THE N-VALUE GAME OVER $\mathbb{Z}$ AND $\mathbb{R}$

YIDA GAO, MATT REDMOND, ZACH STEWARD

ABSTRACT. The n-value game is an easily described mathematical diversion with deep underpinnings in dynamical systems analysis. We examine the behavior of several variants of the n-value game, generalizing to arbitrary polygons and various sets. Key results include the guaranteed convergence of the 4-value game over the integers, the cyclic behavior of the 3-value game, and the existence of infinitely many solutions of infinite length in all real-valued games.

1. INTRODUCTION

The n-value game is a deterministic system based on a simple transition rule: from a polygon with labelled vertices, generate a new polygon by placing labelled vertices on the midpoints of its edges. We describe the $n = 4$ case, other polygons generalize naturally. To begin, draw a square and label its vertices with numbers (a, b, c, d). At the midpoint of each edge, write the absolute value of the difference between the edges' endpoints. Finally, connect these midpoints to form a new square. Repeat until all vertices are zero, with the "length" of the game defined as the number of transitions required to reach the zero game. The transition $(a, b, c, d) \to (|b - a|, |c - b|, |d - c|, |a - d|)$ represents this rule. In this paper, we prove key properties of n-value games over different sets. Section 2, authored by Yida Gao and Matt Redmond, investigates the convergence and behavior of the $\{3, 4\}$-value games over $\mathbb{Z}$, and relates the 4-value games over $\mathbb{Z}$ to those over $\mathbb{Q}$. Section 3, authored by Matt Redmond, investigates the general case of an n-value game over $\mathbb{R}$, and demonstrates the existence of an infinite family of infinite-length solutions. Section 4, authored by Zach Steward, considers a combinatorial approach to counting the equivalence classes of the 4-value game over integers in $[0, n - 1]$ for fixed n. Section 5, authored by Matt Redmond and Zach Steward, presents some interesting empirical results about the distribution of path lengths for 4-value games over integers in $[0, n - 1]$.

Date: March 1, 2013.

2. n-VALUE GAMES OVER $\mathbb{Z}$

2.1. **The convergence of the 4-value game.** In this section, we establish that all 4-value games over $\mathbb{Z}$ converge to (0, 0, 0, 0). We accomplish this by demonstrating that each game eventually reduces to a state in which all of its entries are even, and that games which are constant multiples of each other have the same length. This naturally gives a bound on the maximum length of a game, given its starting state.

Lemma 2.1. *If $r \in \mathbb{R}^+$, (ra, rb, rc, rd) has the same length as (a, b, c, d).*

Proof. Consider the entries after t steps of the (ra, rb, rc, rd) game. These entries are equal to r times the entries of the (a, b, c, d) game after t steps by the linearity of subtraction. Suppose the length of the (a, b, c, d) game is L. There must exist some non-zero entry n in step $L - 1$. This implies that in the (ra, rb, rc, rd) game, $rn \neq 0$ at step $L - 1$, so the (ra, rb, rc, rd) game does not end after $L - 1$ steps. Finally, the (a, b, c, d) game ends on step L, with each entry equal to zero, so we must have $r \cdot q = 0$ for each entry q in the Lth step of the (ra, rb, rc, rd) game. Because $r \neq 0$, $q = 0$ for all entries in the Lth step of the (ra, rb, rc, rd) game, so these games have the same length. □

We introduce new notation: let g_t be the vector corresponding to the game g after transitioning for t steps.

Lemma 2.2. *For any given game g, at least one of $\{g_0, g_1, g_2, g_3, g_4\}$ has all even entries.*

Proof. Proof procedes by case analysis over various parities. Let e represent an even element; let o represent an odd element. It is handy to recall rules for subtraction: $e - e = e, e - o = o, o - e = o, o - o = e$.

There are six potential configurations (up to symmetry over D_8) for the parities of the starting game.

(1) $g = (e, e, e, e)$
(2) $g = (e, e, e, o)$
(3) $g = (e, e, o, o)$
(4) $g = (e, o, e, o)$
(5) $g = (e, o, o, o)$
(6) $g = (o, o, o, o)$

Examining each case in turn:

(1) If all entries are even, g itself satisfies our condition.
(2) After one step, $g_1 = (e, e, o, o)$. After two steps, $g_2 = (e, o, e, o)$. After three steps, $g_3 = (o, o, o, o)$. After four steps, $g_4 = (e, e, e, e)$ and we are done.

(3) $g_1 = (e, o, e, o)$. $g_2 = (o, o, o, o)$. $g_3 = (e, e, e, e)$.
(4) $g_1 = (o, o, o, o)$. $g_2 = (e, e, e, e)$.
(5) $g_1 = (o, e, e, o)$. $g_2 = (o, e, o, e)$. $g_3 = (o, o, o, o)$. $g_4 = (e, e, e, e)$.
(6) $g_1 = (e, e, e, e)$.

Each case becomes (e, e, e, e) after at most four steps. □

Theorem 2.3. *All 4-value games over $\mathbb{Z}$ converge to (0, 0, 0, 0)*

Proof. From any starting configuration $G = (a_1, a_2, a_3, a_4)$, take several steps until the game reaches a state where all entries are even. This will take at most four steps, by Lemma 2.2. The new configuration G^{even} can be written as $(2b_1, 2b_2, 2b_3, 2b_4)$. By Lemma 2.1, the length of G^{even} is exactly the same as the game (b_1, b_2, b_3, b_4). However, we are guaranteed that the maximum element in (b_1, b_2, b_3, b_4) has decreased from the maximum element in (a_1, a_2, a_3, a_4). Proceed inductively, by stepping each new game until all entries are even (at most four steps each time), then factor out another 2. As the maximum element is constantly decreasing, each game must terminate in $(0, 0, 0, 0)$ in a finite number of steps. □

Corollary 2.4. *The path length L of a game (a, b, c, d) is bounded above by $4\lceil \log_2(\max(a, b, c, d)) \rceil$*

2.2. **The orbits of the 3-value game.** In this section, we diverge from the 4-value game and consider the 3-value game over $\mathbb{Z}$. We prove that all non-trivial 3-value games cycle, rather than converging to (0,0,0). We accomplish this proof by examining the five cases which encompass all possible 3-value games.

First, let us imagine the values in the tuple as points on a number line. For example, a starting triangle with (1,3,5) looks like this:

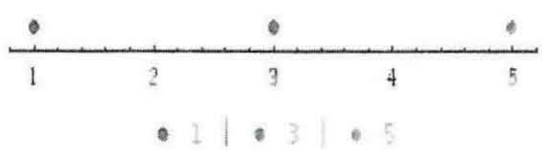


FIGURE 1. A number line with points corresponding to (1,3,5) game state.

Definition 2.5. *Let **range**(G) be defined as the largest positive difference between any two points in a 3-value game tuple:*

$$range(G) = |\max_{g_i \in G}(g_i) - \min_{g_i \in G}(g_i)|$$

Definition 2.6. *A* ***non-trivial 3-value game*** *is one in which the start state is not (x,x,x), where* $x \in \mathbb{Z}$.

Theorem 2.7. *All non-trivial 3-value games over* $\mathbb{Z}$ *cycle in* $(0, x, x)$ *form.*

Proof. The proof is by cases. Consider five possible cases for the non-trivial 3-value game over $\mathbb{Z}$:

(1) One zero and two numbers of the same value $(0, x, x)$: this case enters a cycle that returns a permutation of $(0, x, x)$ on every step.

$$(0, x, x) -> (|0 - x|, |x - x|, |x - 0|) = (x, 0, x)$$

$$(x, 0, x) -> (|x - 0|, |0 - x|, |x - x|) = (x, x, 0)$$

$$(x, x, 0) -> (|x - x|, |x - 0|, |0 - x|) = (0, x, x)$$

(2) One zero and two numbers of different values $(0, x, y)$: in this case, the range decreases by the positive difference of the two non-zero numbers. Without loss of generality, assume $y > x > 0$:

$$(0, x, y) -> (|0 - x|, |x - y|, |y - 0|) = (x, y - x, y)$$

Range of $(0, x, y) = |y - 0| = y$; range of $(x, |x - y|, y) = |y - (y - x)| = x$. In this case, the range decreases by $y - x$.

(3) Two zeros and one non-zero number $(0, 0, x)$: this case only occurs as a start state because two pairs of overlapping points are required to create two zeros and the 3-value game only has three points in total. Range stays the same and the game enters *case 1.*

$$(0, 0, x) -> (0 - 0, |0 - x|, |x - 0) = (0, x, x)$$

Range of $(0, 0, x) = x$; range of $(0, x, x) = x$

(4) Three non-zero values in which two values are the same (x, y, y): The range stays the same and the game transitions to *case 1.*

$$(x, y, y) -> (|x - y|, 0, |y - x|)$$

Range of $(x, y, y) = |x - y|$; range of $(|x - y|, 0, |y - x|) = |x - y|$

(5) Three unique non-zero values (x, y, z): Without loss of generality, $z > y > x$. In this case, the range decreases by $y - x$ or $z - y$, and the new range is $z - y$ or $y - x$. $(x, y, z) \rightarrow (|x - y|, |y - z|, |z - x|) = (y - x, z - y, z - x)$ Original ~~range is $z - y$~~ If $z - y > y - x$, new range $= z - x - (y - x) = z - y$, otherwise new range $= z - x - (z - y) = y - x$. The difference in range is either $z - x - (z - y) = y - x$ or $z - x - (y - x) = z - y$.

For all non-trivial 3-value games, the range is guaranteed to decrease at each step until the game transitions to a (x, y, y) *(case 4)* or $(0, 0, x)$ *(case 3)* state, which both lead to the cycling *case 1* state. Thus, all non-trivial 3-value games over $\mathbb{Z}$ will reduce to *case 1* and cycle. □

2.3. **The equivalence of games over $\mathbb{Z}$ to games over $\mathbb{Q}$.** In this section we use Lemma 2.1, Theorem 2.3, and Theorem 2.6 to extrapolate the behavior of $\{3, 4\}$-value games over $\mathbb{Z}$ to behavior over $\mathbb{Q}$.

Too condensed

See comments

Theorem 2.8. *All 4-value games over $\mathbb{Q}$ converge to (0,0,0,0).*

Proof. Let

$$\left(\frac{n_1}{d_1}, \frac{n_2}{d_2}, \frac{n_3}{d_3}, \frac{n_4}{d_4}\right)$$

represent our 4-value game over $\mathbb{Q}$. By defining a common denominator, $D = d_1d_2d_3d_4$, we can rewrite this as the equivalent game

$$\left(\frac{n_1d_2d_3d_4}{D}, \frac{n_2d_1d_3d_4}{D}, \frac{n_3d_1d_2d_4}{D}, \frac{n_4d_1d_2d_3}{D}\right).$$

In Lemma 2.1 we showed that for any $r \in \mathbb{R}^+$ the two games (a, b, c, d) and (ra, rb, rc, rd) have the same length. In the game above we have $r = \frac{1}{D}$ which if factored out gives us a 4-value game over $\mathbb{Z}$. In Theorem 2.3 we showed that every 4-value game over $\mathbb{Z}$ will converge to $(0, 0, 0, 0)$ and we conclude that by reducing the game over $\mathbb{Q}$ to one over $\mathbb{Z}$ any 4-value game over $\mathbb{Q}$ will converge to (0,0,0,0). □

Theorem 2.9. *All non-trivial 3-value games over $\mathbb{Q}$ cycle in (0,x,x) form.*

Proof. By deduction from Lemma 2.1, we can conclude that a non-trivial 3-value game $\left(\frac{n_1d_2d_3}{D}, \frac{n_2d_1d_3}{D}, \frac{n_3d_1d_2}{D}\right)$ over $\mathbb{Q}$ where $D = d_1d_2d_3$ reduces to a non-trivial 3-value game $(n_1d_2d_3, n_2d_1d_3, n_3d_1d_2)$ over $\mathbb{Z}$, which by Theorem 2.6 cycles in the $(0, x, x)$ form. □

3. n-VALUE GAMES OVER $\mathbb{R}$

Nice intro.

In this section, we consider the properties of the n-valued game over the real numbers. Several questions come to mind: do all real-valued games terminate? If not, does there exist a real-valued game that demonstrates cyclic behavior? If not, does there exist a real-valued game of infinite length? We answer the first question (no) and third question (yes) by proving the existence of infinitely many games with infinite length. We accomplish this by representing a single step of the game as a linear operator (with a restricted domain), then demonstrating the existence of an infinite game for each value of n. Finally,

Since you brought up cyclic behavior, we deserve an answer, even if it's "we don't know."

we show that every infinite length game can be modified to generate infinitely many games of infinite length.

3.1. Linearizing the n-value game. Given an n-value game on $\mathbb{R}$, $G = (a_1, a_2, \dots a_n)$, we produce each step by the transformation rule $G_t \to G_{t+1} = (a_1, a_2, \dots a_n) \to (|a_2 - a_1|, |a_3 - a_2|, \dots |a_1 - a_n|)$. Due to the absolute value, this transformation is not representable as a linear operator; however, if we restrict the domain of the input to the set of vectors $(m_1, m_2, \dots m_n)$ such that $m_1 < m_2 < \dots < m_n$, we can eliminate the use of the absolute value function. $G_t \to G_{t+1} = (a_1, a_2, \dots a_n) \to (a_2 - a_1, a_3 - a_2, \dots a_n - a_1)$. Notice that the last element has had its operands reversed. With this "increasing order" constraint, we can write $G_t \to G_{t+1}$ as an $n \times n$ linear operator T_n:

$$T_n = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & -1 & 1 \\ -1 & 0 & \dots & \dots & 0 & 1 \end{bmatrix}$$

3.2. Identifying an infinite length game for each n. To compute the next element in the game, left-multiply by T_n. As an example, consider the effects of T_4 on $G = (1, 5, 7, 11)$:

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 7 \\ 11 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 4 \\ 10 \end{bmatrix}$$

As this example shows, it is not necessarily the case that the output G_{t+1} maintains the "increasing order" invariant. In general, increasing inputs are not guaranteed to be increasing outputs. For the special case, however, of an increasing eigenvector $\mathbf{v}$ of T_n, we *are* guaranteed that the invariant will hold: the output $\mathbf{v}'$ is guaranteed to be a scalar multiple of $\mathbf{v}$ because $T\mathbf{v} = \lambda\mathbf{v} = \mathbf{v}'$. A scalar multiple of an increasing sequence is an increasing sequence.

If our intial game $\mathbf{v}$ is a real non-zero eigenvector of T_n, then we are guaranteed that $T_n\mathbf{v} = \lambda\mathbf{v} \neq 0$. In general, for all k, $T_n^k\mathbf{v} = \lambda^k\mathbf{v} \neq 0$, so real, increasing eigenvectors of T_n are guaranteed to generate infinite length games.

To demonstrate that there exists an infinite length game for all n, we must demonstrate the existence of a real, increasing, nonzero eigenvector/value pair $\mathbf{v_n}, \lambda_n$ for all n.

3.3. Establishing and bounding a positive real eigenvalue.

$$S_n = T_n - \lambda I_n = \begin{bmatrix} -1-\lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & -1-\lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & -1-\lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & -1-\lambda & 1 \\ -1 & 0 & \dots & \dots & 0 & 1-\lambda \end{bmatrix}$$

Expanding $\det(S_n)$ by cofactors along the bottom row, we see

$$\det(S_n) = -1(-1)^{1+n} \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ -1-\lambda & 1 & 0 & \dots & 0 \\ 0 & -1-\lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1-\lambda & 1 \end{vmatrix} +$$

$$(1-\lambda)(-1)^{n+n} \begin{vmatrix} -1-\lambda & 1 & 0 & \dots & 0 \\ 0 & -1-\lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & -1-\lambda \end{vmatrix}$$

The determinant in the first term reduces to 1, and the determinant in the second term reduces to $(-1-\lambda)^{n-1}$. The characteristic polynomial of T_n is then $(-1)^{2+n}+(1-\lambda)(-1)^{2n}(-1-\lambda)^{n-1} = 0$. Expanding, we have

$$(-1)^{2+n} + (-1-\lambda)^{n-1} - \lambda(-1-\lambda)^{n-1} = 0$$

$$(-1)^{2+n} + \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{n-1-k}(-\lambda)^k - \lambda \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{n-1-k}(-\lambda)^k = 0$$

$$(-1)^{2+n} + \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{n-1}\lambda^k - \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{n-1}\lambda^{k+1} = 0$$

$$(-1)^{2+n} + (-1)^{n-1} \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\lambda^k - \lambda^{k+1}\right) = 0$$

We examine the pattern of signs on this polynomial to determine the number of positive roots. In each case, $\lambda = 0$ is a root, so the coefficient on the constant term is zero.

When n is even, the sign pattern is $(\underbrace{+,\ldots,+}_{\frac{n}{2}},0,\underbrace{-,\ldots,-}_{\frac{n}{2}-1},0)$.

When n is odd, the sign pattern is $(\underbrace{-,\ldots,-}_{\frac{n+1}{2}},\underbrace{+,\ldots,+}_{\frac{n-1}{2}},0)$.

Each case has exactly one change of sign, so there exists exactly one positive real root for each characteristic polynomial by Descartes' Rule of Signs [1]. Let this eigenvalue be λ_n. We claim that $0 < \lambda_n < 1$ for all n - to see this, consider the method for finding a bound on the largest positive real root of a polynomial via synthetic division: dividing a polynomial $P(x)$ by $(x-k)$ will result in a polynomial with all positive coefficients if k is an upper bound for the positive roots [2, Eqn. 15]. Dividing each of the characteristic polynomials by $(\lambda_n - 1)$ (easily done symbolically on a CAS) yields polynomials with all positive coefficients for all n, which demonstrates that 1 is always the least integral upper bound.

3.4. **Identifying an increasing eigenvector.** To determine the corresponding eigenvector $\mathbf{v_n} = (a_1, a_2, \ldots a_n)$, we solve $(T_n - \lambda_n I_n)\mathbf{v_n} = \mathbf{0}$. This produces the following set of equations:

$$\begin{pmatrix} (-1-\lambda_n)a_1 + a_2 = 0 \\ (-1-\lambda_n)a_2 + a_3 = 0 \\ \vdots \\ (-1-\lambda_n)a_{n-1} + a_n = 0 \\ (1-\lambda_n)a_n - a_1 = 0 \end{pmatrix} \text{ or } \begin{pmatrix} (1+\lambda_n)a_1 = a_2 \\ (1+\lambda_n)a_2 = a_3 \\ \vdots \\ (1+\lambda_n)a_{n-1} = a_n \\ (1-\lambda_n)a_n = a_1 \end{pmatrix}$$

Arbitrarily, let $a_n = 1$. This forces $a_1 = (1-\lambda_n)$, which forces $a_2 = (1-\lambda_n)(1+\lambda_n)$. In general, for $1 \le i < n$ we have $a_i = (1-\lambda_n)(1+\lambda_n)^{i-1}$. An eigenvector that corresponds to the eigenvalue λ_n is thus

$$\boxed{\begin{bmatrix} (1-\lambda_n) \\ (1-\lambda_n)(1+\lambda_n) \\ (1-\lambda_n)(1+\lambda_n)^2 \\ \vdots \\ (1-\lambda_n)(1+\lambda_n)^{n-2} \\ 1 \end{bmatrix}}$$

We verify that this eigenvector is in increasing order for all n - given $0 < \lambda_n < 1$, we have $(1-\lambda_n)(1+\lambda_n)^k < (1-\lambda_n)(1+\lambda_n)^{k+1}$ because $(1+\lambda_n)^k < (1+\lambda_n)^{k+1}$ and $(1-\lambda_n) > 0$ when $0 < \lambda_n < 1$. Additionally, we have $(1-\lambda_n)(1+\lambda_n)^{n-2} < 1$ for all $\lambda_n < 1$ because $(1-\lambda_n)(1+\lambda_n)^{n-1} = 1$.

Empirically, for the $n = 4$ case, we have $\lambda_4 \approx 0.839287$, so the eigenvector which generates a game of infinite length is approximately $G = (0.160713, 0.295598, 0.543689, 1)$. The progression of this game after t timesteps results in $G_t = (0.839287)^t \cdot (0.160713, 0.295598, 0.543689, 1)$.

3.5. **Generating infinitely many solutions of infinite length.** Our choice of $a_n = 1$ was arbitrary - the eigenvector we obtained was parametrized only on a_n. Choosing other values of $a_n > 1$ will lead to infinitely many such solutions.

See comments

To see this a different way, consider $w = (a_1, a_2, \ldots a_n) + (k, k, \ldots k) = a + k$ for some constant k. $Tw = (((a_2 + k) - (a_1 + k)), ((a_3 + k) - (a_2 + k)), \ldots ((a_n + k) - (a_1 + k))) = (a_2 - a_1, a_3 - a_2, \ldots, a_n - a_1) = Ta$. Applying the transform T on some starting vector plus a constant yields the same result as applying the transform to the starting vector: $T(a + k) = Ta$. We can choose any value of $k > 0$ and create a different game of infinite length from our starting game.

Finally, we can apply any of the group actions from the symmetry group of the square (D_8) to any 4-value game and preserve its path length, because the actions of D_8 will preserve neighboring vertices. This generates another infinite family of solutions: all cyclic rotations and horizontal/vertical/diagonal reflections of our starting vector.

4. Counting unique 4-value games over $\mathbb{Z}$

In this section we consider a combinatorial approach to determine the number of equivalence classes of a 4-game over the integers from 0 to $n - 1$. For future simulations of empirical cases, we would like to be able to quickly determine the total number of games required for simulation. One may initially think that for any value of n we simply have n^4 possible starting states as we can choose n numbers for each of the four positions. This approach, however, fails to take into account the symmetries of D_8 discussed previously in section 3.5. It is useful for our analysis to recall that the number of ways to choose k elements from a set of n for $n \geq k$ is given by the binomial coefficient

Not a very convincing rationale. I think you wanted this because you wanted to know the length distribution.

if you want to recall this, do it in the proof

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Theorem 4.1. *The number of unique 4-games over the integers from 0 to $n - 1$ as a function of n is given by*

wrong use of this word in math

$$f(n) = \frac{1}{8}\left(n^4 + 2n^3 + 3n^2 + 2n\right)$$

Proof. The proof of $f(n)$ is by cases. Let k define the number of unique integers in a given game and $g(k)$ be the number of unique initial states

distinct

I would say "equivalence classes of" assuming you have set this language up somewhere.

not a useful description

Leaving n out of the notation will probably cause problems later. How about $g_k(n)$?

for a given k. We consider the contributions to $f(n)$ for each case of k and simplify for the explicit expression of $f(n)$.

1. $k = 4$

   When $k = 4$ we are considering a game of the form (a, b, c, d). First we note that there are exactly $\binom{n}{4}$ ways to determine the unique integers a, b, c and d. Given the 4 integers we then have 4! possible orderings. We recall, however, that under symmetry of D_4 there are exactly 8 ways to order the elements (a, b, c, d) that represent the same initial state. There are therefore exactly

$$g(4) = \frac{4!\binom{n}{4}}{8} = 3\binom{n}{4}$$

   unique games for $k = 4$.

2. $k = 3$

   For $k = 3$ we consider games of the form (a, a, b, c). First we have exactly $\binom{n}{3}$ ways to choose the distinct elements a, b and c. We next have 3 ways of choosing which of the 3 elements will be repeated. Now we note that the 4 elements can only be arranged in 1 of 2 possible configurations by considering one of the non repeated elements. Any possible configuration of the elements will leave the unique element b with neighbors of a, a or a, c. We then have exactly

$$g(3) = 2 \cdot 3\binom{n}{3} = 6\binom{n}{3}$$

   unique games for $k = 3$.

3. $k = 2$

   For $k = 2$ there are actually two sub-cases to consider.

   (i) Games of the form (a, a, b, b)
   In this case we will first have $\binom{n}{2}$ ways to determine the unique integers a and b. Next we note that there are only two possible unique configurations of these elements, namely (a, a, b, b) and (a, b, a, b).

   (ii) Games of the form (a, a, a, b)
   In this case we again have $\binom{n}{2}$ ways to determine the unique integers a and b. Next, however, we have to choose which of the integers a or b we wish to repeat 3 times, which there are exactly 2 choices. Finally we note that the only unique configuration is of the form (a, a, a, b).

Each of the two sub-cases contribute a factor of $2\binom{n}{2}$ and we conclude that there are exactly

$$g(1) = 4\binom{n}{2}$$

unique games with $k = 2$.

4. $k = 1$

   In the basic case where we have a game with only 1 unique element it will be of the form (a, a, a, a). It is obvious that any arrangement of the 4 elements will result in the same game and because we have exactly n choices for a we get that there are exactly n games of this form.

$$g(1) = n$$

The total number of unique initial states is then given by

$$f(n) = \sum_{k=1}^{4} g(k) = 3\binom{n}{4} + 6\binom{n}{3} + 4\binom{n}{2} + n$$

By substituting in the definition of the binomial coefficients we have

$$f(n) = \frac{n}{8}(n-1)(n-2)(n-3) + n(n-1)(n-2) + 2n(n-1) + n$$

If we expand each of the terms and collect like terms we find the number of unique initial states is given by

$$f(n) = \frac{1}{8}\left(n^4 + 2n^3 + 3n^2 + 2n\right)$$

□

5. The distribution of game lengths for large n

In this section we make a few empirical observations about path length and consider their implications to gain a better understanding of the dynamics of the 4-game over $\mathbb{Z}$. We first consider the effect of symmetry on the frequency distribution of path length. Next we evaluate the tightness of the bound on path length given in Corollary 2.4 with the computed results. Finally we compare the distribution to the normal probability density function.

I don't understand the point of this page.

5.1. **Accounting for symmetry.** In section 4 we derived an explicit expression for the equivalence classes of a 4-game over the integers from 0 to $n-1$. This, in fact, raises an important question when considering empirical results. Is it really worth it to account for symmetry when approximating the distribution of path lengths for a fixed n? To answer this we let E be the event that the initial state of our game is composed of 4 unique integers (a, b, c, d) and subsequently consider the probability $P(E)$ if we do not account for symmetries about D_8. In order to create a game of this form we will have n choices for a, $n-1$ for b and so on giving us

distinct?

$$P(E) = \frac{n(n-1)(n-2)(n-3)}{n^4}$$

We note that both the numerator and denominator are dominated by a term of n^4 and that the limit for very large n is then given by

$$\lim_{n \to \infty} P(E) = 1$$

Intuitively it makes that as n grows, we become increasingly more likely to choose 4 distinct integers to start our game. From section 4 we know that any game of the specified form (a, b, c, d) is in an equivalence class of size 8 meaning that if we do not account for symmetry on average we will be over counting the number of path lengths by a factor of 8. Now, however, note the relationship between $f(n)$ of section 4 and the total number of games n^4 in the limit

$$\lim_{n \to \infty} \frac{f(n)}{n^4} = \frac{1}{8}$$

So, although we are over counting the vast majority of path lengths by a factor of 8, we are also over counting the total number of games by a factor of 8. The result is that for large enough n we see no qualitative difference in our distribution results and it is therefore not worth the extra computational costs to eliminate the symmetrical cases. As an example of this, consider the two events A and B such that A denotes picking a game of path length 4 from the set of all games not accounting for symmetry and B denotes picking a game of path length 4 from the set of all games *with* symmetry accounted for. The probability that a randomly chosen game from the integers $[0, \dots, n-1]$ has a path length of 4 for various n is shown below

Not what you mean

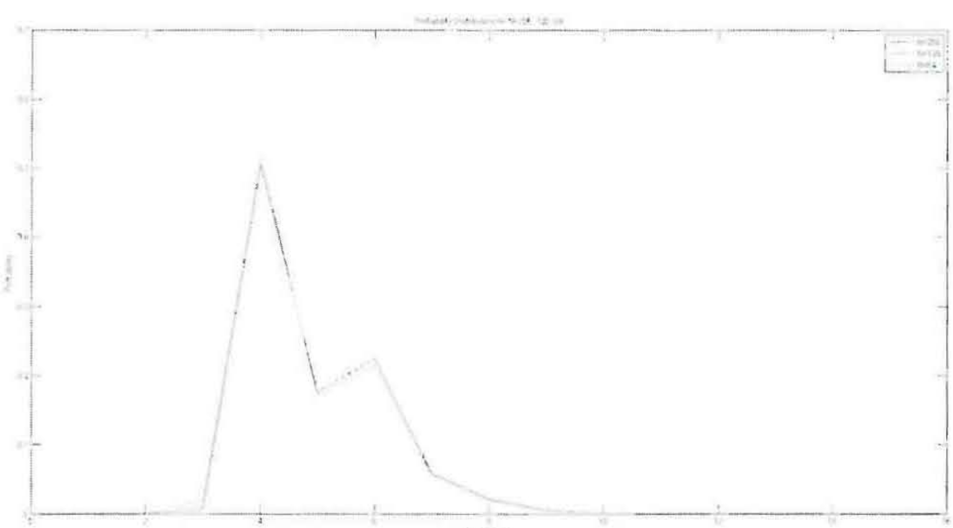
n	$P(A)$	$P(B)$	ϵ
2	0.5000	0.3333	0.1667
4	0.5938	0.1818	0.4119
8	0.5820	0.6066	0.0246
16	0.5513	0.5848	0.0335
32	0.5284	0.5519	0.0235

What is ϵ?

Already for $n = 8$ we are seeing pretty similar results and we conclude that the effects of symmetry for reasonably large n are minor and not worth the additional computation.

5.2. **Theoretical bound for specified path length.** In corollary 2.4 we mention that the path length L of a game (a, b, c, d) can be at most $4\lceil\log_2(\max(a, b, c, d))\rceil$, but we would like to investigate just how good of a bound this really is. In the following to plots we consider the distribution of length over the set of paths computed while *not* accounting for symmetry for reasons mentioned above. Figure 2 plots the path length distribution for $n = 64, 128, 256$ to demonstrate the very close match these distributions have for increasing n.

FIGURE 2. Distribution of Path Length for $n = \{256, 128, 64\}$



First note that for $n = 128$ we have at best $\max(a, b, c, d) = 127$ and therefore have a path length L at most $4 \cdot 7 = 28$, but we are observing a maximum length of only 15. Similarly, for $n = 64$ we observe a maximum length of 13 compared to 24. Furthermore when we increase

You don't know this; all you know is that you can't account for it.

n to 256 and have a new bound on L of at most 32 we observe that in reality we have only gained one more iteration on our maximum path length which is now 16. The reasons for this are non-trivial, but it seems to indicate that our sequences are converging to $(0, 0, 0, 0)$ even faster than the method given in Theorem 2.3.

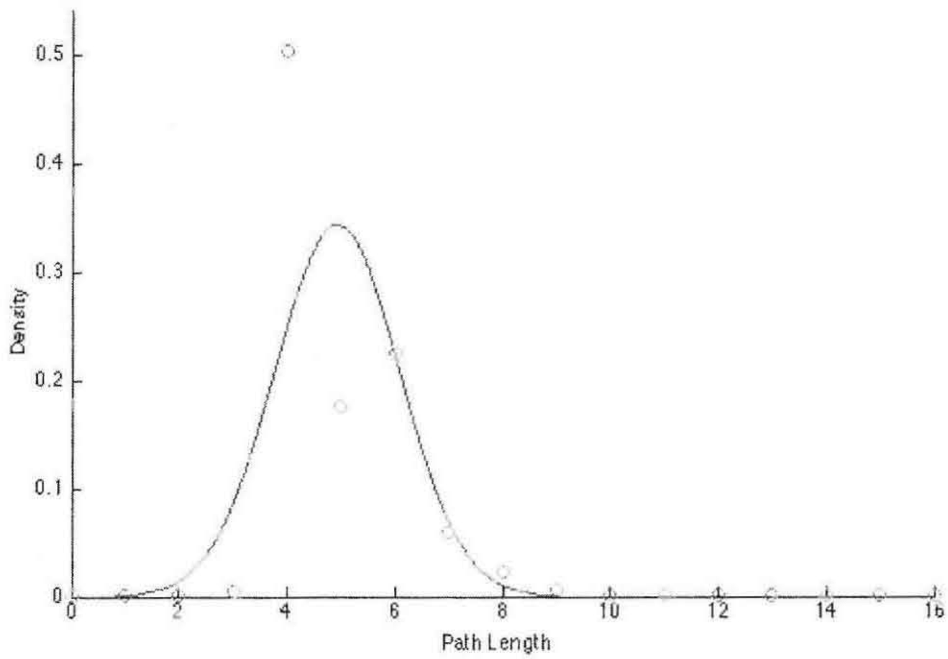
5.3. **Probability.** If we let X be the path length, we can compute the mean and variance of our observations such that

$$E[X] = \sum_{x \in X} p(x) \cdot x = 4.93192197$$

$$Var(X) = E[X^2] - (E[X])^2 = \sum_{x \in X} p(x) \cdot x^2 - \left(\sum_{x \in X} p(x) \cdot x \right)^2 = 1.34398723$$

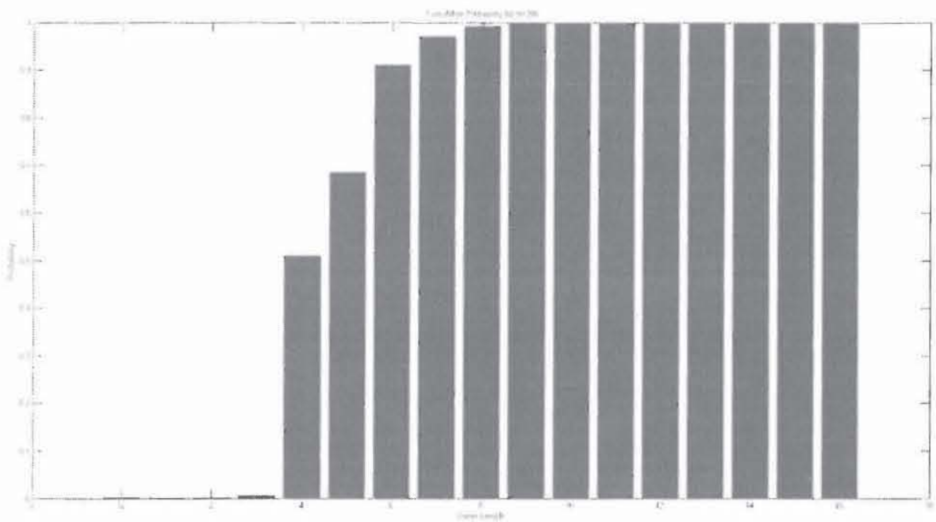
In Figure 3 we now plot the discrete probability distribution of the path length, and this time we include the continuous distribution for a normal random variable with the above specified mean and variance. It is reasonably clear that this data does not follow a normal distribution. Future explorations of this topic may consider modelling the distribution as a mixture of gaussians, or perhaps as a mixture of Poisson distributions.

FIGURE 3. Game Length for $n = 256$ vs. $N(4.931, 1.344)$



We see that the path length data has a much larger right-skew than a gaussian, and maintains a bimodal shape. In Figure 4, is interesting to note that a large number of games converge to the final state $(0, 0, 0, 0)$ after just 4 steps - cumulatively, more than 50% of these games terminate in 4 or fewer steps, and 91% terminate in 6 or fewer steps.

FIGURE 4. Cumulative Distribution of Path Length for $n = 256$



REFERENCES

[1] Wikipedia contributors, "Descartes' rule of signs," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Descartes

[2] Weisstein, Eric W. "Polynomial Roots." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/PolynomialRoots.html