Purely Combinatorial Proofs of Van Der Waerden-Type Theorems

Andy Parrish

December 17, 2007

1 Introduction

In this monograph we give purely combinatorial proofs of several known theorems in Ramsey Theory. They are all be variants of, or generalizations of, van der Waerden's theorem.

Notation 1.1 If $m \in \mathbb{N}$ then $[m]$ is $\{1, \ldots, m\}$.

Def 1.2 If $k \in \mathbb{N}$ then a k-AP is an arithmetic progression of length k.

Van Der Waerden's Theorem: For every $k \geq 1$ and $c \geq 1$ there exists $W = W(k,c)$ such that for every c-coloring $COL : [W] \to [c]$ there exists a monochromatic k-AP. In other words there exists $a, d \in [W]$ such that

$$\{a\} \cup \{a+d, \ldots, a+(k-1)d\} \subseteq [W],$$

$$COL(a) = COL(a+d) = \cdots = COL(a+(k-1)d).$$

We will proof van der Waerden's theorem in Section 3.2.2.

In van der Waerden's theorem we can think of

$$a, a+d, \ldots, a+(k-1)d$$

as

$$a, a+p_1(d), \ldots, a+p_{k-1}(d)$$

where $p_i(x) = ix$. Why these functions?

The following remarkable theorem was first proved by Bergelson and Leibman [1].

Polynomial Van Der Waerden Theorem For any natural number c and any polynomials $p_1(x), \ldots, p_k(x) \in \mathbb{Z}[x]$ such that $(\forall i)[p_i(0) = 0]$, there exists $W = W(p_1, \ldots, p_k; c)$ such that, for any c-coloring $COL : [W] \rightarrow [c]$ there exists a $a, d \in [W]$ such that

$$\{a\} \cup \{a + p_1(d), \ldots, a + p_k(d)\} \subseteq [W]$$

$$COL(a) = COL(a + p_1(d)) = COL(a + p_2(d)) = \cdots = COL(a + p_k(d)).$$

This was proved for $k = 1$ by Furstenberg [2] and (independently) Sarkozy [3]. The original proof of the full theorem by Bergelson and Leibman [1] used ergodic methods. A later proof by Walters [4] uses purely combinatorial techniques. We will present an expanded version of Walters' proof in Section 3.2.2.

The Hales-Jewett theorem is a generalization of van der Waerden's theorem which we will state and prove in Section 5. There is also a polynomial Hales-Jewett theorem, which we will state and prove in Section 6.

The following is a corollary of the Hales-Jewett Theorem:

The Coloring Square Theorem: For any c there exists $W = W(c)$ such that for any c-coloring of $[W] \times [W]$ there exists a square with all four corners the same colors. Formally: for any c-coloring $COL : [W] \times [W] \rightarrow [c]$ there exists $a_1, a_2, d \in [W]$ such that

$$\{(a_1, a_2), (a_1, a_2 + d), (a_1 + d, a_2), (a_1 + d, a_2 + d)\} \subseteq [W] \times [W],$$

$$COL(a_1, a_2) = COL(a_1, a_2 + d) = COL(a_1 + d, a_2) = COL(a_1 + d, a_2 + d).$$

The following is a corollary of the Polynomial Hales-Jewett Theorem:

The Coloring Squared Rectangle Theorem: For any c there exists $W = W(c)$ such that for any c-coloring of $[W] \times [W]$ there exists $d \in [W]$, $d \neq 0$, such that there is a $d \times d^2$ rectangle with all four corners the same color. Formally: for any c-coloring $COL : [W] \times [W] \rightarrow [c]$ there exists $a_1, a_2, d \in [W]$ such that

$$\{(a_1, a_2), (a_1, a_2 + d^2), (a_1 + d, a_2 + d^2), (a_1 + d, a_2 + d^2)\} \subseteq [W] \times [W],$$

$$COL(a_1, a_2) = COL(a_1, a_2 + d^2) = COL(a_1 + d, a_2) = COL(a_1 + d, a_2 + d^2).$$

The next corollary requires some discussion. The polynomial van der Waerden theorem is equivalent (by a compactness argument) to the following statement:

Polynomial Van Der Waerden Theorem over $\mathbb{Z}$: For any natural number c and any polynomials $p_1(x), \ldots, p_k(x) \in \mathbb{Z}[x]$ such that $(\forall i)[p_i(0) = 0]$, for any c-coloring $COL : [\mathbb{Z}] \rightarrow [c]$ there exists a $a, d \in \mathbb{Z}$, $d \neq 0$, such that

$$COL(a) = COL(a + p_1(d)) = COL(a + p_2(d)) = \cdots = COL(a + p_k(d)).$$

Note that the domain $\mathbb{Z}$ appears three times— the polynomials have coefficients in $\mathbb{Z}$, the c-coloring is of $\mathbb{Z}$, and $a, d \in \mathbb{Z}$. What if we replaced $\mathbb{Z}$ by $\mathbb{R}$ or some other integral domain?

The following is a corollary of the Polynomial Hales-Jewett Theorem.

Generalized Polynomial Van Der Waerden Theorem: Let S be any infinite commutative integral domain. For any natural number c and any polynomials $p_1(x), \ldots, p_k(x) \in S[x]$ such that $(\forall i)[p_i(0) = 0]$, for any c-coloring $COL : [S] \rightarrow [c]$ there exists $a, d \in S$, $d \neq 0$, such that

$$COL(a) = COL(a + p_1(d)) = COL(a + p_2(d)) = \cdots = COL(a + p_k(d)).$$

The motivation for this monograph is Walters' paper [4]. He gave the first purely combinatorial proofs of the polynomial van der Waerden Theorem and the Polynomial Hales-Jewett Theorem. However, his techniques can be used to obtain cleaner proofs of the original van der Waerden and Hales-Jewett Theorems. In addition, there are many corollaries of these theorems that, because of his work, now have purely combinatorial proofs. In this monograph we will present (1) clean proofs of van der Waerden's theorem and the Hales-Jewett theorem, (2) an expanded version of Walters' proofs of the polynomial van der Waerden theorem and polynomial Hales-Jewett Theorem, and (3) variants, corollaries, and applications of these results.

Since this is an exposition there will be more figures and examples than is common in a mathematics paper.

2 Technical Summary

Section 3 is on van der Waerden's theorem. We prove it, give upper and lower bounds on the van der Waerden numbers, and give some applications. Section 4 is on the Polynomial van der Waerden's theorem. We prove it, give upper and lower bounds on the polynomial van der Waerden numbers. We do not give any applications because we do not know any.

Section 5 is on the Hales-Jewett Theorem. We prove it two ways. The classical proof yields insanely large upper bounds on the Hales-Jewett numbers. The proof by Shelah gives sanely large upper bounds. We discuss both upper and lower bounds for the Hales-Jewett numbers and then give applications. One of the applications is another proof of van der Waerden's theorem with better bounds. Section 6 is on the Polynomial Hales-Jewett Theorem. We prove it, give upper and lower bounds on the polynomial Hales-Jewett numbers, and give many applications.

Note that this monograph is only about results that can be proven using purely combinatorial techniques. There are many results in this field for which the only proofs known use non-combinatorial techniques.

3 Van der Waerden's theorem

3.1 Introduction

Van der Waerden's theorem states that given any number of colors, c, and any length, k, there is (large) number W so that any c-coloring of the numbers $\{1, 2, \ldots, W\}$ contains a monochromatic k-AP.

Van der Waerden's theorem: For all $k, c \in \mathbb{N}$ there exists $W = W(k, c)$ such that, for all c-colorings $COL : [W] \to [c]$, there exists $a, d \in \mathbb{N}$, $d \neq 0$, such that

$$COL(a) = COL(a + d) = COL(a + 2d) = \cdots = COL(a + (k - 1)d)$$

Example 3.1 The coloring of $\{1, 2, \ldots, 10\}$ given by

1	2	3	4	5	6	7	8	9	10
R	R	B	B	R	R	B	B	B	R

has a blue 3-AP with the numbers {7, 8, 9}, and a red 3-AP with the numbers {2, 6, 10}. It has no monochromatic 4-AP.

Before attempting to prove the full theorem, let's look at a few simple base cases.

- $c = 1$ — $W(k, 1) = k$, because the sequence $1, 2, \ldots, k$ forms a k-AP.
- $k = 1$ — $W(1, c) = 1$, because a 1-AP is any single term.
- $k = 2$ — $W(2, c) = c + 1$, because any 2 terms form a 2-AP.

Alright, not bad so far — we have proven the theorem for ω cases. How many more could there be?

3.2 Proof of van der Waerden's Theorem

3.2.1 The First Interesting Case: $W(3, 2)$

We show that there exists a W such that any 2-coloring of $[W]$ has a monochromatic 3-AP.

We leave the proofs of the following facts to the reader.

Fact 3.2

1. *Let B be a block of 5 consecutive numbers. Let $COL : B \to [2]$ be a 2-coloring of B. Then there exists a, d such that*

$$\{a, a+d, a+2d\} \subseteq B$$

$$COL(a) = COL(a+d)$$

We make no comment on $COL(a+2d)$.

R R ?

Figure 1: A block of five numbers

2. *Let $W = 5(2 \times 2^5 + 1) = 325$. We view W as $(2 \times 2^5 + 1) = 65$ blocks of size 5 which we denote*

$$B_1 B_2 \cdots B_{325}.$$

Let $COL : [W] \to [2]$ be a 2-coloring of $[W]$. **We view COL as a 32-coloring of the blocks. We will use this change of viewpoint over and over again in this paper!** *Let COL^* be the induced 32-coloring of the blocks. Then there exists A, D such that*

$$\{A, A+D, A+2D\} \in [65]$$

$$COL^*(B_A) = COL^*(B_{A+D}).$$

We make no comment on $COL^*(B_{A+2D})$.

R B B R B	R B B R B	? ? ? ? ?

Figure 2: 65 blocks of 5

Theorem 3.3 *Let $W = 5(2 \times 2^5 + 1) = 325$. Let $COL : [W] \to [2]$ be a 2-coloring of $[W]$. Then there exists $a, d \in [W]$ such that*

$$\{a, a+d, a+2d\} \subseteq [W],$$

$$COL(a) = COL(a+d) = COL(a+2d).$$

Proof: We take the colors to be RED and BLUE.

View $[W]$ as being 65 blocks of 5. We denote the blocks

$$B_1 B_2 \cdots B_{65}.$$

Let COL^* be the induced 32-coloring of the blocks. By Fact 3.2.2 there exists A, D such that

$$\{A, A+D, A+2D\} \subseteq [65]$$
$$COL^*(B_A) = COL^*(B_{A+D}).$$

By Fact 3.2.1 there exists a, d such that $a \in B_A$, $d \neq 0$, and $a+d, a+2d \in B_A$.

$$COL(a) = COL(a+d).$$

We will assume the color is RED. Since $COL(a) = COL(a+d)$ and $COL^*(B_A) = COL^*(B_{A+D})$ we have

$$COL(a) = COL(a+d) = COL(a+D) = COL(a+d+D) = RED.$$

Since $COL^*(B_A) = COL^*(B_{A+D})$

$$COL(a+2d) = COL(a+2d+D).$$

We have made no claim about the color of $a+2d$.

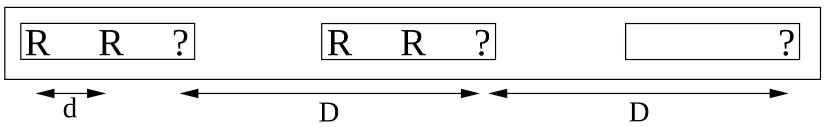

Figure 3: The distances between corresponding points

There are two cases.

1. If $COL(a+2d) = RED$ then $a, a+d, a+2d$ are a RED 3-AP.
2. If $COL(a+2d) = BLUE$ then $COL(a+2d+D) = BLUE$ also.
   (a) If $COL(a+2d+2D) = BLUE$ then $a+2d, a+2d+D, a+2d+2D$ are a BLUE 3-AP.
   (b) If $COL(a+2d+2D) = RED$ then $a, a+d+D, a+2d+2D$ are a RED 3-AP.

∎

Note 3.4 The proof of Theorem 3.3 yields $W(3,2) \leq 365$. A more careful proof using blocks of 3 can get $W(3,2) \leq 42$. A boring proof by cases shows that $W(3,2) = 9$.

3.2.2 The Full Proof

We will prove a lemma from which van der Waerden's theorem will follow easily.

Lemma 3.5 *Fix $k, c \in N$ with $k > 1$. Assume $(\forall c')[W(k-1, c')$ exists]. Then, for all r, there exists $U = U(k, c, r)$ such that for all c-colorings $COL : [U] \to [c]$, one of the following statements holds.*

Statement I: $\exists a, d \in \mathbb{N}, d \neq 0$ *such that*

$$\{a, a+d, a+2d, \ldots, a+(k-1)d)\} \subseteq [U]$$

$$COL(a) = COL(a+d) = COL(a+2d) = \cdots = COL(a+(k-1)d).$$

Statement II: $\exists a, d_1, d_2, \ldots, d_r \in \mathbb{N}, d_i \neq 0\ \forall i$, *such that*

$$\{a, a+d_1, a+2d_1, \ldots, a+(k-1)d_1\} \subseteq [U]$$

$$\{a, a+d_2, a+2d_2, \ldots, a+(k-1)d_2\} \subseteq [U]$$

$$\vdots$$

$$\{a, a+d_r, a+2d_r, \ldots, a+(k-1)d_r\} \subseteq [U]$$

$$COL(a+d_1) = COL(a+2d_1) = \cdots = COL(a+(k-1)d_1)$$

$$COL(a+d_2) = COL(a+2d_2) = \cdots = COL(a+(k-1)d_2)$$

$$\vdots$$

$$COL(a+d_r) = COL(a+2d_r) = \cdots = COL(a+(k-1)d_r)$$

With $COL(a+d_i) \neq COL(a+d_j)$ when $i \neq j$. We refer to a as the anchor. (Informally we are saying that if you c-color $[U]$ either you will have a monochromatic k-AP or you will have many monochromatic $(k-1)$-AP's, all of different colors, and different from a. Once "many" is more than c, then the latter cannot happen, so the former must, and we have van der Waerden's theorem.)

Proof:

We define $U(k,c,r)$ to be the least number such that this Lemma holds. We will prove $U(k,c,r)$ exists by giving an upper bound on it.

Base Case: If $r=1$ then $U = U(k,c,1) \le 2W(k-1,c)$. Let $COL : [U] \rightarrow [c]$ be a c-coloring of U. Consider COL restricted to the last half of U, which is of size $W(k-1,c)$. By the definition of $W(k-1,c)$ there exists $a'\{W(k-1,c),\ldots,2W(k-1,c)\}$ and $d \in [W(k-1,c)]$ such that

$$\{a', a'+d', a'+2d', \ldots, a'+(k-2)d'\} \subseteq \{W(k-1,c),\ldots,2W(k-1,c)\},$$

$$COL(a') = COL(a'+d') = COL(a'+2d') = \cdots = COL(a'+(k-2)d').$$

Let $a = a'-d$ and $d_1 = d'$. Clearly

$$COL(a+d_1) = COL(a+2d_1) = COL(a+3d_1) = \cdots COL(a+(k-1)d)$$

Note that we have a better bound than $d \in [W(k-1,c]$. We easily have $d \in [\lfloor W(k-1,c)/(k-1)\rfloor]$, though all we need is $d \in [W-1]$. Since $a' \ge [W(k-1,c)]$ and $d_1 \in [W-1]$ we have $a = a'-d \ge 1$.

$d < W(k-1,c)$, so $a'-d \ge 1$. Clearly $a' \le a \le U$, so $a' \in [U]$.

The first half of $[U]$ will contain the the anchor, hence (2) holds.

Induction Step: By induction, assume $U(k,c,r)$ exists. We will show that $U(k,c,r+1) \le 2U(k,c,r)W(k-1,c^{U(k,c,r)})$. Let

$$U = 2U(k,c,r)W(k-1,c^{U(k,c,r)}).$$

Let $COL : [U] \rightarrow [c]$ be an arbitrary c-coloring of $[U]$.

We view $[U]$ as being $U(k,c,r)W(k-1,c^{U(k,c,r)})$ numbers followed by $W(k-1,c^{U(k,c,r)})$ blocks of size $U(k,c,r)$. We denote these blocks by

$$B_1, B_2, \ldots, B_{W(k-1,c^{U(k,c,r)})}.$$

The key point here is that we view a c-coloring of the second half of $[U]$ as a $c^{U(k,c,r)}$-coloring of these blocks. Let COL^* be this coloring. By the definition of $W(k-1,c^{U(k,c,r)})$, we get a monochromatic $(k-1)$-AP of blocks. Hence we have A, D' such that.

$$COL^*(B_A) = COL^*(B_{A+D'}) = \cdots = COL^*(B_{A+(k-2)D'})$$

Now, consider the elements of block B_A. It has length $U(k,c,r)$, which tells us that either (1) or (2) from the lemma holds. If (1) holds — we have a monochromatic k-AP — then we are done. If not, then we have the following: $a', d_1, d_2, \ldots, d_r$ with $a \in B_A$, and

$$\begin{gathered}
\{a' + d_1, a' + 2d_1, \ldots, , a' + (k-1)d_1 \subseteq B_A \\
\{a' + d_2, a' + 2d_2, \ldots, , a' + (k-1)d_2 \subseteq B_A \\
\vdots \\
\{a' + d_r, a' + 2d_r, \ldots, , a' + (k-1)d_r \subseteq B_A
\end{gathered}$$

$$\begin{gathered}
COL(a' + d_1) = COL(a' + 2d_1) = \cdots = COL(a' + (k-1)d_1) \\
COL(a' + d_2) = COL(a' + 2d_2) = \cdots = COL(a' + (k-1)d_2) \\
\vdots \\
COL(a' + d_r) = COL(a' + 2d_r) = \cdots = COL(a' + (k-1)d_r)
\end{gathered}$$

where $COL(a' + d_i)$ are all different colors, and different from a' (or else there would already be a monochromatic k-AP). How far apart are corresponding elements in adjacent blocks? Since the blocks viewed as points are D' apart, and each block has $U(k,c,r)$ elements in it, correspoinding elements in adjacent blocks are $D = D' \times U(k,c,r)$ numbers apart. Hence

$$\begin{gathered}
COL(a' + d_1) = COL(a' + D + d_1) = \cdots = COL(a' + (k-2)D + d_1) \\
COL(a' + d_2) = COL(a' + D + d_2) = \cdots = COL(a' + (k-2)D + d_2) \\
\vdots \\
COL(a' + d_r) = COL(a' + D + d_r) = \cdots = COL(a' + (k-2)D + d_r)
\end{gathered}$$

We now note that we have only worked with the second half of $[U]$. Since we know that

$$a > \frac{1}{2}U = U(k,c,r)W(k-1, c^{U(k,c,r)})$$

and

$$D \leq \frac{1}{k-1}U(k,c,r)W(k-1, c^{U(k,c,r)}) \leq U(k,c,r)W(k-1, c^{U(k,c,r)})$$

so we find that $a = a' - D > 0$ and thus $a \in [U]$.

So now we have

$$COL(a+(D+d_1)) = COL(a+2(D+d_1)) = \cdots = COL(a+(k-1)(D+d_1))$$
$$COL(a+(D+d_2)) = COL(a+2(D+d_2)) = \cdots = COL(a+(k-1)(D+d_2))$$
$$\vdots$$
$$COL(a+(D+d_r)) = COL(a+2(D+d_r)) = \cdots = COL(a+(k-1)(D+d_r))$$

With each sequence a different color.

We need an $(r+1)$st monochromatic set of points. Consider

$$\{a+D, a+2D, \ldots, a+(k-1)D\}.$$

These are correspoinding points in blocks that are colored (by COL^*) the same, hence

$$COL(a+D) = COL(a+2D) = \cdots = COL(a+(k-1)D)).$$

In addition, since

$$(\forall i)[COL(a') \neq COL(a'+d_i)]$$

the color of this new sequence is different from the r sequences above.

Hence we have $r+1$ monochromatic $(k-1)$-AP's, all of different colors,and all with projected first term a. Formally the new parameters are $a, (D+d_1), \ldots, (D+d_r)$, and D. ∎

Theorem 3.6 ***(Van der Waerden's theorem)*** *$\forall k, c \in \mathbb{N}, \exists W = W(k,c)$ such that, for all c-colorings $COL : [W] \rightarrow [c]$, $\exists a, d \in \mathbb{N}, d \neq 0$ such that*

$$COL(a) = COL(a+d) = COL(a+2d) = \cdots = COL(a+(k-1)d)$$

Proof:

We prove this by induction on k. That is, we show that

- $(\forall c)[W(1,c)$ exists$]$
- $(\forall c)[W(k,c)$ exists$] \implies (\forall c)[W(k+1,c)$ exists $]$

Base Case: $k=1$ As noted above $W(1,c)=1$ suffices. In fact, we also know that $W(2,c)=c+1$ suffices.

Induction Step: Assume $(\forall c)[W(k-1,c)$ exists $]$. Fix c. Consider what Lemma 3.5 says when $r=c$. In any c-coloring of $U=U(k,c,c)$, either there is a monochromatic k-AP or there are c monochromatic $(k-1)$-AP's which are all colored differently, and a number a whose color differs from all of them. Since there are only c colors, this cannot happen, so we must have a monochromatic k-AP. Hence $W(k,c) \leq U(k,c,c)$.
∎

Note that the proof of $W(k,c)$ depends on $W(k-1,c')$ where c' is quite large. Formally the proof is an ordering on the following order on (k,c)

$$(1,1) \prec (1,2) \prec \cdots \prec (2,1) \prec (2,2) \prec \cdots \prec (3,1) \prec (3,2) \cdots$$

This is an ω^2 ordering. It is well founded, so induction works.

4 The Polynomial Van Der Waerden's Theorem

4.1 Introduction

In this section we state and proof a generalization of van der Waerden's theorem known as *the Polynomial Van Der Waerden's Theorem.* We rewrite van der Waerden's theorem with an eye toward generalizing it.

For all $k, c \in \mathbb{N}$ *there exists* $W = W(k, c)$ *such that, for all c-colorings* $COL : [W] \to [c]$*, there exists* $a, d \in [W]$*, such that*

- $\{a\} \cup \{a + d, a + 2d, \ldots, a + (k-1)d\} \subseteq [W]$,
- $\{a\} \cup \{a + d, a + 2d, \ldots, a + (k-1)d\}$ *is monochromatic.*

In the proof of Lemma 3.5 we needed to take some care to make sure that $a \in [W]$. This needed the fact that $d \le U(k, c, r)$, which was obvious and needed no commentary. For the theorems in this section we will not have the relevant d bounded unless we assume it inductively. Hence we will often have the condition $d \in [W]$ or $d \in [U]$ which will help us show $a \in [W]$. The membership of other elements in $[W]$ will be obvious and not need commentary.

Note that the proof of the ordinary van der Waerden Theorem used the functions $d, 2d, 3d, \ldots, (k-1)d$. Why these functions? Would other functions work? What about polynomials? The following statement is a natural generalization of van der Waerden's theorem; however, it is not true.

Fix $c \in \mathbb{N}$ *and* $P \subseteq \mathbb{Z}[x]$ *finite. Then there exists* $W = W(P, c)$ *such that, for all c-colorings* $COL : [W] \to [c]$*, there are* $a, d \in \mathbb{N}$*,* $d \neq 0$*, such that*

- $\{a\} \cup \{a + p(d) \mid p \in P\} \subseteq [W]$,
- $\{a\} \cup \{a + p(d) \mid p \in P\}$ *is monochromatic.*

The above statement is false since the polynomial $p(x) = 2$ and the coloring

1	2	3	4	5	6	7	8	9	10 $\cdots$
R	R	B	B	R	R	B	B	R	$R \cdots$

provides a counterexample. Hence we need a condition to rule out constant functions. The condition $(\forall p \in P)[p(0) = 0]$ suffices.

The Polynomial Van Der Waerden Theorem (POLYVDW) *Fix* $c \in \mathbb{N}$ *and* $P \subseteq \mathbb{Z}[x]$ *finite, with* $(\forall p \in P)[p(0) = 0]$*. Then there exists* $W = W(P, c)$ *such that, for all c-colorings* $COL : [W] \to [c]$*, there are* $a, d \in [W]$*, such that*

- $\{a\} \cup \{a + p(d) \mid p \in P\} \subseteq [W]$,
- $\{a\} \cup \{a + p(d) \mid p \in P\}$ *is monochromatic.*

(When we apply this theorem to coloring $\{s+1,\ldots,s+W\}$, we will have $d \in [W]$ and $\{a\} \cup \{a+p(d) \mid p \in P\} \subseteq \{s+1,\ldots,s+W\}$.)

Note 4.1 Do we need the condition $d \in [W]$? For the classical van der Waerden Theorem $d \in [W]$ was obvious since

$$\{a\} \cup \{a+d,\ldots,a+(k-1)d\} \subseteq [W] \implies d \in [W].$$

For the polynomial van der Waerden's theorem one could have a polynomial with negative coefficients, hence it would be possible to have

$$\{a\} \cup \{a+p(d) \mid p \in P\} \subseteq [W] \text{ and } d \notin [W].$$

For the final result we do not care where d is; however, in order to prove POLYVDW inductively we will need the condition $d \in [W]$.

Note 4.2 The POLYVDW, as stated, allows the constant function 0 to be in P. Note that this does not matter: If POLYVDW is true for a set P then its true for the set $P \cup \{0\}$. In our main lemma we will need to assume inductively that POLYVDW holds for a set Q which *may have 0 in it*, and then prove POLYVDW for a set P that we insist *does not have 0 in it.* This is not a big deal; however, we are warning you to watch for it.

Note 4.3 The condition $(\forall p \in P)[p(0) = 0]$ is strong enough to make the theorem true. There are pairs (P, c) where $P \subseteq \mathbb{Z}[x]$ (that does not satisfy the condition) and $c \in \mathbb{N}$ such that the theorem is true. Classifying which pairs (P, c) satisfy the theorem is an interesting open problem.

Note 4.4 What happens if instead of polynomials we use some other types of functions? See Section 4.3 for a commentary on that.

Def 4.5 Let $n_e, \ldots, n_1 \in \mathbb{N}$. Let $P \subseteq \mathbb{Z}[x]$. *$P$ is of type $(n_e, \ldots, n_1)$* if the following hold:

1. P is finite.
2. $\forall p \in P, p(0) = 0$
3. The largest degree polynomial in P is of degree $\leq e$.
4. For all i, $1 \leq i \leq e$, There are $\leq n_i$ different lead coefficients of the polynomials of degree i. Note that there may be many more than n_i polynomials of degree i.

Note 4.6

1. Type $(0, n_e, \ldots, n_1)$ is the same as type $(n_e, \ldots, n_1)$.

2. We have no n_0. This is intentional. All the polynomials $p \in P$ have $p(0) = 0$. Hence the only degree 0 (constant) polynomial that we might have in P is 0. By convention P and $P \cup \{0\}$ will have the same type.

Example 4.7

1. The set $\{x, 2x, 3x, 4x, \ldots, 100x\}$ is of type (100).

2. The set
   $\{x^4 + 17x^3 - 65x, x^4 + x^3 + 2x^2 - x, x^4 + 14x^3, -x^4 - 3x^2 + 12x, -x^4 + 78x,$
   $x^3 - x^2, x^3 + x^2, 3x, 5x, 6x, 7x\}$
   is of type $(2, 1, 0, 4)$

3. The set
   $$\{x^4 + b_3x^3 + b_2x^2 + b_1x \mid -10^{10} \le b_1, b_2, b_3 \le 10^{10} \}$$
   is of type $(1, 0, 0, 0)$.

4. If P is of type $(1, 0)$ then there exists $b \in \mathbb{Z}$ and $k \in \mathbb{N}$ such that
   $$P \subseteq \{bx^2 - kx, bx^2 - (k-1)x, \ldots, bx^2 + kx\} \cup \{0\}.$$

5. If P is of type $(1, 1)$ then there exists $b_2, b_1 \in \mathbb{Z}$, and $k \in \mathbb{N}$ such that
   $$P \subseteq \{b_2x^2 - kx, b_2x^2 - (k-1)x, \ldots, b_2x^2 + kx\} \cup \{b_1x\} \cup \{0\}.$$

6. If P is of type (f, g, h) then there exists $b_3^{(1)}, \ldots, b_3^{(f)} \in \mathbb{Z}$, $b_2^{(1)}, \ldots, b_2^{(g)} \in \mathbb{Z}$, $b_1^{(1)}, \ldots, b_1^{(h)} \in \mathbb{Z}$, $k_1, k_2 \in \mathbb{N}$, T_1 of type (k_1), and T_2 of type (k_2, k_1) such that
   $$\begin{aligned} P \subseteq \quad & \{b_3^i x^3 + p(x) \mid 1 \le i \le f, p \in T_2\} \cup \\ & \{b_2^i x^2 + p(x) \mid 1 \le i \le g, p \in T_1\} \cup \\ & \{b_1^i x \mid 1 \le i \le h\} \cup \{0\} \end{aligned}$$

7. (a) Let
   $$P = \{2x^2 + 3x, x^2 + 20x, 5x, 8x\}.$$
   Let
   $$Q = \{p(x) - 8x \mid p \in P\}.$$
   Then
   $$Q = \{2x^2 - 5x, x^2 + 12x, -3x, 0\}.$$
   P is of type $(2, 2)$ and Q is of type $(2, 1)$. Note that the type 'decreases'.

8. Let
$$P = \{2x^2 + 3x, x^2 + 20x, 5x, 8x, 0\}.$$
Let
$$Q = \{p(x) - 8x \mid p \in P\}.$$
Then
$$Q = \{2x^2 - 5x, x^2 + 12x, -3x, -8x\}.$$
P is of type $(2, 2)$ and Q is of type $(2, 2)$. Note that the type stays the same. This is the reason we will later have to assume $0 \notin P$.

9. Let P be of type $(n_e, \ldots, n_i + 1, 0, \ldots, 0)$. Let bx^i be the leading term of some polynomial of degree i in P (note that we are not saying that $bx^i \in P$). Let $Q = \{p(x) - bx^i \mid p \in P\}$.
   (a) If $0 \notin P$ then there are numbers $n_{i-1}, \ldots, n_1$ such that Q is of type $(n_e, \ldots, n_i, n_{i-1}, \ldots, n_1)$. So, in some sense, the type decreases.
   (b) If $0 \in P$ then there are numbers $n_{i-1}, \ldots, n_1$ such that Q is of type $(n_e, \ldots, n_i + 1, n_{i-1}, \ldots, n_1)$. So, in some sense, the type increases or stays the same.

Def 4.8

1. Let $P \subseteq \mathbb{Z}[x]$ such that $(\forall p \in P)[p(0) = 0]$, POLYVDW($P$) means that the following holds:
   For all $c \in \mathbb{N}$, there exists $W = W(P, c)$ such that for all c-colorings $COL : [W] \to [c]$, there exists $a, d \in [W]$ such that
   $$\{a\} \cup \{a + p(d) \mid p \in P\} \text{ is monochromatic.}$$
   (If we use this definition on a coloring of $\{s+1, \ldots, s+W\}$ then the conclusion would have $a \in \{s + 1, \ldots, s + W\}$ and $d \in [W]$.)

2. Let $n_e, \ldots, n_1 \in \mathbb{N}$. POLYVDW($n_e, \ldots, n_1$) means that, for all $P \subseteq \mathbb{Z}[x]$ of type $(n_e, \ldots, n_1)$ POLYVDW(P) holds.

3. Let $(n_e, \ldots, n_i, \omega, \ldots, \omega)$ be the e-tuple that begins with $(n_e, \ldots, n_i)$ and then has $i - 1$ ω's.
   $$\text{POLYVDW}(n_e, \ldots, n_i, \omega, \ldots, \omega)$$
   is the statement
   $$\bigwedge_{n_{i-1}, \ldots, n_1 \in \mathbb{N}} \text{POLYVDW}(n_e, \ldots, n_i, n_{i-1}, \ldots, n_1).$$

4. POLYVDW is the statement

$$\bigwedge_{i=1}^{\infty} \text{POLYVDW}(\omega, \ldots, \omega)(\ \omega \text{ occurs } i \text{ times}).$$

Note that POLYVDW is the complete polynomial van der Waerden theorem.

Example 4.9

1. The statement POLYVDW(ω) is equivalent to the ordinary van der Waerden's theorem.
2. To prove POLYVDW$(1,0)$ it will suffice to prove POLYVDW(P) for all P of the form
$$\{bx^2 - kx, bx^2 - (k-1)x, \ldots, bx^2 + kx\}.$$
3. Assume that you know POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ and that you want to prove POLYVDW$(n_e, \ldots, n_i + 1, 0, \ldots, 0)$. Let P be of type $(n_e, \ldots, n_i + 1, 0 \ldots, 0)$. Let bx^i be the first term of some polynomial of degree i in P. Assume $0 \notin P$.
   (a) Let $Q = \{p(x) - bx^i \mid p \in P\}$. Then there exists $n_{i-1}, \ldots, n_1$, such that Q is of type $(n_e, \ldots, n_i, n_{i-1}, \ldots, n_1)$. Since POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ holds by assumption, we can assert that POLYVDW(Q) holds.
   (b) Let $U \in \mathbb{N}$. Let
   $$Q = \{p(x+u) - p(u) - bx^i \mid p \in P, 0 \le u \le U\}.$$
   Note $q(0) = 0$ for all $q \in Q$. Then there exists $n_{i-1}, \ldots, n_1$, such that Q is of type $(n_e, \ldots, n_i, n_{i-1}, \ldots, n_1)$. Since POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ holds by assumption, we can assert that POLYVDW(Q) holds.
   (c) Note that if $0 \in P$ then the type of Q will be the same as the type of P.

We will prove the Polynomial van der Waerden's theorem by an induction on a complicated structure. We state the implications we need to prove and then the ordering.

1. POLYVDW(1) follows from the pigeon hole principle.
2. We will show that, for all $n_e, \ldots, n_i \in \mathbb{N}$,
$$\text{POLYVDW}(n_e, \ldots, n_i, \omega, \ldots, \omega) \implies \text{POLYVDW}(n_e, \ldots, n_i + 1, 0, 0, \ldots, 0).$$
Note that this includes the case
$$\text{POLYVDW}(n_e, \ldots, n_2, n_1) \implies \text{POLYVDW}(n_e, \ldots, n_2, n_1 + 1).$$

The ordering we use is formally defined as follows:

Def 4.10 $(n_e, \ldots, n_1) \preceq (m_{e'}, \ldots, m_1)$ if either

- $e < e'$, or
- $e = e'$ and, for some i, $1 \le i \le e$, $n_e = m_e$, $n_{e-1} = m_{e-1}$, ..., $n_{i+1} = m_{i+1}$, but $n_i < m_i$.

This is an ω^ω ordering.

Example 4.11 We will use the following ordering on types.

$$(1) \prec (2) \prec (3) \prec \cdots$$

$$(1,0) \prec (1,1) \prec \cdots \prec (2,0) \prec (2,1) \prec \cdots \prec (3,0) \cdots \prec$$

$$(1,0,0) \prec (1,0,1) \prec \cdots \prec (1,1,0) \prec (1,1,1) \prec (1,2,0) \prec (1,2,1) \prec$$

$$(2,0,0) \prec \cdots \prec (3,0,0) \prec \cdots (4,0,0) \cdots .$$

4.2 The Proof of the Polynomial van der Waerden Theorem

4.2.1 POLYVDW($\{x^2, x^2+x, \ldots, x^2+kx\}$)

Def 4.12 Let $k \in \mathbb{N}$.

$$P_k = \{x^2, x^2+x, \ldots, x^2+kx\}.$$

We show POLYVDW(P_k). This proof contains many of the ideas used in the proof of POLYVDW.

We prove a lemma from which POLYVDW(P_k) will be obvious.

Lemma 4.13 *For all $k, c, r \in \mathbb{N}$, there exists $U = U(k,c,r)$ such that for all c-colorings $COL : [U] \to [c]$ one of the following Statements holds.*
Statement I: *There exists $a, d \in [U]$, such that*

- $\{a\} \cup \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\} \subseteq [U]$,
- $\{a\} \cup \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\}$ *is monochromatic.*

Statement II: *There exists $a, d_1, \ldots, d_r \in [U]$ such that the following hold.*

- $\{a+d_1^2, a+d_1^2+d_1, \ldots, a+d_1^2+kd_1\} \subseteq [U]$.
  $\{a+d_2^2, a+d_2^2+d_2, \ldots, a+d_2^2+kd_2\} \subseteq [U]$.
  $$\vdots$$
  $\{a+d_r^2, a+d_r^2+d_r, \ldots, a+d_r^2+kd_r\} \subseteq [U]$.
  (The element a is called the anchor*)*

- $\{a+d_1^2, a+d_1^2+d_1, \ldots, a+d_1^2+kd_1\}$ *is monochromatic.*
  $\{a+d_2^2, a+d_2^2+d_2, \ldots, a+d_2^2+kd_2\}$ *is monochromatic.*
  $$\vdots$$
  $\{a+d_r^2, a+d_r^2+d_r, \ldots, a+d_r^2+kd_r\}$ *is monochromatic.*
  With each monochromatic set being colored differently and differently from a.

Informal notes:

1. *We are saying that if you c-color $[U]$ either you will have a monochromatic set of the form*
$$\{a\} \cup \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\}$$
or you will have many *monochromatic sets of the form*
$$\{a+d^2, a+d^2+d, \ldots, a+d^2+kd\},$$
all of different colors, and different from a. Once "many" is more than c, then the latter cannot happen, so the former must, and we have POLYVDW(P).

2. *If we apply this theorem to a coloring of $\{s+1, \ldots, s+U\}$ then we either have*
$$d \in [U] \text{ and } \{a\} \cup \{a+d^2+d, \ldots, a+d^2+kd\} \subseteq \{s+1, \ldots, s+U\}.$$
or
$$d_1, \ldots, d_r \in [U] \text{ and, for all } i \text{ with } 1 \le i \le r \text{ such that}$$
$$\{a\} \cup \{a+d_i^2+d_i, \ldots, a+d_i^2+kd_i\} \subseteq \{s+1, \ldots, s+U\}, \text{ and}$$
$$\{a+d_i^2+d_i, \ldots, a+d_i^2+kd_i\} \subseteq \{s+1, \ldots, s+U\} \text{ monochromatic for each } i.$$

Proof:

We define $U(k,c,r)$ to be the least number such that this Lemma holds. We will prove $U(k,c,r)$ exists by giving an upper bound on it.

Base Case: $r=1$. $U(k,c,1) \le W(k+1,c) + W(k+1,c)^2$.

Let COL be any c-coloring of $[W(k+1,c)+W(k+1,c)^2]$. Look at the coloring restricted to the last $W(k+1,c)$ elements. By van der Waerden's theorem applied to the restricted coloring there exists

$$a' \in [(W(k+1,c))^2+1, \ldots, (W(k+1,c))^2+W(k+1,c)]$$

and

$$d' \in [W(k+1,c)]$$

such that

$$\{a', a'+d', a'+2d', \ldots, a'+kd'\} \subseteq \{(W(k+1,c))^2+1, \ldots, (W(k+1,c))^2+W(k+1,c)\}.$$

$$\{a', a'+d', a'+2d', \ldots, a'+kd'\} \text{ is monochromatic .}$$

Let $a = a' - (d')^2$ and $d = d'$.

$$\{a', a'+d', a'+2d', \ldots, a'+kd'\} = \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\} \text{ is monochromatic.}$$

If a is the same color then Statement I holds. If a is a different color then Statement II holds. There is one more issue– do we have $a, d \in [(W(k+1,c))^2 + W(k+1,c)]$?

Since

$$a' \geq (W(k+1,c))^2 + 1$$

and

$$d' \leq W(k+1,c)$$

we have that

$$a \geq (W(k+1,c))^2 + 1 - (W(k+1,c))^2 = 1.$$

Clearly

$$a < a' \leq W(k+1,c) + (W(k+1,c))^2.$$

Hence

$$a \in [W(k+1,c) + (W(k+1,c))^2].$$

Since $d = d' \in [W(k+1,c)]$ we clearly have

$$d \in [W(k+1,c) + (W(k+1,c))^2].$$

Induction Step: Assume $U(k,c,r)$ exists, and let

$$X = W(k + 2U(k,c,r), c^{U(k,c,r)}).$$

(X stands for eXtremely large.)

We show that

$$U(k,c,r+1) \leq (X \times U(k,c,r))^2 + X \times U(k,c,r).$$

Let COL be a c-coloring of

$$[(X \times U(k,c,r))^2 + X \times U(k,c,r)].$$

View this set as $(X \times U(k,c,r))^2$ consecutive elements followed by X blocks of length $U(k,c,r)$. Let the blocks be

$$B_1, B_2, \ldots, B_X.$$

Restrict COL to the blocks. We view the restricted c-coloring of numbers as a $c^{U(k,c,r)}$-coloring of the blocks. By the choice of X there exists $A, D' \in [X]$ such that

- $\{A, A+D', \ldots, A+(k+2U(k,c,r))D'\} \subseteq [X]$,
- $\{B_A, B_{A+D'}, \ldots, B_{A+(k+2U(k,c,r))D'}\}$ is monochromatic. How far apart are corresponding elements in adjacent blocks? Since the blocks viewed as points are D' apart, and each block has $U(k,c,r)$ elements in it, correspoinding elements in adjacent blocks are $D = D' \times U(k,c,r)$ numbers apart.

Consider the coloring of B_A. Since B_A is of size $U(k,c,r)$ either there exists $a, d \in U(k,c,r)$ such that

- $\{a\} \cup \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\} \subseteq B_A$,
- $\{a\} \cup \{a+d^2, a+d^2+d, \ldots, a+d^2+kd\}$ is monochromatic

in which case Statement I holds so we are done, or there exists
$a' \in B_A$, $d'_1, \ldots, d'_r \in [U(k,c,r)]$
such that

- $\{a'+{d'_1}^2, a'+{d'_1}^2+d'_1, \ldots, a'+{d'_1}^2+kd'_1\} \subseteq B_A$
  $\{a'+{d'_2}^2, a'+{d'_2}^2+d'_2, \ldots, a'+{d'_2}^2+kd'_2\} \subseteq B_A$
  $\vdots$
  $\{a'+{d'_r}^2, a'+{d'_r}^2+d'_r, \ldots, a'+{d'_r}^2+kd'_r\} \subseteq B_A$
- $\{a'+{d'_1}^2, a'+{d'_1}^2+d'_1, \ldots, a'+{d'_1}^2+kd'_1\}$ is monochromatic.
  $\{a'+{d'_2}^2, a'+{d'_2}^2+d'_2, \ldots, a'+{d'_2}^2+kd'_2\}$ is monochromatic.
  $\vdots$
  $\{a'+{d'_r}^2, a'+{d'_r}^2+d'_r, \ldots, a'+{d'_r}^2+kd'_r\}$ is monochromatic.

  with each monochromatic set colored differently from the others and from a'.

  Since $\{B_A, B_{A+D}, \ldots, B_{A+(k+2U(k,c,r))D}\}$ is monochromatic we also have that, for all j with $0 \le j \le k+2U(k,c,r)$,

$$\{a' + {d'_1}^2 + jD, a' + {d'_1}^2 + d'_1 + jD, \ldots, a' + {d'_1}^2 + kd'_1 + jD \mid 0 \le j \le k + 2U(k,c,r)\}$$

is monochromatic

$$\{a' + {d'_2}^2 + jD, a' + {d'_2}^2 + d'_2 + jD, \ldots, a' + {d'_2}^2 + kd'_2 + jD\} \mid 0 \le j \le k + 2U(k,c,r)\}$$

is monochromatic

$$\vdots$$

$$\{a' + {d'_r}^2 + jD, a' + {d'_r}^2 + d'_r + jD, \ldots, a' + {d'_2}^2 + kd'_r + jD\} \mid 0 \le j \le k + 2U(k,c,r)\}$$

is monochromatic.

with each monochromatic set colored differently from the others and from a', but the same as their counterpart in B_A.

Let $a = a' - D^2$, $d_i = D + d'_i$ for all $1 \le i \le r$, and $d_{r+1} = D$. We first show that these parameters work and then show that $a, d_1, \ldots, d_r \in [U(k,c,r+1)]$.

For $1 \le i \le r$ we need to show that

$$\{a + (D + d'_i)^2, a + (D + d'_i)^2 + (D + d'_i), \ldots, a + (D + d'_i)^2 + k(D + d'_i)\}$$

is monochromatic. Let $0 \le j \le k$. Note that

$$\begin{aligned} a + (D + d'_i)^2 + j(D + d'_i) &= (a' - D^2) + (D^2 + 2Dd'_i + {d'_i}^2) + (jD + jd'_i) \\ &= a' + {d'_i}^2 + jd'_i + (j + 2d'_i)D. \end{aligned}$$

Notice that $0 \le j + 2d'_i \le k + 2U(k,c,r)$. Hence $a + d_i^2 + jd_i \in B_{A+(j+2d'_i)D'}$, the $(j + 2d'_i)$th block. Since B_A is the same color as $B_{A+(j+2d'_i)D'}$,

$$COL(a + d_i^2) = COL(a + d_i^2 + jd_i).$$

So we have that, for all $0 \le i \le r$, for all j, $0 \le j\ lek$, the set

$$\{a + d_i^2, a + d_i^2 + d_i, \ldots, a + d_i^2 + kd_i\}$$

is monochromatic for each i. And, since the original sequences were different colors, so are our new sequences. Finally, if $COL(a) = COL(a + d_i^2)$ for some i, then we have $\{a, a + d_i^2, a + d_i^2 + d_i, \ldots, a + d_i^2 + kd_i\}$ monochromatic, satisfying Statement I. Otherwise, we satisfy Statement II.

We still need to show that $a, d_1, \ldots, d_r \in [X \times U(k,c,r))^2 + X \times U(k,c,r)]$. This is an easy exercise based on the lower bound on a' (since it came from the later $X \times U(k,c,r)$ coordinates) the inductive upper bound on the d_i's, and the upper bound $D \le U(k,c,r)$.

▮

Theorem 4.14 *For all k,* POLYVDW(P_k).

Proof: We show $W(P_k, c)$ exists by bounding it. Let $U(k, c, r)$ be the function from Lemma 4.13. We show $W(P_k, c) \le U(k, c, c)$. If COL is any c-coloring of $[U(k, c, c)]$ then second case of Lemma 4.13 cannot happen. Hence the first case must happen, so there exists $a, d \in [U(k, c, c)]$ such that

- $\{a\} \cup \{a + d^2, a + d^2 + d, \ldots, a + d^2 + kd\} \subseteq [U(k, c, c)]$
- $\{a\} \cup \{a + d^2, a + d^2 + d, \ldots, a + d^2 + kd\}$ is monochromatic.

Therefore $W(P_k, c) \le U(k, c, c)$. ∎

4.2.2 The Full Proof

We prove a lemma from which the implication

$$\text{POLYVDW}(n_e, \ldots, n_i, \omega, \ldots, \omega) \implies \text{POLYVDW}(n_e, \ldots, n_i + 1, 0, 0, \ldots, 0)$$

will be obvious. The lemma will need to assume that $0 \notin P$.

Lemma 4.15 *Let $n_e, \ldots, n_i \in \mathbb{N}$. Assume that* POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ *holds. For all $P \subseteq \mathbb{Z}[x]$ of type $(n_e, \ldots, n_i + 1, 0, \ldots, 0)$ such that $0 \notin P$, for all $c \in \mathbb{N}$, for all r, there exists $U = U(P, c, r)$ such that for all c-colorings $COL : [U] \to [c]$ one of the following Statements holds.*
Statement I: *there exists $a, d \in [U]$, such that*

- $\{a\} \cup \{a + p(d) \mid p \in P\} \subseteq [U]$.
- $\{a\} \cup \{a + p(d) \mid p \in P\}$ *is monochromatic.*

Statement II: *there exists $a, d_1, \ldots, d_r \in [U]$ such that the following hold.*

- $\{a + p(d_1) \mid p \in P\} \subseteq [U]$
  $\{a + p(d_2) \mid p \in P\} \subseteq [U]$
  $\vdots$
  $\{a + p(d_r) \mid p \in P\} \subseteq [U]$
  (The number a is called the anchor)
- $\{a + p(d_1) \mid p \in P\}$ *is monochromatic*
  $\{a + p(d_2) \mid p \in P\}$ *is monochromatic*
  $\vdots$
  $\{a + p(d_r) \mid p \in P\}$ *is monochromatic*

With each monochromatic set being colored differently and differently from a.
Informal notes:

1. *We are saying that if you c-color* $[U]$ *either you will have a monochromatic set of the form*
$$\{a\} \cup \{a + p(d) \mid p \in P\}$$
or you will have many monochromatic sets of the form
$$\{a + p(d) \mid p \in P\},$$
all of different colors, and different from a. Once "many" is more than c, then the latter cannot happen, so the former must, and we have
$$\text{POLYVDW}(n_e, \ldots, n_i, \omega, \ldots, \omega) \implies \text{POLYVDW}(n_e, \ldots, n_i + 1, 0, \ldots, 0).$$

2. *If we apply this theorem to a coloring of* $\{s + 1, \ldots, s + U\}$ *then we either have*
$$d \in [U] \text{ and } \{a\} \cup \{a + p(d) \mid p \in P\} \subseteq \{s + 1, \ldots, s + U\}$$
or
$$d_1, \ldots, d_r \in [U] \text{ and, for all } i \text{ with } 1 \le i \le r$$
$$\{a\} \cup \{a + p(d_i) \mid p \in P\} \subseteq \{s + 1, \ldots, s + U\}.$$

Proof: We define $U(P, c, r)$ to be the least number such that this Lemma holds. We will prove $U(P, c, r)$ exists by giving an upper bound on it. In particular, for each r, we will bound $U(P, c, r)$. We will prove this theorem by induction on r.

One of the fine points of this proof will be that we are careful to make sure that $a \in [U]$. The fact that we have inductively bounded on various d_i's will help that.

Fix $P \subseteq \mathbb{Z}[x]$ of type $(n_e, \ldots, n_i + 1, 0, \ldots, 0)$. Fix $c \in \mathbb{N}$. We can assume P actually has $n_i + 1$ lead coefficents for degree i polynomials (else P is of smaller type and hence POLYVDW(P, c) already holds and the lemma is true). In particular there exists some polynomial of degree i in P. Let bx^i be the first term of some polynomial of degree i in P. We will assume that $b > 0$. The proof when $b < 0$ is very similar.

Base Case: $r = 1$. Let
$$Q = \{p(x) - bx^i \mid p \in P\}.$$
It is easy to show that, since $0 \notin P$, there exists $n_{i-1}, \ldots, n_1$ such that Q is of type $(n_e, \ldots, n_i, n_{i-1}, \ldots, n_1)$, and that $(\forall q \in Q)[q(0) = 0]$. By the assumption that POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ is true, POLYVDW(Q) is true. Hence $W(Q, c)$ exists.

We show that
$$U(P, c, 1) \le bW(Q, c)^i + W(Q, c).$$

Let COL be any c-coloring of $[bW(Q,c)^i+W(Q,c)]$. Look at the coloring restricted to the last $W(Q,c)$ elements. By POLYVDW(Q) applied to the restricted coloring there exists $a' \in \{bW(Q,c)^i+1, \ldots, bW(Q,c)^i+W(Q,c)\}$ and $d' \in [W(Q,c)]$ such that

$$\{a'\} \cup \{a' + q(d') \mid q \in Q\} \text{ is monochromatic .}$$

(Note- we will only need that $\{a' + q(d') \mid q \in Q\}$ is monochromatic.)

Let $a = a' - b(d')^i$ and $d = d'$. (We will use $b > 0$ later to show that $a \in [U(P,c,1) \le bW(Q,c)^i + W(Q,c)]$.)

Then

$$\begin{aligned} \{a' + q(d') \mid q \in Q\} &= \{a' + p(d') - b(d')^i \mid p \in P\} \\ &= \{(a' - b(d)^i) + p(d) \mid p \in P\} \\ &= \{a + p(d) \mid p \in P\} \text{ is monochromatic.} \end{aligned}$$

If a is the same color then Statement I holds. If a is a different color then Statement II holds. There is one more issue– do we have $a, d \in [U(P,c,1)]$.

Since

$$a' \ge bW(Q,c)^i + 1$$

and

$$d' \le W(Q,c) \text{ (Recall that POLYVDW has the restriction } d \in [W].)$$

we have that

$$a \ge bW(Q,c)^i + 1 - bW(Q,c)^i = 1.$$

Clearly

$$a < a' \le bW(Q,c)^i + W(Q,c)$$

Hence

$$a \in [bW(Q,c)^i + W(Q,c)].$$

Since $d = d' \in [W(Q,c)]$ we clearly have

$$d \in [bW(Q,c)^i + W(Q,c)].$$

Induction Step: Assume $U(P,c,r)$ exists. Let

$$Q = \{p(x+u) - p(u) - bx^i \mid p \in P, 0 \le u \le U(P,c,r)\}.$$

Note that

$$\{p(x) - bx^i \mid p \in P\} \subseteq Q.$$

Clearly $(\forall q \in Q)[q(0) = 0]$. It is an easy exercise to show that, since $0 \notin P$, there exists $n_i, \ldots, n_1$ such that Q is of type $(n_e, \ldots, n_{i+1}, n_i, \ldots, n_1)$.

Now, let

$$Q' = \left\{ \frac{q(x \times U(P,c,r))}{U(P,c,r)} \mid q \in Q \right\}$$

Since every $q \in Q$ is an integer polynomial with $q(0) = 0$, it follows that $N|q(Nx)$, so we have $Q' \subseteq \mathbb{Z}[x]$. Moreover, it's clear that Q' has the same type as Q.

Since POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$ holds, we have POLYVDW(Q').

Hence $(\forall c')[W(Q', c')$ exists$]$. We show that

$$U(P, c, r+1) \le b\left(U(P,c,r)W(Q', c^{U(P,c,r)})\right)^i + U(P,c,r)W(Q', c^{U(P,c,r)}).$$

(Note that we are using $b > 0$) here.)

Let COL be a c-coloring of

$$\left[b\left(U(P,c,r)W(Q', c^{U(P,c,r)})\right)^i + U(P,c,r)W(Q', c^{U(P,c,r)})\right].$$

View this set as $b\left(U(P,c,r)W(Q', c^{U(P,c,r)})\right)^i$ elements followed by $W(Q', c^{U(P,c,r)})$ blocks of size $U(P,c,r)$ each. Restrict COL to the blocks. Now view the restricted c-coloring of *numbers* as a $c^{U(P,c,r)}$-coloring of *blocks*. Let the blocks be

$$B_1, B_2, \ldots, B_{W(Q', c^{U(P,c,r)})}.$$

By the definition of $W(Q', c^{U(P,c,r)})$ there exists $A, D' \in [W(Q', c^{U(P,c,r)})]$ such that

$$\{B_{A+q'(D')} \mid q \in Q\} \text{ is monochromatic.}$$

Note that we are saying that the blocks are the same color. Let $D = D' \times U(P,c,r)$ be the distance between corresponding elements of the blocks. Because each block is length $U(P,c,r)$, if we have an element $x \in B_A$, then in block $B_{A+q'(D')}$ we have a point x', where

$$\begin{aligned} x' &= x + q'(D')U(P,c,r) \\ &= x + q'\left(\frac{D'}{U(P,c,r)}\right)U(P,c,r) \\ &= x + q(D) \text{ for some } q \in Q, \text{ by definition of } Q' \end{aligned}$$

This will be very convenient.

Consider the coloring of B_A. Since B_A is of size $U(P,c,r)$ one of the following holds.

I) There exists $a \in B_A$ and $d \in [U(P,c,r)]$ such that

- $\{a\} \cup \{a + p(d) \mid p \in P\} \subseteq B_A$
- $\{a\} \cup \{a + p(d) \mid p \in P\}$ is monochromatic (so we are done).

II) There exists $a' \in B_A$ (so $a' \ge bW(Q', c^{U(P,c,r)})^i + 1$) and $d'_1, \ldots, d'_r \in [U(P,c,r)]$ such that

- $\{a' + p(d'_1) \mid p \in P\} \subseteq B_A$
  $\{a' + p(d'_2) \mid p \in P\} \subseteq B_A$
  $\vdots$
  $\{a' + p(d'_r) \mid p \in P\} \subseteq B_A$
- $\{a' + p(d'_1) \mid p \in P\}$ is monochromatic
  $\{a' + p(d'_2) \mid p \in P\}$ is monochromatic
  $\vdots$
  $\{a' + p(d'_r) \mid p \in P\}$ is monochromatic
  with each monochromatic set being colored differently from each other and from a'.

Since $\{B_{A+q'(D')} \mid q' \in Q'\}$ is monochromatic, and since we know that $x \in B_A$ corresponds to $x + q(D) \in B_{A+q'(D')}$, we discover that, for all $q \in Q$,

$$\begin{array}{c}\{a' + p(d'_1) + q(D) \mid p \in P\} \text{ is monochromatic} \\ \{a' + p(d'_2) + q(D) \mid p \in P\} \text{ is monochromatic} \\ \vdots \\ \{a' + p(d'_r) + q(D) \mid p \in P\} \text{ is monochromatic.}\end{array}$$

with each monochromatic set being colored differently from each other, and from a', but the same as their counterpart in B_A.

Let $a = a' - bD^i$. Note that since

$$a' \geq bW(Q', c^{U(P,c,r)})^i + 1$$

and

$$D \leq W(Q', c^{U(P,c,r)})$$

we have

$$a = a' - bD^i \geq bW(Q', c^{U(P,c,r)})^i + 1 - bW(Q', c^{U(P,c,r)})^i = 1$$

Clearly $a \leq a' \leq bW(Q', c^{U(P,c,r)} + U(P,c,r)W(Q', c^{U(P,c,r)})$. Hence

$$a \in [bW(Q', c^{U(P,c,r)})^i + U(P,c,r)W(Q', c^{U(P,c,r)})].$$

The value a will be our new anchor.

Since

$$\{B_{A+q'(D')} \mid q' \in Q'\}$$

is monochromatic (viewing the coloring on blocks) we know that

$$\{a' + q(D) \mid q \in Q\}$$

is monochromatic (viewing the coloring on numbers). Remember that the following is a subset of Q:

$$\{p(x) - bx^i \mid p \in P\}.$$

Hence the following set is monochromatic:

$$\begin{aligned}\{a' + p(D) - bD^i \mid p \in P\} &= \{a + bD^i + p(D) - bD^i \mid p \in P\} \\ &= \{a + p(D) \mid p \in P\}.\end{aligned}$$

If a is the same color then Statement I holds and we are done. If a is a different color then we have one value of d, namely $d_{r+1} = D$. We seek r additional ones to show that Statement II holds.

For each i we want to find a new d_i that works with the new anchor a. Consider the monochromatic set $\{a' + p(d'_i) \mid p \in P\}$. We will take each element of it and shift it $q(D)$ elements for some $q \in Q$. The resulting set is still monochromatic. We will pick $q \in Q$ carefully so that the resulting set, together with the new anchor a and the new values $d_i = d'_i + D$ work.

For each $p \in P$ we want to find a $q \in Q$ such that $a + p(d'_i + D)$ is of the form $a' + p(d'_i) + q(D)$, and hence the color is the same as $a' + p(d'_i)$.

$$\begin{aligned}a' + p(d'_i) + q(D) &= a + p(d'_i + D) \\ a' + p(d'_i) + q(D) - a &= p(d'_i + D) \\ bD^i + p(d'_i) + q(D) &= p(d'_i + D) \\ q(D) &= p(d'_i + D) - p(d'_i) - bD^i\end{aligned}$$

Take $q(x) = p(x + d'_i) - p(d'_i) - bD^i$. Note that $d'_i \leq U(Q, c, r)$ so that $q \in Q$.

Let $d_i = d'_i + D$ for $1 \leq i \leq r$, and $d_{r+1} = D$.

We have seen that

$$\{a + p(d_1) \mid p \in P\} \text{ is monochromatic}$$

$$\vdots$$

$$\{a + p(d_r) \mid p \in P\} \text{ is monochromatic}$$

AND

$$\{a + p(d_{r+1}) \mid p \in P\} \text{ is monochromatic}$$

The first r are guaranteed to be different colors by the inductive assumption. The $(r+1)^{st}$ is yet another color, because it shares a color with the anchor of our original sequences, which we assumed had its own color. So here we see that the Lemma is satisfied with parameters $a, d_1, \ldots, d_r, d_{r+1}$. ∎

Lemma 4.16 *For all* $n_e, \ldots, n_i$

$$\text{POLYVDW}(n_e, \ldots, n_i, \omega, \ldots, \omega) \implies \text{POLYVDW}(n_e, \ldots, n_i + 1, 0, 0, \ldots, 0).$$

Proof: Assume POLYVDW$(n_e, \ldots, n_i, \omega, \ldots, \omega)$. Let P be of type POLYVDW$(n_e, \ldots, n_i + 1, 0, 0, \ldots, 0)$. Apply Lemma 4.15 to $P - \{0\}$. Let

$$U(P - \{0\}, c, r)$$

be the function from Lemma 4.15. Clearly

$$W(P - \{0\}, c) \leq U(P - \{0\}, c, c)$$

suffices. Clearly

$$W(P, c) = W(P - \{0\}, c)$$

so we are done. ∎

We can now prove the Polynomial van der Waerden theorem.

Theorem 4.17 *For all* $P \subseteq \mathbb{Z}[x]$ *finite, such that* $(\forall p \in P)[p(0) = 0]$*, for all* $c \in \mathbb{N}$*, there exists* $W = W(P, c)$ *such that for all c-colorings* $COL : [W] \to [c]$*, there exists* $a, d \in [W]$ *such that*

- $\{a\} \cup \{a + p(d) \mid p \in P\} \subseteq [W]$,
- $\{a\} \cup \{a + p(d) \mid p \in P\}$ *is monochromatic.*

Proof:

We use the ordering from Definition 4.10. The least element of this set is (0). POLYVDW(0) is the base case. The only sets of polynomials of type (0) are $\emptyset$ and $\{0\}$. For each of these sets, the Polynomial van der Waerden theorem requires only one point to be monochromatic (the anchor), so of course POLYVDW(0) holds.

Lemma 4.15 is the induction step.

This proves the theorem.

∎

Note that our proof of POLYVDW did not use van der Warden's Theorem. The base case for POLYVDW was POLYVDW(0) which is trivial.

4.3 What if we do not use Polynomials?

It is tempting to expect that, if Polynomial van der Waerden's theorem holds, then so should "Exponential" van der Waerden. The question here is, for any choice of $b, c \in \mathbb{N}$, and for every c-coloring of $\mathbb{N}$, can we assure that there are $a, d \in \mathbb{N}$ with $a, a + b^d$ monochromatic?

The answer, as we will show, is no.

Proposition 4.18 *Fix $b \in \mathbb{N}$. Let p be the smallest prime number which is not a factor of b, Then there is a p-coloring $COL : \mathbb{N} \to [p]$ such that, $\forall a, d \in \mathbb{N}, COL(a) \neq COL(a + b^d)$.*

Proof: Fix $b, p \in \mathbb{N}$ with p the smallest prime non-factor of b. Now define the p-coloring $COL : \mathbb{N} \to [p]$ such that $COL(n) = n'$, where n' is the reduction of n modulo p with $n' \in [p]$. Most importantly, $COL(n) \equiv n \pmod{p}$. Thus, $COL(a) = COL(b)$ if and only if $p \mid (b - a)$.

Now let $a, d \in \mathbb{N}$, and consider $COL(a)$ and $COL(a + b^d)$. Well, since p is prime and $p \nmid b$, we know that $p \nmid b^d$. This guarantees that $COL(a) \neq COL(a + b^d)$, which is what was to be shown. ▮

5 Hales-Jewett theorem

5.1 Introduction

Hales-Jewett theorem feels very much like van der Waerden's, despite living in a very different domain. In the case of Hales-Jewett theorem, we replace $[W]$ with a hypercube, and the arithmetic progressions with monochromatic lines, but it will feel very similar. Here's the cast of players in Hales-Jewett:

- The hypercube — Given $c, t, N \in \mathbb{N}$, we will color the elements of the N-dimensional hypercube of length t — namely $[t]^N$.

  When $t = 26$, we can look at $[t]^N$ as strings of letters. For example, PUPPY and TIGER are points in $[26]^5$.

- The lines — In $[t]^N$, a *line* is a collection of points $P_1, P_2, \ldots, P_t$ such that $\exists \lambda \subseteq [n], \lambda \neq \emptyset$ satisfying

  $$\forall s \in \lambda, P_i^s = s \text{ and } \forall s \notin \lambda, \forall i, j, P_i^s = P_j^s$$

  where P_i^s denotes the s^{th} component of the point P_i. We call λ the "moving" coordinates, and the rest are static.

  Example 5.1 The following form a line in $[26]^9$, with $A = \{2, 3, 5, 8\}$:

  GAABARDAA

  GBBBBRDBA

  $\vdots$

  GZZBZRDZA

- The line$^-$. A *line*$^-$ is the first $t-1$ points of a line in $[t]^N$. The line$^-$ corresponding to the previous example is

$$\text{GAABARDAA}$$
$$\text{GBBBBRDBA}$$
$$\vdots$$
$$\text{GYYBYRDYA}$$

  Given a line L, we will refer to L^- as *the line*$^-$ *corresponding to* L.

- Completion — the would-be t^{th} point of a line$^-$. The completion of our line$^-$ is the point GZZBZRDZ. If more than one point would complete the line, we choose the least such point, according to a lexicographical ordering of $[t]^N$.

  Note 5.2 When $t \leq 2$, a line$^-$ may have more than one completion, since in that case a line$^-$ is a single point. For example, $\{\text{BAA}\}$ is a line$^-$ in $[2]^3$. Its completions are BAB, BBA, and BBB, depending on our choice of moving coordinates. However, when $t \geq 3$, a line$^-$ will have at least 2 points, which establishes the set of moving coordinates, and thus the completion of the line. This means, when $t \geq 3$, every line$^-$ has a unique, predetermined t^{th} point. The definition's use of the "least" t^{th} point only matters when $t \leq 2$

We are now ready to present the Hales-Jewett theorem.

Hales-Jewett theorem $\forall t, c, \exists N = HJ(t,c)$ such that, for all c-colorings $COL : [t]^N \to [c], \exists L \subseteq [t]^N, L$ a monochromatic line.

There are some easy base cases:

- $c = 1$ — $HJ(t,1) = 1$; $[t]^1 = [t]$ itself is a monochromatic line
- $t = 1$ — $HJ(1,c) = 0$; when $t = 1$ there is only a single point.

There is also a slightly harder base case:

Proposition 5.3 $HJ(2,c) = c + 1$

Proof:

Let $COL : [2]^{c+1} \to [c]$ be a c-coloring of $[2]^{c+1}$. Consider the following elements of $[2]^{c+1}$

$$\begin{array}{cccccc} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 2 \\ 1 & 1 & 1 & \cdots & 2 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2 & 2 & \cdots & 2 & 2 \\ 2 & 2 & 2 & \cdots & 2 & 2 \end{array}$$

Since there are $c+1$ elements and only c colors, two of these elements are the same color. We can assume they are of the form

$1^i 2^j$ where $i+j=t+1$

$1^{i'} 2^{j'}$ where $i'+j'=t+1$

These two elements form a monochromatic line. ▮

5.2 Proof of the Hales-Jewett Theorem

We prove a lemma from which the theorem will easily follow.

Lemma 5.4 *Fix $t, c \in \mathbb{N}$. Assume $(\forall c')[HJ(t-1, c')$ exists$]$. Then, for all r, there exists $N = U(t, c, r)$ such that, for all c-colorings $COL : [t]^N \to [c]$ one of the following statements holds.*

Statement I: *There exists a monochromatic line $L \subseteq [t]^N$.*

Statement II: *There exists r monochromatic line$^-$'s $L_1^-, L_2^-, \ldots, L_r^- \subset [t]^N, \in [t]^N$, and a point $Q \in [t]^N$, such that each L_i^- has a different color, Q is yet another color, and Q is the completion of every L_i^-. (Informally, we say that if you c-color $[t]^N$ then you will either have a monochromatic line, or many monochromatic line$^-$ structures, each of a different color. Once "many" becomes more than c, we must have a monochromatic line.)*

Proof:

We define $U(t, c, r)$ to be the least number such that this Lemma holds. We will prove $U(t, c, r)$ exists by giving an upper bound on it.

Base Case: If $r = 1$ then $U(t, c, 1) \leq HJ(t-1, c)$ suffices (actually $U(t, c, 1) = HJ(t-1, c)$). We take any c-coloring of $[t]^{HJ(t-1,c)}$, and restrict the domain to a c-coloring of $[t-1]^{HJ(t-1,c)}$ to find a monochromatic line, which it has by definition of HJ. This becomes a line$^-$ in $[t]^{HJ(t-1,c)}$, so we are done.

Induction Step: By induction, assume $U(t, c, r)$ exists. Let

$$X = c^{t^{U(t,c,r)}}. \text{ This is the number of ways to } c\text{-color } [t]^{U(t,c,r)}.$$

(X stands for eXtremely large.)

We will show that

$$U(t, c, r + 1) \leq HJ(t - 1, X) + U(t, c, r).$$

Let $N = HJ(t - 1, X) + U(t, c, r)$. Now we view $[t]^N$ as

$$[t]^{HJ(t-1,X)} \times [t]^{U(t,c,r)}.$$

Define $S = \{\chi \mid \chi : [t]^{U(t,c,r)} \to [c]\}$. Note that $|S| = X$. How convenient.

Let $COL : [t]^N \to [c]$ be our c-coloring. We define, for each $\sigma \in [t - 1]^{HJ(t-1,X)}$,

$$COL'(\sigma) : [t - 1]^{HJ(t-1,X)} \to S.$$

Note 5.5 At this point, it is essential to realize that COL' is a X-coloring of $[t - 1]^{HJ(t-1,X)}$. With every vector in $[t - 1]^{HJ(t-1,X)}$, we associate some $\chi \in S$. Although χ is itself a *coloring*, here we treat it as a *color*.

For example, $COL(\sigma)$ might be the following 3-coloring of $[2]^3$

$$COL'(\sigma)(0, 0, 0) = 1$$

$$COL'(\sigma)(0, 0, 1) = 1$$

$$COL'(\sigma)(0, 1, 0) = 3$$

$$COL'(\sigma)(0, 1, 1) = 2$$

$$COL'(\sigma)(1, 0, 0) = 1$$

$$COL'(\sigma)(1, 0, 1) = 3$$

$$COL'(\sigma)(1, 1, 0) = 2$$

$$COL'(\sigma)(1, 1, 1) = 2$$

Given $\sigma \in [t - 1]^{HJ(t-1,X)}$, $COL'(\sigma)$ will be a c-coloring of $[t]^{U(t,c,r)}$. Accordingly, we define COL' by telling the color of $COL'(\sigma)(\tau)$ for $\tau \in [t]^{U(t,c,r)}$. From here, our choice is clear — we associate to σ the c-coloring $COL'(\sigma) : [t]^{U(t,c,r)} \to [c]$ defined by

$$COL'(\sigma)(\tau) = COL(\sigma\tau).$$

Here $\sigma\tau$ is the vector in $[t]^N$ which is the concatenation of σ and τ.

We treat COL' as an X-coloring of $[t - 1]^{HJ(t-1,X)}$. By definition of $HJ(t - 1, X)$, we are guaranteed a monochromatic line, L, where $L \subseteq [t - 1]^{HJ(t-1,X)} \subset [t]^{HJ(t-1,X)}$. Let $L = \{P_1, P_2, \ldots, P_{t-1}\}$. So we have

$$COL'(P_1) = COL'(P_2) = \cdots = COL'(P_{t-1}) = \chi$$

L is a line in $[t-1]^{U(t,c,r)}$, but it is only a line$^-$ in $[t]^{HJ(t-1,X)}$. Let P_t be its completion.

Of course, χ itself is a c-coloring of $[t]^{U(t,c,r)}$. By definition of $U(t,c,r)$, we get one of two things:

Case 1: If χ gives a monochromatic line $L' = \{Q_1, Q_2, \ldots, Q_t\}$, then our monochromatic line in $[t]^N$ is

$$\{P_1Q_1, P_1Q_2, \ldots, P_1Q_t\}$$

and we are done.

Case 2: We have $L_1^-, L_2^-, \ldots, L_r^-$, each a monochromatic line$^-$ in $[t]^{U(t,c,r)}$, and each with the same completion $Q \in [t]^{U(t,c,r)}$. Note that Q must have an $(r+1)^{st}$ color, or else we would be in case 1. Let Q_i^j denote the j^{th} point on L_i^-. We now have all the components needed to piece together $r+1$ monochromatic line$^-$ structures:

$$\{P_1Q_1^1, P_2Q_1^2, \ldots, P_{t-1}Q_1^{t-1}\}$$
$$\{P_1Q_2^1, P_2Q_2^2, \ldots, P_{t-1}Q_2^{t-1}\}$$
$$\vdots$$
$$\{P_1Q_r^1, P_2Q_r^2, \ldots, P_{t-1}Q_r^{t-1}\}$$
$$AND$$
$$\{P_1Q, P_2Q, \ldots, P_{t-1}Q\}$$

We already know the first r have different colors.

Case 2.1: The line$^-$ $\{P_1Q, P_2Q, \ldots, P_{t-1}Q\}$ is the same color as the sequence $\{P_1Q_i^1, P_2Q_i^2, \ldots, P_{t-1}Q_i^{t-1}\}$ for some i. Then the line given by

$$\{P_1Q_i^1, P_1Q_i^2, \ldots, P_1Q_i^{t-1}, P_1Q\}$$

is monochromatic, so we are done, satisfying Statement I.

Case 2.2: The line$^-$ structures listed are all monochromatic and different colors, so we are still done, satisfying Statement II. ▮

We now restate and prove the Hales-Jewett theorem:

Theorem 5.6 ***Hales-Jewett theorem*** *$\forall t, c, \exists N = HJ(t,c)$ such that, for all c-colorings $COL : [t]^N \to [c], \exists L \subseteq [t]^N, L$ a monochromatic line.*

Proof:

We prove this by induction on t. We show that

- $(\forall c)[HJ(1,c)$ exists$]$
- $(\forall c)[HJ(t-1,c)$ exists$] \implies (\forall c)[HJ(t,c)$ exists$]$

Base Case: $t = 1$ As noted above $HJ(1,c) = 1$ works.
Induction Step: Assume $(\forall c)[HJ(t-1,c)$ exists $]$. Fix c. Consider Lemma 5.4 when $r = c$. In any c-coloring of $[t]^{U(t,c,c)}$, either there is a monochromatic line, or there are c monochromatic line$^-$ structures which are all colored differently, and share a completion Q colored differently. Since there are only c colors, this cannot happen, and we must have a monochromatic line. Hence $HJ(t,c) \le U(t,c,c)$. ▮

6 The Polynomial Hales-Jewett Theorem

Much as van der Waerden's theorem has a generalization to polynomials, so does the Hales-Jewett theorem. To get there, we must first generalize a few definitions, and create some we had no need for in the original version.

Recall that, in Hales-Jewett, we colored elements of $[t]^N$ and looked for monochromatic lines. Of course, we used the ground set $[t]$ only for convenience — we used none of the numerical properties. In that spirit, we may replace $[t]$ with any alphabet Σ of t letters.

Let $\Sigma = (\Sigma_d, \ldots, \Sigma_1)$ be a list of alphabets, and n a natural number. A *Hales-Jewett space* has the form

$$S_\Sigma(n) = \Sigma_d^{n^d} \times \Sigma_{d-1}^{n^{d-1}} \times \cdots \times \Sigma_1^n$$

We view an element $A \in S_\Sigma(n)$ as a collection of structures: a vector with coordinates from Σ_1, a square with coordinates from Σ_2, a cube with coordinates from Σ_3, and so on. In the ase of $d = 1$, and $\Sigma = [t]$, this is precisely the space colored in the ordinary Hales-Jewett theorem.

We define a set of formal polynomials over Σ by

$$\Sigma[\gamma] = \left\{a_d\gamma^d + \cdots + a_1\gamma \mid a_i \in \Sigma_i\right\}$$

Note that every polynomial has exactly d terms — omitting a term is not permitted. This differs from the Polynomial Van der Waerden theorem, where we allowed any polynomials. For example, x^3 is a valid polynomial in when dealing with Polynomial Van der Waerden's theorem. The closest we can come to this in Polynomial Hales-Jewett is $1x^3 + 0x^2 + 0x$. Note that although the coefficients may suggest meaning to the reader, they will have no numerical significance in the context of Hales-Jewett.

Let $A \in S_\Sigma(n)$, $p \in \Sigma[\gamma]$ of the form $p(\gamma) = a_d\gamma^d + \cdots + a_1\gamma$, and $\lambda \subseteq [n]$. Then we define $A + p(\lambda) \in S_\Sigma(n)$ as follows. Take the line from A, and replace the coordinates in λ by a_1. Similarly, replace the coordinates from the square in $\lambda^2 = \lambda \times \lambda$ with a_2, and so on.

Example 6.1 Let $\Sigma_1 = \{a, b, c, d\}, \Sigma_2 = [9]$, and $\Sigma = (\Sigma_2, \Sigma_1)$. Then a typical element $A \in S_\Sigma(3)$ looks like

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 8 & 8 & 9 \\ 4 & 5 & 3 \end{pmatrix} \ (a \quad d \quad c)$$

Note that A consists of a 3×3 block and a 1×3 block together, but they have no mathematical significance as a matrix or a vector.

Now, let $p \in \Sigma[\gamma]$ be given by $p(\gamma) = 1\gamma^2 + b\gamma$, and let $\lambda = \{1, 2\}$. Then

$$A + p(\lambda) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 9 \\ 4 & 5 & 3 \end{pmatrix} \ (b \quad b \quad c)$$

Now, we can restate the Hales-Jewett theorem in this language.

Theorem 6.2 ***Hales-Jewett theorem***
For every c, every finite alphabet Σ, there is some N such that, for any c-coloring $COL : S_\Sigma(N) \to [c]$, there is a point $A \in S_\Sigma(N), \lambda \subseteq [N]$, with $\lambda \neq \emptyset$ such that the set $\{A + \sigma\lambda \mid \sigma \in \Sigma\}$ is monochromatic.

From this terminology, we see a very natural generalization to a polynomial version of the theorem.

Theorem 6.3 ***Polynomial Hales-Jewett theorem***
For every c, every list of finite alphabets $\Sigma = (\Sigma_d, \ldots, \Sigma_1)$, and every collection $P \subseteq \Sigma[\gamma]$, there is a number $N = HJ(\Sigma, P, c)$ with the following property. For any c-coloring $COL : S_\Sigma(N) \to [c]$, there is a point $A \in S_\Sigma(N), \lambda \subseteq [N]$ with $\lambda \neq \emptyset$, such that the set $\{A + p(\lambda) \mid p \in P\}$ is monochromatic.

Example 6.4 Let $d = 2, \Sigma_2 = \{0, \ldots, 9\}, \Sigma_1 = \{a, \ldots, z\}$. Let $P = \{p, q, r\}$ where

$$p(\gamma) = 1\gamma^2 + a\gamma$$

$$q(\gamma) = 1\gamma^2 + b\gamma$$

$$r(\gamma) = 2\gamma^2 + c\gamma$$

If $N = 3$ and $\lambda = \{2, 3\}$, then the following would be an appropriate monochromatic set:

$$\begin{pmatrix} 5 & 2 & 4 \\ 7 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \ (f \quad a \quad a)$$

$$\begin{pmatrix} 5 & 2 & 4 \\ 7 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (f \quad b \quad b)$$

$$\begin{pmatrix} 5 & 2 & 4 \\ 7 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix} \quad (f \quad c \quad c)$$

Note that, in the Polynomial Van der Waerden's theorem, we needed to assume that $p(0) = 0$ for every $p \in P$. We have no such statement here, because we have no notion of a constant term for a polynomial in $\Sigma[\gamma]$.

To prove this theorem, we will do induction on the "type" of the set of polynomials P, as in the Polynomial Van der Waerden theorem. However, each polynomial necessarily has degree d, which makes the notion of type used previously rather unhelpful. In order to get the induction to work, we need to introduce a relative notion of degree, and tweak the definition of type.

Def 6.5 Let Σ be a list of finite languages, and $p, q \in \Sigma[\gamma]$. Then we say the degree of p relative to q is the degree of the highest term on which they differ. Formally, let $p(\gamma) = a_d\gamma^d + \cdots + a_1\gamma^1$, and $q(\gamma) = b_d\gamma^d + \cdots + b_1\gamma^1$. Let k be the largest index such that $a_k \neq b_k$ (or 0 if $p = q$). Then p has degree k with respect to q. We also say that p has leading coefficient a_k with respect to q

Note that the definition is symmetric: the degree of p relative to q is the same as the degree of q relative to p.

Example 6.6 Define

$$f(\gamma) = a\gamma^3 + 3\gamma^2 + \heartsuit\gamma$$
$$g(\gamma) = a\gamma^3 + 3\gamma^2 + \diamondsuit\gamma$$
$$h(\gamma) = b\gamma^3 + 3\gamma^2 + \heartsuit\gamma$$

The we see the following:

- f has degree 1 relative to g.
- f has leading coefficient $\heartsuit$ relative to g, and g has leading coefficient $\diamondsuit$ relative to f.
- h has degree 3 relative to both f and g.
- h has leading coefficient b relative to f and g, which each have leading coefficient a relative to h.

With this definition, we can now define the type of a set of polynomials relative to q virtually the same as we did for Polynomial Van der Waerden.

Def 6.7 Let Σ be a list of d finite alphabets, and $P \subseteq \Sigma[\gamma]$, $q \in \Sigma[\gamma]$. For each index k, let $P_k \subseteq P$ be the subset of polynomials with degree k relative to q. Let n_k be the number of leading coefficients relative to q of polynomials in P_k. Then the type of P relative to q is vector $(n_d, \ldots, n_1)$. We give these type vectors the same ordering as before, as seen in Definition 4.10.

For each $p_i \in P$, let t_i be the type of P relative to p_i. Then we say P has [absolute] type $t = \min t_i$.

Example 6.8 Let $P = \{p_1, p_2, p_3, p_4, p_5\}$, where

$$p_1 = a\gamma^3 + 6\gamma^2 + \diamondsuit\gamma$$

$$\begin{aligned} p_2 &= a\gamma^3 + 6\gamma^2 + \heartsuit\gamma \\ p_3 &= a\gamma^3 + 7\gamma^2 + \heartsuit\gamma \\ p_4 &= b\gamma^3 + 6\gamma^2 + \diamondsuit\gamma \\ p_5 &= b\gamma^3 + 6\gamma^2 + \heartsuit\gamma \end{aligned}$$

Let $Q = Q - \{p_5\}$. We see that:

- The type of P relative to p_1 and p_2 is $(1, 1, 1)$.
- The type of P relative to p_3 is $(1, 1, 0)$.
- The type of P relative to p_4 and p_5 is $(1, 0, 1)$.
- The [absolute] type of P is $(1, 0, 1)$, minimized by p_4 and p_5.
- P and $P - \{p_5\}$ have the same type relative to p_1, p_2, and p_3, the type remains unchanged.
- The type of $P - \{p_5\}$ relative to p_4 is $(1, 0, 0)$ — lower than the type of P.

The next proposition states that this last point always happens — the type of a set always decreases when you remove the polynomial which minimizes it.

Proposition 6.9 *Let P be a set of polynomials, such that $p \in P$ minimizes its type. Then $P - \{p\}$ has lower type.*

Proof: Let P have type $(n_d, \ldots, n_1)$, and let p minimize the type of P. Choose $q \in P$ to have minimal degree with respect to p, and call that degree k. Define $Q = P - \{p\}$. For polynomials in Q of degree greater than k relative to p, the degree is unchanged relative to q. Since the leading coefficients are also unchanged, the first $d - k$ coefficients of the type vector are identical for P and Q.

Now, let $Q_k \subseteq Q$ be the set of polynomials with degree $\leq k$ relative to p. By definition of the type vector, there are [exactly] n_k different leading coefficients of degree k polynomials in this set. Moreover, there are no polynomials of lower degree relative to p, since q was chosen to minimize k. Now, q has one of the n_k leading coefficients relative to p. Thus, relative to q, Q_k has $n_k - 1$ leading coefficients of degree k, with the remaining polynomials reducing in degree, because they agree with q on that coefficient. Thus, the type of Q relative to q is $(n_d, \ldots, n_{k+1}, n_k - 1, n'_{k-1}, \ldots, n'_1)$, for some values of $n'_{k-1}, \ldots, n'_1$. This type is lower than that of P, so the minimum type of Q is lower as well. ∎

Remark: We picked k to be the minimal degree of a polynomial relative to p. This means that the type of P is $(n_d, \ldots, n_k, 0, \ldots, 0)$. If there were any polynomials of degree $< k$, we would have picked one of those rather than q.

Now, in proving the Hales-Jewett theorem, it was important to view Σ^{n+m} as $\Sigma^n \times \Sigma^m$. We will need something similar for the polynomial version.

Proposition 6.10 *Let $n, m \in \mathbb{N}$, and Σ be a list of finite alphabets. Then there is a finite list of alphabets Σ' so that $S_\Sigma(n+m) \cong S_\Sigma(n) \times S_{\Sigma'}(m)$, where Σ' is independent of m.*

The proof of this is rather messy, but is done by manipulating the definition of $S_\Sigma(n+m)$. Rather than prove it in general here, we show the case when $\Sigma = (\Sigma_2, \Sigma_1)$.

$$\begin{aligned} S_\Sigma(n+m) &= \Sigma_2^{(n+m)^2} \times \Sigma_1^{n+m} \cong \Sigma_2^{n^2} \times \Sigma_2^{2nm} \times \Sigma_2^{m^2} \times \Sigma_1^n \times \Sigma_1^m \\ &\cong \left(\Sigma_2^{n^2} \times \Sigma_1^n\right) \times \left(\Sigma_2^{m^2} \times [\Sigma_2^{2n} \times \Sigma_1]^m\right) \end{aligned}$$

By setting $\Sigma' = (\Sigma_2, \Sigma_2^{2n} \times \Sigma_1)$, this comes out to be $S_\Sigma(n) \times S_{\Sigma'}(m)$, as desired. We view the transformation from $S_\Sigma(n+m)$ to $S_\Sigma(n) \times S_{\Sigma'}(m)$ as follows:

- Cut the line of length $n+m$ into two lines — one of length n, and one of length m. The former belongs to $S_\Sigma(n)$, and the latter to $S_{\Sigma'}(m)$.
- Cut the $(n+m) \times (n+m)$ block into four blocks. One is an $n \times n$ square, which belongs to $S_\Sigma(n)$. Another is an $m \times m$ square, which lives in the 2-dimensional portion of $S_{\Sigma'}(m)$. Leftover are blocks of size $n \times m$ and $m \times n$. We view these as "thick" lines of length m, with each entry representing n entries of the original space. In this way, we attach these pieces of the square to the line in $S_{\Sigma'}(m)$.

- Similarly, the k-dimensional block of $S_\Sigma(n+m)$ will be cut into 2^k pieces. One goes to the k-dimensional portion of $S_\Sigma(n)$ and another to the k-dimensional portion of $S_{\Sigma'}(m)$. The rest go to lower-dimensional portions of $S_{\Sigma'}(m)$.

Looking the other direction, let A' be a point in $S_{\Sigma'}(m)$.

- The d-dimensional part of A' comes from the d-dimensional portion of the point in the original space ($S_\Sigma(n+m)$).
- The $(d-1)$-dimensional part has one piece which is "truly" $(d-1)$-dimensional, and the rest of the pieces originally lived in d dimensions.
- The k-dimensional part of A' has one piece which is "truly" k-dimensional, and the other pieces are from higher dimensions.

How does viewing $S_\Sigma(n+m)$ like this affect polynomials? Let $\lambda \subseteq \{1,\ldots,n\}$, and $\kappa \subseteq \{n+1,\ldots,n+m\}$. Consider a polynomial $p(\gamma) = 1\gamma^2 + 2\gamma$. Then, given a point $A \in S_\Sigma(n+m)$, $A + p(\lambda \cup \kappa)$ involves putting a 1 at every point in $(\lambda \cup \kappa)^2$, and a 2 in $\lambda \cup \kappa$. That is, we must put a 1 everywhere in $\lambda \times \lambda, \lambda \times \kappa, \kappa \times \lambda$, and $\kappa \times \kappa$, and a 2 in λ and κ. We may now [nearly] view p as two polynomials: one in $\Sigma[\gamma]$, and the other in $\Sigma'[\gamma]$. The first is just p, since the alphabet has not changed. For the other, we need to know ahead of time what λ will be, to correctly place the 1's in $\lambda \times \kappa$ and $\kappa \times \lambda$, since we have control over all entries in $[n] \times \kappa$ and $\kappa \times [n]$. For this, we define $p|_\lambda$, the restriction of p to the entries of λ, by

$$p|_\lambda(\gamma) = 1\gamma^2 + (2, a_1, \ldots, a_{2n})\gamma$$

Here $a_i = 1$ if $i \in \lambda$ or $i + n \in \lambda$. For all other a_i, we have the freedom to prescribe any entries from Σ_2. For now we will use x as an unspecified symbol from Σ_2 to highlight where the choice lies.

So how do these polynomials work together? Let $A \in S_\Sigma(2+3)$ be all 0's:

$$A = \left(\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad (0 \;\; 0 \mid 0 \;\; 0 \;\; 0)$$

Now, let $\lambda = \{1\}$, and $\kappa = \{3, 4\}$, and let (B, C) be the decomposition of A as an element of $S_\Sigma(2) \times S_{\Sigma'}(3)$. Then we get

$$A + p(\lambda \cup \kappa) = \left(\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad (2 \;\; 0 \mid 2 \;\; 2 \;\; 0)$$

$$A' = (B + p(\lambda), C + p|_\lambda(\kappa)) = \left(\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & x & x & 0 \\ \hline 1 & x & 1 & 1 & 0 \\ 1 & x & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad (2 \;\; 0 \mid 2 \;\; 2 \;\; 0)$$

Note 6.11 A' is a close approximation of $A+p(\lambda\cup\kappa)$ — it agrees on $(\lambda\cup\kappa)^2$ and on $\lambda\cup\kappa$ as required by p. It only differs where x appears, because we could not predict what entries A would have there. Fortunately, in proving the theorem, we will only be interested in controlling the entries of $(\lambda\cup\kappa)^2$ and $(\lambda\cup\kappa)$ and ensuring the rest does not change. Therefore, if we are given a set of polynomials $P \subseteq \Sigma[\gamma]$, we may decompose each $p \in P$ as $(p, p|_\lambda)$ as above, and prescribe constant values for the x's. In proving the Polynomial Hales-Jewett theorem, if we have a sequence

$$\{(B + p(\lambda), C + p|_\lambda(\kappa)) \mid p \in P\}$$

we will fix the x's so that it is equal to

$$\{(B, C) + p(\lambda \cup \kappa) \mid p \in P\}$$

To do this, we will have a fixed polynomial p_0 over Σ which will dictate all these choices. Formally, let $p, p_0 \in \Sigma[\gamma]$ be polynomials, with $p(\gamma) = a_d\gamma^d + \cdots + a_1\gamma$, and $p_0(\gamma) = b_d\gamma^d + \cdots + b_1\gamma$. Then $p|_\lambda(\gamma) = c_d\gamma^d + \cdots + c_1\gamma$ has the following structure:

- $c_d = a_d$
- c_{d-1} is a list of symbols. One of these is a_{d-1}. The rest come from a_d and b_d, but which goes where depends on λ. These coefficients are for the d-dimensional piece of the polynomial. We can therefore define c_{d-1} as $(a_{d-1}, f(a_d, b_d, \lambda))$.
- c_k is a list of symbols. One of these is a_k. The rest are divided up based on which dimension they represent. The coefficients representing dimension j come from a_j or b_j, depending on λ. Thus, we can write

$$c_k = (a_k, f(a_d, \ldots, a_{k+1}, b_d, \ldots, b_{k+1}, \lambda))$$

- If $a_d = b_d, \ldots, a_{k+1} = b_{k+1}$, then λ has no on the k^{th} coefficient, so we can write it as
$$c_k = (a_k, g(a_d, \ldots, a_{k+1}))$$

Def 6.12 Just as in the proof of the Polynomial Van der Waerden's theorem, we define $POLYHJ(n_d, \ldots, n_1)$ to be the statement that the Polynomial Hales-Jewett theorem holds for all sets of polynomials of type $(n_d, \ldots, n_1)$. As in Definition 4.8, we also define $POLYHJ(n_d, \ldots, n_k, \omega, \ldots, \omega)$ to be the analogous statement.

We are now ready to prove a lemma from which the theorem will become trivial.

Lemma 6.13 *Assume $POLYHJ(n_d, \ldots, n_k, \omega, \ldots, \omega)$ holds. Fix a finite list of alphabets Σ and let $P \subseteq \Sigma[\gamma]$ have type $(n_d, \ldots, n_k + 1, 0, \ldots, 0)$, minimized by $p_0 \in P$. Then, for all numbers $c, r > 0$, there is a number $U = U(\Sigma, P, c, r)$ with the following property. For all c-colorings $COL : S_\Sigma(U) \to [c]$, one of the following Statements holds:*
Statement I: *There is a point $A \in S_\Sigma(U)$, $\lambda \subseteq [U], \lambda \neq \emptyset$, where $\{A + p(\lambda) \mid p \in P\}$ is monochromatic, or*
Statement II: *There are points $A_1, \ldots, A_r, A' \in S_\Sigma(U)$, and $\lambda_1, \ldots, \lambda_r \subseteq [U]$ with each $\lambda_i \neq \emptyset$ so that each set $\{A_i + p(\lambda_i) \mid p \in P, p \neq p_0\}$ is monochromatic, each with its own color, and each different from A'. Additionally, $A' = A_i + p_0(\lambda_i)$ as points for each $i \leq r$. We call A' the completion point of the sequences.*

Proof: By induction on r:

Base case ($r = 1$) — Recall that $P - \{p_0\}$ has lower type than P. Thus, Poly HJ holds for $P - \{p_0\}$. Let $U = HJ(\Sigma, P - \{p_0\}, c)$. Take any c-coloring of $S_\Sigma(U)$. By the definition of this number, there is some $A_1 = A \in S_\Sigma(U)$, and $\lambda_1 = \lambda \subseteq [U]$ with $\lambda \neq \emptyset$ so that $\{A_1 + p(\lambda_1) \mid p \in P - \{p_0\}\}$ is monochromatic. If the completion point is the same color, then Statement I holds. If not, Statement II holds.

Inductive case — Assume the lemma holds for r. We show that $U(\Sigma, P, c, r + 1)$ exists by giving an upper bound. In particular,
$$U(\Sigma, P, c, r + 1) \leq U + HJ = U(\Sigma, P, c, r) + HJ(\Sigma', Q, X)$$
where Q will be given shortly, and $X = c^{|S_\Sigma(U)|}$ is the number of c-colorings of $S_\Sigma(U)$. How convenient. By Proposition 6.10, $S_\Sigma(U + HJ) \cong S_\Sigma(U) \times S_{\Sigma'}(HJ)$, for some list of finite alphabets Σ' independent of the value of HJ. Then let
$$Q = \{p|_\lambda \in \Sigma'[\gamma] \,:\, p \in P - \{p_0\}, \lambda \subseteq [U]\}$$
where the free choice of entries is prescribed by p_0. This will ensure that
$$(A + p(\lambda), B + p|_\lambda(\kappa)) = (A, B) + p(\lambda \cup \kappa)$$
for any choice of p, λ, κ.
Claim: Q has type $(n_d, \ldots, n_k, n'_{k-1}, \ldots, n'_1)$ for some choice of $n'_{k-1}, \ldots, n'_1$.

Proof: By Proposition 6.9, $P - \{p_0\}$ has lower type than P, attained by some p_1 of degree k relative to p_0. Thus $P - \{p_0\}$ has type $(n_d, \ldots, n_k, m'_{k-1}, \ldots, m'_1)$ for some choice of $m'_{k-1}, \ldots, m'_1$. We will use this to show that the type of Q relative to $p_1|_\emptyset$ is low enough. In particular, we will show two things:

1. If p has degree $\ell \geq k$ relative to p_1, then $p|_\lambda$ has degree ℓ relative to $p_1|_\emptyset$ for every λ.

2. If p and q both have degree $\ell \geq k$ relative to p_1, and they have the same leading coefficient, then $p|_\lambda$ and $q|_\kappa$ have the same leading coefficient, for all λ and κ.

The two things together will guarantee that the number of distinct leading coefficients in Q will agree with the number in $P - \{p_0\}$, for all degrees $\geq k$, which is exactly what we want.

To see (1), write $p(\gamma) = a_d\gamma^d + \cdots + a_1\gamma$, $p_1(\gamma) = b_d\gamma^d + \cdots + b_1\gamma$, and $p_0(\gamma) = c_d\gamma^d + \cdots + c_1\gamma$. Since p_1 has degree k relative to p_0, the two polynomials agree on $b_d, \ldots, b_{k+1}$. Similarly, since p has degree $\ell \geq k$ relative to p_1, we get that $a_d = b_d = c_d, \ldots, a_{\ell+1} = b_{\ell+1} = c_{\ell+1}$. By Note 6.11, we see that the j^{th} coefficient of $p|_\lambda$ and $p_1|_\emptyset$ are given by $(a_j, f(a_d, \ldots, a_{j+1}))$ and $(b_j, f(b_d, \ldots, b_{j+1}))$ when $j \geq \ell$. Since these are identical when $j > \ell$, and different when $j = \ell$, we see that $p|_\lambda$ has degree ℓ relative to $p_1|_\emptyset$.

To see (2), let p and q have the same degree $\ell \geq k$ relative to p_1 (and thus relative to p_0), and also the same leading coefficient. Let $p(\gamma) = a_d\gamma^d + \cdots + a_1\gamma$, $q(\gamma) = b_d\gamma^d + \cdots + b_1\gamma$, and $p_0(\gamma) = c_d\gamma^d + \cdots + c_1\gamma$. As before, we see that $a_d = b_d = c_d, \ldots, a_{\ell+1} = b_{\ell+1} = c_{\ell+1}$. Fixing $\lambda, \kappa \subseteq [U]$, consider $p|_\lambda$ and $q|_\kappa$. By Note 6.11, for $j \geq \ell$, the j^{th} coefficient of these are given by $(a_\ell, f(a_d, \ldots, a_{\ell+1}))$ and $(b_\ell, f(b_d, \ldots, b_{\ell+1}))$ respectively. Since $a_j = b_j$ for $j \geq \ell$, these coefficents are identical. Thus, the two polynomials share a common leading coefficient relative to $p_1|_\emptyset$. ∎

This gives Q a lower type than P, which will allow us to use $POLYHJ$ as assumed.

Now, Let $COL : S_\Sigma(U + HJ) \to [c]$ be a c-coloring. Then we view COL as a c-coloring of $S_\Sigma(U) \times S_{\Sigma'}(HJ)$. As such, for each $\sigma \in S_{\Sigma'}(HJ)$, define $COL^*(\sigma) : S_{\Sigma'} \to [c]$ as the coloring of $S_\Sigma(U)$ induced by COL — for $\tau \in S_\Sigma(U)$, the map is defined so that $COL^*(\sigma)(\tau) = COL(\tau, \sigma)$. This makes COL^* a map from $S_{\Sigma'}$ to the c-colorings of $S_\Sigma(U)$.

The crucial observation here is there are X possible c-colorings of $S_\Sigma(U)$, so COL^* serves as an X-coloring of $S_{\Sigma'}(HJ)$. Thus, by choice of HJ, there is some point $B \in S_{\Sigma'}(HJ)$, and $\Lambda \subseteq [HJ]$ with $\Lambda \neq \emptyset$ so that

$$\{B + q(\Lambda) \mid q \in Q\} = \{B + p|_\lambda(\Lambda) \mid p \in P - \{p_0\}, \lambda \subseteq [U]\}$$

is monochromatic. This means that each point induces the same coloring χ on $S_\Sigma(U)$.

Now χ is a c-coloring of $S_\Sigma(U)$, so the choice of U allows us to use the inductive hypothesis on χ. Thus, either Statement I or II hold.

Case 1: There is a point $A \in S_\Sigma(U), \lambda \subseteq [U], \lambda \neq \emptyset$, so that $\{A + p(\lambda) \mid p \in P\}$ is monochromatic under χ. Then fix any $q \in Q$. Define $C = B + q(\Lambda)$. Since C induces the coloring χ on $S_\Sigma(U)$, we see that $\{(A + p(\lambda), C) \mid p \in P\}$ is monochromatix under COL. Moreover, viewing λ as a subset of $[U + HJ]$, these points are actually $(A + C) + p(\lambda)$, so we satisfy Statement I.

Case 2: There is are points $A_1, \ldots, A_r, A' \in S_\Sigma(U), \lambda_1, \ldots, \lambda_r \subseteq [U]$ with each $\lambda_i \neq \emptyset$ with the following properties:

$$\{A_1 + p(\lambda_1) \mid p \in P - \{p_0\}\} \text{ is monochromatic under } \chi$$

$$\vdots$$

$$\{A_r + p(\lambda_r) \mid p \in P - \{p_0\}\} \text{ is monochromatic under } \chi$$

and each of these sets has a different color, all different from $\chi(A')$. We also have $A' = A_i + p(\lambda_i)$ for all $i \leq r$

Since each $B + q(\Lambda)$ induces χ on $S_\Sigma(U)$, this gives us very many monochromatic points. For each i, this set is monochromatic under COL:

$$\{(A_i + p(\lambda_i), B + q) \mid p \in P - \{p_0\}, q \in Q\}$$

In particular, the following r sets of points are monochromatic, so that each set has its own color:

$$\{(A_1 + p(\lambda_1), B + p|_{\lambda_1}(\Lambda)) \mid p \in P - \{p_0\}\} = \{(A_1, B) + p(\lambda_1 \cup \Lambda) \mid p \in P - \{p_0\}\}$$

$$\vdots$$

$$\{(A_r + p(\lambda_r), B + p|_{\lambda_r}(\Lambda)) \mid p \in P - \{p_0\}\} = \{(A_r, B) + p(\lambda_r \cup \Lambda) \mid p \in P - \{p_0\}\}$$

Let $B' = B + p_0|_\emptyset$. Then we see that the final point of each of these sequences is given by

$$(A_i + p_0(\lambda_i), B + p_0|_{\lambda_i}(\Lambda)) = (A', B') + p_0(\Lambda)$$

This realization gives us the following choice for the $(r+1)^{\text{st}}$ sequence:

$$\{(A', B' + p|_\emptyset(\Lambda)) \mid p \in P - \{p_0\}\} = \{(A', B') + p(\Lambda) \mid p \in P - \{p_0\}\}$$

Since each $B' + p|_\emptyset(\Lambda)$ induces χ on $S_\Sigma(U)$, each of these has the color $\chi(A')$, so this set is monochromatic. It is also immediate that its completion point is the same as the other r: $(A', B') + p_0(\Lambda)$.

If the completion point has the same color as the i^{th} sequence, then that sequence with its completion satisfies Statement I. If not, then Statement II holds. Either way, we have the goal. ∎

Theorem 6.14 ***Polynomial Hales-Jewett theorem***
For every c, every lits of finite alphabets $\Sigma = (\Sigma_d, \ldots, \Sigma_1)$*, and every collection* $P \subseteq \Sigma[\gamma]$*, there is a number* $N = HJ(\Sigma, P, c)$ *with the following property. For any c-coloring* $COL : S_\Sigma(N) \to [c]$*, there is a point* $A \in S_\Sigma(N), \lambda \subseteq [N]$ *with* $\lambda \neq \emptyset$*, such that the set* $\{A + p(\lambda) \mid p \in P\}$ *is monochromatic.*

Proof: By induction on the type of P. Note that, as in the proof of the Polynomial Van der Waerden theorem, types are well-ordered, so induction is a correct approach.

Base case: Let P have type $(0, \ldots, 0)$, so that $P = \{p\}$ is a single polynomial (p has degree 0 relative to itself). Set $N = 1$, and let $COL : S_\Sigma(1) \to [c]$ be any c-coloring. Then, for any $A \in S_\Sigma(1)$, we have $\{A + p(\{1\})\}$ monochromatic, since it is just one point.

Inductive case: Suppose we know $POLYHJ(n_d, \ldots, n_k, \omega, \ldots, \omega)$. Let P have type $(n_d, \ldots, n_k + 1, 0, \ldots, 0)$. Let $N = U(\Sigma, P, c, c)$ as guaranteed by the lemma above. Let $COL : S_\Sigma(U) \to [c]$ be a c-coloring. Statement II cannot hold, since it requires $c + 1$ different colors. Thus, Statement I holds, which was the goal. ▮

References

[1] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemeredi's theorems. *Journal of the American Mathematical Society*, pages 725–753, 1996.

[2] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemeredi's on arithmetic progressions. *Journal of the Annals of Mathematics*, 31:204–256, 1977.

[3] A. Sarkozy. On difference sets of sequences of integers I. *Acta Math. Sci. Hung.*, 31:125–149, 1977.

[4] M. Walters. Combinatorial proofs of the polynomial van der waerden theorem and the polynomial hales-jewett theorem. *Journal of the London Mathematical Society*, 61:1–12, 2000.

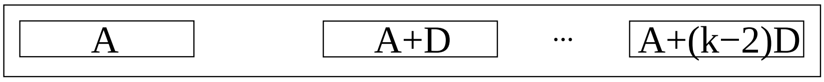

Figure 4: $k-1$ blocks in arithmetic progression

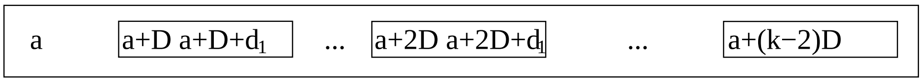

Figure 5: Very many monochromatic points