A RAMSEY-TYPE PROPERTY IN ADDITIVE NUMBER THEORY

by S. A. BURR and P. ERDÖS

Dedicated to Prof. R. A. Rankin on his 70th birthday

1. Introduction. Let A be a sequence of positive integers. Define $P(A)$ to be the set of all integers representable as a sum of distinct terms of A. Note that if A contains a repeated value, we are free to use it as many times as it occurs in A. We call A *complete* if every sufficiently large positive integer is in $P(A)$, and *entirely complete* if every positive integer is in $P(A)$. Completeness properties have received considerable, if somewhat sporadic, attention over the years. See Chapter 6 of [3] for a survey.

One interesting question about completeness properties is how robust they are. For instance, if one is content with a slightly weaker property ("subcompleteness"), it often happens that even rather strong perturbations of a sequence cannot destroy the property [1, 2]. Another sense in which a property can be robust is that, even after the object in question is somehow partitioned into two or more classes, the property remains true for some class. This is the province of *Ramsey theory*, which is heavily studied today; see [4, 5] for surveys of the general field. The most straightforward definition of a "Ramsey-complete" sequence would be that if it were partitioned into two classes A_1 and A_2, either $P(A_1)$ or $P(A_2)$ would contain every large positive integer. Unfortunately, this does not appear to lead to interesting problems. It seems necessary to permit the class to vary with the number represented, as in the following definitions.

DEFINITION. A sequence is *Ramsey-complete* if whenever the sequence is partitioned into two classes A_1 and A_2, every sufficiently large positive integer is a member of $P(A_1) \cup P(A_2)$.

DEFINITION. A sequence is *entirely Ramsey-complete* if whenever the sequence is partitioned into two classes A_1 and A_2, every positive integer is a member of $P(A_1) \cup P(A_2)$.

These definitions lead naturally to many questions. Here we will consider that of the maximum possible growth rate of such sequences. As usual, let $A(x)$ denote the number of terms of the sequence A that are less than or equal to x. It is easy to see that a sequence for which $A(x) < \log_2 x$ infinitely often cannot be complete. Furthermore, a complete sequence with exponential growth also must grow in a fairly regular way. Something similar applies to Ramsey-completeness. Our main object here is to prove the following two theorems. It will be convenient in most of what follows to work with the binary logarithm $\log_2$, which we will denote by lg. To avoid confusion, we use ln for the natural logarithm.

THEOREM 1. *There is an entirely Ramsey-complete sequence A satisfying $A(x) - A(\frac{1}{2}x) < 2 \lg^2 x$ for all sufficiently large x.*

THEOREM 2. *There is an $\varepsilon > 0$ such that no infinite sequence of integers A satisfying $A(x) - A(\frac{1}{2}x) < \varepsilon \lg x$ for all sufficiently large x is Ramsey-complete.*

Theorem 1 gives a lower bound on the maximum growth rate of an (entirely) Ramsey-complete sequence, and Theorem 2 essentially gives an upper bound, along with a smoothness condition. In each of these theorems, it is possible to replace the condition on $A(x) - A(\frac{1}{2}x)$ by a pure growth condition. For the first, this is essentially trivial.

THEOREM 1a. *There is an entirely Ramsey-complete sequence A satisfying $a_x > 2^{(1/2)\sqrt[3]{x}}$ for all sufficiently large x.*

Proof (using Theorem 1). We have $A(x) = (A(x) - A(\frac{1}{2}x)) + (A(\frac{1}{2}x) - A(\frac{1}{4}x)) + \ldots$, where about $\lg x$ terms appear. Therefore, the sequence in Theorem 1 satisfies $A(x) < (2 + \varepsilon) \lg^3 x$; hence $a_x > 2^{\sqrt[3]{(x/(2+\varepsilon))}} > 2^{(1/2)\sqrt[3]{x}}$.

At the cost of considerable effort, it is possible to prove a similar analog of Theorem 1.

THEOREM 2a. *There is a $C > 0$ such that no infinite sequence of integers A satisfying $a_x > 2^{C\sqrt{x}}$ for all sufficiently large x is Ramsey-complete.*

We will not prove this here; the proof is quite complicated, and the small improvement does not seem to justify the effort, in view of the substantial gap between Theorems 1 and 2.

2. The Lower Bound. We now prove Theorem 1, in a form which explicitly exhibits the required sequence.

THEOREM 1b. *Let A be a sequence defined as follows. The sequence begins with 16 copies of 1, followed by blocks B_n for each $n \geq 1$, where B_n consists of $2n + 2$ copies of 2^n, together with $n + 2$ copies of each of the numbers $2^n + 1, 2^n + 2, 2^n + 4, \ldots, 2^n + 2^{n-1}$. Then the sequence A is entirely Ramsey-complete.*

Proof. We will use induction. We will employ a stronger hypothesis, namely that if A is partitioned into two classes, then for each $n > 0$, all the integers $(n - 1)2^{n-1}, \ldots, n2^n$ are representable by a single class, using only terms from $B_{n-1}, B_{n-2}, \ldots$, and the 1's at the beginning. (Of course, the class may vary with the value of n.) Denote the set $\{(n - 1)2^{n-1}, \ldots, n2^n\}$ by D_{n-1}. The hypothesis is trivially true for $n = 1, 2$. Now suppose that the hypothesis holds for $n - 1$, $n \geq 2$. Thus, without loss of generality, all of D_{n-1} is representable by Class 1, without using $B_n, B_{n+1}, \ldots$; we distinguish two cases.

(1) At least $n + 2$ of the members of B_n are in Class 1, say $b_1, \ldots, b_{n+2}$. These all satisfy $2^n \leq b_i \leq 2^n + 2^{n-1}$. In this case, we will show that all of $D_n = \{n2^n, \ldots, (n + 1)2^{n+1}\}$ is representable in Class 1 as well. First consider the numbers representable using just b_1, together with the members of D_{n-1}. The numbers so representable are $(n - 1)2^{n-1} + b_1, \ldots, n2^n + b_1$. If $n \geq 2$, it is clear that this range of values overlaps D_{n-1}, even if b_1 is as large as $2^n + 2^{n-1}$. Hence, we have represented all of $(n - 1)2^{n-1}, \ldots, n2^n + b_1$. In the same way, if we now use b_2 as well, we can represent all of $(n - 1)2^{n-1}, \ldots, n2^n + b_1 + b_2$. Continuing in the same fashion, we see that all of $(n - 1)2^{n-1}, \ldots, n2^n + b_1 + b_2 + \ldots + b_{n+2}$ can be represented by Class 1. Since $n2^n + b_1 + b_2 + \ldots + b_{n+2} \geq n2^n + (n + 2)2^n = (n + 1)2^{n+1}$, this includes all of D_n.

(2) No more than $n+1$ of the members of B_n are in Class 1. In this case, Class 2 contains at least $n+1$ copies of 2^n, and at least one copy of each of $2^n+1, 2^n+2, \ldots, 2^n + 2^{n-1}$. We will show that all of D_n can be represented by Class 2, using only these numbers. Consider any number of the form $n2^n + m$, where $0 \le m < 2^n$. The binary expansion of m requires just n bits. For each zero in this expansion, add in a copy of 2^n. Each 1 corresponds to some value 2^k; add in a copy of $2^n + 2^k$. Thus, we can represent all of $n2^n, \ldots, (n+1)2^n - 1$ by Class 2. At least one copy of 2^n is left over. With this, in similar fashion to Case (1), we can represent all of $n2^n, \ldots, (n+2)2^n - 1$. It is now an easy matter to continue as in Case (1), representing all of D_n. (And a great deal more: observe that at least $n^2 + 3n + 1$ members of B_n are in Class 2 in this case.) This completes the proof.

Theorem 1b uses a sequence in which many values are repeated. Repeated values are convenient, but not necessary.

THEOREM 1c. *There is a strictly increasing sequence A which is entirely Ramsey-complete and satisfies $A(x) - A(\frac{1}{2}x) < \delta \lg^2 x$ for all sufficiently large x, where δ is some constant.*

Sketch of proof. We will describe roughly how to construct such a sequence in a way similar to that of the proof of Theorem 1b. In place of the 16 copies of 1, use $\{1, 2, 3, \ldots, r\}$ for some suitable r. For the blocks B_n, use $\{2^n \pm 1, 2^n \pm 2, 2^n \pm 3, \ldots, 2^n \pm sn\}$ in place of the $2n+2$ copies of 2^n, and $\{2^n + 2^i \pm 1, 2^n + 2^i \pm 2, 2^n + 2^i \pm 3, \ldots, 2^n + 2^i \pm tn\}$ in place of the $n+2$ copies of $2^n + 2^i$, where s and t must be suitably chosen. The definition of D_n must change somewhat. The rest of the proof is then rather similar, using the fact that if both $x \pm a$ are in a class, then $2x$ is representable by that class.

3. The Upper Bound.

We will shortly prove Theorem 2, but we need a lemma. As a technical convenience, we extend the definition of the binomial coefficient $\binom{u}{v}$ to arbitrary nonnegative real u and v using the gamma function. Furthermore, whenever a sum or product involves $\binom{u}{v}$, the condition $0 \le v \le u$ is implicit.

LEMMA. *Let $S(u, v)$ denote the sum $\sum_{i \ge 0} \binom{u}{v-i}$, and let γ be a fixed constant. Then there is an $\alpha = \alpha(\gamma)$, not depending on u, such that for every $u > 1$,*

$$\prod_{1 < 2^i \le \gamma u} S(u, 2^i) < 2^{\alpha u}.$$

Proof. We first consider the range in which $2^i \le \frac{1}{4}u$. Let $v \le \frac{1}{4}u$; then $\binom{u}{v-1} < \frac{1}{2}\binom{u}{v}$ for all such v, so that $S(u, v) < 2\binom{u}{v}$. Taking natural logarithms and using Stirling's

approximation, we have:

$$\ln S(u, v)$$
$$= \ln\binom{u}{v} + O(1)$$
$$= u \ln u - u - (u-v)\ln(u-v) + (u-v) - v \ln v - v + O(\ln u)$$
$$= v \ln u - v \ln v - (u-v)\ln(u-v) + (u-v)\ln u + O(\ln u)$$
$$= v \ln u/v - (u-v)\ln(1 - v/u) + O(\ln u)$$
$$\leq v \ln u/v - (u-v)(-v/u - v^2/2u^2 - \ldots) + O(\ln u)$$
$$\leq v \ln u/v + 2v + O(\ln u).$$

Therefore, returning to binary logarithms,

$$\lg\binom{u}{2^j} \leq 2^j \lg(u/2^j) + 2^{j+1} \lg e + O(\lg u)$$

$$\leq 2^j(\lg u + 2 \lg e) - j2^j + O(\lg u).$$

Let us sum this from 0 to k, where $u/8 < 2^k \leq u/4$, so that $k = \lg u + O(1)$. Using the standard formulas $\sum_{j=0}^{k} 2^j = 2^{k+1} - 1$ and $\sum_{j=0}^{k} j2^j = (k-1)2^{k+1} + 2$, we have

$$\sum_{j=0}^{k} \lg S(u, 2^j)$$
$$= 2^{k+1}(\lg u + 2 \lg e) - (k-1)2^{k+1} + O(\lg u)$$
$$= 2^{k+1}(\lg u + 2 \lg e) - 2^{k+1}(\lg u + O(1))$$
$$\leq \alpha' u$$

for some $\alpha' > 0$. Thus,

$$\prod_{0 \leq j \leq \lg(u/4)} S(u, 2^j) < 2^{\alpha' u}.$$

However, the number of factors omitted from this product is not more than $2 + \lg \gamma + 1$, and each of these is no more than 2^u. Including these factors, we have our desired upper bound on the whole product. This completes the proof.

We now restate and prove Theorem 2.

THEOREM 2. *There is an $\varepsilon > 0$ such that no infinite sequence of integers $A = \{a_1 \leq a_2 \leq \ldots\}$ satisfying $A(X) - A(\frac{1}{2}X) < \varepsilon \lg X$ for all sufficiently large X is Ramsey-complete.*

Proof. Let $\gamma = 5$ and let α be the corresponding value in the lemma. We will prove the desired result for $\varepsilon < 1/(\alpha + 1)$. Let A be any sequence satisfying the given condition. Define the sequence $\{Z_1, Z_2, \ldots\}$ by $Z_1 = 4$, $Z_k = 2^{\lg^2 Z_{k-1}}$. Partition A as follows. For each i, $Z_k < a_i \leq Z_{k+1}$ for some k. When k is odd, place a_i in Class 1; when k is even, place a_i in Class 2. We will show that if k is large enough, there are integers in each interval

$(Z_k, Z_{k+1}]$ that are not representable by either class. Set $Z = Z_k$; without loss of generality, k is even. Set $X = 2\varepsilon Z \lg Z$. We will show that most of the integers in $[X, 2X]$ are not representable by either class. Indeed, no such integer is representable by Class 1, since it is clear that $2X < Z_{k+1}$, and

$$
\begin{aligned}
\sum_{a_i \leq Z} a_i = \sum_{\frac{1}{2}Z < a_i \leq Z} a_i + \sum_{\frac{1}{4}Z < a_i \leq \frac{1}{2}Z} a_i + \ldots \\
\leq \varepsilon Z \lg Z + \varepsilon \tfrac{1}{2} Z \lg \tfrac{1}{2} Z + \ldots \\
< 2\varepsilon Z \lg Z = X.
\end{aligned}
$$

We now must estimate how many integers in $[X, 2X]$ are representable by Class 2. To do this we break up the interval $[Z, 2X]$ into subintervals:

$$
(Z, 2X] \subseteq (X, 2X] \cup (X/2, X] \cup (X/4, X/2] \cup \ldots \cup (X/2^{k+1}, X/2^k],
$$

where the last interval contains the number Z. Hence $2^{k+1} < 2X/Z \leq 4\varepsilon \lg Z \leq 5\varepsilon \lg X$. In each interval $(X \cdot 2^{-j}, X \cdot 2^{-j+1}]$, there are fewer than $\varepsilon \lg (X \cdot 2^{-j+1}) \leq \varepsilon \lg X$ members of A. Moreover, no more than $2X/(X \cdot 2^{-j}) = 2^{j+1}$ summands in each such interval can be used to represent a number less than or equal to $2X$. Hence, no more than

$$
\sum_{i \geq 0} \binom{\varepsilon \lg x}{2^{j+1} - i} = S(\varepsilon \lg X, 2^{j+1})
$$

sums are possible using summands in such an interval. Hence, by the above lemma, with $\gamma = 5$, the total number of sums using all these intervals is less than

$$
\prod_{2 \leq 2^{j+1} \leq 5\varepsilon \lg X} S(\varepsilon \lg X, 2^{j+1}) < 2^{\varepsilon \alpha \lg X}.
$$

Finally, we must take into account any other a_i in Class 2. These are all no greater than Z_{k-1}. As in the proof of Theorem 1a,

$$
\begin{aligned}
A(Z_{k-1}) = (A(Z_{k-1}) - A(\tfrac{1}{2}Z_{k-1})) + (A(\tfrac{1}{2}Z_{k-1}) - A(\tfrac{1}{4}Z_{k-1})) + \ldots \\
\leq \varepsilon \lg Z_{k-1} + \varepsilon \lg (\tfrac{1}{2}Z_{k-1}) + \ldots \\
\leq \varepsilon \lg^2 Z_{k-1}.
\end{aligned}
$$

Therefore, the number of integers representable by $a_i \leq Z_{k-1}$ is no more than

$$
2^{\varepsilon \lg^2 Z_{k-1}} \leq 2^{\varepsilon \lg Z_k} < 2^{\varepsilon \lg X}.
$$

Consequently, the total number of sums in $[X, 2X]$ can be no more than

$$
2^{\varepsilon(\alpha+1)\lg X} < 2^{\lg X} = X.
$$

This completes the proof.

4. Concluding Remarks. Our results leave open many interesting problems. The most obvious of these is that of the true maximum order of growth for Ramsey-complete sequences. Possibly, Theorems 2 and 2a are closer to the truth than Theorems 1 and 1a. It

certainly appears that the proof of Theorem 1b leaves room for improvement. It seems likely that one might be able to improve, say, the condition in Theorem 1a to perhaps $a_x > 2^{c\sqrt[3]{x}\lg x}$, without too much trouble. However, no obvious line of attack has presented itself to narrow the gap, at either end, by a substantial margin.

Another very interesting area to study is that of generalizing the definition of Ramsey-completeness to partitioning the sequence into three (or more) classes. The following would seem to be a natural conjecture, by analogy to Theorem 1a.

CONJECTURE. *There is a β and a sequence A satisfying $a_x > 2^{x^\beta}$ for large x, such that if A is partitioned into three classes A_1, A_2, A_3, then $P(A_1) \cup P(A_2) \cup P(A_3)$ contains all large integers.*

However, serious complications arise when trying to mimic the proof of Theorem 1b, and it is possible that no such β and sequence A exist.

Finally, one can ask whether various "natural" sequences are Ramsey-complete. The most obvious of these would seem to be sequences of polynomial values, for instance the sequence of squares. Indeed, the ideas in the proofs of Theorems 1b and 1c may be useful here.

REFERENCES

1. S. A. Burr, On the Completeness of Sequences of Perturbed Polynomial Values, *Pacific J. Math.* **85** (1979), 355–360.
2. S. A. Burr and P. Erdös, Completeness Properties of Perturbed Sequences, *J. Number Theory* **13** (1981), 446–455.
3. P. Erdös and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monographie 28 (L'Enseignement Mathématique, 1980).
4. R. L. Graham, *Rudiments of Ramsey Theory* (American Mathematical Society, 1981).
5. R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory* (Wiley, 1980).

CITY COLLEGE   HUNGARIAN ACADEMY OF SCIENCES
C.U.N.Y.
NEW YORK NY 10031
U.S.A.