ACTA ARITHMETICA
122.1 (2006)

Behrend-type constructions for sets of linear equations

by

Asaf Shapira (Tel Aviv)

1. Introduction

1.1. *Background.* For an integer n we denote the set $\{1,\dots,n\}$ by $[n]$ and the set $\{-n,\dots,n\}$ by $[-n,n]$. Call an equation E of the form $\sum_{i=1}^{k} a_i x_i = 0$ with integers $a_i \in [-h,h]$ and $\sum_{i=1}^{k} a_i = 0$ a (k,h)-*equation*. In this paper we will be interested in integer sets without non-trivial solutions to sets of (k,h)-equations. We will mainly follow the terminology of [7]. Suppose $[k]$ can be partitioned into disjoint (non-empty) subsets $A_1,\dots,A_g$ such that $\sum_{i\in A_j} a_i = 0$ for every $1 \le j \le g$. Clearly, any sequence $x_1,\dots,x_k$ in which $x_{i_1} = x_{i_2}$ whenever i_1 and i_2 belong to the same set A_j is a solution of E. We call such a solution *trivial*. Obviously, any solution in which all the integers are distinct is non-trivial, and any solution in which all the integers are identical is trivial. For a (k,h)-equation E we let $r_E(n)$ denote the size of the largest subset of $[n]$ without non-trivial solutions of E. In this paper we will be interested in cases where $1 \ll k \ll h \ll n$.

Several special cases of estimating $r_E(n)$ are some of the best studied problems in additive number theory. For example, it is easy to derive from Szemerédi's celebrated theorem about integer sets without long arithmetic progressions [9] that for every fixed k and h and every (k,h)-equation, we have $r_E(n) = o(n)$. When E is the equation $x_1 + x_2 - x_3 - x_4 = 0$ we get Sidon's problem. This problem has been extensively studied and it is known that in this case $r_E(n) = (1+o(1))\sqrt{n}$ (see [7] for proofs and references). Here and throughout the paper, $o(1)$ represents a quantity that approaches 0 as n tends to infinity. When E is the equation $x_1 + x_2 - 2x_3 = 0$ we get the (even better studied) problem of the largest subset of $[n]$ without three-term arithmetic progressions (see [7]). The following lower bound is known (which applies to the three-term arithmetic progressions equation

2000 *Mathematics Subject Classification*: Primary 11P99.

This work forms part of the author's Ph.D. thesis under the supervision of Prof. Noga Alon. Research supported in part by a Charles Clore Foundation Fellowship.

as a special case). Its proof is an easy consequence of Behrend's construction [1].

THEOREM 1 ([1]). *For every h there is $c = c(h) > 0$ such that for every pair of integers $a_1, a_2 \in [-h, h]$, and for every large enough n, there is $X \subset [n]$ of size at least $n/e^{c\sqrt{\log n}} = n^{1-o(1)}$ with no non-trivial solution of $a_1x_1 + a_2x_2 - (a_1 + a_2)x_3 = 0$.*

In [7], Ruzsa gave several general results on $r_E(n)$ based on certain properties of the equation E. It seems that one of the most important properties of an equation is its *genus* which is defined in [7] as the largest integer g such that $[k]$ can be partitioned into disjoint subsets $A_1, \dots, A_g$ such that $\sum_{i\in A_j} a_i = 0$ for every j. The following result is proved in [7].

THEOREM 2 ([7]). *For every (k, h)-equation E of genus g we have $r_E(n) = O(n^{1/g})$.*

The results of [7] suggest the following possibility (though, quoting [7], "there is too little positive support to call it a conjecture").

PROBLEM 1. *Is $r_E(n) \geq n^{1/g-o(1)}$ for every (k, h)-equation E of genus g?*

As the sum of the coefficients of a (k, h)-equation is 0, it is clear (see Lemma 3.3) that there are $\Theta(h^{k-1})$ such equations (where the hidden constant depends only on k; recall that we are interested in the cases where $k \ll h$). Similarly, there are $\Theta(h^{k-2})$ such equations of genus at least 2. It thus follows that if indeed $r_E(n) = n^{1/g-o(1)}$ then all the (k, h)-equations E besides a $c(k)/h$ fraction of them are such that $r_E(n) \geq n^{1-o(1)}$, where $c(k)$ is a constant that depends only on k. To simplify the presentation, whenever we write $c(k)$ we mean a constant that depends only on k. Our main result in this paper is that such a phenomenon is true for large enough sets of linear equations in k unknowns.

1.2. *New results.* Let S be a set of linear equations in k unknowns. Let $a_{i,1}, \dots, a_{i,k}$ denote the coefficients of equation i in S. Suppose $[k]$ can be partitioned into g disjoint subsets $A_1, \dots, A_g$ such that $\sum_{p\in A_j} a_{i,p} = 0$ for all i and j. Clearly any sequence $x_1, \dots, x_k$ in which $x_{i_1} = x_{i_2}$ whenever i_1 and i_2 belong to the same set A_j is a solution of S. We call such a solution *trivial*. For a set S of (k, h)-equations we let $r_S(n)$ denote the size of the largest subset of $[n]$ without non-trivial solutions of S. It is rather easy to show that for some sets of equations all but an $O(1/h)$ fraction of them are such that $r_S(n) > n^{1-o(1)}$. For example, by Theorem 1 for *every* $(3, h)$-equation E, we have $r_E(n) > n^{1-o(1)}$. As another example, consider sets of $k - 2$ (k, h)-equations on the same set of k unknowns. It is easy to show (e.g. using Lemma 3.2 below) that all but an $O(1/h)$ fraction of them have certain linear independence properties that enable one to extract an

equation of type $a_1x_1 + a_2x_2 - (a_1 + a_2)x_3 = 0$, and by Theorem 1 there is a subset of $[n]$ of size $n^{1-o(1)}$ with no non-trivial solution to such an equation. Thus, all but $O(1/h)$ of pairs of equations S in four unknowns are such that $r_S(n) > n^{1-o(1)}$. The same applies for sets of three equations in five unknowns. Our main result is that for larger k one may consider much smaller sets of equations.

THEOREM 3. *For an integer $k \geq 6$ let t be the largest integer satisfying $\binom{t}{2} \leq k$ (hence, $t \geq \lfloor\sqrt{2k}\rfloor$). Then there is a constant $c = c(k,h) > 0$ such that all but $c(k)/h$ of the sets S of $k-t+1$ (k,h)-equations in k unknowns are such that*

$$r_S(n) \geq n/e^{c\sqrt{\log n}} = n^{1-o(1)}.$$

We stress that in proving Theorem 3 we make no assumption regarding the answer to Problem 1. We also make no real effort to optimize the constants $c(k)$ and $c(k,h)$ in Theorem 3. We mention that using the main idea of [5], the lower bound on $r_S(n)$ in Theorem 3 can be improved to $n/e^{c\log^{1/p} n}$ where $p \approx \log k$. As in this paper we are interested in whether $r_S(n) > n^{1-o(1)}$ we will not describe this slightly better lower bound. As we observe at the end of the proof, we actually prove a stronger claim; namely, that all but a small fraction of the sets are such that the only solution is one in which all the integers are identical (which is always a trivial solution). Observe that the well studied problem of sets of integers without k-term arithmetic progressions (see [2], [3], [5], [6] and [9]) is actually the study of the largest subset of $[n]$ with no k integers which satisfy a set of $k-2$ equations of type $x_i + x_{i+2} - 2x_{i+1} = 0$. Though we do not explicitly state it in the course of the proof, the details of the proof of Theorem 3 actually give a simple sufficient criterion to decide whether $r_S(n) \geq n^{1-o(1)}$ for a given set S as in the statement of Theorem 3. See the discussion after the proof of Theorem 3.

One may also try to find properties of sets of equations and their effect on $r_S(n)$, and thus obtain bounds for specific sets of equations. To this end we extend the notion of a genus to sets of equations as follows: define the *genus* of a set S of equations to be the largest integer g such that $[k]$ can be partitioned into g disjoint non-empty subsets $A_1, \ldots, A_g$ such that $\sum_{p\in A_j} a_{i,p} = 0$ for every j and every equation i in S. Note that for a set containing one equation, this is equivalent to the genus previously defined. Our first result is the following:

THEOREM 4. *For every set S of equations with genus g and rank t we have $r_S(n) = O(n^{t/g})$.*

We also prove the following theorem, which shows that Theorem 4 cannot be generally improved.

THEOREM 5. *There are sets S of equations with genus g and rank g which satisfy $r_S(n) = n$.*

An interesting consequence of the above theorem is that unlike the case of single equations, where for every equation E we have $r_E(n) = o(n)$, for sets of equations we may have $r_S(n) = n$. We also raise the following possibility as an open problem.

PROBLEM 2. *Is $r_S(n) \geq n^{1/g-o(1)}$ for every S of genus g?*

We finally prove the following theorem relating Problems 1 and 2.

THEOREM 6. *Problems* 1 *and* 2 *are equivalent.*

The proof of Theorem 3 is given in Sections 2 and 3. In Section 2 we study special sets of linear equations, which we call *diagonalized*. In Section 3 we give the proof of Theorem 3. Using the results on diagonalized sets of equations from Section 2, we show that most sets of equations have certain non-degeneracy properties that allow us to infer that certain points, defined by the set of equations (its *parametric representation* as defined in Section 2), do not all lie on a high-dimensional sphere. To this end, we use certain properties of multi-variate polynomials. In Section 3 we use a version of Behrend's argument [1] (already used in [8]), in which one represents integers as high-dimensional vectors in order to show that some high-dimensional sphere contains many integers with no non-trivial solution of a given set of equations. Our proof of Theorem 3 is also somewhat motivated by the interpretation of Laba and Lacey [5] of the construction of Behrend [1]. In Section 4 we prove Theorems 4–6 and also discuss some open problems and additional observations.

2. Diagonalized sets of equations. In this section we deal with a special kind of sets of linear equations in k unknowns. We start with the following definition:

DEFINITION 2.1. For positive integers $t < k$, let $\mathcal{F}(k, h, t)$ denote the collection of all sets of $k-t+1$ (k, h)-equations in k unknowns $x_0, \ldots, x_{k-1}$.

A set $S \in \mathcal{F}(k, h, t)$ of equations is called *diagonalized* if the only non-zero coefficients of the set are the following: (i) the coefficients of $x_0, \ldots, x_{t-2}$, (ii) the coefficient of x_{t-2+i} in equation i (the coefficients of the second type can be thought of as forming a diagonal). The advantage of diagonalized sets is that they make the proofs and notations very easy compared to general sets of equations. We start with the following simple claim that helps us represent integers that satisfy a diagonalized set of equations as the image of a certain linear function.

CLAIM 2.2. *Suppose that for an integer t we have a set of $k \geq t$ reals $z_0, \ldots, z_{k-1}$ that satisfy the following diagonalized set of $k-t+1$ linear equations (where the solution is obtained by setting $x_0 = z_0, \ldots, x_{k-1} = z_{k-1}$):*

$$a_{i,0}x_0 + a_{i,1}x_1 + \cdots + a_{i,t-2}x_{t-2} - (a_{i,0} + \cdots + a_{i,t-2})x_{t-2+i} = 0, \quad 1 \leq i \leq k-t+1. \tag{1}$$

For $1 \leq i \leq k-t+1$ put $d_i = a_{i,0} + \cdots + a_{i,t-2}$ and $d = \prod_i d_i$. Then for $0 \leq j \leq k-1$ we can write

$$z_j = z_0 + p_{j,1}\frac{z_1 - z_0}{d} + \cdots + p_{j,t-2}\frac{z_{t-2} - z_0}{d}, \tag{2}$$

where the $(t-2)$-tuples $(p_{j,1}, \ldots, p_{j,t-2})$ are the following:

(i) $(p_{0,1}, \ldots, p_{0,t-2}) = (0, \ldots, 0)$.

(ii) *For $1 \leq j \leq t-2$ we have $(p_{j,1}, \ldots, p_{j,t-2}) = de_j$ where e_j is the j^{th} unit vector of size $t-2$.*

(iii) *For $t-1 \leq j \leq k-1$ we have*

$$(p_{j,1}, \ldots, p_{j,t-2}) = \frac{d}{d_{j-t+2}}(a_{j-t+2,1}, \ldots, a_{j-t+2,t-2}).$$

Proof. Observing (2) one can immediately see that the first two assertions are trivial. For the third, note that by (1) for every $t-1 \leq j \leq k-1$ the integer z_j appears only in equation $j-(t-2)$ of (1). To simplify notation set $i = j-(t-2)$ and consider equation i in (1) after substituting $x_0 = z_0, \ldots, x_{k-1} = z_{k-1}$. Note that we can rewrite each such equation as

$$a_{i,0}z_0 + a_{i,1}(z_1 - z_0) + a_{i,1}z_0 + \cdots + a_{i,t-2}(z_{t-2} - z_0) + a_{i,t-2}z_0 - d_i(z_{t-2+i} - z_0) - d_i z_0 = 0. \tag{3}$$

Rearranging the above (recall that $d_i = a_{i,0} + \cdots + a_{i,t-2}$ and that we have set $i = j-t+2$) gives

$$z_j = z_{t-2+i} = z_0 + a_{i,1}\frac{z_1 - z_0}{d_i} + \cdots + a_{i,t-2}\frac{z_{t-2} - z_0}{d_i}.$$

Thus, we can indeed use $(p_{j,1}, \ldots, p_{j,t-2}) = \frac{d}{d_i}(a_{i,1}, \ldots, a_{i,t-2})$ in (2). ∎

For the rest of the proof we need some additional definitions. As will become clear later, we will mainly be interested in the $(t-2)$-tuples that define the integer solution $z_0, \ldots, z_{k-1}$ of a diagonalized set of equations in Claim 2.2. To this end, we define the *parametric representation* of a diagonalized set $S \in \mathcal{F}(k,h,t)$ as the set $\{p_0, \ldots, p_{k-1}\}$ of $(t-2)$-tuples which Claim 2.2 returns, where each p_i is short for $(p_{i,1}, \ldots, p_{i,t-2})$. We will also use the parametric representation of such a set when we regard the coefficients $a_{i,j}$ as unknowns. Note that in such a case, each coordinate $p_{i,j}$ of each of these $(t-2)$-tuples is a linear function in each of the unknowns $a_{i,j}$. In order to

carry out our proof we will also need some basic notions from analytic geometry. A *d-dimensional quadric* is the set of points $(x_1, \ldots, x_d) \in \mathbb{R}^d$ that satisfy an equation of type (for convenience we use $x_0 = 1$)

$$\sum_{0 \le i \le j \le d} c_{i,j} x_i x_j = 0.$$

A quadric is *non-zero* if some $c_{i,j} \neq 0$. Note that a d-dimensional quadric has $\binom{d}{2} + 2d + 1$ coefficients.

DEFINITION 2.3. Fix any $t \ge 4$. Given a set P of

$$k = \binom{t-2}{2} + 2(t-2) + 1 = \binom{t}{2}$$

points $P = \{p_0, \ldots, p_{k-1}\}$ with $p_i = (p_{i,1}, \ldots, p_{i,t-2}) \in \mathbb{R}^{t-2}$, let $A_{t-2}(P)$ be the $k \times k$ matrix of the matrix representation of the set of homogeneous linear equations which force a $(t-2)$-dimensional quadric with coefficients $c_{i,j}$ to pass through $p_0, \ldots, p_{k-1}$. For example, for $t = 4$ we can write this set of equations as

$$\begin{pmatrix} p_{0,1}^2 & p_{0,1}p_{0,2} & p_{0,2}^2 & p_{0,1} & p_{0,2} & 1 \\ p_{1,1}^2 & p_{1,1}p_{1,2} & p_{1,2}^2 & p_{1,1} & p_{1,2} & 1 \\ p_{2,1}^2 & p_{2,1}p_{2,2} & p_{2,2}^2 & p_{2,1} & p_{2,2} & 1 \\ p_{3,1}^2 & p_{3,1}p_{3,2} & p_{3,2}^2 & p_{3,1} & p_{3,2} & 1 \\ p_{4,1}^2 & p_{4,1}p_{4,2} & p_{4,2}^2 & p_{4,1} & p_{4,2} & 1 \\ p_{5,1}^2 & p_{5,1}p_{5,2} & p_{5,2}^2 & p_{5,1} & p_{5,2} & 1 \end{pmatrix} \begin{pmatrix} c_{1,1} \\ c_{1,2} \\ c_{2,2} \\ c_{0,1} \\ c_{0,2} \\ c_{0,0} \end{pmatrix} = 0.$$

DEFINITION 2.4. Fix any integer $t \ge 4$. For a diagonalized set $S \in \mathcal{F}\left(\binom{t}{2}, h, t\right)$ let $p_0, \ldots, p_{\binom{t}{2}-1}$ be its parametric representation. Denote by $A_{t-2}(S)$ the matrix obtained by plugging $p_{i,j}$ in the matrix A_{t-2} of Definition 2.3. The set S is called *degenerate* if $A_{t-2}(S)$ is not invertible.

CLAIM 2.5. *Let $A_{t-2}(S)$ be the matrix from Definition 2.4 when we regard the coefficients $a_{i,j}$ of a diagonalized set $S \in \mathcal{F}\left(\binom{t}{2}, h, t\right)$ as unknowns. Let D be the determinant of this matrix. Then D is a non-zero polynomial of degree at most t^2 in each variable.*

Proof. Call the matrix A for short. Recall that by Claim 2.2 each coordinate of the parametric representation is linear in each of the unknowns $a_{i,j}$. As each entry of A has degree at most 2, the degree of each $a_{i,j}$ in each entry is at most 2. The polynomial D represents a determinant, therefore we can view it as a sum of monomials. As A has fewer than $t^2/2$ rows, each monomial has fewer than $t^2/2$ terms. Therefore, the degree of each $a_{i,j}$ in each monomial, and hence also in D, is at most t^2.

We turn to show that D is not identically zero. It will be simpler to show this by requiring the unknowns $a_{i,0}$ to satisfy $a_{i,0} = 1 - \sum_{j=1}^{t-2} a_{i,j}$ for every $1 \le i \le \binom{t}{2} - t + 1$ (this is clearly a stronger claim). We thus get, in the parametric representation of S, the equalities $d_1 = \cdots = d_{\binom{t}{2}-t+1} = d = 1$. Hence, by Claim 2.2(iii), for $i \ge t-1$ we have $p_{i,j} = a_{i-t+2,j}$. Furthermore, by Claim 2.2(i), (ii), for $0 \le i \le t-2$ the ith row of the matrix contains only 0s and 1s. In fact, the only non-zero entry in the first row is the rightmost. Thus, when computing the determinant of A we may disregard the first row and the rightmost column of A. More importantly, it is easy to see (recall Definition 2.3 and the example of $d = 2$) that we can find distinct columns $i_1, \ldots, i_{t-2}$ whose entries satisfy $A_{1,i_1} = \cdots = A_{t-2,i_{t-2}} = 1$, and there are no other 1s in these columns within rows $1, \ldots, t-2$. Consider row $i \ge t-1$ of the matrix. By our choice of $a_{i,0}$, this row is the only one in which the unknowns $a_{i-t+2,1}, \ldots, a_{i-t+2,t-2}$ appear. Moreover, by definition of A_{t-2} in Definition 2.3, each entry of this row contains a different combination of $a_{i-t+2,1}, \ldots, a_{i-t+2,t-2}$. Thus, any term in the expansion of the determinant which uses $i_1, \ldots, i_{t-2}$ in rows $1, \ldots, t-2$ will have a coefficient precisely 1. Furthermore, the combination of the unknowns $a_{i-t+2,j}$ of this term cannot appear in another one, hence, it will not cancel. Therefore, D is not identically zero. ∎

As a non-zero polynomial cannot vanish everywhere, we immediately infer that for some choice of coefficients $a_{i,j}$, that is, *not* necessarily integers from $[-h, h]$, the polynomial D does not vanish. In other words:

COROLLARY 2.6. *For every integer $t \ge 4$, there is a non-degenerate diagonalized set of $\binom{t}{2} - t + 1$ equations in $\binom{t}{2}$ unknowns (possibly with non-integer coefficients).*

3. Proof of the main theorem. In this section we consider arbitrary sets of equations from $\mathcal{F}(k, h, t)$ as defined in the previous section. As in the previous section, we will mainly look at the coefficients of these equations as unknowns. In order to distinguish general sets of equations from diagonalized sets of equations, we will denote the coefficients of the former by $b_{p,q}$ and of the latter by $a_{p,q}$ as we did in the previous section. We start with the following simple claim, in which a *rational function of degree r* is any quotient of two polynomials of degree at most r in each unknown. In what follows, $r(k)$ will denote any quantity that depends only on k.

CLAIM 3.1. *Let $S \in \mathcal{F}(k, h, t)$ be a set of equations with coefficients $b_{p,q}$, and suppose we consider them as unknowns. Then there is an equivalent diagonalized set S_{diag}, with coefficients $a_{i,j}$, such that each $a_{i,j}$ is a rational*

function of degree $r(k)$ in each of the coefficients $b_{p,q}$. Also, none of the denominators of $a_{i,j}$ is identically zero.

Proof. We just use Gaussian elimination. Initially we have $a_{i,j} = b_{i,j}$ and therefore the degree of each $a_{i,j}$ in each $b_{p,q}$ is at most 1. As we perform at most k operations on each equation, and each operation at most doubles the maximum degree of each $a_{i,j}$ as a function of any other $b_{p,q}$, the resulting $a_{i,j}$ are rational functions in each $b_{p,q}$ of degree bounded by a function $r(k)$ that depends on k only. Finally, note that if one of the denominators of $a_{i,j}$ is identically zero, then for *any* assignments to the unknowns $b_{p,q}$, the resulting diagonalized set after the Gaussian elimination process has some entries which are not well defined. This, however, is clearly not the case. For example, when S is already diagonalized, S_{diag} is just S and all its entries are well defined. ∎

Note that the output of the above claim is a diagonalized set in the sense that all the coefficients $a_{i,j}$ that should be zero in order to satisfy the properties of a diagonalized set (see beginning of Section 2) are given by rational functions in $b_{p,q}$ that are identically zero. Clearly there are some assignments to the unknowns $b_{p,q}$ for which the diagonalized set S_{diag} is not well defined. These are the sets S for which the denominators of some of the rational functions that represent $a_{i,j}$ vanish. S_{diag} is called *well defined* if none of the denominators of $a_{i,j}$ vanish. Note that in such a case S and S_{diag} are equivalent in the sense that every solution of one is also a solution of the other. In what follows we will use the following lemma of Zippel (cf., e.g. [4]).

LEMMA 3.2. *Let F be an arbitrary field, and let $f = f(x_1, \ldots, x_b)$ be a non-zero polynomial in $F[x_1, \ldots, x_b]$. Suppose the degree of f in each variable is at most r. If H is a subset of F with $|H| > r$, then there are at most $|H|^b - (|H| - r)^b = c(r, b)|H|^{b-1}$ assignments $x_1 \in H, \ldots, x_b \in H$ so that $f(x_1, \ldots, x_b) = 0$.*

LEMMA 3.3. *For every $t \geq 4$ and $k \geq \binom{t}{2}$, all but a $c(k)/h$ fraction of $S \in \mathcal{F}(k, h, t)$ are such that:*

(i) *S_{diag} is well defined.*
(ii) *The points of the parametric representation of S_{diag} do not all lie on any non-zero $(t-2)$-dimensional quadric.*

Proof. Denote the number of equations in each set of $\mathcal{F}(k, h, t)$ by $e = k - t + 1$. Note that for any choice of $k-1$ integers from $[-h/k, h/k]$ we can find an integer in $[-h, h]$ such that the sum of the k integers is 0. We thus see that $\mathcal{F}(k, h, t)$ contains at least $(h/k)^{k-1}$ equations. Furthermore, we may conclude that $\mathcal{F}(k, h, t)$ contains at least $c(k)h^{e(k-1)}$ sets of equations. Thus, we may prove the claim by showing that there are at most $c(k)h^{e(k-1)-1}$

sets of equations in $\mathcal{F}(k,h,t)$ for which either (i) or (ii) does not hold. Consider the coefficients $b_{p,q}$ of the equations in $\mathcal{F}(k,h,t)$ as unknowns, and use Claim 3.1 in order to obtain an equivalent diagonalized set of equations with unknowns $a_{i,j}$ that are given by rational functions in $b_{p,q}$. Denote this new set by S_{diag}.

We first show that only $c(k)h^{e(k-1)-1}$ of the equations of $\mathcal{F}(k,h,t)$ are such that (i) does not hold. Let B be the product of the denominators of all the unknowns $a_{i,j}$ in S_{diag}. As each $a_{i,j}$ is a rational function of degree at most $r(k)$ and there are fewer than $tk < k^2$ coefficients $a_{i,j}$, one may conclude that B is a polynomial of degree at most $r(k)$ in each of the unknowns $b_{p,q}$. As by Claim 3.1 none of the denominators is identically zero, B is not identically zero. Clearly if for some assignment to the unknowns $b_{p,p}$ the polynomial B does not vanish, then S_{diag} is well defined. By Lemma 3.2 with $F = R$, $H = [-h,h]$, $r = r(k)$ and $b = e(r-1)$ we find that there are at most $c(k)h^{e(k-1)-1}$ assignments to the coefficients $b_{p,q}$ for which B vanishes, therefore there are at most this many equations in $\mathcal{F}(k,h,t)$ for which S_{diag} is not well defined.

We turn to show that only $c(k)h^{e(k-1)-1}$ of the equations of $\mathcal{F}(k,h,t)$ are such that (ii) does not hold. Let W be the first $\binom{t}{2}-t+1$ equations of S_{diag}, and note that they form $\binom{t}{2}-t+1$ equations in $\binom{t}{2}$ unknowns. We will prove a statement somewhat stronger than needed, namely, that only $c(k)h^{e(k-1)-1}$ of the equations of $\mathcal{F}(k,h,t)$ are such that the points of the parametric representation of W all lie on some non-zero $(t-2)$-dimensional quadric. Consider the set W with unknowns $a_{i,j}$. As in Claim 2.5, let $A = A_{t-2}(W)$ be the matrix in Definition 2.4, and D the polynomial of its determinant in unknowns $a_{i,j}$. By Definition 2.3, the $\binom{t}{2}$ points in $\mathbb{R}^{t-2}$ of the parametric representation of W cannot all lie on a $(t-2)$-dimensional non-zero quadric if A is invertible. Equivalently, these points do not lie on a non-zero quadric if the value of the polynomial D defined above, when evaluated on the coefficients of this set of equations, is defined and non-zero. It thus follows that we can simply show that D is either not defined or vanishes only on $c(k)h^{e(k-1)-1}$ of the assignments (to the unknowns $b_{p,q}$ which determine the value of $a_{i,j}$) which consist of integers from $[-h,h]$.

Now recall that each $a_{i,j}$ is a rational function in the unknowns $b_{i,j}$ of degree at most $r(k)$ and by Claim 2.5 the degree of each $a_{i,j}$ in D is at most t^2. Thus, D is also a rational function of degree at most $r(k)$ in each $b_{p,q}$. Let D_1 and D_2 be its numerator and denominator, respectively. As D_2 is a product of the denominators of $a_{i,j}$, and by Claim 3.1 none of them is identically zero, neither is D_2. We claim that D_1 is also not identically zero. Suppose that D_1 is identically zero. This means that D is identically zero, hence, for every assignment to the unknowns $b_{i,j}$, which does *not* need to

be necessarily of integers from $[-h,h]$), the resulting set is such that if we transform it into an equivalent diagonalized set, we either get a set that is not well defined (in the case where D_2 vanishes) or the first $\binom{t}{2}$ equations of the set are degenerate. This, however, is false, as we can a priori set the first $\binom{t}{2}-t+1$ equations of the set to be the (well defined) non-degenerate diagonalized set of equations whose existence is guaranteed by Corollary 2.6.

Consider now the product of D_1 and D_2. This is a non-zero polynomial in $b_{p,q}$ of degree $r(k)$ in each unknown. By Lemma 3.2, with $F=R$, $H=[-h,h]$, $r=r(k)$ and $b=e(k-1)$, either D_1 or D_2 vanishes on at most $c(k)h^{e(k-1)-1}$ of the possible assignments to $b_{p,q}$. Thus, there are at most this many sets in $\mathcal{F}(k,h,t)$ for which D is either not defined or zero, which is what we wanted to show. ∎

As in Behrend's construction [1], we will represent integers in base g with g being a non-fixed integer, namely depending on n. This will allow us to look at integers as high-dimensional vectors. In Behrend's argument, one shows that a sphere in the high-dimensional vector space in which our integers are represented, does not contain three-term arithmetic progressions. The main step in our proof is to prove an analogous statement, Lemma 3.5 below, based on Lemma 3.3. We will then conclude the proof by showing that for an appropriate choice of g we get a large set with no non-trivial solution of S.

Given a set V of integers we denote by $V+r$ the *translate* of V by r, that is, $V+r=\{z+r : z\in V\}$. Note that if V does not contain any non-trivial solution of a set $S\in\mathcal{F}(k,h,t)$ then neither does any translate of V (this is because the sum of the coefficients is zero). As we will explain shortly, it will be simpler to prove Theorem 3 with respect to the set of integers $[-n/2,n/2]$ rather than $[n]$. To this end, we will consider representations of integers from $[-n/2,n/2]$ in base g. For integer g and b satisfying $n=g^b$, we define, for an integer $w\geq 2$,

$$Q_w=\Big\{z\in\mathbb{Z} : z=\sum_{i=0}^{b-1} z_i g^i,\ z_i\in[-g/w,g/w]\Big\}.$$

In other words, these are the integers whose "digits" in base g belong to $[-g/w,g/w]$. As $Q_w\subseteq[-n/2,n/2]$ for any $w\geq 2$, we may and will construct our sought-after sets from integers belonging to Q_w for an appropriate constant w. Note that, somewhat unconventionally, we allow for negative digits. This representation, however, is well defined in the sense that given $z\in Q_w$, there are unique integers $z_0,\ldots,z_{b-1}\in[-g/w,g/w]$ such that $z=\sum_{i=0}^{b-1} z_i g^i$. Given an integer $z\in Q_w$ we denote by $\overline{z}=(z_0,\ldots,z_{b-1})$ the unique b-dimensional vector in $\mathbb{Z}^b$ such that $z=\sum_{i=0}^{b-1} z_i g^i$. We also write $\|z\|^2=\|\overline{z}\|^2=\sum_{i=0}^{b-1} z_i^2$. Our argument will critically rely on the observa-

tion (first made by Salem and Spencer [8]) that if w is sufficiently large then addition, and more generally linear combinations with small coefficients, of numbers from Q_w is equivalent to linear combinations of their corresponding vectors. For example, if $z_1, z_2, z_3 \in Q_2$, then $z_1 + z_2 = z_3$ if and only if $\overline{z}_1 + \overline{z}_2 = \overline{z}_3$. The reason for that is simply that there is no carry in the base g addition of these numbers. More generally, we have the following

FACT 3.4. *Suppose $\alpha, \alpha_1, \ldots, \alpha_t$ are rational numbers with numerators and denominators bounded in absolute value by some constant c. If w is sufficiently large with respect to c, then for every $z, z_1, \ldots, z_t \in Q_w$,*

$$\alpha z = \sum_{i=1}^{t} \alpha_i z_i \;\Leftrightarrow\; \alpha \overline{z} = \sum_{i=1}^{t} \alpha_i \overline{z}_i. \tag{4}$$

It should be noted that had we chosen to work with the set $[n]$ rather than $-n/2, \ldots, n/2$ and represented integers using positive digits, then (4) would not necessarily hold for negative coefficients. The reason is that the difference of two numbers with small positive digits may contain very large digits. As we also allow for negative digits, the difference also contains small digits. We now arrive at the main step of the proof, where we prove a Behrend-type argument about spheres containing the vector representations of integers from Q_w.

LEMMA 3.5. *For $t \geq 4$ and $k \geq \binom{t}{2}$ let $S \in \mathcal{F}(k, h, t)$ be such that S_{diag} is well defined and the points of the parametric representation of S_{diag} do not all lie on any non-zero $(t-2)$-dimensional quadric. For an integer r let*

$$X_r = \{z \in Q_w : \|z\|^2 = r\}.$$

If w is large enough in terms of h and k, then X_r contains no non-trivial solution of S.

Proof. Suppose to the contrary that X_r contains k integers $z_0, z_1, \ldots, z_{k-1}$, which form a non-trivial solution of S. As S_{diag} is assumed to be well defined, S and S_{diag} are equivalent, thus, $z_0, \ldots, z_{k-1}$ are also a solution of S_{diag}. Also, by Claim 3.1 the coefficients of S_{diag} are rational numbers, whose numerators and (non-zero) denominators are bounded by $m = m(h, k)$ (because these numerators and denominators are polynomials in h of degree bounded by a function of k). By Claim 2.2, there are $(t-2)$-tuples $p_0, \ldots, p_{k-1}$ which form the parametric representation of $z_0, \ldots, z_{k-1}$. That is, for $0 \leq i \leq k-1$ we have

$$z_i = z_0 + p_{i,1} \frac{z_1 - z_0}{d} + \cdots + p_{i,t-2} \frac{z_{t-2} - z_0}{d}. \tag{5}$$

Recall that by Claim 2.2, each $p_{i,j}$ is either d or $da_{i,j}/d_i$ and that $d = \prod d_i$. Thus, as by assumption all the coefficients of S_{diag} are rational numbers whose numerators and denominators are bounded in absolute value by $m =$

$m(h,k)$, the numerators and denominators of the rational numbers $p_{i,j}$ are each bounded in absolute value by a function depending only on m and k. As m depends on h and k, we conclude that the numerators and denominators of $p_{i,j}$ are bounded by a function of h and k. Multiplying (5) by d we deduce that for $0 \le i \le k-1$ we have

$$dz_i = dz_0 + p_{i,1}(z_1 - z_0) + \cdots + p_{i,t-2}(z_{t-2} - z_0). \tag{6}$$

Hence, if w is large enough in terms of h and k, we can use (4) to write (6) as [1]

$$d\overline{z}_i = d\overline{z}_0 + p_{i,1}\overline{z_1 - z_0} + \cdots + p_{i,t-2}\overline{z_{t-2} - z_0}. \tag{7}$$

Define the following $(t-2)$-variate polynomial of degree 2:

$$P(x_1,\dots,x_{t-2}) := \sum_{q=0}^{b-1}(d(\overline{z}_0)_q + x_1(\overline{z_1 - z_0})_q + \cdots + x_{t-2}(\overline{z_{t-2} - z_0})_q)^2,$$

where $(\overline{v})_q$ denotes the q^{th} entry of the vector $\overline{v}$. The key observation now is that by (7), for $0 \le i \le k-1$ we have

$$P(p_{i,1},\dots,p_{i,t-2}) = \|d\overline{z}_j\|^2 = d^2\|z_j\|^2.$$

Therefore, as by assumption all the integers z_i belong to X_r, for $1 \le i \le k-1$ we have

$$P(p_{i,1},\dots,p_{i,t-2}) - d^2 r = 0.$$

Hence, the k points $p_0,\dots,p_{k-1} \in \mathbb{R}^{t-2}$ all lie on the $(t-2)$-dimensional quadric defined by the equation $P(x_1,\dots,x_{t-2}) - d^2 r = 0$. This will contradict our choice of the coefficients of the equations once we show that $P(x_1,\dots,x_{t-2}) - d^2 r$ is not identically zero. To see that, note that P is a sum of squares, hence, the coefficients of the monomials $x_1^2,\dots,x_{t-2}^2$ are sums of squares. Therefore, it is enough to show that for some $1 \le j \le t-2$, and $0 \le q \le b-1$, we have $(\overline{z_j - z_0})_q \ne 0$. This, however, must be true, as otherwise we would get $z_1 - z_0 = z_2 - z_0 = \cdots = z_{t-2} - z_0 = 0$ and thus $z_0 = z_1 = \cdots = z_{k-1}$ by (5), contradicting our assumption that these integers form a non-trivial solution. ∎

As in Behrend's argument, we are now just left with the task of optimizing the values of g and b (recall that we write $n = g^b$) in order to deduce that one of the sets X_r contains many integers.

Proof of Theorem 3. For an integer $k \ge 6$, let t be the largest integer satisfying $k \ge \binom{t}{2}$ (hence, $t \ge 4$). In Lemma 3.3 it is proved that if $t \ge 4$ and $k \ge \binom{t}{2}$, then only a $c(k)/h$ fraction of the sets $S \in \mathcal{F}(k,h,t)$ of equations are

[1] We actually use (4) twice: first, to deduce that $z_1 - z_0,\dots,z_{t-2} - z_0 \in Q_{w'}$ for some $w' \le w$, and second, to show that this implies (7).

such that either S_{diag} is not well defined or the points of the parametric representation of S_{diag} all lie on a non-zero $(t-2)$-dimensional quadric. We claim that for any other set S we can construct a subset $X \subseteq [n]$ of size $n/e^{c\sqrt{\log n}}$ with no non-trivial solution of S, where $c = c(h,k)$. By Lemma 3.5, provided w is large enough in terms of h and k, for any integer r the set X_r contains no non-trivial solution of S. As the absolute value of each digit in Q_w is bounded by g/w, the integer r can take at most $b(g/w)^2 \le bg^2$ values. Similarly, we find that Q_w is of size $(2g/w)^b > (g/w)^b$. As the union of the sets X_r covers the entire set Q_w there must be one r for which $|X_r| \ge (g/w)^b/bg^2 = n/bg^2w^b$. Setting $b = \sqrt{\log n}$, and hence $g = e^{\sqrt{\log n}}$, gives that some X_r is of size at least $n/e^{c\sqrt{\log n}}$ for an appropriate constant $c = c(h,k)$. ∎

As was alluded to in the introduction, the details of Lemma 3.5 actually show that under the conditions of the lemma, X_r contains no solution in which two of the integers are distinct. Therefore, the only solution of the sets S, discussed in the proof of Theorem 3, which use integers from the set constructed in the proof of the theorem is one in which all the integers are identical. This is clearly a much stronger property than not containing non-trivial solutions.

One can also derive, from the details of the proof of Theorem 3 and the claims and lemmas used in its proof, the following simple criterion for a given set S to satisfy $r_S(n) > n^{1-o(1)}$: Let M be the $(k-t+1)\times k$ matrix containing the coefficients of the equations of the set. If M cannot be transformed into a diagonalized set we cannot say anything. If it can be transformed into such a set, then let S_{diag} be the new set of equations, and let $A(S_{\text{diag}})$ be the matrix from Definition 2.4. Then, if $A(S_{\text{diag}})$ is non-singular, we have $r_S(n) > n^{1-o(1)}$.

4. Additional results and open problems. The main result of this paper establishes that most large enough sets S of equations are such that $r_S(n) > n^{1-o(1)}$. As observed at the end of Section 3, the details of the proof actually give a simple sufficient criterion for a specific set S to have $r_S(n) > n^{1-o(1)}$. In this section we give upper and lower bounds for $r_S(n)$ based on specific properties of S.

The first question one may ask is whether there are any sets of equations that do not satisfy $r_S(n) > n^{1-o(1)}$ and more interestingly, how large such a set can be. Of course, one can construct arbitrarily large such sets simply by taking many copies of an equation E for which $r_E(n) \le n^{1-\varepsilon}$. Recall that by Theorem 2 any equation E with genus $g \ge 2$ satisfies $r_E(n) \le \sqrt{n}$. It therefore seems more reasonable to require the equations of the set to be linearly independent. Even with this requirement it is easy to show that there are large sets which satisfy $r_S(n) \le n^{1-\varepsilon}$.

PROPOSITION 4.1. *For every $k \geq 4$ and h, there is a set S of $k-3$ linearly independent (k,h)-equations for which $r_S(n) = O(\sqrt{n})$.*

Proof. For every $k \geq 4$ consider the set S of equations $e_1, \ldots, e_{k-3}$ in unknowns $x_1, \ldots, x_k$ where for $1 \leq i \leq k-3$ equation e_i is $x_1 + x_2 - x_{2+i} - x_{3+i} = 0$. Clearly, these equations are linearly independent. To show that $r_S(n) = O(\sqrt{n})$, recall that for Sidon's equation $E := x_1 + x_2 - x_3 - x_4 = 0$ it is known (see [7]) that $r_E(n) = (1+o(1))\sqrt{n}$. Thus, any set of size at least $(1+o(1))\sqrt{n}$ contains a non-trivial solution of E, namely z_1, z_2, z_3, z_4 such that $z_1 + z_2 - z_3 - z_4 = 0$. Note that if we assign $x_1 = z_1$, $x_2 = z_2$, $x_3 = x_5 = x_7 = \cdots = z_3$ and $x_4 = x_6 = x_8 = \cdots = z_4$ we get a non-trivial solution of S. ∎

In what follows, a set S of equations $e_1, \ldots, e_t$ in $2g$ unknowns is called *symmetric* if every equation $e_i \in S$ is of the form

$$a_{i,1}x_1 + \cdots + a_{i,g}x_g - a_{i,g+1}x_{g+1} - \cdots - a_{i,2g}x_{2g} = 0 \tag{8}$$

where $a_{i,j} = a_{i,j+g}$ for every $1 \leq j \leq g$. In the following proofs, it will be more convenient to write a symmetric equation as

$$a_{i,1}x_1 + \cdots + a_{i,g}x_g = a_{i,g+1}x_{g+1} + \cdots + a_{i,2g}x_{2g}. \tag{9}$$

We now turn to the proof of Theorem 4, which establishes a connection between the genus of a set of equations, its rank and $r_S(n)$.

Proof of Theorem 4. First note that a non-trivial solution of a subset of a set of equations cannot be a trivial solution of the entire set. Thus, it is enough to show that every subset of $[n]$ of size $\Omega(n^{t/g})$ contains a non-trivial solution of the t equations that span the entire set of equations (whose rank is t by assumption). We will thus confine ourselves to sets of t linearly independent equations. The proof will mostly follow the main idea of the proof of Theorem 3.6 in [7].

Consider first any symmetric set S of t equations in $2g$ unknowns $x_1, \ldots, x_{2g}$ as in (9). We claim that for such a set $r_S(n) = O(n^{t/g})$. The proof for general sets of equations will follow by a certain reduction to this special case. Fix any subset $X \subseteq [n]$. Let $s(b_1, \ldots, b_t)$ denote the number of solutions of the set of t equations $a_{i,1}x_1 + \cdots + a_{i,g}x_g = b_i$ with integers x_i taken from X. As all the coefficients $a_{i,j}$ belong to $[-h,h]$, the only feasible b_i are such that $-ghn \leq b_i \leq ghn$. In other words,

$$\sum_{-ghn \leq b_1, \ldots, b_t \leq ghn} s(b_1, \ldots, b_t) = |X|^g. \tag{10}$$

Observe that $\sum s^2(b_1, \ldots, b_t)$ is precisely the number of solutions of S including the trivial solutions (recall that $a_{i,j} = a_{i,j+g}$ for every $1 \leq j \leq g$).

By Jensen's inequality and (10) we get

$$\sum_{-ghn\le b_1,\ldots,b_t\le ghn} s^2(b_1,\ldots,b_t) \ge \frac{|X|^{2g}}{(2ghn)^t}. \tag{11}$$

We now bound the total number of trivial solutions of S. Consider any partition of $[2g]$ into disjoint non-empty subsets $A_1,\ldots,A_w$ such that $\sum_{p\in A_j} a_{i,p} = 0$ for every equation i and all j. As every A_j must be of size at least 2, we have $w \le g$. It follows that for every such partition the total number of solutions in which $x_{i_1} = x_{i_2}$ whenever i_1, i_2 belong to the same set A_j is at most $|X|^g$. Hence, the total number of trivial solutions is at most $c(g)|X|^g$ where $c(g)$ is the total number of partitions of $[2g]$. Now, if X contains no non-trivial solution of S, we must have $|X|^{2g}/(2ghn)^t \le c(g)|X|^g$ and therefore $|X| = O(n^{t/g})$.

Now, consider an arbitrary set of t equations $u_1,\ldots,u_t$ in k unknowns of genus g. As the set has genus g, there is a partition of $[k]$ into $A_1,\ldots,A_g$ such that $\sum_{p\in A_j} a_{i,p} = 0$ for every equation u_i and all j. For every j pick any integer $r(j) \in A_j$ and for every $1 \le i \le t$ and $1 \le j \le g$ set $b_{i,j} = b_{i,j+g} = a_{i,r(j)}$. For each equation u_i consider an auxiliary symmetric equation e_i,

$$b_{i,1}y_1 + \cdots + b_{i,g}y_g = b_{i,1+g}y_{1+g} + \cdots + b_{i,2g}y_{2g}. \tag{12}$$

Observe that any solution of $e_1,\ldots,e_t$ can be transformed into a solution of $u_1,\ldots,u_t$ by setting, for every $1 \le j \le g$, $x_{r(j)} = y_j$ and $x_p = y_{j+g}$ for every $r(j) \ne p \in A_j$ (recall that for every equation u_i and all j we have $\sum_{r(j)\ne p\in A_j} a_{i,p} = -a_{i,r(j)}$). We further claim that a non-trivial solution of the equations $e_1,\ldots,e_t$ translates to a non-trivial solution of $u_1,\ldots,u_t$. Assume this is not the case. Then for any integer ℓ we have

$$\sum_{\{i\,:\,x_i=\ell\}} a_i = 0. \tag{13}$$

By the definition of the transformation and the fact that $b_{i,j} = b_{i,j+g}$ we also have

$$\sum_{\{i\,:\,x_i=\ell\}} a_i = \sum_{\{i\,:\,y_i=\ell\}} b_i. \tag{14}$$

As by assumption the solution of $e_1,\ldots,e_t$ is non-trivial there is some ℓ for which the right side of (14) does not vanish, thus contradicting (13). The proof is now complete, as by the first part of the proof, the largest subset of $[n]$ with no non-trivial solution of $e_1,\ldots,e_t$ is of size $O(n^{t/g})$. ∎

While it seems reasonable that for a single equation with genus g we have $r_E(n) = n^{1/g-o(1)}$, Theorem 5, which we now prove, shows that for sets of equations this is far from being the case. This theorem also shows that the bound of Theorem 4 cannot be generally improved.

Proof of Theorem 5. For every integer $g \geq 2$ consider a symmetric set of g equations $e_1, \ldots, e_g$ in $2g$ unknowns $x_1, \ldots, x_{2g}$ as in (9). Clearly this set has genus g, as in every equation the sum of the coefficients of x_j and x_{j+g} is 0. Consider the matrix G with $G_{i,j} = a_{i,j}$. Clearly if G is non-singular (e.g. when $a_{i,i} = a_{i,i+g} = 1$ and all the other entries are 0) the set of equations has rank g. We claim that if G is non-singular then $r_S(n) = n$. To show this, it is enough to show that every solution of S is trivial. Consider any solution $z_1, \ldots, z_{2g}$ of S and let $b_1, \ldots, b_k$ satisfy $b_i = a_{i,1}z_1 + \cdots + a_{i,g}z_g$. As G is invertible, there are unique $z_1, \ldots, z_g$ for which the value of the right hand side of (9) is b_i for every equation e_i. As the left hand side of (9) has the same coefficients, it must be the case that $z_{j+g} = z_j$ for every $1 \leq j \leq g$. This means that the solution is trivial. ∎

By using Lemma 3.2, one can easily strengthen Theorem 5 by showing that in fact, all but an $O(1/h)$ fraction of the symmetric sets of g equations in $2g$ unknowns are such that $r_S(n) = n$. The reason is simply that for most sets the matrix G is invertible. We omit the details. Given Theorems 4 and 5 one may consider the possibility that for a set S of equations, with genus g and rank t we have $r_S(n) = \min(n, n^{t/g-o(1)})$. However, note that in Proposition 4.1 we construct a set of $k-3$ linearly independent equations of genus 2 which satisfies $O(\sqrt{n})$, thus ruling out this possibility. On the positive side, we now turn to the proof of Theorem 6, which shows that a somewhat weaker lower bound on $r_S(n)$ may hold.

Proof of Theorem 6. Clearly a positive answer to Problem 2 implies a positive answer to Problem 1. So assume that for every linear equation E of genus g we have $r_E(n) > n^{1/g-o(1)}$. Consider any set S of size t of (k,h)-equations with genus b. Pick t integers $c_1, \ldots, c_t$, where each c_i is chosen independently and uniformly at random from $\{1, \ldots, 2^k\}$. Consider a linear combination E of the t equations with coefficients $c_1, \ldots, c_t$. Let $b_1, \ldots, b_k$ be the coefficients of E. We claim that with positive probability, E has the following property: A set $\emptyset \neq A \subseteq [k]$ satisfies $\sum_{p \in A} b_p = 0$ if and only if $\sum_{p \in A} a_{i,p} = 0$ for all equations i. As the "if" part is obvious we prove the other direction. Fix any set A such that for some i we have $\sum_{p \in A} a_{i,p} \neq 0$. Conditioning on any choice of the coefficients $c_1, \ldots, c_t$ other than c_i, the probability that c_i is such that $\sum_{p \in A} b_p = 0$ is at most 2^{-k}. Therefore, the probability that $\sum_{p \in A} b_p = 0$ is also at most 2^{-k}. As there are fewer than 2^k possible choices of the set A, we conclude by the union bound that with positive probability, E has the desired property. In particular, E must have genus exactly b. By assumption, there is a subset of $[n]$ of size $n^{1/g-o(1)}$ with no non-trivial solution of E. As by the property of E discussed above, any non-trivial solution of S is also a non-trivial solution of E, this set contains no non-trivial solution of S. ∎

In light of the results proved in this section, and the gap between the upper bound of Theorem 4 and the (possible) lower bound of Theorem 6, it seems that one would have to define and study other properties of equations, besides their genus and rank, in order to determine the value of $r_S(n)$. It seems very interesting to further explore this problem. It also seems interesting to strengthen Theorem 3 by showing that most sets of equations in k unknowns of size smaller than $k - \lfloor\sqrt{2k}\rfloor + 1$ are such that $r_S(n) > n^{1-o(1)}$.

Acknowledgements. I would like to thank my advisor, Prof. Noga Alon, for a careful reading of this paper, as well as for suggestions that significantly improved its presentation.

References

[1] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. USA 32 (1946), 331–332.

[2] H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math. 31 (1977), 204–256.

[3] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. 11 (2001), 465–588.

[4] S. Jukna, *Extremal Combinatorics With Applications in Computer Science*, Springer, 2001.

[5] I. Laba and M. Lacey, *On sets of integers not containing long arithmetic progressions*, unpublished manuscript; http://arxiv.org/abs/math.CO/0108155.

[6] R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A 65 (1962), 332–344.

[7] I. Z. Ruzsa, *Solving a linear equation in a set of integers I*, Acta Arith. 65 (1993), 259–282.

[8] R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. USA 28 (1942), 561–563.

[9] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.

School of Computer Science
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv, Israel
E-mail: asafico@tau.ac.il

Received on 1.9.2004
and in revised form on 6.10.2005 (4845)