Exponentially Larger Affine and Projective Caps

Christian Elsholtz and Gabriel F. Lipnik

Abstract

In spite of a recent breakthrough on upper bounds of the size of cap sets (by Croot, Lev and Pach (2017) and Ellenberg and Gijswijt (2017)), the classical cap set constructions had not been affected. In this work, we introduce a very different method of construction for caps in all affine spaces with odd prime modulus p. Moreover, we show that for all primes $p \equiv 5 \bmod 6$ with $p \leq 41$, the new construction leads to an exponentially larger growth of the affine and projective caps in $\mathrm{AG}(n,p)$ and $\mathrm{PG}(n,p)$. For example, when $p = 23$, the existence of caps with growth $(8.0875\ldots)^n$ follows from a three-dimensional example of Bose (1947), and the only improvement had been to $(8.0901\ldots)^n$ by Edel (2004), based on a six-dimensional example. We improve this lower bound to $(9 - o(1))^n$.

1 Introduction and Overview

The study of large point sets without three points on any line, in affine or projective spaces, is a classical topic in geometry, and more recently also in additive combinatorics. An introduction and some general information on these sets called *caps*, in particular from a geometric point of view, can be found in several chapters of Hirschfeld's three volumes on projective geometries over finite fields [24, 25, 27], in a survey by Hirschfeld and Storme [26], in relevant papers by Bierbrauer and Edel, e.g. [7, 8, 16, 17], and on Edel's website [15].

A lot of results study the size of complete caps (i.e., caps which cannot be extended) in a fixed dimension over a fixed finite field, see e.g. [1, 2, 3]. It is even an open problem to characterize complete caps in dimension 3 over $\mathbb{F}_q$; see for example Hirschfeld and Thas [23]. Numerous papers give alternative constructions for non-equivalent caps; see e.g. Kroll and Vincenti [28].

An important breakthrough [5, 12, 18, 22] has recently lead to greatly improved upper bounds for the largest possible size of these sets in the affine geometry $\mathrm{AG}(n,p)$.

In this paper, we improve longstanding *lower bounds* for caps when $p \in \{11, 17, 23, 29, 41\}$. In fact, the improvement is actually an *exponential* improvement (in the standard terminology, see [12]). It might be clear that this does not come from a refinement of previous methods but from an entirely different approach.

Christian Elsholtz elsholtz@math.tugraz.at, Institute of Analysis and Number Theory, Kopernikusgasse 24/II, Graz University of Technology, 8010 Graz, Austria

Gabriel F. Lipnik math@gabriellipnik.at, Institute of Analysis and Number Theory, Kopernikusgasse 24/II, Graz University of Technology, 8010 Graz, Austria

Support The authors acknowledge the support of the Austrian Science Fund (FWF): W 1230 and I 4945-N.

2020 Mathematics Subject Classification 51E20, 51E22, 05B25

Key words caps, affine space, projective space, progression-free sets

Previous cap set constructions are (recursively) based on a product construction from good examples in low dimensions, which we think of as a "local" approach. See for example [7, 8, 14, 16, 17, 32]. In contrast we construct a set of vectors with certain constraints with regards to the occurring digits, similar to a construction by Salem and Spencer [36] in the integer case, and we think of this construction as a "global" approach.

In this paper, we describe a new type of cap construction in the affine space $\mathrm{AG}(n, p)$ over the field $\mathbb{Z}_p$ with $p \geq 5$, (and therefore also in the corresponding projective space $\mathrm{PG}(n, p)$) that actually works for all dimensions over $\mathbb{Z}_p$. In its most basic case this includes the simple cap construction $\{0, 1\}^n \subset \mathbb{F}_3^n$. This has been generalized previously to certain product constructions. In this paper we generalize this in a novel way to combine well chosen digit sets with certain conditions. It will be apparent from the construction below that for given n and p, there are usually many non-equivalent caps; see Section 5. For some primes we can even achieve new records of the largest known caps and we will concentrate on this aspect. These appear to be the first improvements over the results of Bose, Bierbrauer and Edel; for details see below.

In the following, we consider the affine space $\mathrm{AG}(n, p)$, where $n \in \mathbb{N}$ is the dimension and p is a prime (and thus, $\mathrm{AG}(n, p) = ((\mathbb{Z}/p\mathbb{Z})^n, +)$, or $\mathbb{Z}_p^n$ for brevity). An affine cap S is a subset of $\mathbb{Z}_p^n$ such that no three points in S are collinear, i.e., for any three pairwise distinct points x, y, $z \in S$, the vectors $y - x$ and $z - x$ are linearly independent over $\mathbb{Z}_p$. This condition is equivalent to the fact that for any $(a, b, c) \in \mathbb{Z}_p^3 \setminus \{(0, 0, 0)\}$ with $a + b + c = 0$, one also has $ax + by + cz \neq 0$.

Projective caps are analogue sets in the projective space $\mathrm{PG}(n, p)$ instead of $\mathrm{AG}(n, p)$. Since affine spaces can be embedded into projective spaces, our improved caps in $\mathrm{AG}(n, p)$ also represent caps in the projective space.

Related Work. It is known that for $m \in \{3, 4, 5\}$, a cap in $\mathbb{Z}_m^n$ is equivalent to a set in which no three distinct points are in an arithmetic progression. (Note that $\mathbb{Z}_4^n \neq \mathrm{AG}(n, p)$.) There were important contributions by Brown and Buhler [11], Frankl, Graham and Rödl [20], Meshulam [31], Lev [30], Bateman and Katz [4], Croot, Lev and Pach [12], Ellenberg and Gijswijt [18] as well as Petrov and Pohoata [34]. Moreover, some readers may recall the case of caps in $\mathbb{Z}_3^4$ from the popular card game *SET* [13].

So far, the best known approach to construct caps *for general prime modulus* is to take a simple product construction of large caps in low dimension. Let $C_{n,p}$ respectively $C_{n,p}^{\mathrm{pr}}$ denote[1] the size of the largest affine respectively projective cap in dimension n. It is known that the largest affine cap in dimension 3 has size p^2, i.e., $C_{3,p} = p^2$; see for example in [7]. These maximal caps are also called ovaloids. In $\mathrm{PG}(3, q)$ with odd q, these maximal caps come from elliptic quadrics, see [33]. A representative for such a cap is the set

$$\{(t^2 + st + as^2, s, t, 1) \mid s, t \in \mathbb{F}_q\} \cup \{(1, 0, 0, 0)\},$$

where $x^2 + x + a$ is irreducible over $\mathbb{F}_q$. In the corresponding affine space, the point $(1, 0, 0, 0)$ is removed. As a consequence, we obtain the bound $C_{n,p} \gg p^{2n/3}$ by simply taking products of this cap. This result can be considered classical, as the determination of the size of caps in $\mathrm{PG}(3, q)$ for odd prime powers q goes back to Bose [10] in 1947.

The refinement by Edel and Bierbrauer is based on the fact that one can form an almost-product of special projective caps, namely if they possess a tangent hyperplane (see [17, Theorem 10]). In particular, this gives $(q^2 + 1)^2 - 1 = q^4 + 2q^2$ points in $\mathrm{PG}(6, q)$, and the reduction to the affine space gives $q^4 + q^2 - 1$ points in $\mathrm{AG}(6, q)$; see [16, Section 1].

[1] Note that for projective caps the size of the largest cap is often denoted by $m_2(n, p)$, and sometimes $m_2^{\mathrm{aff}}(n, p)$ is used in the affine case.

There are several computational results on caps in small dimension; see [15, 26]. However, the only known asymptotic improvement over Bose's result on the lower bound when $p \geq 5$ is due to Bierbrauer and Edel [17, Theorem 11] for projective caps and Edel [16] for affine caps, and is based on a product construction of a large cap in dimension 6. If n is a multiple of 6, then Edel's construction yields $C_{n,p}^{\mathrm{pr}} \geq C_{n,p} \geq (p^4 + p^2 - 1)^{n/6}$. If n is not a multiple of 6, one can modify the construction slightly, but in any case this only influences a constant C_p in $C_{n,p} \geq C_p(p^4 + p^2 - 1)^{\lfloor n/6 \rfloor}$.

It is known that the limit $c_p = \lim_{n\to\infty}(C_{n,p})^{1/n}$ exists and is in the interval $[2, p)$; see for example [19, Proposition 3.8]. Numerically, Edel's construction gives only a small improvement of the earlier bound $p^{2n/3}$. For example, when $p = 17$, then the bound $6.611\ldots$ is improved to $6.615\ldots$. In this paper, we will improve this to 7. When $p = 23$, a lift of Bose's result gives a constant $8.087\ldots$, which Edel improved to $8.090\ldots$. We improve this to 9. However, while Edel's construction works for all primes, our construction has to be optimized for individual primes.

For $p = 5$, Edel's construction gives $c_5 \geq 649^{1/6} = 2.942\ldots$. Recently, Elsholtz and Pach [19] have constructed large progression-free sets and it emerged that in $\mathbb{Z}_5$, their construction is asymptotically better than Edel's bound; Edel's lower bound was improved to $c_5 \geq 3$. In the case modulo 4 (i.e., working in $\mathbb{Z}_4^n$ rather than $\mathbb{F}_4^n$), Elsholtz and Pach [19] gave a much more substantial improvement from $c_4 \geq 2.519\ldots$ to $c_4 \geq 3$. Improvements in the case of a prime base seem to be much more difficult, since the existing construction of Edel seems to be good.

Another important measure for the size of caps is the exponent $\mu(p) = \lim_{n\to\infty}(\log_p C_{n,p})/n$ in the representation of the size as $p^{\mu(p)n}$. The mentioned result $C_{n,p} \gg p^{2n/3}$ clearly implies $\mu(p) \geq 2/3$. The recent breakthrough of Ellenberg and Gijswijt [18] shows that $\mu(p) < 1$. Indeed, their method yields the bound

$$C_{n,p} \leq (J(p)p)^n,$$

where

$$J(p) = \frac{1}{p} \min_{0<t<1} \frac{1 - t^p}{(1 - t)\, t^{(p-1)/3}};$$

see [9]. It is known that $J(s)$ is decreasing and $\lim_{s\to\infty} J(s) = 0.8414\ldots$; see [9, Equation (4.11)].

Besides the mentioned product constructions, also another approach is known: In an unpublished work of Lev [29], he describes an elegant method to "globally" construct large caps in $\mathbb{F}_3^n$. These caps have basically the form

$$S = \{(x, y, x^2 - \lambda y^2) \mid x, y \in \mathbb{F}_q\} \subseteq \mathbb{F}_q^3 \cong \mathbb{F}_3^n,$$

where $\lambda \in \mathbb{F}_q$ is a fixed non-square, n is a multiple of 3 and $q = 3^{n/3}$. However, these sets have size $3^{2n/3} = (2.08008\ldots)^n$, which is of the same quality as Bose's construction lifted to higher dimension.

Overview of our Work. In this paper, we extend the combinatorial method of Elsholtz and Pach [19] from the case of sets avoiding arithmetic progressions to affine caps (with prime modulus larger than 4). In particular, we introduce some new directions for finding good digit sets, which are crucial for our constructions of large caps; see Section 3.

Our results improve the lower bounds of c_p for $p \in \{11, 17, 23, 29, 41\}$, and the improvements in these cases are indeed substantial. Especially the case $p = 23$ with an exponent of $\mu(23) \geq 0.70075\ldots$ comes quite close to the case of $p = 3$, where a construction is known based on a large cap in dimension 480, giving $\mu(3) \geq 0.72485\ldots$; see [16, Section 5].

Table 1 compares our new lower bounds to those by Edel [16].

$\boldsymbol{p}$	**lower bounds for $\boldsymbol{c_p}$**				**exponent $\boldsymbol{\mu(p)}$**
	$p^{2/3}$	$(p^4+p^2-1)^{1/6}$	new	improvement*	
5	2.92401...	2.94243...	**3**	1.9562%	**0.68260...**
7	3.65930...	3.67139...	3		0.56457...
11	4.94608...	4.95282...	**5**	0.9526%	**0.67118...**
13	5.52877...	5.53418...	4		0.54047...
17	6.61148...	6.61528...	**7**	5.8156%	**0.68682...**
19	7.12036...	7.12364...	6		0.60852...
23	8.08757...	8.09012...	**9**	11.2468%	**0.70075...**
29	9.43913...	9.44099...	$\geq$ **10**	$\geq$ 5.9210%	$\geq$ **0.68380...**
31	9.86827...	9.86998...	$\geq$ 8		$\geq$ 0.60554...
37	11.10370...	11.10505...	$\geq$ 10		$\geq$ 0.63767...
41	11.89020...	11.89138...	$\geq$ **12**	$\geq$ 0.9134%	$\geq$ **0.66914...**

*Compared to the best previously known bound $(p^4+p^2-1)^{1/6}$.

Table 1: Comparison of previously known best lower bounds for c_p to our new ones, and new lower bounds for the exponent $\mu(p)$. The $\geq$-sign is meant to indicate cases in which we cannot ensure that our method is not able to produce better results than the stated ones.

2 Results and Construction

In the following, we use Vinogradov's notation, where $f(n) \gg g(n)$ means that there exists some $C > 0$ such that $f(n) \geq Cg(n)$ holds for all $n > 0$.

We directly start by stating our main result.

Theorem 1. *If $C_{n,p}$ denotes the size of the largest affine cap in $\mathbb{Z}_p^n$ and $c_p = \lim_{n\to\infty}(C_{n,p})^{1/n}$, then it holds that*

$$C_{n,11} \gg \frac{5^n}{n^{1.5}}, \quad C_{n,17} \gg \frac{7^n}{n^{2.5}}, \quad C_{n,23} \gg \frac{9^n}{n^{3.5}}, \quad C_{n,29} \gg \frac{10^n}{n^4} \quad \text{and} \quad C_{n,41} \gg \frac{12^n}{n^5},$$

and as a consequence, we have

$$c_{11} \geq 5, \quad c_{17} \geq 7, \quad c_{23} \geq 9, \quad c_{29} \geq 10 \quad \text{and} \quad c_{41} \geq 12.$$

Since every subset of $\mathrm{AG}(n,p)$ can be embedded into $\mathrm{PG}(n,p)$, this directly implies the following corollary.

Corollary 2. *The lower bounds from Theorem 1 also hold for the largest caps in* $\mathrm{PG}(n,p)$.

Moreover, our improved bounds on caps can also be transformed into improved bounds on linear codes. For details we refer to [17, Theorem 1].

These new bounds are based on a "global" construction of affine caps: We take a set of n-dimensional points, where the set depends on n in a much stronger way than taking a tensor product construction of a small (local) cap. The idea is, for a fixed prime p, to find a large set of digits $D \subseteq \mathbb{Z}_p$ and a subset $D' \subseteq D$ such that the set

$$S(D,D',n) := \left\{(a_1,\dots,a_n) \in D^n \,\middle|\, \forall d \in D' \colon a_i = d \text{ for } \frac{n}{|D|} \text{ values of } i\right\} \tag{2.1}$$

is a cap in $\mathrm{AG}(n,p)$ for all $n \in \mathbb{N}$ with $|D| \mid n$. If this is the case, then we say that (D, D') is *admissible*. Moreover, we say that D is admissible if there is some $D' \subseteq D$ such that (D, D') is admissible. Note that if $D_1' \subseteq D_2' \subseteq D$, then the admissibility of (D, D_1') implies the admissiblity of (D, D_2').

Next, we combinatorially determine the cardinality of the set $S(D, D', n)$ and then asymptotically estimate it by applying Stirling's formula, which leads to

$$|S(D,D',n)| = \left(\prod_{\ell=0}^{|D'|-1} \binom{n - \frac{\ell n}{|D|}}{\frac{n}{|D|}}\right)(|D| - |D'|)^{n-\frac{|D'|n}{|D|}} \sim \frac{c|D|^n}{n^{\delta/2}} \tag{2.2}$$

with

$$\delta = \min\{|D'|, |D|-1\} \quad \text{and} \quad c = \frac{1}{\sqrt{1-\delta/|D|}}\left(\frac{|D|}{2\pi}\right)^{\delta/2}.$$

The form of the parameter δ comes from the fact that fixing the frequencies of $|D|$ digits leads to the same result as fixing the frequencies of $|D|-1$ digits, because then the frequency of the last digit is fixed automatically. With the usual interpretation $0^0 = 1$, (2.2) also holds true for $D' = D$. The given cardinality is of order $(|D| - o(1))^n$ as n increases.

In order to obtain a large cap, (2.2) implies that, first of all, we need to

- choose the digit set D as *large* as possible, and then
- find a corresponding set $D' \subseteq D$ of digits with fixed frequencies which is as *small* as possible.

However, the minimization of the set D' is restricted by the fact that the frequency conditions are crucial to ensure that the resulting set is indeed a cap. More details can be found in Section 3.

Finally, we give some additional comments on the construction.

Remark 2.1. (a) For simplicity, the reader can assume in the first reading that $D' = D$. This still covers all the main improvements, and only slightly weakens the exponent of n in the denominators of our results.

(b) It is not crucial for our method that the frequencies of the digits in (2.1) are exactly $n/|D|$. Other constants which can also vary depending on the digit and add up to n also work. However, if we want to maximize the size of the cap $S(D, D', n)$, then $n/|D|$ is the best choice, in view of the multinomial distribution.

(c) If the dimension n is not a multiple of $|D|$, then we can trivially extend the set $S(D, D', n - (n \bmod |D|))$ to a subset of $\mathbb{Z}_p^n$ by filling the remaining coordinates with a good cap in dimension $n \bmod |D|$. As a consequence, (2.2) holds for all $n \in \mathbb{N}$, understood as an asymptotic lower bound with a slightly weaker constant c.

(d) One could also think of restrictions other than fixing the frequency of some digits, e.g., fixing the "radius" of the points (compare Behrend's construction for progression-free sets, the application to the multidimensioanl setting as explained by Petrov and Poahata [34] in the case modulo 8 and by Elsholtz and Pach [19] more generally). Or one could think of fixing the frequency of multiple digits *together* (as mentioned in [19, Proof of Theorem 3.11]). Both approaches do not seem to work for caps *in general*. However, we have refrained from further optimizing the denominators in Theorem 1.

(e) It turned out that if D is admissible, then the corresponding set D' can be chosen in such a way that $|D'| \leq |D| - 2$ holds. We believe that this is always possible.

(f) So far, our method only leads to an improvement for primes p with $p \equiv 5 \bmod 6$. We do not know why it only works for this residue class.

(g) It seems to be possible to add some smaller caps to a large cap constructed in this way so that the union of all points is still a cap. This would improve the constant c by a small factor (probably less than 2). For some details see [19, Theorem 3.2 and Corollary 3.4].

3 Approaches for Finding Admissible Sets

As already mentioned in the introduction, for $p = m \in \{3, 4, 5\}$ the cap set condition can be verified by only ensuring that no three points x, y and z from the set satisfy $x + z = 2y$ (which describes arithmetic progressions). For $p > 5$, the cap set condition is not only based on this equation, but also on the other equations $ax + by + cz = 0$, where a, b, $c \in \mathbb{Z}_p$ with $a + b + c = 0$. If $m = p$ is a prime, then without loss of generality, it is enough to assume that $a = 1$. With $c = -(b + 1)$ we can assume that $b \in \{1, \dots, p - 2\}$. (If $b = 0$, then $c = -1$ simply means that x, y and z are distinct. If $b = p - 1$, then we have $c = 0$ with the same consequence.)

3.1 Modelling the Problem

For the moment, let $b \in \{1, \dots, p - 2\}$ be fixed and $c = -(b + 1)$. Moreover, let

$$P_b(D) = \{(x, y, z) \in D^3 \mid x + by + cz = 0 \text{ and not } x = y = z\}$$

be the set of non-trivial “weighted progressions” corresponding to b. Assume that there is some $n \in \mathbb{N}$ with $|D| \mid n$ such that there are three points $x = (x_1, \dots, x_n)^\top$, $y = (y_1, \dots, y_n)^\top$, $z = (z_1, \dots, z_n)^\top \in S(D, D', n)$ which lie on a line. For each weighted progression $v = (v_1, v_2, v_3) \in P_b(D)$, we introduce a variable χ_v which describes the number of occurrences of v in the components of these three points, i.e.,

$$\chi_v = |\{i \in \{1, \dots, n\} \mid (x_i, y_i, z_i) = v\}|.$$

Because every digit d in D' has to occur the same number of times, we find the equations

$$\sum_{\substack{v \in P_b(D) \\ v_1 = d}} \chi_v = \sum_{\substack{v \in P_b(D) \\ v_2 = d}} \chi_v \quad \text{and} \quad \sum_{\substack{v \in P_b(D) \\ v_1 = d}} \chi_v = \sum_{\substack{v \in P_b(D) \\ v_3 = d}} \chi_v \tag{3.1}$$

for each $d \in D'$.

Now it is easy to see that the non-existence of a non-negative non-trivial integral solution $\chi = (\chi_v \mid v \in P_b(D))$ for the equations above for all $b \in \{1, \dots, p - 2\}$ is equivalent to the non-existence of three points on a line, i.e., the fact that $S(D, D', n)$ is a cap. So in order to prove admissibility of some (D, D'), we have to ensure that the polyhedron

$$\mathcal{P} = \{\chi \in \mathbb{Z}_{\geq 0}^{\ell} \mid A \cdot \chi = 0\}$$

only contains the zero vector for all $b \in \{1, \dots, p-2\}$, where the system of linear equations $Ax = 0$ describes the equations given in (3.1) and clearly depends on b, and $\ell = |P_b(D)|$. This can be done by methods of linear integer programming. In particular, the corresponding integer linear program[2] (IP)

$$\begin{aligned} &\max 0 \\ &\text{s.t. } A \cdot \chi = 0 \\ &\chi \in \mathbb{Z}_{\geq 0}^{\ell} \end{aligned} \tag{3.2}$$

can be checked for feasibility with a usual IP solver. For this article, we have used the MILP packages of SageMath [35] as well as JuMP, an optimization package of Julia [6]. A complete list of all admissible digit sets of maximal size for small p can be found at `https://gitlab.com/galipnik/large-caps`.

One way of ensuring that an admissible digit set has largest size, say size ℓ, among all admissible digit sets for fixed p is to find a feasible solution of the IP for *all*[3] possible digit sets of size $\ell+1$ for at least one b (which implies that all these sets cannot be admissible). We have done this for $p \leq 23$; see also Table 1. In order to give an idea of the computation times, our implementation took about 95 minutes for the case $p = 23$ (and $\ell + 1 = 10$), while it was executed on an Intel(R) Core(TM) i7-7500U CPU at 2.70GHz. In other words, showing the non-admissibility of thousands of individual digit patterns each took only a fraction of a second.

Unfortunately, deciding if a polyhedron contains an integer point is NP-complete [21], which implies, together with the fact that the number of possible digit sets also grows exponentially for increasing p, that checking admissibility for all possible digit sets modulo p can only be done for small p. Hence, it is very natural to look for simpler ways of checking whether digit sets are admissible. Two such approaches are described in the following section.

For an illustration of setting up the equations given in (3.1) as well as the corresponding constraint matrix A, we refer to the case $p = 23$ in Section 4.

3.2 Digit-Reducibility as a Sufficient Condition

Besides the computational method presented in the previous section, we next give a sufficient condition for the admissibility of a digit set, which allows us to verify very easily that a set is admissible.

A pair (D, D') with $D' \subseteq D \subseteq \mathbb{Z}_p$ is said to be *digit-reducible* if for every $b \in \{1, \ldots, p-2\}$ and $c = -(b+1)$ the following recursively defined algorithm results in the empty set: If there exists a position $r \in \{1, 2, 3\}$ and a digit $d \in D'$ such that d does not occur at position r in any of the triples in $P_b(D)$ but it occurs at one of the other positions in at least one of the triples in $P_b(D)$, then remove all weighted progressions from $P_b(D)$ which contain d at *any position*. Recursively apply this rule to the remaining set $P_b(D)$ again. If there do not exist an $r \in \{1, 2, 3\}$ and a digit $d \in D'$ such that d does not occur in any of the triples in $P_b(D)$ at position r but it occurs in at least one triple at any position, then stop the process.

We now explain why the reducibility of (D, D') implies that $S(D, D', n)$ is a cap for all $n \in \mathbb{N}$ with $|D| \mid n$. Assume that (D, D') is reducible and there are three pairwise different vectors $x = (x_1, \ldots, x_n)^\top$, $y = (y_1, \ldots, y_n)^\top$, $z = (z_1, \ldots, z_n)^\top \in S(D, D', n)$ for some $n \in \mathbb{N}$ and $b \in \mathbb{Z}_p \setminus \{0, -1\}$ such that $x + by + cz = 0$ with $c = -(b+1)$. This implies that there exists some i with $1 \leq i \leq n$ such that the component (x_i, y_i, z_i) of the vectors is a non-trivial weighted progression, i.e., it is in $P_b(D)$. However, the test above says that there is no triple in $P_b(D)$ which can occur, due to the fact that every digit in D' has to occur $|D|/n$ times in each vector. This is a

[2] Since we only need to check feasibility, the objective function is irrelevant but needed in an actual implementation.

[3] However, some equivalent digit sets can be neglected. For example, we can always assume without loss of generality that an admissible digit set contains the digits 0 and 1.

contradiction to the assumption that the vectors are pairwise different. Thus, the set $S(D, D', n)$ is a cap for all suitable $n \in \mathbb{N}$, and D is admissible.

For examples we refer to Section 4, cases $p = 11$ and $p = 17$.

3.3 Matrix-Reducibility as a Sufficient Condition

In order to show that a digit set D is admissible for some set $D' \subseteq D$ of digits with fixed frequency, we can also use the following sufficient condition based on the matrix A, which represents the linear constraints given in (3.1) via $Ax = 0$. Again, we consider each equation $x + by + cz = 0$ separately and fix $b \in \{1, \dots, p-2\}$. Let A_r be the reduced row echelon form of the matrix A. For each row of A_r which only contains non-negative respectively non-positive entries, it is clear that the variables corresponding to non-zero entries of this row have to be zero (otherwise, the equation that corresponds to the said row cannot be fulfilled). This is due to the fact that we only search for non-negative solutions x.

Thus, we can delete the columns of A_r that belong to these variables, and proceed with the next non-negative or non-positive row. Note that the deletion of columns can bring out new non-negative or non-positive rows. Naturally, this process determines if no such row is left in A_r. If at the end all columns of A_r are deleted, then all variables x_i have to be zero. If this is the case for all $b \in \{1, \dots, p-2\}$, then we say that (D, D') (or simply D) is *matrix-reducible*, which implies that the digit set D is admissible.

Remark 3.1. This procedure described here and the algorithm that we use for digit-reducibility in the previous subsection are *essentially* of the same shape: While we start with the reduced row echelon form of A here, we can reformulate the algorithm of Subsection 3.2 in such a way that it is the same as this one but with A itself as initial matrix instead of its echelon form. The reason for the different descriptions of the two algorithms is our belief that it is easier and more convenient to handle with digits and weighted progressions instead of the corresponding matrices—at least if one wants to understand it and do it by hand.

One can also think of other transformations of A as initial matrices for the reduction than the reduced row echelon form A_r or A itself, and even combine them. However, we refrained from optimizing this point because it works fine for our purpose.

For an example we refer to Section 4, case $p = 23$.

We remark that reducibility (both via digits or matrices) is only a sufficient condition for D to be admissible, but not necessary. The system of equations involved could have a more sophisticated structure, and there are indeed admissible digit sets which are not reducible. However, it turned out that these algorithmically simple tests are in fact very useful. They help to keep the proofs for the admissibility of digit sets simple and readable.

Moreover, digit- and matrix-reducibility are not equivalent: There exist digit sets which are digit- but not matrix-reducible (see case $p = 17$ in Section 4) and vice versa (see case $p = 23$ in Section 4).

Finally, it is of course also possible to combine the latter two approaches: We can choose between the digit- and matrix-reducibility algorithm depending on the parameter b. Indeed, there are digit sets which are neither digit- nor matrix-reducible, but if we combine the two approaches, then reducibility of the digit set can be shown.

3.4 Elimination of Some Equations

So far, it seems that admissibility (respectively reducibility) of a fixed digit set has to be checked in $p - 2$ cases, namely for all equations $x + by + cz \neq 0$ with b, $c \in \mathbb{Z}_p^n \setminus \{0, -1\}$ and $b + c = -1$.

This is in fact not necessary: The following two observations help to *significantly* reduce the cases that have to be studied later on.

Remark 3.2. Let p be a prime, $D \subseteq \mathbb{Z}_p$ and b, $c \in \mathbb{Z}_p$ with $c \neq 0$ and $b + c = -1$. Then the following assertions are true:

(a) A triple $(x, y, z) \in \mathbb{Z}_p^3$ satisfies $x+by+cz = 0$ if and only if (z, y, x) satisfies $z+c^{-1}by+c^{-1}x = 0$. In particular, this means that $P_b(D)$ contains the same elements as $P_{c^{-1}b}(D)$ but mirror-inverted. Hence, only one of the two equations $x + by + cz = 0$ and $x + c^{-1}by + c^{-1}z = 0$ has to be considered.

(b) The equation $x+by+cz = 0$ implies that $(x, y, z) \in P_b(D)$ holds if and only if $(x, z, y) \in P_c(D)$. In other words, $P_b(D)$ and $P_c(D)$ contain the same elements, but the last two components of the triples are always flipped. Thus, it is enough to consider one of the equations $x + by + cz = 0$ and $x + cy + bz = 0$.

We say that two equations $x + b_1y + c_1z = 0$ and $x + b_2y + c_2z = 0$ are equivalent if either $b_2 = c_1^{-1}b_1$ (case (a) above) or $b_1 = c_2$ (case (b) above). Hence, only representatives of non-equivalent equations have to be tested for the cap set property.

For the primes p considered in Theorem 1, the iterated application of the two cases of Remark 3.2 implies an immense simplification in our proof: It reduces the number of relevant equations from $p - 2$ to $(p + 1)/6$.

4 Proof of Theorem 1

If we find an admissible set of digits D and $D' \subseteq D$ of suitable sizes (depending on p), then the statements of the theorem follow by (2.2). Because of the comments above, it is enough to show reducibility.

Case $p = 11$. We claim that $D = \{0, 1, 3, 4, 5\}$ with fixed digits $D' = \{0, 1, 3\}$ is digit-reducible (as well as matrix-reducible, which is not shown here), and study solutions x, y, $z \in D$ of $x+by+cz = 0$ with $b \in \mathbb{Z}_p \setminus \{0, -1\}$ and $c = -(b + 1)$.

1. Case $x + z = 2y$. We list all triples of digits $(x, y, z) \in \{0, 1, 3, 4, 5\}^3$ that are solutions of $x + z = 2y$, but leave out the trivial solutions $x = y = z$. These are the triples in $P_{-2}(D)$ and are given by
$$(1, 3, 5), (3, 4, 5), (5, 3, 1), (5, 4, 3).$$
We have $1 \in D'$ and thus, the frequency of this digit has to be equal in any of the three positions. However, 1 does not occur in any of the triples in the second position, and as a consequence, the digits 1 can only occur in the trivial progression $(1, 1, 1)$. So the triples $(1, 3, 5)$ and $(5, 3, 1)$ cannot occur in any component of a potential weighted progression in $S(D, D', n)$. Hence, we delete $(1, 3, 5)$ and $(5, 3, 1)$ from the above list and
$$(3, 4, 5), (5, 4, 3)$$
remain. None of these two triples has the digit 3 in the second position. Thus, we delete both of them, and no triple from the set $P_{-2}(D)$ remains.

   By Remark 3.2 (b), this also solves the case $x + 9z = 10y$. Moreover, as 5 is the inverse of 9 modulo 11, also the equation $x + 5z = 6y$ is covered due to Remark 3.2 (a).

2. Case $x + 2z = 3y$. For this equation ($b = -3$) the set of non-trivial weighted progressions $P_{-3}(D)$ is given by

$$(1,0,5),(1,3,4),(1,4,0),(3,0,4),(3,1,0),(4,1,5),(4,5,0),(5,0,3).$$

   As 0 never occurs in the first position and 1 never occurs in the last position, we can remove all triples with any occurrence of 0 or 1. Therefore, again no non-trivial solutions in D remain.

   By Remark 3.2 (a) with $2^{-1} \equiv 6 \bmod 11$, also the equation $x + 6z = 7y$ has no non-trivial solution in D. By Remark 3.2 (b), this moreover solves the cases $x + 8z = 9y$ and $x + 4z = 5y$. Again applying the observation from Remark 3.2 (a) to the latter two equations with $8^{-1} \equiv 7 \bmod 11$ respectively $4^{-1} \equiv 3 \bmod 11$, also the equations $x + 7z = 8y$ and $x + 3z = 4y$ are covered.

Since we have (directly or via Remark 3.2) considered all cases $b \in \{1, \dots, p-2\}$, we conclude that (D, D') is digit-reducible and thus, the appropriate size of $S(D, D', n)$ follows by (2.2). □

Case $p = 17$. We claim that the digit set $D = \{0, 1, 2, 4, 8, 9, 13\}$ is reducible with fixed digits $D' = \{0, 1, 2, 4, 8\}$, and argue in analogy to the case $p = 11$ above.

1. Case $x + z = 2y$. We list all triples of digits in $P_{-2}(D)$, which are

$$(0,1,2),(0,2,4),(0,4,8),(0,9,1),(0,13,9),(1,9,0),(1,13,8),(2,1,0),(4,0,13),$$
$$(4,2,0),(8,0,9),(8,2,13),(8,4,0),(8,13,1),(9,0,8),(9,13,0),(13,0,4),(13,2,8).$$

   Since the digit 8 does not occur in any of the triples on the second position, we can delete all triples that contain any 8 and obtain the remaining list

$$(0,1,2),(0,2,4),(0,9,1),(0,13,9),(1,9,0),$$
$$(2,1,0),(4,0,13),(4,2,0),(9,13,0),(13,0,4).$$

   Next, we observe that no triple of this list has a 4 on the second position. Thus, we delete all triples which contain the digit 4. This yields

$$(0,1,2),(0,9,1),(0,13,9),(1,9,0),(2,1,0),(9,13,0).$$

   Now this list contains no triple with the digits 0 or 2 on the second position. By deleting all triples containing these digits, no non-trivial solution remains, which closes the argument for this case.

   By Remark 3.2, this also solves the cases

   - $x + 15z = 16y$ (as $2 + 16 \equiv 1 \bmod 17$) and
   - $x + 8z = 9y$ (as $15^{-1} \equiv 8 \bmod 17$).

2. Case $x + 2z = 3y$. This equation yields

$$(1,0,8),(1,9,13),(1,13,2),(2,1,9),(2,9,4),(4,1,8),(4,2,1),$$
$$(4,13,9),(8,0,13),(8,4,2),(8,9,1),(9,0,4),(13,0,2),(13,4,8)$$

as triples in $P_{-3}(D)$. As 0 never occurs in the first position and 8 never occurs in the second position, we can remove all triples with any 0 or 8. The remaining list is given by

$$(1,9,13),(1,13,2),(2,1,9),(2,9,4),(4,2,1),(4,13,9).$$

Next, we observe that the digit 4 never occurs in the second position, which leads to the list

$$(1,9,13),(1,13,2),(2,1,9).$$

Here, the digit 1 does not occur in the third position. So all triples can be removed, which implies that there is no non-trivial solution of $x+2z=3y$ in D^3.

Moreover, by repeatedly applying Remark 3.2, this also solves the cases

- $x+14z=15y$,
- $x+9z=10y$,
- $x+7z=8y$,
- $x+11z=12z$,
- $x+5z=6y$.

3. Case $x+3z=4y$. This equation has the triples

$$(1,2,8),(1,13,0),(2,9,0),(4,1,0),(8,2,0),(8,13,9),$$
$$(9,1,4),(9,4,8),(9,8,2),(13,2,4),(13,4,1),(13,9,2)$$

as non-trivial solutions. Since 0 never occurs in the first position, we can remove all triples containing any 0 and obtain

$$(1,2,8),(8,13,9),(9,1,4),(9,4,8),(9,8,2),(13,2,4),(13,4,1),(13,9,2).$$

Furthermore, the digits 2 and 4 do not occur in any triple in the first position, which leads to $(8,13,9)$ as only remaining triple. Surely, a single triple leads to the empty set.

By Remark 3.2, this also solves the cases

- $x+13z=14y$,
- $x+6z=7y$,
- $x+4z=5y$,
- $x+10z=11y$,
- $x+12z=13y$.

Since all cases $b\in\{0,\dots,p-2\}$ are covered, this implies that (D,D') is reducible and thus, also admissible. The appropriate size of the corresponding cap $S(D,D',n)$ follows by (2.2) again. □

Case $p=23$. We claim that the digit set $D=\{0,1,3,4,8,9,10,12,17\}$ with fixed digits $D'=\{0,1,3,4,8,10,17\}$ is admissible. Unfortunately, D is neither digit- nor matrix-reducible for any D' of size 7. So the admissibility has been checked by solving the IP given in (3.2) with appropriate software. This leads to the result

$$C_{n,23}\gg\frac{9^n}{n^{3.5}},$$

as stated in Theorem 1.

As a consolation price, we show that D is matrix-reducible for $D'=D$, i.e., if we fix the frequencies of *all* digits. (This would lead to a lower bound of $9^n/n^4$.) For this purpose, we again study solutions $(x,y,z)\in D^3$ of $x+by+cz=0$ with $b\in\mathbb{Z}_p\setminus\{0,-1\}$ and $c=-(b+1)$.

The equivalent equations with respect to Remark 3.2 are given as follows, where each set represents an equivalence class:

$$\{x+z=2y, x+21z=22y, x+11z=12y\},$$
$$\{x+20z=21y, x+15z=16y, x+12z=13y, x+10z=11y, x+7z=8y, x+2z=3y\},$$
$$\{x+19z=20y, x+17z=18y, x+14z=15y, x+8z=9y, x+5z=6y, x+3z=4y\},$$
$$\{x+18z=19y, x+16z=6y, x+13z=14y, x+9z=10y, x+6z=7y, x+4z=5y\}.$$

Only one representative of each class has to be considered.

Let us have a look at the equation $x+z=2y$. Here, the progressions in $P_{-2}(D)$ are given by

$$(0,4,8),(0,12,1),(1,9,17),(1,12,0),(1,17,10),(3,10,17),(3,17,8),(4,8,12),$$
$$(8,1,17),(8,4,0),(8,9,10),(8,10,12),(8,17,3),(10,9,8),(10,17,1),$$
$$(12,3,17),(12,8,4),(12,10,8),(17,1,8),(17,3,12),(17,9,1),(17,10,3),$$

and we call them $v_1, \ldots, v_{22}$ in the given order. The corresponding constraint matrix A (defined by the equations in (3.1)) then has the form

$$A = \begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\
0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 1 & 1
\end{pmatrix},$$

where the first nine rows represent equations which arise from the first and second position in the vectors of $P_{-2}(D)$ (left equation in (3.1)), and the last nine rows represent the constraints for the positions one and three in the vectors of $P_{-2}(D)$ (right equation in (3.1)).

Let us take a closer look at the construction of A: For the first row of A we consider the first digit of D, which is 0. This digit occurs in the triples $v_1=(0,4,8)$ and $v_2=(0,12,1)$ in the first position, and in none of the triples in the second position. Hence, following the left equation of (3.1) with $d=0$, this leads to the equation

$$x_{v_1}+x_{v_2}=0,$$

which is represented by the first row of A.

As a second, more sophisticated example, we consider the fourteenth row of A and the corresponding fifth digit in D, which is 8. Now the first and the third positions of the progressions are significant (because the row is part of the last nine rows). In the vectors $v_9=(8,1,17)$, $v_{10}=(8,4,0)$, $v_{11}=(8,9,10)$, $v_{12}=(8,10,12)$ and $v_{13}=(8,17,3)$, the digit 8 occurs in the first position. The vectors $v_1=(0,4,8)$, $v_7=(3,17,8)$, $v_{14}=(10,9,8)$, $v_{18}=(12,10,8)$ and $v_{19}=(17,1,8)$ contain 8 in the third position. Hence, due to the right equation in (3.1) with $d=8$, this yields the equation

$$x_{v_9}+x_{v_{10}}+x_{v_{11}}+x_{v_{12}}+x_{v_{13}}=x_{v_1}+x_{v_7}+x_{v_{14}}+x_{v_{18}}+x_{v_{19}}.$$

This es exactly the equation represented by the fourteenth row of A.

The reduced row echelon form A_r of A is given by

$$A_r = \begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&-1&0&17&1&1&0&0&1\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&6&-1&-1&0&0&-1\\
0&0&1&0&0&0&0&0&0&0&0&0&0&1&1&0&17&-1&-1&0&0&-1\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&-1&0&18&1&1&0&0&1\\
0&0&0&0&1&0&0&0&0&0&0&0&0&-1&0&0&17&-1&-1&0&-1&-1\\
0&0&0&0&0&1&0&0&0&0&0&0&0&-1&-1&0&21&1&0&-1&0&1\\
0&0&0&0&0&0&1&0&0&0&0&0&0&1&1&0&4&0&0&0&0&-1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&22&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&6&-1&0&0&-1&-1\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&1&0&5&-1&-1&0&0&-1\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&-1&0&6&1&1&0&1&1\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&2&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&2&1&0&-1&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&2&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{pmatrix}.$$

Now we look for non-zero rows of A_r in which all entries are either non-negative or non-positive. The indices of these rows are given by $R_1 = \{8, 12, 14, 15\}$ (and are framed in green in the following matrix). Next, we delete all columns of A_r in which any row r with $r \in R_1$ has some non-zero entry (symbolized by red lines). This means that we eliminate the variable which corresponds to this column. As a result, the first reduction step looks like

$$\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&-1&0&17&1&1&0&0&1\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&6&-1&-1&0&0&-1\\
0&0&1&0&0&0&0&0&0&0&0&0&0&1&1&0&17&-1&-1&0&0&-1\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&-1&0&18&1&1&0&0&1\\
0&0&0&0&1&0&0&0&0&0&0&0&0&-1&0&0&17&-1&-1&0&-1&-1\\
0&0&0&0&0&1&0&0&0&0&0&0&0&-1&-1&0&21&1&0&-1&0&1\\
0&0&0&0&0&0&1&0&0&0&0&0&0&1&1&0&4&0&0&0&0&-1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&22&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&6&-1&0&0&-1&-1\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&1&0&5&-1&-1&0&0&-1\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&-1&0&6&1&1&0&1&1\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&2&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&2&1&0&-1&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&2&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{pmatrix}
\rightsquigarrow
\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&-1&1&0&1\\
0&1&0&0&0&0&0&0&0&0&0&0&1&-1&0&-1\\
0&0&1&0&0&0&0&0&0&0&0&1&1&-1&0&-1\\
0&0&0&1&0&0&0&0&0&0&0&0&-1&1&0&1\\
0&0&0&0&1&0&0&0&0&0&0&-1&0&-1&-1&-1\\
0&0&0&0&0&1&0&0&0&0&0&-1&-1&0&0&1\\
0&0&0&0&0&0&1&0&0&0&0&1&1&0&0&-1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&-1&-1\\
0&0&0&0&0&0&0&0&1&0&0&0&1&-1&0&-1\\
0&0&0&0&0&0&0&0&0&1&0&0&-1&1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{pmatrix}.$$

The rows r for $r \in R_1$ are zero now. Next, we proceed the same way with the resulting smaller matrix on the right-hand side. The only non-negative respectively non-positive non-zero row in this matrix is row 13. So we delete the corresponding two columns, and again proceed with the smaller matrix, and so on. The full remaining reduction is given by

$$\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&-1&1&0&1\\
0&1&0&0&0&0&0&0&0&0&0&0&1&-1&0&-1\\
0&0&1&0&0&0&0&0&0&0&0&1&1&-1&0&-1\\
0&0&0&1&0&0&0&0&0&0&0&0&-1&1&0&1\\
0&0&0&0&1&0&0&0&0&0&0&-1&0&-1&-1&-1\\
0&0&0&0&0&1&0&0&0&0&0&-1&-1&0&0&1\\
0&0&0&0&0&0&1&0&0&0&0&1&1&0&0&-1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&-1&-1\\
0&0&0&0&0&0&0&0&1&0&0&0&1&-1&0&-1\\
0&0&0&0&0&0&0&0&0&1&0&0&-1&1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{pmatrix}
\rightsquigarrow
\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&-1&1&0\\
0&1&0&0&0&0&0&0&0&0&0&1&-1&0\\
0&0&1&0&0&0&0&0&0&0&1&1&-1&0\\
0&0&0&1&0&0&0&0&0&0&0&-1&1&0\\
0&0&0&0&1&0&0&0&0&0&-1&0&-1&-1\\
0&0&0&0&0&1&0&0&0&0&-1&-1&0&0\\
0&0&0&0&0&0&1&0&0&0&1&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&-1\\
0&0&0&0&0&0&0&0&1&0&0&1&-1&0\\
0&0&0&0&0&0&0&0&0&1&0&-1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{pmatrix}
\rightsquigarrow$$

$$\rightsquigarrow \begin{pmatrix} 1&0&0&0&0&0&0&0&0&1&0\\ 0&1&0&0&0&0&0&0&0&-1&0\\ 0&0&1&0&0&0&0&0&0&-1&0\\ 0&0&0&1&0&0&0&0&0&1&0\\ 0&0&0&0&1&0&0&0&0&-1&-1\\ 0&0&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0&-1\\ 0&0&0&0&0&0&0&1&0&-1&0\\ 0&0&0&0&0&0&0&0&1&1&1\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0&0&0&0&0\\ 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&0&0&0\\ 0&0&1&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \\ \end{pmatrix},$$

where the last step trivially follows.

All other equations can be handled analogously. □

Case $p \in \{29, 41\}$. For these moduli, we only give the digit sets for which the digit-reducibility can be checked analogously to the cases $p = 11$ and $p = 17$: The sets (D, D') with

- $D = \{0, 1, 2, 3, 4, 6, 14, 16, 22, 26\}$ with $D' = \{1, 2, 3, 4, 6, 16, 22, 26\}$ for $p = 29$ and
- $D = \{1, 2, 4, 5, 6, 9, 15, 16, 27, 32, 33, 35\}$ with $D' = \{1, 2, 4, 5, 6, 9, 15, 27, 32, 33\}$ for $p = 41$

are digit-reducible.

Finally, we give the lists of equivalent equations with respect to Remark 3.2; each set of equations represents an equivalence class. In the case $p = 29$, we have

$$\begin{aligned}
&\{x + z = 2y, x + 27z = 28y, x + 14z = 15y\},\\
&\{x + 26z = 27y, x + 19z = 20y, x + 15z = 16y, x + 13z = 14y, x + 9z = 10y, x + 2z = 3y\},\\
&\{x + 25z = 26y, x + 21z = 22y, x + 18z = 19y, x + 10z = 11y, x + 7z = 8y, x + 3z = 4y\},\\
&\{x + 24z = 25y, x + 23z = 24y, x + 22z = 23y, x + 6z = 7y, x + 5z = 6y, x + 4z = 5y\},\\
&\{x + 20z = 21y, x + 17z = 18y, x + 16z = 17y, x + 12z = 13y, x + 11z = 12y, x + 8z = 9y\},
\end{aligned}$$

and for $p = 41$, the classes are given by

$$\begin{aligned}
&\{x + z = 2y, x + 39z = 40y, x + 20z = 21y\},\\
&\{x + 38z = 39y, x + 27z = 28y, x + 21z = 22y, x + 19z = 20y, x + 13z = 14y, x + 2z = 3y\},\\
&\{x + 37z = 38y, x + 30z = 31y, x + 26z = 27y, x + 14z = 15y, x + 10z = 11y, x + 3z = 4y\},\\
&\{x + 36z = 37y, x + 32z = 33y, x + 31z = 32y, x + 9z = 10y, x + 8z = 9y, x + 4z = 5y\},\\
&\{x + 35z = 36y, x + 34z = 35y, x + 33z = 34y, x + 7z = 8y, x + 6z = 7y, x + 5z = 6y\},\\
&\{x + 29z = 30y, x + 25z = 26y, x + 23z = 24y, x + 17z = 18y, x + 15z = 16y, x + 11z = 12y\},\\
&\{x + 28z = 29y, x + 24z = 25y, x + 22z = 23y, x + 18z = 19y, x + 16z = 17y, x + 12z = 13y\}.
\end{aligned}$$

Only one representative of each of these classes has to be considered.

This concludes the proof of Theorem 1. □

5 Non-Equivalent Caps

Two caps are equivalent if there is an affine transformation from one cap to the other. In some cases, two caps in $\mathrm{AG}(n, p)$ based on the above digit constructions but with different digit sets D_1 and D_2 are equivalent, while in other cases they are not. We briefly discuss this in the cases $p = 5$ and $p = 11$.

If a digit set $D_1 \subseteq \mathbb{Z}_p$ can be mapped by an affine transformation $f(x) = ax + b$ to another digit set $D_2 \subseteq \mathbb{Z}_p$, then the corresponding caps are equivalent.

For example, modulo $p = 5$ all digit sets consisting of three distinct digits are eqivalent. One can first map two arbitrary digits to 0 and 1. Then the three remaining digit sets $D_1 = \{0, 1, 2\}$, $D_2 = \{0, 1, 3\}$ and $D_3 = \{0, 1, 4\}$ can be seen to be equivalent: D_3 is mapped to D_1 by $f(x) = x+1$ and D_2 is mapped to D_1 by $f(x) = 3x + 2$.

A simple criterion to see that two digit sets are not equivalent is as follows: For a given digit set write the multiset of differences (including the gap from the largest digit to p). If the multiset of differences of two digit sets D_1 and D_2 contain different frequencies of differences, then the two digit sets are not equivalent.

Applying this modulo 5 to the above digit sets gives twice the set of differences $\{1, 1, 3\}$ and once $\{1, 2, 2\}$. This helps finding the map $f(x) = 3x + 2$.

On the other hand, we easily find many admissible digit sets are not equivalent modulo 11: $D_1 = \{0, 1, 2, 3, 4\}$ with difference multiset $\{1, 1, 1, 1, 7\}$ is non-equivalent to $D_2 = \{0, 1, 2, 3, 6\}$ with difference multiset $\{1, 1, 1, 3, 5\}$, and both are different from $D_3 = \{0, 1, 2, 3, 7\}$ with difference multiset $\{1, 1, 1, 4, 4\}$.

Another criterion is that the order of gaps of the same frequencies must also be preserved: $D_4 = \{0, 1, 2, 6, 7\}$ is different from the earlier three digit sets, as D_4 does not contain four elements in an arithmetic progression, which would be preserved by an affine map. We leave it to the reader to argue why $D_5 = \{0, 1, 2, 6, 8\}$ and $D_6 = \{0, 1, 2, 8, 9\}$ lead to further non-equivalent digit sets.

For larger primes, the number of admissible digit sets is typically much larger than the number $p(p - 1)$ of affine transformations of the digit sets. Hence, our digit-based constructions typically indicate the existence of many non-equivalent caps with the same number of points. (However, we do not formally prove these caps are non-equivalent.) In any case this seems to be of interest even for those primes for which these caps are not larger than previously known ones.

References

[1] Daniele Bartoli, Giorgio Faina, and Massimo Giulietti, *Small complete caps in three-dimensional Galois spaces*, Finite Fields Appl. **24** (2013), 184–191. MR 3093866

[2] Daniele Bartoli, Giorgio Faina, Stefano Marcugini, and Fernanda Pambianco, *Complete caps in AG(N, q) with both N and q odd*, J. Combin. Des. **25** (2017), no. 9, 419–425. MR 3684305

[3] Daniele Bartoli, Massimo Giulietti, Giuseppe Marino, and Olga Polverino, *Maximum scattered linear sets and complete caps in Galois spaces*, Combinatorica **38** (2018), no. 2, 255–278. MR 3800841

[4] Michael Bateman and Nets H. Katz, *New bounds on cap sets*, J. Amer. Math. Soc. **25** (2012), no. 2, 585–613. MR 2869028

[5] Michael Bennett, *Bounds on sizes of generalized caps in AG(n, q) via the Croot-Lev-Pach polynomial method*, J. Combin. Theory Ser. A **168** (2019), 255–271. MR 3974704

[6] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah, *Julia: A fresh approach to numerical computing*, SIAM Rev. Soc. Ind. Appl. Math. **59** (2017), no. 1, 65–98.

[7] Jürgen Bierbrauer, *Large caps*, J. Geom. **76** (2003), no. 1-2, 16–51, Combinatorics, 2002 (Maratea). MR 2005528

[8] Jürgen Bierbrauer and Yves Edel, *Large caps in projective Galois spaces*, Current Research Topics in Galois Geometry (Leo Storme and Jan De Beule, eds.), Nova Science Publishers, New York, 2010, pp. 81–94.

[9] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans, *On cap sets and the group-theoretic approach to matrix multiplication*, Discrete Anal. (2017), Paper No. 3, 27. MR 3631613

[10] Raj C. Bose, *Mathematical theory of the symmetrical factorial design*, Sankhyā **8** (1947), 107–166. MR 26781

[11] Tom C. Brown and Joe P. Buhler, *A density version of a geometric Ramsey theorem*, J. Combin. Theory Ser. A **32** (1982), no. 1, 20–34. MR 640624

[12] Ernie Croot, Vsevolod F. Lev, and Péter P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337. MR 3583357

[13] Benjamin L. Davis and Diane Maclagan, *The card game SET*, Math. Intell. **25** (2003), no. 3, 33–40. MR 2005098

[14] Alexander A. Davydov and Patric R. J. Östergård, *Recursive constructions of complete caps*, vol. 95, 2001, Special issue on design combinatorics: in honor of S. S. Shrikhande, pp. 167–173. MR 1829107

[15] Yves Edel, *Caps: Introduction, Results and Generator Matrices for Caps*, `https://www.mathi.uni-heidelberg.de/~yves/Matritzen/CAPs/CAPMatIndex.html`, accessed on August 13, 2020.

[16] Yves Edel, *Extensions of generalized product caps*, Des. Codes Cryptogr. **31** (2004), no. 1, 5–14. MR 2031694

[17] Yves Edel and Jürgen Bierbrauer, *Recursive constructions for large caps*, Bull. Belg. Math. Soc. Simon Stevin **6** (1999), no. 2, 249–258. MR 1705136

[18] Jordan S. Ellenberg and Dion Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343. MR 3583358

[19] Christian Elsholtz and Péter Pál Pach, *Caps and progression-free sets in $\mathbb{Z}_m^n$*, Des. Codes Cryptogr. **88** (2020), no. 10, 2133–2170. MR 4156230

[20] Peter Frankl, Ronald L. Graham, and Vojtěch Rödl, *On subsets of abelian groups with no 3-term arithmetic progression*, J. Combin. Theory Ser. A **45** (1987), no. 1, 157–161. MR 883900

[21] Michael R. Garey and David S. Johnson, *Computers and intractability*, W. H. Freeman and Co., San Francisco, Calif., 1979, A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences. MR 519066

[22] Joshua A. Grochow, *New applications of the polynomial method: the cap set conjecture and beyond*, Bull. Amer. Math. Soc. (N.S.) **56** (2019), no. 1, 29–64. MR 3886143

[23] J. W. P. Hirschfeld and J. A. Thas, *Open problems in finite projective spaces*, Finite Fields Appl. **32** (2015), 44–81. MR 3293405

[24] James W. P. Hirschfeld, *Finite projective spaces of three dimensions*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1985, Oxford Science Publications. MR 840877

[25] ———, *Projective geometries over finite fields*, second ed., Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1998. MR 1612570

[26] James W. P. Hirschfeld and Leo Storme, *The packing problem in statistics, coding theory and finite projective spaces: update 2001*, Finite geometries, Dev. Math., vol. 3, Kluwer Acad. Publ., Dordrecht, 2001, pp. 201–246. MR 2061806

[27] James W. P. Hirschfeld and Joseph A. Thas, *General Galois geometries*, Springer Monographs in Mathematics, Springer, London, 2016. MR 3445888

[28] Hans-Joachim Kroll and Rita Vincenti, *A new construction of caps*, Discrete Math. **310** (2010), no. 22, 3155–3161. MR 2684085

[29] Vsevolod F. Lev, *A uniform cap in* $\mathrm{AG}(r, 3)$, unpublished manuscript.

[30] ———, *Progression-free sets in finite abelian groups*, J. Number Theory **104** (2004), no. 1, 162–169. MR 2021632

[31] Roy Meshulam, *On subsets of finite abelian groups with no* 3*-term arithmetic progressions*, J. Combin. Theory Ser. A **71** (1995), no. 1, 168–172. MR 1335785

[32] A. C. Mukhopadhyay, *Lower bounds on* $m_t(r, s)$, J. Combinatorial Theory Ser. A **25** (1978), no. 1, 1–13. MR 491256

[33] Christine M. O'Keefe, *Ovoids in* $\mathrm{PG}(3, q)$*: a survey*, vol. 151, 1996, Graph theory and combinatorics (Manila, 1991), pp. 175–188. MR 1391265

[34] Fedor Petrov and Cosmin Pohoata, *Improved bounds for progression-free sets in* C_8^n, Israel J. Math. **236** (2020), no. 1, 345–363. MR 4093890

[35] The SageMath Developers, *SageMath Mathematics Software (Version 9.0)*, 2020, `http://www.sagemath.org`.

[36] Raphaël Salem and Donald C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Natl. Acad. Sci. USA **28** (1942), 561–563. MR 7405