Fast computation of the Nth term of an algebraic series in positive characteristic

Alin Bostan ¹ Gilles Christol ² Philippe Dumas ¹



 1 Inria Saclay

 $^2 {\rm Institut}$ de mathématiques de Jussieu



Functional Equations in Limoges, March 29th 2016



Motivation

- ▶ Ubiquity of algebraic functions (combinatorics, number theory, algebraic geometry)
- ► Confluence of several domains:
 - functional equations
 - automatic sequences
 - complexity theory
- ▶ One of the most difficult questions in modular computations is the complexity of computations mod p for a large prime p of coefficients in the expansion of an algebraic function.

D. Chudnovsky & G. Chudnovsky, 1990

Computer Algebra in the Service of

Mathematical Physics and Number Theory

Problem and main result

Input:

- field $\mathbb{K} = \mathbb{F}_p$
- $-E(x,y) \in \mathbb{K}[x,y]$, with E(0,0) = 0, $E_{y}(0,0) \neq 0$
- $-N \in \mathbb{N}_{>0}$

Output:

- the Nth coefficient of the unique solution $f(x) \in \mathbb{K}[[x]]$ of E(x, f(x)) = 0, f(0) = 0

Main result:

– arithmetic complexity linear in $\log N$ and almost linear in p

First N coefficients

Method	char. 0	char. p
Undetermined coefficients	$O(N^d)^{\dagger}$	√
Fixed point iteration	$\widetilde{O}(N^2)^\P$	√ ¶
Newton iteration	$\widetilde{O}(N)^{\P}$	√ ¶
Linear recurrence	O(N)	√ *

 $[\]dagger_{d \,=\, \deg_y E(x,\,y)} \,\, \P_{\text{FFT used all along }} \, {}^*_{\text{with } p\text{-adic computations}}$

Kung + Traub, 1976 All algebraic functions can be computed fast Chudnovsky + Chudnovsky, 1986 On expansion of algebraic functions in power and Puiseux series, I

Only Nth term (with precomputation)

Method	char. 0	char. p
Binary powering $(d = 1 \text{ only})$	$O(\log_2 N) + O(1)$	✓
Baby steps – Giant steps	$\widetilde{O}(\sqrt{N}) + O(1)$	√ *
Divide and Conquer	×	$O(\log_p N) \times \widetilde{O}(p^{3d})$

Rational series by binary powering

Miller + Brown, 1966, An algorithm for evaluation of remote terms in a linear recurrence sequence Fiduccia, 1985 An efficient formula for linear recurrences

Only Nth term (with precomputation)

Method	char. 0	char. p
Rational series	$O(\log_2 N) + O(1)$	✓
Baby steps – Giant steps	$\widetilde{O}(\sqrt{N}) + O(1)$	√ *
Divide and Conquer	×	$O(\log_p N) \times \widetilde{O}(p^{3d})$

Baby steps – Giant steps after precomputing algebraic equation \longrightarrow differential equation \longrightarrow linear recurrence

Chudnovsky + Chudnovsky, 1988 Approximations and complex multiplication according to Ramanujan

Only Nth term (with precomputation)

Method	char. 0	char. p
Rational series	$O(\log_2 N) + O(1)$	✓
Baby steps – Giant steps	$\widetilde{O}(\sqrt{N}) + O(1)$	√ *
Divide and Conquer	×	$O(\log_p N) \times \widetilde{O}(p^{3d})$

Divide and Conquer after precomputing algebraic equation \longrightarrow Mahler equation \longrightarrow DAC recurrence

Christol + Kamae + Mendès France + Rauzy, 1980, Suites algébriques, automates et substitutions

From algebraic equation to Mahler equation

Algebraic equation $y = 2x + 5xy + 4xy^2 + xy^3$

Mahler equation

$$(2x^{2} + 2x^{3} + x^{4}) f(x)$$

$$+ (1 + x^{2} + 2x^{3} + 2x^{4} + x^{5} + 2x^{6}) f(x^{3})$$

$$+ (2 + 2x^{3} + 2x^{5} + x^{6} + 2x^{9}) f(x^{9}) + x^{9} f(x^{27}) = 0$$

Frobenius $f(x)^p = f(x^p)$

From algebraic equation to Mahler equation

Algebraic equation $y = 2x + 5xy + 4xy^2 + xy^3$ Mahler equation

$$(2x^{2} + 2x^{3} + x^{4}) f(x)$$

$$+ (1 + x^{2} + 2x^{3} + 2x^{4} + x^{5} + 2x^{6}) f(x^{3})$$

$$+ (2 + 2x^{3} + 2x^{5} + x^{6} + 2x^{9}) f(x^{9}) + x^{9} f(x^{27}) = 0$$

Divide-and-conquer recurrence

$$2f_{n-2} + 2f_{n-3} + f_{n-4}$$

$$+ f_{\frac{n}{3}} + f_{\frac{n-2}{3}} + 2f_{\frac{n-3}{3}} + 2f_{\frac{n-4}{3}} + f_{\frac{n-5}{3}} + 2f_{\frac{n-6}{3}}$$

$$+ 2f_{\frac{n}{9}} + 2f_{\frac{n-3}{9}} + 2f_{\frac{n-5}{9}} + f_{\frac{n-6}{9}} + 2f_{\frac{n-9}{9}} + f_{\frac{n-9}{27}} = 0$$

$$f_x = 0 \text{ if } x \notin \mathbb{N}_{\geq 0}$$

```
 \begin{array}{l} \textbf{From} \overset{131}{\text{algebraic}} \overset{129}{\text{equation}} \overset{128}{\text{to}} \overset{128}{\text{to}} \overset{128}{\text{to}} \overset{128}{\text{to}} \overset{126}{\text{to}} \overset{131}{\text{to}} \overset{122}{\text{to}} \overset{123}{\text{to}} \overset{129}{\text{to}} \overset{128}{\text{to}} \overset{129}{\text{to}} \overset
                  +5x^{106} + 6x^{105} + 4x^{104} + 3x^{103} + 4x^{102} + x^{101} + 2x^{100} + 5x^{99} + 3x^{98} + 4x^{97} + 5x^{71} + 4x^{69} + 2x^{68} + 2x^{67} + 2x^{68} + 2x^{68}
                                         +3x^{44} + x^{43} + 6x^{42} + 5x^{41} + 2x^{40} + 5x^{39} + 3x^{38} + 6x^{37} + 3x^{36} + 2x^{35} + x^{34} + 4x^{33} + 3x^{32} + 6x^{31} + 5x^{30} + 6x^{31} + 5x^{30} + 6x^{31} + 5x^{30} + 6x^{31} + 6x^{3
                                         +3x^{29} + 4x^{28} + x^{27} + 3x^{26} + 6x^{25} + 5x^{24} + 5x^{23} + 5x^{22} + 2x^{21} + 4x^{20} + x^{18} + 2x^{17} + 5x^{16} + 5x^{15} + 3x^{14} + 5x^{15} + 3x^{15} + 3x^{14} + 5x^{15} + 3x^{15} + 3x^{1
                                                                                                                                                                                                                                      +4x^{13} + 6x^{12} + 2x^{11} + 4x^{10} + 2x^9 + x^8 + 2x^7 + 5x^6 + 5x^4 + 3x^3 + 4x^2 + 1 \Big) f(x^7)
                                                    + \left(6x^{165} + 5x^{164} + 2x^{163} + 2x^{162} + 4x^{161} + 3x^{160} + 2x^{155} + 3x^{153} + 2x^{151} + 4x^{150} + 3x^{149} + 6x^{147} + 3x^{160} + 3x^{160} + 2x^{163} + 2x^{163} + 3x^{160} + 3x^{1
                                     +x^{144} + 2x^{143} + 5x^{142} + 4x^{141} + 3x^{140} + 6x^{139} + x^{137} + 2x^{136} + 5x^{135} + x^{134} + 3x^{133} + 5x^{132} + 2x^{130}
                                                                   +4x^{129} + 3x^{128} + 4x^{127} + 6x^{126} + 6x^{125} + 6x^{123} + 5x^{122} + 2x^{121} + 2x^{120} + 4x^{119} + 3x^{118} + 5x^{116} + 6x^{125} + 6x^{125}
                                                    +3x^{115} + 4x^{114} + 4x^{113} + x^{112} + 6x^{111} + 4x^{106} + 6x^{104} + 4x^{102} + x^{101} + 6x^{100} + 5x^{98} + 2x^{71} + 3x^{69} + 6x^{104} + 6x
                                     +x^{67} + 2x^{66} + 5x^{65} + 5x^{64} + 3x^{63} + 4x^{62} + 5x^{57} + 4x^{55} + 5x^{53} + 3x^{52} + 4x^{51} + x^{49} + 5x^{46} + 3x^{45} + 4x^{44} + 3x^{45} + 4x^{45} + 3x^{45} + 3x^{4
                                             +6x^{43} + x^{42} + 2x^{41} + 5x^{39} + 3x^{38} + 4x^{37} + 5x^{36} + x^{35} + 4x^{34} + 3x^{32} + 6x^{31} + x^{30} + 6x^{29} + 2x^{28} + 2x^{27} + 2x^{28} + 2x^{27} + 2x^{28} + 2x^{28} + 2x^{27} + 2x^{28} + 2x^{28
       +2x^{25} + 4x^{24} + 3x^{23} + 6x^{21} + 6x^{18} + 5x^{17} + 2x^{16} + 2x^{15} + 4x^{14} + 3x^{13} + 2x^8 + 3x^6 + 2x^4 + 4x^3 + 3x^2 + 6 \Big) f(x^{49})
+ \left(x^{165} + 2x^{164} + 5x^{163} + 5x^{162} + 3x^{161} + 4x^{160} + 5x^{155} + 4x^{153} + 5x^{151} + 3x^{150} + 4x^{149} + x^{147}\right) f(x^{343}) = 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      O(\log_p N) \times p^{O(d)}
```

10/30

Sections

Section operators

$$S_r \sum_{n \ge 0} f_n x^n = \sum_{k \ge 0} f_{pk+r} x^k, \quad 0 \le r < p$$

Lemma

Let f be in $\mathbb{K}[[x]]$ and let $N = (N_{\ell} \cdots N_1 N_0)_p$ be the radix p expansion of N. Then $f_N = (S_{N_{\ell}} \cdots S_{N_1} S_{N_0} f)(0)$.

Sections

$$N = 100000$$

$$p = 7$$

$$= 5 \times 16807 + 6 \times 2401 + 4 \times 343 + 3 \times 49 + 5 \times 7 + 5 \times 1$$

$$= (5, 6, 4, 3, 5, 5)_{7}$$

$$f(x) = f_{0} + f_{1}x + f_{2}x^{2} + f_{3}x^{3} + f_{4}x^{4} + f_{5}x^{5} + \cdots$$

$$S_{5}f(x) = f_{5} + f_{12}x + f_{19}x^{2} + f_{26}x^{3} + f_{33}x^{4} + f_{40}x^{5} + \cdots$$

$$S_{5}S_{5}f(x) = f_{40} + f_{89}x + f_{138}x^{2} + f_{187}x^{3} + f_{236}x^{4} + f_{285}x^{5} + \cdots$$

$$S_{3}S_{5}S_{5}f(x) = f_{187} + f_{530}x + f_{873}x^{2} + f_{1216}x^{3} + f_{1559}x^{4} + f_{1902}x^{5} + \cdots$$

$$S_{4}S_{3}S_{5}S_{5}f(x) = f_{1559} + f_{3960}x + f_{6361}x^{2} + f_{8762}x^{3} + f_{11163}x^{4} + \cdots$$

$$S_{5}S_{6}S_{4}S_{3}S_{5}S_{5}f(x) = f_{15965} + f_{32772}x + f_{49579}x^{2} + f_{66386}x^{3} + f_{83193}x^{4} + \cdots$$

$$S_{5}S_{6}S_{4}S_{3}S_{5}S_{5}f(0) = f_{100000} + f_{217649}x + f_{335298}x^{2} + f_{452947}x^{3} + \cdots$$

$$S_{5}S_{6}S_{4}S_{3}S_{5}S_{5}f(0) = f_{100000}$$

Diagonal

$$F(x,y) = \frac{y(1-5xy-xy^2-3xy^3)}{1-2x-5xy-4xy^2-xy^3} = p=7$$

$$y + 2xy + 4x^2y + x^3y + 2x^4y + 4x^5y + x^6y + 2x^7y + 4x^8y + x^9y + 2x^{10}y + 3x^2y^2 + 5x^3y^2 + x^4y^2 + 5x^5y^2 + 2x^6y^2 + 2x^7y^2 + 6x^9y^2 + 3x^{10}y^2 + 5x^9y^4 + 6x^2y^4 + x^4y^4 + x^5y^4 + x^7y^4 + 3x^8y^4 + 5x^9y^4 + 6x^2y^5 + 2x^3y^5 + 6x^6y^5 + 4x^6y^5 + 5x^7y^5 + 4x^9y^5 + 4x^{10}y^5 + 2x^2y^6 + 3x^3y^6 + 5x^4y^6 + 5x^5y^7 + 6x^3y^7 + 5x^5y^7 + 6x^3y^7 + 5x^3y^9 + x^4y^9 + 3x^5y^8 + 2x^6y^8 + 6x^7y^7 + 3x^8y^8 + 5x^10y^8 + x^3y^9 + x^4y^9 + 3x^5y^9 + 6x^6y^9 + 6x^7y^9 + x^9y^9 + 6x^10y^9 + 5x^3y^{10} + 6x^5y^{10} + 2x^6y^{10} + x^8y^{11} + 4x^8y^{11} + 2x^9y^{11} + 6x^5y^{12} + 2x^6y^{12} + x^7y^{12} + 3x^9y^{12} + 3x^9y^{12} + 3x^{10}y^{12} + 5x^4y^{13} + 5x^6y^{13} + 3x^7y^{14} + 4x^8y^{14} + 2x^9y^{14} + 6x^5y^{15} + x^6y^{15} + 3x^7y^{15} + x^8y^{15} + 2x^9y^{16} + 4x^{10}y^{16} + 6x^5y^{15} + x^6y^{15} + 3x^7y^{15} + x^8y^{15} + 2x^9y^{16} + 4x^{10}y^{16} + 6x^5y^{15} + x^6y^{15} + 3x^7y^{15} + x^8y^{15} + 2x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^5y^{16} + 4x^6y^{16} + 4x^7y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^9y^{16} + 4x^9y^{16} + 4x^{10}y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5x^9y^{16} + 4x^{10}y^{16} + 5$$

Diagonal

Theorem (Furstenberg's theorem)

Every algebraic series is the diagonal of a bivariate rational function.

$$E(x, f(x)) = 0, \quad f(0) = 0, \quad E_y(0, 0) \neq 0$$

$$f(x) = D\frac{a}{b} \quad \text{with} \quad a(x, y) = yE_y(xy, y), \quad b(x, y) = E(xy, y)/y$$

Furstenberg, 1967, Algebraic functions over finite fields

Actions

univariate sections S_r : action on formal series

$$S_r f = S_r D \frac{a}{b}$$
 diagonal and commutation rule
$$S_r D = D S_r$$

bivariate sections S_r : action on bivariate rational functions

$$S_r D \frac{a}{b} = D S_r \frac{a}{b}$$
 Frobenius

pseudo-sections T_r : action on bivariate polynomials

$$DS_r \frac{a}{b} = D \frac{S_r a b^{p-1}}{b} = D \frac{T_r a}{b}$$

Christol, 1979, Ensembles presque périodiques k-reconnaissables

Actions

univariate sections S_r : action on formal series

$$f_N = (S_{N_\ell} \cdots S_{N_1} S_{N_0} f)(0)$$

bivariate sections S_r : action on bivariate rational functions



pseudo-sections T_r : action on bivariate polynomials

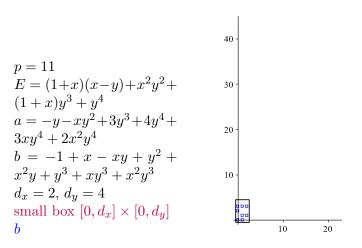
$$f_N = \frac{(T_{N_\ell} \cdots T_{N_1} T_{N_0} a)(0,0)}{b(0,0)}$$

$$DS_r \frac{a}{b} = D \frac{S_r a b^{p-1}}{b} = D \frac{T_r a}{b}$$

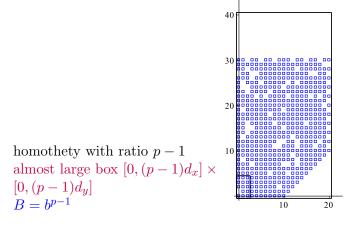
$$\begin{split} d_x &= \max(\deg_x a, \deg_x b), \qquad d_y = \max(\deg_y a, \deg_y b) \\ \mathbb{F}_p[x,y]_{d_x,d_y} &\longrightarrow & \mathbb{F}_p[x,y]_{pd_x,pd_y} &\longrightarrow & \mathbb{F}_p[x,y]_{d_x,d_y} \\ v &\longmapsto & vb^{p-1} = vB &\longmapsto & S_r vB = T_r v \end{split}$$

$$\mathbb{F}_p[x,y]_{d_r,d_y} \text{ left stable by } T_r, \ 0 \leq r$$

$$\mathbb{F}_p[x,y]_{d_x,d_y} \longrightarrow \mathbb{F}_p[x,y]_{pd_x,pd_y} \longrightarrow \mathbb{F}_p[x,y]_{d_x,d_y}$$
$$v \longmapsto vb^{p-1} = vB \longmapsto S_r vB = T_r v$$



$$\mathbb{F}_p[x,y]_{d_x,d_y} \longrightarrow \mathbb{F}_p[x,y]_{pd_x,pd_y} \longrightarrow \mathbb{F}_p[x,y]_{d_x,d_y}$$
$$v \longmapsto vb^{p-1} = vB \longmapsto S_r vB = T_r v$$



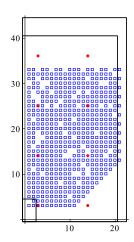
$$\mathbb{F}_p[x,y]_{d_x,d_y} \longrightarrow \mathbb{F}_p[x,y]_{pd_x,pd_y} \longrightarrow \mathbb{F}_p[x,y]_{d_x,d_y}$$
$$v \longmapsto vb^{p-1} = vB \longmapsto S_r vB = T_r v$$

10 20

 $v = xy^3$ in the canonical basis translation by (1,3)large box $[0, pd_x] \times [0, pd_y] xy^3B$

$$\mathbb{F}_p[x,y]_{d_x,d_y} \longrightarrow \mathbb{F}_p[x,y]_{pd_x,pd_y} \longrightarrow \mathbb{F}_p[x,y]_{d_x,d_y}$$

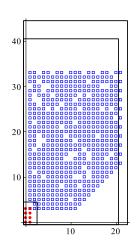
$$v \longmapsto vb^{p-1} = vB \longmapsto S_r vB = T_r v$$



filter r = 3 xv^3B

$$\mathbb{F}_p[x,y]_{d_x,d_y} \longrightarrow \mathbb{F}_p[x,y]_{pd_x,pd_y} \longrightarrow \mathbb{F}_p[x,y]_{d_x,d_y}$$
$$v \longmapsto vb^{p-1} = vB \longmapsto S_r vB = T_r v$$

contraction small box $[0, d_x] \times [0, d_y]$ $T_3xy^3 = S_3xy^3B$



Precomputation of $A_0, A_1, \ldots, A_{p-1}$

All information is in $B = b^{p-1}$.

matrix
$$A_r$$
: $x^n y^m \xrightarrow{\text{translation}} x^n y^m B \xrightarrow{\text{selection}} S_r x^n y^m B$
No computation in $\mathbb{K} = \mathbb{F}_p$, except raising b to the power $p-1$

Cost: $\widetilde{O}(p^2)$ (binary powering, Kronecker substitution, FFT)

Theorem

The Nth coefficient can be computed in time

$$\widetilde{O}(p^2) + O(\log_p N).$$

Only a small part of $B = b^{p-1}$ is enough

Task: Computation of $A_0, A_1, \ldots, A_{p-1}$

row index
$$i = (k, \ell)$$

column index $j = (n, m)$

$$B = \sum_{\alpha,\beta} c_{\alpha,\beta} x^{\alpha} y^{\beta}$$

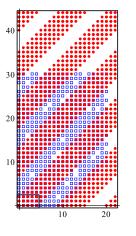
$$x^n y^m \xrightarrow{\text{translation}} x^n y^m B \xrightarrow{\text{selection}} S_r x^n y^m B$$

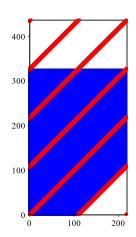
$$x^n y^m \longrightarrow \sum_{\alpha,\beta} c_{\alpha,\beta} x^{n+\alpha} y^{m+\beta} \longrightarrow \sum_{\substack{\alpha,\beta \ (C)}} c_{\alpha,\beta} x^k y^\ell$$

$$(C) \left\{ \begin{array}{l} n + \alpha = pk + r \\ m + \beta = p\ell + r \end{array} \right. \implies \beta - \alpha = p(\ell - k) + n - m$$

Only a small part of $B = b^{p-1}$ is enough

$$p = 11 p = 109$$





Improved precomputation

$$B(x/t,t) = \sum_{\alpha,\beta} c_{\alpha,\beta} x^{\alpha} t^{\beta-\alpha} = \sum_{\delta} \pi_{\delta}(x) t^{\delta}$$

$$B(x/t,t) = b(x/t,t)^{p-1} = \frac{b(x/t,t)^p}{b(x/t,t)} = \frac{b(x^p/t^p,t^p)}{b(x/t,t)}$$

$$\frac{1}{b(x/t,t)} = \sum_{u} c_u(x)t^u \qquad b(x^p/t^p,t^p) = \sum_{v} b_v(x^p)t^{pv}$$

rational series

for free

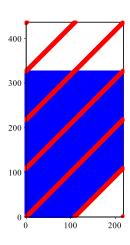
$$\pi_{\delta}(x) = \sum_{u+pv=\delta} c_u(x)b_v(x^p)$$

Improved precomputation

 $\delta = \beta - \alpha = \text{intercept of the strips with}$ the ordinates axis

big leaps from one strip to the next $\widetilde{O}(p)$ (Kronecker substitution, FFT, Newton iteration)

Fiduccia, 1985
An efficient formula for linear recurrences



Main result

Theorem

Let E be in $\mathbb{F}_p[x,y]_{h,d}$ satisfy E(0,0) = 0 and $E_y(0,0) \neq 0$, and let $f \in \mathbb{F}_p[[x]]$ be its unique root with f(0) = 0. One can compute the coefficient f_N of f in

$$h^2(d+h)^2\log_p N + \widetilde{O}(h(d+h)^5p)$$

operations in \mathbb{F}_p .

Timings

Maple implementation; timings on Intel Core i5, 2.8 GHz, 3MB. With p = 9001, $N = 10^{10^k}$, k = 1..6,

$$ComputingTime(N) \simeq 0.0011 \cdot \log_p(N) - 1.3563$$

With $p = NextPrime(2^k)$, k = 1..12,

 $PrecomputingTime(p) \simeq 0.00037 \cdot p \log(p) + 0.04835$

Documents at url http://specfun.inria.fr/dumas/Research/AlgModp/

Theorem

Let E be in $\mathbb{F}_p[x,y]_{h,d}$ satisfy E(0,0) = 0 and $E_y(0,0) \neq 0$, and let $f \in \mathbb{F}_p[[x]]$ be its unique root with f(0) = 0. One can compute the coefficient f_N of f in

$$h^2(d+h)^2\log_p N + \widetilde{O}(h(d+h)^5p)$$

operations in \mathbb{F}_p .

Thanks for your attention!