GREEDILY PARTITIONING THE NATURAL NUMBERS INTO SETS FREE OF ARITHMETIC PROGRESSIONS

JOSEPH GERVER, JAMES PROPP AND JAMIE SIMPSON

(Communicated by Larry J. Goldstein)

ABSTRACT. We describe a "greedy" algorithm for partitioning the natural numbers into sets free of arithmetic progressions of length 3. A recursive formula governing the resulting partition is proved, and some features of its asymptotic behavior are discussed.

Introduction. In 1927, van der Waerden [12] showed that if the set of non-negative integers is partitioned into a finite number of sets, one of these sets must contain arithmetic progressions of length 3, and indeed of arbitrary length. Of course, if one instead allows an infinite number of sets in the partition, it is easy to ensure that each of the sets contains no arithmetic progression of length 3; one can for example partition $\mathbf{N} = \{0, 1, 2, \dots\}$ into singletons $\{0\}, \{1\}, \{2\}, \dots$. A natural idea for constructing a more "economical" partition of $\mathbf{N}$ into sets $S_0, S_1, S_2, \dots$ with the aforementioned property is to assign each successive integer to the set S_i with i as small as possible, subject to the constraint that no S_i may contain three integers in arithmetic progression. Thus:

$0 \in S_0$,
$1 \in S_0$,
$2 \in S_1$ (since $0, 1 \in S_0$),
$3 \in S_0$,
$4 \in S_0$,
$5 \in S_1$ (since $3, 4 \in S_0$),
$6 \in S_1$ (since $0, 3 \in S_0$),
$7 \in S_2$ (since $1, 4 \in S_0$ and $5, 6 \in S_1$),
$8 \in S_2$ (since $0, 4 \in S_0$ and $2, 5 \in S_1$),

$\dots$

This is the *greedy algorithm* for partitioning $\mathbf{N}$ into sets free of length 3 arithmetic progressions. The set S_0 constructed in this way consists (as will be shown) of those integers whose base 3 representations contain only the digits 0 and 1. This set was first considered by Szekeres (see [1]), but the sets $S_1, S_2, \dots$ appear to be new.

Throughout this paper, we will frequently represent the partition $\{S_0, S_1, S_2, \dots\}$ in terms of its "characteristic sequence" $(a_n)_0^\infty$, where

$$a_n = k \quad \text{exactly if} \quad n \in S_k.$$

That is,

$$(a_n)_0^\infty = (0, 0, 1, 0, 0, 1, 1, 2, 2, \ldots).$$

It will prove convenient to pass back and forth between the partition-representation and the sequence-representation of the greedy construction.

The major result proved in this article is the following recurrence formula which allows a_n to be determined quite easily, even for very large values of n:

$$(1) \qquad\qquad a_{3k+r} = [\tfrac{1}{2}(3a_k + r)] \quad \text{for } k \geq 0, \ 0 \leq r \leq 2.$$

The relation (1) was discovered empirically. A computer was used to check the conjecture for values of $n = 3k + r$ up to $3^7 - 1 = 2186$ and this gave us the confidence to search for a proof.

We emphasize that for any natural density-related notion of optimality, our partition is far from optimal. The sets S_k follow the density law $|S_k \cap [0, n-1]| \approx n^\alpha$ with $\alpha = \log 2 / \log 3 \approx .63$, whereas a theorem of Salem and Spencer [10] says that much denser sets free of arithmetic progression exist (for instance, the exponent may be replaced by any number less than 1). Readers interested in results along these lines should read Rankin's article [8].

One direction in which further work seems warranted is the investigation of the greedy partitioning algorithm under different initial conditions, in the spirit of Odlyzko and Stanley [7] (see section E10 of [5]); they grew sets free of arithmetic progressions from various initial "seeds" and observed an interesting mix of orderly and random behavior.

Also worthy of study is the greedy partition of the natural numbers that arises when one forbids only subsets that contain a length k arithmetic progression (with $k > 3$). When k is composite, apparent chaos ensues [4, 3], but for $k = p$ prime, the greedy partition is extremely orderly; for instance, the set S_0 ends up containing all those nonnegative integers that have no $p - 1$'s in their base p representations (this too was known by Szekeres, as described in [1]). No analogue of formula (1) has been found even for the case $p = 5$ (in particular, the plausible conjecture that

$$a_{pk+r} = \left[\frac{1}{p-1}(pa_k + r) \right]$$

seems to fail at $a_{p^3 - p^2 + p - 1}$ whenever $p > 3$), but it appears likely that some recursive formula exists. Indeed, observed regularities in the partition depend in an intriguing way on the prime factorization of $p - 1$. We hope further work by ourselves and others will shed some light on this situation.

In §I, notations and conventions are introduced. In §II, equation (1) is proved. Finally, §III answers some questions concerning the asymptotic behavior of (a_n).

I. Notations and conventions. Inductively define a_n ($n \geq 0$) as the least nonnegative integer such that for no positive integer $d \leq n/2$ does the relation

$$(2) \qquad\qquad a_n = a_{n-d} = a_{n-2d}$$

hold. Set $b_0 = 0$ and define b_n ($n > 0$) according to the recurrence relation

$$(3) \qquad\qquad b_{3k+r} = [\tfrac{1}{2}(3b_k + r)] \quad \text{for } k \geq 0, \ 0 \leq r \leq 2.$$

Put $S_i = \{n \geq 0:\ a_n = i\}$ and $T_i = \{n \geq 0:\ b_n = i\}$ for $i \geq 0$. Equation (1) is tantamount to the claim that $a_n = b_n$ for all n, or equivalently, that $S_i = T_i$ for all i. In proving this, the following notation will be useful:

$$A + B = \{a + b:\ a \in A,\ b \in B\}, \qquad kA = \{ka:\ a \in A\},$$

where k is an integer and A, B are sets of integers. Furthermore, in those cases where a set is a singleton and there is no possibility of confusion, we omit the brackets, so that for example $A + 2$ is an abbreviation for $A + \{2\}$.

We use the term "arithmetic progression" to signify an arithmetic progression of length 3 whose common difference is not zero.

II. The main theorem. The plan of this section is as follows. We split the claim that $S_i = T_i$ for all i into two parts:

THEOREM 1. *None of the sets T_i $(i \geq 0)$ contains an arithmetic progression.*

THEOREM 2. *If $n \in T_i$ and $j < i$, then there exists $0 < d \leq n/2$ such that $n - d, n - 2d \in T_j$.*

The first of these can be settled easily by induction. The other is harder, and will require the use of a supplementary induction hypothesis.

To see why Theorems 1 and 2 together imply that $S_i = T_i$ for all i, fix n and assume that $S_j \cap [0, n-1] = T_j \cap [0, n-1]$ for all j; consider what happens at the nth stage of the greedy algorithm, when it is determined to which of the S_j's we should assign n. Say $n \in T_i$. Theorem 2 says that n may not be assigned to any of the sets $S_0, S_1, \ldots, S_{i-1}$, while Theorem 1 says that n can (and hence will) be assigned to S_i without creating an arithmetic progression. Induction on n completes the argument.

PROOF OF THEOREM 1 (BY CONTRADICTION). Let $\{n - 2d, n - d, n\}$ be a minimal counterexample, so that $n, n - d, n - 2d \in T_i$ for some i; that is, $b_n = b_{n-d} = b_{n-2d} = i$. Put $n = 3k + r$ with $0 \leq r \leq 2$.

Case I: If $3 | d$, write $d = 3e$ so that $b_{3k+r} = b_{3(k-e)+r} = b_{3(k-2e)+r} = i$. By (3), each of the numbers $3b_k + r$, $3b_{k-e} + r$, and $3b_{k-2e} + r$ is either $2i$ or $2i + 1$. Since the three numbers are all congruent to one another modulo 3, no two can be consecutive, so they must all be equal. But then $b_k = b_{k-e} = b_{k-2e}$, contradicting the minimality of n.

Case II: If $3 \nmid d$, then $n, n - d$, and $n - 2d$ are pairwise incongruent modulo 3, so that we can write $b_{3k_0+0} = b_{3k_1+1} = b_{3k_2+2} = i$, where $3k_0 + 0$, $3k_1 + 1$, $3k_2 + 2$ is some permutation of n, $n - d$, $n - 2d$. By (3), each of the numbers $3b_{k_0} + 0$, $3b_{k_1} + 1$, $3b_{k_2} + 2$ is either $2i$ or $2i + 1$. Hence two of them must be equal. But this is impossible, since they are distinct modulo 3. $\square$

Before proving Theorem 2, we need a couple of lemmas.

LEMMA 1. *For all $n > 0$,*

(4) $$b_n \leq b_{n-1} + 1.$$

PROOF. Put $n = 3k + r$ with $0 \leq r \leq 2$.

$$b_{3k+1} = [\tfrac{1}{2}(3b_k + 1)] \leq [\tfrac{1}{2}(3b_k)] + 1 = b_{3k} + 1;$$
$$b_{3k+2} = [\tfrac{1}{2}(3b_k + 2)] \leq [\tfrac{1}{2}(3b_k + 1)] + 1 = b_{3k+1} + 1;$$

and

$$b_{3k+3} = [\tfrac{1}{2}(3b_{k+1})] \le [\tfrac{1}{2}(3(b_k+1))] \quad \text{(by induction, since } k+1 < 3k+3)$$
$$\le [\tfrac{1}{2}(3b_k+2)] + 1 = b_{3k+2} + 1. \quad \square$$

LEMMA 2. (a) $T_{3m} = 3T_{2m} + \{0,1\}$,
(b) $T_{3m+1} = (3T_{2m} + 2) \cup (3T_{2m+1})$,
(c) $T_{3m+2} = 3T_{2m+1} + \{1,2\}$.

PROOF. If $k \in T_{2m}$, then $b_k = 2m$ so that

$$b_{3k} = [\tfrac{1}{2}(3b_k + 0)] = [\tfrac{1}{2}(6m)] = 3m \quad \text{and}$$
$$b_{3k+1} = [\tfrac{1}{2}(3b_k + 1)] = [\tfrac{1}{2}(6m+1)] = 3m,$$

whence $3k, 3k+1 \in T_{3m}$. Conversely, suppose $n \in T_{3m}$. Write $n = 3k + r$ with $0 \le r \le 2$. By (3),

$$3m = b_n = [\tfrac{1}{2}(3b_k + r)]$$

so that $3b_k + r = 6m$ or $6m + 1$. This implies that $r = 0$ or 1 and that $b_k = 2m$, whence $k \in T_{2m}$. Thus (a) is established; the proofs of (b) and (c) are similar. $\square$

Note that (a) implies, in the case $m = 0$, that T_0 is the set of integers whose base 3 representations contain only 0's and 1's.

We now state and prove

THEOREM 2 (STRENGTHENED). *If $n \in T_i$ and $j \le i$, then there exists $d = d(n,j) \ge 0$ such that*

$$n - d \in T_j \quad \text{and} \quad n - 2d \in T_j.$$

If moreover $j < i$, then there exists $e = e(n,j) \ge 0$ such that

$$n - e \in T_{j+1} \quad \text{and} \quad n - 2e - 1 \in T_j.$$

(Note that if $j = i$ we may take $d = 0$, while if $j < i$, d must be positive since T_i and T_j are disjoint. Hence the first half of the claim is equivalent to the original version of Theorem 2.)

PROOF (BY INDUCTION ON i AND n). We will show that if the theorem holds for all i' and n' such that $i' \le i$ and $n' < n$ then it holds for i and n. The basis for the induction is easy: when $i = 0$, the only j to be considered is $j = 0$, for which we put $d = 0$. Now suppose $i \ge 1$. Fix $n = 3k + r \in T_i$ with $0 \le r \le 2$; $n \ge 2$ because $0, 1 \in T_0$ and $i \ne 0$. Since $i = b_n = [\tfrac{1}{2}(3b_k + r)]$ and $i \ge 1$, we have

$$3b_k \le 3b_k + r = 2i \quad \text{or} \quad 2i + 1$$
$$\le 2i + 1 \le 3i,$$

so that $b_k \le i$. Since $n > 0$, we have either $k > 0$ or $r > 0$, so that $k < 3k + r = n$. Hence it will suffice to prove that the theorem holds for i, n given that it holds for all i', n' with $i' \le b_k$, $n' \le k$.

Fix $j \le i$. If $j = i$, we take $d = 0$ and are done. Otherwise, write $j = 3m + t$ with $0 \le t \le 2$. Since

$$3b_k + r = 2i \quad \text{or} \quad 2i + 1$$
$$\ge 2i \ge 2(j+1) = 2(3m + t + 1),$$

we have

(5) $$b_k - 2m \ge \tfrac{1}{3}(2t + 2 - r).$$

Hence $b_k \geq 2m$. Applying our induction hypothesis to $n' = k$ and $i' = b_k$ with $2m$ in the role of j, we may assert that there exists d_0 such that

$$k - d_0 \in T_{2m}, \quad k - 2d_0 \in T_{2m}.$$

The analysis now divides naturally into cases, according to the values of t and r.

As long as $(t, r) \neq (0, 2)$, inequality (5) implies that $b_k > 2m$, or $b_k \geq 2m + 1$. This implies (apply the induction hypothesis to $n' = k$ and $i' = b_k$ with $2m + 1$ in the role of j) that there exists d_1 such that

$$k - d_1 \in T_{2m+1}, \quad k - 2d_1 \in T_{2m+1}.$$

It also implies (use $n' = k$ and $i' = b_k$ with $2m$ in the role of j) that there exists e_0 such that

$$k - e_0 \in T_{2m+1}, \quad k - 2e_0 - 1 \in T_{2m}.$$

Moreover, by Lemma 1, $b_{k-1} \geq b_k - 1 \geq 2m$, and this implies (use $n' = k - 1$ and $i' = b_{k-1}$ with $2m$ in the role of j) that there exists d_0' such that

$$(k - 1) - d_0' \in T_{2m}, \quad (k - 1) - 2d_0' \in T_{2m}.$$

Finally, in the cases $(t, r) = (1, 0), (2, 0), (2, 1)$, and $(2, 2)$, inequality (5) implies that $b_k > 2m + 1$, or $b_k \geq 2m + 2$. This implies (use $n' = k$ and $i' = b_k$ with $2m + 1$ in the role of j) that there exists e_1 such that

$$k - e_1 \in T_{2m+2}, \quad k - 2e_1 - 1 \in T_{2m+1}.$$

Moreover, in these cases Lemma 1 gives $b_{k-1} \geq b_k - 1 \geq 2m + 1$, and this implies (use $n' = k - 1$ and $i' = b_{k-1}$ with $2m + 1$ in the role of j) that there exists d_1' such that

$$(k - 1) - d_1' \in T_{2m+1}, \quad (k - 1) - 2d_1' \in T_{2m+1}.$$

To summarize the preceding paragraphs: in our task of finding $d(n, j)$ and $e(n, j)$, we have available to us

$$
\begin{aligned}
&d_0 = d(k, 2m): &&\text{all cases,} \\
&d_1 = d(k, 2m + 1): &&\text{all cases except } (t, r) = (0, 2), \\
&d_0' = d(k - 1, 2m): &&\text{all cases except } (t, r) = (0, 2), \\
&d_1' = d(k - 1, 2m + 1): &&(t, r) = (1, 0), (2, 0), (2, 1), \text{ or } (2, 2), \\
&e_0 = e(k, 2m): &&\text{all cases except } (t, r) = (0, 2), \\
&e_1 = e(k, 2m + 1): &&(t, r) = (1, 0), (2, 0), (2, 1), \text{ or } (2, 2).
\end{aligned}
$$

We now find $d(n, j)$ and $e(n, j)$ by cases.

$(t, r) = (0, 0)$: Take $d = 3d_0$, $e = 3d_0' + 1$. Then

$$
\begin{aligned}
n - d &= 3k - 3d_0 = 3(k - d_0) \in 3T_{2m} \subset T_{3m} = T_j; \\
n - 2d &= 3k - 6d_0 = 3(k - 2d_0) \in 3T_{2m} \subset T_{3m} = T_j; \\
n - e &= 3k - 3d_0' - 1 = 3(k - 1 - d_0') + 2 \in 3T_{2m} + 2 \subset T_{3m+1} = T_{j+1}; \\
n - 2e - 1 &= 3k - 6d_0' - 3 = 3(k - 1 - 2d_0') \in 3T_{2m} \subset T_{3m} = T_j.
\end{aligned}
$$

The other eight cases are similar.

$(t, r) = (0, 1)$: Take $d = 3d_0$, $e = 3e_0 + 1$.

$(t, r) = (0, 2)$: Take $d = 3d_0 + 1$, $e = 3d_0$.

$(t, r) = (1, 0)$:   Take $d = 3d_1$, $e = 3d_1' + 1$.
$(t, r) = (1, 1)$:   Take $d = 3e_0 + 1$, $e = 3d_1$.
$(t, r) = (1, 2)$:   Take $d = 3d_0$, $e = 3e_0 + 1$.
$(t, r) = (2, 0)$:   Take $d = 3d_1' + 1$, $e = 3e_1$.
$(t, r) = (2, 1)$:   Take $d = 3d_1$, $e = 3e_1 + 1$.
$(t, r) = (2, 2)$:   Take $d = 3d_1$, $e = 3e_1 + 1$.

This completes the construction of d and e, and the theorem now follows by induction. □

Note that in each case, both $d(n, j)$ and $e(n, j)$ are of the form $3x$ or $3x+1$, where x is some earlier constructed d or e. (Of course this recursive construction does not nest indefinitely; ultimately we reach the level of $d(n, 0) = 0$ with $n \in T_0$.) It follows that the d and e given by the proof of Theorem 2 have base 3 representations containing only 0's and 1's and thus lie in the set T_0. In general, there may be many different values of d and e that validate Theorem 2 for a particular pair n, j. However, it appears that the $d(n, j)$ and $e(n, j)$ constructed above are the *only* such values of d and e that lie in T_0. The same goes for $f(n, j)$, where f is required to satisfy

$$n - f - 1 \in T_j \quad \text{and} \quad n - 2f - 1 \in T_j$$

whenever $j < b_n$; the existence of such an f has been proved, but uniqueness in T_0 has not. We have used a computer to verify the uniqueness of d, e, and f in T_0 for $n \leq 3^7 - 1$.

The function $f(n, j)$ mentioned above arose in an earlier version of the proof which used a joint induction on all three functions. It was later seen that f could be eliminated from consideration. It would be nice to have a proof of Theorem 2 requiring no extra hypothesis whatsoever (i.e., to eliminate e from consideration as well). Observe that in our proof of Theorem 2, only the prescription "$d = 3e_0 + 1$" for the case $(t, r) = (1, 1)$ prevents us from dropping the induction hypothesis on e from our proof.

III. Some associated sequences. One question we ask about the limiting behavior of the partition $\{S_0, S_1, S_2, \dots\}$ $(= \{T_0, T_1, T_2, \dots\})$ is, how quickly does $|S_i \cap [0, n-1]|$ grow as a function of n? In [9], Roth showed (as a special case of the later Szemerédi Theorem [11, 2]) that the subsets of $[0, n-1]$ containing no arithmetic progressions of length 3 have density tending to 0 as $n \to \infty$ (for a short and elementary proof, see [6]); this tells us in particular that for fixed i, $|S_i \cap [0, n-1]|$ is $o(n)$.

Let us first consider the set S_0 and define $\sigma_0(n) = |S_0 \cap [0, n-1]|$. Since S_0 consists of those numbers whose base three expansions contain only 0's and 1's, $\sigma_0(3^k) = 2^k$. It follows that for $3^k \leq n < 3^{k+1}$, $2^k \leq \sigma_0(n) \leq 2^{k+1}$, whence

$$\frac{\log \sigma_0(n)}{\log n} \to \frac{\log 2}{\log 3}$$

as $n \to \infty$. In other words, $\sigma_0(n)$ grows like n^α, where $\alpha = \log 2 / \log 3 \approx .63$.

Now let $\sigma_i(n) = |S_i \cap [0, n-1]|$. The functions $\sigma_i(n)$ $(i > 0)$ also grow like n^α; this is a consequence of the following fact.

THEOREM 3. *For all i, and for all k sufficiently large relative to i,*

$$\sigma_i(3^k) = 2^k - \Delta_i$$

where the sequence $(\Delta_i)_0^\infty = (0,1,2,4,6,8,12,14,\dots)$ *satisfies the recurrence*

$$\Delta_i = \Delta_{[\frac{1}{3}(2i)]} + \Delta_{[\frac{1}{3}(2i+1)]}.$$

PROOF. Since $S_i = T_i$ for all i, Lemma 2 of the previous section implies

$$\sigma_i(3^k) = \sigma_{j_0}(3^{k-1}) + \sigma_{j_1}(3^{k-1})$$

where $j_0 = [\frac{1}{3}(2i)]$, $j_1 = [\frac{1}{3}(2i+1)]$. For $i = 1$, the relation

$$\sigma_1(3^k) = 2^{k-1} + \sigma_1(3^{k-1})$$

yields

$$\sigma_1(3^k) = 2^{k-1} + 2^{k-2} + \cdots + 1 = 2^k - 1$$

for all $k \geq 0$. Now fix $i > 1$, so that $j_0, j_1 < i$, and assume there exists K such that

$$\sigma_{j_0}(3^k) = 2^k - \Delta_{j_0} \quad \text{and} \quad \sigma_{j_1}(3^k) = 2^k - \Delta_{j_1}$$

for all $k \geq K$. Then for all $k \geq K + 1$, we have

$$\sigma_i(3^k) = (2^{k-1} - \Delta_{j_0}) + (2^{k-1} - \Delta_{j_1}) = 2^k - (\Delta_{j_0} + \Delta_{j_1}),$$

so that the theorem is true by induction. □

Another kind of asymptotic question we ask concerns the subsequence $(c_k)_0^\infty = (0,1,2,4,7,\dots)$ of $(a_n)_0^\infty$, where $c_k = a_n$ for $n = 3^k - 1$. It may be shown that c_k is the largest of the first 3^k terms of (a_n), and measures how quickly $\max\{a_n: n < N\}$ grows. We have

$$a_{3^{k+1}-1} = [\tfrac{1}{2}(3a_{3^k-1} + 2)],$$

so $c_{k+1} = [\frac{1}{2}(3c_k + 2)]$. That is, for each k, either c_k is even, in which case $c_{k+1} = \frac{3}{2}c_k + 1$ and

$$\frac{c_{k+1}}{(3/2)^{k+1}} = \frac{c_k}{(3/2)^k} + \frac{1}{(3/2)^{k+1}},$$

or c_k is odd, in which case $c_{k+1} = (3/2)c_k + 1/2$ and

$$\frac{c_{k+1}}{(3/2)^{k+1}} = \frac{c_k}{(3/2)^k} + \frac{1/2}{(3/2)^{k+1}}.$$

Hence $c_k/(3/2)^k$ tends toward the limit

$$\beta = 1(2/3) + \tfrac{1}{2}(2/3)^2 + 1(2/3)^3 + 1(2/3)^4 + \tfrac{1}{2}(2/3)^5 + \cdots,$$

where the coefficient of $(2/3)^{k+1}$ is 1 if c_k is even and $1/2$ if c_k is odd. The coefficients alternate between $1/2$ and 1 in a seemingly erratic fashion, and we have no idea whether β can be characterized in terms of the standard algebraic and transcendental functions. Even proving that c_k is odd 50% of the time asymptotically looks hard.

REFERENCES

1. P. Erdös and P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261–264.
2. H. Furstenberg, Y. Katznelson, and D. Ornstein, *The ergodic theoretical proof of Szemerédi's theorem*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), 527–552.
3. J. L. Gerver, *Irregular sets of integers generated by the greedy algorithm*, Math. Comp. **40** (1983), 667–676.

4. J. L. Gerver and L. T. Ramsey, *Sets of integers with no long arithmetic progressions generated by the greedy algorithm*, Math. Comp. **33** (1979), 1353–1359.

5. Richard K. Guy, *Unsolved problems in number theory*, Springer-Verlag, 1980.

6. D. J. Newman, *Sequences without arithmetic progressions*, Analytic Number Theory (Philadelphia, Pa., 1980), Lecture Notes in Math., vol. 899, Springer, Berlin and New York, 1981, pp. 311–314.

7. A. M. Odlyzko and R. P. Stanley, *Some curious sequences constructed with the greedy algorithm*, unpublished Bell Laboratories report, January 1978.

8. 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** (1960/1961), 332–344.

9. K. F. Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris **234** (1952), 388–390.

10. R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **28** (1942), 561–563.

11. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

12. B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw. Arch. Wisk. **15** (1927), 212–216.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, CAMDEN, NEW JERSEY 08102 (Current address of Joseph Gerver)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720

DEPARTMENT OF MATHEMATICS, SOUTH AUSTRALIAN INSTITUTE OF TECHNOLOGY, WHYALLA, S.A., AUSTRALIA (Current addresss of Jamie Simpson)

Current address (James Propp): Department of Mathematics, University of Maryland, College Park, Maryland 20742