APPLICATION OF A METHOD OF SZEMEREDI

by H. HALBERSTAM

To Robert Rankin on his 70th Birthday

1. Let $\mathcal{B} = \{b_i : b_1 < b_2 < \ldots\}$ be an infinite sequence of positive integers that exceed 1 and are pairwise coprime, so that

$$(b_i, b_j) = 1, \qquad i \neq j. \tag{1.1}$$

Assume also that

$$\sum_{i=1}^{\infty} \frac{1}{b_i} < \infty. \tag{1.2}$$

Let $\mathcal{A} = \mathcal{A}_{\mathcal{B}}$ denote the sequence of $\mathcal{B}$-free numbers, that is, of positive integers divisible by *no* element of $\mathcal{B}$. This concept, generalizing square-free and k-free numbers, derives from Erdös [**2**] who proved in 1966 that there exists a constant $c, 0 < c < 1$, independent of $\mathcal{B}$, such that the interval $(x, x + x^c)$ contains elements of $\mathcal{A}$ provided only that x is large enough. This result of Erdös was shown by Szemeredi [**7**] in 1973 to hold with $c = \frac{1}{2} + \varepsilon$, if $x \geq x_0(\varepsilon, \mathcal{B})$, and quite recently Bantle and Grupp [**1**] have sharpened Szemeredi's result to $c = 9/20 + \varepsilon$.

The purpose of this note is to show how the method of Szemeredi can be used to derive, virtually without change, the following result.

THEOREM. *Let k be a positive integer and let h satisfy $1 \leq h < k$, $(h, k) = 1$. Given $\delta > 0$, there exists a $\mathcal{B}$-free number a such that*

$$a \equiv h \bmod k, \qquad a \leq k^{2+\delta}$$

provided only that $k \geq k_0(\varepsilon, \mathcal{B})$.

Define, as (1.2) permits us to do,

$$\beta = \prod_{i=1}^{\infty} \left(1 - \frac{1}{b_i}\right), \tag{1.3}$$

and denote by s the least positive integer so that

$$\sum_{i=s+1}^{\infty} \frac{1}{b_i} < \frac{1}{100} \varepsilon\beta. \tag{1.4}$$

It is easy to see that, without any loss of generality, one may assume $b_1, \ldots, b_s$ to be prime. We shall use the letters p and q, with or without suffices, exclusively to denote primes.

We shall prove the theorem by showing that, actually, there exist at least $(1/20)\varepsilon\beta k^{1+\varepsilon}$ $\mathcal{B}$-free numbers $a \leq k^{2+\varepsilon}$ in the arithmetic progression $h \bmod k$, provided that k is large enough.

Glasgow Math. J. **27** (1985) 81–85.

2. Proof of the theorem. The natural approach would be to develop an argument to show that the set $\{n:1\leqslant n\leqslant k^{2+\varepsilon}, n\equiv h \bmod k\}$ contains elements of $\mathscr{A}$ if k is large enough. Instead, following Szemeredi, we narrow attention at the outset to a subset of these integers each having a large prime factor. It turns out that this surrender of advantage is more than compensated by the increased difficulty of having such integers divisible by large elements of $\mathscr{B}$.

Accordingly, let

$$\mathscr{P}=\{p:2k^{1+\varepsilon/2}<p<k^{1+\varepsilon},\ p\notin\mathscr{B},\ p\nmid k\}, \tag{2.1}$$

and focus on the set of integers

$$\mathscr{C}=\{n:1\leqslant n\leqslant k^{2+\varepsilon},\ n\equiv h \bmod k,\ n \text{ divisible by a prime of } \mathscr{P}\}.$$

If an integer n in $\mathscr{C}$ were to have two prime divisors from $\mathscr{P}$ we should have $k^{2+\varepsilon}\geqslant n>(2k^{1+\varepsilon/2})^2$, a contradiction. Hence $\mathscr{P}$ induces the partition

$$\mathscr{C}=\bigcup_{p\in\mathscr{P}}\mathscr{C}^{(p)}, \tag{2.2}$$

where

$$\mathscr{C}^{(p)}:=\{n:1\leqslant n\leqslant k^{2+\varepsilon},\ n\equiv h \bmod k,\ n\equiv 0 \bmod p\}. \tag{2.3}$$

Moreover, if $\mathscr{C}_1$ now denotes the number of elements of $\mathscr{C}$ divisible by none of $b_1,\ldots,b_s$, then the cardinality $|\mathscr{C}_1|$ of $\mathscr{C}_1$ is given by

$$|\mathscr{C}_1|=\sum_{p\in\mathscr{P}}|\mathscr{C}_1^{(p)}|=\sum_{p\in\mathscr{P}}|\{m:1\leqslant m\leqslant k^{2+\varepsilon}p^{-1},\ m\equiv hp' \bmod k,\ (m,b_1\ldots b_s)=1\}| \tag{2.4}$$

where the interpretation of $\mathscr{C}_1^{(p)}$ is obvious and $p'=p'(k)$ is the inverse of p modulo k, i.e. $pp'\equiv 1 \bmod k$.

LEMMA 1. *If k is sufficiently large,*

$$|\mathscr{C}_1|\geqslant\frac{\varepsilon}{10}\beta k^{1+\varepsilon}.$$

Proof. By (2.4) and the definition of $\mathscr{P}$, which guarantees that $(p,b_1\ldots b_s)=1$ when $p\in\mathscr{P}$, we have

$$\begin{aligned}
|\mathscr{C}_1^{(p)}| &= \sum_{\substack{d\mid b_1\ldots b_s\\(d,k)=1}}\mu(d)\sum_{\substack{1\leqslant m\leqslant k^{2+\varepsilon}/p\\ m\equiv hp' \bmod k\\ m\equiv 0 \bmod d}}1\\
&=\sum_{\substack{d\mid p_1\ldots p_s\\(d,k)=1}}\mu(d)\left(\frac{k^{1+\varepsilon}}{pd}+\theta_{p,d}\right),\quad |\theta_{p,d}|<1,\\
&\geqslant\frac{k^{1+\varepsilon}}{p}\sum_{\substack{d\mid b_1\ldots b_s\\(d,k)=1}}\frac{\mu(d)}{d}-2^s\\
&=\frac{k^{1+\varepsilon}}{p}\prod_{\substack{i=1\\ b_i\nmid k}}^{s}\left(1-\frac{1}{b_i}\right)-2^s\\
&\geqslant\frac{k^{1+\varepsilon}}{p}\beta-2^s.
\end{aligned}$$

https://doi.org/10.1017/S001708950000608X Published online by Cambridge University Press

Hence, by (2.4),

$$|\mathscr{C}_1| \geqslant \beta k^{1+\varepsilon} \sum_{p \in \mathscr{P}} \frac{1}{p} - 2^s \pi(k^{1+\varepsilon})$$
$$\geqslant k^{1+\varepsilon}\left\{\beta \sum_{p \in \mathscr{P}} \frac{1}{p} - \frac{2^{s+1}}{\log k}\right\}.$$

But

$$\sum_{p \in \mathscr{P}} \frac{1}{p} = \sum_{2k^{1+\varepsilon/2} < p < k^{1+\varepsilon}} \frac{1}{p} - \sum_{\substack{p \in \mathscr{B} \\ p > 2k^{1+\varepsilon/2}}} \frac{1}{p} - \sum_{\substack{p \mid k \\ p > 2k^{1+\varepsilon/2}}} \frac{1}{p}$$
$$\geqslant \frac{\varepsilon}{5} + O\left(\frac{1}{\log k}\right) - \sum_{i > s} \frac{1}{b_i} - k^{-1-\varepsilon/2} \omega(k)$$

if k is large enough (to ensure that $k^{1+\varepsilon/2} \geqslant b_{s+1}$), where $\omega(k)$ is the number of distinct prime factors of k; so that, by (1.4),

$$\sum_{p \in \mathscr{P}} \frac{1}{p} \geqslant \frac{\varepsilon}{5} - \frac{\varepsilon\beta}{100} + O\left(\frac{1}{\log k}\right)$$

and

$$|\mathscr{C}_1| \geqslant \beta k^{1+\varepsilon}\left\{\frac{\varepsilon}{5} - \frac{\varepsilon}{100} + O\left(\frac{1}{\log k}\right)\right\} \geqslant \frac{\beta\varepsilon}{10} k^{1+\varepsilon}$$

for all sufficiently large k, as required.

It follows from Lemma 1 that if $\mathscr{C}_0$ denotes the set of elements of $\mathscr{C}$ having no divisors from $\mathscr{B}$, then

$$|\mathscr{C}_0| \geqslant |\mathscr{C}_1| - |\mathscr{C}_2| - |\mathscr{C}_3| \geqslant \frac{\varepsilon\beta}{10} k^{1+\varepsilon} - |\mathscr{C}_2| - |\mathscr{C}_3|,$$

where $\mathscr{C}_2$ is the set of all those integers in $\mathscr{C}$ that are divisible by an element b in $\mathscr{B}$ of 'intermediate' size, i.e. one satisfying

$$b_{s+1} \leqslant b \leqslant k^{1+\varepsilon},$$

and $\mathscr{C}_3$ is the set of all integers belonging to $\mathscr{C}_1$ that have a large factor b from $\mathscr{B}$, i.e. satisfying $b > k^{1+\varepsilon}$. We shall prove that $|\mathscr{C}_2| + |\mathscr{C}_3|$ is relatively small.

Lemma 2. *We have* $|\mathscr{C}_2| \leqslant \dfrac{\varepsilon\beta}{50} k^{1+\varepsilon}$.

https://doi.org/10.1017/S001708950000608X Published online by Cambridge University Press

Proof. We argue quite crudely. We have that

$$|\mathscr{C}_2| \leq \sum_{\substack{b_{s+1} \leq b \leq k^{1+\varepsilon} \\ (b,k)=1}} \sum_{\substack{1 \leq n \leq k^{2+\varepsilon} \\ n \equiv h \bmod k \\ n \equiv 0 \bmod b}} 1$$

$$\leq \sum_{\substack{b_{s+1} \leq b \leq k^{1+\varepsilon} \\ (b,k)=1}} \left(\frac{k^{1+\varepsilon}}{b} + 1 \right)$$

$$\leq 2k^{1+\varepsilon} \sum_{b \geq b_{s+1}} \frac{1}{b} < \frac{\varepsilon\beta}{50} k^{1+\varepsilon}$$

by (1.4), and this completes the proof of the lemma.

LEMMA 3. *We have* $|\mathscr{C}_3| \leq \frac{1}{2}k^{1+\varepsilon/2}$.

Proof. Suppose n is counted in $\mathscr{C}_3$. Then n is divisible by a (unique) prime p from $\mathscr{P}$, and n is divisible also by an element $b > k^{1+\varepsilon}$ from $\mathscr{B}$. This cannot happen if $(b, p) = 1$, for then $n \geq bp > k^{1+\varepsilon} . 2k^{1+\varepsilon/2} = 2k^{2+3\varepsilon/2}$, a contradiction. Hence the b dividing n is composite and divisible by a $p > 2k^{1+\varepsilon/2}$. Writing $b = lp$, we have $1 < l < \frac{1}{2}k^{1+\varepsilon/2}$; and given such an integer l, there is, by (1.1), at most one $b \in \mathscr{B}$ divisible by l. Hence there are at most $\frac{1}{2}k^{1+\varepsilon/2}$ available choices of b. Finally, given such a b, if $1 \leq n \leq k^{2+\varepsilon}$, $n \equiv h \bmod k$ and $n \equiv 0 \bmod b$ with $b > k^{1+\varepsilon}$, there is at most one such n. This proves the lemma.

We are now able to complete the proof of the theorem. By (2.5) and Lemmas 2 and 3 we have

$$|\mathscr{C}_0| \geq \frac{\varepsilon}{10} \beta k^{1+\varepsilon} - \frac{\varepsilon}{50} \beta k^{1+\varepsilon} - \tfrac{1}{2}k^{1+\varepsilon/2} \geq \frac{\varepsilon}{20} \beta k^{1+\varepsilon}$$

if k is large enough. Thus the theorem is proved, in a quantitative form.

3. Some concluding remarks. If we replace condition (1.2) by the more demanding

$$B(x) := |\{b \in \mathscr{B} : b \leq x\}| \ll x^{\theta}, \tag{3.1}$$

where $0 < \theta < 1$, we can, with only a little more trouble, replace the exponent $2 + \varepsilon$ in the theorem by $1 + \theta + \varepsilon$. Thus when $\mathscr{B}$ is the sequence of squares of primes and $\mathscr{A}_{\mathscr{B}}$ is the sequence of squarefree numbers, we obtain the exponent $(3/2) + \varepsilon$ which is very close to the best that was known until the recent work of Heath-Brown [**3**]. While it is unlikely that one can emulate Heath-Brown's delicate argument in the more general situation, I do believe that the theorem itself can be improved a little, in the spirit of [**1**].

The condition (1.1) can be relaxed somewhat in the theorem. For such and other variations of Szemeredi's theorem see Narlikar and Ramachandra [**4**].

REFERENCES

1. G. Bantle and F. Grupp, On a problem of Erdös and Szemeredi, *J. Number Theory*, to appear.

https://doi.org/10.1017/S001708950000608X Published online by Cambridge University Press

2. P. Erdös, On the difference of consecutive terms of sequences defined by divisibility properties, *Acta Arith.* **12** (1966), 175–182.

3. D. R. Heath-Brown, The least square-free number in an arithmetic progression, *J. Reine Angew. Math.* **332** (1982), 204–220.

4. H. J. Narlikar and K. Ramachandra, Contributions to the Erdös-Szemeredi theory of sieved integers, *Acta Arith.* **38** (1980), 157–165.

5. E. Szemeredi, On the difference of consecutive terms of sequences defined by divisibility properties II, *Acta Arith.* **23** (1973), 359–361.

University of Illinois at Urbana-Champaign
Department of Mathematics
1409 West Green Street
Urbana
Illinois 61801
U.S.A.

https://doi.org/10.1017/S001708950000608X Published online by Cambridge University Press