The Hales-Jewett Theorem

1. **Quote.** The Hales–Jewett theorem strips van der Waerden's theorem of its unessential elements and revels the heart of Ramsey theory.

   Ron Graham, Bruce Rothschild, and Joel Spencer

2. **The Hales-Jewett Theorem.** Let $m, k \in \mathbb{N}$ and let A be an symbols on m symbols. There exists an $n \in \mathbb{N}$ such that whenever A^n is k-coloured there exists a monochromatic line.

   Note: The smallest such n is denoted by $HJ(m,k)$.

   Proof:

   (a) **Settings:** Let $m, k \in \mathbb{N}$. As an alphabet on m symbols we take $A = [1, m]$.

   (b) **Reminder:** A root $\tau \in [1,m]_*^n$ is an n-word on $m+1$ symbols, $1, 2, \ldots, m$ and $*$, that contains the symbol $*$. A combinatorial line in $[1,m]^n$ rooted in τ is the set of words

   $$L_\tau = \{\tau_a : a \in [1,m]\}.$$

   Here, for $a \in [1,m]$ and $i \in [1,n]$,

   $$\tau_a(i) = \begin{cases} \tau(i) & \text{if} \quad \tau(i) \neq *, \\ a & \text{if} \quad \tau(i) = *. \end{cases}$$

   (c) **Focussed and Colour-Focussed Lines:**

   - Let $r \in \mathbb{N}$ and let and $\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(r)} \in [1,m]_*^n$ be r roots. We say that the corresponding combinatorial lines are focussed at $f \in [1,m]^n$ if

     $$\tau_m^{(1)} = \tau_m^{(2)} = \cdots = \tau_m^{(r)} = f.$$

     Example: Consider $\tau^{(1)}, \tau^{(2)}, \tau^{(3)} \in [1,4]_*^4$ given by

     $$\tau^{(1)} = * \ * \ 3 \ *, \ \tau^{(2)} = * \ 4 \ 3 \ *, \ \tau^{(3)} = * \ 4 \ 3 \ 4.$$

     Then

     $$\tau_4^{(1)} = \square \ \square \ 3 \ \square, \ \tau_4^{(2)} = \square \ 4 \ 3 \ \square, \ \tau_4^{(3)} = \square \ 4 \ 3 \ 4.$$

     Hence the corresponding combinatorial lines are focussed at $f = 4\ 4\ 3\ 4$:

$L_{\tau^{(1)}}$				$L_{\tau^{(2)}}$				$L_{\tau^{(3)}}$			
1	1	3	1	1	4	3	1	1	4	3	4
2	2	3	2	2	4	3	2	2	4	3	4
3	3	3	3	3	4	3	3	3	4	3	4
4	4	3	4	4	4	3	4	4	4	3	4

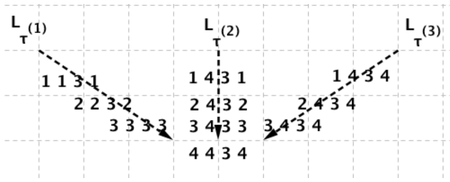


Figure 1: Three focussed lines in $[1,4]^4$.

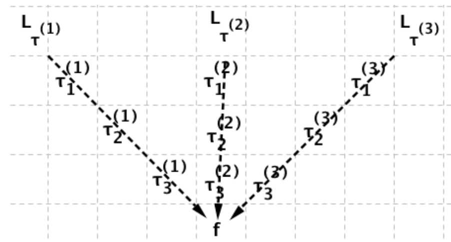


Figure 2: Three lines in $[1,4]^4$ focussed at f.

- **Colour–focussed combinatorial lines.** Let c be a k-colouring of $[1,m]^n$ and let

$$\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(r)} \in [1,m]^n_*$$

be r roots. We say that r combinatorial lines $L_{\tau^{(1)}}, L_{\tau^{(2)}}, \ldots, L_{\tau^{(r)}}$ are colour–focussed if:

  i. For each $i \in [1,r]$, $c(\tau_1^{(i)}) = c(\tau_2^{(i)}) = \cdots = c(\tau_{m-1}^{(i)})$.
  ii. For each $i,j \in [1,r]$, if $i \neq j$ then $c(\tau_1^{(i)}) \neq c(\tau_1^{(j)})$.
  iii. Combinatorial lines $L_{\tau^{(1)}}, L_{\tau^{(2)}}, \ldots, L_{\tau^{(r)}}$ are focussed at some $f \in [1,m]^n$.

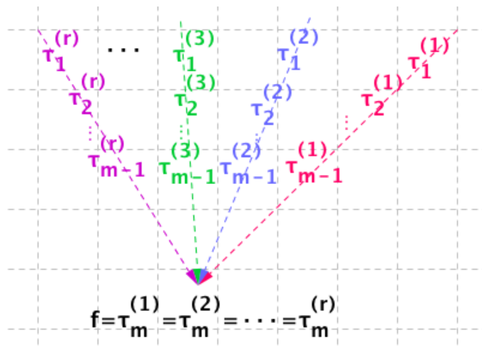


Figure 3: r colour-focussed lines: different colours and $\tau_m^{(1)} = \tau_m^{(2)} = \cdots = \tau_m^{(r)}$.

(d) **Strategy.** Induction on m.

(e) **Reminder: The Hales-Jewett Theorem.** Let $m,k \in \mathbb{N}$ and let A be an alphabet on m symbols. There exists an $n \in \mathbb{N}$ such that whenever A^n is k-coloured there exists a monochromatic line.

(f) **Base Case.** If $m = 1$ then $H(1,k) = 1$ for any number of colours k.

(g) **Inductive step.** Given $m > 1$, we assume that $HJ(m-1,k)$ exists for all k.

i. **Claim.** For all $1 \leq r \leq k$, there exists n such that whenever $[1,m]^n$ is k coloured, there exists **either** a monochromatic line **or** r colour-focussed lines.

ii. **Base Case.** Let $k \in \mathbb{N}$ and let $r = 1$. We take $n = HJ(m-1,k)$. Le c be a k-colouring of $[1,m]^n$.

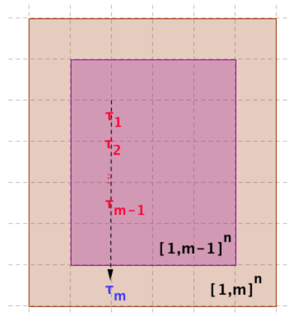
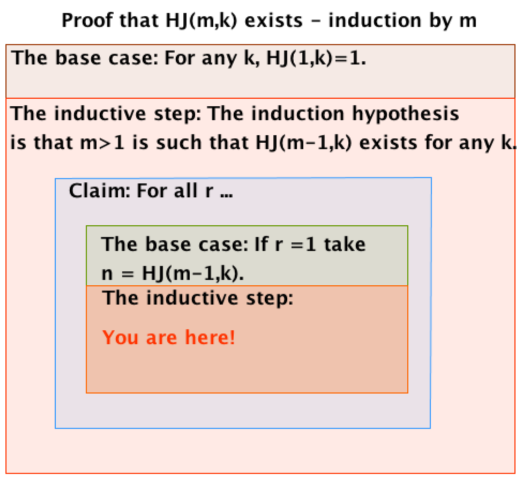


Figure 4: The colouring c of $[1,m]^n$ induces a k-colouring of $[1,m-1]^n$. Our choice of n guarantees the existence of a monochromatic line in $[1,m-1]^n$.

iii. **Where Are You?**



iv. **Inductive Step.** Let $r \in [1, k-1]$ and let $n = n(r)$ be such that whenever $[1, m]^n$ is k coloured, there exists **either** a monochromatic line **or** r colour-focussed lines.

Let $n' = HJ(m-1, k^{m^n})$ and let $N = n + n'$. Let c be a k-colouring of $[1, m]^N = [1, m]^{n+n'}$ **without a monochromatic line**.

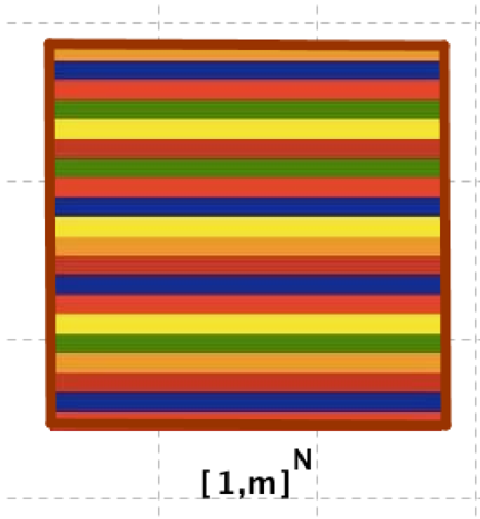


Figure 5: The k-colouring c of $[1, m]^N = [1, m]^{n+n'}$ **without a monochromatic line**.

A. **A** c **induced** k^{m^n}**-colouring of** $[1, m-1]^{n'}$**:**
Step 1.

Figure 6: Choose $b = b_1 b_2 \cdots b_{n'} \in [1, m-1]^{n'}$. Consider c_b, a k-colouring of $[1, m]^n$ such that for $a \in [1, m]^n$, $c_b(a) = c(ab)$.

Step 2. Note that there are k^{m^n} k-colourings of $[1,m]^n$.

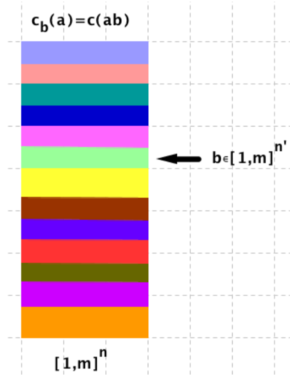


Figure 7: The mapping $\chi : b \mapsto c_b$ is a k^{m^n}-colouring of $[1,m]^{n'}$.

Step 3. There is a χ-monochromatic line in $[1,m-1]^{n'}$.

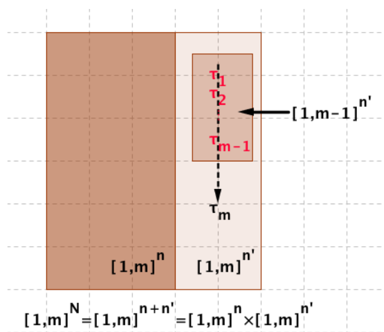


Figure 8: There is a χ-monochromatic line L_τ in $[1,m-1]^{n'}$.

B. **Reminder - Inductive Step.** Let $r \in [1,k-1]$ and let $n = n(r)$ be such that whenever $[1,m]^n$ is k coloured, there exists **either** a monochromatic line **or** r colour-focussed lines.

Let $n' = HJ(m-1, k^{m^n})$ and let $N = n + n'$. Let c be a k-colouring of $[1,m]^N = [1,m]^{n+n'}$ **without a monochromatic line**.

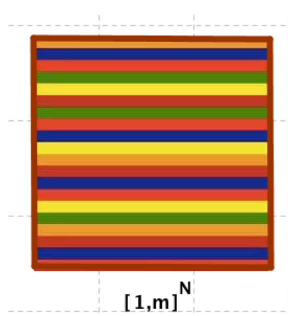


Figure 9: The k-colouring c of $[1,m]^N = [1,m]^{n+n'}$ **without a monochromatic line**.

C. **A c induced k-colouring of $[1,m]^n$:**

Step 1 There is a χ-monochromatic line in $[1,m-1]^{n'}$.

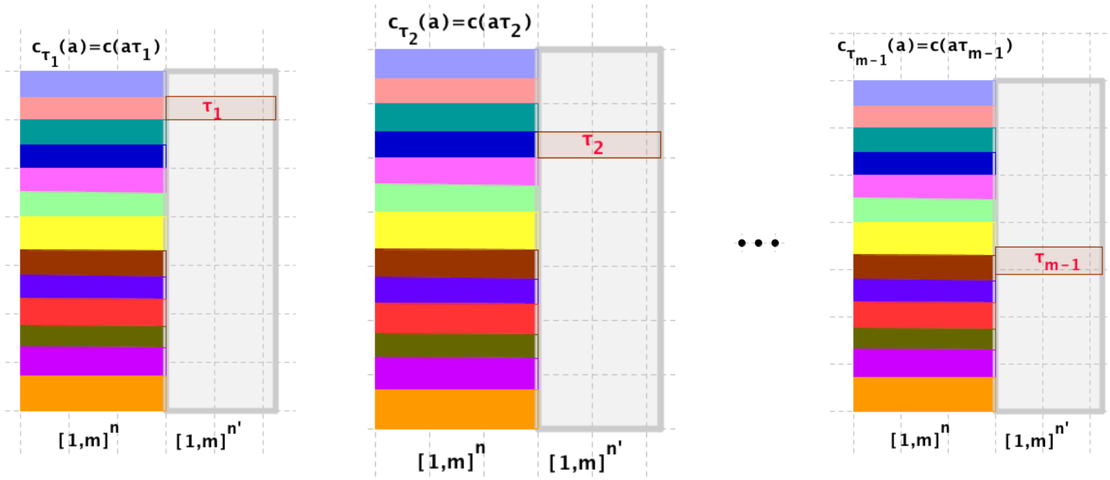


Figure 10: L_τ is monochromatic: $c_{\tau_1} = c_{\tau_2} = \cdots = c_{\tau_{m-1}}$

Step 2. A k-colouring c_τ of $[1,m]^n$ emerges:

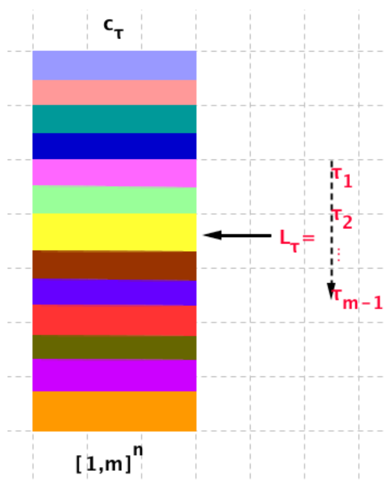


Figure 11: The k-colouring c_τ of $[1,m]^n$ is with the property that, for any $a \in [1,m]^n$ and any $i \in [1,m-1]$, $c_\tau(a) = c(a\tau_i)$.

Step 3. Back to colour-focussed lines:

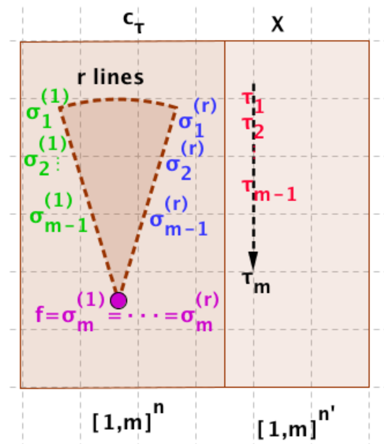


Figure 12: There are r c_τ-coloured-focussed lines $L_{\sigma^{(1)}}, \ldots, L_{\sigma^{(r)}}$ in $[1,m]^n$ with the focus f and one χ-monochromatic line $L\tau$ in $[1,m]^{n'}$ with the focus τ_m. None of the lines $L_{\sigma^{(1)}}, \ldots, L_{\sigma^{(r)}}$ is monochromatic.

D. **Making new roots from old:** We define $r+1$ roots in $[1,m]_*^N$ as follows:

$$\tau^{(1)} = \sigma^{(1)}\tau,\ \tau^{(2)} = \sigma^{(2)}\tau, \ldots, \tau^{(r)} = \sigma^{(r)}\tau,\ \tau^{(r+1)} = f\tau.$$

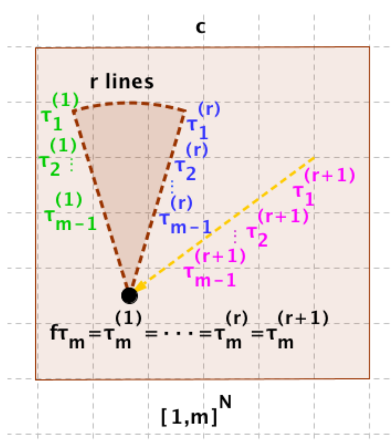


Figure 13: There are $r+1$ c-coloured-focussed lines $L_{\tau^{(1)}}, \ldots, L_{\tau^{(r+1)}}$ in $[1,m]^N$ with the focus $f\tau_m$.

E. Where Are You?

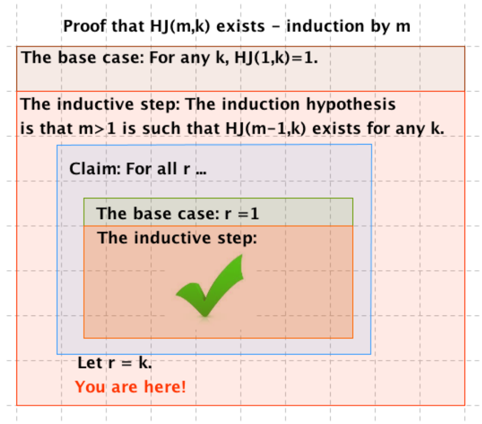


Figure 14: You are here!

F. **Let $r = k$.**

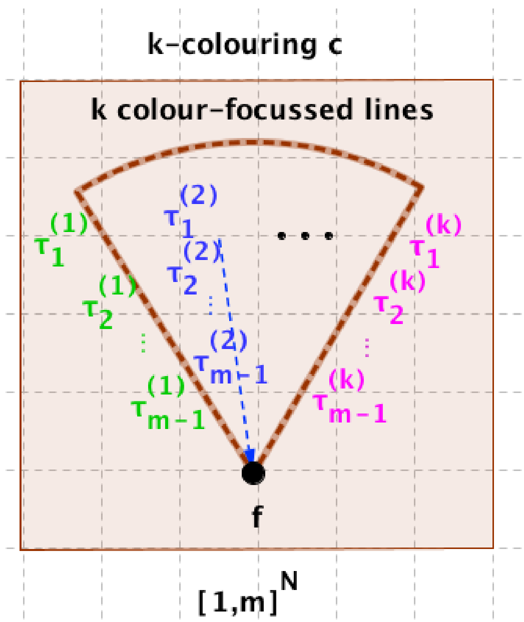


Figure 15: What is the colour of the focus f? There is a monochromatic line!

G. **Done!**

$$HJ(m-1,k) \text{ exists } \Rightarrow HJ(m,k) \text{ exists}$$

3. **The Hales-Jewett Theorem.** Let $m, k \in \mathbb{N}$ and let A be an alphabet on m symbols. There exists an $n \in \mathbb{N}$ such that whenever A^n is k-coloured there exists a monochromatic line.

4. **Exercise:** Use the Hales-Jewett theorem to prove van der Waerden's theorem.

   Solution: Let $l, k \in \mathbb{N}$ be given. Let $c : \mathbb{N} \to \{1, 2, \ldots, k\}$ be a k-colouring of the set of natural numbers. Let $N = HJ(l, k)$.

   We define a k-colouring of the N-cube $[1, l]^N$ as follows

   $$c'(x_1 x_2 \cdots x_N) = c(x_1 + x_2 + \ldots + x_N),\ x_1 x_2 \cdots x_N \in [1, l]^N.$$

   By the Hales-Jewett theorem there is a c'-monochromatic line rooted in the root $\tau \in [1, l]_*^N$. Let $S \subset [1, N]$ be such that

   $$\tau(i) \in [1, l] \text{ if } i \in S \text{ and } \tau(i) = * \text{ if } i \in [1, l] \backslash S.$$

   Let

   $$a = \sum_{i \in S} \tau(i) \text{ and } d = |[1, l] \backslash S|.$$

   Note that

   $$\begin{aligned} \sum_{i=1}^{N} \tau_1(i) &= \sum_{i \in S} \tau_1(i) + \sum_{i \in [1,l] \backslash S} \tau_1(i) = a + \sum_{i \in [1,l] \backslash S} 1 = a + d \\ \sum_{i=1}^{N} \tau_2(i) &= \sum_{i \in S} \tau_2(i) + \sum_{i \in [1,l] \backslash S} \tau_2(i) = a + \sum_{i \in [1,l] \backslash S} 2 = a + 2d \\ &\vdots \\ \sum_{i=1}^{N} \tau_l(i) &= \sum_{i \in S} \tau_l(i) + \sum_{i \in [1,l] \backslash S} \tau_l(i) = a + \sum_{i \in [1,l] \backslash S} l = a + ld. \end{aligned}$$

   On the other hand

   $$c'(\tau_1) = c'(\tau_2) = \cdots = c'(\tau_l)$$

   which together with

   $$c'(\tau_j) = c\left(\sum_{i=1}^{N} \tau_j(i) \right) = c(a + jd),\ \text{for each } i \in [1, l],$$

   implies that $c(a + d) = c(a + 2d) = \cdots = (a + ld)$. Thus, there is a c-monochromatic l-term arithmetic progression.

5. **Exercise. Gallai–Witt theorem.** Let $(V, +, \cdot)$ be a real vector space and let $A = \{a_1, a_2, \ldots, a_m\}$ be a finite subset of V. Prove that, for any r–colouring of the vector space V, there exists a vector $u \in V$ such that the set $u + \lambda \cdot A = \{u + \lambda \cdot a_1, u + \lambda \cdot a_2, \ldots, u + \lambda \cdot a_m\}$ is monochromatic.