JOURNAL OF COMBINATORIAL THEORY, Series A **52**, 275–287 (1989)

Ascending Waves

N. ALON*

Bell Communication Research, Morristown, New Jersey 07960 and
Department of Mathematics, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel

AND

JOEL SPENCER

Department of Mathematics, SUNY at Stony Brook,
Stony Brook, New York 11794

Communicated by the Managing Editors

Received September 15, 1987

A sequence of integers $x_1 < x_2 < \cdots < x_k$ is called an *ascending wave of length k* if $x_{i+1} - x_i \leqslant x_{i+2} - x_{i+1}$ for all $1 \leqslant i \leqslant k-2$. Let $f(k)$ be the smallest positive integer such that any 2-coloring of $\{1, 2, ..., f(k)\}$ contains a monochromatic ascending wave of length k. Settling a problem of Brown, Erdös, and Freedman we show that there are two positive constants c_1, c_2 such that $c_1 k^3 \leqslant f(k) \leqslant c_2 k^3$ for all $k \geqslant 1$. Let $g(n)$ be the largest integer k such that any set $A \subseteq \{1, 2, ..., n\}$ of cardinality $|A| \geqslant n/2$ contains an ascending wave of length k. We show that there are two positive constants c_3 and c_4 such that $c_3(\log n)^2/\log \log n \leqslant g(n) \leqslant c_4(\log n)^2$ for all $n \geqslant 1$.

0. INTRODUCTION

A sequence of integers $x_1 < x_2 < \cdots < x_k$ is called an *ascending wave* (AW) of *length k* if $x_{i+1} - x_i \leqslant x_{i+2} - x_{i+1}$ for all $1 \leqslant i \leqslant k-2$. Let $f(k)$ be the smallest positive integer such that in any 2-coloring of $\{1, 2, ..., f(k)\}$ there is a monochromatic ascending wave of length k. It is easy to see that

$$f(k) \leqslant O(k^3) \tag{0.1}$$

Indeed, let C be a coloring of $\{1, ..., k^3\}$ in red and blue. Assume, without loss of generality, that 1 is colored red. If there are k consecutive integers colored blue we have a monochromatic AW of length k, as needed.

* Research supported in part by Allon Fellowship and by the Fund for basic research administered by the Israel Academy of Sciences.

0097-3165/89 $3.00

Otherwise, put $a_0=0$, $a_1=1$ and define, recursively, $a_{i+1}=\min\{x \mid x-a_i \geqslant a_i-a_{i-1}$ and x is colored red$\}$, $(i \geqslant 1)$. As there are no k consecutive integers colored blue this gives $a_{i+1} \leqslant a_i+a_i-a_{i-1}+k$ and hence $a_k \leqslant 1+k+2k+\cdots+k(k-1)=1+\frac{1}{2}k^2(k-1) \leqslant k^3$. $A=a_1<a_2<\cdots<a_k$ is a red AW of length k, establishing (0.1). Brown, Erdös, and Freedman [BEF] showed that

$$k^2-k+1 \leqslant f(k) \leqslant k^3/3-4k/3+3$$

for all $k \geqslant 1$ and asked if, in fact, the lower bound is the exact value of $f(k)$ for all $k \geqslant 1$. Here we show that this is false by proving the following theorem, which determines the asymptotic behavior of the function f.

THEOREM 1.1. $\Omega(k^3) \leqslant f(k) \leqslant O(k^3)$.

I.e., there are two positive constants $c_1, c_2>0$ such that

$$c_1 k^3 \leqslant f(k) \leqslant c_2 k^3$$

for all $k \geqslant 1$.

Let $g=g(n)$ be the largest positive integer such that any set $A \subseteq \{1, 2, ..., n\}$ of cardinality $|A| \geqslant \frac{1}{2}n$ contains an ascending wave of length g. Brown, Erdös, and Freedman [BEF] showed that

$$\Omega(\log n) \leqslant g(n) \leqslant O(\sqrt{n}).$$

Our next theorem determines almost precisely the asymptotic behavior of the function g.

THEOREM 2.1. $\Omega(\log^2 n/\log\log n) \leqslant g(n) \leqslant O(\log^2 n)$.

The proofs of Theorems 1.1 and 2.1 are given in Sections 1 and 2, respectively. The final Section 3 contains some concluding remarks and open problems.

1. Monochromatic Ascending Waves

In this section we prove Theorem 1.1. Recall that Brown, Erdös, and Freedman [BEF] showed that $f(k) \leqslant k^3/3-4k/3+3$, so it remains to prove the lower bound. We prove the lower bound by a probabilistic construction. Put $b=\lfloor k/40 \rfloor$ and $m=\lfloor 10^{-20}k^2 \rfloor$. Let C be a random 2-coloring of the integers $1, 2, ..., 4bm$ defined as follows; Split these integers into $4m$ blocks $B_1, B_2, ..., B_{4m}$ of b consecutive integers each. I.e., $B_i=\{(i-1)b+1, (i-1)b+2, ..., ib\}$. For each j, $1 \leqslant j \leqslant m$, choose, randomly and independently, an integer $c_j \in \{0, 1, 2, 3\}$ where $\text{Prob}(c_j = l) = \frac{1}{4}$ for all $l \in \{0, 1, 2, 3\}$. Let $A = (a_{ls})_{l,s=0}^{3}$ be the following 4×4 matrix:

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

For all $1 \leqslant j \leqslant m$ and all $0 \leqslant s \leqslant 3$ we color all the members of the block B_{4j-s} by the color $a_{c_j s}$. This defines a random coloring C of $N = \{1, 2, ..., 4bm\}$. We next show that the probability that in this coloring there is a monochromatic ascending wave of length k is smaller than 1 for sufficiently large k. This clearly implies the lower bound in Theorem 1.1 and completes its proof. For convenience, we split the proof into several lemmas. The first one is an immediate consequence of the construction. For $1 \leqslant i \leqslant 4m$ let $\text{col}(B_i)$ denote the common color of the members of the block B_i. Thus $\text{col}(B_i)$ is a random variable whose range is $\{0, 1\}$.

LEMMA 1.2. (i) *No five consecutive blocks B_i have the same color.*

(ii) *For all $1 \leqslant i \leqslant 4b$,*

$$\text{Prob}(\text{col}(B_i) = 0) = \text{Prob}(\text{col}(B_i) = 1) = \tfrac{1}{2}.$$

(iii) *The colors of any pair of consecutive blocks are independent. I.e., for every $1 \leqslant i < 4m$ and every $\varepsilon, \delta \in \{0, 1\}$,*

$$\text{Prob}(\text{col}(B_i) = \varepsilon \quad \textit{and} \quad \text{col}(B_{i+1}) = \delta) = \tfrac{1}{4}.$$

(iv) *The colors of blocks whose pairwise distances are at least four mutually independent. I.e., for every $1 \leqslant i_1 < i_2 < i_3 < \cdots < i_s \leqslant 4m$ that satisfy $i_{j+1} - i_j \geqslant 4$ for $1 \leqslant j < s$ and for every $\varepsilon_1, \varepsilon_2, ..., \varepsilon_s \in \{0, 1\}$,*

$$\text{Prob}(\text{col}(B_{i_j}) = \varepsilon_j \textit{ for } 1 \leqslant j \leqslant s) = 1/2^s.$$

(v) *The colors of pairs of consecutive blocks whose pairwise distances are at least four are mutually independent. Namely, for every $1 \leqslant i_1 < i_1 + 5 \leqslant i_2 < i_2 + 5 \leqslant i_3 < \cdots < i_{s-i} < i_{s-1} + 5 \leqslant i_s < 4m$ and for every $\varepsilon_1, \delta_1, \varepsilon_2, \delta_2, ..., \varepsilon_s, \delta_s \in \{0, 1\}$,*

$$\text{Prob}(\text{col}(B_{i_j}) = \varepsilon_j \textit{ and } \text{col}(B_{i_j - 1}) = \delta_j \textit{ for } 1 \leqslant j \leqslant s) = 1/4^s.$$

An arithmetic progression $x_1 < x_2 < \cdots < x_t$ in N is called *good* (with respect to the coloring C) if for every $\varepsilon \in \{0, 1\}$ there is a member x_i of the progression that belongs to a block B_j and $\text{col}(B_j) = \text{col}(B_{j+1}) = \varepsilon$. A progression which is not good is *bad.*

LEMMA 1.3. *The probability that there is a bad progression with difference $d > b$ of $t = [100 \log k]$ terms is at most* 0.01.

Proof. Let $P = x_1 < x_2 < \cdots < x_t$ be a fixed progression with difference $d > b$. Notice that no two elements of P belong to the same block. For each i, $1 \leqslant i \leqslant t/5$, let C_i be the block that contains x_{5i} and let D_i be the consecutive block. By Lemma 1.2, part (v):

$$\begin{aligned}\text{Prob}(P \text{ is bad})\\ &\leqslant \text{Prob}(\nexists i,\ 1 \leqslant i \leqslant t/5; \text{col}(C_i) = \text{col}(D_i) = 0)\\ &\quad + \text{Prob}(\nexists i,\ 1 \leqslant i \leqslant t/5; \text{col}(C_i) = \text{col}(D_i) = 1)\\ &= 2 \cdot (3/4)\lfloor t/5 \rfloor < 1/k^6.\end{aligned}$$

The total number of arithmetic progressions of length t in N is bounded by $(4bm)^2 < 0.01k^6$. Thus, the expected number of such bad sequences is smaller than 0.01, and the probability that there is one is at most 0.01, as claimed. ▮

LEMMA 1.4. *Suppose the coloring C contains no bad progressions with difference $d > b$ of $t = [100 \log k]$ terms. Then in any monochromatic ascending wave in N of $l = \lceil k/2 \rceil$ terms the last difference is at least $k^2/10^8 \log^2 k$.*

Proof. Let $A = a_1 < a_2 < \cdots < a_l$ be a monochromatic ascending wave. Without loss of generality we may assume that its color is 0. By Lemma 1.2, part (i), C contains no 5 consecutive blocks of color 0. It follows that at most $4b < k/10$ members of A are consecutive integers. After these, A has to jump at least one block and as it is ascending we conclude that for $g = \lfloor k/10 \rfloor$ we have $a_{i+1} - a_i > b$ for all $i \geqslant g$. Suppose $h \geqslant g$ and put $d = a_{h+1} - a_h$. Consider the arithmetic progression $x_0, x_1, \ldots, x_t$ where $t = \lfloor 100 \log k \rfloor$, $x_0 = a_h$ and $x_1 = a_{h+1}$. As this progression is good there is some i such that $x_i \in B_j$ and $\text{col}(B_j) = \text{col}(B_{j+1}) = 1$. As A is an ascending wave it is clear that $a_{h+i} \geqslant x_i$. Moreover, as a_{h+i} is colored 0, it must be bigger than the maximum element of B_{j+1}. It follows that $a_{h+i} - a_h \geqslant id + b$. Therefore, the average difference $a_{j+1} - a_j$ for $h \leqslant j < h + i$ is at least $d + b/i \geqslant d + b/100 \log k$. As A is ascending we conclude that $a_{h+t+1} - a_{h+t} \geqslant a_{h+1} - a_h + b100 \log k$. Summing these inequalities for $h = g + it$ $0 \leqslant i < \lceil k/3t \rceil$ we obtain

$$\begin{aligned}a_l - a_{l-1} &\geqslant a_{g+\lceil k/3t \rceil \cdot t+1} - a_{g+\lceil k/3t \rceil \cdot t}\\ &\geqslant \left\lceil \frac{k}{3t} \right\rceil \cdot \frac{b}{100 \log k} > \frac{k^2}{10^8 \log^2 k}.\end{aligned}$$

This completes the proof. ▮

An immediate consequence of the last two lemmas is the following.

COROLLARY 1.5. *For sufficiently large k, the probability that in the coloring C there is a monochromatic ascending wave of* $\lceil k/2 \rceil$ *terms whose last difference is smaller than* $4b$ ($\leqslant k^2/10^8 \log^2 k$) *is at most* 0.01.

Remark 1.6. Lemmas 1.3 and 1.4 easily imply that $f(k) \geqslant \Omega(k^3/\log^2 k)$, since in any monochromatic ascending wave of length k, each of the last $k/2$ differences will be at least $k^2/10^8 \log^2 k$. Our objective is to improve this bound and get the sharp estimate $f(k) \geqslant \Omega(k^3)$. This requires some additional work, which follows.

Let us call a sequence X of positive real numbers $x_1 < x_2 < \cdots < x_t$ a real ascending t-wave if $0 \leqslant x_1 < x_t \leqslant t^2$, $x_{i+1} - x_i \leqslant x_{i+2} - x_{i+1}$ for all $1 \leqslant i \leqslant t-2$, and $x_t - x_{t-1} \leqslant 10^{-14} t$. For such a sequence X, let $\lfloor X \rfloor$ denote the sequence of the integer parts of the members of X, i.e., $\lfloor X \rfloor = (\lfloor x_1 \rfloor, \lfloor x_2 \rfloor, ..., \lfloor x_t \rfloor)$. Let R_t denote the set of all real ascending t-waves and define $\lfloor R_t \rfloor = \{\lfloor X \rfloor : X \in R_t\}$. Clearly, $\lfloor R_t \rfloor$ is a finite set. The next lemma provides an upper bound on its cardinality.

LEMMA 1.7. *For all* $t \geqslant 1000$,

$$|\lfloor R_t \rfloor| \leqslant 10^5 \cdot t^2 \cdot 4^{\lceil t/1000 \rceil} \cdot \binom{t + \lceil 10^{-6} t \rceil - 1}{t-1} \cdot 2^{\lceil 10^{-3} \cdot t \rceil} \cdot 2^{t/2}.$$

Thus, for all sufficiently large t,

$$|\lfloor R_t \rfloor| \leqslant 2^{2t/3}.$$

Proof. For $X = (x_1 < x_2 < \cdots < x_t) \in R_t$ define $\Delta = \Delta(X)$ by $\Delta = (d_1, d_2, ..., d_{t-1})$, where $d_i = x_{i+1} - x_i$ for $1 \leqslant i < t$. Define also $\Delta^* = \Delta^*(X) = (d_1^*, d_2^*, ..., d_{t-1}^*)$, where $d_i^* = 10^{-8} \lfloor 10^8 d_i \rfloor$. Notice that $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_{t-1} \leqslant 10^{-14} t$ and hence $d_1^* \leqslant d_2^* \leqslant \cdots \leqslant d_{t-1}^* \leqslant 10^{-14} t$. The sequence $10^8 d_1^*, 10^8 d_2^* + 1, 10^8 d_3^* + 2, ..., 10^8 d_{t-1}^* + t - 2$ is a strictly increasing sequence of nonnegative integers, whose largest member is at most $t + \lceil 10^{-6} t \rceil - 2$. The number of such sequences is clearly at most

$$\binom{t + \lceil 10^{-6} t \rceil - 1}{t-1} \tag{1.1}$$

and hence (1.1) is an upper bound on the number of possible sequences Δ^*. We next bound the number of sequences $\lfloor X \rfloor$ corresponding to members $X = (x_1, ..., x_t) \in R_t$ with a given sequence $\Delta^*(X) = (d_1^*, ..., d_{t-1}^*)$. There are at most t^2 possibilities for $\lfloor x_1 \rfloor$. Given $\lfloor x_1 \rfloor$, there are at most 10^5 possibilities for the value of $10^{-5} \lfloor 10^5 x_1 \rfloor$. Notice that knowing this value, we know the values of $x_1, x_2, ..., x_{1001}$, up to an error of 2×10^{-5}, since for $i \leqslant 1001$, $x_i = x_1 + \sum_{j=1}^{i-1} d_j = 10^{-5} \lfloor 10^5 x_1 \rfloor + \sum_{j=1}^{i-1} d_j^* + \varepsilon$, where $0 \leqslant \varepsilon < 10^{-5} + i \cdot 10^{-8} \leqslant 2 \times 10^{-5}$. In particular, there are less than four possibilities for the value of $10^{-5} \lfloor 10^5 x_{1001} \rfloor$, and each of these values gives us the values of $x_{1001}, x_{1002}, ..., x_{2001}$ up to an error of 2×10^{-5}. Continuing in this manner we conclude that for our given set $\Delta^* = (d_1^*, ..., d_{t-1}^*)$ there are at most

$$10^5 \cdot t^2 \cdot 4^{\lfloor t/1000 \rfloor} \tag{1.2}$$

possibilities for the values of $(10^{-5} \lfloor 10^5 x_1 \rfloor, 10^{-5} \lfloor 10^5 x_{1001} \rfloor, ..., 10^{-5} \lfloor 10^5 x_{1000 \lfloor t/1000 \rfloor + 1} \rfloor)$ and knowing these values we know each x_j, up an error of 2×10^{-5} (and hence there are at most two possibilities for $\lfloor x_j \rfloor$).

Let us call a block of 1000 consecutive x_j's of the form $(x_{(x-1)1000+1}, ..., x_{i1000})$ *bad* if $d_{(i-1)1000+1}^* < d_{i \cdot 1000+1}^*$. Otherwise, we call it *good*. Since $d_{t-1} \leqslant 10^{-14} t$ the number of bad blocks is at most $10^{-6} t$ (since the difference between any two nonequal d_j^*'s is at least 10^{-8}). These bad blocks contain at most $\lceil 10^{-3} t \rceil$ x_j's, and there are at most

$$2^{\lceil 10^{-3} t \rceil} \tag{1.3}$$

possibilities for their floors. To complete the proof of the lemma we estimate the number of possibilities for the floors of the members of the good blocks. Let $D = (x_{(i-1)1000+1}, ..., x_{i1000})$ be a good block. The value of each member of D is known, up to an error of 2×10^{-5}, and we also know that the differences between any pair of consecutive members of the block differ from each other by at most 10^{-8}. Let us call a member x_j of the block *sure* if $\lfloor x_j \rfloor$ is determined from our knowledge about the value of x_j. Otherwise, call it *unsure*. If the number of unsure elements in the block is at most 500, then there are at most 2^{500} possibilities for the values of $\{\lfloor x_j \rfloor : x_j \in D\}$. Else, there are two consecutive unsure elements in the block. Each of these two elements differs from an integer by at most 2×10^{-5} and hence their difference differs from an integer d by at most 4×10^{-5}. As D is a good block, the differences between consecutive elements of D satisfy $d_{(i-1)1000+1} \leqslant d_{(i-1)1000+2} \leqslant \cdots \leqslant d_{i1000-1} \leqslant d_{(i-1)1000+1} + 10^{-8}$. It thus follows that each member of the block differs from an integer by not more than 5×10^{-2}. Let $\{x_j : j \in J\}$ be the collection of all unsure elements in C. Each $\lfloor x_j \rfloor$ can be, according to our estimate for x_j, either n_j or $n_j + 1$. Suppose that for a sequence X and for some $i < j$, $i, j \in J$ we have $\lfloor x_i \rfloor = n_i$ and $\lfloor x_j \rfloor = n_j + 1$. Then one can easily check that $d_{j-1} = x_j - x_{j-1}$ is slightly more than the integer d, as $x_j - x_i$ is slightly more than $d \cdot (j - i)$ and the sequence is ascending. It thus follows that in this case, for every $l > j$, $l \in j$, we have $\lfloor x_l \rfloor = n_l + 1$. But this means that the number of possibilities for $(\lfloor x_j \rfloor : j \in J)$ is at most $\binom{|J|+1}{2} + 1 < 1000^2 < 2^{500}$.

Therefore, in any case, the number of possibilities for $(\lfloor x_j \rfloor : x_j \in D)$ for any good block D is at most 2^{500}, and the total number of possibilities for the floors of all the members of the good blocks is at most

$$2^{t/2}. \tag{1.4}$$

We conclude that $|\lfloor R_t \rfloor|$ is bounded by the product of the estimates in (1.1), (1.2), (1.3), and (1.4), as needed. As the expression in (1.1) can be bounded by $2^{H(1/(1+10^{-6}))\cdot(1+10^{-6})t}$, where $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function this implies that for large t, $|\lfloor R_t \rfloor| \leqslant 2^{2t/3}$, as claimed. ∎

It is worth noting that with some additional effort the estimate in Lemma 1.7 can be improved, but for our purposes here the present bound suffice.

We now return to our random coloring C of the integers $1, 2, ..., 4bm$.

LEMMA 1.8. *For sufficiently large k, the probability that there is, in C, a monochromatic ascending wave of length $t = \lfloor k/4 \rfloor$ whose first difference is at least $4b$ and whose last difference is smaller than $10^{-14}t\cdot b = 10^{-14}\lfloor k/4 \rfloor \cdot \lfloor k/40 \rfloor$ is smaller than* 0.01.

Proof. Let Q denote the set of all ascending waves in $N = \{1, 2, ..., 4bm\}$ whose first difference is bigger than $4b$ and whose last difference is smaller than $10^{-14}\cdot t\cdot b$. For each member $A = (a_1 < a_2 < \cdots a_t)$ of Q define an associated real ascending t-wave $X = X(A) \in R_t$ by $X = (x_1 < x_2 < \cdots < x_t)$ where $x_i = a_i/b$. Recall that $\lfloor X \rfloor = (\lfloor x_1 \rfloor, \lfloor x_2 \rfloor, ..., \lfloor x_t \rfloor)$ and notice that A is monochromatic if and only if all blocks $\{B_{\lfloor x_i \rfloor} : 1 \leqslant i \leqslant t\}$ have the same color. By Lemma 1.2 part (iv) for each fixed sequence $\lfloor X(A) \rfloor = (\lfloor x_1 \rfloor, ..., \lfloor x_t \rfloor)$, the probability that all the blocks $\{B_{\lfloor x_i \rfloor} : 1 \leqslant i \leqslant t\}$ have the same color is precisely $2/2^t$. By Lemma 1.7, for sufficiently large t there are at most $2^{2t/3}$ possible sequences $\lfloor X(A) \rfloor$. Therefore, the probability that there is a monochromatic ascending wave A in Q is bounded by $2^{2t/3}\cdot 2/2^t$, which is smaller than 0.01, for all sufficiently large t. ∎

We can now complete the proof of Theorem 1.1. By Corollary 1.5 and Lemma 1.8, for sufficiently large k, there is a coloring C of $N = \{1, 2, ..., 4bm\} = \{1, 2, ..., 4\cdot\lfloor k/40 \rfloor \cdot \lfloor 10^{-20}k^2 \rfloor\}$ that satisfies the following two properties;

The last difference of any monochromatic ascending wave of length $\lceil k/2 \rceil$ is at least $4b$. (1.5)

The last difference of any monochromatic ascending wave of length $\lfloor k/4 \rfloor$ whose first difference is at least $4b$, is at least $10^{-14}\lfloor k/4 \rfloor \cdot \lfloor k/40 \rfloor > 10^{-17}\cdot k^2$. (1.6)

Let us assume that there is a monochromatic ascending wave $A=(a_1<a_2<\cdots<a_k)$ of length k in N. By (1.5), $a_{\lceil k/2\rceil}-a_{\lceil k/2\rceil-1}\geqslant 4b$. Therefore, by (1.6), $a_{\lceil 3k/4\rceil+1}-a_{\lceil 3k/4\rceil}>10^{-17}\cdot k^2$. However, since A is ascending this implies that $a_{i+1}-a_i>10^{-17}k^2$ for all $i>\lceil 3k/4\rceil$. Therefore $a_k>10^{-17}k^2\cdot\lfloor k/4\rfloor>10^{-18}k^3>4bm$, and hence A is not contained in N, contradicting our hypothesis. This shows that $f(k)\geqslant\Omega(k^3)$ and completes the proof of Theorem 1.1. ∎

2. Ascending Waves in Dense Sets

In this section we prove Theorem 2.1.

Upper bound. Our construction is a discrete form of a Cantor set. For $i\geqslant 1$, $1\leqslant r\leqslant 2^i-1$, r odd, let A_i^r denote those integers x,

$$nr2^{-i}\leqslant x<nr2^{-i}+n2^{-i}/2\log n.$$

Let A_i denote the union of A_i^r over these r. As $|A_i^r|\leqslant n2^{-i}/2\log n+1$, $|A_i|\leqslant n/4\log n+2^{i-1}$. Let s be maximal with $2^s<n/2\log n$. Let B equal the union of A_i, $1\leqslant i\leqslant s$ and $S=\{1,...,n\}-B$,

$$|B|\leqslant\sum_{i=1}^{s}(n/4\log n+2^{i-1})\leqslant n/4+2^s,$$

so $|S|\geqslant n/2$.

Now let $a_1, a_2, ...$ be an AW in S. Let $1<i\leqslant s+1$ and suppose $x_{t+1}-x_t\geqslant n2^{-i}/2\log n$. Then $x_{t+8\log n}-x_t\geqslant n2^{-(i-2)}$. Hence $[x_t, x_{t+8\log n}]$ must intersect some A_{i-1}^r and therefore $x_{u+1}-x_u>2^{-(i-1)}/2\log n$ for some u, $t\leqslant u\leqslant t+8\log n$. The hypothesis holds for $i=s+1$, $k=1$ (as $x_2-x_1>1$) so for some u, $1\leqslant u\leqslant 1+s(8\log n)\leqslant 8\log^2 n$, it holds for $i=2$, i.e., $x_{u+1}-x_u>n/4\log n$. The AW can have only $4\log n$ further terms so its total length is less than $8\log^2 n+4\log n=(8+o(1))\log^2 n$.

Lower bound. Fix $S\subseteq[n]$, $|S|=n/2$. For $0\leqslant i\leqslant\log n$ an i-gap is a gap in S with length in $[2^i, 2^{i+1})$. Let u_i be the total length of all i-gaps so

$$u_0+\cdots+u_{\log n}=n/2.$$

Set

$$u_i'=\sum_{j\geqslant 0}u_{i+j}2^{-j}.$$

Then

$$\sum_{i=0}^{\log n} u_i' = \sum_{k=0}^{\log n} u_k \sum_{j \leqslant k} 2^{-j}$$
$$\leqslant 2 \sum_{k=0}^{\log n} u_k = n.$$

Also, let

$$u_i'' = \sum_{j=0}^{\log \log n} u_{i-j}.$$

Then

$$\sum_{i=0}^{\log n} u_i'' \leqslant (\log \log n) \sum_{k=0}^{\log n} u_k = n(\log \log n)/2.$$

We call i OK if

$$u_i' < 4n/\log n \tag{2.1}$$

$$u_i'' < 2n(\log \log n)/\log n \tag{2.2}$$

$$i < \log n - 10 \log \log n. \tag{2.3}$$

Let I denote the set of OK i. At most one quarter of the i fail (2.1), at most one quarter fail (2.2) and $o(1)$ of the i fail (2.3), so $|I| \geqslant 0.49 \log n$.

For $i \in I$ call $x \in [n]$ i-fine if

$$x \in S \tag{2.4}$$

$$x + 2^i \log n/100 \leqslant n \tag{2.5}$$

$[x, x + 2^i \log n/100)$ intersects no $(i+j)$-gap, $j \geqslant 0$. (I.e., no gap of size at least 2^i.) (2.6)

In $(x, x + 2^i \log n/100)$ the sum of all $(i-j)$-gaps, $1 \leqslant j \leqslant \log \log n$ (i.e., of size between $2^i/\log n$ and 2^i), is less than $2^i \log \log n$. (2.7)

CLAIM. *At least* $0.43\, n$ $x \in [n]$ *are i-fine.*

Proof. Precisely $n/2$ $x \in [n]$ fail (2.4) and, by (2.3), $o(n) \leqslant 0.01 n$ $x \in [n]$ fail (2.5). If x satisfying (2.4), (2.5) fails (2.6) then

$$y - 2^i \log n/100 \leqslant x < y,$$

where y is the leftmost element of an $(i+j)$-gap, $j>0$. There are at most $u_{i+j}2^{-i-j}$ such gaps and so at most

$$(2^i \log n)/100 \sum_{j\geq 0} u_{i+j}2^{-i-j} = u_i'(\log n)/100 \leqslant n/25$$

(by (2.1)) such x.

Let A be the union of all $(i-j)$-gaps, $1 \leqslant j \leqslant \log\log n$, so that $|A| = u_i'' \leqslant 2n \log\log n/\log n$ by (2.2). Then

$$\sum_{x=1}^{n} |[x, x+2^i \log n/100] \cap A| \leqslant |A|\, 2^i \log n/100$$

$$\leqslant 2^i n(\log\log n)/50$$

and (2.7) fails for at most $n/50$ $x \in [n]$.

Altogether at most $0.57nx$ are not i-fine, completing the proof of the claim.

Double counting, there exists $x \in [n]$ which is i-fine for at least $0.43|I| \geqslant (\log n)/10$, $i \in I$. Fix this x and let I^* be those $i \in I$ for which x is i-fine.

The AW will be found by "splicing together" AWs for the various $i \in I^*$. The individual AWs are given by the following result.

CLAIM. *Let x be i-fine. Let y satisfy*

$$x \leqslant y \leqslant x + 2^i \log n/200.$$

That is, y is at most "halfway" through the interval $(x, x+2^i \log n/100)$. Then there is an AW

$$y = x_0 < x_1 < \cdots < x_u$$

of length $u = 10^{-5} \log n/\log\log n$ satisfying

$$x_u < x_0 + 2^i \log n/200 \tag{2.8}$$

$$x_i - x_0 \geqslant 10 \cdot 2^i \log \log n \tag{2.9}$$

$$x_u - x_{u-1} \leqslant 20 \cdot 2^i \log\log n. \tag{2.10}$$

Proof. We find this AW by a "greedy algorithm." Set $x = y$, $\Delta_0 = 10 \cdot 2^i \log\log n$. Now, by induction, let x_t be the least element of S which is at least $x_{t-1} + \Delta_{t-1}$ and set $\Delta_t = x_t - x_{t-1}$. Set $D_t = \Delta_t - \Delta_{t-1}$. Each D_t is at most the length of the gap containing $x_{t-1} + \Delta_{t-1}$; when $x_{t-1} + \Delta_{t-1} \in S$, $D_t = 0$.

Assume (2.8). The gaps containing $x_{t-1}+\Delta_{t-1}$, $1 \leqslant t \leqslant u$, then all lie in $(x, x+2^i \log n/100)$. Critically, they are all disjoint as they are separated by elements of the AW. Thus

$$D_1 + \cdots + D_u \leqslant 2^i \log\log n + u(2^i/\log n) \leqslant 2 \cdot 2^i \log\log n.$$

Here $2^i \log\log n$ is a bound by (2.6), (2.7) on the sum of *all* gaps in the interval of size at least $2^i/\log n$ and $u(2^i/\log n)$ bounds the sum of those terms less than $2^i/\log n$. Now

$$x_u - x_{u-1} \leqslant \Delta_0 + D_1 + \cdots + D_u \leqslant 20 \cdot 2^i \log\log n$$

as desired.

We show (2.8) by a reductio ad absurdum. Assume to the contrary that $x_v < x_0 + 2^i \log n/200 \leqslant x_{v+1}$ for some $v < u$. Applying the above argument to v

$$\Delta_v = x_v - x_{v-1} \leqslant 20 \cdot 2^i \log\log n$$

and thus

$$x_v - x_0 \leqslant v\,\Delta_v \leqslant 2 \cdot 10^{-4} \cdot 2^i \log n.$$

But then $x_v + \Delta_v < x_0 + 2^i \log n/200$ so $D_{v+1} \leqslant 2^i$ and

$$\begin{aligned} x_{v+1} &= x_v + \Delta_v + D_{v+1} \\ &\leqslant x_0 + (2 \cdot 10^{-4} \cdot 2^i \log n) + (20 \cdot 2^i \log\log n) + 2^i \\ &< x_0 + 2^i \log n/200 \end{aligned}$$

the desired contradiction.

Now we complete the lower bound argument. Let $i' < i$ be successive elements of I^* and assume, by induction, that an AW

$$x = z_0 < z_1 < \cdots < z_w$$

has been created with

$$\begin{aligned} z_w &< x + 2^{i'} \log n/100 \\ z_w - z_{w-1} &< 20 \cdot 2^{i'} \log\log n. \end{aligned}$$

Since $i' \leqslant i-1$, $2^{i'} \leqslant 2^i/2$ and so

$$\begin{aligned} z_w &< x + 2^i \log n/200 \\ z_w - z_{w-1} &< 10 \cdot 2^i \log\log n. \end{aligned}$$

Apply the claim for i with $y = z_w$ to give an AW that we relabel $z_w < z_{w+1} < \cdots < z_{w+u}$. As

$$z_{w+1} - z_w \geqslant 10 \cdot 2^i \log\log n > z_w - z_{w-1}$$

we may *append* this wave onto the old AW giving an AW $z_0 < \cdots < z_{w+u}$ still satisfying the induction hypothesis.

As each $i \in I^*$ adds $u = 10^{-5} \log n/\log\log n$ elements to the AW, the final AW has

$$u|I^*| \geqslant 10^{-6} \log n/\log\log n$$

elements.

3. Concluding Remarks

A *real* ascending wave of length k is a sequence of k real numbers $0 \leqslant a_1 < a_2 < \cdots < a_k$ such that $a_{i+1} - a_i \leqslant a_{i+2} - a_{i+1}$ for all $1 \leqslant i \leqslant k-2$. Our proof for Theorem 1.1 actually gives the following, somewhat stronger, result: There is a positive constant $c > 0$ such that for every $k \geqslant 1$ there is a 2-coloring of the real interval $[0, ck^3]$ which contains no monochromatic real ascending waves $a_1 < a_2 < \cdots < a_k$ of length k, with $a_2 - a_1 \geqslant 1$.

The results of [BEF] imply that for any α, $0 < \alpha < 1$, there is a positive constant $c = c(\alpha)$ such that any set $A \subset \{1, 2, ..., n\}$ of cardinality $|A| \geqslant n^\alpha$ contains an ascending wave of length at least $c \cdot \log n$. Our Cantor-set-type construction for the upper bound of Theorem 2.1 can be easily modified to show that this is sharp. Namely, for every α, $0 < \alpha < 1$, there is a positive constant $d(\alpha)$ and a set $A \subset \{1, ..., n\}$ of cardinality $|A| \geqslant n^\alpha$ which contains no ascending wave of length greater than $d(\alpha) \log n$.

We close with some imprecise remarks concerning attempts to improve the lower bound of Theorem 2.1. Let x be i-fine. We would like to eliminate the $\log\log n$ factor in the claim. We try to set $u = 10^{-5} \log n$ and $\Delta_0 = x_1 - x_0 = 10 \cdot 2^i$. Say that a gap G in $(x, x + 2^i \log n/100)$ is "hit" if some $x_j + \Delta_j \in G$. As all $x_j - x_{j-1} > 2^i$ an $(i-j)$-gap G has "probability" less than 2^{-j} of being hit. The total size of all $(i-j)$-gaps, $j > 0$, that are hit has "expectation" less than

$$\sum_{j>0} u_{i-j} (2^i \log n/100\, n)\, 2^{-j} = \sum_{j>0} 2^{i-j}/100 = 2^i/100,$$

when all $u_i = n/\log n$. If for a given x gaps with total size less than $2^i/100$ are hit then $D_1 + \cdots + D_u \leqslant 2^i/100$ and the proof succeeds.

There are several problems with this approach. The probability is not clearly defined, do we select x at random? The u_i are not necessarily uniform, though this can probably be dealt with by an averaging argument. Most critically, an $(i-j)$-gap G may have "probability" far greater than 2^{-j} of being hit of earlier gaps force the greedy algorithm AW to "focus in" on G. Nonetheless, we make the following guess as to the true state of affairs.

Conjecture. $g(n) = \Theta(\log^2 n)$.

Reference

[BEF] T. C. Brown, P. Erdös, and A. R. Freedman, Quasi-progressions and descending waves, *J. Combin. Theory Ser. A*, in press.