PARTS AND SUBWORD PATTERNS IN COMPOSITIONS

BRIAN HOPKINS, MARK SHATTUCK, ANDREW V. SILLS,
THOTSAPORN "AEK" THANATIPANONDA, AND HUA WANG

Abstract. We find relationships between subword patterns and residue classes of parts in the set of integer compositions of a given weight. In particular, we show that it is always possible to express the total number of parts in compositions of n that are congruent to i modulo m as a linear combination of the total number of occurrences of subword patterns of length no more than m. We also find an explicit formula enumerating all such parts.

1. Introduction

This study began with the empirical observation that the number of odd parts $OP(n)$ occurring in the set of compositions of the integer n appeared to be equal to the number of "runs" $R(n)$ (see Definition 2.7 in next section) in the compositions of n. We give a direct bijective proof that this is in fact the case, and from there, move on to present relationships between the number of parts congruent to $i \pmod{m}$ and occurrences of various subword patterns, again over the set of compositions of n, using generating functions. We close with enumeration formulas for the number of parts congruent to i (mod m) among all compositions of n.

Our main result (Theorem 4.5 below) is that $P(i, m; n)$ for $1 \leq i \leq m$ may always be expressed as a linear combination of $S(\tau; m)$ taken over subword patterns of length at most m (see Definitions 2.5 and 2.10 for notation). These relations are computed explicitly for all moduli up to six in the appendix. Also, a direct bijective proof can be given in the case $m = 3$. While there has been work done concerning finding formulas (such as for the generating function) counting compositions

---

2010 *Mathematics Subject Classification.* 05A05, 05A19.

Key words and phrases. compositions, subword patterns, residue classes of parts of compositions.

The work of Andrew Sills was partially supported by National Security Agency grant H98230-14-1-0159.

The work of Hua Wang was partially supported by Simons Foundation grant #245307.

according to the number of occurrences of certain subword patterns (see, e.g., [1, 3, 5], the text [4] and references contained therein), the connection discussed here between the total number of occurrences of a specific kind of part within compositions of a given weight and the number of occurrences of subword patterns is one that does not seem to have been previously considered. Perhaps such a relationship is indicative of deeper connections between the underlying distributions for part sizes and subword occurrences on the set of compositions of n.

2. Definitions and notation

We will make subsequent use of the following definitions and notation.

Definition 2.1. A *composition* σ of a positive integer n is an ℓ-tuple of positive integers $(\sigma_1, \sigma_2, \dots, \sigma_\ell)$ such that $n = \sum_{i=1}^{\ell} \sigma_i$. Each σ_i is called a *part* of σ and the number of parts $\ell = \ell(\sigma)$ is the *length* of σ.

When no confusion can arise, we forgo the commas and parentheses when explicitly writing out a composition.

Often it is useful to consider the composition with no parts, $\varnothing = (\ )$, to be the unique composition of 0.

Definition 2.2. The *weight* of a composition σ, denoted $|\sigma|$, is the sum of the parts of σ.

Definition 2.3. Let $\mathscr{C}_n$ denote the set of all compositions of n. E.g.,

$$\mathscr{C}_4 = \{(4), (3,1), (1,3), (2,2), (2,1,1), (1,2,1), (1,1,2), (1,1,1,1)\}.$$

Definition 2.4. Let $P(k;n)$ denote the number of parts in $\mathscr{C}_n$ equal to k. For example, $P(1;4) = 12$.

Definition 2.5. Let $P(i,m;n)$ denote the number of parts in $\mathscr{C}_n$ congruent to $i \pmod{m}$, where i is the least positive residue of n modulo m. That is, we choose to write multiples of m as congruent to m, not 0, modulo m. Note that

$$P(i,m;n) = P(i;n) + P(i+m;n) + P(i+2m;n) + \cdots .$$

For example, $P(1,3;4) = P(1;4) + P(4;4) = 13$.

Modulus $m = 2$ will be of particular interest, so we make the following definitions.

Definition 2.6. Let $OP(n) = P(1,2;n)$ (resp., $EP(n) = P(2,2;n)$) denote the total number of odd (resp., even) parts in $\mathscr{C}_n$.

Table 1 gives values of $OP(n)$ and $EP(n)$ for $1 \le n \le 10$. Note that both sequences match [6, A059570] with different offsets. (That this relationship holds in general will follow from results in §4.)

TABLE 1. Initial values for number of composition parts by parity.

n	1	2	3	4	5	6	7	8	9	10
$OP(n)$	1	2	6	14	34	78	178	398	882	1934
$EP(n)$	0	1	2	6	14	34	78	178	398	882

We may also denote a composition in the compressed "exponential" notation

$$\langle \sigma_1^{r_1}\sigma_2^{r_2}\cdots\sigma_j^{r_j}\rangle = (\overbrace{\sigma_1,\sigma_1,\ldots,\sigma_1}^{r_1},\overbrace{\sigma_2,\sigma_2,\ldots,\sigma_2}^{r_2},\ldots,\overbrace{\sigma_j,\sigma_j,\ldots,\sigma_j}^{r_j}),$$

where we require $\sigma_i \neq \sigma_{i+1}$ for $1 \le i \le j-1$. E.g.,

$$\mathscr{C}_4 = \{\langle 4^1\rangle, \langle 3^1 1^1\rangle, \langle 1^1 3^1\rangle, \langle 2^2\rangle, \langle 2^1 1^2\rangle, \langle 1^1 2^1 1^1\rangle, \langle 1^2 2^1\rangle, \langle 1^4\rangle\}.$$

Definition 2.7. We call each instance of r_i consecutive parts equal to σ_i a *run* in the composition σ. Let $R(n)$ denote the total number of runs in all compositions of n; i.e., $R(n) = \sum j(\sigma)$, where the sum is taken over all compositions $\sigma = \langle \sigma_1^{r_1}\cdots\sigma_{j(\sigma)}^{r_{j(\sigma)}}\rangle \in \mathscr{C}_n$. E.g., $R(4) = 14$.

We shall be concerned with subword pattern matching in compositions and its relationship to parts in congruence classes modulo m.

Definition 2.8. The *reduced form* of a sequence of parts within a composition (i.e., a subcomposition) $\sigma = \sigma_r\sigma_{r+1}\cdots\sigma_{r+h-1}$ is given by $s_1s_2\cdots s_h$, where $s_i = j$ if σ_i is the jth smallest part of subcomposition σ.

Definition 2.9. A composition σ contains a *subword pattern* $\tau = \tau_1\tau_2\cdots\tau_k$ if the reduced form of any subsequence of k consecutive parts of σ equals τ.

Definition 2.10. Let $S(\tau; n)$ denote the number of occurrences of the subword pattern τ in $\mathscr{C}_n$.

In §4, we will make use of integer partitions.

Definition 2.11. A *partition* of n is a composition of n that is nondecreasing (i.e., contains no matches for the 21 subword).

Definition 2.12. If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ is a partition of n, then we define the *conjugate* λ' of λ to be the partition

$$\langle 1^{\lambda_\ell - \lambda_{\ell-1}} 2^{\lambda_{\ell-1}-\lambda_{\ell-2}} 3^{\lambda_{\ell-2}-\lambda_{\ell-3}} \cdots (\ell-1)^{\lambda_2-\lambda_1} \ell^{\lambda_1} \rangle.$$

Definition 2.12 is consistent with the usual definition of conjugation via transposition of Ferrers graphs.

Definition 2.13. Let $\mathscr{P}_n(T)$ denote the set of all partitions of n whose parts lie in the set T and in which each member of T appears at least once as a part.

3. Modulus 2: runs versus parity of parts

We wish to ultimately express $P(i, m; n)$ as a linear combination of $S(\tau; n)$ for various τ where m is general and $1 \leq i \leq m$. The question is motivated by observations in the $m = 2$ case, which we consider in this section in greater detail.

First note that $S(1; n)$ gives the total number of parts within all members of $\mathscr{C}_n$. Then $P(2, 2; n) = S(11; n)$, upon halving even parts within compositions, and thus

$$P(1, 2; n) = S(1; n) - P(2, 2; n) = S(1; n) - S(11; n).$$

Observe further that $S(1; n) - S(11; n) = R(n)$, since subtracting from the total number of parts of all members of $\mathscr{C}_n$, the total number of occurrences of $\tau = 11$ (i.e., the number of parts that equal their predecessor), leaves the total number of parts within members of $\mathscr{C}_n$ that start runs (i.e., the number of runs). Combining the last two observations yields the following result.

Proposition 3.1. *OP(n)=R(n) for* $n \geq 1$.

The following direct bijective proof of the prior result will yield further extensions of it.

Bijective proof of Proposition 3.1.

Let us call a part of a composition *primary* if it is first within its run. Note that then runs are synonymous with primary parts. Let $\mathcal{C}_n^*$ (resp., $\widetilde{\mathcal{C}}_n$) denote the set of compositions of n in which an odd (resp., primary) part is underlined. To prove Proposition 3.1, it suffices to define a bijection f between $\mathcal{C}_n^*$ and $\widetilde{\mathcal{C}}_n$. Let $\lambda \in \mathcal{C}_n^*$. If the underlined part is primary, then let $f(\lambda) = \lambda$. Otherwise, if the underlined part is d, then replace this d and its predecessor with the single part $2d$. If this part $2d$ is primary in the resulting composition, then underline it and stop. Otherwise, replace the new part $2d$ and its predecessor with $4d$. Then either underline $4d$ and stop or replace two parts of size $4d$

with $8d$ and so on. Note that this process must result at some point in a primary part of size $2^i d$ for some $i \geq 1$ which we underline (as the first part of a composition is primary, by definition). Let $f(\lambda)$ be the resulting member of $\widetilde{\mathcal{C}}_n$.

One may verify that f is a bijection from $\mathcal{C}_n^*$ to $\widetilde{\mathcal{C}}_n$ by constructing its inverse g. To do so, consider the underlined part $2^i d$ of $\rho \in \widetilde{\mathcal{C}}_n$, where $i \geq 0$ and d is odd. If $i = 0$, then let $g(\rho) = \rho$. Otherwise, let $g(\rho)$ be obtained from ρ by replacing $2^i d$ with the sequence of parts $2^{i-1}d, 2^{i-2}d, \ldots, 2d, d, d$ and underlining the second part d. One may verify that $f \circ g = g \circ f$ is the identity map. ■

Remark 1: The bijection above shows in fact that the number of parts of size i within all members of $\mathscr{C}_n$ where i is a given odd number equals the number of runs involving parts of size $2^j i$ for some $j \geq 0$.

Remark 2: The preceding bijection also yields a generalization of Proposition 3.1 for any prime number p. Let us call a part of a composition *p-primary* if it occupies the i-th position within the run to which it belongs for some $1 \leq i \leq p-1$. Then the preceding argument can be extended to show that the number of parts not divisible by p equals the number of p-primary parts within all members of $\mathscr{C}_n$. Taking $p = 2$ implies Proposition 3.1.

The bijection for $n = 5$ is given explicitly in Table 2 below.

TABLE 2. Mapping between $\mathcal{C}_n^*$ and $\widetilde{\mathcal{C}}_n$ when $n = 5$ (only non-fixed points are shown).

odd part	$31\underline{1}$	$1\underline{1}3$	$21\underline{1}1$	$211\underline{1}$	$121\underline{1}$	$1\underline{1}21$
primary part	$3\underline{2}$	$\underline{2}3$	$\underline{4}1$	$21\underline{2}$	$1\underline{4}$	$\underline{2}21$
odd part	$1\underline{1}12$	$11\underline{1}2$	$1\underline{1}111$	$11\underline{1}11$	$111\underline{1}1$	$1111\underline{1}$
primary part	$\underline{2}12$	$1\underline{2}2$	$\underline{2}111$	$1\underline{2}11$	$11\underline{2}1$	$111\underline{2}$

4. Generalizing to longer subwords and greater moduli

We have established the relationship between $OP(n) = P(1, 2; n)$ and $R(n)$, which can be expressed in terms of length two subword patterns. In the remainder of the paper, we consider similar relations between $P(i, m; n)$ for some $m > 2$, which requires analysis of longer subword patterns.

Theorem 4.1. *Let $\tau = \tau_1\tau_2\cdots\tau_\ell$ be a subword pattern of length ℓ where*

$$1 = \tau_1 \le \tau_2 \le \cdots \le \tau_\ell \le \ell.$$

Let $S(\tau;n)$ denote the number of matches to the subword pattern τ among all compositions of n. Considering τ as a composition, $|\tau|$ is its weight. Then

$$\sum_{n\ge 0} S(\tau;n)x^n = \frac{x^{|\tau|}(1-x)^2}{(1-2x)^2(1-x^\ell)}\prod_{j=1}^{\ell-1}\frac{1}{(1-x^{\ell-j})^{\tau_{j+1}-\tau_j}}.$$

Proof. Observe that

$$\frac{x^{|\tau|}(1-x)^2}{(1-2x)^2(1-x^\ell)}\prod_{j=1}^{\ell-1}\frac{1}{(1-x^{\ell-j})^{\tau_{j+1}-\tau_j}} = \frac{1-x}{1-2x}\cdot\frac{x^\ell}{1-x^\ell}\prod_{j=1}^{\ell-1}\frac{x^{(\ell-j)(\tau_{j+1}-\tau_j)}}{(1-x^{\ell-j})^{\tau_{j+1}-\tau_j}}\cdot\frac{1-x}{1-2x}. \quad (4.1)$$

Note that the first and last factors in (4.1) are each

$$\frac{1-x}{1-2x} = \sum_{n\ge 0}|\mathscr{C}_n|x^n,$$

while the middle factor is

$$\frac{x^\ell}{1-x^\ell}\prod_{j=1}^{\ell-1}\frac{x^{(\ell-j)(\tau_{j+1}-\tau_j)}}{(1-x^{\ell-j})^{\tau_{j+1}-\tau_j}} = \sum_{n\ge 0}|\mathscr{P}_n(T)|x^n,$$

where T is the set of parts appearing in τ' when τ is interpreted as a partition rather than a subword pattern. (Note further that the form of τ forces τ' to have distinct parts, i.e., that $\tau_{j+1}-\tau_j \in \{0,1\}$ for all $j = 1,2,\ldots,\ell-1$.)

Thus (4.1) is the generating function for the number of ordered triples (γ,λ,ρ), where $\gamma \in \mathscr{C}_{n_1}$, $\lambda \in \mathscr{P}_{n_2}(T)$, $\rho \in \mathscr{C}_{n_3}$, and $n = n_1+n_2+n_3$.

Let σ be a composition of n containing the subword $\mu = \mu_1\mu_2\cdots\mu_\ell$ that matches the subword pattern τ. Thus σ is of the form

$$\sigma = (\gamma_1,\gamma_2,\ldots,\gamma_j,\mu_1,\mu_2,\ldots,\mu_\ell,\rho_1,\rho_2,\ldots,\rho_k),$$

where $(\gamma_1,\ldots,\gamma_j) \in \mathscr{C}_{n_1}$ and $(\rho_1,\ldots,\rho_k) \in \mathscr{C}_{n_3}$ for some nonnegative integers j and k, with $n_1 = |(\gamma_1,\ldots,\gamma_j)|$, $n_3 = |(\rho_1,\ldots,\rho_k)|$, $n_2 = |\mu|$, and $n = n_1+n_2+n_3$.

Observe that σ can be mapped bijectively to

$$((\gamma_1,\gamma_2,\ldots,\gamma_j),\mu',(\rho_1,\rho_2,\ldots,\rho_k)),$$

noting that the parts of μ' must all lie in T. This implies that the generating function $\sum_{n\geq 0} S(\tau;n)x^n$ is the common value of the expressions in equation (4.1), as desired. ∎

Table 3 provides a detailed example of the correspondence.

TABLE 3. Theorem 4.1 correspondence for $S(112223, 13)$.

$\mathscr{C}_{13}$ elt.	11112223	11122231	1112224	11222311
triple	$(11, 146, \varnothing)$	(1, 146, 1)	(1, 1146, $\varnothing$)	($\varnothing$, 146, 11)
$\mathscr{C}_{13}$ elt.	1122232	1122241	112225	2112223
triple	($\varnothing$, 146, 2)	($\varnothing$, 1146, 1)	($\varnothing$, 11146, $\varnothing$)	(2, 146, $\varnothing$)

Theorem 4.2.

$$\sum_{n\geq 0} P(i,m;n)x^n = \frac{x^i(1-x)^2}{(1-2x)^2(1-x^m)}.$$

Proof. Applying Theorem 4.1 to the subword pattern $\tau = 1^m$ gives

$$\sum_{n\geq 0} S(1^m;n)x^n = \frac{x^m(1-x)^2}{(1-2x)^2(1-x^m)}$$

(terms in the product are all 1 for this τ). Upon replacing the parts j comprising an occurrence of τ with the single part mj, we have $P(m,m;n) = S(1^m;n)$, which shows that the claim holds for $i = m$.

Let $\mathscr{C}_{n,k}$ be the set of compositions of n in which one part k is underlined. Then the bijection $\underline{k} \leftrightarrow \underline{k+1}$ between $\mathscr{C}_{n,k}$ and $\mathscr{C}_{n+1,k+1}$ implies $P(k;n) = P(k+1;n+1)$. Summing over remainder classes gives $P(i,m;n) = P(i+1,m;n+1)$ so that summing over all n gives the same result for each i. ∎

Given the similarity in the forms of the generating functions for $S(\tau;n)$ and $P(i,m;n)$, one expects them to be related, hopefully in simple ways. After an example and two general relations, we give derivations for $P(2,3;n)$ and $P(1,3;n)$. Subword pattern expressions for all $P(i,m;n)$ through $m = 6$ are given in the appendix.

Example 4.3. From the generating functions in Theorems 4.1 and 4.2, the equality $P(1,2;n) = S(1;n) - S(11;n)$ amounts to

$$\frac{x}{1-x^2} = \frac{x}{1-x} - \frac{x^2}{1-x^2}. \tag{4.2}$$

Theorem 4.4. *Suppose $P(i, m; n)$ can be expressed as a linear combination of the pattern occurrence counts $S(\tau; n)$. Then $P(ki, km; n)$ is given by the same linear combination (same number of terms, same coefficients, etc.), where each $\tau = (\tau_1, \dots, \tau_\ell)$ is replaced by $(\tau_1^k, \dots, \tau_\ell^k)$.*

Proof. This is a direct consequence of Theorems 4.1 and 4.2. Given a linear combination of $\sum S(\tau; n)x^n$ that represents $\sum P(i, m; n)x^n$, replacing every x with x^k, except for those in the common factor $(1-x)^2/(1-2x)^2$, yields a linear combination of $\sum S(\tau^k; n)x^n$ that represents $\sum P(ki, km; n)x^n$. ■

Theorem 4.5. *$P(i, m; n)$ can be represented as a linear combination of $S(\tau; n)$ for subword patterns τ of length no more than m.*

Proof. From (4.1),

$$\sum_{n\geq 0} S(\tau; n)x^n = \frac{(1-x)^2}{(1-2x)^2} \cdot \frac{x^\ell}{1-x^\ell} \prod_{j=1}^{\ell-1} \frac{x^{(\ell-j)(\tau_{j+1}-\tau_j)}}{(1-x^{\ell-j})^{\tau_{j+1}-\tau_j}}$$

$$= \frac{(1-x)^2}{(1-2x)^2} \prod_{k=1}^{\ell} \frac{x^{t_k}}{1-x^{t_k}},$$

where t_k is the cardinality of the set $\{j \mid 1 \leq j \leq \ell, \tau_j \geq k\}$ and the factor $x^{t_k}/(1-x^{t_k})$ is understood to be 1 if $t_k = 0$. By Theorem 4.2,

$$\sum_{n\geq 0} P(i, m; n)x^n = \frac{(1-x)^2}{(1-2x)^2} \cdot \frac{x^i}{(1-x^m)},$$

so a linear representation of $P(i, m; n)$ where $1 \leq i \leq m$ by $S(\tau; n)$, with each τ nondecreasing and of length at most m, is equivalent to a linear representation of $x^i/(1-x^m)$ by terms of the form

$$\prod_{k=1}^{r} \frac{x^{t_k}}{1-x^{t_k}}, \tag{4.3}$$

where $r \geq 1$, $1 \leq t_1 \leq m$, and $t_1 > t_2 > \cdots > t_r \geq 1$.

We proceed by induction on m; the initial cases $m = 1$ and $m = 2$ discussed earlier. Let $m > 2$ and suppose $2 \leq i \leq m$. If i and m share a common factor larger than 1, the conclusion follows from the induction hypothesis and Theorem 4.4. So assume i and m are relatively prime and let $u = \max\{v \mid vi < m\}$. Now write

$$\frac{x^i}{1-x^m} = \frac{x^i}{1-x^i} + \frac{x^m}{1-x^m} \cdot \frac{x^i}{1-x^i} - \frac{x^{2i}}{(1-x^m)(1-x^i)}. \tag{4.4}$$

Focusing on the last term of (4.4), the only one not of the form (4.3), we have

$$\frac{x^{2i}}{(1-x^m)(1-x^i)} = \frac{x^{2i}\left(1+x^i+\cdots+x^{(u-1)i}\right)}{(1-x^m)(1-x^{ui})}$$
$$= \frac{1}{1-x^m}\cdot\left(\frac{x^{2i}}{1-x^{ui}}+\cdots+\frac{x^{ui}}{1-x^{ui}}\right)+\frac{x^{(u+1)i}}{(1-x^m)(1-x^{ui})}$$
$$= \left(1+\frac{x^m}{1-x^m}\right)\cdot\frac{x^{2i}}{1-x^{ui}}+\cdots+\left(1+\frac{x^m}{1-x^m}\right)\cdot\frac{x^{ui}}{1-x^{ui}}$$
$$+\frac{x^m}{1-x^m}\cdot\frac{x^{(u+1)i-m}}{1-x^{ui}}.$$

By the induction hypothesis, each $x^{ji}/(1-x^{ui})$ term and $x^{(u+1)i-m}/(1-x^{ui})$ can be expressed as linear combinations of products of the form (4.3) since $1 \le (u+1)i-m < ui < m$. Together with (4.4), this establishes the claim for $P(i,m;n)$ when $i \ge 2$. If $i=1$, observe that

$$\frac{x+x^2+\cdots+x^{m-1}}{1-x^m} = \frac{x}{1-x} - \frac{x^m}{1-x^m},$$

where each of the $x^i/(1-x^m)$ terms on the left-hand side for $1 < i < m$ is already of the desired form by the preceding. This completes the induction on m. ■

We now give two detailed examples of determining $P(i,m;n)$ in terms of subword pattern counts.

Example 4.6. The next case is $m=3$; we initially determine $P(2,3;n)$. Leaving out the $(1-x)^2/(1-2x)^2$ factor common to the expressions in Theorems 4.1 and 4.2, the remaining factor for $P(2,3;n)$ is $x^2/(1-x^3)$. Using (4.4), we have

$$\frac{x^2}{1-x^3} = \frac{x^2}{1-x^2}+\frac{x^3}{1-x^3}\cdot\frac{x^2}{1-x^2}-\frac{x^4}{(1-x^3)(1-x^2)}$$
$$= \frac{x^2}{1-x^2}+\frac{x^3}{1-x^3}\cdot\frac{x^2}{1-x^2}-\frac{x^3}{1-x^3}\cdot\frac{x}{1-x^2}$$
$$= \frac{x^2}{1-x^2}+\frac{x^3}{1-x^3}\cdot\frac{x^2}{1-x^2}-\frac{x^3}{1-x^3}\left(\frac{x}{1-x}-\frac{x^2}{1-x^2}\right)$$
$$= \frac{x^2}{1-x^2}+2\cdot\frac{x^3}{1-x^3}\cdot\frac{x^2}{1-x^2}-\frac{x^3}{1-x^3}\cdot\frac{x}{1-x},$$

where the substitution in the third line uses (4.2). We conclude that

$$P(2,3;n) = S(11;n)+2S(122;n)-S(112;n).$$

Example 4.7. We could determine a subword pattern expression for $P(1,3;n)$ using the method in the previous example, but instead we demonstrate a more general approach.

From Theorem 4.5, $P(1,3;n)$ is a linear combination $\sum_\tau c_\tau\, S(\tau;n)$, where τ runs over all nondecreasing subword patterns of length at most 3. Explicitly, using the generating functions and coefficients to be determined, we have

$$\begin{aligned}\frac{(1-x)^2x}{(1-2x)^2\,(1-x^3)} &= \frac{c_1(1-x)x}{(1-2x)^2} + \frac{c_{11}(1-x)^2x^2}{(1-2x)^2\,(1-x^2)} \\ &\quad + \frac{c_{12}(1-x)x^3}{(1-2x)^2\,(1-x^2)} + \frac{c_{111}(1-x)^2x^3}{(1-2x)^2\,(1-x^3)} + \frac{c_{112}(1-x)x^4}{(1-2x)^2\,(1-x^3)} \\ &\quad + \frac{c_{122}(1-x)^2x^5}{(1-2x)^2\,(1-x^2)\,(1-x^3)} + \frac{c_{123}(1-x)x^6}{(1-2x)^2\,(1-x^2)\,(1-x^3)}.\end{aligned}$$

After clearing denominators, we equate coefficients of $x, x^2, \ldots, x^{11}$ to obtain a linear system of eleven equations in the seven c_τ unknowns. The general solution is

$$(c_1, c_{11}, c_{12}, c_{111}, c_{112}, c_{122}, c_{123}) = (1, -1, t, -1-t, 1-t, -2-t, 0),$$

where t is a free variable, and so

$$\begin{aligned}P(1,3;n) = S(1;n) - S(11;n) + tS(12;n) - (t+1)S(111;n) \\ + (1-t)S(112;n) - (t+2)S(122;n).\end{aligned}$$

Letting $t=0$ gives the expression

$$P(1,3;n) = S(1;n) - S(11;n) - S(1^3;n) + S(112;n) - 2S(122;n).$$

We conclude this section with a combinatorial proof of the formulas found above for $P(i,m;n)$ when $m=3$.

Bijective proof of $m = 3$ case:

We first prove the formula for $P(2,3;n)$ from Example 4.6. To do so, let D_n denote the set obtained by underlining exactly one part congruent to 2 (mod 3) within each composition of n containing at least one such part. Let E_n be the set of compositions of n in which an occurrence of the subword pattern 221 is marked (which will be done by underlining the first letter within the occurrence), and let F_n be defined the same way as E_n but with the subword 111 instead. Let G_n denote the set of compositions of n ending in an occurrence of 11 (i.e., the last two parts are equal). Since $S(122;n) = S(221;n)$ and $S(11;n) = S(111;n) + S(112;n) + S(221;n) + |G_n|$, the formula for

$P(2,3;n)$ may be rewritten equivalently as

$$|D_n| = 3|E_n| + |F_n| + |G_n|, \qquad n \geq 1.$$

Note that both sides of this equality are zero when $n = 1$, so we may assume $n > 1$. To show it, we will describe a mapping f transforming members of D_n to members of $E_n \cup F_n \cup G_n$ in which each member of E_n arises exactly three times, while those in F_n and G_n arise once. Let $\lambda \in D_n$. If the underlined part of λ equals $3a-1$ where $a > 1$, to obtain $f(\lambda)$, replace $\underline{3a-1}$ by $\underline{a}, a, a-1$, leaving all other parts the same. So assume that the underlined part of λ equals 2. If λ can be written as $\lambda = \lambda'\underline{2}2^k$ for some $k \geq 0$ where λ' is possibly empty, then let $f(\lambda) = \lambda'(k+1)(k+1)$. Finally, assume $\lambda = \lambda'\underline{2}2^\ell d\lambda''$, where $\ell \geq 0$, $d \geq 1$ with $d \neq 2$, and λ', λ'' are possibly empty. Then let $f(\lambda)$ be obtained from λ in this case via the replacements

$$\underline{2}2^\ell d = \begin{cases} \underline{\frac{d+2}{3}+\ell}, \frac{d+2}{3}+\ell, \frac{d+2}{3}, & \text{if } d \equiv 1 \pmod 3; \\ \underline{\frac{d+4}{3}+\ell}, \frac{d+4}{3}+\ell, \frac{d-2}{3}, & \text{if } d \equiv 2 \pmod 3; \\ \underline{\frac{d}{3}+\ell+1}, \frac{d}{3}+\ell+1, \frac{d}{3}, & \text{if } d \equiv 3 \pmod 3, \end{cases}$$

where all other parts are left undisturbed. Allowing ℓ and d to vary, it is seen that each member of E_n in which the marked occurrence of 221 is of the form a, a, b where $a > b+1$ arises three times. Members of E_n in which the marked 221 is of the form $a, a, a-1$ come about only when $\ell = 0$ and $d \equiv 3 \pmod 3$ or when $\ell = 1$ and $d \equiv 1 \pmod 3$. Taken together with the first case above in which the underlined part of λ was greater than 2, we see that these members of E_n also arise exactly three times. Compositions in F_n correspond to the $\ell = 0$, $d \equiv 1 \pmod 3$ case, while allowing $k \geq 0$ to vary in the second case above gives all members of G_n. This completes the proof of the formula for $P(2,3;n)$.

The formula for $P(1,3;n)$ follows from subtraction, upon noting

$$\begin{aligned} P(1,3;n) + P(3,3;n) &= S(1;n) - P(2,3;n) \\ &= S(1;n) - S(11;n) + S(112;n) - 2S(122;n) \end{aligned}$$

and $P(3,3;n) = S(1^3;n)$. One may also obtain a bijective proof for this by applying the complementary bijection principle to the mapping f above. ■

The appendix gives subword pattern expressions for all $P(i,m;n)$ through $m = 6$.

5. Enumeration

In this last section, we derive various enumeration formulas. We will need the following preliminary result.

Theorem 5.1.

$$P(k;n) = \begin{cases} (n-k+3)2^{n-k-2}, & \textit{if } n > k; \\ 1, & \textit{if } n = k. \end{cases}$$

Proof. This theorem collects results of [2, Question 3]. The $k = 1$ case of the formula is established by combinatorial reasoning and recurrence relations. The proof is completed by observing the equality $P(k;n) = P(1;n-k+1)$ (which follows from the fact $P(k;n) = P(k+1;n+1)$ mentioned earlier). For another approach, see [7, p. 120, Ex. 24] and its solution. ∎

We have the following general explicit formula for $P(i, m; n)$.

Theorem 5.2.

$$P(i,m;n) = \left\lfloor \frac{2^{n+m-i-2}\Big((n-i+3)(2^m-1)-m\Big)}{(2^m-1)^2} \right\rceil,$$

where $\lfloor x \rceil$ is the integer nearest to x.

Proof. We have

$$\begin{aligned} P(i,m;n) &= \sum_{j=0}^{\infty} P(jm+i;n) - \sum_{j=1+\lfloor \frac{n-i}{m} \rfloor}^{\infty} P(jm+i;n) \\ &= \frac{2^{n+m-i-2}\Big((n-i+3)(2^m-1)-m\Big)}{(2^m-1)^2} + E, \end{aligned}$$

using Theorem 5.1, where

$$E = E(i,m;n) = \frac{1}{4}[\![m \mid n-i]\!] - U$$

and $[\![q]\!]$ is the Iverson bracket, i.e., 1 if q is true and 0 if q is false, with

$$U = U(i,m;n) = \sum_{j=1+\lfloor \frac{n-i}{m} \rfloor}^{\infty} (n-jm-i+3)2^{n-jm-i-2}.$$

Thus, we will be done upon demonstrating that $|E| < 1/2$ for all n, m and i.

To this end, we note that U is a discrete periodic function of n with period m since, by direct computation, it can be shown that

$$U(i,m;n) - U(i,m;n+m) = 0.$$

For fixed m, the set of all values assumed by U is the same for all $i = 1, 2, \dots, m$. In fact, $U(i,m;n) = U(i+\ell, m; n+\ell)$, so without loss of generality we take $i = 1$.

By summing the series, we find

$$\begin{aligned} U(1,m;n) &= \frac{2^{n-m\lfloor (n-1)/m \rfloor}}{8(2^m-1)^2} \\ &\quad \times \left((2^m - 1)\left(n - m \left\lfloor \frac{n-1}{m} \right\rfloor \right) - (2 + 2^m(m-2)) \right). \end{aligned}$$

Thus, for all integers $k \geq 0$ and $1 \leq j \leq m$, we have

$$U(1,m;mk+j) = -\frac{2^{j-3}}{(2^m-1)^2}\left(2^m(m-j-2) + (j+2)\right). \qquad (5.1)$$

As the right-hand side of (5.1) is independent of k, without loss of generality, set $k = 0$ and consider $|U(1,m;j)|$ as a real variable x over the closed interval $1 \leq j \leq m$. Using elementary calculus, one may conclude that the function $|U(1,m;x)|$ is maximized at $x = m$ and that this maximum value is less than $1/4$ when $m > 1$, whence $|E| < 1/2$. For the case $m = 1$, we have $U(1,1;n) = 1/4$ and so $|E| = |1/4 - 1/4| = 0 < 1/2$. ∎

Corollary 5.3. *The total number of parts in* $\mathscr{C}_n$ *is* $2^{n-2}(n+1)$.

Proof. Set $i = m = 1$ in Theorem 5.2. For a combinatorial proof, see [7, p. 120, Ex. 23] and its solution. ∎

We conclude with an exact equation for our sequence of interest; see Table 1.

Corollary 5.4.

$$OP(n) = \left\lfloor \frac{2^{n-1}(3n+4)}{9} \right\rceil = \frac{2^{n-1}(3n+4) - 2(-1)^n}{9}.$$

Proof. Setting $i = 1$ and $m = 2$ in Theorem 5.2 gives the first equality and computing $E(1,2;n)$ leads to the second. For the error term, one could also use a partial fraction decomposition directly. ∎

References

[1] L. Carlitz, Enumeration of compositions by rises, falls and levels, Math. Nachr. 77(3), 254–264, (1977).

[2] P. Chinn, G. Coyler, M. Flashman and E. Migliore, Cuisenaire rods go to college, PRIMUS 2, 118–130, (1992).

[3] S. Heubach and T. Mansour, Counting rises, levels, and drops in compositions, Integers 5, Article A11, (2005).

[4] S. Heubach and T. Mansour, Combinatorics of Compositions and Words, CRC Press, Boca Raton, FL (2010).

[5] T. Mansour and B. O. Sirhan, Counting ℓ-letter subwords in compositions, Discrete Math. Theor. Comput. Sci. 8, 285–298, (2006).

[6] N. J. Sloane, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`, 2016.

[7] R. P. Stanley, Enumerative Combinatorics, Vol. 1, second edition, Cambridge University Press (2011).

Appendix: Subword expressions to modulus six

Here, we provide subword pattern expressions for all of the $P(i, m; n)$ through $m = 6$. These were determined by the methods described in §4. Recall that often these expressions are not unique. In the cases where there were free variables, they were assigned values that minimized the number of terms. The following hold for all $n \in \mathbb{N}$.

$$P(1,2;n) = S(1;n) - S(11;n)$$
$$P(2,2;n) = S(11;n)$$

$$P(1,3;n) = S(1;n) - S(11;n) - S(1^3;n) + S(112;n) - 2S(122;n)$$
$$P(2,3;n) = S(11;n) - S(112;n) + 2S(122;n)$$
$$P(3,3;n) = S(1^3;n)$$

$$P(1,4;n) = S(1;n) - S(11;n) - S(1^3;n) + S(1122;n) - S(12^3;n) - S(1223;n) + 2S(1233;n)$$
$$P(2,4;n) = S(11;n) - S(1^4;n)$$
$$P(3,4;n) = S(1^3;n) - S(1122;n) + S(12^3;n) + S(1223;n) - 2S(1233;n)$$
$$P(4,4;n) = S(1^4;n)$$

$$P(1,5;n) = S(1;n) - S(11;n) - S(1^3;n) - S(1^5;n) + 2S(1^42;n) + 3S(1^322;n) - 2S(112^3;n) + S(12^4;n) - S(12^33;n)$$

$$
\begin{aligned}
&+ 2S(12233;n) + S(123^3;n) - 2S(12334;n) + 4S(12344;n)\\
P(2,5;n) &= S(11;n) - S(1^4;n) + S(1^42;n) + 2S(1^322;n) + S(12^4;n)\\
&- S(12233;n) + S(123^3;n) + S(12334;n) - 2S(12344;n)\\
P(3,5;n) &= S(1^3;n) - S(1^42;n) + S(1^322;n) + 2S(112^3;n) - S(12^4;n)\\
&+ S(12^33;n) - 2S(12233;n) - S(123^3;n) + 2S(12334;n)\\
&- 4S(12344;n)\\
P(4,5;n) &= S(1^4;n) - S(112^3;n) + S(12^4;n) + S(12233;n) - S(123^3;n)\\
&- S(12334;n) + 2S(12344;n)\\
P(5,5;n) &= S(1^5;n)
\end{aligned}
$$

$$
\begin{aligned}
P(1,6;n) &= S(1;n) - S(11;n) - S(1^3;n) - S(1^5;n) + S(1^6;n) + S(112^4;n)\\
&- S(12^5;n) - S(1223^3;n) + S(123^4;n) + S(123344;n)\\
&- S(1234^3;n) - S(123445;n) + 2S(123455;n)\\
P(2,6;n) &= S(11;n) - S(1^4;n) - S(1^6;n) + S(1^422;n) - 2S(112^4;n)\\
P(3,6;n) &= S(1^3;n) - S(1^6;n)\\
P(4,6;n) &= S(1^4;n) - S(1^422;n) + 2S(112^4;n)\\
P(5,6;n) &= S(1^5;n) - S(112^4;n) + S(12^5;n) + S(1223^3;n) - S(123^4;n)\\
&- S(123344;n) + S(1234^3;n) + S(123445;n) - 2S(123455;n)\\
P(6,6;n) &= S(1^6;n)
\end{aligned}
$$

Department of Mathematics, Saint Peter's University, Jersey City, NJ, 07306, USA
E-mail address: bhopkins@saintpeters.edu

Department of Mathematics, University of Tennessee, Knoxville, TN, 37919, USA
E-mail address: shattuck@math.utk.edu

Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA, 30458, USA
E-mail address: asills@georgiasouthern.edu

Mahidol University, International College, Highway 3310, Salaya, Phutthamonthon District, Nakhon Pathom 73170, Thailand
E-mail address: thotsaporn@gmail.com

Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA, 30458, USA
E-mail address: hwang@georgiasouthern.edu