# PARTS AND SUBWORD PATTERNS IN COMPOSITIONS

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

Abstract. We find relationships between subword patters and residue classes of parts in the set of integer compositions of a given weight.

## 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 2.8) in the compositions of $n$, and that it was neither obvious nor easy to prove that this was indeed the case. After recording the necessary definitions and notations, we present three proofs that $OP(n) = R(n)$: a bijective proof of a complementary result, a bijective proof of the original result, and a generating function proof.

From there, we move on to present relationships between number of parts congruent to $i \pmod{m}$ and the occurrence 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 $i \pmod{m}$ among all compositions of $n$.

This paper is accompanied by a Maple package available for free download at `home.dimacs.rutgers.edu/~asills/SillsMathemMaple.html` and `thotsaporn.com`. The program implements the bijections given in §3.

## 2. Definitions and notation

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

---

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

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

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 $\sigma$, denoted $|\sigma|$, is the sum of the parts of $\sigma$.

**Definition 2.3.** The *reverse* of a composition $\sigma = (\sigma_1, \ldots, \sigma_l)$ is the composition $v(\sigma) = (\sigma_l, \ldots, \sigma_1)$ also of the same weight.

**Definition 2.4.** 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 2.5.** Let $P(k;n)$ denote the number of parts in $\mathscr{C}_n$ equal to $k$. For example, $P(1;4) = 12$.

**Definition 2.6.** 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 defintions.

**Definition 2.7.** 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 [3, A059570] with different offsets.

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 $1 \leq i \leq j-1$. 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 2.8.** We call each instance of $r_i$ consecutive parts equal to $\sigma_i$ a *run* in the composition $\sigma$. Let $R(n)$ denote the total number of runs in all compositions of $n$; $R(n) = \sum j(\sigma)$ where the sum is taken over all compositions $\sigma = \langle \sigma_1^{r_1} \cdots \sigma_{j(\sigma)}^{r_{j(\sigma)}}\rangle \in \mathscr{C}_n$. Note that $R(4) = 14$; see Table 4.

**Definition 2.9.** A *rise*, *level*, *drop* in a composition $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_l)$ is any instance of $\sigma_i < \sigma_{i+1}$, $\sigma_i = \sigma_{i+1}$, $\sigma_i > \sigma_{i+1}$, respectively. Let $r(n)$, $\ell(n)$, $d(n)$ denote the number of rises, levels, and drops in $\mathscr{C}_n$, respectively. Continuing our example, $r(4) = 3$, $\ell(4) = 6$, and $d(4) = 3$.

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

**Definition 2.10.** The *reduced form* of a sequence of parts within a composition (i.e., a subcomposition) $\sigma = \sigma_r\sigma_{r+1}\ldots\sigma_{r+h-1}$ is given by $s_1 s_2 \ldots s_h$, where $s_i = j$ if $\sigma_i$ is the $j$th smallest part of subcomposition $\sigma$.

**Definition 2.11.** A composition $\sigma$ contains a *subword pattern* $\tau = \tau_1\tau_2\ldots\tau_k$ if the reduced form of any subsequence of $k$ consecutive parts of $\sigma$ equals $\tau$. Thus, rises, levels, and drops are subword pattern matches to 12, 11, and 21 respectively.

**Definition 2.12.** Let $S(\tau; n)$ denote the number of occurrences of the subword pattern $\tau$ in $\mathscr{C}_n$.

The following definitions regarding initial terms of compositions arise in §3.2.

**Definition 2.13.** Let $IO(n)$ denote the number of compositions of $n$ whose initial part is odd. and $IE(n)$ the number of compositions of $n$ whose initial part is even.

**Definition 2.14.** Let $I1(n)$ denote the number of compositions of $n$ whose initial part is 1 and $IB(n)$ the number of compositions of $n$ whose initial part is "big," i.e., greater than 1.

**Definition 2.15.** Let $IL(n)$ denote the number of compositions of $n$ whose first two parts are equal, i.e., an initial level.

**Definition 2.16.** Let $IS(n)$ denote the number of compositions of $n$ whose first two parts are a "step up," i.e., have the form $k, k+1$.

Notice that, while $IL(n)$ counts initial levels, $IS(n)$ counts only the initial rises with difference 1.

In §4 we will make use of integer partitions.

**Definition 2.17.** A *partition* of $n$ is a composition of $n$ that contains no drops.

**Definition 2.18.** If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ is a partition of $n$, then we define the *conjugate* $\lambda'$ of $\lambda$ to be the partition

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

Definition 2.18 is consistent with the usual definition of conjugation via transposition of Ferrers graphs.

**Definition 2.19.** 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: levels and runs versus parity of parts

### 3.1. Bijection between even parts and levels.

The most immediate resolution of our motivating question comes from considering what turns out to be the complementary situation.

A bijection between even parts and levels follows immediately from the following more general yet straightforward result.

**Theorem 3.1.** $P(m, m; n) = S(1^m; n)$.

*Proof.* We establish a bijection between all parts $km$ for $k \geq 1$ among all compositions of $n$ and all occurrences of the sub words $k^m$ among all compositions of $n$ (which are exactly the occurrences of the $1^m$ pattern). The map is simply $km \leftrightarrow k^m$, switching the single part $km$ and the $m$ adjacent $k$'s. ∎

Note that this is not a bijection on compositions, rather between parts divisible by $m$ (which may occur multiple times in a single composition) and $m$ adjacent equal parts (which may overlap in a single composition). Table 2 gives an example.

Table 2. $P(2, 2; 4) = S(1^2; 4)$ via the bijection from Theorem 3.1.

| $2k$ | $\underline{4}$ | $\underline{2}2$ | $2\underline{2}$ | $\underline{2}11$ | $1\underline{2}1$ | $11\underline{2}$ |
|---|---|---|---|---|---|---|
| $k, k$ | $\underline{22}$ | $\underline{11}2$ | $2\underline{11}$ | $\underline{11}11$ | $1\underline{11}1$ | $11\underline{11}$ |

**Corollary 3.2.** $EP(n) = \ell(n)$.

*Proof.* Set $m = 2$ in Theorem 3.1. ∎

**Theorem 3.3.** *$OP(n) = R(n)$.*

*Proof.* All parts of all compositions of $n$ may be partitioned in the following two ways: by parity or by the relation between a part and its successor—any two adjacent parts within a composition make a rise, level, or drop, and there are $|\mathscr{C}_n|$ final parts of compositions not counted yet. Therefore,

$$OP(n) + EP(n) = |\mathscr{C}_n| + r(n) + \ell(n) + d(n). \tag{3.1}$$

Now everything counted by the right-hand side of (3.1) except the levels constitutes the runs, (comparing Definitions 2.8 and 2.9, $\ell(n)$ is the sum of all $r_i - 1$ terms), i.e.,

$$R(n) = |\mathscr{C}_n| + r(n) + d(n).$$

Subtracting $EP(n)$ and $\ell(n)$, equal by Corollary 3.2, from their respective sides of (3.1) gives

$$OP(n) = |\mathscr{C}_n| + r(n) + d(n)$$

and thus $OP(n) = R(n)$. ∎

3.2. **Bijection between odd parts and runs.** Although our motivating question is settled, we found the method somewhat unsatisfying. Similar to the proof of Theorem 3.1, we want a bijection between the objects counted by $OP(n)$ and $R(n)$. While the following bijection is not as direct, it does establish the connection between the desired objects rather than their complements.

The bijection is recursive, thus Definitions 2.13–2.16 concerning the initial parts of compositions. Small values are given in Table 3.

Table 3. Small values of some functions based on initial parts of compositions.

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $IO(n)$ | 1 | 1 | 3 | 5 | 11 | 21 | 43 | 85 | 171 | 341 |
| $IE(n)$ | 0 | 1 | 1 | 3 | 5 | 11 | 21 | 43 | 85 | 171 |
| $IL(n)$ | 0 | 1 | 1 | 3 | 5 | 11 | 21 | 43 | 85 | 171 |
| $IS(n)$ | 0 | 0 | 1 | 1 | 3 | 5 | 11 | 21 | 43 | 85 |
| $I1(n)$ | 1 | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 |
| $IB(n)$ | 0 | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 |

The first four of these composition functions shown in Table 3 are counted by the Jacobsthal numbers [3, A001045] with varying offsets. Rather than connect each function to this common sequence, we prove two of the many relationships among them in the context of compositions.

But first we establish the well-known formula for $|\mathscr{C}_n|$ by a construction that addresses $I1(n)$ and $IB(n)$.

**Proposition 3.4.**

$$|\mathscr{C}_n| = \begin{cases} 2^{n-1} & \textit{if } n > 0, \\ 1 & \textit{if } n = 0. \end{cases}$$

*Proof.* The claim is clearly true for $n = 0, 1$. To construct $\mathscr{C}_{n+1}$, take two copies of $\mathscr{C}_n$. To the first, prefix each composition with a 1. To the second, increase the initial part of each composition by 1. The resulting $2|\mathscr{C}_n|$ compositions all have weight $n + 1$ and the obvious inverse map establishes a bijection. Keeping track of compositions with initial 1's versus initial "big" parts, we also have $I1(n) = IB(n)$ for $n \geq 2$. ∎

See [2, Theorem 1.3] for MacMahon's proof of Proposition 3.4 using binomial coefficients.

**Proposition 3.5.** *$IE(n) = IL(n)$.*

*Proof.* This is Corollary 3.2 restricted to the beginning of the compositions. ∎

**Proposition 3.6.** *$IO(n) - I1(n) = IS(n)$.*

*Proof.* The difference $IO(n) - I1(n)$ counts compositions of $n$ whose first part is an odd $k \geq 3$. Write $k = 2j + 1$ where $j \geq 1$. The correspondence $k, \ldots \longleftrightarrow j, j + 1, \ldots$ gives a bijection to $IS(n)$. ∎

We now give our second combinatorial proof of Theorem 3.3.

*Proof.* We proceed by induction. Certainly $OP(1) = R(1) = 1$ from the unique composition of one.

We establish recursive expressions for both $OP(n+1)$ and $R(n+1)$. Starting with odd parts, we claim that

$$OP(n + 1) = (OP(n) + |\mathscr{C}_n|) + (OP(n) + IE(n) - IO(n)). \tag{3.2}$$

Recall the two-part construction from Proposition 3.4. For the half of $\mathscr{C}_{n+1}$ with an initial 1, all odd parts of $\mathscr{C}_n$ remain odd and there is one new odd part per composition, the initial 1. This accounts for $(OP(n) + |\mathscr{C}_n|)$ in (3.2).

For the half of $\mathscr{C}_{n+1}$ with first part greater than 1, those first parts came from adding 1 to the first part of each of $\mathscr{C}_n$. All odd parts of $\mathscr{C}_n$

other than the first part stay odd. For the first parts, compositions in $\mathscr{C}_n$ that began with an odd part now start with an even part and vice versa. Altogether, this accounts for $(OP(n) + IE(n) - IO(n))$ in (3.2), which is now fully explained.

For runs, we claim that

$$R(n+1) = (R(n) + IB(n)) + (R(n) + IL(n) - IS(n)). \tag{3.3}$$

In the half of $\mathscr{C}_{n+1}$ with an initial 1, the runs of $\mathscr{C}_n$ are maintained and there is one more for each $\mathscr{C}_n$ that does not begin with a 1. This gives $(R(n) + IB(n))$ in (3.3).

For the half of $\mathscr{C}_{n+1}$ with first part greater than 1, all runs of $\mathscr{C}_n$ remain except some of those involving the first part.

- When an element of $\mathscr{C}_n$ has one part or its first two parts satisfy $\sigma_2 - \sigma_1 \notin \{0, 1\}$, these runs are among those already counted by $R(n)$.
- When an element of $\mathscr{C}_n$ begins with a run of at least two parts, i.e., is counted in by $IL(n)$, then the transition from $k, k, \ldots$ to $k+1, k, \ldots$ adds one run.
- When an element of $\mathscr{C}_n$ is counted by $IS(n)$, then the transition from $k, k+1, \ldots$ to $k+1, k+1, \ldots$ subtracts one run.

This explains the remainder of (3.3), the term $(R(n) + IL(n) - IS(n))$.

Using the induction hypothesis $OP(n) = R(n)$ in (3.2) and (3.3) leaves us with the following claim to show that $OP(n+1) = R(n+1)$.

$$|\mathscr{C}_n| + IE(n) - IO(n) = IB(n) + IL(n) - IS(n). \tag{3.4}$$

But it follows from Proposition 3.4 that $|\mathscr{C}_n| - IB(n) = I1(n)$, so that (3.4) is equivalent to

$$IE(n) + IS(n) = IL(n) + IO(n) - I1(n)$$

which is exactly the sum of Propositions 3.5 and 3.6. ■

The bijection for $n = 4$ is given explicitly in Table 4. The reader may download the Maple package mentioned in the introduction to see details of the bijection through $n = 13$.

3.3. **Generating functions.** We outline how our motivating theorem can also be proved using generating functions. First we need the following proposition.

**Proposition 3.7.** *For positive $n$, $r(n) = d(n)$.*

*Proof.* A palindromic composition satisfies $\sigma = v(\sigma)$ (its reverse) and therefore has an equal number of rises and drops. For a nonpalindromic composition $\rho$, the number of rises in $\rho$ equals the number of drops in

Table 4. $OP(4) = R(4)$ via the bijection from the §3.2 proof of Theorem 3.1.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| odd part | $\underline{3}1$ | $3\underline{1}$ | $2\underline{1}1$ | $21\underline{1}$ | $\underline{1}3$ | $1\underline{3}$ | $\underline{1}21$ |
| run | $\underline{2}11$ | $3\underline{1}$ | $\underline{22}$ | $2\underline{11}$ | $\underline{1}3$ | $1\underline{2}1$ | $\underline{1}21$ |
| odd part | $12\underline{1}$ | $\underline{1}12$ | $1\underline{1}2$ | $\underline{1}111$ | $1\underline{1}11$ | $11\underline{1}1$ | $111\underline{1}$ |
| run | $12\underline{1}$ | $\underline{3}1$ | $\underline{11}2$ | $\underline{4}$ | $1\underline{3}$ | $11\underline{2}$ | $\underline{1111}$ |

$v(\rho)$, and the number of drops in $\rho$ equals the number of rises in $v(\rho)$. The set $\mathscr{C}_n$ is the union of palindromic compositions and nonpalinromic pairs $\{\rho, v(\rho)\}$. Summing the numbers of rises and drops over all of these these singletons and pairs gives the result. ∎

We now give a generating function proof of Theorem 3.3.

*Proof.* We show that both $OP(n)$ and $R(n)$ are given by the generating function

$$\frac{x(1-x)}{(1+x)(2x-1)^2}.$$

The statement for $OP(n)$ is [2, Exercise 3.13]. The outline of an argument is to show that

$$\sum_{k=1}^{\infty}\left(\frac{xy+x^2}{1-x^2}\right)^k = \frac{\frac{xy+x^2}{1-x^2}}{1-\frac{xy+x^2}{1-x^2}} = \frac{xy+x^2}{1-xy-2x^2}$$

is the generating function where the coefficient of $y^m x^n$ counts compositions of $n$ containing exactly $m$ odd parts. As the partial derivative with respect to $y$ gives a factor of $m$, show then that

$$\sum_{n\geq 1} OP(n)x^n = \frac{\partial}{\partial y}\left(\frac{xy+x^2}{1-xy-2x^2}\right)\Bigg|_{y=1} = \frac{(1-x)x}{(1+x)(2x-1)^2}.$$

For runs, recall from the (first) proof of Theorem 3.3 that

$$R(n) = |\mathscr{C}_n| + r(n) + d(n) = |\mathscr{C}_n| + 2r(n),$$

the second equality from Proposition 3.7. For $|\mathscr{C}_n|$, we use the generating function $x/(1-2x)$. (Usually, one uses $(1-x)/(1-2x)$ which produces the composition $\varnothing$ of weight 0, but we do not consider the empty composition to have any runs.) For $r(n)$, [2, Example 4.6] derives the generating function $x^3/((1+x)(1-2x)^2)$. Thus

$$\sum_{n\geq 1} R(n)x^n = \frac{x}{1-2x} + \frac{2x^3}{(1+x)(1-2x)^2} = \frac{(1-x)x}{(1+x)(2x-1)^2}. \qquad \blacksquare$$

## 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 $\tau = \tau_1\tau_2\cdots\tau_l$ be a subword pattern of length $l$ where*

$$1 = \tau_1 \le \tau_2 \le \cdots \le \tau_l \le l.$$

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

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

*Proof.* Observe that

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

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

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

while the middle factor of (4.1) is

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

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

Thus (4.1) is the generating function for the number $n$ of ordered triples $(\gamma, \lambda, \rho)$ 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 $\sigma$ be a composition of $n$ containing the subword $\mu = \mu_1\mu_2 \dots \mu_l$ that matches the subword pattern $\tau$. Thus $\sigma$ is of the form

$$\sigma = (\gamma_1, \gamma_2, \dots, \gamma_j, \mu_1, \mu_2, \dots, \mu_l, \rho_1, \rho_2, \dots, \rho_k),$$

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

Observe that $\sigma$ can be mapped bijectively to

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

noting that the parts of $\mu'$ must all lie in $T$. This completes the equality of (4.1) and $\sum_{n\geq 0} S(\tau; n)x^n$. ■

See Table 5 for a detailed example of the correspondence.

TABLE 5. 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 in for this $\tau$). By Theorem 3.1, this shows that the claim holds for $i = m$.

A basic fact of compositions is that $P(k; n) = P(k + 1; n + 1)$ by the bijection $k \leftrightarrow k + 1$ in $\mathscr{C}_n$ and $\mathscr{C}_{n+1}$, respectively. 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.** The trivial subword pattern 1 gives $S(1;n) = |\mathscr{C}_n|$. From Theorem 3.1, we know $P(2,2;n) = S(11;n)$. Therefore $P(1,2;n) = S(1;n) - S(11;n)$.

From the generating functions of Theorems 4.1 and 4.2, this expression for $P(1,2;n)$ implies

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

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

*Proof.* This is a direct consequence from 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 $\tau$ of length no more than $m$.*

*Proof.* From (4.1)

$$\begin{aligned}\sum_{n\geq 0} S(\tau;n)x^n &= \frac{(1-x)^2}{(1-2x)^2}\cdot\frac{x^l}{1-x^l}\prod_{j=1}^{l-1}\frac{x^{(l-j)(\tau_{j+1}-\tau_j)}}{(1-x^{l-j})^{\tau_{j+1}-\tau_j}}\\ &= \frac{(1-x)^2}{(1-2x)^2}\prod_{k=1}^{\tau_l}\frac{x^{t_k}}{1-x^{t_k}},\end{aligned}$$

where $t_j$ is the cardinality of the set $\{\tau_j \mid 1\leq j\leq l, \tau_j\geq k\}$. Regrouping the conclusion of Theorem 4.2, given

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

a linear representation of $P(i,m;n)$ by $S(\tau;n)$ where the length of $\tau$ being no more than $m$ is equivalent to a linear representation of $x^i/(1-x^m)$ with terms of the form

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

such that $t_k < t_1 \leq m$ for any $k\neq 1$.

We proceed by induction on $m$; the initial cases are given in Example 4.3. For any given $i$ and $m$, if they share a common factor, the conclusion follows from Theorem 4.4. Now let $m > i \geq 2$ be relatively prime to $i$ and $u = \max\{u \mid ui < m\}$ . Then $(u+1)i > m$ as $i$ and $m$ are relatively prime. We now have

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

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

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

By induction hypothesis, each $x^{ji}/(1-x^{ui})$ term and $x^{(u+1)i-m}/(1-x^{ui})$ can be represented, with $t_k \leq ui < m$, as linear combinations of the form (4.3). Together with (4.4), this establishes the claim for $P(i,m;n)$ when $i \geq 2$. ■

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

**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

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

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 previous example, but instead we demonstrate a more general approach.

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

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

After clearing denominators, we equate coefficients of $x, x^2, \ldots, x^{11}$ to obtain a linear system of 11 equations in the seven $c_\tau$ 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

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

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.

**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 [1, Question 3]. The $k = 1$ case is established by combinatorial reasoning and recurrence relations. The equality $P(k;n) = P(k+1;n+1)$ mentioned in the proof of Theorem 4.2 completes their argument, as it follows that $P(k;n) = P(1;n-k+1)$. For another approach, see [4, p. 120, Ex. 24] and its solution. ∎

**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.*

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

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 $i = 1, 2, \ldots, m$. In fact, $U(i,m;n) = U(i+j,m;n+j)$, so without loss of generality we take $i = 1$.

By summing the series, we find

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

From here, it can be shown that, for all nonnegative integers $k$,

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

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 over the closed interval $1 \leq j \leq m$. Using elementary calculus, we may conclude that the function $|U(1, m; n)|$ is maximized at $n = m$ and this maximum value is less than 1/4 when $m > 1$, thus $|E| < 1/2$. For the case $m = 1$, $U(1, 1; n) = 1/2$ and so $|E| = |1/4 - 1/2| = 1/4 < 1/2$. ∎

**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 [4, 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. ∎

## Acknowledgments

The work of the 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.

## References

[1] P. Z. Chinn, G. Coyler, M. Flashman, E. Migliore, Cuisenaire rods go to college, *PRIMUS* **2** (1992) 118–130, http://dx.doi.org/10.1080/10511979208965656.

[2] S. Heubach, T. Mansour, *Combinatorics of Compositions and Words*, CRC Press, Boca Raton, FL, 2010.

[3] *The On-Line Encyclopedia of Integer Sequences*, http://oeis.org, 2015.

[4] R. P. Stanley, *Enumerative Combinatorics*, vol. 1, 2nd edition, Cambridge University Press, 2011.

## Appendix: Subword expressions to modulus six

Here we provide subword pattern expressions for all $P(i, m; n)$ through $m = 6$. These were determined by the methods described in Example 4.7 or Theorem 4.4. 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 $n \in \mathbb{N}$.

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

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

$$\begin{aligned} P(1,4;n) &= S(1;n) - S(11;n) - S(1^3;n) + S(1122;n) - S(12^3;n) \\ &\quad - S(1223;n) + 2S(1233;n) \end{aligned}$$
$$P(2,4;n) = S(11;n) - S(1^4;n)$$
$$P(3,4;n) = S(1^3;n) - S(1122;n) + S(12^3;n) + S(1223;n) - 2S(1233;n)$$
$$P(4,4;n) = S(1^4;n)$$

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

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

$$
\begin{aligned}
P(3,6;n) &= S(1^3;n) - S(1^6;n)\\
P(4,6;n) &= S(1^4;n) - S(1^4 22;n) + 2S(112^4;n)\\
P(5,6;n) &= S(1^5;n) - S(112^4;n) + S(12^5;n) + S(1223^3;n) - S(123^4;n)\\
&\quad - S(123344;n) + S(1234^3;n) + S(123445;n) - 2S(123455;n)\\
P(6,6;n) &= S(1^6;n)
\end{aligned}
$$

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

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