RUTCOR
RESEARCH
REPORT

A polynomial algorithm for a two parameter extension of the Wythoff NIM based on the Perron-Frobenius theory.

Endre Boros[a] Vladimir Gurvich[b]
Vladimir Oudalov[c]

RRR 19-2011, November 2011

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ, 08854; e-mail: boros@rutcor.rutgers.edu
[b]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ, 08854; e-mail: gurvich@rutcor.rutgers.edu
[c]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ, 08854; e-mail: voudalov@rutcor.rutgers.edu

Rutcor Research Report
RRR 19-2011, November 2011

A polynomial algorithm for a two parameter extension of the Wythoff NIM based on the Perron-Frobenius theory.

Endre Boros Vladimir Gurvich Vladimir Oudalov

Abstract. For any positive integer parameters a and b, the second author recently introduced a generalization mex_b of the standard minimum excludant $\mathrm{mex} = \mathrm{mex}_1$, along with a game NIM$(a,b)$ that extends further Fraenkel's NIM = NIM$(a,1)$, which in its turn is a generalization of the classical Wythoff NIM = NIM$(1,1)$. It was shown that P-positions (the kernel) of NIM(a,b) are given by the following typical recursion:

$$x_n = \mathrm{mex}_b\{x_i, y_i \mid 0 \le i < n\}, \;\; y_n = x_n + an; \;\; n \ge 0,$$

and conjectured that for all a, b the limits $\ell(a,b) = x_n(a,b)/n$ exist and are irrational algebraic numbers. In this paper we prove this conjecture showing that $\ell(a,b) = \frac{a}{r-1}$, where $r > 1$ is the Perron root of the polynomial

$$P(z) = z^{b+1} - z - 1 - \sum_{i=1}^{a-1} z^{\lceil ib/a \rceil},$$

whenever a and b are coprime; furthermore, it is known that $\ell(ka, kb) = k\ell(a,b)$. In particular, $\ell(a,1) = \alpha_a = \frac{1}{2}(2-a+\sqrt{a^2+4})$. In 1982, Fraenkel (1982) introduced the game NIM(a) = NIM$(a,1)$, obtained the above recursion and solved it explicitly getting $x_n = \lfloor \alpha_a n \rfloor$, $y_n = x_n + an = \lfloor (\alpha_a + a)n \rfloor$. Here we provide a polynomial time algorithm based on the Perron-Frobenius theory solving game NIM(a,b), although we have no explicit formula for its kernel.

Keywords: impartial games, NIM, Wythoff's NIM, Fraenkel's NIM, minimum excludant, algebraic number, asymptotic

Acknowledgements: The authors gratefully acknowledge the partial support by NSF grants IIS 0803444 and CMMI 0856663.

1 Introduction

For any positive integer a and b, a game NIM(a,b) was recently introduced in [5] as follows. Two piles contain x and y matches. Two players take turns. By one move, it is allowed to take x' and y' matches from these two piles whenever

$$0 \le x' \le x, \quad 0 \le y' \le y, \quad 0 < x' + y', \quad \text{and either } |x' - y'| < a \quad \text{or} \quad \min(x', y') < b. \tag{1}$$

In other words, a player can take "approximately equal" (differing by at most $a-1$) numbers of matches from both piles or any number of matches from one pole but at most $b-1$ from the other. Game NIM(a,b) extends further Fraenkel's game NIM(a) = NIM$(a,1)$ [2, 3], which, in its turn, is a generalization of the Wythoff NIM$(1,1)$ [10], see also [1].

A position of NIM(a,b) is a non-negative integer pair (x,y). Due to obvious symmetry, positions (x,y) and (y,x) are equivalent. By default, we assume that $x \le y$.

In [5], game NIM(a,b) was solved by the following standard recursive formula for the kernel, or in other words, for the P-positions (x_n, y_n):

$$x_n = \mathrm{mex}_b\{x_i, y_i \mid 0 \le i < n\}, \quad y_n = x_n + an; \quad n \ge 0, \tag{2}$$

where function mex_b is defined as follows: Given a finite nonempty subset $S \subseteq \mathbb{Z}_+$ of m non-negative integers, let us order S and extend it by $s_{m+1} = \infty$ and by $s_0 = -b$, to get a sequence $s_0 < s_1 < \cdots < s_m < s_{m+1}$. Let us choose the (unique) minimum i such that $s_{i+1} - s_i > b$. Then, by definition, $\mathrm{mex}_b\{S\} = s_i + b$. It is easy to see that we get $\mathrm{mex}_b\{\emptyset\} = 0$.

It follows that mex_b is well-defined and for $b = 1$ it is the classic minimum excludant mex, which assigns to S the (unique) minimum non-negative integer missing in S. Thus, $\mathrm{mex}_1 = \mathrm{mex}$ and (2) turns into recursive solution of NIM$(a,1)$ given by Fraenkel in [2, 3].

Furthermore, Fraenkel solved the recursion for NIM$(a,1)$ and got the following explicit formula for (x_n, y_n): Let $\alpha_a = \frac{1}{2}(2 - a + \sqrt{a^2+4})$ be the (unique) positive root of the quadratic equation $\widehat{z}^2 + (a-2)\widehat{z} - a = 0$ or equivalently $\frac{1}{\widehat{z}} + \frac{1}{\widehat{z}+a} = 1$. In particular, $\alpha_1 = \frac{1}{2}(1+\sqrt{5})$ is the *golden section* and $\alpha_2 = \sqrt{2}$. Then it follows that for all $n \in \mathbb{Z}_+$ we have

$$x_n = \lfloor \alpha_a n \rfloor \quad \text{and} \quad y_n = x_n + an \equiv \lfloor n(\alpha_a + a) \rfloor. \tag{3}$$

By (3) we obtain the asymptotics $\lim\limits_{n\to\infty} \dfrac{x_n(a)}{n} = \alpha_a$ and $\lim\limits_{n\to\infty} \dfrac{y_n(a)}{n} = \alpha_a + a$. As mentioned in [2], the explicit formula (3) solves the game in linear time, in contrast to recursion (2), providing only an exponential algorithm. However, it looks too difficult to solve (2) explicitly when $b > 1$, because of the following bounds obtained in [5].

$$b \le x_{n+1} - x_n \le 2b \quad \text{and} \quad b + a \le y_{n+1} - y_n \le 2b + a. \tag{4}$$

Thus, for $b = 1$ the difference $x_{n+1} - x_n$ is either 1 or 2, and thus $\alpha_a n$ is a good approximation of x_n. Similar estimate seems to be harder to find due to 4.

Although we are not able to give closed form expressions for x_n and y_n in case of $b > 1$, we can compute these values (and, thus, solve NIM(a, b)) by a polynomial time algorithm.

Remark 1 *A somewhat similar situation appears in the recent works Hadad [7], Fraenkel and Peled [4].*

Theorem 1 *The values x_n and y_n can be computed in $O(g(a, b) \log n)$ iterations (each of which involves arithmetical operations with integers of size $O(n)$) for all $n \in \mathbf{Z}_+$, where $g(a, b)$ is a constant depending only on a and b. Furthermore, given $Z \in \mathbf{R}_+$, we can find the largest index n such that $x_n \le Z$ using $O(g(a, b) \log Z)$ operations.*

The proof of this theorem will be given in Sections 2 and 3. This theorem provides a polynomial algorithm to play NIM(a, b). Indeed, given positive integers x and y we can decide in polynomial time whether the position (x, y) is a P-position of NIM(a, b), that is, whether $x = x_n$ and $y = y_n$ for some $n \ge 0$. If yes then there is no winning move from (x, y). Otherwise, we can find in polynomial time a P-position (x_n, y_n) that can be reached from (x, y) by one move, in accordance with the rules of NIM(a, b).

The solutions of both normal and misère versions of the game are based on this algorithm and described in Sections 8 and 9, respectively.

The next step is to show that a linear asymptotic still holds for $b > 1$. It was conjectured in [5] that the limits $\ell(a, b) = \lim\limits_{n\to\infty} \dfrac{x_n(a, b)}{n}$ exist for all a, b and are algebraic numbers. In this paper we prove this conjecture and provide an explicit formula for the value of the limit.

Theorem 2 *The limit $\ell(a, b)$ exists for all a, b, and whenever $(a, b) = 1$, it is given by the formula $\ell(a, b) = \frac{a}{r-1}$, where $r > 1$ is the unique positive real root of the polynomial*

$$P(z) = z^{b+1} - z - 1 - \sum_{i=1}^{a-1} z^{\lceil ib/a \rceil}, \tag{5}$$

which is the characteristic polynomial of a non-negative $(b + 1) \times (b + 1)$ integer matrix associated to NIM(a, b) depending only on parameters a and b.

Remark 2 *Note that by the Perron-Frobenius theorem, we have $|r'| < r$ for any other root r' of $P(z)$. For notational convenience, we use a variable transformation $\widehat{z} = a/(z - 1)$. Thus, in case of $b = 1$ Theorem 2 yields the same as the above cited results of Fraenkel.*

The case $g = \gcd(a, b) > 1$ is also covered by Theorem 2, since, as it was shown in [5], $x_n(a, b)$ (and, hence, $y_n(a, b)$ and $\ell(a, b)$ as well) are uniform functions of a and b, that is, for all positive integer a, b and k, n we have

$$x_n(ka, kb) = kx_n(a, b), \quad y_n(ka, kb) = ky_n(a, b), \quad \text{and} \quad \ell(ka, kb) = k\ell(a, b), \tag{6}$$

We provide the proof for Theorem 2 in Sections 5, 6 and 7. It is derived with the help of the Perron-Frobenius theorem and the Collatz-Wielandt formula for non-negative matrices; see Chapter 8 of the textbook [8].

Remark 3 *Alternatively, Theorem 2 could be derived from the Cauchy-Ostrovsky theorem; see theorems 1.1.3, 1.1.4 in the textbook [9] and verify that our polynomial $P(z)$ satisfies all condition of the latter theorem.*

2 Basic Properties

In the rest of the paper we assume that a and b are relatively prime positive integers. We write $a = \alpha b + \beta$, where $\alpha \geq 0$ and $0 < \beta \leq b$ are integers (so, if $b = 1$ then $\alpha = a - 1$, $\beta = 1$), and define the set $B = \{0, 1, ..., b\}$. In our complexity estimates we shall regard the parameters a, b, (and α, β) as fixed constants.

We denote by $\mathcal{S} = \mathcal{S}(a, b) = (x_n, y_n \mid n = 0, 1, ...)$ the sequence defined by (2), and note that $x_0 = y_0 = 0$ and $x_1 = b$, $y_1 = b + a$, etc. Let us next note that the sequences x_n and y_n are monotone increasing and we have $y_{n+1} - y_n \geq a + b > b$ by (4). Thus, every y_i is followed by some x_j, $j > i$ according to (4). Let us then introduce $\sigma(i) = j$ denoting the index of the x_j following y_i immediately in the sequence $\mathcal{S}$. Clearly, $\sigma(i)$ is well defined for all $i \in \mathbb{Z}_+$.

The following monotonicity property of the σ operator is immediate from the definitions:

Lemma 1 *If $\sigma^p(i) < j < \sigma^p(i+1)$ for some i, j and p, then, for all $t \in \mathbb{Z}_+$, we have*

$$\sigma^{p+t}(i) < \sigma^t(j) < \sigma^{p+t}(i+1).$$

□

The following statement provides a "local" description of the sequence $\mathcal{S}$, which will be instrumental in our proofs and algorithms.

Lemma 2 *For $i \in \mathbb{Z}_+$ we introduce $j = j(i)$ such that $b + j = x_{i+1} - x_i$. Then, we have*

$$\sigma(i+1) - \sigma(i) = \begin{cases} \alpha + 1 & \text{if} \quad 0 \leq j \leq b - \beta, \\ \alpha + 2 & \text{if} \quad b - \beta < j \leq b. \end{cases} \tag{7}$$

Furthermore, the sequence $\mathcal{S} \cap [y_i, x_{\sigma(i+1)}]$ looks like

$$y_i,\ x_{\sigma(i)},\ x_{\sigma(i)+1},\ \ldots,\ x_{\sigma(i)+\ell},\ y_{i+1},\ x_{\sigma(i+1)}$$

where $\ell = \ell(j) \in \{\alpha, \alpha + 1\}$ as indicated by (7), and where

$$x_{\sigma(i+1)} - y_{i+1} = x_{\sigma(i)+\ell} - x_{\sigma(i)+\ell-1} = \ldots = x_{\sigma(i)+1} - x_{\sigma(i)} = x_{\sigma(i)} - y_i = b$$

and $x_{\sigma(i+1)} - x_{\sigma(i)+\ell} = b + \mu(j)$, where

$$\mu(j) \ = \ \begin{cases} \beta + j & \text{if} \quad 0 \leq j \leq b - \beta, \\ \beta + j - b & \text{if} \quad b - \beta < j \leq b. \end{cases} \tag{8}$$

Proof: By (4) we know that $y_{i+1} - y_i = x_{i+1} - x_i + a = b + j + a = (\alpha+1)b + \beta + j$. We also know that there are only x-s between y_{i+1} and y_i, and hence, by the mex_b rule we must have some x-s b-apart, as long as they fit this interval. This implies that we have $\lfloor \frac{(\alpha+1)b+\beta+j}{b} \rfloor = \alpha + 1 + \lfloor \frac{\beta+j}{b} \rfloor$ many x-es between y_{i+1} and y_i. Since we also know that by definition $x_{\sigma(i)} - y_i = b$ for all indices i, the claims follow by elementary calculations. □

By default, all considered vectors are assumed to be column vectors, while all row vectors will be indicated explicitly by the transposition sign.

Let us introduce vectors $\mathbf{e} = (1, 1, ..., 1) \in \mathbb{Z}^B$, $\mathbf{b} = (b, b+1, ..., 2b) \in \mathbb{Z}^B$, and, for an arbitrary pair $i < j$ of indices, let us define the vector $\mathbf{d}(i,j) \in \mathbb{Z}_+^B$, where

$$\mathbf{d}(i,j)_k = |\{s \mid i \leq s < j, \quad x_{s+1} - x_s = b + k\}|$$

is the number of consecutive x-s between x_i and x_j the distance between which is exactly $b + k$ for $k \in B$.

Let us denote by $\mathbf{e}^\ell \in \{0,1\}^B$ the ℓ-th unit vector, for $\ell \in B$ and remark that, by definition, for each index $i \in \mathbb{Z}_+$ there is an $\ell = \ell(i) \in B$ such that

$$\mathbf{d}(i, i+1) \;=\; \mathbf{e}^\ell \tag{9}$$

Corollary 1 *The following equalities hold whenever $i < j < k$:*

$$\mathbf{d}(i,k) \;=\; \mathbf{d}(i,j) + \mathbf{d}(i,k); \tag{10}$$

$$j - i \;=\; \mathbf{e}^T \mathbf{d}(i,j); \tag{11}$$

$$x_j - x_i \;=\; \mathbf{b}^T \mathbf{d}(i,j); \tag{12}$$

$$\mathbf{d}(0,1) \;=\; \mathbf{e}^0. \tag{13}$$

Proof*: It is straightforward* □

Let us also introduce a non-negative integer matrix $M \in \mathbb{Z}_+^{B \times B}$ by defining

$$M_{i,j} = \begin{cases} \alpha = \lfloor \frac{a+j-1}{b} \rfloor & \text{if } i = 0 \text{ and } 0 \leq j \leq b - \beta \\ \alpha + 1 = \lfloor \frac{a+j-1}{b} \rfloor & \text{if } i = 0 \text{ and } b - \beta < j \leq b \\ 1 & \text{if } i > 0 \text{ and } (j + a - i \bmod b) = 0 \\ 0 & \text{if } i > 0 \text{ and } (j + a - i \bmod b) \neq 0 \end{cases} \tag{14}$$

	0	1	$\cdots$	$b-\beta-1$	$b-\beta$	$b-\beta+1$	$\cdots$	$b-1$	b
0	α	α	$\cdots$	α	α	$\alpha+1$	$\cdots$	$\alpha+1$	$\alpha+1$
1	0	0	$\cdots$	0	0	1	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\beta-1$	0	0	$\cdots$	0	0	0	$\cdots$	1	0
β	1	0	$\cdots$	0	0	0	$\cdots$	0	1
$\beta+1$	0	1	$\cdots$	0	0	0	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$b-1$	0	0	$\cdots$	1	0	0	$\cdots$	0	0
b	0	0	$\cdots$	0	1	0	$\cdots$	0	0

Let us notice that if $x_{i+1} - x_i = b + j$, then column j of M provides the distribution of the consecutive x differences between $x_{\sigma(i)}$ and $x_{\sigma(i+1)}$, as shown in Lemma 2. Thus, the next two relations follow readily from Lemma 2.

Corollary 2 *For any $i < j$ we have*

$$\mathbf{d}(\sigma(i), \sigma(j)) = M\mathbf{d}(i, j); \tag{15}$$

□

Let us next note that Lemma 2 and the corollaries above provide us with a computational tool, allowing us to compute indices and x-values, as we move forward by the σ operator.

Corollary 3 *Given a positive integer i, the values x_i, x_{i+1}, and vector $\mathbf{d}(0, i)$, we can compute the index $\sigma(i)$, the values x_k, and vectors $\mathbf{d}(0, k)$ for all $\sigma(i) \le k \le \sigma(i+1)$ in $O((\alpha+1)b^2) = O(1)$ total time.*

Proof: Introduce $j = x_{i+1} - x_i - b$ as in Lemma 2 and set $\mu(j)$ as defined in (8). Then, by the Corollaries 1 and 2 we can compute first

$$\sigma(i) \;=\; \mathbf{e}^T \left(\mathbf{d}(0,1) + \mathbf{d}(1, \sigma(i))\right) \;=\; 1 + \mathbf{e}^T M \mathbf{d}(0, i)$$

and

$$\sigma(i+1) \;=\; \sigma(i) + \mathbf{e}^T M \mathbf{d}(i, i+1) \;=\; \sigma(i) + \mathbf{e}^T M \mathbf{e}^{\mu(j)}.$$

Next, using Lemma 2 we can compute the vectors

$$\mathbf{d}(0, k) \;=\; \mathbf{d}(0,1) + \mathbf{d}(1, \sigma(i)) + (k - \sigma(i))\mathbf{e}^0 \;=\; (k + 1 - \sigma(i))\mathbf{e}^0 + M\mathbf{d}(0, i) \tag{16}$$

for $\sigma(i) \le k < \sigma(i+1)$, and finally

$$\mathbf{d}(0, \sigma(i+1)) \;=\; \mathbf{d}(0, \sigma(i+1) - 1) + \mathbf{e}^{\mu(j)}.$$

Since $x_0 = 0$, by (12), we obtain that $x_k = \mathbf{b}^T\mathbf{d}(0,k)$ for $\sigma(i) \leq k \leq \sigma(i+1)$.

Here, computing $M\mathbf{d}(0,i)$ takes $O(b^2)$ time, computing $M\mathbf{e}^{\mu(j)}$ takes $O(b)$ time, while all other operations are multiplications or additions of vectors of dimension b, and hence, take also $O(b)$ time each. Since, we have $O(\alpha+1)$ such operations by Lemma 2, the claim follows. □

To be able to start our algorithms, described in the next section, we need to draw a few more computational consequences of the above basic results.

Corollary 4 *Given positive integers t, i, $\sigma(i)$, x_i, and the corresponding vector $\mathbf{d}(i,\sigma(i))$ we can compute $\sigma^{2^k}(i)$ and $x_{\sigma^{2^k}(i)}$ for all $k = 0,1,...,t$ in $O(t)$ time.*

Proof: Note first that the powers M^{2^j}, $j = 0,1,...,t$ can be computed in $O(t)$ time. Thus, by Corollaries 1 and 2 we get

$$\mathbf{d}(i,\sigma^{2^{k+1}}(i)) = \mathbf{d}(i,\sigma^{2^k}(i)) + \mathbf{d}(\sigma^{2^k}(i),\sigma^{2^{k+1}}(i)) = \left(I + M^{2^k}\right)\mathbf{d}(i,\sigma^{2^k}(i))$$

for $k = 0,1,...,t-1$, and hence the claim follows by (11) and (12). □

Corollary 5 *Given a positive integer N, we can compute the largest integer n such that $\sigma^n(0) \leq N$ in $O(\log n)$ time.*

Proof: We shall compute such a largest n in its binary representation. Just like in Corollary 4, let us compute first M^{2^j} and $\sigma^{2^j}(0)$ for $j = 0,1,...,t+1$, where $\sigma^{2^t}(0) \leq N < \sigma^{2^{t+1}}(0)$. Note that we also get all the vectors $\mathbf{d}(0,\sigma^{2^j}(0))$ for $j = 0,1,...,t$. As in the previous corollary, we can do all these in $O(t)$ time.

Let us also note that for arbitrary integers m and k we have by (11) and (15) that

$$\sigma^{m+k}(0) = \mathbf{e}^T\mathbf{d}(0,\sigma^{m+k}(0)) = \sigma^m(0) + \mathbf{e}^T M^m \mathbf{d}(0,\sigma^k(0)).$$

Thus, starting with $m = 2^t$ we can find a largest integer $j < t$ for which $\sigma^{m+2^j}(0) \leq N$. Then update $m \leftarrow m + 2^j$, compute $M^{m+2^j} = M^m M^{2^j}$, and repeat, until we have $\sigma^m(0) \leq N < \sigma^{m+1}(0)$. Then we stop, and output $n = m$.

Note that after the initialization, we have at most t iterations, in which we need to try to add 2^j to m exactly once for all indices $j = 0,1,...,t-1$. Since each trial by the above equalities takes $O(1)$ time, the total time is $O(t) = O(\log n)$, as claimed. □

Corollary 6 *Given a positive integer X, we can compute the largest integer n such that $x_{\sigma^n(0)} \leq X$ in $O(\log n)$ time.*

Proof: Perfectly analogous to the previous proof. We need to use also equation (12). □

3 Basic Algorithms

We are now ready to describe our main algorithm(s) with which we can answer a number of different questions about NIM(a, b). The precise complexity estimate will follow in the next section, where we show that $\sigma(i)^n$ is an exponential function of n. We will provide three algorithms:

- COMPUTE-X(N): Given a positive integer N, compute x_N;
- FIND-N(X): Given a positive integer X, find the maximum N such that $x_N \leq X$, return both N and x_N;
- FIND-N(Y): Given a positive integer Y, find the maximum N such that $y_N \leq Y$, return both N and y_N.

Let us observe first that we can solve the last question by computing $X = Y + b$ and finding the largest integer N such that $x_{\sigma(N)} \leq X$. Then, by the definition of the σ operator, the value $y_N = x_N + a$ is the right value to return by FIND-N(Y). This observation makes the algorithmic descriptions very similar and allows us to describe all three algorithms simultaneously.

The key idea in our algorithms is to use the σ operator to derive a series of increasingly smaller and smaller windows of upper and lower bounds on the targeted input parameters N or X. We shall describe all three procedures parallel, with indicating the differences in parentheses.

Initialization: We shall start by computing a largest integer n such that $\sigma^n(0) \leq N$ (or $x_{\sigma^n(0)} \leq X$) by Corollary 5 (or 6), in $O(\log n)$ time. Then we initialize $i = 0$, and set $\xi_i = 0$, $x_{\xi_i} = 0$, $x_{\xi_i+1} = b$ and $\mathbf{d}(0, \xi_i) = (0, 0, \ldots, 0) \in \mathbb{R}^B$. We also compute the matrices M^n in $O(\lceil \log n \rceil)$ time, and M^{-1} (as rational) in $O(1)$ time.

Parameters: We shall compute for every index $i = 0, 1, \ldots, n$ a corresponding index ξ_i, the values x_{ξ_i} and x_{ξ_i+1}, and the vector $\mathbf{d}(0, \xi_i)$, in $O(1)$ time for each index i, satisfying the following properties.

Invariant(s): We maintain for every $i = 0, ..., n-1$ that either

$$\sigma^n(0) \leq \sigma^{n-i}(\xi_i) \leq \sigma^{n-i-1}(\xi_{i+1}) \leq N < \sigma^{n-i-1}(\xi_{i+1} + 1) \leq \sigma^{n-i}(\xi_i + 1) \leq \sigma^{n+1}(0)$$

holds (for the first problem) or

$$x_{\sigma^n(0)} \leq x_{\sigma^{n-i}(\xi_i)} \leq x_{\sigma^{n-i-1}(\xi_{i+1})} \leq X < x_{\sigma^{n-i-1}(\xi_{i+1}+1)} \leq x_{\sigma^{n-i}(\xi_i+1)} \leq x_{\sigma^{n+1}(0)}$$

holds (for the last two problems).

Termination: For $i = n$ we shall have either $\xi_n = N$ (for the first problem) in which case $x_{\xi_n} = x_N$ is the right output for COMPUTE-X(N), or $x_{\xi_n} \leq X < x_{\xi_n+1}$ (for the last two problems), in which case $N = \xi_n$ and $x_N = x_{\xi_n}$ are the right output for FIND-N(X), and $N = \xi_{n-1}$ and $y_N = a + x_{\xi_{n-1}}$ are the right output for FIND-N(Y).

Main Iteration: If $i = n$, then we go to **Termination**, otherwise we compute, as in Corollary 3, the indices $\sigma(\xi_i)$, $\sigma(\xi_i + 1)$ and the values x_k and vectors $\mathbf{d}(0, k)$ for all indices $\sigma(\xi_i) \leq k \leq \sigma(\xi_i + 1)$ in $O(1)$ time. We also compute the matrix $M^{n-i-1} = M^{-1}M^{n-i}$ in $O(1)$ time.

- If $\sigma(\xi_i + 1) = \sigma(\xi_i) + 1$, then we set $\xi_{i+1} = \sigma(\xi_i)$, $i = i + 1$, and repeat the **Main Iteration**.
- Otherwise, we compute the indices $\sigma^{n-i-1}(k) = \mathbf{e}^T M^{n-i-1}\mathbf{d}(0, k)$ and values $x_k = \mathbf{b}^T M^{n-i-1}\mathbf{d}(0, k)$ for $\sigma(\xi_i) \leq k \leq \sigma(\xi_i + 1)$ in $O(1)$ time. These indices (in case of the first problem) or values (in case of the other two problems) subdivide the intervals $[\sigma^{n-i}(\xi_i), \sigma^{n-i}(\xi_i + 1))$ (in case of the first problem) or $\left[x_{\sigma^{n-i}(\xi_i)}, x_{\sigma^{n-i}(\xi_i+1)}\right)$ (in case of the other two problems), and one of these intervals will contain N (or X). Let us then choose index k such that $\sigma^{n-i-1}(k) \leq N < \sigma^{n-i-1}(k + 1)$ (or $x_{\sigma^{n-i-1}(k)} \leq X < x_{\sigma^{n-i-1}(k+1)}$), and set $\xi_{i+1} = k$, $i = i + 1$ and return to the **Main Iteration**.

Theorem 3 *The above algorithm(s) correctly compute the answer to all there problems and terminate in $O(n)$ time.*

Proof: The correctness of the computations follow by Lemma 2 and Corollaries 1,2,5 and 6. The complexity then follows since each step in the algorithm takes constant time, and we repeat only the **Main Iteration**, n times.

The correctness follows by the maintained **Invariants(s)**, and by the definition of the σ operator. □

Let us remark finally that to argue that the above procedures are computationally efficient it is enough to show that $\sigma^n(0)$ is an exponential function of n, which we will prove in the next section.

4 Asymptotic Distribution of $\mathcal{S}(a, b)$

Let us denote by r_j, $j = 0, 1, ..., b$ the eigenvalues of M and by $\mathbf{u}(j)$ and $\mathbf{v}(j)$, $j = 0, 1, \ldots, b$ the corresponding left and right eigenvectors of M. We label these such that

$$|r_0| \geq |r_1| \geq \cdots \geq |r_b|, \tag{17}$$

and we scale the eigenvectors such that $\mathbf{e}^T\mathbf{u}(j) = \mathbf{e}^T\mathbf{v}(j) = 1$ for all $j = 0, 1, ..., b$. Let us recall from matrix theory that if $r_i \neq r_j$, then we must have

$$\mathbf{u}(i)^T\mathbf{v}(j) = 0, \tag{18}$$

as the equalities $r_i\mathbf{u}(i)^T\mathbf{v}(j) = \mathbf{u}(i)^T M\mathbf{v}(j) = r_j\mathbf{u}(i)^T\mathbf{v}(j)$ imply.

Let us next show that M is an irreducible matrix, which is equivalent for a nonnegative matrix with the fact that a finite power of M has only positive entries.

Lemma 3 *Every entry in M^{2b} is a positive integer.*

Proof: Let us note that the last β entries in the first row are always positive (and the whole first row is positive if $\alpha > 0$). Due to the cyclic arrangement of the 1-s in the columns of M (in rows 1...b), we get that the last 2β entries are positive in M^2, the last 3β entries are positive in M^3, etc. Thus, M^b has its first row positive. Note also that once the first row of a power of M is positive then it remains positive for all higher powers. Let us observe next that row β is positive in M^{b+1}, and it remains positive in all higher powers of M. Then row $(2\beta \mod b) + 1$ is positive in M^{b+2}, and it remains positive in all higher powers of M. Iterating this argument, using the cyclic structure of rows 1...b of M, we get that all rows are positive in M^n for all $n \geq 2b$. □

Corollary 7 *$r(a, b) = r_0 > 1$ is the unique largest eigenvalue of M and the corresponding eigenvectors $\mathbf{u}(0)$ and $\mathbf{v}(0)$ have positive real components.*

Proof: By Lemma 3 matrix M is irreducible, and hence primitive. Thus, we can apply the Perron-Frobenius theorem and conclude that r_0 is a positive real eigenvalue with multiplicity one, the corresponding eigenvectors $\mathbf{u}(0)$ and $\mathbf{v}(0)$ have positive real components and that $r_0 > |r_1|$.

To see that $r_0 > 1$, we apply the Collatz-Wielandt formula, which claims for a nonnegative matrix M that

$$r_0 = \max_{\mathbf{z}\geq 0,\ \mathbf{z}\neq 0} \min_{i:z_i\neq 0} \frac{M\mathbf{z}}{z_i}.$$

Observing then that r_0^{2b} is the largest real eigenvalue of M^{2b} and that M^{2b} has positive integer entries by Lemma 3 we can apply the above formula with $\mathbf{z} = \mathbf{e}$ to M^{2b} and obtain that $r_0^{2b} \geq (b + 1)$, from which we get

$$r_0 \geq (b + 1)^{1/2b} > 1. \tag{19}$$

□

Remark 4 *By the above claims it follows that the eigenvalues of M satisfy the inequalities $r_0 > 1$ and $r_0 > |r_1| \geq |r_2| \geq \cdots \geq |r_b|$. Let us also note that all these values depend only on parameters a and b. Let us introduce then the parameter*

$$1 > \delta \ = \ \delta(a, b) \geq \frac{|r_1|}{r_0} \tag{20}$$

such that $\delta r_0 \geq 1$.

A further useful property of these eigenvectors is that they span all unit vectors with a positive real coefficient for $\mathbf{v}(0)$:

Lemma 4 *For every $\ell \in B$ there is a positive real $\gamma_0^\ell = \gamma_0^\ell(a,b) > 0$ and there are complex coefficients γ_j^ℓ, $j = 1, ..., b$ such that*

$$\mathbf{e}^\ell \;=\; \sum_{j\in B} \gamma_j^\ell \mathbf{v}(j). \tag{21}$$

Proof: Since the right eigenvectors of M span the space (21) holds for some complex coefficients γ_j^ℓ, $j \in B$. By Corollary 7 the left eigenvector $\mathbf{u}(0)$ has positive real components, and by (18) we have $\mathbf{u}(0)^T\mathbf{v}(j) = 0$ for all $j \neq 0$. Thus by (21) we get

$$0 \;<\; \mathbf{u}(0)^T\mathbf{e}^\ell \;=\; \gamma_0^\ell \mathbf{u}(0)^T\mathbf{v}(0).$$

Since by Corollary 7 the right eigenvector $\mathbf{v}(0)$ has also positive components, the positivity of γ_0^ℓ follows. □

Using the above, we can prove the following statement, which will be instrumental in our proof of Theorem 2.

Lemma 5 *There is a positive real $C = C(a,b)$ depending only on parameters a and b such that for all integers $n \geq 0$ we have*

$$-C(a,b)(n+1)\delta^n \;\leq\; \frac{\sigma^n(1)}{r_0^n} \;-\; \frac{\gamma_0^0 r_0}{r_0 - 1} \;\leq\; C(a,b)(n+1)\delta^n.$$

Proof: Let us first note that since M depends only on parameters a and b, we have the same for all its eigenvalues and eigenvectors.

Using equations (14) and (11) iteratively, starting with $(i,j) = (0,1)$, we obtain

$$\sigma^n(1) \;=\; \mathbf{e}^T \left(I + M + M^2 + \cdots + M^n\right) \mathbf{d}(0,1). \tag{22}$$

Since $\mathbf{d}(0,1) = \mathbf{e}^0$, by Lemma 4 we get

$$\mathbf{d}(0,1) \;=\; \mathbf{e}^0 \;=\; \sum_{j=0}^{b} \gamma_j^0 \mathbf{v}(j). \tag{23}$$

This, together with (22) yields

$$\sigma^n(1) \;=\; \sum_{j=0}^{b} \gamma_j^0 (1 + r_j + \cdots + r_j^n),$$

implying

$$\sigma^n(1) \;=\; \gamma_0^0 \frac{r_0^{n+1}-1}{r_0-1} + \sum_{j=1}^{b} \gamma_j^0 \sum_{k=0}^{n} r_j^k. \tag{24}$$

Since $|r_j| \le |r_1| < r_0$ for all $j = 1, ..., b$ by Remark 4 and Lemma 4 we get

$$\begin{aligned}
\left| \sum_{j=1}^{b} \gamma_j^0 \sum_{k=0}^{n} r_j^k \right| &\le r_0^n \left(\sum_{j=1}^{b} |\gamma_j^0| \right) \left(\sum_{k=0}^{n} \frac{(\delta r_0)^k}{r_0^n} \right) \\
&= r_0^n \left(\sum_{j=1}^{b} |\gamma_j^0| \right) \delta^n \left(\sum_{k=0}^{n} \frac{1}{(\delta r_0)^{n-k}} \right) \\
&\le (n+1) r_0^n \delta^n \left(\sum_{j=1}^{b} |\gamma_j^0| \right)
\end{aligned}$$

Since $r_0\delta \ge 1$ by Remark 4, for every $n \ge 0$ we have $(n+1)r_0^n\delta^n \ge 1$. Thus, the claim holds with $C(a,b) = \frac{\gamma_0^0}{r_0-1} + \sum_{j=1}^{b} |\gamma_j^0|$, since the coefficients γ_j^0, $j \in B$ and eigenvalue r_0 depend only on a and b. □

Proof of Theorem 1. It follows directly from Lemma 5 and Theorem 3. □

5 Existence of $\lim_{n\to\infty} \dfrac{x_n}{n}$

Let us first show that the subsequence $x_{\sigma^n(1)}/\sigma^n(1)$ has a limit. For this we prove a claim for the distribution of $x_{\sigma^n(1)}$, analogues to Lemma 5.

Lemma 6 *There is a positive real $D = D(a,b)$ depending only on parameters a and b such that for all integers $n \ge 0$ we have*

$$-D(a,b)(n+1)\delta^n \;\le\; \frac{x_{\sigma^n(1)}}{r_0^n} \;-\; \frac{\gamma_0^0 r_0(\mathbf{b}^T\mathbf{v}(0))}{r_0-1} \;\le\; D(a,b)(n+1)\delta^n.$$

Proof: Analogously to the proof of Lemma 5, by using equations (14) and (12) iteratively, starting with $(i,j) = (0,1)$, we obtain

$$x_{\sigma^n(1)} \;=\; \mathbf{b}^T \left(I + M + M^2 + \cdots + M^n \right) \mathbf{d}(0,1). \tag{25}$$

Since $\mathbf{d}(0,1) = \mathbf{e}^0$, by Lemma 4 we get

$$\mathbf{d}(0,1) \;=\; \mathbf{e}^0 \;=\; \sum_{j=0}^{b} \gamma_j^0 \mathbf{v}(j). \tag{26}$$

This, together with (25) yields

$$x_{\sigma^n(1)} \;=\; \sum_{j=0}^{b} \gamma_j^0(\mathbf{b}^T\mathbf{v}(j))(1 + r_j + \cdots + r_j^n),$$

implying

$$x_{\sigma^n(1)} \;=\; \gamma_0^0(\mathbf{b}^T\mathbf{v}(0))\frac{r_0^{n+1}-1}{r_0-1} + \sum_{j=1}^{b} \gamma_j^0(\mathbf{b}^T\mathbf{v}(j)) \sum_{k=0}^{n} r_j^k. \tag{27}$$

Since $|r_j| \le |r_1| < r_0$ for all $j = 1, ..., b$ by Remark 4 and Lemma 4 we get

$$\begin{aligned}
\left| \sum_{j=1}^{b} \gamma_j^0(\mathbf{b}^T\mathbf{v}(j)) \sum_{k=0}^{n} r_j^k \right| &\le r_0^n \left(\sum_{j=1}^{b} |\gamma_j^0(\mathbf{b}^T\mathbf{v}(j))| \right) \left(\sum_{k=0}^{n} \frac{(\delta r_0)^k}{r_0^n} \right) \\
&= r_0^n \left(\sum_{j=1}^{b} |\gamma_j^0(\mathbf{b}^T\mathbf{v}(j))| \right) \delta^n \left(\sum_{k=0}^{n} \frac{1}{(\delta r_0)^{n-k}} \right) \\
&\le (n+1) r_0^n \delta^n \left(\sum_{j=1}^{b} |\gamma_j^0(\mathbf{b}^T\mathbf{v}(j))| \right)
\end{aligned}$$

Since $r_0\delta \ge 1$ by Remark 4, for every $n \ge 0$ we have $(n+1)r_0^n\delta^n \ge 1$. Thus, the claim holds with $C(a,b) = \frac{\gamma_0^0(\mathbf{b}^T\mathbf{v}(0))}{r_0-1} + \sum_{j=1}^{b} |\gamma_j^0(\mathbf{b}^T\mathbf{v}(j))|$, since the coefficients γ_j^0, $j \in B$, eigenvectors $\mathbf{v}(j)$, $j \in B$, and eigenvalue r_0 depend only on a and b. □

Theorem 4

$$\lim_{n\to\infty} \frac{x_{\sigma^n(1)}}{\sigma^n(1)} \;=\; \mathbf{b}^T\mathbf{v}(0).$$

Proof: Let us denote by $\pm C(a,b)$ and $\pm D(a,b)$ quantities between $-C(a,b)$ and $C(a,b)$ (respectively, between $-D(a,b)$ and $D(a,b)$) which guarantee the equality in Lemma 5 and 6. Then, these lemmas imply

$$\begin{aligned}
\frac{x_{\sigma^n(1)}}{\sigma^n(1)} &= \frac{r_0^n \frac{\gamma_0^0 r_0(\mathbf{b}^T\mathbf{v}(0))}{r_0-1} \;\pm\; D(a,b)(n+1)r_0^n\delta^n}{r_0^n \frac{\gamma_0^0 r_0}{r_0-1} \;\pm\; C(a,b)(n+1)r_0^n\delta^n} \\
&= \frac{\mathbf{b}^T\mathbf{v}(0) \;\pm\; \frac{D(a,b)(r_0-1)}{\gamma_0^0 r_0}(n+1)\delta^n}{1 \;\pm\; \frac{C(a,b)(r_0-1)}{\gamma_0^0 r_0}(n+1)\delta^n}.
\end{aligned}$$

Since $\delta < 1$, the factor $(n+1)\delta^n$ goes to zero as $n \to \infty$, and the claim follows. □

Next we show that the range $\sigma^n(i) - \sigma^n(i-1)$ is proportional to r_0^n, if n is large, for all integers $i \ge 1$.

Lemma 7 *For every integer $i \geq 1$ there is an index $\ell = \ell(i) \in B$ and a positive real $E = E(\ell, a, b)$ such that*

$$-E(\ell,a,b)\delta^n \;\leq\; \frac{\sigma^n(i) - \sigma^n(i-1)}{r_0^n} \;-\; \gamma_0^\ell \;\leq\; E(\ell,a,b)\delta^n$$

holds for all $n \geq 0$.

Proof: As we noted in (4), there exists an index $\ell = \ell(i) \in B$ such that $\mathbf{d}(i-1,i) = \mathbf{e}^\ell$. Thus, by (11) and Lemma 4 we get

$$\sigma^n(i) - \sigma^n(i-1) \;=\; \mathbf{e}^T M^n \mathbf{d}(i-1,i) \;=\; \mathbf{e}^T M^n \sum_{j\in B} \gamma_j^\ell \mathbf{v}(j) \;=\; \sum_{j\in B} \gamma_j^\ell r_j^n.$$

Since $\mathbf{e}^T\mathbf{v}(j) = 1$ for all $j \in B$, the above implies

$$\frac{\sigma^n(i) - \sigma^n(i-1)}{r_0^n} \;=\; \gamma_0^\ell + \sum_{j=1}^{b} \gamma_j^\ell \left(\frac{r_j}{r_0}\right)^n.$$

Let us note that γ_0^ℓ is a positive real according to Lemma 4. Since $|r_j|/r_0 \leq \delta$ by Remark 4, and since $E(\ell,a,b) = \sum_{j=1}^b |\gamma_j^\ell|$ is a constant depending only on $\ell = \ell(i)$, a and b according to Lemma 4, the claim follows from the above equality. □

We also prove an analogous claim for the difference $x_{\sigma^n(i)} - x_{\sigma^n(i-1)}$.

Lemma 8 *For every integer $i \in \mathbf{Z}_+$ there is an index $\ell = \ell(i) \in B$ and a positive real $F = F(\ell,a,b)$ such that*

$$-F(\ell,a,b)\delta^n \;\leq\; \frac{x_{\sigma^n(i)} - x_{\sigma^n(i-1)}}{r_0^n} \;-\; \gamma_0^\ell(\mathbf{b}^T\mathbf{v}(0)) \;\leq\; F(\ell,a,b)\delta^n$$

holds for all $n \geq 0$.

Proof: As we noted in (4), there exists an index $\ell = \ell(i) \in B$ such that $\mathbf{d}(i-1,i) = \mathbf{e}^\ell$. Thus, by (11) and Lemma 4 we get

$$x_{\sigma^n(i)} - x_{\sigma^n(i-1)} \;=\; \mathbf{b}^T M^n \mathbf{d}(i-1,i) \;=\; \mathbf{b}^T M^n \sum_{j\in B} \gamma_j^\ell \mathbf{v}(j) \;=\; \sum_{j\in B} \gamma_j^\ell (\mathbf{b}^T\mathbf{v}(j)) r_j^n,$$

implying

$$\frac{x_{\sigma^n(i)} - x_{\sigma^n(i-1)}}{r_0^n} \;=\; \gamma_0^\ell(\mathbf{b}^T\mathbf{v}(0)) + \sum_{j=1}^{b} \gamma_j^\ell (\mathbf{b}^T\mathbf{v}(j)) \left(\frac{r_j}{r_0}\right)^n.$$

Let us note that $\gamma_0^\ell(\mathbf{b}^T\mathbf{v}(0))$ is a positive real according to Lemma 4. Since $|r_j|/r_0 \leq \delta$ by Remark 4, and since $F(\ell,a,b) = \sum_{j=1}^b |\gamma_j^\ell(\mathbf{b}^T\mathbf{v}(j))|$ is a constant depending only on $\ell = \ell(i)$, a and b according to Lemma 4, the claim follows from the above equality. □

Theorem 5 *For every integer $i \geq 1$ we have*

$$\lim_{n\to\infty} \frac{x_{\sigma^n(i)}}{\sigma^n(i)} \;=\; \mathbf{b}^T\mathbf{v}(0).$$

Proof: By Lemma 7 there exists a real $-E(\ell(k),a,b) \leq \varepsilon(k) \leq E(\ell(k),a,b)$ for every integer $k \geq 1$ such that

$$\sigma^n(k) - \sigma^n(k-1) \;=\; \gamma_0^{\ell(k)} r_0^n + \varepsilon(k)\delta^n r_0^n. \tag{28}$$

Analogously, by Lemma 8 there exists a real $-F(\ell(k),a,b) \leq \varphi(k) \leq F(\ell(k),a,b)$ for every integer $k \geq 1$ such that

$$x_{\sigma^n(k)} - x_{\sigma^n(k-1)} \;=\; \gamma_0^{\ell(k)}(\mathbf{b}^T\mathbf{v}(0)) r_0^n + \varphi(k)\delta^n r_0^n. \tag{29}$$

Summing these up for $k = 1, 2, ..., i$ and using Lemmas 5 and 6 for $k = 0$ we get

$$\sigma^n(i) \;=\; \left(\gamma_0^0 \frac{r_0}{r_0-1} + \sum_{k=1}^{i} \gamma_0^{\ell(k)}\right) r_0^n \;+\; \left(\sum_{k=0}^{i} \varepsilon(k)\right) r_0^n \delta^n$$

and

$$x_{\sigma^n(i)} \;=\; \left(\gamma_0^0 \frac{r_0}{r_0-1} + \sum_{k=1}^{i} \gamma_0^{\ell(k)}\right) (\mathbf{b}^T\mathbf{v}(0)) r_0^n \;+\; \left(\sum_{k=0}^{i} \varphi(k)\right) r_0^n \delta^n.$$

Note that $A = \gamma_0^0 \frac{r_0}{r_0-1} + \sum_{k=1}^{i} \gamma_0^{\ell(k)}$ is a positive real according to Lemma 4. Furthermore, we have $E(i) = \sum_{k=0}^{i} \varepsilon(k)$ is majorized in absolute value by $(i+1)\max_{\ell\in B} E(\ell,a,b)$ and similarly $F(i) = \sum_{k=0}^{i} \varphi(k)$ is majorized in absolute value by $(i+1)\max_{\ell\in B} F(\ell,a,b)$, and hence both are linear in i. Consequently, we get

$$\frac{x_{\sigma^n(i)}}{\sigma^n(i)} \;=\; \frac{A(\mathbf{b}^T\mathbf{v}(0))r_0^n + E(i)r_0^n\delta^n}{Ar_0^n + F(i)r_0^n\delta^n}$$

$$\;=\; \frac{\mathbf{b}^T\mathbf{v}(0) + \frac{E(i)}{A}\delta^n}{1 + \frac{F(i)}{A}\delta^n}$$

from which the claim follows. □

Now we are ready to prove the existence of the limit.

Theorem 6

$$\lim_{n\to\infty} \frac{x_n}{n} \;=\; \mathbf{b}^T\mathbf{v}(0).$$

Proof: Let us denote by $m = m(n)$ the largest integer for which $x_n \geq x_{\sigma^m(1)}$, and fix an integer $1 \leq k \leq m$ to be specified later. Note that by Lemma 5 we have $m \approx \log_{r_0} n$ and we shall choose $k = o(m)$. Then, there exists an integer $\sigma^k(1) \leq i \leq \sigma^{k+1}(1)$ such that

$\sigma^{m-k}(i) \leq n < \sigma^{m-k}(i+1)$, by the monotonicity of the σ operator. By Lemma 7 we have with $\ell = \ell(i-1) \in B$

$$\Delta \;=\; n - \sigma^{m-k}(i) \leq \gamma_0^{\ell} r_0^{m-k} + E(\ell, a, b) r_0^{m-k} \delta^{m-k},$$

and consequently

$$x_{\sigma^{m-k}(i)} \leq x_n \leq x_{\sigma^{m-k}(i)} + 2\Delta b \;\leq\; x_{\sigma^{m-k}(i)} + 2b r_0^{m-k} \left(\gamma_0^{\ell} + E(\ell, a, b)\delta^{m-k}\right).$$

Since $\sigma^{m-k}(i) \leq n < \sigma^{m-k}(i+1)$ by our choice of i,

$$\sigma^{m-k}(i) \;\leq\; \sigma^{m-k}(i) + \Delta \;=\; n \;\leq\; \sigma^{m-k}(i) + r_0^{m-k}\left(\gamma_0^{\ell} + E(\ell, a, b)\delta^{m-k}\right).$$

The above inequalities imply, by using $A = \left(\gamma_0^{\ell} + E(\ell, a, b)\delta^{m-k}\right)$, that

$$\frac{x_{\sigma^{m-k}(i)}}{\sigma^{m-k}(i) + r_0^{m-k} A} \;\leq\; \frac{x_n}{n} \;\leq\; \frac{x_{\sigma^{m-k}(i)} + 2b r_0^{m-k} A}{\sigma^{m-k}(i)}. \tag{30}$$

Let us note that by Lemma 5 and by our choice of i we have

$$\sigma^{m-k}(i) \;\geq\; \sigma^m(1) \;\geq\; \frac{\gamma_0^0 r_0}{r_0 - 1} r_0^m - C(a, b)(m+1) r_0^m \delta^m$$

from which

$$\frac{r_0^{m-k} A}{\sigma^{m-k}(i)} \;\leq\; \frac{r_0^{m-k} A}{r_0^m \left(\frac{\gamma_0^0 r_0}{r_0 - 1} - C(a, b)(m+1)\delta^m\right)} \;=\; O\left(\frac{1}{r_0^k}\right)$$

follows. Thus, for $k \approx m/2$ we get from the above and (30) that

$$\frac{\frac{x_{\sigma^{m/2}(i)}}{\sigma^{m/2}(i)}}{1 + O(1/r_0^{m/2})} \;\leq\; \frac{x_n}{n} \;\leq\; \frac{x_{\sigma^{m/2}(i)}}{\sigma^{m/2}(i)} + O(1/r_0^{m/2}),$$

implying the claim by Theorem 6. □

In what follows, we can provide a simpler expression for the limit $\mathbf{b}^T \mathbf{v}(0)$, by computing precisely the positive real eigenvector $\mathbf{v}(0)$.

6 Characteristic polynomial

Lemma 9 *Suppose* $\gcd(a, b) = 1$, $P_{a\,b}$ *is* $b \times b$ *cyclic permutation matrix,* $e_i \mapsto e_{i+a}$ *and* $T(a, b, i)$ *is a minor of* $(P_{a\,b} - zI)$ *without row* a *and column* i. *Then* $\det(T(a, b, i)) = (-1)^{a+b+i+1} z^{(ic \bmod b)}$, *where* $c = -a^{-1}$ *in the ring* $\mathbb{Z}/b\mathbb{Z}$.

Proof: Every row and every column of $(P_{a\,b} - zI)$ contains exactly two non-zero entries. Let us trace what entries are included into the only left permutation. If we start from position (a, b) we get a degree of z: we exclude (a, b), include (b, b), exclude $(b-a, b)$, include $(b-a, b-a)$, exclude $(b-2a, b-a)$,..., include $(b-ka, b-ka)$, exclude $(b-(k+1)a, b-ka)$ etcetera. If we start from another end, we get no zs: exclude (a, a), include $(2a, a)$, exclude $(2a, 2a)$, include $(3a, 2a)$, exclude $(3a, 3a)$,..., exclude (ka, ka), include $((k+1)a, ka)$ etcetera. The sequences end, when they reach the deleted column (i). So, if we look at the first sequence we get an equation for the degree of z d in $\mathbb{Z}/b\mathbb{Z}$: $-ad = i$ or $d = -ia^{-1}$. □

Lemma 10 *If* $\gcd(a, b) = 1$*, the characteristic polynomial of the matrix* $X \in \mathbb{R}^{B\times B}$

$$X_{i,j} = \begin{cases} x_i & \text{if } i = 0 \\ 1 & \text{if } i > 0 \text{ and } (j + a - i \bmod b) = 0 \\ 0 & \text{if } i > 0 \text{ and } (j + a - i \bmod b) \neq 0 \end{cases}$$

equals to

$$z^{b+1} - (x_b - x_0) - x_0 z^b - z - \sum_{i=1}^{b-1} z^{ic \bmod b} x_i, \text{ where } c = -a^{-1} \text{ in } \mathbb{Z}/b\mathbb{Z} \tag{31}$$

Proof: Compute directly the determinant of $(X - zI)$ by column 0, using that the characteristic polynomial of a cyclic permutation matrix $X_{1\ldots b,1\ldots b}$ is $(-1)^b(z^b - 1)$ and the Lemma 9:

$$\begin{aligned} \det(X - zI) &= (x_0 - z)(-1)^b(z^b - 1) + (-1)^a \sum_{i=1}^{b} (-1)^{i+1} x_i T(a, b, i) \\ &= (-1)^{b+1} \left(x_0 + z^{b+1} - x_0 z^b - z - x_b - \sum_{i=1}^{b-1} x_i z^{(ic \bmod b)} \right) \end{aligned}$$

□

Theorem 7 *If* $\gcd(a, b) = 1$*,* $a = \alpha b + \beta$*,* $0 < \beta \leq b$*, the following polynomials are identical and represent the characteristic polynomial of the matrix* M*:*

$$z^{b+1} - z - 1 - \alpha \sum_{i=1}^{b} z^i - \sum_{i=1}^{\beta-1} z^{(ia^{-1} \bmod b)}, \text{ where } a^{-1} \text{ is inverse in } \mathbb{Z}/b\mathbb{Z} \tag{32a}$$

$$z^{b+1} - z - 1 - \sum_{i=1}^{a-1} z^{\lceil \frac{ib}{a} \rceil} \tag{32b}$$

Proof: We get (32a), if we substitute $x_j = \left\lfloor \frac{a+j-1}{b} \right\rfloor$ into (31):

$$\begin{aligned} & z^{b+1} - 1 - \alpha z^b - z - \sum_{i=1}^{b-1} z^{ic \bmod b} \left\lfloor \frac{a+i-1}{b} \right\rfloor \\ =& z^{b+1} - z - 1 - \alpha \sum_{i=1}^{b} z^i - \sum_{i=b+1-\beta}^{b-1} z^{ic \bmod b} \\ =& z^{b+1} - z - 1 - \alpha \sum_{i=1}^{b} z^i - \sum_{i=1}^{\beta-1} z^{ia^{-1} \bmod b} \end{aligned}$$

Let us show that the polynomials (32a) and (32b) are identical. Easy to check that it is true, when either a or b is one. Otherwise we omit $z^{b+1} - z - 1$ in each expression, and we can reformulate the equality in the following way:

Lemma 11 *Let* $\gcd(a, b) = 1$, $a = \alpha b + \beta$, $0 < \beta \le b$ *and we have divided a segment into* b *"small" equal segments numbered from 1 to* b*. Let the segment be also divided into a equal segments with "points" (there are* $a - 1$ *of them). Then every small segment contains either* α *or* $\alpha + 1$ *points and those, which contain* $\alpha + 1$ *points have the numbers* $ia^{-1} \bmod b$*, where* $i = 1 \dots b - 1$ *and* a^{-1} *is the inverse to* β *in the ring* $\mathbb{Z}/b\mathbb{Z}$*.*

Proof: Suppose the initial segment has length ab. Then every "small" segment has length a and the distance between points is b. The distance from the end of segment number k and the previous point is $ka \bmod b$. It contains $\alpha + 1$ points if and only if this distance is positive and less than β. We can write this as $ka = i, i = 1 \dots \beta - 1$ or $k = ia^{-1}$. □

7 The positive eigenvector of M and the value of $\lim_{n\to\infty} \frac{x_n}{n}$

In this section we will find the eigenvector $\mathbf{v} = \mathbf{v}(0)$ with eigenvalue $r = r_0$ as defined in Section 4, and then, the value of $\mathbf{b}^T\mathbf{v}$, which, according to Theorem 6 equals to $\lim_{n\to\infty} \frac{x_n}{n}$. We have

$$r\mathbf{v}_{(a \bmod b)} = \mathbf{v}_0 + \mathbf{v}_b \tag{33a}$$

$$r\mathbf{v}_{(2a \bmod b)} = \mathbf{v}_{(a \bmod b)} \tag{33b}$$

$$\cdots\cdots\cdots\cdots$$

$$r\mathbf{v}_{(ia \bmod b)} = \mathbf{v}_{((i-1)a \bmod b)} \tag{33c}$$

$$\cdots\cdots\cdots\cdots$$

$$r\mathbf{v}_b = \mathbf{v}_{(b-a \bmod b)} \tag{33d}$$

$$r\mathbf{v}_0 = \sum_{i=0}^{b} \left\lfloor \frac{i+a-1}{b} \right\rfloor \mathbf{v}_i \tag{33e}$$

Let us denote $\mu = \mathbf{v}_{(a \bmod b)}$. We will also write $x \operatorname{Mod} y = y - (-x \bmod y)$ (that is $x \operatorname{Mod} y = x \bmod y$ except it takes value y instead of 0. From (33a–33d) it is obvious

$$\mathbf{v}_{(ja \operatorname{Mod} b)} = \frac{\mu}{r^{j-1}} \text{ for all } j = 1, \ldots, b. \tag{34}$$

We get from (33a) and (34)

$$\mathbf{v}_0 = \frac{\mu(r^b - 1)}{r^{b-1}}. \tag{35}$$

Adding together (33a–33d) we get

$$r \sum_{i=1}^{b} \mathbf{v}_i = \sum_{i=0}^{b} \mathbf{v}_i$$

As $\mathbf{e}^T \mathbf{v} = 1$, we get

$$\mathbf{v}_0 = \frac{r-1}{r} \tag{36}$$

Combining (35) and (36) we get

$$\mu = \frac{(r-1)r^{b-2}}{r^b - 1} \tag{37}$$

Thus, the eigenvector is

$$\mathbf{v}_0 = \frac{r-1}{r}, \mathbf{v}_b = \frac{r-1}{(r^b-1)r}, \mathbf{v}_{(ka \bmod b)} = \frac{(r-1)r^{b-k-1}}{r^b - 1}, k = 1, \ldots, b-1. \tag{38}$$

Now we can compute $\mathbf{b}^T \mathbf{v}$. Substituting (38) into $\mathbf{b}^T \mathbf{v}$ we get:

$$\mathbf{b}^T \mathbf{v} = \sum_{k=1}^{b-1} \left(\frac{r^{b-k-1}(r-1)}{r^b - 1} (ka \bmod b) \right) + \frac{b(r-1)}{(r^b-1)r} + b$$

Using that $ka \bmod b = ka - b \left\lfloor \frac{ka}{b} \right\rfloor$, we can see this expression as

$$\begin{aligned} &\sum_{k=1}^{b-1} \left(\frac{r^{b-k-1}(r-1)}{r^b - 1} \left(ka - b \left\lfloor \frac{ka}{b} \right\rfloor \right) \right) + \frac{b(r-1)}{(r^b-1)r} + b \\ &= b + \frac{r^{b-2}(r-1)}{r^b - 1} \cdot \left(a \sum_{i=1}^{b} \frac{i}{r^{i-1}} - b \sum_{i=1}^{a-1} \sum_{j=\lfloor \frac{ib}{a} \rfloor + 1}^{b} \frac{1}{r^{j-1}} \right) \end{aligned} \tag{39}$$

We can find that

$$\sum_{i=1}^{b} \frac{i}{r^{i-1}} = \frac{r^2}{(r-1)^2} - \frac{r^3 - br^2(r-1)}{r^{b+1}(r-1)^2} \tag{40}$$

and

$$\sum_{i=1}^{a-1}\sum_{j=\lfloor\frac{ib}{a}\rfloor}^{b}\frac{1}{r^{j-1}}=\frac{\sum_{i=1}^{a-1}r^{1-\lfloor\frac{ib}{a}\rfloor}}{r-1}-\frac{r^2(a-1)}{r^{b+1}(r-1)} \tag{41}$$

Substituting (40) and (41) into (39), we get:

$$\begin{aligned}&b+\left(\frac{ar^2}{(r-1)^2}\left(1-\frac{(b+1)r-b}{r^{b+1}}\right)-\frac{br^2}{r-1}\left(\sum_{i=1}^{a-1}\frac{1}{r^{\lfloor\frac{ib}{a}\rfloor+1}}-\frac{a-1}{r^{b+1}}\right)\right)\frac{r^{b-2}(r-1)}{r^b-1}\\&=\frac{br(r-1)(r^b-1)+a(r^{b+1}-br-r+b)-b(r-1)\left(\sum_{i=1}^{a-1}r^{b-\lfloor\frac{ib}{a}\rfloor}-(a-1)\right)}{r(r-1)(r^b-1)}\end{aligned} \tag{42}$$

But

$$\sum_{i=1}^{a-1}r^{b-\lfloor\frac{ib}{a}\rfloor}=\sum_{i=1}^{a-1}r^{\lceil\frac{ib}{a}\rceil}=r^{b+1}-r-1$$

because r is a root of (32b). So (42) can be represented as

$$\begin{aligned}&\frac{br(r-1)(r^b-1)+a(r^{b+1}-br-r+b)-b(r-1)\left(r^{b+1}-r-1-(a-1)\right)}{r(r-1)(r^b-1)}\\&=\frac{br^{b+2}-br^{b+1}-br^2+br+ar^{b+1}-abr-ar+ab-br^{b+2}+br^{b+1}+br^2-br+abr-ab}{r(r-1)(r^b-1)}\\&=\frac{ar^{b+1}-ar}{r(r-1)(r^b-1)}=\frac{a}{r-1}\end{aligned}$$

8 Solving game NIM(a,b)

By definition, a game is solved if for any given position v one can decide whether it is a P-position and if it is not then one can find a P-position v' such that it can be reached from v by one move. We will show, that one can solve game NIM(a,b) for a position (x^*,y^*) using algorithms COMPUTE-X(N), FIND-N(X), FIND-N(Y) not more than once each, so it will take $O(\log(\max(x^*,y^*))$ operations.

SOLVE-GAME(x^*,y^*): Given non-negative integers $x^*\le y^*$, find an index N and P-position (x_N,y_N) such that it is either the same as (x^*,y^*) or is reachable with one move.

1: Call FIND-N($X=x^*$). If $x_N\le x^*-b$, go to step **2**. Compute $y_N=x_N+aN$. If $y_N>y^*$, go to step **3**. Otherwise, either (x_N,y_N) is the same as (x^*,y^*) (**P-position**), or can be got with decreasing x^* to x_N by less than b and y^* to y_N.

2: Call FIND-N($Y = x^*$). We know that $y_N > x^* - b$. So, the position (x_N, y_N) can be got with decreasing x^* to y_N by less than b and y^* to x_N.

3: Compute $N' = \lfloor \frac{y^*-x^*}{a} \rfloor$. We can say that $N' \leq N$ from step **1**. So, we can call COMPUTE-X($N = N'$) and $(x_{N'}, y_{N'})$ can be got with decreasing y^* by $(y^* - x^*)$ mod a and both of them by $x^* - x_{N'}$.

9 Solving the misère versions of NIM(a, b)

For joint definitions of the normal and misere versions of an impartial game we refer the reader, for example, to [3] or [6].

In case $a = 1$, the set of P-positions P^N and P^M "almost coincide" for both versions. More precisely, their symmetric difference consists of only six positions:

$P^N \setminus P^M = \{(0,0), (b, b+1), (b+1, b)\}$, while $P^M \setminus P^N = \{(0,1), (1,0), (b+1, b+1)\}$.

This result was obtained in [3] for $b = 1$ and extended to any positive integer b in [5].

It is not difficult also to verify, see [6], that

- (i) from any position of $P^M \setminus P^N$ there is a move to $P^N \setminus P^M$;
- (ii) from any non-terminal position of $P^N \setminus P^M$, that is, from $(b, b+1)$ or $(b+1, b)$ there is a move to $P^N \setminus P^M$;
- (iii) from any position $(x, y) \notin P^N \cup P^M$, either both sets P^N and P^M or none of them can be reached in one move.

Thus, for $a = 1$, the algorithm for the normal version of NIM(a, b) constructed in the previous section is applicable to the misère version, as well; see [6] for more details.

For any integer $a > 1$ (and $b \geq 1$) the kernel is defined by recursion

$$\tilde{x}_n = \text{mex}_b\{\tilde{x}_i, \tilde{y}_i \mid 0 \leq i < n\}, \quad \tilde{y}_n = \tilde{x}_n + an + 1; \quad n \in \mathbf{Z}_+.$$

.

This formula was proven in [3] for $b = 1$ and extended for any positive integer b in [5].

Let us notice that it differs just slightly from the recursion for the normal version of NIM(a, b) given in the abstract. In particular, comparing these two formulas we immediately conclude that, for any integer $a > 1$ and $b \geq 1$, the sets of P-positions of the normal and misere versions are disjoint, in contrast to the case $a = 1$; see [6] for more details.

Moreover, all properties described in Section 2 hold, except that (13) looks as $\mathbf{d}(0,1) = \mathbf{e}^1$. So, we need to modify the algorithms from Section 3, replacing e^0 with e^1 in (16) in Corollary 3 with

$$\mathbf{d}(0,k) \;=\; \mathbf{d}(0,1) + \mathbf{d}(1,\sigma(i)) + (k - \sigma(i))\mathbf{e}^0 \;=\; (k - \sigma(i))\mathbf{e}^0 + \mathbf{e}^1 + M\mathbf{d}(0,i) \qquad (43)$$

and initial value of x_{ξ_i+1} with $b+1$. Then we can use the algorithm SOLVE-GAME(x^*, y^*) without any changes.

Finally, let us remark that cases $a = 1$ and $a > 1$ differ substantially, while almost nothing depends on b.

References

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

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

[3] A.S. Fraenkel, Wythoff games, continued fractions, cedar trees and Fibonacci searches, *Theoretical Computer Science* **29** (1984) 49–73.

[4] A.S. Fraenkel and U. Peled, Harnessing the Unwieldy MEX Function, to appear in Games of No Chance 4.

[5] V. Gurvich, Further generalizations of Wythoff's game and minimum excludant function, RUTCOR Research Report, 16-2010, Rutgers University;

[6] V. Gurvich, Miserable and strongly miserable impartial games, RUTCOR Research Report 18-2011, Rutgers University.

[7] U. Hadad, Polynomializing Seemingly Hard Sequences Using Surrogate Sequences, MS. Thesis, Fac. of Math. Weiz. Inst. of Sci., 2008.

[8] Carl D. Meyer, Matrix analysis and applied linear algebra, SIAM, Philadelphia, PA, 2000.

[9] V. V. Prasolov, Polynomials, Algorithms and computation in Mattheatics 11, Springer Verlag, Berlin-Heidelberg, 2010.

[10] W.A. Wythoff, A modification of the game of Nim, Nieuw Archief voor Wiskunde, 7 (1907), 199-202