Enumeration of 3-Letter Patterns in Compositions

Silvia Heubach
Department of Mathematics
California State University Los Angeles

joint work with
Toufik Mansour
Department of Mathematics
University of Haifa, Haifa, Israel
Enumerating Compositions

• Alladi & Hoggatt - $A = \{1, 2\}$ in connection with Fibonacci Sequence [1]

• Carlitz & various co-authors - # rises, levels, falls in $[n] = \{1, 2, \ldots, n\}$ as generalization of permutations [5],[6],[7],[8],[9]

• Carlitz & Vaughan - # compositions according to specification, rises, falls and maxima [9]

• Carlitz, Scoville, & Vaughan - enumeration of pairs of sequences according to rises, levels and falls [8].

• Rawlings - weak rises and falls in connection with restricted words [15]

• Chinn, Grimaldi & Heubach- # rises, levels, falls in specific sets $A$ [10, 11, 12, 13, 14]
Basic Notions

• $\sigma = \sigma_1 \sigma_2 \ldots \sigma_m =$ composition of $n \in \mathbb{N}$ with $m$ parts where $\sum_{i=1}^{m} \sigma_i = n$

• rise = a summand followed by a larger summand

• level = a summand followed by itself

• fall or drop = a summand followed by a smaller summand

Think of these as 2-letter patterns

• level $\leftrightarrow$ 11

• rise $\leftrightarrow$ 12

• drop $\leftrightarrow$ 21
• Look at **pairs** of levels, rises and drops ↔ 3-letter patterns

• $\tau = \tau_1 \tau_2 \tau_3$; level + rise ↔ 112

• **reversal map** $r(\tau) = r(\tau_1 \tau_2 \tau_3) = \tau_3 \tau_2 \tau_1$; \{\tau, r(\tau)\} = symmetry class of $\tau$

• patterns in the same symmetry class occur equally often

• Only patterns to consider because of symmetry:
  
  level+level ↔ 111  
  rise+rise ↔ 123  
  level+rise ↔ 112  
  rise+drop=peak ↔ 121+132+231  
  level+drop ↔ 221  
  drop+rise=valley ↔ 212+213+312
Notation

• $A = \{a_1, a_2, a_3, \ldots, a_d\}$ or $A = \{a_1, a_2, a_3, \ldots\}$, where $a_1 < a_2 < \ldots$ are positive integers

• $C_\tau(n, r) \ (C_\tau(j; n, r)) = \#$ of compositions of $n$ with parts in $A$ ($j$ parts in $A$) containing pattern $\tau$ exactly $r$ times.

• $C_\tau(\sigma_1 \ldots \sigma_\ell|n, r) \ (C_\tau(\sigma_1 \ldots \sigma_\ell|j; n, r)) = \text{those that start with } \sigma_1, \ldots, \sigma_\ell.$

• Generating functions

- $C_\tau(x, y) = \sum_{n,r \geq 0} C_\tau(n, r)x^n y^r$
- $C_\tau(x, y, z) = \sum_{n,r,j \geq 0} C_\tau(j; n, r)x^n y^r z^j$
- $C_\tau(\sigma_1 \ldots \sigma_\ell|x, y) = \sum_{n,r \geq 0} C_\tau(\sigma_1 \ldots \sigma_\ell|n, r)x^n y^r$
- $C_\tau(\sigma_1 \ldots \sigma_\ell|x, y, z) = \sum_{n,r,j \geq 0} C_\tau(\sigma_1 \ldots \sigma_\ell|j; n, r)x^n y^r z^j$

• $C_\tau(x, y, z) = 1 + \sum_{a \in A} C_\tau(a|x, y, z) \quad (*)$
The pattern 111 (level+level)

**Theorem:** Let $A$ be any ordered (finite or infinite) set of positive integers. Then

$$C_{111}(x, y, z) = \frac{1}{1 - \sum_{a \in A} \frac{x^a z (1 + (1-y)x^a z)}{1 + x^a z (1 + x^a z) (1-y)}}.$$  

**Proof:** Split the compositions that start with $a$ into those that start with $ab$ and $aa$, and then split up the latter into those that start with $aab$ and $aaa$ and set up recursion. 

Thus, gf for # of compositions in $\mathbb{N}$ that avoid 111 is given by

$$C_{111}(x, 0, 1) = \frac{1}{1 - \sum_{i \geq 1} \frac{x^i (1+x^i)}{1 + x^i (1+x^i)}},$$

and values of the corresponding sequence are 1, 1, 2, 3, 7, 13, 24, 46, 89, 170, 324, 618, 1183, 2260, 4318, 8249, 15765, 30123, 57556, 109973, 210137...
The patterns 112 (level+rise) and 221 (level+drop)

**Theorem:** Let $A$ be any ordered subset of $\mathbb{N}$. Then

$$C_{112}(x, y, z) = \frac{1}{1 - \sum_{j=1}^{d} \left( x^{a_j} z \prod_{i=1}^{j-1} (1 - (1 - y)x^{2a_i} z^2) \right)} ,$$

and

$$C_{221}(x, y, z) = \frac{1}{1 - \sum_{j=1}^{d} \left( x^{a_j} z \prod_{i=j+1}^{d} (1 - (1 - y)x^{2a_i} z^2) \right)} .$$

The sequence for the # of compositions in $\mathbb{N}$ which **avoid 112** is given by 1, 1, 2, 4, 7, 13, 24, 43, 78, 142, 256, 463, 838, 1513, 2735, 4944, 8931, 16139, 29164, 52693, 95213, ..., and the one for the # of compositions in $\mathbb{N}$ which **avoid 221** is given by 1, 1, 2, 4, 8, 15, 30, 58, 113, 220, 429, 835, 1627, 3169, 6172, 12023, 23419, 45616, 88853, 173073, 337118,...
Proof: Arguments similar to those in proof for 111 give

\[ C_{112}(a|x, y, z) = \frac{x^{2a}z^2}{1-x^{2a}z^2} + \frac{x^{2a}z^2}{1-x^{2a}z^2} \sum_{b \in A, b < a} C_{112}(b|x, y, x) \]
\[ + \frac{x^{2a}z^2y}{1-x^{2a}z^2} \sum_{b \in A, b > a} C_{112}(b|x, y, z) + \frac{x^az}{1+x^az}C_{112}(x, y, z). \]

Assume \( A \) is finite. Let \( x_0 = C_{112}(x, y, z) \), \( x_i = C_{112}(a_i|x, y, z) \), \( \alpha_i = \frac{x^{2a_i}z^2}{1-x^{2a_i}z^2} \), and \( \beta_i = \frac{x^a_i}{1+x^a_i} \), then with Eq. (*) we get a system of \( d + 1 \) equations

\[ x_i - \alpha_i \sum_{j < i} x_j - \alpha_i y \sum_{j > i} x_j - \beta_i x_0 = \alpha_i \quad \text{for} \quad i = 1, \ldots, d, \]
\[ x_0 - \sum_{i=1}^{d} x_i = 1. \]

Now use Cramer’s rule and messy algebra to compute the determinants. Take limits if \( A \) is infinite. Similarly for 221. \( \blacksquare \)
The pattern 123 (rise+rise)

**Theorem:** Let $A$ be any ordered subset of $\mathbb{N}$, with $|A| = d$. Then

$$C_{123}(x, y, z) = \frac{1}{1 - t^1(A) - \sum_{p=3}^{d} \sum_{j=0}^{p-3} \binom{p-3}{j} t^{p+j}(A)(y-1)^{p-2}},$$

where $t^p(A) = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq d} x^{\sum_{i=1}^{p} a_{i,j}}$.

For $A = \mathbb{N}$, $t^p(\mathbb{N}) = x^{\binom{p+1}{2}} z^p \prod_{j=1}^{p} (1 - x^j)^{-1}$, and the sequence for the # of compositions in $\mathbb{N}$ which avoid 123 is given by 1, 1, 2, 4, 8, 16, 31, 61, 119, 232, 453, 883, 1721, 3354, 6536, 12735, 24813, 48344, 94189, 183506, 357518, \ldots
**Proof:** (Outline) Define

- \( A_k = \{a_{k+1}, a_{k+2}, \ldots, a_d\} = A \setminus \{a_1, \ldots, a_k\} \) (the index of \( A \) indicates the largest element excluded).

- \( D^{A_k}(x, y, z) = \text{gf for } \# \text{ of compositions } \sigma \text{ of } n \text{ with } m \text{ parts in } A_k \text{ such that for } a \notin A_k, a \sigma \text{ contains the pattern } 123 \text{ exactly } r \text{ times.} \)

Two possibilities: \( \sigma \) does not contain \( a_1 \), or \( \sigma = \bar{\sigma}a_1\sigma_{k+1} \ldots \sigma_m \), where \( \bar{\sigma} \) is a composition with parts from \( A_1 \):

\[
C_{123}^A(x, y, z) = C_{123}^{A_1}(x, y, z) + C_{123}^{A_1}(x, y, z)C_{123}^A(a_1|x, y, z).
\]

If \( \sigma \) starts with \( a_1 \), then two cases: either exactly one occurrence of \( a_1 \), or \( a_1 \) occurs at least twice in \( \sigma \), i.e., \( \sigma = a_1\bar{\sigma}a_1\sigma_{k+1} \ldots \sigma_m \), where \( \bar{\sigma} \) is a (possibly empty) composition with parts from \( A_1 \):

\[
C_{123}^A(a_1|x, y, z) = x^{a_1}zD^{A_1}(x, y, z) + x^{a_1}zD^{A_1}(x, y, z)C_{123}^A(a_1|x, y, z).
\]
\[ C_{123}^A(x, y, z) = \frac{C_{123}^A(x, y, z)}{1 - x^{a_1}zD^A_1(x, y, z)} \] (**)

To obtain \( D^A_1(x, y, z) \) look at occurrences of \( a_2 \).

- \( \sigma \) contains no \( a_2 \); or \( \sigma = \bar{\sigma}^1 a_2 \bar{\sigma}^2 a_2 \bar{\sigma}^3 \ldots a_2 \bar{\sigma}^\ell + 2 \) with \( \ell \geq 0 \), where \( \bar{\sigma}^j \) is a (possibly empty) composition with parts in \( A_2 \) for \( j = 1, \ldots, \ell + 2 \).

- Four cases \( (\bar{\sigma}^j = \emptyset \) or \( \neq \emptyset, j = 1, 2 \))

\[ D^A_1 = \frac{(1 - x^{a_2}z(1 - y))D^A_2 + x^{a_2}z(1 - y)}{1 - x^{a_2}zD^A_2} \]

Using induction and lots of messy algebra gives

\[ D^A = \frac{1 + \sum_{p=2}^{d} \sum_{j=0}^{p-2} \binom{p-2}{j} t^{p+j}(A)(y - 1)^{p-1}}{1 - t^1(A) - \sum_{p=3}^{d} \sum_{j=0}^{p-3} \binom{p-3}{j} t^{p+j}(A)(y - 1)^{p-2}}. \]

Similar arguments for (***) finish the proof.
The patterns \{121, 132, 231\} (peak = rise+drop) and \{212, 213, 312\} (valley = drop+rise)

For any \(B \subseteq A\) with \(|A| = d\), and \(s \geq 1\)

- \(P^s(B) = \{(i_1, \ldots, i_s) \mid a_{ij} \in B, j = 1, \ldots, s, \text{ and } i_{2\ell-1} < i_{2\ell} \leq i_{2\ell+1} \text{ for } 1 \leq \ell \leq \lfloor s/2 \rfloor\}\)

- \(Q^s(B) = \{(i_1, \ldots, i_s) \mid a_{ij} \in B, j = 1, \ldots, s, \text{ and } i_{2\ell-1} \leq i_{2\ell} < i_{2\ell+1} \text{ for } 1 \leq \ell \leq \lfloor s/2 \rfloor\}\)

- \(M^s(B) = z^s \sum_{(i_1, \ldots, i_s) \in P^s(B)} \prod_{j=1}^s x^{a_{ij}}\)

- \(N^s(B) = z^s \sum_{(i_1, \ldots, i_s) \in Q^s(B)} \prod_{j=1}^s x^{a_{ij}}\)
**Theorem:** Let $A = \{a_1, \ldots, a_d\}$, $P^s(A)$, $Q^s(A)$, $M^s(A)$, and $N^s(A)$ be defined as on previous slide. Then

$$C^A_{\text{peak}}(x, y, z) = \frac{1 + \sum_{j \geq 1} M^{2j}(A)(1 - y)^j}{1 + \sum_{j \geq 1} M^{2j}(A)(1 - y)^j - \sum_{j \geq 0} M^{2j+1}(A)(1 - y)^j},$$

and

$$C^A_{\text{valley}}(x, y, z) = \frac{1 + \sum_{j \geq 1} M^{2j}(A)(1 - y)^j}{1 + \sum_{j \geq 1} M^{2j}(A)(1 - y)^j - \sum_{j \geq 0} N^{2j+1}(A)(1 - y)^j}.$$

The sequence for the number of compositions in $\mathbb{N}$ which avoid “peak” is given by

$1, 1, 2, 4, 7, 13, 22, 38, 64, 107, 177, 293, 481, 789, 1291, 2110, 3445, 5621, 9167, 14947, 24366, \ldots$ and the one for the number of compositions in $\mathbb{N}$ which avoid “valley” is given by

$1, 1, 2, 4, 8, 15, 28, 52, 96, 177, 326, 600, 1104, 2032, 3740, 6884, 12672, 23327, 42942, 79052, 145528, \ldots$
Proof: Concentrate on where the largest part occurs. Let $\bar{A}_k = \{a_1, \ldots, a_k\}$. Four different cases:

- $\sigma$ does not contain $a_d$
- $\sigma = a_d \sigma'$, $\sigma'$ possibly empty
- $\sigma = \bar{\sigma} a_d$, where $\bar{\sigma}$ is a non-empty composition with parts in $\bar{A}_{d-1}$
- $\sigma = \bar{\sigma} a_d \sigma'$, where $\sigma'$ is a non-empty composition with parts in $A$
  - $\sigma'$ starts with $a_d$
  - $\sigma'$ does not start with $a_d$

Combining all cases and using induction gives
Lemma: For \( A = \{a_1, \ldots, a_d\} \), and \( b_i = x^{a_i} z \),

\[
C^A_{\text{peak}}(x, y, z) = \frac{1}{1 - b_d - G_d}.
\]

where

\[
G_d = \frac{1}{b_d(1 - y) + \frac{1}{b_{d-1}(1 - y) + \frac{1}{\cdots + \frac{1}{b_2(1 - y) + \frac{1}{b_1}}}}}.
\]
Next we prove that
\[
G_d = \frac{\sum_{j \geq 0} M^{2j+1}(A)(1 - y)^j}{1 + \sum_{j \geq 1} M^{2j}(A)(1 - y)^j},
\]
using induction on \(d\) and the recursions below for odd and even \(s\), obtained by conditioning on whether last element is \(a_d\).

- \(s\) odd \(\Rightarrow\) last and second last element can be equal to \(a_d\)

\[
M^{2s+1}(A) = b_d M^{2s}(A) + M^{2s+1}(\bar{A}_{d-1})
\]

- \(s\) even \(\Rightarrow\) second last element can be at most \(a_{d-1}\)

\[
M^{2s}(A) = b_d M^{2s-1}(\bar{A}_{d-1}) + M^{2s}(\bar{A}_{d-1}).
\]

Proof for valley follows similarly, where recursions involve \(M^s(A_k)\) and \(N^s(A_k)\).
\textbf{Asymptotic Behavior}

\textbf{Theorem:} The asymptotic behavior for $\tau$-avoiding compositions with parts in $\mathbb{N}$ is given by

\begin{align*}
C_{111}(n,0) &= 0.499301 \cdot 1.91076^n + O((10/7)^n) \\
C_{112}(n,0) &= 0.692005 \cdot 1.80688^n + O((10/7)^n) \\
C_{221}(n,0) &= 0.545362 \cdot 1.94785^n + O((10/7)^n) \\
C_{123}(n,0) &= 0.576096 \cdot 1.94823^n + O((10/7)^n) \\
C_{\text{peak}}(n,0) &= 1.394560 \cdot 1.62975^n + O((10/7)^n) \\
C_{\text{valley}}(n,0) &= 0.728207 \cdot 1.84092^n + O((10/7)^n).
\end{align*}
Application to Words

- \([k] = \{1, 2, \ldots, k\}\) = (totally ordered) alphabet on \(k\) letters
- word = element of \([k]^n\)
- word \(\sigma\) contains a pattern \(\tau\) if \(\sigma\) contains a subsequence (order) isomorphic to \(\tau\)
- complement \(c(\tau)\) is the pattern obtained when replacing \(\tau_i\) by \(k + 1 − \tau_i\)
- \(\{\tau, r(\tau), c(\tau), c(r(\tau))\}\) symmetry class of \(\tau\)
- \(C^{[k]}_{\tau}(1, y, z) = \text{gf for } \# \text{ of words of length } m \text{ on the alphabet } [k] \text{ with } r \text{ occurrences of } \tau.\)

Obtain known results (see [2],[3]) for patterns 111, 112 (221), and 123 , and new results for peak (valley).
Preprint available from my web site at
sheubac@calstatela.edu
References


