## Introduction

This Power Round is an exploration of numerical semigroups, mathematical structures which appear very naturally out of answers to simple questions. For example, suppose McDonald's sells Chicken McNuggets in boxes containing $a$, $b$, or $c$ McNuggets; can you say which exact quantities of McNuggets you can and cannot buy? The same problem is also often stated in terms of stamps or coins of certain values.

You can imagine that solutions to this problem must have numerous practical applications. What is more surprising is that it also has some interesting applications to more advanced, very abstract mathematics. We won't be able to discuss that here, but be aware, as you work through these elementary tricks and techniques for understanding numerical semigroups, that the same tricks and techniques are being used at the cutting edge of research!

**Remark:** Given sets $A$ and $B$, we will use the notations $A - B$ and $A \setminus B$ interchangeably to refer to the set of elements of $A$ which are not also elements of $B$.

## Defining numerical semigroups

We will develop two different definitions of numerical semigroups, each of which has its intuitive advantages, and prove that they are in fact the same. We will use $\mathbb{N}_0$ to refer to the set of nonnegative integers $0, 1, 2, \ldots$.

Here is our first definition: let $a_1, \ldots, a_n$ be a set of positive integers ($n \geq 2$) such that $\gcd(a_1, \ldots, a_n) = 1$. The *numerical semigroup generated by* $a_1, \ldots, a_n$ is the set $\{c_1a_1 + \cdots + c_na_n \mid c_1, \ldots, c_n \in \mathbb{N}_0\}$, which we sometimes refer to as $\langle a_1, \ldots, a_n\rangle$. For example, $\langle 4, 6, 9\rangle$ is the set $\{0, 4, 6, 8, 9, 10, 12, 13, \ldots\}$, which contains the listed numbers along with all integers after 12.

1. (a) **[6]** (i) Compute all elements of the numerical semigroup $\langle 5, 7, 11, 16\rangle$.

   (ii) Can this numerical semigroup be generated by a set of fewer than 4 integers? Prove your answer.

   (iii) Compute all elements of the numerical semigroup $\langle 3, 7, 8\rangle$.

   (iv) Can this numerical semigroup can be generated by a set of fewer than 3 integers? Prove your answer.

   (b) **[4]** Prove that $\langle a_1, \ldots, a_n\rangle$ is "closed under addition"—that is, if $x, y \in \langle a_1, \ldots, a_n\rangle$, then $x + y \in \langle a_1, \ldots, a_n\rangle$.

   (c) **[8]** Prove that $\langle a_1, \ldots, a_n\rangle$ contains all but a finite number of the nonnegative integers. (Hint: you may use without proof the fact that if $\gcd(a_1, \ldots, a_n) = 1$, then there exist *possibly negative* integers $d_1, \ldots, d_n$ such that $d_1a_1 + \cdots + d_na_n = 1$.)

**Solution to Problem 1:**

(a) (i) $\{0, 5, 7, 10, 11, 12, 14\} \cup \{n \in \mathbb{N}_0 : n \geq 15\}$.

(ii) Yes, $\langle 5, 7, 11, 16\rangle$ can be generated by a set of fewer than 4 elements. Specifically, it is generated by $\{5, 7, 11\}$ because $16 = 11 + 5$ and therefore any 16's in an element of the semigroup can be written using 5's and 11's.

(iii) $\{0, 3, 6, 7, 8\} \cup \{n \in \mathbb{N}_0 : n \geq 9\}$.

(iv) No, $\langle 3, 7, 8\rangle$ cannot be generated by a set of fewer than 3 elements. If this were possible, then we could write $\langle 3, 7, 8\rangle = \langle a, b\rangle$ for two integers $a < b$. For this to work, we must have $a = 3$. (If $a < 3$, then $\langle a, b\rangle$ contains $a \notin \langle 3, 7, 8\rangle$. If $a > 3$, then $\langle a, b\rangle$ doesn't contain anything that can generate a 3. The only possibility left is $a = 3$.)

Furthermore, we must have $b = 7$. (If $b < 6$, then $\langle 3, b\rangle$ contains $b \notin \langle 3, 7, 8\rangle$. If $b > 7$, then $\langle 3, b\rangle$ doesn't contain anything that can generate a 7. And $b = 6$ is not allowed because $\gcd(3, 6) > 1$. The only possibility left is $b = 7$.) So if we can generate $\langle 3, 7, 8\rangle$ using fewer than 3 elements, then $\langle 3, 7, 8\rangle = \langle 3, 7\rangle$. This is not true, because $8 \notin \langle 3, 7\rangle$. Therefore the answer is no, as we claimed.

(b) Suppose $x, y \in \langle a_1, \ldots, a_n\rangle$. Then there are $c_1, \ldots, c_n, d_1, \ldots, d_n \in \mathbb{N}_0$ such that $x = c_1a_1 + \cdots + c_na_n$ and $y = d_1a_1 + \cdots + d_na_n$. Then $x + y = (c_1 + d_1)a_1 + \cdots + (c_n + d_n)a_n$, which is in $\langle a_1, \ldots, a_n\rangle$ by definition.

(c) Let $d_1, \ldots, d_n$ be integers such that $d_1a_1 + \cdots + d_na_n = 1$. Let $M = \max |d_i|$. Let $s = a_1Ma_1 + a_1Ma_2 \cdots + a_1Ma_n$. Then for any $0 \le r < a_1$, all the coefficients in $s + r = (a_1M + rd_1)a_1 + (a_1M + rd_2)a_2 + \cdots + (a_1M + rd_n)a_n$ are positive and therefore $s + r \in \langle a_1, \ldots, a_n\rangle$. Any integer $x \ge s$ can be written as $x = qa_1 + (s + r)$ with $r < a_1$ by letting $q$ be the quotient of $\frac{x-s}{a_1}$ and by letting $r$ be the remainder. Now $qa_1 \in \langle a_1, \ldots, a_n\rangle$ by definition and we have shown $(s + r) \in \langle a_1, \ldots, a_n\rangle$, so the sum $x = qa_1 + (s+r)$ is also in $\langle a_1, \ldots, a_n\rangle$. Since every integer $x \ge s$ is in $\langle a_1, \ldots, a_n\rangle$, there are only finitely many positive integers not in $\langle a_1, \ldots, a_n\rangle$.

Here is our second definition: a *numerical semigroup* is any set $S \subseteq \mathbb{N}_0$ which satisfies all of the following three properties: (i) $S$ contains 0, (ii) $S$ is "closed under addition"—that is, for any $x, y \in S$, we have $x + y \in S$, and (iii) $S$ contains all but a finite number of the nonnegative integers. In Problem 1, you showed that $\langle a_1, \ldots, a_n\rangle$ is indeed a numerical semigroup by this definition.

2. (a) **[8]** Prove that any numerical semigroup $S$, by this definition, is "generated by" a finite set $\{a_1, \ldots, a_n\}$—that is, it can be written in the form $\langle a_1, \ldots, a_n\rangle = \{c_1a_1 + \cdots + c_na_n \mid c_1, \ldots, c_n \in \mathbb{N}_0\}$ where $a_1, \ldots, a_n$ are positive integers with $\gcd(a_1, \ldots, a_n) = 1$.

(b) **[8]** We say that $\{a_1, \ldots, a_n\}$ is a *minimal generating set* of $S$ if $S$ is generated by $\{a_1, \ldots, a_n\}$ and $S$ cannot be generated by any set of positive integers with fewer than $n$ elements. Prove that every numerical semigroup $S$ has a unique minimal generating set.

**Solution to Problem 2:**

(a) We will try to keep adding smallest un-generated elements to our set of generators until we get a set of generators that generate everything. To do this, let $A_0 = \emptyset$. If $S - \langle A_i\rangle$ is nonempty, then let $A_{i+1}$ be $A_i$ unioned with the smallest element of $S - \langle A_i\rangle$. (Where $\langle A\rangle$ denotes the set of all $\mathbb{N}_0$-linear combinations of elements of $A$). If $S - \langle A_i\rangle$ is empty, then let $A_{i+1} = A_i$.

First we claim that every integer in $S$ that is less than $i$ is in $\langle A_i\rangle$. We can show this by induction. For the base case, every integer in $S$ less than 0 is in $\langle A_0\rangle$. For the inductive step, suppose every integer in $S$ that is less than $i$ is in $\langle A_i\rangle$. We would like to show that every integer in $S$ that is less than $i + 1$ is in $\langle A_{i+1}\rangle$. Since $\langle A_i\rangle \subseteq \langle A_{i+1}\rangle$, every integer in $S$ that is less than $i$ is already in $\langle A_{i+1}\rangle$. So we only need to show that if $i \in S$ then $i \in \langle A_{i+1}\rangle$. If $i \in S$ and $i \in \langle A_i\rangle$, then $i \in \langle A_{i+1}\rangle$ and we are done. If $i \in S$ and $i \notin \langle A_i\rangle$, then $i$ is the smallest number in $S - \langle A_i\rangle$ and therefore $i \in \langle A_{i+1}\rangle$ by construction. So we have proven the claim.

Let's show that there is an $n$ such that $S = \langle A_n\rangle$. To do this, let $p < q$ be two distinct primes in $S$ (there are at least two distinct primes in $S$ because all but finitely many positive integers are in $S$). By our claim, $p, q \in A_{q+1}$. Since $\gcd(p, q) = 1$, the set $\langle p, q\rangle$ contains all but finitely many positive integers (we proved this in 1c). Since $\langle p, q\rangle \subset \langle A_{q+1}\rangle$, this means that $\langle A_{q+1}\rangle$ contains all but finitely many positive integers. In particular, $S - \langle A_{q+1}\rangle$ is finite. So there is an integer $n > q+1$ that is bigger than all the elements in $S - \langle A_{q+1}\rangle$. By our above claim, $\langle A_n\rangle$ contains all of $S - \langle A_{q+1}\rangle$. And since $n > q+1$, $\langle A_n\rangle$ also contains all of $\langle A_{q+1}\rangle$. Therefore $\langle A_n\rangle \supseteq S$. By construction, $\langle A_n\rangle \subseteq S$. Therefore we have equality $\langle A_n\rangle = S$ as desired.

So now we have a set of integers $A_n = a_1, \ldots, a_n$ with $\langle a_1, \ldots, a_n\rangle = S$. We have almost shown what we wanted. But we must still show that $\gcd(a_1, \ldots, a_n) = 1$. To do this, assume for a contradiction that $\gcd(a_1, \ldots, a_n) > 1$. Let $d = \gcd(a_1, \ldots, a_n)$. Then $d > 1$ divides everything in $\langle a_1, \ldots, a_n\rangle$, and so there are infinitely many positive integers not in $\langle a_1, \ldots, a_n\rangle$. This is a contradiction and therefore $\gcd(a_1, \ldots, a_n) = 1$.

(b) The $A_n$ constructed above is the unique minimal generating set. To see this, let $a_1 < a_2 < \ldots < a_N$ be the elements of $A_n$ and let $b_1 < b_2 < \ldots < b_m$ be the elements of a minimal generating set. We will prove that the sequences $a_i, b_i$ are equal. First notice that $N \geq m$ because $b_1, \ldots, b_m$ is a minimal generating set. We will therefore start by showing that $a_i = b_i$ for all $i \leq m$.

Assume for a contradiction that there is some $i \leq m$ such that $a_i \neq b_i$. Let $i$ be the minimum such $i$. Since $\{b_1, ..., b_m\}$ is minimal, $b_i \notin \langle b_1, \ldots, b_{i-1}\rangle$. In other words, $b_i \in S - \langle b_1, \ldots, b_{i-1}\rangle$.

Furthermore, we claim that $b_i$ is the smallest element in $S - \langle b_1, \ldots, b_{i-1}\rangle$. To see this, let $r$ be the smallest element in $S - \langle b_1, \ldots, b_{i-1}\rangle$. Then $r$ is some nonnegative linear combination of $b_1, \ldots, b_n$ involving at least one element past $b_{i-1}$. If $b_i > r$, then all elements past $b_{i-1}$ are greater than $r$ and therefore $r$ cannot be made with such a nonnegative linear combination. Therefore $b_i \leq r$, forcing $b_i = r$ as desired. $b_i$ is the smallest element in $S - \langle b_1, \ldots, b_{i-1}\rangle$.

Since we chose $i$ to be the minimum such that $a_i \neq b_i$, we know that $\langle a_1, \ldots, a_{i-1}\rangle = \langle b_1, \ldots, b_{i-1}\rangle$. In particular, $S - \langle b_1, \ldots, b_{i-1}\rangle = S - \langle a_1, \ldots, a_{i-1}\rangle$. So $b_i$ is the smallest element in $S - \langle a_1, \ldots, a_{i-1}\rangle$. But this is exactly how we defined $a_i$! Therefore $a_i = b_i$, contradicting our assumption that $a_i \neq b_i$.

So we have proven by contradiction that $a_i = b_i$ for all $i \leq m$. Since $\{b_1, \ldots, b_m\}$ generates $S$, our construction of $A_n$ stops adding elements once it gets to $a_m$. So the sequence $a_1, \ldots, a_N$ actually has $m$ elements and we are done.

If $a$ is part of the minimal generating set of $S$, we say that $a$ is a *generator* of $S$. This will be important later.

## The genus and Frobenius number of a numerical semigroup

Now that you have two equivalent definitions of numerical semigroups to work with, we can start analyzing them in more detail. The *genus* of a numerical semigroup $S$ is the number of positive integers not contained in $S$. For example, $\langle 4, 6, 9\rangle = \{0, 4, 6, 8, 9, 10, 12, 13, \ldots\}$ has genus 6, because it does not contain $1, 2, 3, 5, 7$, or 11. The *Frobenius number* of a numerical semigroup $S$ is the largest integer that $S$ does not contain. For example, $\langle 4, 6, 9\rangle$ has Frobenius number 11. Given a numerical semigroup $S$, let $g(S)$ be its genus and $F(S)$ its Frobenius number. We will write $g$ and $F$ for $g(S)$ and $F(S)$ respectively when there is no chance of confusion. (Note that $F$ may be negative. Specifically, if $S$ consists of all the non-negative integers, then $F(S) = -1$.)

3. (a) **[4]** Compute the genus and Frobenius number of (i) $\langle 5, 7, 11, 16\rangle$ and (ii) $\langle 3, 7, 8\rangle$.

(b) **[8]** Prove that for any numerical semigroup $S$, we have $F(S) \leq 2g(S) - 1$.

**Solution to Problem 3:**

(a) (i) Genus 8, Frobenius 13. (ii) Genus 4, Frobenius 5.

(b) Let $N$ be any positive integer that is not in the semigroup. Then at least one number from each of the the $\lfloor \frac{N}{2} \rfloor$ pairs $(1, N-1), (2, N-2), \ldots, (\lfloor \frac{N}{2} \rfloor, \lceil \frac{N}{2} \rceil)$ must not be in the semigroup. Also, $N$ is not in the semigroup. So there are at least $\lfloor \frac{N}{2} \rfloor + 1 \geq \frac{N+1}{2}$ positive integers not in the semigroup. Ie, $g(S) \geq \frac{N+1}{2}$. Plugging in $N = F(S)$, we get $g(S) \geq \frac{F(S)+1}{2}$. Rearranging, $F(S) \leq 2g(S) - 1$.

The famous Chicken McNugget Theorem states that if McDonald's sells Chicken McNuggets in boxes of $a$ or $b$ McNuggets where $\gcd(a, b) = 1$, then the largest number of McNuggets one cannot buy is $ab - a - b$.

4. (a) **[1]** Restate the Chicken McNugget Theorem in terms of the numerical semigroup $\langle a, b \rangle$.

(b) **[8]** Prove the Chicken McNugget Theorem. (Possible hint: consider the grid

$$\begin{pmatrix} 1 & 2 & \cdots & a \\ a+1 & a+2 & \cdots & 2a \\ \vdots & \vdots & \ddots & \vdots \\ (b-1)a+1 & (b-1)a+2 & \cdots & ba \end{pmatrix}.$$

Cross out the numbers of McNuggets that you can buy. What do you notice? Try this with actual numbers in place of $a, b$ if you're not comfortable.)

(c) **[10]** Find, with proof, the genus of $\langle a, b \rangle$.

**Solution to Problem 4:**

(a) $F(\langle a, b \rangle) = ab - a - b$.

(b) We will use the following fact: If $(x_0, y_0)$ is an integer solution to $xa + yb = c$, then the set of integer solutions to $xa + yb = c$ is exactly the set of $(x_0 - kb, y_0 + ka)$ for all integers $k$.

To see that no nonnegative combination of $a, b$ makes $ab - a - b$, notice that $(b-1, -1)$ solves $xa + yb = ab - a - b$. So the set of all solutions is $(b - 1 - kb, -1 + ka)$. For solutions with $k \leq 0$, we have $y < 0$. For solutions with $k > 0$ we have $x < 0$. Therefore there are no nonnegative integer solutions. Ie, no nonnegative combination of $a, b$ makes $ab - a - b$.

Now let $N$ be any integer bigger than $ab - a - b$. Then the set

$$S = \{N + b, N + b - a, N + b - 2a, \ldots, N + b - (b-1)a\}$$

is a set of $b$ positive integers because

$$N + b - (b-1)a > ab - a - b + b - (b-1)a = 0.$$

Since $\gcd(a, b) = 1$, none of the integers in this set may be congruent mod $b$. (If the $i$-th and $j$-th terms are congruent mod $b$, then $b \mid (N + b - ia) - (N + b - ja)$ so $b \mid (j - i)a$ so $b \mid j - i$, which implies $i = j$). Therefore we get all the integers $\{0, \ldots, b-1\}$ by reducing $S$ mod $b$. In particular, there is an $i$-th term congruent to 0 mod $b$. This term is divisible by $b$, so there is some $j$ such that $jb = N + b - ia$. Since $N + b - ia > 0$, we have $j > 0$ and in particular $(j-1) \geq 0$. So $N = ia + (j-1)b$ is a nonnegative linear combination of $a$ and $b$ that makes $N$.

Alternatively, we can find all the multiples of $b$ in the given grid and note that the elements in $\langle a, b\rangle$ in the grid are precisely those lying below a multiple of $b$ in their given column. Because $\gcd(a, b) = 1$, the numbers $0, b, 2b, \ldots, (a-1)b$ go through all the residue classes $\pmod{a}$, each exactly once (this can be shown with a standard argument as follows: if $ib \equiv jb \pmod{a}$, then $(i-j)b \equiv 0 \pmod{a}$; since $b$ is relatively prime to $a$ this is only possible if $a$ divides $i - j$, which is only possible if $i = j$ since both are between 0 and $a - 1$). So the largest element left out is just the one lying directly above $(a-1)b$, that is, $(a-1)b - a = ab - a - b$.

(c) First, we claim that for $i = 1, \ldots, a-1$, the smallest number congruent to $ib$ modulo $a$ is $ib$. Any smaller number is writable as $ax + by$ where $x, y \geq 0$ and $y < i$. This number is congruent to $by$ modulo $a$, so if it were also congruent to $bi$, we would have $by \equiv bi \pmod{a}$. But $a$ is coprime to $b$, so this implies $y \equiv i \pmod{a}$, a contradiction since $i < a$ implies $i$ is the smallest positive number congruent to itself mod $a$.

Now, for any $i = 1, \ldots, a-1$, write $ib = q_i a + r_i$ where $0 \leq r_i < a$ (this is the result of dividing $ib$ by $a$ and finding the quotient and remainder). Note that since by similar reasoning as above, $jb \equiv kb \pmod{a} \implies j = k$ when $0 \leq j, k < a$, so the $r_i$ cycle through $1, \ldots, a-1$ as $i$ ranges from $1, \ldots, a-1$.

The numbers congruent to $r_i$ mod $a$ that are not in $\langle a, b\rangle$ are $r_i, a + r_i, \ldots, (q_i - 1)a + r_i$, so there are precisely $q_i$ such numbers. Hence, the genus of $\langle a, b\rangle$ is precisely $\sum_{i=1}^{a-1} q_i$ (we ignore the residue class of 0 modulo $a$ since clearly all positive multiples of $a$ are in $\langle a, b\rangle$). Now, use the fact that

$$\sum_{i=1}^{a-1} ib = \frac{ab(a-1)}{2} = a\sum_{i=1}^{a-1} q_i + \sum_{i=1}^{a-1} r_i = a\sum_{i=1}^{a-1} q_i + \frac{a(a-1)}{2},$$

which implies that

$$\sum_{i=1}^{a-1} q_i = \boxed{\frac{(a-1)(b-1)}{2}}.$$

## The multiplicity, Apéry set, and embedding dimension of a numerical semigroup

The multiplicity of a numerical semigroup $S$ is the smallest positive integer it contains. For example, $\langle 4, 6, 9\rangle = \{0, 4, 6, 8, 9, 10, 12, 13, \ldots\}$ has multiplicity 4. We refer to the multiplicity of $S$ by $m(S)$, or $m$ when there is no possibility of confusion.

The Apéry set of a numerical semigroup $S$ is the set $A(S) = \{n \mid n \in S, n - m(S) \notin S\}$. For example, $\langle 4, 6, 9\rangle$ has Apéry set $\{0, 6, 9, 15\}$. Notice that $A(S)$ always contains 0. As usual, we say $A$ for $A(S)$ when there is no possibility of confusion.

5. (a) **[4]** Compute the multiplicity and the Apéry set of (i) $\langle 5, 7, 11, 16\rangle$ and (ii) $\langle 3, 7, 8\rangle$.

   (b) **[4]** Prove that if numerical semigroup $S$ has multiplicity $m$, then $A(S)$ can be uniquely written in the form $\{0, k_1 m + 1, k_2 m + 2, \ldots, k_{m-1} m + m - 1\}$ where $k_1, \ldots, k_{m-1}$ are positive integers and $k_i m + i$ is the smallest element of $S$ which has a remainder of $i$ when divided by $m$. For example, $A(\langle 4, 6, 9\rangle) = \{0, 9, 6, 15\} = \{0, 2 \cdot 4 + 1, 1 \cdot 4 + 2, 3 \cdot 4 + 3\}$. In the future, we will often refer to $k_1, \ldots, k_{m-1}$ as the Apéry coefficients of $S$.

(c) **[4]** Prove that $S$ is generated by $(A(S) - \{0\}) \cup \{m\}$. (Note that this does not mean $(A(S) - \{0\}) \cup \{m\}$ is a minimal generating set of $S$—in fact, that is not the case for our favorite example $\langle 4, 6, 9 \rangle$.)

(d) **[3]** Write, with proof, the genus of $S$ in terms of its Apéry coefficients.

(e) **[3]** Write, with proof, the Frobenius number of $S$ in terms of its Apéry coefficients.

**Solution to Problem 5:**

(a) (i) $m(S) = 5$, $A(S) = \{0, 7, 11, 14, 18\}$. (ii) $m(S) = 3$, $A(S) = \{0, 7, 8\}$.

(b) Let's prove two statements: (1) every residue class mod $m$ appears in $A(S)$ (and it appears as the smallest element of $S$ in that residue class) and (2) no residue class mod $m$ appears more than once in $A(S)$. It is obvious that the residue class 0 appears in $A(S)$, so we do not need to show that it appears in $A(S)$.

To see (1), let $k$ be any nonzero residue class mod $m$, and let $x \in S$ such that $x \equiv k \mod m$. (Such an $x$ exists because all numbers past some finite point are in $S$). By integer division, there is some $q$ so that $0 < x - qm < m$. Since $x - qm$ is a positive integer smaller than the smallest positive integer in $S$, $x - qm \notin S$. The sequence $x, x - m, x - 2m, \ldots x - qm$ therefore starts with an element in $S$ and ends up with an element not in $S$. So there is some point in the sequence where $x - im \in S$ and $x - (i+1)m \notin S$. Then $x - im \in A(S)$ and therefore the residue class $k$ mod $m$ appears in $A(S)$. Furthermore, $x - im$ is the smallest element of $S$ congruent to $k$ because if $x - jm \in S$ for $j > i$ then $x - (i+1)m = (x - jm) + (j - i - 1)m \in S$, contradicting the fact that $x - (i+1)m \notin S$.

To see (2), assume for a contradiction that $x, y \in A(S)$ with $x \equiv y \mod m$ and $x < y$. Then $x \in S$. Since $x < y$, there is some integer $k \geq 0$ such that $y - m = x + km$. So $y - m \in S$ by additive closure. But $y - m \notin S$ because $y \in A(S)$. So we have a contradiction.

(c) It is sufficient to show that $A(S) \cup \{m\}$ generates $S$ because removing 0 from a set of generators does not change what it generates.

Since $A(S) \cup \{m\} \subset S$, we have $\langle A(S) \cup \{m\} \rangle \subset S$.

To show the reverse inclusion, let $x \in S$. As in the previous proof, the sequence $x, x - m, x - 2m, \ldots$ eventually hits an element of $A(S)$. Thus $x$ is an element of $A(S)$ plus some multiple of $m$. Ie, $x \in \langle A(S) \cup \{m\} \rangle$.

(d) We claim that the set

$$T = \bigcup_{a \in A(S)} T_a = \bigcup_{a \in A(S)} \{a - qm \mid q \geq 1,\ a - qm > 0\}$$

is exactly the set of positive integers not in $S$. Each element $a - qm$ is not in $S$ because otherwise $a - m = (a - qm) + (q-1)m \in S$, contradicting the fact that $a - m \notin S$. Each positive integer $x$ not in $S$ is in $T$ because eventually the sequence $x, x + m, x + 2m, \ldots$ hits some $a \in A(S)$.

Since no elements of $A(S)$ are congruent mod $m$, the sets $T_a$ (for $a \in A(S)$) are disjoint. So we can count $T$ by counting each of the sets $T_a$. The size of $T_a$ is clearly equal to its corresponding Apéry coeffcient. Therefore

$$g(S) = \sum_{i=1}^{m-1} k_i.$$

(e) It is $\max_i((k_i-1)m+i)$, since any other element left out of $S$ is $m$ less than another element which was also left out of $S$, and so cannot be the Frobenius number.

Note that because $S$ is generated by $(A(S)-\{0\})\cup\{m\}$, different numerical semigroups must have different Apéry sets. Hence we can associate each $S$ with a unique sequence of Apéry coefficients $k_1,\dots,k_{m-1}$. The natural next question becomes: when can an arbitrary sequence of positive integers $k_1,\dots,k_{m-1}$ be the Apéry set of a valid numerical semigroup?

6. (a) **[4]** Suppose numerical semigroup $S$ has Apéry coefficients $k_1,\dots,k_{m-1}$. Prove that if $1\le i,j\le m-1$ and $i+j<m$, then $k_i+k_j\ge k_{i+j}$. Also prove that if $1\le i,j\le m-1$ and $i+j>m$, then $k_i+k_j+1\ge k_{i+j-m}$.

   (b) **[8]** Prove that if $k_1,\dots,k_{m-1}$ satisfy the inequalities given in part a, there is a semigroup $S$ with $k_1,\dots,k_{m-1}$ as its Apéry coefficients.

   (c) **[8]** Find, with proof, in terms of $g$ and $m$, the number of numerical semigroups $S$ of genus $g$ and multiplicity $m$ satisfying $F(S)<2m$.

   (d) **[8]** Prove that the number of numerical semigroups $S$ of a fixed genus $g$ (but any multiplicity) satisfying $F(S)<2m(S)$ is a Fibonacci number.

**Solution to Problem 6:**

(a) Suppose $1\le i,j\le m-1$. Then $k_im+i\in S$ and $k_jm+j\in S$ so $(k_i+k_j)m+(i+j)\in S$ by additive closure. Therefore the smallest element of $S$ congruent to $i+j$ mod $m$ is at most $(k_i+k_j)m+(i+j)$. If $i+j<m$, then the smallest element of $S$ congruent to $i+j$ mod $m$ is $k_{i+j}m+(i+j)$ so we get the inequality $k_{i+j}\le k_i+k_j$. If $i+j>m$, then the smallest element of $S$ congruent to $i+j$ mod $m$ is $k_{i+j-m}m+(i+j-m)$ so we get the inequality $k_{i+j-m}-1\le k_i+k_j$.

(b) Suppose $k_1,\dots,k_{m-1}$ satisfy the inequalities given in part a. Let

$$A=\{0,k_1m+1,\dots k_{m-1}m+m-1\}.$$

Let $S=\langle A\cup\{m\}\rangle$. We claim that $A(S)=A$. By 5b, we can do this by showing that the smallest element congruent to $i$ mod $m$ is $k_im+i$. So let $x\in S$ be the smallest element with $x\equiv i \mod m$. Then $x$ is a positive linear combination of the generators in $A\cup\{m\}$. We can write the positive linear combination as follows:

$$x=(k_{j_1}m+j_1)+\dots+(k_{j_n}m+j_n)+cm$$

for some sequence $1\le j_1,\dots,j_n\le m-1$ (which might contain duplicate elements) and some positive integer $c$. If $c>0$, then $x-m\in S$ is a smaller element with $x-m\equiv i$ mod $m$. So $c=0$ and

$$x=(k_{j_1}m+j_1)+\dots+(k_{j_n}m+j_n).$$

Reducing both sides mod $m$, we see that

$$i\equiv j_1+\dots+j_n \mod m.$$

Therefore $j_1+\dots+j_n=i+qm$ for some $q\ge 0$. Repeatedly applying the inequalities to $k_{j_1}+\dots+k_{j_n}$, we get

$$k_{j_1}+\dots+k_{j_n}+q\ge k_i.$$

Multiplying both sides by $m$ and adding $j_1 + \ldots + j_n$ to both sides gives

$$k_{j_1}m + j_1 + \ldots + k_{j_n}m + j_n + qm \geq k_i m + j_1 + \ldots + j_n.$$

Move $qm$ to the other side of the inequality and note that $j_1 + \ldots + j_n - qm = i$ to get

$$k_{j_1}m + j_1 + \ldots + k_{j_n}m + j_n \geq k_i m + i.$$

The left side of this inequality is simply $x$, so we have $x \geq k_i m + i$. Now $x$ is the smallest element in $S$ congruent to $i$, and $k_i m + i$ is an element in $S$ congruent to $i$, so this forces $x = k_i m + i$ as desired.

Alternatively, explicitly list out a general element of the desired semigroup as the Apéry set elements plus multiples of the multiplicity, and verify that this is closed under addition.

(c) By problem 5 part e, the Frobenius number of $S$ is $\max_i((k_i - 1)m + i)$; in order to have $(k_i - 1)m + i < 2m$ for all $i$, we must have all $k_i$ equal to 1 or 2. Furthermore, since $1 + 1 \geq 2$ and $1 + 1 + 1 \geq 2$, any such choice of $k_i$ automatically induces a valid Apéry set. By problem 5 part d, we have $g = \sum_{i=1}^{m-1} k_i$, so $g - (m - 1)$ of the $k_i$s are 2s and the rest are 1s. Hence there are $\binom{m-1}{g-m+1}$ ways to choose those $k_i$s to set to 2, and thus $\binom{m-1}{g-m+1}$ distinct such numerical semigroups.

(d) We prove by induction on $g$ that we have $\sum_{m=1}^{g+1} \binom{m-1}{g-m+1} = F_{g+1}$, where the summand is understood to be 0 if $m - 1 < g - m + 1$, and $F_1 = F_2 = 1$. The base cases are easy to check. Suppose this is true for $g$ and $g + 1$; then we have

$$\begin{aligned}
F_{g+3} = F_{g+1} + F_{g+2} &= \sum_{m=1}^{g+1} \binom{m-1}{g-m+1} + \sum_{m=1}^{g+2} \binom{m-1}{g-m+2} \\
&= \sum_{m=1}^{g+1} \left( \binom{m-1}{g-m+1} + \binom{m-1}{g-m+2} \right) + \binom{g+1}{0} \\
&= \sum_{m=1}^{g+1} \binom{m}{g-m+2} + \binom{g+2}{0} = \sum_{m=1}^{g+2} \binom{m}{g-m+2} \\
&= \sum_{m=2}^{g+3} \binom{m-1}{g-m+3} = \sum_{m=1}^{g+3} \binom{m-1}{g-m+3}
\end{aligned}$$

where we use Pascal's Identity to get from the second line to the third, and the fact that $\binom{1-1}{g-1+3} = 0$ (because $1 - 1 < g - 1 + 3$ for $g > -2$) for the last equality. So we are done.

The *embedding dimension* of a numerical semigroup $S$ is the number of elements in its minimal generating set, which we call $e(S)$ or $e$ when there is no chance of confusion. Note that because $S$ is generated by $(A(S) - \{0\}) \cup \{m\}$, we have $e(S) \leq m(S)$. If $S$ is such that $e(S) = m(S)$, we say that $S$ is a *maximal embedding dimension* numerical semigroup, or MED for short.

7. **[10]** Given a sequence of positive integers $k_1, \ldots, k_{m-1}$, give, with proof, necessary and sufficient conditions for $k_1, \ldots, k_{m-1}$ to be the Apéry coefficients of an MED numerical semigroup.

**Solution to Problem 7:** We claim that $k_1, \ldots, k_{m-1}$ define an MED semigroup if and only if the constraints in problem 6, part a hold without equality: that is, for any $i, j \in \{1, \ldots, m-1\}$, we have $k_i + k_j > k_{i+j}$ if $i + j < m$, and $k_i + k_j + 1 > k_{i+j-m}$ if $i + j > m$.

We first show that these conditions are necessary. Suppose one of the equalities from 6a holds. We write $B = \{A(S) - \{0\}\} \cup \{m\}$. We have two cases.

Case 1: there exist $i, j \in \{1, \ldots, m-1\}$ with $i + j < m$ and $k_i + k_j = k_{i+j}$. Then $(k_i m + i) + (k_j m + j) = (k_i + k_j)m + i + j = k_{i+j}m + (i + j) \in B$, so $B \setminus \{k_{i+j}m + i + j\}$ has $m - 1$ elements and also generates $S$.

Case 2: there exist $i, j \in \{1, \ldots, m-1\}$ with $i + j > m$ and $k_i + k_j + 1 = k_{i+j-m}$. Then $(k_i m + i) + (k_j m + j) = (k_i + k_j + 1)m + i + j - m = k_{i+j-m}m + i + j - m \in B$, so $B \setminus \{k_{i+j-m}m + i + j - m\}$ has $m - 1$ elements and also generates $S$.

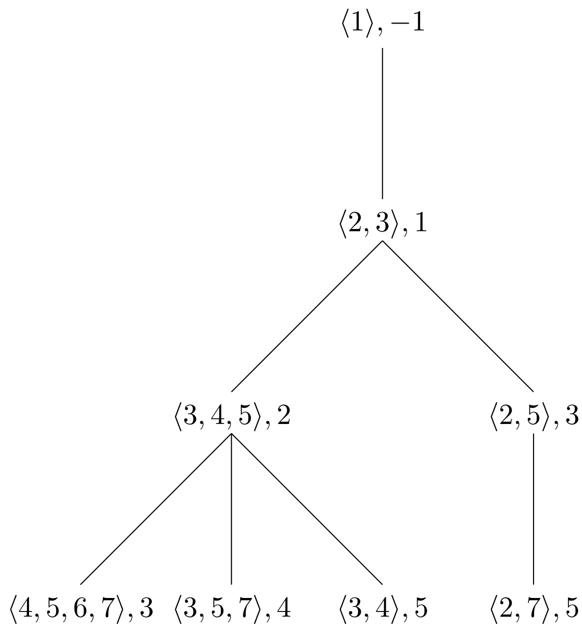
Next we show that these conditions are sufficient. Specifically, we claim that if these conditions are given, then for every $q \in \{1, \ldots, m-1\}$, an element of $\langle B \setminus \{k_q m + q\}\rangle$ that is congruent to $q \pmod m$ must be greater than $k_q m + q$, so all elements of $B$ are needed to generate $S$.

Take any $x \in \langle B \setminus \{k_q m + q\}\rangle$, and let $x = \sum_{i=1}^{n} a_i$ where $a_i \in B \setminus \{k_q m + q\}$ for all $i$. We induct on $n$. We have two base cases: $n = 1$ is obvious, and $n = 2$ is true by our given conditions: if $i + j \equiv q \pmod m$ for some $i, j \in \{1, \ldots, m-1\}$, then the inequalities tell us that $k_i m + i + k_j m + j > k_q m + q$.

Now suppose for the sake of induction that our claim is true for $n$, and consider $\sum_{i=1}^{n+1} b_i \equiv q \pmod m$, $b_i \in B \setminus \{k_q m + q\}$. This can be written as $\sum_{i=1}^{n-1} b_i + b_n + b_{n+1}$. Let $b_n + b_{n+1} \equiv c \pmod m$ where $c \in \{1, \ldots, m-1\}$. Then, by the given conditions, either one of $b_n, b_{n+1}$ is $k_c m + c$ or $b_n + b_{n+1} > k_c m + c$—but we have $b_n + b_{n+1} > k_c m + c$ in both cases. Hence $\sum_{i=1}^{n+1} b_i > \sum_{i=1}^{n-1} b_i + k_c m + c > k_q m + q$ by the inductive hypothesis. This completes the induction.

## The semigroup tree

The semigroup tree is a systematic way of creating numerical semigroups. We start at level 0 of the tree, where we put the unique numerical semigroup of genus 0, that is, $\langle 1\rangle = \mathbb{N}_0$. (By convention, $\mathbb{N}_0$ has Frobenius number $-1$.) If numerical semigroup $S$ appears at level $g$, it has some number of *children* which appear at level $g + 1$. Each child is created by removing from the set $S$ a single element $n$, with the condition that $n$ is a generator (that is, an element of the minimal generating set of $S$) which is larger than the Frobenius number $F(S)$. Hence, we get the only child of $\langle 1\rangle$ by removing 1, which results in $\langle 2, 3\rangle$ of Frobenius number 1 at level 1. Now $2, 3$ are both larger than 1, so $\langle 2, 3\rangle$ has two children at level 2: $\langle 3, 4, 5\rangle$, which we get by removing 2, and $\langle 2, 5\rangle$, which we get by removing 3. The first few levels of the tree are shown below. Each element is given in the format (minimal generating set, Frobenius number).




For convenience, when a semigroup has multiple children, we arrange them from left to right in increasing order of the size of the generator removed from the "parent".

8. (a) **[7]** Compute the next level of the tree, following the format given above. (So write each child in terms of its minimal generating set and give its Frobenius number.) Use a separate sheet of paper, *not* the provided answer sheet.

   (b) **[4]** Prove that as stated, the algorithm which generates the tree really does only create valid numerical semigroups, and that every numerical semigroup $S$ appears exactly once in this tree (at level equal to its genus).

   (c) **[4]** Describe in general, with justification, all elements of the rightmost branch of the tree, including minimal generating set and Frobenius number.

   (d) **[4]** Describe in general, with justification, all elements of the leftmost branch of the tree, including minimal generating set and Frobenius number.

**Solution to Problem 8:**

(a) They are, in the standard order,

- $\langle 5,6,7,8,9\rangle, 4$
- $\langle 4,6,7,9\rangle, 5$
- $\langle 4,5,7\rangle, 6$
- $\langle 4,5,6\rangle, 7$
- $\langle 3,7,8\rangle, 5$
- $\langle 3,5\rangle, 7$
- $\langle 2,9\rangle, 7$

(b) If you remove a generator from a numerical semigroup, then the result is still a numerical semigroup because no two elements in a numerical semigroup can sum to a generator (by our explicit algorithm for finding generators in problem 2). So all elements in the tree are valid numerical semigroups.

Each numerical semigroup is its parent with one element removed, so the genus increases by exactly one at each level of the tree. So by induction, the level is equal to the genus.

To see that we get every numerical semigroup, let $S$ be any numerical semigroup. Let $a_1 < a_2 < \ldots < a_n$ be all the elements of $\mathbb{N}_0 - S$. I claim that the path starting at $\mathbb{N}_0$ and proceeding by removing each of the $a_i$'s in order is a valid path through the tree. We can prove this by induction on the node in the path. The 0-th node $\mathbb{N}_0$ is at the root of the tree, which establishes the base case. For the inductive step, let $i \le n$ and assume that $\mathbb{N}_0$, $\mathbb{N}_0 - \{a_1\}$, ..., $\mathbb{N}_0 - \{a_1, \ldots, a_{i-1}\}$ is a valid path through the tree. We need to show that one may move from $\mathbb{N}_0 - \{a_1, \ldots, a_{i-1}\}$ to $\mathbb{N}_0 - \{a_1, \ldots, a_i\}$ along the tree. Ie, we need to show that $a_i$ is a generator of $\mathbb{N}_0 - \{a_1, \ldots, a_{i-1}\}$ that is greater than its Frobenius number. If $a_i$ is not a generator, then some elements of $\mathbb{N}_0 - \{a_1, \ldots, a_{i-1}\}$ sum to it and therefore $a_i$ cannot not be in $S$. Therefore $a_i$ is a generator of $\mathbb{N}_0 - \{a_1, \ldots, a_{i-1}\}$. $a_i$ is obviously bigger than the Frobenius number because the Frobenius number is $a_{i-1}$. So we are done. We can reach all numerical semigroups through the tree.

The above path is the only path from the root to $S$ because the Frobenius number constraint forces us to remove elements in increasing order. Therefore each numerical semigroup appears exactly once.

(c) $\langle 2, 2g+1 \rangle$ of genus $g$. To see this, we induct on $g$. The base case appears in the diagram. For the inductive step, suppose $\langle 2, 2g+1 \rangle$ of genus $g$ is on the rightmost side of the tree. Its child is $\langle 2, 2g+1 \rangle - \{2g+1\}$ of genus $g+1$. It is easy to see that $\langle 2, 2g+1 \rangle - \{2g+1\} = \langle 2, 2(g+1)+1 \rangle$, completing the inductive step.

(d) $\langle g+1, \ldots, 2g+1 \rangle$ of genus $g$. To see this, we induct on $g$. The base case appears in the diagram. For the inductive step, suppose $\langle g+1, \ldots, 2g+1 \rangle$ of genus $g$ is on the leftmost side of the tree. Its leftmost child is $\langle g+1, \ldots, 2g+1 \rangle - \{g+1\} = \{g+2, g+3, \ldots\}$. By applying the algorithm we described in the solution to 2, we get that this has generators $\{g+2, \ldots, 2(g+1)+1\}$, completing the inductive step.

9. (a) **[6]** Suppose that $S$ is not in the leftmost branch of the semigroup tree and that it has a child $S'$. Then answer—*proof not required*—the following in terms of the multiplicity, Frobenius number, and embedding dimension of $S$ and $S'$: (i) Which generators of $S$ are still generators of $S'$, and which are not? (ii) Which generators of $S'$ were not also generators of $S$? Use a separate sheet of paper, *not* the provided answer sheet.

(b) **[10]** Now, prove your answers to part a.

**Solution to Problem 9:**

Note: We use $m$ below to denote $m(S) = m(S')$, which is valid because we are not in the leftmost branch of the semigroup tree. Therefore, the Frobenius number of $S$ is greater than $m(S)$. Otherwise, $S$ must be of the form $\langle g+1, \ldots, 2g+1 \rangle$ where $g$ is the genus of $S$. We showed in 8d that such a semigroup must lie in the leftmost branch, and by 8b, this is the only place in which it occurs.

(a) Suppose $S$ has Frobenius number $F$ and the generator $F' > F$ is removed from $S$ to give $S'$. Then our answers are: (i) $F'$ is not a generator of $S'$, but all other generators of $S$ are; (ii) $F' + m$ if $e(S) = e(S')$, otherwise none.

(b) Clearly all generators of $S$ except $F'$ are still generators of $S'$, since if they could not be written in terms of other elements before, this will still not be possible when $F'$ is removed. Now, note that a generator $a$ of $S'$ cannot be larger than $F' + m$, since otherwise we would have $a - m > F'$, hence $a - m \in S'$ and $a - m \neq 0$, hence $a$ would be the sum of $m$ and another nonzero element of $S'$. But if $a < F' + m$, and $a$ is not the sum of two nonzero elements of $S'$, note that $a$ is also not the sum of $F'$ and any nonzero element of $S$ (since the smallest such element is $m$) and consequently $a$ is also a generator of $S$. That is, all generators of $S'$, except possibly $F'+m$, were also generators of $S$. Thus we have $e(S') = e(S)$ if $F' + m$ is a generator of $S'$, and $e(S') = e(S) - 1$ otherwise.

## Weights

The *weight* of a numerical semigroup $S$ is the sum of the positive integers not contained in $S$. For example, the weight of $\langle 4, 6, 9\rangle = \{0, 4, 6, 8, 9, 10, 12, 13, \dots\}$ is $1 + 2 + 3 + 5 + 7 + 11 = 29$.

10. (a) **[2]** Compute the weight of (i) $\langle 5, 7, 11, 16\rangle$ and (ii) $\langle 3, 7, 8\rangle$.
    (b) **[4]** Write, with proof, the weight of $S$ in terms of its Apéry coefficients $k_i$. (You may leave your answer as a summation, but only over $i$.)
    (c) **[10]** Find, with proof, the weight of $\langle a, b\rangle$ in terms of $a$ and $b$.

**Solution to Problem 10:**

(a) (i) 46. (ii) 12.

(b) The positive integers not in $S$, in terms of its Apéry coefficients, are just $\{1, m + 1, \dots, (k_1 - 1)m + 1, 2, m + 2, \dots, (k_2 - 1)m + 2, \dots, m - 1, m + m - 1, \dots, (k_{m-1} - 1)m + m - 1\}$. We have

$$\sum_{j=0}^{k_i-1}(jm + i) = m \cdot \sum_{j=0}^{k_i-1} j + ik_i = m \cdot \frac{k_i(k_i - 1)}{2} + ik_i$$

so the weight is

$$\sum_{i=1}^{m-1}\left(m \cdot \frac{k_i(k_i - 1)}{2} + ik_i\right).$$

(c) We extend the computations performed in Problem 4c. Recall that we wrote $ib = aq_i + r_i$ for $i = 1, \dots, a - 1$ and $0 \leq r_i < a$. $r_i$ cycle through the numbers $1, \dots, a - 1$ as $i$ goes from 1 to $a - 1$. Additionally, the numbers congruent to $r_i$ mod $a$ that are not in $\langle a, b\rangle$ are precisely the $q_i$ numbers $r_i, a + r_i, \dots, (q_i - 1)a + r_i$.

From these preliminaries, we see that we wish to compute

$$\sum_{i=1}^{a-1}\sum_{j=0}^{q_i-1} aj + r_i = \sum_{i=1}^{a-1} q_i r_i + \frac{aq_i(q_i - 1)}{2}.$$

Let $S = \sum_{i=1}^{a-1} \frac{1}{2}aq_i^2 + q_i r_i$, so that the desired sum is

$$S - \frac{a}{2}\sum_{i=1}^{a-1} q_i = S - \frac{a(a - 1)(b - 1)}{4}$$

by the computations in 4c.

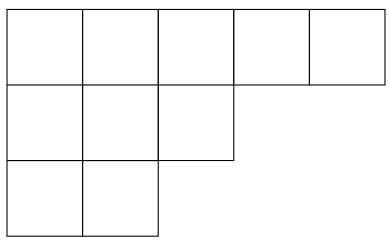
Now, we write $(ib)^2 = (aq_i + r_i)^2 = a^2q_i^2 + 2aq_ir_i + r_i^2$ and sum over $i$ to get

$$b^2\sum_{i=1}^{a-1} i^2 = \frac{b^2a(a-1)(2a-1)}{6} = 2aS + \sum_{i=1}^{a-1} r_i^2 = 2aS + \frac{a(a-1)(2a-1)}{6}.$$

This implies $S = \frac{(b^2-1)(a-1)(2a-1)}{12}$. Plugging back in, we see that the weight of $\langle a, b\rangle$ is

$$\frac{(b^2-1)(a-1)(2a-1)}{12} - \frac{a(a-1)(b-1)}{4} = \frac{(a-1)(b-1)((b+1)(2a-1)-3a)}{12}$$
$$= \boxed{\frac{(a-1)(b-1)(2ab-a-b-1)}{12}}.$$

A *partition* of a positive integer $n$ is a list of positive integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ such that $\lambda_1 + \lambda_2 + \cdots + \lambda_k = n$. For example, the distinct partitions of 4 are $4, 3+1, 2+2, 2+1+1$, and $1+1+1+1$. Each $\lambda_i$ is called a *part* of a partition. Given a partition $\lambda = \lambda_1 + \cdots + \lambda_k$ of $n$, the *Ferrers-Young* diagram of $\lambda$ consists of a row of $\lambda_1$ boxes, underneath which is a left-aligned row of $\lambda_2$ boxes, underneath which is a left-aligned row of $\lambda_3$ boxes, and so on. For example, the following figure is the Ferrers-Young diagram of the partition $5+3+2$ of 10.



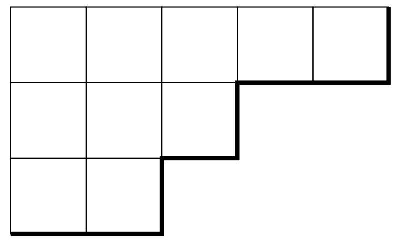
Given a box in a Ferrers-Young diagram, its associated *hook* is itself together with the boxes below it and the boxes to its right. The *size* or *length* of the hook is the number of boxes it contains. For example, the top left box in the Ferrers-Young diagram of $5+3+2$ is associated with a hook of length 7. All the hook lengths for the same partition are shown below.

| | | | | |
|---|---|---|---|---|
| 7 | 6 | 4 | 2 | 1 |
| 4 | 3 | 1 | | |
| 2 | 1 | | | |

The *hookset* of a partition $\lambda$, denoted $H_\lambda$, is the set of hook lengths which appear in the Ferrers-Young diagram of $\lambda$. For example, the hookset of $5+3+2$ is $\{1, 2, 3, 4, 6, 7\}$.

11. (a) **[8]** Let $p(x, y, z)$ be the number of partitions of $x$ into at most $y$ parts, each of size at most $z$. Prove that the number of numerical semigroups $S$ with genus $g$, multiplicity $m$, and weight $w$ satisfying $m < F(S) < 2m$ is exactly $p(w-(g-m+1), g-m+1, 2m-2-g)$.

    (b) **[10]** Prove that given any $\lambda$, the set $\mathbb{N}_0 \setminus H_\lambda$ is a numerical semigroup. (Possible hint: think of the Ferrers-Young diagram of $\lambda$ as a partial grid whose edges one may walk along, and consider the walk starting at the bottom left corner and traversing the lower right edges of the diagram, as shown below.)




(c) **[10]** Prove that given any numerical semigroup $S$, there exists a partition $\lambda$ with $H_\lambda = \mathbb{N}_0 \setminus S$.

**Solution to Problem 11:**

(a) Suppose $\mathbb{N}_0 \setminus S = \{1, 2, \ldots, m-1, m+i_1, \ldots, m+i_{g-m+1}\}$ with $i_a \in [1, m-1]$ for all $a$. We have

$$\begin{aligned} w &= 1 + 2 + \cdots + m - 1 + (m + i_1) + \cdots + (m + i_{g-m+1}) \\ &\quad - (1 + 2 + \cdots + m - 1 + m + \cdots + g) \\ &= \sum_{a=1}^{g-m+1} (i_a - a + 1), \end{aligned}$$

which can be rearranged as

$$w - (g - m + 1) = \sum_{a=1}^{g-m+1} (i_a - a).$$

The $i_a - a$ are nonnegative because $i_1 \geq 1$ and $i_a > i_{a+1}$. They are non-decreasing since $i_{a+1} - (a + 1) \geq i_a + 1 - (a + 1) = i_a - a$. Finally, since $m - 1 \geq i_{g-m+1}$ and $i_{g-m+1} - (g - m + 1) \geq i_a - a$, we have $2m - 2 + g \geq i_a - a$. Thus each distinct choice of these $i_a$ is associated with a unique partition of $w - (g - m + 1)$ into at most $g - m + 1$ parts, each of size at most $2m - 2 - g$. Furthermore, from any such partition $j_1 + \cdots + j_{g-m+1}$, where $0 \leq j_1 \leq \cdots \leq j_{g-m+1}$, it is easy to reconstruct $S$ by setting $i_a = j_a + a$; the resulting $i_a$ will be strictly increasing and bounded above by $m - 1$, as desired.

(b) Let $N$ be the length of the hook associated with the top left square in the Ferrers-Young diagram of $\lambda$. Then the walk described in the hint has $N + 1$ total steps; denote the right steps by $R$ and the up steps by $U$ (so the first step is an $R$ and the last step is a $U$), and number the steps from 0 to $N$. Note that each pair of steps $(i, i + j)$ where $i, j \geq 0$ and $0 \leq i, i + j \leq N$ such that step $i$ is an $R$ and step $i + j$ is a $U$ corresponds uniquely to a hook of the diagram, with the length of the hook being $j$.

Suppose $a, b \in \mathbb{N}_0 \setminus H_\lambda$ but $a + b \in H_\lambda$, and choose $i$ so that there is a hook of length $a + b$ beginning at step $i$ and ending at step $i + a + b$ (that is, there is an $R$ step at $i$ and a $U$ step at $i + a + b$). Since there is no hook of length $a$, step $i + a$ cannot be a $U$ step (otherwise the pair $(i, i + a)$ would give a hook of length $a$); hence step $i + a$ is an $R$ step. But then the pair $(i + a, i + a + b)$ starts with an $R$ step and ends with a $U$ step, hence gives a hook of length $b$, contradiction.

(c) Write $\mathbb{N}_0 \setminus S = \{n_1 < \cdots < n_g\}$, and consider the Ferrers-Young diagram whose walk along the bottom right edges (as described in the hint to part b) has $n_g + 1$ steps, with a $U$ at steps $n_1, \ldots, n_g$ and an $R$ everywhere else. By pairing the $R$ at step 0 with the $U$s at steps $n_1, \ldots, n_g$, we can see that $n_1, \ldots, n_g \in H_\lambda$. We need to check that if $a \in S$, $a \notin H_\lambda$.

Suppose to the contrary that for some $a \in S$ there is a hook of length $a$; that is, there is $i$ so that there is an $R$ at step $i$ and a $U$ at step $i + a$. Then, by the construction of the walk, $i \in S$ and $i + a \in \mathbb{N}_0 \setminus S$. But we assumed that $a \in S$, so $i + a \in S$, contradiction.