

involve

a journal of mathematics

A probabilistic heuristic for counting components of
functional graphs of polynomials
over finite fields

Elisa Bellah, Derek Garton, Erin Tannenbaum and Noah Walton



A probabilistic heuristic for counting components of functional graphs of polynomials over finite fields

Elisa Bellah, Derek Garton, Erin Tannenbaum and Noah Walton

(Communicated by Michael E. Zieve)

Flynn and Garton (2014) bounded the average number of components of the functional graphs of polynomials of fixed degree over a finite field. When the fixed degree was large (relative to the size of the finite field), their lower bound matched Kruskal's asymptotic for random functional graphs. However, when the fixed degree was small, they were unable to match Kruskal's bound, since they could not (Lagrange) interpolate cycles in functional graphs of length greater than the fixed degree. In our work, we introduce a heuristic for approximating the average number of such cycles of any length. This heuristic is, roughly, that for sets of edges in a functional graph, the quality of being a cycle and the quality of being interpolable are "uncorrelated enough". We prove that this heuristic implies that the average number of components of the functional graphs of polynomials of fixed degree over a finite field is within a bounded constant of Kruskal's bound. We also analyze some numerical data comparing implications of this heuristic to some component counts of functional graphs of polynomials over finite fields.

1. Introduction

A (*discrete*) *dynamical system* is a pair (S, f) consisting of a set S and a map $f : S \rightarrow S$. Given such a system, an element $s \in S$ is a *periodic point* of the system if there exists some $k \in \mathbb{Z}^{>0}$ such that $(f \circ \dots \circ f)(s) = s$, where f appears k times; the smallest $k \in \mathbb{Z}^{>0}$ with this property is called the *period* of s . The *functional graph* of such a system, which we denote by $\Gamma(S, f)$, is the directed graph whose vertex set is S and whose edges are given by the relation $s \rightarrow t$ if and only if $f(s) = t$. A *component* of such a graph is a component of the underlying undirected graph. For any $n \in \mathbb{Z}^{>0}$, let $\mathcal{K}(n)$ denote the average number of components of a random

MSC2010: primary 37P05; secondary 05C80, 37P25.

Keywords: arithmetic dynamics, functional graphs, finite fields, polynomials, rational maps.

functional graph on a set of size n ; that is, choose any set S with $|S| = n$ and let

$$\mathcal{K}(n) = n^{-n} \sum_{f: S \rightarrow S} |\{\text{components of } \Gamma(S, f)\}|.$$

Kruskal [1954] proved that

$$\mathcal{K}(n) = \frac{1}{2} \log n + \frac{1}{2}(\log 2 + C) + o(1),$$

where $C = .5772\dots$ is Euler's constant.

Recently, researchers have begun studying the analogous situation for polynomials (and rational maps) over finite fields. More precisely, if q is a prime power and $f \in \mathbb{F}_q[x]$, define $\Gamma(q, f) = \Gamma(\mathbb{F}_q, f)$ (if there is no ambiguity, we will frequently write Γ_f for $\Gamma(q, f)$). Then we can ask the question: for $d \in \mathbb{Z}^{>0}$, what is the average number of components of Γ_f , for f ranging over all polynomials over \mathbb{F}_q of a fixed degree? In particular, if we define

$$\mathcal{P}(q, d) := \frac{1}{|\{f \in \mathbb{F}_q[x] \mid \deg f = d\}|} \cdot \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} |\{\text{components of } \Gamma_f\}|,$$

then we can ask:

Question 1.1. For a prime power q and $d \in \mathbb{Z}^{>0}$, how does $\mathcal{P}(q, d)$ compare to $\mathcal{K}(q)$?

In this paper, we recast these questions in probabilistic terms. Specifically, in Section 2, we define two families of random variables whose interaction determines the answer to Question 1.1. Briefly, both families of random variables have sample space a certain collection of subsets of $\mathbb{F}_q \times \mathbb{F}_q$ — one random variable determines if a collection is a cycle, and the other returns how many polynomials of a given degree pass through every point in a collection.

Our main result, Theorem 3.3, states that if these two families of random variables satisfy a certain “noncorrelation hypothesis”, then

$$\mathcal{P}(q, d) = \mathcal{K}(q) + O(1);$$

see Heuristic 3.1 for an exact formulation of this hypothesis. In Section 2 we define and study these random variables; in particular, we compute their expected values. Next, in Section 3 we use the results from Section 2 to prove the aforementioned Theorem 3.3. Then, in Section 4, we provide numerical evidence in support of Heuristic 3.1. Finally, these results carry over easily to the analogous question for rational functions; these results make up Section 5.

Previous work of Flynn and the second author (see [Flynn and Garton 2014]) provided a partial answer to the question under discussion. In particular, they proved that if $d \geq \sqrt{q}$, then the average number of components of functional graphs of

polynomials (or rational maps) of degree d over \mathbb{F}_q is bounded below by [Flynn and Garton 2014, Corollary 2.3 and Theorem 3.6]

$$\frac{1}{2} \log q - 4.$$

To describe their method, which is the starting point for this paper, we require a definition and an observation. If a map f has a periodic point s of period k , with orbit

$$s = s_1 \xrightarrow{f} \dots \xrightarrow{f} s_k \xrightarrow{f} s_1,$$

then we refer to its orbit as a *cycle* (cycles of length k are called *k-cycles*); see [Vasiga and Shallit 2004] for more exposition and illustrations of the cycle structure of functional graphs. This definition is especially useful since it allows for the following observation.

Observation 1.2. Components of Γ_f are in one-to-one correspondence with the cycles of f .

To obtain their results, Flynn and the second author used Lagrange interpolation to interpolate all the cycles of length smaller than the degree of the maps in question. Since they could not interpolate longer cycles,

- they obtained only a lower bound for $\mathcal{P}(q, d)$, and
- their result required that d be at least \sqrt{q} .

See Remark 2.5 for a discussion on the relationship between the results of this paper and the results of [Flynn and Garton 2014]; for example, they proved that the random variables mentioned above are indeed uncorrelated in certain cases.

The cycle structure of functional graphs of polynomials over finite fields has been studied extensively in certain cases. Vasiga and Shallit [2004] studied the cycle structure of Γ_f for the cases $f = x^2$ and $f = x^2 - 2$, as did Rogers [1996] for $f = x^2$. For any $m \in \mathbb{Z}^{>0}$, the squaring function is also defined over $\mathbb{Z}/m\mathbb{Z}$; Carlip and Mincheva [2008] addressed this situation for certain m . Similarly, Chou and Shparlinski [2004] studied the cycle structure of repeated exponentiation over finite fields of prime size. In the context of Pollard’s rho algorithm for factoring integers (see [Pollard 1975]), researchers have provided copious data and heuristic arguments supporting the claim that quadratic polynomials produce as many “collisions” as random functions, but very little has been proven (see [Pollard 1975; Bach 1991]). For many other aspects of functional graphs besides their cycle structure, see [Flajolet and Odlyzko 1990] for a study of about twenty characteristic parameters of random mappings in various settings.

More recently, Burnette and Schmutz [2017] used the probabilistic point of view to study a similar question to the one we address here. If f is a polynomial (or rational function) over \mathbb{F}_q , define the *ultimate period of f* to be the least common

multiple of the cycle lengths of Γ_f . They found a lower bound for the average ultimate period of polynomials (and rational functions) of fixed degree, whenever the degree of the maps in question, and the size of the finite field, were large enough.

2. Two families of random variables

In this section, we define two families of random variables and compute their expected values. The interaction of these random variables determines the answer to [Question 1.1](#); see [Remark 2.4](#) and the remarks that follow for details about this interaction. For the remainder of the section, fix a prime power q and positive integer d . Now, for any set S and $C \subseteq S \times S$, we say that C is *consistent* if and only if it has the following property: if $(a, b), (a, c) \in C$, then $b = c$. Next, for any $k \in \mathbb{Z}^{\geq 0}$, define

$$\mathfrak{C}(q, k) = \{C \subseteq \mathbb{F}_q \times \mathbb{F}_q \mid C \text{ is consistent and } |C| = k\}.$$

Any element of $C \in \mathfrak{C}(q, k)$ defines a directed graph with vertex set \mathbb{F}_q and edge set $\{s \rightarrow t \mid (s, t) \in C\}$; let $X_{q,k} : \mathfrak{C}(q, k) \rightarrow \{0, 1\}$ be the binary random variable that detects whether or not an element of $\mathfrak{C}(q, k)$ defines a graph that happens to be a k -cycle. If $f \in \mathbb{F}_q[x]$ and $C \in \mathfrak{C}(q, k)$, we say that f *satisfies* C if $f(a) = b$ for all $(a, b) \in C$. Next, we let

$$Y_{q,d,k} : \mathfrak{C}(q, k) \rightarrow \mathbb{Z}^{\geq 0}$$

be the random variable defined by

$$Y_{q,d,k}(C) = |\{f \in \mathbb{F}_q[x] \mid \deg f = d \text{ and } f \text{ satisfies } C\}|.$$

Before computing the expected values of $X_{q,k}$ and $Y_{q,d,k}$, we first mention the size of their sample space.

Remark 2.1. If $k \in \mathbb{Z}^{>0}$, then

$$|\mathfrak{C}(q, k)| = q^k \binom{q}{k}.$$

Proof. Since the elements of $\mathfrak{C}(q, k)$ are consistent, there are $\binom{q}{k}$ possible choices for the sets of abscissas for any choice of ordinates. Since the ordinates of elements of $\mathfrak{C}(q, k)$ are unrestricted, we conclude that $|\mathfrak{C}(q, k)| = \binom{q}{k} q^k$. \square

Remark 2.2. If $k \in \{1, \dots, q\}$, then

$$\mathbb{E}[X_{q,k}] = \frac{q(q-1) \cdots (q-(k-1))}{k|\mathfrak{C}(q, k)|} = \frac{(q-1)!}{q^k}.$$

Proof. Since

$$\mathbb{E}[X_{q,k}] = \frac{|\{C \in \mathfrak{C}(q, k) \mid C \text{ is a cycle}\}|}{|\mathfrak{C}(q, k)|},$$

we only need to count the number of elements in $\mathfrak{C}(q, k)$ that are cycles. Since there are

$$\frac{q(q-1)\cdots(q-(k-1))}{k}$$

cycles of length k , we conclude by [Remark 2.1](#). □

Proposition 2.3. *If $k \in \{1, \dots, q\}$, then $\mathbb{E}[Y_{q,d,k}] = q^{d+1-k} - q^d$.*

Proof. Since

$$\begin{aligned} \sum_{C \in \mathfrak{C}(q,k)} Y_{q,d,k}(C) &= \sum_{C \in \mathfrak{C}(q,k)} |\{f \in \mathbb{F}_q[x] \mid \deg f = d \text{ and } f \text{ satisfies } C\}| \\ &= \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} |\{C \in \mathfrak{C}(q, k) \mid C \text{ is satisfied by } f\}| \\ &= \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} \binom{q}{k} = (q^{d+1} - q^d) \binom{q}{k}, \end{aligned}$$

we see by [Remark 2.1](#) that

$$\begin{aligned} \mathbb{E}[Y_{q,d,k}] &= |\mathfrak{C}(q, k)|^{-1} \cdot \sum_{C \in \mathfrak{C}(q,k)} Y_{q,d,k}(C) \\ &= \frac{(q^{d+1} - q^d) \binom{q}{k}}{q^k \binom{q}{k}} = q^{d+1-k} - q^{d-k}. \end{aligned} \quad \square$$

Remark 2.4. If we assume that $X_{q,d}, Y_{q,d,k}$ are uncorrelated for all $k \in \{1, \dots, q\}$, then $\mathcal{K}(q) = \mathcal{P}(q, d)$.

Proof. Note that for any $k \in \{1, \dots, q\}$,

$$\begin{aligned} \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} |\{k\text{-cycles in } \Gamma_f\}| &= \sum_{C \in \mathfrak{C}(q,k)} X_{q,k} Y_{q,d,k}(C) \\ &= |\mathfrak{C}(q, k)| \mathbb{E}[X_{q,k} Y_{q,d,k}] \\ &= |\mathfrak{C}(q, k)| \mathbb{E}[X_{q,k}] \mathbb{E}[Y_{q,d,k}] \quad \text{by assumption.} \end{aligned}$$

Now we can apply [Remarks 2.1](#) and [2.2](#), along with [Proposition 2.3](#), to see that

$$\begin{aligned} \mathcal{P}(q, d) &= \frac{|\mathfrak{C}(q, k)|}{q^{d+1} - q^d} \cdot \sum_{k=1}^q \mathbb{E}[X_{q,k}] \mathbb{E}[Y_{q,d,k}] \\ &= \sum_{k=1}^q \frac{q(q-1)\cdots(q-(k-1))}{kq^k} \\ &= \mathcal{K}(q) \quad \text{by [Kruskal 1954, Equation 16].} \end{aligned} \quad \square$$

Remark 2.5. Unfortunately, we must face up to the fact that the random variables $X_{q,d}, Y_{q,d,k}$ are not uncorrelated for all $k \in \{1, \dots, q\}$. Indeed, if they were, then the computations from [Remark 2.4](#) would show that

$$\sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f=2}} |\{q\text{-cycles in } \Gamma_f\}| = \frac{q!(q-1)}{q^{q-2}}.$$

But, if $q > 3$, then the quantity on the left is an integer, and the quantity on the right is not! In [Section 3](#), we propose a heuristic that is more reasonable than that these two random variables are uncorrelated.

On the other hand, we should note that the variables $X_{q,d}, Y_{q,d,k}$ are indeed uncorrelated whenever $k \in \{1, \dots, d\}$; this is the content of [\[Flynn and Garton 2014, Lemma 2.1\]](#).

3. The heuristic assumption and its implications

As mentioned in [Remark 2.5](#), the variables $X_{q,d}, Y_{q,d,k}$ are not uncorrelated for all $k \in \{1, \dots, q\}$. In this section, we propose a weaker heuristic for these variables, one which nevertheless implies $\mathcal{P}(q, d) = \mathcal{K}(q) + O(1)$.

Heuristic 3.1. For any $k \in \mathbb{Z}^{>0}$ and any $d \in \mathbb{Z}^{\geq 0}$,

$$\mathbb{E}[X_{q,k}Y_{q,d,k}] = \mathbb{E}[X_{q,k}] \mathbb{E}[Y_{q,d,k}] + O(q^{d-2k}).$$

Here, the implied constant depends only on d .

In fact, [Heuristic 3.1](#) implies more than $\mathcal{P}(q, d) = \mathcal{K}(q) + O(1)$; we state the stronger implication here as a conjecture after one more definition. If $k \in \mathbb{Z}^{>0}$ and any $d \in \mathbb{Z}^{\geq 0}$, let

$$\mathcal{P}(q, d, k) := \frac{1}{|\{f \in \mathbb{F}_q[x] \mid \deg f = d\}|} \cdot \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f=d}} |\{k\text{-cycles in } \Gamma_f\}|.$$

Conjecture 3.2. For any $k \in \mathbb{Z}^{>0}$ and any $d \in \mathbb{Z}^{\geq 0}$,

$$\mathcal{P}(q, d, k) = \frac{q(q-1) \cdots (q-(k-1))}{kq^k} + O(1/q),$$

where the implied constant depends only on d . In particular, $\mathcal{P}(q, d) = \mathcal{K}(q) + O(1)$.

Theorem 3.3. *If [Heuristic 3.1](#) is true, then [Conjecture 3.2](#) is true.*

Proof. As in the proof of [Remark 2.4](#), [Heuristic 3.1](#) immediately implies that

$$\sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f=d}} |\{k\text{-cycles in } \Gamma_f\}| = |\mathfrak{C}(q, k)| (\mathbb{E}[X_{q,k}] \mathbb{E}[Y_{q,d,k}] + O(q^{d-2k})).$$

Next, we can apply Remarks 2.1 and 2.2, along with Proposition 2.3, to see that

$$\begin{aligned} \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} |\{k\text{-cycles in } \Gamma_f\}| &= \frac{q(q-1) \cdots (q-(k-1))}{kq^k} (q^{d+1} - q^d) + \binom{q}{k} q^k \cdot O(q^{d-2k}) \\ &= \frac{q(q-1) \cdots (q-(k-1))}{kq^k} (q^{d+1} - q^d) + O(q^d). \end{aligned}$$

To conclude, note that

$$\begin{aligned} \mathcal{P}(q, d, k) &= \frac{1}{q^{d+1} - q^d} \cdot \sum_{\substack{f \in \mathbb{F}_q[x] \\ \deg f = d}} |\{k\text{-cycles in } \Gamma_f\}| \\ &= \frac{q(q-1) \cdots (q-(k-1))}{kq^k} + O(1/q). \end{aligned} \quad \square$$

Remark 3.4. The available numerical data suggests that the implied constants in Heuristic 3.1 could be quite small. For example, the constant for $d = 2$ seems as if it could be as small as 60 (see Section 4 for more details on the available data).

4. Numerical evidence

In constructing numerical evidence for Conjecture 3.2, we computed the number of cycles of every length for all polynomials in $\mathbb{F}_q[x]$

- of degree 2, up to $q = 241$, and
- of degree 3 up to $q = 73$.

For the remainder of the section, we will address only the quadratic case; a similar analysis works for the cubic case.

Of course, if we let $\mathfrak{Q} = \{q \in \mathbb{Z} \mid q \text{ is a prime power, and } 2 \leq q \leq 241\}$, then for any $k \in \{1, \dots, 241\}$, there is certainly a constant — let’s call it C_k — for which

$$\left| \mathcal{P}(q, 2, k) - \frac{q(q-1) \cdots (q-(k-1))}{kq^k} \right| \leq C_k \cdot 1/q \quad \text{for all } q \in \mathfrak{Q}.$$

There are two obvious questions to ask about these constants, which we will address in turn:

- For any particular k , how plausible is it that

$$\left| \mathcal{P}(q, 2, k) - \frac{q(q-1) \cdots (q-(k-1))}{kq^k} \right| \leq C_k \cdot 1/q$$

for all prime powers q ?

- Even if

$$\mathcal{P}(q, 2, k) = \frac{q(q-1) \cdots (q-(k-1))}{kq^k} + O(1/q)$$

for all $k \in \mathbb{Z}^{>0}$, does it seem likely that the implied constants are bounded, as asserted by [Conjecture 3.2](#)?

To answer the former question, we could plot, for various k ,

$$\mathcal{P}(q, 2, k) \quad \text{and} \quad \frac{q(q-1) \cdots (q-(k-1))}{kq^k} \pm C_k \cdot 1/q.$$

But, as these numbers quickly become minuscule, it is convenient to let

$$\widehat{\mathcal{P}}(q, d, k) = |\{f \in \mathbb{F}_q[x] \mid \deg f = d\}| \cdot \mathcal{P}(q, d, k) = (q^{d+1} - q^d) \cdot \mathcal{P}(q, d, k);$$

that is, $\widehat{\mathcal{P}}(q, d, k)$ is the number of k -cycles appearing in functional graphs of polynomials in $\mathbb{F}_q[x]$ of degree d . [Conjecture 3.2](#) predicts that this quantity is about

$$(q^{d+1} - q^d) \cdot \frac{q(q-1) \cdots (q-(k-1))}{kq^k},$$

which we will denote by $\mathcal{G}(q, d, k)$. By the definition of C_k , we know that for all $q \in \mathfrak{Q}$ and $k \in 1, 2, \dots, 241$,

$$|\widehat{\mathcal{P}}(q, 2, k) - \mathcal{G}(q, 2, k)| \leq C_k(q^2 - q).$$

As two examples of the data we have compiled, we include plots of $\widehat{\mathcal{P}}(q, 2, k)$ and $\mathcal{G}(q, 2, k) \pm C_k(q^2 - q)$ for $k = 6, 10$, where $C_6 = 59$ and $C_{10} = 14$; see [Figure 1](#). These graphs are typical for $k \in \{1, \dots, 241\}$.

To address the second question mentioned above, we plot the various values of C_k in the hopes that they appear to be bounded. This graph is shown in [Figure 2](#).

We should point out that the small values of C_k in [Figure 2](#) are a result of the fact that in our data, we simply found no k -cycles at all for all $k > 82$. So from $k = 82$ onward, the graph is simply plotting

$$\frac{241!}{(241-k)! \cdot k \cdot 241^{k-1}}.$$

This begs two questions:

- As cycles of larger length arise for larger values of q , will the size of C_k increase?
- Conversely, if these cycles do not arise promptly, will this increase the size of C_k ?

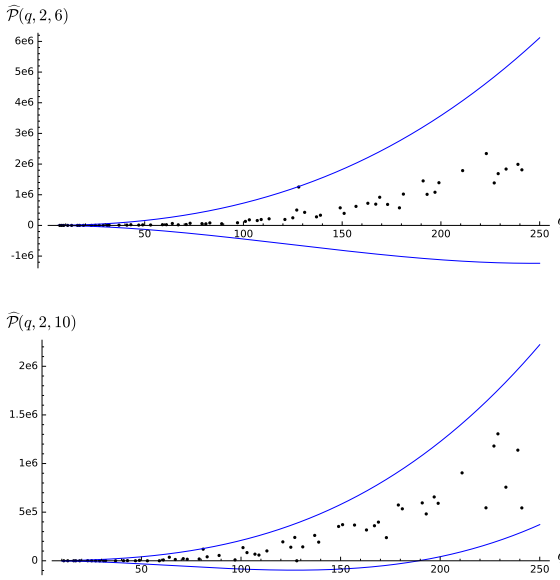


Figure 1. Plots of $\widehat{\mathcal{P}}(q, 2, k)$ and $\mathcal{G}(q, 2, k) \pm C_k(q^2 - q)$ for $k = 6, 10$.
 Top: $C_6 \approx 59.06$; bottom: $C_{10} \approx 14.86$.

Of course, we cannot answer these questions, but note that for the particular value of $k = 82$, the quadratic polynomials we tested yielded exactly 27722 82-cycles (all appearing when $q = 167$), whereas for $k \in \{70, \dots, 81\}$, they yielded exactly zero. That is, this is an example of a cycle of larger length arising without affecting the maximum of the C_k .

As for the second question, the lack of k -cycles will not cause C_k to rise above 60 as long as the first k -cycle appears in a graph for a finite field of size less than $60k$. For example, the smallest q for which 62-cycles appear is $q = 128$ (which is well

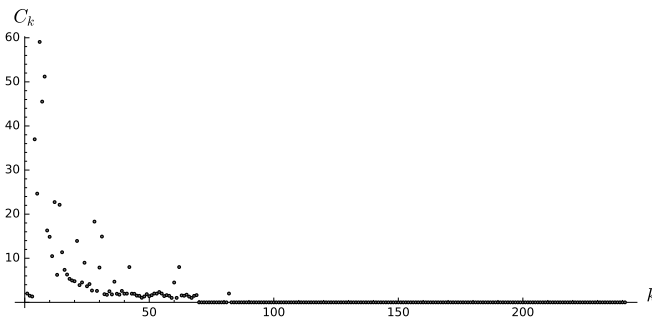


Figure 2. Various values of C_k .

under $60 \cdot 62$). The smallest cycle length that does not appear for $q \in \mathfrak{Q}$ is $k = 43$; if a 43-cycle does not appear by the time $q = 2579$, then C_{43} will rise above 60. It is unfortunately beyond our abilities to determine if a 43-cycle appears by this time.

5. Rational functions

In this section, we briefly mention the results for rational functions, which are analogous to those for polynomials. For any prime power q and $d \in \mathbb{Z}^{\geq 0}$, let

$$\mathcal{R}(q, d) := \frac{1}{|\{f \in \mathbb{P}^1(\mathbb{F}_q)[x] \mid \deg(f) = d\}|} \cdot \sum_{\substack{f \in \mathbb{P}^1(\mathbb{F}_q)[x] \\ \deg(f) = d}} |\{\text{cycles in } \Gamma(\mathbb{P}^1(\mathbb{F}_q), f)\}|.$$

If $k \in \mathbb{Z}^{>0}$, we can define $\mathcal{R}(q, d, k)$ in exactly the same way as $\mathcal{P}(q, d, k)$.

To define our new families of random variables, for any prime power q and $k \in \mathbb{Z}^{>0}$, let

$$\mathfrak{T}(q, k) = \{T \subseteq \mathbb{P}^1(\mathbb{F}_q) \times \mathbb{P}^1(\mathbb{F}_q) \mid T \text{ is consistent and } |T| = k\},$$

and $V_{q,k} : \mathfrak{T}(q, k) \rightarrow \{0, 1\}$ be the binary random variable that detects whether or not an element of $\mathfrak{T}(q, k)$ is a k -cycle. If $d \in \mathbb{Z}^{\geq 0}$, let

$$W_{q,d,k} : \mathfrak{T}(q, k) \rightarrow \mathbb{Z}^{\geq 0}$$

be the random variable defined by

$$W_{q,d,k}(T) = |\{f \in \mathbb{F}_q(x) \mid \deg f = d \text{ and } f \text{ satisfies } T\}|.$$

The rational function analogs of [Remark 2.1](#), [Remark 2.2](#), [Proposition 2.3](#) are proved as above, leading to the following conjecture, which again follows from the heuristic that the random variables $V_{q,k}$, $W_{q,d,k}$ are “uncorrelated enough”.

Conjecture 5.1. For any $k \in \mathbb{Z}^{>0}$ and any $d \in \mathbb{Z}^{\geq 0}$,

$$\mathcal{R}(q, d, k) = \frac{(q+1)q \cdots (q-(k-2))}{k(q+1)^k} + O(1/q),$$

where the implied constant depends only on d . In particular, $\mathcal{R}(q, d) = \mathcal{K}(q+1) + O(1)$.

Heuristic 5.2. If $k \in \{1, \dots, q\}$, and $d \in \mathbb{Z}^{\geq 0}$, then

$$\mathbb{E}[V_{q,k}W_{q,d,k}] = \mathbb{E}[V_{q,k}]\mathbb{E}[W_{q,d,k}] + O(q^{2d-2k}).$$

Here, the implied constant depends only on d .

Theorem 5.3. If [Heuristic 5.2](#) is true, then [Conjecture 5.1](#) is true.

Proof. Similar to the proof of [Theorem 3.3](#). □

Acknowledgments

The authors would like to thank Ian Dinwoodie, Rafe Jones, and Christopher Kramer for their help and advice.

References

- [Bach 1991] E. Bach, “Toward a theory of Pollard’s rho method”, *Inform. and Comput.* **90**:2 (1991), 139–155. [MR](#) [Zbl](#)
- [Burnette and Schmutz 2017] C. Burnette and E. Schmutz, “Periods of iterated rational functions”, *Int. J. Number Theory* **13**:5 (2017), 1301–1315. [MR](#)
- [Carlip and Mincheva 2008] W. Carlip and M. Mincheva, “Symmetry of iteration graphs”, *Czechoslovak Math. J.* **58(133)**:1 (2008), 131–145. [MR](#) [Zbl](#)
- [Chou and Shparlinski 2004] W.-S. Chou and I. E. Shparlinski, “On the cycle structure of repeated exponentiation modulo a prime”, *J. Number Theory* **107**:2 (2004), 345–356. [MR](#) [Zbl](#)
- [Flajolet and Odlyzko 1990] P. Flajolet and A. M. Odlyzko, “Random mapping statistics”, pp. 329–354 in *Advances in cryptology: EUROCRYPT ’89* (Houthalen, 1989), edited by J.-J. Quisquater and J. Vandewalle, Lecture Notes in Comput. Sci. **434**, Springer, Berlin, 1990. [MR](#) [Zbl](#)
- [Flynn and Garton 2014] R. Flynn and D. Garton, “Graph components and dynamics over finite fields”, *Int. J. Number Theory* **10**:3 (2014), 779–792. [MR](#) [Zbl](#)
- [Kruskal 1954] M. D. Kruskal, “The expected number of components under a random mapping function”, *Amer. Math. Monthly* **61** (1954), 392–397. [MR](#) [Zbl](#)
- [Pollard 1975] J. M. Pollard, “A Monte Carlo method for factorization”, *Nordisk Tidskr. Inform.* **15**:3 (1975), 331–334. [MR](#) [Zbl](#)
- [Rogers 1996] T. D. Rogers, “The graph of the square mapping on the prime fields”, *Discrete Math.* **148**:1-3 (1996), 317–324. [MR](#) [Zbl](#)
- [Vasiga and Shallit 2004] T. Vasiga and J. Shallit, “On the iteration of certain quadratic maps over $\text{GF}(p)$ ”, *Discrete Math.* **277**:1-3 (2004), 219–240. [MR](#) [Zbl](#)

Received: 2016-10-17

Accepted: 2016-12-05

ebellah@uoregon.edu

*Department of Mathematics, University of Oregon,
Eugene, OR 97403, United States*

gartondw@pdx.edu

*Fariborz Maseeh Department of Mathematics and Statistics,
Portland State University, PO Box 751, Portland, OR 97207,
United States*

ejt3@pdx.edu

*Fariborz Maseeh Department of Mathematics and Statistics,
Portland State University, PO Box 751, Portland, OR 97207,
United States*

nwalton@pdx.edu

*Fariborz Maseeh Department of Mathematics and Statistics,
Portland State University, PO Box 751, Portland, OR 97207,
United States*

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

MANAGING EDITOR

Kenneth S. Berenhaut Wake Forest University, USA

BOARD OF EDITORS

Colin Adams	Williams College, USA	Suzanne Lenhart	University of Tennessee, USA
John V. Baxley	Wake Forest University, NC, USA	Chi-Kwong Li	College of William and Mary, USA
Arthur T. Benjamin	Harvey Mudd College, USA	Robert B. Lund	Clemson University, USA
Martin Bohner	Missouri U of Science and Technology, USA	Gaven J. Martin	Massey University, New Zealand
Nigel Boston	University of Wisconsin, USA	Mary Meyer	Colorado State University, USA
Amarjit S. Budhiraja	U of North Carolina, Chapel Hill, USA	Emil Minchev	Ruse, Bulgaria
Pietro Cerone	La Trobe University, Australia	Frank Morgan	Williams College, USA
Scott Chapman	Sam Houston State University, USA	Mohammad Sal Moslehian	Ferdowsi University of Mashhad, Iran
Joshua N. Cooper	University of South Carolina, USA	Zuhair Nashed	University of Central Florida, USA
Jem N. Corcoran	University of Colorado, USA	Ken Ono	Emory University, USA
Toka Diagana	Howard University, USA	Timothy E. O'Brien	Loyola University Chicago, USA
Michael Dorff	Brigham Young University, USA	Joseph O'Rourke	Smith College, USA
Sever S. Dragomir	Victoria University, Australia	Yuval Peres	Microsoft Research, USA
Behrouz Emamizadeh	The Petroleum Institute, UAE	Y.-F. S. Pétermann	Université de Genève, Switzerland
Joel Foisy	SUNY Potsdam, USA	Robert J. Plemmons	Wake Forest University, USA
Erin W. Fulp	Wake Forest University, USA	Carl B. Pomerance	Dartmouth College, USA
Joseph Gallian	University of Minnesota Duluth, USA	Vadim Ponomarenko	San Diego State University, USA
Stephan R. Garcia	Pomona College, USA	Bjorn Poonen	UC Berkeley, USA
Anant Godbole	East Tennessee State University, USA	James Propp	U Mass Lowell, USA
Ron Gould	Emory University, USA	József H. Przytycki	George Washington University, USA
Andrew Granville	Université Montréal, Canada	Richard Rebarber	University of Nebraska, USA
Jerrold Griggs	University of South Carolina, USA	Robert W. Robinson	University of Georgia, USA
Sat Gupta	U of North Carolina, Greensboro, USA	Filip Saidak	U of North Carolina, Greensboro, USA
Jim Haglund	University of Pennsylvania, USA	James A. Sellers	Penn State University, USA
Johnny Henderson	Baylor University, USA	Andrew J. Sterge	Honorary Editor
Jim Hoste	Pitzer College, USA	Ann Trenk	Wellesley College, USA
Natalia Hritonenko	Prairie View A&M University, USA	Ravi Vakil	Stanford University, USA
Glenn H. Hurlbert	Arizona State University, USA	Antonia Vecchio	Consiglio Nazionale delle Ricerche, Italy
Charles R. Johnson	College of William and Mary, USA	Ram U. Verma	University of Toledo, USA
K. B. Kulasekera	Clemson University, USA	John C. Wierman	Johns Hopkins University, USA
Gerry Ladas	University of Rhode Island, USA	Michael E. Zieve	University of Michigan, USA

PRODUCTION

Silvio Levy, Scientific Editor


Cover: Alex Scorpan

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2018 is US \$190/year for the electronic version, and \$250/year (+\$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber 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/>

© 2018 Mathematical Sciences Publishers

involve

2018

vol. 11

no. 1

On halving-edges graphs	1
TANYA KHOVANOVA AND DAI YANG	
Knot mosaic tabulation	13
HWA JEONG LEE, LEWIS D. LUDWIG, JOSEPH PAAT AND AMANDA PEIFFER	
Extending hypothesis testing with persistent homology to three or more groups	27
CHRISTOPHER CERICOLA, INGA JOHNSON, JOSHUA KIERS, MITCHELL KROCK, JORDAN PURDY AND JOHANNA TORRENCE	
Merging peg solitaire on graphs	53
JOHN ENGBERS AND RYAN WEBER	
Labeling crossed prisms with a condition at distance two	67
MATTHEW BEAUDOUIN-LAFON, SERENA CHEN, NATHANIEL KARST, JESSICA OHRLEIN AND DENISE SAKAI TROXELL	
Normal forms of endomorphism-valued power series	81
CHRISTOPHER KEANE AND SZILÁRD SZABÓ	
Continuous dependence and differentiating solutions of a second order boundary value problem with average value condition	95
JEFFREY W. LYONS, SAMANTHA A. MAJOR AND KAITLYN B. SEABROOK	
On uniform large-scale volume growth for the Carnot–Carathéodory metric on unbounded model hypersurfaces in \mathbb{C}^2	103
ETHAN DLUGIE AND AARON PETERSON	
Variations of the Greenberg unrelated question binary model	119
DAVID P. SUAREZ AND SAT GUPTA	
Generalized exponential sums and the power of computers	127
FRANCIS N. CASTRO, OSCAR E. GONZÁLEZ AND LUIS A. MEDINA	
Coincidences among skew stable and dual stable Grothendieck polynomials	143
ETHAN ALWASE, SHULI CHEN, ALEXANDER CLIFTON, REBECCA PATRIAS, ROHIL PRASAD, MADELINE SHINNERS AND ALBERT ZHENG	
A probabilistic heuristic for counting components of functional graphs of polynomials over finite fields	169
ELISA BELLAH, DEREK GARTON, ERIN TANNENBAUM AND NOAH WALTON	