Small weight codewords in the codes arising from Desarguesian projective planes

V. Fack, Sz. L. Fancsali, L. Storme, G. Van de Voorde[*] and J. Winne[†]

Research Group on Combinatorial Algorithms and Algorithmic Graph Theory
Department of Applied Mathematics and Computer Science, Ghent University,
Krijgslaan 281-S9, 9000 Ghent, Belgium
Veerle.Fack@UGent.be, Joost.Winne@UGent.be

Department of Pure Mathematics and Computer Algebra, Ghent University,
Krijgslaan 281-S22, 9000 Ghent, Belgium
ls@cage.ugent.be, gvdvoorde@cage.ugent.be

Department of Computer Science, Eötvös Loránd University, Budapest,
Pazmany P. s. 1/c, Hungary, H-1117
nudniq@cs.elte.hu

Abstract

We study codewords of small weight in the codes arising from Desarguesian projective planes. We first of all improve the results of K. Chouinard on codewords of small weight in the codes arising from $PG(2,p)$, p prime. Chouinard characterized all the codewords up to weight $2p$ in these codes. Using a particular basis for this code, described by Moorhouse, we characterize all the codewords of weight up to $2p+(p-1)/2$ if $p \geq 11$. We then study the codes arising from $PG(2, q = q_0^3)$. In particular, for $q_0 = p$ prime, $p \geq 7$, we prove that the codes have no codewords with weight in the interval $[q+2, 2q-1]$. Finally, for the codes of $PG(2,q)$, $q = p^h$, p prime, $h \geq 4$, we present a discrete spectrum for the weights of codewords with weights in the interval $[q+2, 2q-1]$. In particular, we exclude all weights in the interval $[3q/2, 2q-1]$.

[*]This author's research is supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

[†]Supported by the Fund for Scientific Research - Flanders (Belgium).

1 Introduction

We define the incidence matrix $A = (a_{ij})$ of the projective plane $PG(2,q)$, $q = p^h$, p prime, $h \geq 1$, as the matrix whose rows are indexed by lines of the plane and whose columns are indexed by points of the plane, and with entry

$$a_{ij} = \begin{cases} 1 & \text{if point } j \text{ belongs to line } i, \\ 0 & \text{otherwise.} \end{cases}$$

The p-ary code C of the projective plane $PG(2,q)$, $q = p^h$, p prime, $h \geq 1$, is the $\mathbb{F}_p$-span of the rows of the incidence matrix A. The references [1] and [11] contain a lot of information on codes from planes.

In particular, in [1], it is proven that the scalar multiples of the incidence vectors of the lines are the codewords of minimal weight $q+1$ in the code arising from $PG(2,q)$. Chouinard [3] proved that for the code arising from $PG(2,p)$, p prime, there are no codewords of weight in the interval $[p+2, 2p-1]$ and that the only codewords of weight $2p$ are the scalar multiples of the differences of the incidence vectors of two distinct lines.

We will improve the result of Chouinard by characterising the codewords up to weight $2p + (p-1)/2$, for $p \geq 11$. We show that the only possible non-zero weights are $p+1, 2p$, and $2p+1$, and prove that codewords of weight $2p+1$ are a linear combination of two incidence vectors of lines, with the linear combination non-zero in the intersection point of the two lines.

To obtain these results, we will use a particular basis for the code C, found by E. Moorhouse, see [7].

We then concentrate on the codes arising from $PG(2,q)$, $q = q_0^3$, $q_0 = p^h$, p prime, $h \geq 1$. For $h = 1$ and $p \geq 7$, we prove that there are no codewords having weight in the interval $[q+2, 2q-1]$. For $h > 1$ and $p \geq 7$, we exclude the possible weights $q + q^{2/3} + 1$ and $q + q^{2/3} + q^{1/3} + 1$ for the codewords.

For arbitrary Desarguesian projective planes, we give a discrete spectrum for the possible weights of the codewords in the interval $[q+2, 3q/2]$ and exclude all codewords with weight in the interval $[3q/2, 2q-1]$. For all the new results, we rely on links with blocking sets in $PG(2,q)$.

Acknowledgement The authors thank the referees and Simeon Ball for the detailed reading of the article and their helpful suggestions in writing the final version.

2 The Moorhouse basis for $AG(2,p)$, p prime

The rank of the p-ary linear code of the projective plane $PG(2,p)$, p prime, is $\binom{p+1}{2} + 1$ and the rank of the p-ary linear code of the affine plane $AG(2,p)$, p prime, is $\binom{p+1}{2}$. In [7], Moorhouse gives an easy construction for a basis for $AG(2,p)$, p prime, which can be seen as the projective plane $PG(2,p)$, with one line M and its points omitted.

Consider the $(p^2+p+1)\times(p^2+p+1)$ incidence matrix A of $PG(2,p)$ with the line M as the first row:

$$A = \left(\begin{array}{ccccccc} 1 & \dots & 1 & 0 & \dots & 0 \\ * & \dots & * & & & \\ \vdots & & \vdots & & B & \\ * & \dots & * & & & \end{array} \right).$$

The $(p^2+p)\times p^2$ matrix B, obtained by deleting the first row and the first $p+1$ columns of A, is the incidence matrix of $AG(2,p)$. Moorhouse gives the following basis for the row space of B, in which $r_0, r_1, \dots, r_p$ are the points of M:

for $i \in \mathbb{N},\ 0 \le i \le p-1$, take $p-i$ random affine lines through r_i.

These, in total, $\binom{p+1}{2}$ lines form a basis for the row space of B. When we also add the line M, we obtain a basis for the code C of $PG(2,p)$.

This basis will play a crucial role in our arguments. We will refer to this particular basis as the *Moorhouse basis of* $AG(2,p)$.

We present this basis in the next figure. The full lines denote the lines forming the basis of the code C of $PG(2,p)$, while the dotted lines are lines through the points $r_0, \dots, r_i, \dots$, that are not taken as lines for the basis of the code C of $PG(2,p)$.

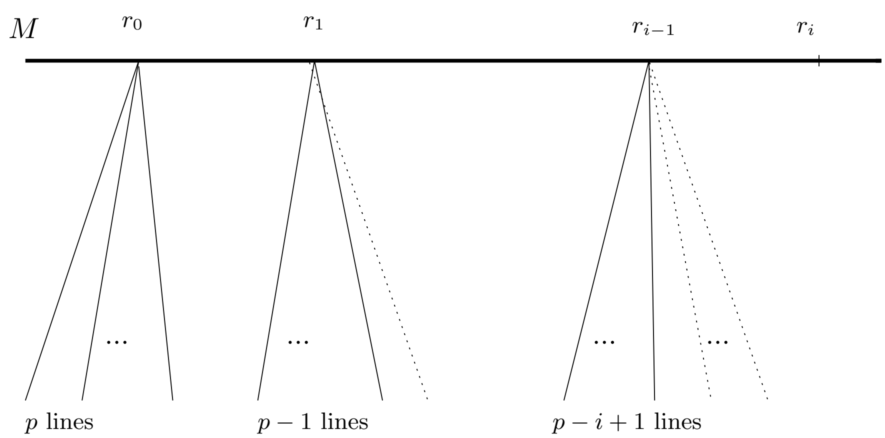

Figure 1: the basis of Moorhouse

We will also use a slight variation on this basis. Inspired by the results of a computer search, we found the following result.

Theorem 1. *The vector space generated by the affine lines of the Moorhouse basis through r_0, r_1, and r_2, can also be generated by choosing $p-1$ affine lines through each of the points r_0, r_1, and r_2, with the restriction that the three non-selected affine lines are not concurrent.*

Proof: Select the affine lines of the Moorhouse basis through the three points r_0, r_1, and r_2. Let M_1 be the non-selected affine line through r_1. Let M_2 and M_3 be the non-selected affine lines through r_2. We first show that we can select M_1, M_2, and M_3 without loss of generality.

We use elations with center r_0 and axis M to select M_1 without loss of generality. Let $r = M_1 \cap M_2$, we then use elations with center r_1 and axis M to fix the point r on M_2 without loss of generality. We finally use homologies with center r and axis M to select M_3 without loss of generality. We now work affinely with coordinates (x, y). We can assume without loss of generality that:

- M is the line at infinity,
- r_0 is the point at infinity of the vertical lines $X = x$,
- r_1 is the point at infinity of the horizontal lines $Y = y$,
- r_2 is the point at infinity of the diagonal lines $Y = X + y - x$, i.e. lines with slope 1.

From the preceding paragraphs, we can assume that the non-selected affine lines through r_1 and r_2 for the Moorhouse basis are $M_1 : Y = 0$, $M_2 : Y = X$, and $M_3 : Y = X + 1$.

We now want to write the incidence vector of the line M_2 as a linear combination of the vertical lines, the horizontal lines (except for M_1) and the diagonal lines (except for M_2 and M_3). The point $(0, 0)$ belongs to M_2, so has coefficient 1. Since we cannot use the lines M_1 and M_2, the line $X = 0$ has coefficient 1 in the linear combination defining M_2. This implies that the point $(0, 1)$ already has coefficient 1, but it should have coefficient 0 because it does not belong to M_2. We can only use the horizontal line $Y = 1$ (M_3 is forbidden), so $Y = 1$ has coefficient -1 in the linear combination defining M_2.

Continuing in this way for the points $(1, 1)$ and $(1, 2)$, $X = 1$ has coefficient 2 and $Y = 2$ has coefficient -2. In general, $Y = i$ has coefficient $-i$ and $X = j$ has coefficient $j + 1$ in the linear combination defining M_2. Then the point (i, i) has coefficient $-i + (i + 1) = 1$. The point (a, b) belongs to the line $Y = X + b - a$; we give this line the coefficient $b - a - 1$. If we use all horizontal, vertical, and diagonal lines for (a, b), then (a, b) gets the coefficient $(a + 1) + (-b) + (b - a - 1) = 0$. But we do not use M_2, so the points of M_2 have coefficient $-i + (i + 1) = 1$. We do not use M_1 and M_3 since both have coefficient 0.

We conclude that to write M_2 as a linear combination, we use all vertical lines, lines through r_0, except for the line $X = -1$ intersecting M_1 and M_3 in the same point (equivalently, we use all affine lines through r_0 except for the line $r_0 r$, with $\{r\} = M_1 \cap M_3$) and we use all affine lines through r_1 and r_2 which are lines of the Moorhouse basis.

Let $N_1, \ldots, N_{p-1}$ be the affine lines through r_0 except for $r_0 r$. We write M_2 as $\sum_{i=1}^{p-1} \epsilon_i N_i + R$, where R is a linear combination of other lines through r_1 and r_2, different from M_1 and M_3. As thus $N_j = (M_2 - R - \sum_{i=1, i \neq j}^{p-1} \epsilon_i N_i)/\epsilon_j$. So if we remove N_j from the basis and add M_2 to it, we still have a vector space of dimension $3(p-1)$ generated by these lines. $\square$

3 Improved results for $PG(2,p)$, p prime

We study from now on codewords in the code arising from $PG(2,q)$. Since $PG(2,q)$ is self-dual, we can describe the incidence matrix A of $PG(2,q)$ in either of the following two ways:

- the columns of the incidence matrix A of $PG(2,q)$ correspond to the points of $PG(2,q)$ and the rows correspond to the lines of $PG(2,q)$,
- the columns of the incidence matrix A of $PG(2,q)$ correspond to the lines of $PG(2,q)$ and the rows correspond to the points of $PG(2,q)$.

In this section, we will use the second correspondence. The following theorem links nicely the non-zero positions in codewords to a rank problem regarding the incidence matrix A of $PG(2,q)$.

Theorem 2. *Let C be the linear code generated by a matrix A.*

Let A' be the matrix obtained from A by deleting a set D of columns of A and let r be the rank of the subspace of codewords of C whose non-zero positions only appear in the columns of D, then $rk(A) - rk(A') = r$.

Consequently, if a set of columns is deleted from A, then the rank of A decreases if and only if there is a non-zero codeword in C with its non-zero positions contained in the set of the deleted columns of A.

Proof: Consider the projection $\varphi : C \to C_T$, where T is the complement of D in the set of columns of A. Then C_T is the code C *punctured* at D, that is, with the coordinates in D deleted. Then A' is a generator matrix for C_T.

The kernel of the projection φ is the set $\{c \in C | supp(c) \subseteq D\}$, so $rk(A) - rk(A') = r$, with $r = rk(\{c \in C | supp(c) \subseteq D\})$. $\square$

We will use this theorem to improve the results of Chouinard who characterized all the codewords in the code of $PG(2,p)$, p prime, of weight at most $2p$ [3].

Since we let the columns of the incidence matrix A correspond to the lines of $PG(2,p)$ and the rows to the points of $PG(2,p)$, deleting columns from the incidence matrix A then corresponds to deleting a set B of lines of $PG(2,p)$. The rank of A only decreases when it is not possible to reconstruct a basis for the column space of A by using the non-deleted lines of $PG(2,p)$.

A possible way for constructing a basis for the column space of A is by trying to construct a Moorhouse basis for an affine space contained in $PG(2,p)$ by using the lines not in B, and then by finding a last line which extends this basis of $AG(2,p)$ to a basis of $PG(2,p)$.

This is the method we will apply.

All codewords of weight up to $2p$ in the code arising from $PG(2,p)$, p prime, are known by the results of Assmus and Key [1], and Chouinard [3]. We characterize all codewords c, with $2p+1 \leq wt(c) \leq 2p+(p-1)/2$, by induction on the weight of the codewords.

In the induction hypothesis, we assume that the codewords of weight smaller than $wt(c)$ are already classified as being either:

1. a codeword of weight $p+1$ which is, up to a scalar multiple, the incidence vector of all lines through one point r,
2. a codeword of weight $2p$ which is, up to a scalar multiple, the difference of the incidence vectors of all lines through two points r and r',
3. a codeword of weight $2p+1$ which is a linear combination $\alpha c_1 + \beta c_2$ of the incidence vectors c_1 and c_2 of all lines through two points r and r', with $\alpha + \beta \neq 0$.

We also rely on a result of Ball and Blokhuis on dual double blocking sets.

Definition 1. *A* dual double blocking set *of $PG(2,q)$ is a set B of lines such that each point of $PG(2,q)$ belongs to at least two lines of B.*

Theorem 3. (Ball and Blokhuis [2]) *A double blocking set in $PG(2,p)$, p prime, has at least size $(5p+5)/2$.*

Suppose now that c is a codeword with $wt(c) = 2p+i$, with $i \in [1, \frac{p-1}{2}]$, where we assume that there are no codewords of weight in the interval $[2p+2, 2p+i-1]$. The non-zero positions in such a codeword define a set B of lines such that if the columns in A corresponding to these lines are deleted, the rank of A decreases (Theorem 2).

We now study all cases in which we delete at most $2p+(p-1)/2$ lines corresponding to the set of non-zero positions of a codeword c of C. The set of deleted lines is denoted by B.

Case 1: Suppose that there is a point r_0 on zero lines of B.

If at most $2p+(p-1)/2$ lines are deleted, we can select and delete two lines through r_0, then at most $2p+(p+3)/2$ lines are deleted. So there remains a point r_1 on at most one deleted line since a dual double blocking set in $PG(2,p)$ has at least $(5p+5)/2$ lines (Theorem 3).

Let $M = r_0 r_1$ and let M be the line at infinity of the corresponding affine plane $AG(2,p)$ of $PG(2,p)$. Note that $M \notin B$. Let $r_0, \ldots, r_p$ be the points of M. We check whether we can reconstruct the Moorhouse basis for $AG(2,p)$. Using the notations of the beginning of Section 2, through the point r_i, there need to pass $p-i$ affine lines of the Moorhouse basis.

The p affine lines through r_0 and the $p-1$ affine lines through r_1 which are necessary for the Moorhouse basis are indeed available. By induction on the index i for r_i, we can select $p-i$ affine lines through a point r_i, $2 \leq i \leq p$, of M

for the Moorhouse basis if $(2p+(p-1)/2)/(p-i+1) < i+1$ since then there is a point in the set $\{r_i, \ldots, r_p\}$ lying on less than $i+1$ lines in B. The previous condition is equivalent to $i+1+(p-1)/(2(i-1)) < p$.

This is satisfied for all $i \leq p-2$ when $p > 5$.

Problems arise when all lines through r_{p-1} and r_p, different from the line $r_{p-1}r_p = M$, belong to B since we need one affine line through r_{p-1} for the Moorhouse basis.

If all affine lines through r_{p-1} and r_p are deleted, then this means that in the corresponding codeword c, the positions corresponding to these $2p$ lines all have non-zero entries. So two out of the p deleted lines through r_{p-1} have the same non-zero entry. We rescale c so that at least these two entries are equal to 1, i.e.

$$c = (\underbrace{0}_{\text{line } r_{p-1}r_p}, \underbrace{1, 1, *, \ldots, *}_{p \text{ affine lines through } r_{p-1}}, \underbrace{*, \ldots, *}_{p \text{ affine lines through } r_p}, *, \ldots, *).$$

The codeword c' of weight $2p$ defined by the $2p$ affine lines through r_{p-1} and r_p is, up to a scalar multiple,

$$c' = (\underbrace{0}_{\text{line } r_{p-1}r_p}, \underbrace{1, \ldots, 1}_{p \text{ affine lines through } r_{p-1}}, \underbrace{-1, \ldots, -1}_{p \text{ affine lines through } r_p}, 0, \ldots, 0).$$

Then

$$c - c' = (\underbrace{0}_{\text{line } r_{p-1}r_p}, \underbrace{0, 0, *, \ldots, *}_{p \text{ affine lines through } r_{p-1}}, \underbrace{*, \ldots, *}_{p \text{ affine lines through } r_p}, *, \ldots, *).$$

So $wt(c-c') < wt(c)$. By induction on $wt(c)$, $2p+1 \leq wt(c) \leq 2p+(p-1)/2$, we can assume that $c - c'$ is already characterized as being either:

1. a codeword of weight $p+1$ which is, up to a scalar multiple, the incidence vector of all lines through one point r,
2. a codeword of weight $2p$ which is, up to a scalar multiple, the difference of the incidence vectors of all lines through two points r and r',
3. a codeword of weight $2p+1$ which is a linear combination $\alpha c_1 + \beta c_2$ of the incidence vectors c_1 and c_2 of all lines through two points r and r', with $\alpha + \beta \neq 0$.

All three possibilities show that c can be written as a linear combination of at most three codewords of weight $p+1$, so a linear combination of at most three incidence vectors of all lines through points r, r', and r''.

Now a linear combination of the incidence vectors of three lines has weight at least $3p-2$ for $p > 2$. Namely, take three non-concurrent lines L_1, L_2 and L_3, then $L_1 - L_2 + L_3$ has weight $3p-2$. Since $wt(c) \leq 2p + (p-1)/2$, we deduce that c is a linear combination of at most two such codewords of weight $p+1$. Hence, c is described as written in one of the three possibilities above.

Now we can assume that not all lines through r_{p-1}, different from $r_{p-1}r_p$, are deleted. We use one of them for the Moorhouse basis. Then select the line $r_0r_1 = M$ through r_p to obtain a basis of size $(p^2+p)/2+1$ for the code of $PG(2,p)$.

In this latter case, we have reconstructed a basis for the column space of A. The rank of A has not decreased, so the set B of deleted lines cannot correspond to a codeword of the code of $PG(2,p)$ (Theorem 2).

Case 2: Suppose that every point of $PG(2,p)$ lies on at least one line of B.

Then there is a point on exactly one deleted line, since a double blocking set in $PG(2,p)$, p prime, has size at least $2p+(p+5)/2$, see Theorem 3.

Case 2.1: Suppose that there is a line $L \in B$ containing two points lying on no other line of B.

Let r_0, r_1 be two points of L lying on no other line of B, thus $L = r_0r_1$.

We try to reconstruct the Moorhouse basis for the affine plane defined by L. As in Case 1, problems only start to arise when all lines through r_{p-1} and r_p belong to B, now including the line L. As in Case 1, we can reduce the codeword c by the codeword c', which corresponds to all affine lines through r_{p-1} and r_p, to a codeword $c-c'$ of lower weight. So these codewords $c-c'$ are classified, leading to the same characterization for c as in Case 1.

So we can assume that at least one affine line through r_{p-1} is not deleted.

Suppose that all lines through r_p belong to B, then, the $p+1$ positions in c corresponding to the lines through r_p are non-zero. At least two of those positions have the same non-zero value; assume that this value is equal to 1.

Consider the codeword $c' = (\underbrace{1,\ldots,1}_{p+1 \text{ times}},0,\ldots,0)$ with 1 in the positions corresponding to the lines through r_p. Then $c-c'$ is a codeword of weight at most $wt(c)-2$. By induction on the weight, we can assume that the codeword c is already characterized. So either we get a basis for the code C, or $c-c'$ is a codeword already characterized as being a linear combination of at most two codewords of minimal weight $p+1$. Then c is a codeword which is a linear combination of at most three codewords of minimal weight. In fact, since $wt(c) \leq 2p+(p-1)/2$, c is a linear combination of at most two codewords of minimal weight.

If *not* all lines through r_p belong to B, we can select a line through r_p, not in B, as the last line for a basis of the code of $PG(2,p)$, p prime. But this then implies that the set B of deleted lines does not correspond to a codeword (Theorem 2).

Case 2.2: Suppose that there is a line $L \in B$ containing at least one point r_0 lying on no other line of B and at least one point r_1 lying on exactly two lines of B.

This case is discussed in the same way as Case 2.1.

Case 2.3: Suppose that there is a line $L \in B$ such that all points of L belong to at least two lines of B, and containing three points r_0, r_1, r_2 lying on exactly two lines of B.

Let M_0, M_1, M_2 be the lines, different from L, lying in B and passing through respectively r_0, r_1, r_2.

Let L be the line at infinity of the corresponding affine plane for which we try to construct the Moorhouse basis.

Case 2.3.1: Suppose that M_0, M_1, M_2 are not concurrent.

From Theorem 1, we know that the affine lines through r_0, r_1, and r_2, not belonging to B, generate the same vector space as the lines of the Moorhouse basis through these points generate. We can find enough lines through the points $r_i,\ i \in \mathbb{N},\ 3 \leq i \leq p$, of L if $(2p + (p-9)/2)/(p-i+1) < i+1$.

Note that M_0, M_1, M_2 and L are not considered in this inequality.

As before, problems only start to arise if all affine lines through r_{p-1} and r_p belong to B. But then it is impossible that all points of L lie on at least two lines of B. Hence, there are no problems to select an affine line through r_{p-1} for constructing the Moorhouse basis for $AG(2,p)$.

If all lines through r_p are deleted, as in Case 2.1, we can again reduce c to a codeword of lower weight (known by induction on the weight).

If not all lines through r_p are deleted, as in Case 2.1, we reconstruct a basis for the code C to obtain the same contradiction.

Case 2.3.2: Suppose that M_0, M_1, M_2 are concurrent in a point r.

Let c be the codeword corresponding to the set B of deleted lines. Let c' be the codeword corresponding to the $p+1$ lines through r. Let c and c' have the same non-zero symbol in the coordinate position corresponding to the line $r_0 r$. Then $c - c'$ is a new codeword of weight at most

$$\underbrace{2p + \frac{p-1}{2}}_{wt(c)} \quad + \quad \underbrace{(p-2)}_{\text{lines } r_i r \ ;\ i=3,\ldots,p} \quad \underbrace{-1.}_{\text{line } r_0 r \text{ is zero}}$$

So $wt(c-c') \leq 3p + (p-7)/2$. When we remove the lines corresponding to $c - c'$, we know that the point r_0 is not on any deleted affine line, and that the points r_1 and r_2 are on at most one deleted line.

A point r_i, $i > 2$, of L is on at most i deleted lines if

$$(3p + (p-7)/2)/(p-i+1) < i+1 \iff i+2+\frac{p-1}{2(i-2)} < p.$$

For $i = p-3$, this inequality reduces to $p > 9$. So if $p > 9$, all essential affine lines for the Moorhouse basis of the affine plane with L as line at infinity can be selected through the points r_i of L for $i \in \mathbb{N}, 3 \leq i \leq p-3$.

We still need two affine lines through one of the points r_{p-2}, r_{p-1}, and r_p, and one affine line through one of the other points among r_{p-2}, r_{p-1}, and r_p. Suppose that at least $p-1$ affine lines are deleted through each of the points r_{p-2}, r_{p-1}, and r_p, so at least $3(p-1)$ affine lines are deleted through these three points.

Since subtracting the codeword c' from c only affects one line through each of the points r_{p-2}, r_{p-1}, and r_p, at least $3p-6$ affine lines of B would necessarily pass through r_{p-2}, r_{p-1}, and r_p. But then $|B| \geq 3p-6+1+(p-2)$ since also the line L belongs to B and the points $r_0, \ldots, r_{p-3}$ still belong to a second line of B. For $p > 3$, this is false since $|B| \leq 2p + (p-1)/2$.

So it is possible to find a point r_{p-2} still lying on at least two affine lines not in B, which then can be selected as lines through r_{p-2} for the Moorhouse basis.

We also need at least one affine line through r_{p-1} or r_p for the Moorhouse basis. Assume that all affine lines through r_{p-1} and r_p have non-zero positions in the codeword $c - c'$. Then at least $2p-2$ of the affine lines through r_{p-1} and r_p have non-zero positions in c, so are lines of B. But then at most $2p + (p-1)/2 - 1 - (2p-2) = (p+1)/2$ other affine lines in B remain. This then contradicts the assumption that every point of L lies on a second line in B.

So we find the requested affine line through r_{p-1} for the construction of the Moorhouse basis for $AG(2,p)$.

If at least one line through r_p has a zero position in $c - c'$, then this line can be used as the last line for the basis of $PG(2,p)$, but then $c - c'$ does not define a codeword of the code of $PG(2,p)$, so also c does not define a codeword of the code of $PG(2,p)$.

So assume that all lines through r_p have non-zero coordinate values in $c - c'$. Add a suitable scalar multiple of the codeword c'' of weight $p+1$ defined by the lines through r_p to $c - c'$ so that some line through r_p has a zero position in $c - c' + c''$. We have a new codeword of C. But at the same time, we can construct a basis for the column space of A by using lines with zero positions in $c - c' + c''$. For, we still can use the previously determined $(p^2+p)/2$ lines of the Moorhouse basis since none of those lines passes through r_p. We now can select a line through r_p having a zero position in $c - c' + c''$ as the final line to construct a basis of the code of $PG(2,p)$. This is however impossible since $c - c' + c''$ is a codeword of C.

Summary: The preceding cases imply the following assumptions on the lines in the set B, for the cases not yet discussed.

- Every point of $PG(2,p)$ belongs to at least one line of B (consequence of Case 1).
- If a line $L \in B$ contains a point r_0 lying on exactly one line L of B, then all other points of L lie on at least three lines of B (consequence of Cases 2.1 and 2.2).
- If all points of a line $L \in B$ lie on at least two lines of B, and there is a point $r_0 \in L$ on exactly two lines of B, then there is at most one other point $r_1 \in L$ on exactly two lines of B. All other points of L lie on at least three lines of B (consequence of Case 2.3).

The preceding cases imply that a line L of B has at most two points that are on at most two lines of B. Let x be the number of points on one line of B, let y be the number of points on two lines of B, then the second bullet implies $2(|B| - x) \geq y$. The number of incidences of the points of $PG(2,p)$ with the lines of B is at least $3(p^2 + p + 1 - x - y) + 2y + x$, which implies $(p+1)|B| \geq 3p^2 + 3p + 3 - 2|B|$, so $(p+3)|B| \geq 3p^2 + 3p + 3$.

But $|B| \leq 2p + (p-1)/2$. This yields that

$$(p+3)(2p + (p-1)/2) \geq 3p^2 + 3p + 3,$$

which is false for $p > 7$.

This brings us to the following new theorem. We state the theorem in the original setting where the rows of A correspond to the incidence vectors of the lines of $PG(2,p)$.

Theorem 4. *The only codewords c, with $0 < wt(c) \leq 2p + (p-1)/2$, in the p-ary linear code C arising from $PG(2,p)$, p prime, $p \geq 11$, are:*

- *codewords with weight $p+1$: the scalar multiples of the incidence vectors of the lines of $PG(2,p)$,*
- *codewords with weight $2p$: $\alpha(c_1 - c_2)$, c_1 and c_2 the incidence vectors of two distinct lines of $PG(2,p)$,*
- *codewords with weight $2p+1$: $\alpha c_1 + \beta c_2$, $\beta \neq -\alpha$, with c_1 and c_2 the incidence vectors of two distinct lines of $PG(2,p)$.*

Remark 1. In [3], the weight enumerators of the linear codes of the projective planes $PG(2,p)$ of order two, three, four, five and eight are listed.

We note that the codewords of smallest weight are equal to the scalar multiples of the incidence vectors of the lines [1], and those of weight $2p$ are equal to the scalar multiples of the differences of the incidence vectors of two distinct lines of $PG(2,p)$ [1, Corollary 6.4.4] (see also Theorem 5).

The code of $PG(2,p)$, $p = 3$, has codewords of weight $2p + 1 = 7$ different from a linear combination of two lines, which is in contrast with the results for $p \geq 11$ of the preceding theorem.

Regarding the code of $PG(2,p)$, $p=5$, all codewords of weight $2p+1$ are a linear combination of the incidence vectors of two lines, which coincides with the results for $p \geq 11$ of the preceding theorem. But the code of $PG(2,p)$, $p=5$, has codewords of weight $2p+(p-1)/2 = 2p+2 = 12$, which is in contrast with the results for $p \geq 11$ of the preceding theorem [3, 5].

4 Codewords of small weight in $PG(2,q)$, $q=q_0^3$

We now consider the p-ary linear code C arising from the projective plane $PG(2,q)$, $q=q_0^3$, $q_0=p^h$, p prime, $h \geq 1$.

Consider first of all the planes $PG(2,p^3)$, $p \geq 7$ prime, $h \geq 1$. To prove that there are no codewords in C of weight between p^3+2 and $2p^3-1$, we first prove that every codeword in the code of $PG(2,q)$, $q=p^h$, p prime, $h \geq 1$, of weight in $[q+2, 2q-1]$ is a minimal blocking set intersecting every line in $1 \pmod p$ points. The following theorem is Corollary 6.4.4 of [1].

Theorem 5. *The codewords of minimal weight in $C \cap C^\perp$ have weight $2q$ and are the scalar multiples of differences of incidence vectors of two distinct lines of $PG(2,q)$.*

Lemma 1. *A codeword $c \in C \backslash C^\perp$ with weight in $[q+2, 2q-1]$ is a scalar multiple of the incidence vector of a minimal blocking set of $PG(2,q)$, $q=p^h$, p prime, $h \geq 1$, intersecting every line in $1 \pmod p$ points.*

Proof: The results of this lemma can also be found in [3]. We prove these results again to make the article self-contained, and because this lemma plays a crucial role in the remaining results of this article.

By Theorem 5, we know that the codewords of C, with weight in $[q+2, 2q-1]$, have to belong to $C \backslash C^\perp$. The scalar product (c, L), with c a codeword and L a line, is constant for all lines L because $(c, L_1 - L_2) \equiv 0 \pmod p$ for all distinct lines L_1, L_2, since $C \cap C^\perp$ is generated by all the differences of two lines of $PG(2,q)$ [1, Theorem 6.3.1]. The codeword $c \in C \setminus C^\perp$ defines via its non-zero positions a blocking set B of $PG(2,q)$ since $(c, L) = a \neq 0$, for all lines L of $PG(2,q)$.

We take a look at the points of the blocking set B defined by the non-zero positions in the codeword c. By the results of T. Szőnyi [14, Section 3], we know that every blocking set of $PG(2,q)$ of size smaller than $2q$ can be reduced in a unique way to a minimal blocking set, namely, by deleting all non-essential points. Let r be an essential point of B, thus lying on a tangent line L to B. We can rescale (c, L) to 1. Because c intersects the line L only in r, c has to take value $c_r = 1$ in the coordinate position corresponding to the point r. Since every essential point of B lies on at least one tangent line to B, and since $(c, L) = 1 \neq 0$ for all lines L of $PG(2,q)$, the coordinate positions in c corresponding to all the essential points r of B have the value $c_r = 1$.

Suppose that B is not minimal. Suppose that the point r' is not essential for the blocking set B. Then r' lies on at least one line containing only 2 points

r and r' of B. Otherwise, the weight of c would be greater than or equal to $1 + 2(q+1)$, a contradiction.

So there is a line intersecting B only in r and r'. Again, we know that every blocking set of $PG(2,q)$ of size smaller than $2q$ can be reduced in a unique way to a minimal blocking set, namely by deleting all non-essential points. Because r' is not essential for B, r is an essential point for B. So the value c_r of c in the coordinate position of the point r is equal to $c_r = 1$. But $(rr', c) = 1 = c_r + c_{r'} = c_{r'} + 1$. We see that $c_{r'}$, the coordinate value in the position of r', has to be equal to zero, but then the point r', which was not essential, is not a point of B.

Hence, all points of B are essential points of B; the blocking set B is minimal.

Since we now know that B is minimal, the non-zero coordinate positions in c all correspond to essential points of B, so are equal to 1. Since we also know already that $(c, L) = 1$ for all lines L of $PG(2,q)$, B necessarily intersects every line in $1 \pmod{p}$ points. $\square$

Remark 2. The minimal blocking sets B of size $q + 2 \leq |B| \leq 2q - 1$ in $PG(2, q = p^3)$, p prime, $p \geq 7$, intersecting every line in $1 \pmod{p}$ points, have been classified [8, 9, 10]. They are projectively equivalent to one of the following two blocking sets (points given with homogeneous coordinates):

$$B_1 = \left\{(x, T(x), 1) | x \in \mathbb{F}_{p^3}\right\} \cup \left\{(x, T(x), 0) | x \in \mathbb{F}_{p^3} \setminus \{0\}\right\},$$

with $T : \mathbb{F}_{p^3} \to \mathbb{F}_p : x \longmapsto x + x^p + x^{p^2}$, or

$$B_2 = \left\{(x, x^p, 1) | x \in \mathbb{F}_{p^3}\right\} \cup \left\{(x, x^p, 0) | x \in \mathbb{F}_{p^3} \setminus \{0\}\right\}.$$

Note that $|B_1| = p^3 + p^2 + 1$ and $|B_2| = p^3 + p^2 + p + 1$.

Lemma 2. [1, Lemma 6.6.1] *Let C be the p-ary linear code defined by the plane $PG(2,q)$, $q = p^h$, p prime, $h \geq 1$.*

A vector v, with constant non-zero symbols, is contained in $C + C^\perp$ if and only if $|supp(v) \cap L| \pmod{p}$ is independent of the line L of $PG(2,q)$.

Lemma 3. [1, Lemma 6.6.2] *Suppose that X is a codeword, with constant non-zero symbols, of the code C of $PG(2,q)$ and Y is a vector, with constant non-zero symbols, of $C + C^\perp$. Rescale X and Y so that every non-zero value is equal to 1. If $|Y \cap L| \equiv |X \cap L| \pmod{p}$ for each line L, then $|X \cap Y| \equiv |X|$ mod p.*

Theorem 6. *In the p-ary linear code of $PG(2,p^3)$, p prime, $p \geq 7$, there are no codewords with weight in the interval $[p^3 + 2, 2p^3 - 1]$.*

Proof: By the preceding lemmas, we know that the only candidates for the codewords with weight in the interval $[p^3 + 2, 2p^3 - 1]$ correspond, up to a scalar multiple, to the incidence vectors of the minimal blocking sets with sizes in the interval $[p^3 + 2, 2p^3 - 1]$ that intersect every line in $1 \pmod{p}$ points.

By the classification results of Polverino and Storme (Remark 2), only two types of blocking sets need to be checked. To show that the incidence vectors of these blocking sets cannot define a codeword in C, Lemmas 1 and 3 show that it is sufficient to find a second blocking set B' of one of the types described in Remark 2 such that $|B \cap B'| \not\equiv 1 \pmod{p}$.

Note that if the incidence vector of a blocking set B defines a codeword of C, then so does every projective image of B, since C is invariant under the collineation group of $PG(2, p^3)$.

We have to distinguish between the two possibilities of Remark 2. We deal with the case $B = B_1$ first.

Case 1: The blocking set B_1 does not define a codeword of C.

Here

$$B_1 = \left\{(x, T(x), 1) | x \in \mathbb{F}_{p^3}\right\} \cup \left\{(x, T(x), 0) | x \in \mathbb{F}_{p^3} \setminus \{0\}\right\}$$

and

$$B'_1 = \left\{(x', 1, T(x')) | x' \in \mathbb{F}_{p^3}\right\} \cup \left\{(x', 0, T(x')) | x' \in \mathbb{F}_{p^3} \setminus \{0\}\right\}.$$

What is $B_1 \cap B'_1$? We check the different possibilities.

Case 1.1. If $(x, T(x), 1) = (x', 0, T(x'))$, then $T(x) = 0$ and $T(x') \neq 0$. Thus $(x, T(x), 1) = (x'/T(x'), 0, 1)$, so that $x = x'/T(x')$. But then as $T(x') \in \mathbb{F}_p$, $T(x) = T(x')/T(x') = 1 \neq 0$.

Case 1.2. Similarly, if $(x, T(x), 0) = (x', 1, T(x'))$, we need $T(x) \neq 0$. Then $x/T(x) = x'$ gives $T(x') = 1 \neq 0$.

Case 1.3 If $(x, T(x), 0) = (x', 0, T(x'))$, then $T(x) = T(x') = 0$, and we get one common point $(1, 0, 0)$.

Case 1.4 Suppose that $(x, T(x), 1) = (x', 1, T(x'))$. Then none of the components can be 0, and $(x/T(x), 1, 1/T(x)) = (x', 1, T(x'))$. Thus $x/T(x) = x'$, which makes $T(x') = 1$; and then $1/T(x) = T(x')$ makes $T(x) = 1$ also. So $x = x'$. Hence, the points in $B_1 \cap B'_1$ of this form are the p^2 points $(x, 1, 1)$ for which $T(x) = 1$.

It follows that $|B_1 \cap B'_1| = p^2 + 1$, and the symmetric difference $B_1 \Delta B'_1$ has size $2p^3$. Suppose that B_1 corresponds to a codeword b_1, so that B'_1 also corresponds to a codeword b'_1. Then because $|B_1 \cap L| \equiv |B'_1 \cap L| \equiv 1 \pmod{p}$ for all lines L, $b_1 - b'_1 \in C \cap C^{\perp}$. As $b_1 - b'_1$ has weight $2p^3$, it is a minimum weight codeword of $C \cap C^{\perp}$ and thus has the form $L - L'$ for two lines L and L', by [1, Corollary 6.4.4] (see also Theorem 5) and the fact that the non-zero coefficients of $b_1 - b'_1$ are ± 1. Now the line $z = 0$ meets $B_1 \backslash B'_1$ in p^2 points, and $y = 0$ meets $B'_1 \backslash B_1$ in p^2 points. Thus it could only be that L is the line $z = 0$ and L' is the line $y = 0$. But these lines don't meet $B_1 \Delta B'_1$ in the required p^3 points.

Case 2: The blocking set B_2 does not define a codeword of C.

For B_2, the proof is analogous. We are looking for a blocking set B'_2 such that $|B_2 \cap B'_2| \not\equiv 1 \pmod{p}$. Set $B'_2 = \left\{(\omega x, \omega x^p, 1) | x \in \mathbb{F}_{p^3}\right\} \cup \left\{(x, x^p, 0) | x \in \mathbb{F}_{p^3} \setminus \{0\}\right\}, \omega \in \mathbb{F}_{p^3} \backslash \mathbb{F}_p$. Then

$$(\omega x, \omega x^p, 1) \in B_2 \cap B'_2$$

$$\Updownarrow$$
$$\omega x^p = \omega^p x^p$$
$$\Updownarrow$$
$$x = 0 \text{ or } \omega^{p-1} = 1.$$

It follows that $|B_2 \cap B_2'| = \underbrace{1}_{(0,0,1)} + \underbrace{p^2+p+1}_{points(x,x^p,0)} \equiv 2 \pmod{p}$, so it is not congruent to $1 \pmod{p}$. $\square$

Remark 3. The same arguments as given in the proof of the preceding theorem eliminate the incidence vectors of the following minimal blocking sets

$$B_1 = \left\{(x, T(x), 1) | x \in \mathbb{F}_{q_0^3}\right\} \cup \left\{(x, T(x), 0) | x \in \mathbb{F}_{q_0^3} \setminus \{0\}\right\},$$

with $T : \mathbb{F}_{q_0^3} \to \mathbb{F}_{q_0} : x \longmapsto x + x^{q_0} + x^{q_0^2}$, and

$$B_2 = \left\{(x, x^{q_0}, 1) | x \in \mathbb{F}_{q_0^3}\right\} \cup \left\{(x, x^{q_0}, 0) | x \in \mathbb{F}_{q_0^3} \setminus \{0\}\right\},$$

as codewords for the code arising from $PG(2, q = q_0^3)$, $q_0 = p^h$, p prime, $p \geq 7$, $h \geq 1$.

Since also the Baer subplanes in $PG(2, q)$, q square, are eliminated as possible codewords in the code of $PG(2, q)$ [1, Proposition 6.6.3], we obtain the following result.

Theorem 7. *The p-ary linear code C corresponding to the plane $PG(2, q = q_0^3)$, $q_0 = p^h$, $p \geq 7$ prime, $h \geq 1$, does not have codewords of weight $q_0^3 + q_0^2 + 1$ or of weight $q_0^3 + q_0^2 + q_0 + 1$; and if q_0 is a square, C has no codewords of weight $q_0^3 + q_0^{3/2} + 1$.*

Remark 4. The next minimal blocking sets of $PG(2, q = q_0^3)$, $q_0 = p^h$, $p \geq 7$ prime, $h \geq 1$, which need to be checked as possible codewords for the code C are minimal blocking sets B intersecting every line in $1 \pmod{p^e}$ points, where e is the largest divisor smaller than h of $3h$. This follows from the recent classification results of Sziklai [13] who proved that all the minimal blocking sets B of $PG(2, q = p^n)$, p prime, of size $|B| < 3(q + 1)/2$, intersect the lines of $PG(2, q = p^n)$ in $1 \pmod{p^e}$ points for some divisor e of n.

5 Codewords in $PG(2, q = p^h)$

We know that a codeword c with weight in the interval $[q + 2, 2q - 1]$ defines a minimal blocking set of $PG(2, q)$, $q = p^h$, p prime, $h \geq 1$, intersecting every line in $1 \pmod{p}$ points (Lemma 1). We wish to exclude as many values as possible as weights for the codewords in the general case $q = p^h$, with p prime, $h \geq 4$.

Consider a minimal blocking set B of size $|B| < 2q$ in $PG(2, q)$, $q = p^h$, p prime, $h \geq 1$, intersecting every line in $1 \pmod{p^e}$ points, with e the maximal integer for which this is true. Let $p^e = E$. Then we can derive the following equations.

$$\sum_{i\geq 0} \tau_{1+iE} = \sum_{i=1}^{q^2+q+1} 1 = q^2+q+1,$$

$$\sum_{i\geq 0}(1+iE)\tau_{1+iE} = |B|(q+1) = \sum_{i=1}^{q^2+q+1} x_i, \text{ and}$$

$$\sum_{i=1}^{q^2+q+1} x_i(x_i-1) = \sum_{i\geq 0}(1+iE)iE\tau_{1+iE} = |B|(|B|-1),$$

with $x_i = |L_i \cap B|$, τ_{1+iE} the number of lines intersecting B in $1+iE$ points, and $L_1, \ldots, L_{q^2+q+1}$ the lines of $PG(2,q)$, $q=p^h$. We get the second equation by counting the number of pairs (point r of B, line L), with $r \in L$, and the third equation by counting the number of triples (r_0, r_1, L), $r_0 \neq r_1$, $r_0, r_1 \in B$, where L contains the points r_0 and r_1.

Since all lines intersect the blocking set B in 1 or in at least $1+E$ points, we have the following inequality:

$$\sum_{i=1}^{q^2+q+1}(x_i-1)(x_i-1-E) \geq 0, \text{ or}$$

$$\sum_{i=1}^{q^2+q+1} x_i(x_i-1) - E\sum_{i=1}^{q^2+q+1} x_i - \sum_{i=1}^{q^2+q+1} x_i + (1+E)\sum_{i=1}^{q^2+q+1} 1 \geq 0.$$

Substituting the first three equations in the last inequality gives the following quadratic inequality:

$$|B|(|B|-1) - (E+1)|B|(q+1) + (1+E)(q^2+q+1) \geq 0. \ (\star)$$

Theorem 8. *There are no codewords with weight in $[3q/2, 2q-1]$ in the p-ary linear code of $PG(2,q)$, $q=p^h$, corresponding to a minimal blocking set intersecting every line in $1 \pmod{E}$ points when $E = p^e \geq 4$.*

Proof: If such a codeword exists, it corresponds to a minimal blocking set B. We will prove that $|B| < 3q/2$ when $|B| \leq 2q-1$. We check, under certain conditions, that when we substitute $|B| = 3q/2$ and $|B| = 2q$ in the quadratic inequality $(\star)$, the value is negative. Since the coefficient of $|B|^2$ is positive, this yields that $|B| < 3q/2$ or $|B| > 2q$.

For $|B| = 3q/2$, we get

$$q^2\left(\frac{7}{4} - \frac{E}{2}\right) + q\left(-2 - \frac{E}{2}\right) + E + 1.$$

This last value is smaller than 0 when $7/4 < E/2$. So when we suppose that $E \geq 4$, we have the desired conclusion.

For $|B| = 2q$, we get

$$q^2(3-E) + q(-3-E) + E + 1.$$

When $E \geq 4$, the last expression is strictly smaller than 0. $\square$

We excluded in Theorem 8 half of the interval $[q+2, 2q-1]$. Our goal is now to find in the other half $[q+2, 3q/2]$ of the interval, smaller pairwise disjoint intervals for the possible values for $|B|$. These intervals will depend on the possible values for $E = p^e$ and, for $p > 3$, will be disjoint for different values of e, further reducing the possibilities of the weights of codewords in $[q+2, 2q-1]$.

From [14, Section 5], we get

$$q + 1 + \frac{q}{p^e + 2} \leq |B| \leq q + \frac{9q}{4p^e} \ .$$

Note that the intervals are disjoint for distinct values of e if $p \neq 2, 3$. We will now derive a different upper bound on $|B|$.

Since we know that the codewords correspond to minimal blocking sets of $PG(2, q)$, $q = p^h$, $h \geq 1$, p prime, of size smaller than $3(q+1)/2$, intersecting every line in $1 \pmod{p^e}$ points, we can use the results of Sziklai [13, Corollary 4.18] which state that the largest integer e for which this is true is equal to a divisor of h. That is why we give the upper bound the form

$$|B| = q + a_0 \tfrac{q}{p^e} + a_1 \tfrac{q}{p^{2e}} + \cdots + a_{h/e-2} p^e + 1, \text{ with } a_0, \ldots, a_{h/e-2} \in \mathbb{N}.$$

Note that $|B| = 1 \pmod{p}$, so the constant term will be equal to 1.

The two roots of the quadratic equation on the left hand side of $(*)$ are

$$\frac{qE}{2} + \frac{q}{2} + \frac{E}{2} + 1 \pm \frac{qE}{2} \left(1 - \frac{2}{E} - \frac{3}{E^2} + \frac{2}{q} + \frac{2}{qE} + \frac{1}{q^2}\right)^{1/2}.$$

Now $|B|$ is at most equal to the smallest of the two roots. We also have that

$$\left(1 - \frac{2}{E} - \frac{3}{E^2} + \frac{2}{q} + \frac{2}{qE} + \frac{1}{q^2}\right)^{1/2} \geq (1 - \frac{2}{E} - \frac{3}{E^2})^{1/2} + \frac{1}{q}.$$

Hence,

$$|B| \leq \frac{qE}{2} + \frac{q}{2} + \frac{E}{2} + 1 - \frac{qE}{2} \left\{(1 - \frac{2}{E} - \frac{3}{E^2})^{1/2} + \frac{1}{q}\right\}.$$

From [12], sequence A001006,

$$\left(1 - \frac{2}{E} - \frac{3}{E^2}\right)^{1/2} = 1 - \frac{1}{E} - \frac{2}{E^2} \sum_{n=0}^{+\infty} a_n \frac{1}{E^n},$$

where the coefficients a_n are the Motzkin numbers: $a_0 = 1, a_1 = 1, a_2 = 2, a_3 = 4, a_4 = 9, \ldots$.

Therefore,

$$|B| \leq 1 + q + \frac{q}{E} \sum_{n=0}^{+\infty} a_n \frac{1}{E^n},$$

which gives the upper bound

$$|B| \leq q + a_0 \frac{q}{p^e} + a_1 \frac{q}{p^{2e}} + \cdots + a_{h/e-2} p^e + 1$$

for large values of the prime number p.

As already indicated, the coefficients $a_0, \ldots, a_{h/e-2}$ are known as the *Motzkin numbers* [12]. The first eight Motzkin numbers are $1, 1, 2, 4, 9, 21,\ 51, 127$. For these numbers a_n, we have in general that $a_{n+2} - a_{n+1} = a_0 a_n + a_1 a_{n-1} + \cdots + a_n a_0$. The general expression for a_n is known and equals

$$a_n = \frac{1}{n+1} \sum_{i=0}^{n+1} (-1)^i \binom{n+1}{i} \binom{2n+2-2i}{n-i}.$$

See [15] for this description.

Motzkin numbers appear in many combinatorial problems; we refer to [12] for more references on the Motzkin numbers.

So we have proven the following result.

Theorem 9. *When B is a minimal blocking set in $PG(2, q = p^h)$, p prime, $h \geq 1$, of size $|B| \leq 2q - 1$, intersecting every line in $1 \pmod{p^e}$ points with e the maximal integer for which this is valid, then for large prime numbers p,*

$$|B| \leq q + a_0 \frac{q}{p^e} + a_1 \frac{q}{p^{2e}} + \cdots + a_{h/e-2} p^e + 1,$$

with a_i the i-th Motzkin number.

6 Computer results

We investigated by computer which removal of columns of the incidence matrix A reduces its rank. We use a standard backtracking algorithm in which any x columns are removed recursively. The rank calculation is done by the explicit construction of the vector space of the remaining columns. Adding a column involves a time-consuming diagonalisation algorithm which adds the line incidence vector to the vector space, so it determines if it increments the rank or not.

We illustrate the behavior of our backtracking strategy by an example. Consider the 13×13 incidence matrix of $PG(2, 3)$, with columns numbered from 1 up to 13. We remove 4 columns exhaustively, therefore we generate all ordered removals (a, b, c, d), with $a < b < c < d$. Suppose that the algorithm removed columns 1 and 4. Now we know that columns 2 and 3 will not be removed in this part of the search, therefore we create the vector space V_1 of columns 2 and 3. Suppose that we remove 8 in the next recursive step. We make a copy of V_1 to V_2 and add columns 5, 6 and 7 to V_2. Finally, 10 is the last removed column, therefore we copy V_2 to V_3 and add columns 9, 11, 12 and 13 to V_3. This way parts of the rank calculation are reused.

What about isomorph rejection? We use the well-known *nauty* software [6] to calculate the orbits of the set of non-removed columns with respect to the set

	16	15	14	13	12
5	$5_1, \ldots$				
6 $(p+1)$	$5_1, \ldots$	6_1			
7	$5_1, \ldots$	6_1			
8	$5_1, \ldots$	6_1			
9	$5_2, \ldots$	6_1			
10 $(2p)$	$5_2, \ldots$	$6_1, 5_2*$			
11 $(2p+1)$	$5_2, \ldots$	$6_1, 5_2*$	6_2		
12	$\ldots$	$6_1, 5_2*, 4_3*$	6_2		
13	$\ldots$	$6_1, 5_2, \ldots$	6_2		
14	$\ldots$	$\ldots$	$6_2, 6_1, 5_2*$		
15 $(3p)$	$\ldots$	$\ldots$	$6_2, 6_1, 4_9$	$6_3, 6_2, 5_3*$	
16		$\ldots$	$\ldots$	$6_3, 6_2, 5_3*, 4_{12}$	6_3

Table 1: Exhaustive line removal in $PG(2,5)$, showing what possible rank (table columns) is left when removing a certain amount (table rows) of lines. The meaning of the numbers is explained in the text.

of removed columns. From each orbit in the set of non-removed columns, we choose only one column to remove in the recursive step. To be compatible with the generation method, we remove only the smallest column from each orbit which is larger than the last removed column.

The results of this algorithm on the smallest $PG(2,p)$'s revealed some properties about the removed set of columns. From now on, we use the term "lines" instead of "columns". As an example, Table 1 shows what rank (table columns) is left when removing a certain amount (table rows) of lines. An empty entry indicates no such line removal leads to the rank. Otherwise, a value is the size of the largest subset of concurrent lines of a certain removal, its subscript is the number of such subsets. We use dots when more possibilities than the listed ones are possible. A star (*) indicates the subsets of concurrent lines are disjoint. From the table, we see that when removing less than $2p$ lines, the rank decreases if and only if we remove all lines through a point. Such a removal corresponds to a codeword of weight $p+1$. When removing $2p$ lines, the rank can also decrease by removing all lines through two points, but not the joining line. Such a removal corresponds to a codeword of weight $2p$.

When removing all lines through three points, the rank sometimes decreases by 3 and sometimes by 4. A closer look at all possibilities when removing all lines through three points revealed the following result.

Theorem 10. *If all lines of $PG(2,p)$, p prime, through three collinear points are deleted, then the rank of the incidence matrix decreases by four. If all lines of $PG(2,p)$, p prime, through three non-collinear points are deleted, then the rank of the incidence matrix decreases by three.*

Proof: We prove this by use of the Moorhouse basis. We use the notations of Section 2, i.e. $r_0, r_1, \ldots, r_p$ are the points of the line M defining the affine plane $AG(2,p)$.

Case 1: We delete all lines through the points r_{p-2}, r_{p-1}, r_p of M. For $i \in \mathbb{N},\ 0 \le i \le p-3$, take all lines (different from M) through r_i. These lines give a matrix of rank $\sum_{i=0}^{p-3}(p-i) = (\binom{p+1}{2}+1)-4$. The rank decreased by four.

Case 2: We delete all lines through the points r_{p-1}, r_p and r (r not on M). For the point r_0, we only have $p-1$ lines available for the Moorhouse basis (not $r_0 r$). For $i \in \mathbb{N},\ 1 \le i \le p-2$, we have $p-i$ lines through r_i available for the Moorhouse basis. So the rank is at least $(p-1)+\sum_{i=2}^{p-1} i = \binom{p+1}{2}-2$.

Suppose that we have rank $\binom{p+1}{2}-1$, then by results of Moorhouse [7, Theorem 6.1], we have the net defined by the directions $r_0, \ldots, r_{p-2}$, including the line $r_0 r$. But it is impossible to have $r_0 r$ as a linear combination of the other chosen $\binom{p+1}{2}-2$ lines, because r is not on any of those lines. So the rank is $(\binom{p+1}{2}+1)-3$. The rank decreased by three. $\square$

When removing $3(p-1)$ lines, the rank can also be reduced by removing $p-1$ lines through three points. A closer look gave the following result.

Theorem 11. *If in $PG(2,p)$, p prime, $p-1$ lines through three collinear points a, b and c, but not their joining line, are deleted, then the rank decreases if the three non-removed lines M_1, M_2 and M_3 ($\neq ab$) through respectively a, b and c are concurrent.*

The unique codeword which corresponds to the removal of these lines is, up to equivalence, given by

$$(\underbrace{1,2,\ldots,p-1}_{\text{lines through } a},\underbrace{1,2,\ldots,p-1}_{\text{lines through } b},\underbrace{1,2,\ldots,p-1}_{\text{lines through } c},0,\ldots,0).$$

Proof: Let $M = ab$ be the line at infinity of the corresponding affine plane $AG(2,p)$. Let a, b and c be the points at infinity of respectively the vertical, horizontal, and diagonal lines. Suppose that M_1, M_2, M_3 all pass through the origin $(0,0)$.

We give the coordinate positions of the $p-1$ remaining affine lines through a, b and c the following values, and we prove that the constructed vector indeed is a codeword.

In the coordinate positions of the lines $X = \alpha$, we put the value α. In the coordinate positions of the lines $Y = \beta$, we put the value $-\beta$, and in the coordinate positions of the lines $Y = X + \beta - \alpha$, we put the value $\beta - \alpha$. All other coordinate positions are zero. Note that the coordinate values of the lines M_1, M_2 and M_3 are indeed zero.

Let the incidence matrix A of $PG(2,p)$ have rows corresponding to the points of $PG(2,p)$. We show first of all that the constructed vector c is orthogonal to all the rows of A.

The vector c is orthogonal to the rows of A corresponding to the points a, b and c, since $\sum_{i=1}^{p-1} i \equiv 0 \pmod{p}$. The vector c is also orthogonal to the rows of A corresponding to the other points at infinity since these points lie on none of the lines with non-zero coordinates.

An affine point (a, b) lies on the lines $X = a$, $Y = b$, and $Y = X + b - a$, so the sum of the corresponding coordinate values is $a - b + b - a = 0$.

We have shown that c is orthogonal to all the rows of A, hence $c \in C^{\perp}$.

But $C^{\perp} \subset C$. This is proven in the following way. The code C is a $[p^2+p+1, (p^2+p)/2+1]$-code, so $C^{\perp}$ is a $[p^2+p+1, (p^2+p)/2]$-code. But $\mathrm{Hull}(C) = C \cap C^{\perp}$ is a code of dimension $(p^2+p)/2$ [1, Theorem 6.3.1]. So this shows that $C^{\perp} \subset C$. Hence, $c \in C^{\perp}$ also implies $c \in C$.

This shows that there is a codeword of C with its non-zero positions in the $3(p-1)$ positions of the deleted lines through a, b and c. So, by Theorem 2, the rank of A decreases when deleting these $3(p-1)$ columns from A. □

Theorem 12. *If in $PG(2,p)$, p prime, $p-1$ lines through three collinear points a, b and c, but not their joining line, are deleted, then the rank does not decrease if the three non-removed lines M_1, M_2 and M_3 ($\neq ab$) through respectively a, b and c are non-concurrent.*

Proof: Let r_0 be a point of the line ab, different from a, b and c. Let $\{r_1\} = M_1 \cap M_2$ and let $M = r_0 r_1$. Let $\{r_2\} = M \cap M_3$.

We construct a Moorhouse basis for the affine plane defined by the line M. Through r_0, we have the p necessary lines for the affine Moorhouse basis. Through r_1, we have the $p-1$ necessary affine lines for the Moorhouse basis since the only line through r_1 that cannot be used is the line $r_1 c$. Through r_2, we have the $p-2$ necessary affine lines for the Moorhouse basis since only the lines $r_2 a$ and $r_2 b$ cannot be used. Through all the remaining points of M, we have $p-3$ affine lines available for the Moorhouse basis. Finally, M can be used to construct the final line for the basis of the code of $PG(2,p)$.

So the rank of the incidence matrix of $PG(2,p)$ does not decrease, the $3(p-1)$ deleted lines are not the non-zero positions of a codeword of the p-ary linear code defined by $PG(2,p)$. □

Corollary 1. *If in $PG(2,p)$, p prime, $p-1$ lines through three collinear points a, b and c, but not their joining line, are deleted, then the rank decreases if and only if the three non-removed lines M_1, M_2 and M_3 ($\neq ab$) through respectively a, b and c are concurrent.*

By a similar (easier) construction, we can show that the codeword corresponding to the removal of p lines through 2 points a and b, but not their joining line is

$$(\underbrace{1, 1, \ldots, 1}_{p \text{ lines through } a}, \underbrace{-1, -1, \ldots, -1}_{p \text{ lines through } b}, 0, \ldots, 0).$$

Now we consider removing $p-2$ lines through 4 collinear points (but not their joining line), in which the 8 non-removed lines can be partitioned in two disjoint sets of concurrent lines. Here again, we assume the codeword is such that the linear combination of the incidence vectors of the corresponding removed lines is the $\overline{0}$ vector. The unique codeword was found by an exhaustive computer search for $PG(2,p)$, p prime, $p \leq 23$. The only remarkable thing about these codewords is that, for every $p-2$ lines through a point, we twice have $(p-3)/2$ occurrences of the same value, and then once some other value.

References

[1] E. F. Assmus Jr. and J. D. Key. Designs and their codes. Cambridge: Cambridge University Press, 1992.

[2] S. Ball and A. Blokhuis. On the size of a double blocking set in $PG(2,q)$. *Finite Fields Appl.* **2** (1996), 125–137.

[3] K. L. Chouinard. Weight distributions of codes from planes. Ph.D Thesis, University of Virginia, 2000.

[4] K. L. Chouinard. On weight distributions of codes of planes of order 9. *Ars Combin.* **63** (2002), 3–13.

[5] G. McGuire and H. N. Ward. A determination of the weight enumerator of the code of the projective plane of order 5. *Note Mat.* **18** (1998), no. 1, 71–99.

[6] B. D. McKay. nauty User's Guide (Version 2.2) Computer Science Department, Australian National University, 2004.

[7] G. E. Moorhouse. Bruck nets, codes, and characters of loops. *Des. Codes Cryptogr.* **1** (1991), no. 1, 7–29.

[8] O. Polverino. Small minimal blocking sets and complete k-arcs in $PG(2,p^3)$. *Discrete Math.* **208/209** (1999), 469–476.

[9] O. Polverino. Small blocking sets in $PG(2,p^3)$. *Des. Codes Cryptogr.* **20** (2000), 319–324.

[10] O. Polverino and L. Storme. Small minimal blocking sets in $PG(2,p^3)$. *European J. Combin.* **23** (2002), 83–92.

[11] H. Sachar. Error-correcting codes associated with finite planes. Ph.D Thesis, Lehigh University, 1973.

[12] N. J. A. Sloane. On-line Encyclopedia of Integer Sequences. http://www.research.att.com/∼njas/sequences

[13] P. Sziklai. On small blocking sets and their linearity. *J. Combin. Theory, Ser. A*, submitted.

[14] T. Szőnyi. Blocking sets in Desarguesian affine and projective planes. *Finite Fields Appl.* **3** (1997), 187–202.

[15] E. W. Weisstein. Motzkin Number.
http://mathworld.wolfram.com/MotzkinNumber.html