AN ALGORITHM TO COUNT THE NUMBER OF CAPS IN $\mathbb{P}^3(\mathbb{F}_q)$

KELLY ISHAM

Abstract. An $n$-cap in $k$-dimensional projective space is a set of $n$ points so that no three lie on a line. In this note, we provide an algorithm to count the number of $n$-caps in $\mathbb{P}^3(\mathbb{F}_q)$, which follows from our recent paper [9]. We then give exact formulas for the number of $n$-caps when $n \leq 7$. The formulas are polynomial in $q$ when $n \leq 6$ and quasipolynomial in $q$ when $n = 7$.

1. Introduction

An $n$-arc in the projective plane is a set of $n$ points so that no three lie on a line. We can generalize this notion to higher-dimensional space in a couple different ways.

First, we define an $n$-arc in $k$-dimensional space to be a set of $n$ points so that no $k + 1$ lie on a hyperplane. This object is studied across mathematics in areas including coding theory, finite geometry, algebraic geometry, and linear algebra. In particular, $n$-arcs in $\mathbb{P}^k(\mathbb{F}_q)$ are closely related to MDS codes, which are codes that achieve the Singleton bound. One important conjecture in coding theory is the MDS conjecture, which, rewritten in the language of arcs, states that if $k \leq q$, the size of the largest arc in $\mathbb{P}^k(\mathbb{F}_q)$ is $q + 2$ if $q$ is even and $k \in \{2, q - 2\}$, and is $q + 1$ otherwise. Researchers from various areas of mathematics have made substantial progress toward proving the MDS conjecture; see [1] for a survey of large arcs.

Second, we define an $n$-cap in $k$-dimensional projective space to be a set of $n$ points so that no three lie on a line. Notice that the definition of $n$-arc and $n$-cap coincide when $k = 2$. In this paper, we will consider $n$-caps.

In 1936, Erdős and Turán [5] conjectured that sets of integers with positive density must have arbitrarily long arithmetic progressions. Famously, Szemerédi [14] proved their conjecture in 1975 using combinatorial methods. The Cap Set problem is an analogue, which asks for the largest $n$ so that an $n$-cap exists in $\mathbb{P}^3(\mathbb{F}_q)$ (e.g. the largest set which does not contain a three-term arithmetic progression). More generally, we can ask for the largest $n$ so that an $n$-cap exists in $\mathbb{P}^k$. Substantial progress on these problems has been made, see [12] for a summary of known results and methods. The best known upper bounds are due to Ellenberg and Gijswijt [4] in a groundbreaking paper from 2017.

Some results about maximal caps are also known in $\mathbb{P}^k(\mathbb{F}_q)$. It is known that the maximal size of a cap in $\mathbb{P}^3(\mathbb{F}_q)$ is $q^2 + 1$ unless $q = 2$. Further, in 1955, Barlotti [2] and Panella [11] independently showed that if $q$ is odd or $q = 4$, there is a one-to-one correspondence between elliptic quadrics and maximal caps in $\mathbb{P}^3(\mathbb{F}_q)$. If $q = 2^r$ and $r$ is odd, there exist maximal caps which are not elliptic quadrics [15]. Much less is known when $k > 3$. For a recent survey on caps in $\mathbb{P}^k(\mathbb{F}_q)$, see [7].

In this paper, we consider the counting version of the Cap Set problem.

**Question 1.1.** Let $c_{n,k}(q)$ denote the number of $n$-caps in $\mathbb{P}^k(\mathbb{F}_q)$. For fixed $k$ and $n$, what is $c_{n,k}(q)$ as a function of $q$?
It may be difficult to give an explicit formula for \( c_{n,k}(q) \), so we can ask to classify the function. We use the hierarchy: polynomial, quasipolynomial, and non-quasipolynomial. A function \( f(q) \in \mathbb{Q}[q] \) is quasipolynomial if there exist finitely many polynomials \( g_0(q), g_1(q), \ldots, g_{N-1}(q) \in \mathbb{Q}[q] \) so that \( f(q) = g_i(q) \) whenever \( q \equiv i \pmod{N} \).

When \( k = 2 \), caps are equivalent to arcs. In this case, \( c_{n,2}(q) \) is known explicitly when \( n \leq 9 \) \([3, 8, 10]\). This function is polynomial in \( q \) when \( n \leq 6 \) and is quasipolynomial in \( q \) when \( n \in \{7, 8, 9\} \). In recent work joint with several authors, we show that \( c_{10,2}(q) \) is non-quasipolynomial in \( q \), though we are unable to give an explicit formula. Most likely these results together give a full classification of \( c_{n,2}(q) \). In fact, in the upcoming joint paper, we conjecture that \( c_{n,2}(q) \) will continue to be non-quasipolynomial for all \( n \geq 11 \).

One way to count \( n \)-caps in \( \mathbb{P}^2(\mathbb{F}_q) \) is to count the opposite — namely sets of \( n \) points so that at least one subset of three points is contained in a line. Glynn provides a formula \([6, \text{Theorem 3.16}]\) that shows we can do much better than this naive inclusion-exclusion count by studying special combinatorial objects that contain several three-point lines.

2. Connection to arcs in \( \mathbb{P}^3(\mathbb{F}_q) \)

Motivated by Glynn’s ideas, in a recent paper \([9]\) we provide a formula to count \( n \)-arcs in \( \mathbb{P}^3(\mathbb{F}_q) \) in terms of a small number of combinatorial objects. In this paper, we will derive a formula to count \( c_{n,3}(q) \) that follows from the formula to count \( n \)-arcs in \( \mathbb{P}^3(\mathbb{F}_q) \). In order to state the algorithm, we first need several definitions from \([9]\).

While our main results hold over \( \mathbb{P}^3(\mathbb{F}_q) \) — the unique finite 3-dimensional projective space of order \( q \) — we can first consider caps in a more general space.

**Definition 2.1.** \([3]\) A linear space \((\mathcal{P}, \mathcal{L})\) is a pair of sets where \( \mathcal{P} \) denotes a set of points and \( \mathcal{L} \) denotes a set of lines satisfying the following properties:

1. Every line in \( \mathcal{L} \) is a subset of \( \mathcal{P} \).
2. Any two distinct points belong to exactly one line in \( \mathcal{L} \).
3. Every line in \( \mathcal{L} \) contains at least 2 points.

Let \( 2^X \) denote the power set of the set \( X \).

**Definition 2.2.** \([9]\) Definition 1.8) A planar space is a triple of sets \((\mathcal{P}, \mathcal{L}, \mathcal{H})\) where \( \mathcal{P} \) is the set of points, \( \mathcal{L} \) is the set of lines, and \( \mathcal{H} \) is the set of planes such that

1. \( \mathcal{L}, \mathcal{H} \subseteq 2^{\mathcal{P}} \).
2. \((\mathcal{P}, \mathcal{L})\) is a linear space.
3. Any three distinct non-collinear points lie on a unique plane.

Linear and planar spaces are very general. For example, affine, projective, and Euclidean spaces satisfy these axioms. In order to count arcs in \( \mathbb{P}^3(\mathbb{F}_q) \), we define maps that embed the combinatorial sets of points, lines, and planes into \( \mathbb{P}^3(\mathbb{F}_q) \).

A strong realization of a planar space \( f = (\mathcal{P}, \mathcal{L}, \mathcal{H}) \) is an injective mapping \( \sigma : \mathcal{P} \hookrightarrow \mathbb{P}^3(\mathbb{F}_q) \) so that

- \( \sigma(L) \) is contained in a line of \( \mathbb{P}^3(\mathbb{F}_q) \) if and only if \( L \subseteq \mathcal{L} \) and
- \( \sigma(H) \) is contained in a plane of \( \mathbb{P}^3(\mathbb{F}_q) \) if and only if \( H \subseteq \mathcal{H} \).

The first condition states that \( \sigma \) preserves lines and the second states that \( \sigma \) preserves planes.

We let \( A_f(4, q) \) denote the number of strong realizations of \( f \) to remain consistent with the notation of \([9]\) where \( \sigma \) maps \( \mathcal{P} \) into \( \mathbb{P}^{k-1}(\mathbb{F}_q) \) for \( k = 4 \).
Two planar spaces $f = (P_1, L_1, H_1)$ and $g = (P_2, L_2, H_2)$ are isomorphic if there exists a bijective mapping $\rho : P_1 \to P_2$ that preserves lines and planes.

In a linear or planar space, every two points determine a line. In a planar space, every three non-collinear points determine a plane. Thus we refer to a line containing at least three points a full line and a plane containing at least four points a full plane. Each planar space is determined uniquely by its full lines and planes, so we drop the word full and ignore two-point lines and three-point planes when writing a planar space. We say a point has index $(i,j)$ if it lies on exactly $i$ (full) planes and $j$ (full) lines.

Example 2.3. Consider the planar space $f = (P, L, H)$ on the points $\{1, 2, 3, 4\}$ such that all four points lie on a common plane and the first three points lie on a line. The planar space can simply be given by $P = \{1, 2, 3, 4\}$, $L = \{\{1, 2, 3\}\}$, and $H = \{\{1, 2, 3, 4\}\}$. Implicitly the line $\{1, 2\}$ is contained in $L$ since every two points determine a line, but we suppress this information when writing the set $L$.

Notice that a cap can be defined as a planar space which contains no full lines. In order to count caps in $P_3(\mathbb{F}_q)$, we will count the number of strong realizations of planar spaces so that $L$ contains at least one full line, then use inclusion-exclusion. The objects remaining in the summation are special combinatorial objects called hyperfigurations.

Definition 2.4. [9, Definition 1.12] A hyperfiguration is a planar space on $n$ points such that for every point $P$, the index $(i,j)$ of $P$ satisfies at least one of the following:

1. $i \geq 4$
2. $j \geq 3$
3. $(i, j) = (3, 0)$.

Example 2.5. The planar space from Example 2.3 is not a hyperfiguration since the index of point 1 is $(1, 1)$.

This definition is not intuitive. The proof of Theorem 2.6 relies on a lemma [9, Lemma 2.3] that allows us to work inductively by removing a point of index $(i,j)$ from a planar space $f$. The lemma assumes that $(i,j)$ does not satisfy any of the three conditions given in the definition of hyperfiguration.

Theorem 2.6. [9, Theorem 1.13] There exist polynomials $p(q)$ and $p_h(q)$ in $\mathbb{Z}[q]$ such that

$$\#\{n\text{-arcs in } \mathbb{P}^3(\mathbb{F}_q)\} = p(q) + \sum_h p_h(q)A_h(4, q)$$

where the summation runs over all isomorphism classes $h$ of hyperfigurations on at most $n$ points. Moreover, there is an algorithm that produces $p(q)$ and $p_h(q)$ for each isomorphism class $h$.

While not explicitly stated in [9], the algorithm is inductive and produces a similar expression for $A_f(4, q)$ for every planar space $f$ on at most $n$ points. We state this stronger fact as a corollary.

Corollary 2.7. For any planar space $f$, there exist polynomials $r(q)$ and $r_h(q)$ in $\mathbb{Z}[q]$ such that

$$A_f(4, q) = r(q) + \sum_h r_h(q)A_h(4, q)$$
where the summation runs over all isomorphism classes \( h \) of hyperfigurations on at most \( n \) points. Moreover, there is an algorithm that produces \( r(q) \) and \( r_h(q) \) for each isomorphism class \( h \).

2.1. Main Results. We begin by introducing an algorithm to count \( c_{n,3}(q) \) in terms of a summation over all planar space functions on at most \( n \) points. Let \( \mathcal{N}_n \) be the set of all planar spaces on at most \( n \) points for which the largest size of a subset in \( \mathcal{L} \) is two, i.e., the planar spaces which do not contain a full line. We can understand \( c_{n,3}(q) \) by counting the strong realizations of each planar space \( f \in \mathcal{N}_n \), then summing. This gives the following algorithm.

**Algorithm 2.8.**

1. Use Corollary 2.7 to compute \( A_f(4,q) \) for each planar space \( f \) on at most \( n \) points.
2. Output \( c_{n,3}(q) = \sum_{f \in \mathcal{N}_n} A_f(4,q) \).

Applying Algorithm 2.8 leads to the following main theorem.

**Theorem 2.9.** There exist polynomials \( s(q) \) and \( s_h(q) \) so that

\[
c_{n,3}(q) = s(q) + \sum_h s_h(q)A_h(4,q)
\]

where the summation runs over all isomorphism classes \( h \) of hyperfigurations on at most \( n \) points. Moreover, Algorithm 2.8 produces \( s(q) \) and \( s_h(q) \) for each isomorphism class \( h \).

It is natural to wonder what the degrees of the polynomials \( s_h(q) \) are. In Propositions 3.1 and 3.2, we will show that for \( n \leq 7 \), the polynomials \( s_h(q) \) in \( c_{n,3}(q) \) are constants when \( h \) is a hyperfiguration on exactly \( n \) points. This is consistent with the findings in Kaplan, Kimport, Lawrence, Peilen, and Weinreich [10] about the coefficients of the function that counts arcs in \( \mathbb{P}^2(\mathbb{F}_q) \).

We implement this algorithm in Sage [13] to count \( n \)-caps in \( \mathbb{P}^3(\mathbb{F}_q) \) for small \( n \). We arrive at the following theorem.

**Theorem 2.10.** Let

\[
a(q) = \begin{cases} 1 & \text{if } q \equiv 2 \pmod{4} \\ 0 & \text{else} \end{cases}
\]

Then

\[
c_{3,3}(q) = (q^2 + q + 1)(q^2 + 1)(q + 1)^2q^3 \\
c_{4,3}(q) = (q^3 + q^2 - 2q + 1)(q^2 + q + 1)(q^2 + 1)(q + 1)^2q^3 \\
c_{5,3}(q) = (q^4 + 4q^3 + q^2 - 5q + 6)(q^2 + q + 1)(q^2 + 1)(q + 1)^2(q - 1)^2q^3
\]
\[ c_{6,3}(q) = \left(q^{11} + 5q^{10} - 5q^9 - 39q^8 + 35q^7 + 96q^6 - 210q^5 + 144q^4 + 169q^3 - 116q^2 \right. \\
- 260q + 240)(q^2 + q + 1)(q^2 + 1)(q + 1)(q - 1)q \\
c_{7,3}(q) = \left(q^{10} + 6q^9 - 13q^8 - 98q^7 + 148q^6 + 629q^5 - 1461q^4 - 686q^3 + 6462q^2 \\
- 11004q + 7470 - 30a(q)\right)(q^2 + q + 1)(q^2 + 1)(q + 1)^2(q - 1)^2q^3. \]

In particular, \(c_{n,3}(q)\) is a polynomial in \(q\) when \(n \leq 6\) and is a quasipolynomial in \(q\) when \(n = 7\).

3. Proof of Main Results

There are no hyperfigurations on \(n \leq 5\) points, so the Algorithm \(2.8\) produces the exact polynomial formulas for \(c_{n,3}(q)\) when \(n \leq 5\) given in Theorem \(2.10\).

We could count 4-caps and 5-caps combinatorially as well. For example, to count 4-caps, we need to count sets of four points so that no three are on a line. We can choose any two points, then pick a third point not on the line determined by those first two points. Lastly, we must pick a fourth point that does not lie on any of the three lines formed by the first three points. This gives the count

\[(q^3 + q^2 + q + 1)(q^3 + q^2 + q)(q^3 + q^2)(q^3 + q^2 - 2q + 1)\]

which matches the formula in Theorem \(2.10\) after factoring. We can count 5-caps similarly, but the count is much more intricate with cases depending on whether or not the first four points lie on a common plane.

There is exactly one hyperfiguration on 6 points, which has planes \{\{0, 1, 2, 3\}, \{0, 1, 2, 4\}, \{0, 1, 2, 5\}, \{0, 3, 4, 5\}, \{1, 3, 4, 5\}, \{2, 3, 4, 5\}\} and lines \{\{0, 1, 2\}, \{3, 4, 5\}\}. It is important to note that these lines must be skew in \(\mathbb{P}^3(\mathbb{F}_q)\). Algorithm \(2.8\) gives the following expression for \(c_{6,3}(q)\) in terms of a \(\mathbb{Z}[q]\)-linear combination of 1 and \(A_6(4, q)\), the number of strong realizations of this hyperfiguration on 6 points.

**Proposition 3.1.** We have

\[ c_{6,3}(q) = q^{18} + 6q^{17} + q^{16} - 39q^{15} - 20q^{14} + 76q^{13} - 80q^{12} + 79q^{11} + 132q^{10} + 115q^9 \\
- 128q^8 - 176q^7 - 133q^6 + 43q^5 + 207q^4 + 136q^3 + 20q^2 + 10A_6(4, q) - 240q. \]

In [9], we show that

\[ A_6(4, q) = (q^2 + q + 1)(q^2 + 1)(q + 1)^2(q - 1)^2q^6. \]

Plugging this into Proposition 3.1 and simplifying gives the count for \(c_{6,3}(q)\) in Theorem 2.10.

Finally, we consider 7-caps in \(\mathbb{P}^3(\mathbb{F}_q)\). There are 6 non-isomorphic hyperfigurations on 7 points, which we list below.
 Proposition 3.3. This error does not affect any results in [9].

We have

\[ A \] combination of 1, \( A_6(4, q) \), and \( A_{h_i}(4, q) \) for \( 1 \leq i \leq 6 \).

**Proposition 3.2.** We have

\[
c_{7,3}(q) = q^{21} + 7q^{20} - 7q^{19} - 112q^{18} - 28q^{17} + 637q^{16} - 119q^{15} - 1840q^{14} + 4676q^{13} - 4949q^{12} - 1436q^{11} + 4136q^{10} - 107q^9 + 7839q^8 - 5908q^7 - 2184q^6 - 4542q^5 - 3534q^4 + 7470q^3 + (70q^3 + 70q^2 - 560q + 280)A_6(4, q) - 105A_{h_5}(4, q) - 30A_{h_6}(4, q).
\]

In [9], we give formulas for the number of strong realizations of each hyperfiguration on seven points. There is a small error in the computation of \( A_{h_5}(4, q) \), which we correct now. This error does not affect any results in [9].

**Proposition 3.3.** [9] Proposition 3.8] Let

\[
a(q) = \begin{cases} 
1 & q \equiv 0 \pmod{2} \\
0 & q \equiv 1 \pmod{2} 
\end{cases}.
\]

The number of strong realizations for each hyperfiguration is given by
\[ A_{h_1}(4, q) = (1 - a(q)) \cdot |\text{PGL}_4(\mathbb{F}_q)| \]
\[ A_{h_2}(4, q) = a(q) \cdot |\text{PGL}_4(\mathbb{F}_q)| \]
\[ A_{h_3}(3, q) = (q - 2) \cdot |\text{PGL}_4(\mathbb{F}_q)| \]
\[ A_{h_4}(4, q) = |\text{PGL}_4(\mathbb{F}_q)| \]
\[ A_{h_5}(4, q) = (q^2 + q + 1)(q^2 + 1)(q + 1)^2(q - 1)^2(q - 2)q^6 \]
\[ A_{h_6}(4, q) = a(q) \cdot (q^3 + q^2 + q + 1) \cdot |\text{PGL}_3(\mathbb{F}_q)|. \]

Proof. We give a corrected proof of the calculation for \( A_{h_6}(4, q) \). Notice that all points in a strong realization of \( h_6 \) lie in the same plane. Fix a particular plane \( H \) in \( \mathbb{P}^3(\mathbb{F}_q) \). Each strong realization of \( h_6 \) projects to a copy of the Fano plane in \( H \). Therefore we can count strong realizations of \( h_6 \) in \( H \) by counting strong realizations of the Fano plane in \( \mathbb{P}^2(\mathbb{F}_q) \). This is well-known to be \( |\text{PGL}_3(\mathbb{F}_q)| \cdot a(q) \). We then multiply by the number of planes in \( \mathbb{P}^3(\mathbb{F}_q) \). \( \square \)

Proof of Theorem 2.10. Apply Proposition 3.3 to Proposition 3.2 and simplify. \( \square \)

4. Future Directions

Our algorithm from [9] that counts arcs in \( \mathbb{P}^3(\mathbb{F}_q) \) leads to an algorithm that counts caps in \( \mathbb{P}^3(\mathbb{F}_q) \). In [9], we also give a road map to count arcs in higher-dimensional projective space. It would be interesting to develop similar algorithms to count arcs and caps in \( \mathbb{P}^4(\mathbb{F}_q) \). This will likely be computationally intensive as the analogue of a planar space would now contain points, lines, planes, and hyperplanes and the interactions between these objects are more complicated.

It is interesting to note that the non-polynomial behavior of \( c_{7,3}(q) \) comes from counting strong realizations of \( h_6 \). Each strong realization of \( h_6 \) projects to a strong realization of the Fano plane in \( \mathbb{P}^2(\mathbb{F}_q) \). Therefore we can say that the non-polynomial behavior of \( c_{7,3}(q) \) comes from the Fano plane. It is unlikely that for large \( n \), all hyperfigurations contributing to the non-polynomial behavior of \( c_{n,3}(q) \) will project to superfigurations in the projective plane. We pose the question: what is the smallest \( n \) for which a hyperfiguration exists with all points not in a common plane?

In a future paper, we will study 8-arcs and 8-caps in \( \mathbb{P}^3(\mathbb{F}_q) \). It is likely that this formula is also quasipolynomial in \( q \). Since the number of 10-arcs in \( \mathbb{P}^2(\mathbb{F}_q) \) is non-quasipolynomial, it seems extremely likely that \( c_{10,3}(q) \) is non-quasipolynomial as well. The remaining open case is whether or not \( c_{9,3}(q) \) is quasipolynomial.

5. Acknowledgments

The author thanks Nathan Kaplan and Max Weinreich for helpful comments and suggestions.

References

[1] S. Ball and M. Lavrauw. Arcs in finite projective spaces. EMS Surv. Math. Sci., 6(1):133–172, 2020.
[2] A. Barlotti. Un’estensione del teorema di segre-kustaanheimo. Bollettino dell’Unione Matematica Italiana, 10:498–506, 1955.
L. Batten and A. Beutelspacher. *The theory of finite linear spaces - combinatorics of points and lines.* Cambridge University Press, 1993.

J. S. Ellenberg and D. Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math.*, 185(1):339–343, 2017.

P. Erdős and P. Turán. On some sequences of integers. *J. London Math. Soc.*, s1-11(4):261–264, 1936, https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/jlms/s1-11.4.261.

D. G. Glynn. Rings of geometries II. *J. Combin. Theory Ser. A.*, 49(1):26 – 66, 1988.

J. W. P. Hirschfeld and J. Thas. Open problems in finite projective spaces. *Finite Fields Appl.*, 32:44–81, 03 2015.

A. V. Iampolskaia, A. N. Skorobogatov, and E. A. Sorokin. Formula for the number of [9,3] MDS codes. *IEEE Trans. Inform. Theory*, 41(6):1667–1671, 1995.

K. Isham. An algorithm for counting arcs in higher-dimensional projective space. *Finite Fields Appl.*, 80:102006, 2022.

N. Kaplan, S. Kimport, R. Lawrence, L. Peilen, and M. Weinreich. Counting arcs in the projective plane via Glynn’s algorithm. *J. Geom.*, 12 2016.

G. Panella. Caratterizzazione delle quadriche di uno spazio (tridimensionale) lineare sopra un corpo finito. *Bollettino Della Unione Matematica Italiana*, 10:507–513, 1955.

S. Peluse. Recent progress on bounds for sets with no three terms in arithmetic progression, 2022, https://arxiv.org/pdf/2206.10037.pdf arXiv:2206.10037.

W. Stein et al. *Sage Mathematics Software (Version 8.9).* The Sage Development Team. http://www.sagemath.org.

E. Szemerédi. On sets of integers containing $k$ elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

J. Tits. Ovoides et groupes de Suzuki. *Arch. Math.*, 13:187–198, 1962.

Department of Mathematics, 214 McGregory Hall, Colgate University, Hamilton, NY

Email address: kisham@colgate.edu