ON A VARIANT OF VAN DER WAERDEN'S THEOREM

Sujith Vijay

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
sujith@math.uiuc.edu

Received: 6/11/08, Revised: 1/13/10, Accepted: 1/19/10, Published: 5/12/10

Abstract

Given positive integers n and k, a k-term quasi-progression of diameter n is a sequence $(x_1, x_2, \ldots, x_k)$ such that $d \leq x_{j+1} - x_j \leq d + n, 1 \leq j \leq k-1$, for some positive integer d. Thus an arithmetic progression is a quasi-progression of diameter 0. Let $Q_n(k)$ denote the least integer for which every coloring of $\{1, 2, \ldots, Q_n(k)\}$ yields a monochromatic k-term quasi-progression of diameter n. We obtain an exponential lower bound on $Q_1(k)$ using probabilistic techniques and linear algebra.

1. Introduction

A cornerstone of Ramsey theory is the theorem of van der Waerden [5], stating that for every positive integer k, there exists an integer $W(k)$ such that any 2-coloring of $\{1, 2, \ldots, W(k)\}$ yields a monochromatic k-term arithmetic progression. It is known that $W(k)$ is at least exponential in k, but the upper and lower bounds are nowhere close to each other. Indeed, the best known upper bound on $W(k)$ is a five-times iterated tower of exponents.

Given positive integers n and k, a k-term quasi-progression of diameter n is a sequence $(x_1, x_2, \ldots, x_k)$ such that for some positive integer d,

$$d \leq x_{j+1} - x_j \leq d + n, \quad 1 \leq j \leq k-1.$$

The integer d is called the *low-difference* of the quasi-progression. Analogous to the van der Waerden number $W(k)$, we can define $Q_n(k)$ as the least integer for which any 2-coloring of $\{1, 2, \ldots, Q_n(k)\}$ yields a monochromatic k-term quasi-progression of diameter n. Note that $Q_n(k) \leq W(k)$ with equality if $n = 0$.

2. An Exponential Lower Bound for $Q_1(k)$

Landman [3] showed that $Q_1(k) \geq 2(k-1)^2+1$. We improve this to an exponential lower bound, using elementary probabilistic techniques (see [1]) and some linear algebra.

Theorem. *Let $k \geq 3$. Then, $Q_1(k) \geq 1.08^k$.*

Proof. Let $S = \{1, 2, \ldots, N\}$. (The value of N will be specified later.) Define $m = \lfloor (k-1)/2 \rfloor$. We group the elements of S from left to right in *zones* of size $2m$, and subdivide each zone into two *blocks* of size m. We color each zone randomly and uniformly in one of two ways: left block red, right block blue; or left block blue, right block red. Let $A \subseteq S$ be a monochromatic k-term quasi-progression of diameter 1 under this coloring. Since the coloring ensures that no three consecutive blocks have the same color, A must consist of elements from different blocks. Thus A is monochromatic only if the associated block sequence is monochromatic.

Observe that there are $N-k+1$ ways to choose the first term of A and at most $N/(k-1)$ ways to choose the low difference. Suppose we are able to show, for a fixed first-term and low difference, that there are at most c^k block sequences corresponding to k-term quasi-progressions of diameter 1, with $c < 2$. Since a block sequence is monochromatic with probability 2^{1-k}, the linearity of expectation implies that the expected number of monochromatic k-term quasi-progressions under a random coloring is at most $2N^2(c/2)^k/(k-1)$. When $N = \lfloor (2/c)^{k/2} \rfloor$, the expected number is less than 1, so that there must exist some coloring under which there are no monochromatic k-term quasi-progressions. Thus $Q_1(k) \geq (2/c)^{k/2}$. From what follows, it will be clear that we may take $c < 1.71$. We remark, in passing, that the number of k-term quasi-progressions of diameter 1 contained in S far exceeds 2^k, dooming the naive approach of randomly coloring the elements themselves.

For $1 \leq j \leq k$, let $B^j_{a,d} = \{(b_1, b_2, \ldots, b_j)\}$ be the set of all possible block sequences corresponding to j-term quasi-progressions $\{a_1, a_2, \ldots, a_j\}$ with first term $a_1 = a$ and low-difference d, where a_i belongs to the block numbered b_i. Since the possible values of a_j lie in an interval consisting of $j \leq k$ integers, there are at most three possible values for each b_j. (In fact, for $j \leq \lceil k/2 \rceil$, there are at most two possible values for each b_j.) We claim that $|B^k_{a,d}| < 1.71^k$.

Given a and d, we can compute $|B^j_{a,d}|$ as follows. Let a_j and a_{j+1} be consecutive terms of a quasi-progression of diameter 1 and low difference d. Note that there are at most two possible values for the difference in block numbers of successive terms of a quasi-progression of diameter 1 and low-difference d.

Consider a k-partite digraph G_k, with three vertices in each part corresponding to possible values of b_j (including dummy vertices if there are fewer than three possible values of b_j), and a directed edge from a vertex in part j to a vertex in part $j+1$ if and only if there exists a block sequence containing the corresponding blocks in positions j and $j+1$. We now assign a unit weight to the non-dummy vertex corresponding to b_1 and recursively define the weight of a vertex v to be the sum of the weights of all vertices w such that there is a directed edge from w to v (dummy vertices have weight 0). It follows that $|B^j_{a,d}|$ equals the sum of weights of vertices in the j^{th} part. We encode the weights of vertices in the j^{th} part with 3×1 column vectors $[x_j, y_j, z_j]$, starting with $[x_1, y_1, z_1] = [1, 0, 0]$. (An example corresponding to $k = 7, a = 15$ and $d = 4$ is shown in Figure 1. Note that $|B^k_{a,d}| = 3 + 8 + 7 = 18$.)

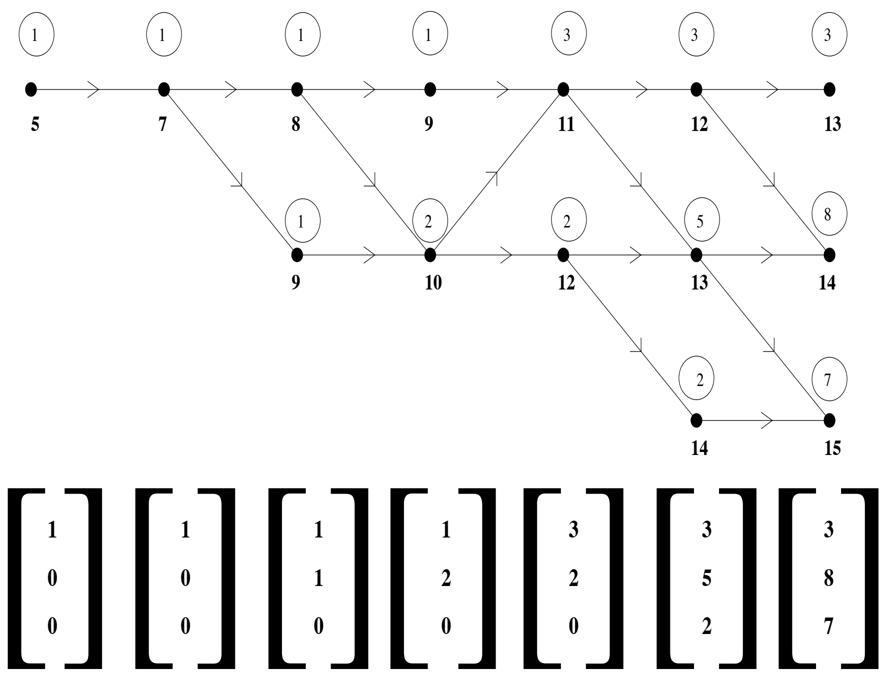

Figure 1: The computation of $|B^7_{15,4}|$ $(m = 3)$.

We will now show that there are only nine labelled digraphs that could be induced on adjacent partite sets of G_k (five of these can be seen in Figure 1). This will imply, in turn, that $[x_{j+1}, y_{j+1}, z_{j+1}]^T = A[x_j, y_j, z_j]^T$ where A is one of the following nine (0,1)-matrices:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = A_3^T \qquad A_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = A_5^T$$

$$A_6 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = A_7^T \qquad A_8 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = A_9^T.$$

Let $I_r = [a+(r-1)d, a+(r-1)(d+1)]$ denote the interval of possible values of the r^{th} term of a quasi-progression of diameter 1 with first term a and low-difference d. In keeping with our division of the set of positive integers into blocks of size m, we say that the interval I_r straddles block $B+1$ if $I_r \cap [Bm+1, (B+1)m] \neq \emptyset$. Note that each interval I_r straddles at most three blocks. That the matrices A_1 through A_9 form an exhaustive list of action matrices is a consequence of the following observations:

- If I_r straddles one block, then I_{r+1} straddles either one block (matrix A_1) or two blocks (matrix A_2).
- If I_r straddles two blocks, then I_{r+1} straddles one (matrix A_3), two (matrices A_4 and A_5) or three (matrix A_6) blocks.
- If I_r straddles three blocks, then I_{r+1} straddles two (matrix A_7) or three blocks (matrices A_8 and A_9).

In other words, $[x_k, y_k, z_k]$ can be written as the product of a sequence of $k-1$ matrices, each selected from the nine matrices A_i listed above, acting on the vector $[1,0,0]$. We now recall the definition of the spectral norm $||A||_2$ of an $n \times n$ matrix A:

$$||A||_2 = \sup_{||\mathbf{x}||_2=1} ||A\mathbf{x}||_2 = \sqrt{\lambda_{max}(A^T A)}$$

where $\mathbf{x}$ varies over all $n \times 1$ column vectors, $||\mathbf{x}||_2$ denotes the Euclidean norm of $\mathbf{x}$, and $\lambda_{max}(M)$ denotes the largest eigenvalue of a symmetric matrix M with non-negative diagonal entries. The following properties of the spectral norm are well-known, and are immediate consequences of the definition:

$$||A\mathbf{x}||_2 \leq ||A||_2 ||\mathbf{x}||_2$$

$$||AB||_2 \leq ||A||_2 ||B||_2.$$

Evaluating the spectral norms, we find that $||A_1||_2 = 1, ||A_2||_2 = ||A_3||_2 = \sqrt{2}, ||A_4||_2 = ||A_5||_2 = (1+\sqrt{5})/2 < 1.619, ||A_6||_2 = ||A_7||_2 = \sqrt{3}$ and $||A_8||_2 = ||A_9||_2 < 1.803$. Note that the matrix that takes $[x_1, y_1, z_1] = [1,0,0]$ to $[x_2, y_2, z_2]$ must be A_1 or A_2. Moreover, A_6, A_7, A_8 and A_9 come into play only if $j > m = \lfloor (k-1)/2 \rfloor$. It follows from the submultiplicativity of the spectral norm that

$$\sqrt{x_k^2 + y_k^2 + z_k^2} < \sqrt{2}\ (1.619)^{m-1}\ (1.803)^{k-m-1}.$$

Finally, by the Cauchy-Schwarz inequality,

$$|B_{a,d}^k| = x_k + y_k + z_k \leq \sqrt{3}\left(\sqrt{x_k^2 + y_k^2 + z_k^2}\right) < 1.71^k.$$

Thus $Q_1(k) > (2/1.71)^{k/2} > 1.08^k$, completing the proof. □

3. Concluding Remarks

While numerical evidence seems to indicate that $Q_1(k) = O(c^k)$ for some absolute constant c, there is no reason to believe that the constant 1.08 is even close to optimal; more delicate computations and an application of the Local Lemma will very likely push it to around 1.2. However, it would be far more interesting to have a reasonable upper bound for quasi-progressions of small diameter, if not for diameter 1. Landman [3] has shown that $Q_{\lceil 2k/3 \rceil}(k) \leq \frac{43k^3}{324} + o(k^3)$, but no upper bound for $Q_n(k)$ is known when $n = o(k)$.

We end with a table of known values of $Q_1(k)$ (see [4]):

k	3	4	5	6	7	8	9
$Q_1(k)$	9	19	33	67	≥ 124	≥ 190	≥ 287

Acknowledgement

I thank the referee for a careful and thorough reading of the article, and for several helpful comments that helped improve the overall clarity of presentation.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*. Academic Press, New York, 1992.

[2] T. C. Brown, P. Erdős and A. R. Freedman, *Quasi-progressions and Descending Waves*, Journal of Combinatorial Theory Series A 53 (1990), 81-95.

[3] B. M. Landman, *Ramsey Functions for Quasi-progressions*, Graphs and Combinatorics 14 (1998), 131-142.

[4] B. M. Landman and A. Robertson, *Ramsey Theory on the Integers*. American Mathematical Society, Providence, 2004.

[5] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Niew Archief voor Wiskunde 15 (1927), 212-216.