Analogues of the $3x + 1$ Problem in Polynomial Rings of Characteristic 2
Supplemental Document

Daniel Nichols

Abstract

This document contains additional proof details which were left out of the main paper for clarity and brevity. These are mostly straightforward calculations.

1 Terras' theorem in $\mathbb{F}_2[t]$

1.1 The map Φ_m

Let $\Phi_m : \mathbb{F}_2[t]/t^N \to \{0,1\}^N$ be defined as the function which maps each element $f \in \mathbb{F}_2[t]$ of degree less than N to the first N terms of its parity sequence.

Lemma 1.1. *The map Φ_m described above is a set bijection. That is, every sequence $\{p_0, p_1, \ldots, p_{N-1}\}$ with $p_i \in \{0,1\}$ is the first N terms of the parity sequence of a unique polynomial $f \in \mathbb{F}_2[t]$ with $\deg f < N$. Specifically, the parity sequence determines the initial polynomial f and its N-th iterate $T^N(f)$ as follows, up to choice of q_N:*

$$\begin{aligned} f &= g_{N-1} + t^N q_N, & \deg g_{N-1} &< N \\ T^N(f) &= h_{N-1} + m^{s(N)} q_N, & \deg h_{N-1} &< ds(N) \end{aligned}$$

where $d = \deg m$ and $s(N) = \sum_{i=0}^{N-1} p_i$. Therefore, parity sequences of polynomials in $\mathbb{F}_2[t]$ of degree $< N$ are distributed uniformly in $\{0,1\}^N$.

In the paper we prove this lemma by induction on N. When we come to the inductive step, there are four cases to consider, depending on the values of $h_{N-1}(0)$ and p_N in $\{0,1\}$. Here we give the full proof for all four cases.

Case 1: $h_{N-1}(0) = 0$, $p_N = 0$. That is, the N-th term of the trajectory is 'even' and q_N is also even. Let $q_N = tq_{N+1}$. Then the next term is

$$\begin{aligned} f_{N+1} = \frac{f_N}{t} &= \frac{h_{N-1} + m^{s(N)} q_N}{t} \\ &= \frac{h_{N-1}}{t} + m^{s(N)} q_{N+1} \end{aligned}$$

We can rewrite the initial polynomial as

$$f = g_{N-1} + t^{N+1} q_{N+1}.$$

Since $\deg h_{N-1}/t < s(N) \deg m$ and $\deg g_{N-1} < N + 1$, the theorem holds in this case.

Case 2: $h_{N-1}(0) = 0$, $p_N = 1$. That is, the N-th term of the trajectory is odd and q_N is also odd. Let $q_N = 1 + tq_{N+1}$. Then the next term is

$$\begin{aligned} f_{N+1} &= \frac{m\left[h_{N-1} + m^{s(N)} q_N\right] + 1}{t} \\ &= \frac{mh_{N-1} + m^{s(N+1)} + 1}{t} + m^{s(N+1)} q_{N+1} \end{aligned}$$

Let $h_N = \frac{mh_{N-1} + m^{s(N+1)} + 1}{t}$. Since $\deg h_{N-1} < 2s(N)$, we have $\deg h_N < (\deg m) s(N + 1)$ as required. We rewrite the initial polynomial as

$$\begin{aligned} f &= g_{N-1} + t^N \left(tq_{N+1} + 1\right) \\ &= \left(g_{N-1} + t^N\right) + t^{N+1} q_{N+1}. \end{aligned}$$

Clearly $\deg(g_{N-1} + t^N) < N + 1$, so the theorem holds in this case.

Case 3: $h_{N-1}(0) = 1$, $p_N = 0$. That is, the N-th term of the trajectory is even and q_N is odd. Let $q_N = 1 + tq_{N+1}$. Then the next term is

$$\begin{aligned} f_{N+1} &= \frac{h_{N-1} + m^{s(N)} q_N}{t} \\ &= \frac{h_{N-1} + m^{s(N)}}{t} + m^{s(N+1)} q_{N+1} \end{aligned}$$

Let $h_N = (h_{N-1} + m^X)/t$. Since $\deg h_{N-1} < 2s(N + 1)$, we have $\deg h_N < s(N + 1) \deg m$ as required. Next we rewrite the initial polynomial as

$$\begin{aligned} f &= g_{N-1} + t^N q_N \\ &= g_{N-1} + t^N + t^{N+1} q_{N+1}. \end{aligned}$$

And we know $g_{N-1} + t^N$, has degree less than $N+1$, so the theorem holds in this case.

Case 4: $h_{N-1}(1) = 1$, $p_N = 1$. That is, the N-th term of the trajectory is odd and q_N is even. Let $q_N = tq_{N+1}$. Then the next term is

$$\begin{aligned} f_{N+1} &= \frac{m\left[h_{N-1} + m^{s(N)} q_N\right] + 1}{t} \\ &= \frac{mh_{N-1} + 1}{t} + m^{s(N+1)} q_{N+1}. \end{aligned}$$

Let $h_N = (mh_{N-1} + 1)/t$. This has degree $< 2s(N+1)$ as required. Lastly, we rewrite the initial polynomial:

$$\begin{aligned} f &= g_{N-1} + t^N q_N \\ &= g_{N-1} + t^{N+1} q_{N+1}. \end{aligned}$$

The theorem is satisfied because $\deg g_{N-1} < N+1$.

1.2 Gambler's Ruin

In the paper, we describe how the problem of determining the probability that a polynomial $f \in \mathbb{F}_2[t]$ will have finite stopping time can be formulated as a version of the well-known "gambler's ruin" problem. We prove the following lemma.

Lemma 1.2. *For $k = 0, \ldots, N-1$, let X_k be IID uniform Bernoulli variables and let P_d be defined*

$$P_d = P\left(\exists N > 0 : \sum_{k=0}^{N-1} X_k < \frac{1}{d} N\right).$$

Then $P_1 = P_2 = 1$, and for $d > 2$, P_d is the unique real root of the polynomial $g_d(z) = z^d - 2z + 1$ lying inside the unit disk.

Here we present some additional details of the proof that were left out of the paper to save space.

1.2.1 Solving a linear recurrence

For $d > 2$, let $\lambda_1, \lambda_2, \ldots, \lambda_d$ be the d distinct complex roots of the polynomial $g_d(z) = z^2 - 2z + 1$. In the paper, we write the probability of ruin in this case as

$$P_d = \lim_{W \to \infty} P_{d,W} = \lim_{W \to \infty} (c_1 + c_2 + \ldots + c_d),$$

where c_j are the solutions of the following linear system:

$$\begin{bmatrix} \lambda_1^{-1} & \lambda_2^{-1} & \lambda_3^{-1} & \cdots & \lambda_d^{-1} \\ \lambda_1^{W} & \lambda_2^{W} & \lambda_3^{W} & \cdots & \lambda_d^{W} \\ \lambda_1^{W+1} & \lambda_2^{W+1} & \lambda_3^{W+1} & \cdots & \lambda_d^{W+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_1^{W+d-1} & \lambda_2^{W+d-1} & \lambda_3^{W+d-1} & \cdots & \lambda_d^{W+d-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This system can be solved analytically using Cramer's rule. Let A be the $d \times d$ matrix above and let b be the column vector on the right-hand side of the system. Using Cramer's rule, we write

$$U_0 = \sum_{i=1}^{d} c_i = \frac{\sum_{i=1}^{d} \det A_i}{\sum_{i=1}^{d} \lambda_i^{-1} A_{1,i}} \tag{1}$$

where A_i is the matrix formed by replacing the i-th column of A with b, and $A_{i,j}$ is the i, j cofactor of A.

Because b in this case is just the first standard basis vector, $\det A_i = A_{1,i}$ for each $1 \leq i \leq d$. We compute $A_{1,1}$ as an example; the others follow the exact same pattern.

$$\begin{aligned} \det A_1 &= \det \begin{bmatrix} 1 & \lambda_2^{-1} & \lambda_3^{-1} & \cdots & \lambda_d^{-1} \\ 0 & \lambda_2^{W} & \lambda_3^{W} & \cdots & \lambda_d^{W} \\ 0 & \lambda_2^{W+1} & \lambda_3^{W+1} & \cdots & \lambda_d^{W+1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \lambda_2^{W+d-1} & \lambda_3^{W+d-1} & \cdots & \lambda_d^{W+d-1} \end{bmatrix} \\ &= \det \begin{bmatrix} \lambda_2^{W} & \lambda_3^{W} & \cdots & \lambda_d^{W} \\ \lambda_2^{W+1} & \lambda_3^{W+1} & \cdots & \lambda_d^{W+1} \\ \vdots & \vdots & & \vdots \\ \lambda_2^{W+d-1} & \lambda_3^{W+d-1} & \cdots & \lambda_d^{W+d-1} \end{bmatrix} \\ &= \prod_{j=2}^{d} \lambda_j^{W} \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_2 & \lambda_3 & \cdots & \lambda_d \\ \vdots & \vdots & & \vdots \\ \lambda_2^{d-1} & \lambda_3^{d-1} & \cdots & \lambda_d^{d-1} \end{bmatrix}. \end{aligned}$$

The matrix in the last row above is a Vandermonde matrix with parameters $\lambda_2, \lambda_3, \ldots, \lambda_d$, so its determinant is $\prod_{2 \leq j < k \leq d} (\lambda_k - \lambda_j)$. More generally, for any $1 \leq i \leq d$, let B_i be the determinant of the $(d-1) \times (d-1)$

Vandermonde matrix with parameters $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_d$. Then

$$B_i = \prod_{\substack{1 \le j < k \le d \\ j,k \ne i}} (\lambda_k - \lambda_j)$$

And since $\prod_{j=1}^{d} \lambda_j = 1$, we can write

$$\begin{aligned} \det A_i &= (-1)^{1+i} \prod_{\substack{1 \le j \le d \\ j \ne i}} \lambda_j^W B_i \\ &= (-1)^{1+i} \lambda_i^{-W} B_i. \end{aligned}$$

We can now rewrite equation (1) as follows:

$$U_0 = \frac{\sum_{i=1}^{d} (-1)^{1+i} \lambda_i^{-W} B_i}{\sum_{i=1}^{d} (-1)^{1+i} \lambda_i^{-W-1} B_i}.$$

This makes it clear that if there exists a root λ_1 of $g_d(z)$ with minimal absolute value, then $\lim_{W \to \infty} U_0 = \lambda_1$.

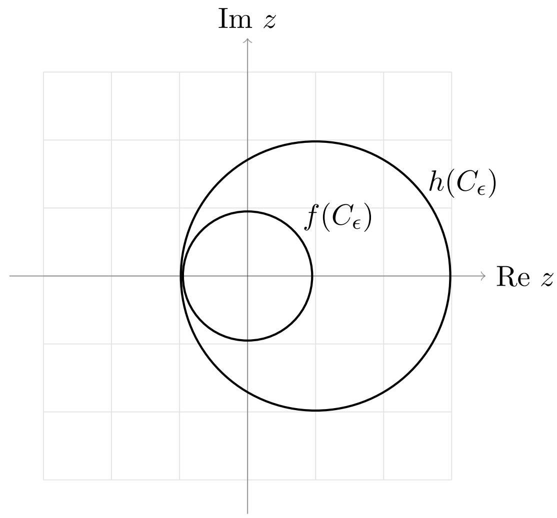
1.2.2 The roots of $g_d(z)$

Here we provide a fully detailed proof that for $d > 2$, $g_d(z) = z^d - 2z + 1$ has a unique root inside the unit disk, and that this root is real and positive. Using Descartes' rule of signs, we determine that there are two positive real roots of $g_d(z)$, one of which is $z = 1$. Since $g_d'(1) = d - 2 > 0$, we know that $g_d(1 - \epsilon) < 0$ for small positive epsilon. On the other hand, $g_d(1/2) = (1/2)^d > 0$, so the other real root must lie in the interval $(1/2, 1)$.

Next, we use Rouche's theorem to prove that there is only one root within the unit circle. Let $f(z) = z^d$ and let $h(z) = -2z + 1$. For small positive ϵ, consider the circle $C_\epsilon = \{z \in \mathbb{C} : |z| = 1 - \epsilon\}$. The function f maps C_ϵ to a smaller circle $|z| = (1 - \epsilon)^d$. Define $m_f(\epsilon) = (1 - \epsilon)^d$. Then $|f(z)| = m_f(\epsilon)$ for all $z \in C_\epsilon$. The other function h maps C_ϵ to a circle of radius $2(1 - \epsilon)$ centered at $z = 1$. The point on this circle closest to the origin is the point $z = -1 + 2\epsilon$, with magnitude $|-1 + 2\epsilon| = 1 - 2\epsilon$. Define $m_h(\epsilon) = 1 - 2\epsilon$. Then for all $z \in C_\epsilon$, $|h(z)| \ge m_h(\epsilon)$. See Figure 1.

We claim that for small positive ϵ, $m_h(\epsilon) > m_f(\epsilon)$ and therefore that $|h(z)| > |f(z)|$ for all $z \in C_\epsilon$. Notice that $m_h(0) = m_f(0) = 1$. Calculating the derivatives of the two functions, we see that $m_h'(0) = -2$ and $m_f'(0) = -d$. By continuity, since $m_h'(0) > m_f'(0)$, $m_h(\epsilon)$ must be greater than $m_f(\epsilon)$ for small positive values of epsilon. Since $|h(z)| > |f(z)|$ for all $z \in C_\epsilon$,

Figure 1:



$g_d(z) = h(z) + f(z)$ must have the same number of roots within C_ϵ as $h(z)$. The function $h(z) = 1 - 2z$ has one root at $z = 1/2$. Therefore, for small positive ϵ, $g_d(z)$ has a unique root inside the circle $|z| = 1 - \epsilon$, which must be the previously mentioned real root lying in the interval $(1/2, 1)$.

2 Terras' theorem in R_r

In the ring $R_r = \mathbb{F}_2[x, t]/(x^2 + tx + r(t))$, we once again formulate the probability that a randomly chosen polynomial has finite $mx + 1$ stopping time as a version of the gambler's ruin problem. We prove the following lemma.

Lemma 2.1. *For $d > 0$, let P_d be defined*

$$P_d = P\left(\exists N > 0 : \sum_{k=0}^{N-1} X_k < \frac{N}{d}\right)$$

where X_i are IID Bernoulli variables taking the value 1 with probability 1/4 and 0 otherwise. If $d \leq 4$, then $P_d = 1$. If $d > 4$, then P_d is the unique root of $g_d(z) = z^d - 4z + 3$ inside the unit disk, which is real and lies in the interval $(3/4, 1)$.

Here we present some additional details of the proof that were left out of the paper to save space.

2.1 Solving a recurrence relation

As in $\mathbb{F}_2[t]$, we first use a recurrence relation to solve the alternate version of the game which ends if the gambler reaches a value of $\$W$. We label U_k the probability of ruin under these conditions given a starting value of $\$k$. Clearly $U_k = -1$ for all $k < 0$ and $U_k = 0$ for all $k \geq W$. For other values of k, we have the following linear recurrence relation.

$$U_k = \frac{3}{4}U_{k-1} + \frac{1}{4}U_{k+d-1}$$

Our goal is to find the value of U_0, representing the probability of ruin (depending on W) starting from a value of 0. If we then take the limit of this quantity as $W \to \infty$, we will learn the actual probability of ruin in a game with no upper limit.

The auxiliary polynomial for the recurrence is $g_d(z) = z^d - 4z + 3$, which is separable as long as $d \neq 4$. When $d = 4$ the root $z = 1$ has multiplicity 2, so we handle this case first. In this case, the solutions to the recurrence equation will take the form $U_k = c_1 + c_2 k + c_3 \lambda^k + c_4 \bar{\lambda}^k$. Since we know that $U_{-1} = 1$ and $U_W = U_{W+1} = U_{W+2} = 0$, we can find the specific solution we need by solving the following linear system:

$$\begin{bmatrix} 1 & -1 & \lambda^{-1} & \bar{\lambda}^{-1} \\ 1 & W & \lambda^{W} & \bar{\lambda}^{W} \\ 1 & W+1 & \lambda^{W+1} & \bar{\lambda}^{W+1} \\ 1 & W+2 & \lambda^{W+2} & \bar{\lambda}^{W+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The quantity we are seeking is then $U_0 = c_1 + c_3 + c_4$. We label the 4×4 matrix above A. Using Cramer's rule, we write

$$\begin{aligned} U_0 &= c_1 + c_3 + c_4 \\ &= \frac{\det A_1}{\det A} + \frac{\det A_3}{\det A} + \frac{\det A_4}{\det A} \\ &= \frac{\det A_1 + \det A_3 + \det A_4}{\det A} \end{aligned}$$

where A_j is the determinant of A with the column j replaced by $[1, 0, \ldots, 0]$. Next, we expand the determinant of A in terms of the cofactors.

$$\det A = A_{1,1} - A_{1,2} + \lambda^{-1} A_{1,3} + \bar{\lambda}^{-1} A_{1,4}.$$

For this linear system, because the right-hand vector b is just the first standard basis vector, the determinant of A with the j-th column replaced by b is the same as the $(1,j)$-cofactor of A. That is, $\det A_i = A_{1,i}$. This allows us to write

$$U_0 = \frac{A_{1,1} + A_{1,3} + A_{1,4}}{A_{1,1} - A_{1,2} + \lambda^{-1} A_{1,3} + \bar{\lambda}^{-1} A_{1,4}}.$$

We argue that $A_{1,1}$ dominates the other terms asymptotically as $W \to \infty$, and therefore that $P_4 = \lim_{W\to\infty} U_0 = 1$. We must express all four cofactors as functions of W.

$$\begin{aligned}
A_{1,1} &= \begin{vmatrix} W & \lambda^W & \bar{\lambda}^W \\ W+1 & \lambda^{W+1} & \bar{\lambda}^{W+1} \\ W+2 & \lambda^{W+2} & \bar{\lambda}^{W+2} \end{vmatrix} = \lambda^W \bar{\lambda}^W \begin{vmatrix} W & 1 & 1 \\ W+1 & \lambda & \bar{\lambda} \\ W+2 & \lambda^2 & \bar{\lambda}^2 \end{vmatrix} \\
&= \lambda^W \bar{\lambda}^W \left(W \begin{vmatrix} \lambda & \bar{\lambda} \\ \lambda^2 & \bar{\lambda}^2 \end{vmatrix} - \begin{vmatrix} W+1 & \bar{\lambda} \\ W+2 & \bar{\lambda}^2 \end{vmatrix} + \begin{vmatrix} W+1 & \lambda \\ W+2 & \lambda^2 \end{vmatrix} \right) \\
&= \lambda^W \bar{\lambda}^W \left[W \left(\lambda\bar{\lambda}^2 - \lambda^2\bar{\lambda} \right) - \left(W\bar{\lambda}^2 + \bar{\lambda}^2 - W\bar{\lambda} - 2\bar{\lambda} \right) + \left(W\lambda^2 + \lambda^2 - W\lambda - 2\lambda \right) \right] \\
&= \lambda^W \bar{\lambda}^W \left[W \left(\lambda\bar{\lambda}^2 - \lambda^2\bar{\lambda} + \lambda^2 - \bar{\lambda}^2 + \bar{\lambda} - \lambda \right) + \lambda^2 - \bar{\lambda}^2 + 2\bar{\lambda} - 2\lambda \right] \\
&\text{Here we use the fact that } \lambda \text{ and } \bar{\lambda} \text{ are the roots of } x^2 + 2x + 3. \\
&= 3^W \left[W \left(3\bar{\lambda} - 3\lambda - 2\lambda - 3 + 2\bar{\lambda} + 3 + \bar{\lambda} - \lambda \right) - 2\lambda - 3 + 2\bar{\lambda} + 3 + 2\bar{\lambda} - 2\lambda \right] \\
&= 6(\bar{\lambda} - \lambda) W\, 3^W + 4(\bar{\lambda} - \lambda) 3^W.
\end{aligned}$$

$$\begin{aligned}
A_{1,2} &= - \begin{vmatrix} 1 & \lambda^W & \bar{\lambda}^W \\ 1 & \lambda^{W+1} & \bar{\lambda}^{W+1} \\ 1 & \lambda^{W+2} & \bar{\lambda}^{W+2} \end{vmatrix} = -\lambda^W \bar{\lambda}^W \begin{vmatrix} 1 & 1 & 1 \\ 1 & \lambda & \bar{\lambda} \\ 1 & \lambda^2 & \bar{\lambda}^2 \end{vmatrix} \\
&= -3^W \left[\left(3\bar{\lambda} - 3\lambda \right) - \left(\bar{\lambda}^2 - \bar{\lambda} \right) + \left(\lambda^2 - \lambda \right) \right] \\
&= -3^W \left[3\bar{\lambda} - 3\lambda + 3\bar{\lambda} + 3 - 3\lambda - 3 \right] \\
&= -6(\bar{\lambda} - \lambda) 3^W.
\end{aligned}$$

$$\begin{aligned}
A_{1,3} &= \begin{vmatrix} 1 & W & \bar{\lambda}^W \\ 1 & W+1 & \bar{\lambda}^{W+1} \\ 1 & W+2 & \bar{\lambda}^{W+2} \end{vmatrix} = \bar{\lambda}^W \begin{vmatrix} 1 & W & 1 \\ 1 & W+1 & \bar{\lambda} \\ 1 & W+2 & \bar{\lambda}^2 \end{vmatrix} \\
&= \bar{\lambda}^W \left(\begin{vmatrix} W+1 & \bar{\lambda} \\ W+2 & \bar{\lambda}^2 \end{vmatrix} - W \begin{vmatrix} 1 & \bar{\lambda} \\ 1 & \bar{\lambda}^2 \end{vmatrix} + \begin{vmatrix} 1 & W+1 \\ 1 & W+2 \end{vmatrix} \right) \\
&= \bar{\lambda}^W \left[\left(W\bar{\lambda}^2 + \bar{\lambda}^2 - W\bar{\lambda} - 2\bar{\lambda} \right) - W \left(\bar{\lambda}^2 - W\bar{\lambda} \right) + (W + 2 - W - 1) \right] \\
&= \bar{\lambda}^W \left(\bar{\lambda}^2 - 2\bar{\lambda} + 1 \right) \\
&= \bar{\lambda}^W \left(-4\bar{\lambda} - 2 \right).
\end{aligned}$$

$$\begin{aligned}
A_{1,4} &= -\begin{vmatrix} 1 & W & \lambda^W \\ 1 & W+1 & \lambda^{W+1} \\ 1 & W+2 & \lambda^{W+2} \end{vmatrix} = -\lambda^W \begin{vmatrix} 1 & W & 1 \\ 1 & W+1 & \lambda \\ 1 & W+2 & \lambda^2 \end{vmatrix} \\
&= -\lambda^W \left(\begin{vmatrix} W+1 & \lambda \\ W+2 & \lambda^2 \end{vmatrix} - W \begin{vmatrix} 1 & \lambda \\ 1 & \lambda^2 \end{vmatrix} + \begin{vmatrix} 1 & W+1 \\ 1 & W+2 \end{vmatrix} \right) \\
&= -\lambda^W \left[\left(W\lambda^2 + \lambda^2 - W\lambda - 2\lambda \right) - W\left(\lambda^2 - W\lambda \right) + (W+2-W-1) \right] \\
&= -\lambda^W \left(\lambda^2 - 2\lambda + 1 \right) \\
&= -\lambda^W \left(-4\lambda - 2 \right).
\end{aligned}$$

To summarize, the asymptotic growth rates of the cofactors are:

$$\begin{aligned}
A_{1,1} &\sim W\,3^W \\
A_{1,2} &\sim 3^W \\
A_{1,3} &\sim \lambda^W \\
A_{1,4} &\sim \bar{\lambda}^W.
\end{aligned}$$

It is clear that $A_{1,1}$ dominates the other cofactors as $W \to \infty$. Since the numerator and denominator have the same dominant term with the same coefficient, the probability of ruin in this case is

$$P_4 = \lim_{W\to\infty} P_{4,W} = 1.$$

For $d \neq 4$ we have $\gcd(f, f') = 1$, so in this case the polynomial is separable. Therefore every solution must have the form $U_k = c_1\lambda_1^k + c_2\lambda_2^k + \ldots + c_d\lambda_d^k$. The linear system we must solve is exactly the same as the one we found in $\mathbb{F}_2[t]$, except that the roots λ_i are now the roots of $z^d - 4z + 3 = 0$.

$$\begin{bmatrix} \lambda_1^{-1} & \lambda_2^{-1} & \lambda_3^{-1} & \cdots & \lambda_d^{-1} \\ \lambda_1^{W} & \lambda_2^{W} & \lambda_3^{W} & \cdots & \lambda_d^{W} \\ \lambda_1^{W+1} & \lambda_2^{W+1} & \lambda_3^{W+1} & \cdots & \lambda_d^{W+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_1^{W+d-1} & \lambda_2^{W+d-1} & \lambda_3^{W+d-1} & \cdots & \lambda_d^{W+d-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

We can solve this system in the same way, using Cramer's rule and Vandermonde determinants. The product of all the roots is still the constant term of $g_d(z)$, which in this case is $\prod_{j=1}^{d} \lambda_j = 3$. So $\det A_i =$

$(-1)^{1+i}3^W\lambda_i^{-W}B_i$, and the solution to the recurrence relation is

$$\begin{aligned} P_{d,W} = U_0 &= \sum_{j=1}^{d} c_j \\ &= \sum_{j=1}^{d} \frac{\det A_j}{\det A} \\ &= \frac{\sum_{j=1}^{d}(-1)^{1+j}3^W\lambda_j^{-W}B_j}{\sum_{j=1}^{d}(-1)^{1+j}3^W\lambda_j^{-W-1}B_j} \end{aligned}$$

where B_j are defined as Vandermonde determinants as before. Just as in $\mathbb{F}_2[t]$, if λ_1 is a real root with strictly smaller absolute value than all of the others, then the limit of the above quantity is

$$P_d = \lim_{W\to\infty} U_0 = \lambda_1.$$

2.2 Roots of $g_d(z)$

We will now examine the polynomial $g_d(z) = z^d - 4z + 3$ and show that when $d \neq 4$, such a root does in fact exist.

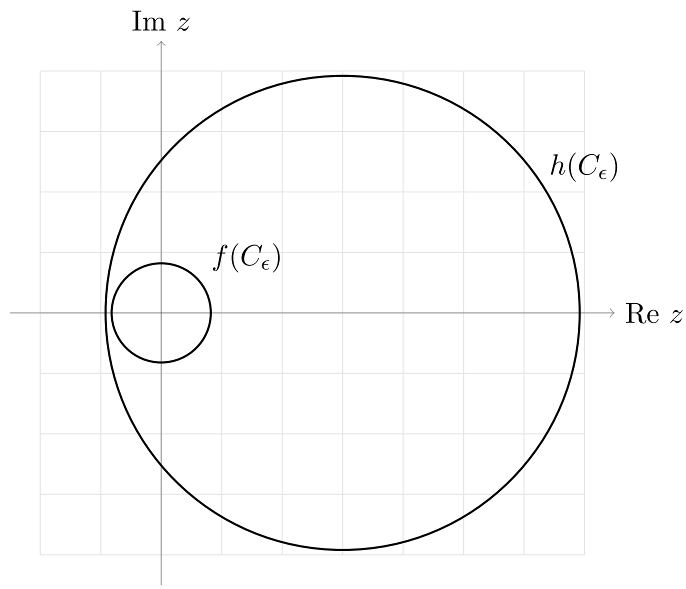
$d = 2$: The polynomial $g_2(z) = z^2 - 4z + 3$ has roots at $z = 1$ and $z = 3$. Since $z = 1$ is the root with the smallest magnitude, the probability of ruin is 1.

$d = 3$: We write $g_3(z) = z^3 - 4z + 3 = (z - 1)(z^2 + z - 3)$. Using the quadratic formula, we find that the roots of z^2+z-3 are $z = -\frac{1}{2} \pm \frac{\sqrt{13}}{2}$, both of which have magnitude > 1. Since $z = 1$ is the root with the smallest magnitude, the probability of ruin is 1.

$d > 4$: Using Descartes' rule of signs, we determine that there are two positive real roots of $g_d(z)$. We know that one of these is $z = 1$. Since $g_d'(1) = d-4 > 0$, we know that $g_d(1-\epsilon) < 0$ for small positive epsilon. On the other hand, $g_d(3/4) = (3/4)^d > 0$. Therefore, the other real root must lie in the interval $(3/4, 1)$.

Next, we use Rouche's theorem to prove that there is only one root within the unit circle. Let $f(z) = z^d$ and let $h(z) = -4z+3$. For small positive ϵ, consider the circle $C_\epsilon = \{z \in \mathbb{C} : |z| = 1 - \epsilon\}$. The function f maps C_ϵ to a smaller circle $|z| = (1 - \epsilon)^d$. Define $m_f(\epsilon) = (1 - \epsilon)^d$. Then $|f(z)| = m_f(\epsilon)$ for all $z \in C_\epsilon$.

The other function h maps C_ϵ to a circle of radius $4(1-\epsilon)$ centered at $z=3$. The point on this circle closest to the origin is the point $z=-1+4\epsilon$, with magnitude $|-1+4\epsilon| = 1-4\epsilon$. Define $m_h(\epsilon) = 1-4\epsilon$. Then for all $z \in C_\epsilon$, $|h(z)| \geq m_h(\epsilon)$.



We claim that for small positive ϵ, $m_h(\epsilon) > m_f(\epsilon)$ and therefore that $|h(z)| > |f(z)|$ for all $z \in C_\epsilon$. Notice that $m_h(0) = m_f(0) = 1$. Calculating the derivatives of the two functions, we see that $m_h'(0) = -4$ and $m_f'(0) = -d$. By continuity, since $m_h'(0) > m_f'(0)$, $m_h(\epsilon)$ must be greater than $m_f(\epsilon)$ for small positive values of epsilon.

Since $|h(z)| > |f(z)|$ for all $z \in C_\epsilon$, $g_d(z) = h(z) + f(z)$ must have the same number of roots within C_ϵ as $h(z)$. The function $h(z) = 3 - 4z$ has one root at $z = 3/4$. Therefore, for small positive ϵ, $g_d(z)$ has a unique root inside the circle $|z| = 1-\epsilon$, which must be the previously mentioned real root lying in the interval $(3/4, 1)$. Since this root has the smallest magnitude among roots of $g_d(z)$, the value of this root is the probability of ruin P_d.