# Tight bounds on additive Ramsey-type numbers

By

Karol Cwalina* and Tomasz Schoen†

**Abstract**

A classical result of Rado states that for every homogenous regular equation with integer coefficients there is the least natural number $\mathsf{R}(n)$ such that if the elements of $[N] = \{1, \ldots, n\}$ are colored into $n$ colors for $N \geqslant \mathsf{R}(n)$, then there is a monochromatic solution to the equation. While density results provide quite accurate bounds for $\mathsf{R}(n)$ in case of invariant equations, the general upper bounds known have been tower-like in nature. In this paper we prove that $\mathsf{R}(n) = \exp(O(n^4 \log^4 n))$ for all homogenous regular equations.

Also, we prove that the Schur-like numbers for the equation $x_1 + x_2 + x_3 = y_1 + y_2$ are at most $O(n^{-c\frac{\log n}{\log\log n}} n!)$, for some absolute constant $c > 0$. This beats a bound following classical Schur's argument. In the last section we establish a new upper bound for the van der Waerden numbers $\mathsf{W}(3, k)$.

## 1 Introduction

It has been a long studied question of when a diophantine equation has a solution in a set of integers and it has been known for long that not all equations are created equal to this respect. The main division line goes between those which can be called *invariant* and those which cannot. A great account of this is two-part Ruzsa's work [16, 17].

**Definition 1.1.** Given a linear equation of integer coefficients

$$a_1x_1 + \cdots + a_kx_k = b\,, \tag{1.1}$$

for $a_i \in \mathbb{Z}$ and $b \in \mathbb{Z}$, we say that:

1. it is *homogenous* if $b = 0$;
2. it is *invariant* if $b = \sum_{i=1}^{k} a_i = 0$;

*The first author is supported by the National PhD Programme in Mathematical Sciences at the University of Warsaw.

†The second author is supported by NCN grant 2012/07/B/ST1/03556

2010 *Mathematics Subject Classication*. 11B30 (primary), and 11D40, 11D72 (secondary).

3. it *contains an equation of genus g* if there are $g$ pairwise disjoint non-empty subsets $I_1, \ldots, I_g \subseteq \{1, \ldots, k\}$ such that $\sum_{i \in I_j} a_i = 0$ for every $j = 1, \ldots, g$.

Note that equations that have genus $g$ in accordance with Ruzsa's [16, Definition 3.5], do also contain an equation of genus $g$ in the sense of the definition above.

While existence of solutions to invariant equations can be guaranteed on density basis alone, it is no longer so for general non-invariant equations. A natural example is the set of odd numbers, which is free of solutions to Schur's equation $x+y=z$ despite of having density $\frac{1}{2}$. For this reason a study of non-invariant equations has to follow slightly different lines, designated long ago by Schur [22] and Rado [15].

**Definition 1.2.** We say that equation (1.1) is *regular* if in every finite coloring $\mathbb{N} = A_1 \cup \cdots \cup A_n$ there exists a monochromatic solution to it.

Rado proved [15] that a homogenous equation is regular if and only if it contains an equation of genus 1. Thus, in a sense, a homogenous regular equation must contain an invariant equation. In this case it follows by the compactness principle that there is the smallest integer $\mathsf{R}(n)$ such that for every $n$-coloring of $\{1, \ldots, \mathsf{R}(n)\}$ there is a monochromatic solution to (1.1).

Since known proofs of Rado's theorem rely on finding long monochromatic arithmetic progressions or similar arithmetic structures, the resulting upper bounds are rather poor. A straightforward application of the van der Waerden theorem would result in an Ackerman-type bound for the Rado numbers and application of the powerful result of Gowers [11, Theorem 18.2] cannot give anything better than roughly

$$\mathsf{R}(n) \leqslant \mathsf{tower}(5n)\,,$$

where

$$\mathsf{tower}(n) = 2^{2^{\cdot^{\cdot^{2}}}} \Big\} \; n \text{ times.}$$

As already mentioned, if (1.1) is invariant, then density results are highly related to Rado numbers. It follows from Behrend's construction [2] composed with a probabilistic covering argument on the one hand, and from Bloom's [3, Theorem 1.1] based on Sanders's work [20] on the other, that for every $k$-variable invariant equation we have

$$2^{O(\log^2 n)} \leqslant \mathsf{R}(n) \leqslant 2^{O(n^{1/(k-2)} \log^5 n)}\,.$$

If, additionally, we assume that $k \geqslant 6$ then from [21, Theorem 1.1] one can deduce that

$$\mathsf{R}(n) \leqslant 2^{O(\log^7 n)}\,.$$

Notice that Behrend's construction still provides the bound $\mathsf{R}(n) \gg 2^{O(\log^2 n)}$ for every *convex* equation

$$a_1 x_1 + \cdots + a_k x_k = (a_1 + \cdots + a_k) y\,,$$

$a_1, \ldots, a_k \in \mathbb{N}$. Hence for all convex equations with $k \geqslant 6$ we have quite tight bounds on Rado numbers. Furthermore, for equations with genus $g \geqslant 2$ we have

$$\mathsf{R}(n) = n^{1+O(1/g)}\,,$$

see Ruzsa's [16, Theorem 3.6].

On the other hand, if (1.1) is non-invariant, then every set of integers contains a subset proportional in size and free of solutions to this equation. Hence, by iterative argument,

$$\mathsf{R}(n) \gg C^n$$

for some $C > 1$ depending on the equation.

The above discussion shows that one of the most widely open questions concerning Rado numbers is that of upper bounds for non-invariant equations. The paper is devoted to these equations only and our main results are the following three theorems proved in Section 2. Many of the proofs presented share the same idea of identifying long monochromatic arithmetic progressions or, in the more involved cases, large Bohr sets. The following results are presented in the order of increasing strength of hypothesis. The more structural the equation considered, the more efficient our methods will be. All implied constants depend only on the equation considered.

**Theorem 1.3.** *Let $a_1x_1 + \cdots + a_kx_k = 0$ be any regular equation with integer coefficients. Then for every $n$*

$$\mathsf{R}(n) \ll 2^{O(n^4 \log^4 n)}.$$

**Theorem 1.4.** *Let $a_1x_1 + \ldots + a_kx_k = 0$ be an equation with integer coefficients that contains an invariant equation with at least 4 variables. Then for every $n$*

$$\mathsf{R}(n) \leqslant 2^{O(n^3 \log^5 n)}.$$

**Theorem 1.5.** *Let $a_1x_1 + \ldots + a_kx_k = 0$ be an equation with integer coefficients that contains an equation of genus 2, then*

$$\mathsf{R}(n) \leqslant 2^{O(n^2 \log^5 n)}.$$

While the above results make a significant progress when compared with tower-like bounds, the gap between lower and upper bounds is still wide. We will keep this issue in mind in Section 3 when a particularly simple class of Schur-like equations will be considered.

A classical theorem of Schur [22], prior to general Rado's result [15], asserts that for every partitioning of the first $\lfloor en! \rfloor$ positive integers into $n$ classes one can always find three numbers in one partition class satisfying the equation $x + y = z$. In other words, a certain class is not sum-free.

Denote by $\mathsf{S}(n)$ the smallest integer $N$ such that for every $n$-coloring of $\{1, \ldots, N\}$ there is a monochromatic solution to $x + y = z$. We know that

$$321^{n/5} \ll \mathsf{S}(n) \leqslant \lfloor (e - 1/24)n! \rfloor .$$

For the lower bound see [8]. The upper one stems from the relation $\mathsf{S}(n) < \mathsf{R}(3, \ldots, 3; 2)$ between Schur and Ramsey numbers, from the classical recurrence relation

$$\mathsf{R}(k_1, \ldots, k_n; 2) \leqslant 2 - n + \sum_{i=1}^{n} \mathsf{R}(\ldots, k_{i-1}, k_i - 1, k_{i+1}, \ldots ; 2)$$

and the bound $\mathsf{R}(3,3,3,3;2) \leqslant 65$ proved in [24]. Abbott and Moser [1] proved that $\lim_{n\to\infty} \mathsf{S}(n)^{1/n}$, although not necessarily finite, does exist. Proving any significantly stronger bound on $\mathsf{S}(n)$ would be highly appreciated but we will only manage to improve it a little bit in a special case only.

We call a set $A \subseteq \mathbb{Z}$ *$k$-sum-free* if it contains no solution to the equation

$$x_1 + \cdots + x_{k+1} = y_1 + \cdots + y_k\,. \tag{1.2}$$

Also we denote by $\mathsf{S}_k(n)$ the analogues of the Schur numbers for equation (1.2). Because every $(k+1)$-sum-free set is also $k$-sum-free, we have

$$\mathsf{S}(n) = \mathsf{S}_1(n) \geqslant \mathsf{S}_2(n) \geqslant \ldots.$$

It is easy to check that for every $k$ we still have $\mathsf{S}_k(n) > C_k^n$, for some constant $C_k > 1$. While such equation is easier to handle for larger $k$, because of the large number of summands involved, a straightforward application of Schur's argument only gives $\mathsf{S}_k(n) \ll \frac{1}{k}n!$. Our result is the following.

**Theorem 1.6.** *For some absolute positive constant $c$, we have*

$$\mathsf{S}_2(n) \ll n^{-c\frac{\log n}{\log\log n}} n!\,.$$

Finally, in Section 4, we shall improve a bound for van der Waerden numbers $\mathsf{W}(3,k)$.

The sections that are about to follow can be read independently.

## Notation

All sets considered in the paper are finite subsets of $\mathbb{Z}$ or $\mathbb{Z}/N\mathbb{Z}$. For sets $A, B$ we write

$$A + B = \{a + b : a \in A, b \in B\},$$

and $a{\cdot}A = \{ax : x \in A\}$. On the other hand, iterated sumsets are defined as $mA = \overbrace{A + \ldots + A}^{m}$, and we shall consistently distinguish between these two notations.

We denote by $\mu_X$ the uniform measure on a nonempty set $X$ and we write $A(x)$ for the indicator function of the set $A$. For $f, g : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ the convolution of $f$ and $g$ is defined by

$$(f * g)(x) = \sum_{t\in\mathbb{Z}/N\mathbb{Z}} f(t)g(x-t).$$

The Fourier coefficients of a function $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ are defined by

$$\widehat{f}(r) = \sum_{x\in\mathbb{Z}/N\mathbb{Z}} f(x)e^{-2\pi i x r/N},$$

where $r \in \mathbb{Z}/N\mathbb{Z}$, and the above applies to the indicator function of a set $A \subseteq \mathbb{Z}/N\mathbb{Z}$, which we will write $A$, as well. We set

$$\operatorname{Spec}_\eta(A) = \{r \in \mathbb{Z}/N\mathbb{Z} : |\widehat{A}(r)| \geqslant \eta|A|\}.$$

Parseval's formula states that $\sum_{r=0}^{N-1} |\widehat{A}(s)|^2 = |A|N$. We will also use the fact that $\widehat{(A * B)}(r) = \widehat{A}(r)\widehat{B}(r)$.

## 2 Rado numbers

This section consists of two parts. In the first one we describe the main ideas of our approach to upper bounds for Rado numbers and compare it with a traditional one, based on identifying long monochromatic arithmetic progressions in structured sets. Since sumsets, nor even sets of the form $2A-2A$ do not have to contain sufficiently long progressions, the results obtained will be rather poor, but still much better than previous bounds. In the second part we will prove the main results of this paper. To this end we will heavily rely on properties of Bohr sets.

### 2.1 Sketch of the argument

To prove our results we try to adopt classical Schur's method, which is originally designed to prove upper bounds on partitions free of solutions to the equation $x+y=z$ and can be described as follows. Suppose that $X_0=[N]=A_1\cup\cdots\cup A_n$ is a sum-free partition. Then, iteratively, for $X_{k-1}\subseteq A_k\cup\ldots\cup A_n$ and $X_{k-1}-X_{k-1}$ disjoint with $A_1\cup\ldots\cup A_{k-1}$ we may assume that $X_{k-1}\cap A_k$ is the largest among $X_{k-1}\cap A_k,\ldots,X_{k-1}\cap A_n$. Let $a=\max X_{k-1}\cap A_k$. Clearly $X_k=a-(X_{k-1}\cap A_k\setminus\{a\})$ satisfies the conditions imposed. Iterating this process, after $n$ steps we find a set $X_n$ of size roughly$N/n^n$ such that $(X_n-X_n)\cap[N]=\emptyset$. This results in the bound $N\ll n^n$.

Now, notice that it is enough in the general case to consider equations of the form

$$ax-ay+bz=0$$

with $a,b$ positive integers, because by Rado's theorem every regular equation can be reduced to such an equation. It is immediate that Schur's argument cannot be directly applied for the above equations. To make it work one can try to locate in $a{\cdot}A_1-a{\cdot}A_1$ a symmetric arithmetic progression disjoint with $b{\cdot}A_1$, and iterate this procedure. To express this idea we recall the following lemma.

**Lemma 2.1** ([7, Corollary 1]). *Let $A\subseteq\{1,\ldots,N\}$ with $|A|\geqslant\delta N$. Then there exist integers $d>0$ and $l\gg\frac{\log N}{\log(1/\delta)}$ such that $d,2d,\ldots,ld\in A-A$.*

**Theorem 2.2.** *For any regular homogenous equation $a_1x_1+\ldots+a_kx_k=0$ with integer coefficients we have*

$$\mathsf{R}(n)\leqslant\mathsf{tower}((1+o(1))n)\,,$$

*where the $o(1)$ term depends only on the equation.*

*Proof.* Let $[\mathsf{R}(n)-1]=A_1\cup\cdots\cup A_n$ be a partition without a monochromatic solution to our equation and set $N=R(n)-1$. Also, following Rado's characterization of regular equations, let $I$ be such that $\sum_{i\in I}a_i=0$. First, observe that there are no monochromatic solutions to the equation

$$ax-ay+bz=0\,,$$

where $a=|a_{i_0}|$ for some $i_0\in I$ and $b=|\sum_{i\notin I}a_i|$. Suppose that $|A_1\cap\{1,\ldots,N/a\}|\geqslant N/na$ and let $A\subseteq A_1$ be any set of elements of $A_1$ belonging to the same residue class modulo $b$ with

$|A| \geqslant |A_1|/b$. We apply Lemma 2.1 to $A$, so that $d, 2d, \ldots, ld \in A - A$ for some $l \gg \frac{\log N}{\log(abn)}$. Notice that $d \equiv 0 \pmod b$ and $ad/b, 2ad/b, \ldots, lad/b \notin A_1$, so that

$$ad/b, 2ad/b, \ldots, lad/b \in A_2 \cup \cdots \cup A_n\,.$$

Whence $\mathsf{R}(n-1) \gg \frac{\log N}{\log(abn)}$ or, equivalently,

$$\mathsf{R}(n) \leqslant (abn)^{O(\mathsf{R}(n-1))}$$

and the assertion follows. ∎

Next we show that Theorem 2.2 can be highly improved provided that the equation considered contains an invariant component of at least three variables, i.e. there exists $I$ such that $\sum_{i\in I} a_i = 0$ and $|I| \geqslant 3$. To this end we need some lemmas. The first one is a deep result due to Sanders, proved in [20, Theorem 1.1], see also [3]. The other can be easily extracted from the proof of [9, Theorem 3].

**Lemma 2.3.** *Let $A \subseteq \{1, \ldots, N\}$ with $|A| \geqslant \delta N$. Then $A$ contains $\exp(-O((1/\delta)\log^5(1/\delta))|A|^{k-1}$ solutions to any invariant equation with $k \geqslant 3$ variables.*

**Lemma 2.4.** *Let $A, B, C \subseteq \{1, \ldots, N\}$ with $|A|, |B|, |C| \geqslant \delta N$. Then every $x$ with at least $\varepsilon|A||B||C|/N$ representations in $A + B + C$ is a middle term of an arithmetic progression of length $\Omega(\varepsilon N^{\varepsilon^2\delta^3})$, fully contained in $A + B + C$.*

Having introduced the lemmas we can prove the following.

**Theorem 2.5.** *Let $a_1x_1 + \ldots + a_kx_k = 0$ be an equation of integer coefficients and $I \subseteq \{1, \ldots, k\}$ be such that $\sum_{i\in I} a_i = 0$. Suppose that $|I| \geqslant 3$, then*

$$\mathsf{R}(n) \leqslant 2^{2^{O(n^2\log^6 n)}}\,. \tag{2.1}$$

*The implied constant depends only on the equation.*

*Proof.* Let $M = \sum_{i\in I} |a_i|/2$ and let $[\mathsf{R}(n) - 1] = A_1 \cup \cdots \cup A_n$ be a partition without a monochromatic solution to our equation. Set $N = \mathsf{R}(n) - 1$, suppose that $|A_1 \cap \{1, \ldots, N/M\}| \geqslant N/Mn$ and for $b = |\sum_{i\notin I} a_i|$ let again $A \subseteq A_1$ consist of all elements of $A_1$ belonging to the same residue class modulo $b$ with $|A| \geqslant |A_1|/b$. Set also $\varepsilon = \exp(-CMbn\log^5(Mbn))$, where $C > 0$ is the constant given by Lemma 2.3.

Suppose that $I = \{a_1, a_2, a_3\}$ and observe that no $A_i$ contains a solution to the equation

$$a_1x_1 + a_2x_2 + a_3x_3 + by = 0\,.$$

By Lemma 2.3 there are at least $\varepsilon|A|^2$ solutions to the invariant equation

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

In other words, 0 has at least $\varepsilon|A|^2$ representations in $a_1{\cdot}A + a_2{\cdot}A + a_3{\cdot}A$, hence by Lemma 2.4 there is a symmetric arithmetic progression $P$ of length $\Omega(\varepsilon N^{\varepsilon^2/(bn)^3})$ contained in $a_1{\cdot}A + a_2{\cdot}A + a_3{\cdot}A \subseteq \{1, \ldots, N\}$. Hence, since the set $A_1$ is free of solutions to the equation considered, $\frac{1}{b}{\cdot}P \subseteq A_2 \cup \cdots \cup A_n$, so that $\mathsf{R}(n-1) \gg \varepsilon \mathsf{R}(n)^{\varepsilon^2/(Mbn)^3}$.

The last inequality implies that $R(n) \leqslant R(n-1)^{O(n^3\exp(O(n\log^5 n)))}$, which proves (2.1). ∎

It is worth mentioning that one can obtain even better upper bound for all equations containing an equation of genus 2. To get still further improvement, instead of arithmetic progressions we make use of Bohr sets. This reduces roughly one exponent in our bounds, but it makes all the proofs more complicated. A crucial additive property of dense sets $A \subseteq \mathbb{Z}/N\mathbb{Z}$ that influences our approach is that one can guarantee existence of a shift of a large, low dimensional Bohr sets in $A+A+A$, but it is just not so for $A+A$. Therefore, we cannot proceed as in the proof of Theorem 2.2. On the other hand, one can show that $A+A$ contains a large proportion of a shift of a low dimensional, large Bohr set, which allows us to overcome this difficulty. Next sections contain rigorous proofs based on the above ideas.

## 2.2 Main results based on Bohr sets analysis

Proving the strongest results of ours requires recalling a more sophisticated concept of Bohr sets, some extra notation and some lemmas. Bohr sets were introduced to modern additive combinatorics, beyond the limited setting of the Freiman-type problems, by Bourgain [4] and since then became a fundamental tool in additive combinatorics. Sanders [18, 19] further developed the theory of Bohr proving many important results.

**Definition 2.6.** Let $G = \mathbb{Z}/N\mathbb{Z}$ be a cyclic group and its dual group be $\widehat{G} \backsimeq \mathbb{Z}/N\mathbb{Z}$. We define the Bohr set with frequency set $\Gamma \subseteq \widehat{G}$ and width parameter $\gamma \in (0,2]$ to be the set

$$B(\Gamma,\gamma) = \{x \in G : \forall_{t\in\Gamma} \left\| \frac{tx}{N} \right\| \leqslant \gamma\}\,.$$

Also, we call $\dim B = |\Gamma|$ the *dimension* of the Bohr set $B$.

An important property of Bohr sets to mention is that $c{\cdot}B(\Gamma,\gamma) = B(c^{-1}{\cdot}\Gamma,\gamma)$. Furthermore, for $\eta > 0$ and a Bohr set $B = B(\Gamma,\gamma)$ by $B_\eta$ we mean the Bohr set $B(\Gamma,\eta\gamma)$.

The above definition and the three lemmas below are pretty standard, hence we refer the reader to [23] for a more complete account.

**Lemma 2.7.** *For every $\gamma > 0$ we have*

$$\gamma^{|\Gamma|} N \leqslant |B(\Gamma,\gamma)| \leqslant 8^{|\Gamma|+1} |B(\Gamma,\gamma/2)|\,.$$

The size of Bohr sets can vary significantly even for small changes of the width function, which is the motivation for the following definition.

**Definition 2.8.** We call a Bohr set $B(\Gamma,\gamma)$ *regular* if for every $\eta$, $|\eta| \leqslant 1/(100|\Gamma|)$, we have

$$(1-100|\Gamma||\eta|)|B| \leqslant |B_{1+\eta}| \leqslant (1+100|\Gamma||\eta|)|B|.$$

Bourgain [4] showed that regular Bohr sets are ubiquitous.

**Lemma 2.9.** *For every Bohr set $B(\Gamma,\gamma)$ there exists $\frac{1}{2}\gamma \leqslant \gamma' \leqslant \gamma$ such that $B(\Gamma,\gamma')$ is regular.*

The most important consequence of regularity of a Bohr set is expressed by the following lemma.

**Lemma 2.10** ([5, Lemma 3.16]). *Let $B$ be a $d$-dimensional, regular Bohr set. Suppose that $S \subseteq B_\varepsilon$ and $\varepsilon < \kappa/(100d)$. Then for every set $A \subseteq B$, we have*

$$\|\mu_B \cdot A - (\mu_B * \mu_S) \cdot A\|_1 < 2\kappa\,. \tag{2.2}$$

An immediate consequence of the above lemma is the following.

**Lemma 2.11.** *Let $B$ be a $d$-dimensional regular Bohr set, let $A \subseteq B$ and $\mu_B(A) = \delta$. Suppose that $S \subseteq B_\varepsilon$ and $\varepsilon < \kappa\delta/(200d)$. Then*

$$\frac{1}{|B|}\sum_{x \in B} \mu_S(A + x) \geqslant (1-\kappa)\delta\,.$$

*Proof.*

$$\begin{aligned}
\delta &= \sum_{x \in B} \mu_B(x) A(x) \\
&\leqslant \|\mu_B \cdot A - \mu_B * \mu_S \cdot A\|_1 + \sum_{x \in B} (\mu_B * \mu_S)(x) A(x) \\
&\leqslant \kappa\delta + \frac{1}{|B|}\sum_{x \in B} (\mu_S * (-A))(x)\,.
\end{aligned}$$

■

The above is pretty standard and we will refer to it in course of proving the theorems.

**Proof of Theorem 1.3.**

The following lemma is due to Sanders.

**Lemma 2.12** ([18, Lemma 6.4]). *Let $B = B(\Gamma, \gamma)$ be a $d$-dimensional regular Bohr set and let $A \subseteq B$ and $\mu_B(A) = \delta_A \geqslant \delta$. Then either $A - A$ contains $(1-\alpha)$ fraction of a regular Bohr set $B_\rho$, where $\rho \gg \delta^4/d$ and $\rho$ does not depend on $A$, or there is a regular Bohr set $B' = B(\Gamma \cup \Lambda, \gamma')$ and $x$ such that $\mu_{B'}(A + x) \geqslant 1.01\delta_A$. Furthermore, $|\Lambda| = O(\delta^{-2}\log(1/\alpha))$ and $\gamma' \gg \gamma\delta^6/(d^3\log(1/\alpha))$.*

It is important to realize that the Bohr set mentioned in the first alternative of the lemma can be chosen universally, i.e. independently of the set $A$. This follows from Sanders's proof of the lemma, although he does not state it this way. We will make use of this property when we apply the lemma to several sets $A_i$ simultaneously.

*Proof of Theorem 1.3.* Clearly, it is enough to consider an equation of the form

$$ax - ay + bz = 0$$

with $a, b > 0$. Suppose that $[N] = A_1 \cup \cdots \cup A_n$ is a solution free partition. Let $p$ be a prime between $(2a+b)N$ and $2(2a+b)N$. Then each color class is solution free in $\mathbb{Z}/p\mathbb{Z}$.

Let $\delta = 1/(3n)$ and $\varepsilon = n^{-2}$. We build the proof around an iterative procedure and during its execution we keep track of several variables: a subset $\mathcal{I} \subseteq [n]$, a regular Bohr set $B = B(\Gamma, \gamma)$, counters $\mathsf{Count}_i$ and the aggregated value $\mathsf{Total} = \sum \mathsf{Count}_i$. Also, we make the following invariants hold:

**I0** $\forall_{i \notin \mathcal{I}}\ \mathsf{Count}_i = 0$

**I1** $\forall_{i \in \mathcal{I}}\ \mu_{ab \cdot B}(a \cdot A_i + x_i) \geqslant 1.01^{\mathsf{Count}_i} \cdot \frac{\delta}{2} + (O(n \log n) - \mathsf{Total})\varepsilon$ for some $x_i$

**I2** $B = B(\Gamma, \gamma)$ is regular, $|\Gamma| = O(\mathsf{Total} \cdot n^2 \log n)$ and $\gamma \gg n^{-O(\mathsf{Total})}$

The aim that the procedure is supposed to pursue is to make the following conditions hold:

**C1** $\forall_{i \in \mathcal{I}}\ \mu_{ab \cdot B}(a \cdot A_i - a \cdot A_i) \geqslant 1 - \delta$

**C2** $\forall_{i \notin \mathcal{I}}\ \mu_{ab \cdot B}(b \cdot A_i) < \delta$

To begin with let $B = B^0 = [-N/a, N/a]$, $\mathcal{I} = \emptyset$ and $\mathsf{Count}_i = 0$ for all $i$. Whenever any of the conditions is violated we perform one of the two operation described below and increase one of the counters by one.

By the invariant it is clear that this procedure stops after at most $O(n \log n)$ steps and, when it stops, we must have both conditions satisfied. For this reason we allow ourselves to plug the bound $\mathsf{Total} = n^{O(1)}$ into the calculations below. Then, since $(a \cdot A_i - a \cdot A_i) \cap b \cdot A_i = \emptyset$, by condition (C1) we have

$$\forall_{i \in \mathcal{I}}\ \mu_{ab \cdot B}(b \cdot A_i) < \delta\,.$$

When combined with condition (C2), we get

$$\mu_{ab \cdot B}(b \cdot [N]) < n\delta = \frac{1}{3},$$

which is a contradiction if $|B| \geqslant 7$, because by the initial choice

$$ab \cdot B \subseteq ab \cdot B^0 \subseteq b \cdot [-N, N]$$

and therefore $\mu_{ab \cdot B}(b \cdot [N]) = \frac{1}{2} - \frac{1}{|B|} > \frac{1}{3}$. Hence $|B| \leqslant 6$ which by Lemma 2.7 implies

$$N = 2^{O(n^4 \log^4 n)}\,.$$

It is now enough to describe what operations are performed in case a condition does not hold and to verify that the invariants are preserved.

If condition (C2) is violated, then there is $i \notin \mathcal{I}$ such that $\mu_{ab \cdot B}(b \cdot A_i) \geqslant \delta$, which is equivalent to $\mu_{a \cdot B}(A_i) \geqslant \delta$. Therefore, by Lemma 2.11 we have

$$\mu_{a \cdot (b \cdot B_\eta)}(A_i + x_i) \geqslant 0.9\delta \geqslant 1.01\frac{\delta}{2} + O(n \log n)\varepsilon$$

for $\eta = \varepsilon\delta/(4000b|\Gamma|) = n^{-O(1)}$ and for some $x_i$. The above implies, for a re-defined $x_i$, that

$$\mu_{ab \cdot (a \cdot B_\eta)}(a \cdot A_i + x_i) \geqslant 1.01\frac{\delta}{2} + O(n \log n)\varepsilon\,.$$

To finalize the operation we update our variables.

1. $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$
2. $B \leftarrow a{\cdot}B_\eta$
3. $\mathsf{Count}_i \leftarrow \mathsf{Count}_i + 1 = 1$.

The only invariant that holds now in a not immediately obvious manner is (I1) for $\mathcal{I}\backslash\{i\}$. However, we know that it held for the old value of $B$, of which the new one is a small subset. Therefore, thanks to the choice of $\eta$ sufficiently small, Lemma 2.11 guarantees that at the expense of one $\varepsilon$ we may have the invariant satisfied.

If condition (C2) is not violated but condition (C1) is so, we need to distinguish two cases. The first is that, for $\rho = \Omega(\delta^4/|\Gamma|) = n^{-O(1)}$, we have condition (C1) satisfied for $B_\rho$ and the family $(A_i)_{i\in\mathcal{I}}$; the second is the opposite, which by Lemma 2.12 implies that some density increment is possible for one of the sets $A_i$ for $i \in \mathcal{I}$.

Let us now consider the first case. If condition (C2) remains satisfied for the resulting Bohr set $B_\rho$ the procedure stops with

1. $B \leftarrow B_\rho$.

Otherwise, for $B_\rho$ and some $i \notin \mathcal{I}$, we repeat the operation described for the case of condition (C2) being violated. This results in the following.

1. $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$
2. $B \leftarrow a{\cdot}B_{\eta\rho}$
3. $\mathsf{Count}_i \leftarrow \mathsf{Count}_i + 1 = 1$.

In the only case remaining there is some $i \in \mathcal{I}$ such that $\mu_{ab\cdot B_\rho}(a{\cdot}A_i - a{\cdot}A_i) < 1 - \delta$. By Lemma 2.12 there is a regular Bohr set $B' = B(\Gamma', \gamma') \subseteq B_\eta$ and $x$ such that $\mu_{ab\cdot B'}(a{\cdot}A_i + x) \geqslant 1.01\mu_{ab\cdot B}(a{\cdot}A_i + x_i)$. Furthermore, $\dim B' = \dim B + O(\delta^{-2}\log(1/\delta)) = \dim B + O(n^2 \log n)$ and $\gamma' \gg \gamma\delta^6/(|\Gamma|^3 \log(1/\delta)) = \gamma \cdot n^{-O(1)}$. Therefore, we update the variables accordingly:

1. $B \leftarrow B'$
2. $\mathsf{Count}_i \leftarrow \mathsf{Count}_i + 1$.

Again, like in the first case considered, Lemma 2.11 guarantees that the invariants keep being satisfied. ■

**Proof of Theorem 1.4.**

The lemma below is what really stands behind proofs of upper bounds for equations with many variables and we will also make use of it in the proof of Theorem 1.5. This is a local variant of, established in [19], Sanders's efective version of Bogolyubov's lemma.

**Lemma 2.13** ([21, Theorem 5.2]). *Let $\varepsilon \in (0,1]$ be a real number. Let $A$ and $S$ be subsets of regular Bohr sets $B$ and $B_\varepsilon$, respectively, where $\varepsilon \leqslant 1/(100d)$ and $d = \dim B$. Suppose that $\mu_B(A), \mu_{B_\varepsilon}(S) \geqslant \delta$. Then $A - A + S - S$ contains a regular Bohr set $\tilde{B} \subseteq B$, such that* $\dim\tilde{B} = d + O(\log^4(1/\delta))$ *and*

$$|\tilde{B}| \geqslant \exp(-O(d\log d + d\log(1/\varepsilon) + \log^4(1/\delta)\log d + \log^5(1/\delta) + d\log(1/\delta)))|B|\,. \tag{2.3}$$

This lemma is proved in [21] for pairs of possibly different sets $S, S'$ and $T, T'$. It is precisely this that stands behind the resulting Bohr set $\tilde{B}$ being translated in the original statement of the lemma. One can check that in the symmetric case of ours a genuine non-translated set $\tilde{B}$ can be found.

Next two lemmas will serve as a main iterative block of Lemma 2.16. The first one was proved by Sanders and is a local version of the Heath-Brown-Szemeredi density increment method.

**Lemma 2.14** ([20, Lemma 3.8])**.** *Let $0 < \eta, \varepsilon \leqslant 1$. Let $A \subseteq B$ and $S \subseteq B_\varepsilon$ be such that $\mu_B(A) = \delta$ and $\mu_{B_\varepsilon}(S) = \tau$ for a $d$-dimensional regular Bohr set $B$. If*

$$\sum_{r \in \mathrm{Spec}_{\,\eta}(S)} |\hat{A}(r)|^2 \geqslant (1+\nu)|A|^2 \,,$$

*then there is a regular Bohr set $B' \subseteq B_\varepsilon$ of dimension* $\mathrm{dim} B' = d + O(\eta^{-2}\log(1/\tau))$ *and cardinality*

$$|B'| \geqslant \left(\frac{\eta}{2d\log(1/\tau)}\right)^{d+O(\eta^{-2}\log(1/\tau))} |B_\varepsilon|$$

*such that $\mu_{B'}(A+x) \geqslant \delta(1+\Omega(\nu))$ for some $x$.*

**Lemma 2.15.** *Let $B \subseteq \mathbb{Z}/N\mathbb{Z}$ be a regular $d$-dimensional Bohr set and $\varepsilon < c/100d$ for $1/64 < c < 1/32$. If $A, A' \subseteq B$ and $S, S' \subseteq B_\varepsilon$, and*

$$\mu_B(A), \mu_B(A'), \mu_{B_\varepsilon}(S), \mu_{B_\varepsilon}(S') \geqslant \delta \,,$$

*then either there is $x \in B_{1+\varepsilon}$ such that*

$$(A * S)(x), (A' * S')(-x) \geqslant \frac{1}{10}\delta^2 |B_\varepsilon| \,,$$

*or there is a regular Bohr set $B' \subseteq B_\varepsilon$ such that* $\mathrm{dim} B' = \mathrm{dim} B + O(\delta^{-1}\log(1/\delta))$,

$$|B'| \gg \left(\frac{\delta}{2d\log(1/\delta)}\right)^{d+O(\delta^{-1}\log(1/\delta))} |B_\varepsilon|$$

*and $\mu_{B'}(A+y) \geqslant (1+\Omega(1))\delta$ or $\mu_{B'}(A'+y) \geqslant (1+\Omega(1))\delta$ for some $y$.*

*Proof.* We have $A+S, A'+S' \subseteq B_{1+\varepsilon}$ and, by regularity, $|B_{1+\varepsilon}| \leqslant (1+c)|B|$. Let us assume that there is no $x$ satisfying the property required, i.e. for all $x \in B_{1+\varepsilon}$ we have either

$$(A * S)(x) < \frac{1}{10}\delta^2|B_\varepsilon| \qquad \text{or} \qquad (A' * S')(x) < \frac{1}{10}\delta^2|B_\varepsilon|.$$

By symmetry we may assume that $(A*S)(x) < \frac{1}{10}\delta^2|B_\varepsilon|$ for at least $\frac{1}{2}|B_{1+\varepsilon}|$ elements $x \in B_{1+\varepsilon}$. Let us denote the set of these $x$'s by $X$.

Therefore

$$\begin{aligned} \sum_{x \in B_{1+\varepsilon} \setminus X} (A*S)(x) &\geqslant \sum_{x \in B_{1+\varepsilon}} (A*S)(x) - (\delta^2/10)|B_{1+\varepsilon}||B_\varepsilon| \\ &\geqslant |A||S| - ((1+c)/10)|A||S| \geqslant \frac{4}{5}|A||S|, \end{aligned}$$

hence

$$\sum_{x\in B_{1+\varepsilon}\setminus X}(A*S)(x)^2 \geqslant \frac{(\frac{4}{5}|A||S|)^2}{\frac{1}{2}|B_{1+\varepsilon}|} \geqslant \frac{(\frac{4}{5}|A||S|)^2}{\frac{1+c}{2}|B|} \geqslant \frac{6|A|^2|S|^2}{5|B|} = \frac{6}{5}\delta|A||S|^2\,.$$

It follows by Parseval's formula that

$$\frac{1}{N}\sum_{r=0}^{N-1}|\widehat{A}(r)|^2|\widehat{S}(r)|^2 \geqslant \sum_{s\in B_{1+\varepsilon}\setminus X}(A*S)(x)^2 \geqslant \frac{6}{5}\delta|A||S|^2$$

and, by the definition of spectrum and Parseval's formula,

$$\frac{1}{N}\sum_{r\notin \mathrm{Spec}\,_\eta(S)}|\widehat{A}(r)|^2|\widehat{S}(r)|^2 \leqslant (\eta|S|)^2\cdot\frac{1}{N}\sum_{r\in\mathbb{Z}_N}|\widehat{A}(r)|^2 = c\delta|S|^2|A|,$$

for $\eta = (c\delta)^{1/2}$. Therefore, as $|\widehat{S}(r)| \leqslant |S|$,

$$\sum_{r\in \mathrm{Spec}\,_\eta(S)}|\widehat{A}(r)|^2 \geqslant \frac{1}{|S|^2}\sum_{r\in \mathrm{Spec}\,_\eta(S)}|\widehat{A}(r)|^2|\widehat{S}(r)|^2 \geqslant \frac{7}{6}|A|^2.$$

The proof concludes with application of Lemma 2.14. ■

**Lemma 2.16.** *Let $B$ be a regular $d$-dimensional Bohr set in $\mathbb{Z}/N\mathbb{Z}$ such that* $\mathrm{dim}B = d$*. Suppose that $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is such that $\mu_B(A) = \delta$ and it contains no solution to the equation*

$$b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + bx = 0$$

*with $b_1 + b_2 + b_3 + b_4 = 0$. Then there exists a regular Bohr set $T \subseteq B$ disjoint from $A$ such that*

$$\mathrm{dim}T = \mathrm{dim}B + O(\delta^{-1}\log^2(1/\delta))$$

*and*

$$|T| \geqslant \exp(-O(d\log d\log^2(1/\delta) + d\log^3(1/\delta) + \log d\cdot\delta^{-1}\log^4(1/\delta) + \delta^{-1}\log^5(1/\delta)))|B|\,. \quad (2.4)$$

*The implied constants depend only on the equation considered.*

*Proof.* Let $M = \max|b_i|$, $1+c$ be the density increment factor given by Lemma 2.15, which we assume to be smaller than 2, and $c_1, c_2 > 0$ be constants small enough for the argument below to work. Set $\varepsilon_1 = c_1\delta/(Md)$ and $\varepsilon_2 = c_2\delta^2/(Md)$.

Let us consider the Bohr sets

$$B^1 = \frac{1}{b_1}\cdot B_{\varepsilon_1}, B^2 = \frac{1}{b_2}\cdot B_{\varepsilon_2}, B^3 = \frac{1}{b_3}\cdot B_{\varepsilon_1}, B^4 = \frac{1}{b_4}\cdot B_{\varepsilon_2}\,.$$

By a proper choice of constants $c_1$ and $c_2$ we may assume that they are all regular Bohr sets of dimension $d$ and, by Lemma 2.7, $|B^i| = \Omega(\delta/d)^{6d+6}|B|$. These sets are all subsets of $B_{c_1\delta/d}$, so by Lemma 2.11 we have

$$\frac{1}{|B|}\sum_{x\in B}\sum_{i=1}^{4}(\mu_{B^i}*A)(x) \geqslant (4-\frac{c}{3})\delta\,.$$

Therefore, either for some $x \in B$ and for all $i = 1, \dots, 4$ we have

$$\mu_{B_{\varepsilon_i}}(b_i \cdot (A + x)) = \mu_{B^i}(A + x) \geqslant (1 - \frac{c}{2})\delta\,,$$

with the convention $\varepsilon_3 = \varepsilon_1$ and $\varepsilon_4 = \varepsilon_2$, or $\mu_{B^i}(A+x) \geqslant (1 + \frac{c}{18})\delta$ for some $i$. In the latter case we repeat the above reasoning for the pair $A + x, B^i$.

Since the density is naturally bounded from above, after at most $O(\log(1/\delta))$ iterative steps we end up with some $A + x$ and some Bohr set $B' \subseteq B$ such that $\mu_{B'}(A + x) \geqslant \delta$ and for some $y$ and all $i = 1, \dots, 4$ we have

$$\mu_{B'_{\varepsilon_i}}(b_i \cdot (A + y)) = \mu_{B'^i}(A + y) \geqslant (1 - \frac{c}{2})\delta\,.$$

Also, $\dim B' = d$ and

$$|B'| = \Omega(\delta/d)^{O(d \log(1/\delta))}|B|\,.$$

Hence, by Lemma 2.15 we have either

$$\mu_{B'_{\varepsilon_2}}(b_2 \cdot (A + y) \cap (x - b_1 \cdot (A + y))) \geqslant 0.1((1 - \frac{c}{2})\delta)^2 \geqslant 0.01\delta^2$$

and

$$\mu_{B'_{\varepsilon_2}}(b_4 \cdot (A + y) \cap (-x - b_3 \cdot (A + y))) \geqslant 0.01\delta^2$$

for some $x$, or there is a regular Bohr set $B'' \subseteq B' \subseteq B$ with the following properties:

$$\dim B'' = \dim B' + O(\delta^{-1} \log(1/\delta)),$$

$$|B''| \geqslant \Big(\frac{\delta}{2d\log(1/\delta)}\Big)^{d+O(\delta^{-1}\log(1/\delta))}|B'| \geqslant \Big(\frac{\delta}{2d\log(1/\delta)}\Big)^{d+O(d\log(1/\delta)+\delta^{-1}\log(1/\delta))}|B|\,.$$

and, for some $z$, we have $\mu_{B''}(A + z) \geqslant (1 + c) \cdot (1 - \frac{c}{2})\delta = (1 + \Omega(1))\delta$.

Repetition of the above procedure at most $O(\log(1/\delta))$ times results in a translate $A + x$ and a regular Bohr set $\tilde{B} \subseteq B$ such that

$$\mu_{\tilde{B}_{\varepsilon_2}}(b_2 \cdot (A + y) \cap (x - b_1 \cdot (A + y))) \geqslant 0.01\delta^2$$

and

$$\mu_{\tilde{B}_{\varepsilon_2}}(b_4 \cdot (A + y) \cap (-x - b_3 \cdot (A + y))) \geqslant 0.01\delta^2$$

for some $x$. Furthermore $\tilde{B}_{\varepsilon_2}$ is a regular Bohr set of $\dim \tilde{B} = d + O(\delta^{-1} \log^2(1/\delta))$, and

$$|\tilde{B}| \geqslant \Big(\frac{\delta}{d}\Big)^{O(d\log^2(1/\delta)+\delta^{-1}\log^2(1/\delta))}|B|\,.$$

Let $A' = b_2 \cdot (A + y) \cap (x - b_1 \cdot (A + y))$ and $S = b_4 \cdot (A + y) \cap (-x - b_3 \cdot (A + y))$. We are almost done but we cannot yet apply Lemma 2.13. One last application of Lemma 2.11 shows that there is some $s$ such that $\mu_{\tilde{B}_{\varepsilon_3}}(S + s) \geqslant \delta^2/101$ for $\varepsilon_3 = \varepsilon_1^3$. Write $S' = (S + s) \cap \tilde{B}_{\varepsilon_3}$.

Applying Lemma 2.13 we obtain a Bohr set $T'$ such that

$$T' \subseteq A' - A' + S' - S' \subseteq 2\tilde{B}_{\varepsilon_2} + 2\tilde{B}_{\varepsilon_3} \subseteq B$$

and

$$T' \subseteq A' - A' + S' - S' \subseteq b_1 A + b_2 A + b_3 A + b_4 A\,.$$

In particular, the above implies that $T = T'_{1/b}$ is disjoint from $A$, because $A$ is free of solutions to the equation by assumption.

The dimension of $T$ is

$$\dim T = \dim\tilde{B} + O(\log^4(1/\delta)) = \dim B + O(\delta^{-1}\log^2(1/\delta))$$

and its cardinality is

$$\begin{aligned} |T| &\geqslant \exp(-O(d\log d + d\log(1/\delta) + \log d \log^4(1/\delta) + \log^5(1/\delta))))|\tilde{B}_{\varepsilon_2}| \\ &\geqslant \exp(-O(d\log d\log^2(1/\delta) + d\log^3(1/\delta) + \log d\cdot\delta^{-1}\log^4(1/\delta) + \delta^{-1}\log^5(1/\delta)))|B|\,. \end{aligned}$$

■

*Proof of Theorem 1.4.* Suppose that $[N] = A_1 \cup \cdots \cup A_n$ is a solution free partition. Let $p$ be a prime between $(|b_1| + |b_2| + |b_3| + |b_4| + b)N$ and $2(|b_1| + |b_2| + |b_3| + |b_4| + b)N$. Then each color class is solution free in $\mathbb{Z}/p\mathbb{Z}$. We start with $T_0 = [-N, N]$ and let $A_1$ be any class with $\mu_{T^0}(A_1) \geqslant 1/(3n)$. Iterative application of Lemma 2.16 gives after $k$ steps a Bohr set $T_k \subseteq T_{k-1}$ that is disjoint from $A_1 \cup \cdots \cup A_k$ and

$$|T_k| \gg \exp(-O(kn\log^5 n))|T_{k-1}| \gg \exp(-O(k^2 n\log^5 n))N\,.$$

Clearly, $T_n$ does not contain any element from $A_1 \cup \cdots \cup A_n$. In particular $T_n = \{0\}$, so that

$$1 = |T_n| \gg \exp(-O(n^3\log^5 n))N,$$

and therefore

$$\mathsf{R}(n) \ll 2^{Cn^3\log^5 n}\,,$$

which completes the proof. ■

**Proof of Theorem 1.5.**

A careful reader might have noticed that a proof of Theorem 1.5 can be deduced from that of Theorem 1.4, because we get sets $S$ and $T$ for free in the genus 2 case when $b_2 = -b_1$ and $b_4 = -b_3$. We extract these essentials here.

The lemma we are about to prove constitutes the main iterative step of the proof of Theorem 1.5.

**Lemma 2.17.** *Let $b_1, b_2$ and $b$ be positive integers and let $B = B(\Gamma, \gamma)$ be a regular Bohr set of dimension $d$. Suppose that $A \subseteq B, \mu_B(A) = \delta$, does not contain any solution to the equation*

$$b_1x_1 - b_1x_2 + b_2x_3 - b_2x_4 + by = 0\,.$$

*Then there exists a regular Bohr set $T \subseteq B$ disjoint from $A$ such that*

$$\dim T = \dim B + O(\log^4(1/\delta))$$

*and*

$$|T| \geqslant \exp(-O(d \log d + \log^4(1/\delta) \log d + \log^5(1/\delta) + d \log(1/\delta)))|B|\,. \tag{2.5}$$

*The implied constants depend only on $b_1, b_2$ and $b$.*

*Proof.* Choose a constant $1/64 \leqslant c \leqslant 1/32$ such that $B_\varepsilon$ is a regular Bohr set, where $\varepsilon = c\delta/(100 b_1 b_2 d)$. Put $B^i = b_i{\cdot}B_\varepsilon, i = 1, 2,$ $B' = b_1 b_2{\cdot}B_\varepsilon$. By Lemma 2.11 we have

$$\frac{1}{|B|} \sum_{x \in B} (\mu_{B^1} * A)(x) \geqslant (1 - 2c)\delta \geqslant \frac{1}{2}\delta\,.$$

Therefore for some $x$

$$(\mu_{B^1} * A)(x) = \mu_{B^1}(A + x) \geqslant \frac{1}{2}\delta\,,$$

hence

$$\mu_{B'}(b_2(A + x)) = \mu_{B^1}(A + x) \geqslant \frac{1}{2}\delta\,.$$

Again choose a constant $1/64 \leqslant c' \leqslant 1/32$ such that $B'_{\varepsilon'}$ is a regular Bohr set, where $\varepsilon' = c'\delta/(100 b_1 b_2 d)$. Using the same argument we find $y$ such that

$$\mu_{B'_{\varepsilon'}}(b_1{\cdot}(A + y)) \geqslant \frac{1}{2}\delta\,.$$

Therefore, by Lemma 2.13

$$b_1{\cdot}A - b_1{\cdot}A + b_2{\cdot}A - b_2{\cdot}A$$

contains a Bohr set $\tilde{B} \subseteq B'$ of dimension $\dim B + O(\log^4(1/\delta))$ that satisfies (2.3). Since $A$ is a solution free set it follows that $b{\cdot}A$ is disjoint from $\tilde{B}$, hence $A$ is disjoint from $T := \tilde{B}_{1/b}$. Observe that

$$T \subseteq B' = b_1 b_2{\cdot}B_\varepsilon \subseteq B_{b_1 b_2 \varepsilon} \subseteq B\,.$$

To finish the proof it is enough to establish (2.5). By Lemma 2.7 and Lemma 2.13 we have

$$\begin{aligned}
|T| &\geqslant \exp(-O(d + \log^4(1/\delta)))|\tilde{B}| \\
&\geqslant \exp(-O(d \log d + d \log(1/\varepsilon) + \log^4(1/\delta) \log d + \log^5(1/\delta) + d \log(1/\delta)))|B'| \\
&\geqslant \exp(-O(d \log d + \log^4(1/\delta) \log d + \log^5(1/\delta) + d \log(1/\delta)))|B|\,.
\end{aligned}$$

Finally the assertion follows by Lemma 2.9 and Lemma 2.7. ■

*Proof of Theorem 1.5.* We proceed similarly as in the proof of Theorem 1.4. Suppose that $\{1,\dots,N\} = A_1\cup\dots\cup A_n$ is a solution free partition and let $p$ be a prime between $(2c_1+2c_2+b)N$ and $2(2c_1+2c_2+b)N$. Then each color class is solution free in $\mathbb{Z}/p\mathbb{Z}$. We start with $T^0 = [-N, N]$ and let $A_1$ be any class with $\mu_{T^0}(A_1) \geqslant 1/(3n)$. Iterative application of Lemma 2.17 gives after $k$ steps a Bohr set $T_k \subseteq T_{k-1}$ of dimension $O(k\log^4 n)$, that is disjoint from $A_1 \cup \dots \cup A_k$ and

$$|T_k| \geqslant \exp(-O(k\log^5 n))|T_{k-1}|\,.$$

Since $T_n \cap (A_1 \cup \dots \cup A_n) = \emptyset$ it follows that $T_n = \{0\}$. Hence

$$1 = |T_n| \geqslant \exp(-O(n^2 \log^5 n))N,$$

and the assertion follows. ■

## 3 Schur-like numbers

In this section we prove Theorem 1.6, which lowers an upper bound on Schur-like numbers below the threshold established by a natural argument presented in the beginning of Subsection 2.1.

To begin with we recall a fundamental result of Plünnecke and Ruzsa.

**Lemma 3.1** (Plünnecke-Ruzsa)**.** *Suppose that $A$ is a subset of an abelian group and $|A+B| \leqslant K|B|$. Then for all natural numbers $k, l \geqslant 0$ we have*

$$|kA - lA| \leqslant K^{k+l}|B|,$$

*where $kA$ and $lA$ denote iterated sumsets.*

We shall also need the following two lemmas.

**Lemma 3.2.** *Suppose that $\{1,\dots,N\} = A_1\cup\dots\cup A_n$ is a partition into sum-free sets such that $|A_1| \geqslant \ldots \geqslant |A_n|$ and set $\sigma_k = \sum_{i>k}|A_i|$. Then we have*

$$|A_k| > \frac{|A_1|}{(k-1)!} - 2(\sigma_k + 1).$$

*Proof.* We use a classical Schur's argument. Let $A_1 = \{a_1,\dots,a_t\}_<$ and notice that all numbers $a_2 - a_1,\dots,a_t - a_1$ belong to $A_2 \cup \dots \cup A_n$. At most $\sigma_k$ of these elements belong to $\bigcup_{i>k} A_i$, hence $B_2 = \{a_2 - a_1,\dots,a_t - a_1\} \cap A_{i_2}$ has

$$|B_2| \geqslant \frac{|A_1| - 1 - \sigma_k}{k-1}$$

elements for some $2 \leqslant i_2 \leqslant k$. Therefore, by repeated application of the above argument, we have

$$|B_k| \geqslant \frac{|A_1|}{(k-1)!} - (\sigma_k + 1)\sum_{i=1}^{k-1}\frac{1}{i!} > \frac{|A_1|}{(k-1)!} - 2(\sigma_k + 1)\,,$$

and the assertion follows, because $|A_k| \geqslant |B_j| \geqslant |B_k|$ for $j$ such that $i_j = k$. ■

**Lemma 3.3.** *Suppose that $N < \mathsf{S}_2(n)$. Then there exists a partition $\{1, \ldots, N\} = A_1 \cup \cdots \cup A_n$ into 2-sum-free sets such that $|A_1| \geqslant \ldots \geqslant |A_n|$ and*

$$\bigcup_{i>k} A_i \subseteq (3A_k - A_k) \cup (2A_k - 2A_k)\,,$$

*for every $1 \leqslant k \leqslant n$.*

*Proof.* Let $\{1, \ldots, N\} = A_1 \cup \cdots \cup A_n$ be any maximal 2-sum-free partition with respect to the lexicographical order of $(|A_1|, \ldots, |A_n|)$. Since no element $a \in \bigcup_{i>k} A_i$ can be added to $A_k$ without spoiling the 2-sum-free property, we have $\bigcup_{i>k} A_i \subseteq (3A_k - A_k) \cup (2A_k - 2A_k)$. ■

The following theorem is the main result of this section.

*Proof of Theorem 1.6.* Assume that our partitioning satisfies the assertion of Lemma 3.3. First, we show that there exist $x_2, \ldots, x_k$ such that

$$|(A_1 - A_1) \cap (A_2 - A_2 + x_2) \cap \cdots \cap (A_k - A_k + x_k)| = \left(\Omega(n^{-\frac{9}{10}})\right)^k N\,, \tag{3.1}$$

for some $k \gg \frac{\log n}{\log\log n}$. To this end we shall prove that the sets $A_i - A_i$, for $i = 1, \ldots, k$, are large. The proof distinguishes two cases.

First, suppose that $|A_1| \leqslant N/n^c$, for some appropriate positive $c$. Then there are at least $\frac{1}{2}n^c$ classes with at least $N/2n$ elements each. By Lemma 3.3, for every $l$, we have

$$|3A_l - A_l| + |2A_l - 2A_l| \geqslant \sum_{i>l} |A_i| \geqslant N - l\frac{N}{n^c}\,.$$

Therefore, by the Plünnecke-Ruzsa inequality,

$$|A_l - A_l| \gg N^{1/4}|A_l|^{3/4} \gg \frac{N}{n^{3/4}}$$

for all $l \leqslant k = \frac{1}{2}n^c$.

Next, we assume that $|A_1| > N/n^c$ and set $k = c\frac{\log n}{\log\log n}$. If $\sigma_k < N/n^{2c}$, then by Lemma 3.2, $|A_k| \gg N/n^{2c}$ and, immediately, $|A_l - A_l| \geqslant |A_k| \gg N/n^{2c}$ for all $l \leqslant k$. If $\sigma_k \geqslant N/n^{2c}$ then it follows that $|A_k| \geqslant |A_{k+1}| \geqslant \sigma_k/n \geqslant N/n^{1+2c}$. Thus, by Lemma 3.3 for every $1 \leqslant l \leqslant k$ we have

$$|3A_l - A_l| + |2A_l - 2A_l| \geqslant \sum_{i>k} |A_i| \geqslant \frac{N}{n^{2c}}\,,$$

so that by the Plünnecke-Ruzsa inequality

$$|A_l - A_l| \geqslant (N/n^{2c})^{1/4}|A_k|^{3/4} \gg \frac{N}{n^{3/4+2c}}\,.$$

In either case we have that $|A_l - A_l| \gg \frac{N}{n^{9/10}}$ for $l \leqslant k = c\frac{\log n}{\log\log n}$. Since the expected size of the set

$$(A_1 - A_1) \cap (A_2 - A_2 + x_2) \cap \cdots \cap (A_k - A_k + x_k) \cap \{1, \ldots, N\},$$

for $x_i$ chosen uniformly at random from $\{-2N+1, \ldots, 2N-1\}$, is at least $(\Omega(n^{-\frac{9}{10}}))^k N$, one obtains (3.1) for some choice of $x_i$'s.

At this point we drop the indexes corresponding to sets $A_{k+1}, \ldots, A_n$ (we will re-enumerate them later on) and we follow Schur's argument again, starting with the set

$$C_k = \{c_1, \ldots, c_q\}_< = (A_1 - A_1) \cap (A_2 - A_2 + x_2) \cap \cdots \cap (A_k - A_k + x_k) \cap \{1, \ldots, N\}$$

given by (3.1). Observe that $c_u - c_v \in 2A_i - 2A_i$, for all $1 \leqslant u < v \leqslant q$ and $1 \leqslant i \leqslant k$. Therefore, as all the sets $A_1, \ldots, A_k$ are 2-sum-free, we have

$$c_2 - c_1, \ldots, c_q - c_1 \notin A_1 \cup \cdots \cup A_k.$$

At least $q' \geqslant (q-1)/(n-k)$ of the above elements, call them $C_{k+1} = \{c'_1, \ldots, c'_{q'}\}_<$, lie in the same partition class, say $A_{k+1}$. It follows by the above argument that

$$c'_2 - c'_1, \ldots, c'_{q'} - c'_1 \notin A_1 \cup \cdots \cup A_{k+1}$$

Iterating this procedure, we obtain in the last step a set $C_n \subseteq A_n$ such that $|C_n| \gg q/(n-k)!$ and $(C_n - \min C_n) \cap (A_1 \cup \cdots \cup A_n) = \emptyset$. Thus $|C_n| \leqslant 1$ and we infer that

$$N \ll n^{\frac{9}{10}k}(n-k)!$$

for some $k \gg \frac{\log n}{\log\log n}$. ■

# 4 Van der Waerden number $\mathsf{W}(3,k)$

Recall, that $\mathsf{W}(3,k)$ is the smallest integer $N$ with the property that for every 2-coloring $\{1, \ldots, N\} = B \cup R$ there is a 3-term arithmetic progression in $B$ or a $k$-term arithmetic progression in $R$. Our last result provides a new upper bound for $\mathsf{W}(3,k)$. The proof of Theorem 4.2 is due to Green [13, Theorem 24]. The improvement relies on the fact that instead of Green's theorem on arithmetic progressions in sumsets we use a result proved in [7].

**Lemma 4.1** ([7, Corollary 1])**.** *Let $A, B \subseteq \{1, \ldots, N\}$ and $|A||B| \geqslant 6N^{2-2/(t-1)}$ for an integer $t$. Then $A + B$ contains an arithmetic progression of length $t$.*

**Theorem 4.2.**

$$\mathsf{W}(3,k) \leqslant 2^{O(k \log k)}.$$

*Proof.* Set $N = \mathsf{W}(3,k) - 1$ and let $\{1, \ldots, N\} = B \cup R$ be a partition such that $B$ does not contain any 3-term arithmetic progression and $R$ does not contain any $k$-term arithmetic progression. Clearly, $|R| \leqslant (1 - 1/k)N$, so that $|B| \geqslant N/k$. Let $S$ be the more numerous parity class in $B$, hence $|S| \geqslant N/2k$. Further, let $S_1$ be the set of $N/4k$ smallest elements of $S$ and set $S_2 = S \setminus S_1$. Since

$$|S_1||S_2| \gg N^{2-\frac{\log k}{\log N}}$$

it follows by Lemma 4.1 that $S_1+S_2$ contains an arithmetic progression of length $c\frac{\log N}{\log k}$. However, $\frac{1}{2}\cdot(S_1+S_2) \subseteq \{1,\ldots,N\} \setminus B = R$, because otherwise there is a 3-term arithmetic progression in $B$. Therefore,

$$\frac{\log N}{\log k} \ll k,$$

which completes the proof. ■

## References


[1] H. L. Abbott, L. Moser, *Sum-Free Sets of Integers,* Acta Arith. 11 (1966), 392–396.

[2] F. A. Behrend, *On sets of integers which contain no three terms in arithmetic progression,* Proc. Nat. Acad. Sci. 23 (1946), 331–332.

[3] T. Bloom, *Translation invariant equations and the method of Sanders,* Bull. London Math. Soc. 44 (2012), 1050–1067.

[4] J. Bourgain, *On triples in arithmetic progression,* Geometric and Functional Analysis 9 (1999), 968–984.

[5] J. Bourgain, *Roth's theorem on progressions revisited,* Journal d'Analyse Mathématique 104 (2008), 155–192.

[6] E. Croot, I. Laba, O. Sisask, *Arithmetic progressions in sumsets and $L^p$-almost-periodicity,* Combinatorics, Probability & Computing, to appear.

[7] E. Croot, I. Ruzsa, T. Schoen, *Arithmetic progressions in sparse sumsets,* Combinatorial number theory, 157–164, de Gruyter, Berlin, 2007.

[8] G. Exoo, *A Lower Bound for Schur Numbers and Multicolor Ramsey Numbers of $K_3$,* Electronic J. Combinatorics 1 (1994), 1–3.

[9] G. Freiman, H. Halberstam, I. Ruzsa, *Integer sum sets containing long arithmetic progressions,* J. London Math. Soc. 46 (1992), 193–201.

[10] A. Geroldinger, I. Ruzsa, *Combinatorial number theory and additive group theory,* Advanced Courses in Mathematics. CRM Barcelona, Birkhäuser Verlag, Basel, 2009, Courses and seminars from the DocCourse in Combinatorics and Geometry held in Barcelona, 2008.

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

[12] R. Graham, B. Rothschild, J. Spencer, *Ramsey theory,* New York: John Wiley and Sons, 1990

[13] B. Green, *Arithmetic progressions in sumsets,* GAFA 12 (2002), 584–597.

[14] B. Landman, A. Robertson, *Ramsey Theory over the integers,* AMS, 2003.

[15] R. Rado, *Studien zur Kombinatorik,* Mathematische Zeitschrift 36 (1933), 424–448.

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

[17] I. Z. Ruzsa, *Solving a linear equation in a set of integers. II,* Acta Arith. 72 (1995), 385–397.

[18] T. Sanders, *Additive structure in sumsets,* Math. Proc. Camb. Philos. Soc. 144 (2008), no. 2, 289–316.

[19] T. Sanders, *On the Bogolubov–Ruzsa Lemma,* Anal. PDE 5 (2012), 627–655.

[20] T. Sanders, *On Roth's Theorem on Progressions,* Ann. of Math. 174 (2011), 619–636.

[21] T. Schoen, I. D. Shkredov, *Roth's theorem in many variables,* Israel J. Math. to appear.

[22] I. Schur, *Über die Kongruenz $x^m + y^m \equiv z^m \pmod p$,* Jahresber. Deutsche Math.-Verein. 25 (1916), 114–116.

[23] T. Tao, V. Vu, *Additive combinatorics,* Cambridge Studies in Advanced Mathematics 105. CUP, 2006.

[24] E. G. Whitehead, *The Ramsey number $N(3,3,3,3;2)$,* Discrete Math. 4 (1974), 389–396.

Faculty of Mathematics, Informatics and Mechanics,
University of Warsaw,
Banacha 2, 02-097 Warszawa, Poland
cwalina@mimuw.edu.pl

Faculty of Mathematics and Computer Science,
Adam Mickiewicz University,
Umultowska 87, 61-614 Poznań, Poland
schoen@amu.edu.pl