot (x^2 - 2\theta^3 x + \theta y + \theta^2 z^2 - 12(\theta^3 + \theta^2 + \theta + 1)) \\ & \cdot (x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x + \theta^3 y + \theta z^2 + 12\theta^2) \mod(17^4, \mathfrak{s}^3).\end{aligned}$$

VIII. There is no need to go to a higher power of $\mathfrak{s}$ since the last congruence is an actual equality over **K**.

5. Univariate Factorization Over $\mathbf{Z}_p(\theta)$. Let $u(x)$ be a polynomial of degree n in $\mathbf{Z}_p(\theta)[x]$, where p is a small prime in $\mathbf{Z}$ and θ is a zero of the minimal polynomial $f(x) \in \mathbf{Z}[x]$. Let $\deg(f) = m$ and $q = p^m$. An algorithm for the complete factorization of $u(x)$ over $\mathbf{Z}_p(\theta)$ is given. All arithmetic is in $\mathbf{Z}_p(\theta)$ which is isomorphic to $\mathrm{GF}(q)$. The main ideas are due to Berlekamp [1] and Zassenhaus [13].

As a first step, a basis $\{v_1(x), v_2(x), \ldots, v_r(x)\}$ for the solution space of

$$v(x)^q \equiv v(x) \mod u(x)$$

is computed by finding the null space of the matrix $Q - I$, where I is the $n \times n$ identity matrix and Q is the $n \times n$ matrix whose ith row is the coefficient vector of the remainder of $x^{q(i-1)}$ divided by $u(x)$. Here the principal computation involved is the triangularization of $Q - I$.

Now, if $m = 1$, i.e., $\mathbf{Z}_p(\theta) = \mathbf{Z}_p$, we can factor $u(x)$ directly from

$$u(x) = \prod_{\alpha \in \mathrm{GF}(q)} \mathrm{GCD}(u(x), v_i(x) - \alpha), \qquad i = 1, 2, \ldots, r;$$

because $q = p$ is small. In the case $m > 1$, the size of GF(q) usually makes this straightforward approach of trying every element in GF(q) unfeasible. What is needed here is a way of finding, for a given $v_i(x)$, all the $\alpha \in \mathrm{GF}(q)$ that make $\mathrm{GCD}(u(x), v_i(x) - \alpha) \neq 1$. We call such an α *nontrivial.* A method to this end has been suggested by Zassenhaus [13]. The residues modulo $u(x)$ of $1, v_i(x), v_i(x)^2, \ldots$ are computed until a power of $v_i(x)$ which is linearly dependent on the previous powers is found. It can be shown that $1, v_i(x), \ldots, v_i(x)^r$ are always linearly dependent modulo $u(x)$. The linear dependence relation is in the form of a monic polynomial $G_i(v_i(x)) = \Sigma g_j v_i(x)^j \equiv 0 \mod u(x)$. It can be shown that

$$G_i(v_i(x)) = \prod (v_i(x) - \alpha), \qquad \alpha \text{ nontrivial.}$$

Hence, $G_i(x)$ splits and its roots are the nontrivial α's for $v_i(x)$. Thus, the problem of finding factors of $u(x)$ is reduced to that of finding the roots of $G_i(x)$ in GF(q). For p small, the roots of a nonlinear polynomial $G(x)$ which splits over GF(p^m) can be computed using an algorithm of Berlekamp [1]. Let

$$\mathrm{Tr}(x) = \sum_{i=0}^{m-1} x^{p^i} \mod G(x).$$

Berlekamp shows that the relations

$$G(x) = \prod_{\beta \in \mathbf{Z}_p} \mathrm{GCD}(G(x), \mathrm{Tr}(\theta^j x) - \beta), \qquad j = 0, 1, \ldots, m-1,$$

lead to all the linear factors of $G(x)$ over GF(q). In actual computation the residues of $x^p, x^{p^2}, \ldots, x^{p^{m-1}}$ generated in computing $\mathrm{Tr}(x)$ are stored for possible later use in calculating $\mathrm{Tr}(\theta^j x)$.

It frequently happens that the prime p chosen in step II causes $u(x)$ to split over GF(q). Thus, in our process for generating the matrix Q, the residues of $x^p, x^{p^2}, \ldots, x^{p^{m-1}} \mod u(x)$ are stored away. If the algorithm finds $r = n$, then the procedure for obtaining nontrivial α's is bypassed and the linear factors of $u(x)$ are found directly by the above root finding procedure.

As an example, let us consider factoring the squarefree polynomial

$$u(x) = x^8 + 2x^7 - 8x^6 - 6x^5 + x^4 + 8x^3 + 4x^2 + 5x - 4$$

over $\mathbf{Z}_{17}(\theta)$, where θ is a zero of $f(x) = x^4 + x^3 + x^2 + x + 1$.

Triangularization of the matrix Q gives (see [3]) $\{x^5 + x^4 - 8x^3 - 3x^2 - 7x, x^6 + 2x^4 - 2x^3 + 8x^2 + 8x, x^7 + 2x^4 + 5x^3 + 8x^2 - 5x, 1\}$ as a basis of the solution space of $v(x)^{83521} \equiv v(x) \mod u(x)$. This means that $u(x)$ has 4 irreducible factors. Taking $v_1(x)$ to be the first polynomial in this basis, we find

$$G_1(x) = x^4 - 3x^3 + 2x^2 + 6x + 2,$$

with the property $G_1(v_1(x)) \equiv 0 \mod u(x)$. The four roots of G_1 are then found to be $3\theta^3 + 4\theta^2 + 4\theta - 5$, $\theta^3 + 3\theta^2 + \theta - 8$, $-\theta^2 - 4\theta + 8$ and $-4\theta^3 - \theta + 8$. These roots turn out to be sufficient for obtaining the four factors of $u(x)$ by GCD computa-

tions:

$$u(x) = (x^2 - 2\theta^2 x - 5\theta)(x^2 + 2(\theta^3 + \theta^2 + \theta + 1)x - 5\theta^2)$$
$$\cdot (x^2 - 2\theta^3 x + 5(\theta^3 + \theta^2 + \theta + 1))(x^2 - 2\theta x - 5\theta^3).$$

6. Coefficient Bound. Factors of U and $\tilde{U}$ can be assumed to have coefficients in $\mathbf{R}$. Thus, coefficients of factors of U or $\tilde{U}$ when written in terms of the chosen integral basis $\{1/\Delta, \theta/\Delta, \ldots, \theta^{m-1}/\Delta\}$ are in the form $\Sigma_{i=0}^{m-1} c_i\theta^i/\Delta$ where $c_i \in \mathbf{Z}$. The factoring algorithm depends on finding upper bounds for the magnitude of the c_i in both the univariate and the multivariate stages of computation.

It is advantageous to have Δ as small as possible. In some cases the smallest Δ is known. If $f(x) = x^2 + a$, a squarefree, we have $\Delta = 1$ when $a \equiv 2$ or $3 \mod (4)$; and $\Delta = 2$ when $a \equiv 1 \mod(4)$. Also, $\Delta = 1$ if $f(x)$ is a cyclotomic polynomial.

If $P(x) = \Sigma_{i=0}^{n} p_i x^i$ is a polynomial with complex coefficients, we define $\|P\| = (\Sigma|p_i|^2)^{1/2}$. Let $z_1, z_2, \ldots, z_k$ (distinct or not) be those zeros of $P(x)$ with absolute value $\geqslant 1$. Mignotte [4] has shown that

$$|p_n| \prod_{i=1}^{k} |z_i| \leqslant \|P\| \quad \text{and} \quad |p_i| \leqslant \binom{n}{i} |z_1 \cdots z_k|\,|p_n|.$$

From this we can deduce the following lemma.

LEMMA. *If b is a coefficient of any primitive factor $g(x)$ of $\tilde{U}(x)$, and if* $\operatorname{lc}(g)$ *is a rational integer, then* $|b| < \binom{n}{n/2}\|\tilde{U}\|$.

Proof. If $g(x) = b_0 + b_1 x + \cdots + b_j x^j$, $b_j \in \mathbf{Z}$, then $|b_i| \leqslant \binom{j}{i}\|\tilde{U}\|$; because $\operatorname{lc}(\tilde{U}) \in \mathbf{Z}$ and $|\operatorname{lc}(\tilde{U})| \geqslant |b_j|$. Since $\deg(\tilde{U}) = n$, the lemma follows.

Recall that θ is a root of the minimal polynomial $f(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$. Let $\|\theta\|$ be the largest absolute value of any of its conjugates: θ, $\theta_2, \ldots, \theta_{m-1}$. $\|\theta\|$ is bounded by the largest positive root of

$$x^m - |a_{m-1}|x^{m-1} - \cdots - |a_1|x - |a_0| = 0.$$

If $c \in \mathbf{R}$ and $|c| < B$, and if c is expressed in the form $c = \Sigma_{i=0}^{m-1} c_i\theta^i/\Delta$, $c_i \in \mathbf{Z}$, then Weinberger [12] shows that

$$\max|c_i| \leqslant \Delta B m!\,\|\theta\|^{m-1}/\det(M),$$

where $\det(M)$ is the determinant of the matrix

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta & \theta_2 & \cdots & \theta_{m-1} \\ \theta^2 & \theta_2^2 & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \theta^{m-1} & \theta_2^{m-1} & \cdots & \theta_{m-1}^{m-1} \end{pmatrix}$$

Thus, a bound for the c_i exists and can be computed in the univariate case. Since the factorization of any multivariate polynomial can be done by univariate factoring using the Kronecker method, a coefficient bound in the multivariate case can also be computed. These theoretical bounds are often too large. In our factoring program we provide the following optional heuristic bounds that are much smaller and easier to compute:

$$\widetilde{B} = \Delta\binom{n}{n/2}\|\widetilde{U}\| \quad \text{and} \quad B = \mathrm{MAX}\left(\widetilde{B}, \Delta h\binom{n}{n/2}v_{\max}\right),$$

where $v_{\max}$ is the greatest absolute value of the integer coefficients in $V(x, y_2, \ldots, y_t)$. Care should be taken in using the heuristic bounds for they may not, though very rarely, be large enough and thus produce reducible factors.

7. **Obtaining True Factors.** Recall that

$$V(x, y_2, \ldots, y_t) = U(x, y_2 + a_2, \ldots, y_t + a_t).$$

We have the factorization

$$V(x, y_2, \ldots, y_t) = V_1(x, y_2, \ldots, y_t) \cdots V_s(x, y_2, \ldots, y_t) \mod(b, \mathfrak{s}^h),$$

where $s \geqslant 2$ and $\mathfrak{s}$ is the ideal $(y_2, y_3, \ldots, y_t)$. The V_i are distinct and irreducible and $V_i \equiv \widetilde{U}_i(x) \mod(b, \mathfrak{s})$. The V_i are unique up to units in the quotient ring $\Phi = \mathbf{R}[y_2, \ldots, y_t]/(b, \mathfrak{s}^h)$.

If U is monic, then V and the V_i are all monic and any irreducible factor $G(x, y_2, \ldots, y_t)$ of V over $\mathbf{R}$ satisfies $G \equiv H \mod(b)$, where H is either some V_i or the product of two or more V_i reduced $\mod(b, \mathfrak{s}^h)$. Then G is computed as

$$G = H^*/\Delta \quad \text{over } \mathbf{R} \text{ where } H^* = \Delta H \mod(b).$$

If U is not monic, then $G \equiv H \mod(b, \mathfrak{s}^h)$ up to units in Φ. Thus, if

$$H^* = \Delta\, \mathrm{lc}(V)\mathrm{lc}(H)^{-1} H \mod(b, \mathfrak{s}^h),$$

then $G = \mathrm{pp}(H^*/\Delta)$ over $\mathbf{R}$. The quantity $\mathrm{lc}(V)\mathrm{lc}(H)^{-1}$ is computed from the leading coefficients of the V_i. For example,

$$\mathrm{lc}(V)\mathrm{lc}(V_i)^{-1} \equiv \prod_{j=1, j\neq i}^{t} \mathrm{lc}(V_j) \mod(b, \mathfrak{s}^h).$$

In actual computation, the H^* are formed in a systematic and efficient manner from the V_i by multiplying an increasing number of them together modulo $(b, \mathfrak{s}^h)$. Any H^* that divides $\mathrm{lc}(V)V$ over $\mathbf{K}$ produces a true factor $\mathrm{pp}(H^*/\Delta)$ of V over $\mathbf{R}$. True factors of $U(x, x_2, \ldots, x_t)$ are obtained from those of V by the substitutions $y_i = x_i - a_i$, $2 \leqslant i \leqslant t$. The reader is referred to [11] for more details.

8. **Univariate Factorization Over K by Multivariate Factorization Over Z.** Rational primes satisfying conditions (1) and (2) in step III exist. They are small and easy to find in almost all cases. However, condition (3) is more difficult to meet because there are polynomials that are irreducible over the rational integers which are reducible modulo any rational prime. If the given minimal polynomial belongs to this class, no prime p can be found such that f is irreducible over $\mathbf{Z}_p$. Our factoring pro-

gram attempts to find a suitable p by trying members from a list of small primes. If a suitable p cannot be found after a set number of trials, the algorithm proceeds with the factorization of $\tilde{U}(x)$ over $\mathbf{K} = \mathbf{Q}(\theta)$ by a different procedure which is not dependent on factorization over $\mathbf{Z}_p(\theta)$.

Let $g(x, \theta)$ be a polynomial in x with coefficients in $\mathbf{R}$ which is given by the minimal polynomial $f(x) \in \mathbf{Z}[x]$. Since $\deg(f) = m$, there are m conjugates of θ: $\theta_1, \theta_2, \ldots, \theta_m$. The *norm* of $g(x, \theta)$, $N(g(x, \theta))$, is defined as

$$N(g(x, \theta)) = \prod_{i=1}^{m} g(x, \theta_i).$$

$N(g(x, \theta))$ is equal to the resultant of $g(x, \theta)$ and $f(\theta)$ with respect to θ, $\operatorname{Res}(g(x, \theta), f(\theta))$ and $N(g(x, \theta)) \in \mathbf{Z}[x]$.

Given $\tilde{U}(x, \theta)$ and $f(x)$, $\tilde{U}$ can be factored over $\mathbf{K}$ by the following procedure:

(a) Compute $V(x, y) = \operatorname{Res}(\tilde{U}(x - y\theta, \theta), f(\theta))$ over $\mathbf{Z}$.

(b) Factor $V(x, y)$ into irreducible factors over $\mathbf{Z}$ (see [11]).

$$V(x, y) = V_1(x, y)V_2(x, y) \cdots V_s(x, y).$$

(c) Compute the contents with respect to the variable y,

$$c_i(x, \theta) = \operatorname{CONT}(V_i(x + y\theta, y)).$$

(d) The irreducible factorization of $\tilde{U}(x)$ over $\mathbf{K}$ is given by

$$\tilde{U}(x, \theta) = \prod_{i=1}^{s} \operatorname{GCD}(\tilde{U}(x, \theta), c_i(x, \theta)),$$

with the GCD computed over $\mathbf{K}$.

The above procedure, when used, replaces steps IV, V and VI in our algorithm. The prime power needed as a modulus in the construction of factors of U can be formed with any prime p that satisfies the first two conditions in step III.

A proof for this procedure can be found in [10, pp. 136–137]. It can be seen that if $\deg(\tilde{U}) = n$ and $\deg(f) = m$, the degree of $V(x, y)$ is mn in either variable. Therefore, almost all the work in this procedure lies in the factorization of $V(x, y)$. Although a rather efficient algorithm for multivariate factoring is available, it is still best to use this method only when a suitable prime cannot be found after considerable effort.

Appendix. Eleven examples of factoring polynomials over algebraic number fields are given. They are done by the MACSYMA system (version 254) at Project MAC, M.I.T. In MACSYMA, the command FACTOR(U, $f(\theta)$) causes the polynomial U to be factored over $\mathbf{Q}(\theta)$ with $f(\theta)$ the given minimal polynomial. If $f(\theta)$ is omitted, then it means factoring over $\mathbf{Q}$. The command GFACTOR is implemented for the convenient use of factoring over Gaussian integers. It is equivalent to FACTOR(U, $A^2 + 1$). In MACSYMA, labels (Ci) and (Di) are used for the ith command and display lines, respectively. The symbol *%I* is used for $\sqrt{-1}$ and *%* for the previous expression. The times indicated are in milliseconds measured on a PDP-10 computer with a memory-cycle time of about two microseconds.

```
                                        4
(D1)                                   X  - 1

(C2) GFACTOR(%);
TIME= 625 MSEC.
(D2)                          (X - 1) (X + 1) (X + %I) (X - %I)

                          4      3      3       2     2
(D3)                     X  + %I X  + 2 X  + 2 %I X  + 5 X  + 2 %I X + 6 X + 6

(C4) GFACTOR(%);
TIME= 1280 MSEC.
                                                              2
(D4)                      (%I + X + 1) (- %I + X + 1) (X %I + X  + 3)

                        4      4        3     3        2     2
(D5)                2 %I X  + 3 X  + 3 %I X  - 2 X  - 2 %I X  - 2 X  + %I X - 1

(C6) GFACTOR(%);
TIME= 11866 MSEC.
                                    2                                 2
(D6)                  (- 2 %I + 13 X  + 3) (2 %I + 3) (X %I + X  - 1)
                      ----------------------------------------------
                                            13


                                        2   2
(D7)                                   Y  + X

(C8) GFACTOR(%);
TIME= 807 MSEC.
(D8)                                (X %I + Y) (Y - X %I)

                                        2
(D9)                                   X  + X - 1

(C10) FACTOR(%,A^2-5);
TIME= 745 MSEC.
                                    (2 X + A + 1) (2 X - A + 1)
(D10)                               --------------------------
                                                4

                                       4     2
(D11)                                 X  + 3 X  + 4

(C12) FACTOR(%,A^2+A+2);
TIME= 2563 MSEC.
(D12)                         (X + A) (X + A + 1) (X - A) (X - A - 1)

                                        6
(D13)                                  64 X  - 4

(C14) FACTOR(%,A^3+2);
TIME= 5938 MSEC.
                                          2          2     2          2
(D14)               (2 X + A) (2 X - A) (4 X  - 2 A X + A ) (4 X  + 2 A X + A )


                                     4     3     2
(D15)                               16 X  + 8 X  + 4 X  + 2 X + 1

(C16) FACTOR(%,A^4+A^3+A^2+A+1);
TIME= 12777 MSEC.
                                        2          3   2              3
(D16)                 (2 X - A) (2 X - A ) (2 X + A  + A  + A + 1) (2 X - A )

                                        4   4
(D17)                                  X  + Y

(C18) FACTOR(%,A^4+1);
TIME= 27702 MSEC.
                                                  3          3
(D18)                       (X + A Y) (X - A Y) (X + A  Y) (X - A  Y)
```

```
                 8      7             2      6              2      5
(D19)     X  + 2 X  + (- Y - Z  - 8) X  + (- 4 Y + 6 Z  - 40) X

     2       2              4       2        4       2       2              4       2       3
 + (Y  + (2 Z  - 48) Y + Z  + 32 Z  + 256) X  + (- 4 Y  + (2 Z  + 32) Y - 4 Z  + 32 Z  + 960) X

      3        2        2       4       2              6       4       2         2
 + (- Y  + (- 3 Z  + 28) Y  + (2 Z  - 4 Z  + 384) Y - Z  - 32 Z  + 144 Z  - 1152) X

      3        2        2       4       2              6       4       2         4
 + (2 Y  + (- 4 Z  + 72) Y  + (6 Z  + 24 Z  - 576) Y + 2 Z  - 48 Z  - 576 Z  + 3456) X + Y

      2        3     4       2       2       6       4       2              8       6       4
 + (- Z  - 12) Y  + (Z  + 24 Z  + 144) Y  + (- Z  + 24 Z  - 432 Z  - 1728) Y + Z  - 12 Z  + 144 Z

          2
 - 1728 Z  + 20736

(C20) FACTOR(%,A^4+A^3+A^2+A+1);
TIME= 118613 MSEC.
          2   3                2      3      2            2
(D20) (A Z  + A  Y + X  + (2 A  + 2 A  + 2 A + 2) X + 12 A )

   2  2          2      3        3      2
 (A  Z  + A Y + X  - 2 A  X - 12 A  - 12 A  - 12 A - 12)

   3  2       3    2              2      2
 (A  Z  + (- A  - A  - A - 1) Y + X  - 2 A  X + 12 A)

      3    2          2  2     2              3
 ((- A  - A  - A - 1) Z  + A  Y + X  - 2 A X + 12 A )


      5          3          3       2 2     2  2     2  2       2 2       3
(D21) X  - 5 V Y X  - 5 U Z X  + 5 U Y  X + 5 Z  Y X + 5 V  Z X + 5 V U  X  - 5 Z Y  X

      2 2                  3          3       2 2       3       5         3       2 2
 + 5 V  Y  X - 5 V U Z Y X  - 5 U  Y X - 5 V Z  X + 5 U  Z  X - 5 V  U X + Y  - 5 V U Y  + 5 V Z  Y

      2   2       3       3         2 2     5     2   2       3     5   5
 + 5 U  Z Y  - 5 U Z  Y - 5 V  Z Y + 5 V  U  Y + Z  + 5 V  U Z  - 5 V U  Z + U  + V

(C22) FACTOR(POLY,A^4+A^3+A^2+A+1);
TIME= 94927 MSEC.
                                 3    2              3      2
(D22) (X + Y + Z + U + V) (V (- A  - A  - A - 1) + U A  + Z A  + Y A + X)

        3    2                3      2                3    2                3      2
 (U (- A  - A  - A - 1) + Y A  + V A  + Z A + X) (Z (- A  - A  - A - 1) + V A  + Y A  + U A + X)

        3    2                3      2
 (Y (- A  - A  - A - 1) + Z A  + U A  + V A + X)
```

Massachusetts Institute of Technology
Laboratory for Computer Science and Department of Mathematics
Cambridge, Massachusetts 02139

1. E. R. BERLEKAMP, "Factoring polynomials over large finite fields," *Math. Comp.*, v. 24, 1970, pp. 713–735. MR 43 #1948.

2. E. R. BERLEKAMP, "Factoring polynomials over finite fields," *Bell System Tech. J.*, v. 46, 1967, pp. 1853–1859. MR 36 #2314.

3. D. E. KNUTH, *The Art of Computer Programming,* Vol. 2: *Seminumerical Algorithms,* Addison-Wesley, Reading, Mass., 1969. MR 44 #3531.

4. M. MIGNOTTE, "An inequality about factors of polynomials," *Math. Comp.*, v. 28, 1974, pp. 1153–1157.

5. D. A. MOON, *MACLISP Reference Manual,* Project MAC, M.I.T., Cambridge, Mass., April 1974.

6. J. MOSES, "Symbolic integration: The stormy decade," *Comm. ACM,* v. 14, 1971, pp. 548–560. MR 46 #8466.

7. D. R. MUSSER, "Multivariate polynomial factorization," *J. Assoc. Comput. Mach.*, v. 22, 1975, pp. 291–308.

8. H. POLLARD, *The Theory of Algebraic Numbers*, Carus Monographs Ser., no. 9, Wiley, New York, 1950. MR 12, 243.

9. R. RISCH, "The solution of the problem of integration in finite terms," *Bull. Amer. Math. Soc.*, v. 76, 1970, pp. 605–608. MR 42 #4530.

10. B. L. VAN DER WAERDEN, *Modern Algebra*. Vol. 1, Springer, Berlin, 1930; English transl., Ungar, New York, 1949. MR 10, 587.

11. P. S. WANG & L. P. ROTHSCHILD, "Factoring multivariate polynomials over the integers," *Math. Comp.*, v. 29, 1975, pp. 935–950.

12. F. WEINBERGER, "Factoring polynomials over algebraic number fields," *Trans. Mathematical Software*. (To appear.)

13. H. ZASSENHAUS, "On Hensel factorization. I," *J. Number Theory*, v. 1, 1969, pp. 291–311. MR 39 #4120.

14. *MACSYMA Reference Manual*, The MATHLAB group, Project MAC, M.I.T., Cambridge, Mass., September 1974.