COUNTING ARCS IN THE PROJECTIVE PLANE VIA GLYNN'S ALGORITHM

NATHAN KAPLAN, SUSIE KIMPORT, RACHEL LAWRENCE, LUKE PEILEN, AND MAX WEINREICH

Abstract. An n-arc in a projective plane is a collection of n distinct points in the plane, no three of which lie on a line. Formulas counting the number of n-arcs in a finite projective plane of order q are known for $n \leq 8$ for general projective planes. In 1995, Iampolskaia, Skorobogatov, and Sorokin counted 9-arcs in the projective plane over a finite field of order q and showed that this count is a quasipolynomial function of q. We present a formula which counts 9-arcs in any projective plane of order q, even those that are non-Desarguesian, deriving Iampolskaia, Skorobogatov, and Sorokin's formula as a special case. We obtain our formula from a new implementation of an algorithm due to Glynn; we give details of our implementation and discuss its consequences for larger arcs.

The following research paper is finished and intended for submission to *Designs, Codes and Cryptography* by December 15, 2016 following a final proofread.

1. Introduction

We begin by recalling the basic definitions needed to discuss n-arcs in finite projective planes.

Definition 1.1. Let q be a positive integer. A *projective plane* Π is a collection of points $\mathcal{P}$ and a collection of lines $\mathcal{L}$, where each $\ell \in \mathcal{L}$ is a subset of $\mathcal{P}$ such that

(1) Every two points are incident with a unique line, that is, given distinct points $p_1, p_2, \in \mathcal{P}$ there exists a unique $\ell \in \mathcal{L}$ such that $\{p_1, p_2\} \subset \ell$.
(2) Every two lines are incident with a unique point, that is, given distinct lines $\ell_1, \ell_2 \in \mathcal{L}$ there exists a unique $p \in \mathcal{P}$ with $p \in \ell_1 \cap \ell_2$.
(3) There exist four points such that no three of them are contained in any line.

We say that Π has *order* q if each line contains exactly $q+1$ points, and if each point is contained in exactly $q+1$ lines.

The projective plane over a field k, denoted $\mathbb{P}^2(k)$, gives a well-studied algebraic class of examples. These are the *Desarguesian* planes, that is, those in which the theorem of Desargues holds. Planes which cannot be coordinatized by a field are called *non-Desarguesian.*

We study functions which count special configurations of points called n-arcs.

Date: December 10, 2016.
2010 *Mathematics Subject Classification.* -.
Key words and phrases. Arcs; Finite Projective Planes; Configurations of Points and Lines; Strong Realizations; Linear Spaces.

Definition 1.2. An *n-arc* in a projective plane Π is a collection of n distinct points, no three of which are collinear.

Arcs are collections of sets in *linear general position*, a fundamental concept in classical and algebraic geometry. In an infinite projective plane, most collections of points form arcs, but in finite projective planes, interesting enumerative problems arise.

For simplicity, throughout this paper we will count ordered n-arcs, that is, n-tuples of points that form an arc, and we often omit the adjective ordered. The number of ordered n-arcs in a projective plane is equal to the number of n-arcs in the plane multiplied by a factor of $n!$.

Definition 1.3. Let Π be a projective plane of order q. Define $C_n(\Pi)$ as the number of ordered n-arcs of Π. In the particular case where Π is the projective plane $\mathbb{P}^2(\mathbb{F}_q)$, we write $C_n(q)$ in place of $C_n(\Pi)$.

For small values of n we can determine $C_n(q)$ using the algebraic structure of $\mathbb{P}^2(\mathbb{F}_q)$. For example, the automorphism group of $\mathbb{P}^2(\mathbb{F}_q)$ is $\mathrm{PGL}_3(\mathbb{F}_q)$, which acts transitively on collections of 4 points, no three of which lie on a line. Therefore,

$$C_4(q) = |\mathrm{PGL}_3(\mathbb{F}_q)| = (q^2+q+1)(q^2+q)q^2(q-1)^2.$$

In fact, exactly the same formula holds in *any* projective plane Π of order q. Similar but more intricate ideas lead to similar polynomial formulas for all $n \leq 6$.

Theorem 1.4. [1, Theorem 4.1] *For any finite projective plane* Π *of order* q*, we have*

(1) $C_1(\Pi) = C_1(q) = q^2+q+1,$

(2) $C_2(\Pi) = C_2(q) = (q^2+q+1)(q^2+q),$

(3) $C_3(\Pi) = C_3(q) = (q^2+q+1)(q^2+q)q^2,$

(4) $C_4(\Pi) = C_4(q) = (q^2+q+1)(q^2+q)q^2(q-1)^2,$

(5) $C_5(\Pi) = C_5(q) = (q^2+q+1)(q^2+q)q^2(q-1)^2(q^2-5q+6),$

(6) $C_6(\Pi) = C_6(q) = (q^2+q+1)(q^2+q)q^2(q-1)^2(q^2-5q+6)(q^2-9q+21).$

Formulas for $C_7(\Pi)$ and $C_8(\Pi)$ are known, but are no longer given by a single polynomial in q for any finite projective plane Π. These formulas involve the number of *strong realizations* of certain special configurations of points and lines. We recall some definitions from [2] and [9].

Definition 1.5. A *linear space* is a pair of sets $(\mathcal{P}, \mathcal{L})$, whose elements are referred to as *points* and *lines* respectively, with the following properties: each line is a subset of $\mathcal{P}$, and no two lines intersect in more than one point. A pair (p, l) consisting of a point p and a line l containing p is called an *incidence* of the linear space. Two linear spaces are *isomorphic* if there is a permutation of the points and a permutation of the lines of the second linear space such that its set of incidences exactly match those of the first.

Linear spaces capture the basic notions of incidence geometry without reference to any contextual projective plane. Thus, a linear space may be thought of as a combinatorial blueprint whose incidence data may or may not be satisfied by any given set of points and lines in a projective plane. For instance, every n-arc encapsulates the data of the linear space $(\{1, 2, ..., n\}, \emptyset)$. To formalize this notion, we introduce the idea of *strong realizations*.

Definition 1.6. A *strong realization* of a linear space S in a projective plane Π is a collection of points and lines of Π such that the set of incidences between them is isomorphic to the set of incidences of S. We write $A_S(\Pi)$ for the number of strong realizations of S in Π, and in the particular case where $\Pi = \mathbb{P}^2(\mathbb{F}_q)$, we write $A_S(q)$.

With this language, we can restate our problem as follows: we wish to give a formula which calculates $A_a(\Pi)$, where a is the linear space $(\{1, ..., n\}, \emptyset)$. We will do so by making use of a special class of linear spaces called *superfigurations*, defined as follows.

Definition 1.7. A (combinatorial) n_k*-configuration* is a linear space such that each line consists of exactly k points, and each point is contained in exactly k lines.

A linear space of n points and some number of lines, not necessarily n, in which each line contains *at least* k points and each point is contained in *at least* k lines is called an n_k-*superfiguration*. Since every configuration is a superfiguration, but not conversely, we state all of our results and definitions in terms of superfigurations.

Throughout this paper all n_k-configurations we encounter will have $k = 3$, so we will refer to n_3-configurations simply as *configurations*. Similarly, we refer to n_3-superfigurations as *superfigurations*. Glynn refers to superfigurations as "variables" [1], and Iampolskaia, Skorobogatov, and Sorokin call them "overdetermined configurations" [4]. We have opted to call them superfigurations for the sake of consistency with Grunbaum's text on classical configurations of points and lines, and to distinguish them from other uses of the terms "variable" and "configuration" [2].

Glynn gives formulas for the number of 7-arcs and the number of 8-arcs in a projective plane Π of order q that involve the number of strong realizations of certain special configurations. Let $A_7(\Pi)$ denote the number of strong realizations of the *Fano plane*, the unique superfiguration with 7 points up to isomorphism. Similarly, let $A_8(\Pi)$ denote the number of strong realizations of the *Möbius-Kantor* configuration, the unique superfiguration with 8 points up to isomorphism.

Theorem 1.8. [1, Theorems 4.2 and 4.4] *Let Π be a projective plane of order q. Then*

(1) $C_7(\Pi) = (q^2+q+1)(q^2+q)q^2(q-1)^2(q-3)(q-5)(q^4-20q^3+148q^2-468q+498)-A_7(\Pi)$,

(2) $$C_8(\Pi) = (q^2 + q + 1)(q^2 + q)q^2(q - 1)^2(q - 5) \\ \cdot (q^7 - 43q^6 + 788q^5 - 7937q^4 + 47097q^3 - 162834q^2 + 299280q - 222960) \\ - (q^2 - 20q + 78)A_7(\Pi) + A_8(\Pi).$$

Glynn demonstrates the value of these formulas by applying them to the problem of classifying finite projective planes. Noting that $A_7(q) \geq 0$ shows that $C_7(\Pi)$ is at most the value of the degree 14 polynomial given in the first part of the theorem. However, for $q = 6$ this polynomial evaluates to -6, which shows that there is no projective plane of order 6 [1]. It is not clear whether formulas for $C_n(\Pi)$ for larger n will have similar consequences for the classification of finite projective planes.

To see another consequence of these formulas, let us consider $A_7(q)$. In planes of odd order, we have $A_7(q) = 0$; however, if q is even, then any 4-arc together with its three diagonal points forms a Fano plane. Thus, $A_7(q) = C_4(q)$. Even in a non-Desarguesian plane, each Fano subplane contains a 4-arc, so $A_7(\Pi) \leq C_4(\Pi)$ in general. Now, Theorem 1.8 shows that the sum of the number of 7-arcs and the number of Fano subplanes of Π depends only on the order of Π. Thus, for any $q = 2^r$, among planes of order q, the Desarguesian plane $\mathbb{P}^2(\mathbb{F}_q)$ is minimal with respect to number of 7-arcs.

Further, a conjecture widely attributed to Neumann [10] states that every non-Desarguesian finite projective plane contains a Fano subplane. Thus, thanks to Theorem 1.8, we may reformulate Neumann's conjecture as follows.

Conjecture 1.9. Let Π be a non-Desarguesian plane. Then the number of 7-arcs in Π is strictly less than $(q^2+q+1)(q^2+q)q^2(q-1)^2(q-3)(q-5)(q^4-20q^3+148q^2-468q+498)$.

And this gives rise to the following equivalent formulation:

Conjecture 1.10. Fix an odd prime power q. Then among planes of order q, the Desarguesian plane $\mathbb{P}^2(\mathbb{F}_q)$ is strictly maximal with respect to number of 7-arcs.

As the above discussion shows, $A_7(q)$ is calculated by one of two polynomials depending on the parity of q. This is an example of a *quasipolynomial* or *PORC* (Polynomial on Residue Classes) formula; that is, there exists some period m and a collection of polynomials $f_1(q), \ldots, f_m(q)$ such that our counting function is given by $f_i(q)$ for all q congruent to i modulo m. Similarly, $A_8(q)$ is given by a quasipolynomial with period 3. Glynn computes these quasipolynomials and derives explicit quasipolynomial formulas for $C_7(q)$ and $C_8(q)$. For details, see Theorem 4.5 of [1]. Glynn did not push his method further, noting "the complexity of the problem as the number of points approaches 10" [1].

In order to study the problem of counting inequivalent linear MDS codes, Iampolskaia, Skorobogatov, and Sorokin give a formula for the number of ordered n-arcs in $\mathbb{P}^2(\mathbb{F}_q)$. There are 10 superfigurations on 9 variables up to isomorphism, and we let $A_{9_3}(\Pi), A_{9_2}(\Pi), \ldots, A_{9_{12}}(\Pi)$ denote the number of strong realizations of the corresponding configuration in a projective plane Π. When Π is $\mathbb{P}^2(\mathbb{F}_q)$ we instead write $A_{9_i}(q)$ for this quantity.

Theorem 1.11. [4, Theorem 1] *The number of* 9*-arcs in the projective plane over the finite field* $\mathbb{F}_q$ *is given by*

$$\begin{aligned}
C_9(q) = {} & (q^2+q+1)(q^2+q)(q^2)(q-1)^2 \\
& (q^{10} - 75q^9 + 2530q^8 - 50466q^7 + 657739q^6 - 5835825q^5 + 35563770q^4 \\
& - 146288034q^3 + 386490120q^2 - 588513120q + 389442480 \\
& - 1080(q^4 - 47q^3 + 807q^2 - 5921q + 15134)a(q) \\
& + 840(9q^2 - 243q + 1684)b(q) \\
& + 30240(-9c(q) + 9d(q) + 2e(q)))
\end{aligned}$$

where

$$
\begin{aligned}
a(q) &= \begin{cases} 1 & \text{if } 2 \mid q, \\ 0 & \text{otherwise}; \end{cases} \\
b(q) &= \#(x \in \mathbb{F}_q : x^2 + x + 1 = 0), \\
c(q) &= \begin{cases} 1 & \text{if } 3 \mid q, \\ 0 & \text{otherwise}; \end{cases} \\
d(q) &= \#\{x \in \mathbb{F}_q : x^2 + x - 1 = 0\} \\
e(q) &= \#\{x \in \mathbb{F}_q : x^2 + 1 = 0\}.
\end{aligned}
$$

The authors further show that each of the functions $a(q), b(q), c(q), d(q), e(q)$ has a quasipolynomial formula. For example, $e(q)$ depends only on the residue of q modulo 4 [4]. As a consequence, the above formula $C_9(q)$ is a quasipolynomial.

In order to prove Theorem 1.11, the authors use the fact that there is a natural way to assign coordinates to the points of $\mathbb{P}^2(\mathbb{F}_q)$. For example, by taking an appropriate change of coordinates, every superfiguration on at most 9 points is projectively equivalent to one where five points are chosen to be $[1:0:0]$, $[0:1:0]$, $[0:0:1]$, $[1:1:1]$, and $[1:1:0]$, and we need only count the number of strong realizations of this configuration where these five points are fixed [4, 8].

In a general projective plane of order q, it is not guaranteed such an assignment of coordinates to points exists; as such, we do not have this technique at our disposal. In particular, we note that Theorem 1.11 is insufficient to describe the count of 9-arcs in non-Desarguesian planes, that is, planes which are not coordinatized by a field.

The main result of this paper is to extend Theorem 1.11 to any projective plane Π of order q.

Theorem 1.12. *The number of* 9*-arcs in a general projective plane* $\mathbb{P}$ *is given by*

$$
\begin{aligned}
C_9(\Pi) &= q^{18} - 75q^{17} + 2529q^{16} - 50392q^{15} + 655284q^{14} \\
&\quad -5787888q^{13} + 34956422q^{12} - 141107418q^{11} + 356715069q^{10} \\
&\quad -477084077q^9 + 143263449q^8 + 237536370q^7 + 52873326q^6 \\
&\quad -2811240q^5 - 588466080q^4 + 389304720q^3 \\
&\quad +(-36q^4 + 1692q^3 - 29052q^2 + 212148q - 539784)A_7(\Pi) \\
&\quad +(9q^2 - 243q + 1647)A_8(\Pi) \\
&\quad -A_{9_3}(\Pi) - A_{9_4}(\Pi) - A_{9_5}(\Pi) + A_{9_7}(\Pi) - 3A_{9_{10}}(\Pi) + 3A_{9_{11}}(\Pi) - 9A_{9_{12}}(\Pi).
\end{aligned}
$$

Theorem 1.12 has several useful consequences. First, to naively count the 9-arcs in a non-Desarguesian plane of order q requires considering all sets of 9 points and testing for collinearity. Because the number of points in such a plane is $q^2 + q + 1$, the direct count takes time $O(q^{18})$. Our formula reduces the problem to counting strong realizations of twelve superfigurations. Each of these superfigurations is so highly determined that at most q^2 sets must be checked following a selection of four initial points. Therefore, our formula improves the time taken to $O(q^{10})$.

Second, Theorem 1.12 demonstrates a non-obvious relationship between 9-arcs and superfigurations such as the Pappus configuration, while showing that superfigurations 9_6, 9_8 and 9_9 have no bearing on the count. These relationships could conceivably produce more arguments along the lines of Conjectures 1.9 and 1.10.

Third, we can derive Theorem 1.11 as a corollary, giving a new proof. We substitute the following values of $A_s(q)$ into the formula above, where the functions a, b, c, d, e are as in the statement of Theorem 1.11.

$$\begin{aligned}
A_7(q) &= 30a(q)C_4(q), \\
A_8(q) &= 840b(q)C_4(q), \\
A_{9_3}(q) &= 3360((q-2-b(q))(q-5)+(q-3)b(q))C_4(q), \\
A_{9_4}(q) &= 40320(q-2-b(q))C_4(q), \\
A_{9_5}(q) &= 30240(q-3)(1-a(q))C_4(q), \\
A_{9_6}(q) &= 30240(q-2)a(q)C_4(q), \\
A_{9_7}(q) &= 60480(e(q)-a(q))C_4(q), \\
A_{9_8}(q) &= 10080(q-2-b(q))C_4(q), \\
A_{9_9}(q) &= 0, \\
A_{9_{10}}(q) &= 1680b(q)C_4(q), \\
A_{9_{11}}(q) &= 90720d(q)C_4(q), \\
A_{9_{12}}(q) &= 30240c(q)C_4(q).
\end{aligned}$$

These are derived in [4], with a minor error in the calculation of $A_{9_3}(\Pi)$. Thus we confirm that despite this error, the end results of [4] are correct.

In Section 3, we describe the algorithm used to prove Theorem 1.12. In Section 5, we describe the computational aspects of the implementation and potential extensions of this work.

2. Motivation for Counting Arcs

The problem of counting arcs in the finite projective plane $\mathbb{P}^2(\mathbb{F}_q)$ has a long history, perhaps best known through its connection to the theory of error-correcting codes.

Definition 2.1. A *linear code* $C \subseteq \mathbb{F}_q^n$ is just a linear subspace of $\mathbb{F}_q^n$. Its dimension as a linear subspace is called the *dimension* of C.

The Hamming distance between two elements $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{F}_q^n$ is the number of $i \in [1, n]$ such that $x_i \neq y_i$, and is denoted $d(x, y)$. The *minimum distance* of a code is the minimum value of $d(x, y)$ taken over all pairs of elements of C with $x \neq y$.

One of the main combinatorial problems in coding theory is to determine, given n, q, and d, the maximum dimension of a linear code $C \subseteq \mathbb{F}_q^n$ with minimum distance at least d. There is a huge body of research on giving constructions of large codes with good minimum distances, and proving upper bounds for the size of a code given its minimum distance. One of the simplest results in this area is known as the Singleton bound, which says that the dimension k of a linear code $C \subseteq \mathbb{F}_q^n$ with minimum distance d satisfies $k \leq n - (d-1)$.

Definition 2.2. A linear code $C \subseteq \mathbb{F}_q^n$ with minimum distance d and dimension $k = n - (d-1)$ is called *MDS*, or *Maximum Distance Separable.*

Two linear codes are equivalent if one can be obtained from the other by permuting some coordinates and scaling some coordinates by a nonzero element of $\mathbb{F}_q$. It is well-known that the problem of classifying n-arcs in $\mathbb{P}^2(\mathbb{F}_q)$ and the problem of classifying inequivalent 3-dimensional MDS codes $C \subseteq \mathbb{F}_q^n$ are equivalent. This is part of a much larger story involving the MDS conjecture, which asks for the largest n such that a k-dimensional MDS code over $\mathbb{F}_q^n$ exists. This leads to the notion of an arc in a higher dimensional projective space, a collection of n distinct points in $\mathbb{P}^{k-1}(\mathbb{F}_q)$, not k of which lie in a hyperplane. The connections between MDS codes and arcs have motivated much of the research about counting arcs and classifying extremal arcs including the paper of Iampolskaia, Skorobogatov, and Sorokinon 9-arcs in $\mathbb{P}^2(\mathbb{F}_q)$ [4]. For more information on the interplay between arcs and MDS codes, see the survey article [3].

The classification of arcs in the plane has also been useful in the study of higher-dimensional arcs; for example, one step in Glynn's classification of 10-arcs in $\mathbb{P}^4(\mathbb{F}_9)$ is to note that projecting this arc gives an 8-arc in $\mathbb{P}^2(\mathbb{F}_9)$, which are counted in Theorem 1.8. Given a count for such arcs it is straightforward to classify them, which aids in classifying the higher-dimensional arcs projecting onto them. We hope that understanding extremal arcs and counting arcs of fixed size in general projective planes of order q will lead to interesting mathematics and perhaps will have similar applications in classification problems.

3. Algorithms

In [1], Glynn gives an inductive algorithm for counting $C_n(\Pi)$ in terms of the number of strong realizations in Π of all superfigurations on at most n points. The following statement is a version of Theorem 3.6 in [1].

Theorem 3.1. *Suppose $n \leq 13$. There exist polynomials $f(q), g_s(q)$ such that for any finite projective plane Π of order q, we have*

$$C_n(\Pi) = f(q) + \sum_s g_s(q) A_s(\Pi), \tag{*}$$

where the sum is taken over all superfigurations s with at most n points.

We call the polynomial $g_s(q)$ the *coefficient of influence* of the superfiguration s, since it measures the degree to which s is relevant in the n-arcs formula. The 13 in this statement comes from the fact that the smallest superfiguration is the Fano plane, which has 7 points, so the corresponding formula for $n = 14$ could involve a quadratic term in $A_7(\Pi)$, related to the number of strong realizations of two disjoint copies of the Fano plane in Π.

In order to prove Theorem 1.12 we return to Glynn's original algorithm from [1]. To our knowledge, this is the first time that the algorithm has actually been implemented to find new enumerative formulas, rather than just used as a theoretical tool to prove that formulas of a certain type exist. The implementation we describe has the potential to give analogues of Theorem 1.12, computing $C_n(\Pi)$ for larger values of n.

The algorithm used to arrive at Theorems 1.3 and 1.4 was first described in [1], and was further clarified in [6]. The latter form of the algorithm was a central component of Iampolskaia, Skorobogatov, and Sorokin's formula for counting 9-arcs in the projective plane over a finite field. We here present an exposition of the algorithm and prove that it works, following [6] and [4]; then we discuss how we altered the algorithm in order to achieve a manageable runtime.

Definition 3.2. A *boolean n-function* is a function taking subsets of $\{1, 2, 3, ..., n\}$ to {0,1}. Two boolean n-functions f and g are *isomorphic* if there is a permutation i of $\{1, ..., n\}$ so that $g = f \circ i$.

Particularly important is that f can be thought of as *labeled*: for instance, we must distinguish the boolean 2-function which just sends $\{1\}$ to 1 from the isomorphic boolean 2-function which just sends $\{2\}$ to 1. We also note that the boolean n-functions are in 1-1 correspondence with the power set of the power set of $\{1, ..., n\}$.

Definition 3.3. For boolean n-functions f and g, we say $f \geq g$ if $f(S) \geq g(S)$ for all $S \subseteq \{1, 2, ..., n\}$. Notice that $\geq$ is a partial order on the set of boolean n-functions.

Example 3.4. Suppose $f : 2^{\{1,...,7\}} \to \{0, 1\}$ sends just the sets {1,3,4,5}, {4,5,6} to 1 and every other set to 0. Then f is a boolean 7-function. If we let $f' : 2^{\{1,...,7\}} \to \{0, 1\}$ be the function that sends just the sets {1,3,4} and {4,5,6} to 1 and all other sets to 0, then we have $f \not\geq f'$.

We are working towards an abstraction of the axioms of geometry, in which the elements of $\{1, ..., n\}$ are points in some geometric setting, and the function f is an indicator function which evaluates to 1 for those subsets of {1,...,n} which are collinear. To make this precise:

Definition 3.5. Call a boolean n-function f a *linear space function* if it satisfies:

(1) If $f(I) = 1$, then $f(J) = 1$ for all $J \subseteq I$.

(2) If $\#(I) \leq 2$, then $f(I) = 1$.

(3) If $f(I) = f(J) = 1$ and $\#(I \cap J) \geq 2$, then $f(I \cup J) = 1$.

If f does not satisfy these laws, we call f *pathological.*

In words, these laws mean: subsets of lines are lines; any set of $0, 1$, or 2 points qualifies as a line; and if two lines intersect in at least two points, then the union of the sets is a line.

Thus, linear space functions capture the notion of collinearity which is common to affine and projective geometry. Indeed, finite affine and projective spaces of any dimension can be thought of as being just the data of some linear space function. Moreover, any subset of such a plane inherits collinearity relations from the plane, so it can also be encoded by some linear space function. Thus, a linear space function f defines a *linear space* with points $\{1, 2, ..., n\}$ and sets of collinear points defined by $f^{-1}(1)$. We can think as of a linear space as an isomorphism class of linear space functions.

For the rest of the section, fix some projective plane Π of order q.

Definition 3.6. Suppose some n-tuple S of distinct points labeled $1, ..., n$ in Π is such the sets which a given linear space function f sends to are exactly the collinear subsets of S. Such a tuple is called a *strong realization* of f.

Definition 3.7. Suppose some n-tuple S of distinct points labeled $1, ..., n$ in Π is such the sets which a given linear space function f sends to are among the collinear subsets of S. Such a tuple is called a *weak realization* of f.

Consider the boolean n-function a for which $f^{-1}(1)$ is the set of subsets of $\{1, ..., n\}$ of size 0, 1, or 2. Then a is a boolean n-function, and is also a linear space function. Every tuple of n distinct points in Π is a weak realization of a. A strong realization of a is an n-arc. Therefore, the goal of the algorithm is to calculate $m_q(a)$. We will do so indirectly, by examining weak realizations and working backwards.

Definition 3.8. For any boolean n-function f, define

$$n_q(f) = \sum_{g \geq f} m_q(g).$$

If f is a linear space function, then this number is the same as the number of weak realizations of f in Π. And if f is pathological, then $n_q(f)$ is still defined, although its interpretation in terms of weak realizations is less clear.

Now we reproduce the method described in [6] to calculate $n_q(g)$ in terms of the n_q numbers for linear space functions on fewer points.

Definition 3.9. Suppose that f is a linear space function. A *full line* of f is a subset $S \subseteq \{1, ..., n\}$ of size 3 or greater, so that $f(S) = 1$ and for all T which contain S, we have $f(T) = 1$ only if $T = S$. In other words, there is no larger set of collinear points which includes S. We say the *index* of a point p of f is the number of full lines which include p.

Notice that we can completely describe a linear space function just by giving its full lines.

Lemma 3.10. *Suppose the linear space function f on n points has a point of index 0, 1, or 2. Then we may construct $n_q(f)$ as a polynomial in q and the values $m_q(g)$, where g ranges over the linear space functions of $n - 1$ points. Further, $n_q(f)$ is linear in the $m_q(g)$.*

Proof. Let f be as stated. Without loss of generality, say point n has index 0, 1, or 2. Define f' to be the linear space function which corresponds to the configuration of just the points $1, 2, ..., n-1$, inheriting collinearity data from f. We count weak realizations of f as follows: each one is a strong realization of some $g \geq f'$, together with point n. (In particular, notice that we do not need to range over choices of point n; just choices of g given a fixed n.) So we shall count, for each g, the number of ways to add point n such that the result is a weak realization of f. We let $\mu(g, f)$ denote the number of ways to add point n to g to get a weak realization of f. We thus get an equation

$$n_q(f) = \sum_{g \geq f'} m_q(g)\mu(g, f).$$

Then, it is enough to give an method which finds μ for any pair g, f as a polynomial in q. This method, described in [6], is subtle in the sense that two isomorphic boolean n-functions g, g' do not necessarily satisfy $\mu(g, f) = \mu(g', f)$.

If the index of n is 0, then n may be placed anywhere in the plane not already occupied by a point of g; thus $\mu(g, f) = q^2 + q + 1 - (n - 1)$.

If the index of n is 1, then let L be the line of f on which n must lie. Then any point on L not already occupied is a valid choice for n, so $\mu(g, f) = q + 1 - |L|$.

If the index of n is 2 (that is, distinct lines L_1 and L_2 of f contain n), then there are multiple cases. If in g, the lines L_1 and L_2 intersect in two or more points, then they are the same full line L, so $\mu(g, f) = q+1-|L|$. If in g, the lines L_1 and L_2 intersect in exactly one point of g, then the only possible spot to place n is already filled, so $\mu(g, f) = 0$. Finally, if L_1 and L_2 do not intersect, then there is a unique place to put n, so $\mu(g, f) = 1$. □

Lemma 3.10 indicates the reason why Glynn's formula is not just a polynomial in q. Since the method for inferring the value of n_q only applies to those linear space functions with a point of index 0, 1, or 2, the algorithm cannot inductively find n_q for linear space functions with all points of index at least 3. In other words, the algorithm cannot determine n_q for superfigurations.

Now the following algorithm will inductively express each m_q and n_q in terms of just the values $m_q(f)$ for superfigurations f.

Algorithm 3.11. (1) Find m_q and n_q for the linear space functions on 1 point.

(2) Assume that we have m_q and n_q for all linear space functions on k points.
(3) Use Lemma 3.10 to find n_q for every linear space function on $k + 1$ points.
(4) Assume that f is a linear space function on $k + 1$ points, and we assume that we know $m_q(g)$ for all linear space functions $g > f$. Then calculate $m_q(f)$, by writing
$$m_q(f) = n_q(f) - \sum_{g>f} m_q(g).$$
(5) Repeat the previous step until we have calculated $m_q(f)$ for all f on $k + 1$ points.
(6) Continue by induction until $k = n$.

Thus, by running this algorithm, we can express the number of strong or weak realizations of any n-point linear space function L in a projective plane Π as
$$f(q) + \sum_{s \in S} g_s(q)A_s(q)$$
where S is the set of superfigurations on up to n points, $A_s(q)$ is the number of realizations of superfiguration s in Π, and $f(q)$ and the $g_s(q)$ are polynomials in q.

We implemented this algorithm in Sage. Running the algorithm up to 9 points, which takes several minutes of computation time, gives us the formula for 9-arcs claimed in Theorem 1.12.

4. Counting larger arcs

Algorithm 3.11 also computes the formula for the count of n arcs for $9 < n \leq 13$. In particular, the formula for 10-arcs in general projective planes is now within reach. However, we run into problems due to the complexity of the algorithm, whose runtime is roughly proportional to the square of the number of linear space functions on n points.

The complete list of n-point linear spaces can be determined by computing the list of hypergraphs on n vertices under the constraints that the minimum set size is 3 and the intersection of any two sets is of size at most 1. To restrict attention to superfigurations, we impose the additional condition that the minimum vertex degree is 3. For $n \leq 11$, McKay's *Nauty* software can quickly compute all such hypergraphs up to isomorphism [11].

Counts of Linear Spaces

n	7	8	9	10	11	12
Linear spaces on n points	24	69	384	5250	232929	28872973
Superfigurations	1	1	10	151	16234	>179000

The fast growth of these figures indicates the increasing difficulty of applying the algorithm. In particular, we found that the prohibitively high runtime comes from of the difficulty of calculating so many values of $n_q(s)$. We now present a new variant of Glynn's algorithm which partially circumvents this problem.

Recall that the number of weak realizations of a n-arc is given by

$$n_q(a) = \sum_{g \geq a} m_q(g)$$

where g ranges over all linear space functions on n points. We may therefore express the strong realizations of the n-arc linear space function as

$$\begin{aligned} m_q(a) &= n_q(a) + \sum_{g>a}(-1)m_q(g) \\ &= n_q(a) + \sum_{g>a}(-1)m_q(g) + \sum_{s>a}(-1)m_q(s), \end{aligned}$$

where the first sum ranges over linear space functions which are *not* superfigurations, and the second sum ranges over superfigurations only.

Choose a linear space function g which is minimal with respect to $>$ among the index set of the first sum. Apply the substitution

$$m_q(g) = n_q(g) - \sum_{h>g} m_q(h).$$

This eliminates the $m_q(g)$ term from our formula, leaving only terms $m_q(h)$ for $h > g$. Thus, by repeated applications of this substitution to the minimal non-superfiguration in the formula, we will arrive at a formula of the form

$$m_q(a) = \sum_{g} k(g)n_q(g) + \sum_{s} l(s)m_q(s),$$

where the $k(g)$ and $l(s)$ are integers.

From this point, we could calculate the Glynn formula for n-arcs by replacing each instance of $n_q(g)$ by a polynomial in q and $m_q(t)$ for supefigurations t on $n-1$ or fewer points. But this substitution does not affect the coefficients of the $m_q(s)$ for superfigurations s on n points. Therefore, the already-calculated values $l(s)$ are the coefficients of influence for the n-point superfigurations.

We also have proven as a consequence the following lemma.

Lemma 4.1. *The coefficient of influence of each n-point superfiguration in the Glynn formula for n-arcs is a constant.*

Let us consider the implications when $n = 10$. Of the 163 superfigurations on up to 10 points, 151 are on exactly 10 points. Therefore, the algorithm just described calculates 151 of the 163 coefficients of influence without finding any values of n_q. The following table states the coefficients of influence for each of the superfigurations $10_{13}, 10_{14}, ..., 10_{163}$, whose definitions may be found on our website [12].

Coefficients of Influence in the Formula for $C_{10}(\Pi)$

s	$g_s(q)$	s	$g_s(q)$	s	$g_s(q)$	s	$g_s(q)$	s	$g_s(q)$	s	$g_s(q)$
10_{13}	27	10_{39}	-3	10_{65}	0	10_{91}	-1	10_{117}	2	10_{143}	-2
10_{14}	27	10_{40}	-3	10_{66}	0	10_{92}	-1	10_{118}	-1	10_{144}	-2
10_{15}	27	10_{41}	-3	10_{67}	0	10_{93}	-1	10_{119}	1	10_{145}	-1
10_{16}	1	10_{42}	-3	10_{68}	0	10_{94}	-1	10_{120}	-1	10_{146}	-2
10_{17}	1	10_{43}	-3	10_{69}	0	10_{95}	-1	10_{121}	-1	10_{147}	-2
10_{18}	1	10_{44}	-3	10_{70}	0	10_{96}	-1	10_{122}	-1	10_{148}	-2
10_{19}	1	10_{45}	-3	10_{71}	0	10_{97}	-1	10_{123}	-1	10_{149}	-2
10_{20}	1	10_{46}	-3	10_{72}	-1	10_{98}	-1	10_{124}	2	10_{150}	1
10_{21}	1	10_{47}	9	10_{73}	-1	10_{99}	3	10_{125}	-1	10_{151}	9
10_{22}	1	10_{48}	9	10_{74}	-1	10_{100}	3	10_{126}	-1	10_{152}	9
10_{23}	1	10_{49}	9	10_{75}	-1	10_{101}	3	10_{127}	2	10_{153}	2
10_{24}	1	10_{50}	9	10_{76}	-1	10_{102}	3	10_{128}	3	10_{154}	1
10_{25}	1	10_{51}	9	10_{77}	-1	10_{103}	3	10_{129}	5	10_{155}	2
10_{26}	-2	10_{52}	4	10_{78}	-1	10_{104}	-1	10_{130}	4	10_{156}	4
10_{27}	-2	10_{53}	4	10_{79}	-1	10_{105}	0	10_{131}	12	10_{157}	5
10_{28}	-2	10_{54}	4	10_{80}	-1	10_{106}	-1	10_{132}	0	10_{158}	6
10_{29}	-2	10_{55}	6	10_{81}	-1	10_{107}	-1	10_{133}	0	10_{159}	18
10_{30}	-2	10_{56}	6	10_{82}	-1	10_{108}	0	10_{134}	0	10_{160}	1
10_{31}	16	10_{57}	6	10_{83}	-1	10_{109}	-1	10_{135}	0	10_{161}	1
10_{32}	-2	10_{58}	6	10_{84}	-1	10_{110}	1	10_{136}	0	10_{162}	10
10_{33}	-2	10_{59}	19	10_{85}	-1	10_{111}	2	10_{137}	-1	10_{163}	-6
10_{34}	-2	10_{60}	-8	10_{86}	-1	10_{112}	1	10_{138}	-1		
10_{35}	-3	10_{61}	19	10_{87}	-1	10_{113}	2	10_{139}	0		
10_{36}	-3	10_{62}	19	10_{88}	-1	10_{114}	2	10_{140}	-1		
10_{37}	-3	10_{63}	-12	10_{89}	-1	10_{115}	1	10_{141}	-2		
10_{38}	-3	10_{64}	-8	10_{90}	-1	10_{116}	1	10_{142}	-1		

These values were obtained by running an implementation of our algorithm in Sage on a High Performance Computing cluster at Yale University. Approximately twenty processors running in parallel completed the algorithm in several hours.

5. Further Questions

The formulas given here, when taken at face value, suggest a conjecture that the arc-counting formulas $C_n(q)$ are quasipolynomial in q for all n. But when $n = 10$, there is strong heuristic evidence that the pattern will break. Whereas there are only 12 superfigurations on up to 9 points, the tenth point forces us to consider an additional 151 superfigurations, including the classical Desargues configuration. We intend to demonstrate in the forthcoming paper [5] that the counting formulas for several of these superfigurations in fact cannot be expressed as a quasipolynomial, which in turn would imply that 10-arcs cannot be counted by means of quasipolynomial formulas. The kinds of functions needed to count larger arcs are worthy of further study, for they demonstrate that the apparent simplicity of incidence geometry belies a subtle reliance on the algebraic structure of the plane.

Acknowledgements

We thank the mathematics department at Yale University and the Summer Undergraduate Research at Yale (SUMRY) program for providing us with the opportunity to conduct this research. SUMRY is supported in part by NSF grant CAREER DMS-1149054. The first author is supported by NSA Young Investigator Grant H98230-16-10305 and by an AMS-Simons Travel Grant.

We would also like to extend our gratitude to Sam Payne and José González for helping to organize SUMRY. We thank the Yale Center for Research Computing for allowing us to use the High Performance Computing resources.

References

[1] D. Glynn, Rings of geometries. II. J. Combin. Theory Ser. A 49 (1988), no. 1, 26-66.
[2] Grunbaum.
[3] Hirschfeld. Thas.
[4] A. Iampolskaia, A. N. Skorobogatov, and E. Sorokin, Formula for the number of $[9, 3]$ MDS codes. IEEE Trans. Inform. Theory 41 (1995), no. 6, part 1, 1667-1671.
[5] N. Kaplan, S. Kimport, R. Lawrence, L. Peilen, and M. Weinreich, The number of 10-arcs in the projective plane is not a quasipolynomial. In preparation, (2016).
[6] R. Rolland and A. N. Skorobogatov, Dénombrement des configurations dans le plan projectif, in Proc. Arithmetic, Geometry, and Coding Theory (Luminy, France 1993 W. de Gruyter), to appear.
[7] A. N. Skorobogatov, Linear codes, strata of Grassmannians, and the problems of Segre, in Lecture Notes in Mathematics, no. 1518 (1992), Coding Theory and Algebraic Geometry 1991, H. Stichtenoth, M. A. Tsfasman, Eds., pp. 210-223.
[8] B. Sturmfels, Computational algebraic geometry of projective configurations. J. Symbolic Computation (1991) 11, 595-618.
[9] L. M. Batten and A. Beutelspacher, The theory of finite linear spaces: combinatorics of points and lines. Cambridge University Press, 2009.
[10] M. Tait, On a problem of Neumann, 2015. (arXiv:1509.06587)

[11] B. D. McKay and A. Piperno, Practical Graph Isomorphism, II, J. Symbolic Computation (2013) 60 94-112.
[12] R. Lawrence and M. Weinreich, Counting 10-arcs in the Projective Plane over a Finite Field. Web. http://rachellawrence.github.io/TenArcs/

Department of Mathematics, University of California, Irvine, CA 92697

E-mail address: nckaplan@math.uci.edu

Department of Mathematics, Stanford University, Stanford, CA 94305

E-mail address: skimport@stanford.edu

Department of Mathematics, Yale University, New Haven, CT 06511

E-mail address: rlaw@aya.yale.edu

Department of Mathematics, Yale University, New Haven, CT 06511

E-mail address: luke.peilen@yale.edu

Department of Mathematics, Yale University, New Haven, CT 06511

E-mail address: maxhweinreich@gmail.com