Progressions in Sequences of Nearly Consecutive Integers

Noga Alon * Ayal Zaks †

Abstract

For $k > 2$ and $r \geq 2$, let $G(k, r)$ denote the smallest positive integer g such that every increasing sequence of g integers $\{a_1, a_2, \ldots, a_g\}$ with gaps $a_{j+1} - a_j \in \{1, \ldots, r\}$, $1 \leq j \leq g - 1$ contains a k-term arithmetic progression. Brown and Hare [4] proved that $G(k, 2) > \sqrt{(k-1)/2}(\frac{4}{3})^{(k-1)/2}$ and that $G(k, 2s-1) > (s^{k-2}/ek)(1+o(1))$ for all $s \geq 2$. Here we improve these bounds and prove that $G(k, 2) > 2^{k-O(\sqrt{k})}$ and, more generally, that for every fixed $r \geq 2$ there exists a constant $c_r > 0$ such that $G(k, r) > r^{k-c_r\sqrt{k}}$ for all k.

A sequence of integers $\{a_1, a_2, \ldots, a_g\}$ is called *nearly consecutive* if $a_{j+1} - a_j \in \{1, 2\}$ for $1 \leq j \leq g - 1$. Let $G(k, 2)$ denote the smallest positive integer g such that every nearly consecutive sequence of length g contains a k-term arithmetic progression. Brown and Hare [4] proved that $G(k, 2) > \sqrt{(k-1)/2}(\frac{4}{3})^{(k-1)/2}$. Their proof is probabilistic: each gap $a_{j+1} - a_j$ is chosen randomly and independently to be either 1 or 2 with equal probability, and the length of the sequence is chosen so that the expected number of arithmetic progressions of length k it contains is smaller than 1.

In this short paper we first show that there exists a nearly consecutive sequence $\{a_i\}_{i=1}^{g}$ where $g > 2^{k-10\sqrt{k}-1}$ that does not contain any arithmetic progression of length k, provided k is large enough. Our proof is also probabilistic, but uses a slightly more sophisticated probabilistic construction. The first idea is to choose the gaps of size 1 with probability p which is much smaller than $\frac{1}{2}$, thus giving the gaps of size 2 a higher probability, to obtain a sequence which is as sparse as possible. The second idea is that the "bad" events of containing potential arithmetic progressions are nearly independent, and thus there should be a way of applying the Lovász Local Lemma to improve the resulting bound. Unfortunately, in the construction based on the Markov process described above, each event does depend on all others. We therefore apply an additional trick, which is similar to the one used in [1], and make our construction in two steps in order to reduce the dependencies between the events. First we choose a random subset of the elements of the sequence with large gaps between them,

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email address: noga@math.tau.ac.il. Research supported in part by a USA Israeli BSF grant and by a grant from the Israel Science Foundation.

†Department of Statistics and Operations Research, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email address: ayalz@math.tau.ac.il.

Mathematics Subject Classification (1991): 05D10.

making sure each potential progression does not contain too many of these elements, and then we fill these large gaps and obtain the desired nearly consecutive sequence. The resulting lower bound is given in the following theorem.

Theorem 1 $G(k,2) > 2^{k-O(\sqrt{k})}$.

Our arguments can be extended to deal with sequences of bigger gaps as well. For any two integers $a < b$, denote the set $\{a, a+1, \ldots, b\}$ by $[a,b]$. A sequence of integers $\{a_1, a_2, \ldots, a_g\}$ is called a $[1,r]$-*gap sequence* if $a_{j+1} - a_j \in [1,r]$ for $j \in [1, g-1]$[1]. Let $G(k,r)$ denote the smallest positive integer g such that every $[1,r]$-gap sequence of length g contains a k-term arithmetic progression. Brown and Hare [4] proved that $G(k, 2s-1) > (s^{k-2}/\mathrm{e}k)(1+o(1))$, where e is the base of the natural logarithm. Their proof uses the following probabilistic construction: each a_j is chosen arbitrarily from the interval $[(j-1)s+1, js]$, thereby generating a $[1, 2s-1]$-gap sequence. Using the Lovász Local Lemma the authors show that with positive probability this sequence contains no k-term progression, provided the length of the sequence does not exceed $(s^{k-2}/\mathrm{e}k)(1+o(1))$.

Extending the proof of Theorem 1, we prove the following.

Theorem 2 *For every fixed $r \geq 2$ there is a constant c_r so that $G(k,r) > r^{k-c_r\sqrt{k}}$ for all $k > 2$.*

In order to prove Theorem 2 we prove the existence of a $[1,r]$-gap sequence $\{a_i\}_{i=1}^{g}$ where $g > r^{k-(2\log r+5)\sqrt{k}-1}$ that does not contain any arithmetic progression of length k, provided k is large enough. As in the proof of Theorem 1 this is done by first choosing a random subset of the elements of the sequence with large gaps between them, making sure each potential progression does not contain too many of these elements, and then by filling these large gaps. This two-step process reduces the dependencies between the "bad" events of containing potential progressions. Inside each large gap we allow gaps of $r-1$ or r only, choosing the gaps of size $r-1$ with probability p which is much smaller than $\frac{1}{2}$, thus giving the gaps of size r a higher probability. This is done to obtain a sequence which is as sparse as possible.

The *van der Waerden number* $W(k,r)$ is the least integer w such that for any covering of $[1,w]$ by r sets $\cup_{i=1}^{r} A_i \supseteq [1,w]$, at least one of the sets A_i contains an arithmetic progression of length k. As proved in [9] (cf., also, [5]) this number is finite for every k and r. Rabung [7] (see also [6]) observed that $G(k,r) \leq W(k,r)$, since the union of any $[1,r]$-gap sequence with $r-1$ shifted copies of itself covers all integers between the smallest and the largest element of the sequence.

The best known lower bound for $W(k,r)$ is $W(k,r) > \frac{r^k}{\mathrm{e}rk}(1+o(1))$, see, e.g., [5], while for $r=2$ and for any prime p it is known that $W(p+1,2) \geq p2^p$, as proved in [3]. Note that both these bounds, as well as our bounds for $G(k,r)$ mentioned in Theorems 1 and 2, are asymptotically $(r+o(1))^k$. Thus there are r $[1,r]$-gap sequences whose union covers a set of almost r^k consecutive integers with no k-term progressions in any of them. Note also that any lower bound for $G(k,r)$

[1]Thus, in particular, a $[1,2]$-gap sequence is a nearly consecutive sequence.

which is significantly bigger than those in Theorems 1 or 2 would improve the known lower bound for $W(k,r)$ as well. As mentioned in [4], the problem of improving the best known upper bound for $G(k,2)$ (which follows from the best known bound for $W(k,2)$, due to Shelah [8]) is also interesting.

In the rest of this note we present the proofs of the two theorems. Note that the assertion of Theorem 2 contains that of Theorem 1, but since the proof of the first theorem is a bit simpler we prefer to describe it separately.

Proof of Theorem 1. We omit all floor and ceiling functions, for the sake of brevity. Let k be a sufficiently large integer, and set $n = 2^{k-10\sqrt{k}}$ (we do not attempt to optimize the constants here and in what follows).

Let $C = \{c_i\}$ be a sequence of integers, and let $\overline{C} = \cup_i \{x : c_i \leq x < c_i + k\}$. An arithmetic progression $D = \{a + dj\}_{j=1}^{k}$ having $d > k$ is called *bad* (with respect to C) if $|D \cap \overline{C}| \geq 3\sqrt{k}$. Finally, C is called *bad* if there exists a bad arithmetic progression. Let $m = n^{1/\sqrt{k}}$ and $l = \frac{n}{m}$. A simple probabilistic argument shows that there exists a sequence $B = \{b_i\}_{i=1}^{l}$ where $(i-1)m < b_i \leq im$ for each i, which is not *bad*, provided k is large enough. Indeed, if each b_i is chosen randomly and independently, the expected number of bad arithmetic progressions is less than $n^2 \binom{k}{3\sqrt{k}} (\frac{k}{m})^{3\sqrt{k}}$ which is smaller than 1, for all sufficiently large k. This holds even if we fix $b_1 = 1$ and $b_l = n$.

We complete such a sequence B into a nearly consecutive sequence $A = \{a_i\}_{i=1}^{g} \subset [1,n]$ in the following way. Let $0 < p < 1$ be some constant, which will be determined later. Start with $a_1 = b_1 = 1$. Suppose $\{a_1, \ldots, a_j\}$ have already been determined. If $a_j + 1 \in B$ set $a_{j+1} = a_j + 1$, so that eventually $A \supset B$. If $a_j + 1 \notin B$, choose a_{j+1} to be either $a_j + 1$ (with probability p) or else $a_j + 2$ (with probability $1-p$), where all choices are mutually independent. If $a_{j+1} = b_l$ stop, and set $g = j+1$. Clearly, $g > n/2$.

For every $a \in [1,n] \setminus B$,

$$\mathrm{Prob}[a \in A] = \mathrm{Prob}[a-1 \notin A] + p \cdot \mathrm{Prob}[a-1 \in A] = 1 - (1-p)\mathrm{Prob}[a-1 \in A]. \tag{1}$$

The boundary condition for (1) is: $\mathrm{Prob}[b_i \in A] = 1$, where $b_i \in B$ such that $b_i < a < b_{i+1}$. Solving (1) yields the following formula: for every $b_i < a < b_{i+1}$,

$$\mathrm{Prob}[a \in A] = \mathrm{Prob}[a \in A | b_i \in A] = \delta_p(a - b_i), \text{ where } \delta_p(x) = \frac{1}{2-p} + \frac{1-p}{2-p}(p-1)^x. \tag{2}$$

Let $\sigma_p(x) = \frac{1}{2-p} + \frac{(1-p)^{x+1}}{2-p}$. Following are simple bounds for the values of $\delta_p(x)$:

$$\begin{aligned}
\delta_p(x) &\leq \delta_p(2) = 1 - p + p^2 \text{ for every } x \geq 1, && (3)\\
\delta_p(x) &\leq \sigma_p(x) \text{ for every } x \geq 1, && (4)\\
\sigma_p(x) &\leq \sigma_p(k) < \frac{1}{2} + \frac{1}{\sqrt{k}} \text{ for } p = \frac{1}{\sqrt{k}} \text{ and every } x \geq k. && (5)
\end{aligned}$$

Before continuing with the proof, we state the asymmetric form of the Lovász local lemma we use (cf., e.g. [2], [5]).

The Lovász local lemma. *Let $A_1, \ldots, A_n$ be events in a probability space Ω, and let $G = (V, E)$ be a graph on $V = [1, n]$ such that for all i, the event A_i is mutually independent of $\{A_j : (i, j) \notin E\}$. Suppose that there exist $x_1, \ldots, x_n$, $0 < x_i < 1$, so that for all i,* $\mathrm{Prob}[A_i] < x_i \prod_{(i,j)\in E}(1-x_j)$. *Then* $\mathrm{Prob}[\wedge \overline{A_i}] > 0$.

For any k-term arithmetic progression $U = \{u_1, u_2, \ldots, u_k\}$ in $[1, n]$ denote by E_U the event "$U \subseteq A$" and let $B_U = \cup_{i=1}^{k}[b_{j(i)}, b_{j(i+1)}]$ where $b_{j(i)}, b_{j(i+1)} \in B$ such that $b_{j(i)} \leq u_i \leq b_{j(i+1)}$. Event E_U is mutually independent of all events $E_{U'}$ such that $U' \cap B_U = \emptyset$. For a fixed U and gap g there are at most $2mk^2$ progressions U' of gap g such that $U' \cap B_U \neq \emptyset$: there are at most k different intervals in B_U which U' can intersect and any such interval contains at most $2m$ elements one of which belongs to U' in one of k possible positions. Let the symbols S, T denote k-term arithmetic progressions in $[1, n]$ having a gap $\leq k$ and a gap $> k$ respectively. Every event E_U is mutually independent of all but at most $d_S = 2mk^3$ events of type E_S, and of all but at most $d_T = 2mk^2 \frac{n}{k} = 2mkn$ events of type E_T. We next show that for an appropriate choice of p, there exist $0 < x_S, x_T < 1$ such that

$$\left\{ \begin{array}{lcl} \mathrm{Prob}[E_S] & < & x_S(1-x_S)^{d_S}(1-x_T)^{d_T} \\ \mathrm{Prob}[E_T] & < & x_T(1-x_S)^{d_S}(1-x_T)^{d_T} \end{array} \right\}.$$

Set $p = \frac{1}{\sqrt{k}}$. We bound the probabilities of each event E_S as follows. Suppose $S = \{s_1, s_2, \ldots, s_k\}$ is an arithmetic progression with $1 \leq s_1 < s_2 < \ldots < s_k \leq n$. Then

$$\mathrm{Prob}[E_S] = \prod_{i=1}^{k} \mathrm{Prob}[s_i \in A | s_1, \ldots, s_{i-1} \in A] = \prod_{i=1}^{k} \mathrm{Prob}[s_i \in A | s_{i-1} \in A].$$

For every $i \in [1, k]$, let $j(i)$ be such that $b_{j(i)} \leq s_i < b_{j(i)+1}$, where $b_{j(i)} \in B$. Similar to the derivation of (2), for every $s_i \notin B$: $\mathrm{Prob}[s_i \in A | s_{i-1} \in A] = \delta_p(s_i - \max\{s_{i-1}, b_{j(i)}\})$. Denote $I_S = \{i : s_i \notin B\}$. Since m is much larger than k, $|I_S| \geq k-2$. Therefore, using (3):

$$\mathrm{Prob}[E_S] \leq \prod_{i \in I} \mathrm{Prob}[s_i \in A | s_{i-1} \in A] \leq (\delta_p(2))^{k-2} < \mathrm{e}^{4-\sqrt{k}}.$$

By a similar reasoning, for any event E_T, denote $I_T = \{i : s_i - b_{j(i)} \geq k\}$. By the choice of B, $|I_T| \geq k - 3\sqrt{k}$. Therefore, using (4) and (5):

$$\mathrm{Prob}[E_T] \leq (\sigma_p(k))^{k-3\sqrt{k}} < 2^{6\sqrt{k}-k}.$$

Set $x_S = 2^{-\sqrt{k}}k^{-3}$ and $x_T = 2^{8\sqrt{k}-k}$. Then $(1-x_S)^{d_S} > 1 - \frac{1}{2^9} > \frac{1}{2}$ and $(1-x_T)^{d_T} > 1 - \frac{k}{2^{\sqrt{k}+9}} > \frac{1}{2}$. The proof is completed by observing that, for sufficiently large k:

$$\begin{array}{rcl} \mathrm{e}^{4-\sqrt{k}} & < & 2^{-\sqrt{k}}k^{-3}/4, \\ 2^{6\sqrt{k}-k} & < & 2^{8\sqrt{k}-k}/4. \qquad \square \end{array}$$

Proof of Theorem 2. To simplify the presentation, some of the technical details are postponed to the appendix. By Theorem 1 we may assume that $r > 2$. Fix such r, let k be a sufficiently large integer, and set $n = r^{k-(2\log r+5)\sqrt{k}}$. As before, define $m = n^{1/\sqrt{k}}$, $l = \frac{n}{m}$ and let $B = \{b_i\}_{i=1}^{l}$ be a sequence which is not *bad*, where $b_1 = 1$, $b_l = n$ and $(i-1)m < b_i \leq im \ \forall i \in [2, l-1]$.

The sequence B is completed into a $[1, r]$-gap sequence $A = \{a_i\}_{i=1}^{g} \subset [1, n]$ in the following way. Let $p = \frac{2\log r}{\sqrt{k}}$. Start with $a_1 = b_1 = 1$. Suppose $\{a_1, \ldots, a_j\}$ have already been determined. If $a_j + x \in B$ for some $x \in [1, r-1]$, set $a_{j+1} = a_j + x$. Otherwise choose a_{j+1} to be either $a_j + r - 1$ (with probability p) or else $a_j + r$ (with probability $1-p$), where all choices are mutually independent. If $a_{j+1} = b_l$ stop, and set $g = j + 1$. Clearly, $g > \frac{n}{r}$.

Let $a \in [1, n] \setminus B$, and $b_i \in B$ be such that $b_i < a < b_{i+1}$. If $a \geq b_i + r$ then

$$\mathrm{Prob}[a \in A] = p \cdot \mathrm{Prob}[a - (r-1) \in A] + (1-p)\mathrm{Prob}[a - r \in A]. \tag{6}$$

The boundary conditions for (6) are:

$$\mathrm{Prob}[x \in A] = \begin{cases} 1 & \text{if } x = b_i \\ 0 & \text{if } x \in [b_i + 1, b_i + r - 2] \\ p & \text{if } x = b_i + r - 1. \end{cases} \tag{7}$$

The corresponding characteristic polynomial is: $f(x) = x^r - px - (1-p) = (x-1)(x^{r-1} + x^{r-2} + \ldots + x^2 + x + 1 - p)$. Let $f_1 = 1, f_2, \ldots, f_r$ be the roots of $f(x)$. For any large k, as $p = \frac{2\log r}{\sqrt{k}}$, it is easy to check that f has no multiple roots (see lemma 6 in the appendix for details). Therefore, solving (6) yields the following formula: for every $a \in [b_i + r, b_{i+1} - 1]$,

$$\mathrm{Prob}[a \in A] = \mathrm{Prob}[a \in A | b_i \in A] = \delta_{p,r}(a - b_i), \ \text{ where } \delta_{p,r}(x) = c_1 + \sum_{i=2}^{r} c_i f_i^x,$$

and $c_1, c_2, \ldots, c_r$ are constants depending only on p and r (and not on x). A simple upper bound of $\delta_{p,r}(x)$ is (see lemma 3 in the appendix):

$$\delta_{p,r}(x) \leq \delta_{p,r}(r) = 1 - p \ \text{ for every } x \geq 1. \tag{8}$$

It is not difficult to see that $|f_i| < 1 \ \forall i \in [2, r]$ (see corollary 8 in the appendix), implying that $\delta_{p,r}(x)$ converges exponentially fast to c_1. It follows that c_1, being the stationary distribution of the Markov process, is equal to the asymptotic density of A, which is $\frac{1}{r-p}$. The values of $\delta_{p,r}(x)$ can be bounded as follows (for a complete proof see appendix, lemma 14), provided k is sufficiently large:

$$\delta_{p,r}(x) \leq \frac{1}{r}\left(1 + \frac{2}{\sqrt{k}}\right) \text{ for every } x \geq k. \tag{9}$$

Let the symbols S, T, U, E_S, E_T, E_U be defined as before. Using (8) and (9):

$$\begin{aligned} \mathrm{Prob}[E_S] &\leq (\delta_{p,r}(r))^{k-2} < r^{6-2\sqrt{k}}, \\ \mathrm{Prob}[E_T] &\leq \left(\frac{1}{r}\left(1 + \frac{2}{\sqrt{k}}\right)\right)^{k-3\sqrt{k}} < r^{5\sqrt{k}-k}. \end{aligned}$$

Again, every event E_U (E_S or E_T) is mutually independent of all but at most $d_S = 2mk^3$ events of type E_S, and of all but at most $d_T = 2mkn$ events of type E_T. Set $x_S = r^{-\sqrt{k}}k^{-3}$ and $x_T = r^{(2\log r+4)\sqrt{k}-k}/k$. Then $(1-x_S)^{d_S} > 1 - \frac{2}{r^{2\log r+5}} > \frac{1}{2}$, and $(1-x_T)^{d_T} > 1 - \frac{2}{r^{2\log r+5}} > \frac{1}{2}$. The proof is completed by observing that, for sufficiently large k:

$$\begin{aligned} r^{6-2\sqrt{k}} &< r^{-\sqrt{k}}k^{-3}/4, \\ r^{5\sqrt{k}-k} &< r^{(2\log r+4)\sqrt{k}-k}/(4k). \end{aligned}$$

□

References

[1] N. Alon and Z. Füredi, Legitimate colorings of projective planes, *Graphs and Combinatorics* **5** (1989), 95–106.

[2] N. Alon and J. H. Spencer, **The Probabilistic Method**, Wiley, 1992.

[3] E. R. Berlekamp, A Construction for partitions which avoid long arithmetic progressions, *Canad. Math. Bull.* **11** (1968), 409–414.

[4] T. C. Brown and D. R. Hare, Arithmetic progressions in sequences with bounded gaps, *J. Combinatorial Theory, Series A* **77** (1997), 222–227.

[5] R. L. Graham, B. L. Rothschild and J. H. Spencer, **Ramsey Theory**, 2nd ed., Wiley, New York, 1990..

[6] M. B. Nathanson, Arithmetic progressions contained in sequences with bounded gaps, *Canad. Math. Bull.* **23**(4) (1980), 491–493.

[7] J. R. Rabung, On applications of van der Waerden's theorem, *Math. Mag.* **48** (1975), 142–148.

[8] S. Shelah, Primitive recursive bounds for van der Waerden numbers, *J. Amer. Math. Soc.* **1** (1988), 683–697.

[9] B. L. van der Waerden, Beweis einer Baudet'schen Vermutung, *Nieuw Arch. Wiskunde* **15** (1927), 212–216.

Appendix

This appendix supplies proofs of several properties used in the proof of Theorem 2. Throughout the appendix we assume that $r > 2$ and that k is large enough so that $p = \frac{2\log r}{\sqrt{k}} < \frac{1}{2r^2}$. Recall that $f_1 = 1, f_2, \ldots, f_r$ are the roots of $f(x) = x^r - px - (1-p)$, and $\delta_{p,r}(x) = c_1 + \sum_{i=2}^{r} c_i f_i^x$ such that:

$$\begin{array}{lll} \text{(7'):} & \text{for } x \in [0, r-1], & \delta_{p,r}(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \in [1, r-2] \\ p & \text{if } x = r-1 \end{cases} \\ \text{(6'):} & \text{for } x \geq r, & \delta_{p,r}(x) = p \cdot \delta_{p,r}(x - r + 1) + (1-p)\delta_{p,r}(x - r). \end{array}$$

Lemma 3 (inequality (8)) $\forall x \geq 1 : \ \delta_{p,r}(x) \leq \delta_{p,r}(r) = 1 - p.$

Proof. The proof is by induction on x. By (7') this holds for all $x \in [1, r-1]$. Note that $\delta_{p,r}(1) = 0$ since $r > 2$, so that $\delta_{p,r}(r) = p \cdot \delta_{p,r}(1) + (1-p)\delta_{p,r}(0) = 1 - p$ by (6'). Assume by induction that $\delta_{p,r}(i) \leq 1 - p$ for all $i \leq x$, where $x \geq r$. By (6') the value of $\delta_{p,r}(x+1)$ is equal to a convex combination of $\delta_{p,r}(x - r + 2)$ and $\delta_{p,r}(x - r + 1)$ which are both at most $1 - p$. Therefore $\delta_{p,r}(x+1) \leq 1 - p$, completing the proof. □

Lemma 4 $\forall j : \ |f_j| \leq 1.$

Proof. Assuming the contrary, suppose $|f_j| > 1$. Then by the triangle inequality $1 - p = |f_j^r - pf_j| = |f_j| \cdot |f_j^{r-1} - p| > 1 - p$, a contradiction. □

Since $\prod_{j=1}^{r} f_j = 1 - p$, the following corollary is a consequence of lemma 4.

Corollary 5 $\forall j : \ |f_j| \geq 1 - p.$ □

The following lemma asserts that $f(x)$ has no multiple roots.

Lemma 6 $\forall i \neq j : \ f_i \neq f_j.$

Proof. Assume the contrary: $f_i = f_j$. Then $f'(f_j) = 0 \Rightarrow f_j^{r-1} = \frac{p}{r}$, and $f(f_j) = 0 \Rightarrow f_j(f_j^{r-1} - p) = 1 - p$, so that $f_j = \frac{1-p}{p(\frac{1}{r} - 1)} < -r$, since $p < \frac{1}{r}$. But $|f_j| \leq 1$ by lemma 4, a contradiction. □

Lemma 7 $|f_j| = 1 \Rightarrow f_j = 1.$

Proof. Let $f_j = a + bi$ and $|f_j| = 1$ so that $a^2 + b^2 = |f_j|^2 = 1$. Let $f_j^r = c + di$, so that $c^2 + d^2 = |f_j^r|^2 = 1$. Then $(c + di) - p(a + bi) = 1 - p$, so that $c = 1 - p + pa$ and $d = pb$. Now,

$$1 = c^2 + d^2 = (1 - p + pa)^2 + (pb)^2 = 1 - 2p + \underbrace{p^2 + p^2a^2 + p^2b^2}_{2p^2} + 2pa - 2p^2a = 1 - 2p(1-p)(1-a).$$

Since $0 < p < 1$ we conclude that $a = 1$, and therefore $f_j = 1$. □

Combining lemmas 4, 6, 7 and the fact that $f_1 = 1$, we obtain the following corollary.

Corollary 8 $\forall j \in [2,r]: \ |f_j| < 1.$ $\square$

The distance between any pair of roots of $f(x)$, which is strictly positive by lemma 6, can be bounded away from zero as follows.

Lemma 9 $\forall i \neq j: \ |f_i - f_j| > \frac{2}{r}.$

Proof. Developing the Taylor series of $f(x)$ around f_j: $f(f_i) = f(f_j) + (f_i - f_j)f'(f_j) + \frac{(f_i-f_j)^2}{2}f''(y)$ where y lies somewhere on the line between f_i and f_j. By lemma 6, $(f_i - f_j) \neq 0$ and $f'(f_j) \neq 0$, implying that $f''(y) = r(r-1)y^{r-2} \neq 0$. Therefore, $0 = (rf_j^{r-1} - p) + (\frac{f_i - f_j}{2})r(r-1)y^{r-2}$ and $f_i - f_j = \frac{-2(rf_j^{r-1}-p)}{r(r-1)y^{r-2}}$. Now, $1 - p \leq |f_j| \leq 1$ by lemma 4 and corollary 5, so that $|rf_j^{r-1}| = r|f_j|^{r-1} \geq r|f_j|^r \geq r(1-p)^r$. Since $p \leq 1$ we can apply Bernoulli's inequality: $(1-p)^r \geq 1 - rp$, so that $r(1-p)^r \geq r(1-rp) > r - \frac{1}{2}$, using the fact that $p < \frac{1}{2r^2}$. Since $|f_i|, |f_j| \leq 1$ it follows that $|y| \leq 1$, and by the triangle inequality,

$$|f_i - f_j| = \frac{2|rf_j^{r-1} - p|}{|r(r-1)y^{r-2}|} \geq \frac{2(|rf_j^{r-1}| - p)}{r(r-1)} > \frac{2(r - \frac{1}{2} - p)}{r(r-1)} > \frac{2}{r}.$$

$\square$

The following corollary is a special case of lemma 9, taking $f_i = f_1 = 1$.

Corollary 10 $\forall j \in [2,r]: \ |f_j - 1| > \frac{2}{r}.$ $\square$

Note that as p tends to 0, the absolute values of the roots of $f(x)$ approach 1. The following lemma bounds the absolute values of the roots of $f(x)$ (except for f_1) away from 1.

Lemma 11 $\forall j \in [2,r]: \ |f_j|^2 < 1 - \frac{p}{r^3}.$

Proof. Similar to the computation presented in the proof of lemma 7, let $f_j = a + bi$ and $|f_j|^2 = a^2 + b^2 = 1 - \epsilon$ where $0 < \epsilon < 1$ (since $|f_j| < 1$ by corollary 8). We will show that $\epsilon > \frac{p}{r^3}$. Let $f_j^r = c + di$, so that $|f_j^r|^2 = c^2 + d^2 = |f_j|^{2r} = (1-\epsilon)^r \geq 1 - r\epsilon$, using Bernoulli's inequality. Then $(c + di) - p(a + bi) = 1 - p$, implying that $c = 1 - p + pa$ and $d = pb$. Now,

$$1 - r\epsilon \leq c^2 + d^2 = (1 - p + pa)^2 + (pb)^2 = 1 - 2p(1-p)(1-a) - \epsilon p^2 < 1 - p(1-a) \Longrightarrow a > 1 - \frac{\epsilon r}{p}.$$

Recall that $|f_j| < 1$ and $|f_j - 1| > \frac{2}{r}$ (by corollaries 8, 10), which by a simple trigonometric argument imply that $a < 1 - \frac{2}{r^2}$. So $1 - \frac{\epsilon r}{p} < a < 1 - \frac{2}{r^2} \Rightarrow \epsilon > \frac{2p}{r^3} > \frac{p}{r^3}$. Therefore $|f_j|^2 = 1 - \epsilon < 1 - \frac{p}{r^3}$. $\square$

Lemma 12 $\forall j: \ |c_j| < r^{r^2}.$

Proof. Let $M = (m_{i,j})$ be an $r \times r$ matrix having $m_{i,j} = f_j^{i-1}$, $(i, j \in [1,r])$, and let $c = (c_1, \ldots, c_r)^T$. Then by (7'), c is a solution of $Mc = d$, where d is the r-vector: $(1, 0, \ldots, 0, p)^T$. Notice that M is a *van der Monde* matrix, so $|\det M| = \prod_{i>j} |f_i - f_j| \neq 0$ by lemma 6. Hence M is nonsingular, and by Cramer's rule, $c_j = \frac{\det M_j}{\det M}$ where M_j is the matrix obtained from M by replacing its jth column by d. Since for every $i, j: \ |m_{i,j}|, |d_j| \leq 1$, it follows that $|\det M_j| \leq r! < r^r$. By lemma 9, $|f_i - f_j| > \frac{1}{r}$, so $|\prod_{i>j}(f_i - f_j)| > r^{-\binom{r}{2}} > r^{-r^2/2}$. Therefore, $|c_j| < r^r \cdot r^{r^2/2} < r^{r^2}$. $\square$

Lemma 13 $\forall x \geq k: \; |\sum_{j=2}^{r} c_j f_j^x| < \frac{1}{2r\sqrt{k}}$ *if k is sufficiently large.*

Proof. $|f_j|^2 < 1 - \frac{p}{r^3}$ and $|c_j| < r^{r^2}$ by lemmas 11, 12. Therefore, $|\sum_{j=2}^{r} c_j f_j^x| \leq \sum_{j=2}^{r} |c_j| \cdot |f_j|^x \leq \sum_{j=2}^{r} |c_j| \cdot |f_j|^k < r^{r^2+1}(1 - \frac{p}{r^3})^{\frac{k}{2}}$. Recall that $p = \frac{2\log r}{\sqrt{k}}$ so that $\frac{k}{2} > \frac{1}{p^2}$. Therefore $(1 - \frac{p}{r^3})^{\frac{k}{2}} < (1 - \frac{p}{r^3})^{\frac{1}{p^2}} \leq \mathrm{e}^{-\frac{1}{pr^3}}$. If k is sufficiently large (with respect to r) then p is small enough so that: $r^{r^2+1}\mathrm{e}^{-\frac{1}{pr^3}} < \frac{p}{r^3} = \frac{2\log r}{r^3\sqrt{k}} < \frac{1}{2r\sqrt{k}}$. □

Lemma 14 (inequality (9)) $\forall x \geq k: \; \delta_{p,r}(x) < \frac{1}{r}(1 + \frac{2}{\sqrt{k}})$ *if k is sufficiently large.*

Proof. $|\sum_{j=2}^{r} c_j f_j^x| < \frac{1}{2r\sqrt{k}}$ by lemma 13, so that $\delta_{p,r}(x) = c_1 + \sum_{j=2}^{r} c_j f_j^x \leq |c_1| + |\sum_{j=2}^{r} c_j f_j^x| < |c_1| + \frac{1}{2r\sqrt{k}}$. Now c_1 is equal to the asymptotic density of 1's which is $\frac{1}{r-p}$. Therefore,

$$c_1 = \frac{1}{r-p} = \frac{1}{r}\left(\frac{1}{1-\frac{p}{r}}\right) = \frac{1}{r}\sum_{i=0}^{\infty}\left(\frac{p}{r}\right)^i < \frac{1}{r}\left(1 + \frac{3p}{2r}\right) = \frac{1}{r}\left(1 + \frac{3\log r}{r\sqrt{k}}\right) < \frac{1}{r}\left(1 + \frac{3}{2\sqrt{k}}\right).$$

Hence $\delta_{p,r}(x) < (\frac{1}{r} + \frac{3}{2r\sqrt{k}}) + \frac{1}{2r\sqrt{k}} = \frac{1}{r}(1 + \frac{2}{\sqrt{k}})$. □