DE GRUYTER
OPEN

TATRA MOUNTAINS
Mathematical Publications

DOI: 10.1515/tmmp-2016-0033
Tatra Mt. Math. Publ. **67** (2016), 93–98

ROTATION-EQUIVALENCE CLASSES OF BINARY VECTORS

Otokar Grošek — Viliam Hromada

ABSTRACT. In this paper we study equivalence classes of binary vectors with regards to their rotation by using an algebraic approach based on the theory of linear feedback shift registers. We state the necessary and sufficient condition for existence of an equivalence class with given cardinality and provide two formulas. The first represents the sharp distribution of cardinalities for given length and Hamming weight of binary vectors and the second enables us to determine the number of different classes with the same cardinality.

1. Introduction

In cryptography and coding theory, there are many algorithms, which use rotation of a binary vector. One interesting example is the McEliece cryptosystem [4], [5] that uses quasi-cyclic codes, e.g., QC-LDPC (low-density parity-check codes) as proposed by B a l d i et al. [1], [2]. Another interesting example, where equivalence classes of rotation of binary vectors are studied, is the rotational cryptanalysis of various cryptosystems [7].

McEliece version with QC-LDPC codes uses quasi-cyclic matrices, which are matrices consisting of blocks of binary circulant matrices. A binary circulant matrix is a matrix, in which each row vector is rotated one element to the right relative to the preceding row. It is therefore helpful to know the corresponding equivalence class of a binary vector with regards to its rotation and the cardinality of this class. These rotations are calculated in the real time in the implementation of these cryptosystems, since it is sufficient to store into memory only the first rows of used binary circulant matrices and the other rows can be computed on-demand by simple rotations, which greatly lowers the memory requirements.

2010 Mathematics Subject Classification: 11T71, 94A60.
Keywords: rotational equivalence classes, binary vectors, binary vector rotation, rotational classes cardinality.
This project is supported by NATO SPS Project G4520.

This paper deals with the sufficient and necessary condition for the existence of a class with given cardinality and the formulas presented in this paper can be used to determine the structure of classes for binary vectors with given length and Hamming weight, i.e., they present the sharp distribution of cardinalities and the number of different classes with the same cardinality.

2. Rotational equivalence classes

Let V_n be n-dimensional vector space over $\mathbb{F}_2$, and $E_t = \{e \mid hw(e) = t\} \subset V_n$, where hw is the Hamming weight. Number of such vectors is equal to $\binom{n}{t}$, i.e., $|E_t| = \binom{n}{t}$. Let $\mathbf{A}$ be the associated $n \times n$ matrix to the characteristic polynomial $f(x) = x^n + 1$ over $\mathbb{F}_2$ of the LFSR as defined in [3]

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

For any $u \in E_t$ let $[u] = \{u, u\mathbf{A}, \dots, u\mathbf{A}^{d-1}\}$ be a class of words (state vectors) obtained from u by consecutive shifts of this vector, where d is the smallest period of this sequence. Hence $u = u\mathbf{A}^d$.

Let ϱ_n be a relation defined on E_t such that $u\varrho_n v$ if and only if u, v belong to the same class. Then ϱ_n is an equivalence relation on E_t and $u\varrho_n = [u]$. The cardinality of such classes may vary from 1 to n as shown in the next example.

Example 1. Here we present several typical cases:

1. If $n = 6, t = 6$, then clearly $|[1,1,1,1,1,1]\varrho_6| = 1$.
2. If $n = 7, t = 3$, then all classes have the same cardinality 7, and there are 5 such classes.
3. If $n = 10, t = 4$, then we have 20 classes of cardinality 10 and 2 classes of cardinality 5, namely,
$$[0,0,1,0,1,0,0,1,0,1]\varrho_{10} \quad \text{and} \quad [0,0,0,1,1,0,0,0,1,1]\varrho_{10}.$$
All divisors of 10 are 2, 5, 10, but there is no class with cardinality 2.

Remark 1. A necessary condition for having the same cardinality for all classes is $n \mid \binom{n}{t}$. From [6] it can be deduced that $\binom{n}{t}$ is divisible by $\frac{n}{\gcd(n,t)}$. Thus, if $\gcd(n,t) = 1$, then $\binom{n}{t}$ is divisible by n. The converse is not true, e.g., $\binom{10}{4} = 210$, and $\gcd(10,4) > 1$, but $10 \mid 210$. As shown in Example 1 in this case there exist 2 classes with 5 elements.

It follows from the theory of LFSR that for any initial state u the cardinality of $[u] = u\varrho_n$ divides the order of $\mathbf{A}$ in the general linear group $GL(n, \mathbb{F}_2)$. Equivalently, the cardinality of $[u]$ divides the order of $f(x)$ in $\mathbb{F}_2[x]$, i.e., the smallest ℓ such that $f(x) \mid x^\ell + 1$. This ℓ coincides with the order of $\mathbf{A}$. Since in our case the order of $\mathbf{A}$ is n, $d \mid n$. Next, we prove a necessary and sufficient condition for having a class of a given cardinality d.

Theorem 1. *Let ϱ_n be the equivalence relation on E_t defined above. Then there exists a class $u\varrho_n$ with cardinality d if and only if $d \mid n$ and $\frac{n}{d} \mid t$.*

Proof. The first condition $d \mid n$ of our claim results from general theory of LFSR (cf. [3]).

Next we concatenate u from smaller parts. Thus we will speak about words over the alphabet $\{0, 1\}$ of a given length, i.e., elements from the free semigroup $\mathfrak{S} = \{0, 1\}^*$. If there is a class with d elements, $|u\varrho_n| = d \leq n$. Then

$$u = u_n u_{n-1} \dots u_1 = u_{n-d} u_{n-d-1} \dots u_1 u_n u_{n-1} \dots u_{n-d+1} \tag{1}$$

and we can concatenate u from words $w_1 \| w_2 \| \dots \| w_z$, where the length of w_i is $|w_i| = d, i = 1, 2, \dots z$, and $z = n/d$. Next we show that

1. all these words are the same, i.e., $w_1 = w_2 = \dots = w_z$;
2. the weight of w_i is $\frac{t}{z} = \frac{td}{n}$ for $i = 1, 2, \dots, z$, providing $n \mid td$.

Clearly, the second claim is a direct consequence of the first one.

From the definition of classes it follows that if $u = w_1 \| w_2 \| \dots \| w_z$, then

$$u = u\mathbf{A}^d = w_z \| w_1 \| w_2 \| \dots \| w_{z-1}.$$

Thus $w_1 = w_z, w_2 = w_1, \dots, w_z = w_{z-1}$ which concludes the first part of the proof.

On the other hand, let $d \mid n$ and $\frac{n}{d} \mid t$. Then we can construct a word w

$$w = \overbrace{00\dots0}^{d-td/n} \| \overbrace{11\dots1}^{td/n} = 0^{d-td/n} 1^{td/n},$$

i.e., $u = w^{n/d}$, and $[u]$ contains precisely d elements. □

Corollary 1. *All classes $u\varrho_n$ have the same cardinality if and only if $t = 0$ or $\gcd(n, t) = 1$.*

Proof. The cases $t = 0$ and $t = n$ are trivial. Let for now $0 < t < n$. From Theorem 1 it follows that the cardinality d of a class must satisfy $n \mid td$ and $d \mid n$. If $\gcd(n, t) = 1$, then $d = n$.

On the other hand, if all classes have the same cardinality d and $\gcd(n, t) = k > 1$, then we can construct two words, namely

1. $u = 0^{n-t}1^t$, which yields $|u\varrho_n| = n$, and
2. $v = w_1 \mid \ldots \mid w_k$ such that $w_i = w, i = 1, \ldots, k, |w| = n/k, hw(w) = t/k$. From the construction it follows that $|v\varrho_n| = n/k \neq n$, a contradiction with our supposition.

This completes the proof. □

Here is a more complex example:

EXAMPLE 2. Let $n = 20$, $t = 10$. Then we have the following distribution of classes:

- 9225 classes with cardinality $d = 20$;
- 25 classes with cardinality $d = 10$;
- 1 class with cardinality $d = 4$;
- 1 class with cardinality $d = 2$.

In this case there is no class with cardinality $d = 5$ since $\frac{n}{d} \nmid t$.

Important question is how many classes with the maximum cardinality $d = n$ exist. Let for given n, t; $C(n, t, d)$ denotes the number of classes with the cardinality d. In Example 2, e.g., $C(20, 10, 4) = 1$, $C(20, 10, 5) = 0$. According to Theorem 1 and definition of ϱ_n we have

$$\binom{n}{t} = \sum_{d|n} dC(n, t, d). \tag{2}$$

By Theorem 1 we can exclude in this formula all summands d for which $C(n, t, d) = 0$.

$$\binom{n}{t} = \sum_{\substack{d|n \\ n/d|t}} dC(n, t, d). \tag{3}$$

There are 2 trivial cases

$$C(n, n, d) = C(n, 0, d) = \begin{cases} 1, & \text{if } d = 1; \\ 0, & \text{if } d > 1. \end{cases} \tag{4}$$

Let for given n, t; $D_{n,t}$ be the set of all d for which summands in Formula 3 are non-zero. Using the proof of Theorem 1 we can easily derive a formula for all non-trivial and non-zero $C(n, t, d)$:

$$C(n, t, d) = \frac{1}{d}\left(\binom{d}{\frac{td}{n}} - \sum_{\substack{k \in D_{n,t} \\ k|d, k<d}} kC(n, t, k)\right). \tag{5}$$

EXAMPLE 3. We apply formula (5) to our examples:
If $n = 10$, $t = 4$, then $D_{10,4} = \{5, 10\}$; $C(10, 4, 5) = \frac{1}{5}\binom{5}{2} = 2$ and
$C(10, 4, 10) = \frac{1}{10}\left(\binom{10}{4} - 5C(10, 4, 5)\right) = 20$.
If $n = 20, t = 10$, then $D_{20,10} = \{2, 4, 10, 20\}$; $C(20, 10, 2) = \frac{1}{2}\binom{2}{1} = 1$,
$C(20, 10, 4) = \frac{1}{4}\left(\binom{4}{2} - 2C(20, 10, 2)\right) = 1$,
$C(20, 10, 10) = \frac{1}{10}\left(\binom{10}{5} - 2C(20, 10, 2)\right) = 25$,
$C(20, 10, 20) = \frac{1}{20}\left(\binom{20}{10} - 2C(20, 10, 2) - 4C(20, 10, 4) - 10C(20, 10, 10)\right) = 9225$.

3. Conclusion

In this paper we studied equivalence classes of binary vectors with regards to their rotation. We used the theory of linear feedback shift registers, since the rotation of a binary vector can be modeled by a register with corresponding characteristic polynomial $f(x) = x^n + 1$. We stated necessary and sufficient condition for the existence of such classes with given cardinalities, and provided a formula that can be used to determine the structure of equivalence classes for binary vectors with given length and Hamming weight. One of the applications of our results are the quasi-cyclic codes used in the McEliece cryptosystems based on QC-LDPC codes, since we are able to determine the existence of square $n \times n$ binary circulant matrices with n distinct rows and the structure of binary circulant matrices, e.g., the number of distinct rows, depending on the length n and Hamming weight t of the first row.

REFERENCES

[1] BALDI, M.—CHIARALUCE, F.: *Cryptanalysis of a new instance of McEliece cryptosystem based on* QC-LDPC *codes,* in: Internat. Symposium on Information Theory—ISIT '07, Nice, France, 2007, IEEE, 2007, pp. 2591–2595.

[2] BALDI, M.—BODRATO, M.—CHIARALUCE, F.: *A new analysis of the McEliece cryptosystem based on* QC-LDPC *codes,* in: 6th Internat. Conf. on Security and Cryptography for Networks—SCN '08, Amalfi, Italy, 2008, (R. Ostrovsky et al., eds.), Lecture Notes in Comput. Sci., Vol. 5229, Springer-Verlag, Berlin, 2008, pp. 246–262.

[3] LIDL, R.—NIEDERREITER, H.: *Finite Fields.* Cambridge University Press, Cambridge, 2008.

[4] MCELIECE, R. J.: *A public-key cryptosystem based on algebraic coding theory,* DSN Progress Report, 1978, pp. 114–116.

[5] REPKA, M.—ZAJAC, P.: *Overview of the McEliece cryptosystem and its security,* Tatra Mt. Math. Publ. **60** (2014), 57–83.

[6] SINGHMASTER, D.: *Divisibility of binomial and multinomial coeficients by primes and prime powers,* in: A Collection of Manuscripts Related to the Fibonacci Sequence, 18th Anniversary Volume of the Fibonacci Association, 1980, pp. 98–113.

[7] ZAJAC, P.—ONDROŠ, M.: *Rotational cryptanalysis of* GOST *with identical* S*-boxes.* Tatra Mt. Math. Publ. **57** (2013), 1–19.

Received November 18, 2016

Institute of Computer Science and Mathematics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
Ilkovičova 3
SK–812-19 Bratislava
SLOVAKIA

E-mail: otokar.grosek@stuba.sk
viliam.hromada@stuba.sk