Sequences of Integers Avoiding 3-term Arithmetic Progressions

Arun Sharma
Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720
asharma@math.berkeley.edu

Submitted: Sep 29, 2011; Accepted: Jan 12, 2012; Published: Jan 21, 2012
Mathematics Subject Classifications: 05A15, 05C55

Abstract. *The optimal length $r(n)$ of a sequence in $[1, n]$ containing no 3-term arithmetic progression is determined for several new values of n and some results relating to the subadditivity of r are obtained. We also prove a particular case of a conjecture of Szekeres.*

A subsequence $S = (a_1, a_2, \ldots, a_k)$ of the sequence $\langle n \rangle = (1, 2, \ldots, n)$ containing no three terms a_p, a_q, and a_r for which $a_q - a_p = a_r - a_q$ (i.e., S contains no three term arithmetic progression) is called an A *sequence* in $\langle n \rangle$. $r(n)$ denotes the maximum number of terms possible in an A sequence in $\langle n \rangle$, and any such sequence is said to be *optimal* in $\langle n \rangle$. Throughout this paper any input variable x in $r(x)$ is assumed to be a positive integer.

The following properties of A sequences and the function r are evident.

(P1) If $S = (a_1, a_2, \ldots, a_k)$ is an A sequence in $\langle n \rangle$, then $(n+1-a_k, n+1-a_{k-1}, \ldots, n-a_1)$ is an A sequence called the *complement* of S in $\langle n \rangle$. Also, for any integer $j < a_1$, a *translate* $(a_1 - j, a_2 - j, \ldots, a_k - j)$ of S is an A sequence.

(P2) For any m and n, $r(m + n) \leq r(m) + r(n)$. In brief, the function r is subadditive.

(P3) For any n, $r(n) \leq r(n + 1) \leq 1 + r(n)$. Whenever $r(n - 1) < r(n)$, we call n a *jump node* for r.

(P4) If $(a_1, a_2, \ldots, a_{k-1}, a_k)$ is an A sequence in $\langle n \rangle$, then $(a_1, a_2, \ldots, a_{k-1}, a_k, 2n-1+a_1, 2n-1+a_2, \ldots, 2n-1+a_{k-1}, 2n-1+a_k)$ is an A sequence in $\langle 3n-1 \rangle$; whence $r(3n - 1) \geq 2r(n)$.

(P5) If $r(n - 1) < r(n)$, then any optimal A sequence in $\langle n \rangle$ contains both 1 and n.

(P6) If $r(n - 1) < r(n) < r(n + 1)$, then any optimal A sequence in $\langle n + 1 \rangle$ contains all four of 1, 2, n, and $n + 1$.

Observe that, by ($P6$), no three consecutive integers can all be jump nodes for r.

The study of A sequences was initiated by Erdős and Turan in [1], and since the appearance of their paper there has been extensive research concerning the asymptotic behavior of the function r and its correspondent that counts the sequences in $\langle n\rangle$ avoiding k-term arithmetic progressions for $k > 3$. A substantial paper by Szemeredi [2] gives many references on this topic. The exact value of $r(n)$ is, however, known for only a few n. In this regard, an error in [1] in computing $r(20)$ has gone undetected and as a consequence, subsequent computations of $r(n)$ for certain $n > 20$ are based on flawed arguments. For example, the evaluations of $r(21)$ and $r(41)$ (and perhaps $r(22)$ and $r(23)$ also) in [1] are founded on incorrect reasoning. The values of $r(n)$ for $n \leq 19$ found in [1] are, however, all correct. We summarize these values by listing only the jump nodes for r:

$$r(2) = 2,\, r(4) = 3,\, r(5) = 4,\, r(9) = 5,\, r(11) = 6,\, r(13) = 7,\, r(14) = 8.$$

The next jump node for r after 14 is 20 and not 21 as claimed in [1]. This is because $r(19) = 8$, and $(1, 2, 6, 7, 9, 14, 15, 18, 20)$ is an A sequence.

There is a sequence $\{T_k\}$ of positive integers with three intriguing questions surrounding it: (a) Is each T_k, $k > 1$, a jump node for r? (b) Is the optimal A sequence in $\langle T_k\rangle$ for each k unique? (c) Is it true that $r(T_k) = 2^k$ for each k? The sequence $\{T_k\}$ is defined recursively as follows:

$$T_k = 3T_{k-1} - 1 \;\; for \;\; k \geq 1;\;\; T_0 = 1.$$

Observe that $T_k = \frac{1}{2}(3^k + 1)$, and that by $(P4)$

$$r(T_k) \geq 2^k. \tag{$*$}$$

One can easily verify that the three questions raised above regarding this sequence are correct for $k = 0$, 1, 2, and 3. Szekeres conjectured that question (c) has an affirmative answer for any k. The proof of this conjecture for $k = 4$ given in [1] is erroneous as it is based an incorrect value of $r(20)$. In this paper we give a correct proof. We also prove some inequalities analogous to $(P2)$ and evaluate $r(n)$ for $21 \leq n \leq 27$ and for $n = 41$, 42, and 43.

If $r(n)$ is known at a jump node n, then one can determine $r(n + 1)$ by listing all the optimal A sequences in $\langle n\rangle$ and then testing if any one amongst them still retains the A property when $n + 1$ is appended to it. This procedure can be suitably modified to test whether $r(n + 1) = c + 1$ given that $r(n) \leq c$. For the convenience of such testing we begin by listing a few A sequences.

(i) By $(P4)$, $(1, 2, 4, 5, 10, 11, 13, 14, 28, 29, 31, 32, 37, 38, 40, 41)$ is an A sequence in $\langle 41\rangle$. Note that the seven terms immediately following the first term in this sequence are all jump nodes.

(ii) There are exactly four optimal A sequences in $\langle 9 \rangle$, namely, $(1, 2, 4, 8, 9)$, $(1, 2, 6, 7, 9)$, $(1, 2, 6, 8, 9)$, and $(1, 3, 4, 8, 9)$. In contrast, there are twenty five such sequences in $\langle 8 \rangle$.

(iii) There are only two optimal A sequences in $\langle 20 \rangle$, namely, $(1, 2, 6, 7, 9, 14, 15, 18, 20)$ and $(1, 3, 6, 7, 12, 14, 15, 19, 20)$.

The following theorem sharpens the inequality in $(P2)$ in a particular case.

Theorem 1. If $r(n-1) < r(n)$, then $r(2n) < r(n)+r(n)$ and $r(2n-1) < r(n)+r(n-1)$

Proof. Let $r(n) = k$. Then, by the given hypothesis, $r(n-1) = k-1$. Now suppose $r(2n) = 2k$ and let $S = (a_1, a_2, \ldots, a_{2k})$ be an optimal A sequence in $\langle 2n \rangle$. Then the first k terms of S are an optimal A sequence in $\langle n \rangle$, and the last k terms are a translate of an optimal A sequence in $\langle n \rangle$. Thus, by $(P5)$, $a_{k+1} = n+1$ and $a_{2k} = 2n$. Now, as $n+1$ and $2n$ both occur in S, therefore $2 \notin S$. Consequently, $a_2 > 2$, whence $(a_2 - 2, a_3 - 2, \ldots, a_k - 2, a_{k+1} - 2)$ is a k term A sequence in $\langle n-1 \rangle$, contradicting that $r(n-1) = k-1$.

To prove the second statement, assume $r(2n-1) \geq 2k-1$, and let $T = (a_1, a_2, \ldots, a_{2k-1})$ be an A sequence in $\langle 2n-1 \rangle$. We may assume that the first k terms of T are in $\langle n \rangle$ (for otherwise we will work with the complement of T in $\langle 2n \rangle$ which then will have this property). Hence $(a_1, \ldots, a_k)$ is an optimal A sequence in $\langle n \rangle$, and so $a_1 = 1$ and $a_k = n$. But then $2n-1 \notin T$, implying that T is also an A sequence in $\langle 2n-2 \rangle$. This is impossible because $r(2n-2) \leq r(n-1) + r(n-1) = 2k-2$. $\diamondsuit$

Theorem 2. $r(21) = r(22) = r(23) = 9$.

Proof. Suppose $r(21) = 10$. Then there exists an A sequence in $\langle 21 \rangle$ having nine terms in $\langle 20 \rangle$. This is impossible because neither of the two nine term A sequences in $\langle 20 \rangle$ retains the A property when 21 is appended to it. Hence $r(21) = 9$.

If $r(22) = 10$, then (after complementing if necessary) there is an optimal A sequence in $\langle 22 \rangle$ having at least five terms in [11]. However, on testing all the A sequences in $\langle 11 \rangle$ of length five and six, we find that not only none of them extends to an A sequence with ten terms in $\langle 22 \rangle$ or but also none so extends to $\langle 23 \rangle$. This proves that $r(22) = 9$ and it also leads us to conclude that $r(23) = 9$ (for if $n = 23$ were a jump node, an optimal A sequence on $\langle 23 \rangle$ would contain both 1 and 23 and exclude 12). $\diamondsuit$

Since arguments similar to those given in the preceding theorem also hold with slight modifications in the next three theorems, we will skip many details.

Theorem 3. $r(24) = r(25) = 10$ and $r(26) = r(27) = 11$.

Proof. As $r(23) = 9$ and $(1, 2, 6, 7, 9, 14, 18, 20, 23, 24)$ is an A sequence, hence $r(24) = 10$. The proof that $r(25) < 11$ can now be completed by examining all A sequences in $\langle 12 \rangle$ having six terms. Next, since $(1, 3, 4, 8, 9, 11, 16, 20, 22, 25, 26)$ is an A sequence, hence $r(26) = 11$. The proof that $r(27) < 12$ can be completed by examining A sequences with at least six terms in $\langle 13 \rangle$. $\diamondsuit$

Theorem 4. $15 \leq r(40) \leq 16$.

Proof. The sixteen term A sequence in [41] listed in (i) shows that $r(40) \geq 15$. On the other hand, by Theorem 1, $r(40) \leq 17$. Now, if $r(40) = 17$, then there is an optimal A sequence in $\langle 40 \rangle$ having nine terms in $\langle 20 \rangle$. However, neither of the two nine term A sequences in $\langle 20 \rangle$ extends to an A sequence with seventeen terms in $\langle 40 \rangle$. This proves that $r(40) \leq 16$. $\diamondsuit$

The next theorem, in part, shows that Szekeres' conjecture holds for $k = 4$.

Theorem 5. $r(41) = r(42) = r(43) = 16$.

Proof. As $r(40) \leq 16$, so $r(41) \leq 17$. Also, as we already know a sixteen term A sequence in $\langle 41 \rangle$, therefore $r(41) \geq 16$. Now if there exists a seveteen term A sequence S in $\langle 41 \rangle$, then it must exclude 21. Thus we may assume (by replacing S by its complement in $\langle 41 \rangle$ if necessary) that S has nine terms in $\langle 20 \rangle$. However, one easily checks that neither of the two nine term A sequences in $\langle 20 \rangle$ extends to a seventeen term A sequence in $\langle 41 \rangle$. Hence $r(41) = 16$. The proof that each of $r(42)$ and $r(43)$ is less than seventeen can be similarly completed by examining all A sequences with nine terms in $\langle 21 \rangle$. $\diamondsuit$

Lemma. If there exists a nonnegative integer c and a positive integer m such that the inequality $r(2n + c) \leq n$ holds for $n = m$, then it also holds for $n = m + 4$.

Proof. Since $r(8) = 4$, therefore $r(2m + 8 + c) \leq r(2m + c) + r(8) \leq m + 4$, which proves the lemma. $\diamondsuit$

The following theorem follows from the preceding lemma and induction on n.

Theorem 6. If there exists a nonnegative integer c and a positive integer m such that the inequality $r(2n + c) \leq n$ holds for $n = m$, $m + 1$, $m + 2$, and $m + 3$, then it holds for all $n \geq m$.

As the hypotheses of Theorem 6 are satisfied for $m = 8$ and $c = 3$, we obtain the following improvement of Theorem 1 in [1].

Corollary. For $n \geq 8$, $r(2n + 3) \leq n$.

The three question listed earlier as (a), (b), and (c) (including Szekeres' conjecture for $k \geq 5$) remain open at the moment.

References

[1] P. Erdös and P. Turan, *On some sequences of integers*, J. London Math. Soc. 11 (1936) 261-264.

[2] E. Szemeredi, *On sets of integers containing no k elements in arithmetic progression*, Acta Mathematica XXVII (1975) 199-245.