MATHEMATICS OF COMPUTATION
VOLUME 36, NUMBER 153
JANUARY 1981

On Polynomial Factorization Over Finite Fields

By Hiroshi Gunji and Dennis Arnon

Abstract. Let $f(x)$ be a polynomial over a finite field F. An algorithm for determining the degrees of the factors of $f(x)$ is presented. As in the Berlekamp algorithm (1968) for determining the factors of $f(x)$, the Frobenius endomorphism on $F[x]/(f(x))$ plays a central role. Little-known theorems of Schwarz (1956) and Cesàro (1888) provide the basis for the algorithm we present. New and stream-lined proofs of both theorems are provided.

1. Introduction. There are a number of computational problems in which one wants the degrees of the factors of a polynomial over a finite field without needing the factors themselves. Factorization of polynomials over the rationals Q provides one example. In recent algorithms for constructing the factors of $g(x)$ in $Q[x]$, one must essentially guess their degrees, check this guess, and repeat until a correct guess is made; see e.g. Musser [4]. For almost all primes p, the degrees of the factors of $g(x) \bmod p$ restrict the possible degrees of factors of $g(x)$. Musser has exploited knowledge of these mod p factor degrees for several primes p to significantly improve the guessing process. A second example of where only factor degrees of a polynomial over a finite field are needed is the technique, due to van der Waerden (see [7, Section 8.10]), of determining subgroups of the Galois group of $g(x)$ in $Q[x]$ from knowledge of its mod p factor degrees for several primes p. Zimmer [9, p. 5] mentions additional examples.

Let F be a given finite field with q elements, where $q = p^s$ for some prime p and some $s \geqslant 1$. Let $f(x)$ be a given monic polynomial of degree $n \geqslant 1$ in $F[x]$. Let

$$f(x) = f_1(x)^{e_1} f_2(x)^{e_2} \cdots f_r(x)^{e_r}, \qquad r \geqslant 1, \tag{1.1}$$

be a complete factorization of f. That is, each f_k has degree $\geqslant 1$ and is irreducible over F; if $k_1 \neq k_2$, then f_{k_1} is not an associate of f_{k_2}, and each e_k is a positive integer. Define σ_j, $1 \leqslant j \leqslant n$, to be the number of f_k's of degree j. We have $0 \leqslant \sigma_j \leqslant r$ for each j, and

$$\sum_{j=1}^{n} \sigma_j = r.$$

Thus, if we can compute $\sigma_1, \sigma_2, \ldots, \sigma_n$ given f, we know the degrees of the (distinct) factors of f.

Let π denote the Frobenius mapping on the F-algebra $R = F[x]/(f(x))$, i.e., for any $h + (f)$ in R, $\pi(h + (f)) = h^q + (f)$. It is easily verified that π is an endomorphism of R. For $1 \leqslant i \leqslant n$, define ν_i to be the dimension of the null space of the

Received March 24, 1980.
1980 *Mathematics Subject Classification.* Primary 12C05; Secondary 10A20.
Key words and phrases. Finite fields, polynomials, factorization, inversion formulas, integer matrices, Frobenius mapping.

0025-5718/81/0000-0026/$02.75

endomorphism $\pi^i - I$ of R, where I is the identity on R. ν_1 plays a crucial role in Berlekamp's algorithm [1] for constructing the factors of a polynomial in $F[x]$.

Let A be the $n \times n$ matrix $\|A_{ij}\|$, where $A_{ij} = (i, j)$, the greatest common divisor of i and j. Let σ and ν denote the column vectors $(\sigma_1, \ldots, \sigma_n)^T$ and $(\nu_1, \ldots, \nu_n)^T$. Schwarz [5] proved

$$A\sigma = \nu.$$

Smith's formula [6] states that

$$\text{Determinant}(A) = \prod_{i=1}^{n} \varphi(i),$$

where φ is the Euler totient function. Hence, as Schwarz observed, A is invertible and σ is uniquely determined by ν. As we will see in Section 5, ν can be computed from $f(x)$.

Schwarz did not discuss the question of obtaining an explicit formula for A^{-1} (as a function of n). Clearly, it would be preferable to use such a formula to compute σ from ν rather than using, say, gaussian elimination. As Carlitz [3] points out, a formula for A^{-1} is implicit in an 1888 result of Cesàro. We give our own rather succinct derivation of the formula in the present paper. We note that Dickson [8] essentially stated the relation $\sigma = A^{-1}\nu$ and the formula for A^{-1}, without proof and without reference to Schwarz or Cesàro. It seems likely, however, that he assumed $f(x)$ to be a separable polynomial (i.e. $e_1 = e_2 = \cdots = e_r = 1$ in (1.1)).

Thus, we have a two-step algorithm for determining the degrees of the factors of $f(x)$: first we compute the length n vector ν, then multiply ν by A^{-1}, an $n \times n$ matrix of rational numbers. We begin this paper with a new proof, in Sections 2 and 3, that $A\sigma = \nu$ whether $f(x)$ is separable or not (we proceed by considering first the separable case). In Section 4 we derive the formula for A^{-1} and note that the derivation yields Smith's formula as a by-product. In Section 5 we give an example of the algorithm.

2. The Separable Case. We assume in this section that the given polynomial $f(x)$ is separable, i.e.

$$f(x) = f_1(x)f_2(x) \cdots f_r(x), \qquad r \geqslant 1,$$

where the f_k's are irreducible nonassociate polynomials of positive degree. Where $n_k \geqslant 1$ is the degree of f_k, we have $n = \sum_{k=1}^{r} n_k$. We write the definition of ν_i, $1 \leqslant i \leqslant n$, in the form

$$\nu_i = \text{dimension}(\text{kernel}(\pi^i - I)) = \dim \ker(\pi^i - I),$$

π the Frobenius on R. By the Chinese Remainder Theorem we have an isomorphism of F-algebras

$$R = \frac{F[x]}{(f(x))} \cong \frac{F[x]}{(f_1(x))} \oplus \frac{F[x]}{(f_2(x))} \oplus \cdots \oplus \frac{F[x]}{(f_r(x))}.$$

For $1 \leqslant k \leqslant r$, let R_k denote $F[x]/(f_k(x))$, π_k the Frobenius on R_k, and I_k the identity on R_k. π_k, in other words, is the map $\pi_k(h + (f_k)) = h^q + (f_k)$ for any $h + (f_k)$ in R_k. Then, for any i, $1 \leqslant i \leqslant n$, we have

$$\nu_i = \sum_{k=1}^{r} \left[\dim \ker(\pi_k^i - I_k)\right].$$

Since each f_k is irreducible, $R_k \cong \mathrm{GF}(q^{n_k})$ for each k. Noting that $h + (f_k)$ is in the kernel of $\pi_k^i - I_k$ if and only if $h^{q^i} - h$ is divisible by f_k, we have, for each k,

$$\ker(\pi_k^i - I_k) \cong \mathrm{GF}(q^i) \cap \mathrm{GF}(q^{n_k}) \cong \mathrm{GF}(q^{(i,n_k)}).$$

Hence, $\dim \ker(\pi_k^i - I_k) = (i, n_k)$. We have thus proved

THEOREM 2.1. *If $f(x)$ is a separable polynomial, then for any i, $1 \leqslant i \leqslant n$,*

$$\nu_i = \sum_{k=1}^{r} (i, n_k).$$

Recalling that, for $1 \leqslant j \leqslant n$, $\sigma_j \geqslant 0$ is the number of n_k's, $1 \leqslant k \leqslant r$, equal to j, we have immediately

COROLLARY 2.1. *If $f(x)$ is a separable polynomial, then, for any i, $1 \leqslant i \leqslant n$*

$$\nu_i = \sum_{j=1}^{n} (i, j)\sigma_j.$$

Thus, using notation from Section 1, we have

COROLLARY 2.2. *If $f(x)$ is a separable polynomial, then $\nu = A\sigma$.*

We also note the following

COROLLARY 2.3. *If $f(x)$ is a separable polynomial, then ν_1 is the number of nonassociate irreducible factors of $f(x)$. In particular, if $f(x)$ is separable, it is irreducible if and only if $\nu_1 = 1$.*

3. The General Case. We now drop the assumption that $f(x)$ is separable, i.e., in the complete factorization of $f(x)$ as given in (1.1), we do not assume that each e_k is one. Letting n_k again denote the degree of f_k, we have now $n = \sum_{k=1}^{r} e_k n_k$. The Chinese Remainder Theorem yields the following isomorphism of F-algebras

$$R = \frac{F[x]}{(f(x))} \cong \frac{F[x]}{(f_1(x)^{e_1})} \oplus \cdots \oplus \frac{F[x]}{(f_r(x)^{e_r})}.$$

For $1 \leqslant k \leqslant r$, let R_k denote $F[x]/(f_k(x)^{e_k})$, π_k the Frobenius on R_k, and I_k the identity on R_k. For any i, $1 \leqslant i \leqslant n$, we have

$$\nu_i = \dim \ker(\pi^i - I) = \sum_{k=1}^{r} \left[\dim \ker(\pi_k^i - I_k)\right].$$

Let $\bar{R}_k$ denote $F[x]/(f_k(x))$, $\bar{\pi}_k$ the Frobenius on $\bar{R}_k$, and $\bar{I}_k$ the identity on $\bar{R}_k$. If we can show for all k and i, $1 \leqslant k \leqslant r$ and $1 \leqslant i \leqslant n$, that

$$\dim \ker(\pi_k^i - I_k) = \dim \ker(\bar{\pi}_k^i - \bar{I}_k), \tag{3.1}$$

then the results of Section 2 will carry over to the present, more general, situation. We establish the validity of (3.1) with three lemmas.

For the lemmas (and only for the lemmas) we will assume given some irreducible $t(x)$ of degree $m \geqslant 1$ in $F[x]$, and a positive integer $e > 1$. We will let S denote $F[x]/(t(x)^e)$ and $\bar{S}$ denote $F[x]/(t(x))$. Again only in the lemmas, π will denote the

Frobenius on S, $\bar{\pi}$ the Frobenius on $\bar{S}$, I the identity on S, and $\bar{I}$ the identity on $\bar{S}$. Elements of S and $\bar{S}$ are cosets $h(x) + (t(x)^e)$ and $h(x) + (t(x))$, respectively, for $h(x)$ in $F[x]$. The map Ψ: $h + (t^e) \to h + (t)$ is an F-algebra homomorphism of S onto $\bar{S}$.

LEMMA 3.1. $\Psi \circ \pi = \bar{\pi} \circ \Psi$.

Proof. For any $h + (t^e)$ in S:

$$\Psi \circ \pi(h + (t^e)) = \Psi(h^q + (t^e)) = h^q + (t).$$

Also,

$$\bar{\pi} \circ \Psi(h + (t^e)) = \bar{\pi}(h + (t)) = h^q + (t).$$

Q.E.D.

COROLLARY 3.1. *For any* i, $1 \leqslant i \leqslant n$, $\Psi \circ (\pi^i - I) = (\bar{\pi}^i - \bar{I}) \circ \Psi$.

Let K^i and $\bar{K}^i$ denote $\ker(\pi^i - I)$ and $\ker(\bar{\pi}^i - \bar{I})$, respectively. Then the inclusion $\Psi(K^i) \subset \bar{K}^i$ is a direct consequence of Corollary 3.1. We shall now prove

LEMMA 3.2. *For any* i, $1 \leqslant i \leqslant n$, $\Psi(K^i) = \bar{K}^i$.

Proof. It is enough to show that $\bar{K}^i \subset \Psi(K^i)$. Let $h + (t)$ be an arbitrary element of $\bar{K}^i$, and let w in $F[x]$ be defined by

$$h^{q^i} - h = wt.$$

If wt is divisible by t^e, then $h + (t^e)$ is in K^i and $\Psi(h + (t^e)) = h + (t)$. Suppose wt is not divisible by t^e. Let

$$\tilde{h} = h + \sum_{k=0}^{c} (wt)^{q^{ik}},$$

where $c \geqslant 0$ is the largest integer such that $(wt)^{q^{ic}}$ is not divisible by t^e. Clearly $\Psi(\tilde{h} + (t^e)) = h + (t)$, and a direct calculation shows that

$$\tilde{h}^{q^i} - \tilde{h} \equiv (wt)^{q^{(c+1)i}} \equiv 0 \pmod{(t^e)},$$

i.e. $\tilde{h} + (t^e)$ is in K^i. Q.E.D.

Finally we have

LEMMA 3.3. *For any* i, $1 \leqslant i \leqslant n$, Ψ *restricted to* K^i *is injective.*

Proof. Suppose for some $h \in F[x]$ that $h + (t^e)$ is a nonzero element of K^i, but $h + (t)$ is zero in $\bar{S}$. Write $h = vt^c$, where $0 < c < e$ and t^c is the largest power of t dividing h, i.e. $v \in F[x]$ is not divisible by t. Since $h + (t^e)$ is in K^i, in $F[x]$ we have

$$h^{q^i} - h = ut^e,$$

for some u in $F[x]$. Hence

$$h = h^{q^i} - ut^e,$$

i.e.

$$vt^c = (vt^c)^{q^i} - ut^e$$

i.e.

$$v = v^{q^i} t^{c(q^i - 1)} + ut^{e-c}.$$

But since $c \geqslant 1$ and $e - c \geqslant 1$, we have inferred that v is divisible by t, a contradiction. Q.E.D.

From Lemmas 3.2 and 3.3 we have

COROLLARY 3.2. *For any* i, $1 \leqslant i \leqslant n$, $\ker(\pi^i - I)$ *and* $\ker(\bar{\pi}^i - \bar{I})$ *are isomorphic as vector spaces.*

Returning to the polynomial $f(x)$ and the notation we had prior to the lemmas, we can now state (3.1) as a theorem:

THEOREM 3.1. *For any* k *and any* i, $1 \leqslant k \leqslant r$ *and* $1 \leqslant i \leqslant n$,

$$\dim \ker(\pi_k^i - I_k) = \dim \ker(\bar{\pi}_k^i - \bar{I}_k).$$

Analogues to the corollaries of Section 2, minus the separability hypothesis, follow immediately:

COROLLARY 3.3. *For any* i, $1 \leqslant i \leqslant n$, $v_i = \sum_{k=1}^{r} (i, n_k)$.

COROLLARY 3.4. *For any* i, $1 \leqslant i \leqslant n$, $v_i = \sum_{j=1}^{n} (i, j)\sigma_j$.

COROLLARY 3.5. $v = A\sigma$.

COROLLARY 3.6. v_1 *is the number of nonassociate irreducible factors of* $f(x)$.

(Note: the irreducibility criterion of Corollary 2.3 is only valid when f is separable.)

Remark. Corollary 3.6 was proved as a theorem by Butler [2]. It is fundamental to the Berlekamp algorithm.

4. The Formula for A^{-1}. The matrices in this section will be $m \times m$, where m is any given positive integer. We view the standard Möbius function μ as being defined on the positive rational numbers; $\mu(c) = 0$ for nonintegral c. We define matrices M, S, A and D:

$$M_{ij} = \mu\left(\frac{i}{j}\right), \quad S_{ij} = \begin{cases} 1, & j \text{ divides } i, \\ 0, & j \text{ does not divide } i. \end{cases}$$

$$A_{ij} = (i, j), \quad D = \text{diagonal}(\varphi(1), \ldots, \varphi(m)).$$

The reader can easily verify the following:

LEMMA 4.1. (1) $MS = I$, *where* I *is the* $m \times m$ *identity matrix.*

(2) Determinant(M) = Determinant(S) = 1.

(3) Determinant$(D) = \prod_{i=1}^{m} \varphi(i)$.

Remark. Part (1) of Lemma 4.1 is a "vector" form of the Möbius inversion formula. That is, if g and h are any arithmetical functions, if

$$g^* = (g(1), g(2), \ldots, g(m))^T \quad \text{and} \quad h^* = (h(1), h(2), \ldots, h(m))^T,$$

then (1) says that $g^* = Sh^*$ if and only if $h^* = Mg^*$, which is just a simultaneous statement of the Möbius formula for the values $g(1), \ldots, g(m), h(1), \ldots, h(m)$ of g and h.

Now let g and h be any two arithmetical functions such that $g(k) = \sum_{d|k} h(d)$, for all k. Let G and H be the following matrices:

$$G_{ij} = g((i,j)), \quad H = \text{diagonal}(h(1), \ldots, h(m)).$$

THEOREM 4.1. $G = SHS^T$.

Proof. For every i, j, $1 \leqslant i,j \leqslant m$,

$$G_{ij} = g((i,j)) = \sum_{d|(i,j)} h(d)$$

$$= \sum_{d|i \text{ and } d|j} h(d) = \sum_{k=1}^{m} S_{ik}h(k)S_{jk} = (SHS^T)_{ij}.$$

Q.E.D.

As Carlitz noted, a different but equivalent version of this theorem was proved by Cesàro in 1888.

COROLLARY 4.1. $A = SDS^T$.

Proof. Set $g(k) = k$ and $h(k) = \varphi(k)$ in Theorem 4.1.

COROLLARY 4.2. $A^{-1} = M^T D^{-1} M$. *Thus, the ijth element of* A^{-1} *is*

$$(A^{-1})_{ij} = \sum_{k=1}^{m} \mu\left(\frac{k}{i}\right)\mu\left(\frac{k}{j}\right)\frac{1}{\varphi(k)}. \tag{4.1}$$

COROLLARY 4.3 (SMITH'S FORMULA). Determinant$(A) = \prod_{i=1}^{m} \varphi(i)$.

5. An Example. We now illustrate how the results of Sections 2, 3, and 4 provide an algorithm for computing the degrees of the irreducible factors of a given $f(x)$. One first constructs a matrix for the Frobenius on $R = F[x]/(f(x))$. One then uses this matrix to obtain $\nu_1, \nu_2, \ldots, \nu_n$, and finally obtains σ as $A^{-1}\nu$. A^{-1} is assumed to have been precomputed by formula (4.1) of Section 4.

We choose $q = 2$ and let F_2 denote the finite field with two elements. We let $f(x) \in F_2[x]$ be $x^8 - x = x^{2^3} - x$; thus $n = 8$. It is well known that $f(x)$ is the product of all monic polynomials in $F_2[x]$ irreducible over F_2 and of degree dividing 3. We assume the standard basis $1, x, x^2, \ldots, x^7$ for $F_2[x]/(f(x))$, so the ith row, $1 \leqslant i \leqslant 8$, of the matrix of π with respect to this basis, is $x^{(i-1)q} \bmod f = x^{2i-2} \bmod f$. We obtain the following 8×8 matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Where I is the 8×8 identity matrix, we compute ν_i, $i = 1, 2, \ldots, 8$, as the dimension of the null space of the matrix $\pi^i - I$. We obtain

i	1	2	3	4	5	6	7	8
ν_i	4	4	8	4	4	8	4	4

Using formula (4.1), we obtain the 8×8 matrix A^{-1}:

$$\begin{bmatrix} \frac{41}{12} & \frac{-3}{2} & -1 & 0 & \frac{-1}{4} & \frac{1}{2} & \frac{-1}{6} & 0 \\ \frac{-3}{2} & 2 & \frac{1}{2} & \frac{-1}{2} & 0 & \frac{-1}{2} & 0 & 0 \\ -1 & \frac{1}{2} & 1 & 0 & 0 & \frac{-1}{2} & 0 & 0 \\ 0 & \frac{-1}{2} & 0 & \frac{3}{4} & 0 & 0 & 0 & \frac{-1}{4} \\ \frac{-1}{4} & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{-1}{2} & \frac{-1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ \frac{-1}{6} & 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 & \frac{-1}{4} & 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}.$$

$A^{-1}v$ yields the vector

$$\sigma = (2, 0, 2, 0, 0, 0, 0, 0)^T.$$

Thus, $f(x)$ has two irreducible factors of degree one, and two irreducible factors of degree three. Since the monic irreducible polynomials of degree one and three over F_2 are x, $x + 1$, $x^3 + x + 1$, and $x^3 + x^2 + 1$, this is just what was expected.

Acknowledgment. We thank Professor I. Martin Isaacs for his comments on a preliminary version of this paper.

Department of Mathematics
University of Wisconsin-Madison
Madison, Wisconsin 53706

Department of Computer Sciences
University of Wisconsin-Madison
Madison, Wisconsin 53706

1. E. R. BERLEKAMP, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

2. M. C. R. BUTLER, "On the reducibility of polynomials over a finite field," *Quart. J. Math. Oxford Ser.*, v. 5, 1954, pp. 102–107.

3. L. CARLITZ, "Some matrices related to the greatest integer function," *J. Elisha Mitchell Sci. Soc.*, v. 76, 1960, pp. 5–7.

4. D. R. MUSSER, "On the efficiency of a polynomial irreducibility test," *J. Assoc. Comput. Mach.*, v. 25, 1978, pp. 271–282.

5. S. SCHWARZ, "On the reducibility of polynomials over a finite field," *Quart. J. Math. Oxford Ser.*, v. 7, 1956, pp. 110–124.

6. S. SMITH, "On the value of a certain arithmetical determinant," *Proc. London Math. Soc.*, v. 7, 1876, pp. 208–212.

7. B. L. VAN DER WAERDEN, *Algebra*. Vol. 1, (transl. of 7th German ed.), F. Blum and J. Schulenberger (Eds.), Ungar, New York, 1970.

→ B. WYMAN, "Correction to 'What is a reciprocity law?'," *Amer. Math. Monthly*, v. 80, 1973, p. 281.

9. H. G. ZIMMER, *Computational Problems, Methods, and Results in Algebraic Number Theory*, Lecture Notes in Math., vol. 269, Springer-Verlag, Berlin, 1972.