Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

An optimal bound on the number of moves for open mancala

Alessandro Musesti^a, Maurizio Paolini^{a,*}, Cesco Reale^b

^a Dipartimento di Matematica e Fisica "Niccolò Tartaglia", Università Cattolica del Sacro Cuore, Via dei Musei 41, 25121 Brescia, Italy ^b Festival Italiano di Giochi Matematici, Italy

ARTICLE INFO

Article history: Received 30 October 2013 Received in revised form 31 March 2015 Accepted 10 April 2015 Available online 5 June 2015

Keywords: Discrete-time games Integer compositions Zero-player game Discrete dynamical system

1. Introduction

ABSTRACT

We determine the optimal bound for the maximum number of moves required to reach a periodic configuration of *open mancala* (also called open owari), inspired by a popular African game. A mancala move can be interpreted as a map from the set of compositions of a given integer in itself, thus relating our result to the study of the corresponding finite dynamical system.

© 2015 Elsevier B.V. All rights reserved.

Mancala is a family of traditional African games played in many versions and under many names. It comprises a circular list of *holes* containing zero or more *seeds*. A move consists of selecting a nonempty hole, taking all its seeds and *sowing* them in the subsequent holes, one seed per hole.

Here we study an idealized version of this called *open mancala* (see [1], where the game is called *open owari*). We assume there is an infinite sequence of holes. Further we assume that the nonempty holes are consecutive; such a configuration is described by a sequence of positive integers λ_i , $i = 1, ..., \ell$, giving the number of seeds in each nonempty hole, starting from the leftmost one. We assume a move always selects the leftmost nonempty hole, sowing to the right. Thus the leftmost nonempty hole always advances of one position. It is clear that the sequence of configurations becomes periodic in a finite time. The periodic configurations have been completely classified [1,3,2]: see Section 2.

This game is related to a card game called *Bulgarian Solitaire*, discussed by Gardner [5] in a 1983 *Scientific American* column. Indeed, that game is isomorphic to the case of open mancala where $\lambda_i \geq \lambda_j$ for $i \leq j$ (monotone mancala), see Section 6. Gardner mentioned a conjecture on the maximum number of moves before the onset of periodicity when the number of cards is of the form k(k + 1)/2. The conjecture was proved independently a few years later by Igusa [7] and Etienne [4]. In 1998, Griggs & Ho [6] revisited the result and proposed a new conjecture in the case of a generic number of cards. Due to the isomorphism between monotone mancala and Bulgarian solitaire, our result is strongly related to this conjecture. However, our results apply to open mancala, not to the monotone variant, hence the conjecture about the Bulgarian Solitaire remains unproven.

The main result of the paper is the optimal bound for the maximum number of moves, before open mancala reaches periodicity, as a function of n, the number of seeds. In Section 3 we prove the lower bound by producing, for any n, a configuration that requires exactly that number of moves. In Section 4 we introduce the main tool in proving the upper

* Corresponding author. E-mail addresses: alessandro.musesti@unicatt.it (A. Musesti), paolini@dmf.unicatt.it (M. Paolini), cescoreale@gmail.com (C. Reale).

http://dx.doi.org/10.1016/j.disc.2015.04.013 0012-365X/© 2015 Elsevier B.V. All rights reserved.







bound, namely *s*-*monotonicity*. In Section 5 we prove that every configuration reaches periodicity within that number of moves.

2. Notations and setting

We denote by \mathbb{N} the set of nonnegative integers and by \mathbb{N}^* the set of strictly positive integers. By T_k , $k \in \mathbb{N}$, we denote the *k*th triangular number

$$T_k = \sum_{i=0}^k i = \frac{k(k+1)}{2},$$

where in particular $T_0 = 0$.

Definition 2.1. A configuration λ is a sequence of nonnegative numbers λ_i , $i \ge 1$, that is $\lambda : \mathbb{N}^* \to \mathbb{N}$. The support supp (λ) of a configuration λ is the set of indices $\{i \in \mathbb{N}^* : \lambda_i > 0\}$.

Definition 2.2. A mancala configuration λ is a configuration having a *connected* support of the form $\{1, \ldots, \ell\}$ for some $\ell = \text{len}(\lambda) \in \mathbb{N}$. The index $\text{len}(\lambda)$ is called the *length* of the configuration. In the special case of the *zero* configuration, we shall conventionally define its length to be zero. We shall denote by Λ the set of all mancala configurations.

Definition 2.3. If $\lambda \in \Lambda$, its *mass* is the number

$$|\lambda| = \sum_{i=1}^{\infty} \lambda_i = \sum_{i=1}^{\operatorname{len}(\lambda)} \lambda_i.$$

We denote by Λ_n the set of all mancala configurations of mass *n*.

The *mancala* game, in our setting, is a discrete dynamical system associated with a function $\mathcal{M} : \Lambda \to \Lambda$ mapping the space of *mancala* configurations Λ in itself.

Definition 2.4 (*Sowing*). The *sowing* of a mancala configuration λ is an operation on λ that results in a new configuration $\mu = \mathcal{M}(\lambda)$ defined as follows:

$$\mu_i = \begin{cases} \lambda_{i+1} + 1 & \text{if } 1 \le i \le \lambda_1, \\ \lambda_{i+1} & \text{if } i > \lambda_1. \end{cases}$$

Conventionally, M maps the empty configuration into itself.

It is clear that \mathcal{M} preserves the mass of a configuration, so that it can be restricted to Λ_n . It is convenient to define the (right) *shift* operator \mathcal{E} , that acts on generic sequences simply by a change in the indices.

Definition 2.5 (*Shift Operator*). If $\lambda : \mathbb{N}^* \to \mathbb{N}$ is a sequence, the sequence $\mathscr{E}(\lambda)$ is defined by

$$\mathscr{E}(\lambda)_i = \begin{cases} 0 & \text{if } i = 1, \\ \lambda_{i-1} & \text{if } i > 1. \end{cases}$$

Clearly, the result of the shift operator is never a mancala configuration (with the exception of the empty configuration). Both \mathcal{M} and \mathcal{E} can be iterated, the symbol \mathcal{M}^k (resp. \mathcal{E}^k) denoting the result of k repeated applications of \mathcal{M} (resp. of \mathcal{E}). The shift operator has a left inverse \mathcal{E}^{-1} , defined by $\mathcal{E}^{-1}(\lambda)_i = \lambda_{i+1}$ and satisfying $\mathcal{E}^{-1}(\mathcal{E}(\lambda)) = \lambda$ for any configuration λ .

Definition 2.6 (*Partial Ordering and Sum*). We say that $\lambda \leq \mu$ if $\lambda_i \leq \mu_i$ for all $i \geq 1$. Moreover we say that $\lambda < \mu$ if $\lambda \leq \mu$ and $\lambda \neq \mu$. If $\lambda, \mu : \mathbb{N}^* \to \mathbb{Z}$ are two integer sequences (in particular if any of them is a configuration in Λ), we define the sum and difference $\lambda \pm \mu$ componentwise: $(\lambda \pm \mu)_i = \lambda_i \pm \mu_i$.

Remark 2.1 (*Comparison*). It is easy to check that if $\lambda, \mu \in \Lambda$ and $\lambda \leq \mu$, then $\mathcal{M}(\lambda) \leq \mathcal{M}(\mu)$.

Definition 2.7. A configuration is called *monotone* if it is weakly decreasing, i.e. if $\lambda_i \ge \lambda_j$ whenever $i \le j$.

Remark 2.2 (*Monotone Mancala*). If λ is a monotone configuration, then so it is $\mathcal{M}(\lambda)$. Moreover, if $\lambda \in \Lambda$ has length ℓ , then $\mathcal{M}^{\ell-1}(\lambda)$ is monotone.

The following special configurations play an important role.

Definition 2.8 (*Marching Group*). For a given $k \in \mathbb{N}$ we define the special mancala configuration \mathfrak{mg}^k with mass T_k and length k, called *marching group*¹ of order k, as follows:

$$\mathfrak{mg}_i^k = \begin{cases} k-i+1 & \text{if } i \leq k, \\ 0. \end{cases}$$

Definition 2.9 (Augmented Marching Group). A mancala configuration λ such that $\mathfrak{mg}^k \leq \lambda < \mathfrak{mg}^{k+1}$ for some $k \in \mathbb{N}$ is called an *augmented marching group* of order k. It satisfies $T_k \leq |\lambda| < T_{k+1}$.

Notice that a marching group is a particular case of an augmented marching group. The following important theorem about augmented marching groups is readily proved. See [1, Theorems 1 and 2].

Theorem 2.1. Augmented marching groups are the only periodic configurations for \mathcal{M} and, if $\mathfrak{mg}^k \leq \lambda < \mathfrak{mg}^{k+1}$, then the period of λ is a divisor of k + 1. Marching groups are the only fixed points for \mathcal{M} .

A remarkable fact about open mancala is that every configuration becomes an augmented marching group, and hence periodic, after a finite number of moves. The paper is devoted to find an optimal bound on such a number of moves which depends only on the mass of the initial configuration.

Definition 2.10 (*Depth and Diameter*). The *depth* of a configuration λ is its distance from the periodic configuration, *i.e.* the number of moves needed to reach periodicity. For $n \ge 0$ we call the *depth of* Λ_n , denoted by D(n), the maximal depth of all configurations λ with $|\lambda| = n$.

The *diameter* of a configuration is the number of moves before the first repetition. It is equal to its depth plus the length of the period minus one. The maximal diameter of all configurations of mass *n* is called the *diameter of* Λ_n .

Clearly if $n = T_k$ the diameter of Λ_n equals its depth D(n), whereas if $n = T_k + r$, $1 \le r \le k$ the diameter of Λ_n is bounded above by D(n) + k.²

The number D(n) can be seen also as the depth of a graph. Indeed, for a given $n \ge 1$ we can construct the directed graph g_n having the configurations in Λ_n as nodes and an arc from λ to μ whenever $\mu = \mathcal{M}(\lambda)$. The graph g_n contains exactly 2^{n-1} nodes. A *cycle* of g_n corresponds to sequences of moves that repeat periodically. A configuration is periodic if it belongs to a cycle.

In the case $n \ge 2$, if we remove all arcs connecting two periodic configurations (arcs that belong to a cycle), the remaining graph \mathscr{G}_n^o is a disjoint union of trees rooted at a periodic configuration, and in each tree all arcs point toward the root. In this setting, D(n) represents the depth of \mathscr{G}_n , *i.e.* the maximal depth of the trees of \mathscr{G}_n^o .

2.1. Energy levels interpretation

We can view the mancala configurations and mancala moves in a way reminiscent of the energy levels for the electrons in an atom.

Let us consider the subset $L \subset \mathbb{N} \times \mathbb{N}$ given by

$$L = \{(i, j) \in \mathbb{N} \times \mathbb{N} : 1 \le i \le j\}.$$

It is convenient to think of the integers *i* and *j* as a numbering of square cells instead of coordinates of points, in a way similar to the familiar sea-battle game. Each square cell in *L* can either be empty or contain a single seed.

The *column* index *i* corresponds to one of the holes in the mancala game, so that λ_i , where $\lambda \in \Lambda$, is the total number of occupied cells in column *i* of *L*. We shall also associate an *energy* to a seed positioned in cell (i, j), given by its second coordinate *j*. The *k*th level of *L* is the set $\{(1, k), \ldots, (k, k)\}$ of cells having energy *k*. Finally we let the seeds free to immediately *fall* down without changing their column, but decreasing their energy (the second coordinate of their current position) provided the new position (and all those in between) are free.

The mancala configuration $\lambda \in \Lambda$ then corresponds to a positioning of seeds in *L* such that (i, j) is occupied by a seed if and only if $1 \le j \le \lambda_i$ (see Fig. 1).

A mancala move can now be viewed in this setting as follows. Firstly, each seed is moved one position to the left without changing its energy level, while seeds in column 1 are *rotated* to the last column of that energy level. Namely, a seed at position (i, j) is moved in the new position (i - 1, j) if i > 1, (j, j) if i = 1. This is just a left rotation of the seeds of each level, so that cells will receive at most one seed. After that, we let all the seeds in each column fall at the lowest energy possible.

¹ Please note that this is *not* a group in the mathematical sense.

² It is also strictly larger than D(n), more precisely a lower bound is given by D(n) + m - 1 where *m* is the smallest divisor of k + 1 larger than 1.



Fig. 1. The energy interpretation of the mancala configuration 5 3 2 (first drawing, t = 0), and some steps of its evolution until 12 mancala moves (last drawing, t = 12). The central dot identifies a *gap*.

We notice that if all energy levels up to k are completely filled, then a mancala move does not change those levels (up to a rearrangement of the seeds), and all the action takes place from the lowest level that is not completely filled, say k + 1, up to the highest level that is not completely empty. The *lowest active level* is the lowest level that is not completely filled.

The overall energy of a configuration (the sum of the individual energy of every seed) does not change after the first part of the move (the rotation within each energy level) whereas it can decrease during the second part (vertical downward shift).

Definition 2.11 (*Gap*). We shall call *gap* an empty cell at the lowest energy level, say *k*, that is not completely filled. We shall think of a gap as an *absent seed*, and in this respect we can *follow* a gap through mancala moves as it rotates left along its energy level.

A gap can be filled when it receives a seed that falls from above after a move, and this can happen only when the gap is in column *k*.

Remark 2.3. The filling of a gap at the lowest energy level that is not completely filled is permanent, since there is no possibility of further energy decrease for seeds at that energy level. All gaps will be eventually filled, provided there are sufficiently many seeds at energy levels larger than *k*.

3. Lower bound

In order to prove a good lower bound $D^*(n)$ for the depth D(n) we need to find appropriate configurations and to compute the number of moves required to reach a periodic configuration (which is an augmented marching group).

3.1. The biaugmented marching group

Let $n = T_k + r$ with $2 \le r \le k + 1$. A *q*-biaugmented marching group is a mancala configuration $\lambda \in \Lambda_n$ of the form $\lambda = \mathfrak{mg}^k + \varepsilon$ where \mathfrak{mg}^k is the marching group of order k and $\varepsilon = (\varepsilon_i)_i$ is an increment of the form

$$\begin{cases} \varepsilon_1 = 2\\ \varepsilon_i = 1 \\ \varepsilon_{q+2} = 0\\ \varepsilon_i \in \{0, 1\} \end{cases} \text{ for } 2 \le i \le q+1$$

for some $0 \le q < k$. The second (resp. fourth) row in the definition is void if q = 0 (resp. if q = k - 1). Clearly such a configuration is not an augmented marching group. Suitable choices of q and ε (compatible with the definition above) allow to obtain all the values of r such that $2 \le r \le k + 1$.

Proposition 3.1. A *q*-biaugmented marching group has depth q(k+2)+2, i.e. it becomes an augmented marching group exactly after q(k+2)+2 moves.

Proof. If q = 0, a direct check shows that after two moves we obtain an augmented marching group. If q > 0, let us *color* the seed³ in the first hole which is at the energy level k + 2 (the highest energy level); we agree that the colored seed is the

³ The idea of the colored seed is due to Bouchet [2, Section 3].

last to be sown in the mancala move, so that it moves at the energy level k + 2 until it falls on the level k + 1, when the configuration becomes an augmented marching group. Notice that the lowest active level is k + 1.

The colored seed is again at the leftmost position after exactly k + 2 moves, while k + 1 is the period of the seeds moving in the lowest active level. Hence the first "hole" at level k + 1, which were in position q + 2 at the beginning, gets one step closer to the colored seed every k + 2 moves. At the end, the configuration becomes a 0-biaugmented marching group in q(k + 2) moves, and in two further moves it becomes periodic. \Box

As an example, in Fig. 1 it is shown the energy-level interpretation of a 2-biaugmented marching group with k = 3 and its evolution. The gray circle in the pictures is the colored seed of the proof.

3.2. The biaugmented marching group of the second kind

Let $n = T_k + r = T_{k-1} + k + r$ with $0 \le r \le k - 2$. A *q*-biaugmented marching group of the second kind is a configuration $\lambda \in \Lambda_n$ of the form $\lambda = \mathfrak{mg}^{k-1} + \varepsilon$, where \mathfrak{mg}^{k-1} is the marching group of order k - 1 and $\varepsilon = (\varepsilon_i)_i$ is an increment of the form

 $\begin{cases} \varepsilon_i \in \{1, 2\} & \text{ for } 1 \leq i < k - q - 1 \\ \varepsilon_{k-q-1} = 2 \\ \varepsilon_i = 1 & \text{ for } k - q \leq i < k \\ \varepsilon_k = 0 \end{cases}$

for some $0 \le q < k - 1$. Again, such a configuration is not an augmented marching group.

Proposition 3.2. The depth of a q-biaugmented marching group of the second kind is given by t = (q + 1)k, i.e. it becomes an augmented marching group precisely after (q + 1)k moves.

Proof. Let us color red the seeds which are at the beginning in the energy level k + 1, and blue the ones in the energy level k, which is the lowest active level. As in the proof of the previous proposition, we agree that the colored seeds are the last to be sown in every mancala move, the red seed (if present) being sown after the blue one. In that way, the blue seeds remain forever at the level k, while one of the red seeds will eventually lose one energy level, when the configuration becomes an augmented marching group.

We track the evolution of the colored seeds, which move in the two highest energy levels. The pattern of red seeds rotates with a period k + 1, while the pattern of blue seeds rotate with a period of k; hence the pattern of red seeds slowly slides one position to the right with respect to the pattern of blue seeds every k moves, and the value of q decreases by one. When q = 0, it is easy to see that after k moves a red seed reaches the lower level. Hence the initial configuration takes exactly (q + 1)k moves in order to reach periodicity (see Fig. 2 for an example).

3.3. Evolution of the Heaviside configuration

Definition 3.1 (*Heaviside Configuration*). We define the *Heaviside configuration* $\mathfrak{h}^n \in \Lambda_n$ by

$$\mathfrak{h}_i^n = \begin{cases} 1 & \text{if } 1 \le i \le n, \\ 0 & \text{if } i > n. \end{cases}$$

We shall also study the special configuration \mathfrak{h}^{∞} with a single seed in all holes $i \geq 1$: although \mathfrak{h}^{∞} is not in Λ since it has unbounded support, the sowing \mathcal{M} is still well-defined.

Proposition 3.3. Let $k \ge 1$ and $t = T_k + r$ with $0 \le r \le k$. Then the configuration after t moves starting from \mathfrak{h}^∞ is given by

$$\lambda_i^{(t)} = \begin{cases} k - i + 2 & \text{if } 1 \le i \le k - r + 1 \\ k - i + 3 & \text{if } k - r + 1 < i \le k + 1 \\ 1 & \text{if } i > k + 1. \end{cases}$$
(3.1)

Proof. It is easy to check that exactly at every move T_{j-1} the configuration becomes greater than the marching group mg^j , hence we have the result for r = 0. In particular, for any $0 \le r \le k$ the leftmost element of the sequence at time $t = T_k + r$ is given by $\lambda_1^{(t)} = k + 1$. Hence in the subsequent move we have to add one seed in the holes with indices i = 2, ..., k + 2 and then shift all piles of one position to the left (dropping the leftmost). If r < k we get

$$\lambda_i^{(t+1)} = \lambda_{i+1}^{(t)} + 1 \quad \forall i = 1, \dots, k+1$$

and the result follows by induction. If r = k, then $t + 1 = T_k + k + 1 = T_{k+1}$ and we have to add one seed to the piles with indices i = 2, ..., k+2, which contain the values k+1, k, k-1, ..., to begin with. Again the result follows immediately. \Box



Fig. 2. Some frames of the evolution of the *q*-biaugmented marching group of the second kind 7 6 4 3 2, where q = 3, n = 22, k = 6. The labels R and B stand for "red" and "blue", resp. In the last diagram the configuration becomes an augmented marching group, exactly at move 24 = (q + 1)k.

Remark 3.1. Notice that at time $\bar{t} = T_{k-1} + r$, $0 \le r \le k$ we moved at most k seeds $(\lambda_1^{(\bar{t}-1)} \le k)$, so that we do not touch the piles at position i > k, which came from \bar{t} shifts to the left. This means that Proposition 3.3 holds for a starting configuration \mathfrak{h}^n with $n = T_k + r$ (we denote it by 1^n) until time \bar{t} , with truncation to zero for elements with index i > k. Moreover, the number of seeds in the leftmost pile (index i = 1) is exactly k at times between $t = T_{k-1}$ and $t = T_k - 1$ and is k + 1 at time $t = T_k$ if r > 0.

Remark 3.2. In the special case $n = T_k + 1$ (r = 1) we can exactly describe the resulting configuration at time $t = T_k$. This is done by observing that the number of seeds in the first pile (Remark 3.1) is the same as for the full \mathfrak{h}^{∞} sequence up to time T_k , meaning that the sowing process is exactly the same. We can thus recover the resulting sequence at time T_k by subtracting from (3.1) the missing seeds in their expected position (2, 3, ...) obtaining

$$\lambda_i^{(T_k)} = \begin{cases} k+1 & \text{if } i=1 \\ k-i+1 & \text{if } 1 < i \le k+1 \\ 0 & \text{if } i > k+1 \end{cases}$$

or equivalently $\lambda^{(T_k)} = \mathfrak{mg}^k + \delta^1$ where δ^1 is the sequence with one in position 1 and zero elsewhere.

3.4. The augmented Heaviside sequences

Now we want to analyze the evolution of the augmented sequence obtained from \mathfrak{h}^n by adding a seed at the position with index *m*. We can write the sequence as

$$\mathfrak{h}^n + \delta^m$$

where δ^m denotes the sequence with a 1 at the position *m* and zero elsewhere. The analysis can be conveniently done by using again Bouchet's trick, see Footnote 3. In this way the evolution of the augmented configuration coincides exactly with the evolution of the non-augmented configuration with the addition of the colored seed, to be positioned appropriately.

At each sowing the colored seed moves one position to the left; when it reaches the leftmost pile (index i = 1), at the next move its new position depends on how large the pile is. We however know precisely how the Heaviside sequence evolves, so that we can explicitly predict the position of the colored seed at each time.

We are particularly interested in the case where $n = T_k + 1$ and $m = T_p + 1$ with $1 \le p \le k$.

Lemma 3.4. The colored seed in the evolution of the augmented Heaviside sequence $\mathfrak{h}^n + \delta^m$ with $n = T_k + 1$ and $m = T_p + 1$, $1 \le p \le k$ is located at position i = k - p + 1 at time $t = T_k$.

Proof. After T_p moves the colored seed reaches the leftmost pile which, according to Proposition 3.3, will contain also p + 1 noncolored seeds. This is true even in the special case p = k, where the colored seed is added at the rightmost nonempty pile of the sequence \mathfrak{h}^n , see Remark 3.1. This covers the case p = k. If p < k, at time $T_p + 1$ the colored seed will go into the hole with index p + 2 and will be again into the leftmost pile at time $t = T_p + p + 2 = T_{p+1} + 1$. If p + 1 < k we can again resort to Proposition 3.3 and conclude that the leftmost pile will contain the colored seed and p + 2 noncolored seeds at that time. This argument can be repeated for q cycles as long as p + q < k with the leftmost pile containing p + q + 1 noncolored seeds at time $t = T_{p+q} + q$. By taking q = k - p - 1, the largest admissible value, we have the colored seed in the leftmost pile at time $t = T_{k-1} + k - p - 1$ together with k noncolored seed. At the next time $t = T_{k-1} + k - p$ the colored seed will be sown in position k + 1 and will move left of one position in the next p moves so that at time $t = T_{k-1} + k = T_k$ it will be in the hole at position k - p + 1, which concludes the proof. \Box

The method of coloring seeds can be also used with more than one seed, provided that we do not place two or more colored seeds in the same hole and as long as the colored seeds do not interact (staying all in different piles). Thanks to Lemma 3.4 we indeed have noninteracting colored seeds for T_k moves if we place them in positions with indices of the form $T_p + 1$ for a set of values $1 \le p \le k$: if two of them would land in the same pile during the evolution, then the expected position given by Lemma 3.4 would be the same, which is not the case. We can thus exactly predict their position at time $t = T_k$, there will be a colored seed in position *s*, if and only if there was a colored seed in position $T_{k-s+1} + 1$. If this is true for $s = 1, \ldots, q + 1$ (with the addition of q + 1 colored seeds), recalling Remark 3.2 we obtain a *q*-biaugmented marching group. By Proposition 3.1, the biaugmented marching group will become periodic after exactly q(k + 2) + 2 further moves.

We want to construct, for a given $n = T_k + r$, a configuration that takes as long as possible to become periodic. In view of the previous discussion to achieve this goal it is mostly convenient to place the colored r seeds in the rightmost positions with index of the form $T_p + 1$.

Definition 3.2. We call *augmented Heaviside sequence* the configuration of mass $n = T_k + r$, $1 \le r \le k + 1$, and length $\ell = T_k + 1$, which is obtained from \mathfrak{h}^{ℓ} by adding a seed in the rightmost r - 1 positions having index of the form $T_p + 1$ (in particular there will be no added seeds if r = 1).

As an example, the augmented Heaviside sequence with mass $n = 19 = T_5 + 4$ is given by

1111112111211112.

The preceding arguments give a proof of the following

Proposition 3.5. Let $n = T_k + r$ with $2 \le r \le k + 1$. Then the corresponding augmented Heaviside sequence becomes an (r-2)-biaugmented marching group after T_k moves. In particular, its evolution becomes periodic after exactly $D_1^*(n)$ moves with

$$D_1^*(n) := T_k + (r-2)(k+2) + 2. \tag{3.2}$$

The choice r = k + 1 is a limit case corresponding to $n = T_{k+1}$, and we have $D_1^*(T_{k+1}) = T_k + (k-1)(k+2) + 2 = 3T_k$. Moreover if we write $n = T_{k+1} - s$ with $0 \le s < k$, then s = k + 1 - r and $D_1^*(n)$ in (3.2) can be equivalently written as $D_1^*(n) = 3T_k - s(k+2)$.

Direct inspection allows to show that for r < k + 1 the particular biaugmented marching group at which we arrive at time T_k evolves into an augmented marching group having period of length exactly k + 1 (and not a proper divisor). Hence the diameter of the evolution is $D_1^*(n) + k$.

3.5. Truncated augmented Heaviside sequences

We now consider the augmented Heaviside sequence corresponding to $n = T_k + k + 1 = T_{k+1}$: it has length k + 1, contains an increment at all the positions of the form $T_p + 1$, $1 \le p \le k$ and the total number n of seeds is itself a triangular number. In the rightmost positions, this sequence has a 2 followed by k - 1 ones and a final 2. Given $0 \le r \le k - 1$, let us remove from this sequence the rightmost k - r piles, obtaining a configuration with mass $T_k + r$. Equivalently, we can obtain the same configuration starting from the augmented Heaviside sequence corresponding to $n = T_{k-1} + k = T_k$ and joining to the right a sequence of r piles with one seed each.

Definition 3.3. Let us call truncated augmented Heaviside sequence such a configuration.

As an example, the truncated augmented Heaviside sequence with mass $n = 17 = T_5 + 2$ is given by

121211211211

Proposition 3.6. The truncated augmented Heaviside sequence with $n = T_k + r$, $0 \le r \le k - 2$, becomes a (k - r - 2)-biaugmented marching group of the second kind after T_{k-1} moves.

Proof. Simply color the rightmost sequence of ones and track their position during the first T_{k-1} moves of the augmented Heaviside sequence of Proposition 3.5 with k - 1 in place of k and r = k. \Box

Corollary 3.7. A truncated augmented Heaviside sequence with $n = T_k + r$, $0 \le r \le k - 2$, becomes periodic after exactly

$$D_2^*(n) := T_k + k(k - r - 2) = 3T_{k-1} - rk$$
(3.3)

moves.

If $n = T_k + k$ or $n = T_k + k - 1$, the function $D_2^*(n)$ is not defined by the previous corollary, and we shall conventionally set it to -1.

As for the augmented Heaviside sequences of Section 3.4, the special biaugmented marching group of the second kind at which we arrive if r > 0 evolves to an augmented marching group with period exactly k + 1, and again we have a diameter given by $D_2^*(n) + k$.

3.6. The special "plateau" sequence

In the special case $n = T_{2r} + r$, $r \ge 1$, we need a further type of configuration.

Definition 3.4. For $n = T_{2r} + r$, $r \ge 1$, we define the *biaugmented Heaviside sequence* in the following way: take the truncated augmented Heaviside sequence for $n - 1 = T_{2r} + r - 1$, and then add a third seed in the pile with two seeds at position $i = T_{2r-1} + 1$.

As an example, $n = 24 = T_6 + 3$ produces the sequence

Proposition 3.8. The biaugmented Heaviside sequence with $n = T_{2r} + r$ becomes periodic after exactly

$$D_3^*(n) := T_{2r} + 2r(r-1) \tag{3.4}$$

moves.

Proof. Let us color the extra added third seed in the pile at position $i = T_{2r-1} + 1$. By construction, without the colored seed the biaugmented Heaviside sequence becomes a truncated augmented Heaviside sequence with mass $T_{2r} + r - 1$, hence by Proposition 3.6, after T_{2r-1} moves we have an (r - 1)-biaugmented marching group of the second kind with an extra seed (corresponding to an increment of +3 with respect to the marching group) in the leftmost pile. The argument we used to study the biaugmented marching groups can be adapted to the present case, proving that the periodic configuration is reached exactly after the same number of moves as for the biaugmented marching group of the second kind: (q + 1)k with q = r - 1 and k = 2r. Hence the initial configuration becomes periodic after

$$T_{2r-1} + 2r^2 = T_{2r-1} + 2r + 2r(r-1) = T_{2r} + 2r(r-1).$$

Also in this case a direct inspection of the resulting augmented marching group allows to show that its period is exactly k + 1, and again the diameter of the special plateau sequence is $D_3^*(n) + k$.

The lower bound $D_3^*(n)$ is only defined for special values of n, of the form $n = T_{2r} + r = 4T_r$, and we conventionally set it to -1 in all the other cases.

We can now gather the three lower bounds and define

$$D^*(n) := \max\{D_1^*(n), D_2^*(n), D_3^*(n)\}$$

so that the previous results can be summarized as the following:

Theorem 3.9 (Lower Bound).

$$D(n) \ge D^*(n).$$

Proof. Simply gather the results obtained in Propositions 3.5, 3.8 and Corollary 3.7.

A plot of the function $D^*(n)$ for $n \le 45$ can be seen in Fig. 3.

Remark 3.3. There is an explicit way to express the lower bound $D^*(n)$. Indeed, writing $n \ge 1$ in the form $n = T_k + r$ with $k \ge 0$ and $1 \le r \le k + 1$, one can prove that

$$D^{*}(n) = \begin{cases} D_{2}^{*}(n) = T_{k} + k(k - r - 2) & \text{if } r < k/2\\ D_{3}^{*}(n) = T_{2r} + 2r(r - 1) & \text{if } r = k/2\\ D_{1}^{*}(n) = T_{k} + (r - 2)(k + 2) + 2 & \text{if } r > k/2. \end{cases}$$

Notice that the middle case occurs only when *k* is even.

(3.6)

(3.5)





4. s-monotonicity

In this section we introduce one of the main notions of the paper, namely *s*-monotonicity. Firstly we define a sequence which is strictly related to the energy levels interpretation of the mancala game.

Definition 4.1 (*Energy Sequence*). Given a mancala configuration $\lambda \in \Lambda$ we define the *energy sequence* $e = \varepsilon(\lambda) : \mathbb{Z} \to \mathbb{N}$ as follows: setting $m = m(\lambda) = \max\{\lambda_1, \operatorname{len}(\lambda)\}$, we define

$$e_i = \lambda_i + i - 1$$
 $i = 1, \dots, m$ (4.1)

and extend it on \mathbb{Z} as a periodic function of period *m*.

Notice that if $\lambda_1 > \text{len}(\lambda)$ we actually use values of λ outside its support, so that we have $e_i = i - 1$ for $\text{len}(\lambda) < i \leq \lambda_1$.

Remark 4.1. We have that $\lambda \geq \mathfrak{mg}^k$ if and only if $\varepsilon(\lambda)_i \geq k$ for all $i \in \mathbb{Z}$.

Definition 4.2 (*s*-monotonicity). Given $s \ge 1$ and $\lambda \in \Lambda$, we say that λ is *s*-monotone if:

(i) $\lambda_1 \leq \text{len}(\lambda) + 1$, and

(ii) its energy sequence $e = \varepsilon(\lambda)$ satisfies the inequality

$$e_j \leq e_i + 1 \quad \forall i < j \leq i + s$$

(notice that the first requirement $\lambda_1 \leq \text{len}(\lambda) + 1$ is redundant if $s \geq 2$).

It is convenient to introduce also the (void) notion of 0-monotonicity, satisfied by all mancala configurations.

Definition 4.3 (*Increasing Plateau*). Let $s \ge 0$. An *increasing plateau of length* s is a subset of (s + 2) consecutive values $e_i, e_{i+1}, \ldots, e_{i+s+1}$ of an energy sequence such that

$$e_i < e_{i+1} = \dots = e_{i+s} < e_{i+s+1}$$
 if $s \ge 1$
 $e_i \le e_{i+1} - 2$ if $s = 0$.

It is readily proved that a configuration is *s*-monotone if and only if each increasing plateau has length at least *s*. Let us give a few examples of *s*-monotone configurations:

- 1 2 is 0-monotone but not 1-monotone
- 1 1 1 is 1-monotone but not 2-monotone
- 4 1 1 and 2 2 1 1 are 2-monotone but not 3-monotone
- 5 2 2 1 and 3 3 2 1 1 are 3-monotone but not 4-monotone
- 6 3 3 2 1 and 4 4 3 2 1 1 are 4-monotone but not 5-monotone.

(4.2)

Remark 4.2. If a mancala configuration is s_2 -monotone, then it is also s_1 -monotone for any $s_1 < s_2$. Moreover, a configuration is an augmented marching group if and only if it is *s*-monotone for all $s \in \mathbb{N}$.

Remark 4.3. The notions of monotonicity, in the sense of Definition 2.7, and of 1-monotonicity, in the sense of the previous definition, are slightly different. For instance, the sequence 4–2 is monotone but not 1-monotone. However, if a configuration λ is monotone, then $\mathcal{M}(\lambda)$ is 1-monotone. Hence, in view also of next theorem, during the evolution of a configuration there is at most one time when it is monotone but not 1-monotone.

In passing, we observe that any 1-monotone sequence has a predecessor that is merely monotone, *i.e.* a 1-monotone configuration can be obtained by applying a move to some monotone configurations. Hence the maximal depth/diameter in the family of monotone configurations is exactly one plus the maximal depth/diameter in the family of 1-monotone configurations.

The following theorem shows that the class of *s*-monotone configurations is closed under mancala moves.

Theorem 4.1. If λ is s-monotone, then $\mathcal{M}(\lambda)$ is s-monotone.

Proof. Let us set $m = \max{\{\lambda_1, \operatorname{len}(\lambda)\}}, \mu = \mathcal{M}(\lambda), e = \varepsilon(\lambda) \text{ and } f = \varepsilon(\mu).$

Clearly any mancala configuration is at least 0-monotone, hence the set of 0-monotone configurations is closed under mancala moves. A direct check shows that the property $\lambda_1 \leq \text{len}(\lambda) + 1$ is preserved under a mancala move whenever λ is a monotone configuration, indeed we have $\lambda_2 \leq \lambda_1 \leq \text{len}(\mu)$, whence $\mu_1 = \lambda_2 + 1 \leq \text{len}(\mu) + 1$. Consequently the set of 1-monotone configurations is closed under mancala moves.

Now assume that s > 1, hence we have to check only the values of the energy sequence. We divide the proof in five cases.

Case $\lambda_1 \ge \text{len}(\lambda)$. We have $m = \text{len}(\mu) = \lambda_1$ and the new period is $m(\mu) = \max\{\mu_1, \text{len}(\mu)\} = \max\{\lambda_2 + 1, \lambda_1\}$. Since s > 1, then λ is monotone and $\lambda_2 \le \lambda_1$, hence we have $m \le m(\mu) \le m + 1$ and the new energy sequence f is obtained by a left-shift of e possibly followed by the insertion of a new value. Indeed $f_i = \mu_i + i - 1 = \lambda_{i+1} + 1 + i - 1 = e_{i+1}$ for $i = 1, \ldots, m - 1$ and $f_m = \mu_m + m - 1 = m$, since $\mu_m = 1$. If $\lambda_2 < \lambda_1$, then $m(\mu) = m$ and the new energy sequence f is exactly a left-shift of e, so that μ is itself s-monotone. If $\lambda_2 = \lambda_1$, then $m(\mu) = m + 1$ and the new value to be inserted is $f_{m+1} = \mu_{m+1} + m + 1 - 1 = m$ (being $\mu_{m+1} = 0$) and it is equal to the previous value f_m . Since inserting a value in the sequence cannot destroy the s-monotonicity property if it is equal to the previous (or the next) value, we can conclude that μ is s-monotone.

Case $\lambda_1 \leq \text{len}(\lambda) - 3$. We have $m = \text{len}(\lambda)$ and $\text{len}(\mu) = \text{len}(\lambda) - 1$. Since $\lambda_2 \leq \lambda_1$ we have $\mu_1 \leq \lambda_1 + 1 \leq \text{len}(\lambda) - 2 = \text{len}(\mu) - 1$, so that $f_1 \leq m - 2$ and $m(\mu) = m - 1$. Moreover, $\lambda_1 < m - 1$ implies that $\mu_{m-1} = \lambda_m$, so that $f_{m-1} = \mu_{m-1} + m - 2 = \lambda_m + m - 2 \geq m - 1$. Using the inequalities above we conclude

$$f_1 < f_{m-1} = f_0. \tag{4.3}$$

The energy sequence *f* is obtained from the energy sequence *e* with the following operations:

- (1) translate the values of e to the left of one position $e'_i = e_{i+1}$. This has no impact on the s-monotonicity;
- (2) decrease by one the values with indices in the range $\lambda_1 + 1, \ldots, m$;
- (3) remove the value at index *m* and extend to \mathbb{Z} with period m 1.

The last two operations can only impact s-monotonicity in the case of indices $i \le 0 < j$ with $j - i \le s$ (up to addition of multiples of m - 1). However using (4.3) we have $f_j - f_i < (f_j - f_1) + (f_0 - f_i) \le 1 + 1$.

Case $\lambda_1 = \text{len}(\lambda) - 2$. The argument is similar to the previous, but (4.3) now becomes

$$f_1 \le f_{m-1} = f_0. \tag{4.4}$$

Moreover the set of indices (within the period $\{1, ..., m\}$) where we decrease the energy by one (step 2 above) reduces to $\{m - 1, m\}$ and is further reduced to the single index $\{m - 1\}$ after removal of the energy value with index m (step 3 above). Now suppose by contradiction that $f_j - f_i \ge 2$ with $i \le 0 < j, j - i \le s$. This is only possible if both $f_j - f_1 = 1$ and $f_0 - f_i = 1$. The latter equation implies i < 0 and, in view of the discussion above, it can be rewritten as $1 = f_{m-1} - f_{m-1+i} = e_m - 1 - e_{m+i}$ which contradicts the *s*-monotonicity of λ .

Case $\lambda_1 = \text{len}(\lambda) - 1$ **and** $\lambda_2 < \lambda_1$. It turns out that $m(\mu) = \text{len}(\lambda) - 1 = m - 1$. Moreover $e_0 = e_m = \lambda_m + m - 1 \ge m$ whereas $e_1 = \lambda_1 < m$, so that $e_1 < e_0$. The resulting energy f can be obtained by a left shift of e followed by removal of element at index m, which is the same as first removing the element at index 1 of e and then performing a left shift. The fact that $e_0 > e_1$ allows to ensure that the removal of the value e_1 does not impact the *s*-monotonicity, which is then unaffected by the final left shift.

Case $\lambda_1 = \text{len}(\lambda) - 1$ **and** $\lambda_2 = \lambda_1$. It turns out that in this case $m(\mu) = m(\lambda) = m = \text{len}(\lambda)$ and that the energy sequence of μ is just a left-shift of the energy sequence of λ : $f_i = e_{i+1}$.

These five cases cover all possible situations so that we conclude the proof. \Box

For particular configurations it is possible to prove that repeated mancala moves actually increase the order of monotonicity.

The following is our first important result for which the distinction between monotonicity and 1-monotonicity is crucial. For instance, the result is false for the sequence 4 2, which is monotone but not 1-monotone.

Lemma 4.2. Let $\lambda \in \Lambda$, $s \in \mathbb{N}$ and $2 \leq q \in \mathbb{N}$ such that

(i) λ is s-monotone;

(ii) $\lambda \geq \mathfrak{mg}^{q-1}$;

(iii) $len(\lambda) = q - 1$, (*i.e.* $\lambda_q = 0$).

Then $\mathcal{M}^{q}(\lambda)$ (the configuration after q moves) is (s + 1)-monotone.

Proof. Denote by μ the configuration obtained after q moves. Since we apply a number of moves that is larger than the length of the sequence, it follows that μ is at least 1-monotone (it is already monotone after q - 2 moves, hence 1-monotone after q - 1 moves by Remark 4.3). This covers the case s = 0, and we can suppose $s \ge 1$. In particular we can assume that $q - 1 \le \lambda_1 \le q$ where the first inequality follows from requirement (ii) and the second inequality comes from (iii) and constraint (i) of Definition 4.2.

Using the energy level interpretation of the game, we want to track the position after q moves of the seeds with energy larger than q. First observe that requirement (ii) implies that all energy levels up to the (q - 1)th are completely filled and that seeds at level q (which is not completely filled) are subjected to q rotations to the left and end up at the same initial position. A seed in column i (notice that requirement (iii) implies that i < q) at level q + j is also rotated q times to the left, however the (q + j)th energy level has length q + j, so that it will end up in column i + j and can possibly decrease its energy (that is, its height) once or more than once. Requirements (i) and (iii) imply that there is no seed at height q + 1 in the first s columns, whereas seeds in column s + 1 and height at least q + 1 get moved to the right after the set of q moves. This allows to conclude that $f_j - f_i \le 1$ for all $i \le 0 < j, j - i \le s + 1$, where $f = \varepsilon(\mu)$ is the energy sequence of μ given by Definition 4.1. If there were no energy decrease, then all increasing plateaus would increase their length by 1 and the (s + 1)-monotonicity would follow. The decrease in energy after each move involves all seeds in the final columns and it can be easily checked that it maintains the (s + 1)-monotonicity. \Box

Definition 4.4 (*q*-canonical Configurations). Let $2 \le q \in \mathbb{N}$, $n \ge T_q$ and $\lambda \in \Lambda_n$. We say that λ is *q*-canonical if

- (1) $\lambda \geq \mathfrak{mg}^{q-1}$;
- (2) $\operatorname{len}(\lambda) = q 1$, in particular λ is not $\geq \mathfrak{mg}^{q4}$;
- (3) the gap (see Definition 2.11) in the rightmost position (column q) in the energy level at height q is the last to be filled after repeated moves. Notice that requirement $n \ge T_q$ implies that the qth energy level will eventually fill up completely.

Due to the central importance of *q*-canonical configurations a few remarks are in order.

Remark 4.4. The value *q* of a *q*-canonical configuration is necessarily the height of the lowest active energy level (the lowest level that is not completely filled). This follows from requirement (1) and (2) of the above definition. Hence a configuration cannot be *q*-canonical with two different values of *q*.

Remark 4.5. In view of the comments following Definition 2.11, in particular the fact that *filling a gap* is permanent, we can indeed define a *filling order* on the gaps at level q. All the gaps (at level q) will be eventually filled due to requirement $n \ge T_q$, and we can identify the gap that will be filled last.

Remark 4.6. An augmented marching group cannot be *q*-canonical, since if $\lambda \in \Lambda_n$ is an augmented marching group with $n \ge T_q$ then $\lambda \ge \mathfrak{mg}^q$. This is incompatible with requirement (2) of Definition 4.4.

Remark 4.7. Since the gap in the rightmost position of a *q*-canonical configuration λ returns in the same position after every sequence of exactly *q* moves, it follows that the *q*th energy level fills up after exactly *rq* moves, for some $r \in \mathbb{N}^*$. In particular $\mathcal{M}^{rq}(\lambda) \geq \mathfrak{mg}^q$. Moreover $\mathcal{M}^{iq}(\lambda)$ is itself a *q*-canonical configuration for all $i = 0, \ldots, r - 1$ and there are no other *q*-canonical configurations in between.

As an example, let us consider the configuration $\lambda = 55362$, it satisfies requirements (1) and (2) of Definition 4.4 with q = 6 (see Fig. 4 for an energy-level interpretation), it has three gaps at the energy level 6 with three seeds at higher level, enough to completely fill the sixth energy level ($|\lambda| = 21 = T_6$). However this configuration is not a 6-canonical configuration because the rightmost gap in column 6 is not the last filled. This can only be seen by letting the configuration evolve and annotating the filling order of the three gaps. In this example only seven mancala moves are sufficient to see that the gap in column 6 is filled first, whereas the gap that is filled last is the one originally in column 3. The configuration $\lambda = 95322$ (obtained from the previous one with three mancala moves) is indeed a 6-canonical configuration. As noted in Remark 4.7, after six mancala moves we have either another 6-canonical configuration or the sixth level becomes completely

⁴ Using requirement (1), the energy level q is the lowest that is not completely filled.



Fig. 4. The configuration 5 5 3 6 2, shown on the left with its energy-level interpretation, is not a 6-canonical configuration. The three gaps at level 6 are filled in the order given by the corresponding numbers. After three mancala moves we obtain the configuration 95322 (middle picture), which is a 6-canonical configuration. Another 6-canonical configuration is obtained after 6 further moves.

filled. It turns out that we have a total of four other 6-canonical configurations, one after each sequence of six moves starting from the 6-canonical configuration 9 5 3 2 2.

Theorem 4.3. A q-canonical configuration cannot be (q - 1)-monotone. In other words it is at most (q - 2)-monotone.

Proof. Let λ be a *q*-canonical configuration, $q \ge 2$; hence $\lambda_1 \ge q - 1$. Now suppose by contradiction that λ is also (q - 1)-monotone; this implies that $\lambda_1 \le q$ by (i) of Definition 4.2. Set $e = \varepsilon(\lambda)$ (the energy sequence), with period $m(\lambda) = \lambda_1$, hence

 $q-1 \leq m(\lambda) \leq q.$

Either $e_1 = \lambda_1 = q - 1$ or $e_q = q - 1$. Energy levels strictly larger than q must also be present, otherwise λ would be an augmented marching group, incompatible with q-canonicity. Hence there must exist an increasing plateau, and its length cannot be longer than $m(\lambda) - 2$ due to the periodicity of e, which is at most q - 2. On the other hand, (q - 1)-monotonicity implies that the length of every increasing plateau is at least q - 1, a contradiction. \Box

5. Upper bound

5.1. Analyzing the evolution of a configuration

In this section we shall fix $n \ge T_k$ for some $k \in \mathbb{N}$ and consider a mancala configuration $\lambda \in \Lambda_n$. We let λ evolve (applying mancala moves) until it reaches an augmented marching group; we denote by $\lambda(t) = \mathcal{M}^t(\lambda)$ the result of $t \in \mathbb{N}$ mancala moves, so that in particular $\lambda(0) = \lambda$. Let \overline{t} be the smallest time t for which $\lambda(t)$ is an augmented marching group, in particular $\lambda(\overline{t}) \ge \mathfrak{mg}^k$. We also denote by $t^* \le \overline{t}$ the smallest time t such that $\lambda(t) \ge \mathfrak{mg}^k$.

During the evolution $\lambda(t)$ will occasionally become a *q*-canonical configuration. Specifically we indicate with $t_1 < t_2 < \cdots < t_s < t^*$ the times before t^* when we have a *q*-canonical configuration for some $q \in \mathbb{N}$, $\lambda(t^*)$ can be itself *q*-canonical, however $\lambda(t^*) \ge \mathfrak{mg}^k$ would imply q > k, hence $n \ge T_{k+1}$. In other words, if we restrict $T_k \le n < T_{k+1}$, then $\lambda(t^*)$ cannot be *q*-canonical. We call such times *critical times*. Specifically we have a sequence $q_1 \le q_2 \le \cdots \le q_s$ such that $\lambda(t_i)$ is a q_i -canonical configuration.

Consider two consecutive critical times t_i and t_{i+1} . Remark 4.7 applied to $\lambda(t_i)$ tells us that after q_i moves, we have either another q_i -canonical configuration, with no other q-canonical configurations in between, in which case $t_{i+1} = t_i + q_i$ and $q_{i+1} = q_i$, or we have just filled the q_i th energy level so that we have $t_{i+1} \ge t_i + q_i$ and $q_{i+1} > q_i$. In both cases we have $t_{i+1} - t_i \ge q_i$. We can now apply Lemma 4.2 to $\lambda(t_i)$ with $q = q_i$ and conclude that the *s*-monotonicity of $\lambda(t_{i+1})$ increases strictly. Since $\lambda(t_1)$ is 0-monotone, we obtain by induction that $\lambda(t_{i+1})$ is *i*-monotone and that $\lambda(t^*)$ is *s*-monotone.

It is possible for the set of critical times to be empty. In this case we shall see that the filling up of the energy levels is very fast.

Lemma 5.1. If $\lambda \ge \mathfrak{mg}^{q-1}$ and $|\lambda| \ge T_q$, then there exists $0 \le i < q$ such that either $\mathcal{M}^i(\lambda) \ge \mathfrak{mg}^q$ or $\mathcal{M}^i(\lambda)$ is a q-canonical configuration. If λ is itself q-canonical, the result holds true with i = q (besides the trivial value i = 0).

Proof. If the *q*th energy level is full, then there is nothing to prove, otherwise locate the gap that will be filled last in column *j*. If j = q we already have a *q*-canonical sequence, otherwise we apply *j* moves in order to rotate the gap position into column *q*, now we either have a *q*-canonical sequence or we filled the *q*th energy level. The second part of the Lemma easily follows by first performing a single move and then reasoning as in the first part. \Box

Corollary 5.2. If the set of critical times is empty (s = 0), then one has $t^* \leq T_{k-1}$, i.e.

 $\lambda(T_{k-1}) \geq \mathfrak{mg}^k$.

Proof. It can be readily seen by repeated applications of the previous Lemma. \Box

Similarly, the q_1 -canonical sequence $\lambda(t_1)$ is reached after at most T_{q_1-1} moves, i.e. $t_1 \leq T_{q_1-1}$. Now let

$$I = \{i \in \mathbb{N} : 1 \le i \le s - 1, q_{i+1} > q_i\}$$

$$J = \{i \in \mathbb{N} : 1 \le i \le s - 1, q_{i+1} = q_i\}$$

(notice that $I \cup J = \{1, \dots, s - 1\}$), and decompose

$$t_s = t_1 + \sum_{i \in I} (t_{i+1} - t_i) + \sum_{j \in J} (t_{j+1} - t_j).$$

If $i \in I$ we first apply one move to $\lambda(t_i)$ and then invoke Lemma 5.1 $(q_{i+1} - q_i)$ times, concluding that

$$t_{i+1} - t_i \le 1 + T_{q_{i+1}-1} - T_{q_i-1}.$$

On the other hand, if $j \in J$ then $t_{j+1} - t_j = q_j \le q_s$ (Remark 4.7). In the end we have the estimate

$$t_s \le T_{q_{s-1}} + |I| + q_s |I| \le T_{q_{s-1}} + q_s (s-1).$$
(5.1)

At each critical time we gain one level of monotonicity, so that $\lambda(t_s)$ is (s - 1)-monotone. The maximal possible value of q_s is $q_s = k$, which gives an upper bound k - 2 for the *s*-monotonicity of $\lambda(t_s)$ (Theorem 4.3) and hence we have $s \le k - 1$. We get

$$t_s \leq T_{q_s-1} + q_s(k-2).$$

To obtain the configuration $\lambda(t^*) \ge \mathfrak{mg}^k$ we need to apply a move to $\lambda(t_s)$ and again invoke Lemma 5.1 (more than once if $q_s < k$) for a total of at most $1 + (k - 1) + \cdots + (q_s - 1)$ moves. We notice that the last term in the previous sum is also present in the sum that defines T_{q_s-1} , so that using (5.1) we arrive at

 $t^* \leq T_{k-1} + 1 + (q_s - 1) + q_s(s - 1) = T_{k-1} + sq_s \leq T_{k-1} + ks.$

We summarize the previous analysis in the following lemma.

Lemma 5.3. Let $n \ge T_k$ and $\lambda \in \Lambda_n$. Then there exists $0 \le s \le k - 1$ such that after at most $t = T_{k-1} + ks$ moves the resulting configuration $\lambda(t)$ satisfies:

(1) $\lambda(t) \geq \mathfrak{mg}^k$;

(2) $\lambda(t)$ is s-monotone.

If $n = T_k$ or $n = T_k + 1$ then the resulting configuration is an augmented marching group.

Definition 5.1. Given $0 < b \le a$ two integers, we define V(a, b) by

$$V(a, b) = \sum_{i \in \mathbb{N}^*: a - ib > 0} (a - ib).$$

The value of *V* can be explicitly computed by first expressing a = qb + p with $q = \lfloor \frac{a}{b} \rfloor$ and $0 \le p < b$ (Euclidean division) and then using the formula for the sum of the first natural numbers. We end up with $V(a, b) = qb + \frac{b}{2}q(q - 1)$. We have the special values

- V(a, a) = 0, V(a, b) = a b if $a \le 2b$ (only one term in the above sum),
- V(a, b) = 2a 3b > a b if $2b < a \le 3b$ (two terms in the above sum),
- V(a, b) > 2a 3b if a > 3b (more than two terms in the above sum).

Moreover, if b < a then V is strictly increasing with respect to a and strictly decreasing with respect to b. We can then "invert" V with respect to b and define the function

 $W(a, v) = \max\{b \in \mathbb{N}^* : b \le a, V(a, b) \ge v\},\$

which is (weakly) decreasing with respect to v and (weakly) increasing with respect to a. We have the special values

- W(a, v) = a v if $1 \le v \le \frac{a}{2} + 1$,
- $W(a, v) = \lfloor \frac{2a-v}{3} \rfloor$, if $\frac{a}{2} \le v \le a+3$.

Referring to Fig. 5 the value of V(a, b) can be graphically interpreted as the number of unit squares completely included in the right triangle having base (long cathetus) of length *a* and height (short cathetus) of length $\frac{a}{b}$.

With the same graphical interpretation in mind function W(a, v) can be viewed as the maximal value of *b* such that the right triangle of sides *a* and $\frac{a}{b}$ contains at least *v* unit squares.

For $n \ge T_k$, let us consider a *k*-canonical configuration. By Lemma 5.1, after a cycle of *k* moves we have that the resulting configuration is either itself *k*-canonical or it has just become larger than \mathfrak{mg}^k . Iterating this procedure, we want to relate the number of cycles needed to become larger than \mathfrak{mg}^k with the degree of monotonicity of the configuration.



Fig. 5. Five unit squares are completely contained in a right triangle of sides 7 and $\frac{7}{3}$, corresponding to a = 7 and b = 3. In other words 5 = V(7, 3) (see Definition 5.1).

Lemma 5.4. Let $\lambda \in \Lambda_n$ with $n = T_k + r$, $r \ge 1$ and suppose that

- (1) λ is a k-canonical configuration;
- (2) λ is s-monotone for some $s \in \mathbb{N}$.

Let $\sigma \geq 1$ denote the number of cycles of k moves that are required to become larger than \mathfrak{mg}^k and suppose that $s + \sigma \geq 2$. Then

$$r \leq V(k-1, s+\sigma - 1) - 1.$$

Equivalently, reasoning by contradiction, we have the upper bound

$$\sigma \le \max\{W(k-1, r+1) + 1 - s, 0\}.$$

Proof. First we analyze the case $\sigma = 1$, so that $s \ge 1$. Since the configuration is *s*-monotone and *k*-canonical, in this case $\lambda_1 \le k$, hence $e_1, \ldots, e_s \le k, e_{s+1}, \ldots, e_{2s} \le k+1$ and so on, where *e* is the energy sequence $e = \varepsilon(\lambda)$. Then the total number of seeds with energy above *k* cannot be larger than V(k-1, s). This is obtained by observing that we have a sequence of *q* plateaus of length *s* and increasing height from 0 to q - 1 and a final top plateau of length *p* and height *q*. This implies that $r \le V(k-1, s) - 1$ where the "-1" is a consequence of the fact that the *k*th energy level is not completely full. For a generic value of σ we simply perform $\sigma - 1$ cycles of *k* moves obtaining a configuration with level of monotonicity $s + \sigma - 1$ with still $\lambda_k = 0$ and we can resort to the special case discussed above. \Box

5.2. Optimal upper bound

The following lemma will be fundamental in proving the optimal upper bound.

Lemma 5.5. Let $\lambda \in \Lambda_n$ with $n = T_k + r$, $0 \le r \le k$, and denote by σ_q the number of *q*-canonical configurations encountered during the evolution of λ , with q = 2, ..., k.

Then the total number $s = \sum_{q} \sigma_{q}$ of canonical configurations encountered, and hence the least level of monotonicity of the first configuration that becomes larger than (or equal to) mg^{k} , is bounded by

$$s \leq W(k-1, r+1) + 1.$$

In particular, for $r < \frac{k}{2}$ we have

$$s \le k - r - 1 \tag{5.2}$$

and for $\frac{k-3}{2} \leq r \leq k$ we have

$$s \le \left\lfloor \frac{2k-r}{3} \right\rfloor,\tag{5.3}$$

where we used the special values of function W computed in Definition 5.1.

Proof. If $\sigma_q = 0$ for all q, then we can resort to Corollary 5.2 and conclude immediately. Otherwise, set $s_q = \sum_{i \le q} \sigma_i$ and let \bar{q} be the largest value of q such that $\sigma_q > 0$. We apply Lemma 5.4 with $k = \bar{q}$, $r = n - T_{\bar{q}}$ and $s = \sum_{i < \bar{q}} \sigma_i = s_{\bar{q}-1}$ and obtain an upper bound for $\sigma_{\bar{q}}$ that, since $\sigma_{\bar{q}} > 0$ can be written as

$$\sigma_{\bar{q}} \le W(\bar{q} - 1, n - T_{\bar{q}} + 1) + 1 - s_{\bar{q} - 1}$$

or equivalently as

$$s = s_{\bar{q}} \le W(\bar{q} - 1, n - T_{\bar{q}} + 1) + 1 \le W(k - 1, r + 1) + 1$$

where the last inequality follows from the monotonicity of W with respect to its arguments. \Box

The evolution of a configuration $\lambda \in \Lambda_n$, with $n = T_k + r$ and $1 \le r \le k$, will be decomposed into two stages:

(1) the evolution from the beginning up to the first time *t* when $\lambda(t) \ge \mathfrak{mg}^k$, where as usual $\lambda(t) = \mathcal{M}^t(\lambda)$;

(2) the subsequent evolution until we reach an augmented marching group.

1840

The duration of each of these stages will be estimated from above with bounds that depend upon a few parameters and we shall find a uniform bound valid for all possible feasible choices of such parameters finally obtaining an upper bound for D(n).

We need some further lemmas.

Lemma 5.6. Let $\lambda \in \Lambda_n$ and let k be the largest integer such that $\lambda \ge \mathfrak{mg}^k$. Write $n = T_k + r$ and suppose that λ is r-monotone. Then λ is an augmented marching group.

Proof. By contradiction suppose that λ is not an augmented marching group, then the energy sequence has an element with value k (otherwise $\lambda \ge mg^{k+1}$) and an element with value k + 2 (otherwise λ would be an augmented marching group). This implies the existence of an increasing plateau at level k + 1, which must have length at least r. An energy level of k + 1 at position i is possible without using one of the r exceeding seeds only if i = k + 1, however if that position is part of the increasing plateau, then the seed at level k + 2 must necessarily stay on top of a seed at level k + 1. In all cases we need at least r + 1 seeds to build such an increasing plateau, which contradicts the assumptions.

Lemma 5.7. Let $\lambda \in \Lambda_n$ with $n = T_k + r$, $0 \le r \le k$, $\lambda \ge \mathfrak{mg}^k$ and s-monotone. If $\mathcal{M}^{t^{\dagger}}(\lambda)$ becomes an augmented marching group, then

$$t^{\dagger} \le \max\{r - 1 + (r - s - 1)(k + 1), 0\}.$$

Proof. Using the energy level interpretation we identify the gap at level k + 1 that will be filled last, and call i_0 its column index.

We consider two cases.

Case $i_0 < r$ or $i_0 = k + 1$. The first term r - 1 in the estimate is then the maximal number of moves that are required to move that gap into the rightmost (k + 1)th column. We observe that starting from that configuration the augmented marching group will be obtained exactly after a number of moves multiple of k + 1

After each cycle of k + 1 moves the level of monotonicity increases at least of one unit (Lemma 4.2), so that after r - s - 1 cycles of k + 1 moves we obtain a configuration μ that is at least (r - 1)-monotone.

Now we want to prove that (r - 1)-monotonicity (and not higher) is actually impossible under the circumstances. This would imply at least *r*-monotonicity that in turn implies that we have an augmented marching group thanks to Lemma 5.6.

Suppose then by contradiction that μ is (r-1)-monotone but not r-monotone and hence it is not an augmented marching group. This means that the gap at level k + 1, column k + 1 is still empty. The energy sequence of μ has then both elements at level k and elements at level k+2, so that there exists at least an increasing plateau at level k + 1 (i.e. a sequence of energy elements as contiguous positions at level k + 1 following an element at level k and preceding an element at level k + 2). Such plateau has length at least r - 1 due to the (r - 1)-monotonicity. Now recall that we have an excess of r seeds with energy larger or equal to k + 1 (levels k and below are completely filled thanks to the assumptions). Column k + 2 is the only one having energy k + 1 with no need to use one of the excess seeds, however we have barely enough seeds to form an increasing plateau of length r - 1 which requires r + 1 seeds that reduce to r only if we take advantage of the energy step at column k + 2.

Recalling that column k + 1 has energy level k (the gap is still empty), the only possible disposition of the r excess seeds is the one with r - 1 of them in columns 1 to r - 1 (energy k + 1) and the last excess seed itself in column r - 1 (then with energy k + 2). This configuration becomes an augmented marching group after exactly r moves, however we know that an augmented marching group is reached after a multiple of k + 1 moves and hence r = k + 1, a contradiction.

Since (r - 1)-monotonicity is impossible for μ , then we have at least *r*-monotonicity and hence we have an augmented marching group.

Case $r \le i_0 < k + 1$. Let us consider the result of the evolution as soon as it becomes an augmented marching group and call it μ , the position of the "marked" gap is then column k + 1 and it has just been filled with the last of the r seeds the were originally at energy levels larger then k + 1. Immediately to the left of this seed there is a contiguous sequence of, say, $\rho - 1$ seeds, $1 \le \rho \le r$ at level k + 1, preceded by a gap in column $k + 1 - \rho$. Clearly this latter gap (let us call it "gap number 2") can be traced back to its corresponding position in the initial configuration λ , and will originally be empty in λ and will continue to be empty during the whole evolution. It is located in column $i_0 - \rho$ of λ , to the left of the marked gap (in column i_0).

No seed at energy larger then k + 1 (they are the "active" seeds, the only ones that "move" with respect to the k + 1 level) can ever cross or reach the column containing gap number 2, so that the active seeds originally to the left of gap number 2 or to the right of the marked gap (call this region the "outer region") can ever interact with the region between gap number 2 and the active gap.

Without loss of generality we can then suppose that the outer region has reached its final configuration already in λ . This actually means that there are no active seeds to the right of the marked gap.

The final step is then to observe that in this situation λ can be interpreted as the result of the evolution (with $k + 1 - i_0$ moves) from a configuration $\overline{\lambda}$ that itself satisfies the hypotheses of this Lemma and the marked gap already in the k + 1 position (in other words we can "steal" a few moves) and we can resort to the first case of the Lemma.

The proof is thus complete. \Box

We now analyze the evolution of $\lambda \in A_n$, $n = T_k + r$, $0 \le r \le k$. We denote with $s \in \mathbb{N}$ the number of *q*-canonical configurations encountered during the evolution with $q \le k$.

Using Lemma 5.5 we obtain the constraint $s \le W(k - 1, r + 1) + 1$, which in particular gives

$$0 \le s \le k - r - 1 \quad \text{if } r < \frac{\kappa}{2},$$

$$0 \le s \le \left\lfloor \frac{2k - r}{3} \right\rfloor \quad \text{if } \frac{k - 3}{2} \le r \le k + 1,$$

$$0 \le s \le r \qquad \text{if } r = \frac{k}{2}.$$
(5.4)

The third estimate is actually just a special case of the second one; the first and second estimates coincide for *r* in the common interval $\frac{k-3}{2} \le r < \frac{k}{2}$.

As usual we denote by $\lambda(t) = \mathcal{M}^t(\lambda)$ and we set

- t_A the smallest time such that $\lambda(t_A) \ge \mathfrak{mg}^k$
- t^{\dagger} the smallest time such that $\lambda(t^{\dagger})$ is an augmented marching group.

Lemma 5.8. We have the following estimates:

(1) $t_A \le T_{k-1} + ks;$ (2) $t^{\dagger} - t_A = 0$ if $s \ge r;$ (3) $t^{\dagger} - t_A \le r - 1 + (r - s - 1)(k + 1)$ if s < r.

Proof. The first point is exactly Lemma 5.3.

```
Point (2) is simply Lemma 5.6.
Point (3) follows from Lemma 5.7
```

We then have a total estimate of $t^{\dagger} = t_A + (t^{\dagger} - t_A)$ as a function of *s* that reads as follows:

$$t^{\dagger} \leq \begin{cases} T_{k-1} + (r-1)(k+2) - s & \text{if } s < r, \\ T_{k-1} + ks & \text{if } s \ge r. \end{cases}$$
(5.5)

Clearly in the case s < r it is most convenient to take s = 0, whereas if $s \ge r$ the maximum value is achieved by selecting the largest possible value of s, which is dictated by the constraints (5.4).

We recall from Section 3 the definition of $D^*(n)$ as

$$D^*(n) = \begin{cases} 3T_{k-1} - kr & \text{if } r < \frac{k}{2}, \\ 3T_{k-1} - k(r-1) & \text{if } r = \frac{k}{2}, \\ T_{k-1} + (k+2)(r-1) & \text{if } r > \frac{k}{2}, \end{cases}$$

Theorem 5.9 (Depth of Λ_n). Let $n = T_k + r$ with $k \ge 2$ and $0 \le r \le k$. Then $D(n) = D^*(n)$. The case $k \le 1$ corresponds to $n \le 2$, easily studied manually.

Proof. In view of the results of Section 3 it suffices to prove the upper bound $D(n) \le D^*(n)$. Let $\lambda \in \Lambda_n$. We invoke Lemma 5.3 so that there exists *s* satisfying the constraints (5.4) such that after at most $T_{k-1} + ks$ moves the resulting configuration is *s*-monotone and $\ge \mathfrak{mg}^k$. Now we enforce Lemma 5.8 and distinguish three cases.

Case $r < \frac{k}{2}$. If $\overline{s} \ge r$, we use the first inequality of (5.4) in (5.5) to conclude

$$t^{\dagger} \leq T_{k-1} + k(k-r-1) = D^{*}(n).$$

If s < r, using $s \ge 0$, and $2r \le k - 1$, i.e. $r \le k - r - 1$ we have

$$t^{\dagger} \le T_{k-1} + (r-1)(k+2) \le T_{k-1} + rk + 2r - k - 2 < T_{k-1} + k(k-r-1) - 2$$

so that in any case $t^{\dagger} \leq D^*(n)$.

Case $r = \frac{k}{2}$. Again we distinguish the case $s \ge r$, together with the third inequality of (5.4) to obtain

 $t^{\dagger} \leq T_{k-1} + kr = 3T_{k-1} - k(k-r-1) = D^{*}(n).$

Whereas if s < r, again using $s \ge 0$ we have

 $t^{\dagger} \leq T_{k-1} + (r-1)(k+2) = 3T_{k-1} - k(r-1) - 2 < D^{*}(n).$

Case $r > \frac{k}{2}$. We start with the case s < r and again use $s \ge 0$ to get

$$t^{\dagger} \leq T_{k-1} + (r-1)(k+2) = D^{*}(n).$$

Finally, if $s \ge r$, enforcing the second inequality in (5.4), using $2r - k \ge 1$ we obtain

$$\begin{split} t^{\dagger} &\leq T_{k-1} + k \frac{2k-r}{3} \\ &= T_{k-1} + (r-1)(k+2) - (2r-k)\left(\frac{2}{3}k+1\right) + 2 < D^*(n). \end{split}$$

We have covered all the possible cases and the proof is complete. \Box

The previous Theorem 5.9, and in particular the upper bound $D(n) \le D^*(n)$ also allows to compute the maximal diameter of A_n :

Corollary 5.10 (Diameter of Λ_n). The diameter of Λ_n (see Definition 2.10) is given by $D^*(n) + P(n) - 1$ where P(n) is the maximal period of an augmented marching group, i.e. P(n) = 1 if $n = T_k$ is a triangular number, otherwise P(n) = k if $n = T_k + r$, with $0 < r \le k$, is not a triangular number.

Proof. That the diameter cannot exceed $D^*(n) + P(n) - 1$ immediately follows from Theorem 5.9. The reverse inequality follows from the constructions of Section 3, in particular we make use of Proposition 3.5, Corollary 3.7, Proposition 3.8 and the discussion about the period of the resulting augmented marching groups following each of these results.

6. Monotone mancala and Bulgarian solitaire

The notion of *s*-monotonicity naturally introduces a sort of graduation on the graph of moves. Indeed, Theorem 4.1 asserts that the subgraph obtained by considering only the *s*-monotone configurations for some $s \in \mathbb{N}$ is closed under mancala moves (no arc leaves this subset). It is then natural to ask what is the depth of such subgraphs.

More precisely, for $n, s \in \mathbb{N}$ let us denote by $\Lambda_{n,s}$ and $\Lambda_{n,\text{mon}}$ the subsets of Λ_n given by

 $\Lambda_{n,s} = \{\lambda \in \Lambda_n : \lambda \text{ is s-monotone}\}, \qquad \Lambda_{n,mon} = \{\lambda \in \Lambda_n : \lambda \text{ is monotone}\}.$

Similarly, with $g_{n,s}$ and $g_{n,mon}$ we denote the subgraphs of g_n having nodes in $\Lambda_{n,s}$ and $\Lambda_{n,mon}$ respectively.

Then the *depth* of $g_{n,s}$, denoted by D(n, s), is the maximal distance from a node of $g_{n,s}$ to a periodic configuration. In this context, recalling that 0-monotonicity is a void notion, our main Theorem 5.9 gives a value for D(n, 0) = D(n). It would be desirable to have a result about the depth D(n, s) for any value of s; unfortunately, at the moment we are not able to extend our result to nontrivial values of s.

Especially the value s = 1 is interesting: indeed, Remark 4.3 relates $g_{n,1}$ to $g_{n,mon}$ and in particular their respective depth. Adding one to D(n, 1) we obtain the depth of the graph of monotone configurations. This is of special interest in view of the following

Remark 6.1. The monotone mancala graph is isomorphic to the graph of the Bulgarian solitaire [4,6,7].

The Bulgarian solitaire, already mentioned in the Introduction, is played as follows. The starting position consists of a number of piles, each with a (possibly different) number of cards. Each move consists in removing one card from each pile and collecting the removed cards in a new pile. Piles that become empty are simply neglected and the order of the piles (and of the cards in a pile) is inessential. We can associate a mancala configuration to a Bulgarian solitaire configuration by defining λ_i equal to the number of piles with at least *i* cards. In this way we obtain a valid mancala configuration which is clearly monotone. It can be readily shown that this correspondence is one to one and that it is consistent with the respective game rules.

Hence, the depth of the Bulgarian solitaire subgraph is given by 1 + D(n, 1). The following conjecture is stated in [6]:

Conjecture 6.1. Let $n = T_k + r$ with $k \ge 2$ and $0 \le r \le k$. The depth of $\mathcal{G}_{n,1}$ is given by

$$D(n,1) = \begin{cases} 2T_{k-1} - rk - 1 & \text{if } 0 \le r < \left\lfloor \frac{k}{2} \right\rfloor \\ T_{k-1} + r - 1 & \text{if } r = \left\lfloor \frac{k}{2} \right\rfloor \text{ or } \left\lfloor \frac{k}{2} \right\rfloor + 1 \\ 2T_k - (k+1-r)(k+2) - 1 & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 < r \le k+1. \end{cases}$$

In [6] the authors proved this result for some special values of n, while the above expression is proved to be a lower bound. Our computer simulations validates the conjecture for values up to n = 180, improving the previous limit of n = 36.

Acknowledgments

The authors are grateful to the anonymous reviewers for their precious suggestions.

References

- A. Bouchet, Owari I. Marching groups and periodical queues, 2005, preprint, http://isgem.rpi.edu/files/1545.
 A. Bouchet, Owari II. Marching groups and Bulgarian Solitaire, 2005, preprint, http://isgem.rpi.edu/files/1546.
 H. Bruhn, Periodical states and marching groups in a closed owari, Discrete Math. 308 (16) (2008) 3694–3698.
- [4] G. Etienne, Tableaux de Young et solitaire bulgare, J. Combin. Theory Ser. A 58 (1991) 181-197.
- [5] M. Gardner, Mathematical games, Sci. Am. 249 (1983) 12–21.
 [6] J.R. Griggs, C.-C. Ho, The cycling of partitions and compositions under repeated shifts, Adv. Appl. Math. 21 (1998) 205–227.
 [7] K. Igusa, Solution of the Bulgarian solitaire conjecture, Math. Mag. 58 (1985) 259–271.