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 $[n] = \{1, 2, ..., n\}$ 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 τ
- 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 $a_1 < a_2 <$ .... 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 τ exactly r times.
- $C_\tau(\sigma_1 \ldots \sigma_\ell | n, r)$ ($C_\tau(\sigma_1 \ldots \sigma_\ell | j; n, r)$) = those that start with $\sigma_1, \ldots, \sigma_\ell$.
- Generating functions
  - $C_\tau(x, y) = \sum_{n,r \geq 0} C_\tau(n, r) x^n y^r$
  - $C_\tau(x, y, z) = \sum_{n,r,j \geq 0} C_\tau(j; n, r) x^n y^r z^j$
  - $C_\tau(\sigma_1 \ldots \sigma_\ell | x, y) = \sum_{n,r \geq 0} C_\tau(\sigma_1 \ldots \sigma_\ell | n, r) x^n y^r$
  - $C_\tau(\sigma_1 \ldots \sigma_\ell | x, y, z) = \sum_{n,r,j \geq 0} C_\tau(\sigma_1 \ldots \sigma_\ell | j; n, r) x^n y^r z^j$
- $C_\tau(x, y, z) = 1 + \sum_{a \in A} C_\tau(a | x, y, z)$ **(*)**

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

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

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 $\beta_i = \frac{x^{a_i}z}{1+x^{a_i}z}$, 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. ■

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 $t^p(A) = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le d} z^p \prod_{j=1}^{p} x^{a_{i_j}}$.

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**, . . .

Proof: (Outline) Define

- $A_k = \{a_{k+1}, a_{k+2}, \ldots, a_d\} = A \backslash \{a_1, \ldots, a_k\}$ (the index of A indicates the largest element excluded).
- $D^{A_k}(x, y, z) =$ gf for # of compositions σ of n with m parts in A_k such that for $a \notin A_k$, $a\,\sigma$ contains the pattern 123 exactly r times.

Two possibilities: σ 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 σ starts with a_1, then two cases: either exactly one occurrence of a_1, or a_1 occurs at least twice in σ, 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:

$$C_{123}^{A}(a_1|x, y, z) = x^{a_1} z D^{A_1}(x, y, z) + x^{a_1} z D^{A_1}(x, y, z) C_{123}^{A}(a_1|x, y, z).$$

$$\Rightarrow \quad C_{123}^{A}(x,y,z) = \frac{C_{123}^{A_1}(x,y,z)}{1 - x^{a_1} z D^{A_1}(x,y,z)} \quad \textbf{(**)}$$

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

- σ 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 $\neq \emptyset, j = 1, 2$)

$$\Rightarrow \quad D^{A_1} = \frac{(1 - x^{a_2} z(1-y))D^{A_2} + x^{a_2} z(1-y)}{1 - x^{a_2} z D^{A_2}}.$$

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 $s \geq 1$

- $P^s(B) = \{(i_1, \ldots, i_s) | \, a_{i_j} \in B, j = 1, \ldots, s,$ and $i_{2\ell-1} < i_{2\ell} \leq i_{2\ell+1}$ for $1 \leq \ell \leq \lfloor s/2 \rfloor\}$
- $Q^s(B) = \{(i_1, \ldots, i_s) | \, a_{i_j} \in B, j = 1, \ldots, s,$ and $i_{2\ell-1} \leq i_{2\ell} < i_{2\ell+1}$ for $1 \leq \ell \leq \lfloor s/2 \rfloor\}$
- $M^s(B) = z^s \sum_{(i_1,\ldots,i_s) \in P^s(B)} \prod_{j=1}^{s} x^{a_{i_j}}$
- $N^s(B) = z^s \sum_{(i_1,\ldots,i_s) \in Q^s(B)} \prod_{j=1}^{s} x^{a_{i_j}}$

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

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

and

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

The sequence for the # 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**, ... and the one for the # 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**,...

Proof: Concentrate on where the largest part occurs. Let $\bar{A}_k = \{a_1, \ldots, a_k\}$. Four different cases:

- σ does not contain a_d
- $\sigma = a_d\sigma'$, σ' 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 σ' is a non-empty composition with parts in A
  - σ' starts with a_d
  - σ' does not start with a_d

Combining all cases and using induction gives

Lemma: For $A = \{a_1, \ldots, a_d\}$, and $b_i = x^{a_i} z$,

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

where

$$G_d = \cfrac{1}{b_d(1-y) + \cfrac{1}{b_{d-1} + \cfrac{1}{b_{d-1}(1-y) + \cfrac{1}{\ddots + \cfrac{1}{b_2(1-y) + \cfrac{1}{b_1}}}}}}.$$

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 $N^s(A_k)$.

■

Asymptotic Behavior

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

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

Application to Words

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

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

[1] K. Alladi and V. E. Hoggatt, Jr. Compositions with ones and twos. *Fibonacci Quart.*, 13(3):233–239, 1975.

[2] A. Burnstein and T. Mansour. Words restricted by 3-letter generalized multipermutation patterns. *Annals of Combinatorics*, 7(1):1–14, 2003.

[3] A. Burnstein and T. Mansour. Counting occurrences of some subword patterns. *Discrete Mathematics and Theoretical Computer Science*, 6(1):1–12, 2003.

[4] L. Carlitz. Enumeration of sequences by rises and falls: a refinement of the Simon Newcomb problem. *Duke Math. J.*, 39:267–280, 1972.

[5] L. Carlitz. Enumeration of up-down sequences. *Discrete Math.*, 4:273–286, 1973.

[6] L. Carlitz. Enumeration of compositions by rises, falls and levels. *Math. Nachr.*, 77:361–371, 1977.

[7] L. Carlitz and R. Scoville. Up-down sequences. *Duke Math. J.*, 39:583–598, 1972.

[8] L. Carlitz, R. Scoville, and T. Vaughan. Enumeration of pairs of sequences by rises, falls and levels. *Manuscripta Math.*, 19(3):211–243, 1976.

[9] L. Carlitz and T. Vaughan. Enumeration of sequences of given specification according to rises, falls and maxima. *Discrete Math.*, 8:147–167, 1974.

[10] P. Chinn, R. Grimaldi, and S. Heubach. Rises, levels, drops and "+" signs in compositions: extensions of a paper by Alladi and Hoggatt. *The Fibonacci Quarterly*, 41(3):229–239, 2003.

[11] P. Chinn and S. Heubach. (1,k)-compositions. *Congressus Numerantium*, 164:183–194, 2003.

[12] P. Chinn and S. Heubach. Compositions of n with no occurrence of k. *Congressus Numerantium*, 164:33–51, 2003.

[13] R. P. Grimaldi. Compositions with odd summands. *Congressus*

Numerantium, 142:113–127, 2000.

[14] R. P. Grimaldi. Compositions without the summand 1. *Congressus Numerantium*, 152: 33–43, 2001.

[15] D. Rawlings. Restricted words by adjacencies. *Discrete Mathematics*, 220:183–200, 2000.