# Van Der Waerten's Theorem

Given k and l, there exists an n(k,l) such that any string S of k different colors contains a single-color arithmetic progression (AP) of length l if $|S| > n(k,l)$.

**Overview:**
Van Der Waerten's key contribution was the idea of an inductive proof where we view the blocks of colors created at step k-1 as a single color at step k. We can create a constructive proof with this idea, but it does not yield particularly good bounds for the VDW numbers (which are defined as the lowest n(k,l) that satisfies the above statement).

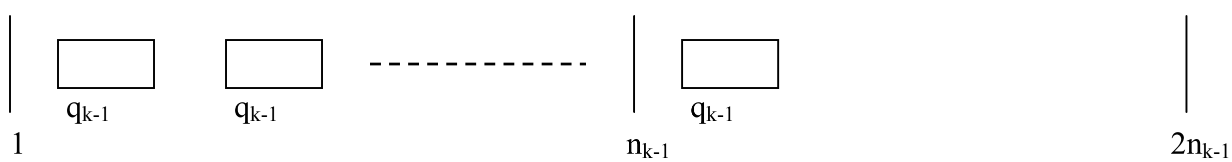
A key idea in the proof we'll call Lemma 1:

**Lemma 1:**
Assume $B [i_1\ i_2 \ldots i_s] = B [j_1\ j_2 \ldots j_s]$ are two blocks of colors with 'sub-blocks' that are defined down k-s levels in the same positions in both blocks, (with the $k\text{-}s^{th}$ level block having size 1 – e.g. being a single color). $(s < k)$
Then for any arbitrary indices $i_{s+1}, i_{s+2}, \ldots i_k$ taken from [1, l+1]
the color $B [i_1\ i_2 \ldots i_s]\ i_{s+1}\ i_{s+2} \ldots i_k$ is in the same position of the block $B [i_1\ i_2 \ldots i_s]$
that the color $B [j_1\ j_2 \ldots j_s]\ i_{s+1}\ i_{s+2} \ldots i_k$ is in the block $B [j_1\ j_2 \ldots j_s]$
Since we know the two outer blocks are equal, the two colors are equal.

**Proof of Van Der Waerten's Theorem:**
Proof by induction on l. Assume n(k,l) exists for all k. Show n(k, l+1) exists for all $k \geq 2$

*Base case*: l=2. n(k,2) = k+1 for all k.

*Inductive step*: We take $n(k,l+1) = 2\ n_{k-1}\ q_{k-1}$ ($n_x$ and $q_x$ are defined below) for any given k.




We take our string of colors and group them into $2n_{k-1}$ blocks, each of size $q_{k-1}$.
Note there are $k^{qk-1}$ different combinations for these blocks. Consider all these different possible blocks as colors themselves.
So by our inductive hypothesis, we have an l-length AP of these blocks in the left half (the first $n_{k-1}$ blocks), where $n_{k-1} = n(k^{qk-1}, l)$. Also note the $(l+1)^{th}$ block in this AP has to fall in the above string.

Then, since each of the blocks in the AP are of length n(x,l) (for some x), we know an l-length arithmetic progression occurs in them, and also that the $(l+1)^{th}$ block in those APs appears in them too, exactly as before.

We simply apply this construction k times. Note the length of the blocks goes from

$q_{k-1} \to q_{k-2} \to \dots q_0$ where the values are recursively defined as follows:

$q_0 = 1$ $\quad n_0 = n(k,l)$
$q_1 = 2n_0q_0$ $\quad n_1 = n(k^{q1}, l)$
$q_2 = 2n_1q_1$ $\quad n_2 = n(k^{q2}, l)$
…
$q_k = 2n_kq_k$ $\quad n_k = n(k^{qk}, l)$

**Intuition:**
By Lemma 1, we can show that:

$B_{1\,1\,..\,1\;\;1\,1\,1\;\;l'..\,l'}$
$B_{1\,1\,..\,1\;\;2\,2\,2\;\;l'..\,l'}$
$B_{1\,1\,..\,1\;\;3\,3\,3\;\;l'..\,l'}$
…
$B_{1\,1\,..\,1\;\;l\,l\,l\;\;l'..\,l'}$ are the same colors (where $l' = l+1$)

(We know that $B_{1\,1\,..\,1\;1} = B_{1\,1\,..\,1\;2}$ since they are part of an AP at one of the levels.
Also, $B_{1\,1\,..\,1\;1\,1} = B_{1\,1\,..\,1\;1\,2}$ for the same reason, and
$B_{1\,1\,..\,1\;1\,2} = B_{1\,1\,..\,1\;2\,2}$ by Lemma 1, etc).

So we can make an l-length AP this way. Now we need to consider the color:

$B_{1\,1\,..\,1\;\;l'\,l'\,l'\;\;l'..\,l'}$ which is certainly not guaranteed to be the same color as above.

Note we can define k+1 colors this way:

$a_0 = B_{l'\,l'\,l'\,l'\,l' \dots l'}$
$a_1 = B_{1\,l'\,l'\,l'\,l' \dots l'}$
$a_2 = B_{1\,1\,l'\,l'\,l' \dots l'}$
…
$a_{k+1} = B_{1\,1\,1\,1\,1 \dots 1}$

Since we only have k colors, there exists r, s (r<s) such that:

$B_{1 \dots 1\;\; l' \dots l'} = B_{1 \dots 1\;\; l' \dots l'}$
(r) (k-r) (s) (k-s)

Now we can look at the $l$+1 colors as we did above:

$B_{1 \dots 1\;\; i \dots i\;\; l' \dots l'}$
(r) (s-r) (k-s) $(1 \le i \le l')$

As we showed above, the first $l$ of these colors are the same by Lemma 1.
And since the $l'$ th color is the same as the 1st color by our choice of s and r, we have an AP of length $l+1$.

The only thing left to show is that these colors are indeed equally spaced from their neighbors. For any $c_j$, $(j<l')$, let $c_k=c_{j+1}$

$$c_j = B_{\underbrace{1 \ldots 1}_{(r)}\ \underbrace{j \ldots j}_{(s-r)}\ \underbrace{l' \ldots l'}_{(k-s)}}$$

$$c_k = B_{\underbrace{1 \ldots 1}_{(r)}\ \underbrace{k \ldots k}_{(s-r)}\ \underbrace{l' \ldots l'}_{(k-s)}}$$

Note that the distance $d_{r+1} = d(B_{1 \ldots 1\ j}, B_{1 \ldots 1\ k})$ does not depend on our choice of j
(it depends on the 'hop size' of the AP at step r+1.
Likewise the $d_{r+2} = d(B_{1 \ldots 1\ j\, j}, B_{1 \ldots 1\ k\, k})$ depends on the hop sizes at steps r+1 and r+2.

So $d(c_j, c_k) = d_{r+1} + d_{r+2} + \ldots d_s$ which is independent of our choice of j. $\blacksquare$

---

tdarling@eng.umd.edu