Polynomial Decomposition Algorithms

Dexter Kozen[1]
Susan Landau[2*]

TR 86-773
August 1986

Department of Computer Science
Cornell University
Ithaca, NY 14853

1 Supported by NSF grant DCR-8602663.
2 Supported by NSF grant DCR-8402175.
* Department of Mathematics, Wesleyan University, Middletown, CO 06457

POLYNOMIAL DECOMPOSITION ALGORITHMS

Dexter Kozen[1]

Department of Computer Science
Cornell University
Ithaca, New York 14853

Susan Landau[2]

Department of Mathematics
Wesleyan University
Middletown, Connecticut 06457

August 5, 1986

ABSTRACT

In a recent paper [BZ], Barton and Zippel examine the question of when a polynomial $f(x)$ over a field of characteristic 0 has a nontrivial decomposition $f(x)=g(h(x))$. They give two exponential-time algorithms, both of which require polynomial factorization. We present an $O(s^2 r \log r)$ algorithm, where $r = \deg g$ and $s = \deg h$. The algorithm does not use polynomial factorization. We also show that the problem is in NC. In addition, we give a new structure theorem for testing decomposibility over any field. We apply this theorem to obtain an NC algorithm for decomposing irreducible polynomials over finite fields and a subexponential algorithm for decomposing irreducible polynomials over any field.

[1]Supported by NSF grant DCR-8602663. [2]Supported by NSF grant DCR-8402175.

1. Introduction

In a recent paper, Barton and Zippel [BZ] examine the question of when a polynomial $f(x)$ over a field K of characteristic 0 has a nontrivial decomposition $f(x)=g(h(x))$. There are a number of reasons one might seek such a decomposition. For example, Barton and Zippel use it to present a quick proof that the polynomial

$$f(x) = x^6+6x^4+x^3+9x^2+3x-5$$

can be expressed in radicals over $\mathbf{Q}$, since f is just the composition $f=g \circ h$, where

$$g(x) = x^2+x-5 \quad \text{and} \quad h(x) = x^3+3x.$$

Barton and Zippel present two algorithms for finding such a decomposition if one exists. Their algorithms require polynomial factorization, and require exponential time in the degree of f in the worst case.

In Section 3, we present two polynomial-time algorithms for this problem, neither of which uses polynomial factorization. The first requires $O(r^3s^2)$ algebraic operations, where $r=\deg g$ and $s=\deg h$, and will work in any commutative ring containing a multiplicative inverse of r. The second uses Fourier transforms to improve the bound to $O(s^2 r\log r)$ algebraic operations, and will work in any commutative ring containing a multiplicative inverse of r and appropriate roots of unity. We also observe that a general result of Valiant *et al.* [VSBR] implies that the problem can be solved in NC.

In Section 4, we give a new and simple structure theorem for determining decomposibility of polynomials. This theorem gives an algorithm for decomposing polynomials over any field. We apply this theorem to obtain an NC algorithm for decomposing irreducible polynomials over finite fields, and a subexponential $O(n^{\log n})$ algorithm for decomposing irreducible polynomials over any field admitting efficient polynomial factorization.

2. Algebraic Preliminaries

Let

$$f(x) = x^n+a_{n-1}x^{n-1}+ \cdots +a_1x+a_0$$

be a monic polynomial of degree greater than one with coefficients in a commutative ring K. A *decomposition* of f is a sequence $g_1, \ldots ,g_k$ such that $f(x)=g_1(g_2(\cdots g_k(x) \cdots))$. The g_i are called *components* of f. For any unit a of the ring K, the linear polynomials $ax+b$ and $x/a-b/a$ are inverses under composition, thus f always has trivial decompositions in which all but one of the components is linear. If all decompositions of f are trivial, then f is said to be *indecomposable.* A *complete decomposition* is one in which each of the components is of degree greater than one and is indecomposable.

Complete decompositions are not unique. Consider the two examples

$$\begin{aligned} x^4-10x^2+1 &= (x^2-10x+1)\circ(x^2) \\ &= (x^2-10x+1)\circ(x-17)\circ(x+17)\circ(x^2) \\ &= (x^2-44x+460)\circ(x^2+17), \end{aligned}$$

$$x^6 = x^2\circ x^3 = x^3\circ x^2.$$

over **Q**. In the first example, one decomposition is obtained from the other by inserting linear components. In the second, one decomposition is obtained from the other by permuting components. Ritt's first theorem (see [BZ]) states that any complete decomposition of f can be obtained from any other by inserting linear factors and permuting components.

If f is monic and has a decomposition $f=g\circ h$, where

$$\begin{aligned} g(x) &= b_r x^r + b_{r-1}x^{r-1} + \cdots + b_0 \\ h(x) &= c_s x^s + c_{s-1}x^{s-1} + \cdots + c_0, \end{aligned}$$

then $g(c_s x)$ and $h(x)/c_s$ are monic and also give a decomposition of f. The leading coefficients of g and h are invertible, since $b_r c_s^r = 1$. It follows that any complete decomposition is equivalent to a decomposition in which all the components are monic.

It will follow from Algorithm 3.4 below that if $f \in K[x]$ is decomposable over an extension of K, then it is decomposable over K, provided K contains a multiplicative inverse of r. Fried and MacRae [FM] have shown that if K is a field and $f \in K[x]$ is decomposable over an algebraic extension of K, then it is decomposable over K.

3. Decomposition Algorithms for Characteristic 0

Barton and Zippel [BZ] noted that the equation $f=g\circ h$, where g, h, and f are monic of degree r, s, and $n=rs$ respectively, results in rs equations in $r+s$ unknowns. The longest of these equations can be shown to have at least $2^{\Omega(\sqrt{s})}$ terms [HR], thus a naive implementation results in an exponential algorithm. This led Barton and Zippel to develop two other algorithms which are more efficient in practice, but still exponential in the degree of f in the worst case. These algorithms involve polynomial factorization and assume that K is a field of characteristic 0.

In this section we give an elementary algorithm requiring at most $O(r^3 s^2)$ algebraic operations in the worst case. The algorithm does not involve polynomial factorization, and works in any commutative ring containing a multiplicative inverse of r. In the presence of appropriate roots of unity, the complexity can be further reduced to $O(s^2 r \log r)$ algebraic operations.

We begin with an elementary lemma.

Lemma 3.1. Let K be a commutative ring, and let $f_1(x)$, $f_2(x)$, and $g(x) \in K[x]$ be monic. If $f_1(x)$ and $f_2(x)$ agree on their first k coefficients, then so do $f_1(x)g(x)$ and $f_2(x)g(x)$. $\square$

Let

$$\begin{aligned} f(x) &= x^{rs} + a_{rs-1}x^{rs-1} + \cdots + a_0 \\ g(x) &= x^r + b_{r-1}x^{r-1} + \cdots + b_0 \\ h(x) &= x^s + c_{s-1}x^{s-1} + \cdots + c_0 \end{aligned}$$

such that $r, s < rs = n$ and $f = g \circ h$. Let g factor in an extension of K as

$$g(x) = \prod_{i=1}^{r} (x - \beta_i).$$

Then

$$f(x) = g(h(x)) = \prod_{i=1}^{r} (h(x) - \beta_i).$$

By absorbing the constant coefficient of $h(x)$ into the β_i, we can assume w.l.o.g. that $c_0 = 0$; that is, if $f = g \circ h$, then $f = g_1 \circ h_1$, where

$$\begin{aligned} g_1(x) &= \prod_{i=1}^{r} (x - (\beta_i + c_0)) \\ h_1(x) &= h(x) - c_0. \end{aligned}$$

The coefficients of g_1 are symmetric functions of the β_i and are hence in K.

By $2(r-1)$ applications of Lemma 3.1, we have

(3.2) $\quad h(x)^r$ and $f(x) = \prod_{i=1}^{r} (h(x) - \beta_i)$ agree on their first s coefficients.

Let

$$q_k(x) = x^s + c_{s-1}x^{s-1} + \cdots + c_{s-k}x^{s-k}, \quad 0 \le k \le s.$$

Then $q_0(x)=x^s$, $q_s(x)=h(x)$, and

$$q_k(x) = q_{k-1}(x)+c_{s-k}x^{s-k}, \quad 1\le k\le s.$$

Again by Lemma 3.1,

(3.3) $\quad h(x)^r$ and $q_k(x)^r$ agree on their first $k+1$ coefficients, $0\le k\le s$.

The $k+1^{st}$ coefficient of $q_k(x)^r$ is the coefficient of x^{rs-k}; but since

$$\begin{aligned} q_k(x)^r &= (q_{k-1}(x)+c_{s-k}x^{s-k})^r \\ &= q_{k-1}(x)^r+rc_{s-k}x^{s-k}q_{k-1}(x)^{r-1}+p(x) \end{aligned}$$

where $\deg p(x)\le rs-2k$, this coefficient is just

$$d_k+rc_{s-k},$$

where d_k is the coefficient of x^{rs-k} in $q_{k-1}(x)^r$ and rc_{s-k} is the coefficient of x^{rs-k} in $rc_{s-k}x^{s-k}q_{k-1}(x)^{r-1}$. By (3.2) and (3.3), this agrees with a_{rs-k}, the $k+1^{st}$ coefficient of f, $1\le k\le s-1$. Thus if the earlier coefficients $c_{s-1},\ldots,c_{s-k+1}$ of h are known, then c_{s-k} can be determined by computing

$$c_{s-k}=\frac{a_{rs-k}-d_k}{r}, \quad 1\le k\le s-1,$$

provided K contains a multiplicative inverse of r. This gives rise to the following algorithm for computing the coefficients of h:

Algorithm 3.4. (Coefficients of h.)

Step 0. Set $q_0(x)^i := x^{is}$, $0\le i\le r$.

Step k, $1\le k\le s-1$. (Assume $q_{k-1}(x)^0,\ldots,q_{k-1}(x)^r$ are known from step $k-1$).

(a) Set $c_{s-k} := \dfrac{a_{rs-k}-d_k}{r}$, where d_k is the coefficient of x^{rs-k} in $q_{k-1}(x)^r$.

(b) Calculate c_{s-k}^j, $0\le j\le r$.

(c) Set $q_k(x)^j := (q_{k-1}(x)+c_{s-k}x^{s-k})^j = \sum_{i=0}^{j}\binom{j}{i}c_{s-k}^i x^{i(s-k)}q_{k-1}(x)^{j-i}$, $0\le j\le r$.

This procedure takes $O(r^3s^2)$ arithmetic operations. Once the coefficients of h are known, the coefficients of g may be obtained by solving the triangular system

(3.5) $$Ab=a$$

where A_{ij} is the coefficient of x^{is} in $h(x)^j$, $0 \le i,j \le r$, $b=(b_0, b_1, \ldots, b_r)^T$, and $a=(a_0, a_s, \ldots, a_{rs})^T$. The entries of A are available from the last step of Algorithm 3.4, since $h=q_s$. The matrix A is invertible over K, since it is triangular with all diagonal elements 1.

The coefficients of g and h are determined uniquely by the above procedure (up to the constant coefficient of h). However, the system is overconstrained, and the g, h produced by the above procedure may not compose to give f. This is easily checked in $O(nr)$ time, using the powers of h computed in Algorithm 3.4. The running time of the algorithm is thus dominated by Algorithm 3.4.

It can be shown that the above procedure produces an expression for each coefficient of h and g that is of degree at most r in each coefficient of f. Since these coefficients are calculated by a straight-line program involving only algebraic operations, it follows from a general simulation result of Valiant *et al.* [VSBR] that the algorithm can be parallelized so as to run in $(\log n)^{O(1)}$ parallel time, using polynomially many processors.

We now show that the $O(r^3s^2)$ sequential bound can be improved to $O(s^2r\log r)$ using the discrete Fourier transform, provided appropriate roots of unity are present.

Assume there exists $m \ge n+1$, $m=O(n)$, such that K contains a principal m^{th} root of unity ζ and a multiplicative inverse of m. The discrete Fourier transform F is represented by the $m \times m$ matrix $F_{ij}=\zeta^{ij}$, $0 \le i,j \le m-1$, and its inverse F^{-1} by $F_{ij}^{-1}=\frac{1}{m}\zeta^{-ij}$, $0 \le i,j \le m-1$. The Fourier transform $F(p)$ of the polynomial

$$p(x) = d_{m-1}x^{m-1}+d_{m-1}x^{m-1}+ \cdots +d_0$$

is the vector

$$(p(1),p(\zeta),p(\zeta^2), \ldots ,p(\zeta^{m-1}))^T = F\cdot(d_0, \ldots ,d_{m-1})^T.$$

Fourier transforms and their inverses can be computed in $O(n\log n)$ time [AHU].

The following algorithm is essentially Algorithm 3.4, except that all computations are carried out in the transformed space. The polynomials $q_k(x)$ and coefficients d_k are the same as in Algorithm 3.4.

Algorithm 3.6. (Coefficients of h.)

Step 0. Set $\hat{f} := F(f)=F\cdot(a_0, \ldots ,a_{n-1}, 1, 0, \ldots ,0)^T$.

Set $e_0 := F(q_0)=(1, \zeta^s, \zeta^{2s}, \ldots, \zeta^{(m-1)s})^T$.

Step k, $1 \le k \le s-1$. (Assume $e_{k-1}=F(q_{k-1})$.)

(a) Set $e := e_{k-1}^r$, the pointwise r^{th} power of the vector e_{k-1}. (Now $e=F(q_{k-1}(x)^r)$, and the coefficient of x^{rs-k} in $q_{k-1}(x)^r=F^{-1}(e)$ is d_k.)

(b) Set $e := \frac{1}{r}(\hat{f}-e)$. (Now the coefficient of x^{rs-k} in $F^{-1}(e)$ is $\frac{1}{r}(a_{rs-k}-d_k)=c_{s-k}$.)

(c) Set $e :=$ the pointwise product of e with the vector

$$(1, \zeta^{-(r-1)s}, \zeta^{-2(r-1)s}, \ldots, \zeta^{-(m-1)(r-1)s}).$$

(This rotates $F^{-1}(e)$ so that c_{s-k} is now the coefficient of x^{s-k} in $F^{-1}(e)$. This operation is equivalent to $e := FR_{-(r-1)s}F^{-1}(e)$, where R_k is the rotation matrix $(R_k)_{ij}=1$ if $i=j+k \bmod m$, 0 otherwise, $0 \le i,j \le m-1$.)

(d) Set $b := e^T \cdot (1, \zeta^{-(s-k)}, \zeta^{-2(s-k)}, \ldots, \zeta^{-(m-1)(s-k)})$;
set $e := (b, b\zeta^{s-k}, b\zeta^{2(s-k)}, \ldots, b\zeta^{(m-1)(s-k)})$. (This projects $F^{-1}(e)$ onto its $s-k+1^{st}$ coordinate, so that now $e=F(c_{s-k}x^{s-k})$. This operation is equivalent to $e := FB_{s-k}F^{-1}(e)$, where B_{s-k} is the projection matrix $(B_{s-k})_{ij}=1$ if $i=j=s-k$, 0 otherwise, $0 \le i,j \le m-1$.)

(e) Set $e_k := e+e_{k-1}$. (Now $e_k=F(q_{k-1}(x)+c_{s-k}x^{s-k})=F(q_k)$.)

Step s. Set $h(x) := F^{-1}(e_{s-1})$.

The Fourier and inverse Fourier transforms in steps 0 and s take $O(n\log n)$ steps each [AHU]; the running time of each of steps 1 through $s-1$ is dominated by the calculation of the pointwise r^{th} power of a vector of length m in (a), which takes $O(n\log r)$ steps. Thus the total cost is $O(s^2 r\log r)$.

4. Decomposition Algorithms for Arbitrary Fields

The algorithms of Section 3 require that the ring K contain certain multiplicative inverses. Barton and Zippel [BZ] consider only fields of characteristic 0. In this section we prove a general theorem that can be used to decompose any polynomial f over any field. For irreducible polynomials over finite fields, this yields an *NC* algorithm (parallel $(\log mnp)^{O(1)}$ time on $(mnp)^{O(1)}$ processors, where p is the characteristic of the field, m is the degree of K over $GF(p)$, and n is the degree of f). For irreducible polynomials over other fields, it gives a subexponential $O(n^{\log n})$ sequential algorithm.

Let K be an arbitrary field. Let $f \in K[x]$ be monic of degree $n=rs$. Let $\hat{K}$ denote the splitting field of f, and let G denote the Galois group of $\hat{K}$ over K.

Definition 4.1. A *block decomposition* for f is a multiset Δ of multisets of elements of $\hat{K}$ such that

(i) $f(x) = \prod_{A \in \Delta} \prod_{\alpha \in A} (x-\alpha)$

(ii) if $\alpha \in A \in \Delta$, $\beta \in B \in \Delta$, and $\sigma \in G$ such that $\sigma(\alpha)=\beta$, then
$$B = \{\sigma(\gamma) \mid \gamma \in A\}.$$

A block decomposition Δ is an $r \times s$ *block decomposition* if $|\Delta|=r$ and $|A|=s$ for all $A \in \Delta$. $\square$

Definition 4.1 generalizes the usual notion of block decomposition for irreducible polynomials over a field to the more general case where f need not be irreducible or square-free.

Let c_k^m denote the k^{th} elementary symmetric polynomial in m variables:
$$c_k^m(x_1, \ldots, x_m) = \sum_{1 \le i_1 < \cdots < i_k \le m} x_{i_1} x_{i_2} \cdots x_{i_k}.$$
We take $c_0^m = 1$ by convention. Since c_k^m is symmetric, we may write $c_k^m(A)$ for $c_k^m(\alpha_1, \ldots, \alpha_m)$ without ambiguity, where $A = \{\alpha_1, \ldots, \alpha_m\}$ is a multiset.

Theorem 4.2. Let $f \in K[x]$ be monic of degree $n=rs$. The following are equivalent:

(i) $f = g \circ h$ for some g, $h \in K[x]$ monic of degree r and s, respectively;

(ii) there exists a multiset Γ of elements of $\hat{K}$ and an $h \in K[x]$ monic of degree s such that
$$f(x) = \prod_{\gamma \in \Gamma} (h(x)-\gamma);$$

(iii) there exists an $r \times s$ block decomposition Δ for f such that
$$c_k^s(A) = c_k^s(B) \in K, \quad A, B \in \Delta, \ 0 \le k \le s-1.$$

Proof. (i) $\to$ (ii). Let Γ be the multiset of roots of g. Then
$$g(x) = \prod_{\gamma \in \Gamma} (x-\gamma),$$
and
$$f(x) = g(h(x)) = \prod_{\gamma \in \Gamma} (h(x)-\gamma).$$

$\Gamma \subseteq \hat{K}$, since each $\gamma \in \Gamma$ can be expressed $c_0 \pm \gamma'$, where c_0 is the constant coefficient of h and γ' is a product of roots of f.

(ii) $\rightarrow$ (iii). For $\gamma \in \Gamma$, let A_γ be the multiset of roots of $h(x)-\gamma$. Let $\Delta = \{A_\gamma \mid \gamma \in \Gamma\}$. Then

$$f(x) = \prod_{\gamma \in \Gamma} (h(x)-\gamma) = \prod_{\gamma \in \Gamma} \prod_{\alpha \in A_\gamma} (x-\alpha).$$

For each $\gamma \in \Gamma$, $(-1)^k c_k^s(A_\gamma)$ is the coefficient of x^{s-k} in $h(x)-\gamma$, $0 \le k \le s-1$, therefore

$$c_k^s(A_\gamma) \in K, \quad 0 \le k \le s-1.$$

It remains to show that Δ is a block decomposition. Suppose γ, $\delta \in \Gamma$, $\alpha \in A_\gamma$, $\beta \in A_\delta$, $\sigma \in G$ with $\sigma(\alpha)=\beta$. Since $h(\alpha)=\gamma$ and $h(\beta)=\delta$,

$$\sigma(\gamma) = \sigma(h(\alpha)) = h(\sigma(\alpha)) = h(\beta) = \delta.$$

Extending $\sigma: \hat{K} \rightarrow \hat{K}$ uniquely to $\sigma: \hat{K}[x] \rightarrow \hat{K}[x]$ by taking $\sigma(x)=x$, we have

$$\begin{aligned}
\prod_{\alpha \in A_\gamma} (x-\sigma(\alpha)) &= \prod_{\alpha \in A_\gamma} (\sigma(x)-\sigma(\alpha)) \\
&= \sigma\Big(\prod_{\alpha \in A_\gamma} (x-\alpha)\Big) \\
&= \sigma(h(x)-\gamma) \\
&= h(\sigma(x))-\sigma(\gamma) \\
&= h(x)-\delta \\
&= \prod_{\beta \in A_\delta} (x-\beta).
\end{aligned}$$

Thus $A_\delta = \{\sigma(\alpha) \mid \alpha \in A_\gamma\}$.

(iii) $\rightarrow$ (i). Given Δ, let $A \in \Delta$ and let

$$h(x) = \sum_{k=0}^{s-1} (-1)^k c_k^s(A) x^{s-k} \in K[x].$$

The choice of A does not matter, by the assumption of (iii). For each $A \in \Delta$, let

$$\gamma_A = (-1)^{s+1} \prod_{\alpha \in A} \alpha = (-1)^{s+1} c_s^s(A) \in \hat{K},$$

and let

$$g(x) = \prod_{A \in \Delta} (x-\gamma_A).$$

Then

$$h(x)-\gamma_A = \sum_{k=0}^{s} (-1)^k c_k^s(A) x^{s-k} = \prod_{\alpha \in A} (x-\alpha),$$

and

$$f(x) = \prod_{A \in \Delta} \prod_{\alpha \in A} (x-\alpha) = \prod_{A \in \Delta} (h(x)-\gamma_A) = g(h(x)).$$

The coefficients of h are in K by the assumption of (iii), and the coefficients of g are in K since they are solutions of the nonsingular system (3.5). $\square$

Theorem 4.2(iii) gives a simple algebraic condition that can be used to test decomposability over any field K over which the splitting field of f can be constructed.

If the polynomial f is irreducible, then in checking the condition of Theorem 4.2(iii), we need only check whether $c_k^s(A) \in K$, $0 \le k \le s-1$, for *some* $A \in \Delta$; if it holds for one such A, then it holds for all, since G is transitive on elements of Δ.

If in addition f has no multiple roots, then the block decomposition of Definition 4.1 is just the standard block decomposition of permutation group theory. In this case, G acts transitively on the set Ω of roots of f, and a *block decomposition* is a set Δ of pairwise disjoint subsets of Ω such that for all $A \in \Delta$ and $\sigma \in G$, either $\sigma A = A$ or $\sigma A \cap A = \varnothing$. The elements of Δ are called *blocks of imprimitivity*. Every group has trivial blocks $\{a\}$ and Ω. If there are no others, then G is said to act *primitively* on Ω.

Block decompositions of irreducible polynomials over finite fields are particularly easy to compute. If K is an extension of degree m over $GF(p)$ and f is irreducible over K, then $\hat{K} = K(\alpha)$ where α is any root of f, and the only automorphisms of $K(\alpha)$ over K are of the form $\beta \to \beta^{p^{km}}$, $0 \le k \le n-1$. It follows that there is a *unique* $r \times s$ block decomposition

$$\Delta = \{A_i | 0 \le i \le r-1\},$$
$$A_i = \{\alpha^{p^{(i+jr)m}} | 0 \le j \le s-1\}, \quad 0 \le i \le r-1$$

for f. In order to compute h and g, it will suffice to show how to compute representations of all the roots $\alpha^{p^{km}}$, $0 \le k \le n-1$ of f, so that one the polynomials

$$h(x)-\gamma_i = \prod_{\beta \in A_i} (x-\beta), \quad 0 \le i \le r-1$$

can be computed and the condition of Theorem 4.2(iii) checked.

We represent $\beta \in K(\alpha)$ as a polynomial $h(x) \in K[x]$, $\deg h \le n-1$, such that $h(\alpha) = \beta$. Each $\beta \in K(\alpha)$ has a unique representation relative to a given α.

Lemma 4.3. Let K be a extension of $GF(p)$ of degree m, and let $f \in K[x]$ be monic and irreducible of degree n. Let $h(x) \in K[x]$ be given, $\deg h \le n-1$, representing $\beta = h(\alpha)$ for some root α of f. Then representations of all the conjugates of β over K can be computed in NC (parallel time $\log^{O(1)} mnp$ on $(mnp)^{O(1)}$ processors).

Proof. First suppose that $\beta=\alpha$ and $h(x)=x$. The conjugates of α over K are

$$a^{p^{km}}, \quad 0\le k\le n-1.$$

Define

$$q_0(x)=x,$$
$$q_1(x)=x^p \bmod f,$$
$$q_{i+j}(x)=q_i(q_j(x)) \bmod f.$$

It is easily shown by induction on k that q_k is well-defined, and that

$$q_k(\alpha) = \alpha^{p^k}.$$

The conjugates of α over K are thus represented by q_{km}, $0\le k\le n-1$. The polynomial q_1 can be calculated in parallel time $O(\log p \log^2 n)$ by repeated squaring and reducing mod f. The polynomial q_{2k} can be computed from q_k in parallel time $O(\log^2 n)$ by composing q_k with itself and then reducing mod f. By iterating $\log mn$ times, we obtain q_{2^i}, $0\le i\le \log mn$, in parallel time $O(\log mn \log^2 n)$. Up to $\log mn$ of these can be composed and reduced mod f in parallel time $O(\log mn \log^2 n)$ to give q_k, $0\le k\le mn-1$. The total parallel time needed to produce all the q_k is then $O(\log p \log^2 n+\log mn \log^2 n)=O(\log mnp \log^2 n)$.

If $\beta\ne\alpha$ but $\beta=h(\alpha)$, where $h\in K[x]$ of degree $\le n-1$, then the conjugates of β over K are

$$h(\alpha)^{p^{km}} = h(\alpha^{p^{km}}), \quad 0\le k\le n-1,$$

and these are represented by

$$h(x^{p^{km}}) = h(q_{km}(x)) \bmod f, \quad 0\le k\le n-1.$$

The compositions $h\circ q_{km}$ can be calculated and reduced mod f in parallel time $O(\log^2 n)$. $\square$

Theorem 4.4. An irreducible polynomial f over a finite field K can be tested for the existence of a nontrivial decomposition $g\circ h$ in NC. If such a decomposition exists, the coefficients of g and h can be computed in NC.

Proof. Using Lemma 4.3, representations of the roots of f can be computed in NC. For any r, s such that $n=rs$, these roots can be arranged in the unique $r\times s$ block decomposition Δ, and for some $A\in\Delta$, the product

$$\prod_{\alpha\in A}(x-\alpha)$$

computed, and tested for the condition of Theorem 4.2(iii). This computation also gives the coefficients of h, if it exists. Finally, the triangular linear system (3.5) can be solved in NC [C] to obtain the coefficients of g. $\square$

Example 4.5. There are exactly two primitive irreducible polynomials of degree 4 over $GF(2)$, namely

$$\begin{aligned} p_1(x) &= x^4+x+1 \\ p_2(x) &= x^4+x^3+1. \end{aligned}$$

We will show that p_1 has a nontrivial decomposition and p_2 does not. First consider p_1. The unique 2×2 block decomposition is

$$\Delta = \{\{\alpha, \alpha^4\}, \{\alpha^2, \alpha^8\}\},$$

where α is a root of p_1. This reduces modulo p_1 to

$$\{\{\alpha, \alpha+1\}, \{\alpha^2, \alpha^2+1\}\}.$$

Computing the two polynomials whose roots are the two multisets of Δ, we obtain

$$\begin{aligned} (x-\alpha)(x-(\alpha+1)) &= x^2+x+(\alpha^2+\alpha) \\ (x-\alpha^2)(x-(\alpha^2+1)) &= x^2+x+(\alpha^2+\alpha+1), \end{aligned}$$

thus the requirements of Theorem 4.2(iii) are met, and we may take

$$h(x) = x^2+x.$$

The coefficients of g are obtained by solving the nonsingular triangular system (3.5). In this case, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

with solution $(1,1,1)^T$, giving

$$g(x) = x^2+x+1.$$

Computing $g(h(x))$, we find

$$g(h(x)) = (x^2+x)^2+(x^2+x)+1 = p_1(x),$$

thus $p_1 = g \circ h$.

For p_2, the unique 2×2 block decomposition is

$$\Delta = \{\{\beta, \beta^4\}, \{\beta^2, \beta^8\}\} = \{\{\beta, \beta^3+1\}, \{\beta^2, \beta^3+\beta^2+\beta\}\}$$

where β is a root of p_2. Computing the two polynomials whose roots are the two multisets of Δ, we obtain

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

neither of whose second coefficients are in $GF(2)$, therefore h does not exist, and p_2 has no nontrivial decomposition. $\square$

Theorem 4.2 also gives an $O(n^{\log n})$ algorithm for the decomposition of separable irreducible polynomials over any field admitting a polynomial-time polynomial factorization algorithm. Separability is not really a problem, since if K is a field of characteristic p, and f is inseparable and irreducible, then $f(x)=g(x^{p^k})$ for some irreducible and separable g, thus we have an immediate nontrivial decomposition. Hence we may assume without loss of generality that f is separable.

In Landau and Miller [LM], it was shown how to determine minimal blocks of the roots of f in polynomial time. A similar procedure can be used to determine *all* blocks.

By Theorem 4.2, any nontrivial decomposition $f=g \circ h$ corresponds to an $r \times s$ block decomposition Δ, where $s=\deg h$, such that $c_i^s(A) \in K$, $0 \leq i < s$, for one (and hence all) $A \in \Delta$. We first use the algorithm of [LM] to compute all minimal blocks and test this condition. For each block A which satisfies this condition, the coefficients of h are obtained immediately from the $c_i^s(A)$, and the coefficients of g are obtained by solving (3.5). The decomposition algorithm is then applied to g. Note that g is irreducible and h is indecomposable; if g were reducible, then f would have been, and if h were decomposable, then the block A would not have been minimal.

If no minimal block decomposition satisfying Theorem 4.2(iii) is found, then the algorithm is repeated using blocks whose only proper subblock is minimal. If any of these satisfy Theorem 4.2(iii), then the appropriate g and h are constructed, and the decomposition algorithm is then performed on g. If not, the next level of minimal blocks is examined. Each time a decomposition is found, the polynimial g is irreducible, and the polynomial h is indecomposable. The algorithm halts when the only block left to examine is Ω or when g is indecomposable.

If there is a polynomial-time algorithm for factoring over the base field K, then the algorithm of [LM] for finding a block is polynomial in the size of f. The difficulty is that there are potentially $n^{\log n}$ blocks, and we must search through all of them to find a decomposition. For each block, the condition of Theorem 4.2(iii) can be tested and the coefficients of g and h obtained in polynomial time. Thus we have

Theorem 4.6. A complete decomposition of an irreducible polynomial f over any field K admitting a polynomial-time polynomial factorization algorithm can be constructed in time $O(n^{\log n})$. $\square$

Finally we observe that over $\mathbf{Q}$, most irreducible polynomials are indecomposable. This is because Theorem 4.2 implies that an irreducible polynomial f over $\mathbf{Q}$ can be decomposed *only if* the Galois group of the splitting field of f over $\mathbf{Q}$ acts imprimitively on the roots of f. But Gallagher [G] has shown that most irreducible polynomials over $\mathbf{Q}$ have Galois group S_n, which acts primitively.

References

[AHU] Aho, A.V., J.E. Hopcroft, and J.D. Ullman. *The Design and Analysis of Computer Algorithms.* Addison Wesley, Reading, Mass., 1975.

[C] Csanky, L. Fast parallel matrix inversion algorithms. *SIAM J. Comput.* 5 (1976), 618-623.

[BZ] Barton, D.R., and R.E. Zippel. Polynomial Decomposition Algorithms. *J. Sym. Comp.* 1 (1985), 159-168.

[FM] Fried, M.D. and R.E. MacRae. On the invariance of chains of fields. *Ill. J. Math.* 13 (1969).

[G] Gallagher, P.X. Analytic Number Theory. In: *Proc. Symp. in Pure Mathematics*, Amer. Math. Soc., Providence, 1972, 91-102.

[HR] Hardy, G.H. and S. Ramanujan, *Proceedings, London Math. Soc. (2)* 17 (1918), 75-115.

[L] Landau, S. Factoring Polynomials over Algebraic Number Fields. *SIAM J. Comput.* 14 (1985), 184-195.

[LLL] Lenstra, A.K., H.W. Lenstra, and L. Lovasz. Factoring Polynomials with Rational Coefficients. *Math. Ann.* 261 (1982), 515-534.

[LM] Landau, S. and G. Miller. Solvability by Radicals is in Polynomial Time. *J. Comput. Syst. Sci.* 30 (1985), 179-208.

[VSBR] Valiant, Skyum, Berkowitz, and Rackoff. Fast Parallel Computation on Polynomials using Few Processors. In: Springer-Verlag Lecture Notes in Computer Science 118 (1981), 132-139.