# How to color shift hypergraphs

Noga Alon \*

Department of Mathematics Raymond and Beverly Sackler Faculty of Exact Sciences Tel Aviv University, Tel Aviv, Israel Igor Kriz

Department of Mathematics University of Chicago, Chicago, IL 60637, USA and Jaroslav Nešetřil Department of Applied Mathematics Charles University, 11800 Praha 1, Czechoslovakia

#### Abstract

Let g(k) denote the minimum integer m so that for every set S of m integers there is a k-coloring of the set of all integers so that every translate of S meets every color class. It is a well known consequence of the Local Lemma that g(k) is finite for all k. Here we present a new proof for this fact, that yields a very efficient parallel algorithm for finding, for a given set S, a coloring as above. We also discuss the problem of finding colorings so that every translate of S has about the same number of points in each color. In addition, we prove that for large k

 $(1 + o(1))k \log k \le g(k) \le (3 + o(1))k \log k.$ 

<sup>\*</sup>Research supported in part by a United States Israel BSF Grant

### 1 Introduction

Strauss (cf. [8]) raised the following problem: Is there a function g(k) ( $< \infty$ ) such that for every set S of at least g(k) integers there is a coloring of the integers by k colors so that every translate of S meets all the colors? This problem was solved by Erdős and Lovász [8], who proved that

$$g(k) \le O(k \log k). \tag{1}$$

The proof of [8] is probabilistic and uses the Lovász Local Lemma, a result that has been used for tackling many other combinatorial problems in numerous subsequent papers. As remarked in [12] there is no known proof for the finiteness of f(k) that does not use the Local Lemma.

Our first result in the present short paper is such a proof, namely a solution of Strauss' problem that does not apply the Local Lemma. Although our basic solution works only for sets S of cardinality at least  $4k^2$  it has the advantage that it is more constructive than the original solution of [8], and yields very efficient deterministic and parallel algorithms for finding a coloring of the integers with the required properties. As is the case with many applications of the Local Lemma, the proof in [8] supplies neither a randomized nor a deterministic polynomial time algorithm for finding, given a set S of a sufficiently large cardinality, a k-coloring so that each translate of S meets every color. The recent technique of J. Beck [5] and its modification in [1] that supply efficient sequential and parallel deterministic algorithms for various applications of the Local Lemma do not seem to apply directly to the problem of Strauss mentioned above, even when the cardinality of S is much larger than  $\Theta(k \log k)$ . Note that here the length of the input is the number of bits in the representation of S whereas it is not even clear if the output can be represented with a finite number of bits, as the output is a k-coloring of the infinite set of integers.

Our technique here yields a very efficient parallel algorithm that produces, for a given set S of at least  $4k^2$  integers a k-coloring of the integers so that every translate of S meets every color class. In fact, the required coloring can be found in *constant* time in the standard model for parallel computation known as a CRCW PRAM with a polynomial number of parallel processors. (See, e.g., [11] for the exact definition of a CRCW PRAM; we assume that each processor is capable of adding, comparing, multiplying or dividing numbers of size as that of the members of S in constant time). Since the required k-coloring is a coloring of an infinite set we agree that the k-coloring is produced successfully if it can be described by a polynomial number of bits that enable us, given

any integer x, to compute the color of x efficiently; in our case this will be done by a constant number of modular additions and multiplications. The basic approach can be combined with the technique in [5] and yield efficient sequential coloring algorithms even if S has only  $ck \log k$  elements for some (large) constant c. Still, we believe that the most interesting consequence of the argument is the new method for solving Strauss' original problem.

Another result we prove here is the fact that the estimate in (1) is sharp.

In order to formulate our results and proofs in a more concise form we introduce two definitions. For a set of integers S, let H = H(S) denote the infinite hypergraph whose set of vertices is the set Z of all integers and whose set of edges is the set of all translates of S, i.e., the set  $\{x + S : x \in Z\}$ . We call H the *shift hypergraph of* S. A k-coloring  $c : Z \mapsto \{1, 2, ..., k\}$  is called *good (for* H), if every edge of H meets every color class, i.e., if for every  $i, 1 \le i \le k$  and for every integer x there is an  $s \in S$  so that c(x + s) = i.

In this notation, our two main results are the following.

**Theorem 1.1** Let S be a set of at least  $4k^2$  integers. Then there exists a good k-coloring c for the shift hypergraph H(S). Such a coloring can be found in constant time, using a polynomial number of parallel processors on a CRCW PRAM. In addition, there exists a positive constant c such that for every set S of at least  $ck \log k$  integers one can find a good k-coloring for the shift hypergraph H(S) in (sequential) polynomial time.

**Theorem 1.2** There exists an absolute positive constant a such that for every k > 1 there is a set  $S = S_k$  of at least  $k \log k$  integers so that there is no good k-coloring for the shift hypergraph H(S).

The proof of Theorem 1.1 is based on the ideas of [3] and is presented in the next section, together with some related extensions. In Section 3 we describe two proofs of Theorem 1.2; a probabilistic one and a constructive one. The final Section 4 contains some concluding remarks.

## 2 Finding a good coloring

**Proof of Theorem 1.1** Let S be a given set of  $m = 4k^2$  distinct integers. Our objective is to find a good k-coloring c for the shift hypergraph H(S). To do so, we first choose a prime p so

that the members of S are pairwise distinct modulo p. Let  $P = \{0, 1, \dots, p-1\}$  be the set of all remainders modulo p and let us split P into k pairwise disjoint intervals of consecutive remainders  $I_1, I_2, \dots, I_k$ , where  $\lfloor p/k \rfloor \leq \lfloor p/k \rfloor$  for all  $1 \leq i \leq k$ .

For two integers a and b in P, let  $c = c_{a,b}$  be the following k-coloring of the set of integers. For every integer y, c(y) is the unique i so that  $(ay + b) \pmod{p} \in I_i$ . Define

$$Y = Y_{a,b} = \{ (as+b) \ (mod \ p) : s \in S \}$$

We claim that if a, b are chosen in such a way that the set Y intersects every (cyclic) interval of length  $\lfloor p/k \rfloor$  in P then every translate of S intersects each color class of c. To see this, observe that if x + S is a translate of S then the set

$$\{(ay+b) \pmod{p} : y \in x+S\}$$

is a cyclic translate of Y and hence it intersects every interval of length  $\lfloor p/k \rfloor$  in P and in particular it intersects every  $I_i$ , implying the desired result. It thus suffices to choose a, b so that  $Y_{a,b}$  has the above property. We next show that this can be done.

Fact: If a and b are chosen randomly and independently in P, according to a uniform distribution, then with positive probability the set  $Y = Y_{a,b}$  intersects every cyclic interval of length at least  $\lfloor p/k \rfloor$  in P.

**Proof:** The argument essentially appears in [3], but since it is very short we repeat it here. Let  $J_1, \ldots, J_{2k}$  be a fixed covering of P by 2k intervals of length  $\lceil p/2k \rceil$  each. Observe that if Y intersects each  $J_i$  then it certainly satisfies the required property, since every cyclic interval of length  $\lfloor p/k \rfloor$  must fully contain at least one interval  $J_i$ . Fix an  $i, 1 \le i \le 2k$  and put  $m = |S| = 4k^2$ . For each element  $s \in S$  let  $X_s^i$  be the indicator random variable whose value is 1 if  $(as + b) \pmod{p} \in J_i$  (and is 0 otherwise). Define  $X^i = \sum_{s \in S} X_s^i$  and observe that  $X^i = 0$  if and only if Y does not intersect  $J_i$ . We estimate the probability of this event by computing the expectation and variance of  $X^i$ . By linearity of expectation

$$E(X^i) = \sum_{s \in S} E(X^i_s) = m \lceil p/2k \rceil / p.$$

The crucial (and simple) fact for computing the variance is the fact that the random variables  $X_s^i$ ,  $(s \in S)$  are pairwise independent. This follows from the fact that if s and t are two distinct

members of S then when the pair (a, b) ranges over P x P so does the pair

$$((as + b)(mod p), (at + b)(mod p)),$$

implying that  $X_s^i$  and  $X_t^i$  are independent. Hence,

$$VAR(X^{i}) = \sum_{s \in S} VAR(X_{s}^{i}) = m \frac{\lceil p/2k \rceil}{p} (1 - \frac{\lceil p/2k \rceil}{p}).$$

Therefore, by Chebyschev's Inequality,

$$Prob(X^{i} = 0) \le Prob(|X^{i} - E(X^{i})| \ge E(X^{i})) \le \frac{VAR(X^{i})}{E(X^{i})^{2}} < 2k/m$$

Since there are 2k possible values of i the probability that  $X^i = 0$  for some i is strictly smaller than  $4k^2/m = 1$ , completing the proof of the fact.  $\Box$ 

Returning to the proof of Theorem 1.1 observe that the last fact implies that randomly chosen a and b supply a good k-coloring with positive probability. In particular, there is at least one such pair a, b and hence a good coloring exists.

For the algorithm, it is essential to choose a small prime p so that all members of S are distinct modulo p, that is, a prime which is smaller than some (fixed) polynomial in the length of the input. Fortunately, the existence of such a prime is simple. A prime p is not good if and only if it divides the product

$$\prod_{a,s'\in S, s>s'} (s-s').$$

If each number in S is at most  $2^n$  (and at least 0, as we may assume since the problem is invariant under any additive shift of S), and S has m members, the last product is certainly at most  $2^{nm^2}$ . Since it is well known that the product of all primes smaller than x is  $e^{(1+o(1))x}$  this showes that there is a prime  $p < nm^2$  that does not divide the product.

Therefore, for the algorithm, we simply check, in parallel, for every prime p up to  $nm^2$  if it is good, i.e., if all the members of S are distinct modulo p. Once we find such a prime we check, in parallel, all pairs a, b and find a pair for which the set  $Y_{a,b}$  intersects every interval of length  $\lfloor p/k \rfloor$  (we can afford checking all the intervals in parallel). One can check that all this can, indeed, be performed in constant time in parallel using polynomially many processors. (Recall that each processor can add, compare, multiply and divide numbers in the required range in constant time). This completes the proof of the first part of the theorem. For the second part of the proof of Theorem 1.1, namely the existence of the constant c, we only need to combine one simple ingredient of the proof above with the technique of [5]. Given a set S of  $m \ge ck \log k$  positive integers, each at most  $2^n$ , we can find efficiently a prime p so that  $p < nm^2$  and all the members of S are distinct modulo p. Consider the hypergraph H whose set of vertices is the set  $Z_p$  of integers modulo p and whose set of edges is the set of all shifts of S modulo p, i.e., the set

$$\{ \{ (x+s) \ (mod \ p) : s \in S \} : x \in Z_p \}.$$

*H* is clearly an *m*-uniform, *m*-regular hypergraph, and the technique in [5] can be applied to obtain in sequential polynomial time, a vertex *k*-coloring of *H* so that each edge meets every color class, provided  $m \ge ck \log k$  for a sufficiently large *c*. Since this is very similar to the examples given in [5], we omit the detailed algorithm. The coloring can now be extended to a good integer *k*-coloring for H(S) simply by letting the color of any integer *y* be the color of  $y(mod \ p)$ . This completes the proof of Theorem 1.1.  $\Box$ 

#### Remarks

1. Combining the basic idea in the last proof with one of the results in [3] we can show that for a sufficiently large set S one can find a coloring with small discrepancy, in the following sense.

**Proposition 2.1** For every set S of m integers, there is a k-coloring of the integers so that the number of points of each color in every translate of S deviates from m/k by at most

$$O(m^{1/2}(\log m)^{3/2}).$$

Such a coloring can be found in constant time with polynomially many processors on a CRCW PRAM.

To prove this result we start, as before, by choosing a polynomially small prime p so that all the members of S are distinct modulo p. Next we use Theorem 2.3 in [3] which asserts that for every set T of cardinality m in  $Z_p$  there is an integer a so that the set aT(mod p)is uniformly distributed in the sense that for every (cyclic) interval of length  $\delta p$  in  $Z_p$  the number of members of aT(mod p) in the interval deviates from its expectation  $\delta m$  by at most  $O(m^{1/2}(\log m)^{3/2})$ . Let  $I_1, \ldots, I_k$  be a partition of  $Z_p$  into k almost equal intervals and define, for every integer y, the color of y as the unique i such that  $ay(mod \ p) \in I_i$ , where a is chosen so that for  $T = S(mod \ p)$ ,  $aT(mod \ p)$  is uniformly distributed as described above. The same reasoning as in the last proof shows that this coloring satisfies the desired requirements. It is also obvious that it can be found in constant time with polynomially many processors in parallel, as before.

We note that one can obtain an even more uniform k-coloring for the shift hypergraph H(S) by applying the local lemma, but we do not know to find such a coloring in constant time in parallel.

2. The basic argument in the proof of Theorem 1.1 can be modified and extended to the real case as we sketch next without discussing the algorithmic issue. Here, too, the Local Lemma yields a sharper result but our argument is more constructive.

**Proposition 2.2** For every set S of at least  $4k^2$  reals, there is a k-coloring c of the set of all real numbers R so that every (real) translate of S intersects each color-class.

To prove this proposition first choose a real number t, so that all members of S are distinct modulo t. Let  $I_1, \ldots, I_k$  be a set of k pairwise disjoint intervals of equal lengths that partition [0, t), defined by  $I_i = [(i - 1)t/k, it/k)$ . Next, let a be chosen randomly and uniformly in the interval (0, M), where M is a large number, to be determined later, and let b be chosen randomly and uniformly in (0, t). Define a coloring c of the real numbers as follows: For a real number y, the color c(y) is the unique i so that  $(ay + b) \pmod{t} \in I_i$ . One can immitate the proof of Theorem 1.1 and show that if S is a set of at least  $4k^2$  reals, and M is sufficiently large, then the probability that c maps S into a set that intersects every cyclic interval of length at least t/k in [0, t) is  $\Omega(1)$ . Such a coloring c will have the desired properties and the assertion of the proposition follows. We omit the details.

3. One can use a simple algebraic idea to extend the first part of Theorem 1.1 even further, to the case of an arbitrary torsion-free Abelian group (such as, e.g., any Euclidean space  $\mathbb{R}^d$ ). Here, too, the Local Lemma yields a sharper result but our argument is more constructive. **Proposition 2.3** Let G be a torsion-free Abelian group. Then for every subset S of G containing at least  $4k^2$  elements there is a k-coloring of G so that every translate of S in G intersects each color class.

**Proof** The set S spans a finitely generated and hence a free abelian subgroup F of G. Select a homomorphism  $\phi : F \mapsto Z$  such that  $\phi$  restricted to S is injective (this is always possible). If  $\chi : Z \mapsto \{1, \ldots, k\}$  is the required coloring of Z with respect to  $\phi(S)$ , then  $\chi \phi$  is the required coloring of F with respect to S. To color G, color each F-coset separately by a translate of the coloring of F.  $\Box$ 

Studying analogues of Strauss' problem mentioned in the introduction for various actions of groups on sets seems to be an interesting program. It is worth mentioning that there is a whole area known as Geometric, or Euclidean Ramsey Theory (see, e.g., [9] for a few examples), which studies questions precisely opposite to Strauss' problem.

## **3** $\Theta(k \log k)$ colors are necessary (and sufficient)

In this section we present two proofs of Theorem 1.2. Let g(k) denote the smallest m so that for every set S of at least m integers there is a good k-coloring for the shift hypergraph H(S). The proof in [8] gives that for large k:

$$g(k) \le (3 + o(1))k \log_e k$$

The next proposition shows that this is sharp, up to a constant factor.

**Proposition 3.1** If q is a prime, and  $q > l^2 2^{2l-2}$  then

$$g(\lceil \frac{2q}{l+1} \rceil) > \frac{q+1}{2}.$$

This implies that for large k:

$$g(k) \ge (\frac{1}{8} + o(1))k \log_2 k.$$

The proof of the above proposition is by a construction that uses the properties of the quadratic residues and non-residues in the field  $Z_q$ . Recall that there are precisely (q+1)/2 quadratic residues modulo q. We need the following simple consequence of the well known theorem of Weil. For a derivation of this lemma from Weil's theorem, see [10] or [6].

**Lemma 3.2** Let  $q > l^2 2^{2l-2}$  be a prime and let  $Z = \{z_1, \ldots, z_s\}$  be a set of  $s \leq l$  members of  $Z_q$ . Then there exists an element  $y \in Z_q$  so that  $z_i - y$  is a quadratic non-residue for all  $1 \leq i \leq s$ .

**Proof of Proposition 3.1** Put m = (q + 1)/2 and let S be the set of all m quadratic residues modulo q, considered here as usual integers. Suppose  $k \ge \frac{2q}{l+1}$  and let  $c : Z \mapsto K = \{1, \ldots, k\}$ be a k-coloring of the integers. To complete the proof we show that there exists a translate of S that misses at least one color class. To this end, consider the colors of the integers in the set  $Q = \{0, 1, \ldots, 2q - 2\}$  and let  $j \in K$  be a color assigned to at most |Q|/k < l + 1 members of this set. Let  $y_1, \ldots, y_s$   $(s \le l)$  be all the members of Q satisfying  $c(y_i) = j$ . Let  $z_i$  be the elements of  $Z_q$  defined by  $z_i \equiv y_i (mod q)$ . By Lemma 3.2 there exists a  $y \in Z_q$  so that  $z_i - y$  is a quadratic non-residue for all  $1 \le i \le s$ . Consider, now, y as a usual integer. We claim that the translate y + S of S does not intersect color class number j. To see this, suppose it is false. Since  $y + S \subset Q$ this means that there is an  $i, 1 \le i \le s$ , and there is an  $x \in S$  so that  $y + x = y_i$ . Reducing this equation modulo q we conclude that  $x = (z_i - y)(mod q)$ . But this is impossible since  $z_i - y$  is a quadratic non-residue whereas x (like all the members of S) is a quadratic residue. Thus the claim holds, and the assertion of the proposition follows.  $\Box$ 

Proposition 3.1 implies Theorem 1.2. We next present another, probabilistic proof of this theorem, that gives a slightly better lower bound for g(k).

**Proposition 3.3** For large k

$$g(k) \ge (1 + o(1))k \log_e k.$$

The main part of the proof is the following somewhat technical lemma.

**Lemma 3.4** For every fixed (small)  $\epsilon > 0$  there exists a (small)  $\delta > 0$  such that for every sufficiently large n the following holds; There exists a subset S of  $N = \{1, 2, ..., n\}$  of cardinality at least  $(1 - \frac{\epsilon}{10})\delta n$  so that for every set T of at most

$$\frac{(1-\frac{\epsilon}{10})\log_e n}{(1+\frac{\epsilon}{10})\delta}$$

positive integers, each at most  $(1 + \frac{\epsilon}{10})n$ , there is an integer  $0 \le y \le \frac{\epsilon}{10}n$  so that y + S does not intersect T.

**Proof** Let  $\delta > 0$  satisfy

$$1 - \delta > e^{-(1 + \frac{\epsilon}{10})\delta}.$$
(2)

(Since

$$e^{-(1+\frac{\epsilon}{10})\delta} = 1 - (1+\frac{\epsilon}{10})\delta + O(\delta^2)$$

any sufficiently small  $\delta > 0$  satisfies (2)). Let *n* be sufficiently large (as a function of  $\epsilon$  and  $\delta$ ), and let *S* be a random subset of  $N = \{1, ..., n\}$  obtained by choosing each  $i \in N$ , randomly and independently, to be a member of *S* with probability  $\delta$ . Denote the cardinality of *S* by *m*. By the standard estimates for Binomial distributions (see, e.g., [4]), for a sufficiently large *n*, with high probability

$$m \ge (1 - \frac{\epsilon}{10})\delta n.$$

Fix a set T of at most

$$\frac{(1-\frac{\epsilon}{10})\log_e n}{(1+\frac{\epsilon}{10})\delta}$$

positive integers, each at most  $(1 + \frac{\epsilon}{10})n$ , and fix an integer  $y \leq \frac{\epsilon}{10}n$ . The probability that y + S does not intersect T is the probability that t - y is not in S for all  $t \in T$ , which is, by (2), at least

$$(1-\delta)^{|T|} > e^{-(1+\frac{\epsilon}{10})\delta\frac{(1-\frac{\epsilon}{10})\log_e n}{(1+\frac{\epsilon}{10})\delta}} = \frac{1}{n^{1-\frac{\epsilon}{10}}}$$

For the above fixed set T consider now all the possible shifts of S by an integer y satisfying  $0 \le y \le \frac{\epsilon}{10}n$ . For each such y we have the estimate above for the probability of the event  $E_y$  that y + S does not intersect T. Moreover, if Y is a set of possible shifts and for every two distinct y and y' in Y, T - y does not intersect T - y' the events  $E_y$   $(y \in Y)$ , are mutually independent. It is easy to see that there is such a set Y of cardinality at least  $\frac{\epsilon}{10}n/|T|^2 = \Omega(n/(\log n)^2)$ , where here the constant in the  $\Omega(\cdot)$  notation depends on  $\epsilon$  and  $\delta$  but not on n. It follows that the probability that there is no shift y in the possible range so that y + S does not intersect T is at most

$$(1 - \frac{1}{n^{1 - \frac{\epsilon}{10}}})^{\Omega(n/(\log n)^2)} = e^{-\Omega(n^{\frac{\epsilon}{10}}/(\log n)^2)}.$$

The total number of choices for a subset T as above is only

$$\sum_{\substack{i \le \frac{(1-\frac{\epsilon}{10})\log_e n}{(1+\frac{\epsilon}{10})\delta}}} \binom{\left(1+\frac{\epsilon}{10}\right)n}{i} \le e^{O((\log n)^2)}.$$

Therefore, the probability that there is a set T so that there is no shift y + S of S that misses it is at most

$$e^{O((\log n)^2)}e^{-\Omega(n\frac{\epsilon}{10}/(\log n)^2)}$$

which tends to 0 as n tends to infinity. This completes the proof of Lemma 3.4.  $\Box$ 

**Proof of Proposition 3.3** Let  $\epsilon$  be a fixed small positive constant. Our objective is to show that for all sufficiently large k,  $g(k) \ge (1 - \epsilon)k \log_e k$ . Let  $\delta$ , n, and S satisfy the assertion of Lemma 3.4. We assume, whenever it is needed, that  $\epsilon$  is sufficiently small and that n is sufficiently large. Put m = |S|, then

$$n \ge m \ge (1 - \frac{\epsilon}{10})\delta n$$

Let k be an integer and let  $c : Z \mapsto K = \{1, 2, ..., k\}$  be a good k-coloring for the hypergraph H(S). Clearly  $k \leq m \ (\leq n)$ . We claim that

$$k < \frac{(1+\frac{\epsilon}{10})^2 \delta n}{(1-\frac{\epsilon}{10}) \log_e n} \le \frac{(1+\epsilon)m}{\log_e n} \le \frac{(1+\epsilon)m}{\log_e k},\tag{3}$$

and hence that

$$m \ge (1 - \epsilon)k \log_e k.$$

Since  $\epsilon > 0$  is arbitrarily small (and for each such  $\epsilon$  any sufficiently large n can be chosen) this, together with the obvious monotonicity of the function g(k), imply the validity of Proposition 3.3. It thus remains to prove the claim. Let Q be the set of all positive integers which do not exceed  $(1 + \frac{\epsilon}{10})n$ . Fix a color  $i \in K$  and let T be the set of all members of Q colored i. If

$$|T| \le \frac{(1 - \frac{\epsilon}{10})\log_e n}{(1 + \frac{\epsilon}{10})\delta}$$

then, since S satisfies the assertion of Lemma 3.4, there is a translate y + S of S contained in Q which misses T, contradicting the assumption that c is a good k-coloring for H(S). Therefore, each of the k colors appears more than

$$\frac{\left(1 - \frac{\epsilon}{10}\right)\log_e n}{\left(1 + \frac{\epsilon}{10}\right)\delta}$$

times in Q and hence

$$(1+\frac{\epsilon}{10})n \ge |Q| > k\frac{(1-\frac{\epsilon}{10})\log_e n}{(1+\frac{\epsilon}{10})\delta}.$$

This implies (3) and completes the proof.  $\Box$ 

## 4 Concluding remarks

1. A sum-free set of integers is a set that contains no (not necessarily distinct) a, b and c so that a + b = c. An old result of Erdős [7] asserts that every set of n nonzero integers contains

a sum-free subset of cardinality at least n/3. A very simple proof for this result (and some extensions of it) is given in [2], where the problem of obtaining a polynomial time deterministic algorithm for finding, for a given set S of n non-zero integers, a sum-free subset of at least n/3 of them, is raised. Our technique here supplies a very simple algorithm (which is also parallelizable), as follows. Given S, find a polynomially small prime p = 3r + 2 so that all the n members of S are distinct modulo p, and are all non-zero modulo p. Next check for every nonzero  $a \in Z_p$  the number of members s of S so that  $as(mod \ p)$  lies in the interval  $r+1, \ldots, 2r+1$ . An easy expectation argument shows that there is an a for which the number of these members is at least  $\frac{r+1}{3r+2}n > n/3$ , and it is easy to see that they form a sum-free subset of S.

2. As shown in Section 3, for large k

$$(1 + o(1))k \log_e k \le g(k) \le (3 + o(1))k \log_e k.$$

It would be interesting to find the correct constant in the expression for g(k).

Acknowledgement We would like to thank Uri Zwick for helpful comments that improved the statement of Theorem 1.1 and the presentation of its proof.

## References

- N. Alon, A parallel algorithmic version of the local lemma, Random Structures and Algorithms 2 (1991), 367-378.
- [2] N. Alon and D. J. Kleitman, Sum-free subsets, in : "A Tribute to Paul Erdős" (A. Baker, B. Bollobás and A. Hajnal eds.), Cambridge University Press, Cambridge, England 1990, 13-26.
- [3] N. Alon and Y. Peres, Uniform dilations, Geometric and Functional Analysis, 2 (1992), 1-28.
- [4] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley, 1991.
- [5] J. Beck, An algorithmic approach to the Lovász Local Lemma, I, Random Structures and Algorithms 2 (1991), 343-365.

- [6] B. Bollobás, Random Graphs, Academic Press, 1985, pp. 318-319.
- [7] P. Erdős, Extremal problems in number theory, Proc. Symp. Pure Math., Vol. VIII, AMS (1965), 181-189.
- [8] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in: Infinite and Finite Sets (A. Hajnal et al., eds.), North-Holland, Amsterdam (1975), pp. 609-628.
- [9] R.L. Graham, B.L. Rothschild and J.H. Spencer, Ramsey Theory, Wiley, New York, 1980.
- [10] R. L. Graham and J. H. Spencer, A constructive solution to a tournament problem, Canad. Math. Bull. 14 (1971), 45-48.
- [11] R. M. Karp and V. Ramachandran, A survey of parallel algorithms for shared memory machines, in: Handbook of Theoretical Computer Science (J. Van Leeuwen Ed.), Chapter 17, MIT Press (1990), 871-941.
- [12] J. Spencer, Ten Lectures on the Probabilistic Method, SIAM, Philadelphia 1987, p. 60.