involve
a journal of mathematics

Iterations of quadratic polynomials over finite fields

William Worden

msp

2013 vol. 6, no. 1

INVOLVE 6:1 (2013)
dx.doi.org/10.2140/involve.2013.6.99

Iterations of quadratic polynomials over finite fields

William Worden

(Communicated by Michael Zieve)

Given a map $f : \mathbb{Z} \to \mathbb{Z}$ and an initial argument α, we can iterate the map to get a finite forward orbit modulo a prime p. In particular, for a quadratic map $f(z) = z^2 + c$, where c is constant, work by Pollard suggests that the forward orbit should have length on the order of $\sqrt{p}$. We give a heuristic argument that suggests that the statistical properties of this orbit might be very similar to the birthday problem random variable X_n, for an $n = p$ day year, and offer considerable experimental evidence that the limiting distribution of the orbit lengths, divided by $\sqrt{p}$, for $p \le x$ as $x \to \infty$, converges to the limiting distribution of $X_n/\sqrt{n}$, as $n \to \infty$.

1. Introduction

Let $f \in \mathbb{Z}[z]$ be a polynomial and let $\alpha \in \mathbb{Z}$. We define the orbit of α under f to be

$$\mathcal{O}_f(\alpha) = \{f^n(\alpha) : n = 0, 1, 2, 3, \dots\},$$

and for each prime p we define the orbit modulo p of α under f to be

$$\mathcal{O}_f^p(\alpha) = \{f^n(\alpha) \bmod p : n = 0, 1, 2, 3, \dots\},$$

where f^n is the n-th iterate of f:

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n},$$

and $f^0(\alpha) = \alpha$. For a fixed f and α and a given prime p, let m_p be the size of $\mathcal{O}_f^p(\alpha)$.

If f is a random map, i.e., a map chosen from the uniformly distributed set consisting of all maps from $\mathbb{F}_p$ into $\mathbb{F}_p$ (see [Harris 1960]), then the values of $f^n(\alpha)$

MSC2010: primary 37P05; secondary 11B37.
Keywords: arithmetic dynamics, birthday problem, forward orbit modulo p, random maps.
This research was conducted while the author was an undergraduate at The City College of New York, CUNY. It was funded by a Rich Summer Internship grant from the Dr. Barnett and Jean-Hollander Rich Scholarship Fund. We are very grateful for the opportunity that this grant allowed, and would like to thank the selection committee and those who have made contributions to the fund.

are uniformly distributed for all n, and all α, and so the probability that $f^0(\alpha)$, $f^1(\alpha)$, $f^2(\alpha), \ldots, f^k(\alpha)$ are all different is

$$1 \cdot \frac{p-1}{p} \cdot \frac{p-2}{p} \cdot \ldots \cdot \frac{p-k}{p} = \frac{(p-1)!}{p^k(p-k-1)!},$$

since, once α is fixed, there are $p-1$ choices for $f^1(\alpha)$, $p-2$ choices for $f^2(\alpha)$, and so on. Therefore, in this case the probability that (at least) two of $f^0(\alpha)$, $f^1(\alpha)$, $f^2(\alpha), \ldots, f^k(\alpha)$ are equal is

$$q_k^{(p)} = 1 - \frac{(p-1)!}{p^k(p-k-1)!}.$$

By an analogous argument, $q_k^{(p)}$ is also the probability that, among k people, two people have the same birthday, where p is the number of days in a year. Framing this a little differently, we let the random variable X_n be the number of times that we must sample (uniformly, with replacement) from the set $\{1, 2, 3, \ldots, n\}$ to get a repetition. Since it is known that the expected value of this variable is on the order of $\sqrt{n}$, we look instead at the variable $X_n/\sqrt{n}$.

In light of the above heuristic, we might expect that, for a fixed polynomial f and initial value α, $m_p/\sqrt{p}$ will, on average, "behave" similarly to $X_n/\sqrt{n}$. In particular, we might guess that the limiting distribution of $m_p/\sqrt{p}$, for $p \le x$, $x \to \infty$, will be similar to the limiting distribution of $X_n/\sqrt{n}$, as $n \to \infty$. We note that the above heuristic is not new; similar arguments have been given by Pollard [1975], Bach [1991], and Brent [1980] to name a few, leading to conjectures that m_p is on average approximately equal to $\sqrt{(\pi/2)\,p}$.

We also consider a related question. For a fixed $f \in \mathbb{Z}[z]$, $\alpha \in \mathbb{Z}$, let

$$\mathcal{Q}_{f,\alpha}(x) = \{p \le x : f^n(\alpha) \equiv 0 \pmod{p} \text{ for some } n = 0, 1, 2, \ldots\}.$$

That is, $\mathcal{Q}_{f,\alpha}(x)$ is the set of primes p less than or equal to x such that 0 appears in the orbit modulo p of α under f. In particular, we are interested in the size of $\mathcal{Q}_{f,\alpha}(x)$. Since, for a given prime p, the proportion of elements mod p in the orbit of α under f is m_p/p, we hypothesize that $|\mathcal{Q}_{f,\alpha}(x)|$ will grow at a rate proportional to m_p/p. Therefore, if we are correct that m_p will grow at a rate proportional to $\sqrt{p}$, we might expect that

$$|\mathcal{Q}_{f,\alpha}(x)| = \sum_{p \le x} \frac{m_p}{p} \approx c \cdot \frac{\sqrt{x}}{\log x}$$

for some constant $c \in \mathbb{R}$. The approximation above is discussed further in Section 3, where we derive the appropriate constant c.

In the following we take an experimental approach to studying properties of the set $m_p/\sqrt{p}$. For selected maps f and initial values α, we compute the orbits

modulo p for all $p \le 2^{25}$. In particular, given these orbits we can find the moments of $m_p/\sqrt{p}$, and the length of $\mathcal{Q}_{f,\alpha}(x)$. As we will demonstrate in the sections to follow, our results give strong support to the above heuristic, and lead us to make the following conjectures:

Conjecture 1. *Let $f(z) = z^2 + c$ and $\alpha \in \mathbb{Z}$ be such that*

(1) $c \in \mathbb{Z} \setminus \{0, -2\}$,

(2) $\alpha \neq \pm\frac{1}{2}(1 \pm \sqrt{1-4c})$, $\alpha \neq \pm\frac{1}{2}(1 \pm \sqrt{-3-4c})$, $\alpha \neq 0, \pm 1$ *when* $c = -1$,

and let the orbit length m_p be as defined above. Then, as $x \to \infty$, the distribution of $m_p/\sqrt{p}$ converges, independent of f and α, to a continuous distribution $F(t) = 1 - e^{-t^2/2}$, $t \ge 0$. In particular, the r-th moments of $m_p/\sqrt{p}$ are given by $\mu_r = r(r-2)(r-4)\cdots 2$ for r even, and $\mu_r = r(r-2)(r-4)\cdots 1 \cdot \sqrt{\frac{\pi}{2}}$ for r odd.

The motivation for the result conjectured above is elaborated upon in Section 2, and the need to include conditions (1) and (2) for both conjectures is explained in Section 4.

Conjecture 2. *Let $f(z) = z^2 + c$ and $\alpha \in \mathbb{Z}$ be such that conditions* (1) *and* (2) *of Conjecture 1 hold, and $\alpha^2 \neq -c$. Define*

$$\mathcal{Q}_{f,\alpha}(x) = \{p \le x : f^n(\alpha) \equiv 0 \pmod{p} \text{ for some } n \ge 0\}.$$

Then

$$\lim_{x\to\infty} |\mathcal{Q}_{f,\alpha}(x)| \frac{\log x}{\sqrt{x}} = \sqrt{2\pi}.$$

2. Length of the orbit modulo p and the birthday problem

Let E_k be the k-th number drawn uniformly from the set $\{1, 2, 3, \ldots, n\}$, with replacement, and let X_n be as defined in Section 1. Then for $k \le n$ we have

$$\begin{aligned} P(X_n > k) &= P(E_1, \ldots, E_k \text{ all take different values}) \\ &= \prod_{j=2}^{k} \big(1 - P(E_j = E_i \text{ for some } i < j)\big) \\ &= \prod_{j=2}^{k} \left(1 - \frac{j-1}{n}\right) = \exp \sum_{j=1}^{k-1} \log(1 - j/n). \end{aligned}$$

So as $n \to \infty$, we have the following for $0 \le t \le \sqrt{n}$:

$$\begin{aligned} \lim_{n\to\infty} P(X_n/\sqrt{n} > t) &= \lim_{n\to\infty} P(X_n > t\sqrt{n}) = \lim_{n\to\infty} \exp \sum_{j=1}^{\lfloor t\sqrt{n}\rfloor} \log(1 - j/n) \\ &= \lim_{n\to\infty} \exp\left(-\sum_{j=1}^{\lfloor t\sqrt{n}\rfloor} \sum_{k=1}^{\infty} \frac{(j/n)^k}{k}\right), \end{aligned}$$

where we have used the power series representation for $\log(1-j/n)$ in the third line. Switching the order of summation, and pulling the first term of the sum over k out of the exponential, we have

$$\begin{aligned}
&\lim_{n\to\infty} P(X_n/\sqrt{n} > t) \\
&\qquad = \lim_{n\to\infty} \exp\left(-\sum_{j=1}^{\lfloor t\sqrt{n}\rfloor} j/n\right) \cdot \lim_{n\to\infty} \exp\left(-\sum_{k=2}^{\infty}\sum_{j=1}^{\lfloor t\sqrt{n}\rfloor} \frac{(j/n)^k}{k}\right) \\
&\qquad \approx \lim_{n\to\infty} \exp\left(-\frac{t\sqrt{n}(t\sqrt{n}+1)}{2n}\right) \cdot \lim_{n\to\infty} \exp\left(-\sum_{k=1}^{\infty} O\left(\frac{t^{k+2}}{k\,n^{k/2}}\right)\right) \\
&\qquad \approx e^{-t^2/2} \cdot \exp \sum_{k=1}^{\infty} \lim_{n\to\infty} O\left(\frac{t^{k+2}}{k\,n^{k/2}}\right) = e^{-t^2/2},
\end{aligned}$$

where the second line follows because, in general, $\sum_{j=1}^{m} j^k$ is a polynomial in m of degree $k+1$, and the third line, where we have brought the limit inside the sum, follows from the monotone convergence theorem. Therefore

$$\lim_{n\to\infty} P(X_n/\sqrt{n} \le t) = 1 - e^{-t^2/2},$$

so we see that the distribution of $X_n/\sqrt{n}$ converges to a distribution function $F(t) = 1-e^{-t^2/2}$, which has an associated density function $f(t) = F'(t) = te^{-t^2/2}$. To support our conjecture in Section 1 — that $F(t)$ is the limiting distribution of $m_p/\sqrt{p}$, as $x \to \infty$ — we compare the moments of $m_p/\sqrt{p}$, which we compute in Section 5 for large x, to the limiting moments of $X_n/\sqrt{n}$, as $n \to \infty$. With the limiting density function $f(t)$ of $X_n/\sqrt{n}$ in hand we can derive a general expression for the r-th moment:

$$\begin{aligned}
\mu_r &= \int_0^\infty t^r f(t)\,\mathrm{d}t = \int_0^\infty t^{r+1} e^{-t^2/2}\,\mathrm{d}t \\
&= -t^r e^{-t^2/2}\Big|_0^\infty + r\int_0^\infty t^{r-1} e^{-t^2/2}\,\mathrm{d}t = r\int_0^\infty t^{r-1} e^{-t^2/2}\,\mathrm{d}t,
\end{aligned}$$

where r applications of l'Hôpital's rule give us 0 for the $-t^r e^{-t^2/2}$ term. We continue applying integration by parts as above until we get

$$\begin{aligned}
\mu_r &= r(r-2)(r-4)\cdots 2 \cdot \int_0^\infty te^{-t^2/2}\,\mathrm{d}t \quad \text{if } r \text{ is even,} \\
\mu_r &= r(r-2)(r-4)\cdots 1 \cdot \int_0^\infty e^{-t^2/2}\,\mathrm{d}t \quad \text{if } r \text{ is odd.}
\end{aligned}$$

The first integral above evaluates to $-e^{-t^2/2}\big|_0^\infty = 1$, and the second integral we evaluate as follows:

$$
\begin{aligned}
I &= \int_0^\infty e^{-t^2/2}\,\mathrm{d}t \\
\Longrightarrow \quad (2I)^2 &= \left(\int_{-\infty}^\infty e^{-t^2/2}\,\mathrm{d}t\right)^2 \\
&= \int_{-\infty}^\infty e^{-x^2/2}\,\mathrm{d}x \cdot \int_{-\infty}^\infty e^{-y^2/2}\,\mathrm{d}y \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{r=0}^\infty \int_{\theta=0}^{2\pi} r e^{-r^2/2}\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi \\
\Longrightarrow \qquad I &= \sqrt{\frac{\pi}{2}}.
\end{aligned}
$$

Therefore the r-th moments of the limiting distribution of $X_n/\sqrt{n}$, as $n \to \infty$, are given by

$$
\begin{aligned}
\mu_r &= r(r-2)(r-4)\cdots 2 && \text{if } r \text{ is even,} \\
\mu_r &= r(r-2)(r-4)\cdots 1 \cdot \sqrt{\pi/2} && \text{if } r \text{ is odd.}
\end{aligned}
$$

For the first four moments this gives us $\mu_1 = \sqrt{\pi/2}$, $\mu_2 = 2$, $\mu_3 = 3\sqrt{\pi/2}$, $\mu_4 = 8$. Therefore, to support our claim in Conjecture 1 we must provide evidence that the moments of $m_p/\sqrt{p}$ are converging, as $x \to \infty$, to the moments μ_r above. In our computations we use the following expression for the r-th moments of $m_p/\sqrt{p}$:

$$
M_r = \frac{1}{|\{p \le x\}|} \sum_{p \le x} \left(\frac{m_p}{\sqrt{p}}\right)^r.
$$

3. Iterates of f congruent to zero modulo p

In this section we consider the quantity $|\mathcal{Q}_{f,\alpha}(x)|(\log x)/\sqrt{x}$, as defined in Section 1. Assuming that the probability that $0 \in \mathcal{O}_f^p(\alpha)$ is m_p/p, and that M_1 will converge to $\sqrt{\pi/2}$, we define

$$
G(x) = \frac{\log x}{\sqrt{x}} \sum_{p \le x} \frac{\sqrt{\pi/2}}{\sqrt{p}},
$$

and make a guess that

$$
\lim_{x\to\infty} |\mathcal{Q}_{f,\alpha}(x)| \frac{\log x}{\sqrt{x}} = \lim_{x\to\infty} G(x). \tag{1}
$$

If we let $\pi(x) = \sum_{k \le x} a(k)$, where $a(k) = 1$ if k is prime and 0 otherwise, and

define $f(x) = 1/\sqrt{x}$, then Stieltjes integration by parts gives

$$\sum_{p\le x} \frac{1}{\sqrt{p}} = \frac{\pi(x)}{\sqrt{x}} - \frac{1}{\sqrt{2}} + \frac{1}{2}\int_2^x \frac{\pi(t)}{t^{3/2}}\,\mathrm{d}t,$$

which implies

$$\lim_{x\to\infty} \frac{\log x}{\sqrt{x}} \sum_{p\le x} \frac{1}{\sqrt{p}} = 1 + \lim_{x\to\infty} \frac{\log x}{2\sqrt{x}} \int_2^x \frac{\pi(t)}{t^{3/2}}\,\mathrm{d}t. \tag{2}$$

Now, for $x \ge 55$, we can bound $\pi(x)$ by the inequalities

$$\frac{x}{\log x + 2} < \pi(x) < \frac{x}{\log x - 4};$$

see [Rosser 1941]. Hence, if we shift the lower limit of integration in (2) to 55, changing the value of the integral only by an additive constant which will vanish in the limit, we can write

$$\begin{aligned}\lim_{x\to\infty} \frac{\log x}{2\sqrt{x}} \int_{55}^x \frac{1}{\sqrt{t}(\log t + 2)}\,\mathrm{d}t &\le \lim_{x\to\infty} \frac{\log x}{2\sqrt{x}} \int_{55}^x \frac{\pi(t)}{t^{3/2}}\,\mathrm{d}t \\ &\le \lim_{x\to\infty} \frac{\log x}{2\sqrt{x}} \int_{55}^x \frac{1}{\sqrt{t}(\log t - 4)}\,\mathrm{d}t. \end{aligned} \tag{3}$$

Consider the limit on the left. The integral diverges, since the integrand exceeds $1/t$ everywhere — indeed, $\sqrt{t}(\log t + 2) < t$ for $t \ge 55$. Hence the limit has the form ∞/∞, where the denominator comes from expressing the quotient before the integral as the inverse of $2\sqrt{x}/\log x$. It follows that the limit on the left equals

$$\lim_{x\to\infty} \frac{1}{\sqrt{x}(\log x + 2)} \cdot \frac{\sqrt{x}\log^2 x}{\log x - 2} = \lim_{x\to\infty} \frac{\log^2 x}{\log^2 x - 4} = 1.$$

An analogous reasoning shows that the rightmost limit in (3) is equal to

$$\lim_{x\to\infty} \frac{1}{\sqrt{x}(\log x - 4)} \cdot \frac{\sqrt{x}\log^2 x}{\log x - 2} = \lim_{x\to\infty} \frac{\log^2 x}{\log^2 x - 6\log x + 8} = 1.$$

Therefore so is the limit in the middle. In other words, $\lim_{x\to\infty} G(x) = 2\sqrt{\pi/2} = \sqrt{2\pi}$, and our guess (1) becomes

$$\lim_{x\to\infty} |\mathcal{Q}_{f,\alpha}(x)| \frac{\log x}{\sqrt{x}} = \sqrt{2\pi}.$$

As we test our hypothesis, it should be kept in mind that $\lim_{x\to\infty} G(x)$ converges very slowly. Since the values of x for which $|\mathcal{Q}_{f,\alpha}(x)|(\log x)/\sqrt{x}$ can actually be computed (in a reasonable amount of time) are relatively small, the largest being 2^{27}, we compare our computations to $G(x)$, rather than the limit $\sqrt{2\pi}$.

4. Some special cases

In this paper we consider polynomials of the form $f(z) = z^2 + c$, with $z, c \in \mathbb{Z}$, and initial argument values $\alpha \in \mathbb{Z}$. But for certain f, α pairs we find that we end up with a finite (over $\mathbb{Z}$) orbit, a condition which is clearly incompatible with our hypotheses outlined in Sections 2 and 3, since m_p will have a fixed bound for all primes p. In this section we classify these exceptional pairs f, α.

Proposition 1. *Let $\mathcal{O}_f(\alpha) = \{f^n(\alpha) : n = 0, 1, 2, 3, \dots\}$ be the orbit of α under f, where $f(z) = z^2 + c$, $c \in \mathbb{Z}$, and $\alpha \in \mathbb{Z}$. Then $\mathcal{O}_f(\alpha)$ is finite if and only if one of the following hold*:

$$\begin{aligned} &\text{(i)} \quad \alpha = \pm\tfrac{1}{2}(1 \pm \sqrt{1-4c}), \\ &\text{(ii)} \quad \alpha = \pm\tfrac{1}{2}(1 \pm \sqrt{-3-4c}), \\ &\text{(iii)} \quad \alpha \in \{0, 1, -1\} \quad \textit{and} \quad c \in \{0, -1, -2\}. \end{aligned}$$

Proof. First we prove the converse, which is easier. Assumption (i) gives us the solutions to $\alpha^2 \pm \alpha + c = 0$, and this equation implies that $\alpha^2 + c = \pm\alpha$, which implies that the orbit is finite. Assumption (ii) gives the solutions to $\alpha^2 \pm \alpha + c + 1 = 0$, and this equation implies that $\alpha^2 + c = \pm\alpha - 1$. With one more iteration we get

$$(\alpha^2 + c)^2 + c = (\pm\alpha - 1)^2 + c = \alpha^2 \mp 2\alpha + 1 - \alpha^2 \pm \alpha - 1 = \mp 2\alpha \pm \alpha = \pm\alpha,$$

which again implies that the orbit is finite. As for (iii), testing all possible α, c combinations will quickly convince the reader that the orbits are finite in all cases.

Now suppose that $\mathcal{O}_f(\alpha)$ is finite. First we make some simplifications. Since the orbits of α and $-\alpha$ will be identical except for the sign of the first element $f^0 = \alpha$, we may consider only nonnegative values of α. Also, since it is obvious that $c \in \{0, -1\}$ will have infinite orbit for $\alpha \geq 2$, and that $c \geq 1$ will have infinite orbit for all α, we consider only $c \leq -2$. We claim that $\mathcal{O}_f(\alpha)$ finite implies $\sqrt{-c} - 1 < \alpha < \sqrt{-c} + 1$. If this were not true, then we would have either $\alpha = \lceil\sqrt{-c}\rceil + b$ or $\alpha = \lfloor\sqrt{-c}\rfloor - b$ for some $b \in \mathbb{N}$, giving us

$$\alpha = \lceil\sqrt{-c}\rceil + b \Longrightarrow \alpha^2 + c = (\lceil\sqrt{-c}\rceil)^2 + 2b\lceil\sqrt{-c}\rceil + b^2 + c > \lceil\sqrt{-c}\rceil + b,$$
$$\alpha = \lfloor\sqrt{-c}\rfloor - b \Longrightarrow \alpha^2 + c = (\lfloor\sqrt{-c}\rfloor)^2 - 2b\lfloor\sqrt{-c}\rfloor + b^2 + c < -2b\lfloor\sqrt{-c}\rfloor + c.$$

The first of these immediately implies that the iterates of f are unbounded since they are strictly increasing. In the second case iterating once more gives us

$$(\alpha^2 + c)^2 + c > 4b^2\lfloor\sqrt{-c}\rfloor^2 - 4bc\lfloor\sqrt{-c}\rfloor + c^2 > \lfloor\sqrt{-c}\rfloor + b,$$

where the inequality reverses since $\alpha^2 + c < -2b\lfloor\sqrt{-c}\rfloor + c < 0$, and the second inequality follows since $c \leq -2$. Again we can conclude that the iterates of f are unbounded, and so we have shown that $\mathcal{O}_f(\alpha)$ finite implies $\sqrt{-c} - 1 < \alpha < \sqrt{-c} + 1$.

For any c, there are at most two integers that satisfy the preceding inequality, $\lfloor\sqrt{-c}\rfloor$ and $\lceil\sqrt{-c}\rceil$, so any member of $\mathbb{O}_f(\alpha)$ must be one of $\pm\lfloor\sqrt{-c}\rfloor, \pm\lceil\sqrt{-c}\rceil$, since otherwise the iterates of f will be unbounded. Since we know $\alpha \in \mathbb{O}_f(\alpha)$, the condition above implies that $\mathbb{O}_f(\alpha) \subset \{\alpha, -\alpha, \alpha-1, -\alpha-1\}$ or $\mathbb{O}_f(\alpha) \subset \{\alpha, -\alpha, \alpha+1, -\alpha+1\}$. However, we can rule out the latter case since

$$\begin{aligned} & \alpha^2 + c = \pm\alpha + 1 \\ \Longrightarrow \quad & (\alpha^2+c)^2 + c = \pm 3\alpha + 2 \\ \Longrightarrow \quad & ((\alpha^2+c)^2+c)^2 + c = 7\alpha^2 \pm 13\alpha + 5 > 2\alpha + 5 > \pm\alpha + 1 > \pm\alpha, \end{aligned}$$

where the first inequality follows since in this case $c \le -2 \Longrightarrow \alpha \ge 2$. Therefore the iterates are unbounded in this case, and we are left with the following:

$$\begin{aligned} & \alpha^2 + c = \pm\alpha && \text{or} \quad \alpha^2 + c = \pm\alpha - 1 \\ \Longrightarrow \quad & \alpha^2 \pm \alpha + c = 0 && \text{or} \quad \alpha^2 \pm \alpha + c + 1 = 0 \\ \Longrightarrow \quad & \alpha = \pm\tfrac{1}{2}(1 \pm \sqrt{1-4c}) && \text{or} \quad \alpha = \pm\tfrac{1}{2}(1 \pm \sqrt{-3-4c}). \end{aligned}$$ □

This proposition is the basis for the second condition necessary for Conjectures 1 and 2; we now turn to the first condition, that $c \notin \{0, -2\}$. These two cases behave strikingly differently from the others studied, because they come from homomorphisms of the multiplicative group. The map $z^2 - 2$ is a Chebyshev polynomial, and so is connected to z^2 via a homomorphism from $\mathbb{C}^*$ to $\mathbb{C}^*/\{z \sim z^{-1}\}$ [Silverman 2007, pp. 29–30]. On a finite field $\mathbb{F}_p$, this means that the behavior of $z^2 - 2$ will be very similar to that of z^2. For $c = 0$ it is clear that $|\mathfrak{Q}_{f,\alpha}(x)|$ will not grow as expected, since we'll have $p \in \mathfrak{Q}_{f,\alpha}(x)$ if and only if p divides α. On the other hand, the length m_p of the orbit modulo p will grow much faster than we expect. Vasiga and Shallit [2004] studied these two cases in some depth, showing that, for a given prime p, if $(p-1)/2$ is prime and 2 is a primitive root modulo $(p-1)/2$, then $\sum_{0 \le \alpha < p} m_p$ is at least on the order of p^2. Heuristics by Hardy and Littlewood [1923], along with Artin's conjecture, suggest that the number of primes less than x that satisfy this property is on the order of $x/(\log x)^2$, and thus the density of these primes is on the order of $1/\log x$. If we sum p^2, for $p \le x$, and multiply by $1/\log x$, we get something on the order of $x^3/(\log x)^2$, and dividing this by the sum $\sum_{p \le x} \sum_{0 \le \alpha < p} 1 \sim x^2/\log x$ gives us an average orbit length on the order of $x/\log x$. Note that this estimate only takes into account primes with the aforementioned property, and assumes that all other primes have orbit length 0, so we should expect this to be a low estimate. Indeed, the limited experimentation we did on this question suggests that the average orbit length is closer to $x/(\log x)^{3/4}$.

Finally, Conjecture 2 requires an additional condition, that $\alpha^2 \ne -c$. If we disregard this condition we will have cases where $0 \in \mathfrak{Q}_{f,\alpha}(x)$ for all p, which

clearly conflicts with our claim. To see that the $f^0 = \alpha$ is the only iterate whose square can be equal to $-c$, suppose that the contrary is true, i.e., that we have $(f^l)^2 = -c$ for some $l \in \mathbb{Z}$; then, letting $f^k = f^{l-1}$, we have

$$\begin{aligned} \left((f^k)^2 + c\right)^2 + c &= 0, \\ (f^k)^4 + 2c(f^k)^2 + c^2 + c &= 0, \\ c^2 + (2(f^k)^2 + 1)c + (f^k)^4 &= 0. \end{aligned}$$

Therefore the quadratic formula gives us

$$c = \frac{-2(f^k)^2 - 1 \pm \sqrt{(2f^k)^2 + 1}}{2},$$

which is not an integer unless $f^k = 0$, in which case $(f^l)^2 = c^2 = -c \Longrightarrow c = -1$. It is easy to see that this implies $\alpha \in \{0, 1\}$, and this case has already been excluded by Proposition 1(iii).

5. Results

First we consider the first four moments of $m_p/\sqrt{p}$, as discussed in Section 2, for $f(z) = z^2 + c$, where $c = \pm 1, +2, \pm 3$, and initial arguments $\alpha = 1, 2, \ldots, 9$. Of these we can exclude $\alpha = 1, 2$ when $f(z) = z^2 - 3$, and $\alpha = 1$ when $f(z) = z^2 - 1$, because these (f, α) combinations have finite orbits, as discussed above. For the other 42 combinations, we find that our experimental results support our hypotheses very well. For the first moment we expected the limit to be $\sqrt{\pi/2} = 1.25331413\ldots$, and for all (f, α) tested, M_1 was between 1.25138 and 1.25351 for $x = 2^{25}$, with an average value of 1.25279. Table 1 gives these figures along with the standard

	mean	stand dev	min	max
M_1	1.252795789	0.000518158	1.251387582	1.253505370
$\lvert\sqrt{\pi/2} - M_1\rvert$	0.000544827	0.000490241	0.000000052	0.001926555
M_2	1.998325027	0.001690776	1.993860194	2.000539507
$\lvert 2 - M_2\rvert$	0.001810403	0.001544894	0.000034079	0.006139806
M_3	3.755044605	0.004998323	3.742341997	3.762285912
$\lvert 3\sqrt{\pi/2} - M_3\rvert$	0.005419269	0.004427558	0.000121838	0.017600415
M_4	7.985456401	0.014915109	7.948531018	8.008817811
$\lvert 8 - M_4\rvert$	0.016278430	0.012999594	0.000149649	0.051468982

Table 1. Moments of $m_p/\sqrt{p}$ for $x = 2^{25}$ and distance from predicted limit. For comparison, $\sqrt{\pi/2} \sim 1.25331413731550$.

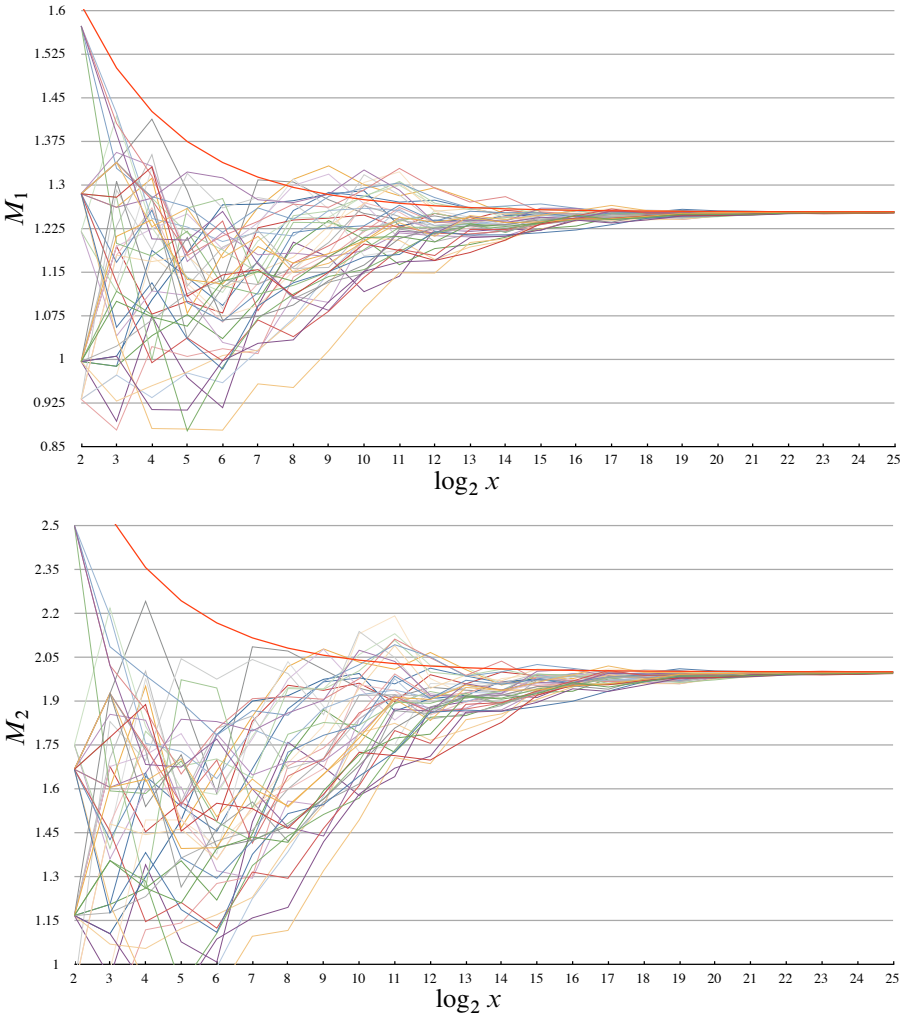

Figure 1. The first and second moments, M_1 and M_2, of $X/\sqrt{n}$ (thicker red lines) and $m_p/\sqrt{p}$ (thin lines) for all (f, α) tested.

deviation of the set of results for each moment. It also shows the mean, standard deviation, minimum, and maximum of the set

$$\{|\sqrt{\pi/2} - M_1| : x = 2^{25}, \text{for } (f, \alpha) \text{ tested}\},$$

and similarly for the second, third and fourth moments. Our complete results are depicted graphically in Figures 1 and 2, for the first, second, third, and fourth moments. In each of these graphs the heavier red curve is the respective moment of $X_n/\sqrt{n}$, for $n = x$. Notice that the y-axes of these graphs are not scaled equally with respect to each other (they are stretched by a factor of two for each subsequent

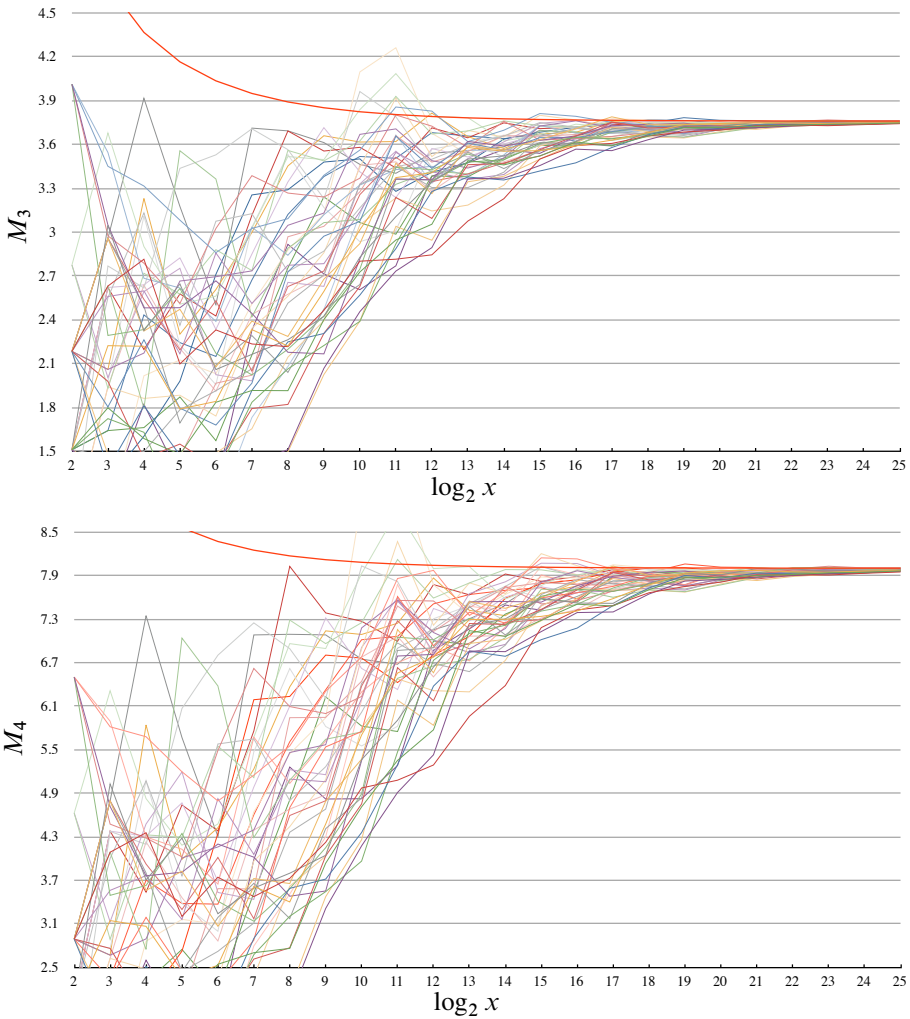

Figure 2. The third and fourth moments, M_3 and M_4, of $X/\sqrt{n}$ (thicker red lines) and $m_p/\sqrt{p}$ (thin lines) for all (f, α) tested.

moment graph), so if we're interested in comparing how quickly two of the moments converge, Table 1 will be more helpful.

The apparent common limit of the moments of $m_p/\sqrt{p}$ and $X/\sqrt{n}$ suggests that the limiting distributions of $m_p/\sqrt{p}$, as $x \to \infty$, and the random variable $X_n/\sqrt{n}$, as $n \to \infty$, are the same. For the variable $X/\sqrt{n}$ we showed in Section 2 that, as $n \to \infty$, the distribution $P(X_n/\sqrt{n} < t)$ converges to the function $F(t) = 1 - e^{-t^2/2}$. As the histogram in Figure 3 shows, the density function $f(t) = F'(t)$ approximates quite well the distribution of $m_p/\sqrt{p}$ for $x = 10^8$, $f(z) = z^2 + 1$, $\alpha = 3$. These results give considerable support to our first conjecture, stated in Section 1.

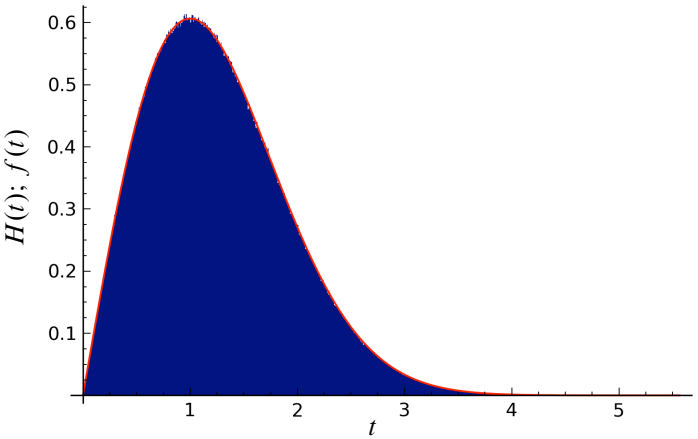

Figure 3. Histogram, $H(t)$, of the distribution of $m_p/\sqrt{p}$ (blue) for $x = 10^8$, $f(z) = z^2 + 1$, $\alpha = 3$, superimposed on the graph of $f(t) = te^{-t^2/2}$ (red). Here

$$H(t : wk \le t < w(k+1)) = \frac{|\{p \le 10^8 : wk \le m_p/\sqrt{p} < w(k+1)\}|}{w \cdot |\{p \le 10^8\}|},$$

for $k \in \mathbb{N}$. Each bar of the histogram has width $w \approx 5.6/800$.

To test the hypothesis discussed in Section 3, we compute $|\mathfrak{Q}_{f,\alpha}(x)|(\log x)/\sqrt{x}$ for (f, α) as described above and $x \in \{2, 2^2, \ldots, 2^{27}\}$. Table 2 shows that, although our results are still fairly widely dispersed at $x = 2^{27}$, the average of the results for this x value is very close to $G(x)$, and the standard deviation is decreasing in general as x increases, as is the error of the mean from $G(x)$. As we mentioned earlier, $\lim_{x\to\infty} G(x)$ converges very slowly, and, as the table shows, even for x as large as 2^{27} we still have $|G(x) - \sqrt{2\pi}| \sim 0.36$, so we are not too surprised to see such a wide range in our results for this x value. That is, intuitively, it seems we should not expect our results to be very tightly grouped until we are close to the limiting value, $\sqrt{2\pi}$. Figure 4 gives a graphical representation of all (f, α) tested, for x from 4 to 2^{27}. On this graph the red and blue lines are $G(x)$ and the mean from Table 2, respectively. From this data, it seems reasonable to suppose that $|\mathfrak{Q}_{f,\alpha}(x)|(\log x)/\sqrt{x}$ will eventually converge to $\sqrt{2\pi}$, independent of f, α, and so we make our second conjecture as stated in Section 1.

Joseph Silverman [2008] carried out computations that lead to a conjecture (in a more general setting) that under certain restrictions the set $\{p : m_p \le p^{1/2-\epsilon}\}$ will have density 0 for $\epsilon > 0$. This conjecture agrees with our own results, and in fact, if Conjecture 1 were proven, a less general version of Silverman's conjecture would readily follow. Computations of a similar nature to ours were also carried out in [Benedetto et al. 2013], with results that are compatible with our own.

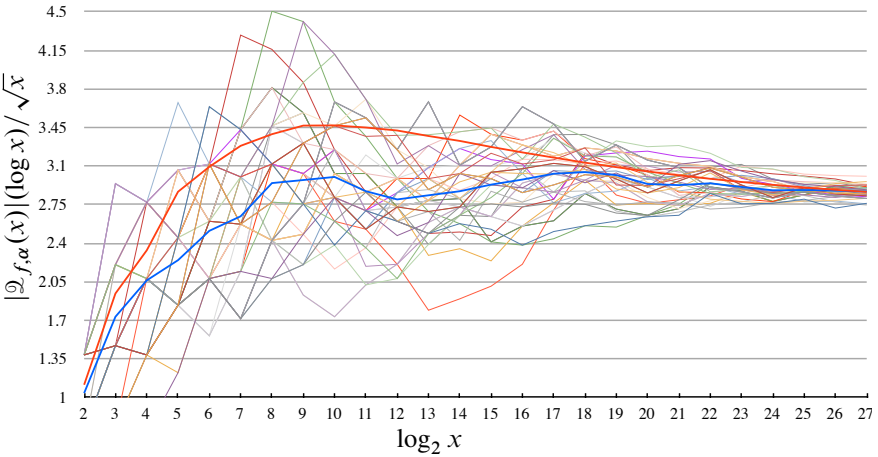

Figure 4. Graphs of $\mathcal{Q}_{f,\alpha}(x)(\log x)/\sqrt{x}$ for all 42 (f,α) combinations tested (thinner lines), the mean of these graphs (thick blue line), and our guess $G(x)$ (thick red line).

Acknowledgements

We would like to thank our advisor Gautam Chinta, whose guidance and assistance throughout the research process were instrumental to the project's success, and whose invaluable feedback during the drafting of this paper improved its quality considerably. In addition, we thank Prof. Hutz for calling our attention to the paper [Benedetto et al. 2013], where the authors carry out similar computations and obtain results compatible with our computations, and we thank Prof. Silverman for directing us to the publication in which his article [Silverman 2008] appeared.

References

[Bach 1991] E. Bach, "Toward a theory of Pollard's rho method", *Inform. and Comput.* **90**:2 (1991), 139–155. MR 92a:11151 Zbl 0716.11065

[Benedetto et al. 2013] R. L. Benedetto, D. Ghioca, B. Hutz, P. Kurlberg, T. Scanlon, and T. J. Tucker, "Periods of rational maps modulo primes", *Math. Ann.* **355**:2 (2013), 637–660. MR 3010142 Zbl 06133902

[Brent 1980] R. P. Brent, "An improved Monte Carlo factorization algorithm", *BIT* **20**:2 (1980), 176–184. MR 82a:10007 Zbl 0439.65001

[Hardy and Littlewood 1923] G. H. Hardy and J. E. Littlewood, "Some problems of "Partitio numerorum", III: On the expression of a number as a sum of primes", *Acta Math.* **44** (1923), 1–70. MR 1555183 Zbl 48.0143.04

[Harris 1960] B. Harris, "Probability distributions related to random mappings", *Ann. Math. Statist.* **31** (1960), 1045–1062. MR 22 #9993 Zbl 0158.34905

[Pollard 1975] J. M. Pollard, "A Monte Carlo method for factorization", *Nordisk Tidskr. Informationsbehandling* (*BIT*) **15**:3 (1975), 331–334. MR 52 #13611 Zbl 0312.10006

		$\lvert\mathcal{Q}_{f,\alpha}(x)\rvert(\log x)/\sqrt{x}$ for all (f,α) tested				
x	$G(x)^a$	mean	stand dev	min	max	$\lvert G(x)-\text{mean}\rvert$
2^{10}	3.46925	3.00157	0.51687	1.73287	4.11556	0.46767
2^{11}	3.45003	2.87221	0.46933	2.02178	3.70660	0.57781
2^{12}	3.42304	2.79734	0.34646	2.07944	3.37909	0.62570
2^{13}	3.37313	2.83502	0.37519	1.79203	3.68363	0.53811
2^{14}	3.32854	2.87187	0.30915	1.89532	3.56321	0.45667
2^{15}	3.27415	2.93202	0.35071	2.01030	3.44622	0.34213
2^{16}	3.22425	2.97785	0.33397	2.20941	3.63902	0.24640
2^{17}	3.17737	3.03080	0.28861	2.44107	3.48260	0.14657
2^{18}	3.13140	3.04548	0.18849	2.55869	3.38722	0.08593
2^{19}	3.09045	3.01797	0.21004	2.54637	3.32847	0.07248
2^{20}	3.05137	2.93775	0.17602	2.63992	3.27620	0.11362
2^{21}	3.01654	2.92737	0.15786	2.65359	3.28683	0.08917
2^{22}	2.98475	2.94397	0.12988	2.71031	3.21664	0.04077
2^{23}	2.95589	2.91037	0.09402	2.72467	3.11548	0.04552
2^{24}	2.92986	2.88060	0.08428	2.76176	3.07449	0.04925
2^{25}	2.90646	2.88289	0.05445	2.77612	3.02741	0.02357
2^{26}	2.88509	2.87751	0.05869	2.71911	3.01390	0.00759
2^{27}	2.86578	2.86821	0.05418	2.74621	3.00790	0.00243

Table 2. A comparison of our experimental results to our guess, $G(x) = \frac{\log x}{\sqrt{x}} \sum_{p \le x} \frac{\sqrt{\pi/2}}{\sqrt{p}}$.

[Rosser 1941] B. Rosser, "Explicit bounds for some functions of prime numbers", *Amer. J. Math.* **63** (1941), 211–232. MR 2,150e Zbl 0024.25004

[Silverman 2007] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics **241**, Springer, New York, 2007. MR 2008c:11002 Zbl 1130.37001

[Silverman 2008] J. H. Silverman, "Variation of periods modulo p in arithmetic dynamics", *New York J. Math.* **14** (2008), 601–616. MR 2009h:11100 Zbl 1153.11028

[Vasiga and Shallit 2004] T. Vasiga and J. Shallit, "On the iteration of certain quadratic maps over GF(p)", *Discrete Math.* **277**:1-3 (2004), 219–240. MR 2004k:05104 Zbl 1045.11086

Received: 2012-03-02 Revised: 2012-04-25 Accepted: 2012-05-10

william.worden@temple.edu *Temple University, Wachman Hall Rm. 517, 1805 N. Broad St., Philadelphia, PA 19122, United States*

involve

msp.org/involve

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2013 is US $105/year for the electronic version, and $145/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve
2013 vol. 6 no. 1