THE STRUCTURE OF COMPLEMENTARY SETS OF INTEGERS: A 3-SHIFT THEOREM

Aviezri S. Fraenkel[1], Dalia Krieger[2]
[1] Dept. of Computer Science and Applied Mathematics
Weizmann Institute of Science, Rehovot 76100, Israel
fraenkel@wisdom.weizmann.ac.il

[2] Dept. of Computer Science and Applied Mathematics
Weizmann Institute of Science, Rehovot 76100, Israel
daliak@wisdom.weizmann.ac.il

Abstract

Let $0 < \alpha < \beta$ be two irrational numbers satisfying $1/\alpha + 1/\beta = 1$. Then the sequences $a'_n = \lfloor n\alpha \rfloor$, $b'_n = \lfloor n\beta \rfloor$, $n \geq 1$, are complementary over $\mathbb{Z}_{\geq 1}$, thus a'_n satisfies: $a'_n = \text{mex}_1\{a'_i, b'_i : 1 \leq i < n\}$, $n \geq 1$ ($\text{mex}_1(S)$, the smallest positive integer not in the set S). Suppose that $c = \beta - \alpha$ is an integer. Then $b'_n = a'_n + cn$ for all $n \geq 1$.

We define the following generalization of sequences a'_n, b'_n: Let c, $n_0 \in \mathbb{Z}_{\geq 1}$, and let $X \subset \mathbb{Z}_{\geq 1}$ be an arbitrary finite set. Let $a_n = \text{mex}_1(X \cup \{a_i, b_i : 1 \leq i < n\})$, $b_n = a_n + cn$, $n \geq n_0$. Let $s_n = a'_n - a_n$. We show that no matter how we pick c, n_0 and X, from some point on the shift sequence s_n assumes either one constant value or three successive values; and if the second case holds, it assumes these values in a very distinct fractal-like pattern, which we describe.

This work was motivated by a generalization of Wythoff's game to $N \geq 3$ piles.

AMS Subj. Classification: 05A17, 91A05

Key Words: complementary sequences, integer part function, 3-shift theorem, Wythoff games

1 Introduction

Two sequences of integers, $S, T \subseteq \mathbb{Z}$, are said to be *complementary over* $\mathbb{Z}_{\geq N}$, or *N-upper complementary*, if neither contains an element more than once, their intersection is empty, and their union covers $\mathbb{Z}_{\geq N}$. A well-known pair of 1-upper complementary sequences is the following:

$$A = \{\lfloor n\varphi \rfloor\}_{n=1}^{\infty}, \quad B = \{\lfloor n\varphi \rfloor + n\}_{n=1}^{\infty},$$

where $\varphi = (1+\sqrt{5})/2$ is the golden section. The pairs $(A, B) \cup \{(0,0)\}$, constitute the set of P-positions of Wythoff's game [25]; see also also [2], [3], [6], [7], [8], [10], [11], [15], [16], [18], [23], [24], [26].

A generalization of Wythoff's game to more than two piles was suggested by Fraenkel in [13], [14], together with two conjectures, listed in Guy and Nowakowski's compendium of 58 research problems [17]. A position in the new game is an N-tuple of nonnegative integers, where N is the number of piles. A convenient way to enumerate the P-positions of the game (second player winning positions) is to fix the first $N-2$ piles, and let the two last ones grow to infinity. Discounting some initial "noise", the last two piles then appear to be a pair of M-upper complementary sequences, $\{A_n\}_{k_0}^{\infty}$, $\{B_n\}_{k_0}^{\infty}$ for some $M \geq 1$. Computer tests done on these sequences suggested that starting at some n_0, the distance between A_n and B_n grows in jumps of 1. A new enumeration of the sequences gives us therefore:

$$B_n = A_n + n \ \ \forall\, n \geq n_0.$$

This enabled us to represent A_n, B_n in the format:

$$A_n = \lfloor n\varphi \rfloor - s_n, \ \ B_n = \lfloor n\varphi \rfloor + n - s_n, \ \ n \geq n_0,$$

where s_n is defined as $\lfloor n\varphi \rfloor - A_n$.

When studying the shift sequence s_n, we encountered a very surprising and unexpected behavior: after somewhat chaotic beginnings, the sequence seems to settle on three successive values, where the middle value is the regular one, and the two other values persist appearing, if at diminishing frequencies. For example, consider Wythoff's game on three piles , with the first pile fixed to 1. In this case, $B_n = A_n + n$ for all $n \geq 23$, and starting from $n = 30$, $s_n \in \{3, 4, 5\}$. The indices n for which $s_n \neq 4$ (*indices of irregular shifts*) grow sparser and sparser as n tends to infinity. Fig. 1 shows the distance between the first 120 or so irregular shift indices. This diagram seems to indicate that the distance between successive indices of irregular shifts grows larger and larger; but it also maintains a very distinct fractal-like pattern! (Further fractal diagrams: Figs 4, 5 in Section 4.)

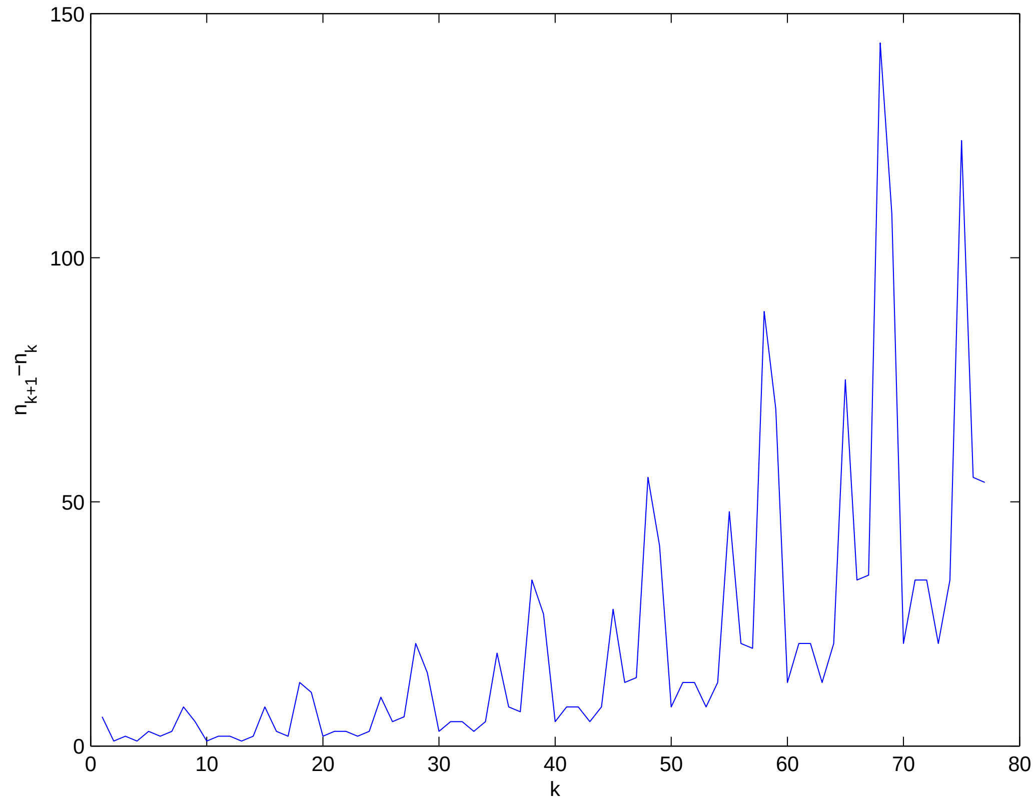

Figure 1: $\{n_k\}$ is the subsequence of indices where $s_{n_k} \neq 4$.

The P-positions of the generalized Wythoff game are but a special case of a large family of complementary sequences. We prove that when a pair of complementary sequences is defined depending on two independent variables, to be specified in Section 2, it exhibits the behavior described above. The main thrust of this work is a precise description of this behavior. In the main Section 3 we prove Theorem 1, which gives the broad structure of the shifts sequence. Its finer structure is given in Section 4, Theorems 2, 3. It is described by means of a second-order recurrence, and a numeration system based on the convergents of the simple continued fraction expansion of an irrational number. The three theorems are formulated at the end of the next section. We also take some steps on the way to showing that the P-positions of the generalized Wythoff game are based on this form. This is done in Section 5, especially in Theorem 7. We end, in the final Section 6, with some further remarks and questions.

2 Preliminaries

Two sequences of positive integers, $S, T \subseteq \mathbb{Z}_{\geq 1}$, are said to be *complementary* over $\mathbb{Z}_{\geq 1}$ if

1. no integer appears more than once in either S or T;
2. $S \cap T = \emptyset$;
3. $S \cup T = \mathbb{Z}_{\geq 1}$.

Definition 1. Let $\alpha < \beta$ be positive real numbers. We define two sequences of positive integers:

$$A' = \{a'_n\}_{n=1}^{\infty}, \qquad a'_n = \lfloor n\alpha \rfloor,\ n \geq 1; \tag{1}$$
$$B' = \{b'_n\}_{n=1}^{\infty}, \qquad b'_n = \lfloor n\beta \rfloor,\ n \geq 1. \tag{2}$$

A well-known result concerning complementary sets of integers is the following:

Lemma 1. *The sequences A' and B' are complementary over $\mathbb{Z}_{\geq 1}$ iff α and β are irrationals, satisfying*

$$\frac{1}{\alpha} + \frac{1}{\beta} = 1. \tag{3}$$

This lemma is usually ascribed to Beatty [1]. In O'Briant [20], it is stated that it was already formulated by Rayleigh [22].

Let $\alpha < \beta$ be two irrational numbers satisfying (3). This implies that

1. $\beta = \alpha/(\alpha - 1)$,
2. $1 < \alpha < 2 < \beta < \infty$.

By the complementarity of A', B', both sequences are strictly increasing, and $b'_n > a'_n$ for all $n \geq 1$. Define the mex_1 operator by

$$X \subsetneq \mathbb{Z}_{\geq 1} \Longrightarrow \text{mex}_1(X) = \min(\mathbb{Z}_{\geq 1} \setminus X). \tag{4}$$

Then

Proposition 1.

$$a'_n = \text{mex}_1(\{a'_i, b'_i : 1 \leq i < n\}). \tag{5}$$

Proof. If $a'_{n_0} \neq x := \text{mex}_1(\{a'_i, b'_i : 1 \leq i < n_0\})$ for some $n_0 \geq 1$, then $x \notin A' \cup B' = \mathbb{Z}_{\geq 1}$, a contradiction. ■

The numbers α and β determine the distance between two consecutive elements of the sequences A' and B', respectively. In general, if x, y are two real numbers, then

$$y - x - 1 < \lfloor y \rfloor - \lfloor x \rfloor < y - x + 1.$$

This implies that

$$\begin{aligned} a'_{n+1} - a'_n &\in \{\lfloor \alpha \rfloor, \lfloor \alpha \rfloor + 1\} = \{1, 2\}, \\ b'_{n+1} - b'_n &\in \{\lfloor \beta \rfloor, \lfloor \beta \rfloor + 1\}. \end{aligned}$$

Therefore, $1 \leq a'_{n+1} - a'_n \leq 2$ for all α, n, whereas $b'_{n+1} - b'_n$ is an appropriate number in $\mathbb{Z}_{\geq 2}$. Also,

$$b'_n - a'_n = \lfloor n\beta \rfloor - \lfloor n\alpha \rfloor \in \{\lfloor n(\beta - \alpha) \rfloor, \lfloor n(\beta - \alpha) \rfloor + 1\}.$$

Note further, that if $\beta - \alpha = c \in \mathbb{Z}_{\geq 1}$, then

$$b'_n = a'_n + cn \quad \forall n \geq 1. \tag{6}$$

In this case, α is given explicitly as a function of c:

$$\frac{1}{\alpha} + \frac{1}{\alpha + c} = 1 \implies \alpha = \frac{2 - c + \sqrt{c^2 + 4}}{2}. \tag{7}$$

For $c = 1$ this number is the golden section, $\varphi = (1 + \sqrt{5})/2$.

We now define a generalization of A' and B', which depends on two independent variables, n_0, X.

Definition 2. Let $\alpha < \beta$ be two irrational numbers satisfying (3), let $n_0 \in \mathbb{Z}_{\geq 1}$, and let $X \subset \mathbb{Z}_{\geq 1}$ be an arbitrary finite set. Define

$$\begin{aligned} a_n &= \text{mex}_1(X \cup \{a_i, b_i : n_0 \leq i < n\}), \quad n \geq n_0, && \text{(8a)} \\ b_n &= a_n + (b'_n - a'_n), \quad n \geq n_0, && \text{(8b)} \\ A &= \{a_n\}_{n=n_0}^{\infty}, && \text{(8c)} \\ B &= \{b_n\}_{n=n_0}^{\infty}. && \text{(8d)} \end{aligned}$$

The special case $n_0 = 1$, $X = \emptyset$ reduces A to A'.

By the mex_1 property, $a_n < a_{n+1}$ for all $n \ge n_0$, i.e., A is a strictly increasing sequence. So is B:

$$b_{n+1} - b_n = a_{n+1} - a_n + (b'_{n+1} - b'_n) - (a'_{n+1} - a'_n) \\ \in a_{n+1} - a_n + \{\lfloor\beta\rfloor, \lfloor\beta\rfloor - 1, \lfloor\beta\rfloor - 2\}. \quad (9)$$

Since $\beta > 2$, we thus have $b_{n+1} - b_n \ge a_{n+1} - a_n \ge 1$. Also, since $a'_n < b'_n$ for all $n \ge 1$, we get that $a_n < b_n$ for all $n \ge n_0$; and together with the mex_1 property of A, this implies that $A \cap B = \emptyset$. The mex_1 property also implies that $\mathbb{Z}_{\ge N} \subset A \cup B$, where $N = \max(X) + 1$. Altogether, we get that A and B are N-upper complementary.

Definition 3. The *shift sequence* of A is the sequence:

$$s_n = a'_n - a_n, \quad n \ge n_0. \quad (10)$$

It follows from this definition that for all $n \ge n_0$,

$$a_n = a'_n - s_n, \quad (11a)$$
$$b_n = b'_n - s_n. \quad (11b)$$

If we could compute the shifts s_n directly, we could also compute a_n, without having to compute $a_{n_0}, \ldots, a_{n-1}$. This would clear a major hurtle in bringing down the computational complexity from exponential to polynomial. The problem is that we don't have any information about s_n: it may assume any value, and doesn't need to be bounded.

In this work we solve this problem for the case $\beta - \alpha = c \in \mathbb{Z}_{\ge 1}$. Before we state the main results, we give an example.

Example 1. $c = 2$, $\alpha = \sqrt{2}$, $n_0 = 6$, $X = \{1, 5\}$. The first few elements of the sequences A, B, s_n, A', B' are displayed in Table 1.

Note that for $n \ge 10$, s_n assumes only the values $7, 8, 9$. For $n \ge 24$, instances of 7 or 9 are always separated by instances of 8. This is a general pattern, as we state now.

Let $c \in \mathbb{Z}_{\ge 1}$, $\alpha = (2 - c + \sqrt{c^2 + 4})/2$, $n_0 \in \mathbb{Z}_{\ge 1}$, $X \subset \mathbb{Z}_{\ge 1}$ a finite set. Let a_n, b_n, s_n be as in Definitions 2, 3.

Theorem 1. *There exist $p \in \mathbb{Z}_{\ge 1}$, $\gamma \in \mathbb{Z}$, such that, either for all $n \ge p$, $s_n = \gamma$; or else, for all $n \ge p$, $s_n \in \{\gamma - 1, \gamma, \gamma + 1\}$. If the second case holds, then:*

1. *s_n assumes each of the three values infinitely often.*

Table 1: $c = 2,\ n_0 = 6,\ X = \{1, 5\}$.

n	a_n	b_n	s_n	a'_n	b'_n
6	2	14	6	8	20
7	3	17	6	9	23
8	4	20	7	11	27
9	6	24	6	12	30
10	7	27	7	14	34
11	8	30	7	15	37
12	9	33	7	16	40
13	10	36	8	18	44
14	11	39	8	19	47
15	12	42	9	21	51
16	13	45	9	22	54
17	15	49	9	24	58
18	16	52	9	25	61
19	18	56	8	26	64
20	19	59	9	28	68
21	21	63	8	29	71
22	22	66	9	31	75
23	23	69	9	32	78
24	25	73	8	33	81
25	26	76	9	35	85
26	28	80	8	36	88
27	29	83	9	38	92
28	31	87	8	39	95
29	32	90	9	41	99
30	34	94	8	42	102
31	35	97	8	43	105
32	37	101	8	45	109
33	38	104	8	46	112
34	40	108	8	48	116
35	41	111	8	49	119

n	a_n	b_n	s_n	a'_n	b'_n
36	43	115	7	50	122
37	44	118	8	52	126
38	46	122	7	53	129
39	47	125	8	55	133
40	48	128	8	56	136
41	50	132	7	57	139
42	51	135	8	59	143
43	53	139	7	60	146
44	54	142	8	62	150
45	55	145	8	63	153
46	57	149	8	65	157
47	58	152	8	66	160
48	60	156	7	67	163
49	61	159	8	69	167
50	62	162	8	70	170
51	64	166	8	72	174
52	65	169	8	73	177
53	67	173	7	74	180
54	68	176	8	76	184
55	70	180	7	77	187
56	71	183	8	79	191
57	72	186	8	80	194
58	74	190	8	82	198
59	75	193	8	83	201
60	77	197	7	84	204
61	78	200	8	86	208
62	79	203	8	87	211
63	81	207	8	89	215
64	82	210	8	90	218
65	84	214	7	91	221

2. *If $s_n \neq \gamma$ then $s_{n-1} = s_{n+1} = \gamma$.*

3. *There exists $K \in \mathbb{Z}_{\geq 1}$, such that the indices $n \geq p$ with $s_n \neq \gamma$ can be partitioned into K disjoint sequences, $\{n_j^{(i)}\}_{j=1}^{\infty}$, $i = 1, \ldots, K$. For each of these sequences, the shift value alternates between $\gamma - 1$ and $\gamma + 1$:*

$$s_{n_j^{(i)}} = \gamma + 1 \Longrightarrow s_{n_{j+1}^{(i)}} = \gamma - 1;$$

$$s_{n_j^{(i)}} = \gamma - 1 \Longrightarrow s_{n_{j+1}^{(i)}} = \gamma + 1.$$

Definition 4. The *indices of irregular shifts* are all the indices $n \geq p$ such that $s_n \neq \gamma$. By Theorem 1, these indices can be partitioned into a family $\mathcal{F} = \{V_1, \ldots, V_K\}$ of $K \geq 0$ disjoint increasing subsequences.

Theorem 2. *Let $V_i = \{n_j\}_{j=1}^{\infty} \in \mathcal{F}$ be a subsequence of irregular shifts. Then it satisfies the following recurrence:*

$$\forall\, j \geq 3,\ n_j = cn_{j-1} + n_{j-2}.$$

Theorem 3. *Let $V_i = \{n_j\}_{j=1}^{\infty} \in \mathcal{F}$ be a subsequence of irregular shifts. Then it satisfies:*

$$\forall\, j \geq 2,\ R_q(n_j) = \mathsf{s}\mathrm{hl}(R_q(n_{j-1})),$$

where $R_q(n)$ is the unique representation of n in the q-numeration system*, and $\mathsf{s}\mathrm{hl}(R_q(n))$ is the* left shift *of this representation (see Section 4).*

3 The value range of s_n

For the rest of this paper, c, n_0 and X are fixed, $\alpha = (2 - c + \sqrt{c^2 + 4})/2$, and A, B, A', B', s_n are as defined in Definitions 1, 2, 3.

Notations:

$$\begin{aligned}
m_0 &= min\{m : a_m > \max(X)\}; && \text{(12a)}\\
\alpha_n &= a_{n+1} - a_n, \quad n \geq m_0; && \text{(12b)}\\
\alpha'_n &= a'_{n+1} - a'_n, \quad n \geq 1; && \text{(12c)}\\
\sigma_n &= s_{n+1} - s_n, \quad n \geq m_0; && \text{(12d)}\\
W &= \{\alpha_n\}_{n=m_0}^{\infty}; && \text{(12e)}\\
W' &= \{\alpha'_n\}_{n=1}^{\infty}; && \text{(12f)}\\
S &= \{\sigma_n\}_{n=m_0}^{\infty}. && \text{(12g)}
\end{aligned}$$

The notation (12a) implies, $\mathbb{Z}_{\geq a_{m_0}} \subseteq A \cup B$.

Proposition 2. *For all $n \geq m_0,\ 1 \leq \alpha_n \leq 2$.*

Proof. Suppose that $\alpha_n \geq 3$ for some $n \geq m_0$. Then $a_n + 1, a_n + 2 \notin A$, but since $\mathbb{Z}_{\geq a_{m_0}} \subseteq A \cup B$, necessarily $a_n + 1, a_n + 2 \in B$. We get that B contains two consecutive numbers, a contradiction:

$$\forall n \geq n_0,\ b_{n+1} - b_n = \alpha_n + c \geq 2.$$

■

Corollary 1. *For all $n \geq m_0,\ -1 \leq \sigma_n \leq 1$.*

Proof. $\sigma_n = \alpha'_n - \alpha_n \in \{1,2\} - \{1,2\} = \{-1,0,1\}$. ■

Proposition 2 and Corollary 1 enable us to regard the difference sequences of A, A', s_n as infinite words over a finite alphabet: W and W' are infinite words over the alphabet $\{1,2\}$; S is an infinite word over the ternary alphabet $\{\bar{1},0,1\}$, where the letter $\bar{1}$ designates -1. Corollary 1 implies the following result:

Corollary 2. *For $n \geq m_0$, there are precisely four possible relationships:*

1. $\alpha_n = 1,\ \alpha'_n = 1,\ \sigma_n = 0$;
2. $\alpha_n = 2,\ \alpha'_n = 2,\ \sigma_n = 0$;
3. $\alpha_n = 1,\ \alpha'_n = 2,\ \sigma_n = 1$;
4. $\alpha_n = 2,\ \alpha'_n = 1,\ \sigma_n = \bar{1}$.

Proof. Follows immediately from $\alpha_n = \alpha'_n - \sigma_n$ and $\alpha_n, \alpha'_n \in \{1,2\}$. ■

Corollary 2 indicates that S depends on the relation between W and W'. The main approach in this paper is to analyze this relation. First, we show how to generate W, W' using a finite prefix. For this, we need a few notations about morphisms on words. These notations can be found e.g. in Lothaire [19].

Let $\sum$ be an alphabet. Then $\sum^*$ is the set of all finite words over $\sum$, ϵ is the empty word, $|w|$ is the length of a word $w \in \sum^*$, and $|w|_x$ is the number of occurrences of the letter $x \in \sum$ in the word $w \in \sum^*$.

The set $\sum^*$ is a *monoid*, i.e., a set equipped with a binary associative operation and an element neutral to this operation: the concatenation operation is associative, and ϵ is a neutral element to concatenation. A function $f : \sum_1^* \to \sum_2^*$ is a *morphism* if it respects the monoid's operation, i.e., $f(xy) = f(x)f(y)$ for all $x, y \in \sum_1^*$. A morphism is uniquely determined by its value on the alphabet. It is *non-erasing* if the image of a letter is always a nonempty word.

Definition 5. $F : \{1,2\}^* \to \{1,2\}^*$ is the non-erasing morphism:

$$F: \begin{array}{l} 2 \to 1^c 2 \\ 1 \to 1^{c-1} 2 \end{array} \tag{13}$$

Clearly, The image of F is the following submonoid of $\{1,2\}^*$:

$$\mathrm{Im}(F) = \{1^{k_1} 2 1^{k_2} 2 \cdots 1^{k_r} 2 \ : \ r \geq 0, \ c-1 \leq k_i \leq c, \ i = 1, \ldots, r\} \ . \tag{14}$$

This submonoid is isomorphic to $\{1,2\}^*$ under F, and every word $w = x_1 \cdots x_r \in \{1,2\}^*$ satisfies:

$$w = F^{-1}(1^{k_1} 2 1^{k_2} 2 \cdots 1^{k_r} 2); \quad k_i = c - 2 + x_i, \ i = 1, \ldots, r \ . \tag{15}$$

Lemma 2 (Generation Lemma). *Let $t \geq m_0$, and let $n > t$ be such that $a_n = b_t + 1$. Let $u = \alpha_t \cdots \alpha_{n-1}$, and let $v = \alpha_{m_0} \cdots \alpha_{t-1}$. Then for all $k \geq 0$,*

$$F^k(u) = \alpha_p \cdots \alpha_{q-1}, \ a_q = b_p + 1 \Longrightarrow \\ F^{k+1}(u) = \alpha_q \cdots \alpha_{r-1}, \ a_r = b_q + 1,$$

and hence,

$$W = vuF(u)F^2(u) \cdots \ .$$

The same holds for W', except that we can start from $t \geq 1$ instead of $t \geq m_0$.

Proof. Note that $b_t + 1$ must be in A, since B contains no consecutive integers. Let $a_m = b_n + 1$, and let $w = \alpha_n \cdots \alpha_{m-1}$. We'll show that $w = F(u)$. The rest will follow by induction.

Suppose $\alpha_t = a_{t+1} - a_t = 1$. Then $b_{t+1} - b_t = c + 1$, i.e., there are c consecutive elements of A in the open interval $I = (b_t, b_{t+1})$: $b_t + 1, \ldots, b_{t+1} - 1$. The distance between each consecutive pair of these elements is 1: otherwise, we would have an element of B between b_t and b_{t+1}, which are consecutive elements of B. The next element of A is $b_{t+1} + 1$, which has a distance of 2 from the previous element. Altogether, we have $c + 1$ consecutive elements of A, $b_t + 1 = a_n, \ldots, a_{n+c} = b_{t+1} + 1$, and the distances between them form a word of length c, $\alpha_n \cdots \alpha_{n+c-1} = 1^{c-1}2 = F(1)$.

Similarly, if $\alpha_t = 2$, there are $c + 1$ consecutive elements of A in I, and altogether we have $c+2$ consecutive elements of A, $b_t + 1 = a_n, \ldots, a_{n+c+1} = b_{t+1}+1$. The distances between them form a word of length $c+1$, $\alpha_n \cdots \alpha_{n+c} = 1^c 2 = F(2)$.

We continue this way to the letter $\alpha_{t+1} = a_{t+2} - a_{t+1}$, which induces the elements $b_{t+1} + 1, \ldots, b_{t+2} + 1$, and so on, until $\alpha_{n-1} = a_n - a_{n-1}$, which induces the elements $b_{n-1} + 1, \ldots, b_n + 1 = a_m$. Altogether,

$$F(u) = F(\alpha_t) \cdots F(\alpha_{n-1}) = \alpha_n \cdots \alpha_{m-1} = w.$$

Let now $a_j = b_m + 1$. By induction, if we set $t = n$, $n = m$, and $m = j$, we get that $F(w) = F^2(u) = \alpha_m \cdots \alpha_{j-1}$, and so on. The same proof applies to W'. ■

Example 2. Figure 2 illustrates Lemma 2. For $c = 2$, the word $\alpha_t \alpha_{t+1} \alpha_n = 212$ is mapped to the word $\alpha_n \cdots \alpha_m = 11212112$. The letter α_t induces the elements $b_t + 1, \ldots, b_{t+1} + 1$; the letter α_{t+1} induces the elements $b_{t+1} + 1, \ldots, b_{t+2} + 1$; the letter $\alpha_{t+2} = \alpha_{n-1}$ induces the elements $b_{t+2} + 1, \ldots, b_n + 1 = a_m$.

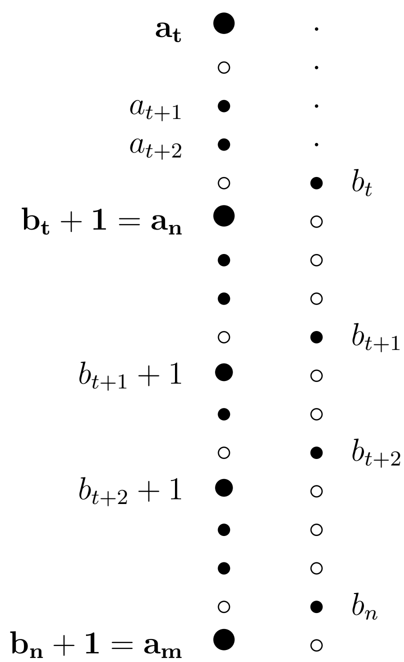

Figure 2: Illustration of Lemma 2. Left column - subset of A; right column - subset of B.

Definition 6. A *generator* for W (W') is a word of the form $u = \alpha_t \cdots \alpha_{n-1}$ ($u' = \alpha'_r \cdots \alpha'_{m-1}$), where $a_n = b_t + 1$ ($a'_m = b'_r + 1$). We say that W, W' are generated *synchronously* if there exist generators u, u', such that $u = \alpha_t \cdots \alpha_{n-1}$, $u' = \alpha'_t \cdots \alpha'_{n-1}$ (same indices t, n), and

$$\forall\, k \geq 0,\ F^k(u) = \alpha_g \cdots \alpha_{h-1} \Longleftrightarrow F^k(u') = \alpha'_g \cdots \alpha'_{h-1},$$

where $a_h = b_g + 1$. Such generators u, u' are called *synchronous generators.*

In order to analyze the relation between W and W', we need to show that we can generate them synchronously. This will enable us to compare $F^k(u)$ and $F^k(u')$ for all $k \geq 0$. But this isn't always the case: if we go back to Example 1, we see that $m_0 = 9$, as $a_9 = 6 = \max(X) + 1$. For $t = m_0$, we get that $b_t + 1 = 25 = a_{24}$, but $b'_t + 1 = 31 = a'_{22}$. Note, however, that for $t = 13$, $b_{13} + 1 = 37 = a_{32}$, and $b'_{13} + 1 = 45 = a'_{32}$.

Proposition 3. *Let $u \in \{1,2\}^*$. Then for all $k \geq 2$,*

$$\begin{aligned} |F^k(u)|_1 &= c|F^{k-1}(u)|_1 + |F^{k-2}(u)|_1, \\ |F^k(u)|_2 &= c|F^{k-1}(u)|_2 + |F^{k-2}(u)|_2, \end{aligned} \tag{16}$$

Proof. Let $g_k = |F^k(u)|_1$, $h_k = |F^k(u)|_2$, $k \geq 0$. By definition of F, for all $k \geq 1$ we have:

$$\begin{aligned} g_k &= (c-1)g_{k-1} + ch_{k-1}, \\ h_k &= g_{k-1} + h_{k-1}. \end{aligned}$$

Thus $g_k - ch_k = -g_{k-1}$ for all $k \geq 1$, i.e., $ch_{k-1} = g_{k-1} + g_{k-2}$ for all $k \geq 2$. The recurrence for h_k holds similarly. ■

The next proposition is phrased for W, but applies for W' as well.

Proposition 4. *Let $n > t \geq m_0$, and consider the subsequence $a_t, \ldots, a_n$ of A. Let $u = \alpha_t \cdots \alpha_{n-1}$. Then*

$$\begin{aligned} |u|_1 &= 2(n-t) - (a_n - a_t), \\ |u|_2 &= (a_n - a_t) - (n-t). \end{aligned}$$

In particular, if $a_n = b_t + 1$, then

$$\begin{aligned} |u|_1 &= 2n - (c+2)t - 1, \\ |u|_2 &= (c+1)t - n + 1. \end{aligned}$$

Proof. The integer interval $[a_t, a_n]$ contains $a_n - a_t + 1$ elements in general, and $n - t + 1$ elements of A. The remaining $(a_n - a_t) - (n - t)$ elements belong to B, and correspond to gaps of size 2 in the subsequence $a_t, \ldots, a_n$ of A. The total number of gaps in $a_t, \ldots, a_n$ is $n - t$, therefore there are $(n-t) - [(a_n - a_t) - (n-t)] = 2(n-t) - (a_n - a_t)$ gaps of size 1. If $a_n = b_t + 1$, then $a_n - a_t = ct + 1$, and hence the special case. ■

Corollary 3. *If for some* $t \geq m_0$, $b_t + 1 = a_n$ *and* $b'_t + 1 = a'_n$, *then the words*

$$\begin{aligned} u &= \alpha_t \cdots \alpha_{n-1}, \\ u' &= \alpha'_t \cdots \alpha'_{n-1}, \end{aligned}$$

are permutations of each other.

Proof. By Proposition 4, $|u|_1 = |u'|_1 = 2n - (c+2)t - 1$ and $|u|_2 = |u'|_2 = (c+1)t - n + 1$. ■

Corollary 4. *If for some* $t \geq m_0$, $b_t + 1 = a_n$ *and* $b'_t + 1 = a'_n$, *then* W, W' *are generated synchronously by* u, u', *respectively.*

Proof. Suppose such t exists. By Lemma 2,

$$\begin{aligned} W &= \alpha_{m_0} \cdots \alpha_{t-1} u F(u) F^2(u) \cdots \\ W' &= \alpha'_1 \cdots \alpha'_{t-1} u' F(u') F^2(u') \cdots . \end{aligned}$$

by Proposition 3, $|F^k(u)|_1$, $|F^k(u)|_2$ depend on $|u|_1$, $|u|_2$, and the same holds for $|F^k(u')|_1$, $|F^k(u')|_2$. Since u and u' are permutations of each other, so are $F^k(u)$, $F^k(u')$ for all $k \geq 0$, i.e.,

$$|F^k(u)| = |F^k(u')| \; \forall \, k \geq 0.$$

Therefore,

$$\forall \, k \geq 0, \; F^k(u) = \alpha_g \cdots \alpha_{h-1} \Longleftrightarrow F^k(u') = \alpha'_g \cdots \alpha'_{h-1},$$

where $a_h = b_g + 1$ and $a'_h = b'_g + 1$. ■

The next key-lemma shows that we can always find generators u, u' that generate W, W' synchronously.

Lemma 3 (Synchronization Lemma). *Let* m_1 *be such that* $a_{m_1} = b_{m_0} + 1$. *Then there exists an integer* $t \in [m_0, m_1]$, *such that* $b_t + 1 = a_n$ *and* $b'_t + 1 = a'_n$.

Proof. If $a'_{m_1} = b'_{m_0} + 1$, then $t = m_0$. Otherwise, we have $a'_{m_1} = b'_{m_0} + 1 + k$ for some $k \neq 0$. In the sequel we use language appropriate for $k \geq 0$, but the arguments hold symmetrically for $k < 0$.

Let $w = \alpha_{m_0} \cdots \alpha_{m_1 - 1}$, $w' = \alpha'_{m_0} \cdots \alpha'_{m_1 - 1}$. By Proposition 4 we get:

$$\begin{aligned} |w|_1 - |w'|_1 = [2(m_1 - m_0) - (a_{m_1} - a_{m_0})] - [2(m_1 - m_0) - (a'_{m_1} - a'_{m_0})] = \\ (b'_{m_0} + 1 + k - a'_{m_0}) - (b_{m_0} + 1 - a_{m_0}) = k. \end{aligned}$$

If we examine the pairs $\left[\begin{smallmatrix}\alpha_i\\ \alpha'_i\end{smallmatrix}\right]$, $i = m_0, \ldots, m_1 - 1$, this means that there are k more pairs of the form $\left[\begin{smallmatrix}1\\ 2\end{smallmatrix}\right]$ than of the form $\left[\begin{smallmatrix}2\\ 1\end{smallmatrix}\right]$. Let us check the influence of each type of pair. First, observe that, for $i \in \{1, 2\}$,

$$a_{m_0+1} - a_{m_0} = i \Longrightarrow b_{m_0+1} - b_{m_0} = c + i \Longrightarrow b_{m_0+1} + 1 = a_{m_1+c+(i-1)}.$$

This is true because there are $c + i - 1$ elements of A between b_{m_0} and b_{m_0+1}. The first of them is $b_{m_0} + 1 = a_{m_1}$, so the last is $a_{m_1+c+i-2}$. The next $A-$element is therefore $b_{m_0+1} + 1 = a_{m_1+c+i-1}$. This gives us:

$$\begin{aligned}\alpha_{m_0} = 1 &\implies b_{m_0+1} + 1 = a_{m_1+c}\ ;\\ \alpha_{m_0} = 2 &\implies b_{m_0+1} + 1 = a_{m_1+c+1}\ .\end{aligned}$$

Analogous results hold for A', B'. Note also, that

$$a'_{m_1} = b'_{m_0} + 1 + k \Longrightarrow b'_{m_0} + 1 = a'_{m_1} - k = a'_{m_1-j} \text{ for some j} \le \text{k}.$$

The last inequality results from the fact that $a'_{m_1-i+1} - a'_{m_1-i}$ might equal 2, so it might take less than k elements to get to $a'_{m_1} - k$. We define the *distance to synchronization* to be j:

$$\begin{aligned}b_{m_0} + 1 &= a_{m_1},\\ b'_{m_0} + 1 &= a'_{m_1-j}.\end{aligned}$$

We will show that this distance must decrease to 0 along the way from m_0 to m_1.

There are four possible values for $\left[\begin{smallmatrix}\alpha_{m_0}\\ \alpha'_{m_0}\end{smallmatrix}\right]$: $\left[\begin{smallmatrix}1\\ 2\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}2\\ 1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}1\\ 1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}2\\ 2\end{smallmatrix}\right]$.

1. $\alpha_{m_0} = 1$, $\alpha'_{m_0} = 2$.

$$\begin{aligned}b_{m_0+1} + 1 &= a_{m_1+c},\\ b'_{m_0+1} + 1 &= a'_{m_1+c-j+1}.\end{aligned}$$

   For the index $m_0 + 1$, the distance to synchronization is now $j - 1$, i.e., a pair of type $\left[\begin{smallmatrix}1\\ 2\end{smallmatrix}\right]$ reduces the distance by 1.

2. $\alpha_{m_0} = 2$, $\alpha'_{m_0} = 1$.

$$\begin{aligned}b_{m_0+1} + 1 &= a_{m_1+c+1},\\ b'_{m_0+1} + 1 &= a'_{m_1+c-j}.\end{aligned}$$

   The distance to synchronization increased to $j + 1$, i.e., a pair of type $\left[\begin{smallmatrix}2\\ 1\end{smallmatrix}\right]$ increases the distance by 1.

3. $\alpha_{m_0} = \alpha'_{m_0} = 1$.

$$\begin{aligned} b_{m_0+1} + 1 &= a_{m_1+c}, \\ b'_{m_0+1} + 1 &= a'_{m_1+c-j}. \end{aligned}$$

The distance to synchronization remains unchanged.

4. $\alpha_{m_0} = \alpha'_{m_0} = 2$.

$$\begin{aligned} b_{m_0+1} + 1 &= a_{m_1+c+1}, \\ b'_{m_0+1} + 1 &= a'_{m_1+c-j+1}. \end{aligned}$$

Again, the distance to synchronization remains unchanged.

In order to achieve synchronization, i.e., t such that $b_t+1 = a_n$ and $b'_t+1 = a'_n$, we need to have j more pairs of type $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$ than of type $\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$. But $j \le k$, and we have k more such pairs between m_0 and m_1. Therefore, we must get to such t along the way. ■

Example 3. As we saw before in Example 1, $m_0 = 9$, $m_1 = 24$, and $a'_{24} = 33 = b'_9 + 1 + 2$, i.e.,

$$b'_9 + 1 = a'_{24} - 2 = 31 = a'_{22} = a'_{m_1-2}.$$

The distance to synchronization is 2, so we need two more pairs of the form $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$ than of the form $\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$. The words u, u' are:

$$\begin{aligned} u &= 111111121212112 = \alpha_9 \cdots \alpha_{23}; \quad |u|_1 = 11. \\ u' &= 211212121121211 = \alpha'_9 \cdots \alpha'_{23}; \quad |u'|_1 = 9. \end{aligned}$$

After 4 letters we achieve synchronization, so at index $m_0 + 4 = 13$ we get:

$$\begin{aligned} b_{13} + 1 &= 37 = a_{32}; \\ b'_{13} + 1 &= 45 = a'_{32}. \end{aligned}$$

The words $v = \alpha_{13} \cdots \alpha_{31}$, $v' = \alpha'_{13} \cdots \alpha'_{31}$ are generators that generate W, W' synchronously.

We will now use Lemma 3 to trace the structure of S (Notations (12g)). In particular, we will show that letters different from 0 must come in pairs (at least from some point on), between which come sequences of 0, that grow longer and longer as n goes to infinity.

Let $u \in \{1,2\}^*$, and let u' be a permutation of u. We apply F successively to pairs of words $\left[\begin{smallmatrix} u \\ u' \end{smallmatrix}\right]$:

$$F \begin{bmatrix} u \\ u' \end{bmatrix} := \begin{bmatrix} F(u) \\ F(u') \end{bmatrix} .$$

By Proposition 3, $F^k(u')$ is a permutation of $F^k(u)$ for all $k \geq 0$.

Example 4. $c = 1$:

$$\begin{bmatrix} 1111222 \\ 2221111 \end{bmatrix} \mapsto \begin{bmatrix} 2222121212 \\ 1212122222 \end{bmatrix} \mapsto \begin{bmatrix} 12121212212212212 \\ 21221221212121212 \end{bmatrix} .$$

Definition 7. A *well-formed string of parentheses* is a string $\vartheta = t_1 \cdots t_n$ over some alphabet which includes the letters '(', ')', such that for every prefix μ of ϑ, $|\mu|_{(} \geq |\mu|_{)}$ (never close more parentheses than were opened), and $|\vartheta|_{(} = |\vartheta|_{)}$ (don't leave opened parentheses).

The *nesting level* $N(\vartheta)$ of such a string is the maximal number of opened parentheses: let $p_1, \ldots, p_n$ satisfy

$$p_i = \begin{cases} 1 & \text{if } \mathrm{t_i} = (\\ -1 & \text{if } \mathrm{t_i} =) \\ 0 & \text{otherwise} \end{cases} ,$$

then

$$N(\vartheta) = \max_{1 \leq k \leq n} \{ \sum_{i=1}^{k} p_i \} .$$

The fact that u, u' are permutations of each other implies that they include the same number of pairs $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ as pairs $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$, so we can regard them as a well-formed string of parentheses: put '$\bullet$' for $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, '$\circ$' for $\left[\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}\right]$; and '(', ')' for $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ alternately, such that the string remains well-formed: if the first non-equal pair we encounter is $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, then '(' stands for $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ and ')' stands for $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$, until all the opened parentheses are closed. Then we start again, by placing '(' for the first occurrence different from $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}\right]$.

Notation: If ϑ is the parentheses string corresponding to u, u', we denote by $F(\vartheta)$ the parentheses string of the application of F to the pair u, u'.

Example 5. $c = 1$, u, u' are as in Example 4. Then

$$(((\bullet))) \mapsto (\circ(\circ \bullet \circ)\circ)\circ \mapsto ()(\circ \bullet \circ() \circ \bullet\circ)() \circ \bullet \circ .$$

Note that the first application of F decreased the nesting level from 3 to 2. The second application left it unchanged, but another application would decrease it to 1:

$$()(\circ \bullet \circ() \circ \bullet\circ)() \circ \bullet\circ \;\mapsto\; () \circ ()(\circ)(\circ \bullet \circ)(\circ)() \circ () \circ \bullet \circ \circ \bullet \circ \,.$$

Example 5 exhibits the general case: the nesting level is monotonically non-increasing under F. Except for one case ($c = 1,\ N(\vartheta) = 2$) it is, in fact, strictly decreasing. To show this, we use an auxiliary result.

Proposition 5. *Let $u \in \{1,2\}^*$. Then the words 1^t, $t > c$, or 21^t2, $t < c - 1$, are not subwords of $F(u)$. In particular, for $c = 1$, 11 is not a subword of $F(u)$.*

Proof. Follows immediately from the definition of F. ∎

Corollary 5. *Let $u \in \{1,2\}^*$. For $c = 1$, 222 is not a subword of $F^2(u)$.*

Proof. The string 222 is the image of 111, or in the image of 211, which are not subwords of $F(u)$. ∎

We are now ready to prove our second key-lemma.

Lemma 4 (Nesting Lemma). *Let $u(0) \in \{1,2\}^*$, let $u'(0)$ be a permutation of $u(0)$. If $c = 1$ and $u(0)$ or $u'(0)$ contains 11, put $u := F(u(0))$, $u' := F(u'(0))$. Otherwise put $u := u(0)$, $u' := u'(0)$. Let $\vartheta \in \{\bullet,\ \circ,\ (,\ )\}^*$ be the parentheses string of $\left[\begin{smallmatrix} u \\ u' \end{smallmatrix}\right]$. Then successive applications of F decrease the nesting level to 1. Specifically,*

- **(I)** *If $c > 1$, then $N(\vartheta) > 1 \Longrightarrow N(F(\vartheta)) < N(\vartheta)$.*
- **(II)** *If $c = 1$,*
  - **(a)** *$N(\vartheta) > 2 \Longrightarrow N(F(\vartheta)) < N(\vartheta)$;*
  - **(b)** *$N(\vartheta) = 2 \Longrightarrow N(F^2(\vartheta)) = 1$.*

Proof. (**I**) Let $w = F(u)$, $w' = F(u')$. Let u_i, $1 \le i \le |u|$, be the prefix of length i of u, and similarly define u'_i, w_i, w'_i. Let $N(\vartheta) = n$, $N(F(\vartheta)) = m$, and suppose that $m \ge n$. Then there exists a smallest $1 \le i \le |w|$ such that $||w_i|_2 - |w'_i|_2| = ||w_i|_1 - |w'_i|_1| = m$. (Since ϑ is well-formed, $i < |w|$.) W.l.o.g., assume that $|w_i|_2 > |w'_i|_2$. We can assume that w_i ends with 2 and

w'_i ends with 1, otherwise there would be shorter prefixes with nesting level m. Therefore by (14),

$$\begin{aligned} w_i &= 1^{k_1}21^{k_2}2\cdots 1^{k_{r+m}}2; \\ w'_i &= 1^{\ell_1}21^{\ell_2}2\cdots 1^{\ell_r}21^t; \\ & r\geq 0,\ k_j,\ell_j\in\{c,c-1\},\ 1\leq t\leq c. \end{aligned}$$

Let

$$\begin{aligned} x &= |\,\{j\in\{1,\ldots,r+m\}\ :\ k_j=c-1\}\,|, \\ y &= |\,\{j\in\{1,\ldots,r\}\ :\ \ell_j=c-1\}\,|. \end{aligned}$$

Then

$$\begin{aligned} i=|w_i| &= xc+(r+m-x)(c+1), \\ i=|w'_i| &= yc+(r-y)(c+1)+t, \end{aligned}$$

i.e.,

$$x-y=mc+m-t. \tag{17}$$

By (15), for all $i=1,\ldots,|u|$, $u_i=F^{-1}(1^{k_1}2\cdots 1^{k_i}2)$, where $1^{k_1}2\cdots 1^{k_i}2$ is a prefix of w. The same applies to u',w'. In particular,

$$F^{-1}(w_i) = u_{r+m}. \tag{18}$$

$$F^{-1}(w'_{i-t}) = u'_r. \tag{19}$$

Therefore,

$$|u_{r+m}|_1 = x, \tag{20}$$

$$|u'_r|_1 = y. \tag{21}$$

From (20), $|u_r|_1\geq x-m$. Therefore, by (17) and since we assumed $m\geq n$,

$$\begin{aligned} n\geq |u_r|_1-|u'_r|_1\geq x-m-y= \\ mc-t=(m-1)c+(c-t)\geq(n-1)c. \end{aligned} \tag{22}$$

This is a contradiction if $(n-1)c>n$. We consider two cases:

$\mathbf{c\geq 3}$: Then $n>1$ implies $(n-1)c>n$. Therefore, if $n>1$, necessarily $m<n$, and the nesting level must decrease to 1 after at most $n-1$ applications of F.

$\mathbf{c = 2}$: Then $n > 2$ implies $(n-1)c > n$, so for $n > 2$, $m < n$. Suppose $n = 2$. By (22), $m > n$ implies $n \geq |u_r|_1 - |u'_r|_1 \geq nc = 4$, a contradiction. Our assumption $m \geq n$ thus implies $m = n$. Again by (22),

$$2 = n \geq |u_r|_1 - |u'_r|_1 \geq |u_{r+m}|_1 - m - |u'_r|_1 = \\ x - m - y = (m-1)c + (c-t) = 2 + (2-t).$$

Thus there is equality throughout, so $t = 2$, $|u_r|_1 - |u'_r|_1 = 2$, and $|u_r|_1 = |u_{r+m}|_1 - m$. The latter equality implies $u_{r+m} = u_r 1^m$. Thus

$$u_{r+1} = u_r 1. \tag{23}$$

Also note that for $t = c$ we have from (19) $w'_i = F(u'_r)1^c$, so

$$u'_{r+1} = u'_r 2 \quad (\text{for } t = c). \tag{24}$$

This yields

$$2 = n \geq |u_{r+1}|_1 - |u'_{r+1}|_1 = |u_r|_1 - |u'_r|_1 + 1 = 3\ ,$$

a contradiction, so $m < n$.

Again, the nesting level must decrease to 1 after at most $n-1$ applications of F, completing the proof of **(I)**.

(IIa). $\mathbf{c = 1, n > 2}$. Then $t = 1$, and by (22), $n \geq m - 1$. If $n = m - 1$, then, again by (22), $n \geq |u_r|_1 - |u'_r|_1 \geq n$. As above, this implies (23) and (24). Hence, $n \geq |u_{r+1}|_1 - |u'_{r+1}|_1 = n + 1$, a contradiction.

So suppose that $m = n$. Then by (22), $n \geq |u_r|_1 - |u'_r|_1 \geq n - 1$. If $|u_r|_1 - |u'_r|_1 = n - 1$, then (22) implies

$$n - 1 = |u_r|_1 - |u'_r|_1 \geq |u_{r+n}|_1 - n - |u'_r|_1 = x - n - y = n - 1.$$

Thus, $u_{r+n} = u_r 1^n$. Since 11 is not a subword of u, this case is not possible for $n > 1$, a contradiction.

If $|u_r|_1 - |u'_r|_1 = n$, then necessarily $u_{r+1} = u_r 2$, otherwise (24) would imply $n \geq |u_{r+1}|_1 - |u'_{r+1}|_1 = n + 1$, a contradiction. By (22),

$$n = |u_r|_1 - |u'_r|_1 \geq x - m - y = |u_{r+n}|_1 - n - |u'_r|_1 = n - 1.$$

Therefore $|u_r|_1 = |u_{r+n}|_1 - n + 1$. This and $u_{r+1} = u_r 2$ imply $u_{r+n} = u_r 21^{n-1}$. Since u doesn't contain the subword 11, this is possible only if $n \leq 2$, a contradiction, completing the proof of subcase **(a)**.

(IIb). $\mathbf{c = 1, n = 2}$. At the beginning of **(IIa)** we concluded that $m = n$ without any assumption on n. Hence the remaining case is $c = 1$, $m = n = 2$.

We saw that this implies $u_{r+2} = u_r21$. In view of (24), we might have $u'_{r+2} = u'_r22$. But then $n \geq |u_{r+2}|_1 - |u'_{r+2}|_1 = n+1$, a contradiction. Therefore, $u'_{r+2} = u'_r21$. By (18), (19),

$$\begin{array}{lllllll} w_i & = & F(u_{r+2}) & = & F(u_r)F(21) & = & 1^{k_1}2\cdots 1^{k_r}2122\,; \\ w'_{i+2} & = & F(u'_{r+2}) & = & F(u'_r)F(21) & = & 1^{\ell_1}2\cdots\cdots 1^{\ell_r}2122\,. \end{array}$$

Since u, u' do not contain 11, w, w' do not contain 222 (Corollary 5). Therefore the next two letters of w must be 12, so

$$\begin{array}{lll} w_{i+2} & = & w_{i-1}\mathbf{21}2\,, \\ w'_{i+2} & = & w'_{i-1}\mathbf{12}2\,. \end{array}$$

The digits in boldface show that m reaches 2 at position i, and it switches back to 1 at position $i+1$. This means that $F(\vartheta)$ has the form,

$$F(\vartheta) = (\cdots()\cdots()\cdots()\cdots)\,. \tag{25}$$

We'll now show that if $F(\vartheta)$ is of the form (25), then $N(F^2(\vartheta)) = 1$. W.l.o.g, assume that '(' stands for $\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$ and ')' stands for $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$. The picture is the following:

$$\begin{array}{lll} w & = & 2v_121v_221\cdots 21v_k1, \\ w' & = & 1v_112v_212\cdots 12v_k2, \end{array}$$

where $v_j \in \{1,2\}^*$, $j = 1,\ldots,k$. Since 11 is not a subword, v_1 and v_k can't be empty , and must begin and end with 2. The other v_j might be empty, but if not they must begin and end with 2 as well. Applying F gives the following picture:

$$\begin{array}{lll} F(w) & = & \mathbf{12}12\cdots 12\mathbf{122}\cdots\cdots\mathbf{122}12\cdots 12\mathbf{2} \\ F(w') & = & \mathbf{2}12\cdots 12\mathbf{212}\cdots\cdots\mathbf{212}12\cdots 12\mathbf{12} \end{array}$$

Let's examine the structure of these words more closely. The 2-letter prefix is a pair (), which can be deleted, since it's a balanced. A prefix of the remaining suffix is:

$$\begin{array}{l} 12\cdots \\ 2\cdots \end{array}\,, \tag{26}$$

which can be continued in precisely two ways, depending on whether v_1 begins with 22 or with 212:

$$\begin{array}{l} 1212\cdots \\ 212\cdots \end{array} \quad\text{or}\quad \begin{array}{l} 12212\cdots \\ 2212\cdots \end{array}\,. \tag{27}$$

In both cases, we can delete a balanced prefix and get back to (26): in the first case () is deleted, in the second (∘). The image of the suffix right after v_1, concatenated to (26), is

$$\begin{matrix} 12\mathbf{122}\cdots \\ 2\mathbf{212}\cdots \end{matrix} \quad ,$$

As long as v_j, $j > 2$ is empty, we continue with

$$\begin{matrix} 12\mathbf{122122}\cdots \\ 2\mathbf{212212}\cdots \end{matrix} \quad .$$

If there is a non-empty v_j ($j < k$), then the picture is:

$$\begin{matrix} 12\mathbf{122}\cdots\mathbf{122}12\cdots12\mathbf{122}\cdots \\ 2\mathbf{212}\cdots\mathbf{212}12\cdots12\mathbf{212}\cdots \end{matrix} \quad .$$

The prefix ($\circ \bullet \circ \cdots \circ \bullet \circ$) can be removed, and we are back to (26). The process ends with

$$\begin{matrix} \cdots\cdot\mathbf{122}12\cdots12\mathbf{2} \\ \cdots\mathbf{212}12\cdots12\mathbf{12} \end{matrix} \quad ,$$

and the nesting level is 1, as claimed. ■

Lemma 5. *Under the hypotheses of the previous lemma, if $N(\vartheta) = 1$, then $F^2(\vartheta)$ has the form:*

$$\cdots()\cdots()\cdots()\cdots \quad , \tag{28}$$

where the dot strings consist of '∘', '•' letters, and might be empty. Further applications of F preserve this form, with the same number of parentheses pairs; the only change is that the dot strings grow longer.

Proof. It's enough to prove the lemma for $\vartheta = (\cdots)$, since

$$F\left[(\cdots)\cdots(\cdots)\right] = F(\cdots)\cdots F(\cdots).$$

Applying F yields

$$F\begin{bmatrix} \mathbf{2}\cdots\mathbf{1} \\ \mathbf{1}\cdots\mathbf{2} \end{bmatrix} = \begin{bmatrix} \mathbf{1^{c-1}12}1^C21^C21\cdots121\mathbf{1^{c-1}2} \\ \mathbf{1^{c-1}2}1^C21^C21\cdots121\mathbf{1^{c-1}12} \end{bmatrix} \quad ,$$

where $C \in \{c-1, c\}$. If $c \geq 2$, we already reached the desired form. If $c = 1$, then a prefix of $F(\vartheta)$ has the form (26), which can be continued in the two ways (27), and then we are back to (26). Therefore, $F(\vartheta)$ is composed of two types of substrings: '()' and '(∘)'. Applying F to '(∘)' gives:

$$F\begin{bmatrix} 221 \\ 122 \end{bmatrix} = \begin{bmatrix} 12122 \\ 21212 \end{bmatrix} \rightarrow\ ()()\circ \quad .$$

Returning to the general case ($c \geq 1$), applying F to () gives:

$$F\begin{bmatrix} 12 \\ 21 \end{bmatrix} = \begin{bmatrix} 1\cdots 1\mathbf{21}\cdots 112 \\ 1\cdots 1\mathbf{12}1\cdots 12 \end{bmatrix} \rightarrow\ \bullet^{c-1}()\bullet^{c-1}\circ \quad .$$

Applying F to a dot section gives a longer dot section by the definition of F. In summary, F preserves the form (28), including the number of parentheses pairs. ■

Proof of Theorem 1. Let u, u' be synchronous generators for W, W' respectively, and let k be the smallest number such that the parentheses string of $w := F^k(u)$, $w' := F^k(u')$ is of the form (28). Let p, q be the indices such that $w = \alpha_p \cdots \alpha_q$, $w' = \alpha'_p \cdots \alpha'_q$, $\omega = \sigma_p \cdots \sigma_q$. By (10) and the notations at the beginning of Section 3, $\sigma_n = \alpha'_n - \alpha_n$. For example,

$$\begin{aligned} w &= \cdots\cdots 12 \cdots\cdots 12 \cdots\cdots \\ w' &= \cdots\cdots 21 \cdots\cdots 21 \cdots\cdots \\ \omega &= \cdots\cdots 1\bar{1} \cdots\cdots 1\bar{1} \cdots\cdots \end{aligned}$$

where the dot strings (some or all of which might be empty) stand for identical letters in w, w'; and for 0 in ω.

Thus, for all $i = p, \ldots, q$, $s_i \in \{\gamma - 1, \gamma, \gamma + 1\}$ for some γ (actually $\gamma = s_p$), namely,

- $\sigma_i = 0 \Longrightarrow s_i = s_{i+1} := \gamma$;
- $\sigma_{i-1}\sigma_i = 1\bar{1} \Longrightarrow s_{i-1} = s_{i+1} = \gamma,\ s_i = \gamma + 1$;
- $\sigma_{i-1}\sigma_i = \bar{1}1 \Longrightarrow s_{i-1} = s_{i+1} = \gamma,\ s_i = \gamma - 1$.

Let K be the number of pairs $1\bar{1}$, $\bar{1}1$ in ω. By Lemma 5, $F(\omega)$ has the same number of pairs: a pair of the form $1\bar{1}$ in ω induces the string $0^{c-1}\bar{1}10^c$ in $F(\omega)$, and a pair of the form $\bar{1}1$ induces the string $0^{c-1}1\bar{1}0^c$. If $K > 0$, this means that each of $\gamma - 1, \gamma, \gamma + 1$ must appear infinitely often, and that the partition into K sequences holds; if $K = 0$, this means that $s_n = \gamma$ for all $n > p$. Assertion 2 follows from the fact that ω consists only of the pairs $1\bar{1}$, $\bar{1}1$, separated by nonempty or empty strings of zeros. ■

Example 6. We saw in Example 3 that $v = \alpha_{13} \cdots \alpha_{31}$, $v' = \alpha'_{13} \cdots \alpha'_{31}$, are synchronous generators for W, W' of Example 1. Examining Table 1, we see that the corresponding parentheses string is $\vartheta = \bullet(\bullet \circ \bullet)()(\bullet)()()() \bullet \circ$, which has nesting level 1. This means that γ can assume only three values, as indeed it does (the values 7,8,9). However, the string doesn't have the form of (28), which means that instances of 7 or 9 appear successively. Indeed, $s_{15} = \cdots = s_{18} = 9$, and $s_{22} = s_{23} = 9$.

To get the form (28), we could take $F(v)$, $F(v')$; or else, we could start with new generators. Since for $n \geq 24$ instances of 7 or 9 do not appear successively, we can start from $n = 24$. The synchronous generators will be $u = \alpha_{24} \cdots \alpha_{57}$, $u' = \alpha'_{24} \cdots \alpha'_{57}$, since $b_{24} + 1 = 74 = a_{58}$ and $b'_{24} + 1 = 82 = a'_{58}$. The corresponding parentheses string is

$$\vartheta = ()()() \bullet \circ \bullet \circ \bullet ()() \bullet ()() \bullet \circ \bullet () \bullet \circ \bullet ()() \bullet \circ,$$

which has the desired format. There are ten parentheses pairs, which correspond to the indices $24 \leq n \leq 58$ such that $s_n \neq 8$. These indices are $25, 27, 29, 43, 46, 50, 53, 60, 67, 70$. Each of them is the first element in one of the ten sequences $V_i \in \mathcal{F}$ of irregular shifts.

4 The structure of s_n

In this section we describe the sequences of irregular shifts (Definition 4), $V_i = \{n_j^{(i)}\}_{j=1}^{\infty} \in \mathcal{F}$, $i = 1, \ldots, K$. We examine V_i for fixed i, hence we can omit i and denote the sequence by $V = \{n_j\}_{j=1}^{\infty}$.

Lemma 6. *Let $t \geq p$ (see Theorem 1), and suppose that $s_t = \gamma$. Then*

$$b_t + 1 = a_n \iff b'_t + 1 = a'_n \ \ (\textit{same } n).$$

Proof. The parentheses pattern of the $n \geq p$ - suffix of the words $\left[\begin{smallmatrix} W \\ W' \end{smallmatrix}\right]$ is $\cdots () \cdots$, where a parentheses pair comes around instances of $\gamma + 1$ or $\gamma - 1$. Therefore, $s_t = \gamma$ implies a nesting level of 0 at point t. Following the proof of Lemma 3, suppose that $a_n = b_t + 1$ and $a'_n = b'_t + 1 + k$, $k > 0$. Then the pairs $\left[\begin{smallmatrix} \alpha_i \\ \alpha'_i \end{smallmatrix}\right]$, $i = t, \ldots, n-1$ contain k more pairs of the form $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ than of the form $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$, i.e., at point n we have a nesting level of k. By the parentheses structure, this implies that $k = 1$ and $\left[\begin{smallmatrix} \alpha_{n-1} \\ \alpha'_{n-1} \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. We get:

$$\begin{aligned} a_n &= b_t + 1, \qquad & a_n - a_{n-1} &= 1; \\ a'_n &= b'_t + 2, \qquad & a'_n - a'_{n-1} &= 2. \end{aligned}$$

Thus $a_{n-1} = b_t$ and $a'_{n-1} = b'_t$, a contradiction. ■

Lemma 7. *Let $\{n_j\}_{j=1}^{\infty}$ be a sequence of irregular shifts. Then for all $j \geq 1$:*

$$\begin{array}{rclcrclcrcl} s_{n_j} & = & \gamma+1 & \Longrightarrow & a_{n_{j+1}} & = & b_{n_j}+1, & a'_{n_{j+1}} & = & b'_{n_j}-1; \\ s_{n_j} & = & \gamma-1 & \Longrightarrow & a_{n_{j+1}} & = & b_{n_j}-1, & a'_{n_{j+1}} & = & b'_{n_j}+1. \end{array}$$

Proof. Let $t = n_j$, $m = n_{j+1}$. Suppose first that $s_t = \gamma + 1$. Then Theorem 1 implies $s_{t-1} = s_{t+1} = \gamma$. Thus $\alpha_{t-1} = 1$, $\alpha'_{t-1} = 2$, $\alpha_t = 2$, $\alpha'_t = 1$. Now $F(\alpha_{t-1}) = 1^{c-1}2$, which are the letters in $[b_{t-1}+1, b_t+1]$. Further, $F(\alpha_t) = 1^c2$, the letters in $[b_t+1, b_{t+1}+1]$ (Lemma 2). The same holds for α'_{t-1}, α'_t with 1^c2, $1^{c-1}2$ respectively. By Lemma 6, $b_{t-1}+1$ and $b'_{t-1}+1$ occur at the same index in the A, A' sequences, therefore the picture is:

$$\begin{array}{lccccccccccccl} b_{t-1}+1 & & & & & & b_t+1 & & & & & & & b_{t+1}+1 \\ & 1 & \cdot & \cdot & 1 & 2 & & 1 & & 1 & \cdot & \cdot & 1 & 2 \\ & 1 & \cdot & \cdot & 1 & 1 & & 2 & & 1 & \cdot & \cdot & 1 & 2 \\ b'_{t-1}+1 & & & & & & & & b'_t+1 & & & & & b'_{t+1}+1 \end{array}$$

The next non-regular shift $s_m = \gamma - 1$ is achieved at the index of $b_t + 1$, i.e., $b_t + 1 = a_m$. Therefore, a'_m is the element of A' that precedes $b'_t + 1$, which is $b'_t + 1 - 2 = b'_t - 1$.

The same argument is used for proving the case $s_t = \gamma - 1$. ■

Example 7. We refer back to Table 1. The first sequence of irregular shifts starts at $n = 25$, where $s_{25} = 9 = \gamma + 1$; its next element is at $n = 60$, since

$$\begin{array}{rclclcl} F(\alpha_{24}\alpha_{25}) & = & F(12) & = & 12112 & = & \alpha_{58}\cdots\alpha_{62}, \\ F(\alpha'_{24}\alpha'_{25}) & = & F(21) & = & 11212 & = & \alpha'_{58}\cdots\alpha'_{62}. \end{array}$$

We see that indeed,

$$\begin{array}{rclclcl} a_{n_2} & = & a_{60} & = & 77 & = & b_{n_1}+1, \\ a'_{n_2} & = & a'_{60} & = & 84 & = & b'_{n_1}-1. \end{array}$$

Lemma 8. *Let* $m_0 = \min\{m : a_m > \max(X)\}$ (*see* (12a)). *Then for all $n > t \geq m_0$:*

$$b_t + 1 = a_n \Longrightarrow b_n - 1 = a_{cn+t}\ ;$$
$$b_t - 1 = a_n \Longrightarrow b_n + 1 = a_{cn+t}\ .$$

Proof. Let $a_n = b_t+1$, and let $u = \alpha_t \cdots \alpha_{n-1}$. Then $F(u)$ is the difference word between $b_t + 1$ and $b_n + 1$. By Propositions 3, 4,

$$\begin{aligned} |F(u)| &= c|u|_1 + (c+1)|u|_2 \\ &= c[2n-(c+2)t-1] + (c+1)[(c+1)t-n+1] \\ &= nc + t + 1 - n. \end{aligned} \tag{29}$$

Therefore, if $b_t + 1 = a_n$, then $b_n + 1 = a_{n+nc+t+1-n} = a_{nc+t+1}$. The element of A previous to $b_n + 1$ is $b_n - 1$, since $b_n \in B$, therefore $b_n - 1 = a_{nc+t}$.

Suppose now that $a_n = b_t - 1$. Then $a_{n+1} = b_t + 1$, and the difference word from a_t to $b_t + 1$ is $u = \alpha_t \cdots \alpha_n$. Therefore, if we put $n+1$ instead of n in (29), we get that the difference word between $b_t + 1$ and $b_{n+1} + 1$ satisfies:

$$|F(u)| = nc + t + c - n.$$

Since $b_t + 1 = a_{n+1}$, $b_{n+1} + 1 = a_{nc+t+c+1}$.

Now, $a_{n+1} - a_n = (b_t + 1) - (b_t - 1) = 2$, therefore $b_{n+1} - b_n = c + 2$, and it takes $c + 1$ elements of A to get from $b_n + 1$ to $b_{n+1} + 1$. We get: $b_n + 1 = a_{(nc+t+c+1)\,-\,(c+1)} = a_{nc+t}$. ■

Example 8. For Example 1 (Table 1) we have $m_0 = 9$, as $a_9 = 6 = \max(\{1, 5\}) + 1$. Take, for example, $t = 9$. We get:

$$b_9 + 1 = 25 = a_{24}; \quad b_{24} - 1 = 72 = a_{57}; \quad 57 = 24 \cdot 2 + 9 = nc + t.$$

The same holds for any indices $n > t \geq 9$ we pick.

Proof of Theorem 2. Let $V = \{n_j\}_{j=1}^{\infty} \in \mathcal{F}$ be a sequence of irregular shifts. Let $j > 2$, and suppose $s_{n_{j-2}} = \gamma + 1$. Then by Theorem 1,

$$\begin{aligned} s_{n_{j-1}} &= \gamma - 1, \\ s_{n_j} &= \gamma + 1. \end{aligned}$$

Therefore, by Lemma 7,

$$\begin{aligned} a_{n_{j-1}} &= b_{n_{j-2}} + 1, \\ a_{n_j} &= b_{n_{j-1}} - 1. \end{aligned}$$

But by Lemma 8,

$$b_{n_{j-2}} + 1 = a_{n_{j-1}} \Longrightarrow b_{n_{j-1}} - 1 = a_{cn_{j-1}+n_{j-2}}.$$

We get:

$$a_{n_j} = a_{cn_{j-1}+n_{j-2}}.$$

Since no element of A appears more than once, necessarily $n_j = cn_{j-1} + n_{j-2}$. The same holds if $s_{n_{j-2}} = \gamma - 1$. ■

The recurrence described in Theorem 2 leads to another representation of $\{n_j^{(i)}\}_{j=1}^{\infty}$: that of *continued fraction numeration system.*

4.1 The continued fraction numeration system.

Let α be an irrational number satisfying $1 < \alpha < 2$, let $[1, a_1, a_2, a_3, \ldots]$ be it's unique expansion into a simple continued fraction, so $a_i \in \mathbb{Z}_{\geq 1}$ for all $i \geq 1$, and let $p_n/q_n = [1, a_1, \ldots, a_n]$ be its convergents. Then p_n, q_n satisfy:

$$p_{-1} = 1,\ p_0 = 1,\ p_n = a_n p_{n-1} + p_{n-2}\ \ (n \geq 1),$$

$$q_{-1} = 0,\ q_0 = 1,\ q_n = a_n q_{n-1} + q_{n-2}\ \ (n \geq 1),$$

(Olds [21]). Fraenkel [10], [12] showed that every positive integer m can be written uniquely in the form

$$m = \sum_{i=0}^{k} r_i p_i;\ \ 0 \leq r_i \leq a_{i+1},\ r_{i+1} = a_{i+2} \Longrightarrow r_i = 0\ (i \geq 0),$$

and also in the form

$$m = \sum_{i=0}^{k} t_i q_i;\ 0 \leq t_0 < a_1,\ 0 \leq t_i \leq a_{i+1},\ t_i = a_{i+1} \Longrightarrow t_{i-1} = 0\ (i \geq 1).$$

These numeration systems are called the *p-system* and *q-system* respectively. For the special case $a_i = 1$ for all i, where $\alpha = \varphi$ is the golden section, the two numeration systems coalesce, and have been known as the Zeckendorf numeration system [27]. If we set $\alpha = (2 - c + \sqrt{c^2 + 4})/2$, then $a_i = c$ for all i, and the convergents satisfy the same recurrence as the sequences of irregular shifts.

Definition 8. (Fraenkel [10].) The *p-representation* of a positive integer n, relative to a simple continued fraction $\alpha = [1, a_1, a_2, \ldots]$, is the $(m+1)-$tuple

$$R_p(n) = (r_m, r_{m-1}, \ldots, r_1, r_0),$$

where $\sum_{i=0}^{m} r_i p_i$ is the unique representation of n in the p-system.

The *q-representation* of n is the $(k+1)-$tuple

$$R_q(n) = (t_k, t_{k-1}, \ldots, t_1, t_0),$$

where $\sum_{i=0}^{k} t_i q_i$ is the unique representation of n in the q-system.

An $(m+1)-$tuple $R = (d_m, d_{m-1}, \ldots, d_1, d_0)$ is a *representation* if it satisfies the coefficient conditions:

$$0 \leq d_i \leq a_{i+1},\ d_{i+1} = a_{i+2} \Longrightarrow d_i = 0\ (i \geq 0).$$

If we know that for some $k \leq m$, $d_k > 0$ and $d_i = 0$ for all $i < k$, we also write $R = (d_m, d_{m-1}, \ldots, d_k)$.

If $R_q(n)$ ends in an even number of zeros (i.e., $R_q(n) = (d_m, d_{m-1}, \ldots, d_{2k})$, $k \geq 0$, $d_{2k} > 0$), we say that $n \in Q_0$. Otherwise, If $R_q(n)$ ends in an odd number of zeros (i.e., $R_q(n) = (d_m, d_{m-1}, \ldots, d_{2k+1})$, $k \geq 0$, $d_{2k+1} > 0$), we say that $n \in Q_1$. Similarly, we define P_0, P_1.

The *p-interpretation* of a representation $R = (d_m, d_{m-1}, \ldots, d_1, d_0)$ is the number

$$I_p(R) = \sum_{i=0}^{m} d_i p_i.$$

The *q-interpretation* of R is the number

$$I_p(R) = \sum_{i=0}^{m} d_i q_i,$$

provided that $d_0 < a_1$; otherwise, R has no q-interpretation.

The *shift left* operator (denoted by shl), when applied to a representation $R = (d_m, d_{m-1}, \ldots, d_1, d_0)$, shifts all the bits leftwards one position, and inserts 0 on the right:

$$\mathsf{shl}(R) = (d_m, d_{m-1}, \ldots, d_1, d_0, 0).$$

Note that, for a non-decreasing sequence of partial quotients $1 \leq a_1 \leq a_2 \leq \cdots$, if R is a representation then so is $\mathsf{shl}(R)$. Moreover, $\mathsf{shl}(R)$ always has a q-interpretation, even if R itself doesn't.

4.2 Representing the irregular shifts in the q-system

Let $V = \{n_j\}_{j=1}^{\infty} \in \mathcal{F}$ be a sequence of irregular shifts, and let $R_q(n_j)$ be the representation of n_j in the q-system. Suppose that for some j_0, $R_q(n_{j_0+1}) = \mathsf{shl}(R_q(n_{j_0}))$. Then, since $\{q_j\}$ and $\{n_j\}$ satisfy the same recurrence, we get that $R_q(n_{j+1}) = \mathsf{shl}(R_q(n_j))$ for all $j \geq j_0$:

$$\begin{aligned}
n_{j_0} &= d_0 q_0 + d_1 q_1 + \cdots + d_k q_k \;; \\
n_{j_0+1} &= d_0 q_1 + d_1 q_2 + \cdots + d_k q_{k+1} \;; \\
n_{j_0+2} &= c n_{j_0+1} + n_{j_0} \\
&= d_0(c q_1 + q_0) + d_1(c q_2 + q_1) + \cdots + d_k(c q_{k+1} + q_k) \\
&= d_0 q_2 + d_1 q_3 + \cdots + d_k q_{k+2} = \mathsf{shl}(R_q(n_{j_0+1})) \;.
\end{aligned}$$

To show that this is indeed the case, we use the following theorem:

Theorem 4. (Fraenkel [10].) *Let* $\alpha = (2-c+\sqrt{c^2+4})/2$, *and let* $p_n/q_n = [1, c^n]$ *be the convergents of* α. *Then:*

1. *For all* $j \ge 1$*:*
   - $j \in P_0 \Longleftrightarrow j = a'_n$ *for some* $n \ge 1$*;*
   - $j \in P_1 \Longleftrightarrow j = b'_n$ *for some* $n \ge 1$*.*
2. *For all* $n \ge 1$, $R_p(b'_n) = \text{shl}(R_p(a'_n))$, *i.e.,*
$$a'_n = \sum_{i=k}^{m} d_i p_i \Longrightarrow b'_n = \sum_{i=k+1}^{m+1} d_{i-1} p_i.$$
3. *For all* $n \ge 1$*:*
   - $n \in Q_0 \Longrightarrow a'_n = I_p(R_q(n))$*;*
   - $n \in Q_1 \Longrightarrow a'_n = I_p(R_q(n)) - 1$*.*

   Thus,
$$n = \sum_{i=2k}^{m} d_i q_i \quad \Longrightarrow \quad a'_n = \sum_{i=2k}^{m} d_i p_i, \tag{30}$$
$$n = \sum_{i=2k+1}^{m} d_i q_i \quad \Longrightarrow \quad a'_n = \sum_{i=2k+1}^{m} d_i p_i - 1. \tag{31}$$

(30), (31) have inverses, as we show now.

Lemma 9. *Let* $n \in P_1$, *and let*
$$R_p(n) = (d_m, \ldots, d_{2k+1}) = (d_m, \ldots, d_{2k+1}, 0^{2k+1}).$$
Then $n-1 \in P_0$, *and*
$$R_p(n-1) = (d_m, \ldots, d_{2k+2}, d_{2k+1}-1, c, (0,c)^k).$$

Proof. Fraenkel ([10]) showed that if $\alpha = [1, a_1, a_2, \ldots]$ is an irrational number, then
$$p_{2k+1} - 1 = a_{2k+1} p_{2k} + a_{2k-1} p_{2k-2} + \cdots + a_1 p_0.$$
Let $\alpha = (2-c+\sqrt{c^2+4})/2$.
$$\begin{aligned} n-1 &= (d_m p_m + \cdots + (d_{2k+1}-1) p_{2k+1}) + (p_{2k+1} - 1) \\ &= (d_m p_m + \cdots + (d_{2k+1}-1) p_{2k+1}) + c p_{2k} + c p_{2k-2} + \cdots + c p_0. \end{aligned}$$
■

Lemma 10. *$R_p(a'_n)$ ends with $d_0 = c$ iff $n \in Q_1$.*

Proof. Let $R_p(a'_n)$ end with c, and suppose that $n \in Q_0$. Then $R_q(n) = (d_m, \ldots, d_{2k})$, so by Theorem 4, $R_p(a'_n) = (d_m, \ldots, d_{2k})$. But by the conditions on the q-representation, $R_q(n)$ cannot end with c, so $R_p(a'_n)$ doesn't end with c. We get two different representations of a'_n in the p-system, a contradiction to uniqueness.

Let now $n \in Q_1$, $n = \sum_{i=2k+1}^{m} d_i q_i$. Then $a'_n = \sum_{i=2k+1}^{m} d_i p_i - 1$ by Theorem 4, so by Lemma 9 $R_p(a'_n)$ ends with $d_0 = c$. ∎

Corollary 6.

$$a'_n = \sum_{i=2k}^{m} d_i p_i,\ d_0 < c \implies n = \sum_{i=2k}^{m} d_i q_i, \tag{32}$$

$$a'_n = \sum_{i=2k+1}^{m} d_i p_i - 1 \implies n = \sum_{i=2k+1}^{m} d_i q_i. \tag{33}$$

Proof. Let $a'_n = \sum_{i=2k}^{m} d_i p_i$, $d_0 < c$. Then by Lemma 10, $n \in Q_0$, $n = \sum_{i=2x}^{y} e_i q_i$, $e_0 < c$. Therefore, by Theorem 4, $a'_n = \sum_{i=2x}^{y} e_i p_i$, and by the uniqueness of representation in the p-system, $(d_m, \ldots, d_{2k}) = (e_y, \ldots, e_{2x})$, and $n = \sum_{i=2k}^{m} d_i q_i$.

Let now $a'_n = \sum_{i=2k+1}^{m} d_i p_i - 1$. Then by Lemma 9, $R_p(a'_n)$ ends with c, so by Lemma 10, $n \in Q_1$, $n = \sum_{i=2x+1}^{y} e_i q_i$. Therefore, by Theorem 4, $a'_n = \sum_{i=2x+1}^{y} e_i p_i - 1$, and by uniqueness of representation we get $(d_m, \ldots, d_{2k+1}) = (e_y, \ldots, e_{2x+1})$, and $n = \sum_{i=2k+1}^{m} d_i q_i$. ∎

Lemma 11. *Let $\{n_j\}_{j=1}^{\infty}$ be a sequence of irregular shifts, and let $R_q(n_j)$ be the representation of n_j in the q-system. Suppose that $s_{n_1} = \gamma + 1$, otherwise start from the second element. Then $R_q(n_2) = \mathsf{shl}(R_q(n_1))$, or $R_q(n_3) = \mathsf{shl}(R_q(n_2))$.*

Proof. Let $x = n_1$, $y = n_2$, $z = n_3$. By Lemma 7, $a'_y = b'_x - 1$, $a'_z = b'_y + 1$.

Case 1: $x \in Q_0$.

$$x = \sum_{i=2k}^{m} d_i q_i \implies a'_x = \sum_{i=2k}^{m} d_i p_i \implies b'_x = \sum_{i=2k+1}^{m+1} d_{i-1} p_i \implies$$
$$a'_y = \sum_{i=2k+1}^{m+1} d_{i-1} p_i - 1 \implies y = \sum_{i=2k+1}^{m+1} d_{i-1} q_i = \mathsf{shl}(x).$$

Case 2: $x \in Q_1$.

$$\begin{aligned}
x &= \sum_{i=2k+1}^{m} d_i q_i \Longrightarrow \\
a'_x &= \sum_{i=2k+1}^{m} d_i p_i - 1 = \sum_{i=2k+2}^{m} d_i p_i + (d_{2k+1} - 1) p_{2k+1} + c p_{2k} + \cdots + c p_0 \Longrightarrow \\
b'_x &= \sum_{i=2k+3}^{m+1} d_{i-1} p_i + (d_{2k+1} - 1) p_{2k+2} + c p_{2k+1} + c p_{2k-1} + \cdots + c p_1 \Longrightarrow \\
a'_y &= b'_x - 1 = \sum_{i=1}^{n} e_i p_i - 1,\ e_1 = c \Longrightarrow y = \sum_{i=1}^{n} e_i q_i = \sum_{i=2}^{n} e_i q_i + c q_1.
\end{aligned}$$

$$\begin{aligned}
a'_y &= \sum_{i=2}^{n} e_i p_i + c p_1 - 1 = \sum_{i=2}^{n} e_i p_i + (c-1) p_1 + c p_0 \Longrightarrow \\
b'_y &= \sum_{i=3}^{n+1} e_{i-1} p_i + (c-1) p_2 + c p_1 \Longrightarrow \\
a'_z &= b'_y + 1 = \sum_{i=3}^{n+1} e_{i-1} p_i + (c-1) p_2 + c p_1 + p_0 = \sum_{i=3}^{n+1} e_{i-1} p_i + c p_2 \Longrightarrow \\
z &= \sum_{i=3}^{n+1} e_{i-1} q_i + c q_2 = \mathsf{shl}(y).
\end{aligned}$$

■

Proof of Theorem 3. Follows directly from Lemma 11. ■

4.3 Computing s_n

We can assume that for all the sequences of irregular shifts, $\{n_j^{(i)}\}_{j=1}^{\infty},\ i = 1, \ldots, K,\ R_q(n_2^{(i)}) = \mathsf{shl}(R_q(n_1^{(i)}))$. Otherwise, we start from the third element.

• To trace a sequence of irregular shifts, compute the q-representation of the first element, and shift left.

• Given some $m \geq 1$, use the following algorithm to compute s_m:

Table 2: $c = 1$, $\alpha = \varphi$, $n_0 = 3$, $X = \{1, 2, 4\}$.

n	a_n	b_n	s_n	a'_n	b'_n
3	3	6	1	4	7
4	5	9	1	6	10
5	7	12	1	8	13
6	8	14	1	9	15
7	10	17	1	11	18
8	11	19	1	12	20
9	13	22	1	14	23
10	15	25	1	16	26

1. Compute the first element of each sequence of irregular shifts, $n_1^{(1)}, \ldots, n_1^{(K)}$. Let $M = \max(n_1^{(1)}, \ldots, n_1^{(K)})$.
2. Compute the q-representation of m.
   - If $R_q(m) = R_q(n_1^{(i)})0^t$ for some $i = 1, \ldots, K$, then $s_m = \gamma + 1$ for even t, and $s_m = \gamma - 1$ for odd t.
   - If $R_q(m) \neq R_q(n_1^{(i)})0^t$ for all $i = 1, \ldots, K$ and $n > M$, then $s_m = \gamma$.
   - If $R_q(m) \neq R_q(n_1^{(i)})0^t$ for all $i = 1, \ldots, K$ and $n < M$, then s_m must be computed manually.

4.4 Examples

We end this section with a few examples.

Example 9. $c = 1$, $\alpha = \varphi$, $n_0 = 3$, $X = \{1, 2, 4\}$. The first few elements of the sequences A, B, s_n, A', B' are listed in Table 2. Here $m_0 = 4$, since $a_4 = 5 = \max(X) + 1$, and synchronization is achieved right from the beginning: $b_4 + 1 = a_7$, $b'_4 + 1 = a'_7$. The generators u, u' satisfy $u = u' = 212$, i.e., we get zero parenthesis pairs. Therefore, there are no irregular shifts, and $s_n = \gamma = 1$ for all $n \geq 4$.

Example 10. The value of γ can be negative: for $c = 2$, $\alpha = \sqrt{2}$, $n_0 = 1$, $X = \{21, 22, 30, 31, 32, 33\}$, $s_n \in \{-4, -5, -6\}$ for all $n \geq 32$. There are $K = 38$ sequences of irregular shifts, which are indicated in Figure 3 and Figure 4. Figure 3 indicates the first 1000 elements of the sequence s_n. Figure 4 shows the distance between two successive indices of the sequence

ω of irregular shifts, of which there are about 180 in the 1000-long prefix of ω.

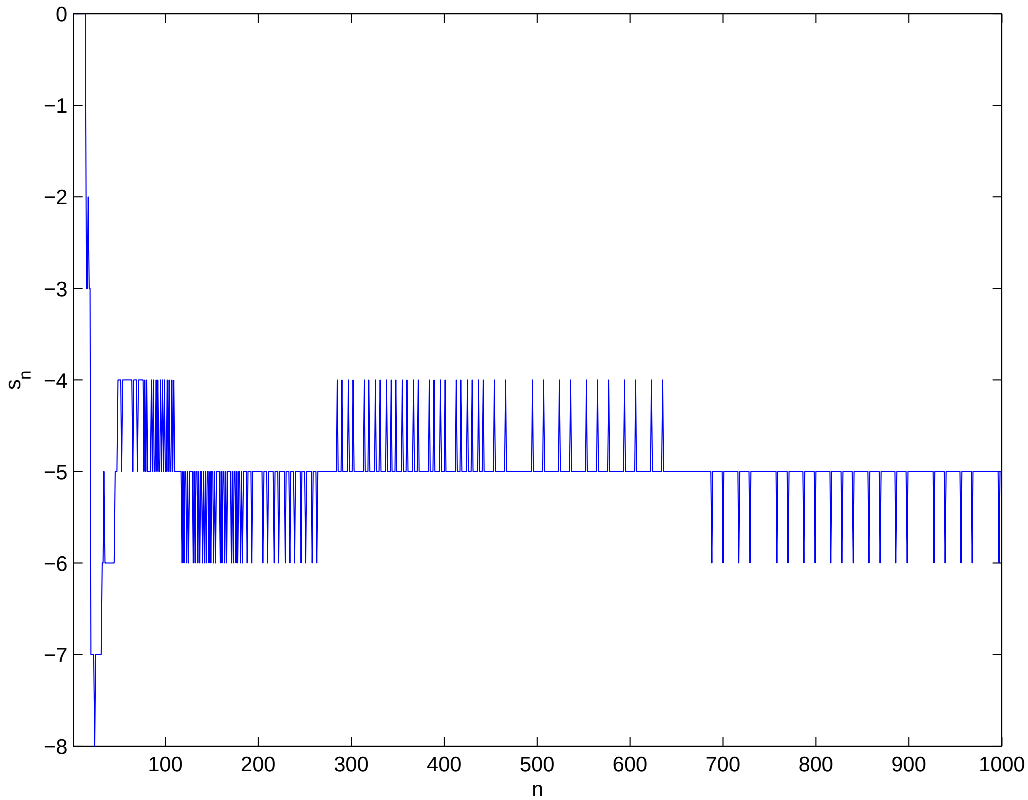

Figure 3: The first 1800 elements of the shift sequence s_n for $c = 2$, $n_0 = 1$, $X = \{21, 22, 30, 31, 32, 33\}$.

Example 11. $c = 2$, $\alpha = \sqrt{2}$, $n_0 = 25$, $X = [15, 25]$. The central shift value is $\gamma = 37$, and for $n \geq 264$ we have $36 \leq s_n \leq 38$. The number of irregular shifts sequences is $K = 108$. Figure 5 shows the distance between two successive indices of irregular shifts.

5 Motivation: Wythoff's game on N piles

Wythoff's game is an impartial 2-player game, consisting of two piles of tokens. There are two types of moves: remove any positive number of tokens from a single pile, or $k > 0$ tokens from one pile and $l > 0$ from the other, such that $|k - l| < c$, where c is a given positive integer. The case $c = 1$ (i.e., a move of the second type consists of removing the same (positive) number

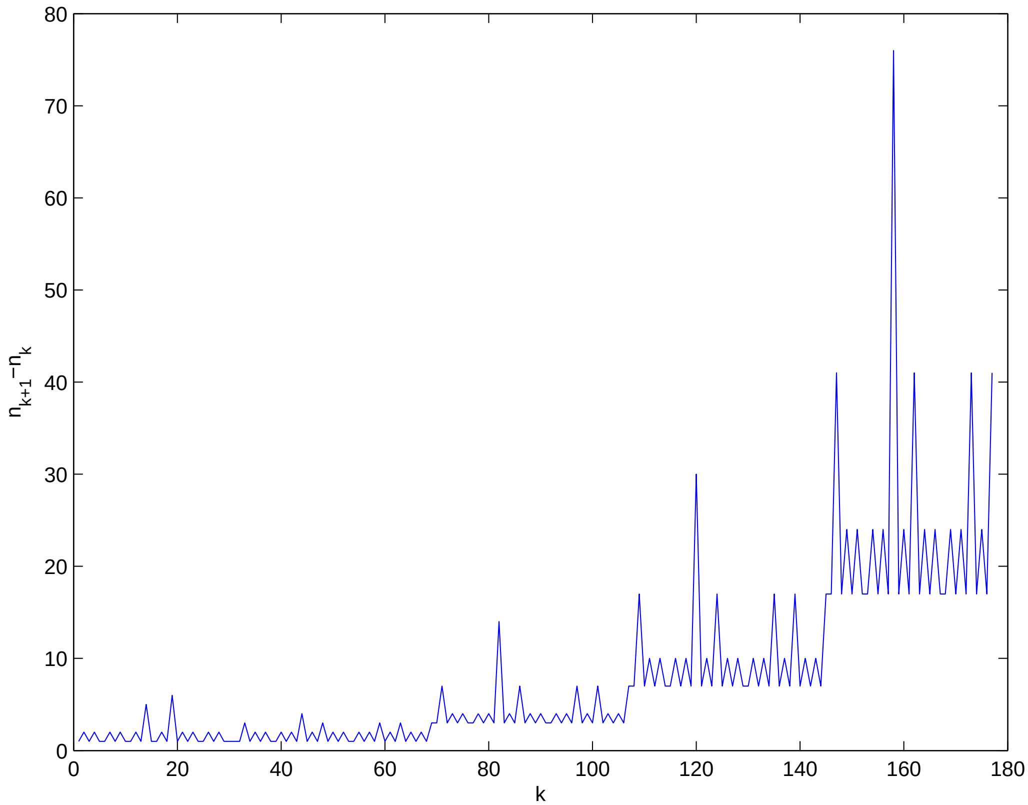

Figure 4: The distance between two successive indices of irregular shifts ($c = 2$, $n_0 = 1$, $X = \{21, 22, 30, 31, 32, 33\}$).

of tokens from both piles) is the classical Wythoff game [25]. The player first unable to move loses, and then the opponent wins.

It is useful to consider the classical Wythoff game as a generalization of the game of *nim*, played on two piles. Nim is an impartial 2-player game, played with $N \geq 1$ piles of tokens, which has only one type of move: that of removing any positive number of tokens from a single pile. A game position in this game is a N-tuple, $(p_1, \ldots, p_N) \in \mathbb{Z}_{\geq 0}^N$, where p_i is the size of pile i. The winning strategy for nim is very simple (Bouton [5]): given a game position $(p_1, \ldots, p_N)$, compute the *nim-sum* (binary xor) of the pile sizes: $\sigma = \sum_{i=1}^{\prime N} p_i$. If $\sigma > 0$, then there's a move that reduces it to 0, and the player who's making it can win. If $\sigma = 0$, then every move results in $\sigma > 0$, and the player who's making it loses. Thus, the classical Wythoff game is nim on two piles, whose move set has been augmented with the P-positions of nim!

The set $\mathcal{P}$ of P-positions for Wythoff's game is well known:

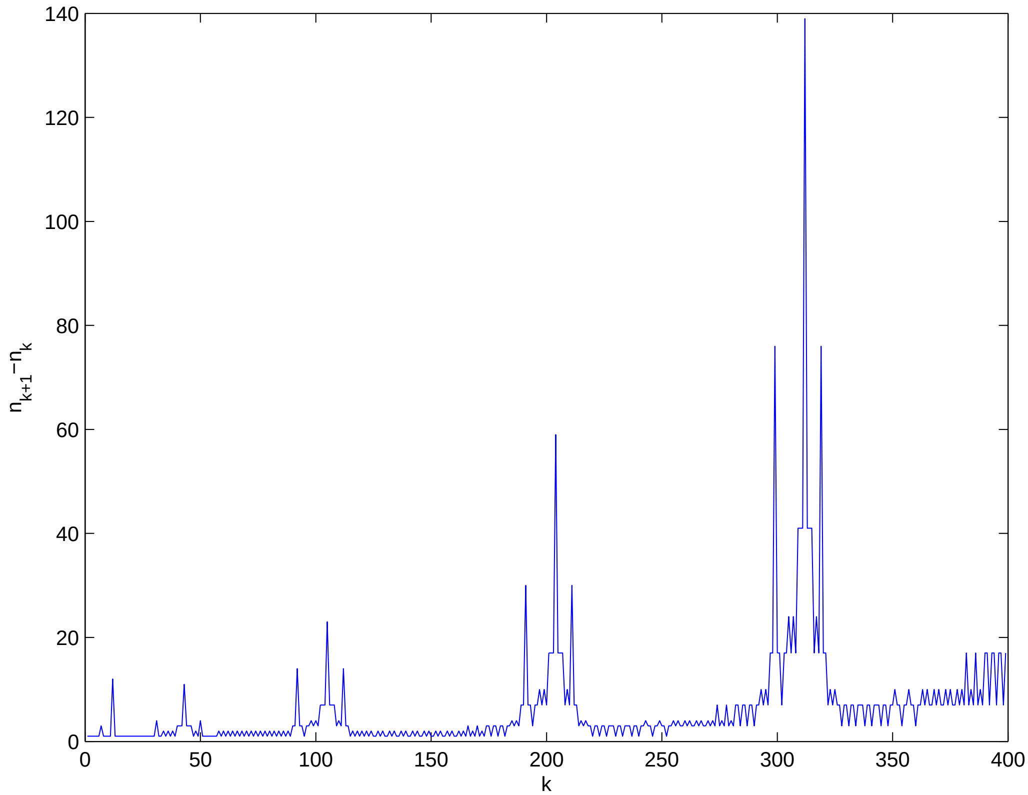

Figure 5: The distance between two successive indices of irregular shifts ($c = 2$, $n_0 = 25$, $X = \{15, 25\}$).

Theorem 5. (Fraenkel [10].) *Let* $\alpha = (2 - c + \sqrt{c^2 + 4})/2$*. Then* $\mathcal{P} = \bigcup_{n=0}^{\infty}\{(a'_n, b'_n)\}$, *where*

$$\begin{array}{rcl} a'_n & = & \text{mex}_0\{a'_i, b'_i : 0 \le i < n\} & = & \lfloor n\alpha \rfloor; \\ b'_n & = & a'_n + cn & = & \lfloor n(\alpha + c) \rfloor. \end{array}$$

The mex_0 operator is defined by $X \subsetneq \mathbb{Z}_{\ge 0} \Longrightarrow \text{mex}_0(X) = \min(\mathbb{Z}_{\ge 0} \setminus X)$ (cf (4)). Thus, the P-positions consist of the pairs (a'_n, b'_n) (Definition 1) together with $(0, 0)$.

We would now like to generalize the classical Wythoff game to more than two piles. In view of the above observation, a natural generalization would be to add the P-positions of nim as moves: The game is played on N piles. A move of the first type consists of taking any positive number of tokens from a single pile; a move of the second type consists of taking a_i tokens from pile i, such that at least two of the a_i are positive, and $\sum_{i=1}^{\prime N} a_i = 0$. This generalization and two conjectures were suggested by Fraenkel [13], [14], mentioned as a research problem in Guy and Nowakowski [17].

One might regard this generalization as unnecessarily complicated. A much simpler generalization of the second move seems to be to take the same number of tokens from each pile, or from any nonempty subcollection of piles, or to take the same number from any pair of piles, etc. Such generalizations have been considered often in the literature, see e.g., Dress, Flammenkamp and Pink [8]. But they are not always interesting, as we show now. Moreover, none of them preserves the central role of the golden section in the strategy.

5.1 How far can nim in disguise be stretched?

Suppose we consider an abstract subtraction game Γ, aimed at finding necessary and sufficient conditions so that Γ will preserve the strategy of nim. Blass, Fraenkel and Guelman investigated this question in [4]. Let $(a_1, \ldots, a_N)$ be a fixed N-tuple of nonnegative integers (not necessarily distinct), such that at least two of the a_i are positive. Let $(p_1, \ldots, p_N)$ be a game position such that $p_i \geq a_i$ for all i. There are two types of moves: a nim-move, i.e., removing a positive number of tokens from a single pile; or removing a_i tokens from pile i, $i = 1, \ldots, N$. The resulting game position after the second type of move is $(p_1 - a_1, \ldots, p_N - a_N)$.

The strategy question for Γ can be divided into three cases:

1. $\sum_{i=1}^{\prime N} p_i \neq \sum_{i=1}^{\prime N}(p_i - a_i)$ for all game positions $(p_1, \ldots, p_N)$, $p_i \geq a_i$. In this case, the strategy remains that of nim, since the nim-sum of a game position is different from the nim-sum of any of its followers. Hence the nim sum is the Sprague-Grundy function of Γ.

2. $\sum_{i=1}^{\prime N} p_i = \sum_{i=1}^{\prime N}(p_i - a_i) := R$ for some $p_i \geq a_i$, $i = 1, \ldots, N$. In this case, the S-G function of Γ cannot be identical to the nim-sum, since there exists a position that has the same nim-sum as one of its followers. However, if R is never 0, the strategy remains that of nim, since positions with nim-sum 0 don't have followers with nim-sum 0. But the strategy of the sum of Γ with other impartial games is affected adversely, because of the harm to the S-G function of Γ.

3. $\sum_{i=1}^{\prime N} p_i = \sum_{i=1}^{\prime N}(p_i - a_i) = 0$ for some $p_i \geq a_i$, $i = 1, \ldots, N$. In this case, a position with nim-sum 0 has a follower with nim-sum 0, and the winning strategy of Γ must be different than that of nim.

Blass, Fraenkel and Guelman gave necessary and sufficient conditions for cases 2 and 3 to hold. Before we state them, we need a few definitions.

Definition 9. Let $S = \{a_1, \ldots, a_N\}$ be a multiset of nonnegative integers, and let $\sigma = \sum_{i=1}^{\prime N} a_i$. Let $a_i(b) = b_i^m b_i^{m-1} \cdots b_i^0$, $b_i^j \in \{0, 1\}$, be the binary representation of a_i, and let k be the least significant bit such that at least one of the a_i satisfies $b_i^k = 1$ (i.e., all the a_i have only 0 to the right of position k). Let $\sigma(b) = \sigma^m \cdots \sigma^0$ be the binary representation of σ.

- If $\sigma^k = 0$, then we say that S is *balanced*.
- If $\sigma^k = \sigma^{k+1} = 0$, then we say that S is *smooth*.
- If $\sigma^j = 0 \; \forall \; j$ (i.e., $\sigma = 0$), then we say that S is *even*.

Note that an even set is also smooth, and a smooth set is also balanced.

Theorem 6. (Blass, Fraenkel and Guelman [4].) *Let $S = (a_1, \ldots, a_N)$ be a multiset of nonnegative integers, with at least two $a_i > 0$. Then :*

- *Case 2 holds iff S is balanced ;*
- *Case 3 holds iff either :*
    1. *N is odd and S is balanced ;*
    2. *N is even and either : S is balanced and some a_i has 0 at position k, or S is smooth and $N \geq 4$, or S is even.*

Let us now return to Wythoff's game. Clearly, the classical game is a special case of the game described above, for which $N = 2$ and $S = \{a, a\}$ for some $a > 0$. Since S is an even multiset, case 3 holds, and the winning strategy is different from that of nim. We see, incidentally, that for $N = 2$, Wythoff's game is unique in the sense that it's the only subtraction game with a strategy different from that of nim. For example, taking $(k, k+1)$ (imbalanced) preserves the strategy of Nim and its S-G function, whereas taking $(1, 3)$ (balanced) preserves the strategy of Nim but not the nonzero values of the S-G function.

What about three piles? If $S = \{a, a, a\}$ for some $a > 0$, then $\sigma = a$. Clearly, S is an imbalanced multiset, so neither case 2 nor case 3 takes place. The winning strategy is therefore that of nim, and the same holds for any odd N.

In the first generalization, on the other hand, S is an even multiset by definition, i.e., the winning strategy is different from that of nim.

5.2 Wythoff's game on three piles

We now consider Wythoff's game on three piles. A game position is a triple (p_1, p_2, p_3) with $0 \leq p_1 \leq p_2 \leq p_3$. A move consists of either taking any number of tokens from a single pile, or taking k_i tokens from pile i, such that $k_1 \oplus k_2 \oplus k_3 = 0$. We wish to characterize the P-positions of this game.

Let $\mathcal{P}$ denote the set of P-positions of the game. A convenient way to arrange the P-positions is to fix the small pile, and to let the two big ones grow to infinity. This way the P-positions are partitioned into classes:

$$\mathcal{P} = \bigcup_{j=0}^{\infty} C_j \; ; \qquad C_j = \{(j, A_n^{(j)}, B_n^{(j)})\}_{n=0}^{\infty},$$

$$j \le A_n^{(j)} \le B_n^{(j)}, \;\; A_n^{(j)} \le A_{n+1}^{(j)} \quad \forall n \ge 0.$$

Let $D_n^{(j)} = B_n^{(j)} - A_n^{(j)}$. In [13], [14] (cited in [17]), Fraenkel made the following two conjectures:

1. For every fixed $j \ge 1$, there exists an integer $n = n_j$ and a finite set of integers $X = X^{(j)}$, such that for all $n \ge n_j$,

$$\begin{aligned} A_n^{(j)} &= \operatorname{mex}_0(X^{(j)} \cup \{A_i^{(j)}, B_i^{(j)} : 0 \le i < n\}), && \text{(34a)} \\ D_{n+1}^{(j)} &= D_n^{(j)} + 1. && \text{(34b)} \end{aligned}$$

2. For every fixed $j \ge 1$, there exist integers $p = p_j$, $\gamma = \gamma_j$ and $k = k_j$, such that for all $n \ge p_j$,

$$\begin{aligned} A_n^{(j)} &\in \{\lfloor n\varphi \rfloor - (\gamma_j + k_j), \lfloor n\varphi \rfloor - (\gamma_j + k_j - 1), \ldots, \lfloor n\varphi \rfloor - (\gamma_j - k_j)\}, \\ D_{n+1}^{(j)} &= D_n^{(j)} + 1. \end{aligned}$$

Suppose that Conjecture 1 holds. Then, in particular, (34a) holds. We shift the index n of $A_n^{(j)}$, $B_n^{(j)}$ by a fixed integer so that $D_n^{(j)} = n$ for all $n \ge n_j$. We also adjoin to $X = X^{(j)}$ the numbers $A_n^{(j)}$, $B_n^{(j)}$ for $n \in \{1, \ldots, n_j - 1\}$. Then Conjecture 1 translates to the following:

$$\begin{aligned} A_n^{(j)} &= \operatorname{mex}_0(X^{(j)} \cup \{A_i^{(j)}, B_i^{(j)} : n_j \le i < n\}), \\ B_n^{(j)} &= A_n^{(j)} + n, \;\; n \ge n_j. \end{aligned}$$

Thus, $A_n^{(j)}, B_n^{(j)}$ satisfy (8a), (8b), where $b_n' - a_n' = n$ for all $n \ge n_j$, i.e., $\beta - \alpha = 1$, and $\alpha = \varphi$ (see (6), (7) for $c = 1$). Conjecture 2 says that $A_n^{(j)} = a_n' - s_n$, where s_n assumes $2k_j + 1$ successive values for all sufficiently large n.

In this paper we have shown that Conjecture 1 implies a much stronger version of Conjecture 2 (that is, if we retain the new enumeration of the

P-positions): Theorem 1 implies that if the first conjecture holds, then the second one holds as well, with $k_j \in \{1, 0\}$ for all j; Theorems 2, 3 describe the pattern of deviations from the central value γ.

We now give a proof of (34a) and an upper bound on $B_n^{(j)}$.

Theorem 7. *For fixed $j \in \mathbb{Z}_{\geq 1}$, let $(j, A_n^{(j)}, B_n^{(j)})$ be the n-th P-position of class C_j $(n \geq 0)$. Let*

$$X^{(j)} = \{x \geq j : \exists\, 0 \leq k < j \quad \text{s.t.} \quad (k, j, x) \in \mathcal{P}\} \cup \{0, \ldots, j-1\}.$$

Then for all $n \geq 0$, $j \geq 1$,

$$\begin{aligned} A_n^{(j)} &= \operatorname{mex}_0(X^{(j)} \cup \{A_i^{(j)}, B_i^{(j)} : 0 \leq i < n\}), \\ B_n^{(j)} &\leq (j+3)A_n^{(j)} + 2j + 2. \end{aligned}$$

Proof. Note that for fixed $j \in \mathbb{Z}_{\geq 1}$, $n \neq m \Longrightarrow A_m^{(j)} \neq A_n^{(j)}, B_m^{(j)} \neq B_n^{(j)}$, otherwise one P-position could be moved to from another. For convenience of notation, let

$$U_n^{(j)} = \{A_i^{(j)}, B_i^{(j)} : 0 \leq i < n\}.$$

First, observe that $X^{(j)}$ is a finite set: for every $0 \leq k < j$ there is at most one x such that $(k, j, x) \in \mathcal{P}$, since two positions that differ only in one pile, one of them is a follower of the other. Therefore, $|X^{(j)}| \leq 2j$.

Secondly, observe that $A_n^{(j)} \notin X^{(j)}$ for all $n \geq 0$: by the definition of $X^{(j)}$, $A_n^{(j)} \geq j$ for all $n \geq 0$, so in particular, $A_n^{(j)} \notin \{0, \ldots, j-1\}$. Therefore, if $A_n^{(j)} \in X^{(j)}$ for some $n \geq 0$, then $(k, j, A_n^{(j)}) \in \mathcal{P}$ for some $0 \leq k < j$. Thus we can move from P-position $(j, A_n^{(j)}, B_n^{(j)})$ to P-position $(k, j, A_n^{(j)})$ by reducing $B_n^{(j)}$ to k, a contradiction.

Finally, for arbitrary but fixed $n \geq 0$, we have $A_n^{(j)} \notin U_n^{(j)}$: (i) $i \neq n \Longrightarrow A_i^{(j)} \neq A_n^{(j)}$ for all $0 \leq i < n$. (ii) If $A_n^{(j)} = B_i^{(j)} := w$ for some $0 \leq i < n$, then $(j, A_i^{(j)}, w)$, $(j, w, B_n^{(j)})$ are both P-positions. But $B_n^{(j)} \neq A_i^{(j)}$, since $i < n$ implies that the i-th and the n-th P-positions are distinct. Therefore we can move from one of them to the other, a contradiction.

Let now $y = \operatorname{mex}_0(X^{(j)} \cup U_n^{(j)})$, and consider game positions of the form (j, y, z), where j is fixed and $z \geq y$. To show that $A_n^{(j)} = y$, it suffices to show that there exists $z \geq y$ such that (j, y, z) is a P-position; i.e., it has no P-followers. We do this by showing that there is only a finite set R of z values such that position (j, y, z) has a P-follower of a form different from (j, y, w), $w \geq y$. Therefore, if we let $T = R \cup \{0, \ldots, y-1\}$ and put $z_0 = \operatorname{mex}_0(T)$, we get $(j, y, z_0) \in \mathcal{P}$.

The followers of position (j, y, z), where $j \geq 1$ is fixed, $y = \operatorname{mex}_0(X^{(j)} \cup U_n^{(j)})$ and $z \geq y$ is arbitrary, are of the following four types:

1. $(j, y, z-k)$, $0 < k \le z$. Since $y \notin U_n^{(j)}$, there are no P-positions of the form $(j, z-k, y)$, $j \le z-k \le y$; otherwise, we would have $y = B_i^{(j)}$ for some $i < n$. Since $y \notin X^{(j)}$ there are no P-positions of the form $(z-k, j, y)$, $z-k \le j \le y$. Finally, a P-position of the form $(j, y, z-k)$, $j \le y \le z-k$ means that $A_n^{(j)} = y$, as required.

2. $(j-k, y, z)$, $0 < k \le j$. For every $k \in \{1, \dots, j\}$ there is at most one $z \ge y$ such that $(j-k, y, z)$ is a P-position, because for any $z_1 > z_2 \ge y$, $(j-k, y, z_2)$ is a follower of $(j-k, y, z_1)$. Denote the z-value of this P-position, if any, by u_k. In order to get from (j, y, z) to the P-position $(j-k, y, u_k)$, if it exists, using the move $j-k$, necessarily $z = u_k$. Therefore, any position of the form (j, y, z), $z \notin \{u_1, \dots, u_j\}$, doesn't have a P-follower via the move $j-k$.

3. $(j, y-k, z)$, $0 < k \le y$. Again, for every $k \in \{1, \dots, y\}$ there is at most one $z \ge y-k$ such that $(j, y-k, z)$ is a P-position. Denote the corresponding value by v_k, and choose $z \notin \{v_1, \dots, v_y\}$. Then (j, y, z) doesn't have a P-follower using the move $y-k$.

4. $(j-k_1, y-k_2, z-k_3)$, $\;k_1 \oplus k_2 \oplus k_3 = 0$, i.e., $k_3 = k_1 \oplus k_2$.

   For every pair (k_1, k_2), $0 \le k_1 \le j$, $0 \le k_2 \le y$, there is at most one w such that $(j-k_1, y-k_2, w)$ is a P-position. Denote the corresponding value by w_{k_1,k_2}. In order to get from position (j, y, z) to the P-position $(j-k_1, y-k_2, z-(k_1 \oplus k_2))$, we must have $z = w_{k_1,k_2} + (k_1 \oplus k_2)$. Choose $z \notin \{w_{k_1,k_2} + (k_1 \oplus k_2) : 0 \le k_1 \le j,\ 0 \le k_2 \le y\}$. Then (j, y, z) doesn't have a P-follower through this type of move.

Thus, the set T has the form,

$$T = \{u_1, \dots, u_j\} \cup \{v_1, \dots, v_y\} \cup \\ \{w_{k_1,k_2} + (k_1 \oplus k_2) : 0 \le k_1 \le j, 0 \le k_2 \le y\} \cup \{0, \dots, y-1\}.$$

Note that actually only those u_k are in T for which $(j-k, y, u_k) \in \mathcal{P}$. Analogously for the terms v_i and w_{k_1,k_2}. We see that T is finite, and

$$|T| \le j + y + (j+1)(y+1) + y = (j+3)y + 2j + 1.$$

Letting $z_0 = \mathrm{mex}_0(T)$, we see that position (j, y, z_0) has no P-followers, so it is a P-position, and $A_n^{(j)} = y$, $B_n^{(j)} = z_0$. Moreover, $z_0 \le |T| + 1$, so $B_n^{(j)} \le (j+3)A_n^{(j)} + 2j + 1$. ■

Table 3: The first P-positions of class C_1.

n	$A_n^{(1)}$	$B_n^{(1)}$	$D_n^{(1)}$	n	$A_n^{(1)}$	$B_n^{(1)}$	$D_n^{(1)}$
0	1	1	0	27	41	68	27
1	3	4	1	28	42	70	28
2	5	9	4	29	44	73	29
3	6	12	6	30	45	75	30
4	7	14	7	31	46	77	31
5	8	11	3	32	47	79	32
6	10	18	8	33	48	81	33
7	13	22	9	34	50	84	34
8	15	20	5	35	53	88	35
9	16	28	12	36	54	90	36
10	17	27	10	37	55	92	37
11	19	30	11	38	57	95	38
12	21	36	15	39	59	98	39
13	23	39	16	40	61	101	40
14	24	37	13	41	62	103	41
15	25	43	18	42	63	105	42
16	26	40	14	43	65	108	43
17	29	49	20	44	66	110	44
18	31	52	21	45	67	112	45
19	32	51	19	46	69	115	46
23	**33**	**56**	**23**	47	71	118	47
24	34	58	24	48	72	120	48
25	35	60	25	49	74	123	49
26	38	64	26	50	76	126	50

Example 12. Table 3 depicts the first 48 P-positions of class C_1. The set $X^{(1)}$ is given by:

$$X^{(1)} = \{x \geq 1 : (0, 1, x) \in \mathcal{P}\} \cup \{0\} = \{x \geq 1 : (1, x) \in \mathcal{P}\} \cup \{0\} = \{0, 2\}.$$

Indeed, we see that $A_0^{(1)} = \text{mex}_0(X^{(1)}) = 1$, $A_1^{(1)} = \text{mex}_0(X^{(1)} \cup \{1\}) = 3$, $A_2^{(1)} = \text{mex}_0(X^{(1)} \cup \{1, 3, 4\}) = 5$, and so on. The bound $B_n^{(1)} \leq 4A_n^{(1)} + 3$ also holds.

For $j = 1$ we have proved the validity of (34b) for $n_1 = 23$, but we omit the proof here. If we increase the indices of $A_n^{(j)}$, $B_n^{(j)}$ by 3 we thus get,

$$B_n^{(1)} = A_n^{(1)} + n \quad \forall\, n \geq 23.$$

This is shown in boldface in Table 3. In order to satisfy (8a), (8b), we adjoin to $X^{(1)}$ the numbers that appeared in the first 20 P-positions:

$$X^{(1)} = \{0, \ldots, 32, 36, 37, 39, 40, 43, 49, 51, 52\}.$$

Then we get:

$$\begin{aligned} A_n^{(1)} &= \text{mex}_0(X^{(1)} \cup \{A_i^{(1)}, B_i^{(1)} : 23 \leq i < n\}), \\ B_n^{(1)} &= A_n^{(1)} + n; \quad n \geq 23. \end{aligned}$$

The question of the distance $B_n - A_n$ is still open, but computer tests done for the 4-pile case suggest that here, too, at some point the distance starts growing in jumps of 1. These sequences were the motivation for this work.

5.3 Wythoff's game on N piles

In the same manner as in the 3-pile case, we fix the first $N - 2$ piles, and let the two largest ones grow to infinity. The set $\mathcal{P}$ is thus partitioned into classes:

$$C_{j_1, \ldots, j_{N-2}} = \{(j_1, \ldots, j_{N-2}, A_n, B_n)\}_{n=0}^{\infty}; \quad 0 \leq j_1 \leq \cdots \leq j_{N-2} \leq A_n \leq B_n.$$

In the same way as in Theorem 7, we can show that there exists a finite set X, such that

$$\begin{aligned} A_n &= \text{mex}_0(X \cup \{A_i, B_i : 0 \leq i < n\}), \\ B_n &\leq 2A_n + \sum_{i=1}^{N-2} j_i + (A_n + 1) \prod_{i=1}^{N-2} (j_i + 1). \end{aligned}$$

The set X is given by:

$$X = \{x \geq j_{N-2} : \exists\, 0 \leq k < j_{N-2} \quad \text{s.t.} \quad (j_1 \ldots, j_{N-2}, k, x) \in \mathcal{P}\} \cup \{0, \ldots, j_{N-2} - 1\}.$$

(Here $(j_1 \ldots, j_{N-2}, k)$ is not nondecreasing, since $0 \leq k < j_{N-2}$.) If we set $y = \text{mex}_0(X \cup \{A_i, B_i : 0 \leq i < n\})$, it can be shown that there is a finite number of values $z \geq y$ such that the position $(j_1, \ldots, j_{N-2}, y, z)$ has a P-follower of a form different from $(j_1, \ldots, j_{N-2}, y, w)$, $w \geq y$. Let R be the set of these z values. The bound on B_n is the bound on $|R \cup \{0, \ldots, A_n - 1\}|$, which is attained in the same manner as in the 3-pile case.

The question of the distance $B_n - A_n$ is still open, but computer tests done for the 4-pile case suggest that here too the distance starts growing in jumps of 1, eventually.

6 Epilogue

In this paper we dealt with a generalization of a special case of complementary Beatty sequences: $\{a_n\}$, $\{b_n\}$ satisfying Definition 2, depending on two independent variables $n_0 \in \mathbb{Z}_{\geq 1}$, $X \subset \mathbb{Z}_{\geq 1}$, where $c = \beta - \alpha$ is an integer. Using the fact that $b_n = a_n + cn$ for all $n \geq n_0$, we were able to show that the difference sequence $a_{n+1} - a_n$, when regarded as a word over the alphabet $\Sigma = \{1, 2\}$, is generated by a finite prefix; and by synchronizing it with the difference word of the homogeneous sequence $a'_n = \lfloor n\alpha \rfloor$, $n \geq 1$, we succeeded in describing the pattern of the shift sequence s_n. Two corner stones in the proof were the Synchronization Lemma and the Nesting Lemma in Section 3. Theorem 7 contributes to settling the Wythoff game conjectures.

Here are some of the remaining questions.

1. How do K (the number of irregular shift sequences) and γ (the regular shift value) depend on c, n_0 and X? Can we compute these values without computing a prefix of W, based on c, n_0, X alone? More important, how does p (the index at which the shifts stabilize) depend on c, n_0 and X? The answer to this question has a great effect on the computational complexity of the problem.

2. What happens when $\delta = \beta - \alpha$ is not an integer? In this case, $b_n - a_n \in \{\lfloor n\delta \rfloor, \lfloor n\delta \rfloor + 1\}$, so we can't generate the difference word W by a straightforward morphism. Moreover, W may contain letters other than 1, 2: recall that $b_{n+1} - b_n \in a_{n+1} - a_n + \{\lfloor \beta \rfloor, \lfloor \beta \rfloor - 1, \lfloor \beta \rfloor - 2\}$ (9). Since

$\beta = \alpha/(\alpha-1)$, $\alpha > 1.5$ implies $\beta < 3$, i.e., $\lfloor\beta\rfloor = 2$, and we might get $b_{n+1} - b_n = a_{n+1} - a_n = 1$. Thus The B sequence might contain consecutive integers, which means that the A sequence contains differences > 2. There are many examples for this: almost every $\alpha > 1.5$ we tested (except, of course, for $\alpha = \varphi \approx 1.618$), generated an A sequence with differences of 3, and sometimes 4.

Nevertheless, computer tests suggest that the same pattern holds: starting at some point, s_n assumes one of three successive values, and the indices at which s_n deviates from the mid-value show a pattern similar to that of the integer case.

3. Can the generalization we made here be extended to other types of complementary sets of integers? For example, sequences of the form $\lfloor n\alpha+\gamma\rfloor$, $\lfloor n\beta+\delta\rfloor$, where $0 < \alpha < \beta$ are real numbers that satisfy (3), and γ, δ are real. For this case it makes sense to consider also α, β rational. In [9], Fraenkel gave necessary and sufficient conditions for such nonhomogeneous sequences to be complementary over $\mathbb{Z}$, or to be N-upper complementary, or N-lower complementary. See also O'Briant [20].

Can the generalization be extended to the case of $m > 2$ sequences whose union is $\mathbb{Z}_{\geq N}$?

4. What are the P-positions of the N-pile Wythoff's game? We would like to prove (34b), i.e., to complete the proof of the first conjecture, and thus prove the second one as well.

7 Acknowledgment

Evgeniya Falkovich, a 3-month summer student of Aviezri Fraenkel in 2002, computed the first few thousand P-positions of Wythoff's game on N piles for several values of N. The results of these computations provided part of the stimulus for this work.

References

[1] S. Beatty, Problem 3173, Amer. Math. Monthly **33** (1926), 159; **34** (1927), 159.

[2] E.R. Berlekamp, J.H. Conway and R.K. Guy, *Winning Ways for your Mathematical Plays*, Vol. I & II, Academic Press, London, 1982.

[3] U. Blass and A.S. Fraenkel [1990], The Sprague–Grundy function for Wythoff's game, *Theoret. Comput. Sci.* (Math Games) **75** (1990) 311–333.

[4] U.Blass, A.S. Fraenkel and R. Guelman, How far can Nim in disguise be stretched?, J. Combin. Theory (Ser. A) **84** (1998) 145–156.

[5] C. L. Bouton, Nim, A Game with a Complete Mathematical Theory, *Ann. Math. Princeton* **3**, 35–39, 1901-1902.

[6] I.G. Connell, A generalization of Wythoff's game, *Canad. Math. Bull.* **2** (1959) 181–190.

[7] H.S.M. Coxeter, The golden section, phyllotaxis and Wythoff's game, *Scripta Math.* **19** (1953) 135–143.

[8] A. Dress, A. Flammenkamp and N. Pink, Additive periodicity of the Sprague-Grundy function of certain Nim games, *Adv. in Appl. Math.* **22** (1999) 249–270.

[9] A.S. Fraenkel, The bracket function and complementary sets of integers, Can. J. Math, **21** (1969), 6-27.

[10] A. S. Fraenkel, How to beat your Wythoff games' opponents on three fronts, Amer. Math. Monthly **89** (1982), 353–361.

[11] A.S. Fraenkel, Wythoff games, continued fractions, cedar trees and Fibonacci searches, *Theoret. Comput. Sci.* **29** (1984) 49–73.

[12] A.S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985) 105–114.

[13] A.S. Fraenkel, Scenic trails ascending from sea-level Nim to alpine chess, in: *Games of No Chance,* Proc. MSRI Workshop on Combinatorial Games, July, 1994, Berkeley, CA, MSRI Publ. (R.J. Nowakowski, ed.), Vol. 29, Cambridge University Press, Cambridge, 1996, pp. 13–42.

[14] A.S. Fraenkel, Complexity, appeal and challenges of combinatorial games, Expanded version of a keynote address at Dagstuhl Seminar "Algorithmic Combinatorial Game Theory", Feb. 2002; to appear in *Theoretical Computer Science*, special issue on Algorithmic Combinatorial Game Theory.

[15] A.S. Fraenkel and I. Borosh, A generalization of Wythoff's game, *J. Combin. Theory* (Ser. A) **15** (1973) 175–191.

[16] A.S. Fraenkel and M. Ozery, Adjoining to Wythoff's game its P-positions as moves, *Theoret. Comput. Sci.* **205** (1998) 283–296.

[17] R.K. Guy and R.J. Nowakowski, Unsolved Problems in Combinatorial Game Theory, *More Games of No Chance,* Proc. MSRI Workshop on Combinatorial Games, July, 2000, Berkeley, CA, MSRI Publ. (R.J. Nowakowski, ed.), Vol. 42, Cambridge University Press, Cambridge, 2002, pp. 457–473, Problem 53.

[18] H. Landman, A simple FSM-based proof of the additive periodicity of the Sprague-Grundy function of Wythoff's game, in: *More Games of No Chance*, Proc. MSRI Workshop on Combinatorial Games, July, 2000, Berkeley, CA, MSRI Publ. (R. J. Nowakowski, ed.), Vol. 42, Cambridge University Press, Cambridge, pp. 383–386, 2002.

[19] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press (2002), chapter 1.

[20] K. O'Briant, Fraenkel's partition and Brown's decomposition, *INTEGERS, Electronic J. Comb. Number Theory* **3** 2003, #A11.

[21] C.D. Olds, Continued Fractions, Random House, New York, 1963.

[22] J. W. S Rayleigh, *The Theory of Sound*, 2nd edition, vol. 1, Macmillan, London, 1894, pp. 122–123.

[23] R. Silber, A Fibonacci property of Wythoff pairs, *Fibonacci Quart.* **14** (1976) 380–384.

[24] R. Silber, Wythoff's Nim and Fibonacci representations, *Fibonacci Quart.* **15** (1977) 85–88.

[25] W. A. Wythoff, A modification of the game of Nim, Nieuw Arch. Wisk., **7** (1907), 199–202.

[26] A.M. Yaglom and I.M. Yaglom, *Challenging Mathematical Problems with Elementary Solutions*, Vol. II, Holden-Day, San Francisco, translated by J. McCawley, Jr., revised and edited by B. Gordon, 1967.

[27] E. Zeckendorf, Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas, *Bull. Soc. Roy. Sci. Liège* **41** (1972) 179–182.