ON THE NUMBER OF DISTINCT FUNCTIONAL GRAPHS OF AFFINE-LINEAR TRANSFORMATIONS OVER FINITE FIELDS

ERIC BACH AND ANDREW BRIDY

Abstract. We study the number of non-isomorphic functional graphs of affine-linear transformations from \((F_q)^n\) to itself, and we prove upper and lower bounds on this quantity as \(n \to \infty\). As a corollary to our result, we prove bounds on the number of conjugacy classes in the symmetric group \(S_{q^n}\) that intersect \(AGL_n(q)\).

1. Introduction

Let \(X\) be a finite set and let \(f : X \to X\) be a function, so that the pair \((X, f)\) defines a discrete dynamical system. We define the \textit{functional graph} of \((X, f)\) to be a directed graph with vertices at each element of \(X\) and an edge from \(x\) to \(y\) if and only if \(f(x) = y\). We denote the functional graph of \((X, f)\) by \(G(X, f)\).

In the setting where \(X = (F_q)^n\) and \(f\) is a linear transformation, the structure of \(G(X, f)\) was explicitly determined by Elspas, who worked in the context of linear sequential networks in electrical engineering [6]. Wang, building on the work of Elspas, provided an explicit description of \(G(X, f)\) in the more general case where \(f\) is affine-linear [20].

Rather than analyze the structure of the functional graph, our problem is to estimate the number of non-isomorphic functional graphs of affine-linear transformations from \((F_q)^n\) to itself. Let \(D_q(n)\) denote this quantity. We will prove the following:

\textbf{Theorem 1.} For \(q\) fixed and \(n \to \infty\),

\[\sqrt{n} \ll \log D_q(n) \ll \frac{n}{\log \log n}.

The upper bound in Theorem 1 is in some sense an improvement on the well-known fact that the number of conjugacy classes in \(GL_n(q)\) is \(q^n + O(q^{n/2})\) [19], as similar linear transformations have isomorphic functional graphs. In fact, Theorem 1 implies that there exist linear transformations which are “similar” under conjugation by a non-linear permutation of \((F_q)^n\). We formalize this notion as follows:

\textbf{Definition 2.} The dynamical systems \((X, f)\) and \((Y, g)\) are \textit{dynamically equivalent} if there exists a bijection \(\sigma : X \to Y\) such that \(\sigma^{-1} \circ g \circ \sigma = f\).

It is easy to check that dynamical equivalence coincides with isomorphism of functional graphs, and we will use both concepts interchangeably. If \((X, f)\) and

\textit{Date:} May 2, 2014.

\textit{2010 Mathematics Subject Classification.} Primary 37P05; Secondary 05C20, 11T55.

\textit{Key words and phrases.} Linear Algebra, Finite Fields, Functional Graphs.
(X, g) are dynamically equivalent and X is understood, we say that f and g are dynamically equivalent, and we write f ∼ g for short. In this language, Theorem 1 can be restated as an estimate of the number of dynamical equivalence classes of affine-linear transformations f : (F_q)^n → (F_q)^n.

Our result also has a group-theoretic interpretation. Let G be the symmetric group of X = (F_q)^n, so that G ∼= S_q^n. Let H = AGL_n(q) be the group of invertible affine-linear transformations of X, so that H ⊆ G. Conjugacy classes of G are cycle types, which are special isomorphism classes of functional graphs on X (for those f : X → X that are bijections), so Theorem 1 yields the following corollary:

Corollary 3. Let A_q(n) be the number of conjugacy classes of G that intersect H. Then

\[ A_q(n) \leq \exp \left( O \left( \frac{n}{\log \log n} \right) \right). \]

We can rephrase Corollary 3 in terms of counting derangements, which are permutations with no fixed points. Let G act on the coset space G/H by left multiplication. The number of conjugacy classes of G that do not contain a derangement in this action is at most \( \exp \left( O \left( \frac{n}{\log \log n} \right) \right). \) As the total number of conjugacy classes of G is \( P(q^n) \), it follows from the Hardy-Ramanujan asymptotic formula for the partition function [2] that the proportion of conjugacy classes of G that contain a derangement approaches 1 as \( n \to \infty \).

There is significant work of Boston and others on the proportion of elements in a permutation group that act as derangements (e.g. [5],[7]). We know of no other results on the proportion of conjugacy classes that contain a derangement.

Remark 4. This paper grew out of work on the number of distinct functional graphs that arise from quadratic polynomials \( f(x) = ax^2 + bx + c \in F_{2^n}[x] \), where \( X = F_{2^n} \). If we fix \( \zeta \in F_{2^n} \) of absolute trace 1, an easy argument using Hilbert’s Theorem 90 shows that every such \( f \) is conjugate by some \( \alpha x + \beta \in F_{2^n}[x] \) to either \( x^2 + bx \) or \( x^2 + bx + (b^2 + 1)\zeta \), and moreover no distinct polynomials in these two families are conjugate by a linear polynomial in \( F_{2^n}[x] \). (However, the two families collapse into one under conjugation by polynomials \( \alpha x + \beta \in F_{2^n}[x] \). This, and the fact that there are exactly two inequivalent maps for each parameter \( b \), are predicted by the cohomological theory of twists of dynamical systems [18].) It follows from these “normal forms” that there at at most \( 2^{n+1} - 1 \) distinct functional graphs in the quadratic family.

Every quadratic polynomial is an affine-linear transformation of X as an \( F_{2^n} \)-vector space, so the improved upper bound in Theorem 1 applies to the special case of quadratic polynomial maps on X, but also to a much broader family of X and f.

2. Bounds on \( D_q(n) \)

The following useful definition is taken from [12] and will play a crucial role in our counting arguments.

Definition 5. The order of \( f \in F_q[x] \) with \( f(0) \neq 0 \) is the smallest positive n such that \( f | x^n - 1 \). We write ord \( f \) for the order of \( f \).

We record for later use a proposition counting orders of degree k irreducible polynomials over \( F_q \).
Proposition 6. Let \( O_q(k) \) be the number of integers that occur as orders of irreducible polynomials over \( \mathbb{F}_q \) of degree \( k \). Then

\[
O_q(k) = \sum_{d|k} \tau(q^d - 1)\mu(k/d),
\]

where \( \tau(n) \) is the number of divisors of \( n \). In particular, \( O_q(k) \leq \tau(q^k - 1) \).

Proof. For \( f \in \mathbb{F}_q[x] \) irreducible of degree \( k \) with \( f(0) \neq 0 \), it is easy to show that \( \text{ord} f \) is the multiplicative order of the element \( [x] \) in the field \( \mathbb{F}_q[x]/(f) \cong \mathbb{F}_{q^k} \) [2]. Each nonzero \( \alpha \in \mathbb{F}_{q^k} \) has an irreducible minimal polynomial over \( \mathbb{F}_q \) of degree dividing \( k \), and the order of this polynomial is the multiplicative order of \( \alpha \). All divisors of \( |\mathbb{F}_{q^k}^\times| = q^k - 1 \) occur as multiplicative orders of some \( \alpha \in \mathbb{F}_{q^k} \), and all irreducible polynomials over \( \mathbb{F}_q \) of degree dividing \( k \) split in \( \mathbb{F}_{q^k} \), so this proves

\[
\sum_{d|k} O_q(d) = \tau(q^k - 1)
\]

and the proposition follows by Möbius inversion. \( \square \)

In Proposition [7] we construct a family of dynamically inequivalent linear transformations of \( (\mathbb{F}_q)^n \) in one-to-one correspondence with partitions of \( n \) with distinct parts. If we denote the number of partitions of \( n \) into distinct parts by \( Q(n) \), we have \( \log Q(n) \sim \pi \sqrt{n/3} \) [2] and the lower bound in Theorem [1] follows.

Proposition 7. There exists a family of dynamically inequivalent linear transformations \( A : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^n \) in bijection with the partitions of \( n \) with distinct parts.

Proof. Let \( \lambda = (\lambda_1, \ldots, \lambda_r) \) be a partition of \( n \) with distinct parts. For \( 1 \leq i \leq r \), let \( A_i : (\mathbb{F}_q)^{\lambda_i} \to (\mathbb{F}_q)^{\lambda_i} \) be a linear transformation given by the companion matrix of any primitive polynomial of degree \( i \) over \( \mathbb{F}_q \), and let \( A_\lambda : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^n \) be defined as \( A = A_1 \oplus \cdots \oplus A_r \). The proposition follows once we show that if \( \lambda \) and \( \mu \) are unequal partitions of \( n \) with distinct parts, then \( A_\lambda \) and \( A_\mu \) are not dynamically equivalent.

Let \( \lambda = (\lambda_1, \ldots, \lambda_r) \) and \( \mu = (\mu_1, \ldots, \mu_r) \) be unequal partitions of \( n \) with distinct parts. There exists some \( m \in \mu \) such that \( m \notin \lambda \). Any primitive polynomial of degree \( i \) has order \( q^i - 1 \), and it follows that the graph \( G_{((\mathbb{F}_q)^n, B)} \) contains a cycle of length \( q^m - 1 \) [12]. We show that \( G_{((\mathbb{F}_q)^n, A)} \) has no cycle of this length.

The graph \( G_{((\mathbb{F}_q)^{\lambda_i}, A_i)} \) consists of one fixed point (the zero vector) and one cycle of length \( q^{\lambda_i} - 1 \) [9]. The cycle lengths in \( G_{((\mathbb{F}_q)^n, A)} \) are equal to least common multiples of the cycle lengths of the \( G_{((\mathbb{F}_q)^{\lambda_i}, A_i)} \) [9]. Therefore every cycle of \( G_{((\mathbb{F}_q)^n, A)} \) has length equal to

\[
\text{LCM}_{i \in S}[q^{\lambda_i} - 1]
\]

for some subset \( S \subseteq \{1, \ldots, r\} \). We need to show that this LCM cannot equal \( q^m - 1 \).

Assume by way of contradiction that \( \text{LCM}_{i \in S}[q^{\lambda_i} - 1] = q^m - 1 \). Each \( q^{\lambda_i} - 1 \) divides \( q^m - 1 \), so by the Euclidean Algorithm, each \( \lambda_i \) divides \( m \). In particular, \( m \geq \lambda_i \) for each \( i \in S \), but we know \( m \neq \lambda_i \) for any \( i \), so \( m > \lambda_i \) for each \( i \in S \). By Zsigmondy's Theorem [17], with the exception of \( (q, m) = (2, 6) \), there exists a prime divisor of \( q^m - 1 \) that does not divide \( q^k - 1 \) for \( k < m \). This prime cannot divide \( \text{LCM}_{i \in S}[q^{\lambda_i} - 1] \), which is a contradiction. In the exceptional case
(q, m) = (2, 6) we can verify directly that 2^6 − 1 = 63 is not the LCM of any collection of smaller numbers of the form 2^{λ_i} − 1.

To prove the upper bound in Theorem 1, we develop a series of preliminary results. Proposition 2 decomposes an affine-linear transformation into the direct sum of an arbitrary linear transformation and an affine transformation whose dynamical equivalence class is determined by an integer partition. Proposition 12 gives a sufficient condition for dynamical equivalence of linear transformations that depends on the factorization of their characteristic polynomials. Counting characteristic polynomials yields the result.

**Lemma 8.** Let V and W be vector spaces, and let the pairs of linear transformations A, B : V → V and C, D : W → W be such that A ∼ B and C ∼ D. Then A ⊕ C ∼ B ⊕ D as maps from V ⊕ W to itself.

**Proof.** We have A = φ^{-1}Bφ and C = ψ^{-1}Dψ for bijections φ : V → V and ψ : W → W. Define θ : V ⊕ W → V ⊕ W as φ on V and ψ on W, extending linearly to V ⊕ W. Then A ⊕ C = θ^{-1}(B ⊕ D)θ. □

**Proposition 9.** Let V = (F_q)^n and let T : V → V be defined as Tx = Ax + b, where A : V → V is linear and b ∈ V. There exists a direct sum decomposition V = V_1 ⊕ V_2 (where it is possible that V_1 = 0) with T = T_1 ⊕ T_2, T_1x = A_1x + b_1 and T_2x = A_2x + b_2, such that

1. T_2 ∼ A_2.

2. The dynamical equivalence class of T_1 is determined by a partition of dim V_1.

**Proof.** This follows from the work of Wang in [20]. Wang proves that if there exists an s ∈ V such that Ts = s, then T ∼ A, which is the V_1 = 0 case of the proposition. If no such s exists, then there exists a decomposition T = T_1 ⊕ T_2 as in the statement of the proposition such that T_2 ∼ A_2. In this case there exists another direct sum decomposition V_t = W_1 ⊕ · · · ⊕ W_r, with T_1 = S_1 ⊕ · · · ⊕ S_r, S_i : W_i → W_i affine-linear, such that if we let t_i = dim W_i, each graph G(W_i, S_i) consists of \frac{q^{t_i}}{ord(x−1)^{λ_i + 1}} cycles of length ord (x − 1)^{λ_i + 1}. Therefore the list (t_1, . . . , t_r), which partitions dim V_1, determines the dynamical equivalence class of each S_i. This determines the dynamical equivalence class of T_1 by Lemma 8. □

**Lemma 10.** Let V = (F_q)^n. Let the linear map A : V → V have characteristic polynomial f^r, where f is irreducible over F_q and f(0) ≠ 0. The dynamical equivalence class of A is determined by deg f, ord f, and a partition of r.

**Proof.** There exist direct sum decompositions V = V_1 ⊕ · · · ⊕ V_m and A = A_1 ⊕ · · · ⊕ A_m such that in some basis of each V_i, A_i : V_i → V_i can be written as the companion matrix of f^{λ_i} for some λ_i, and \sum λ_i = r.

Each graph G(V_i, A_i) is explicitly determined by the data ord f, deg f, and λ_i [9]. This determines the dynamical equivalence class of each A_i, and by Lemma 8 the dynamical equivalence class of A. □

**Remark 11.** The work of Elspas in [6] is intimately connected with the theory of linearly recurrent sequences over finite fields. In Lemma 10 each G(V_i, A_i) consists of cycles that correspond to all linearly recurrent sequences over F_q with characteristic
polynomial $f^{\lambda_i}$, and the lengths of the cycles are the periods of the sequences. The periods of the sequences and the number of sequences with each period can be computed from the data \( \text{ord} f \), \( \deg f \), and \( \lambda_i \) \[12, \text{Theorem 6.63}\].

**Proposition 12.** Let \( V = (\mathbb{F}_q)^n \) and let \( A : V \to V \) be a linear transformation. Let \( p \in \mathbb{F}_q[x] \) be the characteristic polynomial of \( A \) and write its factorization into irreducibles as

\[
p = x^{r_0} \prod_{i=1}^{m} p_i^{r_i}
\]

where the \( p_i \) are distinct and no \( p_i \) equals \( x \). The dynamical equivalence class of \( A \) is completely determined by an integer partition of each \( r_i \) and two lists of \( m \) positive integers: \( \{ \deg p_i \} \) and \( \{ \text{ord} p_i \} \).

**Proof.** By the theory of the Jordan canonical form \[8\] there exist direct sum decompositions \( A = A_0 \oplus \cdots \oplus A_m \) and \( V = V_0 \oplus \cdots \oplus V_m \) where \( A_i : V_i \to V_i \), the characteristic polynomial of \( A_0 \) is \( x^{r_0} \), and the characteristic polynomial of \( A_i \) is \( p_i^{r_i} \) for \( i \geq 1 \). The Jordan form of the nilpotent map \( A_0 \) is specified by a partition of \( r_0 \) in which each part is the size of a Jordan block, so this partition determines the similarity class of \( A_0 \) and hence the dynamical equivalence class (if two linear maps are similar, they are dynamically equivalent).

Suppose that \( \deg p_i \) and \( \text{ord} p_i \) are given for \( i \geq 1 \). Specifying a partition of each \( r_i \) determines the dynamical equivalence class of each \( A_i \) by Lemma \[10\], which in turn determines the dynamical equivalence class of \( A \) by Lemma \[8\]. \(\square\)

Before proceeding with the proof of the upper bound in Theorem \[1\] we record a proposition that will be needed at a key moment in the counting argument.

**Proposition 13.** Let \( i_q(n) \) denote the maximum possible number of distinct irreducible factors of a degree \( n \) polynomial over \( \mathbb{F}_q \). Then

\[
i_q(n) = O\left( \frac{n}{\log n} \right).
\]

**Proof.** For \( q \geq 3 \) it is proved in \[10, \text{Lemma A1}\] that

\[
i_q(n) \leq \frac{n}{\log_q(n) - 3},
\]

which immediately implies the proposition when \( q \neq 2 \). As we only require a weaker big-\( O \) estimate, we present a simplified version of the proof in \[10\] which also works for \( q = 2 \).

For \( f \in \mathbb{F}_q[x] \), let \( \omega(f) \) denote the number of distinct irreducible factors of \( f \). We construct \( f \) such that \( \omega(f) = i_q(n) \) by a greedy algorithm. That is, first multiply together all degree 1 irreducibles, then all degree 2 irreducibles, and so on, until multiplying \( f \) by another irreducible would raise its degree higher than \( n \). Then \( \deg f \leq n \) and no degree \( n \) polynomials have more distinct irreducible factors than \( f \). It suffices to prove the proposition for polynomials of the form

\[
f = g \left( \prod_{\text{irreducible } p \in \mathbb{F}_q[x], \deg p < k} p \right)
\]
where $g$ is a product of $m$ irreducible polynomials of degree $k$, each of which appears with multiplicity 1. Let $N_q(j)$ denote the number of irreducible polynomials over $\mathbb{F}_q$ of degree $j$. We have

$$(1) \quad \omega(f) = \sum_{j=1}^{k-1} N_q(j) + m$$

and

$$(2) \quad \deg(f) = \sum_{j=1}^{k-1} jN_q(j) + mk \leq n.$$ 

We now show that the inequality

$$(3) \quad \sum_{j=1}^{k} N_q(j) \leq \frac{3}{k} \sum_{j=1}^{k} jN_q(j)$$

holds for large $k$. As $N_q(j) \leq q^j / j$ \cite{12} we have

$$\sum_{j=1}^{k} N_q(j) \leq \sum_{j=1}^{k} \frac{q^j}{j} \leq \sum_{1 \leq j \leq k/2} q^j + \sum_{k/2 < j \leq k} q^{j+k/2}$$

$$\leq \frac{q^{k/2+1}}{q-1} + \frac{2q^{k+1} - 2q^{k/2}}{k(q-1)}$$

$$\leq \frac{1}{q-1} \left( q^{k/2+1} + \frac{2q^{k+1}}{k} \right).$$

As $N_q(j) \geq q^j / j - q^{j/2+1} / (j(q-1))$ \cite{12} we have

$$\frac{3}{k} \sum_{j=1}^{k} jN_q(j) \geq \frac{3}{k} \left( \sum_{j=1}^{k} \frac{q^j - q^{j+k/2}}{q-1} \right)$$

$$= \frac{1}{q-1} \left( \frac{3q^{k+1}}{k} - 3q - \frac{3q^{(k+1)/2+1} - 3q^{3/2}}{k(q-1)} \right).$$

Therefore equation \cite{3} holds if

$$q^{k/2+1} + \frac{3q^{k+1}}{k} \leq \frac{3q^{k+1}}{k} - 3q - \frac{3q^{(k+1)/2+1} - 3q^{3/2}}{k(q-1)}.$$

or equivalently

$$q^{k/2+1} + \frac{3q}{k} + \frac{3q^{(k+1)/2+1}}{k(q-1)} \leq \frac{q^{k+1}}{k} + \frac{3q^{3/2}}{k(q-1)}.$$

Comparing powers of $q$ on both sides, it is clear that this inequality holds for large $k$.

Returning to equations \cite{1} and \cite{2} we use \cite{3} to conclude

$$i_q(n) = \omega(f) = \sum_{j=1}^{k-1} N_q(j) + m \leq \frac{3}{k} \left( \sum_{j=1}^{k-1} jN_q(j) + mk \right) \leq \frac{3n}{k}.$$ 

It only remains to show $k \geq C \log n$ for some $C$. By our construction of $f$, the largest that $n$ can be for a given $k$ occurs when $g$ is the product of all degree $k$
irreducible polynomials over $\mathbb{F}_q$. For this $g$, $n \leq k + \sum_{j=1}^{k} jI_j$ because if $n$ exceeded this amount, we could add a degree $k + 1$ irreducible factor to $f$. So

$$n \leq k + \sum_{j=1}^{k} jI_j \leq k + \sum_{j=1}^{k} q^k \leq k + \frac{q^{k+1} - q}{q - 1} \leq k + q^{k+1}.$$ 

When $k$ is large, $k < q^{k+1}$. Recall that $q \geq 2$. These imply that $n \leq k + q^{k+1} \leq q^{k+2}$, so $\log_q(n) \leq k + 2$. For $n \geq q^4$

$$k \geq \log_q(n) - 2 \geq \frac{1}{2} \log_q(n),$$

which completes the proof. □

We now combine Propositions [9] and [12] to give an upper bound on $D_q(n)$, completing the proof of Theorem 1.

**Theorem 14.**

$$D_q(n) = \exp\left(O\left(\frac{n}{\log \log n}\right)\right).$$

**Proof.** Let $V = (\mathbb{F}_q)^n$. Let $T : V \to V$ be defined by $Tx = Ax + b$, where $A$ is linear and $b \in V$. By Proposition [9] $V = V_1 \oplus V_2$ and $T = T_1 \oplus T_2$ such that the dynamical equivalence class of $T_1$ is determined by a partition of dim $V_1$ and $T_2 \sim A_2$ for some linear $A_2 : V_2 \to V_2$. (It may be the case that $V_1 = 0$ and $T_1 = 0$.) Let

$$p = x^{r_0} \prod_{i=1}^{m} p_i^{r_i}$$

be the characteristic polynomial of $A_2$. Note that $\deg \ p \leq n$. By Proposition [12] the dynamical equivalence class of $A_2$ is determined by a partition of each $r_i$ and the two lists of $m$ integers $\{\deg p_i\}$ and $\{\ord p_i\}$. We estimate the number of ways of specifying these data. Assume for the moment that the $\deg p_i$ and the partitions of the $r_i$ are given and that we need to assign orders to the $p_i$.

Let $d_i = \deg p_i$. By Proposition [6] the number of possible ways to assign the $\ord p_i$ is

$$\prod_{i=1}^{m} O_q(d_i) \leq \prod_{i=1}^{m} \tau(q^{d_i} - 1).$$

We split this into two products over the ranges $d_i < d$ and $d_i \geq d$ for some $d$ to be chosen later. First we estimate the quantity

$$C_1 = \prod_{d_i < d} \tau(q^{d_i} - 1).$$

Using the trivial estimate $\tau(x) < x + 1$,

$$C_1 < \prod_{d_i < d} q^{d_i} = q^{\sum_{d_i < d} d_i}.$$ 

Each $d_i$ is the degree of a distinct irreducible polynomial over $\mathbb{F}_q$. As in the proof of Proposition [13] $N_q(k) \leq \frac{q^k}{k}$, so

$$\sum_{d_i < d} d_i \leq \sum_{k=1}^{d-1} kI_k \leq \sum_{k=1}^{d-1} q^k \leq q^d.$$
Therefore $C_1 < q^d$.

Now we estimate

$$C_2 = \prod_{d_i \geq d} \tau(q^{d_i} - 1).$$

By the estimate on $\tau(x)$ in [4, Theorem 8.8.9], there exists $c$ such that

$$\tau(q^{d_i} - 1) \leq \frac{e^{\log(q^{d_i} - 1)}}{\log(q^{d_i} - 1)}.$$  

This implies

$$C_2 \leq \prod_{d_i \geq d} \frac{c d_i \log q}{\log(q^{d_i} - 1)} \leq 2^{\frac{c \log q \sum_{d_i \geq d} d_i}{\log \log q^{d_i - 1}}}$$

where we use the inequality $q^{d_i} - 1 \geq q^{d_i - 1}$, which is true for $q \geq 2$ and $d_i \geq 1$. Also, $\sum_{d_i \geq d} d_i \leq n$, so

$$C_2 \leq \frac{2^{\frac{c \log q \sum_{d_i \geq d} d_i}{\log \log q^{d_i - 1}}}}{\log \log n}.$$  

Putting these estimates together we have

$$\prod_{i=1}^{m} \tau(q^{d_i} - 1) = C_1 C_2 \leq \exp \left( q^d \log q + \log 2 \frac{c \log q n}{\log(d - 1) + \log \log q} \right).$$

Choose $d = \frac{\log n}{\log q}$ and note $\log(d - 1) \geq \log(d/2)$ for $d \geq 1$. Then

$$q^d \log q + \log 2 \frac{c \log q n}{\log(d - 1) + \log \log q} \leq n^{1/2} \log q + \frac{cn \log 2 \log q}{\log \log n} + \log \log q$$

$$= O \left( \frac{n}{\log \log n} \right).$$

Now we estimate the number of ways that the $r_i$ and $\deg p_i$ can occur. Because $\sum_{i=1}^{m} r_i \deg p_i = \dim V_2$, the $\deg p_i$ form a partition of $\dim V_2$ in which each appears $r_i$ times. This is a “factorization pattern” of $\dim V_2$ as in [9] which is specified by first picking a partition of $\dim V_2$ into parts of size $k$, each of which occurs $s_k$ times, and then further dividing each $s_k$ into parts $r_i$. Let $b(n)$ denote the number of factorization patterns of $n$. It is mentioned in [1] and proved in [14] that

$$b(n) = \exp \left( B \sqrt{n \log n} + O(\sqrt{n}) \right).$$

Finally, we need to choose a partition of each $r_i$ and a dynamical equivalence class for $T_1$ given by a partition of $\dim V_1$. If $P(x)$ denotes the partition function, we have $P(x) \leq \exp(K \sqrt{x})$ [2]. The number of ways to specify all these partitions is

$$P(\dim V_1) \prod_{i=0}^{m} P(r_i) \leq P(n) \exp \left( K \sqrt{n + K \sqrt{m + 1} \sum_{i=0}^{m} r_i} \right)$$

$$\leq \exp \left( K \sqrt{n + K \sqrt{m + 1} \sum_{i=0}^{m} r_i} \right) \leq \exp(K \sqrt{n(1 + \sqrt{m + 1})})$$

$$= \exp \left( K \sqrt{n} \left( 1 + O \left( \frac{n}{\sqrt{\log n}} \right) \right) \right) = \exp \left( O \left( \frac{n}{\sqrt{\log n}} \right) \right).$$

The inequality $\sum_{i=0}^{m} \sqrt{r_i} \leq \sqrt{m} \frac{1}{\sqrt{\log n}} \sum_{i=0}^{m} r_i$ follows from the standard fact that the arithmetic mean of the $\sqrt{r_i}$ is at most the root mean square, and $m + 1$ is the
number of distinct irreducible factors of $p$, so $m + 1 = O(n/\log n)$ by Proposition 13.

Putting this all together, the number of ways to choose a dynamical equivalence classes for $T_1$ and $T_2$, and therefore a dynamical equivalence class for $T$ by Lemma 8, is at most

$$b(n) \exp\left( O\left( \frac{n}{\sqrt{\log n}} \right) \right) \exp\left( O\left( \frac{n}{\log \log n} \right) \right) = \exp\left( O\left( \frac{n}{\log \log n} \right) \right)$$

which completes the proof of Theorem 11.

\[\square\]

**Remark 15.** It seems possible that the estimates in Theorem 14 could be improved. The main estimate used for $\tau(q^d - 1)$ is the worst-case estimate on $\tau(x)$ that follows from the prime number theorem. It may be possible to give a better estimate based on the distribution of multiplicative orders of $q$ modulo various integers $n$. (If $n$ divides $q^d - 1$, then $d$ is a multiple of the multiplicative order of $q$ mod $n$.) Questions along these lines tend to be difficult, even for $q = 2$. See [3], [11], [13], [15], and [16] for some related work.

**Acknowledgements.** This research was partly supported by NSF grants no. CCF-0635355 and EMSW21-RTG. The second author would like to thank ICERM for an invitation to the Spring 2012 Semester Program on Complex and Arithmetic Dynamics, which facilitated the development of many of these ideas. We would like to thank Joseph Silverman for helpful comments and observations on a preliminary version of this paper, and Michael Zieve for suggesting the connection to group theory.

**References**

[1] A. K. Agarwal and G. L. Mullen. Partitions with “d(a) copies of a”. *J. Combin. Theory Ser. A*, 48(1):120–135, 1988.

[2] G. E. Andrews. *The Theory of Partitions*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1998. Reprint of the 1976 original.

[3] V. Arnold. Number-theoretical turbulence in Fermat-Euler arithmetics and large Young diagrams geometry statistics. *J. Math. Fluid Mech.*, 7(suppl. 1):S4–S50, 2005.

[4] E. Bach and J. Shallit. *Algorithmic Number Theory. Vol. 1: Efficient Algorithms*. Foundations of Computing Series. MIT Press, Cambridge, MA, 1996.

[5] N. Boston, W. Dabrowski, and T. Foggel et. al. The proportion of fixed-point-free elements of a transitive permutation group. *Comm. Algebra*, 21(9):3259–3275, 1993.

[6] B. Elspas. The theory of autonomous linear sequential networks. In *Linear Sequential Switching Circuits*, pages 21–61. Holden-Day, San Francisco, Calif., 1965.

[7] R. Guralnick and D. Wan. Bounds for fixed point free elements in a transitive group and applications to curves over finite fields. *Israel J. Math.*, 101:255–287, 1997.

[8] K. Hoffman and R. Kunze. *Linear Algebra*. Second edition. Prentice-Hall Inc., Englewood Cliffs, N.J., 1971.

[9] R. A. Hultquist, G. L. Mullen, and H. Niederreiter. Association schemes and derived PBIB designs of prime power order. *Ars Combin.*, 25:65–82, 1988.

[10] M. Kaminski and N. H. Bshouty. Multiplicative complexity of polynomial multiplication over finite fields. *J. Assoc. Comput. Mach.*, 36(1):150–170, 1989.

[11] P. Kurlberg and C. Pomerance. On a problem of Arnold: the average multiplicative order of a given integer. *Algebra and Number Theory*. To appear.

[12] R. Liadh and H. Niederreiter. *Introduction to Finite Fields and their Applications*. Cambridge University Press, Cambridge, first edition, 1994.

[13] F. Luca and I. E. Shparlinski. Average multiplicative orders of elements modulo $n$. *Acta Arith.*, 109(4):387–411, 2003.
[14] D. J. B. Mitchell. Generating functions for various sets of solid partitions. Ph. D. Thesis, Penn. State Univ., 1972.
[15] M. R. Murty, M. Rosen, and J. H. Silverman. Variations on a theme of Romanoff. Internat. J. Math., 7(3):373–391, 1996.
[16] C. Pomerance. On primitive divisors of Mersenne numbers. Acta Arith., 46(4):355–367, 1986.
[17] M. Roitman. On Zsigmondy primes. Proc. Amer. Math. Soc., 125(7):1913–1919, 1997.
[18] J. H. Silverman. The field of definition for dynamical systems on $\mathbb{P}^1$. Compositio Math., 98(3):269–304, 1995.
[19] R. P. Stanley. Enumerative Combinatorics. Volume 1, volume 49 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition, 2012.
[20] K. C. Wang. Transition graphs of affine transformation on vector spaces over finite fields. J. Franklin Inst., 283:55–72, 1967.

ERIC BACH, DEPARTMENT OF COMPUTER SCIENCES, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706, USA
E-mail address: bach@cs.wisc.edu

ANDREW BRIDY, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706, USA
E-mail address: bridy@math.wisc.edu