Another Note on Intervals in the Hales–Jewett Theorem

Nina Kamčev[1]

Department of Mathematics
University of Zagreb
Zagreb, Croatia.

nina.kamcev@math.hr

Christoph Spiegel[2]

AI in Society, Science, and Technology
Zuse Institute Berlin
Berlin, Germany.

spiegel@zib.de

Submitted: Feb 7, 2020; Accepted: Feb 18, 2022; Published: Mar 25, 2022

Abstract

The Hales–Jewett Theorem states that any r–colouring of $[m]^n$ contains a monochromatic combinatorial line if n is large enough. Shelah's proof of the theorem implies that for $m = 3$ there always exists a monochromatic combinatorial line whose set of active coordinates is the union of at most r intervals. For odd r, Conlon and Kamčev constructed r–colourings for which it cannot be fewer than r intervals. However, we show that for even r and large n, any r–colouring of $[3]^n$ contains a monochromatic combinatorial line whose set of active coordinates is the union of at most $r-1$ intervals. This is optimal and extends a result of Leader and Räty for $r = 2$.

Mathematics Subject Classifications: 05D10

1 Introduction

The Hales–Jewett theorem is a cornerstone of Ramsey theory from which many results can be derived, most notably van der Waerden's Theorem. In order to state the theorem we will need to introduce some notation.

Given positive integers m and n, let $[m]^n$ be the collection of all *words* of length n with letters taken from the alphabet $[m] = \{1, \dots, m\}$. We write $[m]^n_\star = \big([m] \cup \{\star\}\big)^n \setminus [m]^n$ and refer to the coordinates in $\boldsymbol{\ell} \in [m]^n_\star$ where the symbol $\star$ occurs as *active*. Let $\boldsymbol{\ell}[\alpha]$ denote the word in $[m]^n$ obtained by substituting each occurrence of the symbol $\star$ in $\boldsymbol{\ell}$ by $\alpha \in [m]$. The set of m words $\{\boldsymbol{\ell}[1], \dots, \boldsymbol{\ell}[m]\} \subset [m]^n$ is referred to as a *combinatorial line* in $[m]^n$, and we will abbreviate it by $\boldsymbol{\ell}$.

[1]Visit to Universitat Politècnica de Catalunya supported by the grant TM2017-82166-P9.

[2]This research was conducted while at the Universitat Politècnica de Catalunya and Barcelona Graduate School of Mathematics and was supported by a Spanish Ministerio de Economía y Competitividad FPI grant under the project MTM2014-54745-P and the María de Maetzu research grant MDM-2014-0445.

 https://doi.org/10.37236/9400

Theorem 1 (Hales–Jewett [14]). *For any positive integers m and r there exists a positive integer n such that any r–colouring of $[m]^n$ contains a monochromatic combinatorial line.*

The smallest such n is called the *Hales–Jewett number* and denoted by $HJ(m, r)$.

We will be imposing additional structural requirements on the desired monochromatic objects, a direction which is often pursued in Ramsey theory. Shelah's celebrated proof of the Hales–Jewett Theorem [20] uses a single induction (on m) and therefore gives primitive recursive bounds for $HJ(m, r)$. Very recently Golshani and Shelah [12] proposed an alternative proof obtaining the same class of bounds. The proof in [20] yields monochromatic combinatorial lines with a specific structure, which has drawn the attention of several researchers. To discuss these results, we describe a combinatorial line as *q–fold* if the set of its active coordinates consists of at most q sub-intervals of $\{1, \dots, n\}$ for some positive integer q. Shelah's argument implies that for sufficiently large n, any r–colouring of $[m]^n$ contains a monochromatic q–fold line with $q \leqslant HJ(m-1, r)$. In particular, when the alphabet size is $m = 3$, we get a monochromatic r–fold line in $[3]^n$. The reason for this is the following: the r intervals are constructed using a pigeonhole-principle argument establishing that the symbols 1 and 2 are interchangeable on those intervals, which effectively reduces the size of the alphabet by one. Conlon and the first author were interested in this additional structure given by the proof as a clue towards understanding the optimality of Shelah's approach. They showed that for odd r and all n, there are r–colourings of $[3]^n$ containing no monochromatic $(r-1)$–fold lines. In other words, Shelah's proof yields monochromatic lines in $[3]^n$ with the simplest possible active set in this sense.

Perhaps surprisingly, Leader and Räty [17] showed that the restriction on the parity of r is necessary by proving that for sufficiently large n, any 2–colouring of $[3]^n$ contains a monochromatic 1–fold line. We show that this case is not an exception by extending their result to any even number of colours.

Theorem 2. *For any even integer $r \geqslant 2$ there exists a positive integer $N = N(r)$ such that any r–colouring of $[3]^N$ contains a monochromatic combinatorial line whose set of active coordinates is a union of at most $r-1$ intervals.*

The colouring from [7] shows that this result is optimal. Given the scarcity of quantitative results in this area of Ramsey theory, it is interesting that we can pin down an exact answer for all r. The fact that the minimal number of intervals depends on the parity of r could also come as a surprise, seeing as the Hales–Jewett Theorem is purely combinatorial.

In the context of van der Waerden's theorem, a number of researchers have looked at analogous questions, where the family of arithmetic progressions is restricted to a specific subclass, usually by putting restrictions on the common difference [2, 5, 9].

The proof of Theorem 2 relies on reducing the problem of finding $(r-1)$–fold lines in arbitrary colourings to considering specific colourings with a certain arithmetic structure. Interestingly, the colourings found in [7] have precisely this structure. The Hales–Jewett Theorem is naturally phrased in terms of proper colourings of a 3-uniform hypergraph, but our reduction turns it largely into a graph colouring problem. Colourings of this type occur elsewhere in literature, e.g. in showing that the Hales–Jewett theorem implies van der Waerden's Theorem.

Structure of the paper. We start Section 2 by introducing some necessary notation and then giving an outline of the proof of Theorem 2. The individual steps of this outline will be executed in Subsections 2.1 to 2.5. Section 3 contains some further remarks and open questions.

2 Proof of Theorem 2

We start this section by fixing some notation and giving an outline of the proof of Theorem 2. Given a word $\mathbf{w} = w_1 \dots w_n \in [m]^n$, let its *contraction* $\overline{\mathbf{w}}$ be obtained by contracting every interval on which $\mathbf{w}$ is constant to a single letter of $[m]$. The word $\mathbf{w} = 11233322$, for example, has the contraction $\overline{\mathbf{w}} = 1232$. We write $\mathcal{P}(m,n) = \{\overline{\mathbf{w}} : \mathbf{w} \in [m]^n\}$ for the set of *patterns* on the alphabet $[m]$ of length at most n. Note that this notion also plays an important role in [17] and was in fact previously introduced by Furstenberg and Katznelson in [10].

Given some pattern $\mathbf{p} = p_1 \dots p_k \in \mathcal{P}(m,n)$ and $i \in [m]$, we also use the notation

$$\varphi_i(\mathbf{p}) = \#\{1 \leqslant j \leqslant k : p_j = i\} \quad \text{and} \quad \varphi(\mathbf{p}) = \big(\varphi_1(\mathbf{p}), \dots, \varphi_m(\mathbf{p})\big).$$

We refer to $\varphi(\mathbf{p})$ as the *count* of $\mathbf{p}$. Lastly, we write

$$\varphi_i^{(q+1)}(\mathbf{p}) = \varphi_i(\mathbf{p}) \pmod{q+1} \quad \text{and} \quad \varphi^{(q+1)}(\mathbf{p}) = \big(\varphi_1^{(q+1)}(\mathbf{p}), \dots, \varphi_m^{(q+1)}(\mathbf{p})\big)$$

and refer to $\varphi^{(q+1)}(\mathbf{p}) \in \mathbb{Z}_{q+1}^m$ as the *reduced count* of $\mathbf{p}$. From now on, the number of colours will be equal to $q+1$, that is

$$r = q + 1.$$

The goal will therefore be to find a monochromatic q–fold combinatorial line. The $(q+1)$–colouring constructed by Conlon and the first author is in fact a function of $\varphi^{(q+1)}(\overline{\mathbf{w}})$. We will show that colourings of this type are inherent to the problem by passing precisely from any colouring of $[3]^n$ to a function of $\varphi^{(q+1)}(\overline{\mathbf{w}})$.

Given some pattern $\mathbf{p} \in \mathcal{P}(3,n)$ and $k \in \mathbb{N}$ (the set of positive integers), the notation

$$(\mathbf{p})^k = \overbrace{\mathbf{p}\mathbf{p}\dots\mathbf{p}}^{k \text{ times}}$$

refers to the k–fold repetition of that pattern. For the rest of the section, we fix

$$k_0 = k_0(q) = 18q + 12 \tag{1}$$

and define the *buffered version* of a pattern $\mathbf{x} = x_1 \dots x_{k'} \in \mathcal{P}(3, k_0)$ satisfying $x_1 \neq 1$ and $x_{k'} \neq 2$ to be

$$\mathbf{x}^+ = 1\,\mathbf{x}\,(23)^{2k_0}\,(13)^{2k_0}\,(21)^{2k_0}\,231 \;\in\; \mathcal{P}(3, 13k_0 + 4).$$

The buffer will help us circumvent the anomaly that certain patterns (e.g. $\mathbf{x} = 3$) are contained in images of very few combinatorial lines under the contraction map. In contrast, $\mathbf{x}^+$ can be found in images of many combinatorial lines.

Using these definitions, we define the following three hypergraphs for any $m \geqslant 2$.

- $\mathbf{H}(m,n,q)$ refers to the m-uniform hypergraph with vertex set $[m]^n$ and edge set consisting of all q–fold combinatorial lines in $[m]^n$.
- $\mathbf{P}(m,n,q)$ refers to the hypergraph obtained from $\mathbf{H}(m,n,q)$ by identifying vertices whose contractions coincide and keeping the edges. Each vertex of $\mathbf{H}(m,n,q)$ is mapped to its contraction in $\mathbf{P}(m,n,q)$ and after this mapping, the edges in this hypergraph are of order $m-1$ or m.
- $\mathbf{C}(n,q)$ refers to the hypergraph on $\mathbb{Z}_{q+1}^3$ obtained by taking the sub-hypergraph of $\mathbf{P}(3,n,q)$ induced by the set of vertices $\mathbf{p} \in \mathcal{P}(3, 13k_0+4+q)$ that are buffered, that is $\mathbf{p} = \mathbf{x}^+$ for some appropriate $\mathbf{x} \in \mathcal{P}(3,k_0)$, and then identifying and labelling vertices based on their reduced count $\varphi^{(q+1)}(\mathbf{x})$. That is, an edge $\{\mathbf{x}^+, \mathbf{u}^+, \mathbf{v}^+\}$ in $\mathbf{P}(3,n,q)$ induces the edge $\{\varphi^{(q+1)}(\mathbf{x}), \varphi^{(q+1)}(\mathbf{u}), \varphi^{(q+1)}(\mathbf{v})\}$ in $\mathbf{C}(n,q)$, assuming $\mathbf{x}^+, \mathbf{u}^+$ and $\mathbf{v}^+$ are of length at most $13k_0+4+q$. Since we are restricting ourselves to the case of $m=3$ for the definition of $\mathbf{C}(n,q)$, all edges in this hypergraph are of order two or three.

Note that the first of these hypergraphs is m-uniform and the other two contain edges of order m as well as $m-1$. We will emphasise the order of an edge whenever relevant. Furthermore, whenever we refer to graph theoretic objects or terminology such as cliques or two adjacent vertices, we are referring only to edges of order two.

We can now describe the idea of the proof as follows: Any colouring of $[3]^n$ that avoids monochromatic q–fold combinatorial lines simply corresponds to a proper vertex-colouring of $\mathbf{H}(3,n,q)$, i.e. a colouring with no monochromatic edges. Using a purely Ramsey-theoretic argument, we first show that for any such colouring and n large enough we can find a sub-hypergraph isomorphic to the set of q–fold combinatorial lines in $[3]^{\tilde{n}}$, where $\tilde{n}$ is significantly smaller than n, with the following important property: any two words in this sub-hypergraph that have the same contraction must also have the same colour. We can therefore identify all the words in this sub-hypergraph that get contracted to the same pattern, meaning we are now considering colourings of $\mathbf{P}(3,\tilde{n},q)$.

We continue by showing that this sub-hypergraph has a rich structure. For instance, the line $\boldsymbol{\ell} = 1{*}{*}22{*}11$ induces the edge $\{121, 13231\}$. More generally, $\mathbf{P}(3,\tilde{n},q)$ contains many interlaced cliques of size $q+1$, i.e. q–powers of arbitrarily long paths. Besides establishing that any proper colouring of the original hypergraph requires at least $q+1$ colours, this structure will imply that, within a significant part of our sub-hypergraph, we can identify patterns with each other if they have the same reduced count $\varphi^{(q+1)}$. This further reduces the problem to colourings of $\mathbf{C}(\tilde{n},q)$. This hypergraph is translation-invariant, has edges between any two vertices which differ in a single coordinate as well as some important additional restrictions. These restrictions imply that it cannot be $(q+1)$–colourable for odd q, from which the main theorem follows. Only edges of order two in $\mathbf{P}(3,\tilde{n},q)$ will be relevant to make the step to $\mathbf{C}(\tilde{n},q)$, but those of order three are crucial for ultimately deriving the lower bound on the chromatic number.

We note that everything up to the bound on the chromatic number of $\mathbf{C}(\tilde{n},q)$ holds for general q (odd or even) with $r=q+1$, and that the initial Ramsey theoretic reduction to the patterns also holds for general m. Therefore, our proof along with the colouring from [7] gives a good intuition on why the function $\mathcal{I}(3,q+1)$ displays the alternating behaviour depending on the parity of q.

In the remainder of this section, we first show in Subsection 2.1 that if $\mathbf{P}(m,n,q)$ is not $(q+1)$–colourable, then neither is $\mathbf{H}(m,N,q)$ for some appropriately large N. Then we assume that $n \geqslant 13k_0+4+q$ and show in Subsection 2.3 that if $\mathbf{C}(n,q)$ is not $(q+1)$–colourable, then neither is $\mathbf{P}(3,n,q)$. We conclude the proof of Theorem 2 in Subsection 2.5 by showing that any $(q+1)$–colouring of $\mathbf{C}(n,q)$ must contain a monochromatic edge.

2.1 From $\mathbf{H}(m,n,q)$ to $\mathbf{P}(m,n,q)$ – reduction to patterns

The notation and idea behind this part are derived from the approach of Leader and Räty [17] for the specific case of $q=1$.

We define the set of *breakpoints* of a given word $\mathbf{w}=w_1\dots w_n \in [m]^n$ to be the set $T(\mathbf{w})=\{a_1,\dots,a_k\}$ for which $w_{a_{i-1}+1}=\cdots=w_{a_i}$ and $w_{a_i}\neq w_{a_i+1}$ for $1\leqslant i\leqslant k+1$ where we set $a_0=0$ and $a_{k+1}=n$. Let $S^{(k)}$ refer to all subsets of size k of some given set S. Given some $N\geqslant n$ and $A=\{a_1<\cdots<a_{n-1}\}\in[N-1]^{(n-1)}$, let $\mathbf{w}^A$ denote the word $\mathbf{w}^A=w_1^A\dots w_N^A\in[m]^N$ defined by $w^A_{a_{i-1}+1}=\cdots=w^A_{a_i}=w_i$ for $1\leqslant i\leqslant n$ where we set $a_0=0$ and $a_n=N$. Note that in general $\overline{\mathbf{w}^A}=\overline{\mathbf{w}}$ and $T(\mathbf{w}^A)\subseteq A$. Specifically $T(w^A)=A$ if and only if $\overline{\mathbf{w}}=\mathbf{w}$.

Example 3. Let $\mathbf{w}=13323$ be given. We have $T(\mathbf{w})=\{1,3,4\}$. If $A=\{2,3,5,6\}$ and $N=8$ then $\mathbf{w}^A=11\,3\,33\,2\,33$.

This notation allows us to make the following statement.

Proposition 4. *For any $n,r\in\mathbb{N}$ there exists $N=N(n,r)\in\mathbb{N}$ so that for any r–colouring χ of $[m]^N$ there exists $A=A(n,N,\chi)\in[N-1]^{(n-1)}$ such that $\chi(\mathbf{w}_1^A)=\chi(\mathbf{w}_2^A)$ for any $\mathbf{w}_1,\mathbf{w}_2\in[m]^n$ satisfying $\overline{\mathbf{w}_1}=\overline{\mathbf{w}_2}$.*

Proof. Give the patterns in $\mathcal{P}(m,n)$ an arbitrary ordering, say $\mathcal{P}(m,n)=\{\mathbf{p}_1,\dots,\mathbf{p}_k\}$, and write $t_i=|\mathbf{p}_i|$ for their respective length. Set $n_0=n-1$ and recursively define $n_i=R^{(t_i-1,r)}(n_{i-1})$ for $1\leqslant i\leqslant k$ where $R^{(t,r)}(s)=R^{(t,r)}(s,\dots,s)$ is the r–colour Ramsey number for t-sets. Lastly, set $N=N(n,r)=n_k+1$.

Now let χ be an arbitrary but fixed r–colouring of $[m]^N$ and let us recursively define sets $T_k\supset\cdots\supset T_1\supset T_0$ satisfying $|T_i|\geqslant n_i$ for $0\leqslant i\leqslant k$ as well as certain properties with respect to that colouring. We start by setting $T_k=[N-1]$. Let $|T_i|\geqslant n_i$ be given and observe that χ induces a colouring χ_i on $T_i^{(t_i-1)}$ given by $\chi_i(A)=\chi(\mathbf{p}_i^A)$ for $A\in T_i^{(t_i-1)}$. Now, since by definition $n_i=R^{(t_i-1,r)}(n_{i-1})$, it follows that there exists $T_{i-1}\subset T_i$ satisfying $|T_{i-1}|\geqslant n_{i-1}$ such that $T_{i-1}^{(t_i-1)}$ is monochromatic with respect to χ_i.

Now fix some $A=\{a_1<\cdots<a_{n-1}\}\in T_0^{(n-1)}$. Consider two words $\mathbf{w}_1,\mathbf{w}_2\in[m]^n$ which satisfy $\overline{\mathbf{w}_1}=\overline{\mathbf{w}_2}=\mathbf{p}_j$ for some $1\leqslant j\leqslant k$. We note that there exist $A_1,A_2\in A^{(t_j-1)}$ such that $\mathbf{w}_1^A=\mathbf{p}_j^{A_1}$ and $\mathbf{w}_2^A=\mathbf{p}_j^{A_2}$, that is $T(\mathbf{w}_1^A)=A_1$ and $T(\mathbf{w}_2^A)=A_2$. Since $T_0\subset T_{j-1}$, we have $A_1,A_2\in T_{j-1}^{(t_j-1)}$ and therefore

$$\chi(\mathbf{w}_1^A)=\chi(\mathbf{p}_j^{A_1})=\chi_j(A_1)=\chi_j(A_2)=\chi(\mathbf{p}_j^{A_2})=\chi(\mathbf{w}_2^A)$$

as desired. □

The lemma states that within $[m]^N$ we can find a 'copy' of $[m]^n$ in which any two words with the same contraction must also have the same colour. The following corollary captures this point.

Corollary 5. *For any $n, r \in \mathbb{N}$ there exists $N = N(n,r) \in \mathbb{N}$ such that if $\mathbf{P}(m,n,q)$ is not r–colourable, then neither is $\mathbf{H}(m,N,q)$.*

Proof. Given a proper r–colouring χ of the vertex set $[m]^N$ of $\mathbf{H}(m,N,q)$, we note that χ' given by $\chi'(\mathbf{w}) = \chi(\mathbf{w}^A)$ for $\mathbf{w} \in [m]^n$ and A as given by Proposition 4 defines a proper colouring of $[m]^n$. This follows since any combinatorial line $\boldsymbol{\ell}$ in $[m]^n$ consisting of $\{\boldsymbol{\ell}[1], \ldots, \boldsymbol{\ell}[m]\}$ corresponds to the combinatorial line $\boldsymbol{\ell}^A$ in $[m]^N$ consisting of $\{\boldsymbol{\ell}[1]^A, \ldots, \boldsymbol{\ell}[m]^A\}$. Proposition 4 now implies that χ' also induces a proper colouring of $\mathbf{P}(m,n,q)$, proving the statement. □

2.2 The structure of edges in $\mathbf{P}(m,n,q)$

Before we proceed with the reduction to the reduced count, let us describe the structure of the edges in $\mathbf{P}(m,n,q)$. Let us write $\mathcal{P}_\star(m,n) = \{\overline{\boldsymbol{\ell}} : \boldsymbol{\ell} \in [m]^n_\star\}$ where the contraction also contracts repeated occurrences of the symbol $\star$. We start with a simple observation which follows from the definition of $\mathbf{P}(m,n,q)$.

Lemma 6. *For any $\boldsymbol{\ell} \in \mathcal{P}_\star(m,n)$, the set $\left\{\overline{\boldsymbol{\ell}[1]}, \ldots, \overline{\boldsymbol{\ell}[m]}\right\}$ forms an edge in $\mathbf{P}(m,n,q)$.*

Proof. Let $\boldsymbol{\ell} = \ell_1 \ldots \ell_k$ where $k \leqslant n$. Clearly $\boldsymbol{\ell}' = \ell_1 \ldots \ell_k (\ell_k)^{n-k} \in [m]^n_\star$ is a combinatorial line in $[m]^n$ so that $\{\overline{\boldsymbol{\ell}[1]}, \overline{\boldsymbol{\ell}[2]}, \ldots, \overline{\boldsymbol{\ell}[m]}\} = \{\overline{\boldsymbol{\ell}'[1]}, \overline{\boldsymbol{\ell}'[2]}, \ldots, \overline{\boldsymbol{\ell}'[m]}\}$ forms an edge in $\mathbf{P}(m,n,q)$. □

In general, the edges described in the previous lemma are of order m or $m-1$. The central observation used for the next reduction is that for the case of $m = 3$ a combinatorial line of the form $\boldsymbol{\ell} = 1 \star 2$ *connects* the patterns 12 and 132 by an edge of order two, since both $\boldsymbol{\ell}[1] = 112$ and $\boldsymbol{\ell}[2] = 122$ get contracted to the same pattern. This is a particularity of that alphabet order and the main reason why this approach does not easily extend to larger m.

Let us derive a precise description of when an edge of order two occurs between two vertices in $\mathbf{P}(3,n,q)$. Edges of order two will be sufficient in realising the reduction to the reduced count, but edges of order three will be crucial at the end of the section when establishing the lower bound on the chromatic number for odd q.

We start by introducing two more necessary notions. Let $\{\alpha_1, \alpha_2, \alpha_3\} = \{1,2,3\}$ and $\mathbf{p} \in \mathcal{P}(3,n)$. An α_3*–insertion* in $\mathbf{p}$ is the operation of inserting a copy of the letter α_3 between an instance of α_1 and an instance of α_2 in $\mathbf{p}$. An α_3*–alteration* of $\mathbf{p}$ is the operation of moving one instance of α_3 whose neighbours in $\mathbf{p}$ are α_1 and α_2 to another part of $\mathbf{p}$ so that its neighbouring letters are again α_1 and α_2.

Example 7. The pattern 13212 can be obtained from 1212 by a 3–insertion and from 12312 by a 3–alteration.

Only the notion of insertion will be needed for the remainder of this subsection, but alterations will become important in the next one.

Lemma 8. *Let $\mathbf{p}_1, \mathbf{p}_2$ be two patterns in $\mathcal{P}(3,n)$ and $\alpha \in [3]$. If $\mathbf{p}_2$ is obtained from $\mathbf{p}_1$ by at most q successive α–insertions, then $\mathbf{p}_1$ and $\mathbf{p}_2$ are adjacent in $\mathbf{P}(3,n,q)$.*

Proof. Assume without loss of generality that $\alpha = 3$. We will construct some $\boldsymbol{\ell} \in \mathcal{P}_\star(3,n)$ satisfying $\overline{\boldsymbol{\ell}[1]} = \overline{\boldsymbol{\ell}[2]} = \mathbf{p}_1$ and $\overline{\boldsymbol{\ell}[3]} = \mathbf{p}_2$. Let $\mathbf{p}_2 = p_1 \ldots p_k$ and let $j_1, \ldots, j_{k'}$ denote the $k' \leqslant q$ indices of the 3–insertions that take one from $\mathbf{p}_1$ to $\mathbf{p}_2$, that is if one removes $p_{j_1}, \ldots, p_{j_{k'}}$ from $\mathbf{p}_2$ then one obtains $\mathbf{p}_1$. We now define $\boldsymbol{\ell} = \ell_1 \ldots \ell_k$ by

$$\ell_i = \begin{cases} p_i & \text{for } i \in \{1, \ldots, k\} \setminus \{j_1, \ldots, j_{k'}\}, \\ \star & \text{for } i \in \{j_1, \ldots, j_{k'}\}. \end{cases}$$

It immediately follows that $\overline{\boldsymbol{\ell}[1]} = \overline{\boldsymbol{\ell}[2]} = \mathbf{p}_1$ and $\overline{\boldsymbol{\ell}[3]} = \mathbf{p}_2$, so by Lemma 6 $\boldsymbol{\ell}$ forms an edge between $\mathbf{p}_1$ and $\mathbf{p}_2$ in $\mathbf{P}(3,n,q)$. □

2.3 From $\mathbf{P}(3,n,q)$ to $\mathbf{C}(n,q)$ – Reduction to the reduced count

Let us introduce one last definition. For $\{\alpha_1, \alpha_2, \alpha_3\} = \{1,2,3\}$ we call a pattern in $\mathcal{P}(3,n)$ α_3*–diverse* if it is of length at most $n-q$ and it contains at least q copies of either of the subwords $\alpha_1\alpha_2$ or $\alpha_2\alpha_1$.

Example 9. The pattern 121 is 3–diverse if $q \leqslant 2$ and $n \geqslant 5$

Note that we are not yet restricting ourselves to buffered patterns for the following remark and the subsequent lemma.

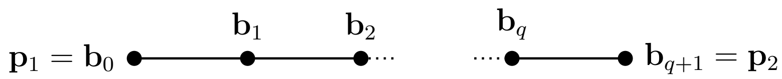
Remark 10. For any given α–diverse pattern $\mathbf{p} \in \mathcal{P}(3,n)$ there exists a sequence $\mathbf{p} = \mathbf{b}_1, \ldots, \mathbf{b}_{q+1}$ in $\mathcal{P}(3,n)$ so that $\mathbf{b}_{i+1}$ can be obtained from $\mathbf{b}_i$ by an α–insertion for $1 \leqslant i \leqslant q$.

Lemma 11. *Let χ be a proper $(q+1)$–colouring of $\mathbf{P}(3,n,q)$, $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}(3,n)$ and $\alpha \in \{1,2,3\}$. We have $\chi(\mathbf{p}_1) = \chi(\mathbf{p}_2)$ if either of the following two cases holds:*

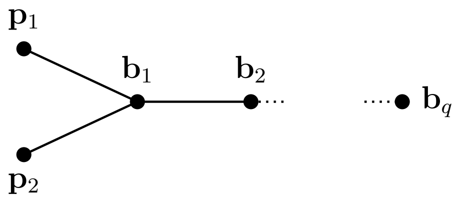
(i) $\mathbf{p}_2$ can be obtained from $\mathbf{p}_1$ by exactly $q+1$ α–insertions,

(ii) $\mathbf{p}_1$ and $\mathbf{p}_2$ are α–diverse patterns and $\mathbf{p}_2$ can be obtained from $\mathbf{p}_1$ by an α–alteration.

Proof. Let us start with case (i). By assumption, there exists a sequence of patterns $\mathbf{p}_1 = \mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_q, \mathbf{b}_{q+1} = \mathbf{p}_2$ in $\mathcal{P}(3,n)$ so that $\mathbf{b}_{i+1}$ is obtained from $\mathbf{b}_i$ by an α–insertion for any $0 \leqslant i \leqslant q$. By Lemma 8, there is an edge of order two connecting $\mathbf{b}_i$ to $\mathbf{b}_j$ if $|i-j| \leqslant q$ and $i \neq j$. It follows that $\{\mathbf{b}_1, \ldots, \mathbf{b}_q\}$ forms a clique of order q that lies in the neighbourhood of both $\mathbf{p}_1$ and $\mathbf{p}_2$. Since χ uses $q+1$ colours, the desired statement follows.



Regarding case (ii), one can easily see that since $\mathbf{p}_2$ can be obtained from $\mathbf{p}_1$ by an α–alteration and $\mathbf{p}_1$ and $\mathbf{p}_2$ are α–diverse, there exist patterns $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \ldots, \mathbf{b}_q$ in $\mathcal{P}(3,n)$ such that $\mathbf{b}_1$ that can be obtained from both $\mathbf{p}_1$ and $\mathbf{p}_2$ by an α–insertion and $\mathbf{b}_{i+1}$ can be obtained from $\mathbf{b}_i$ by an α–insertion for $1 \leqslant i \leqslant q-1$.



Again by Lemma 8, it follows that $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_q\}$ form a clique of order q that lies in the neighbourhood of both $\mathbf{p}_1$ and $\mathbf{p}_2$. The desired statement follows. $\square$

Throughout the remainder of the paper we will assume that $n \geqslant 13k_0 + 4 + q$ and restrict ourselves to the patterns in the vertex set of $\mathbf{P}(3, n, q)$ that are buffered, that is we will consider

$$\mathcal{P}^+ = \{\mathbf{p}^+ : \mathbf{p} = p_1, \ldots, p_{k'} \in \mathcal{P}(3, k_0) \text{ s.t. } p_1 \neq 1 \text{ and } p_{k'} \neq 2\} \subset \mathcal{P}(3, n - q).$$

Note that we have chosen n large enough so that this set is non-empty. Furthermore, since $k_0 = 18q + 12$ and $\mathcal{P}^+ \subset \mathcal{P}(3, n - q)$, every pattern contained in $\mathcal{P}^+$ is α–diverse in $\mathbf{P}(3, n, q)$ for any $\alpha \in [3]$. We can now establish the central lemma that allows us to perform the next reduction.

Lemma 12. *Let χ be a proper $(q + 1)$–colouring of $\mathbf{P}(3, n, q)$ and $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}^+$. If the reduced count of the two patterns is the same, that is $\varphi^{(q+1)}(\mathbf{p}_1) = \varphi^{(q+1)}(\mathbf{p}_2)$, then $\chi(\mathbf{p}_1) = \chi(\mathbf{p}_2)$.*

Proof. We will show that for any $\mathbf{p} \in \mathcal{P}^+$ there exists a sequence of diverse words

$$\mathbf{p} = \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{k_1} = \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k_2}$$

such that $\mathbf{b}_{i+1}$ can be obtained from $\mathbf{b}_i$ for $1 \leqslant i < k_1$ by a single alteration and $\mathbf{c}_i$ can be obtained from $\mathbf{c}_{i+1}$ for $1 \leqslant i < k_2$ by exactly $q + 1$ α–insertions, where the $\alpha \in [3]$ is allowed to depend on the step. Crucially, we will also show that $\mathbf{c}_{k_2}$ is determined by the reduced count of $\mathbf{p}$, so that the two sequences obtained by starting at $\mathbf{p}_1$ and $\mathbf{p}_2$ terminate in the same pattern. By Lemma 11, it follows that $\chi(\mathbf{p}_1) = \chi(\mathbf{p}_2)$.

We start by constructing the sequence $\mathbf{b}_1, \ldots, \mathbf{b}_{k_1}$: let us refer to a copy of any letter $\alpha_3 \in \{1, 2, 3\}$ in a pattern as *movable* if it is positioned between a copy of α_1 and α_2 where as usual $\{\alpha_1, \alpha_2, \alpha_3\} = \{1, 2, 3\}$. In the pattern 12321 for example, both copies of the letter 2 are movable whereas none of the others are. We let $\mathbf{x} \in \mathcal{P}(3, k_0)$ be the pattern for which

$$\mathbf{p} = \mathbf{x}^+ = 1 \mid \mathbf{x} \mid (2 \sqcup 3)^{2k_0} (1 \sqcup 3)^{2k_0} (2 \sqcup 1)^{2k_0} \, 231.$$

Here both the bar as well as the symbol $\sqcup$ only serve as a visual aid to help us with the following definitions: we will refer to the part between the two bars – that is initially $\mathbf{x}$ – as the *core* and to the part to the right of the second bar – that is initially $(2 \sqcup 3)^{2k_0} (1 \sqcup 3)^{2k_0} (2 \sqcup 1)^{2k_0}\, 231$ – as the *buffer*. We will refer to the spaces between the 12s, 13s and 21s in the buffer marked by the symbol $\sqcup$ as *slots*.

We now obtain $\mathbf{b}_{i+1}$ from $\mathbf{b}_i$ by choosing an arbitrary movable letter from the core and moving it by an alteration into an appropriate slot in the buffer. We do so in a canonical fashion by always moving a letter to the left-most available slot. The slot itself, with the corresponding symbol, gets removed. Note that the bars stay in place throughout this process and serve as the reference point for our notions of core and buffer, even as both change in length.

We iterate this until there are no more movable letters in the core and refer to the point at which this happens as k_1'. Since $\mathbf{x}$ is of length at most k_0, we note that we have constructed the buffer large enough to not only contain all of $\mathbf{x}$, but also large enough that all of the $\mathbf{b}_i$ remain α–diverse for any $\alpha \in [3]$. If the core of $\mathbf{b}_{k_1'}$ is empty, then we set $k_1 = k_1'$. If, however, the core of $\mathbf{b}_{k_1'}$ is non-empty, then we must have $\mathbf{b}_{k_1'} = 1 \mid 2121\ldots21 \mid 2\ldots2\mathbf{3}\mathbf{1}$. In this case, we proceed by moving the last 3 in $\mathbf{b}_{k_1'}$ in front of the core so that $\mathbf{b}_{k_1'+1} = 1\,\mathbf{3} \mid 2121\ldots21 \mid 2\ldots21$. We observe that we are now able to recursively move all remaining letters from the core into the buffer until we reach $\mathbf{b}_{k_1-1} = 1\,\mathbf{3} \mid\mid 2\ldots21$. We finish by moving the 3 back to its original position, so that

$$\mathbf{b}_{k_1} = 1 \mid\mid (213)^{\varphi_1}\,(2\sqcup 3)^{2k_0-\varphi_1}\,(123)^{\varphi_2}\,(1\sqcup 3)^{2k_0-\varphi_2}\,(231)^{\varphi_3}\,(2\sqcup 1)^{2k_0-\varphi_3}\,231$$

where $(\varphi_1, \varphi_2, \varphi_3) = \varphi(\mathbf{x})$ is the count of $\mathbf{x}$. Note that the order in which the movable letters were chosen does not affect the outcome of this process, that is, $\mathbf{b}_{k_1}$ is independent of the ordering.

We now proceed to obtain $\mathbf{c}_{i+1}$ from $\mathbf{c}_i$ by removing, for each step, exactly $q+1$ of either the 1s, 2s or 3s from, respectively, the parts (213), (123) or (231). We can continue to do so until we reach

$$\begin{aligned}\mathbf{c}_{k_2} = 1 \mid\mid (213)^{\varphi_1^{(q+1)}}(2\sqcup 3)^{2k_0-\varphi_1^{(q+1)}}(123)^{\varphi_2^{(q+1)}}\ldots\\ \ldots(1\sqcup 3)^{2k_0-\varphi_2^{(q+1)}}(231)^{\varphi_3^{(q+1)}}(2\sqcup 1)^{2k_0-\varphi_3^{(q+1)}}231.\end{aligned}$$

where $(\varphi_1^{(q+1)}, \varphi_2^{(q+1)}, \varphi_3^{(q+1)}) = \varphi^{(q+1)}(\mathbf{x})$. We note that $\mathbf{c}_{k_2}$ only depends on the reduced count of the original core $\mathbf{x}$. Since the original buffer is identical for any core, it follows that $\mathbf{c}_{k_2}$ also only depends on the reduced count of $\mathbf{p}$, as desired. □

Thus we have proved the following statement.

Corollary 13. *If $\mathbf{C}(n,q)$ is not $(q+1)$–colourable, then neither is $\mathbf{P}(3,n,q)$.*

2.4 The structure of edges in $\mathbf{C}(n,q)$

Let us establish the structure of some of the edges that can be found in $\mathbf{C}(n,q)$. In fact, these will be essentially almost all of the edges that can be found in $\mathbf{C}(n,q)$, though we do not provide a formal proof of this. Recall that in $\mathbf{C}(n,q)$ the vertex associated with $\overline{\ell[\alpha]}^{+} \in \mathcal{P}^{+}$ is labelled with the reduced count of $\overline{\ell[\alpha]}$ for any $\alpha \in [3]$. Let $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$ and $\mathbf{e}_3 = (0,0,1)$ be the standard basis vectors in $\mathbb{N}^3$.

Lemma 14. *Given any $\mathbf{x} \in \mathbb{Z}_{q+1}^3$ as well as any $a_1, a_2, a_3 \in \mathbb{Z}$ satisfying $0 < a_1+a_2+a_3 \leqslant q$ and $a_i + a_j \geqslant 0$ for $i \neq j$, the set $\{\mathbf{x} + a_1\,\mathbf{e}_1, \mathbf{x} + a_2\,\mathbf{e}_2, \mathbf{x} + a_3\,\mathbf{e}_3\}$ forms an edge in $\mathbf{C}(n,q)$, where addition is modulo $q+1$.*

Proof. Write $\mathbf{x} = (x_1, x_2, x_3)$ where $0 \leqslant x_1, x_2, x_3 \leqslant q$ are treated as integers. We note that we must have either $a_1, a_2, a_3 \geqslant 0$ or $a_{i_2}, a_{i_3} \geqslant -a_{i_1} > 0$ for $\{i_1, i_2, i_3\} = \{1,2,3\}$. We will distinguish between these two cases.

Case 1. Assume that $a_1, a_2, a_3 \geqslant 0$. Consider

$$\begin{aligned}\boldsymbol{\ell} = {} & (2 \star 3)^{a_1}\,(213)^{x_1}\,(23)^{2(q+1)-x_1-a_1}\ldots \\ & (1 \star 3)^{a_2}\,(123)^{x_2}\,(13)^{2(q+1)-x_2-a_2}\ldots \\ & (2 \star 1)^{a_3}\,(231)^{x_3}\,(21)^{2(q+1)-x_3-a_3} \in \mathcal{P}_\star(3, k_0).\end{aligned}$$

Here the dots merely indicate that the word is continued in the next line. Also note that $x_i + a_j \leqslant 2q$ so that $2(q+1) - x_1 - a_1 > 0$ for any $i, j \in \{1, 2, 3\}$. It is easy to verify that

$$\varphi^{(q+1)}\big(\overline{\boldsymbol{\ell}[1]}\big) = (x_1 + a_1, x_2, x_3) = \mathbf{x} + a_1\,\mathbf{e}_1,$$

where all addition is modulo $q+1$. Similarly, $\varphi^{(q+1)}\big(\overline{\boldsymbol{\ell}[i]}\big) = \mathbf{x} + a_i\,\mathbf{e}_i$ for $i \in \{2, 3\}$. We note that ℓ is of length at most $18q+12 = k_0$ so that $\boldsymbol{\ell}^+ \in \mathcal{P}_\star(3, n)$ and hence by Lemma 6 $\Big\{\overline{\boldsymbol{\ell}^+[1]}, \overline{\boldsymbol{\ell}^+[2]}, \overline{\boldsymbol{\ell}^+[3]}\Big\}$ constitutes an edge in $\mathbf{P}(3, n, q)$ so that $\{\mathbf{x}+a_1\,\mathbf{e}_1, \mathbf{x}+a_2\,\mathbf{e}_2, \mathbf{x}+a_3\,\mathbf{e}_3\}$ is an edge in $\mathbf{C}(n, q)$.

Case 2. Assume that $a_1 < 0$ and $a_2, a_3 \geqslant |a_1|$. The other cases follow likewise. Consider

$$\begin{aligned}\boldsymbol{\ell} = {} & (1 \star 1)^{|a_1|}\,(213)^{x_1}\,(23)^{2(q+1)-x_1+|a_1|}\ldots \\ & (1 \star 3)^{a_2-|a_1|}\,(123)^{x_2}\,(13)^{2(q+1)-x_2-a_2}\ldots \\ & (2 \star 1)^{a_3-|a_1|}\,(231)^{x_3}\,(21)^{2(q+1)-x_3-a_3} \in \mathcal{P}_\star(3, k_0).\end{aligned}$$

It is again easy to verify that $\varphi^{(q+1)}\big(\overline{\boldsymbol{\ell}[i]}\big) = \mathbf{x} + a_i\,\mathbf{e}_i$ for $i \in [3]$ As before we conclude that $\{\mathbf{x} + a_1\,\mathbf{e}_1, \mathbf{x} + a_2\,\mathbf{e}_2, \mathbf{x} + a_3\,\mathbf{e}_3\}$ forms an edge in $\mathbf{C}(n, q)$. □

The edges described in the following easy corollary form a 'Latin cube'-type structure in $\mathbf{C}(n, q)$, that is $\{\mathbf{x}, \mathbf{x} + \mathbf{e}_i, \mathbf{x} + 2\,\mathbf{e}_i, \ldots, \mathbf{x} + q\,\mathbf{e}_i\}$ form a clique of order $q+1$ for any $\mathbf{x} \in \mathbb{Z}^3_{q+1}$ and $i \in \{1, 2, 3\}$.

Corollary 15. *For any $\mathbf{x} \in \mathbb{Z}^3_{q+1}$, $i \in \{1, 2, 3\}$ and $a \in \{1, \ldots, q\}$ there is an edge of order two between the vertices $\mathbf{x}$ and $\mathbf{x} + a\,\mathbf{e}_i$ in $\mathbf{C}(n, q)$ where addition is modulo $q+1$. Therefore, if χ is a proper $(q+1)$–colouring of $\mathbf{C}(n, q)$, then for any $\mathbf{x}$ and $\mathbf{e}_i$, each colour occurs exactly once in the set $\{\chi(\mathbf{x}), \chi(\mathbf{x} + \mathbf{e}_i), \chi(\mathbf{x} + 2\,\mathbf{e}_i), \ldots, \chi(\mathbf{x} + q\,\mathbf{e}_i)\}$.*

2.5 A lower bound on the chromatic number of $\mathbf{C}(n, q)$

Throughout this section we will continue to assume that $n \geqslant 13k_0 + 4 + q$ and simply write $\mathbf{C}_q = \mathbf{C}(n, q)$. The vertex set of $\mathbf{C}_q$ is $\mathbb{Z}^3_{q+1}$ and the edges of $\mathbf{C}_q$ that we will use are described in Lemma 14 and Corollary 15. As already noted, $\mathbf{C}_q$ contains plenty of cliques of order $q+1$ so that $\chi(\mathbf{C}_q) \geqslant q+1$. We know that this bound is sharp for even q, but we wish to show the following:

Proposition 16. *For odd q we have $\chi(\mathbf{C}_q) > q+1$.*

We will prove the proposition by considering a single colour, say 'red', which is assumed to induce no hyperedges of $\mathbf{C}_q$. We will show that the red set is determined by any two vertices which have a common neighbour in $\mathbf{C}_q$ (via edges of size two), and in fact it comes from the zero set of a linear functional. We show this using an inductive argument. Implicitly, it was shown in [7] that $\chi(C_q) \leqslant q+1$ for even q using exactly this type of colouring.

Lemma 17. *Let χ be a proper $(q+1)$–colouring of $\mathbf{C}_q$. Let $a \in \mathbb{Z}$ and $b \in \mathbb{N}_0$ such that $|a| \leqslant b$ as well as $\max(b, a+b) \leqslant q$ and let $\{i_a, i_b, i_0\} = \{1,2,3\}$. If there exists $\mathbf{x} \in \mathbb{Z}_{q+1}^3$ such that*

$$\chi(\mathbf{x}) = \chi(\mathbf{x} - a\,\mathbf{e}_{i_a} + b\,\mathbf{e}_{i_b})$$

then for any $s_0, s_1 \in \mathbb{Z}$, we have

$$\chi\big(\mathbf{p}(s_0, s_1)\big) = \chi(\mathbf{x}), \quad \textit{where} \tag{2}$$

$$\mathbf{p}(s_0,s_1) = \mathbf{p}_{\mathbf{x},a,b,i_0,i_a,i_b}(s_0,s_1) = \mathbf{x} + s_0\,\big(-(a+b)\,\mathbf{e}_{i_0} - b\,\mathbf{e}_{i_b}\big) + s_1\,\big(-a\,\mathbf{e}_{i_a} + b\,\mathbf{e}_{i_b}\big).$$

Here all addition is modulo $q+1$.

Proof. Let us highlight two special cases of (2) that will be needed throughout the proof:

(i) Equation (2) with $(s_0, s_1) = (1,1)$ reads $\chi\big(\mathbf{x} - (a+b)\,\mathbf{e}_{i_0} - a\,\mathbf{e}_{i_a}\big) = \chi(\mathbf{x})$.

(ii) Equation (2) with $(s_0, s_1) = (-1,0)$ reads $\chi\big(\mathbf{x} + (a+b)\,\mathbf{e}_{i_0} + b\,\mathbf{e}_{i_b}\big) = \chi(\mathbf{x})$.

We now prove the statement by induction on

$$d = \max(b, a+b) = \begin{cases} a+b & \text{if } a \geqslant 0 \\ b & \text{if } a < 0. \end{cases} \tag{3}$$

Note that by assumption $0 \leqslant d \leqslant q$. For $d = 0$ we must have $a = b = 0$, for which the statement is trivially true. We therefore assume that the statement of the Lemma holds for $0, 1, \ldots, d-1 < q$ and prove it for d.

We cannot have $a = 0$ since $\mathbf{x}$ and $\mathbf{x} + b\,\mathbf{e}_{i_b}$ are adjacent by Corollary 15. Let us now focus on the the case $a + b > 0$. The case where $a + b = 0$, that is $a = -d$ and $b = d$, will rely on previously having proven the statement for all other cases and therefore has to wait until the end of this proof. That case will in fact turn out to be impossible. Let us further restrict the case $a + b > 0$ by first proving that $\mathbf{p}(1,1)$ is red in a separate claim. This special case will in fact turn out to be essential in establishing (2) for arbitrary s_0, s_1 when $a + b > 0$ by serving both as a base case and as a tool for the inductive step of another induction.

Claim 18. *Let $a' \in \mathbb{Z}\setminus\{0\}$ and $b' \in \mathbb{N}$ satisfy $b' \geqslant a' > -b'$ as well as $\max(b', a'+b') = d$. If there are $\{i'_a, i'_b, i'_0\} = \{1,2,3\}$ and $\mathbf{x}' \in \mathbb{Z}_{q+1}^3$ such that*

$$\chi(\mathbf{x}') = \chi(\mathbf{x}' - a'\,\mathbf{e}_{i'_a} + b'\,\mathbf{e}_{i'_b}), \tag{4}$$

then

$$\chi(\mathbf{x}' - a'\,\mathbf{e}_{i'_a} - (a'+b')\,\mathbf{e}_{i'_0}) = \chi(\mathbf{x}'). \tag{5}$$

Proof of Claim 18. We write

$$\mathbf{c} = \mathbf{x}' - a'\,\mathbf{e}_{i'_a}, \quad \mathbf{y}' = \mathbf{c} + b'\,\mathbf{e}_{i'_b}, \quad \text{and} \quad \mathbf{z} = \mathbf{c} - (a'+b')\,\mathbf{e}_{i'_0}.$$

For the remainder of the proof we simply say that the colour of $\mathbf{x}' = \mathbf{c} + a'\mathbf{e}_{i'_a}$ and $\mathbf{y}' = \mathbf{c} + b'\mathbf{e}_{i'_b}$ is red. The aim will be to show that $\mathbf{z}$ is also red. Specifically, we will proceed by showing that $\mathbf{c} + j\mathbf{e}_{i'_0}$ cannot be red for any $-(a'+b') < j \leqslant q - (a'+b')$, so that by Corollary 15 $\mathbf{z} = \mathbf{c} - (a'+b')\mathbf{e}_{i'_0}$ in fact must be red. To that end, assume to the contrary that $\mathbf{c} + j\mathbf{e}_{i'_0}$ is red for some $-(a'+b') < j \leqslant q-(a'+b')$ and note that we can already exclude the case $j=0$, as $\mathbf{c}$ is adjacent to $\mathbf{x}'$ (as well as $\mathbf{y}'$) by Corollary 15.

Let us also offer some intuition on the proof. The vertices $\mathbf{x}', \mathbf{y}'$ and $\mathbf{z}$ form a right tetrahedron with the apex at $\mathbf{c}$. The idea is to re-orient the axes (keeping the apex at $\mathbf{c}$) so that the corresponding lengths a and b satisfy $\max(b, a+b) < d$, allowing for an application of the inductive hypothesis or Lemma 14.

Case 1. Assume that $a' > 0$. We will reach a contradiction through a case distinction illustrated in Figure 1.

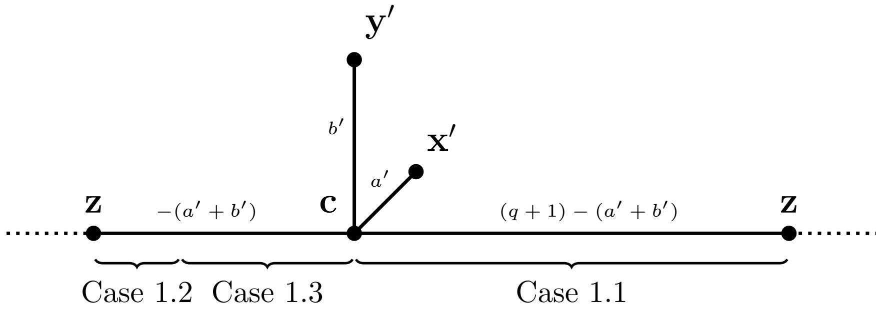


Figure 1: Relative positions of the involved points in the case $a' > 0$.

Case 1.1. If $1 \leqslant j < (q+1)-(a'+b')$, then we note that $\mathbf{x}'$ and $\mathbf{y}'$ form a hyperedge with any of the vertices $\mathbf{c} + j\,\mathbf{e}_{i'_0}$ as described in Lemma 14 and therefore these vertices cannot be red.

Case 1.2. If $-(a'+b') < j < -b'$, we will apply the inductive hypothesis in the form of (ii) with

$$a = -a', \quad b = -j, \quad i_a = i'_a, \quad i_b = i'_0 \quad \text{and} \quad \mathbf{x} = \mathbf{c} + j\,\mathbf{e}_{i'_0},$$

so that $\mathbf{x} + b\,\mathbf{e}_{i_b} - a\mathbf{e}_{i_a} = \mathbf{c} + a' e_{i'_a} = \mathbf{x}'$ and $\mathbf{x} + b\,\mathbf{e}_{i_b} + (a+b)\mathbf{e}_{i_0} = \mathbf{c} - (a'+j)\mathbf{e}_{i'_b}$. Note that we can use the inductive hypothesis as $b = -j > b' \geqslant |a'| = |a|$ and $\max(b, a+b) = b = -j < a'+b' = d$. Since $\mathbf{c} + j\,\mathbf{e}_{i'_0}$ and $\mathbf{c} + a'\,\mathbf{e}_{i'_a}$ are red by assumption, the inductive hypothesis implies that $\mathbf{c} - (a'+j)\,\mathbf{e}_{i'_b}$ is red. But $\mathbf{y}' = \mathbf{c} + b'\,\mathbf{e}_{i'_b}$ is also red, so we have $b' = -j - a'$ by Corollary 15, implying the contradiction $a'+b' = -j < a'+b'$.

Case 1.3. If $-b' \leqslant j \leqslant -1$, then we will use the inductive hypothesis in the form of (i) with

$$a = j, \quad b = b', \quad i_a = i'_0, \quad i_b = i'_b \quad \text{and} \quad \mathbf{x} = \mathbf{c} + j\,\mathbf{e}_{i'_0},$$

so that $\mathbf{x} + b\,\mathbf{e}_{i_b} - a\mathbf{e}_{i_a} = \mathbf{c} + b' e_{i'_b}$ and $\mathbf{x} - a\,\mathbf{e}_{i_a} - (a+b)\mathbf{e}_{i_0} = \mathbf{c} - (j+b')\mathbf{e}_{i'_a}$. Note that the inductive hypothesis can be used since $b = b' \geqslant -j = |a|$ and $\max(b, a+b) = b' < a'+b' = d$. It follows that that $\mathbf{c} - (j+b')\,\mathbf{e}_{i'_a}$ is red. However, since $\mathbf{c} + a'\,\mathbf{e}_{i'_a}$ is also red, we have $a' = -j - b'$ by Corollary 15, giving the contradiction $a'+b' = -j < a'+b'$.

Case 2. Assume that $a' < 0$. The contradiction will be reached through another case distinction that is illustrated in Figure 2.

Case 2.1. If $1 \leqslant j \leqslant |a'| - 1$, we will apply (ii) with

$$a = -j, \quad b = |a'|, \quad i_a = i'_0, \quad i_b = i'_a \quad \text{and} \quad \mathbf{x} = \mathbf{x}' = \mathbf{c} + a'\mathbf{e}_{i'_a},$$

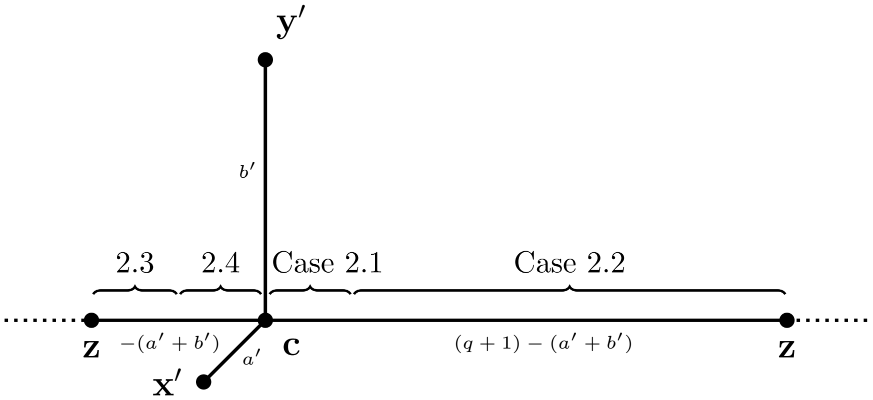


Figure 2: Relative positions of the involved points in the case $a' < 0$.

so that $\mathbf{x}+b\,\mathbf{e}_{i_b}-a\mathbf{e}_{i_a}=\mathbf{c}+j'e_{i'_0}$ and $\mathbf{x}+b\,\mathbf{e}_{i_b}+(a+b)\mathbf{e}_{i_0}=\mathbf{c}+(-a'-j)\mathbf{e}_{i'_b}$. Note that we can use the inductive hypothesis as $b=|a'|>j=|a|$ and $\max(b,a+b)=|a'|<b'=d$. It follows that $\mathbf{c}+(-a'-j)\,\mathbf{e}_{i'_b}$ is red. Since $\mathbf{y}'=\mathbf{c}+\,b'\mathbf{e}_{i'_b}$ is also red, we have $b'=-j-a'$, implying the contradiction $a'+b'=-j$.

Case 2.2. If $|a'|\leqslant j<(q+1)-(a'+b')$, then we argue as in case 1.1. That is, we observe that $\mathbf{x}'$ and $\mathbf{y}'$ form a hyperedge with any of the vertices $\mathbf{c}+j\,\mathbf{e}_{i'_0}$ as described in Lemma 14 and therefore these vertices cannot be red.

Case 2.3. If $-(a'+b')<j\leqslant-|a'|$, then we proceed as in case 1.2. That is, we apply (ii) with

$$a=|a'|,\quad b=-j,\quad i_a=i'_a,\quad i_b=i'_0\quad\text{and}\quad \mathbf{x}=\mathbf{c}+j\,\mathbf{e}_{i'_0},$$

to conclude that $\mathbf{c}+(a+b)\,\mathbf{e}_{i'_b}$ is red. Note that we could use the inductive hypothesis as $b=-j\geqslant|a'|=|a|$ as well as $\max(b,a+b)=|a'|-j<|a'|+a'+b'=b'=d$. It follows that $b'=a+b=|a'|-j$, implying the contradiction $a'+b'=-j<a'+b'$.

Case 2.4. If $\max\big(-|a'|,-(a'+b')\big)<j\leqslant-1$, then we proceed as in case 2.1. That is, using (ii) with

$$a=-j,\quad b=|a'|,\quad i_a=i'_0,\quad i_b=i'_a\quad\text{and}\quad \mathbf{x}=\mathbf{x}',$$

we get that $\mathbf{c}+(a+b)\,\mathbf{e}_{i'_b}$ is red. Note that we could use the inductive hypothesis as $b=|a'|>-j=|a|$ as well as $\max(b,a+b)=|a'|-j<|a'|+a'+b'=b'=d$. Hence $b'=a+b=-j+|a'|$, implying the contradiction $a'+b'=-j<a'+b'$.

As previously in Case 1, we conclude that $\mathbf{z}$ must be red. This concludes the proof of Claim 18. ■

Claim 18 will now be used to show (2) in full generality, so for all $s_0,s_1\in\mathbb{Z}$, when $a+b>0$ and $\max(b,a+b)=d$. Recall that the case $a+b=0$ will be dealt with at the very end of this proof. Also recall that

$$\mathbf{p}(s_0,s_1)=\mathbf{x}+s_0\,\big(-(a+b)\,\mathbf{e}_{i_0}-b\,\mathbf{e}_{i_b}\big)+s_1\,\big(-a\,\mathbf{e}_{i_a}+b\,\mathbf{e}_{i_b}\big).$$

The assumption of our lemma is that $\mathbf{p}(0,0)=\mathbf{x}$ and $\mathbf{p}(0,1)=\mathbf{x}-a\,\mathbf{e}_{i_a}+b\,\mathbf{e}_{i_b}$ are red. Claim 18 established that $\mathbf{p}(1,1)$ is red as well, though its implications will go beyond that. As previously in the proof of the Claim 18, we have to distinguish two cases for a.

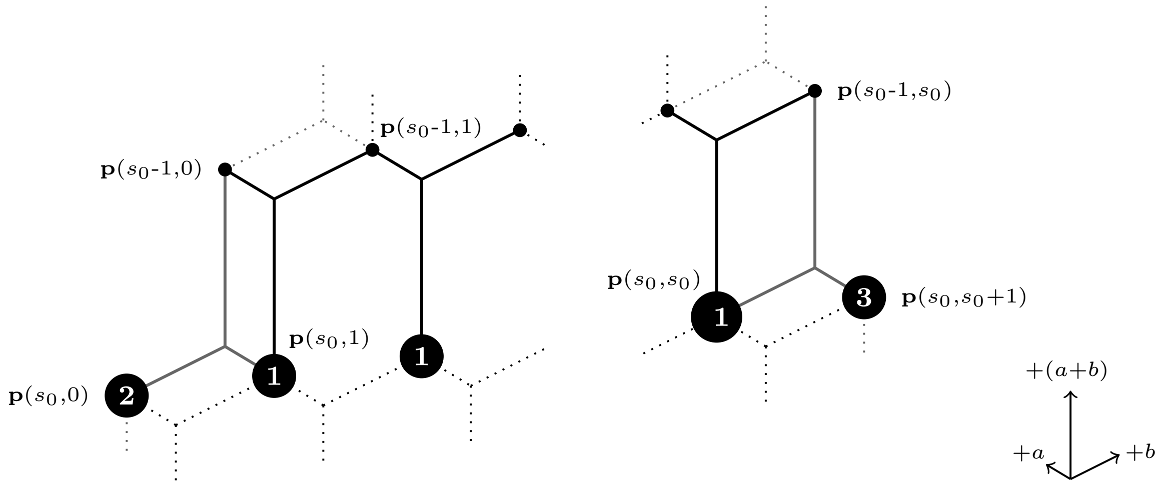


Figure 3: The inductive step over s_0 in the case $a > 0$.

Case I. If $a > 0$, then we start by showing that $\mathbf{p}(s_0, s_1)$ is red for any $s_0, s_1 \in \mathbb{N}_0$ satisfying $s_1 \leqslant s_0 + 1$ by induction on s_0. We note that for $s_0 = 0$ the statement holds as $\mathbf{p}(0, 0) = \mathbf{x}$ and $\mathbf{p}(0, 1) = \mathbf{x} - a\,\mathbf{e}_{i_a} + b\,\mathbf{e}_{i_b}$. Assume therefore that the statement holds for $s_0 - 1$ and let us show that it holds for s_0 and any $0 \leqslant s_1 \leqslant s_0 + 1$. We will do so through another case distinction that is illustrated in Figure 3:

1. We start by proving it for $0 < s_1 < s_0+1$. By inductive assumption, $\mathbf{p}(s_0-1, s_1-1)$ and $\mathbf{p}(s_0 - 1, s_1)$ are red so that we can apply Claim 18 with
$$a' = a, \quad b' = b, \quad i'_a = i_a, \quad i'_b = i_b \quad \text{and} \quad \mathbf{x}' = \mathbf{p}(s_0 - 1, s_1 - 1),$$
to deduce that $\mathbf{p}(s_0, s_1)$ is also red.
2. Secondly, suppose $s_1 = 0$. The inductive hypothesis and Case 1 imply that $\mathbf{p}(s_0 - 1, 0)$ and $\mathbf{p}(s_0, 1)$ are red, so we may apply Claim 18 with
$$a' = -a, \quad b' = a + b, \quad i'_a = i_a, \quad i'_b = i_0 \quad \text{and} \quad \mathbf{x}' = \mathbf{p}(s_0, 1),$$
to deduce that $\mathbf{p}(s_0, 0)$ is also red.
3. Finally, consider $s_1 = s_0 + 1$. Case 1 and the inductive hypothesis imply that $\mathbf{p}(s_0, s_0)$ and $\mathbf{p}(s_0 - 1, s_0)$ are red, so we may apply Claim 18 with
$$a' = -b, \quad b' = a + b, \quad i'_a = i_b, \quad i'_b = i_0 \quad \text{and} \quad \mathbf{x}' = \mathbf{p}(s_0, s_0),$$
to deduce that $\mathbf{p}(s_0, s_0 + 1)$ is also red.

This concludes the inductive step, so we have shown that (2) holds for any $s_0 \in \mathbb{N}_0$ and $0 \leqslant s_1 \leqslant s_0 + 1$ when $a > 0$. However, since the summation is done modulo $(q + 1)$, we have for instance $\mathbf{p}(s_0, s_1) = \mathbf{p}(s_0 - (q + 1), s_1)$. It immediately follows that $\mathbf{p}(s_0, s_1)$ is red for any $s_1, s_0 \in \mathbb{Z}$.

Case II. If $a < 0$, then we note that, by Claim 18, $\mathbf{z} = \mathbf{x} - a\,\mathbf{e}_{i_a} - (a + b)\,\mathbf{e}_{i_0}$ is red. If $a + b \geqslant |a|$, then we rewrite that last equation as $\mathbf{x} = \mathbf{z} - |a|\,e_{i_a} + (a + b)\,\mathbf{e}_{i_0}$ and observe that the position of $\mathbf{x}$ in relation to $\mathbf{z}$ satisfies the conditions of Case I. One can quickly verify that $\mathbf{p}_{\mathbf{x},a,b,i_0,i_a,i_b}(s_0,s_1) = \mathbf{p}_{\mathbf{z},|a|,a+b,i_b,i_a,i_0}(s_0-s_1,-s_1+1)$, where addition is modulo $q+1$, to see that the new orientation does in fact still span the same lattice structure. If $a+b \leqslant |a|$ then we note that $\mathbf{z} = \mathbf{x} - (a+b)\,\mathbf{e}_{i_0} + |a|\,\mathbf{e}_{i_a}$ so that now $\mathbf{z}$ in relation to $\mathbf{x}$ satisfies the conditions of Case I. We can again relate the new orientation to the old one through $\mathbf{p}_{\mathbf{x},a,b,i_0,i_a,i_b}(s_0,s_1) = \mathbf{p}_{\mathbf{z},a+b,|a|,i_b,i_0,i_a}(s_0-s_1,s_0-1)$. In either scenario we can immediately derive (2).

This completes the inductive step over d for $a+b>0$. The remaining case is $-a=b=d$, so let $\mathbf{x}$ and $\mathbf{y} = \mathbf{x}+b\,\mathbf{e}_{i_a}+b\,\mathbf{e}_{i_b}$ be red. By Corollary 15, the vertex $\mathbf{x}+b\,\mathbf{e}_{i_a}$ is adjacent to $\mathbf{x}$ (as well as $\mathbf{y}$) and therefore cannot be red. The vertices $\mathbf{x}+b\,\mathbf{e}_{i_a}+j\,\mathbf{e}_{i_0}$ form an edge of order three with $\mathbf{x}$ and $\mathbf{y}$ for any $|a|=b<j\leqslant q$ as described in Lemma 14 and therefore these vertices also cannot be red. Now assume that $\mathbf{x}+b\,\mathbf{e}_{i_a}+j_0\,\mathbf{e}_{i_0}$ is red for some $1\leqslant j_0\leqslant b-1$. We apply (2) with $-j_0$ in place of a and the same value of b. Note that $|-j_0| = j_0 < b$ as well as $\max(b,-j_0+b) = b = d$ and $j_0 \neq b$. We previously established that (2) holds in this case, so using observation (ii) we get that $\mathbf{x}+b\,\mathbf{e}_{i_a}+(-j_0+b)\,\mathbf{e}_{i_b}$ must be red. Since $\mathbf{y} = \mathbf{x}+b\,\mathbf{e}_{i_a}+b\,\mathbf{e}_{i_b}$ is also red, we must have $b-j_0=b$ in contradiction to $j_0\geqslant 1$. We have shown that the vertices $\mathbf{x}+b\,\mathbf{e}_{i_a}+j\,\mathbf{e}_{i_0}$ cannot be red for any $0\leqslant j\leqslant q$, contradicting Corollary 15. It follows that (2) vacuously holds for the case $-a=b=d$, completing the inductive step over d and proving Lemma 17. □

Let us now derive Proposition 16 from Lemma 17.

Proof of Proposition 16. By Corollary 15, one of the vertices $\{j\,\mathbf{e}_2 : 0\leqslant j\leqslant q\}$ must have the same colour as $-\mathbf{e}_1$, say $\chi(j_0\,\mathbf{e}_2) = \chi(-\mathbf{e}_1)$. If $j_0=0$, then we get an immediate contradiction. For $j_0\neq 0$, we apply Lemma 17 with

$$a=-1, \quad b=j_0, \quad i_a=1, \quad i_b=2 \quad \text{and} \quad \mathbf{x}=-\mathbf{e}_1.$$

With $\mathbf{p}(s_0,s_1)$ as in the statement of Lemma 17, it follows that $\mathbf{p}(s_0,s_1)$ have the same colour as $-\mathbf{e}_1$ and $b\,\mathbf{e}_2$ for any $s_0,s_1\in\mathbb{Z}$. Let A_a,A_b,A_0 respectively denote the set of their projections onto the axes i_a,i_b,i_0 in $\mathbb{Z}_{q+1}^3$. We note that

$$\begin{aligned} |A_a| &= (q+1)/\gcd(1,q+1) = q+1, \\ |A_b| &= (q+1)/\gcd(j_0,q+1), \\ |A_0| &= (q+1)/\gcd(j_0-1,q+1). \end{aligned}$$

However, in order to respect the latin cube structure described by Corollary 15, we must have $|A_a|=|A_b|=|A_0|=q+1$. It follows that

$$1 = \gcd(j_0,q+1) = \gcd(j_0-1,q+1), \tag{6}$$

which immediately gives a contradiction since $q+1$ is even by assumption and at least one of j_0 and j_0-1 must be even as well. □

Theorem 2 now follows as an immediate consequence of Proposition 16, Corollary 13 and Corollary 5, where we set the $N_2=N_2(r)$ of Theorem 2 equal to $N_5(r,13k_0+4+q) = N_5(r,235q+160)$ from Corollary 5, where of course $q=r-1$.

3 Remarks and Open Questions

In this paper, we settle the particular inquiry started by Conlon and the first author into structural properties of the monochromatic combinatorial lines of length three given by the Hales–Jewett theorem. Of course, many related issues remain very much open, which we will outline in the following.

Bounds for larger alphabets

Let $\mathcal{I}(m,r)$ be the minimum q so that for sufficiently large n any r–colouring of $[m]^n$ contains a q–fold combinatorial line. In [7], it was shown that $\mathcal{I}(3,r) = r$ for odd r and conjectured that the result extends to $\mathcal{I}(m,r) = HJ(m-1,r)$. Given the result of Leader and Räty, this conjecture was later retracted, though it is still reasonable to wonder if $\mathcal{I}(m,r)$ is at all linked to $HJ(m-1,r)$ when $m > 3$.

Question 19. Can we improve on either of the immediate bounds

$$\max\left\{\mathcal{I}(m-1,r), \mathcal{I}(m,r-1)\right\} \leqslant \mathcal{I}(m,r) \leqslant HJ(m-1,r)? \tag{7}$$

In particular, $r-1$ is the best explicit lower bound that the present methods can give on $\mathcal{I}(m,r)$, which is probably far from the truth.

An upper bound separating $\mathcal{I}(m,r)$ from $HJ(m-1,r)$ would entail a new argument for the Hales-Jewett theorem. For instance, the reduction to the hypergraph $\mathbf{P}(3,n,r)$ conceived by Leader and Räty, along with the observation that $\mathbf{P}(3,n,r)$ contains an r–clique, already gives an alternative proof for an alphabet of size 3. For $m \geqslant 3$, this strategy essentially reduces the alphabet size by one and therefore may be a good starting point for further inquiries.

Bounds for the Hales–Jewett number

There is currently a large gap in the best bounds on $HJ(3,r)$, see [3, 20, 15, 16, 4, 6]. More specifically, we have

$$r^{c\ln(r)} \leqslant HJ(3,r) \leqslant 2^{2^{cr}} \tag{8}$$

for some $c > 0$, where the upper bound is a recent result of Conlon [6] and the lower bound follows from a lower bound on the van der Waerden number $W(3,r)$ due to Graham, Rothschild and Spencer [13]. With this in mind, we ask the following question.

Question 20. Given $r \geqslant 2$ and $q \geqslant 2\lceil r/2\rceil - 1$, what $N = N(r,q)$ guarantees the existence of q–fold lines in any r–colouring of $[3]^N$?

A better lower bound on $N(r,q)$ may turn out to be more accessible than one for $HJ(3,r)$.

The cyclic setting

It should also be noted that the setup becomes significantly more natural if we take a cyclic ground set for the coordinates, that is $\mathbb{Z}_n$ rather than $[n]$, so that we might have an interval at the border that ‘wraps’ around. In particular, the step to the reduced count in Subsection 2.3 becomes easier while actually resulting in a stronger statement that avoids the need for the “buffer”. We therefore highly recommend that anyone interested in further exploring this topic use that setup instead of the one employed so far in this and the previous papers.

The Density Hales–Jewett Theorem

Much of the recent interest in the Hales–Jewett theorem has been focused at density versions, see for example [18, 1, 19, 8, 11]. It would be interesting if a similar line of inquiry could be followed in this setting. However, to our knowledge, none of the currently stated proofs of the density version allow for similar structural observations as studied in this paper.

Acknowledgements

We would like to thank David Conlon and Oriol Serra for helpful discussions and comments on the topic. The first author would like to thank UPC for their hospitality during her visit to Barcelona. Part of the research was conducted while the first author was at the Department of Mathematics of ETH Zürich. We are grateful to the anonymous referees for significantly improving the presentation of the paper.

References

[1] T. Austin. Deducing the density Hales–Jewett theorem from an infinitary removal lemma. *J. Theor. Probab.*, 24(3):615, 2011.

[2] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden’s and Szemerédi’s theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996.

[3] E. R. Berlekamp. A construction for partitions which avoid long arithmetic progressions. *Canad. Math. Bull*, 11(1968):409–414, 1968.

[4] T. Blankenship, J. Cummings, and V. Taranchuk. A new lower bound for van der Waerden numbers. *Eur. J. Combin.*, 69:163–168, 2018.

[5] T. C. Brown, R. L. Graham, and B. M. Landman. On the set of common differences in van der Waerden’s theorem on arithmetic progressions. *Canad. Math. Bull.*, 42(1):25–36, 1999.

[6] D. Conlon. Monochromatic combinatorial lines of length three. *Proc. Amer. Math. Soc.*, 150(1):1–4, 2022.

[7] D. Conlon and N. Kamčev. Intervals in the Hales–Jewett theorem. *Eur. J. Combin.*, to appear.

[8] P. Dodos, V. Kanellopoulos, and K. Tyros. A simple proof of the density Hales–Jewett theorem. *Int. Math. Res. Notices*, 2014(12):3340–3352, 2014.

[9] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.

[10] H. Furstenberg and Y. Katznelson. Idempotents in compact semigroups and Ramsey theory. *Isr. J. Math.*, 68(3):257–270, 1989.

[11] J. Geelen and P. Nelson. A density Hales–Jewett theorem for matroids. *J. Comb. Theory B*, 112:70–77, 2015.

[12] M. Golshani and S. Shelah. Iterated Ramsey bounds for the Hales-Jewett numbers. arXiv:1912.08643, 2019.

[13] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey theory*, volume 20. John Wiley & Sons, 1990.

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

[15] N. Hindman and E. Tressler. The first nontrivial Hales–Jewett number is four. *Ars Comb.*, 113:385–390, 2014.

[16] M. Lavrov. An upper bound for the Hales–Jewett number HJ(4,2). *SIAM J. Discrete Math.*, 30(2):1333–1342, 2016.

[17] I. Leader and E. Räty. A note on intervals in the Hales–Jewett theorem. *Electron. J. Combin.*, 25(3), #P3.15, 2018.

[18] R. McCutcheon. A variant of the density Hales–Jewett theorem. *B. Lond. Math. Soc.*, 42(6):974–980, 2010.

[19] D. H. J. Polymath. A new proof of the density Hales-Jewett theorem. *Ann. of Math. (2)*, 175(3):1283–1327, 2012.

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