References

[1] P. X. Gallagher, *A large sieve density estimate near $\sigma = 1$*, Inventiones Math. 11 (1970), pp. 329–339.

[2] M. Jutila, *A statistical density theorem for L-functions with applications*, Acta Arith. 16 (1969), pp. 207–216.

[3] — *On character sums and class numbers*, J. Number Theory 5 (1973), pp. 203–214.

[4] — *A density estimate for L-functions with a real character*, Ann. Acad. Sci. Fenn., Ser. A I, 508 (1972), 10 pp.

[5] — *On a density theorem of H. L. Montgomery for L-functions*, Ann. Acad. Sci. Fenn., Ser. A I, 520 (1972), 13 pp.

[6] H. L. Montgomery, *Mean and large values of Dirichlet polynomials*, Inventiones Math. 8 (1969), pp. 334–345.

[7] — *Topics in Multiplicative Number Theory*, Lecture Notes in Math. 227, Berlin, Heidelberg, New York 1971.

[8] K. Prachar, *Primzahlverteilung*, Berlin 1957.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TURKU
Turku, Finland

Received on 17. 7. 1973 (436)

ACTA ARITHMETICA
XXVII (1975)

On sets of integers containing no k elements in arithmetic progression

by

E. Szemerédi (Budapest)

Dedicated to the memory of Yu. V. Linnik

1. Introduction. In 1926 van der Waerden [15] proved the following startling theorem: *If the set of integers is arbitrarily partitioned into two classes then at least one class contains arbitrarily long arithmetic progressions.* It is well known and obvious that neither class must contain an infinite arithmetic progression. In fact, it is easy to see that for any sequence a_n there is another sequence b_n, with $b_n > a_n$, which contains no arithmetic progression of three terms, but which intersects every infinite arithmetic progression. The finite form of van der Waerden's theorem goes as follows: *For each positive integer n, there exists a least integer $f(n)$ with the property that if the integers from 1 to $f(n)$ are arbitrarily partitioned into two* (1) *classes, then at least one class contains an arithmetic progression of n terms.* (For a short proof, see the note of Graham and Rothschild [7].) However, the best upper bound on $f(n)$ known at present is extremely poor. The best lower bound known, due to Berlekamp [3], asserts that $f(n) > n2^n$, which improves previous results of Erdös, Rado and W. Schmidt.

More than 40 years ago, Erdös and Turán [4] considered the quantity $r_k(n)$, defined to be the greatest integer l for which there is a sequence of integers $0 < a_1 < a_2 < \ldots < a_l \leq n$ which does not contain an arithmetic progression of k terms. They were led to the investigation of $r_k(n)$ by several things. First of all the problem of estimating $r_k(n)$ is clearly interesting in itself. Secondly, $r_k(n) < n/2$ would imply $f(k) < n$, i.e., they hoped to improve the poor upper bound on $f(k)$ by investigating $r_k(n)$. Finally, an old question in number theory asks if there are arbitra-

(1) In fact, van der Waerden proved this for partitions into r classes for any positive integer r.

rily long arithmetic progressions of prime numbers. From $r_k(n) < \pi(n)$ this would follow immediately. The hope was that this problem on primes could be attacked not by using special properties of the primes but by only using the fact that they are numerous, a method which is often successful.

Erdös and Turán observed

$$r_k(m+n) \leqslant r_k(m) + r_k(n)$$

from which it follows by a simple argument that

$$\lim_{n\to\infty} \frac{r_k(n)}{n} = c_k$$

exists. Erdös and Turán conjectured that $c_k = 0$ for all k. A few years later Behrend [1] proved that either $c_k = 0$ for every k, or $\lim_{k\to\infty} c_k = 1$. Erdös and Turán also conjectured $r_k(n) < n^{1-\varepsilon_k}$, which was shown to be false by Salem and Spencer [13] who proved

$$r_3(n) > n^{1-c/\log\log n}.$$

In 1946 Behrend [2] proved

$$r_3(n) > n^{1-c/\sqrt{\log n}}$$

which is the best lower bound for $r_3(n)$ currently known. In [8], L. Moser constructed an infinite sequence which contains no arithmetic progression of three terms and which satisfies Behrend's inequality for every n. Behrend's corresponding inequalities for $k > 4$ were improved by Rankin in [9].

The first satisfactory upper bound for $r_3(n)$ was due to Roth [10] who proved

$$r_3(n) < \frac{cn}{\log\log n}.$$

In 1967, I proved that $r_4(n) = o(n)$. The proof used the general theorem of van der Waerden. Roth [11], [12] later gave an analytic proof that $r_4(n) = o(n)$ which did not make use of van der Waerden's theorem (in fact, he proved a much more general theorem) and his method probably gives

$$r_4(n) < \frac{n}{\log_l n}$$

where l is a large fixed integer and $\log_l$ denotes the l-fold iterated logarithm.

In this paper we now prove the general conjecture of Erdös and Turán:

$$c_k = 0 \quad \text{for all } k.$$

Unfortunately, we again have to use the theorem of van der Waerden and so, cannot fulfill the original desiderata of Erdös and Turán. I hope that the proof can be modified by Roth's method so that a weak though not entirely ridiculous upper bound for $r_k(n)$ will be obtained. However, $r_k(n) < \pi(n)$ still seems hopeless at present.

T. Gallai gave the following generalization of van der Waerden's theorem: *Let S be a finite subset of L_k, the set of integer lattice points of k-dimensional space. If L_k is arbitrarily partitioned into two classes, then at least one class contains a set which is similar* (2) *to S by translation.*

Erdös now conjectured: *Let S be a finite subset of L_k. Suppose $R \subseteq L_k$ such that for some $\varepsilon > 0$, R contains at least εn^k lattice points all of whose coordinates are between 0 and n (where $n > n_0(\varepsilon)$). Then R contains a set which is similar to S by translation.*

Unfortunately I cannot prove this conjecture. M. Ajtai and I can prove it in the special case that S is a square. Our proof uses $r_k(n) = o(n)$.

Before presenting the proof that $r_k(n) = o(n)$, which, although using only elementary combinatorial arguments, is rather long and complicated, we first make a few remarks regarding notation. Unless otherwise specified, lower case Greek letters $\alpha, \beta, \ldots$, will denote real numbers strictly between 0 and 1, lower case Roman letters $a, b, \ldots$, will denote non-negative integers and upper case Roman letters $A, B, \ldots$, will usually denote sets. Thus, the union $\bigcup_{i<m} X_i$ will denote $\bigcup_{0\leqslant i<m} X_i$. As usual, $|X|$ denotes the cardinality of X and $[x]$ denotes the greatest integer not exceeding x. Finally, $[a, b)$ will denote the set of integers $\{x:\ a \leqslant x < b\}$.

2. A lemma on bipartite graphs. The first result we prove concerns bipartite graphs. It says essentially that any large bipartite graph can be decomposed into nearly regular bipartite subgraphs. First, we need some notation.

(i) A and B will denote disjoint sets with $|A| = m$, $|B| = n$.

(ii) For $X \subseteq A$, $Y \subseteq B$,

$$[X, Y] = \{\{x, y\}:\ x\in X, y\in Y\}.$$

(iii) The letter I will denote a fixed subset of $[A, B]$.

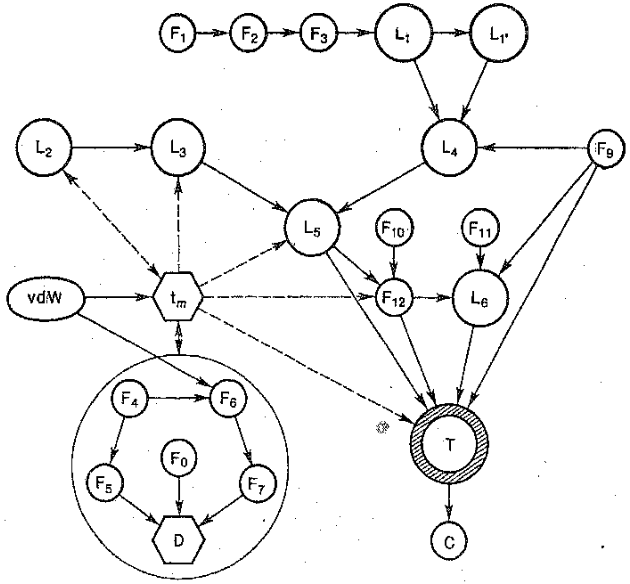
(iv) $k_I(X, Y) = k(X, Y) = |[X, Y] \cap I|$.

(v) For $u \in A\cup B$,

$$k_I(u) = k(u) = \{v \in A\cup B:\ \{u, v\} \in I\}.$$

(vi) $\beta(X, Y) = k(X, Y)|X|^{-1}|Y|^{-1}$.

(2) The corresponding questions for congruence instead of similarity are investigated in the papers of Erdös, et. al. [5], [6].



The diagram represents an approximate flow chart for the accompanying proof of Szemerédi's theorem. The various symbols have the following meanings: F_k = Fact k, L_k = Lemma k, T = Theorem, C = Corollary, D = Definitions of B, S, P, α, β, etc., t_m = Definition of t_m, vdW = van der Waerden's theorem, F_0 = "If $f: \mathbf{R}^+ \to \mathbf{R}^+$ is subadditive then $\lim_{n\to\infty} \frac{f(n)}{n}$ exists".

We require two preliminary facts.

FACT 1. *For all $\varepsilon_1, \varepsilon_2, \delta$ and for all $I \neq \emptyset$, there exist $\bar{X} \subseteq A$, $\bar{Y} \subseteq B$ and $r > 0$ such that:*

(a_1) $r \leqslant 1/\delta$;

(b_1) $|\bar{X}| > \varepsilon_1^r |A|$, $|\bar{Y}| > \varepsilon_2^r |B|$;

(c_1) *For all $S \subseteq \bar{X}$, $T \subseteq \bar{Y}$ with $|S| > \varepsilon_1 |\bar{X}|$, $|T| > \varepsilon_2 |\bar{Y}|$, we have*

$$\beta(S, T) > \beta(\bar{X}, \bar{Y}) - \delta.$$

Proof. By the hypothesis that $0 < \delta < 1$, there exists a positive integer r satisfying (a_1). Also, by taking $\bar{X} = A$, $\bar{Y} = B$, we can find $\bar{X}, \bar{Y}$ satisfying (b_1). *Suppose* $\bar{X} \subseteq A$, $\bar{Y} \subseteq B$ do not satisfy (c_1). Then there must exist $X' \subseteq X$, $Y' \subseteq Y$ with $|X'| > \varepsilon_1 |X|$, $|Y'| > \varepsilon_2 |Y|$ such that $\beta(X', Y') \leqslant \beta(X, Y) - \delta$. We now define sequences X_t and Y_t by induction on t as follows:

$$X_0 = A, \quad Y_0 = B.$$

For $t \geqslant 0$, $X_{t+1} = X_t'$, $Y_{t+1} = Y_t'$, provided $X_t \subseteq A$, $Y_t \subseteq B$, and they do not satisfy (c_1).

Thus,

$$|X_t| > \varepsilon_1 |X_{t-1}| > \varepsilon_1^2 |X_{t-2}| > \ldots > \varepsilon_1^t |X_0| = \varepsilon_1^t |A|,$$

$$|Y_t| > \varepsilon_2 |Y_{t-1}| > \varepsilon_2^2 |Y_{t-2}| > \ldots > \varepsilon_2^t |Y_0| = \varepsilon_2^t |B|.$$

But

$$\begin{aligned} 0 \leqslant \beta(X_t, Y_t) &\leqslant \beta(X_{t-1}, Y_{t-1}) - \delta \\ &\leqslant \beta(X_{t-2}, Y_{t-2}) - 2\delta \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\leqslant \beta(X_0, Y_0) - t\delta = \beta(A, B) - t\delta \leqslant 1 - t\delta. \end{aligned}$$

Thus, $t \leqslant 1/\delta$. However, if for all $s \leqslant 1/\delta$, X_s and Y_s are defined and do not satisfy (c_1), then X_{s+1} and Y_{s+1} are also defined and, since $[1/\delta] + 1 > 1/\delta$, this would contradict $t \leqslant 1/\delta$.

Hence, for some index $r \leqslant 1/\delta$, X_r and Y_r are defined and satisfy (c_1). Thus, letting $\bar{X} = X_r$, $\bar{Y} = Y_r$, we have

$$|\bar{X}| = |X_r| > \varepsilon_1^r |A|, \qquad |\bar{Y}| = |Y_r| > \varepsilon_2^r |B|$$

and so (a_1), (b_1) and (c_1) hold. This proves Fact 1. ∎

FACT 2. *For all ε_1, $0 < \varepsilon_1 < 1/2$, ε_2, δ, there exist integers M, N such that for all I with $|A| = m > M$, $|B| = n > N$, there exist $X \subseteq A$, $Y \subseteq B$ and $r > 0$ such that:*

(a_2) $r \leqslant 1/\delta$;

(b_2) $|X| > \frac{1}{2}\varepsilon_1^r |A|$, $|Y| > \varepsilon_2^r |B|$;

(c_2) *For all $S \subseteq X$, $T \subseteq Y$ with $|S| > 2\varepsilon_1 |X|$, $|T| > \varepsilon_2 |Y|$, we have*

$$\beta(S, T) > \beta(X, Y) - \delta;$$

(d_2) *For all $x \in X$,*

$$|k(x) \cap Y| \leqslant \big(\beta(X, Y) + \delta\big) |Y|.$$

Proof. By Fact 1, there exist $\bar{X} \subseteq A$, $\bar{Y} \subseteq B$ and $r > 0$ such that (a_1), (b_1) and (c_1) hold. Let $\beta = \beta(\bar{X}, \bar{Y})$ and

$$Z = \{x \in \bar{X}: |k(x) \cap \bar{Y}| > (\beta + \delta) |\bar{Y}|\}.$$

We *claim* $|Z| \leqslant \frac{1}{2}|\bar{X}|$. For suppose not, i.e., suppose $|Z| > \frac{1}{2}|\bar{X}|$. Then there exists $Z' \subseteq Z$ such that

$$\tfrac{1}{2}|\bar{X}| < |Z'| \leqslant \tfrac{1}{2}|\bar{X}| + 1.$$

Now,

$$k(\bar{X}, \bar{Y}) = k(Z', \bar{Y}) + k(\bar{X} - Z', \bar{Y})$$

so that

$$\frac{k(\bar{X},\bar{Y})}{|\bar{X}|\,|\bar{Y}|} = \frac{|Z'|}{|\bar{X}|}\frac{k(Z',\bar{Y})}{|Z'|\,|\bar{Y}|} + \frac{|\bar{X}-Z'|}{|\bar{X}|}\frac{k(\bar{X}-Z',\bar{Y})}{|\bar{X}-Z'|\,|\bar{Y}|},$$

i.e.,

$$\beta = \alpha\beta(Z',\bar{Y}) + (1-\alpha)\beta(\bar{X}-Z',\bar{Y})$$

where $\alpha = \frac{|Z'|}{|\bar{X}|}$. By the definition of Z, $k(Z',\bar{Y}) > (\beta+\delta)|Z'|\,|\bar{Y}|$ and so

$$\beta(Z',\bar{Y}) > \beta+\delta.$$

Also, since $|Z'| \leqslant \frac{1}{2}|\bar{X}|+1$, then

$$|\bar{X}-Z'| \geqslant \tfrac{1}{2}|\bar{X}|-1 > \varepsilon_1|\bar{X}|$$

provided $(\frac{1}{2}-\varepsilon_1)|\bar{X}| > 1$, i.e., provided $|\bar{X}| > (\frac{1}{2}-\varepsilon_1)^{-1}$. Since

$$|\bar{X}| > \varepsilon_1^r|A| \geqslant \varepsilon_1^{1/\delta}|A| > \varepsilon_1^{1/\delta}M,$$

then for $M > \left(\varepsilon_1^{1/\delta}(\frac{1}{2}-\varepsilon_1)\right)^{-1}$ we have $|\bar{X}-Z'| > \varepsilon_1|\bar{X}|$. Thus, by (c_1)

$$\beta(\bar{X}-Z',\bar{Y}) > \beta-\delta.$$

Hence,

$$\beta \geqslant \alpha(\beta+\delta)+(1-\alpha)(\beta-\delta) = \beta+(2\alpha-1)\delta.$$

But

$$\alpha = \frac{|Z'|}{|\bar{X}|} > \frac{1}{2}$$

so that $\beta > \beta$ which is *impossible*. Hence, we must have $|Z| \leqslant \frac{1}{2}|\bar{X}|$.

Let $X = \bar{X}-Z$, $Y = \bar{Y}$. Thus,

$$|X| = |\bar{X}-Z| \geqslant \tfrac{1}{2}|\bar{X}| > \tfrac{1}{2}\varepsilon_1^r|A|.$$

By the definition of Z,

$$x \in X \Rightarrow |k(x)\cap\bar{Y}| \leqslant (\beta+\delta)|\bar{Y}|$$

so that (d_2) is satisfied. Finally, to see that (c_2) holds, note that $S \subseteq X$ with $|S| > 2\varepsilon_1|X|$ implies $S \subseteq \bar{X}$ with $|S| > 2\varepsilon_1|X| \geqslant \varepsilon_1|\bar{X}|$ so that (c_1) implies (c_2). ■

We are now ready for the first lemma.

Lemma 1. *For all $\varepsilon_1, \varepsilon_2, \delta, \varrho, \sigma$, there exist m_0, n_0, M, N, such that for all I with $|A| = m > M$, $|B| = n > N$, there exist disjoint $C_i \subseteq A$, $i < m_0$, and, for each $i < m_0$, disjoint $C_{i,j} \subseteq B$, $j < n_0$, such that:*

(a) $|A - \bigcup_{i<m_0} C_i| < \varrho m$, $|B - \bigcup_{j<n_0} C_{i,j}| < \sigma n$ *for any* $i < m_0$;

(b) *For all $i < m_0$, $j < n_0$, $S \subseteq C_i$, $T \subseteq C_{i,j}$, with $|S| > \varepsilon_1|C_i|$, $|T| > \varepsilon_2|C_{i,j}|$, we have*

$$\beta(S,T) \geqslant \beta(C_i, C_{i,j}) - \delta;$$

(c) *For all $i < m_0$, $j < n_0$ and $x \in C_i$,*

$$|k(x)\cap C_{i,j}| \leqslant (\beta(C_i, C_{i,j})+\delta)|C_{i,j}|.$$

Proof. Let $r = 2[1/\delta]+1$. Choose n_0 so that $(1-\varepsilon_2^r)^{n_0} < \sigma$. Define the sequence $\varepsilon(t)$ for $0 \leqslant t \leqslant n_0+1$ as follows:

$$\varepsilon(0) = \varepsilon_1, \quad \varepsilon(t+1) = \varepsilon_1\prod_{i=0}^{t}\left(\frac{\varepsilon(i)}{4}\right)^r, \quad 0 \leqslant t \leqslant n_0.$$

Choose m_0 so that

$$\left(1-\varepsilon(n_0+1)\right)^{m_0} < \varrho.$$

We claim that the following fact holds.

Fact 3. *There exist M, N such that if $m > M$, $n > N$ and $A' \subseteq A$ with $|A'| \geqslant \varrho m$ then there exist $C \subseteq A'$, $\bar{C}_j \subseteq B$ for $j < n_0$ such that $|C| > \varepsilon(n_0+1)|A'|$ and C and the $\bar{C}_j$, $j < n_0$, satisfy the requirements of Lemma 1 if we choose $C_i = C$ and $C_{i,j} = \bar{C}_j$, $j < n_0$.*

Proof of Fact 3. Define the sequences $\bar{C}_j \subseteq B$, $Z_j \subseteq A'$ for $-1 \leqslant j \leqslant n_0$, by induction as follows:

$$Z_{-1} = A', \quad \bar{C}_{-1} = \emptyset.$$

Assume for some $j \leqslant n_0$ that Z_ν and $\bar{C}_\nu$ have been defined for $-1 \leqslant \nu < j$. There are two possibilities:

(i) If $|B - \bigcup_{\nu<j}\bar{C}_\nu| < \sigma n$, set $Z_j = Z_{j-1}$, $\bar{C}_j = \emptyset$;

(ii) Suppose $|B - \bigcup_{\nu<j}\bar{C}_\nu| \geqslant \sigma n$. Apply Fact 2 with $\varepsilon_1 = \frac{1}{2}\varepsilon(n_0-j)$, $\varepsilon_2 = \varepsilon_2$, $\delta = \delta/2$, $A = Z_{j-1}$ and $B = B - \bigcup_{\nu<j}\bar{C}_\nu$. It will be clear that $|Z_{j-1}|$ can be made arbitrarily large by choosing M sufficiently large.

By the hypothesis in (ii), $|B - \bigcup_{\nu<j}\bar{C}_\nu|$ is arbitrarily large if N is sufficiently large. Thus, the hypotheses of Fact 2 are satisfied. Therefore, there exist $X \subseteq Z_{j-1}$, $Y \subseteq B - \bigcup_{\nu<j}\bar{C}_\nu$ and $\bar{r} > 0$ which satisfy the conclusions of Fact 2. Set $Z_j = X$, $\bar{C}_j = Y$. By (a_2),

$$\bar{r} \leqslant 2/\delta < [2/\delta]+1 = r.$$

Also, we have

$$(b_3) \qquad |Z_j| > \frac{1}{2}\left(\frac{1}{2}\varepsilon(n_0-j)\right)^r|Z_{j-1}| > \left(\frac{\varepsilon(n_0-j)}{4}\right)^r|Z_{j-1}|,$$

$$|\bar{C}_j| > \varepsilon_2^r|B - \bigcup_{\nu<j}\bar{C}_\nu| > \varepsilon_2^r|B - \bigcup_{\nu<j}\bar{C}_\nu|;$$

(c_3) For all $S \subseteq Z_j$, $T \subseteq \bar{C}_j$ with $|S| > \varepsilon(n_0-j)|Z_j|$, $|T| > \varepsilon_2|\bar{C}_j|$, we have

$$\beta(S,T) > \beta(Z_j,\bar{C}_j) - \delta/2;$$

(d_3) For all $x \in Z_j$,

$$|k(x)\cap \bar{C}_j| \leqslant \big(\beta(Z_j, \bar{C}_j)+\delta/2\big)|\bar{C}_j|.$$

Once we are forced to use (i) in defining Z_j and $\bar{C}_j$ then we must always use this case until Z_{n_0} and C_{n_0} are defined. Let j_0 be the largest index $j < n_0$ for which (ii) is used to define Z_j and $\bar{C}_j$. By (b_3) we have:

$$|\bar{C}_0| > \varepsilon_2^r|B|,$$

$$|B-\bar{C}_0| < (1-\varepsilon_2^r)|B|,$$

$$|\bar{C}_1| > \varepsilon_2^r|B-\bar{C}_0|,$$

$$|B-\bar{C}_0-\bar{C}_1| < (1-\varepsilon_2^r)|B-\bar{C}_0| < (1-\varepsilon_2^r)^2|B| \quad \text{since } \bar{C}_0\cap\bar{C}_1 = \emptyset,$$

$$\dots\dots\dots\dots\dots\dots$$

$$|B-\bigcup_{\nu<j_0}\bar{C}_\nu| < (1-\varepsilon_2^r)^{j_0}|B|,$$

$$|\bar{C}_{j_0}| > \varepsilon_2^r|B-\bigcup_{\nu<j_0}\bar{C}_\nu|,$$

$$|B-\bigcup_{\nu<j_0+1}\bar{C}_\nu| < (1-\varepsilon_2^r)^{j_0+1}|B|.$$

Therefore,

$$|B-\bigcup_{\nu<n_0}\bar{C}_\nu| \leqslant |B-\bigcup_{\nu\leqslant j_0}\bar{C}_\nu| < (1-\varepsilon_2^r)^{j_0+1}|B|.$$

If $j_0 = n_0-1$ then

$$|B-\bigcup_{\nu<n_0}\bar{C}_\nu| < (1-\varepsilon_2^r)^{n_0}|B| < \sigma n.$$

If $j_0 < n_0-1$ then (i) was used so that

$$|B-\bigcup_{\nu<n_0}\bar{C}_\nu| < \sigma n.$$

Hence, in either case we have

$$|B-\bigcup_{\nu<n_0}\bar{C}_\nu| < \sigma n.$$

Also, it follows that

$$|Z_{n_0}| > \left(\frac{\varepsilon(0)}{4}\right)^r|Z_{n_0-1}|$$

$$> \left(\frac{\varepsilon(0)}{4}\right)^r\left(\frac{\varepsilon(1)}{4}\right)^r|Z_{n_0-2}|$$

$$\dots\dots\dots\dots\dots\dots$$

$$> \prod_{\nu=j+1}^{n_0}\left(\frac{\varepsilon(n_0-\nu)}{4}\right)^r|Z_j| = \prod_{i=0}^{n_0-j-1}\left(\frac{\varepsilon(i)}{4}\right)^r|Z_j|$$

by using either (b_3) or $Z_j = Z_{j-1}$ when (i) is used. Therefore, by the definition of $\varepsilon(n_0-j)$,

$$\varepsilon_1|Z_{n_0}| > \varepsilon(n_0-j)|Z_j|, \quad -1\leqslant j\leqslant n_0.$$

Take $C = Z_{n_0}$. Then

$$|C| = |Z_{n_0}| > \varepsilon(n_0+1)|A'|.$$

Now, suppose we have $S\subseteq C$, $T\subseteq \bar{C}_j$ for some j, $0\leqslant j< n_0$, with $|S| > \varepsilon_1|C|$, $|T| > \varepsilon_2|\bar{C}_j|$. Then

$$|S| > \varepsilon_1|Z_{n_0}| > \varepsilon(n_0-j)|Z_j|.$$

Therefore, by (c_3)

$$\beta(S, T) > \beta(Z_j, \bar{C}_j)-\delta/2.$$

By (d_3), since $C = Z_{n_0}\subseteq Z_j$, we have

$$\beta(C, \bar{C}_j) \leqslant \beta(Z_j, \bar{C}_j)+\delta/2.$$

Thus,

$$\beta(S, T) > \beta(C, \bar{C}_j)-\delta$$

which is just (b) of Lemma 1 with $C_i = C$ and $C_{i,j} = \bar{C}_j$. Since

$$|Z_{n_0}| > \varepsilon(n_0-j)|Z_j|$$

then by (c_3) with $S = Z_{n_0} = C$, $T = \bar{C}_j$, we have

$$\beta(C, \bar{C}_j) > \beta(Z_j, \bar{C}_j)-\delta/2.$$

Thus, by (d_3), for all $x\in C = Z_{n_0}\subseteq Z_j$,

$$|k(x)\cap\bar{C}_j| \leqslant \big(\beta(Z_j, \bar{C}_j)+\delta/2\big)|\bar{C}_j| < \big(\beta(C, \bar{C}_j)+\delta\big)|\bar{C}_j|.$$

This is just (c) of Lemma 1 and the proof of Fact 3 is completed. ■

We now apply Fact 3 recursively to prove Lemma 1.

We begin by setting $A_0 = A$ so that $|A_0| = m > \varrho m$. By Fact 3, there exists $C_0\subseteq A_0$ with $|C_0| > \varepsilon(n_0+1)|A_0|$ such that the conclusions of Fact 3 hold. Let $A_1 = A_0-C_0$. Then

$$|A_1| < \big(1-\varepsilon(n_0+1)\big)|A_0|.$$

Now, if $|A_1| < \varrho m$, then we stop; otherwise, $|A_1|\geqslant \varrho m$ and we can continue, obtaining $C_1\subseteq A_1$ with $|C_1| > \varepsilon(n_0+1)|A_1|$ so that the conclusions of Fact 3 hold. Let $A_2 = A_1-C_1$. Then

$$|A_2| < \big(1-\varepsilon(n_0+1)\big)|A_1| < \big(1-\varepsilon(n_0+1)\big)^2|A_0|, \text{ etc.}$$

By the time we get A_{m_0}, we would have $A_{m_0} = A_{m_0-1}-C_{m_0-1}$ and

$$|A_{m_0}| < \big(1-\varepsilon(n_0+1)\big)^{m_0}|A_0| < \varrho m.$$

Thus, we have to define C_j only for $j < m_0$. This completes the proof of Lemma 1. ■

By applying Lemma 1 to $\bar{I} = [A, B]-I$ we obtain

LEMMA 1'. *For all* $\varepsilon_1, \varepsilon_2, \delta, \varrho, \sigma$, *there exist* m_0, n_0, M, N *such that for all I with* $|A| = m > M$, $|B| = n > N$, *there exist disjoint* $\bar{C}_i \subseteq A$, $i < m_0$, *and, for each* $i < m_0$, *disjoint* $\bar{C}_{i,j} \subseteq B$, $j < n_0$, *such that*:

($\bar{\text{a}}$) $|A - \bigcup_{i<m_0} \bar{C}_i| < \varrho m$, $|B - \bigcup_{j<n_0} \bar{C}_{i,j}| < \sigma n$ *for any* $i < m_0$;

($\bar{\text{b}}$) *For all* $i < m_0$, $j < n_0$, $S \subseteq \bar{C}_i$, $T \subseteq \bar{C}_{i,j}$ *with* $|S| > \varepsilon_1 |\bar{C}_i|$, $|T| > \varepsilon_2 |\bar{C}_{i,j}|$, *we have*

$$\beta(S, T) \leqslant \beta(\bar{C}_i, \bar{C}_{i,j}) + \delta;$$

($\bar{\text{c}}$) *For all* $i < m_0$, $j < n_0$ *and* $x \in \bar{C}_i$,

$$|k(x) \cap \bar{C}_{i,j}| \geqslant (\beta(\bar{C}_i, \bar{C}_{i,j}) - \delta) |\bar{C}_{i,j}|.$$

3. Configurations. We next define certain subsets of integers, called *configurations*, which will be fundamental for remainder of the paper. For each choice of positive integers $l_1, \ldots, l_m$ (where possibly $m = 0$), a set $B(l_1, \ldots, l_m)$ of configurations is defined as follows: $B(\emptyset) = \{\{n\}: n \in N\}$ where N denotes the set of natural numbers. For $m \geqslant 1$, $B(l_1, \ldots, l_m) = \{X \subseteq N: X = \bigcup_{i<l_m} X_i$ where for some $Y \in B(l_1, \ldots, l_{m-1})$ and some $d > 0$ we have $X_i = Y + di$ for $0 \leqslant i < l_m$ and $X_0 < X_1 < \ldots < X_{l_m - 1}\}$. Of course, $Y + di$ denotes $\{y + di: y \in Y\}$ and $X < X'$ denotes $x < x'$ for $x \in X$, $x' \in X'$. If $X \in B(l_1, \ldots, l_m)$ then the meaning of X_i is explained by the definition just given. For example, the elements of $B(l_1)$ are exactly the arithmetic progressions of positive integers of length l_1. The elements X of $B(l_1, l_2)$ are just the sets of l_2 equally spaced nonoverlapping arithmetic progressions X_i of length l_1, etc. Thus, in this case, X can be thought of as a "progression of progressions". The elements of $B(l_1, \ldots, l_m)$ are called *configurations of order* m. In general, if $X \in B(l_1, \ldots, l_m)$ then $|X| = l_1 \ldots l_m$. We say that two configurations $X, Y \in B(l_1, \ldots, l_m)$ are *congruent* if for some d (possibly negative), $X = Y + d$. It is clear that for $X \in B(l_1, \ldots, l_m)$, the $X_i \subseteq X$, $i < l_m$, are congruent configurations which belong to $B(l_1, \ldots, l_{m-1})$.

We assume now that we are given an arbitrary but fixed set R of positive upper density, i.e., such that

$$\overline{\lim_{n\to\infty}} |R \cap \{1, \ldots, n\}| > 0.$$

Our eventual goal is to show that R contains arbitrarily long arithmetic progressions. Our immediate task, however, will be to define a certain sequence $t_1, t_2, \ldots, t_m, \ldots$, certain sets $S(t_1, \ldots, t_m) \subseteq B(t_1, \ldots, t_m)$ and $P(t_1, \ldots, t_m) \subseteq B(t_1, \ldots, t_m)$, as well as a number of other auxiliary sequences of integers and functions. The elements of $S(t_1, \ldots, t_m)$ and $P(t_1, \ldots, t_m)$ will be called *saturated* and *perfect* configurations of order m, respectively.

We set

$$S(\emptyset) = B(\emptyset) = \{\{n\}: n \in N\}, \quad P(\emptyset) = \{\{n\}: n \in R\}.$$

If $X \in B(t_1, \ldots, t_{m-1}, l)$ and $S(t_1, \ldots, t_{m-1})$ and $P(t_1, \ldots, t_{m-1})$ have already been defined then we define

$$s^m(X) = |\{i < l: X_i \in S(t_1, \ldots, t_{m-1})\}|,$$
$$p^m(X) = |\{i < l: X_i \in P(t_1, \ldots, t_{m-1})\}|.$$

Set

$$f_1(l) = \max\{|X \cap R: X \in B(l)|\},$$
$$= \max\{p^1(X): X \in B(l)\},$$

$$a_1 = \overline{\lim_{l\to\infty}} \frac{f_1(l)}{l} = \lim_{l\to\infty} \frac{f_1(l)}{l} > 0.$$

An easy argument, based on the obvious subadditivity of f_1, shows that the limit exists. The assumption that R has positive upper density forces the limit a_1 to be positive. Furthermore put

$$\varepsilon_1(l) = \left| a_1 - \frac{f_1(l)}{l} \right|.$$

Next, we assume that we are also given an arbitrary but fixed integer K (this will play a role throughout the rest of the proof). Define t_1 in such a way that t_1 is sufficiently large depending on K and so that $\varepsilon_1(t_1)$ is sufficiently small. We shall explain precisely what is meant by this later; it plays no role in the structure of the current definitions.

So far, we have defined $S(\emptyset)$, $P(\emptyset)$, $f_1(l)$, a_1, $\varepsilon_1(l)$, and (in principle) t_1. It would now be reasonable to assume that $S(t_1, \ldots, t_{m-1})$, $P(t_1, \ldots, t_{m-1})$, $f_m(l)$, a_m, $\varepsilon_m(l)$, and t_m have been defined, and then to define the corresponding quantities with the indices m replaced by $m+1$. However, the case $m = 1$ does not reflect the situation in its full generality and so it will be useful to describe the case $m = 2$ as well.

Define

$$S(t_1) = \{X \in B(t_1): p^1(X) > (a_1 - \varepsilon_1(t_1)^{1/4}) t_1\},$$

$$g_2(l) = \max\{s^2(X): X \in B(t_1, l)\}, \quad \beta_2 = \overline{\lim_{l\to\infty}} \frac{g_2(l)}{l}, \quad \mu_2(l) = \left| \beta_2 - \frac{g_2(l)}{l} \right|.$$

Note that the corresponding quantities

$$g_1(l) = \max\{s^1(x): X \in B(l)\} = l,$$

$$\beta_1 = \overline{\lim_{l\to\infty}} \frac{g_1(l)}{l} = 1 \quad \text{and} \quad \mu_1(l) = \left| \beta_1 - \frac{g_1(l)}{l} \right| = 0$$

and are thus degenerate. We also point out that it will in fact be shown later that $\mu_2(l) \to 0$ and β_2 is very close to 1.

Consider now two configurations $X, Y \in B(t_1)$. We shall say that X and Y are *R-equivalent* if for any two elements $x \in X$, $y \in Y$ in corresponding positions we have $x \in R$ if and only if $y \in R$. Now, let $X = \bigcup_{i<l} X_i \in B(t_1, l)$. Then the R-equivalence relation partitions the set of $X_i \in S(t_1)$ equivalence classes, in fact, into at most 2^{t_1} equivalence classes. Denote by $\bar{p}^2(X)$ the cardinality of the largest equivalence class. Define

$$\bar{f}_2(l) = \max\{\bar{p}^2(X):\ X \in B(t_1, l) \text{ and } s^2(X) \geqslant (\beta_2 - \sqrt{\mu_2(l)})\,l\},$$

$$a_2 = \overline{\lim_{l\to\infty}} \frac{\bar{f}_2(l)}{l}.$$

Thus, there is a sequence l_n such that

$$\lim_{n\to\infty} \frac{\bar{f}_2(l_n)}{l_n} = a_2.$$

Therefore, there exist $X^{(n)} \in B(t_1, l_n)$ for which

$$s^2(X^{(n)}) \geqslant (\beta_2 - \sqrt{\mu_2(l_n)})\,l_n \quad \text{and} \quad \lim_{n\to\infty} \frac{\bar{p}^2(X^{(n)})}{l_n} = a_2.$$

Now, for infinitely many n, the *same* R-equivalence class occurs in the definition of $\bar{p}^2(X^{(n)})$. Let us choose such an R-equivalence class and denote it by $P(t_1)$. Clearly $P(t_1) \subseteq S(t_1)$. Set

$$f_2(l) = \max\{p^2(X):\ X \in B(t_1, l) \text{ and } s^2(X) \geqslant (\beta_2 - \sqrt{\mu_2(l)})\,l\},$$

$$\varepsilon_2(l) = \left| a_2 - \frac{f_2(l)}{l} \right|.$$

As we shall prove later, we have

$$\lim_{l\to\infty} \frac{g_2(l)}{l} = \beta_2, \quad \lim_{l\to\infty} \frac{f_2(l)}{l} = a_2.$$

We now choose t_2 in such a way that it is sufficiently large depending on a_2, β_2, K, and t_1, and so that $\varepsilon_2(t_2)$ and $\mu_2(t_2)$ are sufficiently small.

Assume now that for some $m \geqslant 1$ the quantities

$$P(t_1, \ldots, t_{m-1}) \subseteq S(t_1, \ldots, t_{m-1}) \subseteq B(t_1, \ldots, t_{m-1}),$$

$g_m(l), f_m(l), a_m, \beta_m, \varepsilon_m(l), \mu_m(l)$, and t_m have been defined in such a way that we have

$$g_m(l) = \max\{s^m(X):\ X \in B(t_1, \ldots, t_{m-1}, l)\},$$

$$\beta_m = \lim_{l\to\infty} \frac{g_m(l)}{l}, \quad \mu_m(l) = \left| \beta_m - \frac{g_m(l)}{l} \right|,$$

$$f_m(l) = \max\{p^m(X):\ X \in B(t_1, \ldots, t_{m-1}, l) \text{ and } s^m(X) \geqslant (\beta_m - \sqrt{\mu_m(l)})\,l\},$$

$$a_m = \lim_{l\to\infty} \frac{f_m(l)}{l}, \quad \varepsilon_m(l) = \left| a_m - \frac{f_m(l)}{l} \right|,$$

t_m is sufficiently large depending on a_m, β_m, K, and t_{m-1}, and $\varepsilon_m(t_m)$ and $\mu_m(t_m)$ are sufficiently small. We give the next step in the definition. Let

$$S(t_1, \ldots, t_m) = \{X \in B(t_1, \ldots, t_m):\ s^m(X) \geqslant (\beta_m - \sqrt{\mu_m(t_m)})\,t_m$$
$$\text{and } p^m(X) > (a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}})\,t_m\},$$

$$g_{m+1}(l) = \max\{s^{m+1}(X):\ X \in B(t_1, \ldots, t_m, l)\}.$$

Define

$$\beta_{m+1} = \overline{\lim_{l\to\infty}} \frac{g_{m+1}(l)}{l}.$$

FACT 4. *For all l and m,*

$$\beta_{m+1} \leqslant \frac{g_{m+1}(l)}{l}.$$

Proof. For $m = 0$ the assertion holds since $g_1(l) = l$ for all l. Suppose for some $m \geqslant 0$, the assertion holds for all values less than or equal to m. We prove it for $m+1$. Let $\varepsilon > 0$ and l be arbitrary. Then by the definition of β_{m+1}, there exists a large integer L such that

$$\frac{g_{m+1}(L)}{L} > \beta_{m+1} - \varepsilon.$$

Thus, there exists $Y \in B(t_1, \ldots, t_m, L)$ such that

$$s^{m+1}(Y) = g_{m+1}(L) > (\beta_{m+1} - \varepsilon)L.$$

Hence, if

$$Y = \bigcup_{i<L} Y_i, \quad Y_i \in B(t_1, \ldots, t_m),$$

denotes the canonical decomposition of Y into L subconfigurations of order m, then at least $g_{m+1}(L)$ of the Y_i belong to $S(t_1, \ldots, t_m)$. Write $L = l\nu + \bar{l}$ where $0 \leqslant \bar{l} < l$. Then at least $g_{m+1}(L) - \bar{l}$ of the configurations $Y_0, Y_1, \ldots, Y_{l\nu-1}$ belong to $S(t_1, \ldots, t_m)$. But of course all ν of the configurations $\bigcup_{j<l} Y_{li+j}$, $i < \nu$, belong to $B(t_1, \ldots, t_m, l)$. Thus, at least one of these elements of $B(t_1, \ldots, t_m, l)$ must have at least $\frac{1}{\nu}(g_{m+1}(L) - \bar{l})$

of its l Y_i's in $S(t_1, \ldots, t_m)$. This implies

$$g_{m+1}(l) \geqslant \frac{1}{\nu}\left(g_{m+1}(L) - \bar{l}\right) > \frac{1}{\nu}\left((\beta_{m+1} - \varepsilon)L - \bar{l}\right) \geqslant \frac{1}{\nu}\left((\beta_{m+1} - \varepsilon)l\nu - l\right)$$

so that

$$\frac{g_{m+1}(l)}{l} > \beta_{m+1} - \varepsilon - \frac{1}{\nu}.$$

Since $\varepsilon > 0$ was arbitrary and ν can be made arbitrarily large (by taking L to be sufficiently large) then we must have

$$\frac{g_{m+1}(l)}{l} \geqslant \beta_{m+1}$$

and Fact 4 is proved. ■

An immediate corollary of Fact 4 is

FACT 5.

$$\beta_{m+1} = \lim_{l\to\infty} \frac{g_{m+1}(l)}{l}.$$

We now continue with the definitions. Set

$$\mu_{m+1}(l) = \left|\beta_{m+1} - \frac{g_{m+1}(l)}{l}\right| = \frac{g_{m+1}(l)}{l} - \beta_{m+1} \text{ by Fact 4.}$$

As before, define an equivalence relation (called R-equivalence) on $B(t_1, \ldots, t_m)$ by calling $X, Y \in B(t_1, \ldots, t_m)$ *R-equivalent* if for any $x \in X$ and $y \in Y$ in corresponding positions, $x \in R$ if and only if $y \in R$.

FACT 6. *We may assume $\beta_m < 1$ for $m \geqslant 2$.*

Proof. Suppose for some $m \geqslant 2$ we have $\beta_m = 1$. One of the conditions on the choice of t_m will be that when $\beta_m > 0$, it is chosen so large that $\beta_m - \sqrt{\mu_m(t_m)} > 0$. (As we shall see, it is always the case that $\beta_m > 0$.) Thus, it follows by induction that if $X' \in S(t_1, \ldots, t_r)$, $1 \leqslant r \leqslant m$, then $|X' \cap R| > 0$. By Fact 4, we have $g_m(l) \geqslant l \cdot \beta_m = l$ and so

$$g_m(l) = l \quad \text{for all } l.$$

Thus, there exists

$$X = \bigcup_{i<l} X_i \in B(t_1, \ldots, t_{m-1}, l)$$

such that $X_i \in S(t_1, \ldots, t_{m-1})$ for all $i < l$. The R-equivalence relation partitions $S(t_1, \ldots, t_{m-1})$ into at most $2^{t_1 t_2 \cdots t_{m-1}}$ R-equivalence classes. This induces a partition of the integers $0, 1, \ldots, l-1$ into at most $2^{t_1 t_2 \cdots t_{m-1}}$ classes where i and j are in the same class if and only if X_i and X_j are

R-equivalent. By the finite form of van der Waerden's theorem, if l is sufficiently large, some class contains an arithmetic progression of length k, say, $a, a+d, \ldots, a+(k-1)d$. This implies that $X_a, X_{a+d}, \ldots, X_{a+(k-1)d}$ are all R-equivalent. Since $|X_a \cap R| > 0$ then by the definitions of X and R-equivalence, R contains an arithmetic progression of length k. Since k was arbitrary, then we are done. ■

Let $X = \bigcup_{i<l} X_i \in B(t_1, \ldots, t_m)$. The R-equivalence relation partitions the set of $X_i \in S(t_1, \ldots, t_m)$ into at most $2^{t_1 t_2 \cdots t_m}$ equivalence classes. Let $\bar{p}^{m+1}(X)$ denote the cardinality of the largest equivalence class. Define

$$\bar{f}_{m+1}(l) = \max\{\bar{p}^{m+1}(X) \colon X \in B(t_1, \ldots, t_m, l) \text{ and } s^{m+1}(X) \geqslant (\beta_{m+1} - \sqrt{\mu_{m+1}(l)})l\},$$

$$\alpha_{m+1} = \overline{\lim_{l\to\infty}} \frac{\bar{f}_{m+1}(l)}{l}.$$

Note that $\bar{p}^{m+1}(X) \geqslant s^m(X) 2^{-t_1 t_2 \cdots t_m}$ for every $X \in B(t_1, \ldots, t_m, l)$. Therefore, taking $Y \in B(t_1, \ldots, t_m, l)$ with $s^{m+1}(Y) = g_{m+1}(l)$, we have (by Fact 4)

$$s^{m+1}(Y) = g_{m+1}(l) \geqslant \beta_{m+1} l$$

so that

$$f_{m+1}(l) \geqslant \bar{p}^{m+1}(Y) \geqslant g_{m+1}(l) 2^{-t_1 t_2 \cdots t_m}$$

and

$$\alpha_{m+1} \geqslant \beta_{m+1} 2^{-t_1 t_2 \cdots t_m}.$$

As before, there must exist an R-equivalence class $P(t_1, \ldots, t_m)$ such that if we define

$$f_{m+1}(l) = \max\{p^{m+1}(X) \colon X \in B(t_1, \ldots, t_m, l) \text{ and } s^{m+1}(X) \geqslant (\beta_{m+1} - \sqrt{\mu_{m+1}(l)})l\}$$

then

$$\alpha_{m+1} = \overline{\lim_{l\to\infty}} \frac{f_{m+1}(l)}{l}.$$

FACT 7.

$$\alpha_{m+1} = \lim_{l\to\infty} \frac{f_{m+1}(l)}{l}.$$

Proof. Let $\varepsilon > 0$. Choose a large integer l. By the definition of α_{m+1}, there exists a very large integer L such that

$$\frac{f_{m+1}(L)}{L} > \alpha_{m+1} - \varepsilon.$$

Thus, there exists $X \in B(t_1, \ldots, t_m, L)$ such that

$$p^{m+1}(X) > (a_{m+1} - \varepsilon)L$$

and

$$s^{m+1}(X) \geqslant \left(\beta_{m+1} - \sqrt{\mu_{m+1}(L)}\right)L.$$

Write $L = l\nu + \bar{l}$ where $0 \leqslant \bar{l} < l$. Let

$$X = \bigcup_{i<L} X_i, \quad X_i \in B(t_1, \ldots, t_m),$$

be the canonical decomposition of X into its L subconfigurations of order m. For $0 \leqslant i < \nu$, let

$$Y_i = \bigcup_{j<l} X_{li+j} \in B(t_1, \ldots, t_m, l)$$

and let

$$X' = \bigcup_{j<\nu l} X_j = \bigcup_{i<\nu} Y_i.$$

Let

$$T = \left\{j < \nu:\ s^{m+1}(Y_j) < \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)l\right\}$$

and let $|T| = a\nu$. For all j, we have

$$s^{m+1}(Y_j) \leqslant \left(\beta_{m+1} + \mu_{m+1}(l)\right)l$$

by the definition of μ_{m+1}.

Thus

$$\sum_{j<\nu} s^{m+1}(Y_j) = \sum_{j\in T} s^{m+1}(Y_j) + \sum_{j\notin T} s^{m+1}(Y_j)$$
$$\leqslant \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)la\nu + \left(\beta_{m+1} + \mu_{m+1}(l)\right)l(1-a)\nu.$$

But

$$\sum_{j<\nu} s^{m+1}(Y_j) = s^{m+1}(X') > s^{m+1}(X) - l \geqslant \left(\beta_{m+1} - \sqrt{\mu_{m+1}(L)}\right)L - l.$$

Therefore,

$$\left(\beta_{m+1} - \sqrt{\mu_{m+1}(L)}\right)L - l \leqslant \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)la\nu + \left(\beta_{m+1} + \mu_{m+1}(l)\right)l(1-a)\nu.$$

Hence,

$$\beta_{m+1} - \sqrt{\mu_{m+1}(L)} - \frac{1}{\nu} \leqslant a\left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right) + (1-a)\left(\beta_{m+1} + \mu_{m+1}(l)\right)$$
$$= \beta_{m+1} + \mu_{m+1}(l) - a\left(\sqrt{\mu_{m+1}(l)} + \mu_{m+1}(l)\right),$$

and

(1) $$a\left(\sqrt{\mu_{m+1}(l)} + \mu_{m+1}(l)\right) \leqslant \mu_{m+1}(l) + \frac{1}{\nu} + \sqrt{\mu_{m+1}(L)}.$$

There are now, two possibilities.

(i) Suppose $\mu_{m+1}(l') = 0$ for all sufficiently large l'. Then by the definition of μ_{m+1},

$$\frac{g_{m+1}(l')}{l'} = \beta_{m+1}$$

for all sufficiently large l'. But since $g_{m+1}(l')$ is an integer between 0 and l' then we must have either $\beta_{m+1} = 0$ or $\beta_{m+1} = 1$. By Fact 6, we may rule out the second possibility and so conclude that $\beta_{m+1} = 0$.

(ii) Suppose $\mu_{m+1}(l') > 0$ for infinitely many l'. We may assume that l was then chosen so that $\mu_{m+1}(l) > 0$. Assuming that case (ii) holds, it follows from (1) that by choosing L sufficiently large, we have

$$a\left(\sqrt{\mu_{m+1}(l)} + \mu_{m+1}(l)\right) < 2\mu_{m+1}(l),$$

and so

$$a < \frac{2\mu_{m+1}(l)}{\sqrt{\mu_{m+1}(l)} + \mu_{m+1}(l)} < 2\sqrt{\mu_{m+1}(l)}.$$

Therefore, we have

(2) $$a \leqslant 2\sqrt{\mu_{m+1}(l)}$$

where we note that (2) holds in case (i) as well, since in this case, $a = 0$.

Suppose, now, that for all $j < \nu$, we have either

$$s^{m+1}(Y_j) < \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)l$$

or

$$p^{m+1}(Y_j) \leqslant (a_{m+1} - 2\varepsilon)l.$$

There are just $a\nu$ indices j such that

$$s^{m+1}(Y_j) < \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)l$$

by the definition of a. Therefore at least $(1-a)\nu$ indices j satisfy $p^{m+1}(Y_j) \leqslant (a_{m+1} - 2\varepsilon)l$. However, the *other* indices j, i.e., the $j \in T$, must each satisfy

$$p^{m+1}(Y_j) < \left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)l$$

since $P(t_1, \ldots, t_m) \subseteq S(t_1, \ldots, t_m)$.

Thus,

$$\sum_{j<\nu} p^{m+1}(Y_j) < a\nu\left(\beta_{m+1} - \sqrt{\mu_{m+1}(l)}\right)l + (1-a)\nu(a_{m+1} - 2\varepsilon)l.$$

On the other hand,

$$\sum_{j<\nu} p^{m+1}(Y_j) = p^{m+1}(X') \geqslant p^{m+1}(X) - l > (a_{m+1} - \varepsilon)L - l.$$

Therefore,

$$(\alpha_{m+1}-\varepsilon)L-l < a\nu\big(\beta_{m+1}-\sqrt{\mu_{m+1}(l)}\big)l+(1-a)\nu(\alpha_{m+1}-2\varepsilon)l$$

and

$$\alpha_{m+1}-\varepsilon-\frac{1}{\nu} < a\big(\beta_{m+1}-\sqrt{\mu_{m+1}(l)}\big)+(1-a)(\alpha_{m+1}-2\varepsilon). \tag{3}$$

The right-hand side of (3) is a convex combination of $\beta_{m+1}-\sqrt{\mu_{m+1}(l)}$ and $\alpha_{m+1}-2\varepsilon$. Since $a\leqslant 2\sqrt{\mu_{m+1}(l)}$ then the coefficient $1-a$ is arbitrarily close to 1 for l sufficiently large. Finally, since $\beta_{m+1}-\sqrt{\mu_{m+1}(l)}$ is bounded then by choosing l (and L) sufficiently large, (3) is *contradicted*.

Hence, there must exist an index $j<\nu$ such that

$$s^{m+1}(Y_j)\geqslant\big(\beta_{m+1}-\sqrt{\mu_{m+1}(l)}\big)l \quad\text{and}\quad p^{m+1}(Y_j)>(\alpha_{m+1}-2\varepsilon)l.$$

Therefore,

$$\frac{f_{m+1}(l)}{l}>\alpha_{m+1}-2\varepsilon$$

for all sufficiently large l. Since $\varepsilon>0$ was arbitrary then the proof of Fact 7 is completed. ∎

Finally, we choose t_{m+1} sufficiently large depending on α_{m+1}, β_{m+1}, K, and t_m and so that $\varepsilon_{m+1}(t_{m+1})$ and $\mu_{m+1}(t_{m+1})$ are sufficiently small (to be made precise later). This completes the inductive step in the definitions.

The following inequality will be essential in what follows.

Lemma 2. *For all* $m\geqslant 1$,

$$\beta_{m+1}\geqslant 1-2\Big(\sqrt{\mu_m(t_m)}+\sqrt{\sqrt{\varepsilon_m(t_m)}+\sqrt{\mu_m(t_m)}}\Big). \tag{4}$$

Proof. Denote the right-hand side of (4) by $\varphi(t_m)$. Note that if $\alpha_1=1$ then for any $\varepsilon>0$,

$$f_1(l)=\max\{|X\cap R|:\ X\in B(l)\}>(1-\varepsilon)l$$

for all sufficiently large l. This implies immediately that R contains arbitrarily long arithmetic progressions. Hence, we may assume $\alpha_1<1$. Of course, we have already noted that $\alpha_1>0$. Exactly as in the proof of Fact 7, it then follows that we cannot have $\varepsilon_1(l)=0$ for all sufficiently large l. Hence, we shall also specify (later) that t_1 is chosen so that $\varepsilon_1(t_1)>0$.

By the definition of ε_m and α_m, for all l there exists $X\in B(t_1,\ldots,t_{m-1},t_ml)$ such that

$$s^m(X)\geqslant\big(\beta_m-\sqrt{\mu_m(t_ml)}\big)t_ml \quad\text{and}\quad p^m(X)\geqslant\big(\alpha_m-\varepsilon_m(t_ml)\big)t_ml.$$

Write

$$X=\bigcup_{i<t_ml}X_i,\qquad X_i\in B(t_1,\ldots,t_{m-1})$$

in the usual way and for all $j<l$, set

$$Y_j=\bigcup_{i<t_m}X_{jt_m+i}\in B(t_1,\ldots,t_m).$$

Let us write X as Y when considered as an element of $B(t_1,\ldots,t_m,l)$. We shall show that

$$s^{m+1}(Y)\geqslant\varphi(t_m)l$$

for l sufficiently large. This will suffice to prove Lemma 2 since in this case we have

$$g_{m+1}(l)\geqslant s^{m+1}(Y)\geqslant\varphi(t_m)l,$$

$$\frac{g_{m+1}(l)}{l}\geqslant\varphi(t_m)\quad\text{for } l \text{ sufficiently large},$$

so that

$$\lim_{l\to\infty}\frac{g_{m+1}(l)}{l}=\beta_{m+1}\geqslant\varphi(t_m)$$

as required. The argument will be essentially the argument of Fact 7 used twice, once for s^m and once for p^m, where the second time cruder approximations will be employed.

Let

$$T_1=\{j<l:\ s^m(Y_j)<\big(\beta_m-\sqrt{\mu_m(t_m)}\big)t_m\},\qquad |T_1|=al.$$

For all $j<l$, the definition of μ_m implies

$$s^m(Y_j)\leqslant\big(\beta_m+\mu_m(t_m)\big)t_m.$$

Therefore,

$$\sum_{j<l}s^m(Y_j)\leqslant\big(\beta_m-\sqrt{\mu_m(t_m)}\big)t_mal+\big(\beta_m+\mu_m(t_m)\big)t_m(1-a)l.$$

But

$$\sum_{j<l}s^m(Y_j)=s^m(X)\geqslant\big(\beta_m-\sqrt{\mu_m(t_ml)}\big)t_ml.$$

Thus,

$$\big(\beta_m-\sqrt{\mu_m(t_m)}\big)t_ml\leqslant\big(\beta_m-\sqrt{\mu_m(t_m)}\big)t_mal+\big(\beta_m+\mu_m(t_m)\big)t_m(1-a)l,$$

$$\beta_m-\sqrt{\mu_m(t_ml)}\leqslant\beta_m-a\sqrt{\mu_m(t_m)}+(1-a)\mu_m(t_m),$$

and

$$a\big(\sqrt{\mu_m(t_m)}+\mu_m(t_m)\big)\leqslant\mu_m(t_m)+\sqrt{\mu_m(t_ml)}. \tag{5}$$

If $m = 1$ so that $\mu_m(l') = \mu_1(l') = 0$ for all l', then since $s^1(Y_j) = t_1$ for all $j < l$, we have $T_1 = \emptyset$, i.e., $a = 0$. Therefore the inequality

$$a \leqslant 2\sqrt{\mu_m(t_m)} \tag{6}$$

holds trivially.

Suppose $m > 1$. If $\mu_m(l') = 0$ for all sufficiently large l', then the argument in the proof of Fact 7 shows that we must have $\beta_m = 0$. The definition of T_1 now implies $a = 0$ so that again (6) holds. On the other hand, if $\mu_m(l') > 0$ for infinitely many l', then we shall specify that t_m is chosen so that $\mu_m(t_m) > 0$. In this case, for l sufficiently large, we have from (5)

$$a \leqslant \frac{2\mu_m(t_m)}{\sqrt{\mu_m(t_m)} + \mu_m(t_m)} < 2\sqrt{\mu_m(t_m)}.$$

Hence, in all cases, (6) holds.

Now, let

$$T_2 = \{j < l\colon\ s^m(Y_j) \geqslant \big(\beta_m - \sqrt{\mu_m(t_m)}\big)t_m$$
$$\text{and } p^m(Y_j) < \big(a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}\big)t_m\},$$
$$|T_2| = bl.$$

Of course, for $j \notin T_1$,

$$p^m(Y_j) \leqslant \big(a_m + \varepsilon_m(t_m)\big)t_m.$$

Therefore,

$$\begin{aligned}\sum_{j<l} p^m(Y_j) &= \sum_{j\in T_1} p^m(Y_j) + \sum_{j\in T_2} p^m(Y_j) + \sum_{j\notin T_1\cup T_2} p^m(Y_j)\\ &\leqslant alt_m + \big(a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}\big)t_m bl + \\ &\qquad + (1-a-b)l\big(a_m + \varepsilon_m(t_m)\big)t_m\\ &\leqslant 2t_m l\sqrt{\mu_m(t_m)} + \big(a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}\big)t_m bl + \\ &\qquad + (1-b)\big(a_m + \varepsilon_m(t_m)\big)t_m l\end{aligned}$$

for l sufficiently large. But

$$\sum_{j<l} p^m(Y_j) = p^m(X) \geqslant \big(a_m - \varepsilon_m(t_m l)\big)t_m l.$$

Thus,

$$a_m - \varepsilon_m(t_m l) \leqslant 2\sqrt{\mu_m(t_m)} + a_m + \varepsilon_m(t_m) - b\Big(\sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}} + \varepsilon_m(t_m)\Big).$$

Therefore,

$$b\Big(\sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}} + \varepsilon_m(t_m)\Big) \leqslant 2\sqrt{\mu_m(t_m)} + \varepsilon_m(t_m) + \varepsilon_m(t_m l). \tag{7}$$

If $m = 1$ then by the choice of t_1, $\varepsilon_m(t_m) > 0$, so that $\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)} > 0$. If $m > 1$ then by induction (for a suitably large choice of t_m) and Fact 6, $0 < \beta_m < 1$. This in turn implies that t_m can be chosen so that $\mu_m(t_m) > 0$. Hence, in all cases

$$\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)} > 0.$$

Thus, since for l sufficiently large,

$$\varepsilon_m(t_m l) \leqslant \varepsilon_m(t_m) \leqslant \sqrt{\varepsilon_m(t_m)},$$

then by (7) we have

$$b \leqslant 2\sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}.$$

But for $j \notin T_1 \cup T_2$,

$$Y_j \in S(t_1, \ldots, t_m).$$

Therefore,

$$s^{m+1}(Y) \geqslant (1-a-b)l \geqslant \varphi(t_m)l$$

and Lemma 2 is proved. ■

We need some further definitions. Define

$$C(t_1, \ldots, t_m, l) = \{X \in B(t_1, \ldots, t_m, l)\colon\ s^{m+1}(X) = l\}.$$

Thus, if

$$X = \bigcup_{i<l} X_i \in C(t_1, \ldots, t_m, l)$$

then for all $i < l$, $X_i \in S(t_1, \ldots, t_m)$.

Fact 8. *If $\beta_{m+1} > 1 - 1/l$ then $C(t_1, \ldots, t_m, l) \neq \emptyset$.*

Proof. By Fact 4, if $\beta_{m+1} > 1 - 1/l$ then

$$g_{m+1}(l) \geqslant l\beta_{m+1} > l(1-1/l) = l-1$$

so that $g_{m+1}(l) = l$. Thus, there exists

$$X \in B(t_1, \ldots, t_m, l) \quad \text{with} \quad s^{m+1}(X) = l.$$

Therefore,

$$X \in C(t_1, \ldots, t_m, l) \neq \emptyset. \ ■$$

Note that by Lemma 2 and the fact that $a_{m+1} \geqslant \beta_{m+1}\cdot 2^{-t_1\ldots t_m}$, $S(t_1, \ldots, t_m)$ and $P(t_1, \ldots, t_m)$ are nonempty for all m (provided t_m is suitable chosen).

For $X = \bigcup_{i<l} X_i \in B(t_1, \ldots, t_{m-1}, l)$ and $C \subseteq [0, l)$, define

$$s^m(X, C) = |\{i < l\colon\ i \in C \text{ and } X_i \in S(t_1, \ldots, t_{m-1})\}|,$$
$$p^m(X, C) = |\{i < l\colon\ i \in C \text{ and } X_i \in P(t_1, \ldots, t_{m-1})\}|.$$

LEMMA 3. *For all δ and τ, there exists l such that if*

$$X = \bigcup_{i<l} X_i \in C(t_1, \ldots, t_m, l), \quad C \subseteq [0, t_m) \quad \text{with} \quad |C| \geqslant \tau t_m,$$

and t_m is sufficiently large depending on l, then there exists an i such that

$$p^m(X_i, C) > (a_m - \delta)|C|$$

and there exists an i' such that

$$p^m(X_{i'}, C) < (a_m + \delta)|C|.$$

Proof. The proof will be similar to that of Lemma 2. By Fact 6 and Lemma 2, we know there must exist arbitrarily large l's for which $\mu_m(l) > 0$. Assume that we have chosen such an l. Write

$$X_i = \bigcup_{j<t_m} X_{i,j} \quad \text{for} \quad i < l$$

and put

$$Y_j = \bigcup_{i<l} X_{i,j} \quad \text{for} \quad j < t_m.$$

Thus, $Y_j \in B(t_1, \ldots, t_{m-1}, l)$. Since $S \in C(t_1, \ldots, t_m, l)$ then for all $i < l$, we have $X_i \in S(t_1, \ldots, t_m)$. Therefore,

$$s^m(X_i) \geqslant \big(\beta_m - \sqrt{\mu_m(t_m)}\big)t_m \quad \text{and} \quad p^m(X_i) > \Big(a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}\Big)t_m$$

for all $i < l$. Thus,

$$\sum_{i<l} s^m(X_i) = \sum_{j<t_m} s^m(Y_j) \geqslant \big(\beta_m - \sqrt{\mu_m(t_m)}\big)t_m l.$$

Let

$$T_1' = \big\{j < t_m:\ s^m(Y_j) < \big(\beta_m - \sqrt{\mu_m(l)}\big)l\big\}, \qquad |T_1'| = a' t_m.$$

By the definition of μ_m, we have

$$s^m(Y_j) \leqslant \big(\beta_m + \mu_m(l)\big)l$$

for all $j < t_m$. Thus,

$$\sum_{j<t_m} s^m(Y_j) = \sum_{j\in T_1'} s^m(Y_j) + \sum_{j\notin T_1'} s^m(Y_j)$$
$$\leqslant \big(\beta_m - \sqrt{\mu_m(l)}\big)t_m a' l + \big(\beta_m + \mu_m(l)\big)(1-a')t_m l.$$

Hence,

$$\big(\beta_m - \sqrt{\mu_m(t_m)}\big)t_m l \leqslant \big(\beta_m - \sqrt{\mu_m(l)}\big)t_m a' l + \big(\beta_m + \mu_m(l)\big)(1-a')t_m l$$

and so

$$a'\big(\mu_m(l) + \sqrt{\mu_m(l)}\big) \leqslant \mu_m(l) + \sqrt{\mu_m(t_m)}.$$

Since $\mu_m(l) > 0$ then

$$a' \leqslant 2\sqrt{\mu_m(l)}$$

provided that t_m is sufficiently large depending on l.

Now, let

$$T_2' = \big\{j < t_m:\ s^m(Y_j) \geqslant \big(\beta_m - \sqrt{\mu_m(l)}\big)l \text{ and}$$
$$p^m(Y_j) < \Big(a_m - \sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}}\Big)l\big\},$$
$$|T_2'| = b' t_m.$$

For $j \notin T_1'$, we have by the definition of ε_m,

$$p^m(Y_j) \leqslant f_m(l) \leqslant \big(a_m + \varepsilon_m(l)\big)l.$$

Therefore,

$$\sum_{j<t_m} p^m(Y_j) = \sum_{j\in T_1'} p^m(Y_j) + \sum_{j\in T_2'} p^m(Y_j) + \sum_{j\notin T_1'\cup T_2'} p^m(Y_j)$$
$$\leqslant a' t_m l + \Big(a_m - \sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}}\Big)b' t_m l +$$
$$+ (1 - a' - b')\big(a_m + \varepsilon_m(l)\big)t_m l$$
$$\leqslant 2t_m l\sqrt{\mu_m(l)} + \Big(a_m - \sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}}\Big)b' t_m l +$$
$$+ (1 - a' - b')\big(a_m + \varepsilon_m(l)\big)t_m l$$

for t_m sufficiently large depending on l. But

$$\sum_{j<t_m} p^m(Y_j) = \sum_{i<l} p^m(X_i) \geqslant \Big(a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}\Big)t_m l.$$

Thus,

$$a_m - \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}$$
$$\leqslant 2\sqrt{\mu_m(l)} + a_m + \varepsilon_m(l) - b'\Big(\sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}} + \varepsilon_m(l)\Big),$$

and

$$b'\Big(\sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}} + \varepsilon_m(l)\Big) \leqslant 2\sqrt{\mu_m(l)} + \varepsilon_m(l) + \sqrt{\sqrt{\varepsilon_m(t_m)} + \sqrt{\mu_m(t_m)}}.$$

Thus, for t_m sufficiently large depending on l, we obtain

$$b' \leqslant 2\sqrt{\sqrt{\varepsilon_m(l)} + \sqrt{\mu_m(l)}}.$$

Hence,

$$(8)\quad |[0, t_m)-T_1'-T_2'|$$
$$= \left|\left\{j<t_m:\ s^m(Y_j)\geqslant \left(\beta_m-\sqrt{\mu_m(l)}\right)l \text{ and } p^m(Y_j)\geqslant \left(\alpha_m-\sqrt{\sqrt{\varepsilon_m(l)}+\sqrt{\mu_m(l)}}\right)l\right\}\right|$$
$$\geqslant (1-a'-b')t_m \geqslant \left(1-2\left(\sqrt{\mu_m(l)}+\sqrt{\sqrt{\varepsilon_m(l)}+\sqrt{\mu_m(l)}}\right)\right)t_m.$$

But for $j\in[0, t_m]-T_1'$, the definition of ε_m implies

$$(9)\qquad p^m(Y_j)\leqslant \left(\alpha_m+\varepsilon_m(l)\right)l.$$

Therefore, for l sufficiently large depending on δ and τ (and t_m sufficiently large depending on l) we have by (8) and (9),

$$\left|\left\{j<t_m:\ \left(\alpha_m-\frac{\delta}{2}\right)l<p^m(Y_j)<\left(\alpha_m+\frac{\delta}{2}\right)l\right\}\right| > \left(1-\frac{\tau\delta}{4}\right)t_m.$$

Now, suppose

$$p^m(X_i, C)\leqslant (\alpha_m-\delta)|C|$$

for all $i<l$. Then

$$\sum_{i<l} p^m(X_i, C) = \sum_{j\in C} p^m(Y_j)\leqslant (\alpha_m-\delta)|C|l.$$

On the other hand, we have

$$\sum_{j\in C} p^m(Y_j)\geqslant \sum_{j\in C-(T_1'\cup T_2')} p^m(Y_j)$$
$$\geqslant \left(\alpha_m-\frac{\delta}{2}\right)l\,|C-(T_1'\cup T_2')| \geqslant \left(\alpha_m-\frac{\delta}{2}\right)l\left(|C|-\frac{\tau\delta}{4}t_m\right).$$

Thus,

$$\left(\alpha_m-\frac{\delta}{2}\right)\left(|C|-\frac{\tau\delta}{4}t_m\right)\leqslant (\alpha_m-\delta)|C|,$$
$$\frac{\delta}{2}|C|\leqslant \left(\alpha_m-\frac{\delta}{2}\right)\frac{\tau\delta}{4}t_m<\frac{\tau\delta}{4}t_m,$$

which contradicts the assumption $|C|\geqslant \tau t_m$. This proves the first assertion of the lemma.

The second assertion is more direct. Suppose

$$p^m(X_i, C)\geqslant (\alpha_m+\delta)|C|$$

for all $i<l$. Then

$$\sum_{i<l} p^m(X_i, C) = \sum_{j\in C} p^m(Y_j)\geqslant (\alpha_m+\delta)|C|l.$$

But by (9)

$$\left|\left\{j<t_m:\ p^m(Y_j)>\left(\alpha_m+\varepsilon_m(l)\right)l\right\}\right|\leqslant |T_1'| = a't_m\leqslant 2\sqrt{\mu_m(l)}\,t_m.$$

Thus,

$$\sum_{j\in C} p^m(Y_j) = \sum_{j\in C\cap T_1'} p^m(Y_j)+\sum_{j\in C-T_1'} p^m(Y_j)\leqslant 2\sqrt{\mu_m(l)}\,t_m l+\left(\alpha_m+\varepsilon_m(l)\right)|C|l.$$

Therefore, comparing the two inequalities for $\sum_{j\in C} p^m(Y_j)$, we have

$$(\alpha_m+\delta)|C|l\leqslant 2\sqrt{\mu_m(l)}\,t_m l+\left(\alpha_m+\varepsilon_m(l)\right)|C|l,$$

and

$$|C|\leqslant \frac{2\sqrt{\mu_m(l)}}{\delta-\varepsilon_m(l)}\,t_m$$

which is impossible for l sufficiently large. This completes the proof of the lemma. ∎

4. Further definitions. Define for $i<K$ the quantities

$$D^i(t_1, \ldots, t_m, K)$$
$$= \left\{X\in C(t_1, \ldots, t_m, K):\ X = \bigcup_{j<K} X_j \text{ and } X_j\in P(t_1, \ldots, t_m) \text{ for all } j<i\right\},$$
$$D^{*i}(t_1, \ldots, t_m, K)$$
$$= \left\{X\in B(t_1, \ldots, t_m, K):\ X = \bigcup_{j<K} X_j \text{ and } X_j\in P(t_1, \ldots, t_m) \text{ for all } j<i\right\}.$$

Clearly

$$D^0(t_1, \ldots, t_m, K) = C(t_1, \ldots, t_m, K),$$
$$D^{*0}(t_1, \ldots, t_m, K) = B(t_1, \ldots, t_m, K),$$

and $D^i\subseteq D^{*i}$ for all $i<K$.

The basic idea of the proof will be to fix K arbitrarily and show by induction on i that $D^i(t_1, \ldots, t_m, K)\neq\emptyset$ for any value of m provided t_m satisfies certain conditions. This will obviously show that R contains arbitrarily long arithmetic progressions (e.g., by choosing $i = K-1$ and $m = 0$).

For $X = \bigcup_{i<K} X_i\in B(t_1, \ldots, t_m, K)$, write

$$X_i = \bigcup_{j<t_m} X_{i,j}\in B(t_1, \ldots, t_m)\quad \text{where}\quad X_{i,j}\in B(t_1, \ldots, t_{m-1}).$$

Define

$$E(t, K) = \{(j_0, \ldots, j_{K-1}):\ j_i<t \text{ for all } i<K$$
$$\text{and } j_0, \ldots, j_{K-1} \text{ forms an arithmetic progression}\}.$$

It is allowed in the definition of $E(t, K)$ for $j_0 = j_1 = \ldots = j_{K-1}$ or $j_0 > j_1 > \ldots > j_{K-1}$ as well as $j_0 < j_1 < \ldots < j_{K-1}$.

FACT 9. *With the notation defined above, we have*

$$(10) \qquad (j_0, \ldots, j_{K-1}) \in E(t_m, K) \quad \text{iff} \quad \bigcup_{i<K} X_{i,j_i} \in B(t_1, \ldots, t_{m-1}, K).$$

Proof. The proof is immediate. If $(j_0, \ldots, j_{K-1}) \in E(t_m, K)$ then certainly $\bigcup_{i<K} X_{i,j_i} \in B(t_1, \ldots, t_{m-1}, K)$. On the other hand, if $\bigcup_{i<K} X_{i,j_i} \in B(t_1, \ldots, t_{m-1}, K)$ then $j_0, \ldots, j_{K-1}$ must form an arithmetic progression of length K with all $j_i < t_m$, i.e., $(j_0, \ldots, j_{K-1}) \in E(t_m, K)$. ■

For $i < K$, $j < t$ define

$$E(t, K, j, i) = \{(j_0, \ldots, j_{K-1}) \in E(t, K) \colon j_i = j\},$$

$$e(t, K, j, i) = |E(t, K, j, i)|.$$

FACT 10. *With the notation defined above, we have*

$$(11) \qquad e(t, K; j, i) \leqslant t,$$

and for $t/4 < j < 3t/4$, $K \geqslant 2$ *and* $t \geqslant 4$,

$$(11') \qquad e(t, K, j, i) \geqslant t/K^2.$$

Proof. If $(j_0, \ldots, j_{K-1}) \in E(t, K, j, i)$ then all $j_s < t$ and $j_i = j$. Thus, if $K > 1$ then the choice of j_{i+1} (or j_{i-1} if $i = K-1$) determines the rest of the arithmetic progression. But there are at most t choices for j_{i+1} (or j_{i-1} if $i = K-1$) so that for $K > 1$, $e(t, K, j, i) \leqslant t$. Of course, if $K = 1$ then $e(t, K, j, i) = 1 \leqslant t$. This proves (11).

To prove (11'), we note that the worst case occurs when $j = [t/2]$ and $i = K-1$. Clearly, $(a, a+d, \ldots, a+(K-1)d) \in E(t, K, [t/2], K-1)$ whenever $a \geqslant 0$, $d \geqslant 0$ and $a+(k-1)d = [t/2]$, i.e., whenever $0 \leqslant d \leqslant \frac{1}{K-1}[t/2]$.

Since

$$\frac{1}{K-1}\left[\frac{t}{2}\right] \geqslant \frac{t}{K^2}$$

for $K \geqslant 2$ and $t \geqslant 4$ then (11') is proved. ■

We shall also need the following result.

FACT 11. *Suppose* $t > l^3 > 1$ *and* $L \subseteq [0, t)$ *with* $|L| > (1-1/l)t$. *Then* L *contains an arithmetic progression of length* l.

Proof. Write $t = t'l + \bar{l}$ with $0 \leqslant \bar{l} < l$ and let I_j denote $[jl, (j+1)l)$ for $j < t'$. Since $|L| > (1-1/l)t$ then

$$|[0, t) - L| < t/l \leqslant t'+1, \quad \text{i.e.,} \quad |[0, t) - L| \leqslant t'.$$

Thus, each I_j must have $|I_j \cap L| = l-1$. For, if $|I_j \cap L| = l$ then we would be done while if $|I_j \cap L| \leqslant l-2$ then for some $j' < t'$, $|I_{j'} \cap L| = l$ and again, we are done. Let $k_j + jl$ be the unique element of I_j not in L. Then we have $k_j \geqslant k_{j+1}$ since otherwise $I_j \cup I_{j+1}$ contains a block of l consecutive integers. But now the sequence

$$l > k_1 \geqslant \ldots \geqslant k_{l^2} \geqslant 0$$

(which is defined since $t' \geqslant l^2$) must have at least l consecutive terms which are equal. The corresponding I_j's clearly contain an arithmetic progression of length l. ■

We next define another subclass of $C(t_1, \ldots, t_m, K)$. For $i \leqslant s < K$ and $X \in C(t_1, \ldots, t_m, K)$, let

$$F(X, j, i, s) = \Big\{(j_0, \ldots, j_{K-1}) \in E(t_m, K, j, s) \colon \bigcup_{i'<K} X_{i',j_{i'}} \in D^{*i}(t_1, \ldots, t_{m-1}, K)\Big\}$$

and let

$$f(X, j, i, s) = |F(X, j, i, s)|.$$

Define $G^i(t_1, \ldots, t_m, K)$ to be the set

$\{X \in C(t_1, \ldots, t_m, K)$: for every s with $i \leqslant s < K$, the condition

$$\text{(a)} \qquad f(X, j, i, s) \leqslant 2^i \alpha_m^i t_m$$

fails for at most $2i\alpha_m^K(1-\beta_m)t_m$ indices $j < t_m$, and the condition

$$\text{(b)} \qquad f(X, j, i, s) \geqslant \frac{1}{K^2}\frac{1}{2^i}\alpha_m^i t_m$$

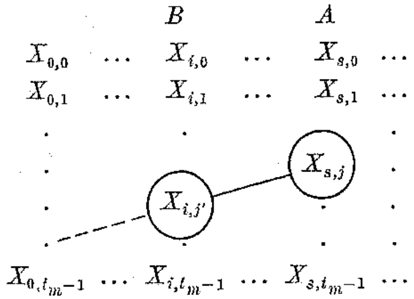
fails for at most $2i\alpha_m^K(1-\beta_m)t_m$ indices j such that $t_m/4 < j < 3t_m/4\}$.

The elements of $G^i(t_1, \ldots, t_m, K)$ will be called *homogeneous K-tuples of type* (m, i). The elements of $D^i(t_1, \ldots, t_m, K)$ will be called simply *K-tuples of type* (m, i). The existence of such K-tuples will be proved by a simultaneous induction. Note that by Fact 10,

$$G^0(t_1, \ldots, t_m, K) = C(t_1, \ldots, t_m, K).$$

Now let $X \in B(t_1, \ldots, t_m, K)$ be arbitrary and let $0 \leqslant i \leqslant s < K$. We shall define a bipartite graph $I(X, i, s)$ which will eventually connect Lemma 1 with the rest of the proof. The vertex sets A and B of $I(X, i, s)$ will both have cardinality t_m. Even though A and B are to be disjoint, we shall use the set $[0, t_m)$ both for A and B (this should cause no confusion). The pair $\{j, j'\}$, $j \in A$, $j' \in B$, is an *edge* of $I(X, i, s)$ if $(j_0, \ldots, j_{K-1}) \in F(X, j, i, s)$ and $j_i = j'$. It is recommended that the reader study this definition carefully since an understanding of the graph $I(X, i, s)$

will be crucial in the remainder of the proof. A useful way to picture $I(X, i, s)$ is the following. Write $X \in B(t_1, \ldots, t_m, K)$ as $X = \bigcup_{i<K} X_i$, $X_i \in B(t_1, \ldots, t_m)$, in the usual way. Also, write $X_i = \bigcup_{j<t_m} X_{i,j}$, $i < K$, in the usual way. We can display these $X_{i,j}$ in the following way.



The set A corresponds to the set of indices j of the $X_{s,j}$, $j < t_m$. The set B corresponds to the set of indices j' of the $X_{i,j'}$, $j' < t_m$. The pair $\{j, j'\}$ is an edge of $I(X, i, s)$ if and only if the arithmetic progression $j_0, j_1, \ldots, j_{K-1}$ defined by letting $j_i = j'$ and $j_s = j$ has all its terms in $[0, t_m)$ *and* $X_{i', j_{i'}} \in P(t_1, \ldots, t_{m-1})$ for all $i' < i$.

We remark that if $X \in G^i(t_1, \ldots, t_m, K)$ then the graph $I(X, i, s)$ satisfies strong valency constraints.

5. The choice of t_m. In this section we specify how the choice of t_m is to be made. We have begun with an arbitrary but fixed sequence R of positive density and an arbitrary but fixed integer $K \geq 2$. Assume now that in the inductive definitions of Section 3, for some $m \geq 1$ the quantities $S(t_1, \ldots, t_{m-1})$, $P(t_1, \ldots, t_{m-1})$, $g_m(l)$, $f_m(l)$, α_m, β_m, $\varepsilon_m(l)$ and $\mu_m(l)$ have already been defined.

5.1. Choose the constants $\varepsilon_1^{(m)}, \varepsilon_2^{(m)}, \delta^{(m)}, \varrho^{(m)}, \sigma^{(m)}$ as follows:

(i) $\varepsilon_1^{(m)} = \alpha_m^K(1-\beta_m)$,

(ii) $\varepsilon_2^{(m)} = \dfrac{\alpha_m}{2}$,

(iii) $\delta^{(m)} = \dfrac{1}{10}\dfrac{\alpha_m^K}{2^K K^2}$,

(iv) $\varrho^{(m)} = \alpha_m^K(1-\beta_m)$,

(v) $\sigma^{(m)} = \dfrac{1}{10}\dfrac{\alpha_m^K}{2^K K^2}$.

5.2. Let us denote by $m_0^{(m)}, n_0^{(m)}, M^{(m)}, N^{(m)}$ the numbers given by Lemma 1 for the corresponding constants chosen in 5.1.

(A) Let $t_m > 2M^{(m)}$, $t_m > N^{(m)}$.

5.3. Let $\tau^{(m)} = \dfrac{\sigma^{(m)}}{n_0^{(m)}}$ and let l'_m be a large number for which the statement of Lemma 3 is satisfied with $\delta = \delta^{(m)}$, $\tau = \tau^{(m)}$ provided t_m is sufficiently large depending on l'_m. Also choose $l'_m > t_{m-1}$.

(B) Let t_m be so large that 5.3 is valid.

5.4. By van der Waerden's theorem, for integers i and j, there exists a least integer $w(i, j)$ such that for any partition of $[0, w(i, j))$ into j classes, at least one class contains an arithmetic progression of length i. Choose l_m so that l_m^{2-K} is an integer exceeding $w(l'_m, 4Km_0^{(m)}n_0^{(m)})$. Thus,

$$2K^3 2^{2K} l_m < \frac{1}{\sqrt{1-\beta_{m+1}}}$$

provided t_m is sufficiently large (by Lemma 2).

(C) Let t_m be so large that 5.4 holds (this also guarantees that $\beta_m > 0$ for $m > 1$).

(D) Let $t_m \geq 4$ (this guarantees that Fact 10 holds with $t = t_m$).

(E) Let $t_m > 4l_{m-1}^3$ for $m \geq 2$.

(F) Let t_m be so large that $\sqrt{\mu_m(t_m)} < \min\{\beta_m, 1-\beta_m\}$ (which is possible since we may assume $0 < \beta_m < 1$).

5.5. As noted in the proof of Lemma 2, there exist arbitrarily large values of l such that $\varepsilon_1(l) > 0$. Using Lemma 2 inductively, we also see that for $m > 1$, there exist arbitrarily large values of l such that $\mu_m(l) > 0$.

(G) Let the t_m be chosen so that $\varepsilon_1(t_1) > 0$ and for $m > 1$, $\mu_m(t_m) > 0$.

In what follows, t_m denotes a fixed number satisfying each of the conditions (A)–(G).

6. Well-saturated K-tuples. Suppose $X \in G^i(t_1, \ldots, t_m, K)$ with $0 \leq i < K-1$. For $i+1 \leq s < K$, let the sets $C_\mu(X, s)$, $\mu < m_0^{(m)}$, and $C_{\mu,\nu}(X, s)$, $\mu < m_0^{(m)}$, $\nu < n_0^{(m)}$, satisfy Lemma 1 with the choice of constants given in 5.1, where the vertex sets A and B of Lemma 1 are defined by

$$A = A^{(m)} = \{j\colon \tfrac{1}{4}t_m < j < \tfrac{3}{4}t_m\},$$
$$B = B^{(m)} = \{j\colon j < t_m\} = [0, t_m),$$

and the bipartite graph I of Lemma 1 is defined by

$$I = I(X, i, s).$$

Similarly, for $i+1 \leq s < K$, let the sets $\bar{C}_\mu(X, s)$, $\mu < m_0^{(m)}$, and $\bar{C}_{\mu,\nu}(X, s)$, $\mu < m_0^{(m)}$, $\nu < n_0^{(m)}$, satisfy Lemma 1′ with the choice of constants given in 5.1, where the vertex sets A and B of Lemma 1′ are defined by

$$A = \bar{A}^{(m)} = [0, t_m), \quad B = \bar{B}^{(m)} = [0, t_m),$$

and the bipartite graph I of Lemma 1′ is defined by

$$I = I(X, i, s).$$

As we have previously noted, the common indexing of the disjoint vertex sets A and B should cause no confusion.

Assume $m \geqslant 1$ and write X in the usual way as

$$X = \bigcup_{i'<K} X_{i'} \in G^i(t_1, \ldots, t_m, K),$$

$$X_{i'} = \bigcup_{j<t_m} X_{i',j} \in B(t_1, \ldots, t_m), \quad i' < K.$$

We shall say that X is *well-saturated* if for all s with $i+1 \leqslant s < K$,

$$\big|p^m\big(X_i, C_{\mu,\nu}(X, s)\big) - a_m |C_{\mu,\nu}(X, s)|\big| \leqslant \delta^{(m)} |C_{\mu,\nu}(X, s)|$$

and

$$\big|p^m\big(X_i, \bar{C}_{\mu,\nu}(X, s)\big) - a_m |\bar{C}_{\mu,\nu}(X, s)|\big| \leqslant \delta^{(m)} |\bar{C}_{\mu,\nu}(X, s)|$$

whenever

$$|C_{\mu,\nu}(X, s)| \geqslant \tau^{(m)} t_m, \quad |\bar{C}_{\mu,\nu}(X, s)| \geqslant \tau^{(m)} t_m,$$

where $\mu < m_0^{(m)}$, $\nu < n_0^{(m)}$.

Lemma 4. *If $X \in G^i(t_1, \ldots, t_m, K)$ is well-saturated with $i+1 < K$ then*

$$X \in G^{i+1}(t_1, \ldots, t_m, K).$$

Proof. Choose a fixed s with $i+1 \leqslant s < K$. We must show:

(a) $f(X, j, i+1, s) \leqslant 2^{i+1} a_m^{i+1} t_m$ fails for at most $2(i+1)\alpha_m^K(1-\beta_m)t_m$ indices $j < t_m$,

(b) $f(X, j, i+1, s) \geqslant \frac{1}{K^2}\frac{1}{2^{i+1}} a_m^{i+1} t_m$ fails for at most $2(i+1) \times$ $\times \alpha_m^K(1-\beta_m)t_m$ indices j with $\frac{1}{4}t_m < j < \frac{3}{4}t_m$.

We first prove (b). For the sake of brevity, set

$$C_\mu = C_\mu(X, s), \quad C_{\mu,\nu} = C_{\mu,\nu}(X, s),$$

$$\bar{C}_\mu = \bar{C}_\mu(X, s), \quad \bar{C}_{\mu,\nu} = \bar{C}_{\mu,\nu}(X, s),$$

$$I = I(X, i, s).$$

Let $Z \subseteq A^{(m)}$ be defined by

$$Z = \Big\{j:\ \tfrac{1}{4}t_m < j < \tfrac{3}{4}t_m \text{ and } j \in \bigcup_{\mu<m_0^{(m)}} C_\mu \text{ and}$$

$$f(X, j, i+1, s) \leqslant \frac{1}{K^2}\Big(\frac{a_m}{2}\Big)^{i+1} t_m \text{ and } f(X, j, i, s) \geqslant \frac{1}{K^2}\Big(\frac{a_m}{2}\Big)^{i} t_m\Big\}.$$

(i) *Assume* $|Z| \leqslant \alpha_m^K(1-\beta_m)t_m$. By Lemma 1 (which applies by condition (A) in the choice of t_m) we know

$$\Big|A^{(m)} - \bigcup_{\mu<m_0^{(m)}} C_\mu\Big| \leqslant \varrho^{(m)}|A^{(m)}| \leqslant \varrho^{(m)} t_m = \alpha_m^K(1-\beta_m)t_m.$$

Also, by the definition of $G^i(t_1, \ldots, t_m, K)$, there are at most $2i\alpha_m^K(1-\beta_m)t_m$ indices j with $f(X, j, i, s) < \frac{1}{K^2}\Big(\frac{a_m}{2}\Big)^i t_m$. Thus

$$\Big|\Big\{j:\ \tfrac{1}{4}t_m < j < \tfrac{3}{4}t_m \text{ and } f(X, j, i+1, s) \leqslant \frac{1}{K^2}\Big(\frac{a_m}{2}\Big)^{i+1} t_m\Big\}\Big|$$

$$\leqslant \alpha_m^K(1-\beta_m)t_m + \alpha_m^K(1-\beta_m)t_m + 2i\alpha_m^K(1-\beta_m)t_m = 2(i+1)\alpha_m^K(1-\beta_m)t_m$$

which is just (b).

(ii) *Assume* $|Z| > \alpha_m^K(1-\beta_m)t_m$. If

$$|Z \cap C_\mu| \leqslant \alpha_m^K(1-\beta_m)|C_\mu| \quad \text{for all } \mu < m_0^{(m)}$$

then

$$|Z| = \Big|Z \cap \bigcup_{\mu<m_0^{(m)}} C_\mu\Big| = \sum_{\mu<m_0^{(m)}} |Z \cap C_\mu| \text{ since the } C_\mu \text{ are disjoint}$$

$$\leqslant \alpha_m^K(1-\beta_m)\sum_{\mu<m_0^{(m)}} |C_\mu| \leqslant \alpha_m^K(1-\beta_m)|A^{(m)}| \leqslant \alpha_m^K(1-\beta_m)t_m$$

which is a contradiction.

Thus, for some $\mu < m_0^{(m)}$

$$|Z \cap C_\mu| > \alpha_m^K(1-\beta_m)|C_\mu| = \varepsilon_1^{(m)}|C_\mu|.$$

Let $Z' = Z \cap C_\mu$. We first show that $k(Z', B^{(m)})$ is not very large.

Let

$$X = B^{(m)} - \bigcup_{\nu<n_0^{(m)}} C_{\mu,\nu} = [0, t_m) - \bigcup_{\nu<n_0^{(m)}} C_{\mu,\nu}.$$

By Lemma 1, $|X| < \sigma^{(m)} t_m$. Also, by Lemma 1, for $x \in C_\mu$, we have

$$|k(x) \cap C_{\mu,\nu}| \leqslant \big(\beta(C_\mu, C_{\mu,\nu}) + \delta^{(m)}\big)|C_{\mu,\nu}|$$

and so,

$$k(Z', C_{\mu,\nu}) \leqslant \big(\beta(C_\mu, C_{\mu,\nu}) + \delta^{(m)}\big)|C_{\mu,\nu}|\ |Z'|.$$

Therefore,

$$(12)\qquad k(Z', B^{(m)}) = \sum_{\nu<n_0^{(m)}} k(Z', C_{\mu,\nu}) + k(Z', X)$$
$$\leqslant \sum_{\nu<n_0^{(m)}} \big(\beta(C_\mu, C_{\mu,\nu}) + \delta^{(m)}\big)|C_{\mu,\nu}||Z'| + |Z'||X|$$
$$\leqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu}||Z'| + \sum_{\nu<n_0^{(m)}} \delta^{(m)}|C_{\mu,\nu}||Z'| + \sigma^{(m)}|Z'|t_m$$
$$\leqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu}||Z'| + (\delta^{(m)} + \sigma^{(m)})|Z'|t_m.$$

Next, let

$$\bar{L} = \{j\colon j < t_m \text{ and } X_{i,j} \in P(t_1, \ldots, t_{m-1})\}.$$

Thus, for all $j \in A^{(m)}$,

$$(13)\qquad |k(j) \cap \bar{L}| = f(X, j, i+1, s)$$

since any edge from j to a vertex $j' \in B^{(m)}$ already has $j_s = j$ and $j_i = j'$ for some arithmetic progression $j_0, j_1, \ldots, j_{K-1}$ such that the sets $X_{0,j_0}, \ldots, X_{i-1,j_{i-1}}$ all belong to $P(t_1, \ldots, t_{m-1})$. Consequently, if $j' \in \bar{L}$ then $X_{i,j'} = X_{i,j_i}$ also belongs to $P(t_1, \ldots, t_{m-1})$ and so

$$(j_0, \ldots, j_{K-1}) \in F(X, j, i+1, s).$$

Since by hypothesis X is well-saturated then for $|C_{\mu,\nu}| \geqslant \tau^{(m)} t_m$ we have

$$|\bar{L} \cap C_{\mu,\nu}| \geqslant (a_m - \delta^{(m)})|C_{\mu,\nu}|$$
$$> \frac{a_m}{2}|C_{\mu,\nu}| \quad \text{since} \quad \delta^{(m)} = \frac{1}{10}\left(\frac{a_m}{2}\right)^K \cdot \frac{1}{K^2} < \frac{a_m}{2}$$
$$= \varepsilon_2^{(m)}|C_{\mu,\nu}|.$$

But we have already seen that $|Z'| > \varepsilon_1^{(m)}|C_\mu|$. Thus, we may apply Lemma 1, obtaining

$$\beta(Z', \bar{L} \cap C_{\mu,\nu}) \geqslant \beta(C_\mu, C_{\mu,\nu}) - \delta^{(m)}$$

for those ν such that $|C_{\mu,\nu}| \geqslant \tau^{(m)} t_m$.

Therefore,

$$k(Z', \bar{L}) \geqslant \sum_{\nu<n_0^{(m)}} k(Z', C_{\mu,\nu} \cap \bar{L}) = \sum_{\nu<n_0^{(m)}} \beta(Z', C_{\mu,\nu} \cap \bar{L})|C_{\mu,\nu} \cap \bar{L}||Z'|$$

$$= \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| \geqslant \tau^{(m)} t_m}} + \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| < \tau^{(m)} t_m}}$$
$$\geqslant \sum_{\nu<n_0^{(m)}} \big(\beta(C_\mu, C_{\mu,\nu}) - \delta^{(m)}\big)|C_{\nu,\mu} \cap \bar{L}||Z'| -$$
$$- \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| < \tau^{(m)} t_m}} \big(\beta(C_\mu, C_{\mu,\nu}) - \delta^{(m)}\big)|C_{\mu,\nu} \cap \bar{L}||Z'|$$
$$\geqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu} \cap \bar{L}||Z'| -$$
$$- \delta^{(m)}|Z'| \sum_{\nu<n_0^{(m)}} |C_{\mu,\nu} \cap \bar{L}| - \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| < \tau^{(m)} t_m}} 1 \cdot |C_{\mu,\nu} \cap \bar{L}||Z'|$$
$$\geqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu} \cap \bar{L}||Z'| - \delta^{(m)}|Z'|t_m - n_0^{(m)}\tau^{(m)}|Z'|t_m$$
$$\geqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu} \cap \bar{L}||Z'| - 2\delta^{(m)}|Z'|t_m$$

since $\tau^{(m)} = \dfrac{\sigma^{(m)}}{n_0^{(m)}}$.

But

$$|\bar{L} \cap C_{\mu,\nu}| = p^m(X_i, C_{\mu,\nu})$$

so that X being well-saturated implies

$$|\bar{L} \cap C_{\mu,\nu}| \geqslant a_m|C_{\mu,\nu}| - \delta^{(m)}|C_{\mu,\nu}|$$

provided $|C_{\mu,\nu}| \geqslant \tau^{(m)} t_m$.

Thus,

$$\sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu} \cap \bar{L}|\,|Z'| \geqslant \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| \geqslant \tau^{(m)} t_m}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu} \cap \bar{L}|\,|Z'|$$
$$\geqslant \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})(a_m - \delta^{(m)})|C_{\mu,\nu}||Z'| -$$
$$- \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}| < \tau^{(m)} t_m}} \beta(C_\mu, C_{\mu,\nu})(a_m - \delta^{(m)})|C_{\mu,\nu}|\,|Z'|$$
$$\geqslant a_m \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu}|\,|Z'| - 2\delta^{(m)}|Z'|t_m$$

as in the preceding inequality. Substituting this inequality for the corresponding term in the preceding inequality for $k(Z', \bar{L})$ we obtain

$$(14)\qquad k(Z', \bar{L}) \geqslant a_m \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu}|\,|Z'| - 4\delta^{(m)}|Z'|t_m.$$

Since $\sigma^{(m)} = \delta^{(m)}$, we may multiply (12) by a_m to obtain

$$a_m k(Z', B^{(m)}) \leqslant a_m \sum_{\nu<n_0^{(m)}} \beta(C_\mu, C_{\mu,\nu})|C_{\mu,\nu}|\,|Z'| + 2\delta^{(m)}a_m|Z'|t_m.$$

Thus, by (14) we have

$$k(Z', \bar{L}) \geqslant a_m k(Z', B^{(m)}) - 6\delta^{(m)}|Z'|t_m.$$

But, by the definition of Z', for each $j \in Z'$,

$$f(X, j, i, s) \geqslant \frac{1}{K^2}\left(\frac{a_m}{2}\right)^i t_m$$

and

$$f(X, j, i+1, s) \leqslant \frac{1}{K^2}\left(\frac{a_m}{2}\right)^{i+1} t_m.$$

Hence, by (13)

$$\begin{aligned}\frac{1}{K^2}\left(\frac{a_m}{2}\right)^{i+1}|Z'|t_m &\geqslant \sum_{j\in Z'} f(X, j, i+1, s) = k(Z', \bar{L})\\ &\geqslant a_m k(Z', B^{(m)}) - 6\delta^{(m)}|Z'|t_m\\ &= a_m \sum_{j\in Z'} f(X, j, i, s) - 6\delta^{(m)}|Z'|t_m\\ &\geqslant \frac{2}{K^2}\left(\frac{a_m}{2}\right)^{i+1}|Z'|t_m - 6\delta^{(m)}|Z'|t_m.\end{aligned}$$

Therefore,

$$\frac{1}{K^2}\left(\frac{a_m}{2}\right)^{i+1} \leqslant 6\delta^{(m)}$$

which *contradicts* the definition of $\delta^{(m)}$ since by 5.1

$$\delta^m = \frac{1}{10}\cdot\frac{1}{K^2}\left(\frac{a_m}{2}\right)^K < \frac{1}{6}\cdot\frac{1}{K^2}\left(\frac{a_m}{2}\right)^{i+1} \quad \text{for} \quad i+1 < K.$$

This proves (b).

Next we prove (a). The proof will be quite similar to the proof of (b). Set

$$\bar{Z} = \{j : j \in \bigcup_{\mu<m_0^{(m)}} \bar{C}_\mu \text{ and } f(X, j, i+1, s) \geqslant 2^{i+1}a_m^{i+1}t_m \text{ and } f(X, j, i, s) \leqslant 2^i a_m^i t_m\}.$$

As before, if $|\bar{Z}| \leqslant a_m^K(1-\beta_m)t_m$ then we would be done. Hence, we may *assume* $|\bar{Z}| > a_m^K(1-\beta_m)t_m$. Again as before, for some $\mu < m_0^{(m)}$, we have

$$|\bar{Z} \cap \bar{C}_\mu| > a_m^K(1-\beta_m)|\bar{C}_\mu| = \varepsilon_1^{(m)}|\bar{C}_\mu|.$$

We let $\bar{Z}' = \bar{Z} \cap \bar{C}_\mu$ so that $|\bar{Z}'| > \varepsilon_1^{(m)}|\bar{C}_\mu|$. By Lemma 1′, for each $x \in \bar{Z}'$, we have

$$|k(x) \cap \bar{C}_{\mu,\nu}| \geqslant \big(\beta(\bar{C}_\mu, \bar{C}_{\mu,\nu}) - \delta^{(m)}\big)|\bar{C}_{\mu,\nu}|.$$

Therefore,

$$k(\bar{Z}', \bar{C}_{\mu,\nu}) = \sum_{x\in\bar{Z}'} |k(x) \cap \bar{C}_{\mu,\nu}| \geqslant \big(\beta(\bar{C}_\mu, \bar{C}_{\mu,\nu}) - \delta^{(m)}\big)|\bar{C}_{\mu,\nu}|\,|\bar{Z}'|$$

and so

$$\begin{aligned}(15)\qquad k(\bar{Z}', \bar{B}^{(m)}) &\geqslant \sum_{\nu<n_0^{(m)}} k(\bar{Z}', \bar{C}_{\mu,\nu}) \geqslant \sum_{\nu<n_0^{(m)}} \big(\beta(\bar{C}_\mu, \bar{C}_{\mu,\nu}) - \delta^{(m)}\big)|\bar{C}_{\mu,\nu}|\,|\bar{Z}'|\\ &\geqslant \sum_{\nu<n_0^{(m)}} \beta(\bar{C}_\mu, \bar{C}_{\mu,\nu})|\bar{C}_{\mu,\nu}|\,|\bar{Z}'| - \delta^{(m)}|\bar{Z}'|t_m.\end{aligned}$$

Also

$$\begin{aligned}(16)\qquad k(\bar{Z}', \bar{L}) &= \sum_{\nu<n_0^{(m)}} k(\bar{Z}', \bar{C}_{\mu,\nu} \cap \bar{L}) + k\Big(\bar{Z}', \big(\bar{B}^{(m)} - \bigcup_{\nu<n_0^{(m)}} \bar{C}_{\mu,\nu}\big) \cap \bar{L}\Big)\\ &\leqslant \sum_{\nu<n_0^{(m)}} k(\bar{Z}', \bar{C}_{\mu,\nu} \cap \bar{L}) + \sigma^{(m)}|\bar{Z}'|t_m \quad \text{by Lemma 1}'\\ &= \sum_{\nu<n_0^{(m)}} \beta(\bar{Z}', \bar{C}_{\mu,\nu} \cap \bar{L})|\bar{C}_{\mu,\nu} \cap \bar{L}|\,|\bar{Z}'| + \sigma^{(m)}|\bar{Z}'|t_m\\ &= \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}|\geqslant\tau^{(m)}t_m}} + \sum_{\substack{\nu<n_0^{(m)}\\ |C_{\mu,\nu}|<\tau^{(m)}t_m}} + \sigma^{(m)}|\bar{Z}'|t_m.\end{aligned}$$

But, by hypothesis, X is well-saturated so that

$$(a_m - \delta^{(m)})|\bar{C}_{\mu,\nu}| \leqslant |\bar{L} \cap \bar{C}_{\mu,\nu}| \leqslant (a_m + \delta^{(m)})|\bar{C}_{\mu,\nu}|$$

provided $|\bar{C}_{\mu,\nu}| \geqslant \tau^{(m)}t_m$. Since

$$a_m - \delta^{(m)} > \frac{a_m}{2} = \varepsilon_2^{(m)}$$

then we can apply Lemma 1′ to the first summand in (16) and obtain

$$k(\bar{Z}', \bar{L}) \leqslant \sum_{\nu<n_0^{(m)}} \big(\beta(\bar{C}_\mu, \bar{C}_{\mu,\nu}) + \delta^{(m)}\big)|\bar{C}_{\mu,\nu}\cap \bar{L}|\,|\bar{Z}'| + \\ + \tau^{(m)}t_m|\bar{Z}'|\,n_0^{(m)} + \sigma^{(m)}|\bar{Z}'|\,t_m$$

$$\leqslant \sum_{\substack{\nu<n_0^{(m)}\\ |\bar{C}_{\mu,\nu}|\geqslant\tau^{(m)}t_m}} \beta(\bar{C}_\mu, \bar{C}_{\mu,\nu})|\bar{C}_{\mu,\nu}\cap \bar{L}|\,|\bar{Z}'| + \\ + \sum_{\substack{\nu<n_0^{(m)}\\ |\bar{C}_{\mu,\nu}|<\tau^{(m)}t_m}} \beta(\bar{C}_\mu, \bar{C}_{\mu,\nu})|\bar{C}_{\mu,\nu}\cap \bar{L}|\,|\bar{Z}'| + 3\delta^{(m)}|\bar{Z}'|\,t_m$$

$$\leqslant \sum_{\nu<n_0^{(m)}} \beta(\bar{C}_\mu, \bar{C}_{\mu,\nu})(a_m+\delta^{(m)})|\bar{C}_{\mu,\nu}|\,|\bar{Z}'| + \\ + n_0^{(m)}\tau^{(m)}|\bar{Z}'|\,t_m + 3\delta^{(m)}|\bar{Z}'|\,t_m$$

$$\leqslant a_m \sum_{\nu<n_0^{(m)}} \beta(\bar{C}_\mu, \bar{C}_{\mu,\nu})|\bar{C}_{\mu,\nu}|\,|\bar{Z}'| + 5\delta^{(m)}|\bar{Z}'|\,t_m.$$

Thus, by (15), we obtain

$$k(\bar{Z}', \bar{L}) \leqslant a_m k(\bar{Z}', \bar{B}^{(m)}) + 6\delta^{(m)}|\bar{Z}'|\,t_m.$$

By the definition of $\bar{Z}'$, for each $j\in\bar{Z}'$,

$$f(X, j, i, s) \leqslant 2^i a_m^i t_m$$

and

$$f(X, j, i+1, s) \geqslant 2^{i+1} a_m^{i+1} t_m.$$

Hence,

$$2^{i+1}a_m^{i+1}|\bar{Z}'|\,t_m \leqslant \sum_{j\in\bar{Z}'} f(X, j, i+1, s) = k(\bar{Z}', \bar{L})$$
$$\leqslant a_m k(\bar{Z}', \bar{B}^{(m)}) + 6\delta^{(m)}|\bar{Z}'|\,t_m$$
$$\leqslant a_m 2^i a_m^i|\bar{Z}'|\,t_m + 6\delta^{(m)}|\bar{Z}'|\,t_m.$$

Therefore,

$$2^i a_m^{i+1} \leqslant 6\delta^{(m)}$$

which *contradicts* the definition of $\delta^{(m)}$. This proves Lemma 4. ∎

LEMMA 5. *Suppose, for some i, $0\leqslant i<K-1$, that*

$$X^{(\xi)}\in G^i(t_1, \ldots, t_m, K), \quad \xi<l, \quad \textit{where} \quad l\geqslant l_m^{2^{-K}}.$$

Assume that

$$X = \bigcup_{\xi<l} X_i^{(\xi)} \in C(t_1, \ldots, t_m, l)$$

and, further, that for each $j<i$, $X_j^{(\xi)}$ and $X_j^{(\eta)}$ are R-equivalent for all ξ, $\eta<l$. Then there exists $\pi<l$ such that

$$X^{(\pi)}\in G^{i+1}(t_1, \ldots, t_m, K).$$

Proof. The assumption that $X_j^{(\xi)}$ and $X_j^{(\eta)}$ are R-equivalent implies

$$I(X^{(\xi)}, i, s) = I(X^{(\eta)}, i, s) \quad \text{for} \quad i\leqslant s<K,\ \xi, \eta<l,$$

since the edges of $I(X, i, s)$ are completely specified by the R-equivalence classes of the X_j for $j<i$. Hence, we may assume

$$C_\mu(X^{(\xi)}, s) = C_\mu(X^{(\eta)}, s) = C_\mu(s),$$
$$C_{\mu,\nu}(X^{(\xi)}, s) = C_{\mu,\nu}(X^{(\eta)}, s) = C_{\mu,\nu}(s),$$

with a similar assumption holding for the $\bar{C}_\mu(s)$ and $\bar{C}_{\mu,\nu}(s)$ as well.

By Lemma 4, it will be enough to show that for some $\pi<l$, $X^{(\pi)}$ is well-saturated. Suppose this is not the case. Then for each $\xi<l$, there exist μ, ν, and s such that either

(i) $|C_{\mu,\nu}(s)|\geqslant\tau^{(m)}t_m$ and

$$\big|p^m(X_i^{(\xi)}, C_{\mu,\nu}(s)) - a_m|C_{\mu,\nu}(s)|\big| > \delta^{(m)}|C_{\mu,\nu}(s)|$$

or

(ii) $|\bar{C}_{\mu,\nu}(s)|\geqslant\tau^{(m)}t_m$ and

$$\big|p^m(X_i^{(\xi)}, \bar{C}_{\mu,\nu}(s)) - a_m|\bar{C}_{\mu,\nu}(s)|\big| > \delta^{(m)}|\bar{C}_{\mu,\nu}(s)|.$$

There are actually four possibilities here, depending upon which way the inequalities go when the absolute value signs are removed. But by the choice of l and l_m (see 5.4), we have

$$l\geqslant l_m^{2^{-K}} > w(l_m', 4Km_0^{(m)}n_0^{(m)}).$$

Since there are at most $m_0^{(m)}$ choices for μ, at most $n_0^{(m)}$ choices for ν and at most K choices for s, then by definition of the van der Waerden function w, there exists an arithmetic progression of length l_m' in $[0, l)$ such that the relevant inequalities in (i) and (ii) involve the same μ, ν and s and so that the inequalities go the same way. Let $\xi_0, \ldots, \xi_{l_m'-1}$ denote this arithmetic progression and suppose for the sake of definiteness that

$$p^m(X_i^{(\xi_j)}, C_{\mu,\nu}(s)) - a_m|C_{\mu,\nu}(s)| > \delta^{(m)}|C_{\mu,\nu}(s)| \tag{17}$$

for all $j<l_m'$. Since, by hypothesis,

$$\bigcup_{\xi<l} X_i^{(\xi)}\in C(t_1, \ldots, t_m, l)$$

then

$$X_i^{(\xi)} \in S(t_1, \ldots, t_m)$$

for all $\xi < l$. But $|C_{\mu,\nu}(s)| \geqslant \tau^{(m)} t_m$ for the special choice of μ, ν, and s above so that we can apply Lemma 3 to $\bigcup_{j<l'_m} X_i^{(\xi_j)}$ (by 5.3) and obtain, for some $j < l'_m$,

$$p^m\big(X_i^{(\xi_j)}, C_{\mu,\nu}(s)\big) < (\alpha_m + \delta^{(m)})\, |C_{\mu,\nu}(s)|.$$

This easily *contradicts* (17). The other cases (for the other choices corresponding to (17)) are treated in exactly the same way. This proves Lemma 5. ∎

We come to the last lemma needed for the proof of the main theorem.

LEMMA 6. *Let $l > 0$, $m \geqslant 0$ and i with $0 \leqslant i < K$ be fixed integers and assume that*

$$X^{(\xi)} = \bigcup_{i'<K} X_{i'}^{(\xi)} \in G^i(t_1, \ldots, t_{m+1}, K)$$

for $\xi < l \leqslant l_m$. Further, assume that for all $j < i$, $\xi, \eta < l$, $X_j^{(\xi)}$ and $X_j^{(\eta)}$ are R-equivalent. Write, as usual,

$$X_{i'}^{(\xi)} = \bigcup_{j'<t_{m+1}} X_{i',j'}^{(\xi)} \quad \textit{for} \quad i' < K.$$

Then there exists a sequence of l_m arithmetic progressions $(j_0^{(\zeta)}, \ldots, j_{K-1}^{(\zeta)}) \in E(t_{m+1}, K)$, $\zeta < l_m$, such that $j_i^{(0)}, \ldots, j_i^{(l_m-1)}$ forms an arithmetic progression of length l_m and such that for all $\xi < l$ and $\zeta < l_m$,

$$\bigcup_{i'<K} X_{i'}^{(\xi)}, j_{i'}^{(\zeta)} \in D^i(t_1, \ldots, t_m, K).$$

Proof. By the R-equivalence assumptions we have

$$F(X^{(\xi)}, j, i, s) = F(X^{(\eta)}, j, i, s)$$

for all $\xi, \eta < l$, $i \leqslant s < k$, $j < t_{m+1}$.

For the sake of brevity write

$$F(X^{(\xi)}, j, i, s) = F(j, i, s), \quad f(X^{(\xi)}, j, i, s) = f(j, i, s).$$

By the definition of $G^i(t_1, \ldots, t_{m+1}, K)$, there exist for each s, $i \leqslant s < K$, sets $Z_s, \bar{Z}_s \subseteq [0, t_{m+1})$ such that

$$|Z_s| \leqslant 2K\alpha_{m+1}^K (1-\beta_{m+1}) t_{m+1}, \quad |\bar{Z}_s| \leqslant 2K\alpha_{m+1}^K (1-\beta_{m+1}) t_{m+1},$$

and

(a) $$f(j, i, s) \leqslant 2^i \alpha_{m+1}^i t_{m+1} \quad \text{if} \quad j \notin Z_s,\ j < t_{m+1},$$

(b) $$f(j, i, s) \geqslant \frac{1}{K^2}\left(\frac{\alpha_{m+1}}{2}\right)^i t_{m+1} \quad \text{if} \quad j \notin \bar{Z}_s,\ \tfrac14 t_{m+1} < j < \tfrac34 t_{m+1}.$$

Define $\bar{Z}$ by

$$\bar{Z} = \{j\colon \tfrac14 t_{m+1} < j < \tfrac34 t_{m+1} \text{ and there is } \textit{no } (j_0, \ldots, j_{K-1}) \in F(j, i, i) \text{ such that } \bigcup_{i'<K} X_{i',j_{i'}}^{(\xi)} \in D^i(t_1, \ldots, t_m, K) \text{ holds for } \textit{all } \xi < l\}.$$

Set

$$P = \bigcup_{j\in\bar{Z}} F(j, i, i), \quad p = |P|.$$

Expressed in different terms, $\bar{Z}$ is the set of all indices j with $\frac14 t_{m+1} < j < \frac34 t_{m+1}$ such that for each $(j_0, \ldots, j_{K-1}) \in F(j, i, i)$, there exists at least one $\xi < l$ so that

(18) $$\bigcup_{i'<K} X_{i',j_{i'}}^{(\xi)} \notin D^i(t_1, \ldots, t_m, K).$$

Note that by the definition of $F(j, i, i)$ we have

$$X_{i',j_{i'}}^{(\xi)} \in P(t_1, \ldots, t_m) \subseteq S(t_1, \ldots, t_m)$$

whenever $i' < i$. Therefore, (18) implies

$$X_{i',j_{i'}}^{(\xi)} \notin S(t_1, \ldots, t_m)$$

for some i' with $i \leqslant i' < K$.

By (b) we have

$$f(j, i, i) \geqslant \frac{1}{K^2}\left(\frac{\alpha_{m+1}}{2}\right)^i t_{m+1} \quad \text{if} \quad j \notin \bar{Z}_i,\ \tfrac14 t_{m+1} < j < \tfrac34 t_{m+1}.$$

Thus, for each $j \in \bar{Z} - \bar{Z}_i$,

$$f(j, i, i) = |F(j, i, i)| \geqslant \frac{1}{K^2}\left(\frac{\alpha_{m+1}}{2}\right)^i t_{m+1}.$$

Since $F(j, i, i)$ and $F(j', i, i)$ are disjoint for $j \neq j'$ and

$$|\bar{Z}_i| \leqslant 2K\alpha_{m+1}^K (1-\beta_{m+1}) t_{m+1}$$

then

(19)

$$p = |P| = \Big|\bigcup_{j\in\bar{Z}} F(j, i, i)\Big| \geqslant \big(|\bar{Z}| - 2K(1-\beta_{m+1})\alpha_{m+1}^K t_{m+1}\big)\frac{1}{K^2}\left(\frac{\alpha_{m+1}}{2}\right)^i t_{m+1}.$$

On the other hand, by the definition of P, it is clear that

$$P \subseteq \bigcup_{i\leqslant s<K} \bigcup_{\xi<l} \bigcup_{\substack{j<t_{m+1}\\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)}} F(j, i, s)$$

$$= \bigcup_{i\leqslant s<K} \bigcup_{\xi<l} \Big(\bigcup_{\substack{j<t_{m+1}\\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)\\ j\in Z_s}} F(j, i, s) \cup \bigcup_{\substack{j<t_{m+1}\\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)\\ j\notin Z_s}} F(j, i, s)\Big)$$

so that

$$p = |P| \leqslant Kl\cdot \underbrace{2K\alpha_{m+1}^{K}(1-\beta_{m+1})t_{m+1}}_{|Z_s|}\cdot \underset{(\max|F(j,i,s)|)}{t_{m+1}} +$$

(domain of s) (domain of ξ)

$$+\,kl\cdot 2^{i}\alpha_{m+1}^{i}t_{m+1}\max_{s,\xi}|\{j:\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)\ \text{and}\ j\notin Z_s\}|$$

$$\leqslant Kl\cdot 2K\alpha_{m+1}^{K}(1-\beta_{m+1})t_{m+1}^{2}+$$

$$+Kl\cdot 2^{K}\alpha_{m+1}^{K}t_{m+1}\max_{s,\xi}|\{j:\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)\}|.$$

But, by hypothesis, for all $\xi<l$,

$$X^{(\xi)}\epsilon G^{i}(t_1,\ldots,t_{m+1},K)\subseteq C(t_1,\ldots,t_{m+1},K).$$

Thus, for all $s<K$, $\xi<l$, we have $X_s^{(\xi)}\epsilon S(t_1,\ldots,t_{m+1})$. Therefore,

$$s^{m+1}(X_s^{(\xi)})\geqslant\big(\beta_{m+1}-\sqrt{\mu_{m+1}(t_{m+1})}\,\big)t_{m+1}$$

so that

$$\max_{s,\xi}|\{j:\ X_{s,j}^{(\xi)}\notin S(t_1,\ldots,t_m)\}|\leqslant\big(1-\beta_{m+1}+\sqrt{\mu_{m+1}(t_{m+1})}\,\big)t_{m+1}$$
$$\leqslant 2(1-\beta_{m+1})t_{m+1}$$

by the choice of t_{m+1} (see (F)). Hence, by (20)

$$p\leqslant Kl\alpha_{m+1}^{K}t_{m+1}^{2}(1-\beta_{m+1})(2K+2^{K+1}).$$

Combining the two inequalities for p we obtain

$$\big(|\bar Z|-2K(1-\beta_{m+1})\alpha_{m+1}^{K}t_{m+1}\big)\frac{1}{K^2}\left(\frac{\alpha_{m+1}}{2}\right)^{i}t_{m+1}$$
$$\leqslant kl\alpha_{m+1}^{K}t_{m+1}^{2}(1-\beta_{m+1})(2K+2^{K+1}),$$

$$|\bar Z|\leqslant 2K(1-\beta_{m+1})\alpha_{m+1}^{K}t_{m+1}+K^{3}l2^{i}\alpha_{m+1}^{K-i}(1-\beta_{m+1})(2K+2^{K+1})t_{m+1}$$
$$\leqslant(1-\beta_{m+1})\cdot 2K^{3}l2^{2K}t_{m+1}\leqslant(1-\beta_{m+1})\cdot 2K^{3}l_m2^{2K}t_{m+1}\leqslant\sqrt{1-\beta_{m+1}}\,t_{m+1}$$

provided

$$K^{3}2^{2K}l_m\leqslant\frac{1}{2\sqrt{1-\beta_{m+1}}}.$$

The last inequality easily holds, however, by the choice of t_{m+1} (see 5.4). Now, since $X^{(0)}\epsilon C(t_1,\ldots,t_{m+1},K)$ then $X_i^{(0)}\epsilon S(t_1,\ldots,t_{m+1})$. Thus,

$$s^{m+1}(X_i^{(0)})\geqslant\big(\beta_{m+1}-\sqrt{\mu_{m+1}(t_{m+1})}\,\big)t_{m+1}$$

and so

$$s^{m+1}\big(X_i^{(0)},(\tfrac14 t_{m+1},\tfrac34 t_{m+1})-\bar Z\big)$$
$$\geqslant\big(\beta_{m+1}-\sqrt{\mu_{m+1}(t_{m+1})}\,\big)t_{m+1}-\tfrac12 t_{m+1}-\big(\sqrt{1-\beta_{m+1}}\big)t_{m+1}$$
$$\geqslant\big(\beta_{m+1}-\tfrac12-2\sqrt{1-\beta_{m+1}}\big)t_{m+1}\ \text{by the choice of } t_{m+1}\ \text{(see (F))}$$
$$>\tfrac12(1-l_m^{-1})t_{m+1}$$

provided

$$\beta_{m+1}-\tfrac12-2\sqrt{1-\beta_{m+1}}>\tfrac12(1-l_m^{-1}).$$

This last inequality, however, follows easily from the choice of t_{m+1} (see (C)). Thus, by (E) and Fact 11, $(\frac14 t_{m+1},\frac34 t_{m+1})-\bar Z$ contains an arithmetic progression $j^{(0)},\ldots,j^{(l_m-1)}$ of length l_m. Therefore by the definition of $\bar Z$, for each $\zeta<l_m$, there exists some arithmetic progression $(j_0^{(\zeta)},\ldots$ $\ldots,j_{K-1}^{(\zeta)})\epsilon F(j^{(\zeta)},i,i)$ such that $\bigcup_{i'<K}X_{i',j_i}^{(\zeta)}(\zeta)\epsilon D^{i}(t_1,\ldots,t_m,K)$ for all $\xi<l$. This is exactly the conclusion of Lemma 6 and the proof is completed. ■

7. The main theorem. We are now almost ready to prove the main result of the paper. However, we first need some notation. For $m\leqslant m'$, suppose

$$X\epsilon B(t_1,\ldots,t_{m'},K),\qquad Y\epsilon B(t_1,\ldots,t_m,K).$$

We write $Y[X$ if for all $i<K$, Y_i is a subconfiguration of X_i of order m. If $Y[X$ and $Y'[X'$, we shall say *the position of Y in X is the same as the position of Y' in X'*, if for each $i<K$, there is a j_i such that Y_i is the j_ith subconfiguration of X_i and Y'_i is the j_ith subconfiguration of X'_i.

FACT 12. *For every $m>0$ and every i with $0\leqslant i<K$, there exists $h(m,i)$ such that if $m'\geqslant h(m,i)$ and*

$$X^{(\xi)}\epsilon D^{i}(t_1,\ldots,t_{m'},K),\qquad \xi<l\leqslant l_m^{2^{-i}},$$

then there exist

$$Y^{(\xi)}\epsilon G^{i}(t_1,\ldots,t_{m'},K),\qquad \xi<l,$$

such that $Y^{(\xi)}[X^{(\xi)}$ and for all $\xi,\eta<l$, the position of $Y^{(\xi)}$ in $X^{(\xi)}$ is the same as the position of $Y^{(\eta)}$ in $X^{(\eta)}$.

Proof. The proof will proceed by induction on i.

$i=0$: We have already seen that

$$D^{0}(t_1,\ldots,t_m,K)=C(t_1,\ldots,t_m,K),$$

$$G^{0}(t_1,\ldots,t_m,K)=C(t_1,\ldots,t_m,K).$$

We show that we can choose $h(m, 0) = m$. Thus, we start with $X^{(\xi)} \in C(t_1, \ldots, t_{m'}, K)$ with $\xi < l \leqslant l_m$ and $m' \geqslant m$. We want to show that there exist $Y^{(\xi)} \in C(t_1, \ldots, t_m, K)$, $\xi < l$, such that $Y^{(\xi)} [X^{(\xi)}$ and for each $\xi < l$, the position of $Y^{(\xi)}$ in $X^{(\xi)}$ is the same.

(i) If $m' = m$ we just take $Y^{(\xi)} = X^{(\xi)}$, $\xi < l$.

(ii) Assume that $m' > m$ and the assertion holds for $m'-1$. The hypothesis of R-equivalence required in Lemma 6 is vacuous here since $i = 0$. Thus, by Lemma 6, there exist

$$Z^{(\xi)} = \bigcup_{i'<K} X^{(\xi)}_{i', j_{i'}^{(0)}} \in C(t_1, \ldots, t_{m'-1}, K).$$

Since $Z_{i'}^{(\xi)} = X^{(\xi)}_{i', j_{i'}^{(0)}}$ then the position of each $Z^{(\xi)}$ in $X^{(\xi)}$ is the same (and, of course, $Z^{(\xi)} [X^{(\xi)}$). Now, apply the induction hypothesis to $Z^{(\xi)} \in C(t_1, \ldots, t_{m'-1}, K)$, $\xi < l$. By induction, there exist $Y^{(\xi)} \in C(t_1, \ldots$ $\ldots, t_m, K)$, $\xi < l$, such that $Y^{(\xi)} [Z^{(\xi)}$ and the position of each $Y^{(\xi)}$ in $Z^{(\xi)}$ is the same. Thus, $Y^{(\xi)} [X^{(\xi)}$ and the position of each $Y^{(\xi)}$ in $X^{(\xi)}$ is the same. This completes the case $i = 0$.

$i > 0$: Assume for some $i \geqslant 0$ that Fact 12 holds. We now prove it for $i+1$. Let $m^* = h(m, i)+1$. We claim we can choose $h(m, i+1) = h(m^*, i)$. So, suppose $m' \geqslant h(m^*, i)$ and

$$X^{(\xi)} \in D^{i+1}(t_1, \ldots, t_{m'}, K) \quad \text{for} \quad \xi < l \leqslant l_m^{2^{-(i+1)}}.$$

Thus, $X^{(\xi)} \in D^i(t_1, \ldots, t_{m'}, K)$. Since $l_m^{2^{-(i+1)}} \leqslant l_m^{2^{-i}}$ then we can apply the induction hypothesis to obtain configurations $Y^{(\xi)} \in G^i(t_1, \ldots, t_{m^*}, K)$, $\xi < l$, such that $Y^{(\xi)} [X^{(\xi)}$ and the position of each $Y^{(\xi)}$ in $X^{(\xi)}$ is the same. Since $X^{(\xi)} \in D^{i+1}(t_1, \ldots, t_{m'}, K)$ then $X_j^{(\xi)} \in P(t_1, \ldots, t_{m'})$ for $j \leqslant i$. Thus, by the definition of $P(t_1, \ldots, t_{m'})$, all the $X_j^{(\xi)}$, $j \leqslant i$, are R-equivalent. Since the position of each $Y^{(\xi)}$ in $X^{(\xi)}$ is the same then all the $Y_j^{(\xi)}$, $j \leqslant i$, are R-equivalent. Since $h(m, i) \geqslant m$ then $l_m \leqslant l_{h(m,i)}$ and we can apply Lemma 6 to the $Y^{(\xi)}$, $\xi < l$. The conclusion of Lemma 6 asserts that there exists a sequence of arithmetic progressions

$$(j_0^{(\zeta)}, \ldots, j_{K-1}^{(\zeta)}) \in E(t_{h(m,i)+1}, K), \quad \zeta < l_{h(m,i)},$$

such that for every $\xi < l$ and every $\zeta < l_{h(m,i)}$ we have

$$\bigcup_{i'<K} Y^{(\xi)}_{i', j_{i'}^{(\zeta)}} \in D^i(t_1, \ldots, t_{h(m,i)}, K)$$

and such that $j_i^{(0)}, \ldots, j_i^{(l_{h(m,i)}-1)}$ forms an arithmetic progression of length $l_{h(m,i)}$. Denote $\bigcup_{i'<K} Y^{(\xi)}_{i', j_{i'}^{(\zeta)}}$ by $Y(\xi, \zeta)$. Since $h(m, i) \geqslant m$ then $Y(\xi, \zeta)$ is defined for $\xi < l$, $\zeta < l_m^{2^{-(i+1)}}$.

Thus, we have

$$Y(\xi, \zeta) \in D^i(t_1, \ldots, t_{h(m,i)}, K) \quad \text{for} \quad \xi < l, \ \zeta < l_m^{2^{-(i+1)}}.$$

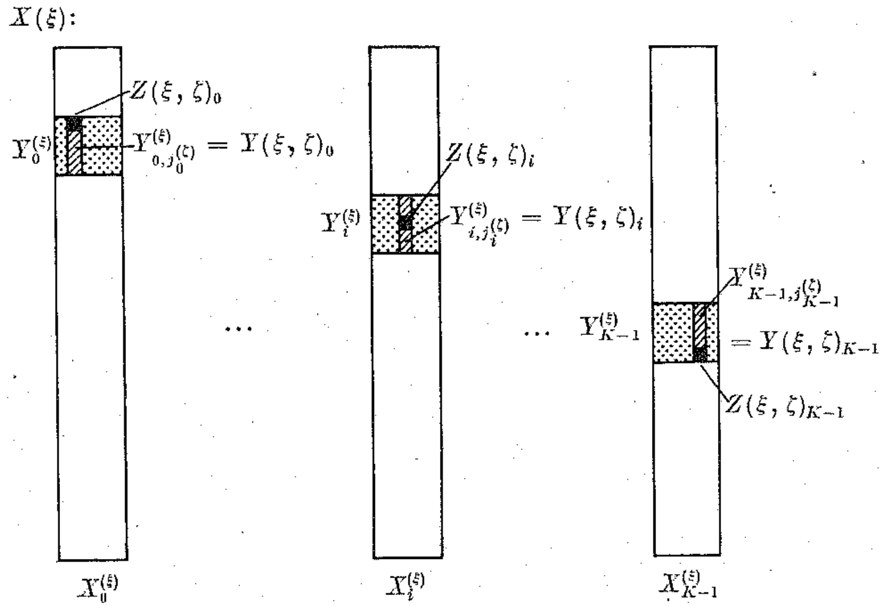
Note that there are at most

$$l \cdot l_m^{2^{-(i+1)}} \leqslant l_m^{2^{-(i+1)}} \cdot l_m^{2^{-(i+1)}} = l_m^{2^{-i}}$$

such $Y(\xi, \zeta)$. Therefore, we may now apply the induction hypothesis to these $Y(\xi, \zeta)$ and obtain a sequence

$$Z(\xi, \zeta) \in G^i(t_1, \ldots, t_m, K)$$

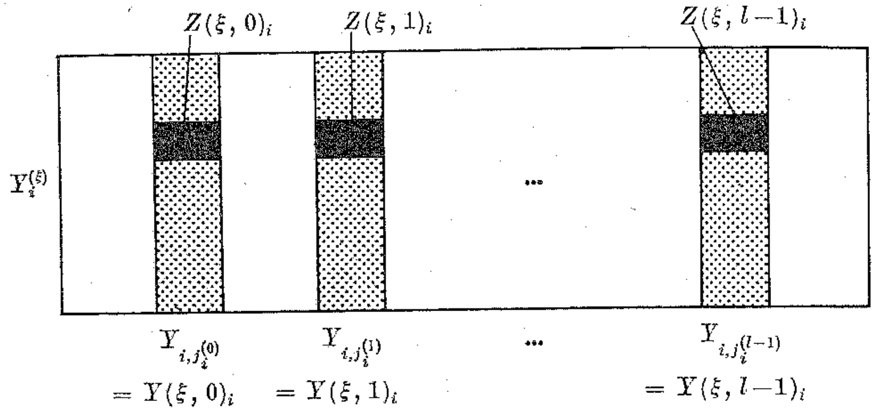
with $Z(\xi, \zeta) [Y(\xi, \zeta)$ for $\xi < l$, $\zeta < l_m^{2^{-(i+1)}}$. The following picture may help to illustrate the situation at this point.



By the way in which the $Y(\xi, \zeta)$ are defined, for fixed ζ, the position of $Y(\xi, \zeta)$ in $X(\xi)$ is the same as the position of $Y(\eta, \zeta)$ in $X(\eta)$ for ξ, $\eta < l$. Since the position of each $Z(\xi, \zeta)$ in $Y(\xi, \zeta)$ is the same (by the induction hypothesis), then for fixed ζ, the position of $Z(\xi, \zeta)$ in $Y(\xi, \zeta)$ and $X(\xi)$ is the same as the position of $Z(\eta, \zeta)$ in $Y(\eta, \zeta)$ and $X(\eta)$ for $\xi, \eta < l$, $\zeta < l_m^{2^{-(i+1)}}$. We should keep in mind that for $j \leqslant i$, all the $X_j^{(\xi)}$ are R-equivalent. Thus, for $j \leqslant i$, all the $Z(\xi, \zeta)_j$, $\xi < l$, $\zeta < l_m^{2^{-(i+1)}}$, are R-equivalent. Since $l \leqslant l_m^{2^{-(i+1)}}$ then $Z(\xi, \zeta)$ is defined for ξ, $\zeta < l$. Let

$$Z = \bigcup_{\zeta<l} Z(\xi, \zeta)_i, \quad \xi < l.$$

The following diagram indicates the current situation.



As part of the conclusion of Lemma 6, we know that $j_i^{(0)}, \ldots, j_i^{(l_{h(m,i)}-1)}$ forms an arithmetic progression of length $l_{h(m,i)}$. Hence, $j_i^{(0)}, \ldots, j_i^{(l-1)}$ forms an arithmetic progression of length l. Also, by the induction hypothesis, the position of each $Z(\xi, \zeta)$ in $Y(\xi, \zeta)$ is the same. Therefore,

$$Z = \bigcup_{\zeta<l} Z(\xi, \zeta)_i \in B(t_1, \ldots, t_m, l).$$

But each $Z(\xi, \zeta) \in G^i(t_1, \ldots, t_m, K)$ so that each $Z(\xi, \zeta)_i \in S(t_1, \ldots, t_m)$. Hence, $Z \in C(t_1, \ldots, t_m, l)$. Finally, since each $Y(\xi, \zeta) \in D^i(t_1, \ldots, t_{h(m,i)}, K)$ then for $j < i$, $Y(\xi, \zeta)_j \in P(t_1, \ldots, t_{h(m,i)})$. Thus, for all $j < i$ and all $\zeta < l_m^{2^{-(i+1)}}$, the $Y(0, \zeta)_j$ are R-equivalent. But for $\zeta < l_m^{2^{-(i+1)}}$, the position of each $Z(0, \zeta)$ in $Y(0, \zeta)$ is the same. Therefore, for all $j < i$ and all $\zeta < l_m^{2^{-(i+1)}}$, the $Z(0, \zeta)_j$ are R-equivalent. Hence, since $l_m^{2^{-K}} \leqslant l_m^{2^{-(i+1)}}$, we may apply Lemma 5 to the sequence $Z(0, \zeta)$, $\zeta < l_m^{2^{-K}}$. The conclusion of Lemma 5 then asserts that for some $\zeta^* < l^{2^{-K}}$,

$$Z(0, \zeta^*) \in G^{i+1}(t_1, \ldots, t_m, K).$$

But we have already seen that for all $\xi < l$, the position of $Z(\xi, \zeta^*)$ in $X^{(\xi)}$ is the same as the position of $Z(0, \zeta^*)$ in $X^{(0)}$. Since all the $X_j^{(\xi)}$, $j \leqslant i$, $\xi < l$, are R-equivalent (they are all in $P(t_1, \ldots, t_{m'})$) then, for each $j \leqslant i$, all the $Z(\xi, \zeta^*)_j$, $\xi < l$, are R-equivalent to $Z(0, \zeta^*)_j$. Therefore, for $\xi < l$,

$$Z(\xi, \zeta^*) \in G^{i+1}(t_1, \ldots, t_m, K).$$

Finally, since

$$Z(\xi, \zeta^*) [Y(\xi, \zeta^*) [X^{(\xi)}, \quad \xi < l,$$

then the induction step is completed. This proves Fact 12.

THEOREM. *For $m > 0$ and $0 \leqslant i < K$,*

$$D^i(t_1, \ldots, t_m, K) \neq \emptyset.$$

Proof. The proof will proceed by induction on i.

$i = 0$: We first recall that

$$D^0(t_1, \ldots, t_m, K) = C(t_1, \ldots, t_m, K).$$

For $m = 0$, $C(K)$ is certainly nonempty since by definition,

$$C(K) = \{X \in B(K): s^1(X) = K\} = B(K).$$

For $m \geqslant 1$, t_m has been chosen (see (C)) so that

$$1 - \beta_{m+1} < 1/K.$$

Hence, by Fact 8, $C(t_1, \ldots, t_m, K) \neq \emptyset$.

Assume now that the assertion holds for a fixed $i < K-1$ and all $m \geqslant 0$. We prove that it also holds for $i+1$ and all $m \geqslant 0$. Let

$$m' = h(m+1, i) \quad \text{and} \quad m'' = h(m'+1, i).$$

Let $X \in D^i(t_1, \ldots, t_{m''}, K)$ which by the induction hypothesis is nonempty. By the definition of h, there exists $Y \in G^i(t_1, \ldots, t_{m'+1}, K)$ with $Y [X$. We now apply Lemma 6 with $l = 1$, i.e., to the "sequence" consisting of a single term Y, where as usual we write

$$Y_{i'} = \bigcup_{j'<t_{m'+1}} Y_{i',j'}, \quad i' < K.$$

By the conclusion of Lemma 6, there exists a sequence of arithmetic progressions

$$(j_0^{(\zeta)}, \ldots, j_{K-1}^{(\zeta)}) \in E(t_{m'+1}, K), \quad \zeta < l_{m'},$$

such that

$$\bigcup_{i'<K} Y_{i',j_{i'}^{(\zeta)}} \in D^i(t_1, \ldots, t_{m'}, K), \quad \zeta < l_{m'}$$

and such that $j_i^{(0)}, \ldots, j_i^{(l_{m'}-1)}$ forms an arithmetic progression of length $l_{m'}$. Let

$$Y^{(\zeta)} = \bigcup_{i'<K} Y_{i',j_{i'}^{(\zeta)}}.$$

Thus, $Y^{(\zeta)} \in D^i(t_1, \ldots, t_{m'}, K)$, $\zeta < l_{m'}$. Since $m+1 \leqslant m'$, then $l_{m+1}^{2^{-i}} \leqslant l_{m'}$ and so we may apply the definition of $h(m+1, i) = m'$ to the sequence $Y^{(\zeta)}$, $\zeta < l_{m+1}^{2^{-i}}$. Hence, there exists a sequence $Z^{(\zeta)} \in G^i(t_1, \ldots, t_{m+1}, K)$, $\zeta < l_{m+1}^{2^{-i}}$, such that $Z^{(\zeta)} [Y^{(\zeta)}$ and the position of each $Z^{(\zeta)}$ in $Y^{(\zeta)}$ is the same. Since each $Z^{(\zeta)} \in G^i(t_1, \ldots, t_{m+1}, K)$ then $Z_i^{(\zeta)} \in S(t_1, \ldots, t_{m+1})$, $i < K$. Thus, letting $l = l_{m+1}^{2^{-(i+1)}}$, we have

$$\bigcup_{\zeta<l} Z_i^{(\zeta)} \in C(t_1, \ldots, t_{m+1}, l)$$

(since $j_i^{(0)}, \ldots, j_i^{(l)}$ forms an arithmetic progression). Since $Y^{(\zeta)} \in D^i(t_1, \ldots \ldots, t_{m'}, K)$, $\zeta < l_{m'}$, then each $Y_j^{(\zeta)}$, $j < i$, belongs to $P(t_1, \ldots, t_{m'})$ and consequently all the $Y_j^{(\zeta)}$, $j < i$, $\zeta < l_{m'}$, are R-equivalent. Since the position of each $Z^{(\zeta)}$ in $Y^{(\zeta)}$ is the same, then for each $j < i$, $Z_j^{(\zeta)}$ is R-equivalent to $Z_j^{(\zeta')}$, $\zeta, \zeta' < l_{m'}$. But $i+1 < K$ so that $l \geqslant l_{m+1}^{2^{-K}}$. Thus, the hypotheses of Lemma 5 are satisfied for the sequence $Z^{(\zeta)}$, $\zeta < l$. Therefore, by the conclusion of Lemma 5, there exists $\xi^* < l$ such that $Z^{(\xi^*)} \in G^{i+1}(t_1, \ldots \ldots, t_{m+1}, K)$. Finally, we apply Lemma 6 to the single configuration $Z^{(\xi^*)}$ (in this case the R-equivalence condition is trivial). The conclusion of Lemma 6 immediately implies

$$D^{i+1}(t_1, \ldots, t_m, K) \neq \emptyset.$$

This completes the induction step and the theorem is proved. ∎

It remains to verify that the theorem does in fact show that R contains arbitrarily long arithmetic progressions. By the theorem, there must exist $X \in D^{K-1}(K)$. By the definition of $D^{K-1}(K)$, this just means that X is an arithmetic progression in which the first $K-1$ terms belong to $P(\emptyset) = R$. Since K was chosen arbitrarily at the beginning of the proof then R does indeed contain arbitrarily long arithmetic progressions.

Finally, the following corollary shows that $c_k = 0$ for all k.

COROLLARY. *For all $\varepsilon > 0$ and k, there exists $N(k, \varepsilon)$ such that if $n > N(k, \varepsilon)$ and $R \subseteq [0, n)$ with $|R| > \varepsilon n$ then R contains an arithmetic progression with k terms.*

Proof. Suppose the Corollary is false. Then there exists $n_1 < n_2 < \ldots$ and $R_i \subseteq [0, n_i)$ with $|R_i| > \varepsilon n_i$, $i \geqslant 1$, such that R_i contains no k-term arithmetic progression. Let $n_1' < n_2' < \ldots$ be a subsequence of the n_i satisfying $n_{i+1}' \geqslant 3n_i'$ for all i. Form the infinite set

$$R' = \bigcup_{i>0} (R_{n_i'} + d_i)$$

where $d_i = \sum_{j<i} n_j'$. Thus,

$$\overline{\lim_{n\to\infty}} \frac{|R' \cap [0, n)|}{n} \geqslant \varepsilon.$$

By the Theorem, R' must contain an arithmetic progression with $3k$ terms, say, $A = \{a + d_i\colon 0 \leqslant i < 3k\}$. Let l satisfy $d_l \leqslant a + (3k-1)d < d_{l+1}$. Since $R_{n_l'} + d_l$ can contain at most $k-1$ terms of A then we have $a + 2kd < d_l$. But it follows from the definition of the n_i and d_i that $d_l \geqslant 3d_{l-1}$. Thus

$$a + kd \geqslant \tfrac{1}{3}\big(a + (3k-1)d\big) > \tfrac{1}{3} d_l \geqslant d_{l-1}.$$

However, this implies

$$|A \cap (R_{n_{l-1}'} + d_{l-1})| > k$$

which is *impossible*. ∎

8. Acknowledgment. My indebtedness to my friends R. L. Graham and A. Hajnal is extremely great. In fact, they wrote the whole paper after listening to my rough oral exposition.

References

[1] F. A. Behrend, *On sequences of integers containing no arithmetic progression*, Časopis Mat. Fys. Praha, 67 (1938), pp. 235–239.

[2] — *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. USA 32 (1946), pp. 331–332.

[3] E. R. Berlekamp, *A construction for partitions which avoid long arithmetic progressions*, Canad. Math. Bull. 11 (1968), pp. 409–415.

[4] P. Erdös and P. Turán, *On some sequences of integers*, J. London Math. Soc. 11 (1936), pp. 261–264.

[5] P. Erdös, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer and E. G. Straus, *Euclidean Ramsey Theorems I*, J. Comb. Th. 14 (1973), pp. 341–363.

[6] — — *Euclidean Ramsey Theorems II, III*, (to appear in Proceedings of Symposium on Infinite and Finite Sets, Keszthély (1973)).

[7] R. L. Graham and B. L. Rothschild, *A short proof of van der Waerden's theorem on arithmetic progressions*, Proc. Amer. Math. Soc. (1974). (To appear.)

[8] L. Moser, *On non-averaging sets of integers*, Canad. J. Math. 5 (1953), pp. 245–252.

[9] R. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh, Section A, 65 (1960/61), pp. 332–344.

[10] K. F. Roth, *On certain sets of integers (II)*, J. London Math. Soc. 29 (1954), pp. 20–26.

[11] — *Irregularities of sequences relative to arithmetic progressions (III)*, J. Number Theory 2 (1970), pp. 125–142.

[12] — *Irregularities of sequences relative to arithmetic progressions (IV)*, Per. Math. Hungar. 2 (1972), pp. 301–326.

[13] R. Salem and C. S. Spencer, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. USA 28 (1942), pp. 561–563.

[14] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. 20 (1969), pp. 89–104.

[15] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw. Arch. Wisk. 15 (1928), pp. 212–216.

Received on 1. 7. 1973 and in revised form 1. 2. 1974 (437)