A Density Version of a Geometric Ramsey Theorem

T. C. Brown and J. P. Buhler

Abstract

Let V be an n-dimensional affine space over the field with p^d elements, $p \neq 2$. Then for every $\varepsilon > 0$ there is an $n(\varepsilon)$ such that if $n = \dim(V) \geq n(\varepsilon)$ then any subset of V with more than $\varepsilon |V|$ elements must contain 3 collinear points (i.e., 3 points lying in a one-dimensional affine subspace).

1   Introduction

A celebrated result due to Van der Waerden says that if the integes in the interal $[1, n]$ are colored with r colors then there is a monochromatic arithmetic progression of length k if n is large enough as a function of k and r. A stronger result, proved for $k = 3$ by Roth [9] and arbitrary k by Szmerédi [12, 13], asserts that for all $\varepsilon > 0$ there is an $n = n(\varepsilon, k)$ such that any set of integers $A \subset [1, n]$ with more than εn elements must contain an arithmetic progression of length k. Alternatively, if $f_k(n)$ denotes the size of the maximal subset of $\{1, \ldots, n\}$ that does not contain a k-term arithmetic progression then $f_k(n) = o(n)$. This statement implies Van der Waerden's Theorem and could be regarded as a "density version" of that result.

Let F be a finite field of odd characteristic and let V be an n-dimensional vector space over F. Three distinct points x, y, z in V are said to be collinear if $z - y$ is a scalar multiple of $y - x$; i.e., $x = u + rv$, $y = u + sv$, $z = u + tv$ for distinct r, s, and t in F. The goal of this paper is to prove a density version of the following known fact (see [4] or [9]): if the points of V are colored with r colors and n is large enough as a function of r then there are 3 collinear points of the same color (in fact there is a whole line, plane, ...; see Section 4). Thus we intend to show that if $\varepsilon > 0$ is given and A is a subset of V with $|A| > \varepsilon |V|$ then A contains 3 collinear points if $n = \dim_F V$ is large enough as a function of ε.

Note that the relation of collinearity depends only on the structure of V as an affine space over F so that we might as well take V to be an n-dimensional affine space over F (i.e., a set that is isomorphic to F^n by an isomorphism that is determined only up to an affine transformation of F^n).

The method of the proof is based on ideas in Graham's exposition in [4] of Szmerédi's proof of Roth's theorem on 3-term arithmetic progressions. The main theorem is proved in the third section of the paper. The second section contains some needed technical results, and the fourth section contains some further remarks and corollaries. Perhaps, as with some other mathematical papers, the easiest approach for the reader is obtained by reading the sections in reverse order.

Our primary interest is in the case in which F is the field with 3 elements. In this special case 3 collinear points form an affine line. Since a combinatorial line is an affine line (but not vice versa) our result could be viewed as a weakening of the density version of the Hales-Jewett Theorem (see Section 4; other geometric Ramsey-type theorems should also have density analogues). It might be worthwhile to point out to the reader that Graham offers a monetary reward [4] for a proof of the density version of the Hales-Jewett result for three-element sets.

An ovaloid in the projective space $\mathbb{P}^r(F)$ is a subset that is maximal with respect to the property of not containing 3 collinear points [8, 10]; they give certain codes that can correct one error and detect three errors. In Section 4 we show how our main result implies that the density of ovaloids must be asymptotically small. Thus if A is an ovaloid then $|A|/|\mathbb{P}^r(F)|$ goes to 0 as r goes to infinity.

2 Lemmas

We need two substantial technical side results for the proof of the main theorem. In order not to disrupt the flow of the main argument these results are presented in this section.

The first lemma is an extension of an analogue of an initial lemma in the proof of Roth's Theorem in Ref. [4]; we incorporate the proof given there.

Let F be a finite field with q elements (q not necessarily odd). If V is an n-dimensional vector space over F and $a, x_1, \ldots, x_k$ are elements of V then define a set

$$Q(a, x_1, \ldots, x_k) = \{a + \alpha_1 x_1 + \cdots + \alpha_k x_k : \alpha_i = 0, 1\}.$$

Thus $Q(a, x_1, \ldots, x_k)$ is a translate of the 0-1 linear combinations of the x_i.

Lemma 1. *For all $\varepsilon > 0$ there is a C depending only on q and ε such that if A is a subset of V that does not contain 3 collinear points and $|A| > \varepsilon|V|$, then there are $a, x_1, \ldots, x_k$ in V such that*

$$Q(a, x_1, \ldots, x_k) \subset A,$$
$$k > \frac{\log\log|V|}{\log(q)} + C = O(\log^2 |V|),$$
$$x_1, \ldots, x_k \text{ are linearly independent.}$$

Remark on notation. Throughout this paper $|A|$ denotes the size of a set A, $\log^m(x)$ denotes the m-fold iterated logarithm $\log(\log(\cdots$, and $o(x)$ and $O(x)$ refer to the behavior of functions as $x \to \infty$.

Proof. First we show by induction on d that if A is a subset of V with $|A| \geq t_d$, where $t_d = 4|V|^{1-1/2^d}$, then there are $a, x_1, \ldots, x_d$ in V with $Q(a, x_1, \ldots, x_d)$ contained in A. For $d = 1$ the result is trivial. Take $d \geq 2$ and assume the result for $d - 1$. Assume that $A = \{a_1, \ldots, a_t\}$ is a subset of V with $|A| = t \geq t_d$. Note that

$$\binom{t_d}{2} \geq t_d^2/4 = 4|V|^{1-2^{-(d-1)}} \cdot |V| = t_{d-1}|V|.$$

Hence more than t_{d-1} of the diffrences $a_i - a_j$, $i > j$, are equal; say, $\{w\} = \{a_{i_k} - a_{j_k} : 1 \leq k \leq t_{d-1}\}$. Now let $A' = \{a_{j_k}\}$ and apply the induction hypothesis to A' to get $a, x_1, \ldots, x_{d-1}$ in V with

$Q(a, x_1, \ldots, x_{d-1}) \subset A'$. Because of the definition of w and A' it follows that if x' is in A' then $x' + w$ is in A. Thus $Q(a, x_1, \ldots, x_{d-1}, w) \subset A$, as required.

Now observe that if

$$d = \frac{\log\log|V| - \log\log(4/\varepsilon)}{\log(2)}$$

then $t_d = \varepsilon|V|$. The above argument then shows that we can find

$$Q(a, x_1, \ldots, x_d) \subset A$$

with $d > \log^2|V|/\log(2) + C$, where C is the generic symbol for a constant that depends on ε and q but not on $n = \dim_F V$.

Now we make use of the assumption that A contains no 3 collinear points.

Claim. The linear combinations $a + \sum \alpha_i x_i$ in the $Q(a, x_1, \ldots)$ constructed above are distinct.

Proof. If two such combinations are equal, say,

$$a + \sum \alpha_i x_i = a + \sum \beta_i x_i, \quad \alpha_i, \beta_i \text{ in } \{0, 1\},$$

then we might as well assume that $\alpha_i \neq \beta_i$ for all i. Find j such that $\alpha_j = 1$. Then

$$x_j + \sum \alpha_i x_i = \sum \beta_i x_i,$$

where the sums now run over all $i \neq j$. Then the three vectors

$$u = a + \sum \beta_i x_i + x_j,$$
$$v = a + \sum \beta_i x_i,$$
$$w = a + \sum \alpha_i + x_i$$

are collinear elements of A; indeed by the definition of x_j

$$w - v = \sum \alpha_i x_i - \sum \beta_i x_i = -x_j = v - u.$$

This finishes the proof of the claim. $\square$

Now we use the Claim to extract a linearly independent subset from $x_1, \ldots, x_d$ of sufficient size.

We now know that $|Q(a, x_1, \ldots, x_d)| = 2^d$. The vectors $x_1, \ldots, x_d$ must therefore span a linear subspace of dimension at least k, where

$$q^k \leq 2^d < q^{k+1},$$

i.e., $k \simeq (\log(2)/(\log(q))d$. Deleting the dependent x_i and renumbering gives us the desired set $x_1, \ldots, x_k$ with

$$k > \frac{\log^2|V|}{\log(q)} + C.$$

This finishes the proof of Lemma 1. $\square$

3

For the next lemma we take V and F as above except that now F must be a prime field with p elements where p is an odd prime. To fix the ideas let W be an m-dimensional linear subspace of the n-dimensional F-vector space V and let W' be a translate of W (one could also let W' be an arbitrary affine space for an F-vector space W). Choose a basis $x_1, \ldots, x_m$ for W and let

$$P = \{\alpha_1 x_1 + \cdots + \alpha_m x_m : \alpha_i = 0, 1\}$$

be the 0-1 linear combinations of the basis elements.

Throughout the following lemma let $f_k(x)$ denote the real-valued function on $[0, 1]$ given by

$$f_k(x) = x + \frac{2}{k-1} x(x-1) = \frac{k+1}{k-1} x - \frac{2}{k-1} x^2.$$

Lemma 2. *If A is a subset of W' with density a then the density of $A + P$ is at least $f_p(a)$; i.e.,*

$$|A| = a|W'| \quad \text{implies} \quad |A + P| \geq f_p(a)|W'|.$$

Remarks The statement is trivially true for the extreme cases $a = 0$ or $a = 1$. The statement is false if W' is an affine space over F_q, $q = p^d$, $d > 1$. Indeed, if A is the $\mathbb{F}_p$ span of P in $W' = W$ then $A + P = A$ which contradicts the bound given.

Proof. The proof is by induction on m. For $m = 1$ and $a \neq 0, 1$, we have

$$|A + P| \geq |A| + 1 = a|W| + 1 = ap + 1 \geq \left(a + \frac{2}{p-1} a(1-a)\right) p$$

since $a(1-a) \leq \frac{1}{4}$, so that

$$\frac{2p}{p-1} a(1-a) \leq \frac{p}{2(p-1)} \leq 1.$$

Now assume that the lemma is true for all spaces of dimensions smaller than m. Write

$$W' = \bigcup W_r, \quad r \text{ in } F = \mathbb{F}_p,$$

where each W_r is an affine hyperplane and

$$W_r = W_0 + r \cdot x_m.$$

(If we think of the x_i as the standard basis vectors in $W' = F^m$ then we could take W_r to be the set of vectors whose last coordinate was equal to r.)

To simplify the notation let P' be the 0-1 linear combinations of $x_1, \ldots, x_{m-1}$ and let $x = x_m$. Put $A_r = A \cap W_r$. Then elements of $A + P$ are of two kinds; either they are in

$$A_r + P' \subset W_r$$

for some r in F, or they are in

$$A_r + P' + x \subset W_{r+1}$$

4

for some r in F. By counting up the elements of $A + P$ that are of the first kind and then including any elements of the second kind left over we can bound $|A + P|$ from below:

$$|A + P| \geq \sum |A_r + P'| + \sum \sup(0, |A_r + P'| - |A_{r+1} + P'|), \tag{1}$$

where both summations are over all r in $F = \mathbb{F}_p$. To bound the second term we need an easy estimate:

Sublemma. *If $k \geq 2$ and $y_1, \ldots, y_k$ are real numbers then*

$$\sum_{i=1}^{k} \sup(0, y_{i+1} - y_i) \geq y_b - y_c$$

for any b and c (where we let y_{k+1} be another name for y_1).

Proof. Use induction on k together with the fact that

$$\sup(0, v - u) + \sup(0, w - v) \geq \sup(0, w - u)$$

(which is easy to check by enumerating the possible cases). End of sublemma. $\square$

Now let $a_r = |A_r|/|W_r|$ be the density of A_r in W_r. Note that the density a of A in W' is the average of the a_r. Apply the sublemma to the second term in (1) to get

$$\begin{aligned}
|A + P| &\geq \sum_{r \in F} |A_r + P'| + |A_b + P'| - |A_c - P'| \\
&= 2|A_b + P'| + \sum_{r \neq b, c} |A_r + P'|.
\end{aligned}$$

Now choose b so that a_b is the largest of the various a_i and choose c so that a_c is the smallest of the a_i. Divide by $|W_r| = |W|/p$ and use the inductive assumption to get

$$p \frac{|A + P|}{|W|} \geq 2f_p(a_b) + \sum_{r \neq b, c} f_p(a_r).$$

Since Lemma 2 just says that $|A + P|/|W|$ is at least as big as $f_p(a)$ it is clear that the proof of the lemma will be finished if we can demonstrate the following property of the function $f_k(x)$:

Claim. If $k \geq 3$ and we are given real numbers in $[0, 1]$ satisfying

$$0 \leq y_1 \leq \cdots \leq y_k \leq 1$$

then

$$f_k(y_2) + f_k(y_3) + \cdots + f_k(y_{k-1}) + 2f_k(y_k) \geq k f_k \left(\frac{y_1 + \cdots + y_k}{k} \right).$$

Proof of the claim. We work backwards. Write the desired inequality as

$$f_k(y_k) - f_k(y_1) \geq k \cdot f_k(\bar{y}) - f_k(y_1) - \cdots - f_k(y_k),$$

where $\bar{y}$ denotes the average of the y_i. If we plug in the definitoin of

$$f_k(x) = \frac{k+1}{k-1}x - \frac{2}{k-1}x^2$$

then all of the linear terms in the y_i on the right hand side cancel. Multiply by $k(k-1)/2$. The result is

$$\frac{k(k+1)}{2}(y_k - y_1) - k(y_k^2 - y_1^2) \geq -(k\bar{y})^2 + ky_1^2 + \cdots + ky_k^2$$
$$= \sum_{i>j}(y_i - y_j)^2.$$

The left hand side can be rewritten, after doing a little juggling, as

$$\frac{k(k-1)}{2}(y_k - y_1)^2 + k(y_k - y_1)\left(\frac{k+1}{2} - \frac{k+1}{2}y_k + \frac{k-3}{2}y_1\right).$$

The desired inequality is now evident since the second term in the preceding expression is positive and $(y_k - y_1)^2$ is greater than each of the $k(k-1)/2$ terms of the form $(y_i - y_j)^2$. End of Lemma 2 $\qquad\square$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3   Theorem

In this section we prove the following result.

Theorem. *Let p be an odd prime and F the field with $q = p^d$ elements. For every $\varepsilon > 0$ there is an $N(\varepsilon)$ such that if V is an affine space over F with $|V| \geq N(\varepsilon)$, and A is a subset of V with $|A| \geq \varepsilon|V|$, then A contains 3 collinear points.*

Proof. Put $N = |V| = q^n$ and let $f(N)$ denote the largest possible size of a subset of V that does not contain 3 collinear points. We shall call such a subset "line-free." The goal is to prove that $f(N) = o(N)$ as N becomes large.

First note that if V is an F vector space then it can be regarded as an $\mathbb{F}_p$ vector space by restriction of scalars. If $\{u, v, w\}$ is a set of collinear points in V regarded as an $\mathbb{F}_p$ vector space then it is certainly a collinear set if V is regarded as an F vector space. Thus any line-free set over F is also a line-free set if V is thought of as an $\mathbb{F}_p$ vector space.

Thus the theorem for arbitrary q follows from the theorem for prime fields so we begin by fixing $q = p =$ an odd prime and considering affine spaces V over $F = \mathbb{F}_p$. At any stage we are free to choose an identification $V \cong F^n$ since the notion of collinearity is independent of the affine space isomorphism that is chosen.

Let

$$V = \bigcup V_r, \quad r \in F = \mathbb{F}_p$$

be a disjoint union of $n-1$ dimensional affine subspaces. (For instance, we could take V_r to be the set of elements of $V \cong F^n$ whose nth coordinate is equal to r.) The intersection of a line-free set in V with

each of the V_r is any $n-1$ dimensional line-free set. Therefore $f(p^n) \le p \cdot f(p^{n-1}))$ so that the sequence $f(p^n)/p^n$ is decreasing to a limit

$$\lim_{N \to \infty} \frac{f(N)}{N} = c.$$

The Theorem asserts that $c = 0$.

Assume that $c > 0$. Choose $\varepsilon > 0$ with $\varepsilon < c^2/8p$ and let $N_0(\varepsilon)$ have the property that if A is a maximal line-free set in V and $|V| \ge N_0(\varepsilon)$ then

$$(c - \varepsilon)|V| < |A| < (c + \varepsilon)|V|, \tag{2}$$

Now choose N to be *very very very* large (i.e., the order of $\exp(\exp(\exp(N_0(\varepsilon))))$) and let A be a line-free set in V that is of maximum size where $|V| = N$. The exact conditions that we need on N are that the estimate (2) should apply to vector spaces with $\log^3 N$ elements and that a certain explicit $o(N)$ term that will arise later should be less than $c^2N/4$.

Write $V = \bigcup V_r$ as above and let $A_r = A \cap V_r$. Even though A_r is not necessarily a maximal line-free set inside V_r we can still get a lower bound on its size

$$|A_r| = |A| - \sum_{s \ne r} |A_s| > (c - \varepsilon)N - (p-1)(c + \varepsilon)\frac{N}{p}$$

$$= (c - (2p-1)\varepsilon)\frac{N}{p}. \tag{3}$$

Note that if u is in V_0 and v is in V_1 then there is a unique w in V_r such that $u, v,$ and w are collinear (in fact $w = (1-r)u + rv$). Thus elements in A_0 and A_1 have the effect of excluding certain elements of V_r from being in A_r.

In order to exploit this idea apply Lemma 1 to the subset A_1 of V_1 (which can be done since the density of A_1 in the affine space V_1 is at least $c - (2p-1)\varepsilon > 0$). We get a in A_1 and $x_1, \ldots, x_k$ such that the x_i are linearly independent, $k = O(\log^2 N)$, and

$$Q(a, x_1, \ldots, x_k) = \{a + \alpha_1 x_1 + \cdots + \alpha_k x_k : \alpha_i = 0, 1\} \subset A_1.$$

Define $Q_0 = \{a\}$ and for $1 \le i \le k$ define

$$Q_i = \{a + \alpha_1 x_1 + \cdots + \alpha_i x_i\} \subset Q_k(a, x_1, \ldots, x_k).$$

For $r \ne 0, 1$ put

$$M_{i,r} = \{w \in V_r : \text{there is a collinear set } u, v, w$$

$$\text{with } u \text{ in } A_0 \text{ and } v \text{ in } Q_1\}.$$

Two immediate observations are that

$$M_{i,r} \cap A_r = \emptyset$$

7

and that
$$|M_{i,r}| \geq |M_{0,r}| = |A_0| > (c - (2p-1)\varepsilon)\frac{N}{p}. \qquad (4)$$

Also we will need the following easy observation.

Fact. If u_0, u_1, and u_r are collinear, where u_i is in V_i, then $u_0, u_1 + v$, and $u_r + rv$ are collinear.

This implies that
$$M_{i+1,r} = M_{i,r} \cup (M_{i,r} + rx_{i+1}).$$

At this point we fix for the duration of the proof an element r of F, $r \neq 0, 1$. To simplify the notation put $M_{i,r} = M_i$.

Since we must certainly have $|M_k| \leq N$ the sequence of numbers $|M_{i+1} \setminus M_i|$ sums to at most N and therefore the "average" value is less than $N/k \simeq O(N/\log^2 N)$. Thus we can find a consecutive run $i+1, i+2, \ldots, i+m$ with
$$|M_{j+1} \setminus M_j| < \frac{N}{\log^3 N}, \quad i+1 \leq j \leq i+m,$$

for m as large as $k/\log^3 N \simeq O(\log^2 N/\log^3 N)$.

In fact we merely take such a run with m such that
$$N_0(\varepsilon) \leq p^m < \log^3 N$$

so that $m = O(\log^4 N)$. Fixing i and m as above let
$$W = \langle x_j : i+1 \leq j \leq i+m \rangle$$

be the linear space spanned by the indicated x_j. Let
$$P = \left\{ \sum_{j=i+1}^{i+m} \alpha_j x_j : \alpha_j = 0, 1 \right\}$$

be the 0-1 linear combinations of the given basis of W. To simplify the notation even further put $M = M_i$.

By the Fact above together with the definitions
$$M + rP \subseteq M_{i+m}.$$

Consequently
$$|M + rP \setminus M| < m\frac{N}{\log^3 N} = O\left(\frac{N\log^4 N}{\log^3 N}\right) = o(N) \qquad (5)$$

since $M_{i+m} \supset M + rP$ is obtained from M by a sequence of m steps of adding at most $N/\log^3 N$ vectors.

Since the coset (=translate) $a + W$ of the subspace W is contained in V_1 there is a coset of W contained in V_r and hence we can find a decomposition
$$V_r = \bigcup W_s$$

in which each of the W_s is a coset of the vector space W. Each of the W_s has $|W| = p^m$ elements and there

are $(N/p) \cdot p^{-m} \simeq O(N/\log^3 N)$ cosets.

We want to bound the size of A_r by an expression that will contradict estimate (3) above. For each coset W_s let $M_s = M \cap W - s$ and let a_s denote the density $|M_s|/|W_s|$. If a_s is large then only a few elements of W_s can belong to A_r (namely, the complement of M_s). If a_s is small then we can use the estimate in (2) applied to W_s since we chose $|W_s| = |W| = p^m \geq N_0(\varepsilon)$. The M_s with intermediate density are bad; all that can be done is to bound the nubmer of such cosets.

With these remarks in mind we define three different types of cosets:

$$
\begin{aligned}
W_s \text{ is } \textit{dense} \qquad && \text{if } a_s > 1 - 1/m, \\
W_s \text{ is } \textit{bad} \qquad && \text{if } 1/m < a_s < 1 - 1/m, \\
W_s \text{ is } \textit{sparse} \qquad && \text{if } a_s < 1/m,
\end{aligned}
$$

Let S, B and D denote the number of elements in, respectively, sparse, bad, and dense cosets so that $S + B + D = |V_r| = N/p$. We use the definitions in two ways. First note that by the remarks above

$$|A_r| < D/M + B + (c + \varepsilon)S = B + (c + \varepsilon)S + o(N). \tag{6}$$

Second, note that the definitions enable us to bound M from above; using (4) gives

$$
\begin{aligned}
(c - (2p - 1)\varepsilon)\frac{N}{p} \leq |M| < D + B + \frac{S}{m} &= D + B + o(N) \\
&= \frac{N}{p} - S + o(N)
\end{aligned}
$$

or

$$S < (1 - c + (2p - 1)\varepsilon)N/p + o(N).$$

Putting this into (6) gives

$$|A_r| < B + (c + \varepsilon)(1 - c + (2p - 1)\varepsilon)N/p + o(N). \tag{7}$$

Now we show that $B = o(N)$ and then show how the upper bound in (7) contradicts the lower bound in (3),

Let W_s be a bad coset and as above let a_s be the density of M_s in W_s. Apply Lemma 2 to the subset M_s of W_s (with rP playing the role of P in the Lemma) to get

$$|M_s + rP| = |M_s| + |M_s + rP \setminus M_s| \geq \left(a_s + \frac{2}{p - 1}a_s(1 - a_s) \right)|W_s|.$$

Since W_s is bad we must have $1/m < a_s < 1 - 1/m$ so that the "new" elements that are in $M_s + rP$ but not in M_s can be bounded by

$$|M_s + rP/M_s| \geq \frac{2}{p - 1}a_s(1 - a_s)|W| \geq \frac{2}{p - 1}\frac{1}{m^2}|W|.$$

9

Using (5) above and summing over each of the $B/|W|$ bad cosets gives

$$\frac{mN}{\log^3 N} > |M + rP \setminus M| \geq \sum_{s \text{ bad}} |M_s + rP \setminus M_s| \geq \frac{B}{|W|} \cdot \frac{2}{p-1} \cdot \frac{1}{m^2} |W|$$

so

$$B < \frac{m^3 N(p-1)}{2\log^3 N} = O\left(\frac{[\log^4 N]^3 N}{\log^3 N}\right) = o(N). \tag{8}$$

Combining (8) with (3) and (7) gives

$$(c - (2p-1)\varepsilon)\frac{N}{p} < |A_r| < (c + \varepsilon)(1 - c + (2p-1)\varepsilon)\frac{N}{p} + o(N),$$

where the $o(N)$ term could be written out explicitly if desired. This can be rearranged to say

$$c^2 < \varepsilon(2pc - 2c + 2p) + (2p-1)\varepsilon^2 + \frac{o(N)}{N}$$
$$< 4p\varepsilon + 2p\varepsilon^2 + o(N)/N < 6p\varepsilon + o(N)/N.$$

It is easy to see that this contradicts the assumptions $\varepsilon < c^2/8p$ and $o(N)/N < c^2/4$ made at the beginning of the proof. Thus the assumption that c is positive is untenable and the proof of the theorem is finished.

$\square$

Remark. The arbitrary element $r \neq 0, 1$ of F can be chosen explicitly (just after the "Fact" above) to give a slightly stronger result. If $r = 2$ then the proof actually shows that sufficiently large subsets of sufficiently large affine spaces contain three "equally spaced" collinear points of the form $x, x+y, x+2y$.

4   Remarks

Possible Generalizations

The well-known Hales-Jewett Theorem [7] says that for all finite sets F and $r \geq 1$ there is an $n = n(r, F)$ such that if the points of F^n are colored using r colors then there is a monochromatic combinatorial line. (A set $L \subset F^n$ is a *combinatorial line* if $|L| = |F|$ and there is a decomposition $\{1, 2, \ldots, n\} = S \cup T$ with the following property. For i in S the projection from L onto the ith coordinate is a constant mapping (with the constant possibly depending on i). For i in T the projection from L onto the ith coordinate is a fixed isomorphism $p : L \mapsto F$. One could visualize a combinatorial line as an $|F| \times n$ array of elements of F such that the columns were either constant (with possibly different constants for different columns) or else run through all elements of F in some fixed order.)

One would like to be able to prove the density version of the Hales-Jewett result to the effect that $f(N) = o(N)$, where $f(|F^n|)$ is the size of the largest subset of F^n not containing a combinatorial line.

If we take $|F| = 3$ in the Theorem above then 3 collinear points form an affine line. An affine line in F^n could be viewed as a $3 \times n$ array with columns that are either constant or else run through all of the elements of F in any order. Perhaps one key to extending the proof given above to combinatorial lines would be a drastically improved version of Lemma 1

A very general Ramsey-type theorem in a geometric context follows from the Hales-Jewett Theorem. This was conjectured by Rota and proved by Graham, Rothschild, and Leeb (see the elegant proof in [11]; also [3,5,6,14]). It says that for any finite field F there is an $n = n(t,k,r,F)$ such that if the t-dimensional affine subspaces of F^n are colored with r colors then some k-dimensional subpsace has all of its t-dimensional subspaces having the same color. The word "affine" can be replaced throughout by "linear" or "projective."

It is natural to ask for "density versions" of this theorem or its special cases.

As a matter of further detail, in the $t = 0$ case the Hales-Jewett Theorem actually gives not just a monochromatic k-dimensional affine subspace, but in fact a monochromatic "combinatorial k-space" whose definition is a natural extension of the definition of a combinatorial line. Density versions of this result would be very strong; they would imply the $t = 0$ case of the density versions of the geometric theorems.

The case in which F is the set, or field, with 2 elements provides some weak evidence in favor of these conjectures. The density version of the Hales-Jewett Theorem is true by using Sperner's Lemma [4] and the density version of the $t = 0$ case of the geometric theorem follows from Lemma 1 immediately.

Corollary to Lemma 1. *For every integer k and every $\varepsilon > 0$ there is an $n(\varepsilon,k)$ such that if V is an n-dimensional affine space over $\mathbb{F}_2$, $n \geq n(\varepsilon,k)$, and A is a subset of V with $|A| > \varepsilon|V|$, then A contains a k-dimensional affine subspace.*

Note that the main result of this paper is still open for field of order 2^n, $n > 1$. It seems that it is necessary to find something to replace the role of Lemma 2.

Finally note that the theorem in Section 3 is a density version of a geometric Ramsey Theorem only in the case $|F| = 3$ (since this is the only case in which three collinear points form a complete affine line). In a similar vein one could look at sets that do not contain 4 collinear (or coplanar?) points.

An Application to Coding Theory

Let $\mathbb{P}^n(F)$ denote the usual n-dimensional projective space over a finite field F. A subset A of $\mathbb{P}^n(F)$ is said to be an *ovaloid* if it is maximal with respect to the property of not containing 3 collinear points.

Ovaloids can be used to produce codes that can correct one error and detect three; in the cases $n = 2$ and $n = 3$ they have a nice geometrical structure. Little is known about ovaloids in $\mathbb{P}^n(F)$ for $n > 3$ (see [8] and [10] and the references therein).

In the case in which F has odd characteristic our main result implies the density of ovaloids must be asymptotically zero. We are grateful to A. Odlyzko and R. Graham for bringing this to our attention.

Proposition. *Let F be the field with q elements, q odd. Then for all $\varepsilon > 0$ there is an $n(\varepsilon)$ such that if n is larger than $n(\varepsilon)$ and A is an ovaloid in $\mathbb{P}^n(F)$ then $|A| < \varepsilon|\mathbb{P}^n(F)|$.*

Proof. Let n_0 be the integer such that if $n \geq n_0$ then any subset of F^n of density bigger than $\varepsilon/2$ must contain 3 collinear points; the existence of n_0 is guaranteed by the theorem in Section 3. Define

$$n(\varepsilon) = n_0 + \log_q(2/\varepsilon),$$

where $\log_q$ denotes logarithm to the base q.

It is a standard fact that projective space can be decomposed into a union of affine spaces as

$$\mathbb{P}^n(F) \cong F^n \cup F^{n-1} \cup \cdots \cup F^1 \cup F^0.$$

If A is an ovaloid in $\mathbb{P}^n(F)$ and $n \geq n(\varepsilon)$ then we can crudely estimate the size of A by breaking it into two parts:

$$|A| = |A \cap (F^n \cup \cdots \cup F^{n_0})| + |A \cap (F^{n_0-1} \cup \cdots \cup F^0)|.$$

Use the definition of n_0 for the first term and include the whole subset in the second term; the result is

$$|A| \leq \frac{\varepsilon}{2}(q^n + \cdots + q^{n_0}) + (q^{n_0-1} + \cdots + 1).$$

After summing the geometric series and using the definition of $n(\varepsilon)$ we get $|A|/|\mathbb{P}^n(F)| < \varepsilon$ as desired. This finishes the proof of the Proposition. $\qquad\square$

It is clear that the estimates on the size of line-free sets implied by the above proposition (which depend on the estimates implicit in the proof of the Theorem in Section 3) are very poor. It would be interesting to describe the size of ovaloids (or line-free sets in F^n) more accurately. In order describe the known values in the case $F = \mathbb{F}_3$ let $f(n) =$ the size of the maximal line-free set in F^n and let $g(n) =$ the size of the maximal line-free set in $\mathbb{P}^n(F)$ (i.e., the size of an ovaloid). By various calculations (see [8] for the case of ovaloids) the following values are known:

$$f(1) = 2, \qquad\qquad g(1) = 2,$$
$$f(2) = 4, \qquad\qquad g(2) = 4,$$
$$f(3) = 9, \qquad\qquad g(3) = 10,$$
$$f(4) = 20, \qquad\qquad g(4) = 20,$$
$$f(5) \geq 45, \qquad\qquad g(5) = 56$$

(the inequality is almost certainly an equality).

References

[1] B. Alspach, T.C. Brown, and P. Hell, *On the density of sets containing no k-element arithmetic progressions of a certain kind*, J. London Math. Soc. (2) **13** (1976), 226–234.

[2] T.C. Brown, *Behrend's theorem for sequences containing no k-element arithmetic progression of a certain type*, J. Combin. Theory Ser. A **18** (1975), 352–356.

[3] R.L. Graham, K. Leeb, and B.L. Rothschild, *Ramsey's theorem for a class of categories*, Adv. in Math. **8** (1972), 417–433, Errata, **10** (1973), 326–327.

[4] Ron L. Graham, *Rudiments of ramsey theory*, Amer. Math. Soc., Providence, RI, 1981.

[5] Ronald L. Graham and Bruce L. Rothschild, *Ramsey's theorem for n-parameter sets*, Trans. Amer. Math. Soc. **159** (1971), 257–292.

[6] _____ , *Rota's geometric analog to Ramsey's theorem*, Proc. Symp. Pure Math. **19** (1971), 101–104.

[7] Alfred W. Hales and Robert I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222–239.

[8] R. Hill, *Caps and codes*, Discrete Math. **22** (1978), 111–137.

[9] Klaus F. Roth, *On certain sets of integers II*, J. London Math. Soc. **29** (1954), 20–26.

[10] B. Segre, *On complete caps and ovaloids in three-dimensional Galois spaces of characteristic 2*, Acta Arith. **5** (1959), 315–332.

[11] Joel H. Spencer, *Ramsey's theorem for spaces*, Trans. Amer. Math. Soc. **249** (1979), 363–371.

[12] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104.

[13] _____ , *On sets of integers containing no k elements in an arithmetic progression*, Acta. Arith. **27** (1975), 199–245, Collection of articles in memory of Jurii Vladimirovic Linnik.

[14] B. Voigt, *A Ramsey theorem for finite geometries*, to appear.