ON WEAKLY COMPLETE SEQUENCES FORMED BY THE GREEDY ALGORITHM

Michael P. Knapp[1]
Department of Mathematics and Statistics, Loyola University Maryland, Baltimore, Maryland
mpknapp@loyola.edu

Michael D. Paul
mdpaul314@gmail.com

Received: 5/3/11, Accepted: 12/9/11, Published: 1/2/12

Abstract

In this article, we consider increasing sequences of positive integers defined in the following manner. Let the initial terms a_1 and a_2 be given, and for any $n > 2$ define a_n to be the smallest integer greater than a_{n-1} which can not be written as a sum of (distinct) previous terms of the sequence. For various parametrized choices of the initial terms, we determine precisely the terms of the sequences obtained by this method. We also conjecture that for all choices of the initial terms, even in a more general setting, the terms of sequences defined in this manner have interesting patterns.

1. Introduction

In this article, we study patterns found in weakly complete sequences constructed by using a greedy algorithm. An increasing sequence $a_1, a_2, \ldots$ of positive integers is called *complete* (see [3, Chapter 8, Section 13]) if every positive integer can be written as a sum of distinct terms of the sequence. Probably the most well-known nontrivial example of such a sequence is the sequence of Fibonacci numbers (with one of the initial 1's deleted): $1, 2, 3, 5, 8, 13, \ldots$. It is well-known that every positive integer can be written as a sum of distinct terms of the Fibonacci sequence. For example, we have

$$\begin{aligned} 19 &= 8+5+3+2+1 \\ 20 &= 13+5+2 \\ 21 &= 21 \\ 22 &= 21+1 \end{aligned}$$

[1]Corresponding author.

and so on. Note that the number 21 can be considered to be a sum of one term of the sequence. The theory of complete sequences has been studied, and a theorem due to Brown [1] gives a simple condition which is both necessary and sufficient for a sequence to be complete.

An increasing sequence of integers is called *weakly complete* if every sufficiently large number can be written as a sum of distinct terms of the sequence, i.e., if there exists an integer N such that every integer $n \geq N$ can be written as a sum of distinct terms of the sequence. For example, it is easy to see that the sequence

$$1, 100, 102, 104, 106, 108, \ldots$$

is weakly complete, with $N = 100$.

In this article, we are interested in specifying the initial terms of the sequence and constructing a weakly complete sequence from this data using the greedy algorithm. To state this more formally, we allow s initial terms $a_1 < a_2 < \cdots < a_s$ to be specified, and also specify $N = a_s + 1$. Then for $n > s$, we define a_n to be the smallest number such that $a_n > a_{n-1}$ and such that a_n cannot be written as a sum of distinct previous terms of the sequence. For example, if we set $a_1 = 1$ and $a_2 = 5$, then we obtain the sequence

$$1, 5, 7, 9, 11, 29, 31, 89, 91, 269, 271, 809, 811, 2429, 2431, \ldots.$$

Similarly, starting with $a_1 = 1$ and $a_2 = 8$ leads to the sequence

$$1, 8, 10, 12, 14, 16, 54, 56, 58, 222, 224, 226, 894, 896, 898, \ldots,$$

and starting with $a_1 = 2$, $a_2 = 3$, and $a_3 = 9$ yields the sequence

$$2, 3, 9, 10, 16, 17, 23, 72, 73, 79, 296, 297, 303, 1192, 1193, 1199, \ldots.$$

In all three of these sequences, the numbers appear to quickly organize themselves into blocks with a common pattern. For example, in the first sequence each block has two elements with a difference of 2. Thus one can think of this sequence as being a union of the "shifted sets" $b_n + \{-1, 1\}$, $n = 0, 1, 2, \ldots$. Looking at the b_n is also interesting, as one quickly guesses the formula $b_n = 10 \cdot 3^n$. Similar formulas can be found for the other two sequences above, and in fact, if one tries any combination of initial terms, formulas like this appear to always exist. Based on these observations, we make the following conjecture.

Conjecture A. Suppose that $a_1, a_2, \ldots$ is an increasing, weakly complete sequence, which has been constructed by the greedy algorithm as described above. Then there exist numbers a and b and a set S such that the terms of the sequence are exactly the set

$$I \cup J \cup \bigcup_{i=0}^{\infty} (a \cdot b^i + S),$$

where I is the set of initial terms, J is a small set of "junk terms," and for any real number r, the set $r + S$ is defined by adding r to each element of S. Moreover, the junk terms all lie between the largest element of I and the smallest element of the union $\bigcup_{i=0}^{\infty}(a \cdot b^i + S)$.

Our goal in this article is to prove this conjecture for various choices of two initial elements of the sequence (i.e., $s = 2$), giving formulas for a, b, and S. In particular, we prove the three theorems below.

Theorem 1. *Suppose that $n \geq 4$ is an integer. Then the conjecture is true for sequences beginning with $a_1 = 1$ and $a_2 = n$. For these sequences, we have $I = \{1, n\}$ and $J = \{n + 2\}$. The numbers a and b are given by*

$$a = n + b + 2 \qquad \text{and} \qquad b = \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Finally, S is the set $\{2j - b + 2 : 0 \leq j \leq b - 2\}$.

Theorem 2. *Suppose that $n \geq 2$. Then the conjecture is true for sequences beginning with $a_1 = n$ and $a_2 = n + 1$. For these sequences, we have $I = \{n, n + 1\}$ and $J = \emptyset$. The numbers a and b are given by*

$$a = \frac{3n+2}{2} \qquad \text{and} \qquad b = n.$$

Finally, we have $S = \{j - \frac{n}{2} + 1 : 0 \leq j \leq n - 2\}$.

Theorem 3. *Suppose that $n \geq 2$. Then the conjecture is true for sequences beginning with $a_1 = n$ and $a_2 = 2n$. For these sequences, we have $I = \{n, 2n\}$ and $J = \{2n + 1, 2n + 2, \ldots, 3n - 1, 4n\}$. The numbers a and b are given by*

$$a = \frac{5n^2 + 6n}{2} \qquad \text{and} \qquad b = n.$$

Finally, we have $S = \{j - \frac{n}{2} + 1 : 0 \leq j \leq n - 2\}$.

We note that while these theorems are only stated for sufficiently large n, they are in fact true for all values of n, as the remaining cases can easily be dealt with individually.

In the proof of each of these theorems, it will be convenient to denote the set $a \cdot b^i + S$ by S_i. To prove the theorems, we first show that after the initial terms, the next terms of the sequence will be the elements of J, and then the elements of S_0. Following this, we proceed by induction on i, assuming that the initial elements of the sequence are $P_i = I \cup J \cup S_0 \cup \cdots \cup S_i$, and showing that the next terms are those in S_{i+1}. In order to do this, the key step is to consider the sets of numbers represented by P_{i-1} and exactly m elements of S_i. We will show that these sets are essentially sets of consecutive integers, and that they typically overlap each other. This will show that there are no additional elements of the sequence smaller than those in S_{i+1}. After this, it will not be difficult to show that the elements of S_{i+1} cannot be written as sums of the previous elements, and that they must therefore be the next block in the sequence.

Sections 2, 3, and 4 of this article are devoted to the proofs of the three theorems. Unfortunately, although all of the proofs use the same methods, we are not able to prove significant parts of the theorems at the same time. In Section 5, we discuss in general terms the sequences that arise from two initial terms, and for a broad range of initial values we conjecture formulae for the numbers a and b and the set S. We note that since this work was completed, this conjecture has been partially proven (see [2]).

We finish this introduction by briefly mentioning some of the notation and terminology used in this article. If S is a finite set, then we write $R(S)$ for the set of all possible sums of distinct elements of S, and write $R_m(S)$ for the set of all sums of (exactly) m distinct elements of S. We will use interval notation to represent sets of integers, so that $[a, b] = \{x \in \mathbb{Z} \ : \ a \leq x \leq b\}$. If we have two intervals $[a, b]$ and $[c, d]$ with $a \leq b, c$ and $b, c \leq d$, then we will say that these intervals are *contiguous* if $[a, b] \cup [c, d] = [a, d]$. This occurs if and only if $c \leq b + 1$. Finally, we will at times refer to "all numbers of the same parity in the interval $[a, b]$." When we do this, the endpoints a and b will have the same parity, and the expression is intended to refer to all numbers in the interval which have the same parity as the endpoints.

2. The Proof of Theorem 1

In our proof of Theorem 1, it will be convenient to define $s_{ij} = ab^i - b + 2 + 2j$ for $0 \leq j \leq b - 2$. Then the sets $S_i = \{s_{i0}, \ldots, s_{i,b-2}\}$ and $a \cdot b^i + S$ coincide. We can trivially see that the set I gives the first two terms of the sequence, and that the set J gives the third term. We claim that the next terms of the sequence are $n+4, n+6, \ldots, n+2b$, the elements of S_0. This follows easily from the observations that if s is an element of the sequence then $s + 1$ is not, and that the smallest possible sum of two elements of the sequence not involving 1 is $2n + 2$.

For any $i \in \mathbb{Z}_{\geq 0}$, define the set P_i as in the introduction, and assume that the sequence begins with P_k for some k. We need to show that the next terms of the sequence are the elements of S_{k+1}. First, we need to show that every integer between $n+1$ and $ab^{k+1} - b + 2$ (the smallest element of S_{k+1}) can be written as a sum of terms of P_k. By hypothesis, we already know that every integer up to $ab^k + b - 2$ (the largest element of S_k) can be represented, and so we only need to show that every integer in the interval $[ab^k + b - 1, ab^{k+1} - b + 1]$ has a representation.

By the definition of the sequence, we have

$$\{1\} \cup \left[n, ab^k - b + 1\right] \subseteq R(P_{k-1}).$$

Now, for any m consider the set $R_m(S_k)$. It is clear that the elements of $R_m(S_k)$ are all elements of the form

$$mab^k - mb + 2m + 2s,$$

where s can be any integer from $0 + 1 + \cdots + (m-1)$ to $(b-m-1) + \cdots + (b-2)$. In this way, we see that $R_m(S_k)$ contains all integers of the same parity in the interval $[mab^k - mb + m^2 + m, mab^k + mb - m^2 - m]$. By adding 1 to each of these integers, we see that every integer in the interval

$$A_m = \left[mab^k - mb + m^2 + m, mab^k + mb - m^2 - m + 1\right]$$

can be represented by elements of P_k.

Next, for any element $s \in R_m(S_k)$, we consider the set $s + R(P_{k-1})$. Clearly this set contains the interval $[s+n, s+ab^k - b+1]$. Keeping in mind that the elements s are consecutive numbers of the same parity, it is easy to see that these sets of sums overlap for consecutive values of s. Combining all these sets of sums, we see that every number in the interval

$$B_m = \left[mab^k - mb + m^2 + m + n, (m+1)ab^k + (m-1)b - m^2 - m + 1\right]$$

can be represented as a sum of elements of P_k.

Now, if we consider the sets A_1 and B_1, we see that we have representations for all numbers in the intervals $[ab^k - b + 2, ab^k + b - 1]$ and $[ab^k - b + 2 + n, 2ab^k - 1]$. We can see that

$$ab^k - b + 2 + n = \begin{cases} ab^k + b + 1, & \text{if } n \text{ is odd;} \\ ab^k + b + 2, & \text{if } n \text{ is even.} \end{cases}$$

Thus we see that we do not yet have a representation for $ab^k + b$, and if n is even then we are also missing a representation for $ab^k + b + 1$. To find representations for these integers, we begin by noting that

$$\sum_{s\in S_i} s = \sum_{j=0}^{b-2}(ab^i - b + 2 + 2j) = (b-1)ab^i, \tag{1}$$

and hence that

$$\begin{aligned} \sum_{s\in P_{k-1}} s &= 1 + n + (n+2) + \sum_{i=0}^{k-1}\sum_{s\in S_i} s \\ &= 2n + 3 + a(b^k - 1) \\ &= \begin{cases} ab^k + b, & \text{if } n \text{ is odd;} \\ ab^k + b + 1, & \text{if } n \text{ is even.} \end{cases} \end{aligned} \tag{2}$$

When n is even, we then see that

$$\sum_{s\in P_{k-1}-\{1\}} s = ab^k + b.$$

Thus we have representations for all numbers in the interval

$$C_1 = \left[ab^k - b + 2, 2ab^k - 1\right].$$

Similarly, if we consider the sets A_{b-2} and B_{b-2}, we have representations for all numbers in the intervals $[(b-2)ab^k - b + 2, (b-2)ab^k + b - 1]$ and $[(b-2)ab^k - b + 2 + n, (b-1)ab^k - 1]$. This time, we have

$$(b-2)ab^k - b + 2 + n = \begin{cases} (b-2)ab^k + b + 1, & \text{if } n \text{ is odd;} \\ (b-2)ab^k + b + 2, & \text{if } n \text{ is even.} \end{cases}$$

This time, we do not yet have a representation for $(b-2)ab^k + b$, and if n is even then we are also missing a representation for $(b-2)ab^k + b + 1$. Note now that $(b-3)ab^k \in R_{b-3}(S_k)$, and so is represented by our sequence. In the same way as above, we can now see that if n is odd, then

$$(b-2)b^k + b = (b-3)ab^k + \sum_{s\in P_{k-1}} s,$$

and if n is even then we have

$$(b-2)b^k + b = (b-3)ab^k + \sum_{s\in P_{k-1}-\{1\}} s$$

and

$$(b-2)b^k + b + 1 = (b-3)ab^k + \sum_{s\in P_{k-1}} s.$$

Thus we have representations for all numbers in the interval

$$C_{b-2} = \left[(b-2)ab^k - b + 2, (b-1)ab^k - 1\right].$$

Continuing, we see from (1) that $\sum_{s \in S_k} s = (b-1)ab^k$. If we add this to each element of $R(P_{k-1})$, then we obtain representations of $(b-1)ab^k + 1$ and every number in the interval $[(b-1)ab^k + n, ab^{k+1} - b + 1]$. To represent the numbers in between, if n is odd, then by (2), adding $\sum_{s \in P_{k-1}} s$ to each element of $R_{b-2}(S_k)$ gives representations of all numbers of the same parity from $(b-1)ab^k + 2$ to $(b-1)ab^k + n - 1$. If we delete the term 1 from each of these representations, then combined with the previous sentence we see that we can represent every integer in the interval $[(b-1)ab^k + 1, (b-1)ab^k + n - 1]$. Thus, we have found representations for every number in the interval

$$C_{b-1} = \left[(b-1)ab^k, ab^{k+1} - b + 1\right].$$

On the other hand, if n is even, then adding $\sum_{s \in P_{k-1}} s$ to each element of $R_{b-2}(S_k)$ gives representations of all numbers of the same parity from $(b-1)ab^k + 3$ to $(b-1)ab^k + n - 1$. Combining this with the integers obtained by deleting the element 1 from each of these representations, we have representations for all numbers in the interval $[(b-1)ab^k + 2, (b-1)ab^k + n - 1]$. Thus we again find that we can represent all elements of C_{b-1}.

If $b - 2 \geq 3$, which happens if and only if $n \geq 9$, we will now show that the intervals A_m and B_m are contiguous for $2 \leq m \leq b - 3$. We see from the above formulae that these intervals are contiguous if and only if

$$(mab^k + mb - m^2 - m + 1) - (mab^k - mb + m^2 + m + n) \geq -1,$$

that is, if and only if

$$2mb - 2m^2 - 2m + 2 - n \geq 0. \tag{3}$$

If we have $m = 1$ or $m = b - 2$, then the expression on the left-hand side equals $2b - 2 - n$, which is negative since $n \geq 9$ and $b \leq (n+1)/2$. On the other hand, if $m = 2$ or $m = b - 3$, then this expression equals $4b - 10 - n$. When we consider the definition of b and the fact that $n \geq 9$, we see that this expression must be non-negative. Considering (3) as a quadratic in m, we then see that one root must lie between 1 and 2, and the other between $b-3$ and $b-2$. Therefore (3) is satisfied for $2 \leq m \leq b-3$, and so the sets are in fact contiguous. Thus for each of these m, we have representations for all numbers in the interval

$$C_m = \left[mab^k - mb + m^2 + m, (m+1)ab^k + (m-1)b - m^2 - m + 1\right].$$

Note that this expression for C_m also holds when $m = 1, b-2$, and $b - 1$.

Finally, we wish to show that for any m, the sets C_m and C_{m+1} are contiguous. Subtracting the smallest element of C_{m+1} from the largest element of C_m, we see that this happens if and only if

$$2mb - 2m^2 - 4m - 1 \geq -1,$$

or in other words if and only if $2m(b-2-m) \geq 0$. Since $1 \leq m \leq b-2$, we see that this is always true. Since the sets C_m are contiguous for consecutive values of m, we see that we have found a representation for every number in the interval

$$C_1 \cup \cdots \cup C_{b-1} = \left[ab^k - b + 2, ab^{k+1} - b + 1\right].$$

Thus every number between $n+1$ and the smallest element of S_{k+1} can be represented by the elements of P_k.

To complete the induction, we need to show that the elements of S_{k+1} are the next elements of the sequence. First, we note that the difference between any two elements of S_{k+1} is at most $2b-4$, which is less than n. Hence the only way to use one element s of this interval to represent another is in the (obvious) representation of $s+1$. From this, it follows that if we can show that no element of S_{k+1} can be represented using only elements of P_k, then the elements of S_{k+1} are the next elements of the sequence.

To accomplish this, consider first a hypothetical representation of an element of S_{k+1} which uses at most $b-2$ elements of S_k. The largest possible sum of these $b-2$ elements is

$$\sum_{j=1}^{b-2}(ab^k - b + 2 + 2j) = (b-2)ab^k + b - 2.$$

Hence the smallest possible remainder, which must be represented using elements of P_{k-1} is

$$\left(ab^{k+1} - b + 2\right) - \left((b-2)ab^k + b - 2\right) = 2ab^k - 2b + 4.$$

However, from (2) we can see that

$$2ab^k - 2b + 4 > \sum_{s \in P_{k-1}} s.$$

Therefore any possible representation of an element of S_{k+1} which uses only elements of P_k must use all $b-1$ elements of S_k. By (1), the sum of these elements is $(b-1)ab^k$. Therefore if we try to represent the element $s_{k+1,j} = ab^{k+1} - b + 2 + 2j$, then we must use only elements of P_{k-1} to represent the remainder

$$s_{k+1,j} - (b-1)ab^k = ab^k - b + 2 + 2j.$$

However, this remainder is an element of S_k, and hence cannot be represented by elements of P_{k-1}. This shows that no element of S_{k+1} can be represented by elements of P_k, and hence these elements are the next terms of the sequence. This completes the proof of the theorem. □

3. The Proof of Theorem 2

This time, after the first two elements n and $n+1$, we clearly cannot represent any number smaller than $2n+1$. Therefore the next elements of the sequence must be $n+2, n+3, \ldots, 2n$, which are the elements of S_0. Defining P_i again as in the introduction, we assume that the sequence starts with P_k for some k. We need to show that the next elements of the sequence are the elements of S_{k+1}.

By the definition of the sequence, we know that all numbers in the interval

$$\left[n, an^k - 1 + \frac{n}{2}\right]$$

can be represented, since the right-hand endpoint is the largest element of S_k. Now, from the definition of the sequence, we have

$$\left[n, an^k - \frac{n}{2}\right] \subseteq R(P_{k-1}),$$

and it is easy to see that for any m we have

$$R_m(S_k) = \left[man^k - m\left(\frac{n-m-1}{2}\right), man^k + m\left(\frac{n-m-1}{2}\right)\right].$$

We now add each element of $R_m(S_k)$ to each element of $\left[n, an^k - \frac{n}{2}\right]$. Since both of these sets are sets of consecutive integers, we find representations for every number in the interval

$$B_m = \left[man^k - m\left(\frac{n-m-1}{2}\right) + n, (m+1)an^k + m\left(\frac{n-m-1}{2}\right) - \frac{n}{2}\right].$$

Consider now the intervals $R_1(S_k)$ and B_1. Comparing the endpoints, we see that these intervals are almost contiguous. We are missing only a representation for the number $an^k + \frac{n}{2}$. To represent this number, we can see that

$$\sum_{s \in S_i} s = (n-1)an^i \tag{4}$$

for all i, and hence

$$\sum_{s \in P_{k-1}} s = n + (n+1) + \sum_{i=0}^{k-1} \sum_{s \in S_i} s = 2n + 1 + \sum_{i=0}^{k-1} (n-1)an^i = an^k + \frac{n}{2}. \tag{5}$$

Thus we now have representations for all elements of the interval

$$C_1 = \left[an^k - \frac{n}{2} + 1, 2an^k - 1\right].$$

Suppose now that $n \geq 4$, and consider the intervals $R_{n-2}(S_k)$ and B_{n-2}. Then

$$R_{n-2}(S_k) = \left[(n-2)an^k - \frac{n}{2} + 1, (n-2)an^k + \frac{n}{2} - 1\right].$$

Adding all elements of $R_{n-2}(S_k)$ to all elements of $\left[n, an^k - \frac{n}{2}\right]$, we see that we can represent the entire interval

$$B_{n-2} = \left[(n-2)an^k + \frac{n}{2} + 1, (n-1)an^k - 1\right].$$

Comparing the last two intervals, we see that we are missing a representation for $(n-2)an^k + \frac{n}{2}$. However, from (5) we can write

$$(n-2)an^k + \frac{n}{2} = (n-3)an^k + \sum_{s \in P_{k-1}} s,$$

and since $(n-3)an^k \in R_{n-3}(S_k)$, this gives us a representation. Hence we can represent all numbers in the interval

$$C_{n-2} = \left[(n-2)an^k - \frac{n}{2} + 1, (n-1)an^k - 1\right].$$

Next, if we add all the terms of S_k together, the sum is $(n-1)an^k$ by (4). Adding this to each element of $\left[n, an^k - \frac{n}{2}\right]$ gives the interval

$$B_{n-1} = \left[(n-1)an^k + n, an^{k+1} - \frac{n}{2}\right].$$

We can represent the missing elements from $(n-1)an^k + 1$ to $(n-1)an^k + n - 1$ by noting that

$$(n-1)an^k + j = \left((n-2)an^k - \frac{n}{2} + j\right) + \sum_{s \in P_{k-1}} s,$$

and that when $1 \leq j \leq n-1$, the first term on the right is an element of $R_{n-2}(S_k)$. Hence we can represent all elements in the interval

$$C_{n-1} = \left[(n-1)an^k, an^{k+1} - \frac{n}{2}\right].$$

Now, if $n \geq 5$, then we still need to consider values of m between 2 and $n-3$. From our formulae for the intervals $R_m(S_k)$ and B_m, we can see that these two sets are contiguous whenever

$$mn - m^2 - m - n + 1 \geq 0. \tag{6}$$

Considering the polynomial above as a quadratic in m, we see that if $m = 1$ or $m = n-2$, then this polynomial is equal to -1, and that if $m = 2$ or $m = n-3$, then this polynomial is equal to $n-5$, which is non-negative. Thus one of the roots of this quadratic satisfies $1 < m \leq 2$ and the other satisfies $n-3 \leq m < n-2$. Since the coefficient of m^2 is negative, this shows that the polynomial is non-negative for $2 \leq m \leq n-3$. Hence the intervals $R_m(S_k)$ and B_m are contiguous for these values of m, whence we have representations for all terms in the interval

$$C_m = \left[man^k - m\left(\frac{n-m-1}{2}\right), (m+1)an^k + m\left(\frac{n-m-1}{2}\right) - \frac{n}{2}\right].$$

Note that this expression for C_m also holds when $m = 1, n-2$, and $n-1$.

To show that the intervals C_m are contiguous for consecutive values of m, we see that C_m and C_{m+1} are contiguous if and only if $mn - 2m - m^2 \geq 0$. Since the polynomial factors as $m(n-2-m)$, we see that this does hold for all m with $0 \leq m \leq n-2$. Since these intervals are contiguous, we may combine them and see that we have representations for every number from $an^k - \frac{n}{2} + 1$ to $an^{k+1} - \frac{n}{2}$. This shows that we can represent all numbers less than the elements of S_{k+1}.

To complete the proof, we must show that the elements of S_{k+1} are the next members of the sequence. First, note that since the difference between any two elements of S_{k+1} is at most $n-2$ and the sequence clearly cannot represent any number smaller than n, no element of S_{k+1} could ever be used in a representation of another member of this set. Therefore any hypothetical representation of an element of S_{k+1} must use only elements of P_k.

Now, let $s = an^{k+1} - \frac{n}{2} + 1 + j$ be an element of S_{k+1}, and consider a hypothetical representation of s which uses at most $n-2$ elements of S_k. Then the remainder, which must be represented using elements of P_{k-1}, is at least

$$\left(an^{k+1} - \frac{n}{2} + 1\right) - \sum_{i=1}^{n-2}\left(an^k - \frac{n}{2} + 1 + i\right) = 2an^k - n + 2.$$

But from (5), the sum of all the elements of P_{k-1} is only $an^k + \frac{n}{2}$, which is too small to represent this remainder. So any possible representation of s must use all of the elements of S_k. However, by (4) the sum of all the elements of S_k is $(n-1)an^k$, which means that the remainder to be represented by elements of P_{k-1} is $an^k - \frac{n}{2} + 1 + j$. But this number is an element of S_k, and so the definition of the sequence implies that it cannot be represented by elements of P_{k-1}. Since none of the elements of S_{k+1} can be represented using previous terms of the sequence, they must actually be the next terms. This completes the induction, and also the proof of the theorem. $\square$

4. The Proof of Theorem 3

While the details are different, this proof proceeds along the same lines as the previous two proofs. Unfortunately, since the "junk" set J is more complicated than before, the proof is a bit longer. Since the first two elements n and $2n$ sum to $3n$, it is clear that no number smaller than $3n$ can be represented, and so the next terms of the sequence are $2n+1, \dots, 3n-1$. By adding n to each of these numbers, we find representations of $3n+1, \dots, 4n-1$, so these numbers will not be in our sequence. To see that $4n$ cannot be represented by the elements we have so far, note that $4n$ cannot be written as a sum of any two of our elements, and that the smallest possible sum of three or more elements is $5n+1$, which is too large.

To show that the next terms of the sequence are the elements of S_0, we first show that the numbers from $4n+1$ to $a - \frac{n}{2}$ all can be represented by the elements of $I \cup J$. Consider the set $D_m = R_m([2n, 3n-1])$, and define $E_m = n + D_m$. Since both D_m and E_m are intervals for all m, we can see as before that we have

$$D_m = \left[2mn + \frac{m(m-1)}{2}, 3mn - \frac{m(m+1)}{2}\right]$$

and

$$E_m = \left[2mn + \frac{m(m-1)}{2} + n, 3mn - \frac{m(m+1)}{2} + n\right].$$

It is then easily seen that D_m and E_m are contiguous if and only if $(n-1-m)(m-1) \geq 0$, and that this is true for all m under consideration. Thus for each m we have a representation for each number in the interval

$$F_m = \left[2mn + \frac{m(m-1)}{2}, 3mn - \frac{m(m+1)}{2} + n\right].$$

Now, the intervals F_m and F_{m+1} are contiguous if and only if $mn - m^2 - m - n + 1 \geq 0$. This is the same as equation (6) from the previous section, where we saw that it holds for $2 \leq m \leq n-3$, provided that $n \geq 5$. Therefore we know that the elements in each of the three intervals

$$F_1 = [2n, 4n-1], \quad F_{\text{mid}} = \bigcup_{m=2}^{n-2} F_m = \left[4n+1, \frac{5n^2-7n}{2} - 1\right], \text{ and}$$

$$F_{n-1} = \left[\frac{5n^2-7n}{2} + 1, \frac{5n^2-3n}{2}\right]$$

are represented by the sequence. Note that if $n=2$ then we only have the interval F_1, and that if $n=3$, then we only have F_1 and F_{n-1}. If we add $4n$ to each element of each of these intervals (noting that this term has not been used in any of our representations so far), we find that we have representations for all elements in each of the intervals

$$G_1 = [6n, 8n-1], \quad G_{\text{mid}} = \left[8n+1, \tfrac{5n^2+n}{2} - 1\right], \text{ and}$$

$$G_{n-1} = \left[\tfrac{5n^2+n}{2} + 1, \tfrac{5n^2+5n}{2}\right] = \left[\tfrac{5n^2+n}{2} + 1, a - \tfrac{n}{2}\right].$$

The three sets $F_1, F_{\text{mid}}, F_{n-1}$ are almost contiguous, except that they do not contain the elements $4n$ and $(5n^2-7n)/2$. Since $4n$ is one of the elements of the sequence, we know that it has a representation. For the other number, we consider various possible values of n. If $n = 2$, then $(5n^2-7n)/2 = 3$, which by definition is not represented by the sequence. If $n = 3$ or $n = 4$, then we can calculate directly that $(5n^2-7n)/2$ is represented by the sequence. Finally, if $n \geq 5$, it is not hard to see that

$$8n+1 \leq \frac{5n^2-7n}{2} \leq \frac{5n^2+n}{2} - 1,$$

and hence that $(5n^2-7n)/2 \in G_{\text{mid}}$. Hence this number is represented by the sequence whenever $n \geq 3$.

Similarly, note that the three sets $G_1, G_{\text{mid}}, G_{n-1}$ together represent every number from $6n$ to $a - \frac{n}{2}$ except for $8n$ and $(5n^2+n)/2$. One can see that

$$\frac{5n^2+n}{2} = n + \sum_{s=2n}^{3n-1} s,$$

and so this number always has a representation. To represent $8n$, we can calculate representations explicitly when $n = 3$ and $n = 4$, and as above, we can show that $8n \in F_{\text{mid}}$ when $n \geq 5$. Thus this number is also represented by the sequence. Putting all these intervals together, we see that every number less than the smallest element of S_0 has a representation.

To show that the elements of S_0 are the next elements of the sequence, we see as before that no element of S_0 can be used in a representation of another element of this set. Now, if we try to represent an element of S_0 without using all of the elements of $[2n, 3n-1]$, the largest sum we can make is $a - \frac{n}{2}$, which is smaller than any element of S_0. Thus any such representation must use all of the elements of this interval. However, if we add n to all of these elements, we obtain $(5n^2+n)/2$, which is still smaller than any element of S_0, and if we add $4n$ to all of these elements, then we obtain $a + \frac{n}{2}$, which is larger than any element of S_0. Thus none of the elements of S_0 can be represented by previous terms of the sequence, and must be the next terms.

Now suppose we know that the sequence begins with P_k for some k. To show that all the numbers between $2n$ and the smallest element of S_{k+1} are represented by the sequence, we begin by noting that by definition, all the numbers from $2n$ to $an^k + \frac{n}{2} - 1$ are represented. We know that we have

$$\{n\} \cup \left[2n, an^k - \frac{n}{2}\right] \subseteq R(P_{k-1})$$

and we also have

$$R_m(S_k) = \left[man^k - m\left(\frac{n-m-1}{2}\right), man^k + m\left(\frac{n-m-1}{2}\right)\right].$$

By adding n to each element of $R_m(S_k)$, we have representations for each element in the interval

$$A_m = \left[man^k - m\left(\frac{n-m-1}{2}\right) + n, man^k + m\left(\frac{n-m-1}{2}\right) + n\right],$$

and by adding each element of $\left[2n, an^k - \frac{n}{2}\right]$ to each element of $R_m(S_k)$, we obtain the interval

$$B_m = \left[man^k - m\left(\frac{n-m-1}{2}\right) + 2n, (m+1)an^k + m\left(\frac{n-m-1}{2}\right) - \frac{n}{2}\right].$$

The above formulas give

$$R_1(S_k) = \left[an^k - \tfrac{n}{2} + 1, an^k + \tfrac{n}{2} - 1\right], A_1 = \left[an^k + \tfrac{n}{2} + 1, an^k + \tfrac{3n}{2} - 1\right],$$

and

$$B_1 = \left[an^k + \frac{3n}{2} + 1, 2an^k - 1\right],$$

but we have not yet found representations for $an^k + \frac{n}{2}$ or $an^k + \frac{3n}{2}$. To represent these, note that we have

$$\sum_{s \in S_i} s = (n-1)an^i \tag{7}$$

for any i. Thus we see that

$$an^k + \frac{3n}{2} = \sum_{s \in P_{k-1}} s \qquad \text{and} \qquad an^k + \frac{n}{2} = \sum_{s \in P_{k-1} - \{n\}} s.$$

Therefore we have representations for all numbers in the interval

$$C_1 = \left[an^k - \frac{n}{2} + 1, 2an^k - 1\right].$$

If $n \geq 4$, then the case $m = n-2$ is not included in the above, and we have

$$R_{n-2}(S_k) = \left[(n-2)an^k - \tfrac{n}{2} + 1, (n-2)an^k + \tfrac{n}{2} - 1\right],$$
$$A_{n-2} = \left[(n-2)an^k + \tfrac{n}{2} + 1, (n-2)an^k + \tfrac{3n}{2} - 1\right],$$

and

$$B_{n-2} = \left[(n-2)an^k + \frac{3n}{2} + 1, (n-1)an^k - 1\right].$$

We have not yet found representations for $(n-2)an^k + \frac{n}{2}$ or $(n-2)an^k + \frac{3n}{2}$. To represent these, note that $(n-3)an^k \in R_{n-3}(S_k)$. This yields the representations

$$(n-2)an^k + \frac{n}{2} = (n-3)an^k + \sum_{s \in P_{k-1} - \{n\}} s$$

and

$$(n-2)an^k + \frac{3n}{2} = (n-3)an^k + \sum_{s \in P_{k-1}} s.$$

Hence we can represent all the elements of the interval

$$C_{n-2} = \left[(n-2)an^k - \frac{n}{2} + 1, (n-1)an^k - 1\right].$$

By (7) with $i = k$, we can add $(n-1)an^k$ to each element of $R(P_{k-1})$ to obtain representations of $(n-1)an^k + n$ and the interval

$$\left[(n-1)an^k + 2n, an^{k+1} - \frac{n}{2}\right].$$

This time, we have not yet represented the numbers $(n-1)an^k + j$ or $(n-1)an^k + n + j$, for $1 \leq j \leq n-1$. To represent these, we note that the number $(n-2)an^k - \frac{n}{2} + j$ is an element of $R_{n-2}(S_k)$ for each j, and so we have

$$(n-1)an^k + j = \left((n-2)an^k - \frac{n}{2} + j\right) + \sum_{s \in P_{k-1} - \{n\}} s$$

and

$$(n-1)an^k + n + j = \left((n-2)an^k - \frac{n}{2} + j\right) + \sum_{s \in P_{k-1}} s.$$

This gives representations for the numbers in the interval

$$C_{n-1} = \left[(n-1)an^k, an^{k+1} - \frac{n}{2}\right].$$

If $n \geq 5$, then we still need to deal with the values of m from 2 to $n-3$, and we will show that for each of these n, the sets $R_m(S_k), A_m$, and B_m are contiguous. Considering our formulae for these sets, we see that the first two are contiguous if and only if we have $mn - m^2 - m - n + 1 \geq 0$. But this is the same as equation (6), which we have already shown holds for $2 \leq m \leq n-3$. Equation (6) also turns out to be the equation to decide whether A_m and B_m are contiguous, and so we know that these sets have this property also. Thus we have representations for all numbers in the interval

$$C_m = \left[man^k - m\left(\frac{n-m-1}{2}\right), (m+1)an^k + m\left(\frac{n-m-1}{2}\right) - \frac{n}{2}\right].$$

Note that this formula for C_m still holds when $m = 1, n-2$, and $n-1$.

Finally, we show that the sets C_m and C_{m+1} are contiguous. We can see that this is true if and only if we have $m(n-m-2) \geq 0$, which is clearly true when $1 \leq m \leq n-2$, as desired. Hence we can combine all the intervals C_m into one interval, and this shows that we have representations for all numbers smaller than the elements of S_{k+1}.

Now we need to show that the elements of S_{k+1} are the next elements of the sequence. As before, no element of this set can be used in a representation of another element from the set. Hence we need to show that no element of S_{k+1} can be represented by the elements of P_k. The largest possible sum of exactly $n-2$ elements of S_k is $(n-2)an^k + \frac{n}{2} - 1$. Hence, if we try to represent an element of S_{k+1} using at most $n-2$ elements of S_k, the remainder to be represented by elements of P_{k-1} is at least

$$\left(an^{k+1} - \frac{n}{2} + 1\right) - \left((n-2)an^k + \frac{n}{2} - 1\right) = 2an^k - n + 2.$$

However, we have seen that the sum of all of the elements of P_{k-1} is $an^k + \frac{3n}{2}$, which is smaller than this remainder. Thus any possible representation of $s = an^{k+1} - \frac{n}{2} + 1 + j \in S_{k+1}$ must use all of the elements of S_k. Since the sum of these elements is $(n-1)an^k$, the remainder to be represented by elements of P_{k-1} is $an^k - \frac{n}{2} + 1 + j$. But this number is an element of S_k, and so cannot be represented by elements of P_{k-1}. This completes the induction and proof of the theorem. $\square$

5. Some General Observations

In addition to proving the theorems above, we spent some time searching for a general formula for the terms of the sequence which would arise for any two initial terms. Unfortunately, this seems to be difficult. We did discover that the "right" way to look at these sequences appears to be to write the first two terms as n and $n+d$, and to consider the relationship between d and n. When $d = n$, then we have Theorem 3 of this article. If $d < n$, then the formulae for the terms of the sequences do appear to be related, and we make the following conjecture.

Conjecture B. If a sequence defined by our method begins with $a_1 = n$ and $a_2 = n + d$, with $d < n$, then the elements of the sequence are exactly the elements of the set

$$I \cup J \cup \bigcup_{i=1}^{\infty} (an^i + S)$$

(note that the indexed union starts with $i = 1$), where $I = \{n, n+d\}$, $J = \{n + d + 1, n + d + 2, \ldots, 2n + d - 1\}$, and

$$a = \frac{3n^2 + (2d-3)n - 4d + 2}{2(n-1)}.$$

Finally, the set S is given by

$$S = \{-c + j \ : \ 0 \leq j \leq n-1, j \neq d-1\},$$

where

$$c = \frac{n^2 - n - 2(d-1)}{2(n-1)}.$$

We note that since the work in this article was completed, this conjecture has been partially (when $d \leq n - 8$) proven by Fox and the first author [2].

The sequences with $d > n$ appear to be more complicated to study. Although for any particular sequence it appears possible to conjecture and prove a formula similar to the ones in this article, it seems to be more difficult to give a general formula in terms of n and d for the parameters a, b, and S in Conjecture A. While the formulae for these parameters do appear to have patterns, we are at this point unable to conjecture a general formula for them.

Acknowledgments. This research was conducted while the second author was a participant in the Hauber Summer Research Program at Loyola College during the summer of 2008 under the direction of the first author, and this article was written while the first author was enjoying the hospitality of the Universidade de Brasília while on sabbatical leave. We would like to thank both the Hauber program and the Universidade de Brasília for their generous support.

References

[1] J. L. Brown, Jr., *Note on complete sequences of integers*, Amer. Math. Monthly **68** (1961), 557-560.

[2] A. Fox and M. P. Knapp, *A note on weakly complete sequences*, submitted.

[3] R. Honsberger, *Mathematical Gems III*, Mathematical Association of America, Washington, D.C., 1985.