Some Quantitative Aspects of Szemerédi's Theorem Modulo n

T. C. Brown

Citation data: T.C. Brown, *Some quantitative aspects of Szemerédi's theorem modulo n*, congressus Numerantium **43** (1984), 169–174.

Abstract

The multiset $P = \{a_1, \ldots, a_k\}$ is a k-term arithmetic progression modulo n if $a_1 \not\equiv a_2 \pmod{n}$ and $a_2 - a_1 \equiv a_3 - a_2 \equiv \cdots \equiv a_k - a_{k-1} \pmod{n}$. For k odd and $k \geq 3$, we find explicit constnats $\varepsilon_k < 1 - 1/k$ such that for any $n \neq k$ and for any subset A of $[0, n-1]$, if $|A| > \varepsilon_k n$ then A contains a k-term arithmetic progression modulo n. ($\varepsilon_3 = .5$ and ε_5 is about .77.)

1 Introduction

For each real number $\varepsilon > 0$ and positive integers k and n_0, let $S(\varepsilon, k, n_0)$ denote the following statement.

$S(\varepsilon, k, n_0)$: For every $n \geq n_0$, and for every subset A of $[0, n-1]$, if $|A| > \varepsilon n$ then A contains a k-term arithmetic progression.

Then Szemerédi's theorem [2] asserts that for every $\varepsilon > 0$ and k, there exists a least positive integer $n_0 = n_0(\varepsilon, k)$ such that $S(\varepsilon, k, n_0)$ holds.

One can ask the following quantitative questions. (Answering them, of course, is something else!)

(a) Given $\varepsilon > 0$ and k, what is $n_0(\varepsilon, k)$, that is, what is the smallest n_0 such that $S(\varepsilon, k, n_0)$ holds?

(b) Given k and n_0, what is the smallest ε such that $S(\varepsilon, k, n_0)$ holds? (We may denote this smallest ε by $\varepsilon(k, n_0)$.)

These questions appear to be simplified if for a given n and a given subset A of $[0, n-1]$ we enlarge the set of arithmetic progressions under consideration. Thus we say that A contains a *k-term arithmetic progression modulo n* if A contains elements $a_0, \ldots, a_{k-1}$ (not necessarily distinct) such that

$$a_j \equiv a_0 + jd \pmod{n}, \quad 0 \leq j \leq k-1,$$

for some integer d with

$$d \not\equiv 0 (\bmod\, n).$$

We can replace statement $S(\varepsilon, k, n_0)$ by the corresponding statement $M(\varepsilon', k, n_0')$, for any real number $\varepsilon' > 0$ and positive integers k and n_0', as follows.

$M(\varepsilon',k,n_0')$: For every $n \geq n_0'$, and for every subset A of $[0,n-1]$, if $|A| > \varepsilon' n$ then A contains a k-term arithmetic progression modulo n.

One can then ask the following questions.

(a') Given $\varepsilon' > 0$ and k, what is $n_0'(\varepsilon,k)$, the smallest n_0' such that $M(\varepsilon',k,n_0')$ holds?

(b') Given k and n_0', what is $\varepsilon'(k,n_0')$, the smallest ε' such that $M(\varepsilon',k,n_0')$ holds?

In this note we obtain bounds what appear to be the easiest cases of these latter two questions. Given a small $\varepsilon > 0$ (namely $\varepsilon \leq \frac{1}{2}$) and arbitrary k, we find a lower bound for $n_0'(\varepsilon,k)$. (Theorem 1 below). Given a small n_0' (namely $n_0' = k+1$) and arbitrary *odd* k, we find an upper bound for $\varepsilon'(k,n_0')$. (Theorem 2 below).

Remark 1. It has been observed in [1] that Szemerédi's theorem is equivalent to the following statemtn: For every $\varepsilon' > 0$ and k, there exists a least positive integer n_0' such that $M(\varepsilon',k,n_0')$ holds.

(In fact,

$$n_0'(\varepsilon,k) \leq n_0(\varepsilon,k) \leq \frac{1}{2}n_0'(\varepsilon/2,k) + \frac{1}{2}.$$

To obtain the second inequality, let $2m \geq n_0'(\varepsilon/2,k)$, and let A be any subset of $[0,m-1]$ such that $|A| > \varepsilon m = (\varepsilon/2)(2m)$. Then regarding A as a subset of $[0,2m-1]$ it follows from the choice of $2m$ that A contains a k-term arithmetic progression modulo $2m$. Since A is a subset of $[0,m-1]$, this k-term arithmetic progression modulo $2m$ is in fact a k-term arithmetic progression. Hence $n_0(\varepsilon,k) \leq m$.)

Remark 2. It is trivial that for any k and n_0', $\varepsilon'(k,n_0') \leq 1-1/k$.

(For if $A \subset [0,n-1]$ and $|A| > (1-1/k)n$, then the average value of $|A \cap [i,i+k-1]|$ is greater than $1-1/k$, hence for some i, A contains $i, i+1, \ldots, i+k-1$ (modulo n). Note, however, that this argument fails for $\varepsilon(k,n_0)$: $A = \{0,1,3\} \subset [0,3]$ and $|A| > (1-1/3)\cdot 4$, but A contains no 3-term arithmetic progression.)

2 Results

From now on, we abbreviate "k-term arithmetic progression" to "k-progression".

Theorem 1. *For $s \geq 2, k \geq 3$,*

$$n_0'(1/s,k) > \sqrt{2}s^{k/2} - 2s + 1. \tag{1}$$

Proof. Fix $s \geq 2, k \geq 3$, and consider the $(m+1)$-element subsets of $[0,ms]$. Note that $m+1 > (1/s)(ms+1)$, so that if one of these subsets contains no k-progression modulo $ms+1$, then $n_0'(1/s,k) > ms+1$.

Given a fixed k-progression P (modulo $ms+1$) in $[0,ms]$, the number of $(m+1)$-element subsets of $[0,ms]$ which contain P is at most $\binom{ms+1-k}{m+1-k}$. The total number of distinct k-progressions P (modulo $ms+1$) in $[0,ms]$ is at most $(ms+1)(ms)/2$. Therefore

$$\binom{ms+1-k}{m+1-k}(ms+1)(ms)/2 < \binom{ms+1}{m+1} \tag{2}$$

implies

$$n_0'(1/s,k) > ms+1. \tag{3}$$

When $m+1 \geq k$, (2) is equivalent to

$$m(m+1) < 2 \cdot \left(\frac{ms-1}{m-1}\right)\left(\frac{ms-2}{m-2}\right) \cdot \left(\frac{ms-k+2}{m-k+2}\right), \tag{4}$$

and each factor on the right hand side of (4) is greater than s. Therefore when $m+1 \geq k$, (2) holnds provided $m(m+1) \leq 2 \cdot s^{k-2}$, which in turn holds provided $(m+1)^2 \leq 2 \cdot s^{k-2}$, or

$$m \leq \sqrt{2}s^{k/2-1} - 1. \tag{5}$$

Now when $k \leq \sqrt{2}s^{k/2-1}$, we can find an integer m such that $k \leq m+1 \leq \sqrt{2}s^{k/2-1}$ and $m > \sqrt{2}s^{k/2-1} - 2$, which gives (1).

Only a small number of pairs (s,k) have $k > \sqrt{2}s^{k/2-1}$ (namely $(s,k) = (2,3)$, $(2,4)$, $(2,5)$, $(2,6)$, $(3,3)$, $(4,3)$), and these can be checked separately, giving (1) in all cases. □

Theorem 2. *Define the numbers ε_k, for odd $k \geq 3$, as follows. Let $\varepsilon_3 = 1/2$. For $k = 2m+1$, $m \geq 2$, let*

$$\varepsilon_k = 1 - \frac{k+1}{k+2}\left(\sqrt{m^2 + \frac{k+2}{k+1}} - m\right). \tag{6}$$

Then $\varepsilon_k < 1 - 1/k$, and for every $n \neq k$ and every subset A of $[0, n-1]$, if $|A| > \varepsilon_k n$ then A contains a k-progression modulo n.

Lemma 1. *In proving Theorem 2, we may assume that $n > k$.*

Proof of Lemma 1. For $k = 3$, the assertion of the lemma is obviously true. For $k > 3$, one can check that $\varepsilon_k > 1 - 1/(k-1)$. From this it follows that if $n < k$ and A is any subset of $[0, n-1]$ such that $|A| > \varepsilon_k n$, then $A = [0, n-1]$ and hence A contains a k-progression modulo n. □

Lemma 2. *In proving Theorem 2, we may assume that n is prime.*

Proof of Lemma 2. Assume that if p is prime, $A \subset [0, p-1]$, $|A| > \varepsilon_k p$, then A contains a k-progression modulo p. Now let n be arbitrary, let $A \subset [0, n-1]$, $|A| > \varepsilon_k n$, and let p be a prime divisor of n. Identify $[0, n-1]$ with the cyclic group Z_n. Then Z_n contains a copy H of Z_p, and for some coset $a + H$ of H, $|A \cap (a+H)| > \varepsilon_k H$, or

$$|(A-a) \cap H| > \varepsilon_k p. \tag{7}$$

Therefore $A - a$ contains a k-progression as a subset of H; since H is a subgroup of Z_n, this k-progression is a k-progression as a subset of Z_n. □

Remark. The same argument shows that in Theorem 2, Z_n can be replaced by an arbitrary abelian group, except for $Z_p \times \cdots \times Z_p$ when $k = p =$prime. In particular, Theorem 2 is true even for $n = k$, provided k is not prime.

Proof of Theorem 2. Case 1. The case $k = 3$. Let p be prime, $p > 3$, $A \subset [0, p-1]$, $|A| = \alpha p$, and assume that A contains no 3-progression modulo p. We need to show that $\alpha \leq 1/2$.

For each pair $x, x+y$ $(y \neq 0)$ of elements of A, the (distinct) elements $w_1 = x - y$, $w_2 = x + 2y$ are excluded from A, since A contains no 3-progression modulo p. (All arithmetic operations here are modulo p.)

Also, given distinct elements w_1, w_2 in $[0, p-1]$, there are unique x, y $(y \neq 0)$ in $[0, p-1]$ such that $x - y = w_1$ and $x + 2y = w_2$.

It easily follows that each excluded pair $\{w_1, w_2\}$ is excluded only once, so that the $\binom{\alpha p}{2}$ pairs of elements of A exclude $\binom{\alpha p}{2}$ distinct pairs $\{w_1, w_2\}$ from A. The union of these $\binom{\alpha p}{2}$ distinct pairs of elements has at least αp elements.

Thus $\alpha p = |A| \leq p - \alpha p$, and $\alpha \leq 1/2$, as required. □

Case 2. The case $k > 3$. From now on, for convenience, we abbreviate "k-progression modulo p" to "k-progression".

Let $k = 2m+1$, $m \geq 2$. Let p be prime, $p > k$, $A \subset [0, p-1]$, $|A| = \alpha p$, and assume that A contains no k-progression.

We need to show that $\alpha \leq \varepsilon_k$. (One can check directly that $\varepsilon_k < 1 - 1/k$. ε_5 is about 0.77.)

The argument proceeds essentially as in the case $k = 3$:

Each $(k-1)$-progression contained in A eliminates a pair $\{w_1, w_2\}$ of elements from A, and each eliminated pair $\{w_1, w_2\}$ is eliminated exactly once.

Let t be the number of $(k-1)$-progressions contained in A. Then the union of the t excluded pairs $\{w_1, w_2\}$ has at least w elements, where w is the smallest integer such that $\binom{w}{2} \geq t$. Then $w > \sqrt{2t}$, so that $\alpha p = |A| < p - \sqrt{2t}$, or

$$(1-\alpha)^2 p^2 > 2t. \tag{8}$$

Now we estimate t from below. The set $[0, p-1] - A$ contains $(1-\alpha)p$ elements, and each of these belong to exactly $m(p-1)$ $(k-1)$-progressions. Thus $[0, p-1] - A$ meets at most $(1-\alpha)pm(p-1)$ $(k-1)$-progressions. Since the total number of $(k-1)$-progressions contained in $[0, p-1]$ is exactly $p(p-1)/2$, it follows that

$$t \geq p(p-1)/2 - p(p-1)(1-\alpha)m,$$

or

$$2t \geq p(p-1)(1 - (1-\alpha)2m). \tag{9}$$

Combining (9) and (8) gives

$$\frac{(1-\alpha)^2}{1 - (1-\alpha)2m} > 1 - 1/p. \tag{10}$$

Since $p \geq k + 2 = 2m + 3$, this gives

$$\frac{(1-\alpha)^2}{1 - (1-\alpha)2m} > 1 - \frac{1}{2m+3}. \tag{11}$$

Using $\alpha \leq 1$, it follows from (11) that $\alpha \leq \varepsilon_k$ as required. □

□

(When k is even, all of the above remains valid except for $\varepsilon_k < 1 - 1/k$. Hence, according to Remark 2 above, the application of this method for even k gives no result. Perhaps some modified version of this method will work for even k.)

References

[1] T.C. Brown and J.P. Buhler, *Lines imply spaces in density Ramsey theory*, J. Combin. Theory Ser. A **36** (1984), 214–220.

[2] E. Szemerédi, *On sets of integers containing no k elements in an arithmetic progression*, Acta. Arith. **27** (1975), 199–245, Collection of articles in memory of Jurii Vladimirovic Linnik.