A remark on the Berlekamp algorithm for binomials $x^n - a$

Ryuichi Harasawa

Nagasaki University

(Joint work with Y. Sueyoshi, A. Kudo and L. Cui)

1. Polynomial factorization over finite fields

Input: $f(x) \in \mathbb{F}_q[x]$

Output: pair(s) $(f_i(x), e_i)$ with $f(x) = \prod f_i(x)^{e_i}$
($f_i(x)$: irreducible polynomial)

Application to cryptography

- **Construction of extension fields:**
  $f(x)$: irre. poly./$\mathbb{F}_q$ of degree $n \Rightarrow \mathbb{F}_{q^n} = \mathbb{F}_q[x]/(f(x))$
- **Index calculus(-like) attack:**
  Check if $f(x)$ is $\mathcal{B}$-smooth.
  If so, factor $f(x)$, where $\mathcal{B}(\subset \mathbb{F}_q[x])$ is a factor base.

2. Procedure of factorization

We first perform the squarefree factorization.
After the procedure, we factor squarefree polynomial(s).

$f(x)$: squarefree $\overset{\text{def}}{\Longleftrightarrow}$ $f(x)$ has no repeated factors
(i.e., $g(x)|f(x)$ $(\deg(g(x)) \geq 1)$ $\Rightarrow$ $g(x)^2 \not| f(x)$)

[Squarefree factorization]

Input: $f(x) \in \mathbb{F}_q[x]$

Output: $g_i(x)$'s: squarefree (possibly $g_i(x) = 1$)
with $f(x) = \prod_{i \geq 1} g_i(x)^i$

3. Squarefree factorization

$f(x) = \prod_i g_i(x)^i$, $g_i(x)$: (unknown) squarefree poly.
$p = \mathrm{char}\mathbb{F}_q$, $f'(x)$: the formal derivation of $f(x)$

[Key facts]

- $f'(x) = 0 \Rightarrow f(x) = g(x)^p$ $(\exists g(x) \in \mathbb{F}_q[x])$,
  more precisely $f(x) = \sum_j a_{jp} x^{jp} = (\sum_j a_j^{(1/p)} x^j)^p$
- $\gcd(f(x), f'(x)) = \prod_{p \not| i} g_i(x)^{i-1} \cdot \prod_{p|i} g_i(x)^i$
- $f(x)/\gcd(f(x), f'(x)) = \prod_{p \not| i} g_i(x)$: squarefree

We compute $g_i(x)$'s using the facts repeatedly.

4. Factorization of squarefree polynomial

$f(x)$: squarefree polynomial over $\mathbb{F}_q$

Two popular methods to factor squarefree poly.;

1. Berlekamp method:
   Using the kernel of the linear mapping, say $\phi_{f(x)}$,
   defined by $g(x) \mapsto (g(x)^q - g(x)) \bmod f(x)$.
2. Cantor/Zassenhaus method:
   Using Distinct-degree & Equal-degree factorizations.

We focus on the Berlekamp method in this talk.

5. Berlekamp Algorithm

- Assume that q is odd.
- $f(x) \in \mathbb{F}_q[x]$: squarefree polynomial.

We consider the linear mapping $\phi_{f(x)}$ from $\mathbb{F}_q[x]/(f(x))$ to itself defined by $g(x) \mapsto (g(x)^q - g(x)) \bmod f(x)$.

Step 1: Compute the kernel of $\phi_{f(x)}$, say $\mathcal{N}$.

Step 2: For a random element $g(x) \in \mathcal{N}$,
we find non-trivial factors of $f(x)$ by computing
$\gcd(f(x), g(x))$ and $\gcd(f(x), g(x)^{(q-1)/2} - 1)$.

We note that #(irre. factor(s) of $f(x)$) $= \dim_{\mathbb{F}_q} \mathcal{N}$

6. The purpose of this talk

Main theme: The computation of the kernel $\mathcal{N}$ for $f(x) = x^n - a$ deined over $\mathbb{F}_q$ with $p = \mathrm{char}\mathbb{F}_q$

Previous work: $a = 1 \Rightarrow$ Eugene Prange (1959)

We extend the method to general a.

Assumption on $f(x) = x^n - a$:

- $a \neq 0$ (otherwise, obvious)
- $p \nmid n$ (otherwise, $f(x) = (x^{n/p} - a^{1/p})^p$ and $f(x) \leftarrow x^{n/p} - a^{1/p}$).

$\Rightarrow f(x)$: squarefree (because $\gcd(f(x), f'(x)) = 1$)

7. The kernel $\mathcal{N}$ of $\phi_{f(x)}$

- $f(x)$: squarefree poly. of degree n to be factored.
- $Q = (q_{ij})_{0 \le i,j \le n-1}$: $n \times n$ matrix with
  $(x^j)^q \equiv \sum_{0 \le i \le n-1} q_{ij} x^i \pmod{f(x)}$.

$$\Downarrow$$

- $Q - I_n$: the matrix representation of $\phi_{f(x)}$
  (I_n: $n \times n$ identity matrix)
  $\Rightarrow \mathcal{N}$: the solution space of $(Q - I_n)X = 0$

(I think we generally apply the Gaussian elimination.)

8. The computation of $\mathcal{N}$ for $x^n - a$ (1/2)

Let $p = \mathrm{char}\mathbb{F}_q$, $f(x) = x^n - a$ ($p \not| n$, $a \neq 0$).

Notation

- For $q \bmod n \neq 0$, $\langle q \rangle := \{q^i \bmod n \mid i = 0, 1, 2, \ldots\}$
- $\bar{\alpha} := \{\alpha q^i \bmod n \mid i = 0, 1, 2, \ldots\}$: the orbit containing $\alpha \in \mathbb{Z}/n\mathbb{Z}$ with respect to $\langle q \rangle$. (Let $\ell = \#\bar{\alpha}$)
- $\alpha_i := \alpha q^i \bmod n$ (note that $\alpha q^\ell \bmod n = \alpha (= \alpha_0)$)
- $T_{\bar{\alpha}} := \{\beta_0 x^{\alpha_0} + \beta_1 x^{\alpha_1} + \cdots + \beta_{\ell-1} x^{\alpha_{\ell-1}} \mid \beta_i \in \mathbb{F}_q\}$ for $\bar{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_{\ell-1}\}$

9. The computation of $\mathcal{N}$ for $x^n - a$ (2/2)

Then we have

- $\mathbb{F}_q[x]/(f(x)) = \bigoplus_{\bar{\alpha}} T_{\bar{\alpha}}$
  ($\bar{\alpha}$ runs over all orbits in $\mathbb{Z}/n\mathbb{Z}$ with respect to $\langle q \rangle$)
- $\phi_{f(x)}(T_{\bar{\alpha}}) \subseteq T_{\bar{\alpha}}$

$$\Downarrow$$

$$\mathcal{N} = \bigoplus_{\bar{\alpha}} (\mathcal{N} \cap T_{\bar{\alpha}})$$

We restrict the domain of $\phi_{f(x)}$ to the subspace $T_{\bar{\alpha}} (\subseteq \mathbb{F}_q[x]/(f(x)))$.

10. The computation of $\mathcal{N} \cap T_{\bar{\alpha}}$ (1/2)

For $h(x) = \beta_0 x^{\alpha_0} + \beta_1 x^{\alpha_1} + \cdots + \beta_{\ell-1} x^{\alpha_{\ell-1}}$ in $T_{\bar{\alpha}}$,
we consider the equation $\phi_{f(x)}(h(x)) = 0$
(in other words, $h(x)^q \equiv h(x) \pmod{f(x)}$).

$$\Rightarrow \begin{cases} \beta_0 = a^{\gamma_{\ell-1}} \beta_{\ell-1} \\ \beta_1 = a^{\gamma_0} \beta_0 \\ \quad \vdots \\ \beta_{\ell-1} = a^{\gamma_{\ell-2}} \beta_{\ell-2}, \end{cases}$$

where $q\alpha_i = \gamma_i n + \alpha_{i+1 \bmod \ell}$.

$$\Rightarrow \beta_0 = a^{\gamma_{\ell-1}} \beta_{\ell-1} = a^{\gamma_{\ell-1}} (a^{\gamma_{\ell-2}} \beta_{\ell-2})$$
$$= \cdots$$

11. The computation of $\mathcal{N} \cap T_{\bar{\alpha}}$ (2/2)

Therefore, we have $\beta_0 = a^{\gamma_0+\gamma_1+\cdots+\gamma_{\ell-1}}\beta_0$.

$$\Rightarrow \mathcal{N} \cap T_{\bar{\alpha}} = \begin{cases} \{\beta(x^{\alpha_0} + a^{\gamma_0}x^{\alpha_1} + a^{\gamma_0+\gamma_1}x^{\alpha_2} \\ \quad + \cdots + a^{\gamma_0+\gamma_1+\cdots+\gamma_{\ell-2}}x^{\alpha_{\ell-1}}) \mid \beta \in \mathbb{F}_q\} \\ \quad \text{(if } a^{\gamma_0+\gamma_1+\cdots+\gamma_{\ell-1}} = 1\text{)}, \\ \{0\} \quad \text{(otherwise)}. \end{cases}$$

In the former case,

$x^{\alpha_0} + a^{\gamma_0}x^{\alpha_1} + a^{\gamma_0+\gamma_1}x^{\alpha_2} + \cdots + a^{\gamma_0+\gamma_1+\cdots+\gamma_{\ell-2}}x^{\alpha_{\ell-1}}$

: an element in an $\mathbb{F}_q$**-basis of** $\mathcal{N}$.

12. Example: $f(x) = x^{22} - 2$ defined over $\mathbb{F}_5$ (1/5)

We consider $q = 5$, $f(x) = x^{22} - 2$ ($n = 22, a = 2$).

With respect to $\langle 5 \rangle$, we partitions $\mathbb{Z}/22\mathbb{Z}$ into six orbits:

$\bar{0} = \{0\}$, $\bar{1} = \{1, 5, 3, 15, 9\}$, $\bar{2} = \{2, 10, 6, 8, 18\}$

$\bar{4} = \{4, 20, 12, 16, 14\}$, $\bar{7} = \{7, 13, 21, 17, 19\}$,

$\overline{11} = \{11\}$

We compute the subspace $\mathcal{N}$ by considering $\mathcal{N} \cap T_{\bar{\alpha}}$ for each orbit $\bar{\alpha}$. (Recall that $\mathcal{N}$ is the kernel of the mapping $g(x) \mapsto (g(x)^q - g(x)) \bmod f(x)$.)

13. Example: $f(x) = x^{22} - 2$ defined over $\mathbb{F}_5$ (2/5)

- [The case of $\bar{0} = \{0\}$ ($\ell = 1$):]

$$(q \cdot \alpha_i = \gamma_i \cdot n + \alpha_{i+1})$$

$$5 \cdot \alpha_0 = 5 \cdot 0 \quad = \underline{0} \cdot 22 + 0 \qquad \rightarrow \underline{\gamma_0 = 0},\ \alpha_1 = 0 = \alpha_0$$

So, we have $a^{\gamma_0} = 2^0 = 1$ in $\mathbb{F}_5$.

$$\Rightarrow \mathcal{N} \cap T_{\bar{0}} = \{\beta x^{\alpha_0} \mid \beta \in \mathbb{F}_5\} = \{\beta \cdot 1 \mid \beta \in \mathbb{F}_5\} = \mathbb{F}_5$$

14. Example: $f(x) = x^{22} - 2$ defined over $\mathbb{F}_5$ (3/5)

- [The case of $\bar{1} = \{1, 5, 3, 15, 9\}$ ($\ell = 5$):]

$$
\begin{cases}
 & (q \cdot \alpha_i = \gamma_i \cdot n + \alpha_{i+1}) & \\
5 \cdot \alpha_0 = 5 \cdot 1 & = \underline{0} \cdot 22 + 5 & \rightarrow \underline{\gamma_0 = 0},\ \alpha_1 = 5 \\
5 \cdot \alpha_1 = 5 \cdot 5 & = \underline{1} \cdot 22 + 3 & \rightarrow \underline{\gamma_1 = 1},\ \alpha_2 = 3 \\
5 \cdot \alpha_2 = 5 \cdot 3 & = \underline{0} \cdot 22 + 15 & \rightarrow \underline{\gamma_2 = 0},\ \alpha_3 = 15 \\
5 \cdot \alpha_3 = 5 \cdot 15 & = \underline{3} \cdot 22 + 9 & \rightarrow \underline{\gamma_3 = 3},\ \alpha_4 = 9 \\
5 \cdot \alpha_4 = 5 \cdot 9 & = \underline{2} \cdot 22 + 1 & \rightarrow \underline{\gamma_4 = 2},\ \alpha_5 = 1 = \alpha_0
\end{cases}
$$

So, we have

$a^{\gamma_0+\gamma_1+\gamma_2+\gamma_3+\gamma_4} = 2^{0+1+0+3+2} = 2^6 = -1 \neq 1$ in $\mathbb{F}_5$.

$\Rightarrow \mathcal{N} \cap T_{\bar{1}} = \{0\}$

15. Example: $f(x) = x^{22} - 2$ defined over $\mathbb{F}_5$ (4/5)

- **[The case of $\bar{2} = \{2, 10, 6, 8, 18\}$ ($\ell = 5$):]**

$$\begin{cases} 5 \cdot \alpha_0 = 5 \cdot 2 = \underline{0} \cdot 22 + 10 \rightarrow \underline{\gamma_0 = 0}, \ \alpha_1 = 10 \\ 5 \cdot \alpha_1 = 5 \cdot 10 = \underline{2} \cdot 22 + 6 \rightarrow \underline{\gamma_1 = 2}, \ \alpha_2 = 6 \\ 5 \cdot \alpha_2 = 5 \cdot 6 = \underline{1} \cdot 22 + 8 \rightarrow \underline{\gamma_2 = 1}, \ \alpha_3 = 8 \\ 5 \cdot \alpha_3 = 5 \cdot 8 = \underline{1} \cdot 22 + 18 \rightarrow \underline{\gamma_3 = 1}, \ \alpha_4 = 18 \\ 5 \cdot \alpha_4 = 5 \cdot 18 = \underline{4} \cdot 22 + 2 \rightarrow \underline{\gamma_4 = 4}, \ \alpha_5 = 2 = \alpha_0 \end{cases}$$

So, we have

$$a^{\gamma_0+\gamma_1+\gamma_2+\gamma_3+\gamma_4} = 2^{0+2+1+1+4} = 2^8 = 1 \text{ in } \mathbb{F}_5.$$

$$\begin{aligned} \Rightarrow \ \mathcal{N} \cap T_{\bar{2}} &= \{\beta(x^{\alpha_0} + a^{\gamma_0} x^{\alpha_1} + a^{\gamma_0+\gamma_1} x^{\alpha_2} \\ &\qquad + a^{\gamma_0+\gamma_1+\gamma_2} x^{\alpha_3} + a^{\gamma_0+\gamma_1+\gamma_2+\gamma_3} x^{\alpha_4}) \mid \beta \in \mathbb{F}_5\} \\ &= \{\beta(x^2 + x^{10} + 4x^6 + 3x^8 + x^{18}) \mid \beta \in \mathbb{F}_5\} \end{aligned}$$

16. Example: $f(x) = x^{22} - 2$ defined over $\mathbb{F}_5$ (5/5)

Performing the same procedure as the previous one for the remainder orbits, we obtain

- $\mathcal{N} \cap T_{\overline{4}} = \{\beta(x^{20} + 4x^{16} + 2x^{14} + x^{12} + x^4) \mid \beta \in \mathbb{F}_5\}$
- $\mathcal{N} \cap T_{\overline{7}} = \mathcal{N} \cap T_{\overline{11}} = \{0\}$.

(Dividing by the leading coefficient, we take monic polynomials.)

$$\Rightarrow \mathbb{F}_5\text{-basis of } \mathcal{N} : \{1, x^{18} + x^{10} + 3x^8 + 4x^6 + x^2, \\ x^{20} + 4x^{16} + 2x^{14} + x^{12} + x^4\}$$

$$(\ \#(\text{irre. factors of } f(x)) = \dim_{\mathbb{F}_5}\mathcal{N} = 3)$$

17. Getting factors using the kernel $\mathcal{N}$ (1/2)

We assume that $\dim_{\mathbb{F}_q}\mathcal{N} \geq 2$ and that q is odd.
($\dim_{\mathbb{F}_q}\mathcal{N} = 1 \Rightarrow f(x)$: irreducible)

$g(x)$: random element in $\mathcal{N}$

$$\Rightarrow \quad g(x)\cdot(g(x)^{(q-1)/2}-1)\cdot(g(x)^{(q-1)/2}+1)$$
$$= g(x)^q - g(x) \equiv 0 \pmod{f(x)}$$

We get a non-trivial factor of $f(x)$
when $\gcd(f(x), g(x)) \neq 1, f(x)$
or $\gcd(f(x), g(x)^{(q-1)/2}-1) \neq 1, f(x)$ (or both).

Repeatedly we perform this procedure
until #(our getting factors of $f(x)$)$= \dim_{\mathbb{F}_q}\mathcal{N}$.

18. Getting factors using the kernel $\mathcal{N}$ (2/2)

Note that, for $g(x) \in \mathcal{N}$ with $\gcd(f(x), g(x)) = 1$,
Prob{getting non-trivial factors} $= 1 - (\frac{1}{2})^{\dim_{\mathbb{F}_q}\mathcal{N}-1} \geq \frac{1}{2}$

Memo: $f(x) = \prod_{1 \leq i \leq k} f_i(x)$: factorization of $f(x)$.

$$\begin{array}{ccc} \mathbb{F}_q[x]/(f(x)) & \simeq & \mathbb{F}_q[x]/(f_1(x)) \times \cdots \times \mathbb{F}_q[x]/(f_k(x)) \\ \cup & & \cup \\ \mathcal{N} & \simeq & \{(a_1, \ldots, a_k) \mid a_i \in \mathbb{F}_q\} \end{array}$$

For each i, we have $g(x)^{(q-1)/2} \equiv \pm 1 \bmod f_i(x)$.

"not getting a non-trivial factor of $f(x)$
$\iff$ $g(x)^{(q-1)/2} \leftrightarrow (1, 1, \ldots, 1)$ or $(-1, -1, \ldots, -1)$"

19. Experimental results (1/3)

We list the running time for factoring $x^n - 3$ over $\mathbb{F}_5$ and $\mathbb{F}_7$. (Average value among 100 times.)

- **For the computation of the kernel $\mathcal{N}$,**
  - **Method 1: Using the Gaussian elimination.**
  - **Method 2: Using the method described in this presentation.**
- **After getting the kernel $\mathcal{N}$, the both methods perform the same procedure.**

20. Experimental results (2/3)

- Running time (ms) for factorning $x^n - 3$ over $\mathbb{F}_5$

the value of n (# of irreducible factors)	804 (21)	901 (7)	1002 (9)	1103 (2)
Method 1	31.7	60.7	80.3	27.9
Method 2	20.1	46.4	58.5	2.8

- Running time (ms) for factorning $x^n - 3$ over $\mathbb{F}_7$

the value of n (# of irreducible factors)	801 (2)	904 (68)	1007 (21)	1100 (78)
Method 1	15.3	111.5	113.5	169.8
Method 2	2.8	102.3	93.5	152.9

21. Experimental results (3/3)

- Method 2 (our method) is faster than Method 1 (using the Gausian elimination) for all cases.

- Mehod 1 $\approx$ Method 2 for almost cases.
  (procedure for getting $\mathcal{N}$) $\ll$ (procedure after getting $\mathcal{N}$).

- (# of irreducible factors)$= 2$
  $\Rightarrow$ (Method 2) $\ll$ (Method 1).
  (procedure after getting $\mathcal{N}$) $\ll$ (procedure for getting $\mathcal{N}$).

22. Future works

- **Analysis of computational complexity.**
- **Comparison/Combination with other methods (e.g., Cantor/Zassenhaus method).**
- **Extension to more general cases (e.g., trinomial polynomials and large size of base fields).**

Thank you!