# From Hypercubes to Sparse Graphs: Structural Translation of the Hales–Jewett Theorem via Monochromatic Stars

Y. Mouhib
Supervisor: Prof. Dr. L. Halbeisen
Department of Mathematics, ETH Zürich

December 5, 2025

**Abstract**

In this paper, we investigate possible approaches to improving the known bounds on the Hales–Jewett numbers. In the first part, we present some variants of the Hales–Jewett numbers corresponding to a restricted number of combinatorial lines, obtained by imposing strengthened conditions, and explore how these variants may be used to derive improved upper bounds for the classical Hales–Jewett numbers.

In the second part, we introduce an induced construction that produces an edge coloring of a sparse graph from a vertex coloring of a hypercube. We study how the previously defined numbers behave under this construction and how this may lead to a connection with Ramsey-type theorems for monochromatic stars in sparse graphs.

## Contents



# 1 Introduction

The Hales-Jewett theorem [1] is one of the central results of Ramsey theory. The theorem says that for any positive integers $t$ and $r$ there is an integer $n = HJ(t, r)$ such that for any coloring of the vertices of the hypercube $[t]^n$ using $r$ color, we can always find a monochromatic combinatorial line. In order to understand what a combinatorial line is, we need the following: a *word* of length $n$ is simply an element in $[t]^n$. $[t] = \{1, \ldots, t\}$ is referred as the *alphabet* of size $t$. We call *root* an element in $([t] \cup \{*\})^n \setminus [t]^n$, and write $[t]^n_*$ for the set of all roots of length $n$ on the alphabet $[t]$. Given a root $\tau$ and a letter $\alpha$, we denote by $\tau(\alpha)$ the word obtained by replacing every $*$ in $\tau$ with an $\alpha$. We refer to the coordinates in $\tau$ where the symbol $*$ appears as *active* coordinates. The set $L_\tau = \{\tau(1), \ldots, \tau(t)\}$ is the *combinatorial line* generated by $\tau$. $L([t]^n)$ is the set of all combinatorial lines in $[t]^n$.

Currently, there is little concrete knowledge about the Hales-Jewett numbers. The proof of Shelah [2] produces a better bound than the one derived from the original proof of Hales and Jewett. However, the bound is still very large and it is natural to try to improve it. More recently, Shelah and Golshani slightly improved this bound [3], but substantial progress remains elusive. Regarding small Hales–Jewett numbers (for 2-colorings), on can trivialy see that $HJ(2,2) = 2$, but $HJ(3,2)$ is already non-trivial: a proof that $HJ(3,2) = 4$ has recently been published,see [4]. Regarding $HJ(4,2)$, we only know $HJ(4,2) < 10^{11}$.), see [4].

Inspired by Shelah's proof that produces monochromatic lines with a specific structure, one sees that it is sometimes possible to restrict the class of combinatorial lines under consideration, for example, by bounding the number of intervals of active coordinates. Some work has been done in that direction by Kamcev and Spiegel; see [5], showing that for even $r$ and large $n$, any $r-$ coloring of $[3]^n$ contains a monochromatic combinatorial line, whose set of active coordinates is the union of at most $r-1$ intervals.

From these observations, it is very natural to ask the following questions: can we obtain a better upper bound of $HJ(t,r)$ by imposing the right restrictions on the set of combinatorial lines?

In this paper, we investigate possible approaches to improving the known bounds on the Hales–Jewett numbers. In the first part, we present some variants of the Hales–Jewett numbers corresponding to a restricted number of combinatorial lines, obtained by imposing strengthened conditions, and explore how these variants may be used to derive improved upper bounds for the classical Hales–Jewett numbers.

In the second part, we introduce a construction that produces an edge coloring of a sparse graph from a vertex coloring of a hypercube: starting from an $r$-coloring of the hypercube $[t]^n$, we build an edge-colored graph whose vertices correspond to combinatorial lines and whose edges correspond to their intersections. In this setting, a monochromatic star corresponds precisely to a monochromatic combinatorial line. We will discuss how the previously defined numbers behave under this construction and how this may lead to a connection with Ramsey-type theorems for monochromatic stars in sparse graphs and improvements for the Hales-Jewett numbers.

Some attempts have been made to express the Hales-Jewett theorem in terms of other structure language, see, for example, Prömel [6], or the approach in [7]. These works suggest that the problem is more difficult than one might expect, these translations of the Hales-Jewett theorem into graphs or other algebraic structures being rather unwiedly or too complicated (using, for instance, strong arguments such as Stone-Čech compacification in [7]). For this reason, we do not expect our own transformation to provide a much better translation of the Hales-Jewett theorem. But we hope that it could open the way for a better one, and still, it is worthwhile to explore this new approach for the following reason: Even if our graph-theoretic interpretation does not align perfectly with the original setting (for instance, because many edges may correspond to the same word), this mapping of the high-dimensional cube $[t]^n$ to a sparse edge-colored graph isolates structural features (monochromatic star-like configurations). These features could reveal some structural mechanisms in Ramsey theory and the theory of sparse graphs.

# 2 Hales-Jewett Theorem

## 2.1 Classical Hales-Jewett Numbers

**Example 1.** Let $A = \{1,2,3,4\}$, be an alphabet of 4 letters.
42 is a word of length 2, $\tau_1 = **and\tau_2 = *2$ are roots of length 2.
$\tau_1(3) = 33$.
$L_{\tau_2} = \{12, 22, 32, 42\}$ is the combinatorial line generated by $\tau_2$.
$\tau = 2*1*$ is a root of length 4 in $A$.
$L_\tau = \{2111, 2212, 2313, 2414\}$.

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

The least such $n$ is called the *Hales-Jewet number* and is denoted by $\mathrm{HJ}(t,r)$. In this paper, we will only consider 2-colorings, so we denote $HJ(t,2)$ by $HJ(t)$ in order to simplify the notation.

**Example 2.** It is easy to see that HJ(2)=2. Figure 1 shows a coloring of [2] with no monochromatic combinatorial line, so $HJ(2) > 1$, and in figure 2, we can see that any coloring of $[2]^2$ leads to a monochromatic combinatorial line, so $HJ(2) \leq 2$.

Figure 1: Coloring of [2] with no monochromatic combinatorial line.

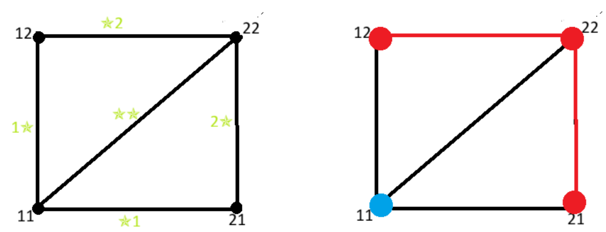


Figure 2: Words and combinatorial lines in the cube $[2]^2$ (left), and a 2-coloring of $[2]^2$ with two monochromatic combinatorial lines (right).

**Example 3.** $HJ(3) = 4$. In figure 3, we see that $HJ(3) > 2$, in figure 4 we see that $HJ(3) > 3$ (where the coloring was taken from [8]). To understand why colorings of $[3]^4$ always contain a monochromatic combinatorial line (and therefore $HJ(3) = 4$) we refer the reader to the proof by Neil Hindman and Eric Tressler [4]. In their paper, the authors provide two different proofs. The first is algorithmic: they consider all possible colorings and, using certain observations, exclude those for which the observations already guaranty the existence of a monochromatic combinatorial line. For the remaining colorings, they run their algorithm to test them all, and in every case they find a monochromatic combinatorial line. The second method relies solely on combinatorial observations, allowing them to directly conclude that any coloring of $[3]^4$ contains a monochromatic combinatorial line.

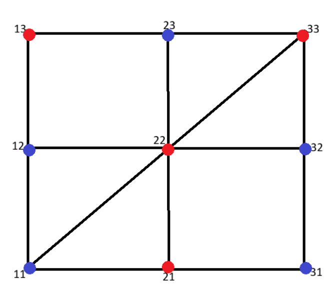


Figure 3: Coloring of $[3]^2$ with no monochromatic combinatorial line.

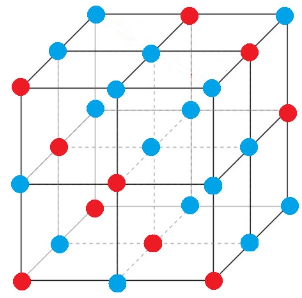

Figure 4: Coloring of $[3]^3$ with no monochromatic combinatorial line.

Very little is known about $HJ(4)$, except $HJ(4) < 10^{11}$, see [9]. Thus, a good starting point would be to try to improve this particular bound.

## 2.2 Some variants of the Hales-Jewett Numbers obtained by imposing stronger conditions on roots

Imposing additional structural requirements on the desired monochromatic objects is a direction which is very common in Ramsey theory. We start by introducing some strong conditions on the set of lines we are looking at. These are interesting to study on their own and will also help us illustrate certain methods later, when talking in the language of graphs. But these conditions might be too strong to properly bound the Hales Jewett numbers, therefore, we also introduce some weaker conditions that do not remove "too many" lines from the original set of combinatorial lines.

For a combinatorial line $l \in L([t]^n)$, we denote $\tau(l)$ its corresponding root. For a root $\tau \in [t]^n_*$, we write $I(\tau)$ for the set of its active coordinates, i.e. $I(\tau) \subseteq [n]$ is the set of coordinates of $\tau$ for which a symbol $*$ appears. We define

$$L^{[k]}([t]^n) := \{l \in C_t^n : |I(\tau(l))| \leq k\},$$

and

$$L^{(q)}([t]^n) := \{l \in C_t^n : I(\tau(l)) \text{ consist of at most } q \text{ subintervals of } [n]\}.$$

A line $l$ in $L^{(q)}([t]^n)$ is called $q$*−fold*, where we follow the name given in [5].

**Example 4.** $1***2*$ is 2-fold since the sub-intervals of active coordinates are $[2,4]$ and $\{6\}$.

Let $HJ^{[k]}(t)$ be the smallest integer $n$ such that any $2-coloring$ of $[t]^n$ contains a monochromatic line $l \in L^{[k]}([t]^n)$.
Let $HJ^{(q)}(t)$ be the smallest integer $n$ such that any $2-coloring$ of $[t]^n$ contains a monochromatic line $l \in L^{(q)}([t]^n)$.

**Remark.** Of course we have as expected $HJ(t) \leq HJ^{[k]}(t)$, for any $k$, and $HJ(t) \leq HJ^{(q)}(t)$, for any $q$ (and $HJ^{[q]}(t) \leq HJ^{(q)}(t)$).

**Remark.** We also note that $L^{[1]}([t]^n)$ behaves very nicely. It consists of only the hypercube's lines without any diagonals, but for that very reason it provides a terrible upper bound.

**Proposition 1.** $HJ^{[1]}(t) = \infty$ *for any positive integer* $t$.

*Proof.* In figure 5 and 6, we see already that $HJ^{[1]}(2) > 2$, $HJ^{[1]}(2) > 3$. By extending the same type of coloring, we are able to color any hypercube $[t]^n$ without having any monochromatic line in $L^{[1]}([t]^n)$. Consider the following coloring: for a word $w$, color $w$ in blue if the sum of its letters is even, and in red otherwise. Let $l \in L^{[1]}([t]^n)$. $\tau(l)(1)$ $\tau(l)(2)|$ must have different colors. □

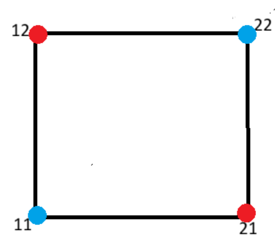


Figure 5: Coloring of $[2]^2$ with no monochromatic line in $L^{[1]}([2]^2)$

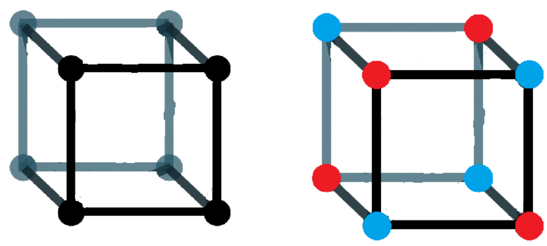

Figure 6: hypercube $[2]^3$ together with the lines from $L^{[1]}([2]^3)$ only (left). Coloring of $[2]^3$ with no monochromatic line in $L^{[1]}([2]^3)$ (right).

**Example 5.** In Figure 7 we see all the combinatorial lines in $[2]^3$. Figure 8 illustrates why it is not possible to color $[2]^3$ without getting a monochromatic line in $L^{[2]}([2]^3)$. Therefore, we also get $HJ[2](2) \leq 3$. In fact $2 = HJ(2) = HJ^{[2]}(2)$, therefore it is very natural to hope that maybe $HJ^{[2]}(3)$ could be a good upper for $HJ(3)$. We already know $HJ(3) = 4$, but if $HJ^{[2]}(3) = 4$ then it could lead to another simple proof. And the same apply to $HJ^{[2]}(4)$ in order to bound $HJ(4)$.

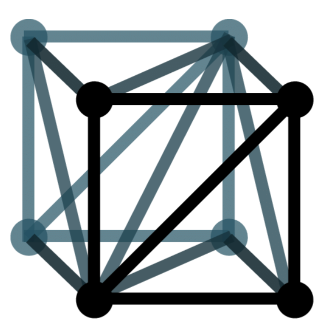

Figure 7: all combinatorial lines in $[2]^3$

In Shelah's proof [2] , the argument implies that for sufficiently large $n$, any $r-$coloration of $[t]^n$ contains a monochromatic line $q - fold$ with $q \leq HJ(t-1, r)$. (In particular, $HJ(4) \leq HJ^{(4)}(4) < \infty$). Colon and Kamčev investigated the extra structure exposed by the proof and found the following result: for $r$ odd, in any dimension, there exists a coloring of $[3]^n$ with no monochromatic $r-1$fold lines, see [10]. (So, $HJ^{(r-1)}(3, r) = \infty$, for $r$ odd). One year later, a result of Leader and Räty, see [11], was extended by Kamcev and Spiegel in [5]: For any even integer $r \geq 2$, there exists a positive integer $N = N(r)$ such that any $r-$coloring of $[3]^N$ contains a monochromatic combinatorial line whose set of active coordinates is a union of at most $r - 1$ intervals. (So, $HJ^{(r-1)}(3, r) < \infty$, for $r$ even. And in particular $HJ^{(1)}(3) < \infty$

**Remark.** We have $HJ(3) \leq HJ^{(1)}(3) < HJ^{[1]}(3) = \infty$. It will be therefore natural to start by exploring $HJ^{(1)}(4)$ later to bound $HJ(4)$.

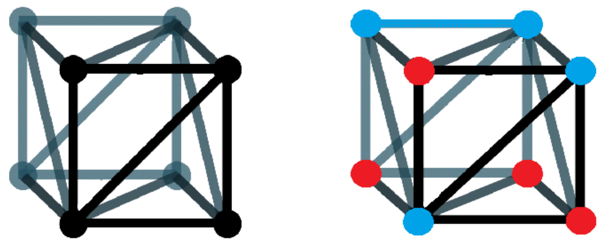

Figure 8: hypercube $[2]^3$ together with the lines from $L^{[2]}([2]^3)$ (left). Coloring of $[2]^3$ with no monochromatic line in $L^{[2]}([2]^3)$ (right).

### 2.3 Table of some trivial or known bounds for small alphabet size

| | | |
|---|---|---|
| $HJ(2) = 2$ | $HJ(3) = 4$ | $4 < HJ(4) = 2 < 10^{11}$ |
| $HJ^{[4]}(2) = 2$ | $HJ^{[4]}(3) = 4$ | $HJ^{[4]}(4) =?$ |
| $HJ^{[3]}(2) = 2$ | $HJ^{[3]}(3) =?$ | $HJ^{[3]}(4) =?$ |
| $HJ^{[2]}(2) = 2$ | $HJ^{[2]}(3) =?$ | $HJ^{[2]}(4) =?$ |
| $HJ^{[1]}(2) = \infty$ | $HJ^{[1]}(3) = \infty$ | $HJ^{[1]}(4) = \infty$ |
| $HJ(2) = 2$ | $HJ(3) = 4$ | $4 < HJ(4) = 2 < 10^{11}$ |
| $HJ^{(4)}(2) = 2$ | $HJ^{(4)}(3) = 4$ | $HJ^{(4)}(4) =?$ |
| $HJ^{(3)}(2) = 2$ | $HJ^{(3)}(3) = 4$ | $HJ^{(3)}(4) =?$ |
| $HJ^{(2)}(2) = 2$ | $HJ^{(2)}(3) = 4$ | $HJ^{(2)}(4) =?$ |
| $HJ^{(1)}(2) = 2$ | $HJ^{(1)}(3) =?$ | $HJ^{(1)}(4) =?$ |

### 2.4 Generalized version of the Hales-Jewett theorem

It is possible to generalize the Hales-Jewett theorem by considering colorings of higher-dimensional combinatorial subspaces. Instead of points and combinatorial lines, we may, for example $2 - color$ the combinatorial lines and be looking for a monochromatic "combinatorial plane" see figure 9, taken from [12].

We call *order-k root* an element in $([t] \cup \{*_1, \ldots, *_k\})^n$ in which every symbol appears at least once. We write $[t]^n_{*_k}$ for the set of all the order-$k$ roots of length $n$ in the alphabet $[t]$. Given an order$-k$ root $\tau$ and a $k$ letters $\alpha_1, \ldots, \alpha_k$, we denote by $\tau(\alpha_1, \ldots, \alpha_k)$ the word obtained by replacing every $*_i$ in $\tau$ with an $\alpha_i$. The set

$$S_\tau = \{\tau(\alpha_1, \ldots, \alpha_k), \alpha_i \in [t], \text{ for } i = 1, \ldots, k\},$$

is the *combinatorial $k-$space* generated by $\tau$.

Graham-Rothschild theorem (in a simplified version), see [13], asserts that for any fixed alphabet size $t$, any integers $m > k \geq 0$, and any number of colors $r$, there exists a dimension $n$ such that every $r$-coloring of the combinatorial $k-$spaces of $[t]^n$ contains a monochromatic combinatorial $m-$space.

Following the same ideas in this section, it is also possible to impose more conditions on the sets of combinatorial $m$-spaces. In the case our construction in next section reveals to be helpful somehow, then it would be very interesting to generalize the ideas to the Graham Rothschilds numbers.

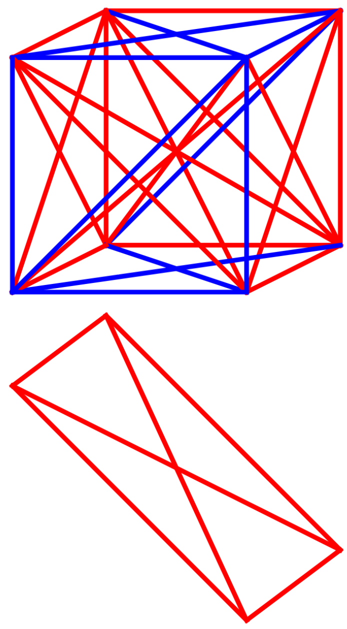

Figure 9: A 2-coloring of the combinatorial lines of a three-dimensional binary cube, and a monochromatic combinatorial plane in this colored cube.

# 3 Sparse graphs and monochromatic stars

## 3.1 Presentation of the construction through a simple example

**Example 6.** Consider the hypercube $[3]^2$ together with its combinatorial lines ($L([3]^2)$). Define the graph $G = (V, E)$, $V = L([3]^2)$ and $l_1 l_2$ is an edge in $E$ if, and only if, $l_1$ and $l_2$ intersect in the hypercube $[3]^2$. In particular, this construction sends vertices of the hypercube to edges in this graph, with possibly many edges representing the same word; see figure 10. We call this graph the graph induced by $L([3]^2)$ and denote it by $G([3]^2)$. Now, any $2-$coloring of $[3]^2$ naturally induces a edge-coloring of the induced graph. For example, figure 11 shows a coloring of $[3]^2$ without monochromatic combinatorial lines and the induced coloring on $G([3]^2)$. Note that if we change the color of the word 33 to blue, $3*$ is now a monochromatic combinatorial line (see figure 12) and the induced graph now contains a monochromatic *star* that corresponds to $3*$ (figure 13).

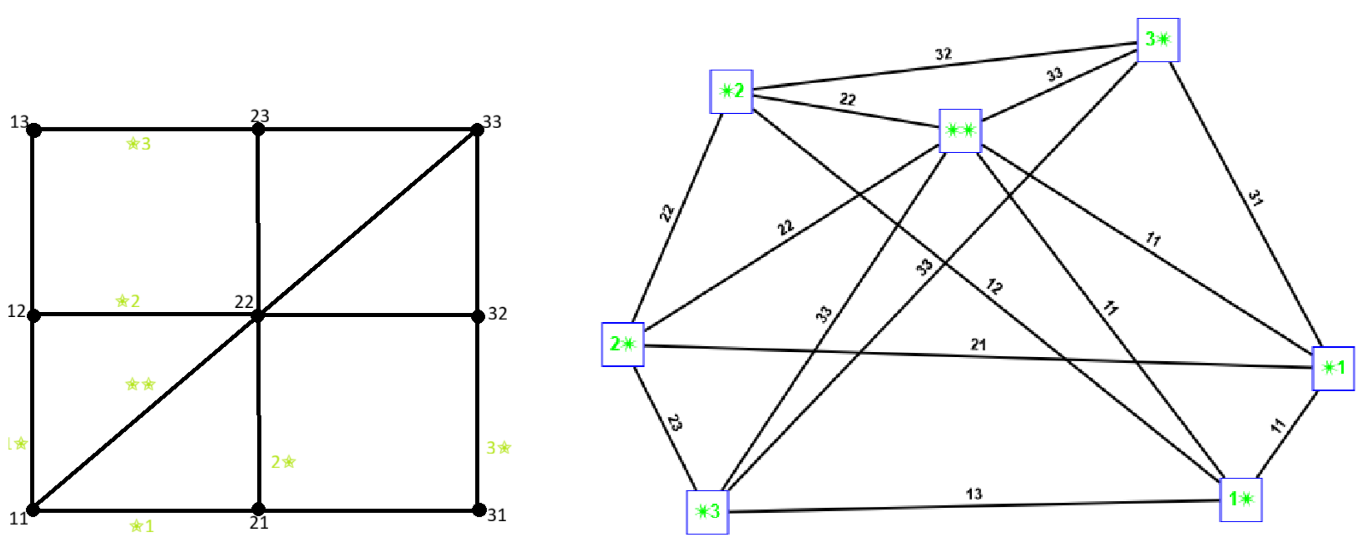


Figure 10: $[3]^2$ together with its combinatorial lines (left), the induced graph $G([3]^2])$ (right).

We now formalize the previous example and discuss it in the context of the previous section on Hales–Jewett numbers.

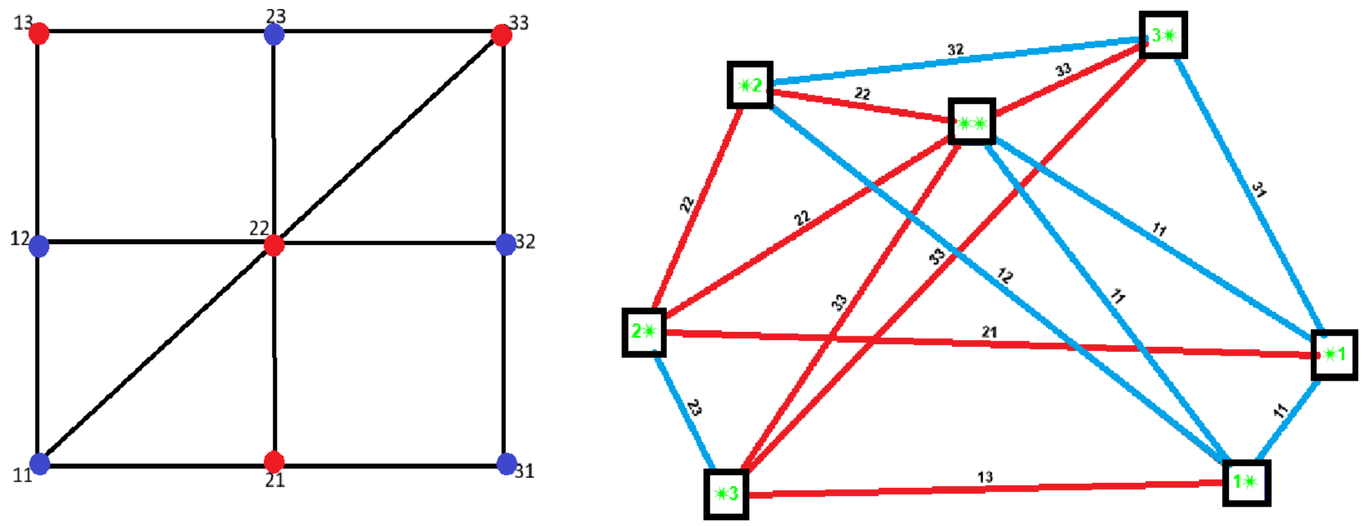


Figure 11: 2-coloring of $[3]^2$ with no monochromatic combinatorial line (left). Corresponding induced coloring on $G([3]^2])$: no monochromatic stars (right).

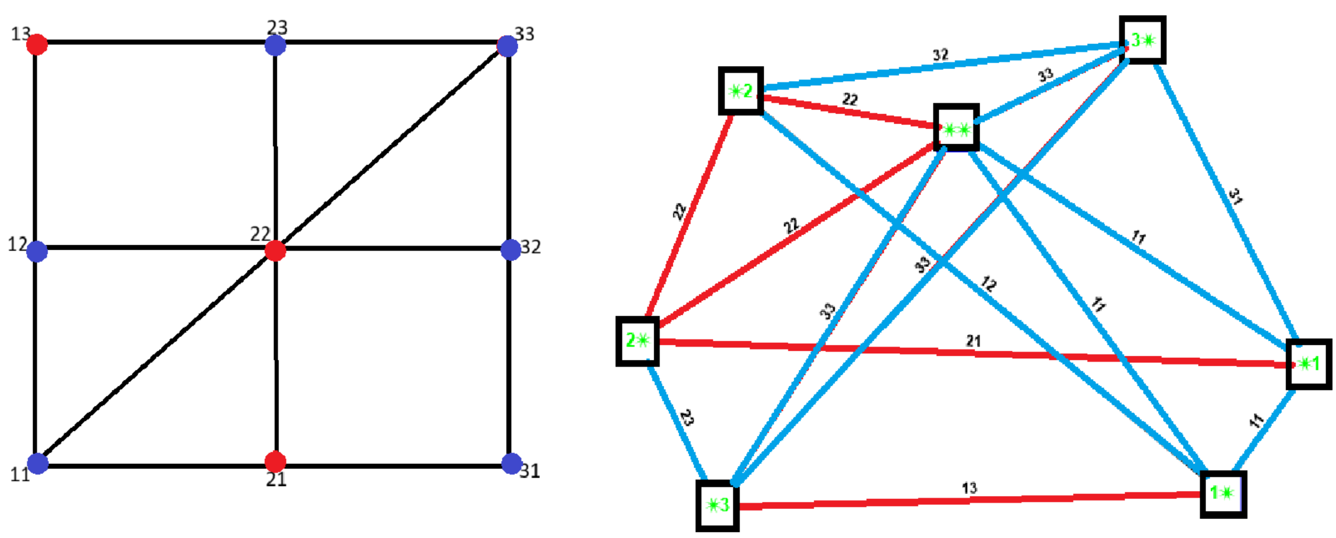


Figure 12: 2-coloring of $[3]^2$ with $3*$ being monochromatic (left). Corresponding induced coloring on $G([3]^2])$: with a monochromatic star (right).

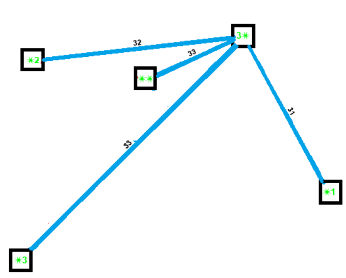

Figure 13: monochromatic star that corresponds to the combinatrial line $3*$.

## 3.2 General construction

In general, to any hypercube $[t]^n$ together with its combinatorial lines $(L([t]^n))$, we define the graph $G([t]^n) := (V, E)$, $V = L([t]^n)$ and $l_1 l_2$ is an edge in $E$ if, and only if, $l_1$ and $l_2$ intersect in the hypercube $[t]^n$.

Moreover, we define the following grpahs associated to $[t]^n$: $G^{[k]}([t]^n)$ is the graph induced by looking only at the lines in $L^{[k]}([t]^n)$, and $G^{(q)}([t]^n)$ is the graph induced by looking only at $q$-fold lines in $L^{(q)}([t]^n)$.

**Example 7.** In figure 14, we can see that removing the root $**$ (i.e. by looking at $L^{[1]}([3]^2)$ instead of $L([3]^2)$), the induced graph is $3-$regular graph.

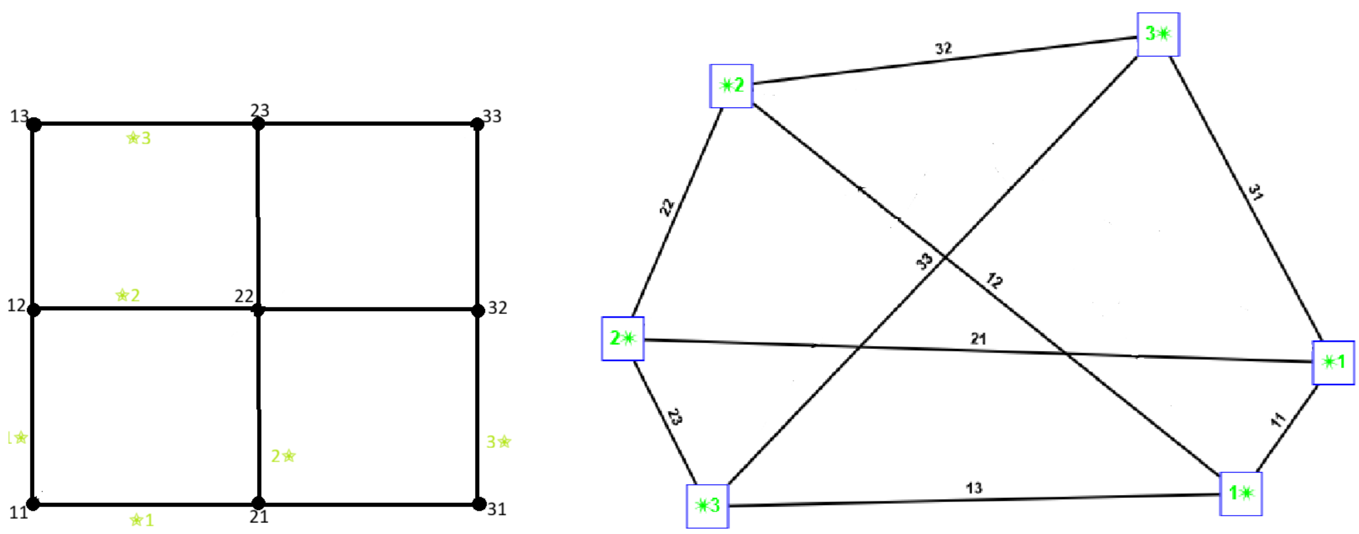


Figure 14: $[3]^2$ with the lines from $L^{[1]}([3]^2)$ (left). induced graph $G^{[1]}([3]^2])$ (right).

**Remark.** We notice that the graphs obtained by the induced graphs on sets of lines are in general very sparse, moreover for certain sets, the conditions on lines impose strong conditions on the graph as well as we saw in the previous example with regularity.

## 3.3 Traduction of the previously defined Hales-Jewett Numbers in terms of this construction

This is as far as we have gotten in this semester project, but many questions have been raised, and this work provides a good starting point, since the formalism and definitions are now fixed and could help anyone that wishes to explore this connection between lines in hypercubes and star-like structures in some defined sparse graphs rigorously.

# References

[1] A. W. Hales and R. Jewett, "Regularity and positional games," in T. A. M. Soc, Ed. 106(2):222-229, 1935.

[2] S. Shelah, "Primitive recursive bounds for van der waerden numbers," *Journal of the American Mathematical Society*, vol. 1, no. 3, pp. 683–697, Jul. 1988, ISSN: 0894-0347. DOI: 10.1090/S0894-0347-1988-0929498-X.

[3] M. Golshani and S. Shelah, *Iterated ramsey bounds for the hales-jewett numbers*, 2022. arXiv: 1912.08643 [math.CO]. [Online]. Available: https://arxiv.org/abs/1912.08643.

[4] N. Hindman and E. Tressler, *The first nontrivial hales-jewett number is four*, 2014. ArsCombinatoria: volume113.385-390.

[5] N. Kamčev and C. Spiegel, *Another note on intervals in the hales-jewett theorem*, 2018. arXiv: 1811.04628 [math.CO]. [Online]. Available: https://arxiv.org/abs/1811.04628.

[6] H. Prömel, *Ramsey Theory for Discrete Structures*. Dec. 2013, ISBN: 978-3-319-01314-5. DOI: 10.1007/978-3-319-01315-2.

[7] N. Hindman, D. Strauss, and L. Q. Zamboni, *Combining extensions of the hales-jewett theorem with ramsey theory in other structures*, 2018. arXiv: 1807.01485 [math.CO]. [Online]. Available: https://arxiv.org/abs/1807.01485.

[8] *Stackexchang*. [Online]. Available: https://math.stackexchange.com/questions/2203520/proof-of-hales-jewett-number.

[9] M. Lavrov, "An upper bound for the hales–jewett number $HJ(4,2)$," *SIAM Journal on Discrete Mathematics*, vol. 30, no. 2, pp. 1333–1342, 2016. DOI: 10.1137/15M1016485.

[10] D. Conlon and N. Kamcev, *Intervals in the hales-jewett theorem*, 2018. arXiv: 1801.08919 [math.CO]. [Online]. Available: https://arxiv.org/abs/1801.08919.

[11] I. Leader and E. Raty, *A note on intervals in the hales-jewett theorem*, 2018. arXiv: 1802.03087 [math.CO]. [Online]. Available: https://arxiv.org/abs/1802.03087.

[12] *Wikipedia graham-rothschild theorem*. [Online]. Available: https://en.wikipedia.org/wiki/Graham%E2%80%93Rothschild_theorem.

[13] R.L.Graham and B.L.Rothschild, "Ramsey's theorem for $n$-parameter sets," *Trans.Amer.Math.Soc.*, vol. 159, no. 257-292, 1971. DOI: https://doi.org/10.1090/S0002-9947-1971-0284352-8.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

## Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. **In consultation with the supervisor**, one of the following two options must be selected:

☐ I hereby declare that I authored the work in question independently, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].

☒ I hereby declare that I authored the work in question independently. In doing so I only used the authorised aids, which included suggestions from the supervisor regarding language and content and generative artificial intelligence technologies. The use of the latter and the respective source declarations proceeded in consultation with the supervisor.

**Title of paper or thesis**:

From Hypercubes to Sparse Graphs:
Structural Translation of the Hales-Jewett theorem via monochromatic stars

**Authored by**:
*If the work was compiled in a group, the names of all authors are required.*

| **Last name(s):** | **First name(s):** |
|---|---|
| Mouhib | Younes |

With my signature I confirm the following:

- I have adhered to the rules set out in the Citation Guidelines.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

| **Place, date** | **Signature(s)** |
|---|---|
| Zürich, 13.12.2025 | |

*If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.*

[1] For further information please consult the ETH Zurich websites, e.g. https://ethz.ch/en/the-eth-zurich/education/ai-in-education.html and https://library.ethz.ch/en/researching-and-publishing/scientific-writing-at-eth-zurich.html (subject to change).