

Algorithms for Finite Fields
September 12, 2006
Notes by: Kim Hopkins

1. BERLEKAMP'S ALGORITHM

Berlekamp's Algorithm is used to factor a polynomial $f(x)$ over a finite field \mathbb{F}_q . Given a squarefree polynomial $f(x) \in \mathbb{F}_q[x]$ of degree n , the goal is to find distinct irreducible polynomials $g_i(x) \in \mathbb{F}_q[x]$ so that

$$f(x) = g_1(x) \cdots g_r(x).$$

If a polynomial $u(x) \in \mathbb{F}_q[x]$ satisfies

$$(1.1) \quad u(x)^q \equiv u(x) \pmod{f(x)}$$

then it can be shown that

$$\prod_{c \in \mathbb{F}_q} \gcd(f(x), u(x) - c) = f(x)$$

which may provide us with a factorization of f .

Polynomials that satisfy (1.1) form a vector space of dimension r over \mathbb{F}_q and are the kernel of a certain linear map on $\mathbb{F}_q[x]/(f)$. By computing the $n \times n$ matrix of this map and then reducing it to row-echelon form, we can easily determine a basis for this subspace and hence polynomials of type (1.1). The gcd's are then computed with the Euclidean algorithm (since $\mathbb{F}_q[x]$ is a Euclidean domain).

1.1. Details. Now we describe this process in detail. By the Chinese Remainder Theorem we have

$$(1.2) \quad \frac{\mathbb{F}_q[x]}{(f(x))} \cong \frac{\mathbb{F}_q[x]}{(g_1(x))} \oplus \cdots \oplus \frac{\mathbb{F}_q[x]}{(g_r(x))}$$

via the map that sends

$$u(x) \pmod{f(x)} \mapsto u(x) \pmod{g_1(x)}, \dots, u(x) \pmod{g_r(x)}$$

for any polynomial $u(x) \in \mathbb{F}_q[x]$.

Clearly \mathbb{F}_q^r embeds into the right hand side of (1.2); in fact at *most* r copies embed, and \mathbb{F}_q^r is isomorphic to the space spanned by polynomials satisfying (1.1).

Theorem 1.1. *The Frobenius map*

$$(1.3) \quad Fr : \frac{\mathbb{F}_q[x]}{(f(x))} \longrightarrow \frac{\mathbb{F}_q[x]}{(f(x))}$$

$$(1.4) \quad u \longmapsto u^q$$

is and \mathbb{F}_q -linear map, and $\ker(Fr - I) \cong \mathbb{F}_q^r$ so in particular

$$\dim_{\mathbb{F}_q} \ker(Fr - I) = r.$$

Proof. Linear:

$$Fr(g + h) = (g + h)^q = g^q + h^q = Fr(g) + Fr(h)$$

since q is a power of the characteristic of \mathbb{F}_q .

If $\lambda \in \mathbb{F}_q$, then $\lambda^q = \lambda$ so

$$Fr(\lambda g) = (\lambda g)^q = \lambda g^q = \lambda Fr(g).$$

Kernel:

$$\begin{aligned}
& (Fr - I)(h) = 0 \pmod{f} \\
\Leftrightarrow & (Fr - I)(h_1, \dots, h_r) = (0, \dots, 0) \quad (\text{where } h_i = h \pmod{g_i}) \\
\Leftrightarrow & ((Fr - I)h_1, \dots, (Fr - I)h_r) = (0, \dots, 0) \\
\Leftrightarrow & h_i \in \mathbb{F}_q
\end{aligned}$$

since elements satisfying $h^q = h$ are precisely the elements of \mathbb{F}_q . So

$$\ker(Fr - I) \cong \mathbb{F}_q \oplus \dots \oplus \mathbb{F}_q \hookrightarrow \frac{\mathbb{F}_q[x]}{(g_1(x))} \oplus \dots \oplus \frac{\mathbb{F}_q[x]}{(g_r(x))}$$

□

Next we find a matrix representation of $(Fr - I)$. $1, x, x^2, \dots, x^{n-1}$ is an \mathbb{F}_q -basis for $\mathbb{F}_q[x]/(f(x))$, so we compute a_{ij} 's in \mathbb{F}_q such that

$$Fr(x^i) = x^{qi} = \sum_{j=0}^{n-1} a_{ij}x^j \pmod{f(x)}, \quad i = 0, \dots, n-1.$$

Then the matrix $(a_{ij} - \delta_{ij})$ (where δ_{ij} is the Kronecker delta) is the matrix for $(Fr - I)$ with respect to this basis.

Note: x^{qi} is an exponentiation in $\mathbb{F}_q[x]/(f(x))$ so it can be computed in polynomial time in $n \log q \log q^p \leq n \log q \log q^n \leq O((n \log q)^c)$.

Now we want to show how elements of the kernel actually provide us with factors of f . Suppose $u \in \ker(Fr - I) \setminus \mathbb{F}_q$ (where \mathbb{F}_q is being viewed as the set of r -tuples, (c, c, \dots, c) in \mathbb{F}_q^r with $c \in \mathbb{F}_q$) so $u = (c_1, \dots, c_r)$, $c_i \in \mathbb{F}_q$ not all equal, say $c_i \neq c_j$.

Then $u - c_i$ can be viewed as a polynomial in $\mathbb{F}_q[x]/(f(x))$ via (1.2). By construction, $u \equiv c_i \pmod{g_i}$ so $u - c_i \equiv 0 \pmod{g_i}$, but modulo g_j , $u - c_i \equiv c_j - c_i \not\equiv 0 \pmod{g_j}$. Hence $g_i \mid u - c_i$ but $g_j \nmid u - c_i$ which implies $\gcd(u - c_i, f) \neq 1, f$. Thus u gives a proper factor of f and we get the following result:

Theorem 1.2. *Let $f(x) \in \mathbb{F}_q[x]$ be monic and squarefree. If $u \in \mathbb{F}_q[x]$ is such that $u^q \equiv u \pmod{f}$ but $u \not\equiv c \pmod{f} \forall c \in \mathbb{F}_q$ (ie $u \in \ker(Fr - I) \setminus \mathbb{F}_q$), then*

$$\prod_{c \in \mathbb{F}_q} \gcd(u - c, f) = f.$$

(Note: None of the factors are equal to f since we required that $u \notin \mathbb{F}_q$, however some of the factors might be trivial since not every c is a c_i in the decomposition of u from above.)

Proof. As noted, if $c \neq c_i$ for any i in the decomposition of u , then $g_i \nmid u - c$ for all i , hence $\gcd(u - c, f) = 1$ and so we are reduced to showing

$$\prod_{c_i \in \mathbb{F}_q} \gcd(u - c_i, f) = f.$$

where $u = (c_1, \dots, c_r)$. From above we have that g_i divides $u - c_i$ and of course g_i divides f , so g_i divides $\gcd(u - c_i, f)$ hence

$$f = \prod_{i=1}^r g_i \left| \prod_{c_i \in \mathbb{F}_q} \gcd(u - c_i, f) \right.$$

On the other hand, for any $c, c' \in \mathbb{F}_q$, $c \neq c'$, we have $\gcd(u - c, u - c') = 1$ because $(u - c) - (u - c') = c - c' \in \mathbb{F}_q^*$ (so there exists $k \in \mathbb{F}_q$ such that $k(c - c') = 1 = k(u - c) - k(u - c')$). Thus $\gcd(u - c, f)$ and $\gcd(u - c', f)$ are coprime, and clearly divide f , hence

$$\prod_{c_i \in \mathbb{F}_q} \gcd(u - c_i, f) \mid f.$$

□

Example. Let $q = 2$ with $\mathbb{F}_2 = \{0, 1\}$ and let $u = (c_1, \dots, c_r)$. Then by the theorem above

$$f = \gcd(u, f) \cdot \gcd(u - 1, f).$$

We summarize:

Berlekamp's Algorithm.

1. Construct the matrix $(Fr - I)$, (an $n \times n$ matrix over \mathbb{F}_q)
2. Compute the kernel (using Gaussian elimination). If $\dim(\ker(Fr - I)) = 1$, then $r = 1$ and f is irreducible, so stop.
3. If $\dim(\ker(Fr - I)) > 1$, find $u \in \ker(Fr - I) \setminus \mathbb{F}_q$ (linear algebra) and compute the gcd's in Theorem 1.2 (using the Euclidean algorithm) to split $f(x)$.

Remark 1.3. This is a deterministic polynomial time factoring algorithm if q is small. For q large we have to do something else...

If $f(x)$ splits into linear factors over \mathbb{F}_q , then $\ker(Fr - I) \cong \mathbb{F}_q^n \cong \mathbb{F}_q[x]/(f(x))$ and

$$f = \prod_{c \in \mathbb{F}_q} \gcd(x - c, f).$$

Example. Let $q = 2$ and $f(x) = x^5 + x + 1$. $\mathbb{F}_2[x]/(f)$ has as a basis $1, x, x^2, x^3, x^4$. The Frobenius applied to the basis elements gives

$$\begin{aligned} Fr(1) &= 1 \\ Fr(x) &= x^2 \\ Fr(x^2) &= x^4 \\ Fr(x^3) &= x^6 = x^2 + x \\ Fr(x^4) &= x^8 = x^4 + x^3 \end{aligned}$$

We need to find a polynomial $u \in \mathbb{F}_2[x]/(f)$ so that $Fr(u) = u$. Let

$$u = \alpha x + \beta x^2 + \gamma x^3 + \delta x^4$$

Then

$$\begin{aligned} Fr(u) &= \alpha x^2 + \beta x^4 + \gamma(x^2 + x) + \delta(x^4 + x^3) \\ &= \gamma x + (\alpha + \gamma)x^2 + \delta x^3 + (\beta + \gamma)x^4 \end{aligned}$$

so take $\alpha = \gamma = \delta = 1$, $\beta = 0$ and then

$$u = x + x^3 + x^4$$

is in $\ker(Fr - I)$.

Computing the gcd's gives:

$$\gcd(u, f) = x^3 + x^2 + 1, \quad \gcd(u + 1, f) = x^2 + x + 1$$

so

$$f = \gcd(u, f) \cdot \gcd(u + 1, f) = (x^3 + x^2 + 1)(x^2 + x + 1),$$

and in this case, both of the polynomials above happen to be irreducible, so we're done.

Remark 1.4. Note that

$$x + x^3 + x^4 \equiv x + (x^2 + 1) + (x^2 + 1)x \equiv x + x^2 + 1 + x^2 + 1 + x \equiv 0 \pmod{(x^3 + x^2 + 1)}$$

and

$$x + x^3 + x^4 \equiv x + (x + 1)x + (x + 1)^2 = x + x^2 + x + x^2 + 2x + 1 \equiv 1 \pmod{(x^2 + x + 1)}$$

which demonstrates the isomorphism between $\ker(Fr - I)$ and \mathbb{F}_2^2 .

We give a couple definitions,

Definition 1.5. Any u satisfying the hypotheses of Theorem 1.2 is called a splitting polynomial for f . The values $c \in \mathbb{F}_q$ for which $\gcd(u - c, f) \neq 1$ are called splitting values for f and u .

We will now focus on different methods for computing splitting polynomials and their splitting values.

Theorem 1.6. If m is an integer such that $\deg g_i | m \forall i = 1, \dots, r$ then

$$T_j := x^j + x^{jq} + x^{jq^2} + \dots + x^{jq^{m-1}}$$

satisfies

$$T_j^q \equiv T_j \pmod{f}.$$

Furthermore $\exists j \leq n$ such that T_j is a splitting polynomial (i.e. such that $T_j \not\equiv c \pmod{f} \forall c \in \mathbb{F}_q$).

Proof. The second part of the theorem is due to McEliece and its proof will be omitted.

For the first part, observe that

$$(1.5) \quad \mathbb{F}_q[x]/(g_i(x)) \cong \mathbb{F}_{q^k} \quad \text{where } \deg g_i =: k$$

$$(1.6) \quad \subseteq \mathbb{F}_{q^m} \quad \Leftrightarrow k | m$$

so if $\deg g_i | m$ then the roots of g_i are in \mathbb{F}_{q^m} and $x^{q^m} \equiv x \pmod{g_i}$ which we use to compute:

$$\begin{aligned} T_j^q &\equiv x^{jq} + x^{jq^2} + \dots + x^{jq^m} \\ &\equiv x^{jq^{m-1}} + \dots + x^{jq} + x^j \\ &\equiv T_j \pmod{g_i} \end{aligned}$$

□

Remark 1.7. To compute the degrees of the g_i 's, recall that $x^{q^d} - x$ is equal to the product of all irreducible polynomials in $\mathbb{F}_q[x]$ of degree k as k runs through the divisors of d . This gives a way to find all the factors of f of a certain degree;

for example $\gcd(x^q - x, f)$ gives all linear factors of f and $\gcd(x^{q^2} - x, f)$ gives all linear *and* quadratic factors of f , so

$$\frac{\gcd(x^{q^2} - x, f)}{\gcd(x^q - x, f)}$$

gives only the quadratic factors of f . Using this method we can write

$$f = h_1 \cdots h_L$$

where each h_i is the product of the irreducible factors of f of degree i (and if $\deg h_i \mid m$ then so does i).

Here is a method for constructing a splitting polynomial in the special case of factoring a cyclotomic polynomial $f(x) = x^l - 1$ modulo \mathbb{F}_q .

Theorem 1.8. *If l is prime, $l \nmid q$, H is a subgroup of $(\mathbb{Z}/l\mathbb{Z})^\times$ containing q , and C is a coset of H , then*

$$u := \sum_{c \in C} x^c$$

satisfies

$$u^q \equiv u \pmod{(x^l - 1)}$$

Proof.

$$(1.7) \quad u^q = \sum_{c \in C} x^{qc} = \sum_{c \in C} x^c = u \pmod{(x^l - 1)}$$

since $q \in H$. □

To apply this theorem we let H be the subgroup of squares in $(\mathbb{Z}/l\mathbb{Z})^\times$, C_1 the set of nonsquares, and then take

$$u := \sum_{j \in H} \binom{j}{l} x^j = \sum_{j \in H} x^j - \sum_{j \in C_1} x^j.$$

By the theorem, u satisfies $u^q \equiv u \pmod{(x^l - 1)}$ since it is a linear combination of elements of the form (1.7). Note that we must have $\binom{q}{l} = 1$ to use this, (so that q is an element of H as required by (1.8)).

Alternatively let m be the order of $q \pmod{l}$; then the factors of $x^l - 1$ have degree dividing m and we can write T_j from (1.6) as

$$T_j = \sum_{c \in \langle j \rangle} x^c$$

where $H := \langle q \rangle = \{1, q, \dots, q^{m-1}\}$, and this will also satisfy $T_j^q \equiv T_j \pmod{(x^l - 1)}$.

We now turn to the task of computing the splitting values for f .

Theorem 1.9. *If $f(x)$ is a monic squarefree polynomial in $\mathbb{F}_q[x]$ of degree n and u is a splitting polynomial for f , then we can compute in deterministic polynomial time in $n \log q$ the polynomial whose roots are exactly the splitting values for u .*

Remark 1.10. Note that if f splits into linear factors, then we can take $u = x$ and then the polynomial whose roots are the splitting values of u is precisely f .

To show how to compute the splitting values we need to use the resultant.

1.2. Resultants. Let K be a field, $f(x) = a_0 + a_1x + \cdots + a_nx^n$, and $g(x) = b_0 + b_1x + \cdots + b_mx^m$ both in $K[x]$. The resultant of f and g , denoted $Res(f, g)$ is defined as the determinant of an $(n + m) \times (n + m)$ matrix as follows,

$$Res(f, g) = \det \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_1 & a_0 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & a_n & \cdots & \cdots & a_0 \\ b_m & b_{m-1} & b_{m-2} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & b_{m-1} & \cdots & b_1 & b_0 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & b_m & \cdots & \cdots & b_0 \end{vmatrix}$$

One of the most important results regarding the resultant is that

$$Res(f, g) = a_n^m b_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j)$$

where $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m$ are the roots of f, g respectively. From this we can conclude,

$$Res(f, g) = 0 \Leftrightarrow f \text{ and } g \text{ have a common root}$$

Now we can prove Theorem 1.9,

Proof. Let

$$h(y) := Res_x(u(x) - y, f(x)) \in \mathbb{F}_q[y].$$

We claim that the roots of $h(y)$ are precisely the splitting values of u (with respect to f): Let $c \in \mathbb{F}_q$, then

$$\begin{aligned} h(c) = 0 &\Leftrightarrow Res_x(u(x) - c, f(x)) = 0 \\ &\Leftrightarrow u(x) - c \text{ and } f(x) \text{ have a common root} \\ &\Leftrightarrow u(x) - c \text{ and } f(x) \text{ have a common (nontrivial) factor} \\ &\Leftrightarrow c \text{ is a splitting value.} \end{aligned}$$

h can be computed by linear algebra in $\mathbb{F}_q[y]$ and hence in polynomial time. \square

Here is a better way to compute $h(y)$ (in the case $n + 1 \leq q$): Choose distinct $c_1, \dots, c_{n+1} \in \mathbb{F}_q$, and compute $h(c_i) = Res_x(u(x) - c_i, f)$. This is just linear algebra in \mathbb{F}_q (as opposed to $\mathbb{F}_q[y]$).

Since $\deg h \leq r \leq n$ (there are at most r splitting values), $n + 1$ is guaranteed to be enough c_i to find an appropriate polynomial h , which we construct using Lagrange interpolation:

$$h(x) := \sum_{j=1}^{n+1} h(c_j) \prod_{i=0, i \neq j}^{n+1} \frac{x - c_i}{c_j - c_i}.$$

Here is the "best way" to compute h (due to Zassenhaus):

Theorem 1.11. $h(y)$ is the monic polynomial of smallest degree such that

$$h(u(x)) \equiv 0 \pmod{f(x)}$$

Proof. We'll prove this later in the course. \square

Since $h(u(x))$ is a polynomial in $u(x)$, we want to find the smallest k such that

$$1, u(x), u(x)^2, \dots, u(x)^k$$

is a linearly *dependent* set modulo f , and then apply the theorem to show this linear combination is equal to h .

Example. Let $f(x) = x^4 + 3x^2 + 2$ in $\mathbb{F}_7[x]$. First we try $T_1 = x + x^7$ but it turns this is not a splitting polynomial since

$$x^7 \equiv x^4 x^3 \equiv (4x^2 + 5)x^3 \equiv 4(4x^2 + 5)x + 5x^3 \equiv 21x^3 + 20x \pmod{f}$$

so

$$x^7 + x = 21x^3 + 21x \equiv 0 \pmod{f}$$

(and $0 \in \mathbb{F}_q$).

Next we try $T_2 = x^2 + x^{14}$. Since $x^4 \equiv 4x^2 + 5$, we have $x^8 \equiv 2x^4 + 5x^2 + 4 \equiv 6x^2$ so

$$x^{14} \equiv x^2 x^8 x^4 \equiv x^2 (6x^2) x^4 \equiv 6x^8 \equiv 36x^2 \equiv x^2 \pmod{f}$$

therefore

$$T_2 \equiv 2x^2 \pmod{f}$$

is a splitting polynomial. We now compute h . The set $\{1, u\}$ are linearly independent but the set

$$\{1, u, u^2\} = \{1, 2x^2, 4x^4\}$$

satisfies

$$2u^2 + 5u + 2 \equiv 0 \pmod{f}.$$

Multiply through by 4, i.e. 2^{-1} , to make this monic, then

$$h(y) := y^2 - y + 1$$

is our h . Its roots, $c_1 = 3/2 \equiv 5 \pmod{7}$ and $c_2 = -1/2 \equiv 3 \pmod{7}$ are the splitting values and when we compute the gcd's we get,

$$\gcd(2x^2 - 5, f) = x^2 + 1, \quad \gcd(2x^2 - 3, f) = x^2 + 2.$$

Both factors are clearly irreducible mod 7 (the only squares mod 7 are 1, 2, and 4), and multiply to f as wanted.