Cumulative subtraction games

Gal Cohensius *[1], Urban Larsson †[1], Reshef Meir ‡[1] and David Wahlstedt§[2]

[1]Technion–Israel Institute of Technology, Haifa, Israel
[2]Göteborg, Sweden

May 30, 2018

Abstract

We study zero-sum games, a variant of the classical combinatorial Subtraction games (studied for example in the monumental work "Winning Ways", by Berlekamp, Conway and Guy), called Cumulative Subtraction (CS). Two players alternate in moving, and get points for taking pebbles out of a joint pile. We prove that the outcome in optimal play (game value) of a CS with a finite number of possible actions is eventually periodic, with period $2s$, where s is the size of the largest available action. This settles a conjecture by Stewart in his Ph.D. thesis (2011). Specifically, we find a quadratic bound, in the size of s, on when the outcome function must have become periodic. In case of two possible actions, we give an explicit description of optimal play. We generalize the periodicity result to games with a so-called reward function, where at each stage of game, the change of 'score' does not necessarily equal the number of pebbles you collect.

*galcohensius@technion.ac.il
†urban031@gmail.com
‡reshefm@ie.technion.ac.il
§david.wahlstedt@gmail.com

1 Introduction

In Subtraction Games [2, 1], two players alternately take pebbles from a common pile until one of the players cannot take any action and loses. At each turn, a player chooses from a finite set of possible actions the number of pebbles to remove from the pile. Therefore a subtraction game is defined by the amount of pebbles x in the pile (known as the position) and the set of possible actions (or subtractions) S.

We are interested in a zero-sum variation of this class of games, defined by Stewart in his PhD thesis [5], here dubbed Cumulative Subtraction (CS); for extensions to general sum variations, see [3]. In this game, the two players, called Positive and Negative, accumulate the pebbles they remove, and they compete in getting the largest number, when the game ends. Each pebble collected by Positive increases the result by a point while each pebble collected by Negative decreases the result by one point.

Therefore an instance of CS, $(S; x, p)$, is composed of an action set S and a position (x, p), where the number of pebbles is $x \in \mathbb{Z}_{\geq 0}$, and the current score is $p \in \mathbb{Z}$.[1] A Positive's move is of the form $(x, p) \mapsto (x - s, p + s)$, for some $s \in S$, provided that $x - s \geq 0$. A Negative's move is of the form $(x, p) \mapsto (x - s, p - s)$, for some $s \in S$, provided that $x - s \geq 0$. A game state (t, p_{t}) is terminal if $t < \min S$. Positive wants to maximize the terminal score, p_{t}, whereas Negative seeks to minimize it.

We are interested in the *optimal* play of CS, which is a zero-sum game. Note that, in standard terminology, optimal play in zero-sum extensive form games defines a pure subgame perfect equilibrium.

Definition 1 (Optimal actions). *Given a game,* optimal play*,* opt*, is a mapping from positions to actions, such that the current player does not have a beneficial deviation from* opt*.*

Definition 2 (Outcome). *The outcome of the game S at position x is*

$$o(x) = \begin{cases} \max_{s \in S}\{-o(x - s) + s\}, & \text{if } x \geq \min S \\ 0, & \text{otherwise.} \end{cases}$$

Optimal play is independent of the accumulated (current) score, and $o(x)$ corresponds to Positive's terminal score, when both players play optimally,

[1] Sometimes, such as in Definition 2, we denote by 'a position' with just the heap size x. See also a discussion in the beginning of Section 4.

and Positive starts from position $(x, 0)$. This requires a proof, and we prove this in a slightly more general situation in Section 4.

Definition 3 (Sequence convergence)**.** *A sequence (x_i) of integers converges at j if, for all $i \geq j$, $x_j = x_i$ is constant, but $x_{j-1} \neq x_j$.*

(If there is no such j, then a sequence of integers does not converge.) We use the following optimal play convention: if in a given position, there are multiple optimal actions, then the current player plays the maximum optimal action. By *the sequence of optimal actions*, we mean a function, which maps heap sizes to unique (maximum) optimal actions.

Definition 4 (Game convergence)**.** *We say that a game S converges at $\xi(S) = j$, if its sequence of optimal actions converges at j.*

The game S has *slower* convergence than the game T if S converges at x and T converges at y, with $x > y$.

A sequence (x_i) is *periodic* if there is a p such that, for all i, $x_i = x_{i+p}$. If p is the smallest such number, then the sequence is periodic with period p.

We will be interested in outcome sequences that become periodic for sufficiently large heaps.

Definition 5 (Eventual periodicity)**.** *A function $g : \mathbb{Z}_{\geq 0} \to \mathbb{Z}$ is eventually periodic, if there is a $\gamma \in \mathbb{Z}_{>0}$, such that $g(x) = g(x + \gamma)$, for all sufficiently large $x \in \mathbb{Z}$. It is eventually periodic with period γ', if γ' is the smallest γ such that $g(x) = g(x + \gamma')$, for all sufficiently large $x \in \mathbb{Z}$.*

Lemma 1. *If the sequence of optimal actions converges, then the sequence of outcomes is eventually periodic.*

Proof. If both players optimally play the same action s from all sufficiently large heap sizes x, then $o(x) = s - o(x-s) = s - s + o(x-2s) = o(x-2s)$. $\square$

2 Results

Those readers familiar with CGT know that there is a standard argument for periodicity of the outcomes of classical subtraction games. This argument can be transfered to our setting with slight modifications, and we explain how to do this in Section 4, in a somewhat more general situation, where players gain various 'rewards' for their actions. We begin with more detailed structural results on the CS games with identity rewards.

1. In [5], Stewart conjectures that CS on finite subtraction sets S, are eventually periodic with period $2s_1$ where $s_1 = \max S$ (although he uses a somewhat different definition of outcomes). That is to say, $o(x+2s_1) = o(x)$, for any large enough position x. In this paper, we prove this conjecture (Section 3.4, Corollary 11), and generalize it in Section 4 to games with reward functions.

2. We characterize the optimal play as well as the precise outcomes in CS for two special cases: CS with support size of two and CS with full support, i.e. $|S| = \max S$. We provide a conjecture in the case of 'truncated support' (Section 3.3).

3. We give an upper bound on the heap size for which CS games converge. The bound is quadratic in $\max S$ (Section 3.4).

3 Cumulative Subtraction

Throughout the paper, we usually consider optimal play by both players in a game $(S; x, p)$, and Positive starts. Our first observation regards the relative number of actions players will have throughout the game.

Observation 1. *Either Positive and Negative play the same number of actions, or Positive has an extra turn.*

This is true because Positive starts and turns alternate between the players. The only reason for a player to play non-maximal action is to get to play an extra turn (getting the last action).

Lemma 2. *Consider optimal play. If a player does not play the maximal action, then she obtains the last move (a parity advantage).*

Proof. Suppose that Positive plays a non-maximal action without getting the last move, when Negative plays optimally. Since Negative gets the last move, and Positive cannot have less actions than Negative, if Positive exchange a non-maximal action for a maximal action she accumulates a higher total score. Hence, the first strategy was not optimal. The same proof can be used for Negative, looking at a game that starts after the first move of Positive. □

Corollary 3. *In optimal play, at least one of the players plays only maximum actions.*

Proof. This follows directly from Lemma 2, because both players cannot get the last move. □

3.1 CS with full support

Consider a CS where the set of possible actions contains all the integers from 1 up to s_1, i.e., $S = \{1, 2, \ldots, s_1\}$. We call this game CS with *full support*. In this game, optimal-play is to play the maximal action available at each position.

Theorem 4. *In CS with* full support*, the optimal play is x for any position $x < s_1$ and s_1 for any position $x \geq s_1$. That is, each CS with full support converges at s_1, and moreover its outcome is periodic with the pattern*

$$(0, 1, \ldots, s_1, s_1 - 1, \ldots, 1). \tag{1}$$

Proof. The proof is by induction. For the base case, consider $0 \leq x \leq s_1$: when playing from position x, Positive takes all the pebbles, and thus $o(x) = x$. When playing from positions $x + s_1$, Positive's optimal play is to take s_1 and negative takes the rest, thus $o(x + s_1) = s_1 - o(x) = s_1 - x$. It is Positive's optimal play since if she takes less than s_1 then Negative can take more than x.

Assume $k > 0$ repetitions of the pattern (1). We study the next s_1 positions and show that the outcome in those positions will be exactly as in (1).

$$\begin{aligned}
o(x + 2(k+1)s_1) &= s_1 - o(x + 2(k+1)s_1 - s_1) \\
&= s_1 - o(x + 2ks_1 + s_1) \\
&= s_1 - o(x + s_1) \\
&= s_1 - (s_1 - x) \\
&= x
\end{aligned}$$

For the following s_1 positions the outcome is $o(x + s_1 + 2(k+1)s_1) = s_1 - o(x + s_1 + 2(k+1)s_1 - s_1) = s_1 - o(x + 2(k+1)s_1) = s_1 - x$. □

3.2 CS with two possible actions

In this section, we consider a CS where the set of possible actions contains just two actions, denoted by $S = \{s_2, s_1\}$, where $s_1 > s_2$. We characterize the optimal play in each position and prove the convergence of the game.

In the simplest case, when $2s_2 \leq s_1$, S is periodic and converge at s_1. This is because it is never beneficial to play s_2 if s_1 is possible, because even if by playing s_2 a player will get an extra turn where she can play s_2, it is still lower than a single s_1. For example, in the game $S = \{2, 5\}$, the outcome for the first 20 positions are:

$$0, 0, 2, 2, 0, 5, 5, 3, 3, 5,$$
$$0, 0, 2, 2, 0, 5, 5, 3, 3, 5 \ldots,$$

and this pattern of the first 10 outcomes repeats.

Theorem 5. *Suppose $S = \{s_2, s_1\}$, with $2s_2 \leq s_1$. Then the game S has a periodic outcome function, with period $2s_1$, and the sequence of optimal actions converges at position $x = s_1$.*

Proof. The proof is analogous to the case of full support, Theorem 4. $\square$

Observation 2. *In the setting of Theorem 5, the first s_1 outcomes are of the form: $o(0) = \cdots = o(s_2 - 1) = 0$, $o(s_2) = \cdots = o(2s_2 - 1) = s_2$, $o(2s_2) = \cdots = o(3s_2 - 1) = 0, \ldots$ until $o(s_1 - 1)$, which is then 0 or s_2 (depending on the numbers a and b in the division algorithm, with $s_1 = as_2 + b$ where $0 \leq b < a$). The following s_1 outcomes then takes the values $o(s_1) = \cdots = o(s_1 + s_2 - 1) = s_1$, $o(s_1 + s_2) = \cdots = o(s_1 + 2s_2 - 1) = (s_1 - s_2)$, $o(s_1 + 2s_2) = \cdots = o(s_1 + 3s_2 - 1) = s_1, \ldots$ until $o(2s_1 - 1)$, which is then analogously s_1 or $s_1 - s_2$. These outcomes repeat indefinitely.*

For example, the outcomes of the game $\{2, 9\}$ are periodic, with the first 18 outcomes: $0, 0, 2, 2, 0, 0, 2, 2, 0, 9, 9, 7, 7, 9, 9, 7, 7, 9$.

For the rest of the section, we assume that $2s_2 > s_1$.

Next we wish to give the reader a better feeling of the game by presenting an example, consider the game $S = \{5, 7\}$, table 1 stats the optimal first action and the outcome, given the position.

Note that for CS with $S = \{5, 7\}$ the only positions where $s_2 = 5$ is better than $s_1 = 7$ are $X^* = \{5, 6, 17, 18, 29, 30\}$. This means that the game converges at $\xi(\{5, 7\}) = 31$.

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13
opt	-	-	-	-	-	5	5	7	7	7	7	7	7	7
$o(x)$	0	0	0	0	0	5	5	7	7	7	7	7	2	2

x	14	15	16	17	18	19	20	21	22	23	24	25	26	27
opt	7	7	7	5	5	7	7	7	7	7	7	7	7	7
$o(x)$	0	0	0	3	3	5	5	7	7	7	4	4	2	2

x	28	29	30	31	32	33	34	35	36	37	38	39	40	41
opt	7	5	5	7	7	7	7	7	7	7	7	7	7	7
$o(x)$	0	1	1	3	3	5	5	7	6	6	4	4	2	2

x	42	43	44	45	46	47	48	49	50	51	52	53	54	55
opt	7	7	7	7	7	7	7	7	7	7	7	7	7	7
$o(x)$	0	1	1	3	3	5	5	7	6	6	4	4	2	2

Table 1: Positive's optimal play (not necessary unique), and the outcome for CS with action set $S = \{5, 7\}$, starting from position x.

Let $\alpha = s_1 - s_2 > 0$ be such that $s_2 > \alpha$, and let $\Delta = \{0, 1, \dots, \alpha - 1\}$. Define

$$X^*(i) = \{is_2 + (i-1)s_1 + \delta \mid \delta \in \Delta\},$$

for each $i \in \mathbb{Z}_{>0}$, such that

$$is_2 > (i-1)s_1, \tag{2}$$

and otherwise $X^*(i) = \varnothing$. We will show that, the set of positions where the unique optimal move is to play s_2 is

$$X^* = \bigcup_{i \in \mathbb{Z}_{>0}} X^*(i)$$

The special case whenever $2s_2 \le s_1$ is treated in Theorem 5. Then equation (2) only holds for $i = 1$, in which case $X^* = X^*(1) = \{s_2, \dots, s_1 - 1\}$. Another example is if $2s_2 > s_1$, but $3s_2 \le 2s_1$. Then equation (2) holds for $i = 1, 2$, in which case $X^* = X^*(1) \cup X^*(2) = \{s_2, \dots, s_1 - 1, 2s_2 + s_1, \dots, 2s_2 + s_1 + \alpha - 1\}$.

A more compact representation of (2) is $\alpha i < s_1$. Therefore, in particular, (2) implies $i < s_1$.

Let $i_{\max}$ denote the largest i such that (2) holds. That is, $\alpha i_{\max} < s_1$, but $\alpha(i_{\max} + 1) \ge s_1$. We show that if the difference between a couple of positions is exactly α, and the larger position is not in X^*, then the outcome difference is bounded from above by α. Later, in the proof of Theorem 7, we will see that the converse holds, if the larger position is in all but the largest subset of X^*. That is, if the larger heap size is in

$$\bigcup_{1\le i<i_{\max}} X^*(i),$$

then the outcome difference is greater or equal than α.

The following lemma holds for any s_2, s_1, but we will only use it for the cases $s_2 > \alpha$ (since we already treated the other cases in Theorem 5). Note in particular if $s_2 \le \alpha$ then $i_{\max} = 1$, so all positions will satisfy inequality (3) in Lemma 6.

Lemma 6. *Consider $S = \{s_2, s_1\}$ and $0 < \alpha = s_1 - s_2$. If $x \notin X^*$, and $x \ge \alpha$ then*

$$o(x) - o(x-\alpha) \le \alpha. \tag{3}$$

Proof. We study the function

$$\eta(x) := \alpha + o(x-\alpha) - o(x),$$

and show that $\eta(x) \ge 0$, if $x \notin X^*$ and $x \ge \alpha$. We think about $o(x)$ as the outcome in optimal play when Positive starts, and $-o(x)$ as the outcome in optimal play when Negative starts. It suffices to show that, for all plays by Negative from x, there is a response by Positive such that the inequality holds.

If there is no move from x (because $s_2 > x$) then $\eta(x) = \alpha \ge 0$.

If there is (an optimal) move from x, but no move from $x-\alpha$, then $x < s_1$; thus $x \in X^*(1)$, which is not part of the positions in the statement.

1. If Negative plays optimally s_1 from x, and Positive plays s_2 from $x-\alpha$, we get

$$\begin{aligned}\eta(x) &\ge \alpha + o(x-s_1) - s_1 - o(x-\alpha-s_2) + s_2\\ &= o(x-s_1) - o(x-s_1)\\ &= 0\end{aligned}$$

2. If Negative plays optimally s_2 from x, and Positive plays s_2 from $x-\alpha$, we get
$$\eta(x) \geq \alpha + o(x-s_2) - o(x-\alpha-s_2),$$
and we note that, if Negative has no move from $x-\alpha-s_2 = x-s_1$, then this implies $\eta(x) \geq 0$. Assume Negative has a move then there are two cases:

   2.1 On the second move, if Negative plays optimally s_2, and Positive plays s_1, we get
   $$\eta(x) \geq \alpha - s_2 + o(x-s_1-s_2) + s_1 - o(x-s_2-s_1) = 2\alpha > 0$$

   2.2 On the second move, if Negative's optimal move is s_1, and Positive responds with s_1, we get
   $$\eta(x) \geq \alpha + o(x-s_1-s_1) - o(x-s_2-s_1) = \eta(x-s_2-s_1)$$
   and since, by definition of X^*, if $x \notin X^*$ then $x-s_1-s_2 \notin X^*$. Therefore $\eta(x-s_2-s_1) \geq 0$ by induction.

This concludes the proof of inequality (3). □

Next we prove that the positions where s_1 is not optimal are exactly the X^* positions, which are of the form $s_2+\delta, 2s_2+s_1+\delta, 3s_2+2s_1+\delta, \ldots$, that is the positions congruent with $s_2+\delta \pmod{s_1+s_2}$, until equation (2) fails to hold. The optimal actions before convergence are shown in Figure 1 (pile sizes modulo s_2+s_1 have the same optimal actions before convergence), and in Figure 2, we illustrate the outcomes (pile sizes modulo $2s_1$ have the same outcomes at convergence).

Observation 3. *If $x \in X^*$, then neither $x-s_1$ and $x-s_2$ is in X^*. This follows because $x \in X^*$ is equivalent with $x \equiv \delta - s_1 \pmod{s_1+s_2} \in \{s_2, \ldots, s_1-1\}$. Then $x-s_1 \in_{(\mathrm{mod}\ s_1+s_2)} \{2s_2, \ldots s_2+s_1-1\}$ and $x-s_2 \in_{(\mathrm{mod}\ s_1+s_2)} \{0, \ldots s_1-s_2-1\}$. Namely, if $2s_2 > s_1$, then these sets are disjoint. Similarly, if $x \notin X^*$, then $x-s_1-s_2 \notin X^*$, because these two numbers are congruent modulo s_1+s_2.*

Theorem 7. *Let the action set be $S = \{s_2, s_1\}$, with $s_1 > s_2$. The action s_2 is the unique optimal action, that is $\mathrm{opt}(x) = \{s_2\}$ if and only if $x \in X^*$.*

Proof. If $s_2 \leq \alpha$, then $X^* = X^*(1)$ and Theorem 5 covers this case. Therefore, we assume $s_2 > \alpha$.

Recall $x \in X^*(i)$ if $x = is_2 + (i-1)s_1 + \delta$, $\delta \in \Delta = \{0, 1, \ldots, \alpha - 1\}$, for $i \geq 1$, with $\alpha i < s_1$. For these positions, we define Positive's '$\star$-strategy'. Positive plays only s_2, unless Negative responds with s_2 in which case Positive plays s_1. Claim: this is the correct optimal strategy and Negative responds optimally with $(i-1)$ s_1 actions. If Positive would have played s_1 (at first), then Positive can obtain at most 0 points, since Positive loses the desired parity advantage. This follows because if Negative plays $(i-1)$ s_1 actions, and Positive plays at least one s_1 action, by inequality (2), we get the following contradiction for a terminal heap size t, $t \leq is_2 + (i-1)s_1 + \delta - ((i-1)s_1 + s_1 + (i-1)s_2) = \delta - s_1 + s_2 < 0$.

This proves that s_2 is the unique optimal action for positions in X^*, so we are done with the case $x \in X^*$. Moreover, the argument implies that $o(x) = is_2 - (i-1)s_1 > 0$ for all $x \in X^*(i)$, that is we get, for $x \in X^*$,

$$o(x) = s_1 - i\alpha. \tag{4}$$

A consequence of this is the following claim.

Claim 1: Suppose that y is such that $x - o(x) \leq y < x$, with $x \in X^*$. Then

$$o(y) = 0, \tag{5}$$

and the optimal action is s_1.

Proof of Claim 1. Consider $x \in X^*(i)$. By (4), we have $x - s_1 + i\alpha \leq y < x$. That is,

$$(2i-2)s_1 + \delta \leq y < x. \tag{6}$$

The upper bound implies that Positive cannot obtain an extra move by playing s_2. Therefore the action s_1 dominates s_2. If Negative removes s_1 pebbles, then we remain in a position of the same form (until the game ends and the score is 0). So, assume that Negative plays s_2 from $y - s_1$. But the position $y - s_1 - s_2$ is again of the same form as in (6). Negative will not play s_2, with a relative loss of α, without gaining a parity advantage. This proves the claim.

For the other direction, we must prove that for each position $x \notin X^*$, then s_1 is an optimal move if $x \geq s_1$. We begin by stating the full base case.

Consider $x \in \{0, \ldots, 2s_1 - 1\}$. If $x < s_2$, no action is available. (For positions $x \in X^*(1)$, only action s_2 is available, so it is optimal.) For positions $x \in \{s_1, \ldots, 2s_1 - 1\} \subset \mathbb{Z} \setminus X^*$, s_1 is the unique optimal action, since it can be countered with at most one s_2 action before the end of play, and $2s_2 + s_1 \geq 2s_1$, by assumption.

Assume next that $x \geq 2s_1$. It suffices to prove:

$$\text{if } x \notin X^*, \text{ then } -o(x - s_1) + s_1 \geq -o(x - s_2) + s_2,$$

or equivalently

$$\text{if } x \notin X^*, \text{ then } \alpha \geq o(x - s_1) - o(x - s_2).$$

There are three cases, depending on whether $x - s_1$ or $x - s_2$ belongs to X^* respectively. Note that both cannot belong to X^*, because $x - s_2 - (x - s_1) = \alpha$, and, for all i, $X^*(i)$ contains at most $\alpha - 1$ consecutive numbers (and more than s_1 numbers separate two disjoint sets $X^*(i)$ and $X^*(j)$).

1. $x - s_1 \in X^*$, $x - s_2 \notin X^*$
2. $x - s_1 \notin X^*$, $x - s_2 \notin X^*$
3. $x - s_1 \notin X^*$, $x - s_2 \in X^*$

For 1., use induction to conclude $s_2 \in \text{opt}(x - s_1)$ and $s_1 \in \text{opt}(x - s_2)$. We get

$$\begin{aligned} o(x - s_1) - o(x - s_2) &= -o(x - s_1 - s_2) + s_2 + o(x - s_2 - s_1) - s_1 \\ &= -\alpha \\ &< \alpha. \end{aligned}$$

For 2., use induction to conclude $s_1 \in \text{opt}(x - s_1)$ and $s_1 \in \text{opt}(x - s_2)$. We get

$$\begin{aligned} o(x - s_1) - o(x - s_2) &= -o(x - s_1 - s_1) + s_1 + o(x - s_2 - s_1) - s_1 \\ &= -o(x - s_1 - s_2 - \alpha) + o(x - s_2 - s_1) \\ &\leq \alpha, \end{aligned}$$

by Lemma 6, using that $x \notin X^*$ implies $x - s_2 - s_1 \notin X^*$, unless perhaps $x - s_2 - s_1 \in X^*(i_{\max})$. This latter case has to be treated separately. By (4) and (5), we know that, in this case, $o(x - s_2 - s_1) - o(x - s_2 - s_1 - \alpha) = s_1 - \alpha i_{\max} - 0 < \alpha$, where the inequality is by definition of $i_{\max}$.

For 3., in case $i < i_{\max}$, we use again the 'duality' (5) between outcomes and number of consecutive positions with outcome 0 below $X^*(i)$, which gives $o(x - s_2) - o(x - s_1) = s_1 - i\alpha - 0 > \alpha$, since $i < i_{\max}$, so

$$o(x - s_1) - o(x - s_2) < -\alpha \leq \alpha.$$

The remaining case is when $x - s_2 \in X^*(i_{\max})$.

This means that $x - s_2 \in_{(\mathrm{mod}\ 2s_1)} \{-i_{\max}\alpha - s_1 + \delta\}$, with

$$-s_1 < -i_{\max}\alpha \leq -s_2. \tag{7}$$

Hence, $x - s_2 \in_{(\mathrm{mod}\ 2s_1)} \{1, \dots, 2\alpha - 1\}$ and so $x - s_1 \in_{(\mathrm{mod}\ 2s_1)} \{1 - \alpha, \dots, \alpha - 1\}$.

Since $x - s_2 \in X^*$, we know that $o(x - s_2) = s_1 - i_{\max}\alpha$ (where $0 < s_1 - i_{\max}\alpha \leq \alpha$). So, in this case, we must prove that $\alpha \geq o(x - s_1) - (s_1 - i_{\max}\alpha)$, or equivalently that $o(x - s_1) \leq \alpha - i_{\max}\alpha + s_1$. Hence, by (7) it suffices to prove that $o(x - s_1) \leq \alpha + 1$. We have two cases:

(i) $x - s_1 \in_{(\mathrm{mod}\ 2s_1)} \{0, \dots, \alpha - 1\}$

(ii) $x - s_1 \in_{(\mathrm{mod}\ 2s_1)} \{1 - \alpha, \dots, -1\}$

We use the previous parts of the proof. For case (i), it is optimal to play all remaining actions s_1. Namely, $x - ks_1 \notin X^*$, for any $k \geq 0$. Since, there is an even number of actions, the outcome is 0. For case (ii), the only action in X^* is the last action, so the outcome is $i_{\max}s_1 - (i_{\max} - 1)s_1 - s_2 = \alpha$. □

Corollary 8. *The game converges to action s_1 at $\max X^* + 1$.*

Proof. This follows from Theorem 7. □

Observation 4. *From the proof of Theorem 7, we get that the outcomes are:*

- $o(x) = s_1 - i\alpha$, *if* $x \in X^*(i)$;
- $o(x) = 0$, *if* $y - s_1 + i\alpha \leq x < y$, *where* $y \in X^*(i)$;
- $o(x) = o(y)$, *if there is a* $y \equiv x \pmod{2s_1}$, *such that* $y \in X^*(j)$, *with* $j < i$;

- $o(s_1 + x) = s_1 - o(x)$, *for all* $x \in_{(\mathrm{mod}\ 2s_1)} \{0, \ldots, s_1 - 1\}$.

In particular, the periodic outcome pattern, at convergence, is obtained by applying $i_{\max} = \lfloor \frac{s_1-1}{\alpha} \rfloor$ *in the items.*

Note that the first three items concern the outcomes of the positions in the congruence classes $0, \ldots, s_1 - 1 \pmod{2s_1}$ *and the last item concerns the 'symmetric' part among the heap sizes* $s_1, \ldots, 2s_1 - 1 \pmod{2s_1}$*. The third item shows that once the outcomes for positions in* $X^*(i)$ *have been computed, then they stabilize, for equivalent larger heap sizes modulo* $2s_1$*.*

Another consequence of the arguments is that if $x \equiv_{(\mathrm{mod}\ 2s_1)} \{s_1, \ldots, 2s_1 - 1\}$*, then* $s_1 \in \mathrm{opt}(x)$*.*

In Figures 1 and 2 we sketch the optimal actions modulo $s_2 + s_1$ and the outcomes modulo $2s_1$, of the two-action games with $2s_2 \geq s_1$.

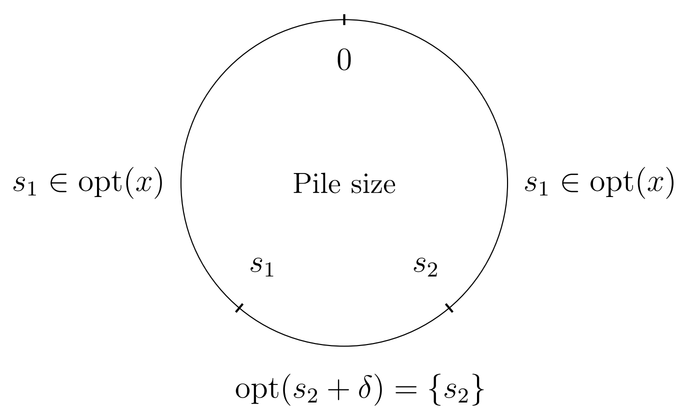
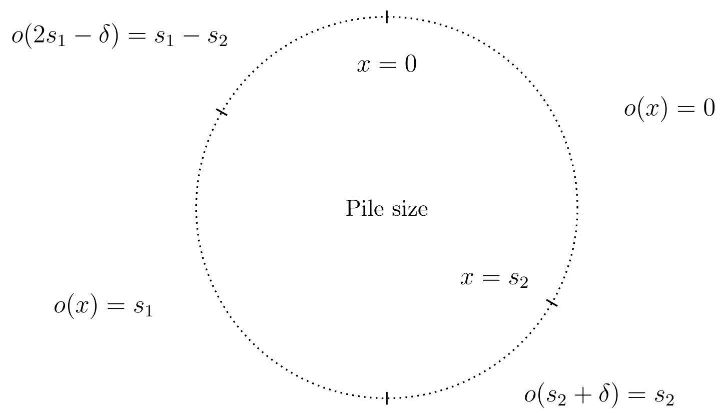


Figure 1: Optimal actions before convergence, for pile sizes of x modulo $s_2 + s_1$, where $x \geq s_1$. The positions in X^*, where the unique optimal action is s_2, are of the form $x + (s_1 + s_2)i$, for $i \geq 0$ and $s_1 - 1 \geq x \geq s_2$, until convergence

The first $2s_1$ outcomes



The outcomes at convergence

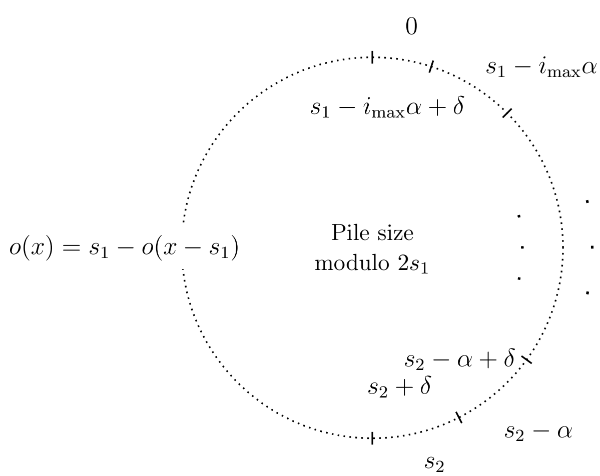


Figure 2: Initial outcomes (top) and outcomes at convergence (bottom) for pile sizes modulo $2s_1$, for 2-action games (the relation between the actions s_2 and s_1 is different in the two pictures). Note that $s_2-(i_{\max}-1)\alpha = s_1-i_{\max}\alpha$.

3.3 Games with truncated support

Consider a game S, with $m = \max S \geq 2$, of the form $S = \{i, i+1, \ldots, m\}$, where $i \in \{1, m-1\}$, so that $|S| = m - i + 1$. We call this the class of *truncated support* games. It includes as special cases the games with full support ($i = 1$) and the games with the slowest convergence ($i = m-1$). We estimate in which $2m$ interval, tr_i^m, optimal play converges to the maximal action m. Let tr^m denote the sequence of the form $\mathrm{tr}^m = (\mathrm{tr}_i^m)_{i=1}^{m-1}$.

For example, when $m = 5$, then the sequence is $\mathrm{tr}^5 = (1, 2, 2, 4)$. Here, the first entry $\mathrm{tr}_1^5 = 1$ shows that when $S = \{1, 2, 3, 4, 5\}$ is the game of full support, then the convergence to maximal action in optimal play occurs already in the *first* interval of length 10. The last entry, $\mathrm{tr}_4^5 = 4$, concerns the game $S = \{4, 5\}$, and, as evidenced, convergence occurs by the 4^{th} 10-interval.

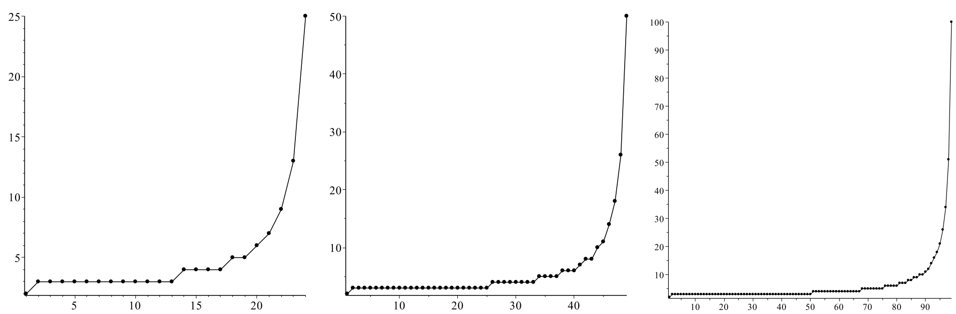
The i^{th} column shows the convergence for $(i-1)$-truncated support games, for $m \in \{2, \ldots, 10\}$. (We will explain the $\#x$ column below.)

i	1	2	3	4	5	6	7	8	9	$\#x$
tr^2	1									1
tr^3	1	2								2
tr^4	1	2	3							3
tr^5	1	2	2	4						3
tr^6	1	2	2	3	5					4
tr^7	1	2	2	2	3	6				4
tr^8	1	2	2	2	3	4	7			5
tr^9	1	2	2	2	2	3	4	8		5
tr^{10}	1	2	2	2	2	3	3	5	9	5

From this table alone, for $i \geq 2$, it might appear that the sequence of number of occurrences is non-increasing. To obtain some more insight, we plot the entries below for $m = 25, 50, 100$.

The pictures seem to converge to some function of the form $\frac{A}{\sqrt{B-x}}$. The appearing 'symmetry' has a precise formulation, explained in the below conjecture.

For each $m = \max S$, shrink the tr^m sequence to the set $x^m = x = \{\mathrm{tr}_i^m\}$ and enumerate the elements in increasing order; we interpret x as a sequence $x^m = (x_i)$ with $x_1 = 1$ (by the theorem for full support) and $\max x^m = m - 1$ (by the support size 2 result). We have, for all $i \geq 1$, $x_i < x_{i+1}$. But, what



is the number of elements in x, for each m? The initial sizes of these sets are displayed in the last column of the table, as $\#x$. Study the first differences $\Delta_i^m = x_{i+1}^m - x_i^m$, $i \geq 1$. Define, for all $m \geq 3$, and for all $1 \leq j \leq \#x$, $M_j := \#\{i \mid x_j^m = \mathrm{tr}_i^m\}$.

One can prove the following result by combining methods and results in Theorem 4 and Theorem 7.

Theorem 9. *For $i \in \{2, \ldots, \lceil m/2 \rceil\}$, $\mathrm{tr}_i^m = 2$, and moreover, $\Delta_{m-1}^m = \mathrm{tr}_m^m - \mathrm{tr}_{m-1}^m = \lfloor m/2 \rfloor = \#\{i \mid \mathrm{tr}_i^m = 2\} = M_2$.*

This result reflects an emerging 'duality' between individual games and sequences of games, which appears to continue in the inner regions of the pictures. We make the following conjecture.

Conjecture 1 (Duality)**.** *Consider any truncated CS.*

- *For all $m \geq 2$, $\#x^m = \left\lfloor \sqrt{4m-7} \right\rfloor$ (corresponding to sequence OEIS: A000267).*
- *The first differences, Δ^m, equal in reverse order the number of multiplicities of the numbers in tr^m. That is, for all i, $M_{m+1-i} = \Delta_i^m$.*

Consider for example tr^{10}. Then $\Delta = (1, 1, 2, 4)$, and $M = (4, 2, 1, 1)$. Careful inspection reveals that the pictures for $m = 25, 50, 100$ satisfy this precise correspondence (and we checked many cases up to $m = 200$), but we have up to date no means of explaining this proposed 'duality'.

3.4 Games with arbitrary support - a tight convergence bound

We prove that, for each game, convergence is no slower than for the game consisting exactly of the two largest actions in that game. Moreover, for each

$s \geq 2$, the game $S = \{s-1, s\}$ has the slowest convergence of all games with $s = \max S$.

Recall Definition 4 (of game convergence). By Theorem 7, we get, for example:

(i) For all $s \geq 2$: $\xi(\{s-1, s\}) = 2(s-1)^2$.

(ii) For all $s \geq 3$: $\xi(\{s-2, s\}) = (s-2)^2 + s - 1$ if s is odd; $\xi(\{s-2, s\}) = (s-2)^2$ if s is even.

We will prove the following two general convergence results.

Theorem 10. *Let $S = \{s_2, s_1\}$, with $s_1 > s_2$. Then $\xi(S) \geq \xi(T)$, for all $S \subset T$ with $s_1 = \max T$ and $s_2 = \max(T \setminus \{s_1\})$.*

Corollary 11. *Let $S = \{s-1, s\}$. Then $2(s-1)^2 \geq \xi(S) \geq \xi(T)$, for all $T \subset \mathbb{Z}_{>0}$ with $s = \max T$.*

That is, we want to prove that item (i) above gives an upper bound of the convergence, for any given game S with $s = \max S$. We generalize X^* (from Section 3) and inequality (2) as follows. Consider a fixed game $S = \{s_1, \ldots, s_n\}$, with $n \geq 2$, and $s_1 > \cdots > s_n$. Let ϕ be a multiset, with elements from $\{2, \ldots, n\}$, such that $\sum_{i \in \phi} a_i = m$ maximizes:

$$\nu(\phi) = \sum_{i \in \phi} a_i s_i - (m-1)s_1. \tag{8}$$

If $\nu(\phi) > 0$, the actions induced by ϕ (played in any order) give player Positive a parity advantage, playing from any position in the set

$$X^\phi = \{\nu(\phi) + \delta \mid \ 0 \leq \delta < \min\{s_1 - s_2, s_n\}\}.$$

Proof of Theorem 10. By Lemma 2, $\nu(x) > 0$, with $x \in X^\phi$ is the only possibility for a non-max S action to be optimal, and Negative's (non-winning) optimal strategy is to play s_1 until the game ends. Therefore, by (8), no situation with $|S| \geq 2$ can decrease $\max X^\phi = \xi(S) - 1$, which is hence bounded above by

$$\begin{aligned} \max X^\phi &\leq m s_2 + (m-1)s_1 + \delta \\ &\leq (m-1)(s_1 - 1) + m s_1 \\ &= 2m s_1 - s_1 - m + 1 \\ &\leq 2s_1^2 - 4s_1 + 2 \\ &= 2(s_1 - 1)^2, \end{aligned}$$

since $\nu(\phi) > 0$ and (8) imply $m \leq s_1 - 1$. $\square$

Proof of Corollary 11. The statement holds trivially for $T = \{s\}$. Moreover, the inequality (2) implies $\xi(S) \geq \xi(T)$, for any $T = \{s_2, s_1\}$ with $s_2 < s_1 = s$. Therefore the result follows from Theorem 10. $\square$

4 Generalization to games with a reward function

We generalize the CS game in the following way: each action $s \in S$ subtracts s pebbles but is worth $r(s)$ points. A function $r : S \to \mathbb{R}$ determines the *reward* for each action. Thus, each generalized game is of the form $(S, r; x, p)$.

Postive's move options are of the form $(S, r; x, p) \mapsto (S, r; x - s, p + r(x))$, for some $s \in S$ such that $x - s \geq 0$, whereas Negative's move options are of the form $(S, r; x, p) \mapsto (S, r; x - s, p - r(x))$.

Definition 1 will be used in this slightly more general setting. We generalize the outcome function, and then justify that it computes a game's value, shifted by a constant (corresponding to the initial score p).

Definition 6 (Generalized outcome)**.** *The outcome of the game, $(S, r; x, p)$, from position x is*

$$o(x) = \begin{cases} \max_{s \in S}\{-o(x - s) + r(s)\} & \text{for } x \geq \min S \\ 0 & \text{for } x < \min S. \end{cases}$$

By von Neumann's classical minimax (maximin) theorem [4], for each 0-sum game G, there exists a unique optimal play game value $\nu(G)$, and moreover, because CS are perfect information games, this value is achieved via pure strategies, i.e. standard backward induction. In CS with a reward function, for the game $G = (S, r; x, p)$, we let

$$\nu(G) = p + \sum_{i=0}^{\tau} (-1)^i \, r(a_i), \tag{9}$$

where $(a_i)_{i=0}^{\tau}$ is any fixed sequence of actions (a path) in optimal play, a_0 is Positive's first action, and a_τ is the terminating action.

We show that for CS with a reward function, equilibrium play does not depend on the current score.

Lemma 12. *The set of paths in optimal play (induced by the PSPE strategy profile) does not depend on the current score.*

Proof. Assume that player Positive starts. The reward function takes as input an action $s \in S$. Suppose we have two games, where only the current score differs, say $P = (S, r; x, p)$ and $Q = (S, r; x, q)$, and that x is the smallest heap size such that $\nu(P) - p > \nu(Q) - q$. This requires that x is such that at least two move options are available. By assumption, optimal play *after* Positive's first move does not depend on the current score. Hence, Positive can deviate in the game Q and use instead an optimal path as induced by $\nu(P)$, which gives the final score $q + \nu(P) - p > \nu(Q)$, a contradiction to Definition 1. $\square$

Theorem 13. *For all games $G = (S, r; x, p)$, $\nu(G) - p = o(x)$.*

Proof. The proof is by induction. If there is at most one available action, then the result trivially holds.

Suppose $G' = (S, r; x - s, p + r(s))$. Then, by definition (9), $\nu(G) = \nu(G')$, because $\nu(G') = p + r(s) + \sum_{i=1}^{\tau} (-1)^i r(a_i)$, with $s = a_0$ in optimal play, which here means maximizing. Since the indexing of the sum starts with an odd number (namely 1), assume by induction that $o(x - s) = -\nu(G') + r(s) + p$. This gives, for a maximizing action s,

$$\begin{aligned} o(x) &= r(s) - o(x - s) \\ &= r(s) - (-\nu(G') + r(s) + p) \\ &= \nu(G) - p. \end{aligned}$$

$\square$

Therefore we have an efficient and compact representation of the game value in form of the outcome function. Namely, a generic set of terminal game positions is exponential in the size of the game tree, whereas the initial state of the outcome function is simply 0. From another perspective, this means that a game is independent of its 'history', which is recorded as a 'current score'. The players may ignore the 'current score' in finding an optimal move from a given position.

4.1 Rewards and eventually periodic outcomes

Given a game (S, r),[2] we call an action $m \in S$, such that $r(m) = \max\{r(s) \mid s \in S\}$, a *maximal action* m.

Interesting situations occur for reward functions, when there is more than one maximal action, or in case the maximal action is not the *maximum action*, i.e. whenever the maximal action $m \neq \max S$. Note that in case r is increasing, then 'maximal' and 'maximum' coincide.

Minimal and minimum actions are defined similarly. An action $n \in S$, such that $r(n) = \min\{r(s) \mid s \in S\}$, is a *minimal action*. Note that both minimal and maximal actions may map to any real reward, and if $r(n) = r(m)$, then the game's outcome depends only on the parity of number of moves, and if there are negative rewards a player may not want the last move (in particular if all actions map to negative rewards, which then becomes 'misére play' in standard jargon).

Example 1. *Let $x = 6$, with $S = \{2, 3, 4\}$, $r(2) = -100$, $r(3) = 50$ and $r(4) = -10$. Positive starts and prefers to take the punishment -10, rather than immediately cashing in 50 points, since Negative will be forced to increase the collective cumulation by 100 points, and the final score becomes 90, a great victory for Positive (otherwise the score would be 0 if Negative plays optimally). So here Positive* avoids *the parity 'advantage' which is no longer an advantage. On the other hand, if Positive starts from $x = 2$, then the outcome is -100, so the range of possible outcomes is no longer bounded from below by 0; see Lemma 14.*

We define game convergence similar to previous sections. We say that the game (S, r) *converges* at $x' \geq 0$ if the same maximal action is optimal for all positions $x \geq x'$.

In experiments we have encountered games that do not converge, but are still periodic, and this observation is proved at last in Theorem 19.

Lemma 14. *Consider a game $(S, r; x, p)$. Let $\rho = \min\{r(n), 0\}$ and $\nu = \max\{-r(n), r(m)\}$, where n (m) is a minimal (maximal) action. Then*

$$\rho \leq o(x) \leq \nu, \tag{10}$$

[2] We sometimes omit the heap size x and the current score p from the description of a game, if the context requires it.

Moreover, if there is no negative reward, then the lower bound holds with $\rho = 0$, and the upper bound holds with $\nu = r(m)$. That is (10) simplifies to

$$0 \le o(x) \le r(m).$$

Even simpler, if the reward function is the identity reward, then, for all x, $0 \le o(x) \le \max S$.

Proof. Since players alternate turns and Positive plays first, either both players have the same number of turns (when Negative plays last) or Positive has one extra turn over Negative (when Positive plays last).

Consider the case where there is no negative reward. If Positive plays a maximal action at each stage of game, then the lowest possible outcome is 0, which happens if Negative also plays exclusively maximal actions, and gets the last move.

The highest possible outcome is obtained when both players play maximal actions and Positive plays both first and last; in this case the outcome is $r(m)$. Positive cannot do any better, unless Negative plays non-maximal move(s), without getting the last move. But then Negative can deviate and play instead only maximal actions, to reduce the result.

In case there are negative rewards, note that no player will play them unless they are forced to because there is no other move, or because they can force the other player to play an even worse negative reward. Moreover, Positive cannot use a negative reward to gain from a parity advantage. Instead Positive should simply play the maximal move directly, even without gaining a parity advantage (or possibly use a smaller positive reward to gain parity advantage). $\square$

Note that the lower bound in Lemma 14, with a negative reward and a minimal action, can be achieved (trivially) as we saw in Example 1.

Theorem 15. *Every game (S, r) has an eventually periodic outcome function. Specifically, let $\ell = \max S$, and let ρ and ν be defined as in Lemma 14. The outcome sequence $o(x'), o(x'+1), \ldots$ is periodic for $x' = (1+\nu-\rho)^\ell$, with period no more than $(1+\nu-\rho)^\ell$.*

Proof. By Lemma 14, the maximal number of outcomes is $\nu - \rho + 1$. The number of combinations of $\nu+\rho+1$ outcomes among ℓ positions is $(1+\nu+\rho)^\ell$.

Therefore a repetition of a sequence of ℓ consecutive outcomes is forced after $(1+\nu+\rho)^\ell$ positions. But optimal play after two such identical outcome sequences is necessarily identical. This implies both results. $\square$

Lemma 16. *Consider a game* (S, r)*. For all sufficiently large pile sizes* x*, a maximal action is optimal.*

Proof. Claim: there can be at most finitely many positions, for each equilibrium strategy profile, for which a non-maximal action $s \neq m$ satisfies $o(x-m) < o(x-s)$. For suppose that one of the players, say Positive, plays $r(m)+1$ non-optimal moves. Then Negative, can play only maximal moves, which will assure that the outcome becomes $< \rho$, as defined in Lemma 14, for sufficiently large heap sizes, a contradiction.

Without loss of generality, we may assume that the finitely many non-maximal actions will be played at last for any such strategy profile. In conclusion, there are at most finitely many pile sizes that require non-maximal actions in optimal play. □

As a consequence of Lemma 16, we have the following result on eventual periodicity.

Theorem 17. *Consider any CS with a unique maximal action* m*. Then the outcome sequences is eventually periodic, with period* $2m$*.*

Proof. If there is a unique maximal action, then by two consecutive such moves, the outcome is the same. The result follows, since, by Lemma 16, both players play maximal actions for all sufficiently large heap sizes. □

As a corollary, we reconfirm Stewart's conjecture (which we proved with more technical details in Section 3), for the case where $r = \text{id}$, that is a game with $r(s) = s$, for all $s \in S$.

Corollary 18. *Consider any game* $(S, id; x, p)$*. It has an eventually periodic outcome function with period* $2 \max S$*, and maximal action* $\max S$*.*

Proof. Since there is a unique maximal action, the results follows by Theorem 17 and Lemma 16. □

In cases where there is not a unique maximal action, the game does not necessarily converge, but the outcome is still eventually periodic, and where the period is some finite linear combination of maximal actions.

Theorem 19. *For any CS, the outcome is eventually periodic, with a period that is a linear combination of the maximal actions.*

Proof. By Theorem 15, the outcome sequence is eventually periodic. Suppose that one of the non-maximal actions contributes to the periodicity. Then, this action is repeated for arbitrary many starting positions. But, this contradicts, the goal of using a non-maximal action, to obtain a parity advantage. □

5 Generalization and open problems

Using similar techniques as in this paper, for 2 player games, one can prove that the eventually periodicity results in the first sections generalize for both general sum and partizan CSs, if the set of maximal actions is the same for both players. Eventually, only maximal action will be optimal for both players. In fact, even if the maximal actions are different for the two players, one can see that it is still optimal for both players to play their respective maximal actions for all but finitely many positions.

Open problem: describe outcomes and convergence properties of partizan CSs.

References

[1] E. Berlekamp, J. H. Conway, R. K. Guy, *Winning Ways for Your Mathematical Plays*, A K Peters Ltd. 2nd edition (2001-2004).

[2] C. L. Bouton, Nim, a game with a complete mathematical theory, *Annals of Mathematics* 3 (1901-1902).

[3] U. Larsson, R. Meir, Properties of cumulative games, in preparation.

[4] J. Von Neumann, Zur theorie der gesellschaftsspiele, *Mathematische Annalen*, 100, 295-320 (1928).

[5] F. Stewart, *Scoring Play Combinatorial Games*, Ph.D. Thesis, University of Dundee (2011), also to appear in: Games of no Chance 5, MSRI, Cambridge University Press (2018).