The Kakeya problem: a gap in the spectrum and classification of the smallest examples

A. Blokhuis[*], M. De Boeck[†], F. Mazzocca[‡], L. Storme

December 16, 2011

Abstract

Kakeya sets in the affine plane $\mathrm{AG}(2,q)$ are point sets that are the union of lines, one through every point on the line at infinity. The finite field Kakeya problem asks for the size of the smallest Kakeya sets and the classification of these Kakeya sets. In this article we present a new example of a small Kakeya set and we give the classification of the smallest Kakeya sets up to weight $\frac{q(q+2)}{2}+\frac{q}{4}$, both in case q even.

1 Introduction

Let $\mathrm{GF}(q)$ be the Galois field with q elements, q a prime power. We denote by $\mathrm{PG}(n,q)$ the n-dimensional projective space over $\mathrm{GF}(q)$, by $H_\infty = \mathrm{PG}(n-1,q)$ a fixed hyperplane of $\mathrm{PG}(n,q)$ (but we write L_∞ if $n=2$) and by $\mathrm{AG}(n,q) = \mathrm{PG}(n,q) \setminus H_\infty$ the n-dimensional affine space over $\mathrm{GF}(q)$.

For every point P on H_∞, let L_P be a line on P not contained in H_∞. The point set

$$\mathcal{K} = (\bigcup_{P\in H_\infty} L_P) \setminus H_\infty$$

is called a *Kakeya set*, or a *minimal Besicovitch set*. The *finite field Kakeya problem* asks for the smallest size $k(n,q)$ of a Kakeya set in $\mathrm{AG}(n,q)$. It is the finite field version of the classical Euclidean *Kakeya problem* (see [15, Section 1.3] for a short survey) and was first posed by *Wolff* in his influential paper [16] of 1996. In the same paper, he conjectured that $k(n,q) \geq c_n q^n$, where $c_n > 0$ is a constant depending only on n. Despite the fact that the conjecture was intensively studied, it remained open for more than ten years and

[*] The research of this author was supported by the Mathematics Department of the Seconda Università degli Studi di Napoli and by GNSAGA of the Italian Istituto Nazionale di Alta Matematica.

[†] The research of this author is supported by FWO-Vlaanderen (Fund for Scientific Research - Flanders).

[‡] The research of this author was supported by the Italian national research project "Geometrie di Galois e Strutture di Incidenza" (COFIN 2008), by the Mathematics Department of the Seconda Università degli Studi di Napoli and by GNSAGA of the Italian Istituto Nazionale di Alta Matematica.

was finally proved by *Dvir* ([4], 2009), using a beautiful argument involving polynomial techniques over finite fields.

Theorem 1.1 (Z. Dvir, 2009)**.** If $\mathcal{K}$ is a Kakeya set in $\mathrm{AG}(n,q)$, then

$$|\mathcal{K}| \geq \binom{q+n-1}{n} \geq \frac{1}{n!}q^n. \tag{1}$$

Dvir's lower bound (1) is not sharp in general and was recently improved in [5] and [13]. The problem of finding the exact value of $k(n,q)$ seems to be very hard and gets more difficult as the dimension n increases. At this moment, it is completely solved only in dimension two and we will give a brief account of this.

Example 1.2. Assume q is odd and consider in $\mathrm{PG}(2,q)$ a dual oval $\mathcal{O}$ (i.e. $q+1$ lines, no three concurrent) and assume $H_\infty = L_\infty$ is a line in $\mathcal{O}$. Under these assumptions, every point $P \in L_\infty$, but one, belongs to a second line $L_P \in \mathcal{O}$ other than L_∞. If A is this remaining point on L_∞, let L_A be any line through it, different from L_∞. Then the Kakeya set

$$(\bigcup_{P\in L_\infty} L_P) \setminus L_\infty$$

has size $\frac{1}{2}q(q+1) + \frac{1}{2}(q-1)$.

In [3], *Blokhuis* and *Mazzocca* characterized the Kakeya sets described in the previous example as the smallest ones in $\mathrm{AG}(2,q)$, q odd.

Theorem 1.3 (A. Blokhuis, F. Mazzocca, 2008)**.** If q is odd, then

$$|\mathcal{K}| \geq \frac{q(q+1)}{2} + \frac{q-1}{2},$$

for every Kakeya set $\mathcal{K}$ in $\mathrm{AG}(2,q)$. Equality holds if and only if $\mathcal{K}$ is associated with a dual oval in $\mathrm{PG}(2,q)$ as in Example 1.2.

Now we describe two ways to obtain a "small" Kakeya set in $\mathrm{AG}(2,q)$, with q even.

Example 1.4. Assume q is even and consider in $\mathrm{PG}(2,q)$ a dual hyperoval $\mathcal{H}$ (i.e. a set of $q+2$ lines, no three concurrent) and assume L_∞ is a line in $\mathcal{H}$. For every point $P \in L_\infty$, let L_P be the line of $\mathcal{H}$ on P other than L_∞. Then the Kakeya set

$$K(\mathcal{H}) = (\bigcup_{P\in L_\infty} L_P) \setminus L_\infty$$

has size

$$|K(\mathcal{H})| = \frac{q(q+1)}{2}.$$

Example 1.5. With the same assumptions and notations of Example 1.4, fix a point $A \in L_\infty$ and a line L' through A different from L_A and L_∞. Then the Kakeya set

$$K(\mathcal{H}, L') = (K(\mathcal{H}) \setminus L_A) \cup (L' \setminus L_\infty)$$

has size

$$|K(\mathcal{H}, L')| = \frac{q(q+2)}{2}\,.$$

When q is even, it is easy to prove that $k(2,q) = \frac{1}{2}q(q+1)$ and equality occurs only for the Kakeya sets described in Example 1.4. Moreover, in [2], *Blokhuis* and *Bruen* proved the following result (stated in its dual form).

Theorem 1.6 (A. Blokhuis, A.A. Bruen, 1989)**.** There are no Kakeya sets $\mathcal{K}$ in $\mathrm{AG}(2,q)$, with $\frac{1}{2}q(q+1) < |\mathcal{K}| < \frac{1}{2}q(q+2)$. Furthermore, all Kakeya sets of size $\frac{1}{2}q(q+2)$ are given by Example 1.5.

The aim of the present article is to determine the Kakeya sets $\mathcal{K}$ with $\frac{1}{2}q(q+2) < |\mathcal{K}| \leq \frac{1}{2}q(q+2) + \frac{1}{4}q$. We will prove that in $\mathrm{AG}(2,q)$, q even, there are no Kakeya sets whose size belongs to the corresponding open interval and we will characterize those of size $\frac{1}{2}q(q+2) + \frac{1}{4}q$.

2 Preliminaries

Definition 2.1. A *k-arc* (or simply arc) in $\mathrm{PG}(2,q)$ is a set of k points, no three of which are collinear. An arc is called *complete* if it is not contained in a larger arc. A *tangent* line to an arc is a line intersecting the arc in precisely one point.

Arcs have been intensively studied in the past decades and many results are known. For an overview, see for example [8]. We mention some results about arcs, that we will need. The first one is given in [9]. In $\mathrm{PG}(2,q)$, q even, a *hyperoval* is a complete $(q+2)$-arc.

Theorem 2.2 (B. Segre)**.** Every k-arc in $\mathrm{PG}(2,q)$, q even, with $k > q-\sqrt{q}+1$, is contained in a hyperoval and hence not complete if $k < q+2$.

In [8, Section 10], *the tangent envelope* of an arc is introduced. This is the algebraic envelope (dual curve) containing all the lines tangent to this arc. The tangent envelope of a k-arc is of class $q+2-k$ if q is even. Dualizing this, we find a tangent curve to a dual k-arc, containing all points which are covered precisely once by the lines of the dual arc. This is an algebraic curve of degree $q+2-k$ if q is even. The following theorem is proved in [8] in the setting of arcs and tangent envelopes, but we state it immediately in the setting of dual arcs and tangent curves.

Theorem 2.3 ([8, Corollary 10.3])**.** Let $\mathcal{A}$ be a dual k-arc in $\mathrm{PG}(2,q)$, q even and $k > \frac{1}{2}q+1$, and let Γ_t be the tangent curve to this dual arc. The line L extends $\mathcal{A}$ if and only if L is a component of Γ_t.

The following lemma uses this theorem about the tangent curve to a dual arc.

Lemma 2.4. Let $\mathcal{A}$ be a dual k-arc in $\mathrm{PG}(2,q)$, q even, with $k > \frac{1}{2}q + 1$. A line, not extending $\mathcal{A}$, contains at least $\frac{1}{2}q$ points not lying on lines of $\mathcal{A}$.

Proof. Let Γ_t be the tangent curve of $\mathcal{A}$. Then Γ_t is an algebraic curve of degree $t = q+2-k$. By Theorem 2.3, a line extending $\mathcal{A}$ is a component of Γ_t and vice versa. Consider a line L not extending $\mathcal{A}$. It intersects Γ_t in x points, with $x \leq t$. These points are the ones lying on precisely one line of $\mathcal{A}$. Consequently, $\frac{1}{2}(k-x)$ points of L are lying on two lines of $\mathcal{A}$. Hence, the number of points not on $\mathcal{A}$ equals $(q+1) - x - \frac{1}{2}(k-x) = q - \frac{1}{2}(k+x) + 1$. Using the bound on x, we find that $q - \frac{1}{2}(k+x) + 1 \geq q - \frac{1}{2}(k+t) + 1 = \frac{1}{2}q$. The lemma follows. □

After its introduction, the concept of arcs has been generalised. We will not discuss this in general, but we mention the following special type, introduced by *Korchmáros* and *Mazzocca* in 1990 ([11]).

Definition 2.5. A $(q+t,t)$-*arc of type* $(0,2,t)$ in $\mathrm{PG}(2,q)$, q even (and $t|q$), is a set of $q+t$ points intersecting any line in 0, 2 or t points.

In [7], it is proved that a $(q+t,t)$-arc of type $(0,2,t)$ has a t-*nucleus*, the common point of all its t-secants. However, it remains an open problem for which pairs (q,t) they exist and how they can be classified. For example, for $t=4$, we only know examples for $q \leq 32$ (for $q = 8, 16$: see [11]; $q = 32$: see [10]).

There are important links between finite geometry and coding theory. Important for us is the code of the plane.

Definition 2.6. Consider the plane $\mathrm{PG}(2,q)$, q even. Let $M_{2,q}$ be the $\mathrm{GF}(2)$-matrix whose rows are labelled by the lines and whose columns are labelled by the points of $\mathrm{PG}(2,q)$ and such that

$$(M_{2,q})_{i,j} = \begin{cases} 1 & \text{if line } i \text{ contains point } j, \\ 0 & \text{otherwise.} \end{cases}$$

This matrix is called *the incidence matrix* of the plane. The binary code generated by the rows of this matrix will be denoted by $C(2,q)$. It is called *the code generated by the points and lines of* $\mathrm{PG}(2,q)$.

This code has been the subject of a lot of research. For a survey, see for example [1, 12]. We will need the following theorem.

Theorem 2.7 ([1, Corollary 6.4.4])**.** Let C be the code $C(2,q)$, q even. The minimum weight of $C \cap C^{\perp}$ is $2q$ and the minimum-weight vectors are obtained by taking the difference of the incidence vectors of two lines.

3 Classifying the next example

Now we describe a Kakeya set, which we will prove to be the (theoretical) third smallest example (provided that it exists).

Example 3.1. Let $\mathcal{A}$ be a dual $(q{+}4)$-arc of type $(0,2,4)$ in $\mathrm{PG}(2,q)$, and let L_0, L_1, L_2, L_∞ be four concurrent lines of $\mathcal{A}$. Consider the affine plane $\mathrm{AG}(2,q) = \mathrm{PG}(2,q) \setminus L_\infty$. Let $\mathcal{A}'$ be the line set $\mathcal{A} \setminus \{L_1, L_2\}$. Consider the set

$$K(\mathcal{A}, L_1, L_2) = \bigcup_{L \in \mathcal{A}'} (L \setminus L_\infty) .$$

This is a Kakeya set since there is precisely one line of $K(\mathcal{A}, L_1, L_2)$ through every point of L_∞. It has size $\frac{1}{2}q(q+2) + \frac{1}{4}q$.

Lemma 3.2. Let $\mathcal{K} = (\cup_{i=0}^{q} L_i) \setminus L_\infty$ be a Kakeya set in $\mathrm{AG}(2,q)$, such that its corresponding line set $\mathcal{L} = \{L_0, \ldots, L_q\}$ contains a dual x-arc, but no dual $(x+1)$-arc. Then $\mathcal{K}$ contains at least $\frac{1}{2}(q+4)(q+1) - \left\lfloor \frac{1}{2}x \right\rfloor - 2x$ points.

Proof. Let $\mathcal{A} = \{L_0, \ldots, L_{x-1}\}$ be a dual x-arc contained in $\mathcal{L}$. If we construct the Kakeya set line by line in the order $L_0, \ldots, L_q$, adding the $(i+1)$-th line L_i increases the number of points in $\mathcal{K}$ by $q - i + m_i$, $m_i \geq 0$. Since $\{L_0, \ldots, L_{x-1}\}$ is a dual x-arc, $m_i = 0$ for $i = 0, \ldots, x-1$. For the line $L_i \in \{L_x, \ldots, L_q\}$, we know $m_i \geq 1$ since none of these lines extends $\mathcal{A}$.

Let L_k be a line of $\mathcal{K}$, with $m_k = 1$. Then L_k contains precisely one intersection point $L_i \cap L_j$, $i, j < k$, of previously added lines. Assume one of these two lines, say L_i, is not contained in $\mathcal{A}$, or equivalently $i \geq x$. Then the line L_k extends $\mathcal{A}$ because it does not go through an intersection point of two lines of $\mathcal{A}$. This is a contradiction since $\mathcal{A}$ does not contain a dual $(x+1)$-arc. Hence, each line $L_k \in \mathcal{K}$, with $m_k = 1$, contains precisely one intersection point of two lines of $\mathcal{A}$. Let $\mathcal{B}$ be the set $\{L_j \mid m_j = 1\}$. The points lying on two lines of $\mathcal{A}$ and a line of $\mathcal{B}$ are called *complete* points.

Let L_a, L_b, L_c be three lines of $\mathcal{A}$ and let L_k, L_l be two lines of $\mathcal{B}$ such that L_k goes through $L_a \cap L_b$ and L_l goes through $L_a \cap L_c$. In other terms, $L_a \in \mathcal{A}$ contains two different complete points. Consider the line set $(\mathcal{A} \setminus \{L_a\}) \cup \{L_k, L_l\}$. This line set is a dual arc since $\mathcal{A}$ is a dual arc, and the lines L_k and L_l each contain precisely one intersection point of the lines of $\mathcal{A}$, both lying on L_a. However, this line set contains $x+1$ lines and is a subset of $\mathcal{L}$. This is a contradiction since we know $\mathcal{L}$ contains no dual $(x+1)$-arc. Hence, a line of $\mathcal{A}$ contains at most one complete point. Consequently, $|\mathcal{B}| \leq \left\lfloor \frac{1}{2}x \right\rfloor$.

From the previous arguments, it follows that $|\{L_j \mid m_j \geq 2\}| = (q+1) - x - |\mathcal{B}|$. So, we conclude

$$\begin{aligned}|\mathcal{K}| = \sum_{i=0}^{q}(q-i) + \sum_{i=0}^{q} m_i &\geq \frac{q(q+1)}{2} + |\mathcal{B}| + 2\cdot((q+1) - x - |\mathcal{B}|) \\ &= \frac{(q+4)(q+1)}{2} - |\mathcal{B}| - 2x \\ &\geq \frac{(q+4)(q+1)}{2} - \left\lfloor \frac{x}{2} \right\rfloor - 2x.\end{aligned}$$

□

Lemma 3.3. Let $\mathcal{K} = (\cup_{i=0}^{q} L_i) \setminus L_\infty$ be a Kakeya set in $\mathrm{AG}(2,q) = \mathrm{PG}(2,q) \setminus L_\infty$, $q > 8$ even, with $|\mathcal{K}| \leq \frac{1}{2}q(q+2) + \frac{1}{4}q$, and assume that the line set $\mathcal{T} = \{L_0, \dots, L_q, L_\infty\}$ is not a dual hyperoval of $\mathrm{PG}(2,q)$. Then $\mathcal{T} \setminus \{L_\infty\}$ contains a dual q-arc or a dual $(\frac{1}{2}q+1)$-arc, not extendable to a larger arc by the remaining lines of $\mathcal{T} \setminus \{L_\infty\}$.

Proof. In the following, for every $j \in \{0, 1, \dots, q\}$, we set

$$\mathcal{L} = \mathcal{T} \setminus \{L_\infty\} = \{L_0, \dots, L_q\}, \quad S_j = \left(\bigcup_{i=0}^{j} L_i\right) \setminus L_\infty, \quad |\mathcal{K}| = |S_q| = \frac{q(q+1)}{2} + \varepsilon,$$

and we assume

$$0 < \varepsilon \leq \frac{3}{4}q\,.$$

Then, $|S_j \setminus S_{j-1}| = q - j + m_j$, with $m_j \geq 0$, for $j = 1, 2, \dots, q$. In other terms, passing from S_{j-1} to S_j by the addition of the $(j+1)$-th line L_j, the number of covered points increases by $q - j + m_j$. Moreover, a direct computation shows that

$$\frac{q(q+1)}{2} + \varepsilon = \sum_{i=0}^{q}(q - i + m_i) = \frac{q(q+1)}{2} + \sum_{i=0}^{q} m_i. \tag{2}$$

Denote by k, $k < q+1$, the maximal integer for which L_∞ and k lines in $\mathcal{L}$ form a dual $(k+1)$-arc in $\mathrm{PG}(2,q)$ and, without loss of generality, assume that $\overline{\mathcal{A}} = \mathcal{A} \cup \{L_\infty\}$, with $\mathcal{A} = \{L_0, \dots, L_{k-1}\}$, is such a dual $(k+1)$-arc. Under this assumption, because each of the lines in $\mathcal{A}$ intersects the union of the remaining ones in exactly $k-1$ affine points, we have $m_i = 0$ for $i = 0, 1, \dots, k-1$. Moreover, because of the maximality of $\overline{\mathcal{A}}$ as a dual arc contained in $\mathcal{T}$, for $j \geq k$, no line L_j extends $\overline{\mathcal{A}}$ and consequently $m_j \neq 0$ for $j \geq k$.

Now, we distinguish two cases: $k \leq \frac{1}{2}q$ and $k \geq \frac{1}{2}q + 1$. For $k \leq \frac{1}{2}q$, we apply Lemma 3.2. We find that $|\mathcal{K}| \geq \frac{1}{2}(q+4)(q+1) - \frac{1}{4}q - q = \frac{1}{2}q(q+2) + \frac{1}{4}q + 2$. Hence, this possibility cannot occur. Now, we look at the case $k \geq \frac{1}{2}q + 1$. Because $k + 1 > k \geq \frac{1}{2}q + 1$, we can apply Lemma 2.4. Each of the lines $L_k, L_{k+1}, \dots, L_q$ contains at least $\frac{1}{2}q$ points not on $\overline{\mathcal{A}}$.

Moreover, setting $\mathcal{K}' = (\cup_{i=0}^{k-1} L_i) \setminus L_\infty$ and counting the number of points in $\mathcal{K}$, we find

$$\begin{aligned}\frac{q(q+1)}{2} + \varepsilon = |\mathcal{K}| &= |\mathcal{K}'| + |(\cup_{j=k}^{q} L_j) \setminus (\mathcal{K}' \cup L_\infty)| \\ &\geq [q + (q-1) + \cdots + (q-k+1)] + \left[\frac{q}{2} + (\frac{q}{2} - 1) + \cdots + (\frac{q}{2} - (q-k))\right] \qquad (3) \\ &= \frac{k(3q-2k+2)}{2} = f(k).\end{aligned}$$

Note that $k = q+1$ would imply that $\overline{\mathcal{A}}$ is a dual hyperoval and that $\varepsilon = 0$. For $k \in [\frac{1}{2}q + 2, q-1]$, we find $f(k) \geq \frac{1}{2}q(q+3) - 2 > \frac{1}{2}q(q+2) + \frac{1}{4}q \geq \frac{1}{2}q(q+1) + \varepsilon$. Consequently, $k \in \{\frac{1}{2}q + 1, q\}$. □

Note that $f(\frac{1}{2}q+1) = f(q) = \frac{1}{2}q(q+2)$, with f as in the previous proof. This proves that $|\mathcal{K}| \notin \left[\frac{1}{2}q(q+1)+1, \frac{1}{2}q(q+2)-1\right]$, which is part of the result of Blokhuis and Bruen (Theorem 1.6).

Lemma 3.4. Let $\mathcal{K} = (\cup_{i=0}^{q} L_i) \setminus L_\infty$ be a Kakeya set in $\mathrm{AG}(2,q) = \mathrm{PG}(2,q) \setminus L_\infty$, $q > 8$ even. Then $|\mathcal{K}| \notin \left[\frac{1}{2}q(q+2)+1, \frac{1}{2}q(q+2) + \frac{1}{4}q - 1\right]$.

Proof. We use the notation introduced in Lemma 3.3. Assume $\mathcal{K}$ covers precisely $\frac{1}{2}q(q+2) + \varepsilon'$ points, $0 \leq \varepsilon' < \frac{1}{4}q$. By Lemma 3.3, there are two cases: $k = q$ or $k = \frac{1}{2}q + 1$. In the first case, $\overline{\mathcal{A}}$ is a dual $(q+1)$-arc in $\mathrm{PG}(2,q)$ containing L_∞. By Theorem 2.2, this dual arc is contained in a unique dual hyperoval $\mathcal{H} = \overline{\mathcal{A}} \cup \{M\}$. Then $M \cap L_\infty = L_q \cap L_\infty$ since both $\mathcal{H}$ and $\mathcal{K}$ are Kakeya sets containing $\overline{\mathcal{A}}$. Obviously, $M \neq L_q$. Hence, the Kakeya set $\mathcal{K}$ is of the type given in Example 1.5 and $|\mathcal{K}| = \frac{1}{2}(q+2)q$. Note that in this case, the inequality in (3) is an equality, $\epsilon = \frac{1}{2}q$ and $\epsilon' = 0$.

Now, we look at the case $k = \frac{1}{2}q + 1$. We apply Lemma 3.2 and we find

$$|\mathcal{K}| \geq \frac{(q+4)(q+1)}{2} - \left\lfloor \frac{1}{4}q + \frac{1}{2} \right\rfloor - 2\left(\frac{1}{2}q + 1\right) = \frac{q(q+2)}{2} + \frac{q}{4}$$

The lemma follows from these observations. □

Lemma 3.5. Let $\mathcal{K} = (\cup_{i=0}^{q} L_i) \setminus L_\infty$ be a Kakeya set in $\mathrm{AG}(2,q) = \mathrm{PG}(2,q) \setminus L_\infty$, $q > 8$ even, with $|\mathcal{K}| = \frac{1}{2}q(q+2) + \frac{1}{4}q$. Then $\mathcal{K}$ is a Kakeya set of the type given in Example 3.1.

Proof. We use the notation we introduced in Lemma 3.3 and Lemma 3.4. We recall that $\mathcal{L}$ is the line set $\{L_0, \ldots, L_q\}$. By the results of these lemmata and the arguments used in their proofs, we know that $\mathcal{L}$ contains a dual $(\frac{1}{2}q+1)$-arc $\mathcal{A} = \{L_0, \ldots, L_{\frac{q}{2}}\}$. Furthermore, $m_j = 1$ for $\frac{1}{2}q + 1 \leq j \leq k'$ and $m_j \geq 2$ for $k'+1 \leq j \leq q$. Just as in the preceding lemmata, every line L_j, $\frac{1}{2}q + 1 \leq j \leq k'$, contains precisely one intersection point of the lines of $\mathcal{A}$. Those intersection points were called complete points. Again arguing as in Lemma 3.2, we know every line of $\mathcal{A}$ contains at most one complete point, hence, $k' \leq \frac{1}{4}q$. Using (2), we then obtain $k' = \frac{1}{4}q$ and $m_j = 2$ for $\frac{3}{4}q + 1 \leq j \leq q$. Thus, there are precisely

$\frac{1}{4}q$ complete points and all but one of the lines in $\mathcal{A}$ contain a complete point. Let L_0 be the line without a complete point and let $\mathcal{A}'$ be the line set $\mathcal{A} \cup \{L_{\frac{q}{2}+1}, \ldots, L_{\frac{3q}{4}}\}$.

For a line in $\mathcal{B} = \{L_j \mid \frac{3}{4}q + 1 \leq j \leq q\}$, there are two possibilities. Either, such a line contains a complete point and no other intersection point of two lines of $\mathcal{A}'$, or else it does not contain a complete point, but it contains two intersection points of two pairs of lines of $\mathcal{A}'$. Let $\mathcal{B}^* = \{L_j \mid \frac{3}{4}q + 1 \leq j \leq \frac{3}{4}q + y\}$ be the set of the former lines and $\mathcal{B}^- = \{L_j \mid \frac{3}{4}q + y + 1 \leq j \leq q\}$ be the set of the latter lines. Remark that we first add the lines of $\mathcal{B}^*$. The complete points lying on a line of $\mathcal{B}^*$ will be called *hypercomplete* points, and the intersection points of two lines of $\mathcal{A}'$, that are not complete points, but are lying on a line of $\mathcal{B}^-$, are called *new complete* points. It follows that there are y hypercomplete points, $\frac{1}{4}q - y$ complete points that are not hypercomplete, and $2(\frac{1}{4}q - y) = \frac{1}{2}q - 2y$ new complete points.

Since a line of $\mathcal{A}' \setminus \{L_0\}$ contains precisely one complete point before adding the lines of $\mathcal{B}$, it contains precisely one complete point, which is possibly hypercomplete. We will prove some properties of the hypercomplete and new complete points. Note that a point cannot be (hyper)complete and new complete at the same time.

- Firstly, we prove that a line of $\mathcal{A}'$ cannot contain a hypercomplete point and a new complete point. Let $L_i \in \mathcal{A} \setminus \{L_0\}$ be a line containing a hypercomplete point $L_i \cap L_j \cap L_n \cap L_p$ and a new complete point $L_i \cap L_s \cap L_r$, with $L_j \in \mathcal{A}$, $L_n \in \mathcal{A}' \setminus \mathcal{A}$, $L_p \in \mathcal{B}^*$, $L_s \in \mathcal{A}'$ and $L_r \in \mathcal{B}^-$. Consider now the ordering

$$\sigma = L_0, \ldots, L_{i-1}, L_{i+1}, \ldots, L_{\frac{1}{2}q}, L_n, L_{\frac{1}{2}q+1}, \ldots, L_{\frac{3}{4}q}, L_p, L_r, L_{\frac{3}{4}q+1}, \ldots, L_q, L_i.$$

  Remark that it is not indicated where L_n, L_p and L_r are removed, but this can easily be seen. Using this alternative ordering, we can define m_a^σ for the line L_a, the same way we defined m_i in the proof of Lemma 3.3. We find that $m_0^\sigma = \ldots = m_{i-1}^\sigma = m_{i+1}^\sigma = \ldots = m_{\frac{1}{2}q}^\sigma = m_n^\sigma = 0$, $m_{\frac{1}{2}q+1}^\sigma = \ldots = m_{n-1}^\sigma = m_{n+1}^\sigma = \ldots = m_{\frac{3}{4}q}^\sigma = m_p^\sigma = m_r^\sigma = 1$ and $m_i^\sigma = 3$. This is a contradiction since also for this ordering $m_a^\sigma \leq 2$, $0 \leq a \leq q$. Now, let $L_i \in \mathcal{A}' \setminus \mathcal{A}$ be a line containing a hypercomplete point $L_i \cap L_j \cap L_n \cap L_p$ and a new complete point $L_i \cap L_s \cap L_r$, with $L_j, L_n \in \mathcal{A}$, $L_p \in \mathcal{B}^*$, $L_s \in \mathcal{A}'$ and $L_r \in \mathcal{B}^-$. Consider the ordering

$$\tau = L_0, \ldots, L_{i-1}, L_{i+1}, \ldots, L_{\frac{3}{4}q}, L_p, L_r, L_{\frac{3}{4}q+1}, \ldots, L_q, L_i.$$

  Again, we can define m_a^τ for a line L_a. We find $m_i^\tau = 3$, which is a contradiction.

- Similarly, we can also prove that a line cannot contain a complete point, which is possibly hypercomplete, and two new complete points. It is obvious that a line of $\mathcal{B} \cup \{L_0\}$ cannot contain a (hyper)complete point and two new complete points. Let $L_i \in \mathcal{A} \setminus \{L_0\}$ be a line containing a complete point $L_i \cap L_j \cap L_{j'}$ and two new complete points $L_i \cap L_n \cap L_{n'}$ and $L_i \cap L_p \cap L_{p'}$, with $L_j \in \mathcal{A}$, $L_{j'} \in \mathcal{A}' \setminus \mathcal{A}$, $L_p, L_n \in \mathcal{A}'$ and $L_{n'}, L_{p'} \in \mathcal{B}^-$. Consider the ordering

$$\sigma' = L_0, \ldots, L_{i-1}, L_{i+1}, \ldots, L_{\frac{1}{2}q}, L_{j'}, L_{\frac{1}{2}q+1}, \ldots, L_{\frac{3}{4}q}, L_{n'}, L_{p'}, L_{\frac{3}{4}q+1}, \ldots, L_q, L_i.$$

As before it is not indicated where $L_{j'}$, $L_{n'}$ and $L_{p'}$ are removed. We define $m_a^{\sigma'}$ for L_a using this ordering σ'. There are $\frac{1}{2}q+1$ lines with $m_a^{\sigma'}=0$ (the lines of $(\mathcal{A}\setminus\{L_i\})\cup\{L_{j'}\}$). However, $m_i^{\sigma'}=3$, a contradiction. If the complete point on L_i is hypercomplete, then there is a line $L_{j''}\in\mathcal{B}^*$ through $L_i\cap L_j\cap L_{j'}$. In the new ordering, we then insert $L_{j''}$ between $L_{p'}$ and $L_{\frac{3}{4}q+1}$. Then we find $m_i^{\sigma'}=4$, a contradiction. Now let $L_i\in\mathcal{A}'\setminus\mathcal{A}$ be a line containing a complete point $L_i\cap L_j\cap L_{j'}$ and two new complete points $L_i\cap L_n\cap L_{n'}$ and $L_i\cap L_p\cap L_{p'}$, with $L_j,L_{j'}\in\mathcal{A}$, $L_p,L_n\in\mathcal{A}'$ and $L_{n'},L_{p'}\in\mathcal{B}^-$. In this case, we consider the ordering

$$\tau'=L_0,\dots,L_{i-1},L_{i+1},\dots,L_{\frac{3}{4}q},L_{p'},L_{n'},L_{\frac{3}{4}q+1},\dots,L_q,L_i.$$

Defining as before $m_a^{\tau'}$ for the line L_a, we find $m_i^{\tau'}=3$, a contradiction. Also in this case, the complete point is allowed to be hypercomplete.

- Finally, we prove that the line L_0 cannot contain new complete points. Remark first that the lines of $\mathcal{A}'\setminus\{L_0\}$ can be partitioned in $\frac{1}{4}q$ sets of 3 lines going through a common complete point. Two of these lines belong to $\mathcal{A}$ and one belongs to $\mathcal{A}'\setminus\mathcal{A}$. Let C_a be the set of three lines containing the complete point on $L_a\in\mathcal{A}'\setminus\{L_0\}$. By swapping their positions in the ordering of the lines in $\mathcal{A}'$, each of the lines can be chosen to be the one in $\mathcal{A}'\setminus\mathcal{A}$.

  Now, assume that L_0 contains a new complete point $L_0\cap L_i\cap L_r$, with $L_i\in\mathcal{A}'\setminus\{L_0\}$ and $L_r\in\mathcal{B}^-$. Let $L_j\cap L_k\cap L_r$ be the second new complete point on L_r, with $L_j,L_k\in\mathcal{A}'\setminus\{L_0\}$. Since $L_j\cap L_k$ is not a complete point, the sets C_j and C_k are different. So, at most one of them equals C_i. Without loss of generality, we can assume that C_j and C_i are different (and hence disjoint). Thus, by the above, we can choose simultaneously both L_i and L_j to be in $\mathcal{A}'\setminus\mathcal{A}$. However, then the set $\mathcal{A}\cup\{L_r\}$ is a dual $(\frac{1}{2}q+2)$-arc contained in $\mathcal{L}$, a contradiction to Lemma 3.3.

Define the set $S'=\{(p,L)\mid p \text{ a hypercomplete point}, L\in\mathcal{L}\setminus\{L_0\}, p\in L\}$. We count the number of elements in this set in two ways. On the one hand, we find $|S'|=4y$ since every hypercomplete point lies on precisely four lines of $\mathcal{L}$, none of which is L_0. On the other hand, we find $|S'|\le y+\left(\frac{3}{4}q-2\cdot(\frac{1}{2}q-2y)\right)$ since every line of $\mathcal{B}^*$ contains one hypercomplete point, none of the lines of $\mathcal{B}^-$ contains a hypercomplete point and none of the $\frac{3}{4}q$ lines in $\mathcal{A}'\setminus\{L_0\}$ contains a hypercomplete and a new complete point. Moreover, every line in $\mathcal{A}'\setminus\{L_0\}$ contains a complete point (possibly hypercomplete), hence contains at most one new complete point. Consequently, all the $2\cdot(\frac{1}{2}q-2y)$ lines of $\mathcal{A}'$ through a new complete point are different and none of them is equal to L_0 by the last of the above properties.

Thus, we find $4y\le 5y-\frac{1}{4}q$. Hence, $y\ge\frac{1}{4}q$; consequently $\mathcal{B}^-$ is empty, $|\mathcal{B}^*|=\frac{1}{4}q$, there are no new complete points and all $\frac{1}{4}q$ complete points are hypercomplete. Since a line of $\mathcal{A}'$ contains at most one complete point regarding the lines of $\mathcal{A}'$, a line of $\mathcal{L}$ contains at most one hypercomplete point. Hence, the lines of $\mathcal{L}\setminus\{L_0\}$ can be partitioned in $\frac{1}{4}q$ groups of four lines, each going through a common (hypercomplete) point. Furthermore, there are precisely $\frac{1}{4}q$ points lying on 4 lines of $\mathcal{L}$ (the hypercomplete ones), there are $2q$ points lying on precisely one line of $\mathcal{L}$ (2 on each line through a hypercomplete point and none on L_0) and there are $\frac{1}{2}q(q-2)$ points on precisely two lines of $\mathcal{L}$.

Consider the binary code $C = C(2,q)$ generated by the lines and points of $\mathrm{PG}(2,q)$ (the points correspond to the positions). Let c be the codeword which is the sum of the (incidence vectors of) lines of $\mathcal{L} \cup \{L_\infty\}$. This corresponds to the set of points which are covered precisely once by the lines of $\mathcal{L} \cup \{L_\infty\}$. By the previous arguments, this is a codeword of weight $2q$. Moreover, c is also a codeword of $C^\perp$ since it can be written as the sum of $\frac{1}{2}q+1$ differences of incidence vectors of two lines. Using Theorem 2.7, we find that c is the difference of the incidence vectors of two lines. Thus, the points covered only once by the lines of $\mathcal{L}$ are lying on two lines. Denote these two lines by M and M'. Then, M and M' intersect each of the lines $L_1, \dots, L_q$ in an affine point since $L_1, \dots, L_q$ each contain two points lying on precisely one line of $\mathcal{L}$. Consequently, $M \cap L_\infty = M' \cap L_\infty = L_0 \cap L_\infty$.

Now, we consider the line set $\{L_0, \dots, L_q, L_\infty, M, M'\}$. This is a set of $q+4$ lines in $\mathrm{PG}(2,q)$ such that every point is contained in 0, 2 or 4 lines of the set. Hence, this is a dual $(q+4,4)$-arc of type $(0,2,4)$. We conclude that the Kakeya set is of the type described in Example 3.1. □

We summarize the known results about the smallest Kakeya sets in the next theorem.

Theorem 3.6. Let $\mathcal{K}$ be a Kakeya set in $\mathrm{AG}(2,q) = \mathrm{PG}(2,q) \setminus L_\infty$, $q > 8$ even. Then, only the following possibilities can occur.

- $|\mathcal{K}| = \frac{1}{2}q(q+1)$ and $\mathcal{K}$ arises from a dual hyperoval.
- $|\mathcal{K}| = \frac{1}{2}q(q+2)$ and $\mathcal{K}$ is a Kakeya set of the type given in Example 1.5.
- $|\mathcal{K}| = \frac{1}{2}q(q+2) + \frac{1}{4}q$ and $\mathcal{K}$ is a Kakeya set of the type given in Example 3.1.
- $|\mathcal{K}| \geq \frac{1}{2}q(q+2) + \frac{1}{4}q + 1$.

We have a look at the smallest cases for q, that are not covered by this theorem.

For $q = 2$, Theorem 1.6 classifies all Kakeya sets since $\frac{1}{2}q(q+2) = 4 = |\mathrm{AG}(2,q)|$ in this case.

For $q = 4$, Theorem 1.6 classifies the Kakeya sets of size 10 and 12, and excludes size 11. Moreover, Lemma 3.4 is trivially valid because all other Kakeya sets have size at least $13 = \frac{1}{2}q(q+2) + \frac{1}{4}q$. However, the Kakeya sets of size 13 have not been classified.

For $q = 8$, Theorem 1.6 classifies the Kakeya sets of size 36 and 40, and excludes the sizes 37, 38 and 39. In this case, $\frac{1}{2}q(q+3) - 2 = 42 = \frac{1}{2}q(q+2) + \frac{1}{4}q$, so the proof of Lemma 3.3 does not continue. However, it does follow that a Kakeya set of size 41 contains a dual 8-arc or a dual 5-arc that is not extendable to a dual 6-arc with an affine line of $\mathcal{K}$. This is enough for the proof of Lemma 3.4 and hence we can exclude the size 41. Kakeya sets of the type given in Example 3.1 have size 42, but it is not proved that this is the only possibility for a Kakeya set of that type.

References

[1] E.F. Assmus, J.D. Key: Designs and their codes, Cambridge University Press, New York, NY, (1992).

[2] A. Blokhuis, A.A. Bruen: The minimal number of lines intersected by a set of $q+2$ points, blocking sets, and intersecting circles, *J. Combin. Theory Ser. A,* **50**, 308-315, (1989).

[3] A. Blokhuis, F. Mazzocca: The Finite Field Kakeya Problem, in Building Bridges Between Mathematics and Computer Science, *Bolyai Society Mathematical Studies,* Vol. 19, M. Grötschel, G.O.H. Katona, (Eds.), 205-218, (2008).

[4] Z. Dvir: On the Size of Kakeya Sets in Finite Fields. *J. AMS,* **22**, 1093-1097, (2009).

[5] Z. Dvir, S. Kopparty, S. Saraf, M. Sudan: Extensions to the method of multiplicities, with applications to Kakeya sets and mergers, in FOCS 09 (to appear), (2009).

[6] X.W.C. Faber: On the Finite Field Kakeya Problem in Two Dimensions, *J. Number Theory*, **117**, 471-481, (2006).

[7] A. Gács, Zs. Weiner: On $(q+t)$-arcs of type $(0,2,t)$, *Des. Codes Cryptogr.*, **29(1-3)**, 131-139, (2003).

[8] J.W.P. Hirschfeld: Projective Geometries over Finite Fields, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, second edition, 1998, xiv+555pp.

[9] J.W.P. Hirschfeld, G. Korchmáros: Arcs and curves over a finite field, *Finite Fields Appl.*, **5**, 393-408, (1999).

[10] J.D. Key, T.P. McDonough, V.C. Mavron: An upper bound for the minimum weight of the dual codes of desarguesian planes, *European J. Combin.*, **30(1)**, 220-229, (2009).

[11] G. Korchmáros, F. Mazzocca: On $(q+t)$-arcs of type $(0,2,t)$ in a desarguesian plane of order q, *Math. Proc. Camb. Phil. Soc.*, **108(3)**, 445-459, (1990).

[12] M. Lavrauw, L. Storme, G. Van de Voorde: Linear codes from projective spaces, in Error-Correcting Codes, Finite Geometries, and Cryptography, *AMS Contemporary Mathematics (CONM) book series*, vol. 523, A.A. Bruen, D.L. Wehlau (Eds.), 185-202, (2010).

[13] S. Saraf, M. Sudan: Improved lower bound on the size of Kakeya sets over finite fields, *Analysis and PDE*, **1(3)**, 375-379, (2008).

[14] B. Segre: Ovals in Finite Projective Planes, *Canad. J. Math.*, **7**, 414-416, (1955).

[15] T. Tao: Poincarés legacies: pages from year two of a mathematical blog, American Mathematical Society, Vol. I, (2009).

[16] T. Wolff: Recent Work Connected with the Kakeya Problem, *Prospects in Mathematics* (Princeton, NJ, 1996), Amer. Math. Soc., Providence, RI, (1999), 129-162.

Addresses of the authors:

A. Blokhuis
Department of Mathematics
Eindhoven University of Technology
P. O. Box 513
5600 MB Eindhoven
The Netherlands
(aartb@win.tue.nl, http://www.win.tue.nl/∼aartb/)

M. De Boeck
Department of Mathematics
Ghent University
Krijgslaan 281-S22
9000 Gent
Belgium
(mdeboeck@cage.ugent.be)

F. Mazzocca
Dipartimento di Matematica
Seconda Università degli Studi di Napoli
Via Vivaldi 43
81100 Caserta
Italy
(francesco.mazzocca@unina2.it, http://francesco.mazzocca.name/)

L. Storme
Department of Mathematics
Ghent University
Krijgslaan 281-S22
9000 Gent
Belgium
(ls@cage.ugent.be, http://cage.ugent.be/∼ls)