Cryptography and Communications
https://doi.org/10.1007/s12095-022-00562-x

A number theoretic view on binary shift registers

George Petrides[1]

Received: 28 July 2021 / Accepted: 31 January 2022

Abstract
We describe a number theoretic view on binary shift registers. We illustrate this approach on some basic shift registers by revisiting known and obtaining new results, which we prove using tools from basic number theory, including modular arithmetic.

Keywords Shift Registers · PCR · CCR · PSR · CSR · Adjacency Graph

Mathematics Subject Classification (2010) 94A55 · 11A07

1 Introduction

The cycles produced by binary shift registers, as well as the representative member used to denote them, are for reasons of brevity sometimes given in decimal notation instead of as binary sequences, see for example [3, 6, 12]. This simple observation piqued our interest to investigate an alternative view on the theory of shift registers where we move away from the traditional approach of binary sequences and work entirely with the corresponding integers.

After providing basic information on the necessary theory using the new approach in Section 2, we apply it on some basic registers to study their cycle structure (Section 3) and adjacency graphs (Section 4) using tools from basic number theory, including modular arithmetic. Apart from re-obtaining known results, our contributions include simple criteria for determining cycle lengths, the enumeration of cycles of fixed length, and a new connection between the pure and complemented cycling and summing registers.

Our main aim is to provide a unified description of this number theoretic approach. Its general usefulness as a tool for studying shift registers beyond the examples given here remains to be seen.

This work was supported by The Research Council of Norway under project 247742/O70. An extended abstract version was presented at *Sequences and Their Applications* 2020.

✉ George Petrides
g.petrides@yahoo.com

[1] University of Bergen, Bergen, Norway

2 From binary sequences to modular arithmetic

Any non-singular binary feedback shift register of order n can be defined in terms of a bijective map $g : \mathbb{F}_2^n \to \mathbb{F}_2^n$ given by

$$g(s_0, \ldots, s_{n-1}) = (s_1, \ldots, s_{n-1}, s_0 \oplus F(s_1, \ldots, s_{n-1})), \tag{1}$$

for some Boolean function $F : \mathbb{F}_2^{n-1} \to \mathbb{F}_2$ [4].

We can shift from binary tuples to modular arithmetic by considering each k-tuple $(s_0, \ldots, s_{k-1}) \in \mathbb{F}_2^k$ as the binary representation of the integer $\sum_{i=0}^{k-1} s_i 2^{k-1-i} \in \mathbb{Z}_{2^k}$. The corresponding functions will be $F : \mathbb{Z}_{2^{n-1}} \to \mathbb{Z}_2$ and $g : \mathbb{Z}_{2^n} \to \mathbb{Z}_{2^n}$ given by

$$g(x) = \begin{cases} 2x + F(x) \mod 2^n & \text{if } x < 2^{n-1} \\ 2x + 1 - F(x - 2^{n-1}) \mod 2^n & \text{if } x \geq 2^{n-1}. \end{cases}$$

Rewriting this in terms of the *support* of F, namely the set $\mathcal{D} \subseteq \mathbb{Z}_{2^{n-1}}$ such that $x \in \mathcal{D}$ if and only if $F(x) = 1$, we obtain

$$g_{n,\mathcal{D}}(x) = \begin{cases} 2x + 1 \mod 2^n & \text{if } x \in \mathcal{D}, \text{ or } x \geq 2^{n-1} \text{ and } x - 2^{n-1} \notin \mathcal{D} \\ 2x \mod 2^n & \text{otherwise.} \end{cases} \tag{2}$$

We can also define the *complemetary* map of $g_{n,\mathcal{D}}$ as $\overline{g}_{n,\mathcal{D}} = g_{n,\mathbb{Z}_{2^{n-1}} \setminus \mathcal{D}}$.

Example 1 Two basic, yet important maps are $g_{n,\emptyset}$, called the *Pure Cycling Register* of order n (PCR_n), and its complementary map $g_{n,\mathbb{Z}_{2^{n-1}}}$, called the *Complemented Cycling Register* of order n (CCR_n). For brevity we will be respectively denoting them by g_{p_n} and g_{c_n}. They are given by

$$g_{p_n}(x) = \begin{cases} 2x \mod 2^n & \text{if } x < 2^{n-1} \\ 2x + 1 \mod 2^n & \text{if } x \geq 2^{n-1} \end{cases} = \overline{g}_{c_n}(x) \tag{3}$$

and

$$g_{c_n}(x) = \begin{cases} 2x + 1 \mod 2^n & \text{if } x < 2^{n-1} \\ 2x \mod 2^n & \text{if } x \geq 2^{n-1} \end{cases} = \overline{g}_{p_n}(x)\,. \tag{4}$$

Cycle structure For each $x \in \mathbb{Z}_{2^n}$, the smallest $i \in \mathbb{Z}$ such that $x = g_{n,\mathcal{D}}^i(x)$ is called its *period with respect to* $g_{n,\mathcal{D}}$ and denoted by $p_{g_{n,\mathcal{D}}}(x)$, where $g_{n,\mathcal{D}}^i$ denotes the composition of $g_{n,\mathcal{D}}$ with itself i times. Each map $g_{n,\mathcal{D}}$ partitions $\mathbb{Z}_{2^n}$ into *cycles*. We say $x_1, x_2 \in \mathbb{Z}_{2^n}$ belong to the same cycle if and only if $x_2 = g_{n,\mathcal{D}}^i(x_1)$ for some i such that $1 \leq i < p_{g_{n,\mathcal{D}}}(x_1)$. We shall denote each cycle by C_t where t is its smallest member. The number of elements in a cycle is called its *length* and equals their period. In case there is a single cycle we call it a *maximal length* or *full* or *de Bruijn* cycle. Mykkeltveit [9] proved the conjecture of Golomb [4] that no more than $Z(n) = \frac{1}{n} \sum_{d|n} \phi(d) 2^{n/d}$ cycles can be obtained from any map $g_{n,\mathcal{D}}$, where ϕ is Euler's Totient function.

Example 2 The 8 cycles from PCR_5 are $C_0 = \{0\}$, $C_1 = \{1,2,4,8,16\}$, $C_3 = \{3,6,12,24,17\}$, $C_5 = \{5,10,20,9,18\}$, $C_7 = \{7,14,28,25,19\}$, $C_{11} = \{11,22,13,26,21\}$, $C_{15} = \{15,30,29,27,23\}$,

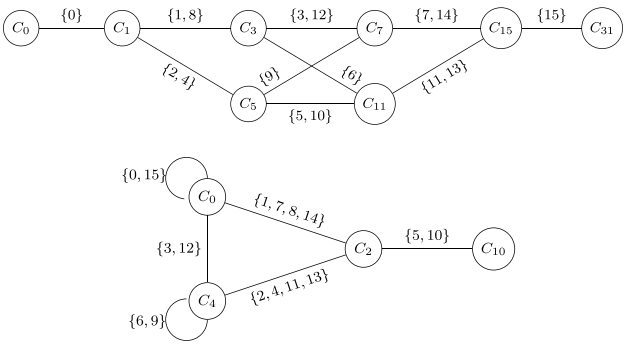

Fig. 1 The adjacency graph of PCR_5 (above) and CCR_5 (below)

and $C_{31} = \{31\}$. The 4 cycles from CCR_5 are $C_0 = \{0,1,3,7,15,31,30,28,24,16\}$, $C_2 = \{2,5,11,23,14,29,26,20,8,17\}$, $C_4 = \{4,9,19,6,13,27,22,12,25,18\}$ and $C_{10} = \{10,21\}$.

Complements and weights For $x \in \mathbb{Z}_{2^n}$, its *complement* is $\overline{x} = 2^n - 1 - x$ and its *weight*, denoted by $\mathrm{wt}(x)$, is the number of ones in its binary representation. By (1) we can deduce that $\mathrm{wt}(x) - 1 \leq \mathrm{wt}(g_{n,\mathcal{D}}(x)) \leq \mathrm{wt}(x) + 1$, and clearly

$$\mathrm{wt}(\overline{x}) = n - \mathrm{wt}(x). \tag{5}$$

Example 3 The function F in (1) for PCR_n is 0 and for CCR_n 1 for all inputs. Thus, for any $x \in \mathbb{Z}_{2^n}$, PCR_n simply cyclically shifts the binary representation of x and therefore

$$\mathrm{wt}(g_{p_n}(x)) = \mathrm{wt}(x). \tag{6}$$

CCR_n also complements the last bit after shifting, thus weights differ by one:

$$\mathrm{wt}(g_{c_n}(x)) = \begin{cases} \mathrm{wt}(x) + 1 & \text{if } x < 2^{n-1} \\ \mathrm{wt}(x) - 1 & \text{if } x \geq 2^{n-1}. \end{cases} \tag{7}$$

If the complements of a cycle's elements are all on one cycle of the same length, we say the two cycles are the *complement* of each other. A *self-dual* cycle is one that is its own complement, and thus has to have even length.

Example 4 In PCR_5, the complements of C_0, C_1, C_3 and C_5 are respectively C_{31}, C_{15}, C_7 and C_{11}. Every CCR_5 cycle is self-dual.

Adjacency graphs The adjacency graph of a map $g_{n,\mathcal{D}}$ is the undirected connected graph with vertices the map's cycles, and for each $x \in \mathbb{Z}_{2^{n-1}}$ an edge labelled x between the cycle containing x and the cycle containing $\overline{g}_{n,\mathcal{D}}(x)$. An edge from a cycle to itself is called *intracyclic*, and *extracyclic* otherwise. For brevity, we represent multiple edges between two cycles by a single edge labelled by the set of the corresponding labels that we call the *adjacency set*.

Example 5 The adjacency graphs of PCR_5 and CCR_5 are given in Fig. 1.

The study of adjacency sets provides guidelines for joining and splitting cycles from a map $g_{n,\mathcal{D}}$. Adding an element of an adjacency set to $\mathcal{D}$ if it does not exist, or removing it if it does, affects the cycles connected by the edge it labels. If the edge is extracyclic then the two cycles sharing it merge into a single cycle, otherwise the corresponding cycle splits into two cycles. By joining all cycles, we obtain a de Bruijn cycle.

Two distinct $x_1, x_2 \in \mathbb{Z}_{2^{n-1}}$ belong to the same adjacency set in the graph of a map $g_{n,\mathcal{D}}$ if either

A. x_1 and x_2 belong to the same cycle and $\overline{g}_{n,\mathcal{D}}(x_1)$ and $\overline{g}_{n,\mathcal{D}}(x_2)$ belong to the same cycle, in which case we shall call x_1 and x_2 an *intracyclic pair*, or
B. x_1 and $\overline{g}_{n,\mathcal{D}}(x_2)$ belong to the same cycle and x_2 and $\overline{g}_{n,\mathcal{D}}(x_1)$ belong to the same cycle, in which case we shall cal x_1 and x_2 an *extracyclic pair*.

Example 6 An intracyclic pair in the graph of PCR_5 are 2 and 4 which are on C_1 while $g_{c_n}(2) = 5$ and $g_{c_n}(4) = 9$ are on C_5. An extracyclic pair in the graph of CCR_5 are 5 and 10 since 5 and $g_{p_n}(10) = 20$ are on C_2, and $g_{p_n}(5) = 10$.

The two conditions for intracyclic pairs can be expressed formally as (a_1) $x_2 = g^i_{n,\mathcal{D}}(x_1)$, and (a_2) $\overline{g}_{n,\mathcal{D}}(x_1) = g^j_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(x_2))$, for some i, j such that $1 \le i < p_1$ and $1 \le j < p_2$, where $p_1 = p_{g_{n,\mathcal{D}}}(x_1)$ and $p_2 = p_{g_{n,\mathcal{D}}}(\overline{g}_{n,\mathcal{D}}(x_1))$. Together they imply

$$\overline{g}_{n,\mathcal{D}}(x_1) = g^j_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(g^i_{n,\mathcal{D}}(x_1))). \tag{8}$$

Remark 1 Conditions (a_1) and (a_2) are equivalent to $x_1 = g^{p_1-i}_{n,\mathcal{D}}(x_2)$ and $\overline{g}_{n,\mathcal{D}}(x_2) = g^{p_2-j}_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(x_1))$ respectively. Hence, if one member of an intracyclic pair satisfies (8) with the pair of exponents (i, j), the other member satisfies it with the pair of exponents $(p_1 - i, p_2 - j)$.

Similarly, the two extracyclic pair conditions can be expressed as (b_1) $x_1 = g^i_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(x_2))$ and (b_2) $x_2 = g^j_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(x_1))$, for some i, j such that $1 \le i < p_1$ and $1 \le j < p_2$, where p_1 and p_2 are as above. Together they imply

$$x_1 = g^i_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(g^j_{n,\mathcal{D}}(\overline{g}_{n,\mathcal{D}}(x_1)))).$$

3 On the cycle structure of some basic shift registers

3.1 The pure cycling register

PCR_n was defined in Example 1. Since $g_{p_n}(2^n - 1) = 2^n - 1$, the PCR_n cycle C_{2^n-1} is of length 1. For any $x \in \mathbb{Z}_{2^n} \setminus \{2^n - 1\}$, (3) can be expressed as

$$g_{p_n}(x) = 2x \mod 2^n - 1. \tag{9}$$

The length of the cycle containing $x \in \mathbb{Z}_{2^n} \setminus \{2^n - 1\}$ is equal to the period $p_{g_{p_n}}(x)$, the smallest positive exponent i such that $2^i x \equiv x \mod 2^n - 1$, or equivalently

$$2^i \equiv 1 \mod \frac{2^n - 1}{\gcd(x, 2^n - 1)}, \tag{10}$$

the *multiplicative order* of 2 modulo this ratio. It follows that when x is coprime to $2^n - 1$, its period is equal to n, the maximum possible. Therefore, there are at least $\frac{\phi(2^n-1)}{n}$ cycles of length n.

Proposition 1 *The length of any PCR_n cycle divides n.*

Proof Suppose a PCR_n cycle has length k not dividing n, in which case $n = ak + b$ for positive integers a and $b < k$. For every element x in the cycle we have $g^n_{p_n}(x) = 2^n x \equiv x \mod 2^n - 1$, where $2^n x = 2^{ak+b} x = 2^b 2^{ak} x \equiv 2^b x \mod 2^n - 1$ since $2^k x \equiv x \mod 2^n - 1$. Thus $2^b x \equiv x \mod 2^n - 1$, a contradiction on the minimality of k. □

Let $\zeta_n(k)$ denote the number of PCR_n cycles of length k. This number is in fact equal to the number of binary Lyndon words and irreducible polynomials of degree k over $\mathbb{Z}_2$ [2, 13] restricted to divisors of n:

$$\zeta_n(k) = \begin{cases} \frac{1}{k} \sum_{d|k} \mu(d) 2^{\frac{k}{d}} & \text{if } k \mid n \\ 0 & \text{otherwise,} \end{cases}$$

where μ is the Möbius function.

Golomb [4] proved that PCR_n partitions $\mathbb{Z}_{2^n}$ into exactly $Z(n)$ cycles. Summing $\zeta_n(k)$ over all divisors of n, an alternative formula can be obtained.

Corollary 1 [13] *The number of PCR_n cycles is*

$$Z(n) = \sum_{d|n} \frac{1}{d} \sum_{d'|d} \mu(d') 2^{\frac{d}{d'}}.$$

Remark 2 This viewpoint reveals that the PCR_n cycles (excluding C_{2^n-1}) are in fact the same as the cyclotomic classes of 2 modulo $2^n - 1$ defined by Carlet et al. [1] in a different context. These authors also showed that the cardinality of these classes is given by (10) and that their number is $\sum_{d|2^n-1} \phi(d)/p_{g_{p_n}}(d)$.

It is easy to check that $g_{p_n}(\overline{x}) = \overline{g_{p_n}(x)}$, hence every PCR_n cycle has a complement (that $p_{g_{p_n}}(x) = p_{g_{p_n}}(\overline{x})$ also follows from (10)). By (6), all elements in a cycle have the same weight, henceforth the *weight of the cycle*. These two observations together with (5) imply that for each cycle of weight w there exists a distinct cycle of weight $n - w$ of the same length, with the exception of self-dual cycles, which might exist if n is even and $w = n/2$. Since such a (self-dual) cycle contains both x and $\overline{x}$, we have $2^i x \equiv -x \mod 2^n - 1$ for some positive exponent i. Simplifying and combining with (10) gives

$$2^{\frac{p_{g_{p_n}}(x)}{2}} \equiv -1 \mod \frac{2^n-1}{\gcd(x, 2^n-1)}.$$

This condition cannot hold if x is coprime to 2^n-1 as $2^{n/2}+1 \not\equiv 0 \mod 2^n-1$.

Let $Z^O(n)$ and $Z^E(n)$ respectively denote the number of PCR_n cycles of odd and even weight, and $\zeta_n^O(k)$ and $\zeta_n^E(k)$ those of odd and even weight and length k. From the above we can deduce that for odd n, complementary cycles have different weight parity, thus $\zeta_n^O(k) = \zeta_n^E(k) = \zeta_n(k)/2$ and $Z^O(n) = Z^E(n) = Z(n)/2$. We will complete the picture by determining these numbers also for even n, in which case complementary cycles have the same weight parity. Knowing $\zeta_n(k)$ and $Z(n)$, it in fact suffices to obtain $\zeta_n^O(k)$ and $Z^O(n)$.

Proposition 2 *For any $x \in \mathbb{Z}_n$, its weight* $\mathrm{wt}(x)$ *is divisible by* $n/p_{g_{p_n}}(x)$.

Proof Consider $x \in \mathbb{Z}_n$ of period $p_{g_{p_n}}(x) = k$. We have $2^k x \equiv x \mod 2^n-1$ which means 2^n-1 divides $(2^k-1)x$, and since k divides n, $(2^n-1)/(2^k-1) = \sum_{i=1}^{n/k} 2^{n-ki}$ divides x. In other words, $x = a\sum_{i=1}^{n/k} 2^{n-ki}$ for some integer $0 \le a \le 2^k-1$. Letting $a = \sum_{j=0}^{k-1} a_j 2^j$, where the $a_j \in \{0,1\}$, we get $x = \sum_{i=1}^{n/k}\sum_{j=0}^{k-1} a_j 2^{n-ki+j}$. The weight of x would be a multiple of n/k provided the powers of 2 in this sum are all distinct. This is indeed the case, as otherwise we would have $2^{n-ki+j} = 2^{n-ki'+j'}$ for distinct $1 \le i, i' \le n/k$ and $0 \le j, j' \le k-1$, implying $k(i'-i) = j'-j$, a contradiction since $j-j' \le k-1$. □

Corollary 2 *The weight of any PCR_n cycle of length k is divisible by n/k.*

Lemma 1 *For positive integers n and k,*

$$\zeta_n^O(k) = \begin{cases} \frac{1}{k}\sum\limits_{d \mid \frac{k}{2^{\nu_2(n)}}} \mu(d) 2^{\frac{k}{d}-1} & \text{if } k \mid n \text{ and } \nu_2(n) = \nu_2(k) \\ 0 & \text{otherwise,} \end{cases}$$

where ν_2 is the dyadic valuation, the highest power of 2 *that divides an integer.*

Proof By Corollary 2, the weight of a PCR_n cycle of length k is of the form wn/k, for $0 \le w \le k$. It is odd when both w and n/k are odd, the latter implying $\nu_2(n) = \nu_2(k)$, or equivalently that $k = 2^{\nu_2(n)}d$ for any divisor d of $n/2^{\nu_2(n)}$. Also, since $\mathbb{Z}_{2^n}$ contains equally many elements of odd and even weight, $\sum_{k \mid n} k\, \zeta_n^O(k) = 2^{n-1}$. Therefore, $\sum_{d \mid n/2^{\nu_2(n)}} 2^{\nu_2(n)} d\, \zeta_n^O(2^{\nu_2(n)}d) = 2^{n-1}$. Substituting $n' = n/2^{\nu_2(n)}$ in the sum's range and the RHS, and applying Möbius inversion we obtain $\zeta_n^O(2^{\nu_2(n)}d) = \frac{1}{2^{\nu_2(n)}d}\sum_{d' \mid d} \mu(d') 2^{2^{\nu_2(n)}\frac{d}{d'}-1}$. Finally, using $k = 2^{\nu_2(n)}d$ gives the required form. □

Summing $\zeta_n^O(2^{\nu_2(n)}d)$ over all divisors d of $n/2^{\nu_2(n)}$ we obtain $Z^O(n)$. Then, we can obtain a closed formula for $Z^E(n) = Z(n) - Z^O(n)$, despite being unable to obtain a nice one for $\zeta_n^E(k) = \zeta_n(k) - \zeta_n^O(k)$. The proof is a straightforward simplification of this difference, using $\sum_{d \mid n} f(d) = \sum_{d \mid n/2^{\nu_2(n)}} \sum_{l=0}^{\nu_2(n)} f(2^l d)$.

Corollary 3

$$Z^O(n) = \frac{1}{2^{\nu_2(n)+1}} \sum_{d \mid \frac{n}{2^{\nu_2(n)}}} \frac{1}{d} \sum_{d' \mid d} \mu(d') 2^{2^{\nu_2(n)} \frac{d}{d'}}, \tag{11}$$

and

$$Z^E(n) = \frac{1}{2^{\nu_2(n)+1}} \sum_{d \mid \frac{n}{2^{\nu_2(n)}}} \frac{1}{d} \sum_{d' \mid d} \mu(d') \sum_{l=0}^{\nu_2(n)} 2^{2^{\nu_2(n)-l} \frac{d}{d'}+l} = \sum_{l=0}^{\nu_2(n)} Z^O\left(\frac{n}{2^l}\right). \tag{12}$$

3.2 The complemented cycling register

CCR_n was defined in Example 1. For any $x \in \mathbb{Z}_{2^n}$, (4) can be expressed as

$$g_{c_n}(x) = 2x + 1 \mod 2^n + 1. \tag{13}$$

It is easy to check that for any positive integer k we have

$$g_{c_n}^k(x) = 2^k(x+1) - 1 \mod 2^n + 1. \tag{14}$$

The length of the CCR_n cycle containing $x \in \mathbb{Z}_{2^n}$ is equal to the period $p_{g_{c_n}}(x)$, the smallest positive exponent i such that $g_{c_n}^i(x) = x$. Using (14) this is equivalent to $2^i(x+1) \equiv x + 1 \mod 2^n + 1$, or

$$2^i \equiv 1 \mod \frac{2^n + 1}{\gcd(x+1, 2^n+1)},$$

the multiplicative order of 2 modulo this ratio. It follows that when $x + 1$ is coprime to $2^n + 1$, the length of the cycle containing it is equal to $2n$, the maximum possible. Hence, there are at least $\frac{\phi(2^n+1)}{2n}$ cycles of length $2n$.

Hauge [5] proved that the length of each cycle is even and divides $2n$ with an odd quotient. We reformulate this as follows.

Proposition 3 *Any CCR_n cycle has even length that divides $2n$ but not n.*

Proof From (14) we can see that $g_{c_n}^{2n}(x) \equiv x \mod 2^n + 1$, and since $2^n \equiv -1 \mod 2^n + 1$, also that $g_{c_n}^n(x) \equiv \bar{x} \not\equiv x \mod 2^n + 1$ for all $x \in \mathbb{Z}_{2^n}$. If a cycle had length k dividing n, that would contradict the inequality, and if it did not divide $2n$ then we would reach a contradiction as in the proof of Proposition 1. Consequently, k is even. □

Proposition 4 *All CCR_n cycles are self-dual.*

Proof Consider any CCR_n cycle and let $2p$ be its length and x one of its elements. We need to show that it also contains $\bar{x}$. By Proposition 3, p must divide n with an odd quotient, say

$n = (2a+1)p$ for some positive integer a. Then, $\overline{x} \equiv g_{c_n}^n(x) \equiv g_{c_n}^{2pa+p}(x) \equiv g_{c_n}^p(x) \mod 2^n+1$, thus $\overline{x}$ is indeed on the cycle. □

In fact, each cycle has the form $\{x_1, \ldots, x_p, \overline{x_1}, \ldots, \overline{x_p}\}$, where $2p$ is its length. From this, or alternatively by (13), we can also deduce that for any $x \in \mathbb{Z}_{2^n}$,

$$g_{c_n}(\overline{x}) = \overline{g_{c_n}(x)} = g_{c_n}^{\frac{p_{g_{c_n}}(x)}{2}+1}(x). \tag{15}$$

Let $\zeta_n^*(k)$ denote the number of CCR_n cycles of length $2k$.

Lemma 2 $\zeta_n^*(k) = \zeta_n^O(k)$.

Proof By Proposition 3, $2k$ must divide $2n$ but not n. It is not difficult to check that for any integer n, the divisors of $2n$ that are not divisors of n are of the form $2^{\nu_2(n)+1}d$ where d divides $n/2^{\nu_2(n)}$. Hence, $\sum_{d|n/2^{\nu_2(n)}} 2^{\nu_2(n)+1} d\, \zeta_n^*(2^{\nu_2(n)}d) = |\mathbb{Z}_{2^n}| = 2^n$, which is exactly what we had in the proof of Lemma 1 for $\zeta_n^O(2^{\nu_2(n)}d)$ and thus Möbius inversion would yield the same expression. □

Golomb [4] states that CCR_n generates $Z^*(n) = \frac{1}{2n}\sum_{\text{odd } d|n} \phi(d) 2^{n/d}$ cycles. An alternative formula for this is (11), since by Lemma 2, $Z^*(n) = Z^O(n)$.

3.3 The pure and complemented summing registers

Another two well-known examples of mutually complemetary shift registers are the *Pure Summing* (PSR_n) and *Complemented Summing* (CSR_n) registers of order n. As the name suggests, the function F in (1) for PSR_n is the sum of its inputs modulo 2. Under the number theoretic viewpoint, the corresponding support is $\mathcal{D}_{psr} = \{x \in \mathbb{Z}_{2^{n-1}} \mid \text{wt}(x) \text{ is odd}\}$. Using it in (2) we get the maps

$$g_{psr_n}(x) = \begin{cases} 2x \mod 2^n & \text{if wt}(x) \text{ is even} \\ 2x+1 \mod 2^n & \text{if wt}(x) \text{ is odd} \end{cases} = \overline{g}_{csr_n}(x), \tag{16}$$

and

$$g_{csr_n}(x) = \begin{cases} 2x+1 \mod 2^n & \text{if wt}(x) \text{ is even} \\ 2x \mod 2^n & \text{if wt}(x) \text{ is odd} \end{cases} = \overline{g}_{psr_n}(x).$$

At a first glance, a way to unify their description as we did for g_{p_n} and g_{c_n} in (9) and (13) is not evident. As a start, we can eliminate the current split (which is according to the parity of the weight) by mapping the elements of $\mathbb{Z}_{2^n}$ to the elements of $\mathbb{Z}_{2^{n+1}}$ of either odd or even weight. One obvious way of doing this is by mapping elements with the chosen weight parity to themselves, and mapping the rest of the elements to themselves plus 2^n. This yields two invertible maps, one per weight parity, namely $h_n : \mathbb{Z}_{2^n} \to \{x \in \mathbb{Z}_{2^{n+1}} \mid \text{wt}(x) \text{ is even}\}$ and $h'_n : \mathbb{Z}_{2^n} \to \{x \in \mathbb{Z}_{2^{n+1}} \mid \text{wt}(x) \text{ is odd}\}$ given by

$$h_n(x) = \begin{cases} 2^n + x & \text{if wt}(x) \text{ is odd} \\ x & \text{if wt}(x) \text{ is even} \end{cases} \text{ with } h_n^{-1}(x) = \begin{cases} x & \text{if } x < 2^n \\ x - 2^n & \text{if } x \geq 2^n \end{cases} \tag{17}$$

and

$$h'_n(x) = \begin{cases} x & \text{if wt}(x) \text{ is odd} \\ 2^n + x & \text{if wt}(x) \text{ is even} \end{cases} \quad \text{with } h'^{-1}_n(x) = \begin{cases} x & \text{if } x < 2^n \\ x - 2^n & \text{if } x \geq 2^n. \end{cases}$$

Note that by construction, for all $x \in \mathbb{Z}_{2^n}$, $\text{wt}(h_n(x))$ is even and $\text{wt}(h'_n(x))$ odd. The next result establishes a connection between the four basic registers.

Theorem 1 *For any $x \in \mathbb{Z}_{2^n}$ we have*

1. $g_{p_{n+1}}(h_n(x)) = h_n(g_{psr_n}(x)) = g_{c_{n+1}}(h'_n(x))$
2. $g_{p_{n+1}}(h'_n(x)) = h'_n(g_{csr_n}(x)) = g_{c_{n+1}}(h_n(x))$

Proof Since all equalities can be proven in a similar way, we only provide proof for the first one. In fact, we will show that for any $x \in \mathbb{Z}_{2^n}$,

$$g_{p_{n+1}}(h_n(x)) = h_n(g_{psr_n}(x)) = \begin{cases} 2x & \text{if wt}(x) \text{ is even} \\ 2x + 1 & \text{if wt}(x) \text{ is odd.} \end{cases}$$

By (17), for $g_{p_{n+1}}(h_n(x))$ we need to consider two cases, namely $\text{wt}(x)$ even and odd. For $\text{wt}(x)$ even, $g_{p_{n+1}}(h_n(x)) = g_{p_{n+1}}(x) = 2x \mod 2^{n+1} = 2x$, by (3) and since $x < 2^n$ means there is no modular reduction. Similarly, for $\text{wt}(x)$ odd, $g_{p_{n+1}}(h_n(x)) = g_{p_{n+1}}(2^n + x) = 2x + 2^{n+1} + 1 \mod 2^{n+1} = 2x + 1$.

Next, by (16), we need to consider two cases for $h_n(g_{psr_n}(x))$ as well: $\text{wt}(x)$ even and odd. For $\text{wt}(x)$ even, $h_n(g_{psr_n}(x)) = h_n(2x \mod 2^n)$, which by (17) further depends on the parity of $\text{wt}(2x \mod 2^n)$. This also being even means $x < 2^{n-1}$, and being odd means $x \geq 2^{n-1}$. In the first case, $h_n(2x \mod 2^n) = 2x \mod 2^n = 2x$, since there is no modular reduction. In the second case, $h_n(2x \mod 2^n) = (2x \mod 2^n) + 2^n = 2x$, since for the range of x we have $2x \mod 2^n = 2x - 2^n$. The case for $\text{wt}(x)$ odd is almost identical, giving $h_n(g_{psr_n}(x)) = 2x + 1$ for both subcases, as required. □

The length of the PSR_n cycle containing $x \in \mathbb{Z}_{2^n}$ is equal to the period $p_{g_{psr_n}}(x)$, the smallest positive exponent i such that $g^i_{psr_n}(x) = x$. By repeated application of Theorem 1 this is equivalent to $2^i h_n(x) \equiv h_n(x) \mod 2^{n+1} - 1$, or

$$2^i \equiv 1 \mod \frac{2^{n+1} - 1}{\gcd(h_n(x), 2^{n+1} - 1)},$$

the multiplicative order of 2 modulo this ratio. It follows that when $h_n(x)$ is coprime to $2^{n+1} - 1$, $p_{g_{psr_n}}(x) = n + 1$, the maximum possible. Similarly, the length of the CSR_n cycle containing x is the smallest positive i such that

$$2^i \equiv 1 \mod \frac{2^{n+1} - 1}{\gcd(h'_n(x), 2^{n+1} - 1)},$$

the multiplicative order of 2 modulo this ratio, and when $h'_n(x)$ is coprime to $2^{n+1} - 1$, $p_{g_{csr_n}}(x) = n + 1$, the maximum possible.

From Theorem 1 and (6), we can deduce that there is a one-to-one correspondence respectively between the PSR_n and CSR_n cycles, and the PCR_{n+1} cycles of even and odd weight. For example, given a PSR_n cycle $\{x_1, \ldots, x_p\}$, then $\{h_n(x_1), \ldots, h_n(x_p)\}$ is a PCR_{n+1}

cycle of even weight. In the other direction, given for instance a PCR_{n+1} cycle $\{y_1, \ldots, y_l\}$ of odd weight, then $\{h_n'^{-1}(y_1), \ldots, h_n'^{-1}(y_l)\}$ is a CSR_n cycle.

Let $\sigma_n(k)$ and $\sigma_n^*(k)$ respectively denote the number of PSR_n and CSR_n cycles of length k. The following corollaries stem from the observation above.

Corollary 4 *The length of any PSR_n and CSR_n cycle divides $n+1$.*

Corollary 5 $\sigma_n(k) = \xi_{n+1}^E(k)$ *and* $\sigma_n^*(k) = \zeta_{n+1}^O(k)$.

Golomb [4] states that PSR_n generates $S(n) = Z(n+1) - Z^*(n+1)$ cycles and CSR_n generates $S^*(n) = Z^*(n+1)$ cycles, thus $S(n) + S^*(n) = Z(n+1)$, all being later verified using the D-morphism introduced by Lempel [7]. We have just observed the latter in the discussion above, and can obtain the former two from Lemma 2 and Corollary 5, with (12) being an alternative formula to $S(n) = \frac{1}{2(n+1)} \sum_{d|n+1} \phi(2d) 2^{\frac{n+1}{d}}$ given in [11].

4 On the adjacency sets of the pure cycling register

In this section we will show how the new approach can also be used for studying adjacency sets. We do so by revisiting those of PCR_n.

We begin with the fact that by (15), if an edge exists between two PCR_n cycles, then an edge exists between their complements, with the corresponding adjacency sets containing the complements of each other's elements.

Next, by (7), edges can only exist between cycles whose weights differ by one. As a consequence, no intracyclic edge exists in PCR_n [5]. A contradiction with respect to weights asserts that no extracyclic pairs exist either: On one hand, since $x_2 < 2^{n-1}$, $g_{c_n}(x_2)$ being on the same cycle as x_1 implies $\mathrm{wt}(x_1) = \mathrm{wt}(g_{c_n}(x_2)) = \mathrm{wt}(x_2) + 1$. On the other hand, since $x_1 < 2^{n-1}$, $g_{c_n}(x_1)$ being on the same cycle as x_2 implies $\mathrm{wt}(x_2) = \mathrm{wt}(g_{c_n}(x_1)) = \mathrm{wt}(x_1) + 1$.

Regarding intracyclic pairs, 0 and $2^n - 1$, which have period 1, and $2^{n-1} - 1$, for which $g_{c_n}(2^{n-1} - 1) = 2^n - 1 \in C_{2^n-1}$, need not be considered. For any $x \in \mathbb{Z}_{2^{n-1}} \setminus \{2^{n-1} - 1\}$, (4) can be expressed as

$$g_{c_n}(x) = 2x + 1 \mod 2^n - 1.$$

Using this with (9) and rearranging, (8) for PCR_n becomes

$$-2(2^{i+j} - 1)x_1 \equiv 2^j - 1 \mod 2^n - 1, \tag{18}$$

for some i, j such that $1 \le i < p_{g_{p_n}}(x_1)$ and $1 \le j < p_{g_{p_n}}(g_{c_n}(x_1))$. In fact, Lemma 3 below asserts that $1 \le i, j \le n-1$.

We note that we must have $i + j \neq n$, otherwise the LHS of (18) would be congruent to 0, leading to a contradiction as the RHS can never be congruent to 0. Then, the congruence is solvable if and only if $\gcd(2^{i+j} - 1, 2^n - 1) = 2^{\gcd(n, i+j)} - 1$ divides $2^j - 1$, which implies $\gcd(n, i+j)$ divides j.

Lemma 3 *In PCR_n, intracyclic pairs label edges between length n cycles only.*

Proof Let x_1 and x_2 be an intracyclic pair in PCR_n, and denote the length of the cycle containing them by p_1, and that of the cycle containing $g_{c_n}(x_1)$ and $g_{c_n}(x_2)$ by p_2. To prove the lemma it suffices to show $p_1 = p_2 = n$.

First, multiplying both sides of (18) by 2^{p_1}, using that $2^{p_1}x_1 \equiv x_1 \mod 2^n - 1$ and applying (18) on the LHS, and rearranging, we obtain

$$2^{p_1} + 2^j \equiv 2^{p_1+j} + 1 \mod 2^n - 1.$$

We must have that each of the summands on the LHS is congruent to a distinct summand on the RHS modulo $2^n - 1$. Such pairwise congruences are equivalent to pairwise congruences in the exponents modulo n. The range of j implies $j \not\equiv 0 \mod n$, hence the only possibility left is $p_1 \equiv 0 \mod n$ giving $p_1 = n$ as required.

Next, we multiply both sides of (18) by 2^{p_2}. On the LHS we have

$$-(2^{i+j} - 1)2^{p_2}(2x_1 + 1 - 1) \equiv -(2^{i+j} - 1)(2x_1 + 1 - 2^{p_2}) \equiv 2^j - 1 + (2^{i+j} - 1)(2^{p_2} - 1) \mod 2^n - 1,$$

where in the second step we used $2^{p_2}(2x_1 + 1) \equiv (2x_1 + 1) \mod 2^n - 1$, and in the third we applied (18). Combining this with the RHS and rearranging, we obtain

$$2^{p_2+i+j} + 2^j \equiv 2^{p_2+j} + 2^{i+j} \mod 2^n - 1.$$

Working as above, since $i \not\equiv 0 \mod n$, we are left with $p_2 = n$ as required. □

Magleby [8] and Fredricksen (as acknowledged in [5]) proved in different ways that the adjacency sets in PCR_n have size at most 2. The number of adjacency sets of this maximal size was determined in [10, 12] and later on in [5], each using a different method. We provide an alternative proof for both results.

Lemma 4 *All intracyclic pairs in PCR_n are disjoint.*

Proof Suppose on the contrary that there exist two non-disjoint intracyclic pairs in PCR_n, say x_1 with x_2 and x_1 with x_3. Apart from the exponent pair (i, j) that connects x_1 and x_2 as above and yields (18), there exists an exponent pair (i', j'), $1 \le i', j' \le n - 1$, connecting x_1 and x_3 and yielding

$$-2(2^{i'+j'} - 1)x_1 \equiv 2^{j'} - 1 \mod 2^n - 1.$$

Multiplying both sides by $2^{i+j} - 1$, applying (18) on the LHS and rearranging yields

$$2^{i+j+j'} + 2^j + 2^{i'+j'} \equiv 2^{i'+j'+j} + 2^{j'} + 2^{i+j} \mod 2^n - 1.$$

Considering pairwise congruences as in the proof of Lemma 3, three cases arise:

First, $i + j + j' \equiv i' + j' + j \mod n$, which implies $i \equiv i' \mod n$. Then, as $j \not\equiv i + j \mod n$, we are left with $j \equiv j' \mod n$. Given that $1 \le i, j, i', j' \le n - 1$, we must have $i = i'$ and $j = j'$, yielding $x_2 = x_3$ and contradicting that they are distinct.

Second, $i + j + j' \equiv j' \mod n$, implying $i + j \equiv 0 \mod n$, which is impossible as we have seen that $i + j \neq n$.

Third, $i + j + j' \equiv i + j \mod n$ implies $j' \equiv 0 \mod n$ and contradicts the range of j'.

Since all cases lead to a contradiction, our assumption must be false and all intracyclic pairs in PCR_n must be disjoint. □

Corollary 6 [8] *The size of adjacency sets in PCR_n is at most* 2.

Proof Suppose on the contrary that there exists an adjacency set in PCR_n containing more than two distinct elements, and consider three of them. Since no extracyclic pairs exist in PCR_n, pairwise these three elements form non-disjoint intracyclic pairs, in contradiction to Lemma 4. □

Corollary 7 [10, 12] *In PCR_n, adjacency sets of size* 2 *label edges between cycles of length n only.*

Proof This is a direct consequence of Lemma 3 and Corollary 6. □

Theorem 2 [10, 12] *In PCR_n, the number of adjacency sets of size* 2 *is*

$$p(n) = \frac{1}{2} \sum_{\substack{d \mid n \\ d \neq n}} \phi\left(\frac{n}{d}\right)\left(\frac{n}{d} - 2\right)2^{d-1}.$$

Proof Adjacency sets of size 2 in PCR_n correspond to intracyclic pairs. Therefore, we begin by counting the number of suitable pairs of exponents (i, j) that render (18) solvable. As we have seen, we must have $i + j \neq n$ and $\gcd(n, i + j)$ dividing j. Any proper divisor d of n is a possible gcd, and any integer m such that $\gcd(n, m) = d$ and $1 \leq m \leq n - 1$ (due to reduction modulo n in the exponents) is a possibility for $i + j$. There are $\phi(n/d)$ of them. The possibilities for j are the multiples of d excluding $i + j$ (since $i \neq 0$) such that $1 \leq j \leq n - 1$. There are $n/d - 2$ of them.

Next, for each suitable d, i and j there are $2^d - 1$ possible solutions to (18) given by $x_1 = x_0 + \frac{2^n-1}{2^d-1}k$ for each integer k in the interval $0 \leq k \leq 2^d - 2$, where

$$x_0 = -2^{-1}\left(\frac{2^{i+j} - 1}{2^d - 1}\right)^{-1}\left(\frac{2^j - 1}{2^d - 1}\right) \mod \frac{2^n - 1}{2^d - 1}.$$

We are only interested in those solutions such that both $x_1 < 2^{n-1}$ and $x_2 = g^i_{p_n}(x_1) < 2^{n-1}$. When $d = 1$, there is a single solution, given by $x_1 \equiv \sum_{l=1}^{-j(i+j)^{-1} \bmod n} 2^{-1-l(i+j)} \bmod 2^n - 1$. It is in fact straightforward to verify that it satisfies (18). If $x_1 > 2^{n-1}$ then 2^{n-1} must appear as one of the summands, and we would have $-1 - l(i + j) \equiv n - 1 \mod n$ which is only possible if either 0 or n were in the range of the sum. This however does not happen as $j(i + j)^{-1} \not\equiv 0 \mod n$. Thus all x_1 are acceptable, and by Remark 1 and the fact that if the congruence is solvable for the pair of exponents (i, j) then it is also solvable for the pair $(n - i, n - j)$, so are all the corresponding x_2.

For $d > 1$, we have $x_0 < \frac{2^n-1}{2^d-1} < 2^{n-1}$ due to the modulus, hence $k = 0$ is suitable. Since x_0 is between 0 and the modulus, the maximum suitable value of k is k_m such that $k_m\left(\frac{2^n-1}{2^d-1}\right) < 2^{n-1}$ and $(k_m + 1)\left(\frac{2^n-1}{2^d-1}\right) > 2^{n-1}$. After simple operations, this becomes $2^{d-1} - 1 - \frac{2^{n-1}-2^{d-1}}{2^n-1} \leq k_m < 2^{d-1} - \frac{2^{n-1}-2^{d-1}}{2^n-1}$. Since the fraction is less than one, $k_m = 2^{d-1} - 1$. Hence, the suitable solutions are for $0 \leq k \leq 2^{d-1} - 1$, which means that only 2^{d-1} out of the $2^d - 1$ possible solutions are suitable. With the same arguments as for the case $d = 1$, $x_2 = g^i_{p_n}(x_0)$ corresponding to $k = 0$ is also suitable. The suitability of those corresponding to the remaining values of k is established by noticing that

$g^i_{p_n}(\frac{2^n-1}{2^d-1}) = 2^i\frac{2^n-1}{2^d-1} = (2^n-1)\frac{2^i-1}{2^d-1} + \frac{2^n-1}{2^d-1} \equiv \frac{2^n-1}{2^d-1} \mod 2^n - 1$, where in the third step we used the fact that d divides i (as it divides j) and thus $2^d - 1$ divides $2^i - 1$.

Finally, putting everything together gives us the number of suitable solutions to (18). The required number of distinct intracyclic pairs is half this number as by Remark 1 we would be otherwise counting each pair twice. □

5 Conclusions

We have presented an alternative view on binary feedback shift registers where we convert binary sequences to integers and use number theory to treat the related problems. We have tried to illustrate the approach by some examples on basic shift registers, and it seemed to be useful in determining cycle lengths and describing conditions for adjacency sets of size 2 in the pure cycling register via simple congruences. It also led to uncovering a previously unknown connection between the pure cycling register and the cyclotomic classes defined in [1].

One of the differences with the traditional approach is that shift registers are now represented by two equations that depend on the support of their feedback function and the input. What facilitated the detailed study of the four basic examples presented here was our ability to obtain unified descriptions for them as a single equation irrespective of the input. This was made possible by the trivial support of two of them, and the simple support of the other two that lead to the discovery of an isomorphism between their description and that of the first two.

Determining more supports of feedback functions that can lead to unified descriptions, and more generally finding out whether this viewpoint will prove to be a useful tool in studying other shift registers or not is a natural direction for future work, as is generalising it to non-binary alphabets.

Acknowledgements The author would like to thank Prof. Tor Helleseth for valuable discussions and encouragement. He would also like to express his appreciation towards the reviewers and the associate editor for comments that helped improve the manuscript.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital).

References

1. Carlet, C., Goubin, L., Prouff, E., Quisquater, M., Rivain, M.: Higher-order masking schemes for s-boxes. In: Proceedings of FSE, LNCS, vol. 7549, pp. 366–384 (2012)
2. Elspas, B.: The theory of autonomous linear sequential networks. IRE Transactions on Circuit Theory, 45–60 (1959)

3. Fredricksen, H.: A survey of full length nonlinear shift register cycle algorithms. SIAM Review **24**(2), 195–221 (1982)
4. Golomb, S.W.: Shift Register Sequences. Aegean Park Press (1981)
5. Hauge, E.R.: On the cycles and adjacencies in the complementary circulating register. Discrete Mathematics **145**, 105–132 (1995)
6. Jansen, C.J.A., Franx, W.G., Boekee, D.E.: An efficient algorithm for the generation of DeBruijn cycles. IEEE Transaction on Information Theory **37**(5), 1475–1478 (1991)
7. Lempel, A.: On a homomorphism of the de Bruijn graph and its applications to the design of feedback shift registers. IEEE Transaction on Computers **C–19**(12), 1204–1209 (1970)
8. Magleby, K.B.: The synthesis of nonlinear feedback shift registers. Technical Report 6207-1, Stanford Electronics Laboratory (1963)
9. Mykkeltveit, J.: A proof of Golomb's conjecture for the de Bruijn graph. Journal of Combinatorial Theory **13**(1), 40–45 (1972)
10. Mykkeltveit, J.: Generating and counting the double adjacencies in a pure circulating shift register. IEEE Transactions on Computers **C–24**(3), 299–304 (1975)
11. Sloane, N.J.: On single-deletion-correcting codes. In: Codes and Designs, pp. 273–291 (2002)
12. Van Lantschoot, E.J.: Double adjacencies between cycles of a circulating shift register. IEEE Transactions on Computers **C–22**(10), 244–955 (1973)
13. Walker, E.A.: Non-linear recursive sequences. Canadian Journal of Math **11**, 370–378 (1959)