ARTICLE IN PRESS

European Journal of Combinatorics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

On solution-free sets of integers

Robert Hancock, Andrew Treglown

School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

ARTICLE INFO

Article history:
Available online xxxx

ABSTRACT

Given a linear equation $\mathcal{L}$, a set $A \subseteq [n]$ is $\mathcal{L}$-free if A does not contain any 'non-trivial' solutions to $\mathcal{L}$. In this paper we consider the following three general questions:

(i) What is the size of the largest $\mathcal{L}$-free subset of $[n]$?
(ii) How many $\mathcal{L}$-free subsets of $[n]$ are there?
(iii) How many maximal $\mathcal{L}$-free subsets of $[n]$ are there?

We completely resolve (i) in the case when $\mathcal{L}$ is the equation $px + qy = z$ for fixed $p, q \in \mathbb{N}$ where $p \geq 2$. Further, up to a multiplicative constant, we answer (ii) for a wide class of such equations $\mathcal{L}$, thereby refining a special case of a result of Green (2005). We also give various bounds on the number of maximal $\mathcal{L}$-free subsets of $[n]$ for three-variable homogeneous linear equations $\mathcal{L}$. For this, we make use of container and removal lemmas of Green (2005).

1. Introduction

Let $[n] := \{1, \dots, n\}$ and consider a fixed linear equation $\mathcal{L}$ of the form

$$a_1x_1 + \dots + a_kx_k = b \tag{1}$$

where $a_1, \dots, a_k, b \in \mathbb{Z}$. If $b = 0$ we say that $\mathcal{L}$ is *homogeneous*. If

$$\sum_{i\in[k]} a_i = b = 0$$

then we say that $\mathcal{L}$ is *translation-invariant*. Let $\mathcal{L}$ be translation-invariant. Then notice that $(x, \dots, x)$ is a 'trivial' solution of (1) for any x. More generally, a solution $(x_1, \dots, x_k)$ to $\mathcal{L}$ is said to be *trivial* if there exists a partition $P_1, \dots, P_\ell$ of $[k]$ so that:

E-mail addresses: rah410@bham.ac.uk (R. Hancock), a.c.treglown@bham.ac.uk (A. Treglown).

http://dx.doi.org/10.1016/j.ejc.2017.06.018

(i) $x_i = x_j$ for every i, j in the same partition class P_r;
(ii) For each $r \in [\ell]$, $\sum_{i \in P_r} a_i = 0$.

A set $A \subseteq [n]$ is *$\mathcal{L}$-free* if A does not contain any non-trivial solutions to $\mathcal{L}$. If the equation $\mathcal{L}$ is clear from the context, then we simply say A is *solution-free*.

The notion of an $\mathcal{L}$-free set encapsulates many fundamental topics in combinatorial number theory. Indeed, in the case when $\mathcal{L}$ is $x_1 + x_2 = x_3$ we call an $\mathcal{L}$-free set a *sum-free set*. This is a notion that dates back to 1916 when Schur [34] proved that, if n is sufficiently large, any r-colouring of $[n]$ yields a monochromatic triple x, y, z such that $x + y = z$. *Sidon sets* (when $\mathcal{L}$ is $x_1 + x_2 = x_3 + x_4$) have also been extensively studied. For example, a classical result of Erdős and Turán [16] asserts that the largest Sidon set in $[n]$ has size $(1 + o(1))\sqrt{n}$. In the case when $\mathcal{L}$ is $x_1 + x_2 = 2x_3$ an $\mathcal{L}$-free set is simply a *progression-free* set. Roth's theorem [27] states that the largest progression-free subset of $[n]$ has size $o(n)$. In [28,29], Ruzsa instigated the study of solution-free sets for general linear equations.

In this paper we prove a number of results concerning $\mathcal{L}$-free subsets of $[n]$ where $\mathcal{L}$ is a homogeneous linear equation in *three variables*. In particular, our work is motivated by the following general questions:

(i) What is the size of the largest $\mathcal{L}$-free subset of $[n]$?
(ii) How many $\mathcal{L}$-free subsets of $[n]$ are there?
(iii) How many *maximal* $\mathcal{L}$-free subsets of $[n]$ are there?

We make progress on all three of these questions. For each question we use tools from graph theory; for (i) and (ii) our methods are somewhat elementary. For (iii) our method is more involved and utilises container and removal lemmas of Green [18].

1.1. The size of the largest solution-free set

As highlighted above, a central question in the study of $\mathcal{L}$-free sets is to establish the size $\mu_{\mathcal{L}}(n)$ of the largest $\mathcal{L}$-free subset of $[n]$. It is not difficult to see that the largest sum-free subset of $[n]$ has size $\lceil n/2 \rceil$, and this bound is attained by the set of odd numbers in $[n]$ and by the interval $[\lfloor n/2 \rfloor + 1, n]$.

When $\mathcal{L}$ is $x_1 + x_2 = 2x_3$, $\mu_{\mathcal{L}}(n) = o(n)$ by Roth's theorem. In fact, very recently Bloom [9] proved that there is a constant C such that every set $A \subseteq [n]$ with $|A| \geq Cn(\log\log n)^4/\log n$ contains a three-term arithmetic progression. On the other hand, Behrend [7] showed that there is a constant $c > 0$ so that $\mu_{\mathcal{L}}(n) \geq n \exp(-c\sqrt{\log n})$. See [15,19] for the best known lower bound on $\mu_{\mathcal{L}}(n)$ in this case.

More generally, it is known that $\mu_{\mathcal{L}}(n) = o(n)$ if $\mathcal{L}$ is translation-invariant and $\mu_{\mathcal{L}}(n) = \Omega(n)$ otherwise (see [28]). For other (exact) bounds on $\mu_{\mathcal{L}}(n)$ for various linear equations $\mathcal{L}$ see, for example, [28,29,6,14,21].

In this paper we mainly focus on $\mathcal{L}$-free subsets of $[n]$ for linear equations $\mathcal{L}$ of the form $px+qy = z$ where $p \geq 2$ and $q \geq 1$ are fixed integers. Notice that for such a linear equation $\mathcal{L}$, the interval $[\lfloor n/(p+q) \rfloor + 1, n]$ is an $\mathcal{L}$-free set. Our first result implies that this is the largest such $\mathcal{L}$-free subset of $[n]$. Let $\min(S)$ denote the smallest element in a finite set $S \subseteq \mathbb{N}$.

Theorem 1. *Let $\mathcal{L}$ denote the equation $px + qy = z$ where $p \geq q$ and $p \geq 2$, $p, q \in \mathbb{N}$. Let S be an $\mathcal{L}$-free subset of $[n]$, and let $\min(S) = \lfloor \frac{n}{p+q} \rfloor - t$ where t is a non-negative integer.*

(i) *If $0 \leq t < (\frac{p+q-1}{p+q+p/q})\lfloor \frac{n}{p+q} \rfloor$ then $|S| \leq \lceil \frac{(p+q-1)n}{p+q} \rceil - \lfloor \frac{p}{q} t \rfloor$.*
(ii) *If $t \geq (\frac{p+q-1}{p+q+p/q})\lfloor \frac{n}{p+q} \rfloor$ then $|S| \leq \frac{(q^2+1)n}{q^2+q+1}$ provided that*

$$n \geq \max\left\{ \frac{3(q^2+q+1)(q^3+p(q^2+q+1))}{q^2+1}, \frac{5(q^2+q+1)(q^5+p(q^4+q^3+q^2+q+1))}{q^4+(p-1)q^3+q^2+1} \right\}.$$

In both cases of Theorem 1 we observe that $|S| \leq n - \lfloor \frac{n}{p+q} \rfloor$, hence the following corollary holds.

Please cite this article in press as: R. Hancock, A. Treglown, On solution-free sets of integers, European Journal of Combinatorics (2017), http://dx.doi.org/10.1016/j.ejc.2017.06.018.