JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 9, Number 1, January 1996

DISCREPANCY IN ARITHMETIC PROGRESSIONS

JIŘÍ MATOUŠEK AND JOEL SPENCER

1. Results and history

Let $\mathcal{A}$ be a family of subsets of a finite set Ω. By a two-coloring of Ω we shall mean a map $\chi\colon \Omega \to \{-1,+1\}$. For any $X \subseteq \Omega$ we define $\chi(X) = \sum_{x \in X} \chi(x)$. The *discrepancy* of $\mathcal{A}$ is defined by

$$\operatorname{disc}(\mathcal{A}) = \min_{\chi} \max_{A \in \mathcal{A}} |\chi(A)|. \tag{1}$$

Let $\Omega = \{1, \ldots, n\}$, which we denote by $[n]$. Let $\mathcal{A}$ denote the set of arithmetic progressions on $[n]$. The discrepancy of this set system was investigated in 1964 by K. F. Roth [7]. If we define the function $ROTH(n) = \operatorname{disc}(\mathcal{A})$, his result can be written

$$ROTH(n) \geq cn^{1/4}$$

for c a positive absolute constant. That is, for any two-coloring χ of the first n integers there will be an arithmetic progression A on which the "imbalance" $|\chi(A)|$ is at least $cn^{1/4}$.

It is interesting that Roth himself did not believe his result to be best possible and speculated that perhaps $ROTH(n) = n^{1/2-o(1)}$. Indeed a bound $ROTH(n) = O(\sqrt{n \ln n})$ follows by elementary probabilistic considerations. In the early 1970s Sárközi (see [3]) showed $ROTH(n) \leq n^{1/3+o(1)}$. A breakthrough was given in 1981 by Beck [2] who showed $ROTH(n) \leq cn^{1/4} \ln^{5/2} n$. Here we show

Theorem 1.1. *$ROTH(n) \leq Cn^{1/4}$ with C an absolute constant.*

In words, we show the *existence* of a two-coloring χ of the first n integers so that all arithmetic progressions A have imbalance $|\chi(A)| \leq Cn^{1/4}$. We remark that the proof does not give a construction of χ in the usual sense and is indeed not satisfactory from an algorithmic point of view. The methods of §2 (see comments in [5]) are such that we have not been able to obtain an algorithm that would output this coloring χ in time polynomial in n. Our proof involves variants of the probabilistic method; we give [1] as a general reference. The technique of our proof combines methods of [2], [5], [6].

Throughout the paper, we'll use the symbols c, c', etc. generically for denoting absolute constants, and in order to limit the number of symbols, we reuse them freely.

Received by the editors February 18, 1994 and, in revised form, December 29, 1994.
1991 *Mathematics Subject Classification.* Primary 11B25, 11N37.
The first author was supported by Charles University grant No. 351 and Czech Republic Grant GAČR 201/93/2167. Part of this research was done during a visit to Princeton University supported by DIMACS.

2. Entropy

Let $A_1, \ldots, A_v \subseteq \Omega$. A *partial coloring* is a map $\chi\colon \Omega \to \{-1, 0, +1\}$. When $\chi(x) = 0$ we call x uncolored, otherwise x is called colored. We define, for $A \subseteq \Omega$, $\chi(A) = \sum_{x \in A} \chi(x)$. Our object will be to give a general condition under which there exists a partial coloring χ with the $|\chi(A_i)|$ "small" and "few" $x \in \Omega$ uncolored.

For any positive integer b define the *b-roundoff function* $R_b(x)$ as that i so that $2bi$ is the nearest multiple of $2b$ to x. In case of ties take the larger. Thus

$$\begin{aligned} R_b(x) = 0 &\quad \text{if and only if} \quad -b \le x < b, \\ R_b(x) \ge i &\quad \text{if and only if} \quad x \ge (2i-1)b, \\ R_b(x) \le -i &\quad \text{if and only if} \quad x < -(2i-1)b. \end{aligned} \tag{2}$$

Let X be any discretely valued random variable. We use the standard definition of the *entropy function* $H(X)$:

$$H(X) = \sum_i -p_i \log_2(p_i)$$

where $p_i = \Pr[X = i]$, the summation is over the possible values of X, and $0 \log_2 0$ is interpreted as 0. We shall use the following well-known facts about entropy:

- Entropy is subadditive, i.e., if $X = (X_1, \ldots, X_v)$, then $H(X) \le \sum_{i=1}^v H(X_i)$.
- When X takes on at most K values it has entropy at most $\log_2 K$, the extreme case being a uniformly chosen value from a K-set. Moreover $\sum_{i \in I} -p_i \log_2(p_i) \le \log_2 |I|$ for any subset of values of X.
- When X has entropy less than K it takes on some value with probability at least 2^{-K}.

Let S_n, as standard, denote the sum of n independent random variables, each uniform on $\{-1, +1\}$. When $\chi\colon \Omega \to \{-1, +1\}$ is uniform and $A \subseteq \Omega, |A| = n$, then $\chi(A)$ has distribution S_n. Now we come to a key definition:

$$ENT(n, b) = H(R_b(S_n)).$$

With this definition we give our general criterion.

Lemma 2.1. *Let $S_1, \ldots, S_v \subseteq \Omega$ with $|\Omega| = n$ and $|S_i| = n_i$. Suppose b_i, ε and $\gamma \le \frac{1}{2}$ are such that*

$$\sum_{i=1}^v ENT(n_i, b_i) \le \varepsilon n$$

and

$$\sum_{j=0}^{\gamma n} \binom{n}{j} < 2^{n(1-\varepsilon)}. \tag{3}$$

Then there is a partial coloring χ of Ω with

$$|\chi(S_i)| \le b_i \quad \text{for all } i$$

and more than $2\gamma n$ points $x \in \Omega$ colored.

Proof. Consider the uniform probability space of all $\chi\colon \Omega \to \{-1, +1\}$ and define the random variable

$$L(\chi) = (R_{b_1}(\chi(S_1)), \ldots, R_{b_v}(\chi(S_v))).$$

By subadditivity of entropy

$$H(L) \le \sum_{i=1}^{v} H(R_{b_i}(\chi(S_i))) = \sum_{i=1}^{v} ENT(n_i, b_i) \le \varepsilon n.$$

Hence some value of L has probability at least $2^{-\varepsilon n}$ of being achieved. As all χ have probability 2^{-n} there is a set Γ of at least $2^{n(1-\varepsilon)}$ colorings χ so that if $\chi_1, \chi_2 \in \Gamma$, then $L(\chi_1) = L(\chi_2)$.

We naturally associate such colorings χ with points on the Hamming Cube $\{-1,+1\}^n$. (With $\Omega = \{1, \dots, n\}$ associate χ with $(\chi(1), \dots, \chi(n))$.) A theorem of Kleitman [4] (basically an isoperimetric inequality) states that any $\Gamma \subseteq \{-1,+1\}^n$ of size bigger than $\sum_{j=0}^{l} \binom{n}{l}$ with $l \le \frac{n}{2}$ contains two points at Hamming distance (i.e., the number of different coordinates) at least $2l$. (This is "best possible" as Γ may be the set of all sequences with at most l coordinates $+1$.) Thus there are $\chi_1, \chi_2 \in \Gamma$ at Hamming distance at most $2\gamma n$. Set

$$\chi(x) = \frac{\chi_1(x) - \chi_2(x)}{2} \quad \text{for all } x \in \Omega.$$

Then χ is a partial coloring. The number of colored points is precisely the Hamming distance which is at least $2\gamma n$. For each i the values $\chi_1(S_i)$, $\chi_2(S_i)$ have the same b_i-roundoff and therefore lie in a common open interval of length less than $2b$. Thus

$$|\chi(S_i)| = \left| \frac{\chi_1(S_i) - \chi_2(S_i)}{2} \right| < b_i$$

as desired. □

We note that (3) holds for, say, $\gamma = \frac{1}{4}$ and $\varepsilon = 0.2$; this value will suffice for our purposes. Also, we shall always use a bound on $|\chi(S)|$ dependent only on $|S|$. We'll use the lemma in the following simpler form.

Corollary 2.2. *Let $\mathcal{A}$ be a family of subsets of an n-set Ω consisting of at most $f(s)$ sets of size s. If $b(s)$ satisfies*

$$\sum_{s} f(s) ENT(s, b(s)) \le \frac{n}{5},$$

then there is a partial coloring χ with $|\chi(S)| \le b(|S|)$ for all $S \in \mathcal{A}$ and fewer than half the points of Ω uncolored.

In applying Corollary 2.2 we need upper bounds on $ENT(n, b)$. The correct parameterization is $b = \lambda\sqrt{n}$. Roughly S_n is like $\sqrt{n}N$ where N is standard normal so that $ENT(n, \lambda\sqrt{n})$ should be like $g(\lambda) = H(R_\lambda(N))$. Analysis gives that $g(\lambda) = \Theta(\lambda^2 e^{-\lambda^2/2})$ for λ large ($i = \pm 1$ giving the dominant terms) while $g(\lambda) = \Theta(\ln(\lambda^{-1}))$ as $\lambda \to 0$, the major contribution being $p_i = \Theta(\lambda^{-1})$ for $i = O(\lambda^{-1})$. The following results are somewhat weaker and certainly not best possible but have the advantage of holding for all n, λ.

Lemma 2.3. *There is an absolute constant c so that $ENT(n, \lambda\sqrt{n}) \le G(\lambda)$ where we define*

$$G(\lambda) = \begin{cases} ce^{-\lambda^2/9} & \text{if } \lambda \ge 10, \\ c & \text{if } 0.1 \le \lambda \le 10, \\ c\ln(\lambda^{-1}) & \text{if } \lambda < 0.1. \end{cases} \tag{4}$$

Proof (outline). We employ the universal bound

$$\Pr[S_n \geq \tau\sqrt{n}] \leq e^{-\tau^2/2}.$$

Set $g_i = \exp(-\lambda^2(2i-1)^2/8)$, $i \geq 1$, and $g_0 = 1 - 2\exp(-\lambda^2/8)$. From (2) $p_i, p_{-i} \leq g_i$ and $p_0 \geq g_0$. On $[0,1]$ the function $-x\log_2 x$ increases to $x = e^{-1}$ and then decreases. When $\lambda \geq 10$, $g_i \leq e^{-1}$ for all $i \geq 1$ and $g_0 \geq e^{-1}$ so

$$ENT(n, \lambda\sqrt{n}) \leq -g_0 \log_2 g_0 + 2\sum_{i=1}^{\infty} -g_i \log_2 g_i.$$

This is a continuous function of λ which is $O(\lambda^2 e^{-\lambda^2/8})$ or, giving ground, $O(e^{-\lambda^2/9})$. When $0.1 \leq \lambda \leq 10$ set $I = \{-100, \ldots, +100\}$. The contribution to $ENT(n,b)$ from $i \in I$ is at most $\log_2 |I| \leq 8$. For $i \notin I$ certainly $g_i < e^{-1}$ so

$$ENT(n, \lambda\sqrt{n}) \leq 8 + 2\sum_{i=101}^{\infty} -g_i \log_2 g_i \leq 9.$$

For $\lambda < 0.1$ set $I = \{i \colon |i| < \lambda^{-20}\}$. Again for $i \notin I$ we have $g_i \leq e^{-1}$ and

$$ENT(n, \lambda\sqrt{n}) \leq \log_2(2\lambda^{-20} + 1) + 2\sum_{|i| > \lambda^{-20}} -g_i \log_2 g_i \leq 40\ln(\lambda^{-1})$$

by computation. $\square$

We may now further reexpress Corollary 2.2.

Corollary 2.4. *Let $\mathcal{A}$ be a family of subsets of an n-set Ω consisting of at most $f(s)$ sets of size s. If $b(s) = \sigma(s)\sqrt{s}$ where $\sigma(s)$ satisfies*

$$\sum_s f(s) G(\sigma(s)) \leq \frac{n}{5}, \tag{5}$$

then there is a partial coloring χ of Ω with $|\chi(A)| \leq b(|A|)$ for all $A \in \mathcal{A}$ and at least half the points of Ω colored. $\square$

With these bounds we can already give a result which is interesting in its own right and may give significant insight into the somewhat technical computations to come.

Theorem 2.5. *There is an absolute constant c so that the following holds for all n, s. If $A_1, \ldots, A_n \subseteq \Omega$, $|\Omega| = n$ and all $|A_i| \leq s$, then there is a partial coloring χ of Ω with less than half the points of Ω uncolored and with*

$$|\chi(A_i)| \leq c\sqrt{s}$$

for all $1 \leq i \leq n$.

Proof. From Lemma 2.3 we may pick c so that $G(c) \leq 0.2$. Now apply Corollary 2.4. $\square$

The monotonicity of G allows a further generalization of Corollary 2.4. Suppose $\mathcal{A}$ is a family of subsets of an n-set Ω which breaks into subfamilies consisting of at most $f(s)$ sets of size *at most* s. When (5) holds, the conclusion of Corollary 2.4 then holds. In particular, given any $A_1, \ldots, A_n \subseteq \{1, \ldots, n\}$ we have n sets of size at most n, and we may pick c so that $G(c) < 0.2$. Then there exists a partial coloring χ of $\{1, \ldots, n\}$ with all $|\chi(A_i)| \leq c\sqrt{n}$ and at least half the points colored. This result was the core of [5].

3. The first partial coloring

Let $\mathcal{A}$ denote the family of arithmetic progessions contained in $\Omega = \{1, \ldots, n\}$. Here we show:

Lemma 3.1. *There is a partial coloring of* Ω *so that* $|\chi(A)| \leq cn^{1/4}$ *for all* $A \in \mathcal{A}$ *and at least half the points of* Ω *are colored.*

3.1. **The decomposition.** Let $X = \{x_1, \ldots, x_l\}$ be any set of integers with $x_1 < \cdots < x_l$. Define $INT(X)$ to be the family of intervals—i.e., all sets $\{x_u \colon i \leq u \leq j\}$ where $1 \leq i \leq j \leq l$. Now define $CINT(X)$ (the *canonical intervals* on X) by taking, for all powers of two, $s = 2^i \leq l$, all sets $\{x_{(j-1)s+1}, \ldots, x_{js}\}$ with $js \leq l$. That is, we split X into consecutive intervals of length $s = 2^i$, ignoring the "extra". The following observation is standard:

Lemma 3.2 (Decomposition lemma)**.** *Any* $A \in INT(X)$ *can be written as* $A = B\backslash C$ *with* $C \subset B$ *and with* B *and* C *both decomposable into disjoint unions of sets in* $CINT(X)$ *of different sizes.*

Proof. With $A = \{x_u \colon i \leq u \leq j\}$ set $B = \{x_u \colon 1 \leq u \leq j\}$ and $C = \{x_u \colon 1 \leq u \leq i-1\}$. Take the binary expansion $j = 2^{b_1} + 2^{b_2} + \cdots$, $b_1 > b_2 > \cdots$, of j. Decompose B into the first 2^{b_1} elements of X union the next 2^{b_2} elements of $X, \ldots$, and do likewise with C. □

We can think of any arithmetic progression as a subinterval of an entire residue class so that

$$\mathcal{A} = \bigcup_{1\leq d\leq n} \bigcup_{0\leq i<d} INT[\{x \in [n] \colon x \equiv i \mod d\}].$$

We define the "canonical arithmetic progressions"

$$\mathcal{C} = \mathcal{C}_n = \bigcup_{1\leq d\leq n} \bigcup_{0\leq i<d} CINT[\{x \in [n] \colon x \equiv i \mod d\}]. \tag{6}$$

Lemma 3.3. *If* χ *is a partial coloring of* $[n]$ *so that*

$$\chi(X) \leq b(|X|)$$

for all $X \in \mathcal{C}$*, then*

$$\chi(A) \leq 2 \sum_{s; s=2^i \leq n} b(s) \tag{7}$$

for all $X \in \mathcal{A}$*.* □

3.2. **The coloring.** For $s = 2^i \leq n$ how many s-sets are in $\mathcal{C}_n$? We restrict $1 \leq d \leq \frac{n-1}{s-1}$ (otherwise the residue classes have fewer than s elements) and for each d the s-sets are disjoint so there are at most $\frac{n}{s}$ of them, giving an upper bound of $\frac{n(n-1)}{s(s-1)}$ of them. For $s = 1$ there are only n distinct singletons. Ignoring asymptotically insignificant terms we'll say that $\mathcal{C}_n$ has at most $n^2 s^{-2}$ sets of size s.

Remark. For $s \sim \sqrt{n}$ we have $\sim n$ sets of size s and Corollary 2.4 gives a partial coloring with $|\chi(A)| \leq cn^{1/4}$ for all such sets. We need to simultaneously color the larger and smaller sets. To avoid a logarithmic term in applying (7) we'll need a slightly better bound on $|\chi(A)|$ when $|A|$ is not near $\sqrt{n}$.

We parameterize $s = \tau\sqrt{n}$ so that we have $n\tau^{-2}$ sets of size s. We'll assume for convenience that $\sqrt{n}$ is a power of two so that $\tau = 2^i$, i integral. We set

$$b(\tau\sqrt{n}) = \sqrt{\tau\sqrt{n}}\sigma(\tau\sqrt{n}) \quad \text{where } \sigma(\tau\sqrt{n}) = \begin{cases} c'\tau^{-1} & \text{if } \tau \geq 1, \\ c'\tau^{-0.1} & \text{if } \tau < 1. \end{cases}$$

We claim that, for an appropriately large constant c', (5) is now satisfied. We need to show

$$\sum_{\tau \geq 1} \tau^{-2} G(c'\tau^{-1}) + \sum_{\tau < 1} \tau^{-2} G(c'\tau^{-0.1}) < \tfrac{1}{5} \tag{8}$$

where τ in the sums runs over integral powers of 2 and G is given by (4). We will insist that $c' \geq 1$ so that $G(c'y) \leq G(y)$. Both $\tau^{-2}G(\tau^{-1}) = O(\tau^{-2}\ln(\tau))$ (τ large) and $\tau^{-2}G(\tau^{-0.1}) = O(\tau^{-2}\exp(-\tau^{-0.2}/9))$ (τ small) give convergent sums so we find an absolute constant T for which

$$\sum_{\tau \geq T} \tau^{-2} G(\tau^{-1}) + \sum_{\tau < T^{-1}} \tau^{-2} G(\tau^{-0.1}) < 0.1.$$

As $\lim_{x\to\infty} G(x) = 0$ we may now select $c' \geq 1$ sufficiently large so that the finite sum

$$\sum_{T^{-1} < \tau < T} \tau^{-2} G(c\tau^{-0.1}) < 0.1,$$

yielding (8). Hence by Lemma 3.3 and Corollary 2.4 there is a partial coloring of $[n]$ with at least half of the points colored and with

$$|\chi(A)| \leq 2\sum_{\tau} b(\tau\sqrt{n}) \leq 2c'n^{1/4}\left[\sum_{\tau\geq 1}\tau^{-1/2} + \sum_{\tau<1}\tau^{0.4}\right] \tag{9}$$

for all $A \in \mathcal{A}$. As the bracketed sums both converge this gives Lemma 3.1.

4. Number theory

Let $X \subseteq \{1, \ldots, n\}$ be an m-element set. Let s be an integer, $1 \leq s \leq n$. For an integer d, let $U(d)$ denote the set of all $x \in X$ in residue classes modulo d for which at least s elements of X lie in that residue class. We are interested in the quantity

$$U = \sum_{d} |U(d)|.$$

We can clearly restrict ourselves to the range $1 \leq d \leq n/s$ (for larger $d, U(d) = \emptyset$). Also, for each d, $|U(d)| \leq m$, and thus we get $U \leq nm/s$. This is tight for $s = 1$ but, for large enough s, the following theorem gives an improvement. The intuition behind it is that while for some individual value of d, the members of X can be distributed among very few residue classes modulo d only, such a distribution cannot occur for too many values of d at once.

Set $\rho = m/n$. We have

Proposition 4.1. *Suppose that $5\sqrt{m} \leq s \leq m$. Then*

$$U \leq c\frac{nm}{s}\sqrt{\rho}$$

for an absolute constant c.

Lemma 4.2. *For any pair d, d' of distinct natural numbers, we have*

$$|U(d) \cap U(d')| \leq \frac{|U(d)| \cdot |U(d')|}{s^2} \left\lceil \frac{n}{\operatorname{lcm}(d, d')} \right\rceil .$$

Proof. A number $x \in U(d) \cap U(d')$ can be specified by giving the number $r = \lfloor x/\operatorname{lcm}(dd') \rfloor$ plus the residue classes of x modulo d and modulo d', by the Chinese Remainder Theorem. The number r can be chosen in at most $\lceil n/\operatorname{lcm}(d, d') \rceil$ ways, and we note that $U(d)$ may intersect at most $|U(d)|/s$ residue classes modulo d, and similarly for $U(d')$. □

Lemma 4.3. *Let $I \subseteq \{1, 2, \ldots, n\}$ be a set such that $d \geq d_0$ for all $d \in I$, and $\gcd(d, d') \leq M$ for all distinct $d, d' \in I$. Suppose that $d_0^2/n \leq M \leq s^2 d_0^2/(9mn)$. Then*

$$\sum_{d \in I} |U(d)| \leq 2m.$$

Proof. If not, add indices to I one by one until the sum first gets over $2m$. Stopping then would give a set I with the same assumptions where $x = \sum_{d \in I} |U(d)|$ satisfies $2m < x \leq 3m$. We use Inclusion-Exclusion:

$$m \geq \left| \bigcup_{d \in I} U(d) \right| \geq \sum_{d \in I} |U(d)| - \sum_{d, d' \in I, d < d'} |U(d) \cap U(d')|. \tag{10}$$

By Lemma 4.2 and by the assumptions on I, we have

$$\begin{aligned} \sum_{d < d'} |U(d) \cap U(d')| &\leq \sum_{d < d'} \frac{|U(d)| \cdot |U(d')|}{s^2} \left\lceil \frac{nM}{d_0^2} \right\rceil \\ &\leq \frac{1}{2s^2} \left(\sum_{d \in I} |U(d)| \right)^2 \left(\frac{nM}{d_0^2} + 1 \right). \end{aligned}$$

The assumption on M implies $nM/d_0^2 \geq 1$. Thus, from (10), we further get

$$m \geq x - \frac{x^2}{2s^2} \frac{2nM}{d_0^2} > 2m - \frac{9m^2 nM}{s^2 d_0^2} \geq 2m - m = m$$

(using the upper bound on M in the assumption of the lemma), a contradiction. □

Proof of Proposition 4.1. We may suppose that m, n, s, ρ^{-1} are all sufficiently large (otherwise the claim is satisfied trivially). We fix a parameter $\varepsilon = 5\sqrt{\rho}$. We let J be the interval

$$J = \left[\varepsilon \frac{n}{s}, \frac{n}{s} \right]$$

(we may also suppose that $\varepsilon mn/s$ is an integer). We note that the d lying outside the interval J only contribute at most $\varepsilon mn/s = 5(nm/s)\sqrt{\rho}$ to U. Hence it suffices to bound $\sum_{d \in J} |U(d)|$.

We want to partition the interval J into consecutive intervals $I_1, I_2, \ldots, I_k$, in such a way that Lemma 4.3 can be applied to each of them, giving the bound $\sum_{d \in I_i} |U(d)| \leq 2m$. It remains to calculate how small k can be made. If we denote $I_i = [d_i, d_{i+1})$, then we have $\gcd(d, d') \leq d_{i+1} - d_i$ for any two distinct numbers

$d, d' \in I_i$. Thus, in order to apply Lemma 4.3, it is enough to have

$$d_{i+1} - d_i \geq \frac{d_i^2}{n}, \tag{11}$$

$$d_{i+1} - d_i \leq \frac{s^2 d_i^2}{9mn}. \tag{12}$$

The upper bound (12) suggests we define the d_i's by the initial condition $d_1 = \varepsilon n/s$ and by the recurrence

$$d_{i+1} = d_i + \left\lfloor \frac{s^2 d_i^2}{9mn} \right\rfloor .$$

One may check that with our choice of parameters, $s^2 d_1^2/(9mn) \geq 2$, and therefore $d_{i+1} \geq d_i + s^2 d_i^2/(18mn)$. We need to check the validity of (11), but this follows by calculation from the assumption $s \geq 5\sqrt{m}$.

It remains to estimate the smallest k such that $d_{k+1} \geq n/s$. Set $\alpha = s^2/(18mn)$. Then $d_{i+1} \geq d_i(1 + \alpha d_i)$. Given i, let us estimate the number j of steps needed so that $d_{i+j} \geq 2d_i$. We have $d_{i+j} \geq d_i(1+\alpha d_i)^j \geq d_i(j\alpha d_i)$, so $j \geq 1/(2\alpha d_i)$ suffices for the doubling. Therefore, the first doubling (from d_1 to at least $2d_1$) needs

$$\frac{1}{2d_1\alpha} = \frac{9m}{s\varepsilon} = O\left(\frac{n}{s}\sqrt{\rho}\right)$$

steps. Then the successive doubling times decrease geometrically, until the ratio of two successive members of the sequence of the d_i's exceeds 2. The number of remaining steps needed for reaching n/s after this happens is at most $\log_2((n/s)/d_1) = \log_2(1/\varepsilon)$. Therefore $k = O((n/s)\sqrt{\rho} + \log(1/\rho)) = O((n/s)\sqrt{\rho})$, and $\sum_{d\in J} |U(d)| = O((mn/s)\sqrt{\rho})$ as claimed. □

Remark. The set $S = \{1, \ldots, m\}$ gives a value $U \sim \frac{m^2}{s} = \frac{nm}{s}\rho$. Finding the maximal value of $U = U(n, m, s)$ is an intriguing problem we do not pursue here but we conjecture that our Proposition 4.1 is *not* best possible.

5. The end of the hunt

Let $X \subseteq [n]$, $|X| = m = \rho n$ with $n^{-3/4} \leq \rho \leq 1$. Our object is to find a partial coloring χ of X so that $|\chi(X \cap A)|$ is small for all $A \in \mathcal{A}$ and at least half the points of X are colored. Once successful, we'll apply this process iteratively beginning with $X = [n]$ (which we did in §3), resetting X to be the uncolored points, until $|X| < n^{1/4}$ at which time the remaining points may be colored arbitrarily.

Following (6) set

$$\mathcal{C} = \mathcal{C}_X = \bigcup_{1\leq d\leq n} \bigcup_{0\leq i<d} CINT[\{x \in X : x \equiv i \mod d\}].$$

For any $A \in \mathcal{A}$ we may, as in §3.1, decompose $A \cap X = B \backslash C$ with $C \subset B$ and B, C both disjoint unions of sets of $\mathcal{C}_X$ of different cardinalities. Lemma 3.3 now generalizes.

Lemma 5.1. *If χ is a partial coloring of X so that $|\chi(Y)| \leq b(|Y|)$ for all $Y \in \mathcal{C}_X$, then*

$$|\chi(A \cap X)| \leq 2 \sum_{s=2^i} b(s)$$

for all $A \in \mathcal{A}$.

Let $f(m,s)$ denote the number of s-sets in $\mathcal{C}_X$ so that $f(m,s) \le s^{-1}U$ with U as in §4. We first apply the elementary bound

$$f(m,s) \le \frac{mn}{s^2}.$$

Lemma 5.2. *There is a partial coloring χ of X with*

$$|\chi(A \cap X)| \le cn^{1/4}$$

for all $A \in \mathcal{A}$ and with more than half the points of X colored.

Proof. We follow the proof of Lemma 3.1 precisely. For $s = \tau\sqrt{n}$ we set

$$b(\tau\sqrt{n}) = \sqrt{\tau\sqrt{n}}\sigma(\tau\sqrt{n}) \quad \text{with } \sigma(\tau\sqrt{n}) = \begin{cases} c'\tau^{-1} & \text{if } \tau \ge 1, \\ c'\tau^{-0.1} & \text{if } \tau < 1. \end{cases}$$

Again we need (8) and the bound (9) is the same. □

We iterate this result, beginning at $X = [n]$, resetting X to be the uncolored points at each iteration, stopping when $|X| < \rho_0 n$, with ρ_0 a sufficiently small (as determined later) absolute constant. This is a constant number of iterations (recall the number of uncolored points is at least halved at each iteration) so together we have a partial coloring χ with $|\chi(A)| \le cn^{1/4}$ for all $A \in \mathcal{A}$ and a set X of fewer than $\rho_0 n$ points uncolored.

Remark. Continuing this process until $|X| < n^{1/4}$ and then coloring the remaining points arbitrarily would give a full coloring with all $|\chi(A)| \le cn^{1/4} \ln n$. Our "slight" improvement of §4 will allow a slight improvement as X becomes smaller so that the sum converges to $O(n^{1/4})$.

Now fix X with $|X| = \rho n$, $n^{-3/4} \le \rho \le \rho_0$. We set $b(\tau\sqrt{n}) = \sqrt{\tau\sqrt{n}}\sigma(\tau\sqrt{n})$ with

$$\sigma(\tau\sqrt{n}) = \begin{cases} \tau^{-0.1} & \text{if } \tau < \rho^{1/5}, \\ \rho & \text{if } \rho^{1/5} \le \tau < 1, \\ \tau^{-1}\rho & \text{if } 1 \le \tau, \end{cases} \tag{13}$$

and set

$$f(\tau\sqrt{n}) = \begin{cases} m\tau^{-2} & \text{if } \tau < \rho^{1/5}, \\ cm\tau^{-2}\sqrt{\rho} & \text{if } \tau \ge \rho^{1/5}, \end{cases}$$

which, by a slight weakening of Proposition 4.1, is an upper bound on the number of $\tau\sqrt{n}$-sets in $\mathcal{C}_X$.

We first claim that for ρ appropriately small

$$m^{-1} \sum_{\tau} f(\tau\sqrt{n}) G(\sigma(\tau\sqrt{n})) \le 0.2 \tag{14}$$

(we recall the convention from §3—τ in summation runs through integral powers of 2). We split the sum by the ranges of (13). As functions of ρ

$$\sum_{\tau<\rho^{1/5}} \tau^{-2}G(\tau^{-0.1}) = O(\rho^{-2/5}e^{-\rho^{-1/30}}),$$

$$\sum_{\rho^{1/5}\le\tau<1} c\tau^{-2}\sqrt{\rho}G(\rho) = O(\rho^{1/2-2/5}\ln^2(\rho^{-1})),$$

$$\sum_{1\le\tau} c\tau^{-2}\sqrt{\rho}G(\tau^{-1}\rho) \le c' \sum_{1\le\tau}\tau^{-2}\sqrt{\rho}(\ln\tau+\ln(\rho^{-1})) = O(\sqrt{\rho}\ln(\rho^{-1}))$$

which are all $o(\rho)$ so that (14) holds when ρ_0 is picked sufficiently small.

We bound

$$2n^{-1/4}\sum_{\tau} b(\tau\sqrt{n}) \le \sum_{\tau<\rho^{1/5}}\tau^{0.4} + \sum_{\rho^{1/5}\le\tau<1}\tau^{1/2}\rho + \sum_{1\le\tau}\tau^{-1/2}\rho.$$

The first sum dominates and this is $O(\rho^{4/25})$ as $\rho\to 0$. We have shown:

Lemma 5.3. *There are absolute positive constants ρ_0, c so that if $|X|=\rho n$, $\rho<\rho_0$, then there exists a partial coloring χ so that $|\chi(A\cap X)|\le cn^{1/4}\rho^{4/25}$ for all $A\in\mathcal{A}$ and with at least half the points of X colored.*

The exponent $\frac{4}{25}$ clearly could be improved by more careful calculation but it does not matter. We are done. Begin with $X=[n]$. Apply Lemma 3.1 and then Lemma 5.1 until $|X|<\rho_0 n$, then apply Lemma 5.3 until $|X|<n^{1/4}$ and then color the remaining points arbitrarily. The final coloring χ has

$$|\chi(A)| \le cn^{1/4} + \sum_{i=0}^{\infty} c'n^{1/4}(\rho_0 2^{-i})^{4/25} + n^{1/4} \le c^* n^{1/4}$$

for all $A\in\mathcal{A}$ and has *no* points uncolored.

References

1. N. Alon and J. Spencer, *The probabilistic method*, Wiley, New York, 1992. MR **93h:**60002
2. J. Beck, *Roth's estimate on the discrepancy of integer sequences is nearly sharp*, Combinatorica **1** (1981), 319–325. MR **83i:**05040
3. P. Erdős and J. Spencer, *Probabilistic methods in combinatorics*, Academic Press, New York, 1974. MR **52:**2895
4. D. Kleitman, *On a combinatorial problem of Erdős*, J. Combin. Theory **1** (1966), 209–214. MR **34:**78
5. J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. **289** (1985), 679–706. MR **86k:**05004
6. J. Matoušek, *Tight upper bounds for the discrepancy of halfspaces*, Discrete Comput. Geom. (to appear).
7. K. F. Roth, *Remark concerning integer sequences*, Acta Arith. **9** (1964), 257–260. MR **29:**5806

Abstract. It is proven that there is a two-coloring of the first n integers for which all arithmetic progressions have discrepancy less than const.$n^{1/4}$. This shows that a 1964 result of K. F. Roth is, up to constants, best possible.

Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic

E-mail address: matousek@kam.mff.cuni.cz

Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, New York 10012

E-mail address: spencer@cs.nyu.edu