## SHORT CYCLE ROTATION BOOLEAN FUNCTIONS: GENERATORS AND COUNT

JOSÉ E. CALDERÓN-GÓMEZ, LUIS A. MEDINA, AND CARLOS MOLINA-SALAZAR

ABSTRACT. Rotation symmetric Boolean functions were introduced by Pieprzyk and Qu in late 1990's. These functions are useful, among other things, in the design of fast hashing algorithms with strong cryptographic properties. A monomial rotation Boolean function is called long cycle if the number of terms in its algebraic normal form coincides with the number of variables and short cycle if the number of terms is less than the number of variables. In this article, we characterize a family of generators of short cycles. We then use such family to count the number of short cycles.

#### 1. INTRODUCTION

Boolean functions are fascinating combinatorial objects with applications to scientific fields like coding theory, cryptography and information theory. An *n*-variable Boolean function is a map from  $\mathbb{F}_2^n \to \mathbb{F}_2$  where  $\mathbb{F}_2$  represents the field of two elements. The set of all *n*-variable Boolean functions is usually denoted by  $\mathcal{B}_n$ .

It is well-established that every Boolean function  $f \in \mathcal{B}_n$  can be identified with a multi-variable polynomial

(1.1) 
$$f(X_1,\ldots,X_n) = \bigoplus_{\boldsymbol{a}=(a_1,\ldots,a_n)\in\mathbb{F}_2^n} \lambda_{\boldsymbol{a}} \prod_{j=1}^n X_j^{a_j},$$

where  $\lambda_{a} \in \mathbb{F}_{2}$  for every  $a \in \mathbb{F}_{2}^{n}$  and  $\oplus$  represents addition mod 2. This polynomial is known as the *algebraic* normal form (or ANF for short) of the Boolean function f. The *algebraic degree* of  $f \in \mathcal{B}_{n}$  is the degree of its ANF.

The Hamming weight of a vector  $\mathbf{x} \in \mathbb{F}_2^n$ , usually denoted by  $wt(\mathbf{x})$ , is the number of its entries that are equal to 1. Let  $\mathbf{x}_0 = (0, 0, \dots, 0, 0), \mathbf{x}_1 = (0, 0, \dots, 0, 1), \mathbf{x}_2 = (0, 0, \dots, 1, 0), \dots, \mathbf{x}_{2^n-1} = (1, 1, \dots, 1, 1)$ . The truth table of  $f \in \mathcal{B}_n$  is the vector  $[f(\mathbf{x}_0), f(\mathbf{x}_1), \dots, f(\mathbf{x}_{2^n-1})]$ . The weight (or Hamming weight) of a Boolean function  $f \in \mathcal{B}_n$ , denoted by wt(f), is the number of 1 in its truth table.

A very important property in some cryptographic applications is balancedness. An *n*-variable Boolean function  $f \in \mathcal{B}_n$  is called *balanced* if the number of zeros and the number of ones in its truth table are the same. Balancedness of Boolean functions is usually studied via Hamming weights or via exponential sums.

The exponential sum of a Boolean function  $f \in \mathcal{B}_n$  is defined as

(1.2) 
$$S(f) = \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{f(\mathbf{x})}.$$

Observe that a Boolean function is balanced if and only if S(f) = 0. The Hamming weight of a Boolean function and its exponential sum are linked by the equation

(1.3) 
$$wt(f) = 2^{n-1} - \frac{1}{2}S(f).$$

For more comprehensive information about Boolean functions, please refer to [2, 15].

Balancedness of special families like symmetric and rotation symmetric Boolean functions has been the subject of several studies [3, 4, 5, 6, 10, 11, 12, 14, 16, 18, 27]. A Boolean function  $f \in \mathcal{B}_n$  is called symmetric if it is invariant under the action of the symmetric group  $S_n$  of n symbols, that is, if

$$f(X_{\sigma(1)},\ldots,X_{\sigma(n)}) = f(X_1,\ldots,X_n)$$

Date: September 6, 2022.

<sup>2020</sup> Mathematics Subject Classification. 05E05, 11T23.

Key words and phrases. rotation monomial Boolean functions, short cycles, recurrences.

for every  $\sigma \in S_n$ . On the other hand, a Boolean function  $f \in \mathcal{B}_n$  is called rotation symmetric if it is invariant under the action of the cyclic group  $C_n$  of n elements, that is, if

$$f\left(X_{\sigma(1)},\ldots,X_{\sigma(n)}\right) = f(X_1,\ldots,X_n)$$

for every  $\sigma \in C_n$ .

It is well-known that ANF of a symmetric Boolean function  $f \in \mathcal{B}_n$  has the form

(1.4) 
$$f = e_{n,k_1} \oplus e_{n,k_2} \oplus \cdots \oplus e_{n,k_s}$$

where  $0 \le k_1 < \cdots < k_s$  are integers and  $e_{n,k}$  represents the *n*-variable elementary symmetric polynomial of degree k. For simplicity, we often use the notation  $e_{n,[k_1,\ldots,k_s]}$  to represent the right-hand side of (1.4). It is known that if  $0 \le k_1 < \cdots < k_s$  are fixed integers, then the sequence  $\{S(e_{n,[k_1,\ldots,k_s]})\}_n$  satisfies a linear recurrence with constant coefficients [1, 3], in other words, it is a C-finite sequence. To be more specific, the sequence  $\{S(e_{n,[k_1,\ldots,k_s]})\}_n$  satisfies the recurrence whose characteristic polynomial is given by

(1.5) 
$$(X-2)\Phi_4(X-1)\Phi_8(X-1)\cdots\Phi_{2^r}(X-1),$$

where  $r = \lfloor \log_2(k_s) \rfloor + 1$  and  $\Phi_n(X)$  represents the *n*-th cyclotomic polynomial. The fact that  $\{S(e_{n,[k_1,\ldots,k_s]})\}_n$  is a *C*-finite sequence is very important, as it implies that the value of the exponential sum can be calculated efficiently provided some initial conditions. This result was extended to Walsh-Hadamard transformations of symmetric Boolean functions in [6] and to every finite field in [8].

Rotation symmetric Boolean functions were introduced by Pieprzyk and Qu [27] (although, they did appear before in the work of Filiol and Fontaine [17] as idempotents). They showed that these functions are useful, among other things, in the design of fast hashing algorithms with strong cryptographic properties. Let  $1 < j_1 < \cdots < j_s$  be integers. Rotation symmetric Boolean functions of the form

(1.6) 
$$R_{j_1,\dots,j_s}(n) = X_1 X_{j_1} \cdots X_{j_s} \oplus X_2 X_{j_1+1} \cdots X_{j_s+1} \oplus \cdots \oplus X_n X_{j_1-1} \cdots X_{j_s-1},$$

where none of the terms overlap or

(1.7) 
$$R_{j_1,\ldots,j_s}(n) = X_1 X_{j_1} \cdots X_{j_s} \oplus \cdots \oplus X_k X_{j_1+k} \cdots X_{j_s+k}$$

where  $X_{k+1}X_{j_1+k+1}\cdots X_{j_s+k+1}$  is the first term overlapping one of the previous terms; are called a monomial rotation symmetric Boolean function (the indices are taken modulo n and the complete system of residues mod n is  $\{1, 2, \ldots, n\}$ ). We say that  $R_{j_1,\ldots,j_s}(n)$  is *long cycle* if the period is n and *short cycle* if the period is less than n. The rotation

$$R_{2,3}(4) = X_1 X_2 X_3 \oplus X_2 X_3 X_4 \oplus X_3 X_4 X_1 \oplus X_4 X_1 X_2$$

is an example of a long cycle, while

$$R_3(4) = X_1 X_3 \oplus X_2 X_4$$

is an example of a short cycle. In [10], T. Cusick proved that, as in the case of symmetric Boolean functions, sequences of weights of rotation symmetric Boolean functions are C-finite. Cusick's result was later generalized to Walsh-Hadamard transformations [6] and to every finite field [7].

In this article, we study monomial rotation symmetric Boolean functions that are short cycle. We provide explicit generators for them and use these generators to count the number of short cycles given a fixed number of variables and a fixed degree.

Since we are working with Boolean functions that are fixed under the action of certain groups ( $C_n$  in our case), then it is a good idea to introduce some notations related to group actions. Suppose that you have group G acting on a set X. Let  $g \cdot x$  denote the action of  $g \in G$  on  $x \in X$ . The orbit of  $x \in X$  under the action of G will be denoted by  $\mathcal{O}(x)$ . In other words,

$$\mathcal{O}(x) = \{g \cdot x : g \in G\}.$$

The stabilizer of  $x \in X$  is denoted by  $\operatorname{Stab}(x)$ , that is,

$$Stab(x) = \{g \in G : g \cdot x = x\}.$$

It is not hard to see that the stabilizer of any element of X is in fact a subgroup of G. One of the classic results of the theory of group actions is called the Orbit-Stabilizer Theorem. It states that

$$|\mathcal{O}(x)| = [G : \operatorname{Stab}(x)]$$

In other words, the length of the orbit of an element of X is the group index of the stabilizer of such element. Thus, in the case of a finite group G, the length of the orbit of an element of X divides the order of the group.

### 2. Generators for short cycles

In the introduction, we define long cycle and short cycle rotation symmetric Boolean functions. To be more explicit, let

(2.1) 
$$\sigma_n = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \end{pmatrix}$$

and let  $\ell, m, n \in \mathbb{N}$  with  $\ell \leq n$  and  $m \leq n$ . If

(2.2) 
$$\sum_{i=1}^{m} X_{k_1+i} X_{k_2+i} \cdots X_{k_{\ell}+i}$$

is fixed by the rotation  $\sigma_n$  and m is the smallest number that satisfies this condition, then (2.2) is called

- (1) a short cycle of size m if m < n,
- (2) a long cycle if m = n.

We will abuse notation and use  $R_{k_1,k_2,\dots,k_\ell}(n)$  to represent (2.2). The monomial  $X_{k_1}X_{k_2}\cdots X_{k_\ell}$  is called a monomial generator of  $R_{k_1,k_2,\dots,k_\ell}(n)$ . Our goal in this section is to characterize short cycles.

Our first result is the next one. It is not new, but we provide a proof for completeness.

**Lemma 2.1.** Let  $\ell, n \in \mathbb{N}$  with  $\ell < n$ . If  $R_{k_1,k_2,\cdots,k_\ell}(n)$  is a short cycle of size m, then m|n.

*Proof.* Recall that rotation symmetric Boolean functions are those functions in  $\mathfrak{B}_n$  that are fixed under the action of the cyclic group  $C_n = \langle \sigma_n \rangle$ . The terms of the rotation generated by the monomial  $X_{k_1} X_{k_2} \cdots X_{k_\ell}$  are precisely the elements of  $\mathcal{O}(X_{k_1} X_{k_2} \cdots X_{k_\ell})$ . Therefore,

$$m = |\mathcal{O}(X_{k_1}X_{k_2}\cdots X_{k_\ell})| = [C_n : \operatorname{Stab}(X_{k_1}X_{k_2}\cdots X_{k_\ell})]$$

by the Orbit-Stabilizar Theorem. Thus m|n, as claimed.

We can be more precise about m. Let j the least positive integer such that

$$\sigma_n^j \cdot X_{k_1} X_{k_2} \cdots X_{k_\ell} = X_{k_1} X_{k_2} \cdots X_{k_\ell}.$$

Then, it is clear that  $\operatorname{Stab}(X_{k_1}X_{k_2}\cdots X_{k_\ell}) = \langle \sigma_n^j \rangle$  and therefore  $|\operatorname{Stab}(X_{k_1}X_{k_2}\cdots X_{k_\ell})| = n/j$ . That implies

$$m = |\mathcal{O}(X_{k_1}X_{k_2}\cdots X_{k_\ell})| = [C_n : \text{Stab}(X_{k_1}X_{k_2}\cdots X_{k_\ell})] = \frac{n}{n/j} = j.$$

Therefore, m is just the minimal positive integer such that  $\sigma_n^m$  generates  $\operatorname{Stab}(X_{k_1}X_{k_2}\cdots X_{k_\ell})$ .

A natural step now is to find a way to detect a short cycle rotation by analyzing its generators. The next result is about an explicit generator for a short cycle rotation of length m.

**Theorem 2.2.** Let  $\ell$  and n be natural numbers with  $\ell < n$ . If  $R_{r_1, \dots, r_\ell}(n)$  is short cycle of length m, then n/m divides  $\ell$  (the degree of  $R_{r_1, \dots, r_\ell}(n)$ ) and there are integers  $1 < k_2 < \dots < k_q \leq m$ , with  $q = m\ell/n$ , such that

(2.3) 
$$\prod_{j=0}^{\frac{n}{m}-1} X_{1+jm} X_{k_2+jm} \cdots X_{k_q+jm}$$

is a generator of  $R_{r_1,\ldots,r_\ell}(n)$ .

*Proof.* Suppose that  $R_{r_1,\ldots,r_\ell}(n)$  is a short cyle of length m. Then  $m \mid n$ . Let d = n/m. Since  $R_{r_1,\ldots,r_\ell}(n)$  is fixed by the rotation  $\sigma_n$ , then there is a term that has  $X_1$  in it. Suppose that  $X_1 X_{k_2} \cdots X_{k_\ell}$  is such term.

This term is a generator of  $R_{r_1,...,r_\ell}(n)$  with  $1 < k_2 < \cdots < k_\ell \leq n$ . Note that the terms of  $R_{r_1,...,r_\ell}(n)$  are given by

$$(2.4) X_1 X_{k_2} \cdots X_{k_{\ell-1}} X_{k_{\ell}} \\ X_2 X_{k_2+1} \cdots X_{k_{\ell-1}+1} X_{k_{\ell}+1} \\ X_3 X_{k_2+2} \cdots X_{k_{\ell-1}+2} X_{k_{\ell}+2} \\ \vdots \\ X_{m-1} X_{k_2+m-2} \cdots X_{k_{\ell-1}+m-2} X_{k_{\ell}+m-2} \\ X_m X_{k_2+m-1} \cdots X_{k_{\ell-1}+m-1} X_{k_{\ell}+m-1}$$

where it is understood that the indices of variables are to be taken mod n with reduced residue system  $\{1, 2, \ldots, n\}$ . These terms can be obtained from  $X_1 X_{k_2} \cdots X_{k_{\ell-1}} X_{k_\ell}$  by applying  $\sigma_n^j$ ,  $j = 0, 1, \cdots, m-1$ , to it. The next term, after applying  $\sigma_n^m$  to  $X_1 X_{k_2} \cdots X_{k_{\ell-1}} X_{k_\ell}$ , is

(2.5) 
$$X_{1+m}X_{k_2+m}\cdots X_{k_{\ell-1}+m}X_{k_{\ell}+m}.$$

However,  $R_{r_1,\ldots,r_\ell}(n)$  is a short cycle of length m (hypothesis), thus term (2.5) is equal to one of the terms on the list (2.4). In other words, there is  $i \in \{2, 3, \ldots, \ell\}$  such that

(2.6) 
$$k_{i} = 1 + m$$
$$k_{i+j} = k_{1+j} + m \text{ for } j \in \{1, 2, \dots, \ell - i\}$$
$$n + k_{j} = k_{\ell-i+1+j} + m \text{ for } j \in \{1, 2, \dots, i - 1\}$$

Therefore, the term (2.5) is given by

(2.7) 
$$X_{k_i}X_{k_{i+1}}\cdots X_{k_{\ell}}\underbrace{X_1X_{k_2}\cdots X_{k_{i-1}}}_{i-1 \text{ variables}}$$

and it is also a generator for  $R_{r_1,\ldots,r_\ell}(n)$ .

The terms of  $R_{r_1,\ldots,r_\ell}(n)$  can now be re-written as

Again, the next term, after applying  $\sigma_n^m$  to  $X_{k_i}X_{k_{i+1}}\cdots X_{k_\ell}X_1X_{k_2}\cdots X_{k_{i-1}}$  is (2.9)  $X_{k_i+m}X_{k_{i+1}+m}\cdots X_{k_\ell+m}\underbrace{X_{1+m}X_{k_2+m}\cdots X_{k_{i-1}+m}}_{i-1 \text{ variables}}$ 

and arguing as before (see (2.6)) implies that this term has the form

(2.10) 
$$X_{k_{2i-1}} \cdots X_{k_{\ell}} \underbrace{X_1 X_{k_2} \cdots X_{k_{i-1}}}_{i-1 \text{ variables}} \underbrace{X_{1+m} X_{k_2+m} \cdots X_{k_{i-1}+m}}_{i-1 \text{ variables}}$$

and it is also a generator for  $R_{r_1,\ldots,r_\ell}(n)$ .

Continue applying this argument to get that

(2.11)  $X_1 X_{k_2} \cdots X_{k_{i-1}} X_{1+m} X_{k_2+m} \cdots X_{k_{i-1}+m} \cdots X_{1+(d-1)m} X_{k_2+(d-1)m} \cdots X_{k_{i-1}+(d-1)m}$ is a generator of  $R_{r_1,r_2,\ldots,r_\ell}(n)$ . Observe that

(2.12) 
$$\ell = d(i-1) = \frac{n}{m}(i-1).$$

Thus, let  $q := i - 1 = m\ell/n$ . Since  $k_{q+1} = k_i = 1 + m$ , then  $1 < k_1 < \dots < k_q \le m$  and

(2.13) 
$$\prod_{j=0}^{\frac{m}{m}-1} X_{1+jm} X_{k_2+jm} \cdots X_{k_q+jm}$$

is a generator of  $R_{r_1,\ldots,r_\ell}(n)$ . This concludes the proof.

The following results are inmediate consequences of this theorem.

**Corollary 2.3.** Each generator of a short cycle of size m in n variables has the form

(2.14) 
$$\prod_{j=0}^{\frac{n}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm}$$

for some  $1 \le k_1 < k_2 < \cdots < k_q < k_1 + m$  where  $q = m\ell/n$  and it is understood that the indices of the variables are to be taken mod n with reduced residue system  $\{1, 2, \ldots, n\}$ .

**Corollary 2.4.** There are not short cycles in n variables and of degree  $\ell$  if and only if  $gcd(n, \ell) = 1$ . In particular, if  $gcd(n, \ell) > 1$ , then every divisor d of  $gcd(n, \ell)$  produces a short cycle of length n/d. The converse is also true, that is, if there is a short cycle in n variables of degree  $\ell$  of length m, then n/m divides  $gcd(n, \ell)$ .

Let us consider the expression

(2.15) 
$$\prod_{j=0}^{\frac{n}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm},$$

with the same conditions as in Corollary 2.3. Observe that, by taking the indices mod n, we have

$$(2.16) \qquad \sigma_n^m \cdot \prod_{j=0}^{\frac{m}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm} = \prod_{j=0}^{\frac{m}{m}-1} X_{k_1+(j+1)m} X_{k_2+(j+1)m} \cdots X_{k_q+(j+1)m}$$
$$= \prod_{j=1}^{\frac{m}{m}} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm}$$
$$= \prod_{j=1}^{\frac{m}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm}$$
$$\times X_{k_1+(n/m)m} X_{k_2+(n/m)m} \cdots X_{k_q+(n/m)m}$$
$$= \prod_{j=0}^{\frac{m}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm}.$$

Therefore,

$$\sigma_n^m \in \text{Stab}\left(\prod_{j=0}^{\frac{n}{m}-1} X_{k_1+jm} X_{k_2+jm} \cdots X_{k_q+jm}\right),\,$$

which implies that (2.15) produces a short cycle of length a divisor of m (regardless of the values of the  $k_i$ 's). **Example 2.5.** Consider the case when n = 18,  $\ell = 12$  and m = 6. In this case,  $q = m\ell/n = 4$ . Then, the monomial

(2.17) 
$$\prod_{j=0}^{2} X_{k_1+6j} X_{k_2+6j} X_{k_3+6j} X_{k_4+6j},$$

where  $1 \le k_1 < k_2 < k_3 < k_4 \le 6$ , generates a short cycle of length a divisor of 6. Indeed, if, for example,  $k_1 = 1, k_2 = 3, k_3 = 5$  and  $k_4 = 6$ , then the short cycle generated by

(2.18) 
$$\prod_{j=0}^{2} X_{k_1+6j} X_{k_2+6j} X_{k_3+6j} X_{k_4+6j} = X_1 X_3 X_5 X_6 X_7 X_9 X_{11} X_{12} X_{13} X_{15} X_{17} X_{18}$$

has length 6. Explicitly,

$$\begin{aligned} R_{1,3,5,6}(18) &= & X_1 X_3 X_5 X_6 X_7 X_9 X_{11} X_{12} X_{13} X_{15} X_{17} X_{18} \oplus \\ & X_1 X_2 X_4 X_6 X_7 X_8 X_{10} X_{12} X_{13} X_{14} X_{16} X_{18} \oplus \\ & X_1 X_2 X_3 X_5 X_7 X_8 X_9 X_{11} X_{13} X_{14} X_{15} X_{17} \oplus \\ & X_2 X_3 X_4 X_6 X_8 X_9 X_{10} X_{12} X_{14} X_{15} X_{16} X_{18} \oplus \\ & X_1 X_3 X_4 X_5 X_7 X_9 X_{10} X_{11} X_{13} X_{15} X_{16} X_{17} \oplus \\ & X_2 X_4 X_5 X_6 X_8 X_{10} X_{11} X_{12} X_{14} X_{16} X_{17} X_{18}. \end{aligned}$$

On the other hand, if  $k_1 = 1, k_2 = 3, k_3 = 4$  and  $k_4 = 6$ , then the short cycle generated by

(2.19) 
$$\prod_{j=0}^{2} X_{k_1+6j} X_{k_2+6j} X_{k_3+6j} X_{k_4+6j} = X_1 X_3 X_4 X_6 X_7 X_9 X_{10} X_{12} X_{13} X_{15} X_{16} X_{18}$$

has length 3. Explicitly,

$$R_{1,3,4,6}(18) = X_1 X_3 X_4 X_6 X_7 X_9 X_{10} X_{12} X_{13} X_{15} X_{16} X_{18} \oplus X_1 X_2 X_4 X_5 X_7 X_8 X_{10} X_{11} X_{13} X_{14} X_{16} X_{17} \oplus X_2 X_3 X_5 X_6 X_8 X_9 X_{11} X_{12} X_{14} X_{15} X_{17} X_{18}.$$

An explanation of why, in this case, the cycle length is 3 and not 6 is given by the fact that  $k_3 = k_1 + 3$  and  $k_4 = k_2 + 3$ . Thus,  $X_1 X_3 X_4 X_6 X_7 X_9 X_{10} X_{12} X_{13} X_{15} X_{16} X_{18}$  can be identified with

(2.20) 
$$X_1 X_3 X_4 X_6 X_7 X_9 X_{10} X_{12} X_{13} X_{15} X_{16} X_{18} = \prod_{j=0}^5 X_{1+3j} X_{3+3j}$$

and therefore, it must generate a short cycle of length a divisor of 3.

# 3. Count of short cycles

In this section we count the number of short cycles of length m in n variables of degree  $\ell$ . We start with the following definitions.

**Definition 3.1.** Suppose that  $\ell$  and n are positive integers with  $\ell < n$ . Let m be a divisor of n such that n/m divides  $\ell$ . We define the following:

(3.1) 
$$q_m(n,\ell) = \frac{m\ell}{n},$$
$$D_m(n,\ell) = \{1 < d < m : d | m \text{ and } q_d(n,\ell) \in \mathbb{N}\},$$
$$C_m(n,\ell) = \{R_{r_1,\dots,r_\ell}(n) : R_{r_1,\dots,r_\ell}(n) \text{ is a cycle of length } m\}$$

Observe that, in particular,  $C_n(n,\ell)$  is the set of all long cycles in n variables of degree  $\ell$ .

Theorem 2.2 can be used to provide a recursive definition for  $\#C_m(n, \ell)$  (if A is a set, #A represents its cardinality).

**Theorem 3.2.** Suppose that  $\ell$  and n are positive integers with  $\ell < n$ . Let m be a divisor of n such that n/m divides  $\ell$ . Then,

(3.2) 
$$\#\mathbf{C}_m(n,\ell) = \frac{1}{q_m(n,\ell)} \left[ \binom{m-1}{q_m(n,\ell)-1} - \sum_{d \in D_m(n,\ell)} q_d(n,\ell) \, \#\mathbf{C}_d(n,\ell) \right]$$

*Proof.* Recall that every cycle of length m has a generator of the form

(3.3) 
$$\prod_{j=0}^{\frac{m}{m}-1} X_{1+jm} X_{k_2+jm} \cdots X_{k_q+jm},$$

where  $q = q_m(n, \ell)$  and  $1 < k_2 < \cdots < k_q \le m$ . Observe that the most important part of this generator is (3.4)  $X_1 X_{k_2} \cdots X_{k_q}$  because the rest of it can be obtained from this piece. Thus, there are

$$\binom{m-1}{q_m(n,\ell)-1}$$

different generators of short cycles of type (3.3). It is not guaranteed that everyone of these generators generates a cycle of length m. In fact, some of these generators may generate cycles of length a divisor of m (such divisors must belong to  $D_m(n, \ell)$ ). Therefore, if we want to count how many cycles of length m we have in this setting, we must eliminate the cycles with length a divisor of m. Also, since every term of cycle  $R_{r_1,\ldots,r_\ell}(n)$  is a generator, then it is possible that some of the generators of type (3.3) generate the same cycle. So we must also take that into account.

Suppose that  $R_{r_1,...,r_{\ell}}(n)$  is a cycle of length m. Two generators of type (3.3) generate  $R_{r_1,...,r_{\ell}}(n)$  if and only if they appear as terms of it. Therefore, we must count how many terms of  $R_{r_1,...,r_{\ell}}(n)$  have  $X_1$  as one of its variables. This is actually not hard. In fact, the cycle  $R_{r_1,...,r_{\ell}}(n)$  has exactly  $q_m(n,\ell)$  terms that have  $X_1$  as one of its variables. Therefore, each cycle of length m has exactly  $q_m(n,\ell)$  generators of type (3.3).

We know that there are

$$\binom{m-1}{q_m(n,\ell)-1}$$

different generators of cycles of type (3.3). We want those that generate cycles of length m. Therefore, we must eliminate the ones that generate cycles of length a divisor of m. By the discussion above, for each  $d \in D_m(n, \ell)$ , the amount of generators of type (3.3) that generate a cycle of length d is given by the number of cycles of length d times  $q_d(n, \ell)$ , that is, it is given by  $q_d(n, \ell) \# C_d(n, \ell)$ . Therefore, there are

$$\binom{m-1}{q_m(n,\ell)-1} - \sum_{d \in D_m(n,\ell)} q_d(n,\ell) \# \mathcal{C}_d(n,\ell)$$

generators of type (3.3) that generate cycles of length m. Since each cycle of length m has exactly  $q_m(n, \ell)$  generators of type (3.3), then it is clear that

(3.5) 
$$\#\mathbf{C}_m(n,\ell) = \frac{1}{q_m(n,\ell)} \left[ \binom{m-1}{q_m(n,\ell)-1} - \sum_{d \in D_m(n,\ell)} q_d(n,\ell) \, \#\mathbf{C}_d(n,\ell) \right].$$

This concludes the proof.

Observe that the formula given by the previous theorem can be re-written as

(3.6) 
$$\# C_m(n,\ell) = \frac{1}{q_m(n,\ell)} \binom{m-1}{q_m(n,\ell)-1} - \sum_{d \in D_m(n,\ell)} \frac{q_d(n,\ell)}{q_m(n,\ell)} \# C_d(n,\ell)$$
$$= \frac{1}{m} \binom{m}{q_m(n,\ell)} - \sum_{d \in D_m(n,\ell)} \frac{d}{m} \# C_d(n,\ell).$$

**Corollary 3.3.** Let n and  $\ell$  be integers with  $\ell < n$ . Suppose that  $gcd(n, \ell) = 1$ . Then,

(3.7) 
$$\#C_n(n,\ell) = \frac{1}{\ell} \binom{n-1}{\ell-1}.$$

In particular, in the case that the number of variables is 2n + 1 and the degree is  $\ell = n + 1$ , the number of cycles of length 2n + 1 (long cycles) is given by the n-th Catalan number

(3.8) 
$$\#C_{2n+1}(2n+1,n+1) = \frac{1}{n+1} \binom{2n}{n}$$

*Proof.* Observe that since  $gcd(n, \ell) = 1$ , then  $D_n(n, \ell) = \emptyset$ . Clearly,  $q_n(n, \ell) = \ell$ . Therefore,

(3.9) 
$$\#C_n(n,\ell) = \frac{1}{\ell} \binom{n-1}{\ell-1}.$$

This concludes the proof.

Observe that

$$\frac{1}{\ell} \binom{n-1}{\ell-1} = \frac{1}{\ell} \frac{(n-1)!}{(\ell-1)!(n-\ell)!} = \frac{1}{n} \frac{n!}{\ell!(n-\ell)!} = \frac{1}{n} \binom{n}{\ell}$$

Therefore, Corollary 3.3 tells us that, when  $gcd(n, \ell) = 1$ ,

(3.10) 
$$\#\mathbf{C}_n(n,\ell) = \frac{1}{n} \binom{n}{\ell}.$$

This formula appears in [32, Th. 9]. Also, since  $gcd(n, \ell) = 1$ , this number counts the different necklaces using  $n - \ell$  black pebbles and  $\ell$  white pebbles (two necklaces are considered the same if one can be obtained from the other by a rotation). This was already pointed out in [32]. There are different combinatorial interpretations for  $\#C_{2n+1}(2n+1, n+1)$  (it is the *n*-th Catalan number).

**Corollary 3.4.** Let n and  $\ell$  be integers with  $\ell < n$ . Let m be a divisor of n such that n/m divides  $\ell$ . Then, for every positive integer a, we have

$$q_m(an, a\ell) = q_m(n, \ell)$$
 and  $D_m(an, a\ell) = D_m(n, \ell).$ 

In particular,

$$#\mathbf{C}_m(an, a\ell) = #\mathbf{C}_m(n, \ell).$$

**Example 3.5.** Consider n = 96,  $\ell = 12$  and m = 32. Theorem 3.2 implies that

$$\#C_{32}(96, 12) = 1120$$

Therefore,

$$\begin{aligned} \#C_{32}(192,24) &= 1120, \\ \#C_{32}(288,36) &= 1120, \end{aligned}$$

and, in general,

$$\#C_{32}(96a, 12a) = 1120,$$

for every  $a \in \mathbb{N}$ .

Theorem 3.2 can be used to provide bounds for  $\#C_m(n, \ell)$ . Recall that if n is a natural number, then  $\Gamma(n+1) = n!$  where

(3.11) 
$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, dx, \ \Re(z) > 0$$

is the gamma function. The gamma function can be extended to the whole complex plane except the non-positive integers in several ways. One of them is its definition as an infinite product due to Euler

(3.12) 
$$\Gamma(z) = \frac{1}{z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^{-1} \left(1 + \frac{1}{n}\right)^{z}$$

Another one is Weiestrass definition

(3.13) 
$$\Gamma(z) = \frac{e^{-\gamma z}}{z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^{-1} e^{z/n}$$

where  $\gamma$  is the Euler-Mascheroni constant

(3.14) 
$$\gamma = \lim_{n \to \infty} \left( \sum_{k=1}^{n} \frac{1}{k} - \log(n) \right) \approx 0.5772156649$$

The gamma function can be used to extend the binomial numbers to non-integer values via

(3.15) 
$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}.$$

**Corollary 3.6.** Suppose that  $\ell$  and n are positive integers with  $\ell < n$ . Let m be a divisor of n such that n/m divides  $\ell$ . Then,

(3.16) 
$$\frac{1}{q_m(n,\ell)} \left[ \binom{m-1}{q_m(n,\ell)-1} - \sum_{\substack{d \mid m \\ 1 < d < m}} \binom{d-1}{\frac{d\ell}{n}-1} \right] \le \# \mathcal{C}_m(n,\ell) \le \frac{1}{q_m(n,\ell)} \binom{m-1}{q_m(n,\ell)-1}.$$

*Proof.* The upper bound follows directly from (3.2). For the lower bound, observe that if  $d \in D_m(n, \ell)$ , then

(3.17) 
$$q_d(n,\ell) \# \mathcal{C}_d(n,\ell) \le \begin{pmatrix} d-1\\ \frac{d\ell}{n} - 1 \end{pmatrix}$$

Thus,

$$(3.18) \qquad \sum_{d \in D_m(n,\ell)} q_d(n,\ell) \# \mathcal{C}_d(n,\ell) \leq \sum_{\substack{d \mid m \\ 1 < d < m}} \frac{d\ell}{n} \# \mathcal{C}_d(n,\ell)$$
$$\leq \sum_{\substack{d \mid m \\ 1 < d < m}} \left(\frac{d-1}{\frac{d\ell}{n}-1}\right).$$

This concludes the proof.

**Example 3.7.** Let us study  $\#C_{12}(24, 12)$ , that is, the number of short cycles of length 12 in 24 variables and of degree 12. Observe that

$$q_{12}(24,12) = \frac{12 \times 12}{24} = 6$$

Corollary 3.6 tells us that the value of  $\#C_{12}(24, 12)$  lies between

$$\frac{1}{6} \left[ \binom{11}{5} - \binom{1}{0} - \binom{2}{1/2} - \binom{3}{1} - \binom{5}{2} \right] = \frac{1}{6} \left( 448 - \frac{16}{3\pi} \right) \approx 74.3837245456144$$

and

$$\frac{1}{6}\binom{11}{5} = 77.$$

Let us compute the exact value of  $\#C_{12}(24, 12)$ . Note that  $D_{12}(24, 12) = \{2, 4, 6\}$ . Observe that

$$q_{6}(24, 12) = \frac{6 \times 12}{24} = 3$$

$$q_{4}(24, 12) = \frac{4 \times 12}{24} = 2$$

$$q_{2}(24, 12) = \frac{2 \times 12}{24} = 1.$$

These numbers can be used to calculate the following three values:

$$\#C_{2}(24, 12) = {\binom{1}{0}} = 1 \#C_{4}(24, 12) = \frac{1}{2} \left[ {\binom{3}{1}} - \#C_{2}(24, 12) \right] = \frac{1}{2}(3-1) = 1 \#C_{6}(24, 12) = \frac{1}{3} \left[ {\binom{5}{2}} - \#C_{2}(24, 12) \right] = \frac{1}{3}(10-1) = 3.$$

Therefore,

$$\begin{aligned} \#C_{12}(24,12) &= \frac{1}{6} \left[ \binom{11}{5} - 3 \#C_6(24,12) - 2 \#C_4(24,12) - \#C_2(24,12) \right] \\ &= \frac{1}{6} [462 - 3 \times 3 - 2 \times 1 - 1] \\ &= 75, \end{aligned}$$

which lies between 74.3837245456144 and 77, as predicted.

Let us study now  $\#C_{24}(48, 24)$  i.e. the number of short cycles of length 24 in 48 variables and of degree 24. Since  $q_{24}(48, 24) = 12$ , then Corollary 3.6 tells us that the value of  $\#C_{24}(48, 24)$  lies between

$$\frac{1}{12} \left[ \binom{23}{11} - \binom{1}{0} - \binom{2}{1/2} - \binom{3}{1} - \binom{5}{2} - \binom{7}{3} - \binom{11}{5} \right] = \frac{1}{12} \left( 1351567 - \frac{16}{3\pi} \right) \\ \approx \quad 112630.441862$$

and

$$\frac{1}{12}\binom{23}{11} = \frac{676039}{6} \approx 112673.166667.$$

Let us compute the exact value of  $\#C_{24}(48, 24)$ . Observe that  $D_{24}(48, 24) = \{2, 4, 6, 8, 12\}$ . Since  $48 = 2 \times 24$  and  $24 = 2 \times 12$ , then Corollary 3.4 implies

$$\begin{array}{rcl} q_{12}(48,24) &=& q_{12}(24,12)=6\\ q_6(48,24) &=& q_6(24,12)=3\\ q_4(48,24) &=& q_4(24,12)=2\\ q_2(48,24) &=& q_2(24,12)=1 \end{array}$$

and

$$\begin{aligned} \#C_2(48,24) &= \#C_2(24,12) = 1 \\ \#C_4(48,24) &= \#C_4(24,12) = 1 \\ \#C_6(48,24) &= \#C_6(24,12) = 3 \\ \#C_{12}(48,24) &= \#C_{12}(24,12) = 75. \end{aligned}$$

These take care of most of the elements of  $D_{24}(48, 24)$ . We are still missing  $8 \in D_{24}(48, 24)$ . Note that  $q_8(48, 24) = 4$ , thus

$$\#C_8(48,24) = \frac{1}{4} \left[ \binom{7}{3} - 2 \#C_4(48,24) - \#C_2(48,24) \right]$$
$$= \frac{1}{4} [35 - 2 \times 1 - 1]$$
$$= 8.$$

Finally,

$$\#C_{24}(48,24) = \frac{1}{12} \left[ \binom{23}{11} - 6 \times 75 - 4 \times 8 - 3 \times 3 - 2 \times 2 - 1 \right]$$
$$= \frac{1}{12} [1352078 - 494]$$
$$= 112632,$$

which lies between 112630.441862 and 112673.166667, as predicted.

Observe that the cumbersome part of the recursive formula given by Theorem 3.2 for  $\#C_m(n,\ell)$  is to control the divisors in  $D_m(n,\ell)$ . The previous example illustrated that. Of course, there are some instances, like the next one, on which we know every possible divisor and can provide an explicit formula for  $\#C_m(n,\ell)$ .

**Proposition 3.8.** Let p be a prime. Let  $\beta, \alpha$ , s and t be positive integers such that  $sp^{\alpha} < p^{\beta}$ ,  $0 \le t < \alpha$  and p does not divide s. Then,

(3.19) 
$$\# C_{p^{\beta-\alpha+t}}(p^{\beta}, sp^{\alpha}) = \frac{1}{sp^{t}} \left[ \binom{p^{\beta-\alpha+t}-1}{sp^{t}-1} - \binom{p^{\beta-\alpha+t-1}-1}{sp^{t-1}-1} \right].$$

The next result is similar in the sense that we have control over the divisors of m. Incidentally, it also shows that the upper bound in Corollary 3.6 can be attained.

**Proposition 3.9.** Let n,  $\ell$  and m be positive integers such that  $\ell < n$  and n/m divides  $\ell$ . Suppose that

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}, 
\ell = a p_r^{\alpha_r}, 
m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_{r-1}^{\alpha_{r-1}}$$

where  $p_j$ , j = 1, ..., r are different primes,  $\alpha_j$ , j = 1, ..., r are non-negative integers and a is a positive integer satisfying  $gcd(a, p_j) = 1$  for all j. Then,

(3.20) 
$$\#\mathbf{C}_m(n,\ell) = \frac{1}{q_m(n,\ell)} \binom{m-1}{q_m(n,\ell)-1} = \frac{1}{a} \binom{m-1}{a-1}$$

*Proof.* Observe that

(3.21) 
$$q_m(n,\ell) = \frac{m\ell}{n} = \frac{ap_1^{\alpha_1}p_2^{\alpha_2}\cdots p_{r-1}^{\alpha_{r-1}}p_r^{\alpha_r}}{p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_{r-1}^{\alpha_{r-1}}} = a$$

It is clear that  $D_m(n, \ell) = \emptyset$  in this case. Theorem 3.2 implies

(3.22) 
$$\#\mathbf{C}_m(n,\ell) = \frac{1}{q_m(n,\ell)} \binom{m-1}{q_m(n,\ell)-1} = \frac{1}{a} \binom{m-1}{a-1}$$

This concludes the proof.

Observe that the hyptheses of Proposition 3.9 implies gcd(m, a) = 1. Also,

(3.23) 
$$\frac{1}{a}\binom{m-1}{a-1} = \frac{1}{m}\binom{m}{a}.$$

Thus, if

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$
  

$$\ell = a p_r^{\alpha_r},$$
  

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_{r-1}^{\alpha_{r-1}}$$

where  $p_j$ , j = 1, ..., r are different primes,  $\alpha_j$ , j = 1, ..., r are non-negative integers and a is a positive integer satisfying  $gcd(a, p_j) = 1$  for all j, then  $\#C_m(n, \ell)$  counts the different necklaces using m - a black pebbles and a white pebbles.

## 4. Concluding Remarks

We characterized a family of generators for short cycles. We used these generators to provide a recursive formula for  $\#C_m(n,\ell)$  i.e. for the number of short cycles in n variables of degree  $\ell$  that have length m, where m is a divisor of n such that n/m divides  $\ell$ . We also provided bounds  $\#C_m(n,\ell)$ . We hope and expect to see applications of our results.

Acknowledgments. The research of the second author was supported by UPR-FIPI 7240022.00 and by The Puerto Rico Science, Technology and Research Trust (PRST) under agreement number 2020-00124. This content is only the responsibility of the authors and does not necessarily represent the official views of The Puerto Rico Science, Technology and Research Trust. The first author was supported as a student was supported as a student by PRST 2020-00124. The third author was supported as a student by UPR-FIPI 7240022.00.

#### References

- [1] J. Cai, F. Green, and T. Thierauf. On the correlation of symmetric functions. Math. Systems Theory 29 (1996) 245-258.
- [2] C. Carlet. Boolean functions for cryptography and error correcting codes, in: Boolean Methods and Models. Cambridge University Press, Cambridge (2010),257–397.
- [3] F. Castro and L. A. Medina. Linear Recurrences and Asymptotic Behavior of Exponential Sums of Symmetric Boolean Functions. Elec. J. Combin. 18 (2011), #P8.
- [4] F. Castro and L. A. Medina. Asymptotic Behavior of Perturbations of Symmetric Functions. Annals of Combin. 18 (2014) 397–417.
- [5] F. Castro, O. E. González and L. A. Medina. Diophantine equations with binomial coefficients and perturbations of symmetric Boolean functions. *IEEE Trans. Inform. Theory* 64(2) (2018) 1347–1360.
- [6] F. Castro, L.A. Medina, P. P. Stănică. Generalized Walsh transforms of symmetric and rotation symmetric Boolean functions are linear recurrent. Applicable Algebra in Engineering, Communication and Computing (2018) 1–21.
- [7] F. Castro, R. Chapman, L.A. Medina, L. B. Sepúlveda Recursions associated to traprezoid, symmetric and rotation symmetric functions over Galois fields. *Discrete Mathematics* 341(2018)1915–1931.
- [8] F. Castro, L.A. Medina, L. B. Sepúlveda Closed formulas for exponential sums of symmetric polynomials over Galois fields. J. Algebr. Comb. (2018) DOI 10.1007/s10801-018-0840-4.
- [9] K. Conrad. Roots on a circle. Expository note available at http://www.math.uconn.edu/ kconrad/blurbs/.

- [10] T. W. Cusick. Weight recursions for any rotation symmetric Boolean functions. IEEE Trans. Inf. Theory 64(4) (2018) 2962–2968.
- [11] T. W. Cusick and B. Johns. Recursion orders for weights of Boolean cubic rotation symmetric functions. Discr. Appl. Math. 186 (2015) 1–6.
- [12] T. W. Cusick and Y. Li. k-th order symmetric SAC Boolean functions and bisecting binomial coefficients. Discrete Appl. Math. 149 (2005) 73–86.
- [13] T. W. Cusick, Y. Li, and P. Stănică. Balanced Symmetric Functions over GF(p). IEEE Trans. Inf. Theory 5 (2008) 1304–1307.
- [14] T. W. Cusick and P. Stănică. Fast evaluation, weights and nonlinearity of rotation symmetric functions. Discr. Math. 258 (2002) 289–301.
- [15] T. W. Cusick, P. Stănică. Cryptographic Boolean Functions and Applications Academic Press, (Ed. 2), San Diego, CA, 2017.
- [16] D. K. Dalai, S. Maitra, and S. Sarkar. Results on rotation symmetric bent functions. Second International Workshop on Boolean Functions: Cryptography and Applications, BFCA'06, publications of the universities of Rouen and Havre (2006) 137–156.
- [17] E. Filiol, C. Fontaine. Highly nonlinear balanced Boolean functions with a good correlation immunity. In: Eurocrypt 1998, LNCS 1403, Springer, Berlin, 1998, pp. 475–488.
- [18] M. Hell, A. Maximov, and S. Maitra. On efficient implementation of search strategy for rotation symmetric Boolean functions. Ninth International Workshop on Algebraic and Combinatorial Coding Theory, ACCT 2004, Black Sea Coast, Bulgaria, 2004.
- [19] S. Hodžić and E. Pasalic. Generalized bent functions: Some general construction methods and related necessary and sufficient conditions. *Cryptograph. Commun.*, 74 (2015) 469–483.
- [20] P. V. Kumar, R. A. Scholtz and L. R. Welch. Generalized bent functions and their properties. J. Combin Theory -Ser. A 40 (1985) 90–107.
- [21] Y. Li and T.W. Cusick. Linear Structures of Symmetric Functions over Finite Fields. Inf. Processing Letters 97 (2006) 124–127.
- [22] Y. Li and T. W. Cusick. Strict Avalanche Criterion Over Finite Fields. J. Math. Cryptology 1(1) (2007) 65-78.
- [23] M. Liu, P. Lu and G.L. Mullen. Correlation-Immune Functions over Finite Fields. IEEE Trans. Inf. Theory 44 (1998), 1273–1276.
- [24] T. Martinsen, W. Meidl, S. Mesnager, P. Stănică. Decomposing generalized bent and hyperbent functions. IEEE Trans. Inf. Theory 63 (2017) 7804–7812.
- [25] L. A. Medina, L. B. Sepúlveda and C. Serna-Rapello. Value distribution of elementary symmetric polynomials and its perturbations over finite fields. *Finite Fields Th. App.* 63 (2020) 1–21.
- [26] M. G. Parker and A. Pott. On Boolean functions which are bent and negabent. Proc. Int. Conf. Sequences, Subsequences, Consequences, LNCS-4893 (2007) 9–23.
- [27] J. Pieprzyk and C.X. Qu. Fast hashing and rotation-symmetric functions. J. Universal Comput. Sci. 5(1) (1999) 20-31.
- [28] C. Riera, M. G. Parker. Generalized bent criteria for Boolean functions. IEEE Trans. Inf. Theory 52(9) (2006), 4142–4159.
- [29] O. S. Rothaus. On bent functions. J. Combin. Theory Ser. A 20 (1976) 300–305.
- [30] P. Stănică. Weak and strong 2<sup>k</sup>-bent functions. *IEEE Trans. Inform. Theory* **62** 2827–2835 (2016)
- [31] P. Stănică, S. Gangopadhyay, B. K. Singh. "Some Results Concerning Generalized Bent Functions.". Available at http:// eprint.iacr.org/2011/290.pdf
- [32] P. Stănică, S. Maitra. Rotation symmetric Boolean functions—Count and cryptographic properties. Discr. Appl. Math. 156(10) (2008) 1567–1580.
- [33] P. Stănică, T. Martinsen, S. Gangopadhyay, B. K. Singh. Bent and generalized bent Boolean functions. Des. Codes Cryptogr., 69(2013) 77–94.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PUERTO RICO, SAN JUAN, PR 00925 Email address: joseemilio.calderon@upr.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PUERTO RICO, SAN JUAN, PR 00925 *Email address*: luis.medina17@upr.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PUERTO RICO, SAN JUAN, PR 00925 Email address: carlos.molina2@upr.edu