6.966 Algebra and Computation	September 23, 1998

Lecture 5

Lecturer: Madhu Sudan — *Scribe: Anna Lysyanskaya*

In this lecture we will see an efficient algorithm for factoring polynomials over $GF(p^k)[x]$ (p is any prime). Recall that in the last lecture we saw how to factor a polynomial of the form $x^2 + a$ in $GF[p](x)$, where $p = 3 \bmod 4$. In this lecture we will first develop an algorithm for factoring such polynomial for any odd prime p, and then generalize our algorithm to factor any polynomial in $GF(p^k)[x]$.

1 An Algorithm to Factor $x^2 + a$

Consider the polynomial $f(x) = x^2 + a$ in $GF(p)[x]$, where $p = 3 \bmod 4$. Since $f(x)$ is of degree 2, we know that unless it is irreducible, it has two factors of degree 1, $f_1(x) = x - \alpha$ and $f_2(x) = x + \alpha$. Note that, since these are of degree 1, they are irreducible. In Lecture 4, we have seen that if these factors exist and they are distinct, the following algorithm will always find them:

- Compute $g(x) = x^{\frac{p-1}{2}} - 1(\mod f(x))$.
- Compute $h(x) = GCD(g(x), f(x))$.
- Output $h(x)$, $\frac{f(x)}{h(x)}$.

The reason that this algorithm works, as was stated in lecture 4, is that exactly one of $f_1(x)$, $f_2(x)$, divides the polynomial $g(x)$ defined above. Recall the proof of this fact: Any polynomial of degree 1 will divide the polynomial $x^p - x$ by Fermat's little theorem. Then we have:

$$f_1(x)|x^p - x,$$

$$f_2(x)|x^p - x$$

It follows that

$$f_1(x)f_2(x)|x^p - x = x(x^{\frac{p-1}{2}} - 1)(x^{\frac{p-1}{2}} + 1)$$

Suppose one of the factors of $f(x)$ divides $x^{\frac{p-1}{2}} + 1$. Without loss of generality, let $f_1(x)|x^{\frac{p-1}{2}} + 1$. Then, since $f(\alpha) = 0$, $\alpha^{\frac{p-1}{2}} = -1$. On the other hand, since $p = 3 \bmod 4$ and consequently $\frac{p-1}{2}$ is odd,

$$\begin{aligned}
(-1)^{\frac{p-1}{2}} &= -1 \Rightarrow \\
(-1)^{\frac{p-1}{2}}\alpha^{p-1} &= -1 \Rightarrow \\
(-1)^{\frac{p-1}{2}}\alpha^{\frac{p-1}{2}}\alpha^{\frac{p-1}{2}} &= -1 \Rightarrow \\
(-\alpha)^{\frac{p-1}{2}}\alpha^{\frac{p-1}{2}} &= -1 \Rightarrow \\
(-\alpha)^{\frac{p-1}{2}} &= 1 \Rightarrow \\
x + \alpha|x^{\frac{p-1}{2}} - 1.
\end{aligned}$$

The last equation follows because $-\alpha$ is the root of both the polynomial on the right and the one on the left.

Thus we have showed that

$$f_1(x) | x^{\frac{p-1}{2}} + 1,$$
$$f_2(x) | x^{\frac{p-1}{2}} - 1.$$

The other direction can be showed in a similar manner.

The fact that this is true when $p = 3 \bmod 4$ looks like a happy coincidence; however it turns out that a very similar factoring algorithm will work when $p = 1 \bmod 4$. First observe that factoring $f(x) = x^2 - a$ can be reduced to factoring $f^{(1)}(x) = (x-d)^2 - c^2 a$ for any $c, d \in GF(p)$. This is true because if $f(x) = (x - \alpha_1)(x - \alpha_2)$, then $f^{(1)}(x) = (x-d)^2 - c^2 a = ((x-d) - c\alpha_1)((x-d) - c\alpha_2)$, and so if we discover that $f^{(1)}(x) = (x - \beta_1)(x - \beta_2)$, we can compute α_1 from β_1 and α_2 from β_2.

Now suppose that c, d are picked at random. What are the chances that exactly one of the factors of $f^{(1)}(x)$ divides $x^{\frac{p-1}{2}}$? It turns out that it is approximately $\frac{1}{2}$ by the following argument:

$$\forall \alpha_1 \neq \alpha_2, \beta_1, \beta_2 \Pr_{c,d,\in F}[c\alpha_1 + d = \beta_1 \text{ and } c\alpha_2 + d = \beta_2] = \frac{1}{|F|^2}$$

On the other hand,

$$\Pr_{\beta_1}[(x - \beta_1) \text{ divides } (x^{\frac{p-1}{2}} - 1)] = \frac{\frac{p-1}{2}}{p} \approx \frac{1}{2}$$
$$\Pr_{\beta_2}[(x - \beta_2) \text{ does not divide } (x^{\frac{p-1}{2}} - 1)] = \frac{\frac{p-1}{2}}{p} \approx \frac{1}{2}$$

Therefore, uniform choice of c, d guarantees that β_1, β_2 are independent of each other, and so the probability that $(x - \beta_1)$ and $(x - \beta_2)$ divide different factors of $(x^{p-1} - 1)$ is approximately $\frac{1}{2}$.

Then the algorithm proceeds as follows:

1. Compute $g(x) = x^{\frac{p-1}{2}} - 1(\bmod f(x))$.
2. Pick random $c, d \in F$ and compute $f^{(1)}(x) = (x-d)^2 - c^2 a$.
3. Compute $h(x) = GCD(f^{(1)}(x), g(x))$.
4. If $h(x) = 1$, go back to step 2.
5. We now know that $f^{(1)}(x) = h(x)\frac{f^{(1)}(x)}{h(x)} = (x - \beta_1)(x - \beta_2)$. Compute $\alpha_i = \frac{\beta_i - d}{c}$ for $i \in \{1, 2\}$.

2 Extending the Algorithm

2.1 Challenges

Even though the algorithm above can only factor a special case of polynomials, it can be extended to an algorithm that factors polynomials in $GF(p^k)[x]$, where p is prime. Furthermore, it finds a non-trivial factorization of any polynomial of the form $(x - \alpha_1) \cdots (x - \alpha_l)$; and thus by recursing on such a factorization, one can get a complete factorization of any such polynomial.

Now let us enumerate the problems with applying the algorithm as it is to the more general case.

1. The algorithm works only for polynomials whose factorization into irreducible factors is square-free, i.e., for all $g(x)$ that divide $f(x)$, $g^2(x)$ does not divide $f(x)$.
2. Only the first degree irreducible polynomials divide $(x^q - x)$, where $q = |F|$. We have to find polynomials that will be divisible by irreducibles of higher degrees.
3. The trick of classifying factors of a polynomial into those that divide $x^{\frac{q-1}{2}} - 1$ and those that don't needs to be generalized to work for non-linear irreducible factors.
4. The algorithm is designed to work over odd prime fields.

2.2 Square-Free Factorization

Let us (re)introduce the notion of the derivative of a polynomial in $GF(p^k)[x]$ (see also Lecture 4).

Definition 1 *For a polynomial $f(x) = \sum_{i=0}^{n} c_i x^i$, the derivative $f'(x)$ is defined as*

$$f'(x) = \sum_{i=1}^{n} c_i i x^{i-1}.$$

The properties of this derivative often coincide with the properties of the more conventional derivative in the reals. In particular, the derivative of a sum is the sum of the derivatives. This property is straightforward to show. Also, if $u(x)$ and $v(x)$ are two polynomials, then the derivative of their product $(uv)'(x) = u'(x)v(x) + u(x)v'(x)$. This property can be derived by induction on the number of monomials of $u(x)$ and $v(x)$ in a fairly straightforward fashion. In dealing with eliminating repeated factors, the derivative of a polynomial is useful due to the following theorem:

Theorem 2 *Suppose there exist non-zero degree polynomials in $GF(p^k)[x]$, $f(x)$ and $g(x)$, and a polynomial $h(x)$ such that $f(x) = g^2(x)h(x)$. Then $GCD(f(x), f'(x)) \neq 1$.*

Proof Let us show that $g(x)$ divides $f'(x)$: $f'(x) = (g^2 h)'(x) = (g^2)'(x)h(x) + g^2(x)h'(x) = 2g(x)g'(x)h(x) + g^2(x)h'(x) = g(x)(2g'(x)h(x) + g(x)h'(x))$. Clearly, since both $f(x)$ and $f'(x)$ are divisible by $g(x)$, their greatest common divisor is different from 1. ■

Since the greatest common divisor is different from 1, it is us a non-unit factor of $f(x)$. Of course, the factor is non-trivial only when $GCD(f, f') \neq f$. What happens in the case when $GCD(f, f') = f$? Observe that this can only happen when $f' = 0$, because the degree of f' is always smaller than the degree of f. What does the fact that $f' = 0$ tell us about f? Since the coefficients of f come from an integral domain, the only way that the coefficients of f' can be 0's is if all the monomials present in f are of some degree that is divisible by p. But that fact alone gives us a factorization of $f(x)$: $f(x) = \sum a_i x^{ip} = (\sum a_i^{p^{k-1}} x^i)^p$, because over $GF(p^k)$, $(a+b)^p = a^p + b^p$ (recall that the other terms cancelled because their coefficients are divisible by p, which is the characteristic of the field).

2.3 Distinct Degree Factorization

In order to generalize our present algorithm, we need to find an analogue of the polynomial $x^q - x$. Recall that $x^q - x$ has the property that any element of the field is its root, and therefore this polynomial is divisible by all the linear polynomials in $F[x] = GF(p^k)[x]$.

Theorem 3 *The polynomial $x^{q^d} - x$ is a a multiple of any degree d irreducible polynomial. Furthermore if an irreducible polynomial $f(x)$ divides $x^{q^d} - x$, then the degree of f divides d.*

Proof We only prove the first part. We postpone the proof of the "furthermore" part to the next lecture.

Let $f_1(x)$ be an irreducible polynomial of degree d. Let $K = F[x]/(f_1(x))$. K is a field that contains all polynomials in variable x of degree $\leq d-1$. Therefore, for all $\alpha \in K$, $\alpha^{|K|} - \alpha = 0$. Since the polynomial x also belongs to the field K, $x^{|K|} - x = 0 \bmod f_1(x)$, and so $f_1(x)$ divides $x^{|K|} - x$. What is left to note is that $|K| = q^d$. That is true because K consists of all polynomials in variable x of degree $\leq d-1$. Any such polynomial has d coefficients, one per degree, from F. Thus $|K| = |F|^d = q^d$. ■

Owing to this theorem, we now have an algorithm for factoring $f(x)$ into $g_1(x), ..., g_l(x)$, where $g_i(x)$ is a product of irreducible polynomials of degree i. The algorithm goes as follows:

- Set $i = 1$.
- While $f(x)$ is non-unit,
  - $g_i(x) := GCD(f(x), x^{q^i} - x)$
  - $f(x) := f(x)/g_i(x)$
- Output $g_1(x), ..., g_l(x)$.

Note that the above algorithm is also an irreducibility test: if it outputs $f(x)$ in the end, we know that $f(x)$ must be irreducible.

2.4 Generalizing the Trick of Factor Splitting

At this point, we may assume that we have a polynomial $f(x)$ over $F[x] = GF(p^k)[x]$ (p is an odd prime) that is a product of distinct irreducible factors of degree d. This is a situation similar to the one we have seen, where we have to find a polynomial that is a multiple of exactly at least one, but not all factors of $f(x)$. We focus on the case $f(x) = f_1(x).f_2(x)$ and show how the algorithm works in this case. (The algorithm remains the same for $f(x) = f_1(x) \cdots f_l(x)$, only it only produces a non-trivial factorization, on which we may recurse.)

We have already seen that for any $T(x)$, $g(x) = T(x)^{q^d} - T(x)$ should be a multiple of any irreducible polynomial $f_1(x)$ of degree d. Observe that

$$g(x) = T(x)(T(x)^{\frac{q^d-1}{2}} - 1)(T(x)^{\frac{q^d-1}{2}} + 1) = T(x)A_1(T(x))A_2(T(x))$$

.

Theorem 4 *Let $f_1(x) \in F[x] = GF(p^k)[x]$ be an irreducible polynomial of degree d and let $d' > d$. Then*

$$\Pr_{T(x)}[f_1(x)|A_1(T(x))] \approx \frac{1}{2}.$$

where $T(x)$ is chosen uniformly from the space of polynomials of degree d' over F[x].

Proof Notice first that picking $T(x)$ at random over polynomials of degree $d' > d$, corresponds to picking $T'(x) = T(x) \bmod f_1(x)$ at random from the field $K = F[x]/(f_1(x))$. Furthermore $f_1(x)|A_1(T(x))$ if and only if $f_1(x)|A_1(T'(x))$. Thus the question of whether $f_1(x)|(T(x)^{\frac{q^d-1}{2}} - 1)$ is really the same as the question of whether $T'(x)$ is a quadratic residue in the field $K = F[x]/(f_1(x))$. Since $K = GF(q^d)$, and $GF(q^d) \setminus \{0\}$ is a cyclic group of even order $q^d - 1$ under multiplication, exactly half of its elements are quadratic residues. Then there are $\frac{q^d-1}{2}$ quadratic residues modulo $f_1(x)$, and so

$$\Pr_{T(x)}[f_1(x)|A_1(T'(x))] \approx \frac{q^d - 1}{2}.$$

Notice that in the above argument we could have used any other factorization of the polynomial $x^{q^d} - x$ into $A_1(x).A_2(x)$ and the final probability depends only on the degree of A_1. ■

Now we need to show that the event that a polynomial $T(x)$ is a quadratic resudue modulo $f_1(x)$ is independent from the event that it is a quadratic residue modulo $f_2(x)$ for a randomly chosen $T(x)$ of degree $\leq 2d - 1$.

Theorem 5 *For any $T_1(x), T_2(x)$ over $GF(q)[x]$ of degree $\leq d-1$, and irreducible polynomials f_1, f_2 over $GF(q)[x]$, of degree d,*

$$\Pr_{T(x)}[T(x) = T_1(x) \ mod\ f_1(x) \ and\ T(x) = T_2(x) \ mod\ f_2(x)] = \frac{1}{(q^d)^2}$$

where $T(x) \in GF(q)[x]$, of degree $\leq 2d-1$.

Proof This result is a direct consequence of the Chinese Remainder Theorem. ■

The previous two theorems, combined, give us the desired result that if f_1 and f_2 are two factors of f, then with probability approximately $\frac{1}{2}$ one of them divides $A_1(T(x))$ while the other does not, and so by computing $GCD(f, A_1(T(x))$, we get a factor of f with probability approximately $\frac{1}{2}$.

2.5 The Remaining Case: $p = 2$

The only remaining problem with the approach we have to factoring polynomials in $GF(p^k)[x]$ is that the polynomial $T(x)^{p^k d} - T(x)$ can be factored into $T(x)(T(x)^{\frac{p^{kd}-1}{2}} - 1)(T(x)^{\frac{p^{kd}-1}{2}} + 1)$ only if p is odd. What happens if $p = 2$?

Note that $T(x)^{2^{kd}} - T(x) = A(T(x))(A(T(x)) + 1)$ where $A(T(x)) = T(x)^{2^{kd-1}} + T(x)^{2^{kd-2}} + \ldots + T(x) + T(x)^0$. This fact is straightforward to verify keeping in mind that the characteristic of the field is 2.

Theorem 6 *For any irreducible polynomial $f_1(x) \in F[x] = GF(2^k)[x]$ of degree d,*

$$\Pr_{T(x) \in F[x]/f_1}[A(T(x)) = 0 \ mod\ f_1] \approx \frac{1}{2}.$$

Proof We first observe the following (can be verified straightforwardly):

Lemma 7 *If $T_1(x) \neq T_2(x) \in F[x]/f_1$ are such that $A(T_i(x)) = 0 \ mod\ f_1$ ($i \in \{1, 2\}$), then $A(T_1(x) + T_2(x)) = 1 \ mod\ f_1$. On the other hand, if $T_1(x) \neq T_2(x) \in F[x]/f_1$ are such that $A(T_i(x)) = 1 \ mod\ f_1$ ($i \in \{1, 2\}$), then $A(T_1(x) + T_2(x)) = 0 \ mod\ f_1$.*

Now, suppose $\Pr_{T(x) \in F[x]/f_1}[A(T(x)) = 0 \text{ mod } f_1] \geq \frac{1}{2}$. Let

$$S_1 = \{T(x) \text{ such that } A(T(x)) = 0 \text{ mod } f_1\}$$

By our assumption, S_1 must occupy at least a half of $F[x]/f_1$. Let $T_1(x) \neq 0$ be a polynomial from S_1. Define

$$S_2 = \{T(x) \text{ such that } T(x) = T_1(x) + T'(x), \text{ where } 0 \neq T'(x) \neq T_1(x) \in S_1\}$$

By the lemma, all the polynomials in $T(x) \in S_2$ are such that $A(T(x)) = 1 \text{ mod } f_1$. On the other hand, by construction, $|S_2| = |S_1| - 2$. Therefore, we get:

$$\Pr_{T(x) \in F[x]/f_1}[A(T(x)) = 0 \text{ mod } f_1] \geq \frac{1}{2} \Rightarrow \Pr_{T(x) \in F[x]/f_1}[A(T(x)) = 0 \text{ mod } f_1] \approx \frac{1}{2}$$

The other direction can be showed in exactly the same way. ■

This theorem combined with the theorem from the previous section, guarantees that the greatest common divisor of $f(x)$ and $A(T(x)))$ for a randomly chosen $T(x)$ of degree $2d-1$, is a non-trivial factor of $f(x)$ with probability $\geq \frac{1}{2}$.

3 Summary

As a result, we have the following algorithm to factor any polynomial $f(x) \in F[x] = GF(p^k)[x]$[1]:

1. Use $f'(x)$ to factor out repeated factors.
2. Use distinct degree factorization to obtain $g_1(x), \ldots, g_l(x)$, where $g_i(x)$ is a multiple of distinct irreducibles of degree i.
3. Factor each $g_d(x)$ by the following procedure:
   - Pick a random $T(x)$ of degree $2d - 1$.
   - Compute $g(x) = GCD(f(x), A(T(x)))$, where $A(T(x)) = T(x)^{\frac{p^{kd}-1}{2}} - 1(\bmod f(x))$ if $p \neq 2$, $A(T(x)) = T(x)^{2^{kd-1}} + T(x)^{2^{kd-2}} + \ldots + T(x) + T(x)^0(\bmod f(x))$ if $p = 2$.
   - Repeat for $g(x)$, if its degree $\geq d$, and for $f(x)/g(x)$, if its degree $\geq d$.

[1] The first randomized polynomial time algorithm for factoring univaraite polynomials was given by Berlekamp ('70). The algorithm we have seen is different and is due to Cantor and Zassenhaus. The primary difference in the algorithm is the part addressed in Section 2.4 (i.e., how it addresses the splitting of degree d factors). See Knuth (v. 2)/Cohen for more details.