Computing and Informatics, Vol. 22, 2003, 417–426

SEMIGROUP OF MATRICES OVER $GF(2^S)$ AND ITS RELATION TO AES*

Otokar Grošek

Slovak University of Technology
812 19 Bratislava, Slovakia
e-mail: grosek@kmat.elf.stuba.sk

Július Šiška

KPMG Slovensko
Mostová 2, 811 02 Bratislava, Slovakia
e-mail: jsiska@kpmg.sk

Communicated by Juraj Hromkovič

Abstract. We introduce to readers a new matrix C for `MixColumn` operation for AES algorithm for discussion. This matrix has significantly larger multiplicative order, $\mathrm{ord}(C) = 340$, than the used one which is 4 only. This makes so called XSL attack less effective. It is possible to find such a matrix due to our new Euler-Fermat-like theorem and its corollaries for regular circulant matrices over $GF(p^s)$.

Keywords: AES, MixColumn operation, Euler-Fermat theorem

1 INTRODUCTION

After a very careful discussion, the new symmetric block cipher standard referred to as AES was published. Its description can be found in [12]. Another reference book is [3]. It is not surprising that this discussion is still going on. In this paper we contribute to this aim from the point of view of the semigroup of matrices over $GF(2^s)$.

* This material is based upon the work supported under Grant No. VEGA 1/0161/03.

For convenience of a reader, who is not interested in algebraic details, this paper is organized as follows. In Section 2 we present a brief description of `MixColumn` operation which is under our careful algebraic consideration. In Section 3 we present our proposal to modify AES which consists of changing the matrix `MixColumn` to another one satisfying all but one conditions from the design criteria of AES, and one more. Algebraic details are left to the Appendix. There we present new results, so called Euler-Fermat-like theorem for matrices over $GF(2^s)$, namely for semigroup $\mathcal{C}_n$ of all $n \times n$ circulant matrices over $GF(p^s)$. In a special case we get a result which can enlarge the multiplicative order of `MixColumn` matrix significantly from 4 to 340. This makes so called XSL attack less effective. In Section 4 we present a discussion of effectivity of proposed matrix as well as results of our computer search for such matrices.

2 AES AND ITS MIXCOLUMN OPERATION

In `MixColumn` operation multiplication by polynomial $c(x) = 03x^3+01x^2+01x+02$ is used, and the result is reduced by polynomial x^4+1 over $GF(2^8)$. In AES algorithm, $GF(2^8)$ is represented as an extension $GF(2)(\theta)$ of the field $GF(2)$ where θ is the root of primitive polynomial $x^8+x^4+x^3+x+1$. Then, for example, $05 \in GF(2^8)$ is, in fact, θ^2+1. This operation can be viewed as a multiplication by 4×4 matrix over $GF(2^8)$, thereto denoted as A. Its elements are written as two digit hexadecimal numbers:

$$A = \begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix}.$$

To be clear, let $a(x) = 04x^3+42x^2+01x+0C$, then $a(x)*c(x) = 5Fx^3+85x^2+CCx+5D$. In matrix notation with polynomials as columns, $a(x) = a_3x^3+a_2x^2+a_1x+a_0$ is represented as $(a_0, a_1, a_2, a_3)^T$

$$A \begin{pmatrix} 0C \\ 01 \\ 42 \\ 04 \end{pmatrix} = \begin{pmatrix} 5D \\ CC \\ 85 \\ 5F \end{pmatrix}.$$

Matrix A is circulant and invertible (regular). In AES it represents transformation of four 8-bit words to another four 8-bit words. The design criteria were as follows [3]:

1. *Regularity* — The transformation must be invertible.
2. *Symmetry* — The transformation must be symmetric for all columns.

3. *Dimensions* — The transformation is a bricklayer transformation operating on 4-byte columns.
4. *Linearity* — The transformation is preferably linear over $GF(2)$.
5. *Diffusion* — The transformation has to have *relevant* diffusion.
6. *Performance on 8-bit processors* — The performance of the transformation on 8-bit processors has to be high.

Matrix A satisfies all of these criteria. The condition 5 for a linear transformation, where $w_b(a)$ is is used to denote the number of non-zero values of four 8-bit numbers a (weight of a), is based on the following notion.

Definition 1. The differential branch number of a linear transformation λ over $GF(2)$ [1] is given by

$$B_d(\lambda) = \min_{a \neq 0}\{w_b(a) + w_b(\lambda(a))\}. \tag{1}$$

It is known [2] that for any linear transformation λ given by an $n \times n$ matrix the branch number is upper bound by

$$B_d(\lambda) \leq n + 1.$$

Matrix A reaches upper bound with $B_d(\lambda) = 5$.

3 OUR PROPOSAL FOR A NEW MIXCOLUMN MATRIX

Criterion 2 from the previous section implies to use circulant matrices only (matrices with the same rows are clearly not acceptable). A non-circulant matrix is more vulnerable to so called "timing analysis" and "differential power analysis".

Let $\mathbb{F} = GF(2)(\theta)$. Murphy and Robshaw [11] observed that it is possible to map AES into the field $\mathbb{F}^{128}$ and then to use simple operations with elements of $GF(2^8)$ instead of bit represented S-boxes. The new algorithm is called BES. The BES processes 1024 bit blocks using a key of the same size. They invented so called XSL attack applicable to the BES as well, using a system of sparse equations over $GF(2^8)$. Applicability of this attack in practice is in question. A description of the BES algorithm is very simple. For $b \in \mathbb{F}^{128}$ the round function is as follows:

$$b \mapsto M_B \cdot b^{(-1)} + (k_B)i, \tag{2}$$

where inversion is componentwise in $\mathbb{F}$, $(k_B)i$ are round keys and M_B is matrix of the type 128×128 over $\mathbb{F}$. The order of this matrix is small, namely 16, since the order of A is 4 only. This weakness simplifies cryptanalysis of AES too.

Recently Courtois and Pieprzyk [1] published their new XSL attack on AES. S-box of the algorithm can be described by quadratic boolean functions, and then by solving this system of equations over $GF(2)$ the key can be found.

[1] Hence $\lambda(a) \oplus \lambda(b) = \lambda(a \oplus b)$.

Our proposal to modify AES consists of changing matrix A to another one satisfying conditions 1–5, and one more. Unfortunately, performance on 8-bit processors (condition 6) is less effective in this case.

7. *High multiplicative order of matrix* — The multiplicative order of `MixColumn` matrix must be at least the same as the number of rounds in AES, i.e. 10, 12, or 14.

By adding this condition, our aim is to change A with order 4, since for any four-tuple of 8-bit numbers a_0, a_1, a_2, a_3 the following is valid:

$$A^4 \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix}. \qquad (3)$$

This might allow us to simplify AES, and introduce cryptanalysis like in [11]. Hence the higher order could bring higher security of AES. When searching for appropriate matrix the difficulties are mostly with condition 5. Next we show that there are matrices (almost) satisfying all of required 7 conditions. Namely, matrix C associated to the polynomial $05x^3 + 01x^2 + 01x + 02$, and its Jordan's form J_C is

$$C = \begin{pmatrix} 02 & 05 & 01 & 01 \\ 01 & 02 & 05 & 01 \\ 01 & 01 & 02 & 05 \\ 05 & 01 & 01 & 02 \end{pmatrix} \quad J_C = \begin{pmatrix} 07 & 00 & 00 & 00 \\ 01 & 07 & 00 & 00 \\ 00 & 01 & 07 & 00 \\ 00 & 00 & 01 & 07 \end{pmatrix}.$$

From Corollary 2 in the Appendix below we know the upper bound for the order which is $4(2^8 - 1) = 1020$ in this case. Thus, one can find the order of C simply by finding C, C^2, C^3, C^4 while the resulting matrix is not diagonal with some $c \in GF(2^s)$ on the diagonal. Let $C^r, r \leq 4$ be the first such exponent, and t the order of c in the field $GF(2^s)$. Then, clearly, $rt \leq 1020$ is the order of the matrix C. In our case $\mathrm{ord}(C) = 340$.

Lemma 1. The differential branch number of C is 5 (i.e. the same as for A).

Proof. If $x = (x_3, x_2, x_1, x_0)^T$ is of weight 4, then $b = (b_3, b_2, b_1, b_0)^T = C \cdot x$ has weight at least 1, otherwise C would be singular.

Let for now x be of weight 3. We must show that b is of weight at least 2. First observe that since C is nonsingular, b has weight greater than 0. Let b be of weight 1 now, and without loss of generality let $b_3 \neq 0$. If $D = C^{-1}$, then associated polynomial is

$$d(x) = 29x^3 + DAx^2 + 85x + A3.$$

Matrix $D = (d_{ij})_{i=1..4}^{j=1..4}$ contains no zero and

$$D\begin{pmatrix} b_3 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} b_3d_{11} \\ b_3d_{21} \\ b_3d_{31} \\ b_3d_{41} \end{pmatrix} = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \\ x_0 \end{pmatrix}.$$

Since x does not contain zero as well, this is a contradiction to our supposition that x is of weight 3.

Analogously one could prove that for x of weight 2 or 1 respectively, b is of weight at least 3 or 4. □

4 CONCLUSIONS

The sum of Hamming weights of binary written coefficients of the polynomial $d(x)$ is greater than the sum the same weights for the polynomial $c(x)^{-1} = 0Bx^3 + 0Dx^2 + 09x + 0E$. Hence, deciphering in AES with replaced matrix C is less effective than that with the original matrix A (for description of effectivity see [2]). Next we shall discuss this in details.

Let `xtime` denote the multiplication of an element of $GF(2^8)$ by 02, and by `EXOR` we denote the time for 8-bit operation XOR. Then multiplication by A^{-1} takes 4×18 operations `xtime` and 4×7 operations `EXOR`. Multiplication by $D = C^{-1}$ takes 4×21 operations `xtime` and 4×11 operations `EXOR`. Hence deciphering with D is less effective than that with A^{-1}, approximately by $7/25 = 0.28$. But $\mathrm{ord}(A) < \mathrm{ord}(C) = 340$, i.e. condition 7 is satisfied.

Thus we found a matrix which is less effective in high performance on 8-bit processors (condition 6 above) but satisfies one new criterion for its multiplicative order. This new criterion makes the XSL attack less effective.

It is possible to perform exhaustive search for suitable matrices in two steps. Firstly we found matrices

1. with no zeroes;
2. with the differential branch number equal to 5;
3. column multiplication by matrix C should take at most 4×18 `xtime` operations and 4×14 `EXOR` operations.

Then we selected the matrices where column multiplication by C, and by its inverse matrix D should take at most 4×19 `xtime` operations and 4×18 `EXOR` operations together.

We found 176 matrices satisfying these criteria with different multiplicative order. Only 16 out of them have their order 340. (Other matrices have their order 4 which we are trying to avoid.) The polynomials for matrices of order 340 are listed in the following Table 1.

Polynomial $c(x)$	xtime	EXOR	Inverse $d(x)$	xtime	EXOR
$01x^3 + 05x^2 + 03x^1 + 23$	8	4	$58x^3 + 0ax^2 + 01x^1 + 06$	11	4
$01x^3 + 06x^2 + 58x^1 + 0a$	11	4	$03x^3 + 23x^2 + 01x^1 + 05$	8	4
$01x^3 + 0ax^2 + 58x^1 + 06$	11	4	$03x^3 + 05x^2 + 01x^1 + 23$	8	4
$01x^3 + 23x^2 + 03x^1 + 05$	8	4	$58x^3 + 06x^2 + 01x^1 + 0a$	11	4
$03x^3 + 05x^2 + 01x^1 + 23$	8	4	$01x^3 + 0ax^2 + 58x^1 + 06$	11	4
$03x^3 + 23x^2 + 01x^1 + 05$	8	4	$01x^3 + 06x^2 + 58x^1 + 0a$	11	4
$05x^3 + 01x^2 + 23x^1 + 03$	8	4	$06x^3 + 01x^2 + 0ax^1 + 58$	11	4
$05x^3 + 03x^2 + 23x^1 + 01$	8	4	$06x^3 + 58x^2 + 0ax^1 + 01$	11	4
$06x^3 + 01x^2 + 0ax^1 + 58$	11	4	$05x^3 + 01x^2 + 23x^1 + 03$	8	4
$06x^3 + 58x^2 + 0ax^1 + 01$	11	4	$05x^3 + 03x^2 + 23x^1 + 01$	8	4
$0ax^3 + 01x^2 + 06x^1 + 58$	11	4	$23x^3 + 01x^2 + 05x^1 + 03$	8	4
$0ax^3 + 58x^2 + 06x^1 + 01$	11	4	$23x^3 + 03x^2 + 05x^1 + 01$	8	4
$23x^3 + 01x^2 + 05x^1 + 03$	8	4	$0ax^3 + 01x^2 + 06x^1 + 58$	11	4
$23x^3 + 03x^2 + 05x^1 + 01$	8	4	$0ax^3 + 58x^2 + 06x^1 + 01$	11	4
$58x^3 + 06x^2 + 01x^1 + 0a$	11	4	$01x^3 + 23x^2 + 03x^1 + 05$	8	4
$58x^3 + 0ax^2 + 01x^1 + 06$	11	4	$01x^3 + 05x^2 + 03x^1 + 23$	8	4

Table 1. Table of polynomials for matrices of order 340

REFERENCES

[1] Courtois, N.—Pieprzyk, J.: Cryptanalysis of Block Ciphers with Overdefined Systems of Equations. Proceedings of Asiacrypt'02, Lecture Notes in Computer Science, Springer-Verlag, 2002.

[2] Daemen, J.—Rijmen, V.: The Rijndael Block Cipher, AES Proposal: Rijndael, 1999.

[3] Daemen, J.—Rijmen, V.: The Design of Rijndael, AES — The Advanced Encryption Standard. Springer-Verlag, Berlin, 2002.

[4] Ecker, A.: Finite Semigroups and the RSA-Cryptosystem. EuroCrypt'82, LNCS 149, Springer-Verlag, pp. 353–369, 1983.

[5] Grošek, O.—Šiška, J.: Signature Schemes Based on Matrices. Accepted for publication in Congressus Numerantium.

[6] Jungnickel, D.: Finite Fields. Structure and Arithmetics, Wissenschaftverlag, 1992.

[7] Laššák, M.—Porubský, Š.: Fermat-Euler Theorem in Algebraic Number Fields. J. Number Theory 60, pp. 254–290, 1996.

[8] Lidl, R.—Niederreiter, H.: Introduction to Finite Fields and Their Applications. Oxford University Press, 1986.

[9] Lidl, R.—Niederreiter, H.: Finite Fields. Cambridge University Press, Cambridge, 1997.

[10] Marshall, I. B.: On the Extension of Fermat's Theorem to Matrices of Order n. Proceedings of the Edinburgh Math. Soc. 6, pp. 85–91, 1939–41.

[11] Murphy, S.—Robshaw, M. J. B.: Essential Algebraic Structure within the AES. Advances in Cryptology — CRYPTO 2002, LNCS 2442, Springer-Verlag, Berlin 2002, pp. 1–16.

[12] National Institute of Standards and Technology, NIST FIPS PUB 197, "Advanced Encryption Standard". U. S. Department of Commerce, 26th November 2001.

[13] Niven, I.: Fermat's Theorem for Matrices. Duke Math. J. 15, pp. 823–826, 1948.

[14] Schwarz, Š.: On the Semigroup of Binary Relations on a Finite Set. Czech. Math. J., Vol. 20, 1970, No. 95, pp. 632–679.

[15] Schwarz, Š.: Fermat's Theorem for Matrices Revisited. Math. Slovaca, Vol. 4, 1985, No. 35, pp. 343–347.

[16] Varadharajan, V.—Odoni, R.: Extension of RSA Cryptosystems to Matrix Rings. Cryptologia, Vol. 2, 1985, No. 9, pp. 140–153.

[17] Varadharajan, V.: Trapdoor Rings and Their Use in Cryptography. Advances in Cryptology — CRYPTO '85, LNCS 218, Springer-Verlag, pp. 369–395, 1986.

Otokar Grošek graduated at the Comenius University (1973), assigned to Professor Š. Schwarz as a graduate student (PhD. – 1978), (Prof. – 1998). He is working at the Department of Mathematics, FEI STU in Bratislava. Since 1983 he is working in cryptology. Member of American Mathematical Soc., Society for Industrial and Appl. Mathematics, Slovak Mathematical Soc., Editor of the Tatra Mountains Mathematical Publication.

Július Šiška graduated at the Comenius University (1999), assigned to Professor O. Grošek as a graduate student (PhD. – 2003). Since 1997 he is working in the field of cryptography and computer security. Since 2003 he has joined KPMG, where he is working at Information Risk Management department (IRM).

5 APPENDIX

5.1 Periodic Semigroups

Multiplicative semigroups of "trapdoor rings" seem to have been first considered by Ecker [4]. By this he meant a ring with addition and multiplication, where addition is mainly used for calculations, and there exists Euler-Fermat-like theorem for elements under multiplication. Later Varadharajan [17] used this notion exactly for rings where $k = 1$ only (see below). In [5] we generalized this approach to an arbitrary $k > 0$. For reader's convenience, we start to review some basic facts (see e.g. [14, 7]).

Let S be a finite semigroup. Then for any element $x \in S$, in the sequence x, x^2, $\dots$, for some $1 \le s < t$ $x^s = x^t$ must hold. Let $k(x) = k$, and $d(x) = d$ be the least exponent for which $x^k = x^{k+d}$. It is well known, and easy to prove, that

$$\{x^k, \dots, x^{k+d-1}\} \tag{4}$$

forms a cyclic group of order d, and this group is determined by the (unique) idempotent $e = x^r,\ k \le r \le k + d - 1$ belonging to this group.

Definition 2. Let S be a finite semigroup, and define numbers K, D and R as follows:

$$\begin{aligned} K &= \max\{k(x) \,|\, x \in S\} \\ D &= \mathrm{lcm}\{d(x) \,|\, x \in S\}, \end{aligned}$$

and R is uniquely determined integer such that $K \le R < K + D$ and $D|R$.

Euler-Fermat theorem for finite semigroups is as follows:

Theorem 1 ([14])**.** For any $x \in S$ and K, D, R defined as above, holds

$$x^{K+D} = x^K,$$

and x^R is an idempotent. Moreover, K, D and R are the least positive integers having this property.

There is a limited number of semigroups for which the "universal exponents" K, D and R are known (see e.g. [5]). Recall also that for RSA algorithm $K = 1, D = \mathrm{lcm}\{p-1, q-1\}$, although $D = (p-1)(q-1)$ is commonly used.

5.2 Matrices Over $GF(p^s)$

Let $GF(q)$ be the finite field with $q = p^s$ elements ($s \ge 1, p$ a prime), and S_n be the multiplicative semigroup of all $n \times n$ matrices over $GF(q)$. In 1948 I. Niven [13] has proved a similar result for S_n like Schwarz in Theorem 1. This result has been strengthened for singular matrices by Schwarz in 1978, and published in 1985 [15].

Let p be a prime. Then by $p\{y\}$ we denote the least value in the sequence 1, p, p^2, p^3, $\dots$ such that $p^s \ge y$.

Theorem 2. Let $A \in S_n$ be any $n \times n$ matrix over $GF(q)$. Then we have

1. If $rank(A) = h, h < n$ then

$$A^{h+1} = A^{h+1+\lambda(h)}.$$

2. If $rank(A) < n$ then
$$A^n = A^{n+\lambda(n-1)}.$$

3. For any $A \in S_n$ we have
$$A^n = A^{n+\lambda(n)},$$

This implies that $K = n, D = \lambda(n) = p\{n\}\text{lcm}\{q^n - 1, q^{n-1} - 1, \ldots, q - 1\}$.

Especially, for $n = 4, q = 2^8, p\{n\} = 2, \lambda(n) = \text{lcm}\{2^{32} - 1, 2^{24} - 1, 2^{16} - 1, 2^8 - 1\}$ we have $D = 1130315132959740$. The magnitude of D is so high since it includes all matrices $A \in S_n$.

When focusing on large D, matrices over $GF(2)$ or $GF(2^s)$ could be of particular interest, e.g. for n where $2^n - 1$ or $2^{sn} - 1$ is a prime.

Next we prove Euler-Fermat-like theorem for regular circulant matrices over $GF(p^s)$. The universal exponent D for this special group of matrices is remarkably smaller than exponent used in Theorem 2. Denote by $\mathcal{C}_n$ the set of all $n \times n$ circulant matrices over $GF(p^s)$.

Theorem 3. Let $M \in \mathcal{C}_n$ be a matrix over $GF(p^s)$ and $\text{rank}(M) = n$. Then for $n = p^k$ we have $D = n(p^s - 1)$.

Proof. First observe that $\mathcal{C}_n$ can be assumed as a ring where addition and multiplication of a_i is over $GF(p^s)$, P is the permutation matrix of size $n \times n$

$$P = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix},$$

and $M = \sum_{i=0}^{n-1} a_i P^i$. As usual, $P^0 = I$. This ring is exactly the same as the ring of all polynomials of the degree less than n over $GF(p^s)$ where addition and multiplication is performed $\mod (x^n - 1)$. The isomorphism is given by

$$\Phi : M = \sum_{i=0}^{n-1} a_i P^i \mapsto \sum_{i=0}^{n-1} a_i x^i. \tag{5}$$

Clearly, $\Phi(M_1 M_2) = \Phi(M_1)\Phi(M_2) \mod (x^n - 1)$ and $\Phi(M_1 + M_2) = \Phi(M_1) + \Phi(M_2)$. Thus images of regular matrices are precisely polynomials coprime to $x^n - 1$. Next we show that the upper bound follows from arithmetic over the ring of polynomials $\mod (x^n - 1)$.

Let $f(x) = \sum_{i=0}^{n-1} a_i x^i$ be a polynomial over $GF(p^s)$ associated with the matrix M of the size $n \times n$ given by (5). Then for $n = p^k$

$$f(x)^n = \sum_{i=0}^{n-1} a_i^n x^{in \mod n} = \sum_{i=0}^{n-1} a_i^n = c.$$

Since in $GF(p^s)$ we have $c = \sum_{i=0}^{n-1} a_i^n = (\sum_{i=0}^{n-1} a_i)^n$, one can write a primitive element $g \in GF(p^s)^*$ as $g = \sum_{i=0}^{n-1} a_i$. According to $\gcd(n, p^s - 1) = 1$, $c = g^n$ is also an element of $\mathrm{ord}(p^s - 1)$, and this completes the proof. □

A straightforward corollary states that circulant matrices over $GF(2^s)$ have a relatively small D.

Corollary 1. Let $M \in \mathcal{C}_4$ be a 4×4 circulant matrix over $GF(2^s)$. Then M^4 is a diagonal matrix such that $M^4 = cI$ for a suitably chosen $c \in GF(2^s)$.

For further analysis of `MixColumn` matrix the exact value of c from Corollary 1 and other properties of 4×4 regular circulant matrices are needed.

Let M be a 4×4 circulant matrix over $GF(2^s)$, $M = a_0 I + a_1 P + a_2 P^2 + a_3 P^3$, $a_0, a_1, a_2, a_3 \in GF(2^s)$, where

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then $M^2 = b_0 I + b_1 P^2$, where $b_0 = a_0^2 + a_2^2, b_1 = a_1^2 + a_3^2 \in GF(2^s)$. Since $P^4 = I$, this yields $M^4 = M^2.M^2 = cI$, where $c = a_0^4 + a_1^4 + a_2^4 + a_3^4 \in GF(2^s)$.

Moreover, $M \in \mathcal{C}_4$ is regular iff M^4 is regular, or equivalently, iff $c \neq 0$. The product of two such matrices $M_1 = a_0 I + a_1 P + a_2 P^2 + a_3 P^3$, $M_2 = b_0 I + b_1 P + b_2 P^2 + b_3 P^3$ yields

$$\begin{array}{ll} M_1.M_2 = & \\ (a_0 I + a_1 P + a_2 P^2 + a_3 P^3)(b_0 I + b_1 P + b_2 P^2 + b_3 P^3) & = \\ (a_0 b_0 + a_1 b_3 + a_2 b_2 + a_3 b_1) + (a_0 b_1 + a_1 b_0 + a_2 b_3 + a_3 b_2) P & + \\ (a_0 b_2 + a_1 b_1 + a_2 b_0 + a_3 b_3) P^2 + (a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0) P^3, & \end{array}$$

again a circulant matrix. Hence, matrices for which $c \neq 0$ form a subgroup of $\mathcal{C}_4$ belonging to the idempotent I, and $c = c_1 c_2$. Moreover, M^2 is all zero matrix iff $b_0 = 0, b_1 = 0$. Since the underlying field is of characteristic 2, this is if and only if $a_0 + a_2 = a_1 + a_3 = 0$. Thus we have

Corollary 2. For any $M \in \mathcal{C}_4$ over $GF(2^s)$, and $\mathrm{rank}(M) = 4$

$$M = M^{1+D}, \quad D = 4(2^s - 1).$$

Especially, for $s = 8$ we have $D = 1020$.