THE STRUCTURE THEOREM FOR SETS OF LENGTH FOR NUMERICAL SEMIGROUPS

GILAD MOSKOWITZ AND CHRISTOPHER O'NEILL

Abstract. For sufficiently nice families of semigroups and monoids, the structure theorem for sets of length states that the length set of any sufficiently large element is an arithmetic sequence with some values omitted near the ends. In this paper, we prove a specialized version of the structure theorem that holds for any numerical semigroup S. Our description utilizes two other numerical semigroups S_{M} and S_{m}, derived from the generators of S: for sufficiently large $n \in S$, the Apéry sets of S_{M} and S_{m} specify precisely which lengths appear in the length set of n, and their gaps specify which lengths are "missing". We also provide an explicit bound on which elements satisfy the structure theorem.

1. Introduction

Throughout this document, we let S denote a *numerical semigroup* (that is, an additively closed subset of $\mathbb{Z}_{\geq 0}$), and denote by $n_1, \ldots, n_k$ a generating set of S, i.e.,

$$S = \langle n_1, \ldots, n_k \rangle = \{q_1 n_1 + q_2 n_2 + \cdots + q_k n_k \mid q_1, \ldots, q_k \in \mathbb{Z}_{\geq 0}\}.$$

It is known that a numerical semigroup S is cofinite in $\mathbb{Z}_{\geq 0}$ if and only if $\gcd(S) = 1$, and it is common practice to assume this holds. It is also common practice to assume $n_1, \ldots, n_k$ comprise the unique minimal generating set of S. However, in this paper, we do **not** make either of these assumptions.

A *factorization* of $n \in S$ is an expression

$$n = q_1 n_1 + \cdots + q_k n_k$$

of n as a sum of generators of S, and the *length* of a factorization is the sum $q_1 + \cdots + q_k$. The *length set* of n is the set

$$\mathsf{L}_S(n) = \{q_1 + \cdots + q_k : q_1, \ldots, q_k \in \mathbb{Z}_{\geq 0} \text{ with } n = q_1 n_1 + \cdots + q_k n_k\}$$

of all possible factorization lengths of n. Define

$$\mathsf{M}_S(n) = \max \mathsf{L}_S(n) \qquad \text{and} \qquad \mathsf{m}_S(n) = \min \mathsf{L}_S(n).$$

When there can be no confusion, we often omit the subscripts and simply write $\mathsf{L}(n)$, $\mathsf{M}(n)$, and $\mathsf{m}(n)$, respectively.

The structure theorem for sets of length [21], a cornerstone of the factorization theory of atomic rings and semigroups, states that for any sufficiently large element n, the length set of n will be an almost arithmetical progression (that is, an arithmetic sequence with a few elements missing towards the beginning and end of the sequence).

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

The scope of the structure theorem goes well beyond that of numerical semigroups; it is known to hold for a broad family of semigroups and monoids, including finitely presented monoids, large families of Krull monoids, and others; see the monograph [23] for a thorough overview. In fact, one of the central themes in factorization theory is determining for which families of semigroups the structure theorem holds; see [22] for a detailed account.

We now state the structure theorem in the current context of numerical semigroups.

Structure Theorem for Sets of Lengths. *There exist integers t, t' and d such that for sufficiently large $n \in S$, there exist $A \subseteq [1, t]$ and $A' \subseteq [1, t']$ with the property that*

$$\mathsf{L}(n) = \{\mathsf{m}(n), \mathsf{m}(n) + d, \ldots, \mathsf{M}(n) - d, \mathsf{M}(n)\} \setminus \big((dA' + \mathsf{m}(n)) \cup (-dA + \mathsf{M}(n))\big).$$

In recent years, there has been an effort to specialize the structure theorem for semigroups of sufficiently high interest, stemming in part from its connections to open problems in additive combinatorics, such as the long-standing Narkiewicz conjecture [30]; see [14, 38] for recent progress and some related problems. These specializations generally concern which length sets are possible [2, 25, 27, 37], while others focus on refinements of the structure theorem, such as the unions of all sets of length [40], or a description of the "missing lengths", both locally for elements [9, 26] and globally for the semigroup as a whole [11, 24].

The main result of the present paper is Theorem 4.2, a refined structure theorem for sets of length for numerical semigroups, wherein we characterize the values d, t and t' in the theorem, and identify bijections between the sets $A \subseteq [1, t]$ and $A' \subseteq [1, t']$ of missing factorization lengths and sets of gaps in the semigroups

$$S_{\mathsf{M}} = \langle n_2 - n_1, n_3 - n_1, \ldots, n_k - n_1 \rangle \quad \text{and} \quad S_{\mathsf{m}} = \langle n_k - n_1, n_k - n_2, \ldots, n_k - n_{k-1} \rangle$$

respectively. This is best illustrated with an example.

Example 1.1. Let $S = \langle 5, 9, 12 \rangle$. Figure 1(b) depicts the "top" of the length sets $\mathsf{L}(100)$, ..., $\mathsf{L}(104)$, with filled black boxes indicating the "missing" lengths (the "A" sets in the structure theorem). Figure 1(a) depicts the elements of the semigroup

$$S_{\mathsf{M}} = \langle 9 - 5, 12 - 5 \rangle = \langle 4, 7 \rangle$$

with filled black boxes indicating the gap set $\mathbb{Z}_{\geq 0} \setminus S_{\mathsf{M}}$. Notice the identical positioning of the filled black boxes in each depiction; this relationship is the heart of Theorem 4.2. Figures 1(c) and 1(d) depicts a similar phenomenon (after a reflection) for the sets A' in the structure theorem for the numerical semigroup $S' = \langle 4, 6, 9 \rangle$.

Our result comes as part of a recently flurry of papers examining the factorization properties of large numerical semigroup elements, many of which turn out to be eventually periodic or quasipolynomial [5, 6, 7, 32]; see the survey [33] for details and [17, 18] for computational applications. The primary strength of our result is that it characterizes the missing lengths in terms of gap sets [1, 36, 34], which have been a central focus in the study of numerical semigroups since their inception [39].

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

0	**1**	**2**	**3**	**4**
0				4
		7	8	
	11	$\underline{12}$		$\underline{14}$
$\underline{15}$	$\underline{16}$		$\underline{18}$	19
20	21	22	23	24

(a) Elements of $S_{\mathsf{M}} = \langle 4, 7\rangle$ below 25, arranged by equivalence class mod 5.

100	**101**	**102**	**103**	**104**
20				20
		19	19	
	18	18		18
17	17		17	17
16	16	16	16	16

(b) Tops of the length sets of the elements $100, \ldots, 104 \in S = \langle 5, 9, 12\rangle$.

0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
0			3		5	6		8
9	10	11	12	13	14	15	16	17

(c) Elements of $S'_{\mathsf{m}} = \langle 3, 5\rangle$ below 18, arranged by equivalence class mod 9.

91	**92**	**93**	**94**	**95**	**96**	**97**	**98**	**99**
11		11	11		11			11
12	12	12	12	12	12	12	12	12

(d) Bottoms of the length sets of the elements $91, \ldots, 99 \in S' = \langle 4, 6, 9\rangle$.

FIGURE 1

The paper is organized as follows. After introducing a generalization of the Apéry set in Section 2, we prove in Section 3 that for sufficiently large n, the set A from the structure theorem is identical for $\mathsf{L}(n)$ and $\mathsf{L}(n + n_1)$, and the set A' is identical for $\mathsf{L}(n)$ and $\mathsf{L}(n + n_k)$ (Theorems 3.3 and 3.4, respectively). In Section 4, we prove Theorem 4.2, characterizing the sets A and A' in terms of the gaps of the semigroups S_{M} and S_{m}, respectively, as well as obtain an explicit bound on the $n \in S$ for which the structure theorem holds (Theorem 4.7). We also draw conclusions about realization questions akin to those considered in [9, 11, 24, 26] for other families of semigroups and monoids; see the discsussion in Remark 4.5.

2. A generalization of the Apéry set

The Apéry set of a numerical semigroup T is central to both theoretical [35] and computational [29] aspects of numerical semigroups; see [3] for a thorough overview. Usually defined with respect to an element $n \in T$, the Apéry set

$$\mathsf{Ap}(T; n) = \{m \in T : m - n \notin T\}$$

can be shown to consist of the first element of T in each equivalence class modulo n. In this section, we define a generalization of the Apéry set that allows $n \in \mathbb{Z}_{\geq 0}$ and $\gcd(T) > 1$. Other generalizations of the Apéry set have been studied, and while some are similar to our definition [12], most allow $\mathsf{Ap}(T; n)$ to contain more than one element of each equivalence class modulo n if $n \notin T$ [13, 15, 16, 19]. Moreover, none that the authors were able to find allowed $\gcd(T) > 1$. After verifying some basic properties

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

of $\mathsf{Ap}(T;n)$, we introduce a collection of sets that partition T, with $\mathsf{Ap}(T;n)$ as its foundation, that will play a key role in subsequent sections.

Notation 2.1. Throughout this section, let $S \subseteq \mathbb{Z}_{\geq 0}$ denote a cofinite numerical semigroup, let $d \in \mathbb{Z}_{\geq 1}$, and $T = dS \subseteq \mathbb{Z}_{\geq 0}$.

Definition 2.2. Fix $n \in \mathbb{Z}_{\geq 1}$. For each $i \in \{0, 1, \ldots, n-1\}$, let

$$a_i = \begin{cases} 0 & \text{if } T \cap \{i, i+n, i+2n, \ldots\} = \emptyset; \\ \min(T \cap \{i, i+n, i+2n, \ldots\}) & \text{otherwise.} \end{cases}$$

The *Apéry set* of T with respect to n as

$$\mathsf{Ap}(T;n) = \{a_i \mid i = 0, 1, \ldots, n-1\}.$$

Note that if T has finite complement and $n \in T$, then $\mathsf{Ap}(T;n)$ coincides with the usual definition of the Apéry set [3].

We briefly verify that under mild hypotheses, $\mathsf{Ap}(T;n)$ has some familiar properties.

Proposition 2.3. *For any $n \in \mathbb{Z}_{\geq 1}$, the elements of $\mathsf{Ap}(T;n)$ are distinct modulo n. Moreover, if $\gcd(d,n) = 1$, then $|\mathsf{Ap}(T;n)| = n$ and $\mathsf{Ap}(T;dn) = \mathsf{Ap}(T;n)$.*

Proof. The first claim follows from the definition of $\mathsf{Ap}(T;n)$ since $a_0 = 0$ and each nonzero a_i satisfies $a_i \equiv i \bmod n$. Next, fixing $z \in \mathbb{Z}$ so that $z + d\mathbb{Z}_{\geq 0} \subseteq T$, we see $y = nz + d$ satisfies $y \in T$ and $y \equiv d \bmod n$. This means if $\gcd(d,n) = 1$, then taking integer multiples of y reaches each equivalence class modulo n, so $\mathsf{Ap}(T;n)$ contains an element from each equivalence class modulo n, and thus $|\mathsf{Ap}(T;n)| = n$. For the final claim, it suffices to observe $\mathsf{Ap}(T;n) \subseteq \mathsf{Ap}(T;dn)$ (since $a - dn \equiv a - n \bmod n$ for each $a \in \mathsf{Ap}(T;n)$) and $|\mathsf{Ap}(T;dn)| \leq n$ (since each $a \in \mathsf{Ap}(T;dn)$ satisfies $d \mid a$). $\square$

Definition 2.4. Fix $n \in \mathbb{Z}_{\geq 1}$ and $j \geq 1$. The *j-th Apéry set* of T with respect to n is the set $\mathsf{Ap}_j(T;n)$ consisting of the j-th element of $\{a, a+n, \ldots\} \cap T$ for each $a \in \mathsf{Ap}(T;n)$. In particular,

$$\mathsf{Ap}_j(T;n) = \{a + kn \in T : a \in \mathsf{Ap}(T;n) \text{ and } |\{a, a+n, \ldots, a+kn\} \cap T| = j\},$$

where for each $a \in \mathsf{Ap}(T;n)$, there is a unique $k \in \mathbb{Z}_{\geq 0}$ such that $a + kn \in \mathsf{Ap}_j(T;n)$.

Example 2.5. Let $T = \langle 4, 7 \rangle$, whose first few elements are

$$T = \{0, 4, 7, 8, 11, 12, 14, 15, 16, 18, 19, 20, \ldots\}.$$

Under Definition 2.2, we have

$$\mathsf{Ap}(T;5) = \{0, 11, 7, 8, 4\},$$

and under Definition 2.4, we have $\mathsf{Ap}_1(T;5) = \mathsf{Ap}(T;5)$,

$$\mathsf{Ap}_2(T;5) = \{15, 16, 12, 18, 14\}, \quad \text{and} \quad \mathsf{Ap}_3(T;5) = \{20, 21, 22, 23, 19\}.$$

These comprise the first, second, and third integers in each column in Figure 1(a), respectively. For each $j \geq 3$, we see that $\mathsf{Ap}_{j+1}(T;5) = \mathsf{Ap}_j(T;5) + 5$ since $\mathbb{Z}_{\geq 19} \subseteq T$ and $\min \mathsf{Ap}_3(T;5) = 19$.

Lemma 2.6. *For each $n \in \mathbb{Z}_{\geq 1}$, $\mathsf{Ap}_1(T;n) = \mathsf{Ap}(T;n)$, and*

$$T = \bigcup_{j \geq 1} \mathsf{Ap}_j(T;n),$$

where the right hand side in the above equality is a disjoint union. If $n \in T$, then

$$\mathsf{Ap}_j(T;n) = \mathsf{Ap}(T;n) + (j-1)n = \{a + (j-1)n : a \in \mathsf{Ap}(T;n)\}.$$

Proof. All claims follow from induction on j and the fact that for each $a \in \mathsf{Ap}_j(T;n)$, choosing $k \in \mathbb{Z}_{\geq 1}$ minimal so that $a + kn \in T$ ensures $a + kn \in \mathsf{Ap}_{j+1}(T;n)$. $\square$

We close this section with one final definition and lemma we will use in Section 4.

Definition 2.7. The *Frobenius number* of T is

$$\mathsf{Frob}(T) = d(\max \mathsf{Ap}(S;n_1) - n_1).$$

where n_1 is the smallest generator of S. When $d = 1$, we obtain

$$\mathsf{Frob}(S) = \max\big(\mathsf{Ap}(S;n_1)\big) - n_1$$

which coincides with the traditional definition of the Frobenius number.

Lemma 2.8. *Suppose $T = \langle n_1, \ldots, n_k \rangle$. If $n \in T$, then*

$$\tfrac{1}{n_k} n \leq \mathsf{m}(n) \leq \tfrac{1}{n_k} n + (n_k - n_1) \qquad \textit{and} \qquad \mathsf{M}(n) \leq \tfrac{1}{n_1} n.$$

Proof. The first and last inequalities above follow from the fact that

$$(q_1 + \cdots + q_k)n_1 \leq q_1 n_1 + \cdots + q_k n_k \leq (q_1 + \cdots + q_k)n_k$$

for any factorization $n = q_1 n_1 + \cdots + q_k n_k$.

To prove the remaining inequality, first suppose $d = 1$, so that $S = T$. We consider two cases. If $n \leq n_1 n_k$, then

$$\frac{1}{n_1} n - \frac{1}{n_k} n = \frac{n(n_k - n_1)}{n_1 n_k} \leq n_k - n_1,$$

so every $\ell \in \mathsf{L}(n)$ satisfies the desired inequality. Next, suppose $n \geq n_1 n_k$, and write $n = n_1 n_k + q n_k - r$ with $q, r \in \mathbb{Z}_{\geq 0}$ and $0 \leq r < n_k$ by the division algorithm. We have $n_1 n_k - r \geq \mathsf{Frob}(S)$ by [34, Theorem 3.1.1], so since $\mathsf{M}(n_1 n_k - r) \leq n_k$, there exists a factorization of n of length $\ell \leq n_k + q$. As such,

$$\ell \leq n_k + q = n_1 + q + (n_k - n_1) = \tfrac{1}{n_k}(n + r) + (n_k - n_1) < \tfrac{1}{n_k} n + 1 + (n_k - n_1),$$

and since $\ell \in \mathbb{Z}$, we have $\ell \leq \frac{1}{n_k} n + (n_k - n_1)$.

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

Lastly, if $d > 1$, then applying the above argument to $S = \frac{1}{d}T$, there exists a factorization of $\frac{1}{d}n \in S$ of length at most

$$\tfrac{1}{n_k}n + \tfrac{1}{d}(n_k - n_1) \le \tfrac{1}{n_k}n + (n_k - n_1),$$

so there must also exist a factorization of this length for $n \in T$. □

3. Properties of maximum and minimum factorization length

The main results of this section are Theorems 3.3 and 3.4, wherein we classify the j-th maximum and minimum factorization lengths, respectively (Definition 3.2) for sufficiently large $n \in S$. These form the crux of our proof of Theorem 4.2, which makes explicit the phenomenon discussed in Example 1.1 and depicted in Figure 1. Although there is symmetry between the proofs of these two results, we include a proof for each, as there are some subtle differences in the arguments.

Notation 3.1. For the remainder of this paper, unless otherwise stated, fix a numerical semigroup $S = \langle n_1, n_2, \ldots, n_k \rangle$ that is cofinite in $\mathbb{Z}_{\ge 0}$. Write

$$S_{\mathsf{M}} = \langle n_2 - n_1, n_3 - n_1, \ldots, n_k - n_1 \rangle \qquad \text{and} \qquad \mathsf{Ap}_j(S_{\mathsf{M}}; n_1) = \{b_{0j}, b_{1j}, \ldots\}$$

where each $b_{ij} \equiv i \bmod n_1$. Analogously, write

$$S_{\mathsf{m}} = \langle n_k - n_1, n_k - n_2, \ldots, n_k - n_{k-1} \rangle \qquad \text{and} \qquad \mathsf{Ap}_j(S_{\mathsf{m}}; n_k) = \{c_{0j}, c_{1j}, \ldots\}$$

where each $c_{ij} + i \equiv 0 \bmod n_k$. Lastly, let $d = \gcd(S_{\mathsf{M}}) = \gcd(S_{\mathsf{m}})$, which can be shown to be equal by an elementary number theory argument.

Definition 3.2. Fix $j \in \mathbb{Z}_{\ge 1}$, and suppose $n \in S$ with $|\mathsf{L}(n)| \ge j$. Define $\mathsf{M}_j(n)$ and $\mathsf{m}_j(n)$ as the j-th largest and j-th smallest factorization lengths of n, respectively. In particular, $\mathsf{M}_1(n) = \mathsf{M}(n)$ and $\mathsf{m}_1(n) = \mathsf{m}(n)$.

Theorem 3.3. *If $j \ge 1$, then for all sufficiently large $n \in S$ with $n \equiv i \bmod n_1$,*

$$\mathsf{M}_j(n) = \frac{n - b_{ij}}{n_1}.$$

Proof. Fix $n \in S$, and write $n = pn_1 + i$ for $p, i \in \mathbb{Z}$ with $0 \le i < n_1$. Consider a factorization

$$n = q_1 n_1 + q_2 n_2 + \cdots + q_k n_k$$

of n, whose length is $\ell = q_1 + q_2 + \cdots + q_k$. Letting

$$b = (p - \ell)n_1 + i = n - \ell n_1 = q_2(n_2 - n_1) + \cdots + q_k(n_k - n_1), \tag{3.1}$$

we see $b \in S_{\mathsf{M}}$ and $b \equiv i \bmod n_1$, so $b = b_{ij}$ for some $j \ge 1$ by Lemma 2.6. Note i and j only depend on n and ℓ, and not on the specific values of $q_1, \ldots, q_k$. In particular, we have obtained a map

$$f : \mathsf{L}(n) \to \{b_{ij} : j \ge 1\} \qquad \text{given by} \qquad \ell \mapsto n - \ell n_1,$$

which this associates, to each length $\ell \in \mathsf{L}(n)$, an element $b_{ij} \in S_{\mathsf{M}}$.

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

Now, write $\mathsf{L}(n) = \{\ell_1 > \ell_2 > \cdots\}$. We claim that, for each fixed $h \geq 1$, if n is sufficiently large the map f induces a bijection

$$\{\ell_1, \ldots, \ell_h\} \rightsquigarrow \{b_{i1}, \ldots, b_{ih}\}. \tag{3.2}$$

Indeed, fix $j \leq h$. For any factorization

$$b_{ij} = Q_2(n_2 - n_1) + \cdots + Q_k(n_k - n_1)$$

of b_{ij} in S_{M}, we have

$$z = b_{ij} + (Q_2 + \cdots + Q_k)n_1 = Q_2 n_2 + \cdots + Q_k n_k \in S.$$

As such, choosing ℓ so that $n - \ell n_1 = b_{ij}$, if $n \in n_1\mathbb{Z}_{\geq 0} + z$, then

$$n = b_{ij} + \ell n_1 = (\ell - Q_2 + \cdots + Q_k)n_1 + Q_2 n_2 + \cdots + Q_k n_k$$

is a factorization of length $\ell \in \mathsf{L}(n)$, and $f(\ell) = b_{ij}$. In particular, this proves (3.2) is a bijection when n is sufficiently large. As a final step, choosing $\ell = \mathsf{M}_j(n)$ and solving (3.1) for ℓ then yields the desired equality. □

Theorem 3.4. *If $j \geq 1$, then for all sufficiently large $n \in S$ with $n \equiv i' \bmod n_k$,*

$$\mathsf{m}_j(n) = \frac{n + c_{i'j}}{n_k}.$$

Proof. Fix $n \in S$, and write $n = pn_k + i'$ for $p, i' \in \mathbb{Z}$ with $0 \leq i' < n_k$. If

$$n = q_1 n_1 + q_2 n_2 + \cdots + q_k n_k$$

is a factorization of n with length $\ell = q_1 + q_2 + \cdots + q_k$, then letting

$$c = (\ell - p)n_k - i' = \ell n_k - n = q_1(n_k - n_1) + \cdots + q_{k-1}(n_k - n_{k-1}), \tag{3.3}$$

we see $c \in S_{\mathsf{m}}$ and $c + i' \equiv 0 \bmod n_k$, so $c = c_{i'j}$ for some $j \geq 1$ by Lemma 2.6. This yields a map

$$f : \mathsf{L}(n) \to \{c_{i'j} : j \geq 1\} \qquad \text{given by} \qquad \ell \mapsto \ell n_k - n,$$

which this associates, to each length $\ell \in \mathsf{L}(n)$, an element $c_{i'j} \in S_{\mathsf{m}}$. Now, writing $\mathsf{L}(n) = \{\ell_1 < \ell_2 < \cdots\}$, we can show by a similar argument to the proof of Theorem 3.3 that for each fixed $h \geq 1$, if n is sufficiently large the map f induces a bijection

$$\{\ell_1, \ldots, \ell_h\} \rightsquigarrow \{c_{i'1}, \ldots, c_{i'h}\}.$$

Solving (3.3) for $\ell = \mathsf{m}_j(n)$ completes the proof. □

Remark 3.5. It was proven in [5, Theorems 4.2 and 4.3] that

$$\mathsf{M}(n + n_1) = \mathsf{M}(n) + 1 \qquad \text{and} \qquad \mathsf{m}(n + n_k) = \mathsf{m}(n) + 1$$

for sufficiently large $n \in S$. Corollary 3.6 (below) is a generalization of this result.

Another way to state this result is that there exist n_1- and n_k-periodic functions $f_S(n)$ and $g_S(n)$, respectively, such that for all sufficiently large $n \in S$,

$$\mathsf{M}(n) = \tfrac{1}{n_1}n + f_S(n) \qquad \text{and} \qquad \mathsf{m}(n) = \tfrac{1}{n_k}n + g_S(n).$$

19 Jun 2024 10:55:44 PDT
231116-ONeill Version 2 - Submitted to J. Comm. Alg.

The question was posed in [8, Project 3] to characterize the functions f_S and g_S in terms of the generators of S. Theorems 3.3 and 3.4 answer this question, expressing f and g in terms of the elements of $\mathsf{Ap}(S_\mathsf{M}; n_1)$ and $\mathsf{Ap}(S_\mathsf{m}; n_k)$, respectively. It was also asked in [8] whether it is possible $f_S = f_{S'}$ and $g_S = g_{S'}$ for distinct numerical semigroups S and S'; in addition to identfying when this occurs in terms of Apéry sets, our results provide a rubric for constructing examples. For instance, consider

$$S = \langle 10, 16, 44, 49, 51\rangle \qquad \text{and} \qquad S' = \langle 10, 16, 38, 44, 49, 51\rangle.$$

It is not hard to check $S_\mathsf{m} = S'_\mathsf{m} = \langle 2, 7\rangle$, so $\mathsf{Ap}(S_\mathsf{m}; 51) = \mathsf{Ap}(S_\mathsf{m}; 51)$ and thus $g_S = g_{S'}$. However, $S'_\mathsf{M} \setminus S_\mathsf{M} = \{28, 56, 67\}$, even though

$$\mathsf{Ap}(S_\mathsf{M}; 10) = \mathsf{Ap}(S'_\mathsf{M}; 10) = \{0, 41, 12, 53, 24, 45, 6, 47, 18, 39\}$$

and thus $f_S = f_{S'}$.

Corollary 3.6. *Fix $j \ge 1$. For all sufficiently large $n \in S$, we have*

$$\mathsf{M}_j(n + n_1) = \mathsf{M}_j(n) + 1 \qquad \textit{and} \qquad \mathsf{m}_j(n + n_k) = \mathsf{m}_j(n) + 1.$$

Proof. Apply Theorems 3.3 and 3.4. □

4. The refined structure theorem for numerical semigroups

In this section, we prove our main result: a refinement of the structure theorem for sets of length for numerical semigroups (Theorem 4.2). We also give an explicit bound on when the structure theorem holds (Theorem 4.7) and discuss the ramifications of this bound (Remark 4.9).

Notation 4.1. For each $i \in \{0, 1, \ldots, n_1 - 1\}$, let

$$A_i = \{r \in \mathbb{Z}_{\ge 1} : b_{i1} + rdn_1 \notin S_\mathsf{M}\}$$

and for each $i' \in \{0, 1, \ldots, n_k - 1\}$, let

$$A'_{i'} = \{r' \in \mathbb{Z}_{\ge 1} : c_{i'1} + r'dn_k \notin S_\mathsf{m}\}.$$

Theorem 4.2. *For all sufficiently large $n \in S$ with $n \equiv i \bmod n_1$ and $n \equiv i' \bmod n_k$,*

$$\mathsf{L}(n) = \{\mathsf{m}(n), \mathsf{m}(n) + d, \ldots, \mathsf{M}(n) - d, \mathsf{M}(n)\} \setminus \big((dA'_{i'} + \mathsf{m}(n)) \cup (-dA_i + \mathsf{M}(n))\big)$$

Proof. By the structure theorem for sets of length and [6, Proposition 2.9], there exist $t, t' \in \mathbb{Z}_{\ge 1}$ such that for all sufficiently large $n \in S$,

$$(\mathsf{m}(n) + d\mathbb{Z}) \cap [\mathsf{m}(n) + t'd, \mathsf{M}(n) - td] \subseteq \mathsf{L}(n)$$

Fix $n \in S$ with $n \equiv i \bmod n_1$ and $n \equiv i' \bmod n_k$ large enough that (i) the above holds, (ii) Theorem 3.3 holds for $j \le t$, and (iii) Theorem 3.4 holds for $j \le t'$.

First, suppose $\ell = \mathsf{M}(n) - rd$ for some $r \le t$, and let $b = b_{i1} + rdn_1$. If $r \notin A_i$, then $b \in S_\mathsf{M}$, meaning $b = b_{ij}$ for some j, and thus

$$\ell = \mathsf{M}(n) - rd = \frac{n - b_{i1} - rdn_1}{n_1} = \frac{n - b_{ij}}{n_1} = M_j(n) \in \mathsf{L}(n)$$

by Theorem 3.3. Conversely, if $\ell \in \mathsf{L}(n)$, then since $r \le t$, Theorem 3.3 implies

$$\ell = \mathsf{M}_j(n) = \frac{n - b_{ij}}{n_1}$$

for some j. Rearranging, we find

$$b_{ij} = n - \ell n_1 = n - (\mathsf{M}(n) - rd)n_1 = n - (n - b_{i1}) + rdn_1 = b_{i1} + rdn_1 = b$$

which means $b \in S_{\mathsf{M}}$ and thus $r \notin A_i$.

Now, by an analogous argument, if $\ell = \mathsf{m}(n) + dr'$ for some $r' \le t'$, then Theorem 3.4 implies $\ell \in \mathsf{L}(n)$ if and only if $r' \notin A'_{i'}$. This completes the proof. □

Remark 4.3. It was shown in [31, Corollary 5.5] that

$$|\mathsf{L}(n + n_1 n_k)| = |\mathsf{L}(n)| + \tfrac{1}{d}(n_k - n_1)$$

for sufficient large $n \in S$. This also an immediate consequence of Theorem 4.2.

In the remainder of this section, we identify an explicit bound on the "sufficiently large $n \in S$" in the statement of Theorem 4.2. First, we obtain the constants t and t' in the (unrefined) structure theorem for sets of length.

Proposition 4.4. *For each i and i', we have $A_i \subseteq A_0$ and $A'_{i'} \subseteq A'_0$. In particular,*

$$t = \max(A_0) \qquad \text{and} \qquad t' = \max(A'_0),$$

are the minimal values so that $A_i \subseteq [1, t]$ and $A'_{i'} \subseteq [1, t']$ for all i and i', respectively.

Proof. If $r \in \mathbb{Z}_{\ge 1} \setminus A_0$, then

$$b = b_{01} + dn_1 r = dn_1 r \in S_{\mathsf{M}}.$$

This means, for any i, we have

$$b + b_{i1} = b_{i1} + dn_1 r \in S_{\mathsf{M}},$$

so $r \in \mathbb{Z}_{\ge 1} \setminus A_i$. This proves $A_i \subseteq A_0$. An analogous argument proves each $A'_{i'} \subseteq A'_0$, and the remaining claims follow from the fact that

$$t = \max(A_0) \qquad \text{and} \qquad t' = \max(A'_0),$$

and from applying Theorems 3.3 and 3.4. □

Remark 4.5. In addition to yielding upper bounds

$$t \le \tfrac{1}{dn_1}\mathsf{Frob}(S_{\mathsf{M}}) \qquad \text{and} \qquad t' \le \tfrac{1}{dn_k}\mathsf{Frob}(S_{\mathsf{m}})$$

on t and t' in the structure theorem, Proposition 4.4 has implications on questions concerning of which combinations of "missing" lengths can occur, which have been considered for other families of semigroups [9, 11, 24, 26]. Letting

$$\mathcal{A}_S = \{A_i : 0 \le i \le n_1 - 1\} \qquad \text{and} \qquad \mathcal{A}'_S = \{A'_{i'} : 0 \le i' \le n_k - 1\},$$

Proposition 4.4 implies $\bigcup \mathcal{A}_S \in \mathcal{A}_S$ and $\bigcup \mathcal{A}'_S \in \mathcal{A}'_S$, the first known restrictions on $\mathcal{A}_S$ and $\mathcal{A}'_S$ for numerical semigroups. Additionally, under the mild assumption $\gcd(n_1, n_k) = 1$, there are infinitely many $n \in S$ for which

$$\mathsf{L}(n) = \{\mathsf{m}(n), \mathsf{m}(n) + d, \ldots, \mathsf{M}(n) - d, \mathsf{M}(n)\} \setminus \big((dA' + \mathsf{m}(n)) \cup (-dA + \mathsf{M}(n))\big)$$

for each pair $A \in \mathcal{A}_S$ and $A' \in \mathcal{A}'_S$. In particular, in order to classify the possible combinations of "missing" lengths from the "top" and "bottom" of the length sets of large $n \in S$, it suffices to classify $\mathcal{A}_S$ and $\mathcal{A}'_S$ independently.

In view of Remark 4.5, we state the following question, posed by Geroldinger in private communication with the second author and answered in the affirmative in [37] for the family of Krull monoids with finite class group.

Question 4.6. *Given $d, t, t' \in \mathbb{Z}_{\geq 1}$, does there exist a numerical semigroup S such that $\mathcal{A}_S$ and $\mathcal{A}'_S$ equal the power sets of $[1, t]$ and $[1, t']$, respectively, and $d = \gcd(S_{\mathsf{M}})$?*

Theorem 4.7. *Theorem 4.2 holds for all $n \geq n_k^2 - n_1^2$.*

Proof. Suppose $n \geq n_k^2 - n_1^2$. Fix $\ell \in [\frac{1}{n_k}n, \frac{1}{n_1}n] \cap \mathbb{Z}$. First, suppose

$$\ell \geq \frac{n}{(n_k + n_1)/2} = \frac{2n}{n_k + n_1},$$

and let $b = n - \ell n_1$. If $b \notin S_{\mathsf{M}}$, then $\ell + q \notin \mathsf{L}(n + qn_1)$ for all sufficiently large q by Theorem 4.2, so $\ell \notin \mathsf{L}(n)$. If $b \in S_{\mathsf{M}}$, then applying Lemma 2.8 to S_{M}, there exists a factorization of $b \in S_{\mathsf{M}}$ of length at most ℓ since

$$\begin{aligned}\ell(n_k - n_1) &\geq \frac{2nn_k}{n_k + n_1} - \ell n_1 = n - \ell n_1 + \frac{n(n_k - n_1)}{n_k + n_1} \\ &\geq n - \ell n_1 + (n_k - n_1)^2 \geq n - \ell n_1 + (n_k - n_1)(n_k - n_2)\end{aligned}$$

implies

$$\ell \geq \frac{n - \ell n_1}{n_k - n_1} + (n_k - n_2) = \frac{b}{n_k - n_1} + \big((n_k - n_1) - (n_2 - n_1)\big).$$

As such, we apply the correspondence in the proof of Theorem 3.3: if

$$b = q_2(n_2 - n_1) + \cdots + q_k(n_k - n_1)$$

is a factorization of $b \in S_{\mathsf{M}}$ of length at most ℓ, then

$$n = b + \ell n_1 = (\ell - q_2 - \cdots - q_k)n_1 + q_2 n_2 + \cdots + q_k n_k$$

is a factorization of $n \in S$ of length exactly ℓ, so $\ell \in \mathsf{L}(n)$.

Next, suppose

$$\ell \leq \frac{2n}{n_k + n_1},$$

and let $c = \ell n_k - n$. If $c \notin S_{\mathsf{m}}$, then $\ell + q \notin \mathsf{L}(n + qn_k)$ for all sufficiently large q by Theorem 4.2, so $\ell \notin \mathsf{L}(n)$. If $c \in S_{\mathsf{m}}$, then applying Lemma 2.8 to S_{m}, there exists a factorization of $c \in S_{\mathsf{m}}$ of length at most ℓ since

$$\begin{aligned}\ell(n_k - n_1) &\geq \ell n_k - \frac{2nn_1}{n_1 + n_k} = \ell n_k - n + \frac{n(n_k - n_1)}{n_k + n_1} \\ &\geq \ell n_k - n + (n_k - n_1)^2 \geq \ell n_k - n + (n_k - n_1)(n_k - n_{k-1})\end{aligned}$$

implies

$$\ell \geq \frac{\ell n_k - n}{n_k - n_1} + (n_k - n_{k-1}) = \frac{c}{n_k - n_1} + (n_k - n_{k-1})$$

Thus, as before, $\ell \in \mathsf{L}(n)$ by the correspondence in the proof of Theorem 3.4. □

Remark 4.8. If $n \geq n_k^2 - n_1^2$, the "top" and "bottom" of the length set (as described in Proposition 4.4) do not overlap. Indeed, by Theorem 4.2, if $\ell \in [\frac{1}{n_k}n, \frac{1}{n_1}n] \cap \mathbb{Z}$ with $\ell \notin \mathsf{L}(n)$, then

$$\ell \in [\mathsf{m}(n), \mathsf{m}(n) + \tfrac{1}{n_k}\mathsf{Frob}(S_{\mathsf{m}})] \cup [\mathsf{M}(n) - \tfrac{1}{n_1}\mathsf{Frob}(S_{\mathsf{M}}), \mathsf{M}(n)],$$

from which we obtain

$$\begin{aligned}\frac{n}{n_1} - \frac{n}{n_k} &= n\frac{n_k - n_1}{n_1 n_k} \geq (n_2 n_k - n_1 n_{k-1})\frac{n_k - n_1}{n_1 n_k} = \big(n_k(n_2 - n_1) + n_1(n_k - n_{k-1})\big)\frac{n_k - n_1}{n_1 n_k} \\ &= \tfrac{1}{n_1}(n_k - n_1)(n_2 - n_1) + \tfrac{1}{n_k}(n_k - n_1)(n_k - n_{k-1}) \geq \tfrac{1}{n_1}\mathsf{Frob}(S_{\mathsf{M}}) + \tfrac{1}{n_k}\mathsf{Frob}(S_{\mathsf{m}}),\end{aligned}$$

where the final inequality follows from [34, Theorem 3.1.1].

Remark 4.9. Given $n \in S$ and writing $\mathsf{L}(n) = \{\ell_1 < \cdots < \ell_r\}$, the *delta set* of n is

$$\Delta(n) = \{\ell_i - \ell_{i-1} : i \leq r\}.$$

It is known that $\Delta(n) = \Delta(n + \operatorname{lcm}(n_1, n_k))$ for all $n \geq 2kn_2n_k^2$ [10], and some effort has been made to refine this bound [18] and to compute delta sets explicitly [4, 20]. Theorem 4.7, in addition to providing an explicit bound for Corollary 3.6, identifies a bound on the start of periodicity for the delta set. Our bound appears to be better on average than the one obtained in [18] (in a sample of 10000 randomly selected numerical semigroups with $k \leq 10$ and $n_k \leq 10000$, our bound was better in roughly 75% of cases), as well as more concise (the one in [18] takes the better part of a page to write down).

Remark 4.10. If $n_2 - n_1 = d$, then S_{M} has no gaps, and so $A_i = \varnothing$ for all i. Analogously, if $n_k - n_{k-1} = d$, then S_{m} has no gaps and thus $A'_{i'} = \varnothing$ for all i'. In particular, if both of these are satisfied, then, for n sufficiently large, every length set is an arithmetic sequence with step size d. Note that the "sufficiently large n" is necessary even in this special case. For example, if $n = 26 \in S = \langle 5, 6, 13, 14 \rangle$, then $\mathsf{L}(n) = \{2, 5\}$. This is an improvement on [28, Corollary 3.6], which relates the length sets of element of a numerical semigroup generated by an arithmetic sequence to one in which "middle generators" are omitted.

Acknowledgements

The authors would like to thank Scott Chapman, Alfred Geroldinger, and Vadim Ponomarenko for their feedback and helpful conversations.

References

[1] F. Aicardi and L. Fel, *Gaps in nonsymmetric numerical semigroups*, Israel J. Math. **175** (2010), 85–112.

[2] J. Amos, S. Chapman, N. Hine, J. Paixão, *Sets of lengths do not characterize numerical monoids*, Integers **7** (2007), #A50.

[3] A. Assi and P. García-Sánchez, *Numerical semigroups and applications*, RSME Springer Series, 1. Springer, [Cham], 2016.

[4] T. Barron, C. O'Neill, and R. Pelayo, *On dynamic algorithms for factorization invariants in numerical monoids*, Mathematics of Computation **86** (2017), 2429–2447.

[5] T. Barron, C. O'Neill, and R. Pelayo, *On the set of elasticities in numerical monoids*, Semigroup Forum **94** (2017), no. 1, 37–50.

[6] C. Bowles, S. Chapman, N. Kaplan, D. Reiser, *On delta sets of numerical monoids*, J. Algebra Appl. **5** (2006) 1–24.

[7] S. Chapman, M. Corrales, A. Miller, C. Miller, and D. Patel, *The catenary and tame degrees on a numerical monoid are eventually periodic*, J. Aust. Math. Soc. 97 (2014), no. 3, 289–300.

[8] S. Chapman, R. Garcia, and C. O'Neill, *Beyond coins, stamps, and Chicken McNuggets: an invitation to numerical semigroups*, A Project-Based Guide to Undergraduate Research in Mathematics (ed. P. Harris, E. Insko, A. Wootton) Foundations of Undergraduate Research in Mathematics Series, Birkhäuser, Cham.

[9] S. Chapman, F. Gotti, and R. Pelayo, *On delta sets and their realizable subsets in Krull monoids with cyclic class groups*, Colloq. Math. 137 (2014), no. 1, 137–146.

[10] S. Chapman, R. Hoyer, and N. Kaplan, *Delta sets of numerical monoids are eventually periodic*, Aequationes mathematicae 77 **3** (2009) 273–279.

[11] S. Colton and N. Kaplan, *The realization problem for delta sets of numerical semigroups*, J. Commut. Algebra **9** (2017), no. 3, 313–339.

[12] M. de Abreu, M. Hernandes, and E. Marcelo, *On the value set of 1-forms for plane branches*, Semigroup Forum **105** (2022), no. 2, 385–397.

[13] M. Delgado, J. Farrán, P. García-Sánchez, and D. Llena, *On the weight hierarchy of codes coming from semigroups with two generators*, IEEE Trans. Inform. Theory **60** (2014), no. 1, 282–295.

[14] Y. Fan and S. Tringali, *Power monoids: a bridge between factorization theory and arithmetic combinatorics*, J. Algebra **512** (2018), 252–294.

[15] J. Farrán, P. García-Sánchez, B. Heredia, and M. Leamer, *The second Feng-Rao number for codes coming from telescopic semigroups*, Des. Codes Cryptogr. **86** (2018), no. 8, 1849–1864.

[16] J. Farrán and C. Munuera, *Goppa-like bounds for the generalized Feng-Rao distances*, International Workshop on Coding and Cryptography (WCC 2001) (Paris), Discrete Appl. Math **128** (2003), no. 1, 145–156.

[17] J. García-García, M. Moreno-Frías, and A. Vigneron-Tenorio, *Computation of the ω-primality and asymptotic ω-primality with applications to numerical semigroups*, Israel J. Math. 206 (2015), no. 1, 395–411.

[18] J. García-García, M. Moreno-Frías, and A. Vigneron-Tenorio, *Computation of delta sets of numerical monoids*, Monatshefte für Mathematik **178** (2015), no. 3 457–472.

[19] P. García-Sánchez and M. Leamer, *Huneke-Wiegand Conjecture for complete intersection numerical semigroup*, J. Algebra, 391 (2013), 114–124.

[20] P. García-Sánchez, C. O'Neill, and G. Webb, *On the computation of factorization invariants for affine semigroups*, Journal of Algebra and its Applications **18** (2019), no. 1, 1950019, 21 pp.

[21] A. Geroldinger, *A structure theorem for sets of lengths*, Colloq. Math., 78 (1998), pp. 225–259.

[22] A. Geroldinger, *Sets of lengths*, Amer. Math. Monthly **123** (2016), no. 10, 960–988.

[23] A. Geroldinger and F. Halter-Koch, *Nonunique factorization: Algebraic, Combinatorial, and Analytic Theory*, Chapman & Hall/CRC, Boca Raton, FL, 2006.

[24] A. Geroldinger and W. Schmid, *A realization theorem for sets of distances*, Journal of Algebra (2017) 481, 188–198.

[25] A. Geroldinger and W. Schmid, *A realization theorem for sets of lengths in numerical monoids*, Forum Mathematicum. **30** (2018), no. 5, 1111–1118.

[26] A. Geroldinger and P. Yuan, *The set of distances in Krull monoids*, Bull. Lond. Math. Soc. **44** (2012), no. 6, 1203–1208.

[27] A. Geroldinger and Q Zhong, *A realization result for systems of sets of lengths*, Israel J. Math. **247** (2022), no. 1, 177–193.

[28] S. Hyup Lee, C. O'Neill, and B. Van Over, *On arithmetical numerical monoids with some generators omitted*, Semigroup Forum **98** (2019), no. 2, 315–326.

[29] G. Márquez-Campos, I. Ojeda, and J. Tornero, *On the computation of the Apéry set of numerical monoids and affine semigroups* Semigroup Forum 91 (2015), no. 1, 139–158.

[30] W. Narkiewicz, *Finite Abelian groups and factorization problems*, Colloq. Math., 42 (1979) 319–330.

[31] C. O'Neill, *On factorization invariants and Hilbert functions*, Journal of Pure and Applied Algebra **221** (2017), no. 12, 3069–3088.

[32] C. O'Neill and R. Pelayo, *On the Linearity of ω-primality in Numerical Monoids*, J. Pure and Applied Algebra **218** (2014) 1620–1627.

[33] C. O'Neill and R. Pelayo, *Factorization invariants in numerical monoids*, Contemporary Mathematics **685** (2017), 231–249.

[34] J. Ramírez Alfonsín, *The Diophantine Frobenius problem*, Oxford Lecture Series in Mathematics and its Applications, 30. Oxford University Press, Oxford, 2005. xvi+243 pp. ISBN: 978-0-19-856820-9; 0-19-856820-7

[35] J.L. Ramírez Alfonsín and Ø. Rødseth, *Numerical semigroups: Apéry sets and Hilbert series*, Semigroup Forum 79 (2009), no. 2, 323–340.

[36] J. Rosales, P. García-Sánchez, J. García-García, J. Jiménez Madrid, *Fundamental gaps in numerical semigroups*, J. Pure Appl. Algebra 189 (2004), no. 1–3, 301–313.

[37] W. Schmid, *A realization theorem for sets of lengths*, J. Number Theory. 129 no. 5 (2009) 990–999.

[38] W. Schmid, *Characterization of class groups of Krull monoids via their systems of sets of lengths: a status report*, Number Theory and Applications: Proceedings of the International Conferences on Number Theory and Cryptography, Hindustan Book Agency, 2009, pp. 189–212.

[39] J. Sylvester, *Question 7382*, Mathematical Questions from the Educational Times **41** (1884) 21.

[40] S. Tringali, *Structural properties of subadditive families with applications to factorization theory*, Israel J. Math. **234** (2019), no. 1, 1–35.

Mathematics Department, San Diego State University, San Diego, CA 92182
Email address: gilad.moskowitz@gmail.com

Mathematics Department, San Diego State University, San Diego, CA 92182
Email address: cdoneill@sdsu.edu