Journal of Integer Sequences, Vol. 22 (2019),
Article 19.1.2

A Sequence of Quasipolynomials Arising from Random Numerical Semigroups

Calvin Leng
Department of Mathematics
University of California Davis
Davis, CA 95616
USA
calleng@ucdavis.edu

Christopher O'Neill
Department of Mathematics and Statistics
San Diego State University
San Diego, CA 92182
USA
cdoneill@sdsu.edu

Abstract

A numerical semigroup is a cofinite subset of the non-negative integers that is closed under addition. For a randomly generated numerical semigroup, the expected number of minimum generators can be expressed in terms of a doubly-indexed sequence of integers, denoted $h_{n,i}$, that count generating sets with certain properties. We prove a recurrence that implies the sequence $h_{n,i}$ is eventually quasipolynomial when the second parameter is fixed.

1 Introduction

A *numerical semigroup* is an additive subsemigroup of $\mathbb{Z}_{\geq 0}$ with finite complement. The *semigroup generated by a set* $A = \{a_1, a_2, \ldots, a_k\}$ is the smallest additive subsemigroup of

$\mathbb{Z}_{\geq 0}$ containing A, namely

$$S = \langle A \rangle = \langle a_1, \ldots, a_k \rangle = \{a_1 x_1 + \cdots + a_k x_k : x_i \in \mathbb{Z}_{\geq 0}\}.$$

A generating set A is *minimal* if for all $x \in A$, we have $\langle A \rangle \neq \langle A \setminus \{x\} \rangle$. Every additive subsemigroup $S \subset \mathbb{Z}_{\geq 0}$ has a unique minimal generating set, and the *embedding dimension* of S, denoted $e(S)$, is the size of its minimal generating set (see [3] for a thorough introduction).

The authors of [2] introduce a model for randomly selecting a subsemigroup of $\mathbb{Z}_{\geq 0}$ that is similar to the Erdős-Renyi model for random graphs. Their model takes two inputs $M \in \mathbb{Z}_{\geq 1}$ and $p \in [0, 1]$, and randomly selects a generating set A that includes each integer $n = 1, 2, \ldots, M$ with independent probability p. Note that it is possible for a semigroup produced by this model to lack the "finite complement" property, but [2, Theorem 5] implies that asymptotically this happens with probability 0. This justifies the use of the term *random numerical semigroup model*.

As an example, if $M = 40$ and $p = 0.1$, then one possible set is $A = \{6, 9, 18, 20, 32\}$ (this is not unreasonable, as on average one would expect 4 generators to be selected). However, only 3 elements of A are minimal generators, since $18 = 9 + 9$ and $32 = 20 + 6 + 6$. As such, the resulting semigroup $S = \langle A \rangle = \langle 6, 9, 20 \rangle$ has embedding dimension 3.

One of the main results in [2] is that the expected number of minimal generators of a numerical semigroup S sampled with the above model can be expressed as

$$\mathbb{E}[e(S)] = \sum_{n=1}^{M} p(1-p)^{\lfloor n/2 \rfloor} \left(h_{n,0} + h_{n,1} p + h_{n,2} p^2 + \cdots \right),$$

where $h_{n,i}$ equals the number of sets $A \subset [1, n/2) \cap \mathbb{Z}$ with $|A| = i$ that minimally generate an additive subsemigroup of $\mathbb{Z}_{\geq 0}$ not containing n. Of interest is the asymptotic behavior of $\mathbb{E}[e(S)]$ for fixed p as $M \to \infty$. Although this is currently out of reach, $\mathbb{E}[e(S)]$ can be approximated for fixed M using the above formula, so long as $h_{n,i}$ is known for $n \leq M$.

$n = 68$:	1,	29,	249,	888,	1705,	2014,	1599,	888,	347,	91,	14,	1	
$n = 69$:	1,	31,	301,	1181,	2414,	2939,	2365,	1335,	535,	147,	25,	2	
$n = 70$:	1,	28,	248,	1012,	2218,	2873,	2431,	1414,	569,	155,	26,	2	
$n = 71$:	1,	34,	359,	1577,	3615,	4945,	4481,	2878,	1348,	453,	105,	15,	1
$n = 72$:	1,	25,	222,	893,	1923,	2498,	2138,	1267,	526,	147,	25,	2	
$n = 73$:	1,	35,	383,	1764,	4252,	6139,	5883,	4008,	2004,	725,	181,	28,	2
$n = 74$:	1,	34,	337,	1456,	3361,	4694,	4365,	2853,	1345,	453,	105,	15,	1
$n = 75$:	1,	32,	346,	1582,	3810,	5567,	5428,	3758,	1888,	684,	172,	27,	2
$n = 76$:	1,	33,	334,	1448,	3413,	5005,	4992,	3559,	1863,	705,	181,	28,	2

Figure 1: Values of $h_{n,i}$ for $n = 68$ through $n = 76$.

The doubly-indexed sequence $h_{n,i}$ is awaiting approval on OEIS as A319608, computed for $n \leq 90$. The sequence can also be viewed at the following URL:

https://gist.github.com/coneill-math/c2f12c94c7ee12ac7652096329417b7d

Figure 1 contains the values of $h_{n,i}$ for $n = 68, \ldots, 76$, where each row is comprised of $h_{n,0}, h_{n,1}, \ldots, h_{n,d_n}$ from left to right. The following facts about the sequence $h_{n,i}$ are known:

- $h_{n,i}$ is nonzero if and only if $n \geq 1$ and $0 \leq i \leq d_n = \lfloor n/2 \rfloor - \lfloor n/3 \rfloor$;
- $h_{n,0} = 1$;
- $h_{n,1} = \lfloor (n+1)/2 \rfloor - \tau(n)$, where $\tau(n)$ denotes the number of divisors of n; and
- The sum of the n^{th} row equals the number of irreducible numerical semigroups with Frobenius number n [1, 4], which appears in OEIS as A158206 [5].

Currently, computing the values of $h_{n,i}$ for large n is time-intensive; the fastest known algorithm computes the n^{th} row by first computing the set of irreducible numerical semigroups with Frobenius number n and utilizing the last bullet point above [4]. This computation takes 3 days for $n = 89$ on the authors' machines. The more values of $h_{n,i}$ that are known, the more accurately $\mathbb{E}[e(S)]$ can be approximated. Due to the limited known values of $h_{n,i}$, approximations computed with the currently known values still differ drastically from those obtained from experimental data; see [2, Table 2] for several sample estimates.

In this paper, we examine the combinatorics of the sequence $h_{n,i}$. Our main result is Corollary 2, which follows from the following recurrence and states that for fixed k the sequence h_{n,d_n-k} coincides with a polynomial in $n \gg 0$ whose coefficients are 6-periodic functions of n.

Theorem 1. *Fix $k \in \mathbb{Z}_{\geq 0}$, $b \in \{0,1,2\}$, and $m > 24k + 12 - 8b$ with $m \equiv b \pmod 3$. The recurrence*

$$h_{n,d_n-k} = \sum_{l=0}^{k} h_{m,d_m-l} \binom{d_n - d_m}{k - l}$$

holds for all $n \geq m$ satisfying $n \equiv b \pmod 3$.

A *quasipolynomial* is a function $q : \mathbb{Z} \to \mathbb{Z}$ such that

$$q(x) = c_0(x) + c_1(x)x + c_2(x)x^2 + \cdots + c_d(x)x^d$$

where each $c_i(x)$ is a periodic function. The *degree* of q, denoted $\deg q$, is the largest integer d for which c_d is not identically 0, and the *period* of q is the smallest integer p such that $c_i(x + p) = c_i(x)$ for every x and i.

Corollary 2. *For fixed k, the function $n \mapsto h_{n,d_n-k}$ coincides with a quasipolynomial*

$$c_k(n)n^k + \cdots + c_1(n)n + c_0(n)$$

with degree k, period 6, and leading coefficient

$$c_k(n) = \begin{cases} \frac{2}{k!6^k}, & \text{if } n \equiv 0, 1 \pmod 3; \\ \frac{1}{k!6^k}, & \text{if } n \equiv 2 \pmod 3), \end{cases}$$

for all $n > 24k + 12 - 8b$, where $b \in \{0, 1, 2\}$ with $n \equiv b \pmod 3$.

In the development of the proof of Theorem 1, we obtain an algorithm for computing the values $h_{n,i}$ appearing in Corollary 2 (Algorithm 19). Our algorithm has obtained $h_{n,i}$ values that were previously unknown. With the improved algorithm and Theorem 1, explicit quasipolynomials have been provided for h_{n,d_n-k} for each $k \leq 7$ (see Figure 2 for the quasipolynomials up to $k = 4$). Computing the quasipolynomial coefficients of h_{n,d_n-7} requires computing the value of e.g., $h_{183,d_{183}-7} = h_{183,23} = 6423209$, a task that would have been impossible with existing methods.

2 Setup

Unless otherwise stated, throughout the rest of the paper assume $n \in \mathbb{Z}_{\geq 1}$ and $b_n \in \{0, 1, 2\}$ with $n \equiv b_n \pmod 3$. Let

$$X_n = \left(\tfrac{n}{3}, \tfrac{n}{2}\right) \cap \mathbb{Z}.$$

Definition 3. Fix a set $A \subset \mathbb{Z}_{\geq 1}$. We say A *works* for $n \in \mathbb{Z}_{\geq 1}$ if

(i) $n \notin \langle A \rangle$,

(ii) $x < n/2$ for all $x \in A$, and

(iii) A minimally generates an additive subsemigroup of $\mathbb{Z}_{\geq 0}$.

In particular, $h_{n,i}$ equals the number of sets A with $|A| = i$ that work for n.

To motivate the next several definitions, recall from Figure 2 that for $n \geq 13$,

$$h_{n,d_n} = \begin{cases} 2, & \text{if } n \equiv 0, 1 \pmod 3; \\ 1, & \text{if } n \equiv 2 \pmod 3). \end{cases} \tag{1}$$

The set X_n works for n and $|X_n| = d_n$. Let $E_{0,n}$ and $E_{1,n}$ denote the remaining working sets for n of size d_n when $b_n = 0$ and $b_n = 1$, respectively. The key observation is that for all n,

$$\begin{aligned} X_n - \left\lfloor \tfrac{n}{3} \right\rfloor &= \{1, 2, \ldots, d_n\}, \\ E_{0,n} - \left\lfloor \tfrac{n}{3} \right\rfloor &= \{-1, 1, 3, 4, \ldots, d_n\}, \text{ and} \\ E_{1,n} - \left\lfloor \tfrac{n}{3} \right\rfloor &= \{0, 2, 3, \ldots, d_n\}. \end{aligned}$$

Since X_n contains every integer in the interval $(n/3, n/2)$, any other elements in a set counted by $h_{n,-}$ must lie in $\{1, 2, \ldots, \lfloor n/3 \rfloor\}$. Thus we thought of $\lfloor n/3 \rfloor$ as a sort of cutoff point. From this, it felt natural to express sets in terms of how offset the elements are from $\lfloor n/3 \rfloor$. This motivates the following.

$$h_{n,d_n} = \begin{cases} 2, & \text{if } n \equiv 0 \pmod 6 \text{ and } n \geq 18; \\ 2, & \text{if } n \equiv 1 \pmod 6 \text{ and } n \geq 7; \\ 1, & \text{if } n \equiv 2 \pmod 6 \text{ and } n \geq 2; \\ 2, & \text{if } n \equiv 3 \pmod 6 \text{ and } n \geq 15; \\ 2, & \text{if } n \equiv 4 \pmod 6 \text{ and } n \geq 10; \\ 1, & \text{if } n \equiv 5 \pmod 6 \text{ and } n \geq 5. \end{cases}$$

$$h_{n,d_n-1} = \begin{cases} \frac{1}{3}(n+3), & \text{if } n \equiv 0 \pmod 6 \text{ and } n \geq 42; \\ \frac{1}{3}(n+11), & \text{if } n \equiv 1 \pmod 6 \text{ and } n \geq 31; \\ \frac{1}{6}(n+16), & \text{if } n \equiv 2 \pmod 6 \text{ and } n \geq 26; \\ \frac{1}{3}(n+6), & \text{if } n \equiv 3 \pmod 6 \text{ and } n \geq 39; \\ \frac{1}{3}(n+8), & \text{if } n \equiv 4 \pmod 6 \text{ and } n \geq 34; \\ \frac{1}{6}(n+19), & \text{if } n \equiv 5 \pmod 6 \text{ and } n \geq 23. \end{cases}$$

$$h_{n,d_n-2} = \begin{cases} \frac{1}{36}(n^2+108), & \text{if } n \equiv 0 \pmod 6 \text{ and } n \geq 66; \\ \frac{1}{36}(n^2+16n+19), & \text{if } n \equiv 1 \pmod 6 \text{ and } n \geq 55; \\ \frac{1}{72}(n^2+26n+160), & \text{if } n \equiv 2 \pmod 6 \text{ and } n \geq 50; \\ \frac{1}{36}(n^2+6n+117), & \text{if } n \equiv 3 \pmod 6 \text{ and } n \geq 63; \\ \frac{1}{36}(n^2+10n-20), & \text{if } n \equiv 4 \pmod 6 \text{ and } n \geq 58; \\ \frac{1}{72}(n^2+32n+247), & \text{if } n \equiv 5 \pmod 6 \text{ and } n \geq 47. \end{cases}$$

$$h_{n,d_n-3} = \begin{cases} \frac{1}{648}(n^3-9n^2+342n-3240), & \text{if } n \equiv 0 \pmod 6 \text{ and } n \geq 90; \\ \frac{1}{648}(n^3+15n^2-69n+5885), & \text{if } n \equiv 1 \pmod 6 \text{ and } n \geq 79; \\ \frac{1}{1296}(n^3+30n^2+264n-1952), & \text{if } n \equiv 2 \pmod 6 \text{ and } n \geq 74; \\ \frac{1}{648}(n^3+315n-2268), & \text{if } n \equiv 3 \pmod 6 \text{ and } n \geq 87; \\ \frac{1}{648}(n^3+6n^2-132n+6200), & \text{if } n \equiv 4 \pmod 6 \text{ and } n \geq 82; \\ \frac{1}{1296}(n^3+39n^2+471n-863), & \text{if } n \equiv 5 \pmod 6 \text{ and } n \geq 71. \end{cases}$$

$$h_{n,d_n-4} = \begin{cases} \frac{1}{15552}(n^4-24n^3+828n^2-17280n+419904), & \text{if } n \equiv 0 \pmod 6 \text{ and } n \geq 114; \\ \frac{1}{15552}(n^4+8n^3-282n^2+24728n+413225), & \text{if } n \equiv 1 \pmod 6 \text{ and } n \geq 103; \\ \frac{1}{31104}(n^4+28n^3+204n^2-10256n+454912), & \text{if } n \equiv 2 \pmod 6 \text{ and } n \geq 98; \\ \frac{1}{15552}(n^4-12n^3+666n^2-12852n+374949), & \text{if } n \equiv 3 \pmod 6 \text{ and } n \geq 111; \\ \frac{1}{15552}(n^4-4n^3-300n^2+26528n-490112), & \text{if } n \equiv 4 \pmod 6 \text{ and } n \geq 106; \\ \frac{1}{31104}(n^4+40n^3+510n^2-8168n+426817), & \text{if } n \equiv 5 \pmod 6 \text{ and } n \geq 95. \end{cases}$$

Figure 2: Quasipolynomial expressions for h_{n,d_n-k} with $k = 0, 1, \ldots, 4$.

Definition 4. The *offset form* of a set $A = \{x_1, x_2, \ldots, x_k\} \subset \mathbb{Z}_{\geq 1}$ is the set

$$A_{(n)} = A - \lfloor n/3 \rfloor = \{x_1 - \lfloor n/3 \rfloor, x_2 - \lfloor n/3 \rfloor, \ldots, x_k - \lfloor n/3 \rfloor\}.$$

After expressing the sets we computed in offset form, we noticed that we could go one step further. We noticed that if we instead expressed sets in terms of how different they are from X_n and then take the offset form of the result, the expressions would be equal. This motivates the following.

Definition 5. A set $I \subseteq \mathbb{Z}$ is an *inserting set* for $n \in \mathbb{Z}_{\geq 1}$ if

$$I_{(n)} \subseteq \{-\lfloor n/3 \rfloor, \ldots, -1, 0\},$$

and a set $R \subseteq \mathbb{Z}$ is a *removing set* for n if

$$R_{(n)} \subseteq \{1, 2, \ldots, d_n\}.$$

An *RI-pair* for n is a pair (R, I) of a removing set R and an inserting set I.

There is a natural bijection between RI-pairs for n and the power set of $\{1, 2, \ldots, d_n\}$ given by the map

$$\varphi_n(R, I) = (X_n \setminus R) \cup I.$$

The inverse map is given by

$$A \mapsto (X_n \setminus A, A \setminus X_n).$$

Since φ_n gives a bijection between the two objects, we say the set *corresponding to* an RI-pair (R, I) is the set $\varphi_n(R, I)$, and vice-versa.

Theorem 1 follows from the fact that for fixed k and large n, every RI-pair (R, I) corresponding to a working set for n of size $d_n - k$ satisfies $I_{(n)} \subseteq \{p_n(k), \ldots, -1, 0\}$, where

$$p_n(k) = b_n - 2k - 1$$

only depends on n modulo 3 (Theorem 15). As a consequence, the restrictions on removal sets corresponding to a given insertion set are independent of the size of n in this case.

Example 6. If $n = 11$ and $k = 1$, then $h_{n,d_n-k} = h_{11,1} = 4$ and $X_n = \{4, 5\}$. The sets A with $|A| = 1$ that work for 11 are

$$\begin{aligned} A &= \{2\} = (X_n \setminus \{4,5\}) \cup \{2\}, & A &= \{4\} = (X_n \setminus \{5\}) \cup \{\}, \\ A &= \{3\} = (X_n \setminus \{4,5\}) \cup \{3\}, \text{ and} & A &= \{5\} = (X_n \setminus \{4\}) \cup \{\}. \end{aligned}$$

Theorem 11 classifies the possible RI-pairs that correspond to working sets for large n.

3 Strongly bounded sets

We begin by classifying the working sets for n that are strongly n-bounded (Definition 7). As it turns out, for k fixed and large n, every working set for n with size $d_n - k$ is strongly n-bounded (Theorem 16). Note that any strongly n-bounded set automatically satisfies parts (ii) and (iii) of Definition 3.

In what follows, we utilize the notation $mB = \{mb : b \in B\}$ for $B \subset \mathbb{Q}$ and $m \in \mathbb{Q}$.

Definition 7. We say a set $A \subset \mathbb{Z}_{\geq 1}$ is *strongly n-bounded* if $A \subset (n/4, n/2)$.

Proposition 8. *A strongly n-bounded set A works for n if and only if $b_n \notin 3A_{(n)}$.*

Proof. Any strongly n-bounded set automatically satisfies part (ii) and (iii) of Definition 3 since $x + y > \frac{n}{2} > z$ for any $x, y, z \in A$. As such, A works for n if and only if $n \notin \langle A \rangle$. Moreover, since A is strongly n-bounded, we have $x < n < y$ for any $x \in 2A$ and $y \in 4A$, so $n \in \langle A \rangle$ if and only if $n \in 3A$. The claim now follows from the fact that $n = 3\lfloor n/3 \rfloor + b_n$. $\square$

Definition 9. An RI-pair (R, I) is *compatible* for n (or, equivalently, R is *compatible with* I) if the corresponding set A satisfies $b_n \notin 3A_{(n)}$. The *removal degree* of an inserting set I, denoted $r(I)$, is given by

$$r(I) = \min\{|R| : (R, I) \text{ is compatible}\}$$

and the *removal degree* of an integer $\alpha \leq 0$ is given by $r(\alpha) = r(\{\alpha\})$.

Remark 10. Note that Proposition 8 does **not** imply that an RI-pair (R, I) compatible for n corresponds to a set A that works for n, as A need not be strongly n-bounded in general.

Theorem 11 classifies the RI-pairs compatible for n in terms of $I_{(n)}$ and $R_{(n)}$ by examining the different ways for three integers to sum to $b_n \in \{0, 1, 2\}$.

Theorem 11. *If A is a set and (R, I) is the corresponding RI-pair, then $b_n \notin 3A_{(n)}$ if and only if for all $\alpha \in I_{(n)}$, the following hold:*

(i) $b_n - 2\alpha \in R_{(n)}$*;*

(ii) $(b_n - \alpha)/2 \in R_{(n)}$ *if* $\alpha \equiv b_n \pmod 2$*;*

(iii) $b_n - \alpha - \beta \in R_{(n)}$ *for all* $\beta \in I_{(n)}$ *with* $\beta \neq \alpha$*; and*

(iv) $y \in R_{(n)}$ *or* $b_n - \alpha - y \in R_{(n)}$ *for all y satisfying* $1 \leq y < b_n - \alpha - y$*.*

Proof. If any of (i)-(iv) is violated for some $\alpha \in I_{(n)}$, then it is easy to check that $b_n \in 3A_{(n)}$. Conversely, suppose $b_n \in 3A_{(n)}$, meaning $\alpha + y + z = b_n$ for some $\alpha, y, z \in 3A_{(n)}$ with $\alpha \leq y \leq z$. Since $b_n \in \{0, 1, 2\}$, we must have $\alpha \leq 0$ and thus $\alpha \in I_{(n)}$. Let $S = \{\alpha, y, z\}$. If every element of S is nonpositive, then $\alpha = y = z = b_n = 0$ so (i) fails to hold and we are done. As such, at most 2 elements of S are nonpositive, so $z > 0$. Similarly, if $|S| = 1$, then $\alpha = y = z = b_n = 0$ so (i) fails to hold and we are done. This leaves four distinct cases:

- $|S| = 2$ and $y \leq 0$, in which case $y = \alpha$ and (i) fails to hold;
- $|S| = 2$ and $y > 0$, in which case $y = z$ and (ii) fails to hold;
- $|S| = 3$ and $y \leq 0$, in which case $\alpha \neq y$ and (iii) fails to hold; or
- $|S| = 3$ and $y > 0$, in which case $y \neq z$ and (iv) fails to hold.

This completes the proof. □

Remark 12. Given an inserting set I, Theorem 11 provides a systematic way to construct a removing set R such that the set A corresponding to (R, I) satisfies $b_n \notin 3A_{(n)}$. Most applications of Theorem 11 will involve starting with a set $R = \emptyset$ and systematically putting elements into R; see Example 20. Moreover, Theorem 11 yields a better-than-brute-force method of computing h_{n,d_n-k} for large n; see Algorithm 19.

Lemma 13. *We have*

$$r(\alpha) = 1 + \left\lceil \frac{b_n - \alpha - 1}{2} \right\rceil$$

for any integer $\alpha \leq 0$.

Proof. Fix $\alpha \leq 0$, let $I = \{\alpha\}$, and suppose R is a removing set that is minimal among all removing sets compatible with I. We will apply Theorem 11, noting that for fixed α, parts (i)-(iv) each require distinct elements to lie in $R_{(n)}$. Theorem 11(i) requires 1 element to lie in R, and Theorem 11(iv) forces $\lfloor (b_n - \alpha - 1)/2 \rfloor$ additional elements to lie in R. Since $|I| = 1$, Theorem 11(iii) is vacuously satisfied. This leaves Theorem 11(ii), which only requires an additional element to lie in R if $\alpha \equiv b_n \pmod 2$. This completes the proof. □

Lemma 14. *If* $A \subset \mathbb{Z}_{\geq 1}$ *corresponds to an RI-pair* (R, I) *that is compatible for* n*, then*

$$|A| \leq d_n + 1 - r(m),$$

where $m = \min I_{(n)}$.

Proof. Let $m = \min I_{(n)}$, and apply Theorem 11 to $\alpha = m$. Following the proof of Lemma 13, Theorem 11(i), (ii), and (iv) require $r(m)$ elements to lie in R, and Theorem 11(iii) requires R to contain an additional $|I| - 1$ elements. We conclude

$$|A| = d_n + |I| - |R| \leq d_n + |I| - r(m) - |I| + 1 = d_n + 1 - r(m),$$

as desired. □

Theorem 15. *Fix* $k \in \mathbb{Z}_{\geq 0}$*, and suppose* $A \subset \mathbb{Z}_{\geq 1}$ *corresponds to an RI-pair* (R, I) *that is compatible for* n*. If* $|A| \geq d_n - k$*, then*

$$I_{(n)} \subset \{p_n(k), p_n(k) + 1, \ldots, -1, 0\}.$$

Proof. Let $m = \min I_{(n)}$. By Lemma 14, we have

$$d_n - k \leq |A| \leq d_n + 1 - r(m),$$

meaning $k \geq r(m) - 1$. Applying Lemma 13, we obtain

$$k \geq 1 + \left\lceil \frac{b_n - m - 1}{2} \right\rceil - 1 \geq \frac{b_n - m - 1}{2}$$

which can then be rearranged to yield $m \geq b_n - 2k - 1 = p_n(k)$. $\square$

Theorem 16. *If $n > 24k + 12 - 8b_n$, then every set A with $|A| = d_n - k$ that works for n is strongly n-bounded.*

Proof. Fix a set A with $|A| = d_n - k$ that works for n. Theorem 15 implies

$$\min A - p_n(k) \geq \left\lfloor \frac{n}{3} \right\rfloor = \frac{n - b_n}{3} = \frac{\frac{1}{4}n - b_n}{3} + \frac{n}{4} > \frac{6k + 3 - 3b_n}{3} + \frac{n}{4} = 2k + 1 - b_n + \frac{n}{4}$$

meaning A is strongly n-bounded. $\square$

4 Proof of Theorem 1

We now have enough machinery to prove Theorem 1 and Corollary 2.

Proof of Theorem 1. Fix $n, m \in \mathbb{Z}$ satisfying $n \equiv m \pmod 3$ and $n \geq m > 24k + 12 - 8b_n$. Let $\mathcal{S}_k$ denote the set of k-subsets of $\{d_m + 1, d_m + 2, \ldots, d_n\}$. We will prove the claim combinatorially by constructing a bijection between

$$\mathcal{A} := \{A : A \text{ works for } n \text{ and } |A| = d_n - k\}$$

and

$$\mathcal{B} := \bigcup_{l=0}^{k} \big(\{A : A \text{ works for } m \text{ and } |A| = d_m - l\} \times \mathcal{S}_{k-l}\big).$$

Fix $A \in \mathcal{A}$, and let (R, I) denote the corresponding RI-pair. Write $R = R_1 \cup R_2$ with

$$(R_1)_{(n)} = \{\alpha \in R_{(n)} : 1 \leq \alpha \leq d_m\} \text{ and}$$
$$(R_2)_{(n)} = \{\alpha \in R_{(n)} : d_m + 1 \leq \alpha \leq d_n\},$$

and define $f : \mathcal{A} \to \mathcal{B}$ by

$$f(A) = (\varphi_m(R_1, I), (R_2)_{(n)}).$$

We first show f is well-defined. Let $l = |R_1| - |I|$. It is clear that $(R_2)_{(n)} \in \mathcal{S}_{k-l}$, and (R_1, I) is an RI-pair for m, so it remains to show that (R_1, I) is compatible for m. By Theorem 16, A is strongly n-bounded, so Theorem 15 implies $\min I_{(n)} \geq p_n(k) = p_m(k)$. The key observation is that the criteria in Theorem 11(i)-(iv) only involve $I_{(n)}$ and $R_{(n)}$, so tracing through each part, the fact that (R, I) is compatible for n implies (R_1, I) is compatible for m. Hence, f is well-defined.

To prove f is a bijection, we observe that basic set-theoretic arguments verify the map

$$((R_1, I), R_2) \mapsto (R_1 \cup R_2, I),$$

is the inverse function of f, thereby completing the proof. $\square$

Proof of Corollary 2. Fix $k \geq 0$ and $b \in \{0, 1, 2\}$, and let

$$m = \min\{x > 24k + 12 - 8b : x \equiv b \pmod 3\}.$$

For any $n \geq m$ satisfying $n \equiv b \pmod 3$, we obtain the expression

$$h_{n,d_n-k} = h_{m,d_m-k}\binom{d_n - d_m}{0} + h_{m,d_m-k+1}\binom{d_n - d_m}{1} + \cdots + h_{m,d_m}\binom{d_n - d_m}{k} \tag{2}$$

from Theorem 1, wherein each binomial coefficient is a polynomial in d_n of degree at most k. Since d_n is a quasilinear function of n with period 6, we conclude h_{n,d_n-k} is a quasipolynomial in n of degree k and period 6.

It remains to verify the leading coefficient of h_{n,d_n-k} has the desired form. The highest degree term in (2) is

$$h_{m,d_m}\binom{d_n - d_m}{k} = h_{m,d_m}\frac{(d_n - d_m) \cdot (d_n - d_m - 1) \cdots (d_n - d_m - k)}{k!}.$$

Combined with the fact that d_n has constant leading coefficient $1/6$, we obtain the leading coefficient $h_{m,d_m}/k!6^k$, and the claim now follows from examination of Figure 2. $\square$

Example 17. For fixed $k \geq 1$, the proof of Corollary 2 provides a slightly optimized method of computing the eventual quasipolynomial form of h_{n,d_n-k}. For example, let $k = 3$ and $b = 0$. In this case, $m = 87$, and consulting A319608 we see

$$h_{87,d_{87}-0} = 2, \qquad h_{87,d_{87}-1} = 31, \qquad h_{87,d_{87}-2} = 228, \qquad \text{and} \qquad h_{87,d_{87}-3} = 1055.$$

From here, we obtain

$$\begin{aligned} h_{n,d_n-3} = 2 \cdot \binom{d_n - d_{87}}{3} + 31 \cdot \binom{d_n - d_{87}}{2} \\ + 228 \cdot \binom{d_n - d_{87}}{1} + 1055 \cdot \binom{d_n - d_{87}}{0} \end{aligned}$$

for all $n \equiv 0 \pmod 3$ such that $n \geq 87$. Expanding binomial coefficients yields

$$h_{n,d_n-3} = \frac{1}{6}(2d_n^3 + 3d_n^2 + 19d_n - 12),$$

and substituting

$$d_n = \left\lfloor \frac{n-1}{2} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor = \begin{cases} \frac{1}{6}n - 1, & \text{if } n \equiv 0 \pmod 6; \\ \frac{1}{6}n - \frac{1}{2}, & \text{if } n \equiv 3 \pmod 6. \end{cases}$$

into the expression for h_{n,d_n-3}, we arrive at

$$h_{n,d_n-3} = \begin{cases} \frac{1}{648}(n^3 - 9n^2 + 342n - 3240), & \text{if } n \equiv 0 \pmod 6, n \geq 87; \\ \frac{1}{648}(n^3 + 315n - 2268), & \text{if } n \equiv 3 \pmod 6, n \geq 90. \end{cases}$$

Repeating this process for $b = 1, 2$ yields the function given in Figure 2.

Remark 18. It is interesting to note that the eventual quasipolynomial form of h_{n,d_n-3} would not be impossible to compute using the "standard" method of finding polynomial coefficients. Indeed, the values of $h_{n,i}$ have only been successfully computed for $n \leq 90$, and since the quasipolynomial behavior of h_{n,d_n-3} only holds for $n \geq 87$, the standard methods of finding the coefficients of a cubic require knowing $h_{87,d_{87}-3}, h_{90,d_{90}-3}, \ldots$, most of which have yet to be computed. The above method, on the other hand, only relies on $h_{87,d_{87}-i}$ for $0 \leq i \leq 3$.

Algorithm 19. The theory developed in Section 3 yields an algorithm to compute h_{m,d_m-k} for $m > 24k + 12 - 8b_m$. In particular, for each possible inserting set $I \subset \{p_m(k), \ldots, -1, 0\}$ for m, Theorem 11 determines precisely which removal sets R are compatible with I. Example 20 demonstrates the main idea of the algorithm.

The authors used a C++ implementation, now posted on Github at the following URL,

https://github.com/calvinleng97/rnsg-qp-coeffs ,

to compute the quasipolynomial functions in Corollary 2 up to $k = 7$, the last of which took 6 hours to complete.

Example 20. Suppose $n = 60$, and consider the insertion set $I = \{17, 18\}$. Theorem 11 provides a systematic method of constructing all removal sets R that are compatible with I. Since $\min I > n/4$, the resulting set will be strongly n-bounded. This ensures the resulting sets A corresponding to (R, I) will work for n.

We check every item of Theorem 11 with every element $\alpha \in I_{(n)}$ to construct $R_{(n)}$. We first compute the offset form

$$I_{(n)} = \{-3, -2\}$$

and initialize $R_{(n)} = \emptyset$. Note that $b_n = 0$ since $60 \equiv 0 \pmod 3$. We begin by applying Theorem 11(i)-(iii) to each $\alpha \in I_{(n)}$, since Theorem 11(iv) requires additional decisions.

For $\alpha = -2$, we see that $1, 4, 5 \in R_{(n)}$, and for $\alpha = -3$, we must have $R_{(n)} = \{1, 4, 5, 6\}$. Lastly, we deal with Theorem 11(iv), which is vacuously satisfied for $\alpha = -2$, and for $\alpha = -3$ implies either $1 \in R_{(n)}$ or $2 \in R_{(n)}$, the first of which is already required from above. As such, $R_{(n)} = \{1, 4, 5, 6\}$ yields a removal $R = \{21, 24, 25, 26\}$ that is compatible with I. Moreover, any removal set $R' \supset R$ is also compatible with I.

5 Future work

Although the quasipolynomials in Corollary 2 only hold for n sufficiently large, the machinery developed in Section 3 describes the sets counted by h_{n,d_n-k} and the relations between them as n varies. For n just below the start of quasipolynomial behavior, computations indicate the sets counted by h_{n,d_n-k} are simply those predicted by Theorem 11 that still minimally generate an additive subsemigroup of $\mathbb{Z}_{\geq 0}$. A better understanding of this phenomenon could allow Algorithm 19 to be extended to all $n \geq 1$, rather than just sufficiently large n.

Problem 21. Characterize the sets counted by h_{n,d_n-k} for all n in terms of those counted by h_{n,d_n-k} for n sufficiently large.

Algorithm 19 has the potential to be parallelized (with different threads handling different insertion sets), but the current implementation does not take advantage of this fact. Doing so would likely extend the current limits of computation, which would be especially useful if Problem 21 has a positive answer.

Problem 22. Write a parallelized implementation of Algorithm 19.

References

[1] J. Backelin, On the number of semigroups of natural numbers, *Math. Scand.* **66** (1990), 197–215.

[2] J. De Loera, C. O'Neill, and D. Wilbourne, Random numerical semigroups and a simplicial complex of irreducible semigroups, preprint, 2017, `https://arxiv.org/abs/1710.00979`.

[3] J. Rosales and P. García-Sánchez, *Numerical Semigroups*, Developments in Mathematics, Vol. 20, Springer-Verlag, 2009.

[4] J. Rosales, P. García-Sánchez, J. García-García, J. Jiménez Madrid, Fundamental gaps in numerical semigroups, *J. Pure Appl. Algebra* **189** (2004), 301–313.

[5] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`.

2010 *Mathematics Subject Classification*: Primary 20M14; Secondary 05E40.
Keywords: numerical semigroup, quasipolynomial.

(Concerned with sequences A158206 and A319608.)

Received October 9 2018; revised version received August 7 2019. Published in *Journal of Integer Sequences*, August 24 2019.

Return to Journal of Integer Sequences home page.