

On Sets of Integers Containing No k Elements in Arithmetic Progression

E. Szemerédi

In 1926 van der Waerden [13] 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 classes, then at least one class contains an arithmetic progression of n terms. (For a short proof, see the note of Graham and Rothschild [5].) 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 $f(n) < n2^n$, for n prime, which improves previous results of Erdős, Rado and W. Schmidt.

More than 40 years ago, Erdős and Turán [4] considered the quantity $r_k(n)$, defined to be the greatest integer l for which there is a sequence of integers $0 < a_1 < a_2 < \dots < a_l \leq n$ 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 arbitrarily 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.

Erdős and Turán observed $r_k(m+n) \leq r_k(m) + r_k(n)$ from which it follows by a simple argument that

$$\lim_{n \rightarrow \infty} r_k(n)/n = c_k$$

exists. Erdős 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 $\lim_{k \rightarrow \infty} c_k = 1$. Erdős and Turán also conjectured $r_k(n) < n^{1-\epsilon_k}$, which was shown to be false by Salem and Spencer [11] who proved $r_3(n) > n^{1-c/\log \log n}$. In 1946 Behrend [2] proved $r_3(n) > n^{1-c/(\log n)^{1/2}}$ which is the best lower bound for $r_3(n)$ currently known. In [6], 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 [7].

The first satisfactory upper bound for $r^3(n)$ was due to Roth [8] who proved $r_3(n) < cn/\log \log n$. In 1967, I proved that $r_4(n) = o(n)$. The proof used the general theorem of van der Waerden. Roth [9], [10] 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 $r_4(n) < n/\log_l n$ where l is a large fixed integer and \log_l denotes the l -fold iterated logarithm.

In this article we give a brief outline of a proof of the general conjecture of Erdős and Turán: $c_k = 0$ for all k .

The proof is rather long and complicated although it uses only elementary combinatorial arguments. Space limitations do not permit us to outline the proof here so we shall just restrict ourselves to mentioning several of the key ideas.

An important lemma used in the proof states essentially that any finite graph can be partitioned into relatively few "nearly regular" subgraphs. The basic objects with which the proof of the main theorem deals are not just arithmetic progressions themselves but rather generalizations of arithmetic progressions called *m-configurations*. Roughly speaking, a 1-configuration is just an arithmetic progression; an m -configuration is an "arithmetic progression" of $(m-1)$ -configurations. In a nutshell, one can show that for any given set of integers R of positive upper density, a very long m -configuration which intersects R in a moderately regular way must always contain a shorter (but still quite long) $(m-1)$ -configuration which intersects R in an even more regular way. In this way, we eventually conclude that R must contain arbitrarily long 1-configurations, i.e., arithmetic progressions, and we are done.

References

1. F. A. Behrend, *On sequences of integers containing no arithmetic progression*, Casopis Mat. Fys. Praha **67** (1938), 235-239.
2. ———, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331-332. MR **8**, 317.

3. E. R. Berlekamp, *A construction for partitions which avoid long arithmetic progressions*, *Canad. Math. Bull.* **11** (1968), 409–414. MR **38** #1066.
4. P. Erdős and P. Turán, *On some sequences of integers*, *J. London Math. Soc.* **11** (1936), 261–264.
5. R. L. Graham and B. L. Rothschild, *A short proof of van der Waerden's theorem on arithmetic progressions*, *Proc. Amer. Math. Soc.* **42** (1974), 385–386.
6. L. Moser, *On non-averaging sets of integers*, *Canad. J. Math.* **5** (1953), 245–252. MR **14**, 726, 1278.
7. R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, *Proc. Roy. Soc. Edinburgh Sect. A* **65** (1960/61), 332–344. MR **26** #95.
8. K. F. Roth, *On certain sets of integers. II*, *J. London Math. Soc.* **29** (1954), 20–26. MR **15**, 288.
9. ———, *Irregularities of sequences relative to arithmetic progressions. III*, *J. Number Theory* **2** (1970), 125–142. MR **41** #6810.
10. ———, *Irregularities of sequences relative to arithmetic progressions. IV*, *Per. Math. Hungar.* **2** (1972), 301–326.
11. R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetic progression*, *Proc. Nat. Acad. Sci. U.S.A.* **28** (1942), 561–563. MR **4**, 131.
12. E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, *Acta. Math. Acad. Sci. Hungar.* **20** (1969), 89–104. MR **39** #6861.
13. B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, *Nieuw. Arch. Wisk.* **15** (1928), 212–216.

HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY

STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305, U.S.A.

Section 19

Applied Statistics, Mathematics in the Social and Biological Sciences

