A Note on Intervals in the Hales-Jewett Theorem

Imre Leader Eero Räty

Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WB, U.K.

i.leader@dpmms.cam.ac.uk, epjr2@cam.ac.uk

Submitted: Feb 9, 2018; Accepted: May 16, 2018; Published: Jul 27, 2018

Abstract

The Hales-Jewett theorem for alphabet of size 3 states that whenever the Hales-Jewett cube $[3]^n$ is r-coloured there is a monochromatic line (for n large). Conlon and Kamcev conjectured that, for any n, there is a 2-colouring of $[3]^n$ for which there is no monochromatic line whose active coordinate set is an interval. In this note we disprove this conjecture.

Mathematics Subject Classifications: 05D10

1 Introduction

In order to state the Hales-Jewett theorem we need some notation. Given positive integers k and n let $[k]^n$ be the set of all words in symbols $\{1, \dots, k\}$ of length n. A set $L \subset [k]^n$ is called a *combinatorial line* if there exist a nonempty set $S \subset [n]$ and integers $a_i \in [k]$ for all $i \notin S$ such that $L = \{(x_1, \dots, x_n) : x_i = a_i \text{ for all } i \notin S, x_i = x_j \text{ for all } i, j \in S\}$. The set S is called the set of *active coordinates* of L.

Theorem 1. *(Hales-Jewett [2]). For any k and r there exists n such that whenever $[k]^n$ is r-coloured there is a monochromatic combinatorial line.*

As noted by Conlon and Kamcev in [1], by following Shelah's proof [3] of the Hales-Jewett theorem it can be shown that one can always find a monochromatic combinatorial line whose active coordinate set S is an union of at most $HJ(k-1, r)$ intervals, where $HJ(k-1, r)$ is the smallest integer n for which the Hales-Jewett theorem holds for $k-1$ and r.

In the case $k = 3$, since $HJ(2, r) = r$, this says that one can always find a monochromatic line whose active coordinate set is a union of at most r intervals. Conlon and Kamcev proved in [1] that this bound is tight for r odd: in other words, they showed that for each odd r there is an r-colouring of $[3]^n$ (for any n) for which every monochromatic line has active coordinate set made up of at least r intervals. They conjectured that this would also be the case for r even. In particular, for $r = 2$, they conjectured that for all n there exists a 2-colouring of $[3]^n$ for which there exists no monochromatic combinatorial line whose active coordinate set is an interval.

In this note we will prove that, perhaps surprisingly, their conjecture is false. This can be stated in the following form.

Theorem 2. *For all sufficiently large n, whenever $[3]^n$ is 2-coloured there exists a monochromatic combinatorial line whose active coordinate set is an interval.*

2 The proof of Theorem 2

The idea of the proof is as follows. By applying Ramsey's theorem, we will pass to a subspace on which the colour of a word depends only on its 'pattern' of intervals, and not on its 'breakpoints' (the places where the word changes from one letter to another). Once this is done, we can consider some particular small patterns.

For a word w let $\overline{w}$ be obtained from w by contracting every interval on which w is a constant to a single letter. We will consider the particular words $s_1 = 132$, $s_2 = 1232$, $s_3 = 1312$, $s_4 = 13232$ and $s_5 = 13132$. Set t_i to be the length of the word s_i. Put $n_0 = 4$ and for $1 \leqslant i \leqslant 5$ let $n_i = R^{(t_i-1)}(n_{i-1})$ where $R^{(t)}(s) = R^{(t)}(s,s)$ is the t-set Ramsey-number. Finally set $n = n_5 + 1$.

Let c be any 2-colouring of $[3]^n$. For a word w we define the set of *breakpoints* $T(w)$ by $T(w) = \{a_1, \ldots, a_m\}$ if $w_{a_{i-1}+1} = \cdots = w_{a_i}$ and $w_{a_i} \neq w_{a_i+1}$ for all $1 \leqslant i \leqslant m+1$, with the convention $a_0 = 0$, $a_{m+1} = n$. For example, $w = 1122333111$ has breakpoints $T(w) = \{2, 4, 7\}$.

Let s be a sequence of length t and $T_1 = \{a_1, \ldots, a_m\} \subset [n-1]$ with $|T_1| = t - 1$. We say that $w \in [3]^n$ has breakpoints in T_1 with *pattern* s if $T(w) = T_1$ and $\overline{w} = s$. For example $w = 1122333111$ has breakpoints $T(w) = \{2, 4, 7\}$ with pattern $s = 1231$. Note that if $\overline{w} = s$ then there exists a unique set T_1 of size $|s| - 1$ for which w has breakpoints T_1 with pattern s.

Set $T_5 = [n-1]$. Suppose that $|T_i| \geqslant n_i$ is given, and recall that t_i is the length of the word s_i defined at the start of the proof. For all $A \in [n-1]^{(t_i-1)}$ define w^A to be the unique sequence which has breakpoints A with pattern s_i.

Now c induces a 2-colouring c_i on the set $T_i^{(t_i-1)}$ given by $c_i(A) = c(w^A)$. Hence by Ramsey's theorem and the choice of n_i's it follows that there exists $T_{i-1} \subset T_i$ with $|T_{i-1}| \geqslant n_{i-1}$ such that $T_{i-1}^{(t_i-1)}$ is monochromatic for the colouring c_i, say with colour d_i.

Thus we obtain sets $T_0 \subset T_1 \subset \cdots \subset T_5$ with $|T_0| \geqslant 4$ and colours $d_1 \ldots, d_5$ such that c_i restricted to $T_i^{(t_{i+1}-1)}$ is constant with value d_{i+1}. Note that it is impossible to choose colours $d_1, \ldots, d_5$ without at least one of the following sets

$$N_1 = \{d_1, d_2\}$$
$$N_2 = \{d_1, d_3\}$$
$$N_3 = \{d_2, d_4\}$$

$$N_4 = \{d_3, d_5\}$$
$$N_5 = \{d_1, d_4, d_5\}$$

having just one element (i.e. all colours being equal). Indeed, if all the sets N_1, N_2, N_3, N_4 contain both colours we must have $d_2 = d_3$ and $d_1 = d_4 = d_5$, which implies that $|N_5| = 1$.

Let $a_1 < a_2 < a_3 < a_4$ be elements of T_0. We will use the shorthand $w = [b_1b_2b_3b_4b_5]$ for the word which has $w_i = b_j$ for all $a_{j-1} < i \leqslant a_j$ with the convention $a_0 = 0$ and $a_5 = n$. Note that we will allow $b_i = b_{i+1}$. Hence $T(w) \subset \{a_1, \ldots, a_4\} \subset T_0$ and $\overline{w} = \overline{b_1b_2b_3b_4b_5}$.

Set $w_1 = [13332]$, $w_2 = [12232]$, $w_3 = [13112]$, $w_4 = [13232]$ and $w_5 = [13132]$. It is easy to verify that for all i we have $\overline{w_i} = s_i$ and also $T(w_i) \subset T_0 \subset T_{i-1}$. Furthermore set $v_1 = [11132]$, $v_2 = [11232]$, $v_3 = [13122]$ and $u_1 = [13222]$. As before it is easy to verify that $\overline{v_i} = s_i$, $\overline{u_1} = s_1$, and by construction $T(v_i) \subset T_0 \subset T_{i-1}$ and $T(u_1) \subset T_0$. Thus by definition of the sets T_{i-1} it follows that $c(w_i) = d_i$, $c(v_i) = d_i$ and $c(u_1) = d_1$.

It is straightforward to verify that

- v_1, w_2, w_1 forms a combinatorial line L_1 with $S_1 = \{a_1 + 1, \ldots, a_3\}$
- w_3, u_1, w_1 forms a combinatorial line L_2 with $S_2 = \{a_2 + 1, \ldots, a_4\}$
- v_2, w_2, w_4 forms a combinatorial line L_3 with $S_3 = \{a_1 + 1, \ldots, a_2\}$
- w_3, v_3, w_5 forms a combinatorial line L_4 with $S_4 = \{a_3 + 1, \ldots, a_4\}$
- w_5, w_4, w_1 forms a combinatorial line L_5 with $S_5 = \{a_2 + 1, \ldots, a_3\}$

It is clear that the colours used to colour elements of the line L_i are exactly the colours in the set N_i. As observed earlier, one of the sets N_i contains only one colour, which implies that the associated line L_i is monochromatic. Since all the sets S_i are intervals, this completes the proof.

References

[1] D. Conlon and N. Kamcev, Intervals in the Hales-Jewett theorem, arXiv:1801.08919.

[2] A. W. Hales and R. I. Jewett, Regularity and positional games, Trans. Amer. Math. Soc. **106** (1963), 222–229.

[3] S. Shelah, Primitive recursive bounds for van der Waerden numbers, J. Amer. Math. Soc. **1** (1988), 683–697.