

THE STABILISED DIAGONALS OF $n(g, t)$ AND IDEALS OF NUMERICAL SEMIGROUPS

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. Let $n(g, t)$ be the number of numerical semigroups of genus g and type t . Chappelon, Ramírez Alfonsín and Stamate proved that each diagonal $t = g - \ell$ of this array is eventually constant, $n(g, g - \ell) = |V(\ell)|$ for $g \geq 3\ell - 1$, where $V(\ell)$ is a set of balanced binary words of length 2ℓ prescribed by a condition on their difference set; they tabulated $|V(\ell)|$ for $\ell \leq 11$, asked for $|V(\ell)|$ in general, and proved only $|V(\ell)| \geq \ell^2 - 3\ell + 10$ — which at $\ell = 11$ reads 98 against the true value 3436. We show that the defining condition is a statement about numerical semigroups after all: the cotype of a word is the genus of the multiplier semigroup of an associated ideal, so that $V(\ell)$ is the set of subsets $I \subseteq \mathbb{N}$ with ℓ gaps whose multiplier semigroup has genus ℓ . From the largest fibre of this description we deduce $|V(\ell)| \geq 2^{k(\lfloor \ell/(k+1) \rfloor - 1)}$ for every $k \geq 1$, hence $|V(\ell)| \geq 2^{\ell - 2\sqrt{\ell} + O(1)}$, and a sharper union bound $2^{\ell - \log_2 \ell - O(1)}$; the same fibre turns out to be a sequence counted by Marzuola and Miller in 2010, in a paper on a different indexing of numerical sets, which upgrades the answer to $|V(\ell)| \geq \gamma_\infty \cdot 2^{\ell - 1}$ for their density constant $\gamma_\infty = 0.4844 \pm 0.0051$. We also compute $|V(12)| = 7234$, the first value beyond the published table, so that $n(g, g - 12) = 7234$ for every $g \geq 35$, and prove $n(35, 23) \geq 7234$ from explicit numerical semigroups; since the published table gives $n(33, 21) = 7214$, the value 7234 is not recoverable from it. All statements are machine-checked in Lean 4.

1. INTRODUCTION

Write $\mathbb{N} = \{0, 1, 2, \dots\}$. A *numerical semigroup* is a subset $S \subseteq \mathbb{N}$ with $0 \in S$, closed under addition, and with finite complement. The elements of $\mathbb{N} \setminus S$ are the *gaps* of S , written $\text{Gaps}(S)$; their number $g(S) = |\text{Gaps}(S)|$ is the *genus* and the largest of them is the *Frobenius number* $F(S)$. The *pseudo-Frobenius numbers* of S are

$$\text{PF}(S) = \{x \notin S : x + s \in S \text{ for every } s \in S \setminus \{0\}\},$$

and the *type* of S is $t(S) = |\text{PF}(S)|$. Finally

$$n(g, t) = \#\{S \text{ a numerical semigroup} : g(S) = g, t(S) = t\}.$$

The type is the number of generators of the canonical module of the associated semigroup ring, so the array $n(g, t)$ refines the classical count $n_g = \sum_t n(g, t)$ of numerical semigroups by genus (OEIS A007323 [4]) by an invariant of ring-theoretic origin.

Chappelon, Ramírez Alfonsín and Stamate [1] computed $n(g, t)$ for $1 \leq t \leq g \leq 33$ (their Table 4) and the analogous counts for the leaves of the semigroup tree (their Table 5), and studied the shape of the array. Their structural theorem concerns the diagonals $t = g - \ell$. To a numerical semigroup Λ of genus g and type $g - \ell$ with $g \geq 3\ell - 1$ they attach the *gap vector* $v(\Lambda) = (v_1, \dots, v_{2\ell}) \in \{0, 1\}^{2\ell}$ defined by $v_i = 1$ if and only if $g - \ell + i$ is a gap of Λ , and they set

$$\text{cotype}(v) = \#\{j - i : 1 \leq i < j \leq 2\ell, v_i = 0, v_j = 1\}.$$

By [1, Prop. 3.3] the vectors arising this way are exactly the *balanced* words (those with ℓ ones and ℓ zeros) of cotype ℓ ; write $V(\ell)$ for that set, whose elements are called *stable gap vectors*.

Theorem 1.1 ([1, Thm. 3.7]). *For every $\ell \geq 1$ and every $g \geq 3\ell - 1$ we have $n(g, g - \ell) = |V(\ell)|$.*

So each diagonal of the array freezes, and the frozen value is a purely combinatorial count with no semigroup and no genus g in it. The authors of [1] print $|V(\ell)|$ for $\ell \leq 11$ (their Table 1),

$$1, 3, 7, 15, 35, 78, 161, 367, 757, 1632, 3436,$$

and ask for $|V(\ell)|$ in general: “*It would be interesting to determine in terms of ℓ the stabilized value of $n(g, g - \ell)$... this process goes hand in hand with finding the set $V(\ell)$ of stable gap vectors for any ℓ .*” Towards it they classify five explicit families of stable gap vectors ([1, Prop. 3.9]) and count them ([1, Thm. 3.10]), obtaining

$$(1) \quad n(g, g - \ell) = |V(\ell)| \geq \ell^2 - 3\ell + 10 \quad (\ell \geq 6, g \geq 3\ell - 1),$$

a quadratic bound which at $\ell = 11$ reads 98 against the true value 3436. They close by inviting a computation pushed further than theirs.

Results. This paper answers the structural half of the question and replaces (1) by an exponential bound of the correct base, and it settles the first diagonal beyond the published table.

Stable gap vectors are ideals (Section 5). For $I \subseteq \mathbb{N}$ let $\mathcal{M}(I) = \{d \in \mathbb{N} : I + d \subseteq I\}$ be its *multiplier semigroup*, which is a numerical semigroup as soon as I contains a tail of $\mathbb{N}$. Reading a word v of length 2ℓ as the subset $I(v) \subseteq \mathbb{N}$ consisting of its zero positions together with everything from 2ℓ on, we prove that $\text{cotype}(v) = g(\mathcal{M}(I(v)))$, and hence that $V(\ell)$ is precisely the set of subsets of $\mathbb{N}$ with ℓ gaps whose multiplier semigroup has genus ℓ (Theorem 5.2). Combined with Theorem 1.1 this says: for $g \geq 3\ell - 1$, the numerical semigroups of genus g and type $g - \ell$ are in bijection with the pairs (T, J) where T is a numerical semigroup of genus ℓ and J is an ideal $T \subseteq J \subseteq \mathbb{N}$ whose multiplier semigroup is exactly T . The genus drops from g to ℓ . This is not a closed formula, and $|V(\ell)|$ “in terms of ℓ ” remains open; but it says which objects to count, and it is what makes the bounds below findable.

Exponential lower bounds (Section 6). The description locates the largest fibre: the semigroup $T = \langle \ell + 1, \dots, 2\ell + 1 \rangle$ with $\text{Gaps}(T) = \{1, \dots, \ell\}$, the maximal-type semigroup of [1, Prop. 3.8(a)], whose relevant ideals are indexed by the subsets $A \subseteq [1, \ell]$ satisfying a single closure-avoidance condition. A periodic family of such subsets gives

$$|V(\ell)| \geq 2^{k(\lfloor \ell/(k+1) \rfloor - 1)} \quad (k \geq 1), \quad \text{hence} \quad |V(\ell)| \geq 2^{\ell - 2\sqrt{\ell} + O(1)}$$

(Proposition 6.4), and a union bound over the top half of the window improves the exponent to $\ell - \log_2 \ell - O(1)$ (Proposition 6.5); at $\ell = 28$ these read 262 144 and 11 534 336 against the 710 of (1). The counting problem for that fibre is, as it turns out, not new: it is the “no small atoms” count of Marzuola and Miller [2], OEIS A158291, from a 2010 paper about numerical sets indexed by Frobenius number which [1] does not cite and which is not about the type distribution at all. Their density theorem then gives $|V(\ell)| \geq \gamma_\infty \cdot 2^{\ell-1}$ with $\gamma_\infty = 0.4844 \pm 0.0051$ — an exponential answer, of the right base, to the question posed in [1], obtained by connecting two papers sixteen years apart. We re-prove the two structural ingredients we need from [2] (Lemma 6.1 and Theorem 6.2) from the definitions.

The diagonal $\ell = 12$ (Section 7). We compute $|V(12)| = 7234$ (Theorem 7.1), one term past Table 1 of [1], so that by Theorem 1.1

$$n(g, g - 12) = 7234 \quad \text{for every } g \geq 35.$$

This value is not recoverable from the published data: Table 4 of [1] stops at $g = 33$, the $\ell = 12$ diagonal freezes only at $g = 3 \cdot 12 - 1 = 35$, and the published entry at the end of the table is $n(33, 21) = 7214 \neq 7234$. Independently of Theorem 1.1 we exhibit 7234 pairwise distinct numerical semigroups of genus 35 and type 23, giving $n(35, 23) \geq 7234$ (Theorem 7.2) — the lower half of the predicted equality, by a route that shares nothing with the stable-gap-vector computation.

Sections 3 and 4 contain the enumeration machinery and its calibration against the published tables, and Section 8 the sharpness statements and negative controls. Every statement in this paper has been formally verified; see Section 10.

2. PRELIMINARIES

Throughout, S and Λ denote numerical semigroups and g, t, ℓ are non-negative integers. Two elementary facts about the array $n(g, t)$ are used constantly.

Proposition 2.1. *For every numerical semigroup S we have $t(S) \leq g(S)$. Consequently $n(g, t) = 0$ whenever $t > g$: the array $n(g, t)$ is supported on the triangle $1 \leq t \leq g$.*

Proof. $\text{PF}(S) \subseteq \text{Gaps}(S)$ directly from the definition, and $\text{Gaps}(S)$ is finite. $\square$

Lemma 2.2. *Every gap of a numerical semigroup S of genus g is smaller than $2g$; equivalently $F(S) \leq 2g - 1$. The bound is attained at $g = 1$.*

Proof. Let x be a gap. If $i \in S$ and $i \leq x$ then $x - i$ is a gap, since otherwise $x = i + (x - i) \in S$. So $i \mapsto x - i$ is an injection of $S \cap [0, x]$ into $\text{Gaps}(S) \cap [0, x]$, whence $|S \cap [0, x]| \leq |\text{Gaps}(S) \cap [0, x]| \leq g$. Adding the two parts of the partition of $[0, x]$ gives $x + 1 \leq 2g$. At $g = 1$ the semigroup $\{0, 2, 3, 4, \dots\}$ has the gap $1 = 2 \cdot 1 - 1$. $\square$

Lemma 2.2 makes a numerical semigroup of genus g a subset of the finite window $[0, 2g)$, so we may and do encode it by its *gap mask*: the natural number m whose n -th binary digit is 1 exactly when n is a gap. Distinct masks give distinct semigroups.

The second ingredient is the semigroup tree of Rosales; the monograph of Rosales and García-Sánchez [3] is the standard reference for the theory of numerical semigroups, and the form of the tree used here is the one recalled in the introduction of [1].

Lemma 2.3 (parent step). *Let S be a numerical semigroup of genus $g \geq 1$ with Frobenius number F . Then $S \cup \{F\}$ is a numerical semigroup of genus $g - 1$. On gap masks: the mask of S is $2^F + m_0$ with $m_0 < 2^F$ the mask of $S \cup \{F\}$.*

Proof. Closure is the only point: for $s \in S$ with $s \neq 0$ we have $F + s > F$, hence $F + s \in S$; and $F + F > F$ likewise. Removing the largest gap drops the genus by one, and on masks it clears the top bit, F being $\lfloor \log_2 m \rfloor$. $\square$

Iterating Lemma 2.3 arranges all numerical semigroups into a rooted tree whose level g consists of those of genus g , the children of a node being obtained by deleting a minimal generator larger than the Frobenius number. We use it in the mask form: level g is obtained from level $g - 1$ by prepending a new top bit in every admissible position and re-testing the semigroup property.

3. $n(g, t)$ AS A FINITE COMPUTATION

Proposition 3.1. *For all g, t , the number $n(g, t)$ is finite and equals the number of masks in level g of the semigroup tree whose associated semigroup has type t . The level is generated without repetitions, and its length is n_g .*

Proof sketch. Soundness of the generation procedure — every mask it produces lies below 2^{2g} and encodes a numerical semigroup of genus g — is an induction on g using Lemma 2.2. Completeness is Lemma 2.3: a genus- g mask is $2^F + m_0$ with m_0 a genus- $(g - 1)$ mask, so it is produced from its parent. Absence of repetitions follows from the fact that the decomposition $m = 2^F + m_0$ with $m_0 < 2^F$ is unique, so distinct parents give disjoint sets of children and a fixed parent gives distinct children for distinct new top bits. Finally, every numerical semigroup of genus g is the semigroup of exactly one mask below 2^{2g} , again by Lemma 2.2. $\square$

The content of Proposition 3.1 is a change of search space. A blind scan of all gap sets inside the window of Lemma 2.2 costs 2^{2g} ; the tree visits n_g masks, which in the measured range grows by a factor of about 1.62 per genus. In the verified setting this is the difference between $g \leq 12$ and $g \leq 20$.

4. CALIBRATION AGAINST THE PUBLISHED TABLES

Before any value past the published range is asserted, the machinery of Section 3 is checked against everything in [1] that it can be compared with.

Proposition 4.1. *All 210 values $n(g, t)$ with $1 \leq t \leq g \leq 20$ agree with Table 4 of [1], and the 20 row totals $n_g = \sum_t n(g, t)$ agree with OEIS A007323. For instance*

$$(n(20, t))_{t=1}^{20} = (227, 812, 2810, 4422, 6967, 6909, 6024, 4043, 2523, 1307, \\ 709, 342, 161, 78, 35, 15, 7, 3, 1, 1),$$

with total $n_{20} = 37396$.

Method. One exhaustive search per genus $g \leq 20$: build level g of the tree as in Proposition 3.1, compute the type of each of its n_g masks once, and histogram. Each cell is then read off the resulting row. The row totals are an independent check, A007323 being known to $g = 80$ by other methods. $\square$

Proposition 4.2. $|V(\ell)| = 1, 3, 7, 15, 35, 78, 161, 367, 757, 1632, 3436$ for $\ell = 1, \dots, 11$, in agreement with Table 1 of [1].

Method. For each ℓ , an exhaustive search over all 4^ℓ words v of length 2ℓ , keeping those with ℓ ones and $\text{cotype}(v) = \ell$. The cotype is evaluated by the bitmask identity

$$(2) \quad d \in D(v) \iff ((2^{2\ell} - 1) \oplus v) \wedge (v \gg d) \neq 0,$$

where $\oplus$ is bitwise exclusive or, $\wedge$ bitwise and, and $\gg$ the right shift; the left factor is the mask of the zero positions of v . So $\text{cotype}(v)$ costs 2ℓ machine operations and no search. This computation shares nothing with the semigroup enumeration of Proposition 4.1 except the encoding of a set as a bit mask, and the two agree wherever Theorem 1.1 says they must: for instance $|V(4)| = 15 = n(11, 7)$ and $|V(7)| = 161 = n(20, 13)$, at the genera $11 = 3 \cdot 4 - 1$ and $20 = 3 \cdot 7 - 1$ where the respective diagonals freeze. $\square$

Proposition 4.3. *The worked examples printed in [1] reproduce: the semigroup $\langle 3, 7, 8 \rangle$ of the introduction, with $F = 5$ and its two descendants $\langle 3, 8, 10 \rangle$ and $\langle 3, 7, 11 \rangle$; the semigroup $\langle 7, 9, 10, 12, 13, 15 \rangle$ of Section 4, with $g = 8$, $t = 5$, $F = 11$ and descendants of types $4, 4, 2$; the three non-symmetric genus-6 leaves $\langle 5, 6, 8 \rangle$, $\langle 5, 6, 7 \rangle$, $\langle 4, 7, 9 \rangle$ with pseudo-Frobenius sets $\{7, 9\}$, $\{8, 9\}$, $\{5, 10\}$; the characterisation of [1, Prop. 3.8(a)] at genus 6; and the three cotype computations of [1, Ex. 3.4], namely $\text{cotype}(1^\ell 0^\ell) = 0$, $\text{cotype}(0^\ell 1^\ell) = 2\ell - 1$ and $\text{cotype}((01)^\ell) = \ell$.*

5. STABLE GAP VECTORS ARE IDEALS

The set $V(\ell)$ is defined by a condition on difference sets of binary words, and it is tempting to read that as a question with no semigroups left in it. The opposite is true, and this section makes it precise.

Definition 5.1. For $I \subseteq \mathbb{N}$ put $\mathcal{M}(I) = \{d \in \mathbb{N} : I + d \subseteq I\}$, the *multiplier semigroup* of I . If I contains every sufficiently large integer then $\mathcal{M}(I)$ is a numerical semigroup. For a word v of length 2ℓ put

$$I(v) = \{n \in \mathbb{N} : n \geq 2\ell \text{ or } v_n = 0\},$$

the set of zero positions of v together with the whole tail from 2ℓ on. (Positions are indexed from 0 here; the difference set, and hence the cotype, is unaffected.)

The gaps of $I(v)$ are exactly the one positions of v , so $I(v)$ has ℓ gaps precisely when v is balanced. The point of the section is that the second defining condition on $V(\ell)$ is also a genus.

Theorem 5.2. *For every word v of length 2ℓ ,*

$$\text{cotype}(v) = g(\mathcal{M}(I(v))).$$

Consequently $v \in V(\ell)$ if and only if $I(v)$ has exactly ℓ gaps and $\mathcal{M}(I(v))$ has genus ℓ , and

$$|V(\ell)| = \#\{I \subseteq \mathbb{N} : \mathbb{N} \setminus I \text{ is finite, } |\mathbb{N} \setminus I| = \ell, g(\mathcal{M}(I)) = \ell\}.$$

Proof sketch. For the first identity, $d \notin \mathcal{M}(I(v))$ means that some element of $I(v)$ leaves $I(v)$ when shifted by d ; inside the window that is exactly a position i with $v_i = 0$ and $v_{i+d} = 1$, and outside the window nothing can leave, because $I(v)$ contains the whole tail. So $\text{Gaps}(\mathcal{M}(I(v))) = D(v)$, and taking cardinalities gives $\text{cotype}(v) = g(\mathcal{M}(I(v)))$. The second statement follows because v balanced is the same as $I(v)$ having ℓ gaps.

For the third, the map $v \mapsto I(v)$ is injective on words of length 2ℓ , and it remains to see that the window $[0, 2\ell]$ is not a restriction: if $I \subseteq \mathbb{N}$ has at most ℓ gaps and $g(\mathcal{M}(I)) \leq \ell$, then every gap of I is smaller than 2ℓ . The argument is the one of Lemma 2.2 with the multiplier semigroup in the role of S : if x is a gap of I and $m \in \mathcal{M}(I)$ with $m \leq x$, then $x - m$ is again a gap of I , since otherwise $x = (x - m) + m \in I$. So $m \mapsto x - m$ injects $\mathcal{M}(I) \cap [0, x]$ into $\text{Gaps}(I) \cap [0, x]$, giving $|\mathcal{M}(I) \cap [0, x]| \leq \ell$, while $|[0, x] \setminus \mathcal{M}(I)| \leq g(\mathcal{M}(I)) \leq \ell$; adding gives $x + 1 \leq 2\ell$. Hence every I counted on the right is $I(v)$ for a unique word v of length 2ℓ , necessarily balanced and of cotype ℓ . $\square$

Remark 5.3. Since I is an ideal of $\mathcal{M}(I)$, and translating I so that $\min I = 0$ changes neither the number of gaps nor the multiplier semigroup, Theorem 5.2 can be restated as a classification: $|V(\ell)|$ counts the pairs (T, J) with T a numerical semigroup of genus ℓ and J an ideal $T \subseteq J \subseteq \mathbb{N}$ whose multiplier semigroup is exactly T . Combined with Theorem 1.1, for $g \geq 3\ell - 1$ the numerical semigroups of genus g and type $g - \ell$ are in bijection with those pairs. The genus drops from g to ℓ , which is the whole gain: the objects to be counted no longer grow with g .

Remark 5.4. The zero set of a stable gap vector is not itself a numerical semigroup — for $v = (1, 0, 0, 1) \in V(2)$ it is $\{1, 2\}$, which does not contain 0. A semigroup appears only after the tail is adjoined and the multiplier is taken.

6. LOWER BOUNDS FOR $|V(\ell)|$

6.1. The largest fibre. Grouping the pairs of Remark 5.3 by T suggests looking for the numerical semigroup of genus ℓ with the most ideals. Both the data and [1, Prop. 3.8(a)] point at the maximal-type semigroup

$$T_\ell = \langle \ell + 1, \ell + 2, \dots, 2\ell + 1 \rangle, \quad \text{Gaps}(T_\ell) = \{1, \dots, \ell\},$$

the unique numerical semigroup of genus ℓ and type ℓ . Its normalised ideals are $\{0\} \cup A \cup [\ell + 1, \infty)$ for an arbitrary $A \subseteq [1, \ell]$ — no closure condition survives, since any sum of two elements of $T_\ell \setminus \{0\}$ already exceeds ℓ — and the multiplier semigroup of such an ideal is exactly T_ℓ , rather than something larger, if and only if

$$(*) \quad \text{every } d \in A \text{ admits } s \in A \text{ with } s + d \leq \ell \text{ and } s + d \notin A.$$

Call such an A *good* and let $\iota(\ell)$ be the number of good subsets of $[1, \ell]$. By Theorem 5.2, $\iota(\ell) \leq |V(\ell)|$, and the enumeration of the 2^ℓ subsets of $[1, \ell]$ is quadratically cheaper than the enumeration of the 4^ℓ words used in Proposition 4.2. The values are

$$\iota(\ell) = 1, 2, 3, 6, 10, 20, 37, 74, 140, 280, 542, 1084, 2118, 4236, \dots$$

6.2. This fibre is a known sequence. The sequence ι is OEIS A158291 [4]: the number of *numerical sets* S — subsets of $\mathbb{N}$ containing 0 with finite complement, no closure assumed — whose *atom monoid* $\{n : n + S \subseteq S\}$ is $\{0\} \cup (\ell, \infty)$, i.e. which have *no small atoms*. The atom monoid is our multiplier semigroup and the criterion isolated in [2, Thm. 10] is verbatim (*). Marzuola and Miller determine this count recursively ([2, Thm. 11]) and prove that $\gamma_\ell = \iota(\ell)/2^{\ell-1}$ is decreasing ([2, Cor. 6]), hence that $\iota(\ell) \geq \gamma_\infty 2^{\ell-1}$ for its limit γ_∞ , which they locate at $\gamma_\infty = 0.4844 \pm 0.0051$. Their paper predates [1] by sixteen years, is indexed by Frobenius number rather than by genus or type, and is not cited there; what is new is the connection, which through Theorem 5.2 reads

$$(3) \quad n(g, g - \ell) = |\mathbf{V}(\ell)| \geq \iota(\ell) \geq \gamma_\infty \cdot 2^{\ell-1} \quad (g \geq 3\ell - 1),$$

an exponential lower bound for the quantity [1] asks about. We re-prove from the definitions the two structural statements of [2] that our formal development uses.

Lemma 6.1 ([2, Lem. 3]). *Call $d \in A$ bad if $1 \leq d \leq \ell$ and $s + d \in A$ for every $s \in A$ with $s + d \leq \ell$; thus A is good if and only if no d is bad. The set of bad elements is closed under addition inside $[1, \ell]$. Consequently, if A has a bad element then its largest bad element e satisfies $2e > \ell$; so A is good if and only if no e with $2e > \ell$ is bad.*

Proof. If d and e are bad and $d + e \leq \ell$ then $d + e \in A$ (apply badness of d to $s = e$) and for $s \in A$ with $s + d + e \leq \ell$ we get $s + d \in A$ and then $s + d + e \in A$; so $d + e$ is bad. If now e is the largest bad element and $2e \leq \ell$, then $e + e$ is bad and larger, a contradiction. $\square$

Theorem 6.2 ([2, Thm. 11]). $\iota(2m) = 2 \iota(2m - 1)$ for every $m \geq 1$.

Proof sketch. Let $\ell = 2m$. By Lemma 6.1 only $e > m$ can be bad. For such an e , the witnesses s that badness quantifies over satisfy $s \leq 2m - e < m$, and the elements $s + e$ it produces satisfy $s + e > m$: the middle position m occurs neither as a witness nor as a value, so it is invisible to badness. Two consequences. First, A is good if and only if $A \triangle \{m\}$ is good, a fixed-point-free involution of the good sets, so $\iota(2m)$ is even. Second, deleting the position m altogether — shifting the part of A above m down by one — matches badness at e for the window $[1, 2m]$ with badness at $e - 1$ for the window $[1, 2m - 1]$, hence is a bijection from the good subsets of $[1, 2m]$ not containing m onto the good subsets of $[1, 2m - 1]$. Together, $\iota(2m) = 2 \cdot \#\{A \text{ good}, m \notin A\} = 2 \iota(2m - 1)$. $\square$

Proposition 6.3. $\iota(19) = 130787$, and hence $|\mathbf{V}(20)| \geq 261574$; equivalently $n(g, g - 20) \geq 261574$ for every $g \geq 59$.

Proof. An exhaustive search over the 2^{19} subsets of $[1, 19]$, testing (*), gives $\iota(19) = 130787$. Theorem 6.2 at $m = 10$ turns this into $\iota(20) = 261574$, and $\iota(20) \leq |\mathbf{V}(20)|$. The bound (1) gives 350 here. (The value $\iota(19)$ is also recorded in OEIS A164048.) $\square$

6.3. Two self-contained bounds. The bound (3) is the strongest we know, but it is asymptotic and imported. The following two are effective, and both are proved outright rather than evaluated.

Proposition 6.4. For every $k \geq 1$ with $q = \lfloor \ell/(k+1) \rfloor \geq 2$,

$$|\mathbf{V}(\ell)| \geq 2^{k(q-1)}.$$

Choosing $k+1 \approx \sqrt{\ell}$ gives $|\mathbf{V}(\ell)| \geq 2^{\ell-2\sqrt{\ell}+O(1)}$, so $|\mathbf{V}(\ell)| = 2^{\ell(1-o(1))}$.

Proof. Fix $k \geq 1$ and put $q = \lfloor \ell/(k+1) \rfloor \geq 2$. For an arbitrary subset

$$F \subseteq \{j(k+1) + a : 1 \leq j \leq q-1, 1 \leq a \leq k\}$$

set $A = [1, k] \cup F$. No element of A is divisible by $k+1$. Given $d \in A$, let $s = (k+1) - (d \bmod (k+1))$; then $1 \leq s \leq k$, so $s \in A$, and $s + d$ is the least multiple of $k+1$ exceeding d , so $s + d \notin A$, while $s + d = (k+1)(\lfloor d/(k+1) \rfloor + 1) \leq (k+1)q \leq \ell$. So every such A satisfies (*). There are $2^{k(q-1)}$ choices of F , they give distinct sets A , and distinct good sets give distinct elements of $\mathbf{V}(\ell)$ by Theorem 5.2. The exponent $k(q-1) \approx \ell - \ell/(k+1) - k$ is maximised at $k+1 \approx \sqrt{\ell}$. $\square$

The deficit in the exponent of Proposition 6.4 is not slack that can simply be deleted: $|V(4)| = 15 < 16 = 2^4$, so no bound of the bare shape $|V(\ell)| \geq 2^\ell$ holds.

Proposition 6.5. *For every $r \geq 1$ with $2r + 1 \leq \ell$,*

$$|V(\ell)| \geq 2^{\ell-r} - (\ell - r - \lfloor \ell/2 \rfloor) 2^{\ell-2r-1},$$

hence $|V(\ell)| \geq 2^{\ell-\log_2 \ell - O(1)}$ on taking $2^r \approx \ell$. In particular $|V(28)| \geq 11\,534\,336$ and $|V(40)| \geq 34\,359\,738\,368$.

Proof. Count the subsets $X \subseteq [1, \ell - r]$, i.e. leave the top r positions of the window empty; there are $2^{\ell-r}$ of them. By Lemma 6.1 such an X can fail (*) only at some e with $2e > \ell$, and $e \leq \ell - r$ since $e \in X$. If it does fail at e , then X must avoid the whole window $[\ell - r - e + 1, \ell - e]$: an element i of that window would force $i + e \in [\ell - r + 1, \ell]$, which lies outside X by construction, so i would witness e . Hence the failing X are covered by $\ell - r - \lfloor \ell/2 \rfloor$ families — one for each admissible e — each of size at most $2^{\ell-2r-1}$, one element and r window positions being pinned in each. Subtracting and applying Theorem 5.2 gives the bound. For $\ell = 28$ and $\ell = 40$ take $r = 4$. $\square$

Table 1 compares the available bounds. Proposition 6.5 beats Proposition 6.4 at every $\ell \geq 4$ and both beat (1) by an arbitrarily large factor; all three are beaten by (3), whose merit is a matching base and whose defect is that it is neither effective nor, in the form proved in [2], a statement about $V(\ell)$.

ℓ	(1)	Prop. 6.4	Prop. 6.5	$\gamma_\infty 2^{\ell-1}$	$ V(\ell) $
11	98	16	256	496	3436
20	350	4096	73728	$2.5 \cdot 10^5$	2,376,348 [†]
28	710	262144	11,534,336	$6.5 \cdot 10^7$	681,114,026 [†]
40	1490	268,435,456	34,359,738,368	$2.7 \cdot 10^{11}$	—

TABLE 1. Lower bounds for $|V(\ell)| = n(g, g - \ell)$, $g \geq 3\ell - 1$. The γ_∞ column is evaluated at the central estimate $\gamma_\infty = 0.4844$ of [2]. Entries marked [†] come from the unverified enumeration of Remark 7.4.

7. THE DIAGONAL $\ell = 12$

Table 1 of [1] stops at $\ell = 11$ and its Table 4 at $g = 33$. The $\ell = 12$ diagonal freezes at $g = 3 \cdot 12 - 1 = 35$, so its stabilised value lies beyond both.

Theorem 7.1. $|V(12)| = 7234$. *Consequently, by Theorem 1.1,*

$$n(g, g - 12) = 7234 \quad \text{for every } g \geq 35.$$

Method. An exhaustive search over the $2^{24} = 16\,777\,216$ words v of length 24, keeping those with exactly 12 ones whose cotype, evaluated by (2), equals 12. The count is 7234. The same procedure at $\ell \leq 11$ returns Table 1 of [1] (Proposition 4.2), which is what licenses the twelfth term. $\square$

The value is not derivable from the published table: its last entry on this diagonal is $n(33, 21) = 7214$, which differs from 7234, so the diagonal has demonstrably *not* stabilised inside the published range. The next theorem confirms 7234 at $g = 35$ by a route that shares nothing with Theorem 7.1 beyond the encoding of a set of integers as a bit mask.

Theorem 7.2. $n(35, 23) \geq 7234$.

Method. We exhibit a list of 7234 gap masks, strictly increasing — hence pairwise distinct, as certified by a single linear scan — each of which is separately checked to encode a numerical semigroup of genus 35 and type 23. Distinct masks encode distinct numerical semigroups, so $n(35, 23)$ is at least the length of the list. No completeness claim is made or needed: a lower bound of this kind costs

the size of the cell, not the size of the level of the tree, which is why it is available at a genus far beyond the reach of Proposition 4.1. By Theorems 1.1 and 7.1 the true value is exactly 7234, so the list is in fact complete; that is not what is proved here. $\square$

Proposition 7.3. $n(25, 1) \geq 1182$, $n(30, 21) \geq 757$ and $n(33, 21) \geq 7214$, by the same method. All three bounds equal the corresponding entries of Table 4 of [1].

Proposition 7.3 is the calibration of Theorem 7.2: the witness machinery is run first at three cells whose exact values are published, and in each case the number of witnesses produced equals the published count exactly. The cell $n(25, 1)$ counts the symmetric numerical semigroups of genus 25, equivalently the irreducible numerical semigroups of Frobenius number $2 \cdot 25 - 1 = 49$, so it is also OEIS A158206(49) = 1182; and $n(30, 21) = 757 = |V(9)|$ is a third, independent route to the same number through Theorem 1.1 and Proposition 4.2.

Remark 7.4. Outside the formally verified development, two independent single-threaded C programs — a depth-first walk of the semigroup tree carrying 128-bit gap masks, and a pruned depth-first search over balanced words that abandons a prefix as soon as its difference set exceeds ℓ — extend the two computations to $g \leq 43$ and $\ell \leq 28$ in about 1.6 core-hours in total. They give

$$(|V(\ell)|)_{\ell=12}^{28} = 7234, 14982, 31492, 64842, 134097, 276555, 567101, 1161082, 2376348, \\ 4840723, 9860575, 20038261, 40667616, 82388901, 166856560, 337210832, 681114026,$$

with $|V(\ell + 1)|/|V(\ell)|$ settling near 2.065, and $n(g, t)$ for $34 \leq g \leq 43$ whose row totals reproduce A007323 and whose $t = 1$ column reproduces A158206 at Frobenius number $2g - 1$. The two programs cross-check each other exactly where Theorem 1.1 predicts: $|V(12)| = n(35, 23)$, $|V(13)| = n(38, 25)$ and $|V(14)| = n(41, 27)$, at $g = 3\ell - 1$ in each case, while $n(40, 26) = 31447 \neq 31492$ one genus earlier. These values are recorded here as data only; none of them is among the machine-checked statements of this paper except $|V(12)| = 7234$ (Theorem 7.1).

8. SHARPNESS AND NEGATIVE CONTROLS

A computation that agrees with a published table proves nothing about the values past its end unless the same computation also refuses the statements that are false. We record the controls.

Proposition 8.1. *The threshold $g \geq 3\ell - 1$ of Theorem 1.1 is sharp: at $\ell = 2, 3, 4$ the diagonal $n(g, g - \ell)$ really changes between $g = 3\ell - 2$ and $g = 3\ell - 1$, namely*

$$n(4, 2) = 2 \neq 3 = n(5, 3), \quad n(7, 4) = 5 \neq 7 = n(8, 5), \quad n(10, 6) = 12 \neq 15 = n(11, 7).$$

In particular no threshold of the form $g \geq 2\ell$ is available. From the stated threshold on, the diagonal is constant as far as computed: $n(11, 7) = n(14, 10) = n(17, 13) = n(20, 16) = 15 = |V(4)|$.

Remark 8.2. Two further features of the array are worth recording, both of which a careless strengthening would contradict. The column $t = 1$ is not monotone in g : $n(7, 1) = 8$ while $n(8, 1) = 7$. This is exactly why the monotonicity conjecture of [1, Conj. 4.3] is posed for $t \geq 2$. And the rows are not strictly unimodal on the way down: $n(8, 7) = n(8, 8) = 1$.

Proposition 8.3. *The following refute specific over-readings of Section 6.*

- (1) *The doubling of Theorem 6.2 is special to even windows: $\iota(5) = 10 \neq 12 = 2\iota(4)$ and $\iota(7) = 37 \neq 40 = 2\iota(6)$.*
- (2) *No position other than the middle may be toggled freely. At $\ell = 6$, whose middle is 3, toggling 2 turns the non-good $\{4\}$ into the good $\{2, 4\}$, and toggling 4 turns the good $\emptyset$ into the non-good $\{4\}$.*
- (3) *The hypothesis of Lemma 6.1 is satisfiable: 4 is bad for $A = \{4\}$ inside $[1, 6]$.*
- (4) *Being balanced is not sufficient for stability: the word $(1, 1, 0, 0)$ has cotype 0, not 2.*
- (5) *The maximal-type fibre is not all of $V(\ell)$: $\iota(6) = 20$ against $|V(6)| = 78$.*
- (6) *The correction term of Proposition 6.5 is not slack: $\iota(4) = 6 < 8 = 2^{4-1}$.*

(7) *A count is pinned from both sides: $\iota(14) = 4236$ is neither 4235 nor 4237.*

The enumeration of Section 3 carries its own controls. The gap sets $\{2\}$ and $\{1, 4\}$ have the right cardinality for genus 1 and 2 but are not gap sets of numerical semigroups ($1 + 1 = 2$, $2 + 2 = 4$), and neither is produced by the tree; a gap set containing 0 is rejected outright; $n(6, 3) = 6$ is neither 5 nor 7 and $n(10, 3) = 56$ is neither 55 nor 57, so an enumerator that drops or repeats a single node fails; the side condition $s \neq 0$ in the definition of $\text{PF}(S)$ is load-bearing, since dropping it empties $\text{PF}(S)$; and the witness check of Theorem 7.2 rejects a genus-25 semigroup when asked for type 2, and rejects a non-semigroup mask of the right weight, so a witness list cannot be padded.

9. WHAT REMAINS OPEN

Theorem 5.2 identifies the objects but does not count them. The question of $[1] - |\mathbf{V}(\ell)|$ in terms of ℓ — is open, and in the vocabulary of [2] it asks for the number of numerical sets whose atom monoid has a prescribed *genus*, whereas the generating-function machinery of [2] is indexed by the Frobenius number. Summing the fibre size over the n_ℓ numerical semigroups T of genus ℓ is the natural route, and the per-monoid upper bound [2, Thm. 8] is the natural input; we have not attempted it. On the upper side nothing better than the trivial $|\mathbf{V}(\ell)| \leq \binom{2\ell}{\ell}$ appears to be recorded, against a measured growth rate near 2.065; an upper bound of the form $c\lambda^\ell$ with $\lambda < 4$ would already be new.

10. VERIFICATION

Every statement of this paper has been formally verified in Lean 4 against *Mathlib*, and the development is free of `sorry`. The reductions and the bounds are proved outright and depend on no axioms beyond `propext`, `Classical.choice` and `Quot.sound`: this covers Lemmas 2.2, 2.3 and 6.1, Propositions 2.1, 3.1, 6.4 and 6.5 (including its two numerical instances at $\ell = 28$ and $\ell = 40$), and Theorems 5.2 and 6.2. What is delegated to compiled evaluation, through Lean’s `native_decide`, is exactly the finite searches: one per genus for Proposition 4.1, one per ℓ for Proposition 4.2 and Theorem 7.1, one per witness list for Theorem 7.2 and Proposition 7.3, one for the value $\iota(19)$ underlying Proposition 6.3 — whose passage from $\ell = 19$ to $\ell = 20$ is then axiom-free — and one per item for Proposition 8.3 and the controls quoted after it. Each such evaluation contributes a single named axiom recording precisely which computation the theorem rests on. Every numerical value quoted was produced first by an independent implementation in C and the formal statement was written against it; the two implementations of $|\mathbf{V}(\ell)|$ — one over difference sets, one over multiplier semigroups — agree at $\ell = 5, 7, 8$, and $\ell = 8$ agrees with Table 1 of [1]. Repository reference to be added at submission.

REFERENCES

- [1] J. Chappelton, J. L. Ramírez Alfonsín and D. I. Stamate, *Numerical semigroup tree: type-representation*, Experimental Mathematics, published online 6 January 2026, doi:10.1080/10586458.2025.2604786; arXiv:2507.15006. Results and tables are cited by the numbering of the arXiv version.
- [2] J. Marzuola and A. Miller, *Counting numerical sets with no small atoms*, J. Combin. Theory Ser. A **117** (2010), no. 6, 650–667, doi:10.1016/j.jcta.2010.03.002; arXiv:0805.3493. Results are cited by the numbering of the arXiv version.
- [3] J. C. Rosales and P. A. García-Sánchez, *Numerical Semigroups*, Developments in Mathematics, vol. 20, Springer, New York, 2009.
- [4] The On-Line Encyclopedia of Integer Sequences, <https://oeis.org>, sequences A007323 (numerical semigroups by genus), A158206 (irreducible numerical semigroups by Frobenius number) and A158278 (its odd-index part, the symmetric ones), A158291 and A164048 (numerical sets with no small atoms).