NATURAL EXACT COVERING SYSTEMS AND THE REVERSION OF THE MÖBIUS SERIES

I. P. GOULDEN, ANDREW GRANVILLE, L. BRUCE RICHMOND, AND JEFFREY SHALLIT

Abstract. We prove that the number of natural exact covering systems of cardinality k is equal to the coefficient of x^k in the reversion of the power series $\sum_{k\geq 1}\mu(k)x^k$, where $\mu(k)$ is the usual number-theoretic Möbius function. Using this result, we deduce an asymptotic expression for the number of such systems.

1. Introduction

A *covering system* or *complete residue system* is a collection of finitely many residue classes such that every integer belongs to at least one of the classes. The concept was introduced by Erdős [E50, E52], who used the system of 6 congruences

$$\begin{array}{ll} x\equiv 0 \pmod 2 & x\equiv 0 \pmod 3\\ x\equiv 1 \pmod 4 & x\equiv 3 \pmod 8\\ x\equiv 7 \pmod{12} & x\equiv 23 \pmod{24} \end{array}$$

to prove that there exists an infinite arithmetic progression of odd integers, each of which is not representable as the sum of a prime and a power of two. Since 1950, hundreds of papers have been written on covering systems; for surveys of the topic, see, for example [P81, Z82, PS02, S05].

A covering system is called *exact* if every integer belongs to one and only one of the given congruences. (The system of Erdős above is *not* exact, because the integer 19 belongs to both residue classes 3 (mod 8) and 7 (mod 12).) Exact covering systems, or ECS, are sometimes also called *exactly covering systems* or *disjoint covering systems* in the literature (e.g., [N71, F72, F73, NZ74, Z75, S86]).

Among the exact covering systems, one particular subclass that has received attention consists of the so-called *natural exact covering systems*, or NECS. These are the exact covering systems that can be obtained, starting from the single congruence $x \equiv 0 \pmod 1$, by a finite number of applications of the following transformation: for some $r \geq 2$, remove a single congruence $x \equiv a \pmod n$ from the system and replace it with the r new congruences $x \equiv a + jn \pmod{rn}$, $j = 0, 1, \ldots, r-1$; the value of r can vary for the different applications of this transformation.

Date: December 8, 2017.

2010 *Mathematics Subject Classification.* Primary 05A15, 11A07, 11A25; Secondary 05A16, 11N37, 68R15.

Key words and phrases. exact covering system, disjoint covering system, natural exact covering system, Möbius function, congruence, constant gap sequence, generating series, asymptotics.

The work of IPG and JS was supported by NSERC Discovery Grants.

The first paper that we could find describing natural exact covering systems is Porubský [P74], but he credits Znám with introducing them in an earlier unpublished manuscript. Since then, they have been studied by a number of others, e.g., Burshtein [B76a, B76b], Znám [Z82] and Korec [K84], but up to now it appears that nobody has enumerated them. Let a_k denote the number of NECS consisting of k congruences, $k \geq 1$. Define the formal power series

$$A(x) = \sum_{k\geq 1} a_k x^k, \qquad M(x) = \sum_{k\geq 1} \mu(k) x^k, \tag{1}$$

where μ is the usual number-theoretic *Möbius function* defined by $\mu(1) = 1$ and, for $n \geq 2$,

$$\mu(n) = \begin{cases} 0, & \text{if } n \text{ is divisible by a square} > 1; \\ (-1)^e, & \text{if } n \text{ is the product of } e \text{ distinct primes.} \end{cases}$$

The initial terms of these series are given by

$$\begin{aligned} A(x) &= x + x^2 + 3x^3 + 10x^4 + 39x^5 + 160x^6 + 691x^7 + 3081x^8 + \dots, \\ M(x) &= x - x^2 - x^3 - x^5 + x^6 - x^7 + x^{10} - x^{11} - x^{13} + x^{15} + \dots, \end{aligned}$$

where the coefficients in the generating series $A(x)$ were obtained by counting the NECS with at most 8 congruences.

Our main result is that the series A, the generating function for the number of NECS, is the reversion (compositional inverse) of the Möbius series M. Of course, equivalently, this means that M is the reversion of A.

Theorem 1. *The series $A(x)$ is the unique solution to the functional equation*

$$M(A(x)) = x,$$

with initial condition $A(0) = 0$.

Remark 1. The coefficients of the reversion of the Möbius series are given by sequence A050385 in the On-Line Encyclopedia of Integer Sequences [SL].

Remark 2. We remark that, instead of the Möbius power series $M(x)$ defined in (1), it is more usual to study the Dirichlet series $\sum_{k\geq 1} \mu(k) k^{-s}$. Indeed Hardy and Wright [HW60, p. 257] refer to series such as $M(x)$ as "extremely difficult to handle". The generating series $M(e^{-y})$ was mentioned by Hardy and Littlewood [HL16, p. 122], and its unusual analytic properties were studied by Fröberg [F66].

Remark 3. A priori, it is not even obvious that the coefficients in the reversion of $M(x)$ are all positive. This fact follows from our results.

We have not been able to use Theorem 1 to determine a useful explicit expression for the number a_k of NECS with k congruences. However, we are able to determine the precise asymptotic form for a_k, as a corollary to Theorem 1.

Theorem 2. *Let α be the zero of M' of smallest absolute value in $(-1, 1)$, so that $M'(\alpha) = 0$. Also, let*

$$c = \sqrt{-\frac{M(\alpha)}{2\pi M''(\alpha)}}, \qquad \gamma = M(\alpha)^{-1}.$$

Then the kth coefficient in the reversion of the Möbius series (which is also equal to the number of NECS with k congruences) is asymptotically

$$a_k \sim c\,\gamma^k\,k^{-3/2}.$$

Remark 4. Evaluations to 7 decimal places of the constants in Theorem 2 are given by $\alpha \doteq 0.3229939$, $c \doteq 0.0809423$ and $\gamma \doteq 5.4874522$.

In Sections 2 and 3 of this paper, we describe basic notation and terminology for ECS, NECS and for a set of rooted trees that arises in the study of NECS. The proof of Theorem 1 is given in Section 4. Recurrences and numerical results appear in Section 5. The proof of Theorem 2 is in Section 6. In addition, we prove related results for the number of NECS with k congruences in which the gcd and lcm of the congruences are also specified. In Section 7 we give some combinatorial results about the coefficients. Some open problems are described in Section 8. We make some final remarks in Section 9.

2. Basic notation and definitions for exact covering systems

2.1. **Exact covering systems.** For integers $n \geq 1$ and $0 \leq a < n$, let $\langle a, n\rangle$ denote the residue class $\{x \in \mathbb{Z} \;:\; x \equiv a \pmod{n}\}$. Let $\mathcal{U} \subseteq \mathbb{Z}$. The set of $k \geq 1$ residue classes

$$C = \{\langle a_i, n_i\rangle \;:\; i = 1, \ldots, k\}, \tag{2}$$

is called an *exact covering system* (ECS) of $\mathcal{U}$ when

- the sets $\langle a_i, n_i\rangle$, $i = 1, \ldots, k$ are pairwise *disjoint*, and
- the sets $\langle a_i, n_i\rangle$, $i = 1, \ldots, k$ *cover* the set $\mathcal{U}$, i.e.,

$$\dot{\bigcup_{i=1}^{k}} \langle a_i, n_i\rangle = \mathcal{U},$$

where the symbol $\dot\cup$ indicates a disjoint union. Given an ECS C as in (2), suppose that $\gcd\{n_i \;: i = 1, \ldots, k\} = d$, and that $\operatorname{lcm}\{n_i \;:\; i = 1, \ldots, k\} = m$. Then we say that C has *size* k, *gcd* d, and *lcm* m, written

$$|C| = k, \qquad \gcd(C) = d, \qquad \operatorname{lcm}(C) = m.$$

In the case that $\mathcal{U} = \mathbb{Z} = \langle 0, 1\rangle$, we say more simply that C is an ECS (without mentioning the set $\mathbb{Z}$). Table 1 lists all ECS of size at most 4, together with their gcd and lcm.

Remark 5. Note that the gcd of an ECS need not equal its smallest modulus, and the lcm need not equal its largest modulus. The smallest counterexample to both of these claims is

$$\{\langle 1, 4\rangle, \langle 3, 4\rangle, \langle 0, 6\rangle, \langle 2, 6\rangle, \langle 4, 6\rangle\},$$

an ECS of size 5, gcd 2, and lcm 12.

size	exact covering system	gcd	lcm
1	$\{\langle 0,1\rangle\}$	1	1
2	$\{\langle 0,2\rangle,\langle 1,2\rangle\}$	2	2
3	$\{\langle 0,3\rangle,\langle 1,3\rangle,\langle 2,3\rangle\}$	3	3
	$\{\langle 0,2\rangle,\langle 1,4\rangle,\langle 3,4\rangle\}$	2	4
	$\{\langle 1,2\rangle,\langle 0,4\rangle,\langle 2,4\rangle\}$	2	4
4	$\{\langle 0,4\rangle,\langle 1,4\rangle,\langle 2,4\rangle,\langle 3,4\rangle\}$	4	4
	$\{\langle 0,2\rangle,\langle 1,6\rangle,\langle 3,6\rangle,\langle 5,6\rangle\}$	2	6
	$\{\langle 0,3\rangle,\langle 1,3\rangle,\langle 2,6\rangle,\langle 5,6\rangle\}$	3	6
	$\{\langle 0,3\rangle,\langle 2,3\rangle,\langle 1,6\rangle,\langle 4,6\rangle\}$	3	6
	$\{\langle 1,2\rangle,\langle 0,6\rangle,\langle 2,6\rangle,\langle 4,6\rangle\}$	2	6
	$\{\langle 1,3\rangle,\langle 2,3\rangle,\langle 0,6\rangle,\langle 3,6\rangle\}$	3	6
	$\{\langle 0,2\rangle,\langle 1,4\rangle,\langle 3,8\rangle,\langle 7,8\rangle\}$	2	8
	$\{\langle 0,2\rangle,\langle 3,4\rangle,\langle 1,8\rangle,\langle 5,8\rangle\}$	2	8
	$\{\langle 1,2\rangle,\langle 0,4\rangle,\langle 2,8\rangle,\langle 6,8\rangle\}$	2	8
	$\{\langle 1,2\rangle,\langle 2,4\rangle,\langle 0,8\rangle,\langle 4,8\rangle\}$	2	8

Table 1. The ECS of size at most 4.

We now define two constructions for ECS. First, consider an ECS

$$C=\{\langle a_i,n_i\rangle\ :\ i=1,\dots,k\},$$

and a residue class $\langle b,c\rangle$. Define

$$E_{\langle b,c\rangle}(C)=\{\langle b+c\,a_i,c\,n_i\rangle\ :\ i=1,\dots,k\},\tag{3}$$

which we refer to as the $\langle b,c\rangle$-*expansion* of C. Note that $E_{\langle b,c\rangle}(C)$ is itself an ECS of $\langle b,c\rangle$.

Second, consider an ECS C, a residue class $\langle a,n\rangle\in C$, and an integer $r\geq 2$. Let C' be the set of residue classes obtained by removing $\langle a,n\rangle$ from C and replacing it by the r residue classes

$$\langle a+jn,rn\rangle,\qquad\qquad j=0,1,\dots,r-1.\tag{4}$$

Note that the residue classes in (4) are pairwise disjoint, and that they contain between them all integers in $\langle a,n\rangle$; indeed, they are an ECS of the residue class $\langle a,n\rangle$ of size r . An immediate consequence of this is that C' is also an ECS, with size given by $|C'|=|C|+r-1$. We say that the collection of classes in (4) is the r-*split* of $\langle a,n\rangle$, and that C' is an r-split of C. Equivalently, we say that the collection of classes in (4) is obtained by r-*splitting* $\langle a,n\rangle$, and that C' is obtained by r-splitting C. We also use the terms *split* and *splitting* in these same contexts, when we don't choose to specify the value of $r\geq 2$.

2.2. **Natural exact covering systems.** Let $\mathcal{A}$ be the set of exact covering systems that can be obtained by a finite sequence (possibly empty) of splits applied to $\{\langle 0,1\rangle\}=\{\mathbb{Z}\}$; we call this a *split sequence*. It is important to note that if this split sequence is an r_1-split, an r_2-split, … , an r_m-split, for some $m\geq 0$, then the values of $r_1,\dots,r_m$ need not be the same, and indeed can vary arbitrarily over $r_1,\dots,r_m\geq 2$ (and m can be any non-negative integer).

The exact covering systems in $\mathcal{A}$ are called *natural exact covering systems* (NECS). It is well known that in general (see, e.g., [K84, p. 392]) elements of $\mathcal{A}$ can be obtained by a split sequence in more than one way. As an easy example of this, the NECS of size 6

$$\{\langle 0,6\rangle,\langle 1,6\rangle,\langle 2,6\rangle,\langle 3,6\rangle,\langle 4,6\rangle,\langle 5,6\rangle\} \tag{5}$$

can be obtained by 6-splitting $\{\langle 0,1\rangle\}$, but it can also be created by the following split sequence of length 3: first 2-split $\{\langle 0,1\rangle\}$, then 3-split $\{\langle 0,2\rangle\}$, then 3-split $\{\langle 1,2\rangle\}$.

It has also long been known that not every ECS is an NECS, e.g., Porubský [P74]. The smallest size for an ECS that is not an NECS is 13, [S15, Example 3.1], which might seem surprisingly large when one first encounters the study of ECS. In particular, this means that all of the ECS in Table 1 are also NECS (equivalently, the caption for Table 1 could also be "The NECS of size at most 4.").

3. Rooted ordered trees and NECS

3.1. **Rooted ordered trees.** Let $\mathcal{T}$ be the set of rooted ordered trees with a finite (nonempty) set of vertices, in which each vertex has r ordered *children*, for some r in $\{0,2,3,4,\ldots\}$. The rooted tree consisting of a single (root vertex) is in $\mathcal{T}$, and we denote this rooted tree by ε. We regard the trees in $\mathcal{T}$ as being embedded in the plane, with the root vertex at the bottom, and the children of each vertex above that vertex, ordered from left to right. Hence we refer to a vertex with r children as a vertex of *up-degree* r, where $r=0$ or $r\geq 2$. We denote the set of trees in which the up-degree of the root vertex is r by $\mathcal{T}^{(r)}$, so we have $\mathcal{T}=\mathcal{T}^{(0)}\cup\mathcal{T}^{(2)}\cup\mathcal{T}^{(3)}\cup\cdots$, and $\mathcal{T}^{(0)}=\{\varepsilon\}$.

A vertex of up-degree 0 in a tree is called a *leaf*. Thus the root vertex in the tree ε is a leaf (even though it has degree 0 in the graph sense). For a tree T in $\mathcal{T}$ rooted at vertex w, the *height* of w in T is 0, and the height of any other vertex v in T is the edge-length of the unique path in T from the root w to v. The *height* of the tree T itself, denoted by $\mathrm{ht}(T)$, is equal to the maximum height among the leaves in T. Also, the number of leaves in T is denoted by $\lambda(T)$. For example, we have $\mathrm{ht}(\varepsilon)=0$ and $\lambda(\varepsilon)=1$, since the single vertex in ε is a leaf at height 0.

For vertices $u\neq v$ in tree $T\in\mathcal{T}$, we say that u is a *descendant* of v when either u is a child of v, or is (recursively) the descendant of any child of v. We denote the subtree of T whose vertices consist of v and its descendants by T_v. For any vertex v of $T\in\mathcal{T}$, we have $T_v\in\mathcal{T}$. In particular, if v has up-degree r in T, then $T\in\mathcal{T}^{(r)}$. Finally, if v is the root vertex of T, then we have $T_v=T$.

In Figure 1 we display an example of a tree T in $\mathcal{T}$ with root vertex w. The root vertex has up-degree 3 (so $T\in\mathcal{T}^{(3)}$), and its three children, in order, are vertices x,y,z. The tree has $\lambda(T)=10$ leaves; 1 (namely y) at height 1, 3 at height 2, 4 at height 3, and 2 at height 4. Hence the height of the tree is $\mathrm{ht}(T)=4$. For the subtrees T_x, T_y, T_z of T rooted at x,y,z, respectively, we have $\lambda(T_x)=4$, $\mathrm{ht}(T_x)=2$, $T_y=\varepsilon$ (so $\mathrm{ht}(T_y)=0$, $\lambda(T_y)=1$), and $\lambda(T_z)=5$, $\mathrm{ht}(T_z)=3$.

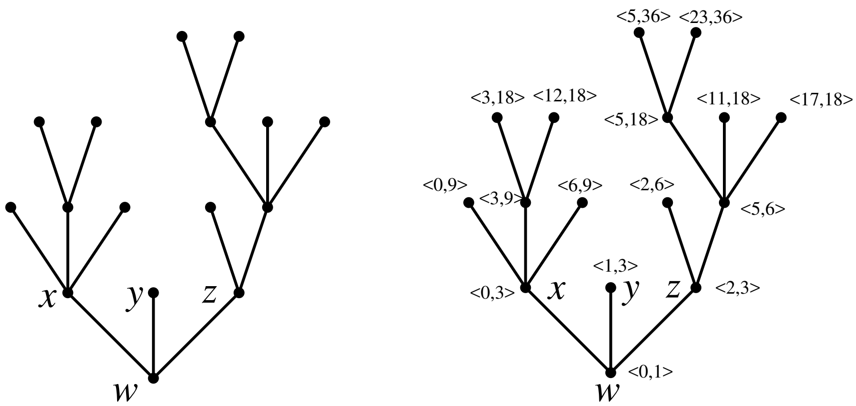

FIGURE 1. Left: a rooted ordered tree in the set $\mathcal{T}$. Right: the same tree with ρ values assigned to vertices as in §3.2.

3.2. **The relationship between trees and NECS.** The reason for introducing the set of trees $\mathcal{T}$ is that every NECS can be represented by at least one tree in $\mathcal{T}$. To see this, consider a tree T in $\mathcal{T}$, and for each vertex v in $\mathcal{T}$, assign a residue class $\rho(v)$ (which we'll refer to as "assigning a ρ value" to v), using the following iterative assignment procedure:

- Initially, assign $\rho(w) = \langle 0, 1\rangle$, where w is the root vertex of T;
- At every stage, find a non-leaf vertex u of T that has been assigned a ρ value, but whose children have not, and suppose that $\rho(u) = \langle a, n\rangle$, for some $0 \le a < n$. Then, denoting the children of u by $c_1, \ldots, c_r$, in order (i.e., u has up-degree r for some $r \ge 2$), assign

$$\rho(c_{j+1}) = \langle a + jn, rn\rangle, \qquad j = 0, 1, \ldots, r-1.$$

  Stop when there is no such vertex u.

Compare the assignments $\rho(c_{j+1})$ given above for the children of a vertex with up-degree r to the r-split of residue classes defined in (4). Clearly, the successive stages for assigning ρ values to vertices of a tree in $\mathcal{T}$ in the iterative assignment procedure above, starting with assigning the ρ value $\langle 0, 1\rangle$ to the root, is a realization of a split sequence (one split for each vertex that is not a leaf) applied to $\langle 0, 1\rangle$, and hence corresponds to an element of $\mathcal{A}$. The order of the splits in the split sequence corresponds to any total order of the set of non-root, non-leaf vertices in the tree that is a linear extension of the following partial order: $u \prec v$ when v is a descendant of u.

Recall that when we apply an r-split in an element of $\mathcal{A}$, we replace one residue class by a set of r residue classes. Hence, in the corresponding tree in $\mathcal{T}$, when we assign a label to a vertex u, we want the residue class $\rho(u)$ to disappear, and to be replaced by the residue classes of its children. This means that if the leaves in a tree $T \in \mathcal{T}$ are given by $\ell_1, \ldots, \ell_k$, $k \ge 1$, then the NECS corresponding to T is exactly the set of leaf labels $\{\rho(\ell_1), \ldots, \rho(\ell_k)\}$, and we write

$$\chi(T) = \{\rho(\ell_1), \ldots, \rho(\ell_k)\}.$$

For example, for the tree T in $\mathcal{T}$ displayed on the left side of Figure 1, using the assignment procedure we initially assign $\rho(w) = \langle 0, 1\rangle$, then at the first stage we assign $\rho(x) = \langle 0, 3\rangle$, $\rho(y) = \langle 1, 3\rangle$, $\rho(z) = \langle 2, 3\rangle$. Upon completion of all stages of the assignment procedure, each vertex of T has a ρ value, which is displayed beside the vertex on the right side of Figure 1. We thus obtain

$$\chi(T) = \{\langle 0,9\rangle, \langle 3,18\rangle, \langle 12,18\rangle, \langle 6,9\rangle, \langle 1,3\rangle, \langle 2,6\rangle, \langle 5,36\rangle, \langle 23,36\rangle, \langle 11,18\rangle, \langle 17,18\rangle\}, \tag{6}$$

where this set of 10 residue classes gives the assigned ρ values for the 10 leaves in T.

The correspondence that we have denoted by χ, between the set of rooted trees in $\mathcal{T}$ and the set $\mathcal{A}$ of NECS, is standard, and has appeared in the literature on NECS from the beginning, e.g., Porubský [P74] and Znám [Z82]. In [P74] each vertex is identified with the residue class that we "assign" in our description above, and in [Z82] the trees are treated in a slightly different but equivalent way using the notion of *product-distance*, and are referred to as $\mathbb{Z}$-*trees*.

It is also standard that this correspondence between $\mathcal{T}$ and $\mathcal{A}$ is not one-to-one; this is simply a restatement of the fact we mentioned in Section 2.2 above, that in general the elements of $\mathcal{A}$ can be obtained by a split sequence in more than one way. We summarize this situation in the following result.

Proposition 1. *The function*

$$\chi\colon \mathcal{T} \to \mathcal{A}\colon T \mapsto C$$

is a surjection, in which $\lambda(T) = |C|$.

Remark 6. Proposition 1 makes it clear that the NECS of size k correspond to certain equivalence classes of trees in $\mathcal{T}$ with k leaves. The problem of counting the total number of trees in $\mathcal{T}$ with k leaves is well known. Let t_k denote the number of trees in $\mathcal{T}$ with k leaves, $k \geq 1$, and $T(x) = \sum_{k\geq 1} t_k x^k$. Then the t_k are called *Schröder* numbers, and the generating function has the closed form

$$T(x) = \frac{1}{4}\Big(1 + x - \sqrt{1 - 6x + x^2}\Big),$$

see, e.g., [FS09, p. 69, 70] for a detailed description. An asymptotic form for t_k also appears in [FS09, p. 474, 475]:

$$t_k \sim \omega\left(3 + 2\sqrt{2}\right)^k k^{-3/2}, \qquad \omega = \frac{1}{4\sqrt{\pi\left(3 + 2\sqrt{2}\right)}}.$$

Of course, t_k is an upper bound for a_k, reflecting the fact that $3+2\sqrt{2} \doteq 5.828$, the asymptotic growth rate for t_k, is larger than $\gamma \doteq 5.487$, the asymptotic growth rate for a_k appearing in Theorem 2.

3.3. **Subtrees rooted at children of the root and NECS.** For $n \geq 2$ and $T \in \mathcal{T}^{(n)}$, suppose that the children of the root vertex of T are $x_1, \ldots, x_n$, in order. When we apply our iterative assignment procedure to T, we obtain $\rho(x_i) = \langle i-1, n\rangle$, for $i = 1, \ldots, n$. Then the residue classes that are assigned to the leaves in T_{x_i} form an NECS of $\langle i-1, n\rangle$, and it is easy to check that this NECS is simply $E_{\langle i-1,n\rangle}(\chi(T_{x_i}))$, using the expansion notation defined in (3). Putting the residue classes for all subtrees together, we thus obtain the relationship

$$\chi(T) = \dot{\bigcup_{i=1}^{n}} E_{\langle i-1,n\rangle}(\chi(T_{x_i})). \tag{7}$$

For example, for the tree T in $\mathcal{T}$ in Figure 1, the root vertex w has up-degree 3, and the children of the root are x, y, z, in order. For the subtrees rooted at x, y, z, we have

$$\chi(T_x) = \{\langle 0,3\rangle, \langle 1,6\rangle, \langle 4,6\rangle, \langle 2,3\rangle\},$$
$$\chi(T_y) = \{\langle 0,1\rangle\},$$
$$\chi(T_z) = \{\langle 0,2\rangle, \langle 1,12\rangle, \langle 7,12\rangle, \langle 3,6\rangle, \langle 5,6\rangle\},$$

and the appropriate expansions of these NECS are given by

$$E_{\langle 0,3\rangle}(\chi(T_x)) = \{\langle 0,9\rangle, \langle 3,18\rangle, \langle 12,18\rangle, \langle 6,9\rangle\},$$
$$E_{\langle 1,3\rangle}(\chi(T_y)) = \{\langle 1,3\rangle\},$$
$$E_{\langle 2,3\rangle}(\chi(T_z)) = \{\langle 2,6\rangle, \langle 5,36\rangle, \langle 23,36\rangle, \langle 11,18\rangle, \langle 17,18\rangle\}.$$

Comparing these with (6), we have

$$\chi(T) = E_{\langle 0,3\rangle}(\chi(T_x)) \ \dot\cup\ E_{\langle 1,3\rangle}(\chi(T_y)) \ \dot\cup\ E_{\langle 2,3\rangle}(\chi(T_z)),$$

confirming that relationship (7) holds for the tree T in Figure 1.

In the following results, we record some useful properties for subtrees of children of the root that follow immediately from (7).

Proposition 2. *For $T \in \mathcal{T}^{(n)}$, $n \geq 2$, suppose that the children of the root vertex of T are $x_1, \ldots, x_n$, in order, and that we have*

$$\chi(T) = C, \qquad \chi(T_{x_i}) = C_i, , \quad i = 1, \ldots, n.$$

Then

(a) $|C| = |C_1| + \cdots + |C_n|$,
(b) $\gcd(C) = n \cdot \gcd\{\gcd(C_1), \ldots, \gcd(C_n)\}$,
(c) $\operatorname{lcm}(C) = n \cdot \operatorname{lcm}\{\operatorname{lcm}(C_1), \ldots, \operatorname{lcm}(C_n)\}$.

Proposition 3. *For $P, Q \in \mathcal{T}^{(n)}$, $n \geq 2$, suppose that the children of the root vertex of P (respectively, Q) are $y_1, \ldots, y_n$ (respectively, $z_1, \ldots, z_n$), in order. Then $\chi(P) = \chi(Q)$ if and only if $\chi(P_{y_i}) = \chi(Q_{z_i})$, $i = 1, \ldots, n$.*

Now we turn to a different type of result for rooted trees T, in which we give a bijective construction that preserves the corresponding NECS $\chi(T)$. For compactness in stating the result, let $\mathcal{G}_{a,b}$ denote the set of trees $T \in \mathcal{T}^{(a)}$ such that all children of the root vertex have up-degree b, $a, b \geq 2$.

Lemma 1. *For $a, b \geq 2$, there is a bijection*

$$\mathcal{T}^{(ab)} \to \mathcal{G}_{a,b} \colon T \mapsto S$$

with $\chi(T) = \chi(S)$.

Proof. Consider $T \in \mathcal{T}^{(ab)}$, and let the children of the root vertex w of T be $x_1, \ldots, x_{ab}$, in order. We construct a tree $S \in \mathcal{G}_{a,b}$ as follows: The root vertex u of S, and its children $y_1, \ldots, y_a$, in order, are newly created vertices (i.e., they are not vertices in T). For $i = 1, \ldots, a$, vertex y_i at height 1 in S has b children, given by vertices $x_i, x_{i+a}, \ldots, x_{i+(b-1)a}$ of T, in order. For $m = 1, \ldots ab$, the construction of S is completed by rooting subtree T_{x_m} of T at vertex x_m of S (to become the subtree S_{x_m} of S).

Note that when we apply our iterative assignment procedure to T, we obtain $\rho(x_m) = \langle m-1, ab\rangle$, for $m = 1, \ldots, ab$ (in which x_m is regarded as a vertex in T). Also, it is easy to check that when we apply our iterative assignment procedure to S, we obtain $\rho(x_m) = \langle m-1, ab\rangle$, for $m = 1, \ldots, ab$ (in which x_m is regarded as a vertex in S). But this implies that all vertices in the subtrees $T_{x_m} = S_{x_m}$, $m = 1, \ldots, ab$, will be assigned the same ρ values, and hence that $\chi(T) = \chi(S)$.

This construction is clearly reversible, and thus gives the required bijection. $\square$

We now apply Lemma 1 to prove a result that will be key for our proof of the main result in Section 4. Again for compactness in stating the result, let $\mathcal{D}_n$ denote the set of trees $T \in \mathcal{T}$ such that $n \mid \gcd(\chi(T))$, $n \geq 2$.

Proposition 4. *For $n \geq 2$,*

$$\{\chi(S) \colon S \in \mathcal{D}_n\} = \{\chi(T) \colon T \in \mathcal{T}^{(n)}\}.$$

Proof. Let $\Omega_1 = \{\chi(S) \colon S \in \mathcal{D}_n\}$ and $\Omega_2 = \{\chi(T) \colon T \in \mathcal{T}^{(n)}\}$. We will prove that $\Omega_1 = \Omega_2$ by proving both containments $\Omega_1 \subseteq \Omega_2$ and $\Omega_2 \subseteq \Omega_1$.

First, to prove $\Omega_2 \subseteq \Omega_1$. Proposition 2(b) implies that $\mathcal{T}^{(n)} \subseteq \mathcal{D}_n$, from which $\Omega_2 \subseteq \Omega_1$ follows immediately.

Second, to prove $\Omega_1 \subseteq \Omega_2$. It is sufficient to prove the following implication: For all $n \geq 2$ and $S \in \mathcal{D}_n$, there exists $T \in \mathcal{T}^{(n)}$ such that $\chi(T) = \chi(S)$. We prove this by induction on the height of S. For the base case, consider $S \in \mathcal{D}_n$ of height equal to 1. Hence, from Proposition 2(b), $S \in \mathcal{T}^{(nb)}$ for some $b \geq 1$, in which the nb children of the root in S are leaves. If $b = 1$, then $S \in \mathcal{T}^{(n)}$, giving the result immediately. If $b \geq 2$, then Lemma 1 with $a = n$ implies that there exists a tree $R \in \mathcal{G}_{n,b} \subseteq \mathcal{T}^{(n)}$ with $\chi(R) = \chi(S)$, proving the result in this case.

Assume that the implication is true for all $n \geq 2$ and trees in $\mathcal{D}_n$ of height at most k, for some $k \geq 1$. Consider $n \geq 2$ and an arbitrary tree $S \in \mathcal{D}_n$ with $\mathrm{ht}(S) = k + 1$. Using Proposition 2(b), there are three cases:

- $S \in \mathcal{T}^{(n)}$, which gives the result immediately.
- $S \in \mathcal{T}^{(nb)}$ for some $b \geq 2$. Then Lemma 1 with $a = n$ implies that there exists a tree $R \in \mathcal{G}_{n,b} \subseteq \mathcal{T}^{(n)}$ with $\chi(R) = \chi(S)$, proving the result in this case.
- $S \in \mathcal{T}^{(a)}$ for some proper divisor $a \geq 2$ of n, so $n = ab$, $a, b \geq 2$. In this case, in addition, if $x_1, \ldots, x_a$ are the children of the root vertex of S, then from Proposition 2(b) we have

  $$b \mid \gcd\{\gcd(\chi(S_{x_1})), \ldots, \gcd(\chi(S_{x_a}))\},$$

  and hence $b \mid \gcd(\chi(S_{x_i}))$ for $i = 1, \ldots, a$. Equivalently, $S_{x_i} \in \mathcal{D}_b$ for $i = 1, \ldots, a$. But $\mathrm{ht}(S_{x_i}) \leq k$ for $i = 1, \ldots, k$. Hence, for $i = 1, \ldots, k$, from the induction hypothesis, there exists $R_{x_i} \in \mathcal{T}^{(b)}$ (also rooted at vertex x_i) such that $\chi(R_{x_i}) = \chi(S_{x_i})$. Now construct a tree R by removing the subtree S_{x_i} from S, and replacing it by the subtree R_{x_i}, for $i = 1, \ldots, a$. Note that $R \in \mathcal{G}_{a,b}$, and that $\chi(R) = \chi(S)$, from Proposition 3. Then, Lemma 1 implies that there exists a tree $Q \in \mathcal{T}^{(ab)} = \mathcal{T}^{(n)}$ with $\chi(Q) = \chi(R) = \chi(S)$, proving the result in this case.

This completes the inductive proof of the implication, and thus that $\Omega_1 \subseteq \Omega_2$. $\square$

4. Proof of the main result

It will be convenient to partition the set $\mathcal{A}$ of NECS according to gcd. Hence let $\mathcal{A}_m$ denote the set of NECS with gcd m, for $m \geq 1$. Recall that a_k is the number of NECS of size k, $k \geq 1$, and let $a_{k,m}$ denote the number of NECS of size k and gcd m, $k, m \geq 1$. As in (1), we have the generating function

$$A(x) = \sum_{k\geq 1} a_k x^k = \sum_{C\in\mathcal{A}} x^{|C|}, \tag{8}$$

and we define the additional generating functions

$$A_m(x) = \sum_{k\geq 1} a_{k,m} x^k = \sum_{C\in\mathcal{A}_m} x^{|C|}, \qquad m \geq 1. \tag{9}$$

Of course, these generating functions are related by

$$A(x) = \sum_{m\geq 1} A_m(x). \tag{10}$$

Note that the situation for $C \in \mathcal{A}$ with $\gcd(C) = 1$ is particularly simple: We must have $C = \chi(T)$ for some $T \in \mathcal{T}^{(0)}$, from Proposition 2(b). But the only tree in $\mathcal{T}^{(0)}$ is the single-vertex tree ε, and we have $\chi(\varepsilon) = \langle 0, 1\rangle$. Since $\langle 0, 1\rangle$ has both size and gcd equal to 1, we conclude that $\mathcal{A}_1 = \{\langle 0, 1\rangle\}$, and hence from (9) that

$$A_1(x) = x. \tag{11}$$

In part (a) of the following result, we prove a functional equation for the generating functions $A(x)$ and $A_m(x), m \geq 1$, that is a generalization of equation (10) above. The proof that we give for part (b) of the result is to apply Möbius inversion to part (a), which is the reason that the Möbius series M defined in (1) appears in the statement of part (b).

Theorem 3.

(a) For $n \geq 1$ we have

$$A(x)^n = \sum_{d\geq 1} A_{nd}(x).$$

(b) For $m \geq 1$ we have

$$M\big(A(x)^m\big) = A_m(x).$$

Proof. (a) For $n = 1$, the result is given by (10) above.

For $n \geq 2$, we begin the proof by defining $\mathcal{U}(\mathcal{A}, n) = \cup_{d\geq 1}\mathcal{A}_{nd}$, the set of all $C \in \mathcal{A}$ such that $n \mid \gcd(C)$. Now, from Proposition 4 and Proposition 1, we have

$$\mathcal{U}(\mathcal{A}, n) = \{\chi(T) \colon T \in \mathcal{T}^{(n)}\}.$$

But removing the root vertex from a tree in $\mathcal{T}^{(n)}$ to obtain an ordered list of n rooted trees in $\mathcal{T}$ (namely, the subtrees rooted at the n children of the root vertex), yields the usual bijection between $\mathcal{T}^{(n)}$ and $\mathcal{T}^n$ (the set of n-tuples of elements of $\mathcal{T}$). Together with Proposition 3, as well as Proposition 1, this implies that there is a bijection between $\{\chi(T) \colon T \in \mathcal{T}^{(n)}\}$ and the set $\mathcal{A}^n$ of n-tuples of NECS in $\mathcal{A}$. Putting these pieces together and eliminating the set $\{\chi(T) \colon T \in \mathcal{T}^{(n)}\}$ yields a bijection

$$\mathcal{U}(\mathcal{A}, n) \to \mathcal{A}^n \colon C \mapsto (C_1, \ldots, C_n). \tag{12}$$

Moreover, in the above bijection, from Proposition 2, C and its image $(C_1, \dots, C_n)$ are related by the equations

$$|C| = |C_1| + \ldots + |C_n|, \tag{13}$$

$$\gcd(C) = n \cdot \gcd\{\gcd(C_1), \dots, \gcd(C_n)\}, \tag{14}$$

$$\operatorname{lcm}(C) = n \cdot \operatorname{lcm}\{\operatorname{lcm}(C_1), \dots, \operatorname{lcm}(C_n)\}. \tag{15}$$

Now we turn to generating functions. Applying bijection (12) to the range of summation below, and using relation (13), we obtain

$$\sum_{C \in \mathcal{U}(\mathcal{A},n)} x^{|C|} = \sum_{(C_1,\dots,C_n)\in\mathcal{A}^n} x^{|C_1|+\ldots+|C_n|}.$$

But

$$\sum_{C \in \mathcal{U}(\mathcal{A},n)} x^{|C|} = \sum_{d\geq 1} \sum_{C\in\mathcal{A}_{nd}} x^{|C|} = \sum_{d\geq 1} A_{nd}(x),$$

from (9), and

$$\sum_{(C_1,\dots,C_n)\in\mathcal{A}^n} x^{|C_1|+\cdots+|C_n|} = \prod_{i=1}^{n} \sum_{C_i\in\mathcal{A}} x^{|C_i|} = A(x)^n,$$

from (8), completing the proof of the result for $n \geq 2$.

(b) For fixed $m \geq 1$, replace n in part (a) of this result by mn, multiply on both sides by $\mu(n)$, and sum over $n \geq 1$, to obtain

$$\sum_{n\geq 1} \mu(n) A(x)^{mn} = \sum_{n\geq 1} \mu(n) \sum_{d\geq 1} A_{mnd}(x).$$

But the right-hand side of this equation can be rewritten as

$$\sum_{n\geq 1} \mu(n) \sum_{d\geq 1} A_{mnd}(x) = \sum_{i\geq 1} A_{mi}(x) \sum_{n|i} \mu(n) = A_m(x),$$

using the standard fact (see, e.g., [HW60, p. 235]) that

$$\sum_{n|i} \mu(n) = \begin{cases} 1, & \text{if } i = 1; \\ 0, & \text{if } i \geq 2. \end{cases} \tag{16}$$

Hence we obtain the equation

$$\sum_{n\geq 1} \mu(n) A(x)^{mn} = A_m(x), \tag{17}$$

or $M\big(A(x)^m\big) = A_m(x)$, as required.

□

We are now able to prove the main result, as a simple consequence of the above Theorem.

Proof of Theorem 1. Specializing Theorem 3(b) to the case $m = 1$ gives $M(A(x)) = A_1(x)$. The result follows immediately from (11). □

5. Recurrences specifying size, gcd and lcm

In the following result we give a recurrence equation for the numbers $a_{k,n}$ of NECS with size k and gcd m.

Proposition 5. *For $k, n, d \geq 1$,*

$$\sum a_{j_1,m_1} \cdots a_{j_n,m_n} = a_{k,nd}, \tag{18}$$

where the summation on the left-hand side is over $j_1, \ldots, j_n \geq 1$ and $m_1, \ldots, m_n \geq 1$ such that

$$j_1 + \cdots + j_n = k, \qquad \textit{and} \qquad \gcd\{m_1, \ldots, m_n\} = d. \tag{19}$$

Proof. For $n = 1$, the result is simply the identity $a_{k,d} = a_{k,d}$.

For $n \geq 2$, consider bijection (12). The number of elements in the set $\mathcal{U}(\mathcal{A}, n)$ with size k and gcd nd is give by $a_{k,nd}$. But from the bijection, this is equal to the number of n-tuples $(C_1, \ldots, C_n) \in \mathcal{A}^n$ in which, from (13) and (14), we have

$$|C_1| + \cdots + |C_n| = k, \qquad \text{and} \qquad \gcd\{\gcd(C_1), \ldots, \gcd(C_n)\} = d.$$

The result for $n \geq 2$ follows immediately. □

Remark 7. Proposition 5, with $d = 1$, can be used to list all elements of $\mathcal{A}_{k,n}$, the set of NECS with size k and gcd n, and also to count their number $a_{k,n}$, using either dynamic programming, or recursion together with memoization. We briefly describe this second approach. The base cases of the recursion are $k = n$ (for which the only NECS is $\{\langle 0, k\rangle, \langle 1, k\rangle, \ldots, \langle k-1, k\rangle\}$) and $n = 1$ (for which the only NECS is $\{\langle 0, 1\rangle\}$ corresponding to $k = 1$). Given k and n as input, we can easily compute all $\binom{k-1}{n-1}$ compositions of k into n positive parts (using, for example, the algorithm in [NW78, Chap. 5]). We now discard those compositions whose gcd is greater than one. For each composition $(j_1, j_2, \ldots, j_n)$ that remains, we consider all $j_1 j_2 \cdots j_n$ of the n-tuples $(m_1, m_2, \ldots, m_n)$ satisfying $1 \leq m_i \leq j_i$ for $i = 1, 2, \ldots n$. For each element (j_i, m_i) in the list of pairs $((j_1, m_1), \ldots, (j_n, m_n))$ we recursively compute all the NECS C_i in $\mathcal{A}_{j_i,m_i}$. Using the expansion construction given in (3), we finally create the NECS

$$\dot{\bigcup_{i=1}^{n}} E_{\langle i-1,n\rangle}(C_i).$$

If we are only interested in counting these NECS, we sum all the products $a_{j_1,m_1} \cdots a_{j_n,m_n}$ instead.

Using an implementation of this algorithm written in APL, we computed $a_{k,n}$ for $1 \leq n \leq k \leq 22$. We report the results for $1 \leq n \leq k \leq 13$ in Table 2.

$k\backslash n$	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1												
2	0	1											
3	0	2	1										
4	0	6	3	1									
5	0	22	12	4	1								
6	0	88	48	18	5	1							
7	0	372	207	80	25	6	1						
8	0	1636	918	366	120	33	7	1					
9	0	7406	4188	1700	580	170	42	8	1				
10	0	34276	19488	8026	2810	864	231	52	9	1			
11	0	161436	92199	38384	13710	4356	1232	304	63	10	1		
12	0	771238	442056	185644	67330	21936	6454	1698	390	75	11	1	
13	0	3728168	2143329	906472	332825	110562	33523	9232	2277	490	88	12	1

TABLE 2. Table of values for $a_{k,n}$, $1 \le k, n \le 13$

Now let $a_{k,m,\ell}$ denote the number of NECS of size k, gcd m and lcm ℓ, for $k, m, \ell \ge 1$. Of course, $a_{k,m,\ell} = 0$ unless $m \mid \ell$. The following is a version of Proposition 5 that records the lcm as well as size and gcd. It can be proved in the same way, by considering the bijection (12) together with relations (13), (14) and (15).

Proposition 6. *For $k, n, d, D \ge 1$,*

$$\sum a_{i_1,j_1,\ell_1} \cdots a_{i_n,j_n,\ell_n} = a_{k,nd,nD},$$

where the sum on the left-hand side is over $i_1, \dots, i_n \ge 1$, $j_1, \dots, j_n \ge 1$ and $\ell_1, \dots, \ell_n \ge 1$ such that

$$i_1 + \cdots + i_n = k, \qquad \gcd\{j_1, \dots, j_n\} = d, \quad \text{and} \quad \operatorname{lcm}\{\ell_1, \dots, \ell_n\} = D.$$

6. Asymptotic growth of coefficients

We now turn to asymptotics, and begin with the proof of Theorem 2.

Proof of Theorem 2. All coefficients in the series $M(y) = \sum_{k\ge1} \mu(k) y^k$ have absolute value equal to 0 or 1, so this series converges for all $|y| < 1$ (e.g., using comparison with the geometric series $\sum_{n\ge1} y^{n-1}$, in which all coefficients are equal to 1, and which converges for all $|y| < 1$). Hence $M(y)$ is analytic for $|y| < 1$. Then for every real $y_0 \in (-1, 1)$, from the Taylor series we obtain

$$M(y) = M(y_0) + M'(y_0)(y - y_0) + \frac{M''(y_0)}{2}(y - y_0)^2 + \frac{M'''(y_0)}{6}(y - y_0)^3 + \mathcal{O}((y - y_0)^4). \tag{20}$$

Note that the zero of M' of smallest absolute value in $(-1, 1)$ is given by

$\alpha \doteq 0.3229939133028335399812256469630856932020517484175227624423337334463495349 9$

(so we have $M'(\alpha) = 0$). Also, note that if we define $\rho = M(\alpha)$, $\delta = M''(\alpha)$, then we have

$$\begin{aligned} \rho = M(\alpha) &\doteq 0.18223393401633630828235226904174072905168066104, \\ \delta = M''(\alpha) &\doteq -4.4268862524695752516745518331111866104593741941617 38. \end{aligned}$$

Now in (20) substitute $y_0 = \alpha$ (noting that $\alpha \in (-1, 1)$), and

$$y = A(x). \tag{21}$$

Applying Theorem 1, which gives $M(y) = x$, we obtain

$$x = \rho + \frac{\delta}{2}(y-\alpha)^2 + c_1(y-\alpha)^3 + \mathcal{O}((y-\alpha)^4),$$

where c_1 is some constant. Rearranging, we get

$$-\frac{2\rho}{\delta}(1 - \frac{x}{\rho}) = (y-\alpha)^2 + c_2(y-\alpha)^3 + \mathcal{O}((y-\alpha)^4),$$

where c_2 is some constant. But $\rho > 0$ and $\delta < 0$, so we have $-\frac{2\rho}{\delta} > 0$, and taking square roots, we obtain

$$-\sqrt{-\frac{2\rho}{\delta}}\left(1 - \frac{x}{\rho}\right)^{\frac{1}{2}} = y - \alpha + c_3(y-\alpha)^2 + \mathcal{O}(y-\alpha)^3),$$

where we have selected the negative square root since y is increasing as x approaches ρ from below. Solving this for $y - \alpha$, and applying (21) to eliminate y, we obtain

$$A(x) - \alpha = -\sqrt{-\frac{2\rho}{\delta}}\left(1 - \frac{x}{\rho}\right)^{\frac{1}{2}} + c_4\Big(1 - \frac{x}{\rho}\Big) + \mathcal{O}\left(\left(1 - \frac{x}{\rho}\right)^{\frac{3}{2}}\right), \tag{22}$$

where c_4 is some constant.

In order to determine the asymptotic behaviour of the coefficients in $A(x)$ from (22), we follow the treatment in [FS09, Chapter VI], referred to as *singularity analysis*. From Theorems VI.1 on page 381 and VI.3 on page 390, we obtain

$$a_k \sim -\sqrt{-\frac{2\rho}{\delta}}\,\rho^{-k}\frac{k^{-\frac{3}{2}}}{\Gamma(-\frac{1}{2})}\left(1 + \mathcal{O}\Big(\frac{1}{k}\Big)\right),$$

where Γ is the usual Gamma function. Now, recalling that $\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}$, the result follows, with

$$\gamma = \rho^{-1} \qquad \text{and} \qquad c = \sqrt{-\frac{\rho}{2\delta\pi}}\ .$$

Decimal approximations to the constants γ and c follow:

$$\begin{aligned} \gamma &\doteq 5.4874521882974621475674452932303092553200429102 4 \\ c &\doteq 0.080942294186097300358615771233555317510353812 67. \end{aligned}$$

□

Remark 8. We have found that $M'(x)$ has two real zeros in the open interval $(-\frac{2}{3}, \frac{2}{3})$, which are α given in the proof above, and

$$\beta \doteq -0.562976540744649358189645954216416402249939799218087618317349878994076506622.$$

Eight-digit approximations were previously given by Fröberg [F66].

Theorem 3(b) gives a closed form for the generating series $A_m(x) = \sum_{k\geq 1} a_{k,m}x^k$, in terms of the series $A(x)$ and $M(x)$. Once again, we have not been able to use this result to determine a useful explicit expression for the kth coefficient $a_{k,m}$ in $A_m(x)$. However, in the next result, we are able to determine a precise asymptotic form for the coefficients $a_{k,m}$, following on from the proof of Theorem 2 above.

Theorem 4. *For each fixed $m \geq 2$, the number $a_{k,m}$ of NECS of size k with gcd m is asymptotically*

$$a_{k,m} \sim m\alpha^{m-1}M'(\alpha^m)\, c\, \gamma^k\, k^{-3/2},$$

where $\alpha \doteq 0.3229939$, $c \doteq 0.0809423$ and $\gamma \doteq 5.4874522$.

Proof. From (22), taking the mth power, we have

$$A(x)^m = \alpha^m - m\alpha^{m-1}\sqrt{-\frac{2\rho}{\delta}}\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + c_5\Big(1-\frac{x}{\rho}\Big) + \mathcal{O}\left(\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}}\right), \tag{23}$$

where c_5 is some constant, and α, ρ, δ are specified in the proof of Theorem 2. In particular, since $\alpha \in (0,1)$, then $\alpha^m \in (0,1)$ for every positive integer m. Now using the linear expansion

$$M(a+z) = M(a) + M'(a)z + \mathcal{O}(z^2),$$

and (23), we obtain

$$M(A(x)^m) = M(\alpha^m) - m\alpha^{m-1}M'(\alpha^m)\sqrt{-\frac{2\rho}{\delta}}\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + c_6\Big(1-\frac{x}{\rho}\Big) + \mathcal{O}\left(\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}}\right),$$

where c_6 is some constant. But $M\big(A(x)^m\big) = A_m(x)$ from Theorem 3(b), and we now determine the asymptotic behaviour of the coefficients in $A_m(x)$ from the above expansion. We again use the technique of singularity analysis as described in [FS09, Chapter VI]. From Theorems VI.1 on page 381 and VI.3 on page 390, we thus obtain

$$a_{k,m} \sim -m\alpha^{m-1}M'(\alpha^m)\sqrt{-\frac{2\rho}{\delta}}\,\rho^{-k}\frac{k^{-\frac{3}{2}}}{\Gamma(-\frac{1}{2})}\left(1+\mathcal{O}\Big(\frac{1}{k}\Big)\right),$$

where $\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}$. The result follows, with

$$\gamma = \rho^{-1} \qquad \text{and} \qquad c = \sqrt{-\frac{\rho}{2\delta\pi}}\,.$$

$\square$

Remark 9. Comparing the asymptotic forms in Theorems 2 and 4, we observe that

$$a_{k,m} \sim m\alpha^{m-1}M'(\alpha^m)\, a_k.$$

But $\sum_{m\geq 2} a_{k,m} = a_k$ for $k \geq 2$, so we should have

$$\sum_{m\geq 2} m\alpha^{m-1}M'(\alpha^m) = 1. \tag{24}$$

Here is a direct proof of (24) (which therefore provides a consistency check on our asymptotic results): A well known series identity (see, e.g., [HW60, p. 258]) that follows immediately from (16) is given by

$$\sum_{k\geq 1} \mu(k)\, \frac{x^k}{1-x^k} = x.$$

The summation on the left-hand side is referred to as the *Lambert series* for the Möbius function. Rewriting the left-hand side in terms of the series M itself, we obtain

$$\sum_{m\geq 1} M(x^m) = x.$$

Differentiating on both sides of this equation with respect to x gives

$$\sum_{m\geq 1} m x^{m-1} M'(x^m) = 1, \tag{25}$$

and this holds for any $x \in (-1, 1)$. Our proof of (24) is then completed by substituting $x = \alpha$ in (25), and noting that $M'(\alpha) = 0$.

Remark 10. Along similar lines, note that Theorem 4 also holds for $m = 1$, but the result is trivial in this case, since $M'(\alpha) = 0$. Hence the result for $m = 1$ states that the number $a_{k,1}$ is asymptotically 0. This is consistent with (11), which states that $A_1(x) = \sum_{k\geq 1} a_{k,m} x^k = x$, and hence $a_{k,1} = 0$ for all $k \geq 2$.

7. A formula for $a_{g+n,n}$

Inspection of the downward-sloping diagonals in Table 2 suggests that for each n there is a polynomial $f_n(x)$ such that $a_{g+n,n} = f_n(g)$ for $g > n$. Furthermore, it seems that $f_n(x)$ is a polynomial of degree n, with leading coefficient $x^n/n!$ and constant term 0, and all other coefficients positive. The first few such polynomials seem to be as follows (expressed in the basis of binomial coefficients):

$$\begin{aligned}
f_1(x) &= \binom{x}{1} \\
f_2(x) &= \binom{x}{2} + 3\binom{x}{1} \\
f_3(x) &= \binom{x}{3} + 6\binom{x}{2} + 10\binom{x}{1} \\
f_4(x) &= \binom{x}{4} + 9\binom{x}{3} + 29\binom{x}{2} + 39\binom{x}{1} \\
f_5(x) &= \binom{x}{5} + 12\binom{x}{4} + 57\binom{x}{3} + 138\binom{x}{2} + 160\binom{x}{1} \\
f_6(x) &= \binom{x}{6} + 15\binom{x}{5} + 94\binom{x}{4} + 324\binom{x}{3} + 654\binom{x}{2} + 691\binom{x}{1}.
\end{aligned}$$

Furthermore it appears that the ℓ'th differences of the coefficients of the ℓ'th column above is 3^ℓ.

We now explain these empirical observations. We start with the following lemma about formal power series.

Lemma 2. *Suppose $F(x) = 1 + c_1x + c_2x^2 + \cdots$ is a formal power series with constant coefficient 1. The coefficient of x^m in $F(x)^n$ equals*

$$\sum_{k=1}^{m} c_{m,k}\binom{n}{k}$$

where $c_{m,k}$ is the coefficient of x^m in $(F(x)-1)^k$. We have $c_{m,m} = c_1^m$ and $c_{m,1} = c_m$. Furthermore, if $c_j \geq 0$ for all j, then $c_{m,k} > 0$ for all k, $1 \leq k \leq m$.

Proof. We have

$$F(x)^n = (1 + (F(x)-1))^n = \sum_{k=0}^{n}\binom{n}{k}(F(x)-1)^k.$$

As $F(x) - 1$ begins with an x-term, we see that the coefficient of x^m in $(F(x)-1)^k$ is 0 for all $k > m$. The first part of the result follows . By the multinomial theorem,

$$c_{m,k} = \sum_{\substack{r_1,r_2,\ldots,r_m \in \mathbb{N} \\ r_1+2r_2+\cdots+mr_m=m \\ r_1+\cdots+r_m=k}} \binom{k}{r_1,\ldots,r_m}\prod_{j=1}^{m} c_j^{r_j}.$$

When $k = m$ then $r_1 = m$ and the other $r_j = 0$, so $c_{m,m} = c_1^m$. When $k = 1$ then $r_m = 1$ and the other $r_j = 0$, so $c_{m,1} = c_m$. □

Now let $f_n(g) = a_{g+n,g}$, the number of NECS of size $g + n$ and gcd g, so that

$$\sum_{n\geq 0} f_n(g)x^n = A_g(x)/x^g = \sum_{m\geq 1}\mu(m)(A(x)^m/x)^g,$$

where we have used Eq. (17). Now $A(x)$ equals x plus higher order terms. Hence, if $m \geq 2$ and $g > n$, then the coefficient of x^n in $(A(x)^m/x)^g$ is 0. Therefore

If $g > n$ then $f_n(g)$ is the coefficient of x^n in $(A(x)/x)^g$.

Now $A(x)/x = 1 + x + 3x^2 + \cdots$ has all positive coefficients (cf. Remark 3). Therefore we may apply Lemma 2 with $F(x) = A(x)/x$ to obtain the following:

Corollary 1. *Fix an integer $n \geq 1$. If $g > n$ then the number of NECS with size $g + n$ and gcd g is given by a polynomial*

$$\sum_{k=1}^{n} c_{n,k}\binom{g}{k}.$$

in which the coefficients $c_{n,k}$ are all positive and, written as a polynomial, the leading term is $g^n/n!$ and the constant term is 0. We also have $c_{n,1} = a_{k+1}$.

Let us now turn to a better formula for $c_{m+\ell,m}$ for fixed $\ell \geq 1$, and with $m \geq \ell$. To do so, we write $A(x)/x = 1 + x + xB(x)$ where x divides $B(x)$. Now $c_{m+\ell,m}$ is the coefficient of $x^{m+\ell}$ in $(\frac{A(x)}{x} - 1)^m = (x(1 + B(x)))^m$, which equals the coefficient of x^ℓ in $(1 + B(x))^m$. As x divides $B(x)$ this implies that

$c_{m+\ell,m}$ equals the coefficient of x^ℓ in $\sum_{h=0}^{\ell}\binom{m}{h}B(x)^h$.

The ℓth backward difference of $(c_{m+\ell,m})$ is

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j c_{m+\ell-j,m-j},$$

which equals the coefficient of x^ℓ in

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j \sum_{h=0}^{\ell} \binom{m-j}{h} B(x)^h = \sum_{h=0}^{\ell} \left(\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j \binom{m-j}{h} \right) B(x)^h.$$

From the theory of finite differences we know, for all m, that

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j \binom{m-j}{h} = \begin{cases} 0, & \text{if } 0 \le h < l; \\ 1, & \text{if } h = l. \end{cases}$$

Therefore the ℓth backward difference of $(c_{m+\ell,m})$ equals the coefficient of x^ℓ in $B(x)^\ell$, which is the leading coefficient, and so equals c_2^ℓ. In our special case this gives, if $\ell \ge 1$ and $m \ge \ell$, then

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j c_{m+\ell-j,m-j} = 3^\ell,$$

as observed in the data.

8. Open Problems

In this section we list three related problems for which we currently have no solution.

Problem 1. Suppose that, instead of counting distinct NECS of size k, we count equivalence classes under "shift". That is, we consider two NECS to be identical if one can be transformed into the other by a transformation of the form $x = x' + C$, for some integer constant C. How many equivalence classes are there? We list the number $s(k)$ for $1 \le n \le 12$:

k	1	2	3	4	5	6	7	8	9	10	11	12
$s(k)$	1	1	2	4	10	26	75	226	718	2368	8083	28367

What is a good formula for $s(k)$? What is the asymptotics of $s(k)$?

Problem 2. Suppose we consider those NECS of size k, and ask how many distinct values of the lcm parameter they can take on. Call the resulting sequence $t(k)$. The first few values are given below.

k	1	2	3	4	5	6	7	8	9	10	11	12
$t(k)$	1	1	2	3	6	8	15	18	31	35	56	62

What is a good formula for $t(k)$? What is the asymptotics of $t(k)$?

Problem 3. Suppose we consider the enumeration of ECS instead of NECS. Hence let b_k denote the number of ECS of size k, $k \ge 1$, and $b_{k,m}$ denote the number of ECS of size k and gcd m, $k, m \ge 1$. Define the generating functions

$$B(x) = \sum_{k \ge 1} b_k x^k, \qquad B_m(x) = \sum_{k \ge 1} b_{k,m} x^k, \quad m \ge 1.$$

It is reasonably straightforward to prove that the analogue of Theorem 3(a) holds for these series, so we have $B(x)^n = \sum_{d\geq 1} B_{nd}(x)$, for $n \geq 1$. Then, applying Möbius inversion, we obtain $M(B(x)^m) = B_m(x)$, $m \geq 1$, the analogue of Theorem 3(b). Specializing to the case $m = 1$ gives

$$M(B(x)) = B_1(x), \tag{26}$$

where $B_1(x)$ is the generating function for the ECS with gcd equal to 1. But it turns out that

$$B_1(x) = x + 30\,x^{13} + \mathcal{O}\big(x^{14}\big),$$

a substantially different situation from the NECS, where we had $A_1(x) = x$. In fact, the $b_{13,1} = 30$ ECS with gcd 1 are the only ECS of size at most 13 that are not also ECS. Thus, in Table 2, the values of $a_{k,n}$ that appear in row k and column n are equal to $b_{k,n}$ everywhere except for row 13 and column 1. However, for larger values of k the gap between the numbers of ECS and NECS grows rapidly, though we have no idea how rapidly. What would we have to know about the growth rate for the coefficients of $B_1(x)$ in order to deduce asymptotics for the numbers b_k from the functional equation (26)? Is it possible that the asymptotics for b_k is the same as for a_k?

9. Comments

This paper was originally motivated by a problem dealing with infinite periodic sequences of constant gap. These are maps from $\mathbb{N}$ to a finite alphabet of size k, say $\Sigma_k := \{0, 1, \ldots, k-1\}$, with the property that for each $i \in \Sigma_k$ there exists a constant c_i such that the occurrences of i lie in an arithmetic progression of difference c_i. For example, the infinite periodic sequence $(0102)^\omega = 010201020102\cdots$ is of constant gap with $k = 3$. These sequences have been studied, e.g., in [G73, H96, H00, AGH00].

David W. Wilson [W17] and JS independently conjectured, on the basis of numerical evidence, that the reversion of the Möbius series M counts the number of ECS. This is incorrect; as we have seen, this reversion instead counts the (strict) subclass of NECS. The first place where these two sequences differ is at $k = 13$, where the number of NECS is 7266979 (e.g., this is the total of the entries in row $k = 13$ of Table 2), but the number of ECS is 7267009 (which is 30 larger—see the discussion in Problem 3 above).

Acknowledgments

We thank Cam Stewart for helpful remarks. Vincent Jost kindly informed us by e-mail on November 21 2017 that an earlier version of our paper was flawed, because we did not distinguish correctly between ECS and the special case of NECS. We gratefully thank him for his comments.

References

[AGH00] E. Altman, B. Gaujal, and A. Hordijk, *Balanced sequences and optimal routing*, J. ACM **47** (2000), 752–775.

[B76a] N. Burshtein, *On natural exactly covering systems of congruences having moduli occurring at most M times*, Discrete Math. **14** (1976), 205–214.

[B76b] N. Burshtein, *On natural exactly covering systems having moduli occurring at most twice*, J. Number Theory **8** (1976), 251–259.

[E50] P. Erdős, *On integers of the form $2^k + p$ and related problems*, Summa. Brasil. Math. **2** (1950), 192–210.
[E52] P. Erdős, *Egy kongruenciarendszerekről szóló problémáról*, Mat. Lapok **3** (1952), 122–128.
[FS09] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
[F73] A. S. Fraenkel, *A characterization of exactly covering congruences*, Discrete Math. **4** (1973), 359–366.
[F72] J. Friedlander, *On exact coverings of the integers*, Israel J. Math. **12** (1972), 299–305.
[F66] C.-E. Fröberg, *Numerical studies of the Möbius power series*, BIT **6** (1966), 191–211.
[G73] R. Graham, *Covering the positive integers by disjoint sets of the form* $\{[n\alpha + \beta]\} : n = 1, 2, \ldots$, J. Combin. Theory **15** (1973), 354–358.
[HL16] G. H. Hardy and J. E. Littlewood, *Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes*, Acta Arith. **41** (1916), 119–196.
[HW60] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 4th edition, Oxford University Press, 1960.
[H96] P. Hubert, *Propriétés combinatoires des suites définies par le billard dans les triangles pavants*, Theoret. Comput. Sci. **164** (1996), 165–183.
[H00] P. Hubert, *Suites équilibrées*, Theoret. Comput. Sci. **242** (2000), 91–108.
[K84] I. Korec, *Irreducible disjoint covering systems*, Acta Arith. **44** (1984), 389–395.
[N71] M. Newman, *Roots of unity and covering sets*, Math. Ann. **191** (1971), 279–282.
[NW78] A. Nijenhuis and H. S. Wilf, *Combinatorial Algorithms*, 2nd edition, Academic Press, 1978.
[NZ74] B. Novak and S. Znám, *Disjoint covering systems*, Amer. Math. Monthly **81** (1974), 42–45.
[P74] S. Porubský, *Natural exactly covering systems of congruences*, Czechoslovak Math. J. **24** (1974), 598–606.
[P81] S. Porubský, *Results and problems on covering systems of residue classes*, Mitt. Math. Sem. Giessen **150** (1981), 1–85.
[PS02] S. Porubský and J. Schönheim, *Covering systems of Paul Erdős: past, present and future*, in Paul Erdős and his Mathematics, Vol. I, Bolyai Society Mathematical Studies **11** (2002), 581–627.
[S15] O. Schnabel, *On the reducibility of exact covering systems*, INTEGERS **15** (2015), Paper #A34.
[S86] R. J. Simpson, *Exact coverings of the integers by arithmetic progressions*, Discrete Math. **59** (1986), 181–190.
[SL] N. J. A. Sloane et al., *The On-Line Encyclopedia of Integer Sequences*, available at `http://oeis.org`.
[S05] Z. W. Sun, *Problems and results on covering systems*, a talk given in 2005, available at `http://maths.nju.edu.cn/~zwsun/Cover.pdf`.
[W17] David W. Wilson, remark in [SL], sequence A050385, May 17 2017.
[Z75] S. Znám, *A simple characterization of disjoint covering systems*, Discrete Math. **12** (1975), 89–91.
[Z82] S. Znám, *A survey of covering systems of congruences*, Acta Math. Univ. Comenianae **40/41** (1982), 59–79.

(1) Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1 Canada
E-mail address: `ipgoulde@uwaterloo.ca`

(2) Département de mathématiques et de statistiques, Université de Montréal, CP 6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada
E-mail address: `andrew@dms.umontreal.ca`

(3) Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1 Canada
E-mail address: `lbrichmond@uwaterloo.ca`

(4) School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1 Canada
E-mail address: `shallit@uwaterloo.ca`