# A Generatingfunctionology Approach to a Problem of Wilf

Paweł Hitczenko
Department of Mathematics and Computer Science
Drexel University
Philadelphia, PA 19104-2875
phitczen@mcs.drexel.edu

Cecil Rousseau
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240
ccrousse@memphis.edu

Carla D. Savage
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
savage@csc.ncsu.edu

October 10, 2000

**Abstract**

Wilf posed the following problem: determine asymptotically as $n \to \infty$ the probability that a randomly chosen part size in a randomly chosen composition of $n$ has multiplicity $m$. One solution of this problem has been given by two of the authors [3]. In this paper, we study this question using the techniques of generating functions and singularity analysis.

# 1 Introduction

Let $n$ be a positive integer. A *composition* of $n$ with $p$ parts is a solution of the equation $n = \kappa_1 + \kappa_2 + \cdots + \kappa_p$ in positive integers $\kappa_1, \kappa_2, \ldots, \kappa_p$. We shall write $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_p)$ to symbolize the composition. For example, there are 16 compositions of 5, namely

| | | | |
|---|---|---|---|
| $(5)$ | $(4,1)$ | $(1,4)$ | $(3,2)$ |
| $(2,3)$ | $(3,1,1)$ | $(1,3,1)$ | $(1,1,3)$ |
| $(2,2,1)$ | $(2,1,2)$ | $(1,2,2)$ | $(2,1,1,1)$ |
| $(1,2,1,1)$ | $(1,1,2,1)$ | $(1,1,1,2)$ | $(1,1,1,1,1).$ |

The terms $\kappa_1, \ldots, \kappa_p$ are called the *parts* of the composition. The *multiplicity* of a part size is the number of parts with that size. For example, in the composition $(1,1,1,2)$ the multiplicity of 1 is 3 and the multiplicity of 2 is 1. A *partition* of $n$ with $p$ parts is a solution of $n = \lambda_1 + \lambda_2 + \cdots + \lambda_p$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. In [1] Corteel, Pittel, Savage, and Wilf proved that for every fixed $m \geq 1$, the probability that a randomly chosen part size in a random partition of $n$ approaches $1/(m(m+1))$ as $n \to \infty$. Wilf then posed the corresponding problem for compositions: determine asymptotically (as $n \to \infty$) the probability that a randomly chosen part size in a randomly chosen composition of $n$ has multiplicity $m$. One solution of this problem has been given by two of the authors [3]. In this note, we address the same question using generating functions and singularity analysis.

It is well known that there are $2^{n-1}$ compositions of $n$. One way to arrive at this result uses generating functions. The generating function for compositions with $p$ parts is

$$(z + z^2 + z^3 + \cdots)^p = \left(\frac{z}{1-z}\right)^p,$$

and summing over $p$ we have the generating function for all compositions:

$$G(z) = \sum_{p=1}^{\infty} \left(\frac{z}{1-z}\right)^p = \frac{z}{1-2z}.$$

The coefficient of $z^n$ in the expansion of $G(z)$, denoted by $[z^n]G(z)$, is the number of compositions of $n$, and clearly $[z^n]G(z) = 2^{n-1}$. From an analytic point of view, the fact that the number of compositions of $n$ is asymptotically (as well as exactly) $2^{n-1}$ is associated with the fact that the generating function is a rational function for which the pole nearest the origin (the only pole in this case) is simple and located at $z = \frac{1}{2}$. Our solution of Wilf's problem uses the same principle.

We shall use the following notation. The probability of the event $A$ is denoted by $\mathbb{P}(A)$, and the expected value of a random variable $X$ is denoted by $\mathbb{E}(X)$. The natural logarithm and base 2 logarithm are denoted by $\log n$ and $\log_2 n$, respectively.

To state the problem more precisely, suppose that a composition $\kappa$ is selected uniformly at random from the set of all $2^{n-1}$ compositions of $n$. Then out of the set of part sizes in $\kappa$, a part size $k$ is chosen uniformly at random. Let $A_n^{(m)}$ denote the event in which $k$ has multiplicity $m$. For example, inspection of the 16 partitions of 5 shown above yields

$$\mathbb{P}(A_5^{(1)}) = \frac{5}{8}, \qquad \mathbb{P}(A_5^{(2)}) = \frac{3}{16}, \qquad \mathbb{P}(A_5^{(3)}) = \frac{1}{8}, \qquad \mathbb{P}(A_5^{(5)}) = \frac{1}{16},$$

and otherwise $\mathbb{P}(A_5^{(m)}) = 0$. The object is to determine $\mathbb{P}(A_n^{(m)})$ asymptotically as $n \to \infty$. We shall find that $\mathbb{P}(A_n^{(m)})$ tends to 0 at the rate $1/\log n$. It then turns out that $\log n \cdot \mathbb{P}(A_n^{(m)})$ does not have a limit, but oscillates about the value $1/m$ as $n \to \infty$.

## 2 Results

The answer to Wilf's question is given in the following theorem, first proved in [3].

**Theorem 1.** *Let $A_n^{(m)}$ be the event in which a randomly selected part size in a randomly selected composition of $n$ has multiplicity $m$. Then*

$$\log n \cdot \mathbb{P}(A_n^{(m)}) = (1 + o(1)) \left( \frac{1}{m} + F(\{\log_2 n\}) \right), \qquad n \to \infty,$$

*where $\{a\} = a - \lfloor a \rfloor$ is the fractional part of $a$ and*

$$F(x) = \frac{2}{m!} Re \sum_{p=1}^{\infty} e^{-2\pi i p x}\, \Gamma\left(1 + i\,\frac{2\pi p}{\log 2}\right),$$

*with $\Gamma$ denoting the gamma function.*

Using well-known facts about the gamma function ($\Gamma(1+z) = z\Gamma(z)$ and $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$), we obtain

$$F(x) = \frac{2}{m!}\sum_{p=1}^{\infty}\left(\frac{p\alpha}{\sinh(p\alpha)}\right)^{1/2}\cos(2\pi p x - \phi_p),$$

where $\alpha = 2\pi^2/\log 2$ and $\phi_p$ is the argument of $\Gamma(1 + i\,2\pi p/\log 2)$. This series converges quite rapidly, and its sum may be approximated by the first term. But even the first term is quite small since $2(\alpha/\sinh\alpha)^{1/2} \approx 10^{-5}$. Thus for large $n$ one finds that $\log n \cdot \mathbb{P}(A_n^{(m)})$ is quite close to $1/m$, but there is a residual dependence on $\{\log_2 n\}$. In the treatment given here using generating functions and singularity analysis, the proof of Theorem 1 will reduce to a well-known calculation after we have established the appropriate sequence of lemmas.

Let $\kappa$ be a composition of $n$. Then $\mathcal{D}(\kappa)$ will denote the set of distinct part sizes in $\kappa$, and $\mathcal{M}_m(\kappa)$ will denote the set of part sizes of $\kappa$ that have multiplicity $m$.

**Lemma 1.** *For a random composition of $n$,*

$$\mathbb{P}(k \in \mathcal{M}_m(\kappa)) = \frac{1}{2^{n-1}}[z^n]\frac{z^{km}(1-z)^{m+1}}{(1-2z+z^k(1-z))^{m+1}}.$$

*Proof.* Let $G_k(z,w)$ be the two-variable generating function in which $[z^n w^m]G_k(z,w)$ is the number of compositions of $n$ in which $k$ has multiplicity $m$. To construct such a generating function, we first note that the contribution made by compositions with $p$ (not necessarily distinct) parts is

$$(z + z^2 + \cdots + wz^k + z^{k+1} + \cdots)^p = \left(\frac{z}{1-z} + (w-1)z^k\right)^p.$$

Thus

$$G_k(z,w) = \sum_{p=1}^{\infty}\left(\frac{z}{1-z} + (w-1)z^k\right)^p$$
$$= \frac{z+(w-1)z^k(1-z)}{1-2z-(w-1)z^k(1-z)}$$
$$= \frac{1-z}{1-2z-(w-1)z^k(1-z)} - 1$$
$$= \frac{1-z}{1-2z+z^k(1-z)-wz^k(1-z)} - 1.$$

Since there are $2^{n-1}$ compositions of $n$, we then have

$$\mathbb{P}(k \in \mathcal{M}_m(\kappa)) = \frac{1}{2^{n-1}}[z^n w^m]\frac{1-z}{1-2z+z^k(1-z)-wz^k(1-z)}$$
$$= \frac{1}{2^{n-1}}[z^n]\frac{z^{km}(1-z)^{m+1}}{(1-2z+z^k(1-z))^{m+1}},$$

as claimed. □

**Lemma 2.** *The polynomial* $1-2z+z^k(1-z)$ *has precisely one zero* $z=\rho$ *satisfying* $|z| \le 1$. *This zero is given by*

$$\rho = \frac{1}{2} + \frac{1}{2^{k+2}} + O\left(\frac{k}{2^{2k}}\right), \qquad k \to \infty.$$

*For all* $k \ge 1$,

$$\exp\left(-\frac{n}{2^k}\right) < \frac{1}{(2\rho)^n} < \exp\left(-\frac{n}{2^{k+2}}\right).$$

*Proof.* For the first part, simply observe that if $z = e^{i\theta}$ then $|1-2z|^2 = 5 - 4\cos\theta$ and $|1-z|^2 = 2 - 2\cos\theta$, so $|1-2z| > |1-z|$ for all $z$ with $|z| = 1$. Apply Rouché's theorem. To get the approximate location of $\rho$, write $\rho = \frac{1}{2} + \epsilon$ and substitute into $1 - 2\rho + \rho^k(1-\rho) = 0$. This yields

$$\epsilon = \frac{1}{2^{k+2}} + \frac{k-1}{2^{k+1}}\,\epsilon + \frac{k(k-3)}{2^k}\,\epsilon^2 + \cdots ,$$

and thus the stated result by iteration. Next we prove

$$\frac{1}{2-2^{-(k+1)}} < \rho < \frac{1}{2} + \frac{1}{2^{k+1}}.$$

A simple calculation shows that $Q(x) = 1 - 2x + x^k(1-x)$ decreases on $(0,1)$. Set $a = 1/(2-2^{-(k+1)})$ and $b = \frac{1}{2} + 2^{-(k+1)}$. Then we find that $Q(a) > 0$ and $Q(b) < 0$, so $a < \rho < b$. Then since $(1+x)^n < \exp(nx)$ for $x > -1$, we have

$$\frac{1}{(2\rho)^n} < \left(1 - \frac{1}{2^{k+2}}\right)^n < \exp\left(-\frac{n}{2^{k+2}}\right) \qquad \text{and} \qquad \frac{1}{(2\rho)^n} > \frac{1}{(1+2^{-k})^n} > \exp\left(-\frac{n}{2^k}\right).$$

□

We shall show that the number of distinct part sizes $|\mathcal{D}(\kappa)|$ of a random composition of $n$ satisfies $|\mathcal{D}| \sim \log_2 n$ with probability $o(1)$. The underlying probabilistic considerations are given in the following lemma.

**Lemma 3.** *Let $X = \sum I_j$ where $(I_j)$ are indicator random variables. Suppose that $\mathbb{P}(I_k) = p_{k,n} = p_k$. If $a$ and $b$ are chosen so that both $\sum_{j\le a}(1-p_j)$ and $\sum_{j>b} p_j$ are $o(1)$, then*

$$\mathbb{P}(a \le X \le b) \ge 1 - o(1).$$

*Proof.* For all $a \le b$ we have

$$\mathbb{P}(a \le X \le b) = 1 - \mathbb{P}(\{X < a\} \cup \{X > b\}) \ge 1 - \mathbb{P}(X < a) - \mathbb{P}(X > b).$$

Now, denoting for simplicity a set and its indicator by the same symbol,

$$\mathbb{P}(X > b) \le \mathbb{P}\left(\bigcup_{j>b} I_j\right) \le \sum_{j>b} \mathbb{P}(I_j) = \sum_{j>b} p_j,$$

and

$$\mathbb{P}(X < a) \le \mathbb{P}\left(\bigcup_{j\le a} I_j^c\right) \le \sum_{j\le a}(1 - \mathbb{P}(I_j)) = \sum_{j\le a}(1-p_j).$$

Hence, if $a$ and $b$ are chosen so that both $\sum_{j\le a}(1-p_j)$ and $\sum_{j>b} p_j$ are $o(1)$ we get

$$\mathbb{P}(a \le X \le b) \ge 1 - o(1).$$

□

**Lemma 4.** *Let $\kappa$ be a random composition of $n$. As $n \to \infty$ the number of distinct part sizes $|\mathcal{D}(\kappa)|$ satisfies $|\mathcal{D}(\kappa)| \sim \log_2 n$ with probability $1 - o(1)$.*

*Proof.* As a special case of Lemma 1, the probability that $k$ has multiplicity 0 in the random composition $\kappa$ is

$$\mathbb{P}(k \in \mathcal{M}_0(\kappa)) = \frac{1}{2^{n-1}}[z^n]\frac{1-z}{1-2z+z^k(1-z)}, \qquad n \geq 1.$$

Hence

$$\mathbb{P}(k \in \mathcal{D}(\kappa)) = 1 - \frac{1}{2^{n-1}}[z^n]\frac{1-z}{1-2z+z^k(1-z)}.$$

From Lemma 2, the rational function $(1-z)/(1-2z+z^k(1-z))$ is analytic for $|z| \leq 1$ except for a simple pole at $z = \rho \approx \frac{1}{2}$. The residue is $-(1-\rho)/(2+(k+1)\rho^k - k\rho^k)$. By standard arguments [6, §5.2],

$$[z^n]\frac{1-z}{1-2z+z^k(1-z)} = \left(\frac{1-\rho}{2+(k+1)\rho^k - k\rho^{k-1}}\right)\frac{1}{\rho^{n+1}} + O(1).$$

By Lemma 2,

$$\frac{1-\rho}{2+(k+1)\rho^k - k\rho^{k-1}} = \frac{1}{4}\left(1 + O\left(\frac{k}{2^k}\right)\right).$$

Hence we have

$$\frac{1}{2^{n-1}}[z^n]\frac{1-z}{1-2z+z^k(1-z)} = \frac{1}{(2\rho)^{n+1}}\left(1 + O\left(\frac{k}{2^k}\right)\right).$$

Using the general bound from Lemma 2

$$\exp\left(-\frac{n}{2^k}\right) < \frac{1}{(2\rho)^n} < \exp\left(-\frac{n}{2^{k+2}}\right),$$

we see that

$$\begin{aligned}
\mathbb{P}(k \in \mathcal{D}(\kappa)) &= 1 - \frac{1}{(2\rho)^{n+1}}\left(1 + O\left(\frac{k}{2^k}\right)\right) \\
&\leq 1 - \exp\left\{-\frac{n+1}{2^k}\right\}\left(1 + O\left(\frac{k}{2^k}\right)\right) \\
&\leq \frac{n+1}{2^k} + \exp\left\{-\frac{n+1}{2^k}\right\} \cdot O\left(\frac{k}{2^k}\right),
\end{aligned}$$

so that letting $b = \lfloor \log_2 n \rfloor + \log\log n$ we get

$$\sum_{k>b} \mathbb{P}(k \in \mathcal{D}(\kappa)) = O\left(\frac{1}{\log n}\right).$$

Similarly,

$$\mathbb{P}(k \in \mathcal{D}(\kappa)) \geq 1 - \exp\left\{-\frac{n+1}{2^{k+1}}\right\}\left(1 + O\left(\frac{k}{2^k}\right)\right),$$

that is

$$1 - \mathbb{P}(k \in \mathcal{D}(\kappa)) \leq \exp\left\{-\frac{n+1}{2^{k+1}}\right\}\left(1 + O\left(\frac{k}{2^k}\right)\right).$$

Consequently, for any positive $a$,

$$\begin{aligned}
\sum_{1\leq k\leq a} (1 - \mathbb{P}(k \in \mathcal{D}(\kappa))) &\leq C \sum_{1\leq k\leq a} \exp\left\{-\frac{n+1}{2^{k+1}}\right\} \\
&= C \sum_{0\leq r<a} \exp\left\{-2^r \frac{n+1}{2^{a+1}}\right\} \\
&\leq C \sum_{r\geq 0} \exp\left\{-(r+1)\frac{n+1}{2^{a+1}}\right\} \\
&= C \, \frac{\exp\{-(n+1)/2^{a+2}\}}{1 - \exp\{-(n+1)/2^{a+2}\}},
\end{aligned}$$

and thus

$$\sum_{1\leq k\leq a} (1 - \mathbb{P}(k \in \mathcal{D}(\kappa))) = O\left(\frac{1}{\log n}\right),$$

provided $a \leq \lfloor \log_2 n \rfloor - \log\log n$. Hence, by Lemma 3 applied to $I_k = \{k \in \mathcal{D}(\kappa)\}$, $|D(\kappa)| \sim \log_2 n$ with probability $1 - o(1)$. □

Given a random composition $\kappa$, the probability that a randomly selected part thereof has multiplicity $m$ is $|\mathcal{M}_m(\kappa)|/|\mathcal{D}(\kappa)|$. Lemma 4 greatly simplifies the basic problem. Since so doing amounts to the neglect of a set of compositions with total probability measure $o(1)$, we may assume that as $n \to \infty$ the randomly selected composition $\kappa$ satisfies $|\mathcal{D}(\kappa)| \sim \log_2 n$. Thus $\mathbb{P}(A_n^{(m)}) \sim \mathbb{E}(|\mathcal{M}_m|)/\log_2 n$.

Now we wish to study the asymptotic behavior of $\mathbb{P}(k \in \mathcal{M}_m(\kappa))$, with the aim of estimating

$$\mathbb{E}(|\mathcal{M}_m|) = \sum_k \mathbb{P}(k \in \mathcal{M}_m(\kappa)).$$

**Lemma 5.** *The expected value of $|\mathcal{M}_m|$ is given by*

$$\mathbb{E}(|\mathcal{M}_m|) = (1+o(1))\,\frac{n^m}{m!}\sum_k 2^{-km}\exp(-n/2^k).$$

*Proof.* As we found in Lemma 1, the relevant generating function is

$$G(z) = \frac{1}{2^{n-1}}\frac{P(z)}{Q^{m+1}(z)}, \quad \text{where} \quad P(z) = z^{km}(1-z)^{m+1},\ Q(z) = 1-2z+z^k(1-z).$$

Recall that $Q$ has a simple zero at $z=\rho\approx\frac{1}{2}$ and no other zeros in $\{z:\ |z|\le 1\}$. In a deleted neighborhood of $\rho$, we have the Laurent expansion

$$\frac{P(z)}{Q^{m+1}(z)} = \sum_{r=1}^{m+1}\frac{c_{-r}}{(z-\rho)^r} + \sum_{s=0}^{\infty} c_s(z-\rho)^s.$$

The asymptotic behavior of $[z^n]P(z)/Q^{m+1}(z)$ is governed by the principal part, more specifically by the $r=m+1$ term. In view of

$$[z^n](1-z)^{-\alpha} = \binom{n+\alpha-1}{n},$$

a simple calculation gives

$$\mathbb{P}(k\in\mathcal{M}_m(\kappa)) = \binom{n+m}{m}\frac{2P(\rho)}{(-\rho Q'(\rho))^{m+1}}\,\frac{1}{(2\rho)^n}\left(1+O\left(\frac{1}{n}\right)\right).$$

Set

$$q(n) = \frac{\log n - \log\log n - \log(4(m+1))}{\log 2},$$

and note that if $k < q(n)$ then $2^k < n/(4(m+1)\log n)$, so

$$\frac{n^m}{(2\rho)^n} < n^m \exp\left(-\frac{n}{4\cdot 2^k}\right) < n^m \exp(-(m+1)\log n) = \frac{1}{n}.$$

Hence we have

$$\sum_{k\le q(n)} \mathbb{P}(k \in \mathcal{M}_m(\kappa)) = O\left(\frac{\log n}{n}\right), \qquad n\to\infty.$$

In view of the fact just noted, in estimating $\sum_k \mathbb{P}(k \in \mathcal{M}_m)$, we can limit ourselves to cases where $k > q(n)$. In that case

$$\frac{1}{(2\rho)^n} = \exp\left(-\frac{n}{2^{k+1}}\right)\left(1 + O\left(\frac{(\log n)^3}{n^2}\right)\right), \qquad n\to\infty.$$

Now

$$P(\rho) = 2^{-km}2^{-(m+1)}\left(1+O\left(\frac{\log n}{n}\right)\right) \qquad \text{and} \qquad \rho\, Q'(\rho) = -1 + O\left(\frac{\log n}{n}\right),$$

so

$$\mathbb{P}(k \in \mathcal{M}_m(\kappa)) = \left(1+O\left(\frac{\log n}{n}\right)\right)\frac{n^m}{m!}\, 2^{-(k+1)m}\exp\left(-\frac{n}{2^{k+1}}\right).$$

It is now evident that the contribution to the sum $\sum_k \mathbb{P}(k \in \mathcal{M}_m(\kappa))$ from those terms with $k > \log_2 n + \log\log n$ is $o(1)$, so there are $O(\log\log n)$ terms in the sum that make a nontrivial contribution. Thus the bound on the error for an individual term suffices to give the correct asymptotic result for the sum. Replacing $k+1$ by $k$ in the sum, we have the stated result. □

*Proof of Theorem 1.* The computational problem that remains is the asymptotic evaluation of

$$\frac{n^m}{m!}\sum_{k=1}^{\infty} 2^{-km}\exp(-n/2^k).$$

Problems of this kind occur frequently in probability theory and the analysis of algorithms, and now there are different methods available for their study, and these methods are described in several excellent references [2, chapter 7]. We sketch an approach due to N. G. de Bruijn, which is described in [4, pp. 131–134] and elsewhere. A special case ($m = 1$) of the above sum is treated in [5]. The starting point is Mellin's formula

$$\exp(-w) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} w^{-z}\Gamma(z)\,dz, \qquad w, \sigma > 0.$$

Substituting this representation (with $\sigma = m - \frac{1}{2}$) for $\exp(-n/2^k)$ and using uniform convergence, one obtains

$$\sum_{k=1}^{\infty} 2^{-km}\exp(-n/2^k) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{n^{-z}\,\Gamma(z)}{2^{m-z}-1}\,dz.$$

Then by the residue theorem,

$$\frac{n^m}{m!}\sum_{k=1}^{\infty} 2^{-km}\exp(-n/2^k)$$
$$= \frac{1}{m!\log 2}\left\{(m-1)! + 2\mathrm{Re}\sum_{p=1}^{\infty} e^{-2\pi i p\log_2 n}\,\Gamma\left(m + i\,\frac{2\pi p}{\log 2}\right)\right\}(1+o(1)).$$

The stated result for $\log n \cdot \mathbb{P}(A_n^{(m)})$ follows. □

# References


[1] S. Corteel, B. Pittel, C. D. Savage, H. S. Wilf, On the multiplicity of parts in a random partition, *Random Structures Algorithms* **14** (1999), pp. 185–197.

[2] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, to be published. In part available at <http://pauillac.inria.fr/algo/flajolet/Publications/publist.html>.

[3] P. Hitczenko, C. D. Savage, On the multiplicity of parts in a random composition of a large integer, preprint. Available at <http://www.csc.ncsu.edu/faculty/savage/>.

[4] D. E. Knuth, *The Art of Computer Programming. Vol. III: Sorting and Searching*, Addison-Wesley, Reading, Massachusetts, 1973.

[5] C. C. Rousseau, Solution of Problem 322, *Canad. Math. Bull* **26** (1983), pp. 375–377.

[6] H. S. Wilf, *Generatingfunctionology*, 2nd ed., Academic Press, San Diego, 1994. Available at <http://www.cis.upenn.edu/∼wilf/>.