On The Probability that Certain Compositions Have The Same Number Of Parts

Miklós Bóna
Department of Mathematics
University of Florida
Gainesville FL 32611-8105
USA

Arnold Knopfmacher
School of Mathematics
University of the Witwatersrand
Johannesburg
South Africa

November 6, 2008

Abstract

We compute the asymptotic probability that two randomly selected compositions of n into parts equal to a or b have the same number of parts. In addition we provide bijections in the case of parts of sizes 1 and 2 with weighted lattice paths and central Whitney numbers of fence posets. Explicit algebraic generating functions and asymptotic probabilities are also computed in the case of pairs of compositions of n into parts at least d, for any fixed natural number d.

1 Introduction

A *composition* of the positive integer n is a sequence $(a_1, a_2, \cdots, a_k)$ of positive integers so that $\sum_{i=1}^{k} a_i = n$. The a_i are called the *parts* of the composition. It is well-known [1] that the number of compositions of n into k parts is $\binom{n-1}{k-1}$. If we choose an ordered pair (X, Y) of compositions of the integer n at random, the probability that X and Y both have k parts is $\frac{\binom{n-1}{k-1}^2}{4^{n-1}}$. If we express this probability p_n in terms of asymptotics as $n \to \infty$, we get that

$$p_n = \sum_{k=0}^{n} \frac{\binom{n-1}{k-1}^2}{4^{n-1}} = \frac{\binom{2n-2}{n-1}}{4^{n-1}} \sim \frac{1}{\sqrt{n\pi}}. \tag{1}$$

In this paper, we will first study the probability $p_{a,b}(n)$ that two compositions of n that only have part sizes a and b have the same number of parts. We begin by studying the special case of $a = 1$ and $b = 2$. We will obtain the generating function for the number of such pairs of compositions with the same number of parts and we point out connections with another class of restricted compositions, and, surprisingly, central Whitney numbers of the fence poset. Then we obtain a precise asymptotic expression for $p_{1,2}(n)$.

We then turn to the general case, and find the generating function for the number of pairs of compositions of n into parts equal to a or b that have the same number of parts. Again we obtain a precise asymptotic expression for this probability.

Thereafter we use the technique of diagonalisation of bivariate power series to obtain explicit generating functions and asymptotic estimates for the number of pairs of compositions with the same of parts, where all parts are at least d, for any fixed natural number d.

In all cases we find that the probability is asymptotic to $C/\sqrt{n}$ for some constant C which depends on the sizes of the permitted parts. We point out that this probability is significantly higher than what straightforward estimates would suggest. Indeed, if q_i is the probability that a randomly selected composition of n (with a fixed set of allowed parts) has exactly i parts, then $\sum_{i=1}^{n} q_i = 1$, while the probability that two randomly selected such compositions have the same number of parts is $p = \sum_{i=1}^{n} q_i^2$. Applying the Cauchy-Schwarz inequality, we only get that

$$p = \sum_{i=1}^{n} q_i^2 \geq \frac{1}{n} \left(\sum_{i=1}^{n} q_i \right)^2 = \frac{1}{n},$$

which is significantly less than the $C/\sqrt{n}$ asymptotic value that we are going to prove.

2 Compositions with parts equal to 1 or 2

2.1 At most two or at least two?

It is well-known [1] that the number of compositions of n into parts equal to 1 or 2 is the Fibonacci number F_{n+1}, with $F_0 = 0$, $F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. Indeed, using induction, such a composition either has 1 for its first part, and then can be continued in F_{n-1} ways, or has 2 for its first part, and then can be continued in F_{n-2} ways. It is worth pointing out that the number of these compositions also equals $\sum_{k=0}^{\lceil n/2 \rceil} \binom{n-k}{k}$, since if such a composition has k parts equal to 2, then it has $n - 2k$ parts equal to 1. When arranging these $n - k$ parts in a line, there are $\binom{n-k}{k}$ ways to choose the positions of the parts equal to 2.

Interestingly, the number of compositions of n into parts that are *at least* two is also a Fibonacci number, namely F_{n-1}. Indeed, for $n = 1$, there are no such compositions, and for $n = 2$, there is one such composition. For larger values of n, we can use induction again. Such a composition either has a 2 for its first part, and then it can be continued in F_{n-3} ways, or has a first part larger than 2, in which case subtracting 1 of that first part, we get one of F_{n-2} compositions of n into parts at least two.

Even more interestingly, there is a one-to-one correspondence between these classes of compositions even if we specify the number of parts. Indeed, the number of compositions of n into $n-k$ parts that are at most 2 is $\binom{n-k}{k}$ since such compositions must consist of k parts equal to 2 and $n-2k$ parts equal to 1. The number of compositions of $n+2$ into $k+1$ parts that are at least two is also $\binom{n+2-(k+1)-1}{k} = \binom{n-k}{k}$ since these compositions are in bijection with the compositions of $n-k+1$ into $k+1$ parts (just add 1 to each part). Therefore, if we can compute the probability that two randomly selected compositions have the same number of parts for one of these two classes of compositions, the result will also apply for the other class.

We note that this line of research fits into an incipient interest in similar questions [2], [7].

2.2 A bijection with lattice paths

Let $\mathbf{C_n}$ be the set of ordered pairs (X, Y) of compositions of n into parts equal to 1 or 2 so that X and Y have the same number of parts. Then it follows from what we said in the first paragraph of the previous subsection that $|\mathbf{C_n}| = \sum_{k=0}^{\lceil n/2 \rceil} \binom{n-k}{k}^2$. It is always interesting to see a set counting pairs (like $\mathbf{C_n}$) being equinumerous to a set counting single objects that are seemingly unrelated. In this subsection, we will provide such an example.

A *weighted lattice path* is a lattice path whose edges are associated a weight. The weight of a lattice path is the sum of the weights of its edges.

Let $\mathbf{L_n}$ be the set of lattice paths of weight n that start in $(0, 0)$ and end on the horizontal axis whose steps are of the following four kinds.

- An $(1, 0)$-step (horizontal step) with weight 1.
- An $(1, 0)$-step (horizontal step) with weight 2.
- A $(1, 1)$-step ("up step") with weight 2.
- A $(1, -1)$-step ("down step ") with weight 1.

Proposition 1 *There is a natural bijection* $f : \mathbf{C_n} \to \mathbf{L_n}$.

Proof: Let $(X,Y) \in \mathbf{C_n}$, and let $X = (x_1, x_2, \cdots, x_k)$, and let $Y = (y_1, y_2, \cdots, y_k)$. Then $f(X,Y)$ will have k steps. For each i, compare x_i and y_i, and define the ith step s_i of $f(X,Y)$ as follows.

- If $x_i = y_i = 1$, then let s_i be a horizontal step of weight 1.
- If $x_i = y_i = 2$, then let s_i be a horizontal step of weight 2.
- If $x_i = 2$ and $y_i = 1$, then let s_i be an "up" step, necessarily with weight 2.
- If $x_i = 1$ and $y_i = 2$, then let s_i be a "down" step, necessarily with weight 1.

Note that for each i, the weight of s_i is equal to x_i, so the total weight of $f(X,Y)$ is indeed $\sum_{i=1}^k x_i = n$. Furthermore, $f(X,Y)$ ends on the horizontal axis, since the fact that X and Y have the same number of parts implies that they have the same number of parts equal to 1, and the same number of parts equal to 2. Therefore, f maps into $\mathbf{L_n}$.

In order to show that f is a bijection, it suffices to show that it has an inverse. Let $L \in \mathbf{L_n}$, and let $s_1, s_2, \cdots, s_k$ be the steps of L. Then, using the four rules above, we can recreate the unique pair (X,Y) for which $f(X,Y) = L$ could hold. What is left to see is that this (X,Y) is indeed in $\mathbf{C_n}$. Clearly, X and Y have both k parts since they were recovered from L. On the other hand, the sum of the parts of X is n, since this sum is equal to the weight of L. Finally, the sum of the parts of Y is also n. Indeed, L ends on the horizontal axis, so it has as many up steps as down steps, so the sum of the parts of Y is equal to the sum of the parts of X. $\diamond$

See Figure 1 for an illustration of this bijection.

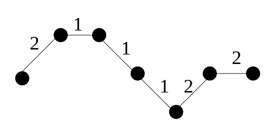

Figure 1: The path $f(X,Y)$ for $X = (2,1,1,1,2,2)$ and $Y = (1,1,2,2,1,2)$.

2.3 Generating functions

In what follows, we will enumerate the elements of $\mathbf{L_n}$ instead of $\mathbf{C_n}$. Let d_n denote the number of elements of $\mathbf{L_n}$, set $d_0 = 1$, and let $D(x) = \sum_{n \geq 0} d_n x^n$.

Furthermore, let b_n be the number of lattice paths in $\mathbf{L_n}$ consisting of at least two steps which do not touch the horizontal axis, except for their starting and ending point. Note that $b_0 = b_1 = b_2 = 0$. Also note that paths counted by b_n can be above or below the axis, so $b_3 = 2$. Figure 2 shows the four paths of L_5 that are enumerated by b_5.

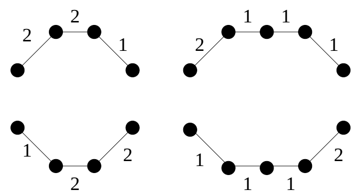

Figure 2: The four paths enumerated by b_5.

Set $B(x) = \sum_{n \geq 0} b_n x^n$. The relation between the sequences b_n and d_n is given in the following lemma.

Lemma 1 *Set $d_n = 0$ if $n < 0$. Then for all positive integers n, the equality*

$$d_n = d_{n-1} + d_{n-2} + \sum_{i=0}^{n} b_i d_{n-i} \tag{2}$$

holds.

Proof: There are three possibilities for the way a path belonging to $\mathbf{L_n}$ can start. It can start with a horizontal step of weight 1, and then finish in one of d_{n-1} ways; it can start with a horizontal step of weight 2, and then finish in one of d_{n-2} ways; or it can start with a non-horizontal step, touch the horizontal axis first at the end of a subpath of weight i (this is possible in b_i ways), and then finish in one of d_{n-i} ways. $\diamond$

Corollary 1 *The generating functions $B(x)$ and $D(x)$ are connected by the equation*

$$D(x) = \frac{1}{1 - x - x^2 - B(x)}. \tag{3}$$

Proof: Multiply both sides of (2) by x^n and sum over all $n \geq 1$ to get

$$D(x) - 1 = xD(x) + x^2 D(x) + B(x)D(x).$$

Expressing $D(x)$ from this equation proves our claim. $\diamond$

So once we find an explicit formula for $B(x)$, we will have an explicit formula for $D(x)$. In order to find an explicit formula for $B(x)$, note that half of the paths enumerated by b_n are over the horizontal axis, and the other half are below that axis. Let us count those that are above the axis. These paths start with an up step and end with a down step. In between, they consist of a subpath of weight $n-3$ that never goes below the $y=1$ line. Therefore, if c_n denotes the number of paths of weight n that start and end on the horizontal axis and never go below that axis, then $b_n = 2c_{n-3}$ for all n. (Here we set $c_n = 0$ if $n < 0$, and $c_0 = 1$.) If $C(x) = \sum_{n\geq 0} c_n x^n$, then this leads to the equality

$$B(x) = 2x^3 C(x). \tag{4}$$

However, there is another relation between $B(x)$ and $C(x)$. Note that a path counted by c_n can either start with a horizontal step of weight 1, and then finish in one of c_{n-1} ways, or start with a horizontal step of weight 2, and then finish in one of c_{n-2} ways, or start with an up step, touch the horizontal axis first at the end of a subpath of weight i, which can happen in $b_i/2$ ways, and then finish in one of c_{n-i} ways. This leads to the recurrence relation

$$c_n = c_{n-1} + c_{n-2} + \frac{1}{2}\sum_{i=0}^{n} b_i c_{n-i}.$$

Or, in terms of generating functions,

$$C(x) - 1 = xC(x) + x^2C(x) + \frac{1}{2}B(x)C(x). \tag{5}$$

Comparing this with (4) and rearranging yields

$$x^3C^2(x) - (1 - x - x^2)C(x) + 1 = 0, \tag{6}$$

which is a quadratic equation for $C(x)$. Solving this equation, one sees easily by verifying the value of the obtained power series at $x = 0$ that the *negative* square root provides the correct solution, that is,

$$C(x) = \frac{1 - x - x^2 - \sqrt{1 - 2x - x^2 - 2x^3 + x^4}}{2x^3}.$$

Therefore, by (4),

$$B(x) = 2x^3C(x) = 1 - x - x^2 - \sqrt{1 - 2x - x^2 - 2x^3 + x^4}.$$

Finally, by (3),

$$D(x) = \frac{1}{\sqrt{1-2x-x^2-2x^3+x^4}} = \frac{1}{\sqrt{(1-3x+x^2)(1+x+x^2)}}. \qquad (7)$$

2.4 Asymptotics

We are going to evaluate the probability $p_{1,2}(n)$ that two randomly selected compositions of n with parts equal to one or two have the same number of parts. Recall from the Introduction that the number of such compositions is the Fibonacci number F_{n+1}. Therefore, $p_{1,2}(n) = \frac{d_n}{F_{n+1}^2}$. It is well-known [1] that the Fibonacci numbers are given by the explicit formula

$$F_n = \frac{1}{\sqrt{5}}\left(\alpha^n - \beta^n\right),$$

where $\alpha = (\sqrt{5}+1)/2$ and $\beta = (\sqrt{5}-1)/2$. Therefore, the number of pairs of compositions of n with parts of size 1 or 2 is

$$F_{n+1}^2 = \frac{1}{5}\left(\alpha^{2n+2} - \beta^{2n+2} - 2\right) \sim \frac{(3+\sqrt{5})^{n+1}}{5 \cdot 2^{n+1}}. \qquad (8)$$

Now we need to determine the asymptotics of d_n. We are going to use the singularity analysis method of Flajolet and Odlyzko to find these asymptotics. This relevant theorem can be found in many sources on analytic combinatorics, such as [3] or [5]. (Equivalently one could use Darboux's theorem, as found for example in [8].)

Essentially the method requires that we expand the relevant function in the neighbourhood of its dominant singularities. From this, singularity analysis allows us to transfer the asymptotic expressions for the function to that of its coefficients.

A routine computation shows that of the four singularities of $D(x)$ in (7), the closest one to the origin is $r = \frac{3-\sqrt{5}}{2}$. minimal modulus is $r = \frac{3-\sqrt{5}}{2}$. An expansion of $D(x)$ around $x = r$, yields

$$\frac{2}{\sqrt[4]{5}\left(-1+\sqrt{5}\right)^2}\frac{1}{\sqrt{1-\frac{x}{r}}}.$$

By singularity analysis it follows that the coeffcients d_n of $D(x)$ satisfy as $n \to \infty$,

$$d_n \sim \frac{2}{\sqrt[4]{5}\left(-1+\sqrt{5}\right)^2}\frac{r^{-n}}{\sqrt{\pi n}}. \qquad (9)$$

Crucially, $r^{-1} = \frac{3+\sqrt{5}}{2}$. Therefore, (9) implies

$$d_n \sim \frac{2}{\sqrt[4]{5}\left(-1+\sqrt{5}\right)^2} \frac{1}{\sqrt{\pi n}} \cdot \left(\frac{3+\sqrt{5}}{2}\right)^n. \tag{10}$$

Comparing (8) and (10), we find that the probability we are looking for is

$$p_{1,2}(n) = \frac{d_n}{F_{n+1}^2} \sim \frac{5^{3/4}}{2} \frac{1}{\sqrt{\pi n}}. \tag{11}$$

So just as for unrestricted compositions in the Introduction, we get a probability that is equal to $C/\sqrt{n}$ for some constant C. Once again, the event that two such compositions of n have the same number of pairs is likely in the logarithmic sense.

2.5 A connection with central Whitney number of the fence poset

The *fence poset* A_n is a $2n$-element poset with whose vertex set consists of vertices $\{x_1, x_2, \cdots, x_n\}$ and $\{y_1, y_2, \cdots y_n\}$, and whose set of relations consists of $x_i \le y_i$ for all i, and $x_i \le y_{i-1}$ for $i \ge 2$. The fence poset A_5 is shown in Figure 3.

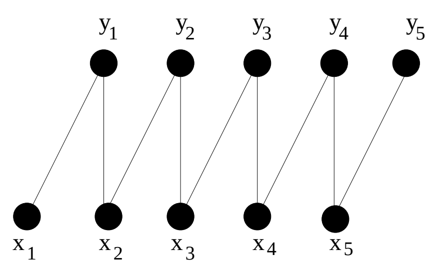

Figure 3: The fence poset A_5.

In a partially ordered set P, the subposet I of P is called an *ideal* if $y \in I$ and $x \le y$ imply $x \in I$. If, in an n-element poset, $i_0, i_1, \cdots, i_n$ denote the number of ideals that have $0, 1, \cdots, n$ elements, then the numbers i_j are called the *Whitney numbers* of the poset. For the fence poset A_n, it is known [4] that the chain of inequalities

$$i_0 \le i_1 \cdots \le i_n \ge i_{n+1} \ge \cdots \ge i_{2n}$$

holds. So the numbers i_n are not simply at the middle of the sequence, they are also maximal. The numbers i_n are called the *central Whitney numbers* of A_n. This provided motivation for the authors of [4] to study these numbers further. They showed that if the number of n-element ideals of A_n is w_n, with $w_0 = 1$, then $\sum_{n \geq 0} w_n x^n = D(x)$, where $D(x)$ is our $D(x)$ defined earlier in this section. Therefore, $d_n = w_n$ for all n. This remarkable fact calls for a combinatorial proof, especially knowing that the way $\sum_{n \geq 0} w_n x^n$ was found in [4] was *not* based on a combinatorial recurrence; in that paper, that generating function was obtained as the diagonal series of a power series in two variables.

We will now exhibit a bijection g from to set $\mathbf{I_n}$ of n-element ideals of A_n and the set $\mathbf{L_n}$ of lattice paths defined earlier in this section. The bijection is not obvious, but it is a true bijection in that it requires no recursive argument.

Consider the two-element subposets $\{x_i, y_i\}$ of A_n. If I is an ideal of A_n, and $S = \{x_i, y_i\}$, and we know the *size* of $I \cap S$, then we know the *set* $I \cap S$ itself. Indeed, if $|I \cap S| = 1$, then $I \cap S = \{x_i\}$, since I is an ideal, and if $|I \cap S| \neq 1$, then the statement is obvious.

For an ideal I of A_n, we define the vector $\mathbf{z(I)} = (z_1(I), z_2(I), \cdots, z_n(I))$, where $z_i = |I \cap \{x_i, y_i\}|$. See Figure 4 for an example.

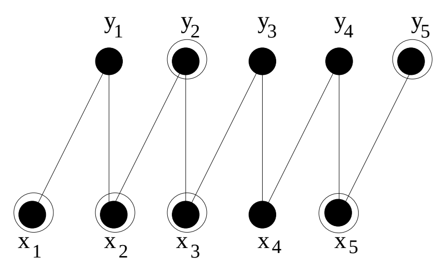

Figure 4: If I is the ideal of the encircled elements, then $\mathbf{z(I)} = (1, 2, 1, 0, 2)$.

The first step in achieving the goal of this subsection is the following observation. A similar technique has been used in [4] to provide exact formulae for the Whitney numbers of fence posets.

Proposition 2 *Let $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ be a vector of length n, with coordinates equal to 0, 1, or 2. Then there is an ideal $I \in \mathbf{I_n}$ so that $\mathbf{z(I)} = \mathbf{v}$ if and only if*

1. $\sum_{i=1}^{n} v_i = n$, *and*

2. if $v_i = 2$, then $v_{i+1} \neq 0$.

Proof: Let us first assume that there exists an $I \in \mathbf{I_n}$ so that $\mathbf{z(I)} = \mathbf{v}$. Then $|I \cap \{x_i, y_i\}| = v_i$, so the sum of the v_i must be n as this sum is equal to the size of I. Furthermore, if $v_i = 2$, then $y_i \in I$, so, as I is an ideal, $x_{i+1} \in I$, and therefore, $v_{i+1} \neq 0$. This proves that the conditions are necessary.

Now let $\mathbf{v}$ be a vector satisfying the conditions. Define I_v as follows.

1. If $v_i = 0$, then $x_i \notin I$, and $y_i \notin I$,

2. if $v_i = 1$, then $x_i \in I$, and $y_i \notin I$, and

3. if $v_i = 2$, then $x_i \in I$, and $y_i \in I$.

Then $I \in A_n$. Indeed, $|I| = \sum_{i=1}^{n} v_i = n$. Furthermore, I is an ideal since on the one hand, the above construction implies that if $y_i \in I$, then $x_i \in I$, and, on the other hand, since $v_i = 2$ implies $v_{i+1} \neq 0$, if $y_i \in I$, then x_{i+1} in I. $\diamond$

So there is a natural bijection between the set $\mathbf{I_n}$ and the set of vectors of length n that have length n, coordinates 0, 1, or 2, and in which a 2 is never followed by a 0. In what follows, we will identify the elements of $\mathbf{I_n}$ with their vectors $\mathbf{z(I)}$. Now we are ready to define the bijection $g : \mathbf{I_n} \rightarrow \mathbf{L_n}$.

Let $I \in \mathbf{I_n}$. Consider $\mathbf{z(I)}$. Say k of the coordinates of $\mathbf{z(I)}$ are equal to 2, and, therefore, $n-2k$ are equal to 1 (this also implies that k are equal to 0). Then $g(I)$ will be a lattice path consisting of $n-k$ steps, k of which will have weight 2, and, therefore, $n-2k$ of which will have weight 1. The relative order of these steps will be given by the relative order of the coordinates of $\mathbf{z(I)}$ that are equal to 2 or 1. For instance, if among the $n-k$ coordinates of $\mathbf{z(I)}$ that are equal to 2 or 1, the first, sixth, and eighth are equal to 2, then it will be the first, sixth, and eighth steps of $g(I)$ that have weight 2.

There remains the question of which steps of $g(I)$ will be horizontal, which will be up steps, and which will be down steps. This information will be gained from the positions of the coordinates of $\mathbf{z(I)}$ that are equal to 0. This part of the proof will remind some readers to the proof of the identity $\binom{n-k}{k} = \sum_{i=0}^{k} \binom{n-2k}{i}\binom{k}{k-i} = \sum_{i=0}^{k} \binom{n-2k}{i}\binom{k}{i}$. Recall that in $\mathbf{z(I)}$, no coordinate equal to 2 is immediately followed by a coordinate equal to 0. Now consider all coordinates of $\mathbf{z(I)}$ that are equal to 0 or 1. There are k coordinates equal to 0, and $n-2k$ equal to 1. These $n-k$ coordinates can be in $\binom{n-k}{k}$ different orders, each defined by the k-element subset $Z \subseteq$

$\{1, 2, \cdots, n-k\}$ of the positions of the 0s in this string of length $n-k$. Let $Z_1 = Z \cap \{1, 2, \cdots, n-2k\}$ and let $Z_2 = Z \cap \{n-2k+1, \cdots, n-k\}$. If $r \in Z_1$, then let the rth weight-1 step of $g(I)$ be a down step; otherwise, let it be a horizontal step. Furthermore, if $j + (n-2k) \in Z_2$, then let the jth weight-2 step of $g(I)$ be horizontal; otherwise, let it be an up step. This completes the definition of $g(I)$.

Theorem 1 *The map g described above is a bijection from $\mathbf{I_n}$ onto $\mathbf{L_n}$.*

Proof: First, g indeed maps into $\mathbf{L_n}$. The only thing that needs explanation is that $g(I)$ indeed always ends on the horizontal axis, that is, $g(I)$ has as many up steps as down steps. Let i be the number of down steps of $g(I)$; then $i = |Z \cap \{1, 2, \cdots, n-2k\}|$. The number of up steps is the number of weight-2 steps that are not horizontal. The number of horizontal weight-2 steps is $|Z_2| = |Z \cap \{n-2k+1, \cdots, n-k\}| = |Z| - i = k - i$. Therefore, the number of up steps (necessarily of weight 2) is $k - (k-i) = i$ as well.

In order to show that g is a bijection, we show that it has an inverse. Let $L \in \mathbf{L_n}$. We will recover the unique ideal $I \in \mathbf{I_n}$ for which $g(I) = L$. Clearly, it suffices to recover $\mathbf{z(I)}$. Let us assume that L has i up steps, then L has i down steps. Furthermore, let us say that L has k steps of weight 2, and $n-2k$ steps of weight 1.

Then if $g(I) = L$, then $\mathbf{z(I)}$ must have k coordinates equal to 2, $n-2k$ coordinates equal to 1, and k coordinates equal to 0. Furthermore, the relative order of the coordinates of $\mathbf{z(I)}$ that are equal to 2 or 1 has to be the same as the relative order of the steps of L that are of weight 2 or 1. (This is possible in exactly one way.) Finally, because of the construction of $g(I)$, the relative order of the coordinates of $\mathbf{z(I)}$ that are equal to 1 or 0 must be as follows. In the string of length $n-k$ of these coordinates of $\mathbf{z(I)}$, the set Z of positions of the 0s must be equal to $Z_1 \cup Z_2$, where $Z_1 \subseteq \{1, 2, \cdots, n-2k\}$ is the set of positions of the down steps among the $n-2k$ weight-1 steps, and $Z_2 \subseteq \{n-2k+1, \cdots, n-k\}$ is the set of positions of the form $(n-2k)+j$, where the jth weight-2 step of L is a horizontal step. This is again possible in exactly one way. Finally, knowing the positions of the 0s and 1s in the string of all 0s and 1s, and the positions of the 1s and 2s in the string of all 1s and 2s completely determines $\mathbf{z(I)}$, since in $\mathbf{z(I)}$, no 2 is ever directly followed by a 0. $\diamond$

3 Compositions with all parts equal to a and b

Let $a < b$ be two positive integers, and let $f_{a,b}(n)$ denote the number of compositions of n into parts equal to a or b. We can assume without loss of generality that a and b are relatively prime to each other, since $f_{ka,kb}(n) = f_{a,b}(n/k)$ (just divide each part by k). Then it is straightforward to prove the recurrence relation $f_{a,b}(n) = f_{a,b}(n-a) + f_{a,b}(n-b)$ for $n \geq 1$, with $f_{a,b}(0) = 1$. Let $F_{a,b}(x)$ denote the ordinary generating function of the sequence $f_{a,b}(n)$; then this yields

$$F_{a,b}(x) = \frac{1}{1 - x^a - x^b}. \tag{12}$$

Fortunately, the denominator of $F_{a,b}(x)$ is not difficult to handle. This is the content of the next lemma.

Lemma 2 *For any two relatively prime positive integers $a < b$, the polynomial $p(x) = 1 - x^a - x^b$ has a unique root $\rho = \rho_{a,b}$ of smallest modulus.*

Proof: We claim that $p(x)$ always has a unique positive real root ρ, and *that* ρ is the unique root of smallest modulus. First, on the set of positive real numbers, the function $q(x) = x^a + x^b$ is strictly monotone increasing, so it cannot equal 1 twice. Second, q is a polynomial, so it is continous. Third, $q(0) = 0$ and $q(1) = 2$, so there exist a unique positive real number ρ (which is between 0 and 1) so that $q(\rho) = 1$, and so $p(\rho) = 0$.

Now let z be any other root of $p(x)$; note that this means that z is not a positive real number. Then $z^a + z^b = 1$, therefore $|z^a + z^b| = 1$, so

$$|z|^a + |z|^b = q(|z|) \geq 1. \tag{13}$$

If this inequality is strict, then by the previous paragraph, $|z| > \rho$, and we are done. The only way that the triangle-inequality of (13) could be not strict is by both z^a and z^b being positive real numbers (do not forget that their sum is 1!). That would mean that the z^{ka} and z^{lb} are both positive real numbers, for all integers k and l. However, since a and b are relatively prime, we can choose k and l so that $ka + lb = 1$, implying that $z^{ka+lb} = z$ is a positive real number, which is a contradiction. $\diamond$

Now by expanding $F_{a,b}(x)$ around the dominant singularity $z = \rho$ we find the following.

$$F_{a,b}(x) \sim \frac{1}{a\rho^{a-1} + b\rho^{b-1}} \cdot \frac{1}{(\rho - x)}.$$

Then by applying singularity analysis [3] we obtain

Corollary 2 *Let a and b be positive integers, and assume without loss of generality that they are relatively prime to each other. Then with ρ as defined above,*

$$f_{a,b}(n) \sim \frac{\rho^{-n-1}}{a\rho^{a-1} + b\rho^{b-1}}.$$

Now we return to pairs of compositions with the same number of parts. Let $d_{a,b}(n)$ denote the number of pairs of (X, Y) of compositions of n into parts equal to a or b so that X and Y have the same number of parts. Set $d_{a,b}(0) = 1$. Again, we can assume that a and b are relatively prime to each other.

Let $D_{a,b}(x) = \sum_{n\geq 0} d_{a,b}(n)x^n$ be the ordinary generating function of the sequence of the numbers $d_{a,b}(n)$. Then an argument analogous to that seen in the previous section for the special case of $a = 1$ and $b = 2$ yields the equality

$$D_{a,b}(x) = \frac{1}{\sqrt{(1 - x^b + x^a)^2 - 4x^{a+b}}}. \tag{14}$$

We are interested in determining the ratio $p_{a,b}(n) = \frac{d_{a,b}(n)}{f_{a,b}(n)^2}$; in other words, the probability that a randomly selected ordered pair of compositions of n into parts a or b have the same number of parts.

Denote by $d_{a,b}(x)$ the polynomial under the square roots sign in the denominator of $D_{a,b}(x)$; that is, $D_{a,b}(x) = \frac{1}{\sqrt{d_{a,b}(x)}}$.

In order to discuss the factorization of $d_{a,b}(x)$, we need the following simple but useful observation.

Proposition 3 *Let $h(x) = \sum_{i=0}^{k} a_i x^i$ be a polynomial, and let $h^*(x) = \sum_{i=0}^{k} a_i x^{2i}$. Then the complex number z^2 is a root of $h(x)$ if and only if z and $-z$ are both roots of $h^*(x)$.*

In particular, for any positive real number z, it holds that $h^*(x)$ has $2m$ roots of modulus z if and only if $h(x)$ has m roots of modulus z^2.

Therefore, when looking for the root(s) of smallest modulus of $d_{a,b}(x)$, we may look for the roots of $d^*(x)$ first. The latter are easier to find since

$$\begin{aligned} d^*(x) &= (1 - x^{2a} - x^{2b})^2 - 4x^{2(a+b)} \\ &= (1 - x^{2a} - x^{2b} - 2x^{a+b})(1 - x^{2a} - x^{2b} + 2x^{a+b}) \\ &= (1 - x^a - x^b)(1 + x^a + x^b)(1 - x^a + x^b)(1 + x^a - x^b). \end{aligned}$$

It now follows from the same argument that we used to prove Lemma 2 that $d^*(x)$ always has two roots of smallest modulus, namely ρ and $-\rho$, where ρ is the unique root of smallest modulus of $1 - x^a - x^b$. Therefore, Proposition 3 implies the following.

Theorem 2 *The polynomial $d_{a,b}(x)$ has a unique root of smallest modulus, and that root is ρ^2.*

So the numbers $d_{a,b}(n)$ and $f_{a,b}(n)^2$ have similar growth rates in the sense that $\lim_{n\to\infty} (d_{a,b}(n))^{1/n} = \lim_{n\to\infty} \left(f_{a,b}^2(n)\right)^{1/n} = \rho^{-2}$.

A more precise computation using singularity analysis, shows that as $z \to \rho^2$,

$$D_{a,b}(x) \sim \frac{1}{\sqrt{4(a+b)\rho^{2a+2b-2} + 2\left(-\rho^{2a} - \rho^{2b} + 1\right)\left(a\rho^{2a-2} + b\rho^{2b-2}\right)}\rho} \frac{1}{\sqrt{1 - \frac{x}{\rho^2}}}$$

from which we deduce that

$$d_{a,b}(n) \sim \frac{\rho^{-2n-1}}{\sqrt{4(a+b)\rho^{2a+2b-2} + 2\left(-\rho^{2a} - \rho^{2b} + 1\right)\left(a\rho^{2a-2} + b\rho^{2b-2}\right)}\sqrt{\pi n}}.$$

We conclude that the probability that two compositions of n into parts a and b have the same number of parts is

$$p_{a,b}(n) = \frac{d_{a,b}(n)}{f_{a,b}^2(n)} \sim \frac{C}{\sqrt{\pi n}},$$

where

$$C = \frac{\rho\left(a\rho^{a-1} + b\rho^{b-1}\right)^2}{\sqrt{4(a+b)\rho^{2a+2b-2} + 2\left(-\rho^{2a} - \rho^{2b} + 1\right)\left(a\rho^{2a-2} + b\rho^{2b-2}\right)}}.$$

In the special case $a = 1$ and $b = 2$, $\rho = \frac{1}{2}\left(-1 + \sqrt{5}\right)$ and we recover the asymptotic estimate for $p_{1,2}(n)$.

4 Diagonals of bivariate generating functions

Let $c(n_1, n_2, k)$ denote the number of pairs of compositions both with k parts in $\mathbb{N}$, where the first part of the pair is a compositions of n_1 and second part is a composition of n_2. Then we have the trivariate generating function

$$\frac{1}{1 - y\frac{s}{1-s}\frac{t}{1-t}} = \sum_{n_1, n_2, k \geq 0} c\left(n_1, n_2,\ k\right) y^k s^{n_1} t^{n_2}.$$

For our purposes the value of k is not important so we can concentrate on the bivariate generating function for $c(n_1, n_2)$, the number of pairs of compositions of n_1 and n_2, respectively, with the same number of parts, given by

$$F(s,t) := \sum_{n_1,n_2\geq 0} c\left(n_1,n_2\right) s^{n_1} t^{n_2} = \frac{1}{1-\frac{st}{(1-s)(1-t)}}.$$

More generally if we restrict the parts in the compositions to lie in some fixed subset A of $\mathbb{N}$ then we must consider

$$F_A(s,t) := \sum_{n_1,n_2\geq 0} c_A\left(n_1,n_2\right) s^{n_1} t^{n_2} = \frac{1}{1-\left(\sum_{a\in A} s^a\right)\left(\sum_{a\in A} t^a\right)}.$$

What we are actually interested in is the *diagonal* of $F_A(s,t)$ given by the single variable generating function $G_A(x) := \sum_{n\geq 0} c_A\left(n,n\right) x^n$.

Firstly, we note that for any finite set A, $F_A(s,t)$ is a rational function of s and t and hence by Theorem 6.3.3 of Stanley [6], $G_A(x)$ is an algebraic function of x.

In the sequel we will study certain classes of infinite sets A for which we can obtain explicit algebraic expressions for $G_A(x)$.

We apply the diagonalisation technique as described in Stanley to find explicit algebraic generating functions for pairs of compositions with the same number of parts where all parts are at least d, with d any positive integer. In this case the bivariate generating function will be denoted by

$$F_d(s,t) := \sum_{n_1,n_2\geq 0} c\left(n_1,n_2\right) s^{n_1} t^{n_2} = \frac{1}{1-\frac{s^d t^d}{(1-s)(1-t)}}.$$

Note that for $d = 2$, compositions of $n \geq 3$ with all parts at least 2 are equal in number to compositions of $n-2$ with all parts either 1 or 2.

Using the approach detailed in Stanley

$$G_d(x) = [s^0]F_d(s, x/s) = \frac{1}{2\pi i}\int_{|s|=\alpha} F(s,x/s)\frac{ds}{s}$$

for some $\alpha > 0$.

By the residue theorem it follows that $G_d(x)$ is equal to the sum of residues of the integrand summed over all singularities inside the circle $|s| = \alpha$. In the case of $G_d(x)$ the integrand $F_d(s,x/s)/s$ becomes

$$\frac{(s-1)(s-x)}{s\left(sx^d - sx + x + s^2 - s\right)}.$$

The poles are at $s=0$ and at the root of the quadratic equation $s^2-s(1+x-x^d)+x=0$ that approaches 0 as $x\to 0$, namely

$$s_0 := \frac{1}{2}\left(1+x-x^d-\sqrt{\left(x^d-x-1\right)^2-4x}\right).$$

The residue at s_0 is

$$\frac{(s_0-1)(s_0-x)}{s_0\left(x^d-x+2s_0-1\right)} = \frac{x^d}{\sqrt{\left(1+x-x^d\right)^2-4x}}.$$

The residue at $s=0$ is 1, so that

$$G_d(x) = 1+\frac{x^d}{\sqrt{\left(1+x-x^d\right)^2-4x}}.$$

4.1 Asymptotic estimates

Firstly we consider the number of compositions $c_d(n)$ of n with parts at least d, which has generating function

$$\frac{1}{1-\frac{s^d}{1-s}} = \frac{s-1}{s^d+s-1}.$$

Let ρ denote the positive root of least modulus of $s^d+s-1=0$. Then by singularity analysis $c_d(n)\sim\frac{\rho-1}{\rho\left(d\rho^{d-1}+1\right)}\rho^{-n}$ and consequently the number of all pairs of such compositions of n is

$$c_d(n)^2 \sim \frac{(\rho-1)^2}{\rho^2\left(d\rho^{d-1}+1\right)^2}\rho^{-2n}.$$

Now consider the discriminant of the square-root in the expression for $G_d(x)$, namely $\left(x^d-x-1\right)^2-4x$. For odd values of d this factors as

$$\left(1-x-2x^{\frac{d+1}{2}}-x^d\right)\left(1-x+2x^{\frac{d+1}{2}}-x^d\right),$$

and for even d it factors as

$$\left(1-x-2x^{d/2}+x^d\right)\left(1-x+2x^{d/2}+x^d\right).$$

First we consider the case of odd d. Here the dominant pole comes from the discriminant factor $1-x-2x^{\frac{d+1}{2}}-x^d$. Setting $x=z^2$ we can factor this further as $\left(-z^d - z + 1\right)\left(z^d + z + 1\right)$ which has real roots of least modulus for $z = \rho$ and $z = -\rho$, that is, for $x = \rho^2$. We must therefore compute the singular expansion of $G_d(x)$ as $x \to \rho^2$.

Now as $x \to \rho^2$,

$$\frac{x^d}{\sqrt{1 - x - x^d + 2x^{\frac{d+1}{2}}}} \sim \frac{\rho^{2d}}{\sqrt{(-\rho^d + \rho + 1)(\rho^d - \rho + 1)}},$$

whereas

$$\frac{1}{\sqrt{1 - x - x^d - 2x^{\frac{d+1}{2}}}} \sim \frac{1}{\rho\sqrt{(\rho^{d-1} + 1)(d\rho^{d-1} + 1)}} \frac{1}{\sqrt{1 - \frac{x}{\rho^2}}}.$$

Therefore

$$G_d(x) \sim \frac{\rho^{2d-1}}{\sqrt{(\rho^{d-1} + 1)(d\rho^{d-1} + 1)}\sqrt{(-\rho^d + \rho + 1)(\rho^d - \rho + 1)}} \frac{1}{\sqrt{1 - \frac{x}{\rho^2}}}.$$

Hence by applying singularity analysis, for d odd

$$c_d(n, n) \sim \frac{\rho^{2d-1}}{\sqrt{(\rho^{d-1} + 1)(d\rho^{d-1} + 1)}\sqrt{(-\rho^d + \rho + 1)(\rho^d - \rho + 1)}} \frac{\rho^{-2n}}{\sqrt{\pi n}}.$$

Next we consider the case of d even. Here the dominant pole comes from the discriminant factor $1 - x - 2x^{d/2} + x^d$. Setting $x = z^2$ we can factor this further as $\left(z^d - z - 1\right)\left(z^d + z - 1\right)$ which again has real roots of least modulus for $z = \rho$ and $z = -\rho$. Then as $x \to \rho^2$ we find

$$G_d(x) \sim \frac{\rho^{2d-1}}{\sqrt{(\rho^d - \rho + 1)(\rho^d + \rho + 1)}\sqrt{1 - d\rho^{d-2}(\rho^d - 1)}} \frac{1}{\sqrt{1 - \frac{x}{\rho^2}}}.$$

Hence by applying singularity analysis we have for d even,

$$c_d(n, n) \sim \frac{\rho^{2d-1}}{\sqrt{(\rho^d - \rho + 1)(\rho^d + \rho + 1)}\sqrt{1 - d\rho^{d-2}(\rho^d - 1)}} \frac{\rho^{-2n}}{\sqrt{\pi n}}.$$

In conclusion, we have shown the following.

Theorem 3 *The asymptotic proportion of pairs of compositions of n with the same number of parts and all parts at least d, is given when d is odd by*

$$\frac{\rho^{2d+1}\left(d\rho^{d-1} + 1\right)^2}{(\rho - 1)^2\sqrt{(\rho^{d-1} + 1)(d\rho^{d-1} + 1)}\sqrt{(-\rho^d + \rho + 1)(\rho^d - \rho + 1)}} \frac{1}{\sqrt{\pi n}}$$

and for d even by

$$\frac{\rho^{2d+1}\left(d\rho^{d-1}+1\right)^2}{(\rho-1)^2\sqrt{\left(\rho^d-\rho+1\right)\left(\rho^d+\rho+1\right)}\sqrt{1-d\rho^{d-2}\left(\rho^d-1\right)}}\frac{1}{\sqrt{\pi n}},$$

where ρ denotes the smallest positive root of the equation $s^d+s-1=0$.

Remark: By subtracting $d-1$ from each part of a partition of n with parts at least d, we can also obtain the exact expressions

$$c_d(n) = \sum_{k=1}^{\left\lfloor\frac{n-1}{d-1}\right\rfloor}\binom{n-(d-1)k-1}{k-1}$$

and

$$c_d(n,n) = \sum_{k=1}^{\left\lfloor\frac{n-1}{d-1}\right\rfloor}\binom{n-(d-1)k-1}{k-1}^2.$$

4.2 Some further explicit algebraic generating functions

In a similar manner we can apply the diagonalisation technique to find algebraic generating functions for pairs of compositions with the same number of parts, where all parts are *odd and at least d*, with d any odd positive integer. In this case the bivariate generating function to consider is

$$\sum_{n_1,n_2\geq 0} c_o\left(n_1,n_2\right)s^{n_1}t^{n_2} = \frac{1}{1-\frac{s^d t^d}{(1-s^2)(1-t^2)}}.$$

The diagonal is then computed to be

$$\sum_{n\geq 0} c_o\left(n,n\right)x^n = 1+\frac{x^d}{\sqrt{\left(1+x^2-x^d\right)^2-4x^2}}.$$

Then using singularity analysis as before, we find that the asymptotic proportion of pairs of compositions of n with the same number of parts, all parts odd and at least d, is given by

$$\frac{\rho^{2d+1}\left(d\rho^{d-1}+2\rho\right)^2}{\left(\rho^2-1\right)^2\sqrt{d\rho^{2(d-1)}-2\rho^2+2}\sqrt{1-\rho^{2d}+\rho^4+2\rho^2}}\frac{1}{\sqrt{\pi n}},$$

where ρ denotes the smallest positive root of the equation $s^d+s^2-1=0$.

Acknowledgment

We are indebted our to anonymous referees for improving our presentation, as well as to Ákos Seress for an interesting remark.

References

[1] M. Bóna, A Walk Through Combinatorics, *World Scientific*, 2002.

[2] M. Bóna, R. Pemantle, H. Wilf, On the probability that two randomly selected permutations have the same number of cycles, *in preparation.*

[3] P. Flajolet, A. M. Odlyzko, Singularity Analysis of Generating Functions, *SIAM J. Discrete Math.* **3** (1990), 216–240.

[4] E. Munarini, N. Z. Salvi, On the rank polynomial of the lattice of order ideals of fences and crowns, *Discrete Mathematics* **259** (2002), 163-177.

[5] A. M. Odlyzko, Asymptotic Enumeration Methods, in *Handbook of Combinatorics*, vol. 2, Elsevier, 1995.

[6] R. Stanley, Enumerative Combinatorics, Volume 2, *Cambridge University Press*, 1997.

[7] H. Wilf, On The Variance of The Stirling Cycle Numbers, *preprint.*

[8] H. Wilf, Generatingfunctionology, 3rd edition, *A. K. Peters*, 2006.