

# Enumeration and asymptotics of restricted compositions having the same number of parts

Cyril Banderier\* and Paweł Hitczenko<sup>†‡§</sup>

April 22, 2010

**Abstract:** We study pairs and  $m$ -tuples of compositions of a positive integer  $n$  with parts restricted to a subset  $\mathcal{P}$  of positive integers. We obtain some exact enumeration results for the number of tuples of such compositions having the same number of parts. Under the uniform probability model, we obtain the asymptotics for the probability that two or, more generally  $m$ , randomly and independently chosen compositions of  $n$  have the same number of parts. Our results extend an earlier work of Bóna and Knopfmacher.

Keywords: integer composition, pairs of compositions, local limit theorem

## 1 Introduction

In this note we study compositions of positive integers. Let us recall that a composition of a positive integer  $n$  is any  $k$ -tuple  $(\kappa_1, \dots, \kappa_k)$ ,  $k \geq 1$ , of positive integers that sum up to  $n$ . The  $\kappa_j$ 's are called the parts (or summands) of a composition. It is elementary and well-known (see, e.g. [1]) that there are  $\binom{n-1}{k-1}$  compositions of  $n$  with  $k$  parts, and thus there are  $2^{n-1}$  compositions of  $n$ . By restricted compositions we mean compositions whose parts are confined to be in a fixed subset  $\mathcal{P}$  of  $\mathbb{N}$ .

The main motivation for this work is a recent paper [4] in which the authors studied pairs of compositions with the same number of parts. More precisely, for two randomly and independently chosen compositions of  $n$  they studied the asymptotic probability that they have the same number of parts. Furthermore, relying on the generating function approach, for a few *specific* subsets  $\mathcal{P}$  they addressed the same question for pairs of restricted compositions. In each of

---

\*Laboratoire d'Informatique de Paris Nord, UMR CNRS 7030, Institut Galilée, Université Paris 13, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France. <http://www-lipn.univ-paris13.fr/~banderier/>

†Departments of Mathematics and Computer Science, Drexel University, Philadelphia, PA 19104, USA. <http://www.math.drexel.edu/~phitczen/>

‡supported in part by the NSA grant #H98230-09-1-0062

§corresponding author

these cases this probability is asymptotic to  $C/\sqrt{n}$  with  $C$  depending on  $\mathcal{P}$ . Our main aim here is to extend these results. First, we show that this asymptotics is universal. That is, we show that for *arbitrary* subset  $\mathcal{P}$  containing two relatively prime elements the probability that two independently chosen random compositions of  $n$  with parts restricted to  $\mathcal{P}$  have the same number of parts, is asymptotic to  $C/\sqrt{n}$ . The value of  $C$  depends, generally, on  $\mathcal{P}$  and is explicit. (As it turns out, the value of one of the constants given in [4] is erroneous and we supply the correct value.) Secondly, we consider the same question for  $m > 2$  and we show that in this case the sought probability is asymptotic to  $C/\sqrt{n^{m-1}}$  for an explicitly given constant  $C$  whose value depends on  $\mathcal{P}$  and  $m$  only.

Our approach is more probabilistic and is based on a local limit theorem for discrete random variables. This approach seems much more universal than that of Bóna and Knopfmacher who relied on analytic techniques and singularity analysis. The universality of using more probabilistic technique was first noticed in [3] where certain types of random trees were studied. The authors obtained, in particular (see [3, Lemma 6]), a general statement indicating how local limit theorem can help in evaluating probabilities that two independently chosen random structures of the same size have the same number of components (our Lemma 5.2 is essentially the same as [3, Lemma 6] for Gaussian density with a bit different proof; it was obtained independently, but later than Lemma 6 of [3]). As we will see, these statements remain true if one considers more than two random structures.

## 2 Generating functions for pairs of compositions having the same number of parts

Let us consider compositions with parts in a set  $\mathcal{P}$  (a fixed subset of  $\mathbb{N}$ ). To avoid trivial complications caused by the fact that there may be no compositions of a given  $n$  with all parts from  $\mathcal{P}$ , we assume that  $\mathcal{P}$  has at least two elements that are relatively prime (except when explicitly stated otherwise).

We introduce the generating function of the parts  $p(z) = \sum_{j \in \mathcal{P}} p_j z^j$ , ( $p_j$  is not necessarily 0 or 1, it can then be seen as the possible colors or the weight of the part  $j$ ). We thus assume that the  $p_j$ 's are non-negative real numbers such that  $\sum_{j \in \mathcal{P}} p_j > 1$ . This last condition is to ensure supercriticality of our scheme (see Section 4 below for more details). In the classical situation when  $p_j$  is 0 or 1 this condition holds automatically. Denote by

$$P(z, u) = \sum_{n \geq 0, k \geq 0} P_{n,k} u^k z^n = \frac{1}{1 - up(z)} \quad (1)$$

the bivariate generating function of compositions of  $n$  where  $k$  encodes its number of parts, and where the “size” of the composition is  $n$ .

With a slight abuse of notation, the corresponding univariate generating

function is

$$P(z) = \sum_{n \geq 0} P_n z^n = \frac{1}{1 - p(z)}. \quad (2)$$

This terminology is classical. For example, here are all the compositions of 5 in 3 parts restricted to be in the set  $\mathcal{P} = \{1, 2, 3, 4, 10\}$ :  $5 = 1 + 1 + 3 = 1 + 3 + 1 = 3 + 1 + 1 = 1 + 2 + 2 = 2 + 1 + 2 = 2 + 2 + 1$ , accordingly,  $P_{5,3} = 6$ .

Let  $X_n^{\mathcal{P}}$  be the random variable giving the number of parts in a random composition of  $n$  with parts belonging to  $\mathcal{P}$ . Random means that we consider the uniform distribution among all compositions of  $n$  with parts belonging to  $\mathcal{P}$ .

Given two subsets  $\mathcal{P}_1$  and  $\mathcal{P}_2$  of  $\mathbb{N}$ , we consider the probability  $\pi_n := \Pr(X_n^{\mathcal{P}_1} = X_n^{\mathcal{P}_2})$  that a random composition of  $n$  with parts in  $\mathcal{P}_1$  has the same number of parts as a random composition with parts in  $\mathcal{P}_2$ . We assume throughout that whenever two such compositions are chosen, they are chosen independently and from now on we will not be explicitly mentioning it. We then introduce the generating function  $D(z)$  of the number of pairs of compositions (the first one with parts in  $\mathcal{P}_1$ , the second one with parts in  $\mathcal{P}_2$ ) having the same size and the same number of parts. ( $D$  stands for “double” or “diagonal”, as  $D(z)$  can be obtained as a diagonal of multivariate function).

That is, we consider all  $k$ -tuples of elements of  $\mathcal{P}_1$  and all  $k$ -tuples of elements of  $\mathcal{P}_2$  such that their sum is  $n$ . For a fixed  $n$ , let  $D_n$  be the total number of such configurations (i.e., we sum over all  $k$ ).

In the next section, we deal with some interesting examples for which we get explicit formulas.

### 3 Some closed form formulas

#### 3.1 An example on tuples of domino tilings

Consider tilings of a  $2 \times n$  strip by dominoes. Any tiling is thus a sequence of either one horizontal domino or 2 vertical dominoes. The generating function of domino tilings is thus  $P(z) = \text{Seq}(z + z^2) = \frac{1}{1 - z - z^2}$ , which is the generating function of  $P_n = F_{n+1}$ , where  $F_n$  is the Fibonacci number  $F_n = F_{n-1} + F_{n-2}$ ,  $F_0 = 0$ ,  $F_1 = 1$ . (Equivalently, the Fibonacci recurrence reflects the fact that removing a horizontal domino on the top of an existing  $2 \times n$  tiling leads to a  $2 \times (n - 1)$  tiling, while removing 2 vertical dominoes on the top leads to a  $2 \times (n - 2)$  tiling).

**Puzzle:** Each of  $m$  children makes a tiling of a  $2 \times n$  strip. What is the probability  $\pi_n$  that these  $m$  tilings all have the same number of vertical dominoes, when  $n$  gets large?

For  $m = 2$ , the number of pairs is given by  $D(z) = [t^0] \sum_{k \geq 0} p^k(zt)p^k(1/t)$ , where  $p(z) = z + z^2$ , and the Cauchy formula gives

$$D(z) = \frac{1}{2i\pi} \oint \frac{1}{1 - p(zt)p(1/t)} \frac{dt}{t} = \frac{1}{2i\pi} \oint \frac{\text{Num}(z, t)}{\text{Den}(z, t)} dt, \quad (3)$$

where  $Num$  and  $Den$  are polynomials in  $z, t$ . Let  $Z(z)$  be any root of  $Den$ , i.e.  $Den(z, Z) = 0$ , such that  $Z$  is inside the contour of integration for  $z \sim 0$ . Then, a residue computation gives:

$$D(z) = \sum_Z \frac{Num(z, Z)}{\partial_t Den(z, Z)} = \frac{1}{\sqrt{z^4 - 2z^3 - z^2 - 2z + 1}}$$

$$= 1 + z + 2z^2 + 5z^3 + 11z^4 + 26z^5 + 63z^6 + 153z^7 + 376z^8 + 931z^9 + O(z^{10}).$$

This is the sequence A051286 from [9]  $D_n = \sum_{k=0}^n \binom{n-k}{k}^2$ , which also is the Whitney number of level  $n$  of the lattice of the ideals of the fence of order  $2n$ . The probability that 2 tilings of a  $2 \times n$  strip have the same number of vertical dominoes is therefore (via singularity analysis, which can be done automatically with some computer algebra systems, e.g. via the Algolib Maple package available at <http://algo.inria.fr/librairies>, and more precisely via the *equivalent* command of Bruno Salvy):

$$\pi_n = D_n/P_n^2 \sim \frac{5^{3/4}}{2\sqrt{\pi}\sqrt{n}} + \frac{5^{1/4}(\frac{11}{32} - \frac{\sqrt{5}}{4})}{\sqrt{\pi n^{3/2}}} + O\left(\frac{1}{\sqrt{n^5}}\right) \approx \frac{.9432407854}{\sqrt{n}}. \quad (4)$$

Note that this is consistent with the constant  $C$  given in Equation (11) in [4]. Our computations are available at <http://lipn.fr/~cb/Pawel/Maple/> but, as Maple does not always simplify algebraic numbers like humans would do (some denesting options are missing), we used here some of our own denesting recipes so that these nested radicals become more readable for human eyes.

For  $m = 3$ , it is possible to compute the diagonal via creative telescoping (as automated in Maple via the *MGfun* package of Frédéric Chyzak), this leads to the following differential equation:

$$\begin{aligned} 0 &= (4z^7 + 7z^6 + 7z^5 + 15z^4 + 41z^2 + z + 1) D(z) \\ &+ (5z^8 + 12z^7 + 7z^6 + 62z^5 + 88z^3 + z^2 + 6z - 1) \frac{d}{dz} D(z) \\ &+ z(z^2 + 1)(z^4 - z^3 + 5z^2 + z + 1)(z^2 + 4z - 1) \frac{d^2}{dz^2} D(z). \end{aligned}$$

Using the Frobenius method (i.e., a basis of local formal solutions of the ODE at its singularity) and a numerical scheme of our own in order to know which linear combination of this basis to take (i.e., in order to get the so-called “connection constants”)<sup>1</sup>, and then using singularity analysis leads to

$$\pi_n \sim \frac{5\sqrt{15}}{6} \frac{1}{\pi n} + \frac{5(10\sqrt{3} - 9\sqrt{15})}{54} \frac{1}{\pi n^2} + O(1/n^3) \approx \frac{1.027340740}{n}.$$

<sup>1</sup>Note that, as typical with the Frobenius method (or also with the Birkhoff-Tritjinski method, see [11]), it is not always possible to decide the connection constant(s); in the next sections, we give a rigorous probabilistic approach which allows to get this constant, and therefore full asymptotics by coupling it with the Frobenius method!

This asymptotics also proves that  $D(z)$  is not an algebraic function (the local basis of the differential equation involves an  $\ln$  term).

For  $m = 4$ , creative telescoping leads to the following differential equation:

$$\begin{aligned} & 2(132z^{16} - 3563z^{15} + \dots + 110)D(z) + 2(209474z^{14} + \dots - 1581z)\frac{d}{dz}D(z) \\ & + (704z^{18} + \dots - 10143z^2)\frac{d^2}{dz^2}D(z) + z(165z^{18} + \dots - 55)\frac{d^3}{dz^3}D(z) \\ & + z^2(z-1)(z+1)(z^2+z+1)(z^2-7z+1)(z^2-z+1)(z^4+2z^3+19z^2+2z+1)(11z^6+\dots+11)\frac{d^4}{dz^4}D(z). \end{aligned}$$

Using the Frobenius method and a numerical scheme of ours, this leads to

$$\pi_n \sim \frac{25}{8} \frac{5^{1/4}\sqrt{2}}{\sqrt{\pi n^3}} + \frac{5}{256} \frac{5^{1/4}\sqrt{2}(47\sqrt{5}-240)}{\sqrt{\pi^3 n^5}} + O\left(\frac{1}{\sqrt{n^7}}\right) \approx \frac{1.186814138}{\sqrt{n^3}}.$$

It is noteworthy that this asymptotics is compatible with the fact that  $D(z)$  could be an algebraic function. However, a guess based on Padé approximants fails to find any such algebraic equation; what is more, the index of nilpotence mod  $2, 3, 5, 7, 11$  of  $D(z)$  is  $3$  (i.e. the smallest  $i$  such that  $(d/dz)^i = L \pmod{p}$  is  $i = 3$  for primes  $p = 2, 3, 5, 7, 11, \dots$  and  $L$  is the above irreducible linear differential operator cancelling  $D(z)$ ), therefore, according to a conjecture of Grothendieck on the  $p$ -curvature (see [5]), the function is not algebraic.

For  $m = 5$ , one gets a holonomic non algebraic function satisfying a differential equation of order 6 and of degree 38, which leads to

$$\pi_n \sim \frac{25\sqrt{5}}{4} \frac{1}{\pi^2 n^2} \approx \frac{1.416006588}{n^2}.$$

For larger  $m$ , one still observes experimentally that  $D(z)$  satisfies a differential equation  $0 = \sum_{i=0}^{2m-4} \ell_i(z) \frac{d^i}{dz^i} D(z)$  which has few nice striking phenomena, indeed where  $\ell_i(z)$  is a polynomial in  $z$  of degree one less than  $\ell_{i+1}(z)$ , and where the leading term  $\ell_{2m-4}(z)$  somehow factorizes nicely, having degree roughly  $5 \times 2^{m-2}$ . The closed form of the coefficients is  $D_n(m) = \sum_{k=0}^n \binom{n-k}{k}^m$ . It is also possible to get their asymptotics via the Laplace method.

### 3.2 Other nice explicit formulas

For sure, it is clear from the previous subsection, that we could play the same game for any  $m$ -tuple of compositions with parts restricted to  $m$  different sets, encoded by  $p_1(z), \dots, p_m(z)$ .

**Proposition 3.1** *The generating function for the number of  $m$ -tuples of compositions having the same number of parts is given by*

$$D(z) = \frac{1}{(2i\pi)^{m-1}} \oint \frac{1}{1 - p_1(zt_2 \dots t_m) p_2(1/t_2) \dots p_m(1/t_m)} \frac{dt_2}{t_2} \dots \frac{dt_m}{t_m}.$$

Therefore, we should not expect any nice close form solution for  $D(z)$  whenever  $m > 2$ ; while for  $m = 2$ , whenever all the  $p_i(z)$ 's are polynomials or rational functions,  $D(z)$  will always be an algebraic function, the coefficients of which can be expressed by nested sums of binomial coefficients (using Lagrange inversion).

For example, if  $p_1 = p_2 = 2z + z^2$  (which can be considered as tilings with bicolored horizontal dominoes), one gets  $D_{n+1}(2) = \sum_{k=0}^n \sum_{j=0}^k \binom{k}{j} \binom{k+j}{j}$  which is [9, Sequence A089165], namely partial sums of the central Delannoy numbers<sup>2</sup>, which also counts the number of peaks at odd level in Schroeder paths of length  $2n+2$ , it also appears in some computations of the resulting resistance between two nodes of an infinite lattice of unit resistors.

If  $p_i(z) = \frac{z}{1-z}$  (that is, we consider compositions with any parts), then  $D_{n+1}(m) = \sum_{k=0}^n \binom{n}{k}^m$  is the  $m$ -th Franel number, and the probability that  $m$  unrestricted compositions of  $n$  have the same number of parts is thus  $\pi_n \sim C_m / \sqrt{(\pi n)^{m-1}}$ , with  $C_2 = 1, C_3 = 2/\sqrt{3}, C_4 = \sqrt{2}$ , etc. As we will see in a next section, for general  $m$  we have  $C_m = \sqrt{2^{m-1}/m}$ .

If  $p_i(z) = \frac{z}{1-z^2}$  (that is, we consider compositions of  $n$  with odd parts only), one has  $D_{n+1}(m) = D_n^{\{1,2\}}(m) = \sum_{k=0}^n \binom{n-k}{k}^m$  (that is compositions of  $n-1$  with parts 1 or 2).

If  $p_i(z) = \frac{z^d}{1-z^d}$  (that is, we consider compositions of  $n$  with parts multiple of  $d$  only), one has  $D_{dn}(m) = \sum_{k=0}^n \binom{n}{k}^m$ , again the  $n$ -th Franel numbers (all the others coefficients being 0, as we have here an example for which the gcd of the parts is not 1).

### 3.3 Explicit formulas for pairs with two types of parts

Let us add few examples for which parts are in two different sets  $\mathcal{P}_1$  and  $\mathcal{P}_2$ . If  $p_1(z) = z + z^2$  and  $p_2(z) = z + 2z^2$ , then  $D_n = \sum_{k=0, k \text{ even}}^n \binom{n+k-1}{n-k} \binom{k}{k/2}$ . This is [9, Sequence A101500], registered as the Chebyshev transform of the central binomial numbers. This case is interesting as we have here

$$\pi_n \sim \frac{\sqrt{72 + 42\sqrt{3}(\sqrt{5} - 5)}(\sqrt{2} - 2)}{12\sqrt{\pi n}} \left( \frac{1 - \sqrt{2} - \sqrt{5} + \sqrt{10}}{2(2 - \sqrt{3})} \right)^n \approx 1.62 \frac{.95^n}{\sqrt{\pi n}},$$

which is therefore exponentially smaller than the order of magnitude of our previous examples. We will comment later on this fact. Going to a slightly more general case  $p_i(z) = \alpha_i z + \beta_i z^2$ , one has for  $m = 2$ :

$$D(z) = \frac{1}{\sqrt{1 - 2\alpha_1\alpha_2z + (\alpha_1^2\alpha_2^2 - 2\beta_1\beta_2)z^2 - 2\alpha_1\alpha_2\beta_1\beta_2z^3 + \beta_1^2\beta_2^2z^4}}$$

<sup>2</sup>We don't have the general formula for  $m > 2$  in this case. One may suspect that it is again a partial (nested) sum of binomial numbers, like in the other examples from this Section. It is a nice challenge for computer algebra to design an algorithm which could output such a closed form (taking in input the differential equation). Note that from a guessing point of view, it helps to know the asymptotics (which are given for free via the Frobenius method), as, if we restrict to closed forms (allowing nested sums of positive terms), it gives some upper and lower bounds on the powers of binomials in this (possible) closed-form.

which factorizes nicely when  $\beta_1 = \beta_2 = 1$ :

$$D(z) = 1/\sqrt{(\alpha_1\alpha_2z - 1 - 2z - z^2)(\alpha_1\alpha_2z - 1 + 2z - z^2)},$$

if additionally  $\alpha_1\alpha_2 = \pm 4$ , it gives back the sum of Central Delannoy numbers:

$$D(z) = 1/(1-z) \times 1/\sqrt{1 + (2 + |\alpha_1\alpha_2|)z + z^2}.$$

If  $p_1(z) = z/(1-z)$  and  $p_2(z) = z + z^2$ , then  $D_{2n} = \binom{3n+1}{n} [9, A045721]$ . If  $p_1(z) = z/(1-z)$  and  $p_2(z) = 2z + z^2$ , then  $D_n = \frac{3n-1}{n} \binom{2n-2}{n-1}$ . If  $p_1(z) = \alpha z + \beta z^2$  and  $p_2(z) = z^k/(1-z^k)$  ( $k > 1$ ), then  $D_{2n} = \beta^k$ . A nice surprise is given by  $p_1(z) = z/(1-z)$  and  $p_2(z) = z + z^2$ , for which we get  $D(z) = 1/2 - 1/2\sqrt{\frac{1+z}{1-3z}}$ , which is known to be the generating function of directed animals ([9, A005773], see also the numerous other combinatorial structures counted by this sequence: variants of Dyck/Motzkin paths, pattern avoiding permutations, base 3  $n$ -digit numbers with digit sum  $n, \dots$ ), we leave to the reader the pleasure of finding a bijective proof of all of this. (Note that this can lead to efficient uniform random generation algorithms).

In summary, it may seem possible to compute everything in all cases; however, for a generic  $\mathcal{P}$ , in order to compute the constant  $C_m$  involved in  $\pi_n \sim C_m \frac{1}{\sqrt{(\pi n)^{m-1}}}$ , we need heavier and heavier computations if the degrees of the  $p_i(z)$ 's get large or if  $m$  is large. Current state of the art algorithms will take more than one day for  $m = 6$ , and gigabytes of memory, so this ‘‘computer algebra’’ approach (may it be via guessing or via holonomy theory) has some intrinsic limitations on actual computers. What is more, for a given  $\mathcal{P}$ , it remains a nice challenge to get a rigorous (Zeilbergerian computer algebra) proof for all  $m$  *at once*.

In the next sections, we show that the technical conditions to get a local limit law hold, and that this allows to get the constant  $C$ , for any  $\mathcal{P}$ , for all  $m$ .

## 4 Local limit theorem for the number of parts in restricted compositions

The discussion in this section pretty much gathers what has been developed in various parts of the compendium on Analytic Combinatorics by Flajolet & Sedgewick [7].

Our main generating function (see Equation 1) is a particular case of a more general composition<sup>3</sup> scheme considered in Flajolet and Sedgewick, namely  $F(z, u) = g(uh(z))$ . In our case  $g(z) = 1/(1-z)$  and  $h(z) = p(z)$ . According to terminology used in [7, Definition IX.2, p. 629, Sec. IX.3], under our assumption that  $\sum_{j \in \mathcal{P}} p_j > 1$  the scheme is supercritical. As a consequence, the number of parts  $X_n^{\mathcal{P}}$  is asymptotically normal as  $n \rightarrow \infty$ , with both the mean and the

<sup>3</sup>We cannot escape to this polysemy: Compositions are enumerated by a composition!

variance linear in  $n$ . We now briefly recapitulate the statements from [7]. The equation

$$p(z) = 1,$$

has the unique positive root  $\rho \in (0, 1)$ . As a consequence,  $F(z, 1)$  has a dominant simple pole as its singularity and thus the number of compositions of  $n$  with all parts in  $\mathcal{P}$  is

$$[z^n]F(z, 1) \sim \frac{1}{\rho p'(\rho)} \rho^{-n} (1 + O(\varepsilon^n)),$$

where  $\varepsilon$  is a positive number less than 1, see [7, Theorem V.1, p. 294]. The probability generating function of  $X_n^{\mathcal{P}}$  is given by

$$f_n(u) = \frac{[z^n]F(z, u)}{[z^n]F(z, 1)}.$$

In a sufficiently small neighborhood of  $u = 1$ , as a function of  $z$ ,  $F(z, u)$  given in (1) has a dominant singularity  $\rho(u)$  which is the unique positive solution of the equation

$$up(\rho(u)) = 1.$$

Consequently,

$$f_n(u) = \frac{[z^n]F(z, u)}{[z^n]F(z, 1)} \sim \frac{p'(\rho(1))}{p'(\rho(u))} \cdot \left(\frac{\rho(u)}{\rho(1)}\right)^{-n-1}.$$

It follows from the analysis of supercritical sequences given in [7, Proposition IX.7, p. 652] that the number of parts  $X_n^{\mathcal{P}}$  satisfies

$$\frac{X_n^{\mathcal{P}} - EX_n^{\mathcal{P}}}{\sqrt{\text{var}(X_n^{\mathcal{P}})}} \xrightarrow{d} N(0, 1),$$

where  $N(0, 1)$  denotes a standard normal random variable whose distribution function is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt,$$

and where the symbol “ $\xrightarrow{d}$ ” denotes the convergence in distribution. The asymptotic expressions for the expected value and the variance of  $X_n^{\mathcal{P}}$  are given by

$$EX_n^{\mathcal{P}} = \frac{n}{\rho p'(\rho)} + O(1),$$

and

$$\text{var}(X_n^{\mathcal{P}}) = \frac{\rho p''(\rho) + p'(\rho) - \rho(p'(\rho))^2}{\rho^2(p'(\rho))^3} n + O(1). \quad (5)$$

[Note that the expression for the coefficient of the variance given in Proposition IX.7 in [7] is incorrect; the correct version is  $(\rho h''(\rho) + h'(\rho) - \rho h'(\rho)^2)/(\rho h'(\rho)^3)$ , as given in many other places in the book.]

We now note that the central limit theorem can actually be strengthened to the local limit theorem, pretty much as discussed in [7, Theorem IX.14 and the remarks following its proof on p. 697]. Let us recall the following

**Definition 4.1** Let  $(X_n)$  be a sequence of integer valued random variables with  $EX_n = \mu_n$  and  $\text{var}(X_n) = \sigma_n^2$ . Let  $(\varepsilon_n)$  be a sequence of positive numbers going to 0. We say that  $(X_n)$  satisfies a local limit theorem (of Gaussian type) with speed  $\varepsilon_n$  if

$$\sup_{x \in \mathbb{R}} \left| \sigma_n \Pr(X_n = \lfloor \mu_n + x\sigma_n \rfloor) - \frac{e^{-x^2/2}}{\sqrt{2\pi}} \right| \leq \varepsilon_n.$$

As was discussed in [7, p. 697] to see that the local limit theorem holds for restricted compositions it suffices to check that  $\rho(u)$  when restricted to the unit circle uniquely attains its minimum<sup>4</sup> at  $u = 1$ . This is what we prove in the following lemma.

**Lemma 4.2** Let  $p$  be a power series with nonnegative coefficients, of radius of convergence  $\rho_p > 0$  (possibly  $\rho_p = +\infty$ ). Let  $\rho(u)$  be as above the positive root<sup>5</sup> of  $up(\rho(u)) = 1$ . If  $p$  is aperiodic<sup>6</sup>, then for  $0 < \frac{1}{R} < \rho_p$  and  $t \in ]0, 2\pi[$ , we have

$$\rho(R) < |\rho(Re^{it})|,$$

i.e. the minimum on each circle is on the positive real axis. In particular, if the radius of convergence of  $p$  is larger than 1, then for  $|u| = 1$  and  $u \neq 1$  we have

$$\rho(1) < |\rho(u)|.$$

**Proof.** First,  $p$  has nonnegative real coefficients, therefore the triangle inequality gives  $p(|\rho(u)|) \geq |p(\rho(u))|$ , with equality only if  $p(\rho(u))$  has just nonnegative terms, which is not possible if  $\rho(u) \notin \mathbb{R}^+$  as  $p$  is aperiodic with nonnegative coefficients. Hence one has a strict triangle inequality:  $p(|\rho(u)|) > |p(\rho(u))| = |1/u| = 1/R$  (the middle equality is just the definition of  $\rho$  and the last equality comes from the fact we are on the circle  $|u| = R$ ). As  $p$  is increasing on  $[0, 1/R]$ , we can apply  $p^{-1}$  to  $p(|\rho(u)|) > 1/R$  which gives  $p^{-1}(p(|\rho(u)|)) > p^{-1}(1/R)$ , that is  $|\rho(u)| > \rho(R)$ .  $\square$

Note that the aperiodicity condition is important, e.g. for  $p(z) = z^2 + z^6$  (i.e.  $\rho(u)$  is the radius of convergence of  $P(z, u) = 1/(1 - u(z^2 + z^6))$ ), one has  $\rho(-1) = i\rho(1)$ ; however some periodic cases have a unique minimum on the unit circle, e.g.  $p(z) = z^2 + z^4$ . Note also that (in either periodic or aperiodic case), the uniqueness of the minimum on the circle  $|u| = 1$  at  $u = 1$  does not hold in general for the other roots of  $up(\rho(u)) = 1$ .

<sup>4</sup>There is a typo in [7] p.697: the inequality direction is wrong.

<sup>5</sup> $p(z)$  has nonnegative real coefficients and is thus increasing in a neighborhood of 0, i.e. on  $z \in [0, +\epsilon]$ .  $p$  being analytic near 0, it is continuous and for any  $x \in \mathbb{R}$  small enough,  $p(z) = x$  will therefore have a real positive root  $z_x$ , and this root will be analytic in  $x$ . This is the root that we call “real positive”.

<sup>6</sup>A power series  $p$  is said to be periodic if and only if there exists a power series  $q$  and an integer  $g > 1$  such that  $p(z) = q(z^g)$ . Equivalently, the gcd of the support (=the ranks of nonzero coefficients) of the power series  $p$  is  $g \neq 1$ . If this gcd  $g$  equals 1, then  $p$  is said to be aperiodic.

## 5 Asymptotic probability that restricted compositions have the same number of parts

Our motivation for including the results from [7] in the preceding section is the following theorem which considerably extends the main results of [4]. We single out the case  $m = 2$  since in some cases it has been already studied in the literature.

### 5.1 Pairs of compositions

**Theorem 5.1** *Let  $\mathcal{P} \subset \mathbb{N}$ . The probability that two random compositions with parts in  $\mathcal{P}$  have the same number of parts is, asymptotically as  $n \rightarrow \infty$ ,*

$$\pi_n \sim \frac{C}{\sqrt{\pi}\sqrt{n}},$$

where the value of  $C$  is given by

$$C = \frac{\rho(p'(\rho))^{3/2}}{2\sqrt{\rho p''(\rho) + p'(\rho) - \rho(p'(\rho))^2}}. \quad (6)$$

Before proving this theorem let us make some comments.

**Remarks and examples:**

- (i) Some special cases were considered in [4]. They include unrestricted compositions ( $\mathcal{P} = \mathbb{N}$ ),  $\mathcal{P} = \{1, 2\}$ , or more generally  $\mathcal{P} = \{a, b\}$  with  $a, b$  relatively prime (although the formula given there is not correct – see below), compositions with all parts of size at least  $d$  ( $\mathcal{P} = \{n \in \mathbb{N} : n \geq d\}$ ), and compositions with all parts odd and at least  $d$ . The arguments of [4] rely on the analysis of the asymptotics of the bivariate generating functions, which is sometimes difficult and does not seem to be easily amenable to the analysis in the case of a general subset  $\mathcal{P}$  of positive integers. Our approach is much more probabilistic and relies on a local limit theorem for the number of parts in a random composition with parts in  $\mathcal{P}$ . This turned out to be a much more universal tool.
- (ii) To illustrate the principle behind our approach, consider the unrestricted compositions. As was observed in [8], in that case  $X_n^{\mathcal{P}}$  is distributed like  $1 + \text{Bin}(n - 1, 1/2)$  random variable. Therefore,

$$\pi_n = \Pr(\text{Bin}(n - 1, 1/2) = \text{Bin}'(n - 1, 1/2))$$

where  $\text{Bin}$  and  $\text{Bin}'$  denote two independent binomial random variables with specified parameters. Since the second parameter is  $1/2$  we have

$$\text{Bin}(n - 1, 1/2) \stackrel{d}{=} n - 1 - \text{Bin}(n - 1, 1/2).$$

Therefore, by independence we get

$$\pi_n = \Pr(\text{Bin}(n - 1, 1/2) + \text{Bin}'(n - 1, 1/2) = n - 1).$$

Finally, since

$$\text{Bin}(n-1, 1/2) + \text{Bin}'(n-1, 1/2) \stackrel{d}{=} \text{Bin}(2(n-1), 1/2),$$

we obtain that

$$\pi_n = \Pr(\text{Bin}(2n-2, 1/2) = n-1) = \frac{\binom{2n-2}{n-1}}{2^{2n-2}} \sim \frac{1}{\sqrt{\pi n}},$$

by Stirling's formula. This is consistent with (6) (and with [4]) as for unrestricted compositions  $p(z) = \sum_{k \geq 1} z^k = z/(1-z)$ , so that  $\rho = 1/2$ ,  $p'(z) = 1/(1-z)^2$ , and  $p''(z) = 2/(1-z)^3$  which gives  $C = 1$ .

- (iii) Although the above argument may look very special and heavily relying on the properties of binomial random variables, our point here is that it is actually quite general. The key feature is that the number of parts (whether in unrestricted or arbitrarily restricted compositions) satisfies the local limit theorem of Gaussian type, and this is enough to asymptotically evaluate the probability in Theorem 5.1.
- (iv) By means of another example consider compositions of  $n$  into two parts, i.e.  $\mathcal{P} = \{a, b\}$  with  $a, b$  relatively prime. Then Theorem 5.1 holds with

$$C = \frac{(a\rho^a + b\rho^b)^{3/2}}{2|a-b|\sqrt{\rho^{a+b}}}, \quad (7)$$

where  $\rho$  is the unique root of  $z^a + z^b = 1$  in the interval  $(0, 1)$ . In this case  $p(z) = z^a + z^b$  so that  $p'(z) = az^{a-1} + bz^{b-1}$  and  $p''(z) = a(a-1)z^{a-2} + b(b-1)z^{b-2}$ . Thus, writing the numerator of (6) as

$$\rho(P'(\rho))^{3/2} = \frac{1}{\sqrt{\rho}}(\rho P'(\rho))^{3/2} = \frac{1}{\sqrt{\rho}}(a\rho^a + b\rho^b)^{3/2},$$

we only need to check that

$$\rho P''(\rho) + P'(\rho) - \rho(P'(\rho))^2 = (a-b)^2 \rho^{a+b-1}. \quad (8)$$

But

$$\rho P''(\rho) + P'(\rho) = a^2 \rho^{a-1} + b^2 \rho^{b-1}$$

so that the left-hand side of (8) is

$$a^2 \rho^{a-1} + b^2 \rho^{b-1} - a^2 \rho^{2a-1} - b^2 \rho^{2b-1} - 2ab\rho^{a+b-1}.$$

Factoring and using  $\rho^a + \rho^b = 1$ , we see that this is

$$a^2 \rho^{a-1} (1 - \rho^a) + b^2 \rho^{b-1} (1 - \rho^b) - 2ab\rho^{a+b-1} = (a-b)^2 \rho^{a+b-1},$$

as claimed.

When  $a = 1$  and  $b = 2$  we have the Fibonacci numbers relation so that  $\rho = (\sqrt{5} - 1)/2$  and (7) becomes

$$C = \frac{(\rho + 2\rho^2)^{3/2}}{2\rho^{3/2}} = \frac{1}{2}(1 + 2\rho)^{3/2} = \frac{5^{3/4}}{2},$$

which agrees with (4) above and also with the expression (11) given in [4]. However, in the case of general  $a$  and  $b$ , the value of  $C$  was given in the last display of Section 3 in [4] as

$$\frac{\rho(a\rho^{a-1} + b\rho^{b-1})^2}{\sqrt{4(a+b)\rho^{2a+2b-2} + 2(1-\rho^{2a}-\rho^{2b})(a\rho^{2a-2} + b\rho^{2b-2})}}. \quad (9)$$

This is incorrect as it is lacking a factor  $|a - b|$  in the denominator (so that it gives the correct value of  $C$  when  $|a - b| = 1$  but not otherwise). To see this and also to reconcile (9) with (7) (up to a factor  $|a - b|$ ) we simplify (9) by noting that  $\rho^a + \rho^b = 1$  implies

$$1 - \rho^{2a} - \rho^{2b} = 1 - (\rho^a)^2 - \rho^{2b} = (1 + \rho^a)\rho^b - \rho^{2b} = \rho^b(1 + \rho^a - \rho^b) = 2\rho^{a+b}$$

so that the expression under the square root sign in (9) becomes

$$4\rho^{a+b-2}((a+b)\rho^{a+b} + a\rho^{2a} + b\rho^{2b}) = 4\rho^{a+b-2}(a\rho^a + b\rho^b)(\rho^a + \rho^b).$$

Using again  $\rho^a + \rho^b = 1$  (9) is seen to be

$$\frac{\rho^2(a\rho^{a-1} + b\rho^{b-1})^2}{2\sqrt{\rho^{a+b}(a\rho^a + b\rho^b)}} = \frac{(a\rho^a + b\rho^b)^{3/2}}{2\sqrt{\rho^{a+b}}},$$

which, except for the factor  $|a - b|$  in the denominator, agrees with (7).

- (v) Other examples from [4] can be rederived in the same fashion, but we once again would like to stress universality of our approach. As an extreme example, we can only repeat after [7]; even if we consider compositions into twin primes,  $\mathcal{P} = \{1, 3, 5, 7, 11, 13, 17, 19, 29, 31, \dots\}$  we know that the probability of two such compositions having the same number of parts is of order  $1/\sqrt{n}$ . This is rather remarkable, considering the fact that we don't even know whether this  $\mathcal{P}$  is finite or not.

**Proof of Theorem 5.1.** This will follow immediately from the following lemma applied to  $X_n = X_n^{\mathcal{P}}$  and formula (5) which gives the expression for  $\sigma_n$ . This lemma should be compared with a more general Lemma 6 of [3]. We include our proof to illustrate that seemingly very special arguments used in item (ii) are actually quite general.

**Lemma 5.2** *Let  $(X_n)$  with  $EX_n = \mu_n$  and  $\text{var}(X_n) = \sigma_n^2 \rightarrow \infty$  as  $n \rightarrow \infty$ , be a sequence of integer valued random variables satisfying a local limit theorem (of Gaussian type) with speed  $\varepsilon_n$  as described in Definition 4.1. Let  $(X'_n)$  be an independent copy of  $(X_n)$  defined on the same probability space. Then*

$$\pi_n = \Pr(X_n = X'_n) = \frac{1}{2\sqrt{\pi}\sigma_n} + O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right).$$

**Proof.** For  $X_n$  and  $X'_n$  as in the statement we have

$$\begin{aligned}\pi_n = \Pr(X_n = X'_n) &= \sum_{k \geq 1} \Pr(X_n = k = X'_n) = \sum_{k \geq 1} \Pr^2(X_n = k) \\ &= \sum_{\ell = -\infty}^{\infty} \Pr(X_n = \lfloor \mu_n \rfloor + \ell) \Pr(X_n = \lfloor \mu_n \rfloor - \ell). \quad (10)\end{aligned}$$

Now,

$$\Pr(X_n = \lfloor \mu_n \rfloor + \ell) = \Pr(X_n = \lfloor \mu_n \rfloor - \ell) + \left\{ \Pr(X_n = \lfloor \mu_n \rfloor + \ell) - \Pr(X_n = \lfloor \mu_n \rfloor - \ell) \right\}.$$

To estimate the term in the curly brackets take  $x_+$  and  $x_-$  such that

$$\lfloor \mu_n \rfloor + \ell = \lfloor \mu_n + x_+ \sigma_n \rfloor, \quad \text{and} \quad \lfloor \mu_n \rfloor - \ell = \lfloor \mu_n - x_- \sigma_n \rfloor.$$

By elementary considerations,  $-2\{\mu_n\}/\sigma_n \leq x_+ - x_- \leq 2(1 - \{\mu_n\})/\sigma_n$  (where  $\{z\}$  is the fractional part of  $z$ ), hence  $|x_+ - x_-| \leq 2/\sigma_n$ . Then

$$\begin{aligned}\Pr(X_n = \lfloor \mu_n \rfloor + \ell) - \Pr(X_n = \lfloor \mu_n \rfloor - \ell) &= \left( \Pr(X_n = \lfloor \mu_n + x_+ \sigma_n \rfloor) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{x_+^2}{2}} \right) \\ &\quad + \frac{1}{\sqrt{2\pi}\sigma_n} \left( e^{-\frac{x_+^2}{2}} - e^{-\frac{x_-^2}{2}} \right) \\ &\quad - \left( \Pr(X_n = \lfloor \mu_n - x_- \sigma_n \rfloor) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{x_-^2}{2}} \right).\end{aligned}$$

The absolute value of the first term is

$$\frac{1}{\sigma_n} \left| \sigma_n \Pr(X_n = \lfloor \mu_n + x_+ \sigma_n \rfloor) - \frac{1}{\sqrt{2\pi}} e^{-\frac{x_+^2}{2}} \right| \leq \frac{\varepsilon_n}{\sigma_n},$$

and similarly with the third term. Applying the inequality  $|f(x) - f(y)| \leq |x - y| \sup |f'(t)|$  to the middle term gives  $e^{-\frac{x_+^2}{2}} - e^{-\frac{x_-^2}{2}} \leq |x_+ - x_-| = O(1/\sigma_n)$  and so, the middle term is  $O(1/\sigma_n^2)$ . Therefore,

$$\left| \Pr(X_n = \lfloor \mu_n \rfloor + \ell) - \Pr(X_n = \lfloor \mu_n \rfloor - \ell) \right| = O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right).$$

Coming back to equation (10), we see that

$$\begin{aligned}\sum_{\ell = -\infty}^{\infty} \Pr(X_n = \lfloor \mu_n \rfloor + \ell) \left( \Pr(X_n = \lfloor \mu_n \rfloor - \ell) + O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right) \right) \\ = \left( \sum_{\ell = -\infty}^{\infty} \Pr(X_n = \lfloor \mu_n \rfloor + \ell, X'_n = \lfloor \mu_n \rfloor - \ell) \right) + 1 \times O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right) \\ = \Pr(X_n + X'_n = 2\lfloor \mu_n \rfloor) + O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right).\end{aligned}$$

Note that  $X_n + X'_n$ , being a sum of two i.i.d. random variables has mean  $2\mu_n$  and the variance  $2\sigma_n^2$ . Furthermore, since each of the summands satisfies the local limit theorem of Gaussian type, so does the sum (its probability generating function, being a square of  $f_n(u)$  falls into quasi-power category, just as  $f_n(u)$  does). Since  $2\lfloor\mu_n\rfloor = \lfloor 2\mu_n + x\sqrt{2}\sigma_n \rfloor$  for some  $x = O(1/\sigma_n)$ , just as before we have

$$\left| \sqrt{2}\sigma_n \Pr(X_n + X'_n = \lfloor 2\mu_n \rfloor) - \frac{1}{\sqrt{2\pi}} \right| = O\left(\varepsilon_n + \frac{1}{\sigma_n}\right).$$

Consequently,

$$\pi_n = \Pr(X_n = X'_n) = \Pr(X_n + X'_n = \lfloor 2\mu_n \rfloor) = \frac{1}{2\sqrt{\pi}\sigma_n} + O\left(\frac{\varepsilon_n}{\sigma_n} + \frac{1}{\sigma_n^2}\right),$$

which completes the proof of Lemma 5.2 and of Theorem 5.1  $\square$

## 5.2 Tuples of compositions

Here we sketch a proof of the following extension of Theorem 5.1.

**Theorem 5.3** *Let  $\mathcal{P} \subset \mathbb{N}$  and let  $m \geq 2$  be fixed. Then, the probability  $\pi_n$  that  $m$  randomly and independently chosen compositions with parts in  $\mathcal{P}$  all have the same number of parts is, asymptotically as  $n \rightarrow \infty$ ,*

$$\pi_n \sim \frac{C_m}{\sqrt{(\pi n)^{m-1}}},$$

where

$$C_m = \frac{1}{\sqrt{2^{m-1}m}} \left( \frac{\rho^2(p'(\rho))^3}{\rho p''(\rho) + p'(\rho) - \rho(p'(\rho))^2} \right)^{m-1}.$$

**Remark.** For unrestricted compositions, the expression in the big parentheses is 2 (see (i) in Remarks above). This gives  $C_m = \sqrt{2^{m-1}/m}$  as stated in Section 3.1.

*Proof of Theorem 5.3.* This follows immediately from the following statement which itself is a straightforward extension of Lemma 6 of [3] with essentially the same proof. We will be using it for Gaussian density in which case

$$\int_{-\infty}^{\infty} g^m(x) dx = \frac{1}{\sqrt{(2\pi)^m}} \int_{-\infty}^{\infty} e^{-\frac{mx^2}{2}} dx = \frac{1}{\sqrt{(2\pi)^{m-1}}} \frac{1}{\sqrt{m}}.$$

**Lemma 5.4** (Bóna–Flajolet). *Let  $(X_n)$  be integer valued with  $\mu_n = EX_n$ ,  $\sigma_n^2 = \text{var}(X_n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Let  $g$  be the probability density function and suppose that*

$$\lim_{n \rightarrow \infty} \sup_x |\Pr(X_n = \lfloor \mu_n + x\sigma_n \rfloor) - g(x)| = 0.$$

*Let further  $(X_n^{(k)})$ ,  $k = 1, \dots, m$  be independent copies of the sequence  $(X_n)$  defined on the same probability space. Then*

$$\sigma_n^{m-1} \Pr(X_n^{(1)} = X_n^{(2)} = \dots = X_n^{(m)}) \longrightarrow \int_{-\infty}^{\infty} g^m(x) dx, \quad \text{as } n \rightarrow \infty.$$

To see this we just follow the argument in [3, Lemma 6] with obvious adjustments: the left hand side above is

$$\sigma_n^{m-1} \sum_{k=1}^{\infty} \Pr^m(X_n = k) = \sigma_n^{m-1} \sum_{k=1}^{\infty} \Pr^m(X_n = \lfloor \mu_n + x_k \sigma_n \rfloor),$$

with  $\frac{k-\mu_n}{\sigma_n} \leq x_k < \frac{k+1-\mu_n}{\sigma_n}$ . This is further equal to

$$\frac{1}{\sigma_n} \sum_k (\sigma_n \Pr(X_n = \lfloor \mu_n + x_k \sigma_n \rfloor))^m \sim \frac{1}{\sigma_n} \sum_k g^m(x_k) \sim \frac{1}{\sigma_n} \int_{-\infty}^{\infty} g^m\left(\frac{x - \mu_n}{\sigma_n}\right) dx,$$

where the first approximation holds by the assumption of the lemma (after having first restricted the range of  $x_k$ 's to a large compact set) and the second by the Riemann sum approximation of the integral. Since the expression on the right is  $\int_{-\infty}^{\infty} g^m(x) dx$ , the result follows.  $\square$

## 6 Concluding remarks

1. In this article, we restricted our attention to compositions (giving first several new closed form formulas, and then going to the asymptotics), but it is clear that Lemma 5.4 can be applied to many other combinatorial structures, e.g. the probability that  $m$  random permutations of size  $n$  have the same number of cycles (see [10] for the case  $m = 2$ ), or the probability that  $m$  such permutations have the longest increasing subsequence of the same length, or the probability that  $m$  random planar maps have the largest component of same size. This leads to interesting analytic/computational considerations, as it will involve evaluating the integral of  $g^m(x)$  where  $g(x)$  will be the Tracy-Widom distribution density (provided the local limit theorem holds, which has not been proven yet), or the map-Airy distribution density (for which a local limit theorem was established, see [2]).
2. A similar approach can be also applied to tuples of combinatorial structures following  $m$  different local limit laws (with  $m$  densities having fast decreasing tails), as long as they have the same mean.
3. When the means are not the same, the probability of the same number of parts is generally of much smaller order. This is because if  $X_n$  has mean  $cn$  and  $X'_n$  has mean  $c'n$  and both have linear variances, then assuming w.l.o.g.  $c > c'$  and choosing  $\alpha < \frac{c-c'}{2}$  we note that if  $|X_n - cn| < \alpha n$  and  $|X'_n - c'n| < \alpha n$  then

$$X_n - X'_n > cn - \alpha n - (c'n + \alpha n) = (c - c' - 2\alpha)n > 0,$$

so that  $X_n \neq X'_n$ . Therefore,

$$\pi_n = \Pr(X_n = X'_n) \leq \Pr(|X_n - cn| \geq \alpha n) + \Pr(|X'_n - c'n| \geq \alpha n).$$

Since both  $X_n$  and  $X'_n$  converge to Gaussian law and  $\sigma_n = \sigma n$ , the first probability is roughly (with  $\beta = \alpha/\sqrt{\sigma}$ )

$$\Pr\left(\frac{|X_n - cn|}{\sqrt{\sigma n}} \geq \beta\sqrt{n}\right) \sim \frac{1}{\sqrt{2\pi}} \int_{\beta\sqrt{n}}^{\infty} e^{-t^2/2} dt \sim \frac{1}{\sqrt{2\pi}\beta\sqrt{n}} e^{-\frac{\beta^2 n}{2}},$$

by the well-known bound on the tails of Gaussian random variables (see, e.g. [6, Chapter VII, Lemma 2]). This is consistent with an example discussed in Section 3.1. The difficulty with making this argument rigorous is that the error in the first approximation is usually of much bigger (typically  $1/\sqrt{n}$ ) magnitude than the quantities that are approximated. However, a slightly weaker bound, namely,  $e^{-\beta n}$  (with a generally different value of  $\beta$ ) can be obtained by using Theorem IX.15 in [7] which asserts that tail probabilities of random variables falling in the scheme of quasi-powers are decaying exponentially fast. While this theorem is stated for the logarithm of  $\Pr(|X_n - cn| > \alpha n)$ , it is clear from its proof that one actually gets exponential bound on the tail probabilities (see Equation (88) on p. 701 in [7] and a few sentences following it).

4. The Gaussian local limit law explains the universality of the  $1/(\pi n)^{(m-1)/2}$  appearance for numerous combinatorial problems in which we would force  $m$  combinatorial structures of size  $n$  to have an extra parameter of the same value. We also wish to point out yet another insight provided by the probabilistic approach. As we mentioned in Section 3.1 (see Footnote 1 in), it allows to solve the connection constant problem intrinsic to the Frobenius method, and therefore, a combination of these two approaches (local limit laws plus Frobenius method) gives access to full asymptotics in numerous cases.

**Acknowledgements.** Part of this work was done during Paweł Hitczenko's sojourn at LIPN (Laboratoire d'Informatique de Paris Nord), thanks to the invited professor position funded by the university of Paris 13 and the Institute Galilée. He would like to thank the members of LIPN for their hospitality.

## References

- [1] George E. Andrews. *The theory of partitions*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.
- [2] Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michèle Soria. Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Struct. Algorithms*, 19(3-4):194–246, 2001.
- [3] Miklós Bóna and Philippe Flajolet. Isomorphism and symmetries in random phylogenetic trees. *J. Appl. Probab.*, 46(4):1005–1019, 2009.

- [4] Miklós Bóna and Arnold Knopfmacher. On the probability that certain compositions have the same number of parts. *Ann. Combinatorics, to appear*. Preprint available at: <http://www.math.ufl.edu/~bona/comp.pdf>.
- [5] Alin Bostan and Éric Schost. Fast algorithms for differential equations in positive characteristic. In John May, editor, *ISSAC'09*, pages 47–54, 2009.
- [6] William Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons Inc., New York, 1968.
- [7] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [8] Paweł Hitczenko and Carla D. Savage. On the multiplicity of parts in a random composition of a large integer. *SIAM J. Discrete Math.*, 18(2):418–435 (electronic), 2004.
- [9] Neil J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. Published electronically at [www.research.att.com/~njas/sequences/](http://www.research.att.com/~njas/sequences/), 2008.
- [10] Herbert S. Wilf. The variance of the Stirling cycle numbers. Available at <http://arxiv.org/abs/math/0511428v2>, 2005.
- [11] Jet Wimp and Doron Zeilberger. Resurrecting the asymptotics of linear recurrences. *J. Math. Anal. Appl.*, 111:162–176, 1985.