Pistol-Dueling and the Thue-Morse Sequence

Tarkan Al-Kazily

Spring 2017

Contents

1 Introduction

Picture for yourself a pistol duel between two equally poor shots, neither of whom are content to be put at a disadvantage. In the interest of fairness, they agree to take turns based on who currently has the lowest probability of winning the duel by the current round. Indeed, that is what Cooper and Dutle did in their paper "Greedy Galois Games," [2] named in reference to the tragic tale of the young mathematician Évariste Galois. From examining how the sequence of turns behaves as the probability of any one shot successfully felling the opponent vanishes, the Thue-Morse sequence appears and can be used to tell us more than just about pistol dueling. The authors show that while it is not the fastest approximation of the zero function, its sum is exactly zero. Furthermore, they show that it is possible to use the results of the analysis of the Greedy Galois dueling order to produce expansions of of rational numbers in non-standard bases.

2 The Pistol Duel Rules

Without loss of generality, let Alice and Bob be our two duelists, and let Alice go first. Second, let the probability of hitting for each duelist be $p > 0$. Once Alice takes her first shot, regardless of the outcome, she will have a total probability of having won of p. Since Bob has had no opportunity to put Alice out of her misery, Bob's probability of winning will be 0 at the start of the second round. Since Bob has a lower probability of having won than Alice, he has the next go. After taking his first shot, Bob will have a probability of winning of $(1 - p) * p$. This is still less than p, indicating that he gets to shoot again. This goes back and forth until someone succumbs to the their mortal coil.

As you can probably predict, there is a better way of determining the turn order than just recalculating the total probability of victory after each shot until someone dies. Let the turn order be defined by the sequence $\{a_i\}_{i=0}^n$ where $a_n = 1$ means that Alice shoots in round n and $a_n = -1$ means that Bob shoots. The probability $\mathbb{P}(\mathbf{X}_n)$ that a player X wins by round n is given by

$$\mathbb{P}(\mathbf{X}_n) = p \sum_{i \in S_{n,X}} (1 - p)^i$$

where $S_{n,X} = \{i \leq n \mid \text{player } X \text{ shoots in round } i\}$. We get this since a miss has probability $1 - p$, so the probability of player X winning on round n is $p(1-p)^n$, since they not only have to hit, but every shot prior must have missed. Then the probability of winning by round n is the sum of winning on each round the player shoots in up to round n.

We define a sequence of functions $f_n(q)$ where $q = 1 - p$ such that

$$f_n(q) = a_n\left(\sum_{j=0}^{n} a_j q^j\right). \tag{1}$$

These functions can be used to determine the next term of the firing sequence, since they relate the difference in the cumulative probability of winning between Alice and Bob by round n.

$$\begin{aligned}
f_n(q) &= a_n(\sum_{j=0}^{n} a_j q^j) \\
&= \frac{a_n}{p}(p \sum_{i \in S_{n,\mathbf{A}}} q^i - p \sum_{i \in S_{n,\mathbf{B}}} q^i) \\
&= \frac{a_n}{p} \cdot (\mathbb{P}(\mathbf{A}_n) - \mathbb{P}(\mathbf{B}_n)).
\end{aligned}$$

The next player's shot is then found from the relation

$$a_{n+1} = \begin{cases} -a_n & \text{if} f_n(q) > 0 \\ a_n & \text{otherwise} \end{cases} \tag{2}$$

since when $f_n(q) > 0$ the current player has a greater cumulative probability of winning, indicating that the other player gets the next shot. In order to reassure you that this sequence does converge, we will prove the following proposition, and then it will be time to introduce to you the Thue-Morse sequence!

Proposition 2.1. *For each $n \in \mathbb{N}$, there is an $\epsilon > 0$ so that the sequence $\{a_i\}_{i=0}^n$ is the same for all $q \in (1-\epsilon, 1)$.*

Proof. The proof is by induction. Since Alice always goes first, the base case in which $n = 0$ is trivial. Assume by induction that for $n \geq 0$ that for all $q \in (1-\epsilon_0, 1)$, the sequence $\{a_i\}_{i=0}^n$ is the same. a_{n+1} is determined using (2) and $f_n(q)$, which is now fixed for all $q \in (1-\epsilon_0, 1)$ since the first n coefficients do not depend on q. Because f_n is a polynomial of degree n, it has at most n roots. Then we can always find $0 < \epsilon_1 < \epsilon_0$ such that none of its roots occur in $(1-\epsilon_1, 1)$. In this interval, f_n is nonzero, and does not change sign regardless of $q \in (1-\epsilon_1, 1)$, meaning that by (2) a_{n+1} is also independent of q inside this interval, completing the inductive step and the proof.

□

3 Enter the Thue-Morse Sequence

The Thue-Morse sequence is a binary sequence that was discovered, among other individuals, by Norwegian mathematician Axel Thue, who in 1906 found it without any desired goal for its utility [1]. In his work which would become the start of what is known today as *combinatorics on words*, he noted that any binary sequence with length of at least 4 must contain at least one square, i.e. two matching sub sequences that are adjacent. Since it was impossible to find an infinite binary sequence without a square, was it also impossible to find an infinite binary sequence without a cube, in which the same sub sequence repeats three times consecutively?

The answer to his question is the Thue-Morse sequence, an infinite, cube-free, binary sequence of which here are the first twenty terms on the alphabet $\{1, -1\}$

$$1, -1, -1, 1, -1, 1, 1, -1, -1, 1, 1, -1, 1, -1, -1, 1, -1, 1, 1, -1...$$

However, it should be mentioned that there are numerous ways for the Thue-Morse sequence to be defined. They are intimately related to the binary representations of the natural numbers and can be computed by their parities. Alternatively, they can be computed by the substitution

$$\begin{aligned} 1 &\to 1, -1 \\ -1 &\to -1, 1. \end{aligned}$$

In fact, one of the most interesting uses of the Thue-Morse sequence was related to chess. To prevent stalemates, there was a chess rule that would state that a draw would occur when the same sequence of moves occurs three times in succession. However, in 1929, Machgielis Euwe, a Dutch chess grandmaster at the time, constructed the Thue-Morse sequence to show that the rule still allowed the possibility of an infinite game, since the Thue-Morse sequence is cube free. The rule has since been changed to call a draw when the same board state occurs three times, eliminating this loophole.

For our purposes, we will use the following definition when discussing the Thue-Morse sequence, from [4].

Definition 3.1 (**The Thue-Morse Sequence**). *The Thue-Morse sequence* $\{t_i\}_{i=0}^{\infty}$ *on the alphabet* $\{1, -1\}$ *is defined by the following recurrences:*

$$\begin{aligned} t_0 &= 1, \\ t_{2i} &= t_i, \\ t_{2i+1} &= (-1)t_{2i}. \end{aligned}$$

The following proposition from [4] and its corollary will also be useful for our main theorem.

Proposition 3.1. *The sequence* $\{(t_{2i}, t_{2i+1})\}_{i=0}^{\infty}$ *is the Thue-Morse sequence on the alphabet* $\{(1,-1),(-1,1)\}$.

Corollary 3.1. *For any* $n \in \mathbb{N}$, *we have* $\sum_{i=0}^{2n+1} t_i = 0$.

We can now prove the major theorem that the sequence of turns converges to the Thue-Morse sequence.

Theorem 3.1. *The sequence* $\{a_i\}_{i=0}^{\infty}$ *as defined in Section 2 tends to the Thue-Morse sequence (on the alphabet* $\{1,-1\}$*) as* $q \to 1^-$.

Proof. From Proposition 2.1, q can always be taken arbitrarily close to 1, and the sequence will be the same for the first n terms. The proof is by induction again. In [2], the authors calculated for certain n what threshold of probabilities fixed the first n terms of the sequence, by finding the roots of the respective $f_n(q)$'s. They found that for $n = 4$, the first 4 terms of the sequence were fixed for all $q \in (0.67, 1)$. Doing some basic calculations, we see that the start of the sequence of turns is $1, -1, -1, 1$ where 1 is Alice's turn and -1 is Bob's. Since these are equal to the first four terms of the Thue-Morse sequence, we have satisfied our base cases.

For the inductive hypothesis, assume $n > 2$ and that the two sequences agree for all $i \leq n$. We split the work into two cases.

Case 1: $n = 2m$ *is even.* Consider $g(q) = \sum_{i=0}^{n-1} a_i q^i$. Since the a_i are the Thue-Morse sequence, Corollary 3.1 implies that $g(1) = 0$. Since g is continuous, and since we can take q arbitrarily close to 1, we can ensure that $-1/2 < g(q) < 1/2$ for all q under consideration, and for our argument we will restrict $q > (1/2)^{1/n}$. From (1), we know that $f_n(q) = q^n \pm g(q)$, so for all q under consideration near 1,

$$f_n(q) = q^n \pm g(q) > 1/2 - 1/2 \geq 0.$$

Therefore (2) mandates that $a_{n+1} = (-1)a_n$. Since n is even, by induction and the definition of the Thue-Morse sequence

$$a_{n+1} = (-1)a_n = (-1)t_n = t_{n+1}.$$

Case 2: $n = 2m + 1$ *is odd.* Since n is odd, Corollary 3.1 gives that $f_n(1) = 0$. Therefore

$$f_n(q) = (q-1)g(q)$$

where g is a polynomial of degree $2m$.

We will now show that $g(q) = f_m(q^2)$. From our inductive hypothesis the sequence $\{a_i\}_{i=0}^n$ matches the Thue-Morse sequence, so $a_{2i+1} = (-1)a_{2i}$ and $a_{2i} = a_i$ which applies to all of the coefficients in our polynomial. Using these,

$$\begin{aligned}
f_n(q) &= a_{2m+1} \sum_{i=0}^{2m+1} a_i q^i \\
&= a_{2m+1} \sum_{1=0}^{m} (a_{2i}q^{2i} + a_{2i+1}q^{2i+1}) \\
&= (-1)a_{2m} \sum_{i=0}^{m} (a_{2i}q^{2i} - a_{2i}q^{2i+1}) \\
&= (-1)a_{2m}(1-q) \sum_{i=0}^{m} a_{2i}q^{2i} \\
&= (q-1)a_m \sum_{i=0}^{m} a_i (q^2)^i \\
&= (q-1) f_m(q^2).
\end{aligned}$$

Since we can take q arbitrarily close to 1, we can take it close enough such that q and q^2 are such that they are past the threshold for which all the coefficients of $\{a_i\}_{i=0}^{n+1}$ are fixed. Additionally, $q < 1$, so $q - 1 < 0$. Therefore one of $f_n(q)$ and $f_m(q^2)$ is positive, and the other must be negative. Now using our result from (2), this means that for some $j \in \{0, 1\}$, we have that $a_{n+1} = (-1)^j a_n$ and $a_{m+1} = (-1)^{j+1} a_m$. From the inductive hypothesis, the first n terms of our sequence agree with the Thue-Morse sequence, so

$$a_{n+1} = (-1)^j a_n = (-1)^j a_{2m+1} = (-1)^{j+1} a_{2m} = (-1)^{j+1} a_m = a_{m+1}.$$

Since $a_{m+1} = t_{m+1} = t_{2(m+1)} = t_{n+1}$, we have successfully proven that $a_{n+1} = t_{n+1}$ and finished the inductive proof. □

4 The Fairest Sequence of Them All?

Now it begs the question - are our two duelists really getting their fairest shot at survival? In general, is there a turn sequence such that each duelist's chance of survival is as close to $1/2$ as possible, regardless of the probability of hitting? Brought up in [3], we are looking for a sequence $\mathbf{b} = \{b_i\}_{i\geq 0}$ of ± 1's that results in the bias function $D_{\mathbf{b}}(p)$ vanishing as $p \to 0$ as fast as possible, where

$$D_{\mathbf{b}}(p) = \mathbb{P}(\mathbf{A} \text{ survives}) - \mathbb{P}(\mathbf{B} \text{ survives}). \tag{3}$$

Drawing from our earlier representation (1) of the difference in probability of players **A** and **B** winning by round n, we can see that to get the bias function to disappear as quickly as possible, we want our sequence $\mathbf{b}$ to be so that the power series $\sum_{i=0}^{\infty} b_i z^i$ is a close approximation to the zero function for z near 1.

$$\begin{aligned}
p\sum_{i=0}^{\infty} b_i q^i &= p \lim_{n\to\infty} b_n f_n(q) \\
&= \lim_{n\to\infty} \frac{pb_n^2}{p} \cdot (\mathbb{P}(\mathbf{A} \text{ wins by round } n) - \mathbb{P}(\mathbf{B} \text{ wins by round } n)) \\
&= \lim_{n\to\infty} [\mathbb{P}(\mathbf{A} \text{ survives round } n)) - \mathbb{P}(\mathbf{B} \text{ survives round } n))] \\
&= \mathbb{P}(\mathbf{A} \text{ survives}) - \mathbb{P}(\mathbf{B} \text{ survives}) \\
&= D_{\mathbf{b}}(p).
\end{aligned}$$

While the Thue-Morse sequence does provide a good approximation of zero, it only does so within $\exp(-c(\log p)^2)$. In [3], Gunturk proved that there exists a better ± 1 sequence that approximates the zero function within $\exp(-c/p)$.

However, the Thue-Morse sequence still does have some good things going for it in terms of the zero function. It turns out that as $p \to 0$, the bias function of the Thue-Morse sequence is exactly zero. In other words, both players have equal probability of surviving the duel.

Lemma 4.1. *For any probability $p \leq 1/2$, the sequence of coefficients $\{a_i\}_{i=0}^{\infty}$ obtained from the greedy Galois duel defined in Section 2 with hitting probability p satisfies*

$$\sum_{i=0}^{\infty} a_i q^i = 0,$$

where $q = 1 - p$*. Let* $S_{p,X} = \{i \in \mathbb{N} : \text{player X shoots in round } i \text{ given probability of hitting } p\}$ *and this is equivalent to*

$$\sum_{i \in S_{p,A}} q^i = \sum_{i \in S_{p,B}} q^i. \tag{4}$$

Proof. Let $p \leq 1/2$, so that $q \geq p$. We note that

$$\begin{aligned}
\lim_{n\to\infty}(\mathbb{P}(A_n) + \mathbb{P}(B_n)) &= \lim_{n\to\infty}(p \sum_{i \in S_{n,A}} q^i + p \sum_{i \in S_{n,B}} q^i) \\
&= p \sum_{i=0}^{\infty} q^i \\
&= p \frac{1}{1-q} \\
&= 1.
\end{aligned}$$

Since the probabilities of each player winning the duel by round n, $\mathbb{P}(A_n)$ and $\mathbb{P}(B_n)$, cannot decrease as n increases, and since their sum tends to 1, we only need to prove that neither player's probability exceeds $1/2$. This will show that the probability of each player surviving is equal, so that

$$\sum_{i \in S_{p,A}} q^i - \sum_{i \in S_{p,B}} q^i = 0.$$

Suppose not. Without loss of generality, let Alice be the player for which the cumulative probability of having won the duel $\mathbb{P}(A_n)$ is greater than $1/2$ in round n. Let $a = \mathbb{P}(A_{n-1})$ and $b = \mathbb{P}(B_{n-1})$. Since Alice shoots in round n, by the rules of the game $b \geq a$. $\mathbb{P}(A_n) = a + pq^n > 1/2$, and since $\mathbb{P}(A_n) + \mathbb{P}(B_n) = 1$ for all n, Bob's probability will always be less than Alice's from round $n + 1$ and on.

Therefore

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(B_n) &= b + p\sum_{i=n+1}^{\infty} q^i \\ &= b + p\frac{q^{n+1}}{1-q} \\ &= b + q^{n+1} \\ &\geq a + q\cdot q^n \\ &\geq a + pq^n.\end{aligned}$$

However, this results in a contradiction since

$$1 = \lim_{n\to\infty}(\mathbb{P}(A_n) + \mathbb{P}(B_n)) \geq 2(a + pq^n) > 1$$

and we have proven the lemma. □

This lemma provides a fascinating result. From (4), we can see that

$$p^{-1} = \frac{1}{1-q} = \sum_{i=0}^{\infty} q^i = 2\sum_{i\in S_{p,B}} q^i.$$

To get an idea of what this means, here is a specific numerical example. Let $p = 1/16$, implying that $q = 15/16$ and $(2p)^-1 = 8$. Then we have that

$$8 = \sum_{i\in S_{p,B}} (\frac{15}{16})^i = \sum_{i\in S_{p,B}} (\frac{16}{15})^{-i}.$$

If we let $\mathbf{d} = \{d_i\}_{i\geq 0}$ be the sequence such that $d_i = 1$ if $i \in S_{p,B}$ and $d_i = 0$ otherwise, then this means that d_i is a positional binary numeral expansion of 8 using "base" 16/15. For a more precise definition of what an expansion means, an *expansion of* $x \in \mathbb{R}^+$ *in the base* $\beta \geq 1$ (not necessarily an integer) is any right-infinite string of the form

$$c_n c_{n-1}...c_1 c_0.c_{-1}c_{-2}...$$

where n is some nonnegative integer, and for each $k \leq n$, $c_n \in \{0, 1, ..., \lfloor\beta\rfloor\}$, and

$$x = \sum_{k=0}^{\infty} c_{n-k}\beta^{n-k}.$$

This can be an infinite string or a finite string, which occurs when $c_j = 0$ for all $j <$ some $-N$, in which we write

$$c_n c_{n-1} ... c_1 c_0 . c_{-1} c_{-2} ... c_{-N}.$$

These expansions were initially introduced by Rényi, called "β-expansions" [5]. Thus the $\beta = 16/15$ expansion we obtained for 8 is

$$.01101001100101101001 01...$$

Even using just the first 40 digits, we obtain 7.394774. With only 70 digits we obtain an estimate of 7.912343, which is within 1% of 8.

5 Conclusions and Possible Generalizations

These results and their connection to the Thue-Morse sequence are impressive, but a natural question is can this be expanded to a greater numbers of players? The Thue-Morse sequence certainly does have a generalized definition [1], based on the definition from the binary expansions of the natural numbers.

Definition 5.1 (The Generalized Thue-Morse Sequence)**.** *The generalized Thue-Morse sequence is* $\mathbf{t}_{k,m} = (s_k(n) \mod m)_{n \geq 0}$ *for fixed integers* $k \geq 2$ *and* $m \geq 1$.

Using this definition, our version of the Thue-Morse sequence is $\mathbf{t}_{2,2}$. These new Thue-Morse sequences also have the property that they are overlap free, as long as $m \geq k$. For example, here is $\mathbf{t}_{3,3}$

$$012120201120201012201012120120...$$

Now lets suppose that Candide joins Alice and Bob in what would become a Greedy Galois Standoff. Keeping in line with the original turn ordering criteria, here are the rules of this three person "truel":

- The player whose turn is next is the player with the current least probability of having won the duel up to that point.
- In the case that two or more players have the least probability of having won the duel up to that point, turn precedence follows in the order of Alice, then Bob, and finally Candide. Thus, Alice goes first.

- For the sake of simplicity, each participant has brought a shotgun to the truel. If anyone makes their shot (with probability p), then that person hits both their opponents and promptly wins. Because of this, the calculation for determining if a player has won up to their current turn is the same as for a two person duel, namely that:

$$\mathbb{P}(\mathbf{X}_n) = p \sum_{S_{n,X}} (1-p)^i$$

where $\mathbb{P}(\mathbf{X}_n)$ is the probability that player $\mathbf{X}$ wins by round n and $S_{n,X} = \{i \leq n \mid \text{player } X \text{ shoots in round } i\}$.

Following these rules, we get that the first 30 turns for the truel between Alice, Bob and Candide in which $p = 0.001$ and $p = 0.000001$ respectively is

$$(p = 0.001)\ 012210210012120021021120210012...$$
$$(p = 0.000001)\ 012210210012120021210012210102...$$

where 0 indicates that it is Alice's turn, 1 indicates Bob's turn and 2 Candide's. We can see that the sequence does start to converge for smaller probabilities still, and that as it converges every player has the same number of turns after $3k$ rounds, for positive integers k. We can see that this sequence is not square-free, as it contains the square "210210" in the first nine characters, but it is unclear if it will converge to a sequence that features the cube-free property of the Thue-Morse sequence.

Compared to the Thue-Morse sequence $\mathbf{t}_{3,3}$ we see that they are quite different. Both sequences have that every three characters some ordering of 0 1 and 2 appear. However, I'd like to draw parallels between the truel turns and the original Thue-Morse sequence $\mathbf{t}_{2,2}$ since it features an interesting alternation property. Much like how Proposition 3.1 shows that every pair of elements (t_{2i}, t_{2i+1}) of $\mathbf{t}_{2,2}$ are opposite, I conjecture that the sequence of turns in the truel has the following property.

Conjecture 5.1. *Given that $\{j\}_0^\infty$ is the sequence of turns for our truel, for any positive integer k*

$$\begin{aligned} j_{6k} &= j_{6k+5} \\ j_{6k+1} &= j_{6k+4} \\ j_{6k+2} &= j_{6k+3}. \end{aligned}$$

While these rules for a three person duel do not produce the Thue-Morse sequence, other rule sets could provide a better approach. A likely problem with the current rules is that with three players, it is unclear how to account for hitting one of the two players. When any single hit results in a win for the current player, it becomes a game in which each player has twice the number of opportunities to lose than to win, and becomes a more lopsided version of a two player duel. Perhaps a modification to these rules in which a player aims to shoot only the person who would be going next (or alternatively the person with the next lowest probability of winning).

The Thue-Morse sequence and fair turn ordering will always be linked together. It naturally produces a fair distribution between two players, even when the value of each turn is known to be different. The Thue-Morse sequence produces a turn order in which players "take turns taking turns". Even in a pistol duel with equal probabilities of hitting, going first has a distinct advantage over going second when players alternate turns normally. When players take turns shooting first, both players' chances of success become more similar. And when the players follow this pattern of alternation recursively, the Thue-Morse sequence appears.

References

[1] Jean-Paul Allouche and Jeffrey Shallit. "The ubiquitous prouhet-thue-morse sequence". In: *Sequences and their applications*. Springer, 1999, pp. 1–16.

[2] Joshua Cooper and Aaron Dutle. "Greedy Galois Games". In: *The American Mathematical Monthly* 120.5 (2013), pp. 441–451. ISSN: 00029890, 19300972.

[3] C. Sinan Güntürk. "Approximation by power series with 1 coefficients". In: *International Mathematics Research Notices* 2005.26 (2005), p. 1601. DOI: 10.1155/IMRN.2005.1601. eprint: /oup/backfile/content_public/journal/imrn/2005/26/10.1155/imrn.2005.1601/2/2005-26-1601.pdf. URL: +%20http://dx.doi.org/10.1155/IMRN.2005.1601.

[4] M. Lothaire. *Algebraic Combinatorics on Words*. Encyclopedia of Mathematics an. Cambridge University Press, 2002.

[5] A. Rényi. "Representations for real numbers and their ergodic properties". In: *Acta Mathematica Academiae Scientiarum Hungarica* 8.3 (1957), pp. 477–493. ISSN: 1588-2632. DOI: 10.1007/BF02020331. URL: http://dx.doi.org/10.1007/BF02020331.