Compositions and samples of geometric random variables with constrained multiplicities

Margaret Archibald[1] and Arnold Knopfmacher[2†] and Toufik Mansour[3]

[1]*Laboratory of Foundational Aspects of Computer Science, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa*
`margaret.archibald@uct.ac.za`
[2]*The John Knopfmacher Centre for Applicable Analysis and Number Theory, Department of Mathematics, University of the Witwatersrand, P. O. Wits, 2050, Johannesburg, South Africa*
`arnold.knopfmacher@wits.ac.za`
[3]*Department of Mathematics, University of Haifa, 31905 Haifa, Israel*
`toufik@math.haifa.ac.il`

Abstract. We investigate the probability that a random composition (ordered partition) of the positive integer n has no parts occurring exactly j times, where j belongs to a specified finite *'forbidden set' A* of multiplicities. This probability is also studied in the related case of samples $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ of independent, identically distributed random variables with a geometric distribution.

Résumé. Nous examinons la probabilité qu'une composition faite au hasard (une partition ordonnée) du nombre entier positif n n'a pas de partie qui arrivent exactement j fois, où j appartient à une série interdite, finie et spécifié A de multiplicités. Cette probabilité est aussi étudiée dans le cas des suites $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ de variables aléatoires identiquement distribués et indépendants avec une distribution géométrique.

Keywords: compositions, generating functions, geometric random variable, Mellin transform, Poisson transform, multiplicity

1   Introduction

In this paper we derive generating functions for random compositions (ordered partitions) of a positive integer n in which no parts occur exactly j times, where j belongs to a specified finite *'forbidden set' A* of multiplicities. For notational convenience we shall refer to such compositions as being *'A-avoiding'*. We go on to find the probabilities that compositions and samples of geometric random variables are A-avoiding.

As a simple example of a forbidden set, we may wish to consider a sample where none of the n elements occur exactly a times. In this case $A = \{a\}$. Another example is when a letter can occur only a times or more (or not at all), i.e., $A = \{1, 2, \ldots, a-1\}$, for $a \geq 2$. Note that we do not allow 0 in the forbidden set.

Previously in [6, 12], geometric samples with the multiplicity constraint that certain values must occur at least once were studied. These were called 'gap-free' and 'complete' samples. A gap-free sample has elements whose values form an interval, namely if elements 2 and 6 are in the sample, then so are 3, 4 and 5. A complete sample is gap-free with minimal element 1.

In this paper we drop the 'interval' restriction, hence no value 0 in our forbidden sets. Here we are more interested in the number of times the elements do occur than in the values of the elements. However, in Section 2, the idea of forbidden sets is generalised even further when we allocate each value a different forbidden set. For example, one could provide the restriction that the value 2 is not allowed to occur once, but that the number of times that 5 can occur is anything except $2, 3$ or 6 times. We denote the forbidden set for the value i by A_i, so in this case, we have $A_2 = \{1\}$ and $A_5 = \{2, 3, 6\}$.

The paper begins with a discussion on compositions (Section 2), where explicit generating functions are derived for A-avoiding compositions and particular forbidden sets are highlighted. In Section 3, the link between compositions and samples of geometric random variables is explained. Section 4 is devoted to geometric samples, and Theorem 2 gives the probability that a geometric sample is A-avoiding, along with some further examples of specific forbidden sets. Finally in Section 5, we state the result for compositions - i.e., the probability that a random composition of n is A-avoiding. Some of the longer proofs, in particular, the proof of Theorem 2 in Section 4 will be detailed in the full version of this paper.

2   Compositions

In this section we investigate the generating function for the number of A-avoiding compositions of n, that is the number of compositions of n such that each part does not appear exactly j times, where $j \in A$. We then go on to generalise this by allowing a different forbidden set for each value, as described in the introduction.

Let $C_{A,d}(x; m)$ be the generating function for the number of A-avoiding compositions of n with exactly m parts from the set $[d] = \{1, 2, \ldots, d\}$. If σ is any A-avoiding composition with m parts in $[d]$, then σ contains the part d exactly j times with $j \notin A$ and $0 \leq j \leq m$. Deleting the parts that equal to d from σ we get an A-avoiding composition σ' of $m - j$ parts in $[d - 1]$. Thus, rewriting the above rule in terms of generating functions we get that

$$C_{A,d}(x; m) = \sum_{\substack{j=0 \\ j \notin A}}^{m} \binom{m}{j} x^{dj} C_{A,d-1}(x; m - j),$$

which is equivalent to

$$\frac{C_{A,d}(x; m)}{m!} = \sum_{\substack{j=0 \\ j \notin A}}^{m} \frac{x^{dj}}{j!} \frac{C_{A,d-1}(x; m - j)}{(m - j)!}. \tag{1}$$

We denote the exponential generating function for the sequence $C_{A,d}(x;m)$ by $C_{A,d}(x,y)$, that is,

$$C_{A,d}(x,y) = \sum_{m \geq 0} C_{A,d}(x;m) \frac{y^m}{m!}.$$

Therefore, the recurrence in (1) can be written as

$$C_{A,d}(x,y) = C_{A,d-1}(x,y) \left(e^{x^d y} - \sum_{j \in A} \frac{x^{dj} y^j}{j!} \right),$$

which implies that

$$C_{A,d}(x,y) = \prod_{k=1}^{d} \left(e^{x^k y} - \sum_{j \in A} \frac{x^{kj} y^j}{j!} \right),$$

for all $d \geq 1$. Hence, we can state the following result.

Proposition 1 *The generating function $C_A(x,y) = \sum_{m \geq 0} C_A(x;m) \frac{y^m}{m!}$ is given by*

$$C_A(x,y) = \prod_{k \geq 1} \left(e^{x^k y} - \sum_{j \in A} \frac{x^{kj} y^j}{j!} \right),$$

where $C_A(x;m)$ is the generating function for the number of A-avoiding compositions of n with exactly m parts in $\mathbb{N}$.

Let $C_A(n,m)$ be the number of A-avoiding compositions of n with m parts and $C_A(n) = \sum_{m \geq 1} C_A(n,m)$ be the number of A-avoiding compositions of n.

Corollary 1 *The generating function $C_A(x) = \sum_{n \geq 0} C_A(n) x^n$ is given by*

$$C_A(x) = \int_0^\infty e^{-y} \prod_{k \geq 1} \left(e^{x^k y} - \sum_{j \in A} \frac{x^{kj} y^j}{j!} \right) dy.$$

Proof: We use the fact that $\int_0^\infty e^{-y} y^m dy = m!$. Then

$$\int_0^\infty e^{-y} C_A(x,y) dy = \sum_{n \geq 0} x^n \sum_{m \geq 0} \frac{C_A(n,m)}{m!} \int_0^\infty y^m e^{-y} dy = \sum_{n \geq 0} C_A(n) x^n.$$

$\square$

Example 1 *Let $A_i = \{1\}$ for all i, then the above proposition gives that*

$$C_{\{1\}}(x,y) = \prod_{k \geq 1} (e^{x^k y} - x^k y).$$

and Corollary 1 *gives*

$$C_{\{1\}}(x) = \int_0^\infty e^{-y} \prod_{k \geq 1} \left(e^{x^k y} - x^k y \right) dy.$$

Similar techniques as before show the following general result.

Proposition 2 *The generating function $D_{A_1,A_2,\ldots}(x,y) = \sum_{m\geq0} D_{A_1,A_2,\ldots}(x;m)\frac{y^m}{m!}$ is given by*

$$D_{A_1,A_2,\ldots}(x,y) = \prod_{k\geq1}\left(e^{x^k y} - \sum_{j\in A_k}\frac{x^{kj}y^j}{j!}\right),$$

where $D_{A_1,A_2,\ldots}(x;m)$ is the generating function for the number of compositions σ of n with exactly m parts in $\mathbb{N}$ such that if σ contains the part i exactly d_i times, then $d_i \notin A_i$. Furthermore,

$$D_{A_1,A_2,\ldots}(x) = \int_0^\infty e^{-y}\prod_{k\geq1}\left(e^{x^k y} - \sum_{j\in A_k}\frac{x^{kj}y^j}{j!}\right)dy.$$

Example 2 *For instance, let $A_1 = \{1\}$ and $A_i = \emptyset$ for $i \geq 2$, then the above proposition gives that*

$$F(x,y) = D_{\{1\},\emptyset,\emptyset,\ldots}(x,y) = (e^{xy} - xy)e^{\frac{x^2 y}{1-x}}.$$

If we expand $F(x,y)$ as a power series at $x = y = 0$, then we obtain that

$$F(x,y) = \sum_{j\geq0}\frac{x^j y^j}{j!(1-x)^j} - xy\sum_{j\geq0}\frac{x^{2j}y^j}{j!(1-x)^j},$$

which implies that

$$D_{\{1\},\emptyset,\emptyset,\ldots}(x;m) = \frac{x^m}{(1-x)^m} - m\frac{x^{2m-1}}{(1-x)^{m-1}}.$$

Summing over all $m \geq 0$, we get that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part 1 in σ does not equal 1 is given by

$$\frac{1-x}{1-2x} - \frac{x(1-x)^2}{(1-x-x^2)^2}.$$

Note that it is not hard to generalize the above enumeration to obtain that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part 1 in σ does not equal ℓ is given by

$$\frac{1-x}{1-2x} - \ell!\frac{x^\ell(1-x)^{\ell+1}}{(1-x-x^2)^{\ell+1}}.$$

Example 3 *For instance, let $A_1 = A_2 = \{1\}$ and $A_i = \emptyset$ for $i \geq 3$, then the above proposition gives that*

$$G(x,y) = D_{\{1\},\{1\},\emptyset,\emptyset,\ldots}(x,y) = (e^{xy} - xy)(e^{x^2 y} - x^2 y)e^{\frac{x^3 y}{1-x}}.$$

If we expand $G(x,y)$ as a power series at $x = y = 0$, then we find that

$$D_{\{1\},\{1\},\emptyset,\emptyset,\ldots}(x;m)$$
$$= \frac{x}{(1-x)^m} - m\frac{x^{m+1}(1-x+x^2)^{m-1}}{(1-x)^{m-1}} - m\frac{x^{2m-1}}{(1-x)^{m-1}} - m(m-1)\frac{x^{3m-3}}{(1-x)^{m-2}}.$$

Summing over all $m \geq 0$, we get that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part i, $i = 1, 2$, in σ does not equal 1 is given by

$$\frac{1-x}{1-2x} - \frac{x(1-x)^2}{(1-x-x^2)^2} - \frac{x^2(1-x)^2}{(1-2x+x^2-x^3)^2} + \frac{2x^3(1-x)^3}{(1-x-x^3)^3}.$$

Theorem 1 *Fix $a \in \mathbb{N}$. Let $A_i = \{a\}$ for all $i = 1, 2, \ldots, \ell$ and $A_{\ell+i} = \emptyset$ for all $i \geq 1$. The ordinary generating function for the number of compositions π of n such that π does not contain part i exactly a times for all $i = 1, 2, \ldots, \ell$ is given by*

$$\sum_{m \geq 0} D_{A_1, A_2, \ldots}(x; m) = \frac{1-x}{1-2x} + \sum_{j=1}^{\ell} \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq \ell} (-1)^j \frac{j! x^{a \sum_{k=1}^{j} i_k}}{\left(1 - \frac{x}{1-x} + \sum_{k=1}^{j} x^{a i_k}\right)^{j+1}},$$

The proof of this result will be given in the full version of this paper. From the theorem we can deduce the following result.

Corollary 2 *The ordinary generating function for the number of $\{a\}$-avoiding compositions of n is given by*

$$\frac{1-x}{1-2x} + \sum_{j \geq 1} \sum_{B \subseteq \mathbb{N}, |B|=j} (-1)^j \frac{(aj)! (x^a/a!)^{\sum_{b \in B} b}}{\left(1 - \frac{x}{1-x} + \sum_{b \in B} x^b\right)^{aj+1}}.$$

Even in this simple case of $A = \{a\}$ it does not seem easy to find asymptotic estimates for the coefficients from the generating functions appearing in either Corollary 1 or Corollary 2. Instead we will exploit the correspondence between compositions and geometric random variables of parameter $p = 1/2$, as detailed in the next section.

3 Reduction of compositions to geometric samples

In order to derive asymptotic estimates, it will be convenient to adopt a probabilistic viewpoint. That is, rather than think of the proportion of A-avoiding compositions we will equip the set of all compositions of n with the uniform probability measure and will be interested in the probability that a randomly chosen composition of n is A-avoiding. In that setting, compositions of n are closely related to the special case for geometric random variables when $p = 1/2$, as shown in [7, 8] and again in this section.

The starting point for reducing compositions to samples of geometric random variables is the following representation of compositions of n (see e.g., [2]). Consider sequences of n black and white dots subject to the following constraints

(i) the last dot is always black

(ii) each of the remaining $n - 1$ dots is black or white.

Then there is a 1-1 correspondence between all such sequences and compositions of n. Namely, part sizes in a composition correspond to "waiting times" for occurrences of black dots. For example, the sequence

represents the composition of 12 into parts $(1, 3, 2, 1, 1, 2, 2)$. As discussed e.g. in [7, 8] this leads to the following representation of random compositions. Let $p = 1/2$ and define

$$\tau = \tau_n = \inf\{k \geq 1 : \Gamma_1 + \Gamma_2 + \cdots + \Gamma_k \geq n\}.$$

Then a randomly chosen composition κ of n has distribution given by

$$\kappa = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{\tau-1}, n - \sum_{j=1}^{\tau-1} \Gamma_j) := (\tilde{\Gamma}_1, \tilde{\Gamma}_2, \ldots, \tilde{\Gamma}_\tau).$$

Furthermore, τ has known distribution, namely,

$$\tau \overset{d}{=} 1 + \mathrm{Bin}(n - 1, \frac{1}{2}),$$

where $\mathrm{Bin}(m, p)$ denotes a binomial random variable with parameters m and p and $\overset{d}{=}$ stands for equality in distribution. Hence, τ is heavily concentrated around its mean. Specifically, since $\mathrm{var}(\tau) = \mathrm{var}(\mathrm{Bin}(n - 1, 1/2)) = (n - 1)/4$, for every $t > 0$ we have (see [1, Section A.1])

$$\mathbb{P}(|\tau - \mathsf{E}\tau| \geq t) \leq 2 \exp\{-\frac{2t^2}{n - 1}\}.$$

In particular, for $t_n \sim \sqrt{cn \ln n}$,

$$\mathbb{P}(|\tau - \mathsf{E}\tau| \geq t_n) = O\left(\frac{1}{n^{2c}}\right),$$

for any $c > 0$.

Let $\mathbb{P}(\kappa \in \mathcal{C})$ be the probability that a random composition is A-avoiding. We proceed by series of refinements exactly as in [6]. Set m_n^- to be

$$m_n^- = \left\lfloor \frac{n + 1}{2} - t_n \right\rfloor.$$

As shown in [6], with overwhelming probability, κ is A-avoiding if and only if the first m_n^- of its parts are A-avoiding. In [6] the property considered is "complete" rather than "A-avoiding", but the arguments remain unchanged.

Ultimately we obtain, exactly as in [6],

$$\mathbb{P}(\kappa \in \mathcal{C}) = \mathbb{P}((\Gamma_1, \ldots, \Gamma_{m_n^-}) \in \mathcal{C}) + O\left(\frac{\ln^{3/2} n}{\sqrt{n}}\right),$$

thereby reducing the problem to samples of geometric random variables.

4 Geometric random variables

Following the discussion in Section 3 above it is natural to start the investigation for the probability that a composition is A-avoiding with samples of geometric random variables with arbitrary parameter p, where $0 < p < 1$. There is now an extensive literature on the combinatorics of geometric random variables and its applications in Computer Science which includes [3, 5, 6, 11, 12, 13, 14].

Let $\Gamma = (\Gamma_1, \Gamma_2, \dots, \Gamma_n)$ be a sample of independent identically distributed (i.i.d.) geometric random variables with parameter p, that is, $\mathbb{P}(\Gamma_i = k) = pq^{k-1}$, with $p + q = 1$, where $k = 1, 2, \dots$ and $i = 1, 2, \dots, n$. We shall restrict the multiplicity of elements in a sample of length n by prohibiting any occurrences of exactly j entries of a given size, for j a natural number belonging to a specified finite set of excluded numbers A, the forbidden set. We also call such a random sample of n geometric variables A-avoiding.

The method used in [6] can be applied to the problem described above. We start with a recursion for the probabilities that depends on the set A and then use Poissonisation and Mellin transforms followed by de-Poissonisation to obtain our asymptotic estimates.

Using this approach, the following main result for geometric random variables will be proved in the full version of this paper. We define $\chi_k := \frac{2k\pi i}{\ln(1/q)}$.

Theorem 2 *Let A be any finite set of positive integers. The probability p_n that a geometric sample of length n has no letter appearing with multiplicity j, for any $j \in A$ is (asymptotically as $n \to \infty$)*

$$p_n = 1 - \frac{T^*(0)}{\ln(1/q)} - \delta\big(\log_{1/q}(n/q)\big) + O(n^{-1}),$$

with

$$T^*(0) = \sum_{j \in A} p^j \sum_{n \geq 0} p_n q^n \frac{1}{n+j} \binom{n+j}{j} \tag{2}$$

and

$$\delta(x) = \frac{1}{\ln(1/q)} \sum_{k \neq 0} T^*(\chi_k) e^{-2k\pi i x}$$

where

$$T^*(\chi_k) = \sum_{j \in A} \frac{p^j}{j!} \sum_{n \geq 0} p_n \frac{q^n}{n!} \Gamma(n+j+\chi_k), \quad for \quad k \in \mathbb{Z} \setminus \{0\}. \tag{3}$$

Here $\delta(x)$ is a periodic function of x with period 1, mean 0 and small amplitude.

The corresponding result for compositions of n is given in Section 5.

4.1 *Examples of finite forbidden sets A*

In the sections above we mentioned a few specific examples that would satisfy this definition of the forbidden set. Here we simplify the $T^*(0)$ and $T^*(\chi_k)$ formulae from Theorem 2 for a few specific cases. The simplest case for A is a singleton set consisting of one value a. If $A = \{a\}$, then

$$T^*(0) = p^a \sum_{n \geq 0} p_n q^n \frac{1}{n+a} \binom{n+a}{a} \quad \text{and} \quad T^*(\chi_k) = \frac{p^a}{a!} \sum_{n \geq 0} p_n \frac{q^n}{n!} \Gamma(n+a+\chi_k).$$
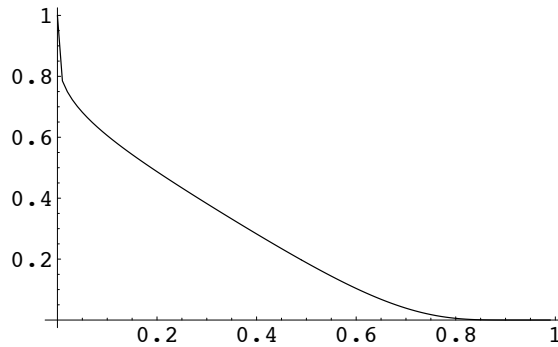
If we consider the case where $A = \{1, \ldots, a-1\}$, then

$$T^*(0) = \sum_{j=1}^{a-1} p^j \sum_{n \geq 0} p_n q^n \frac{1}{n+j} \binom{n+j}{j} \quad \text{and} \quad T^*(\chi_k) = \sum_{j=1}^{a-1} \frac{p^j}{j!} \sum_{n \geq 0} p_n \frac{q^n}{n!} \Gamma(n+j+\chi_k).$$

In particular if we want the probability that no element occurs exactly once (all elements must occur at least twice if they occur at all), we have a main term for p_n of
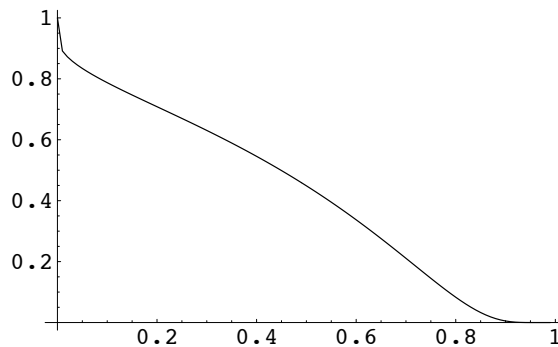
$$1 - \frac{p}{\ln(1/q)} \sum_{n \geq 0} p_n q^n.$$

This main term is plotted as a function of p in Figure 1.

Fig. 1: *Plot of the non-oscillating limit term for p_n for $0 \leq q \leq 1$.*

The corresponding picture for the probability that no element occurs exactly twice is given in Figure 2.

Fig. 2: *Plot of the non-oscillating limit term for p_n for $0 \leq q \leq 1$.*

In spite of what the Figures 1 and 2 tend to suggest for q near 1, the main term here is strictly greater

than zero for every $0 < p < 1$ as

$$T^*(0) = p^a \sum_{n \geq 0} p_n q^n \frac{1}{n+a} \binom{n+a}{a}$$

$$\leq p^a \sum_{n \geq 0} q^n \frac{1}{n+a} \binom{n+a}{a} = p^a \frac{(1-q)^{-a}}{a}$$

$$\leq 1 < \ln(1/q).$$

We observe also that the sequences (p_n) in this section do not have a limit, but exhibit small oscillations where both the period and amplitude of the oscillations depend on p. Such oscillations are almost ubiquitous in problems solved using Mellin transform techniques. For example, Figures 3 and 4 (Section 5) show these oscillations in the case that no element occurs exactly once (twice) when $p = 1/2$.

5   Compositions revisited

From Section 3, we conclude that probabilities for compositions can be reduced to probabilities for samples of geometric random variables. This result together with the special case $p = q = \frac{1}{2}$ in Theorem 2 leads to the following corollary.

Corollary 3 *Let A be any finite set of positive integers. The probability p_n that a composition of n has no part appearing with multiplicity j, for any $j \in A$ is (asymptotically as $n \to \infty$)*

$$p_n = 1 - \frac{T^*(0)}{\ln 2} - \delta(\log_2 n) + O\left(\frac{\ln^{3/2} n}{\sqrt{n}}\right),$$

with

$$T^*(0) = \sum_{j \in A} \left(\frac{1}{2}\right)^j \sum_{n \geq 0} p_n \left(\frac{1}{2}\right)^n \frac{1}{n+j} \binom{n+j}{j} \tag{4}$$

and

$$\delta(x) = \frac{1}{\ln 2} \sum_{k \neq 0} T^*(\chi_k) e^{-2k\pi i x}$$

where $\chi_k = \frac{2k\pi i}{\ln 2}$ and

$$T^*(\chi_k) = \sum_{j \in A} \frac{1}{j!} \left(\frac{1}{2}\right)^j \sum_{n \geq 0} \frac{p_n}{n!} \left(\frac{1}{2}\right)^n \Gamma(n+j+\chi_k), \quad for \quad k \in \mathbb{Z}\backslash\{0\}. \tag{5}$$

As in Theorem 2, $\delta(x)$ is a periodic function of x with period 1, mean 0 and small amplitude. In Figures 3 and 4 we plot the probabilities that no element occurs exactly once (twice) in compositions of n.

In particular, we see that the probabilities p_n that a composition is A-avoiding, do not converge to a limit as $n \to \infty$, but instead oscillate around the value $1 - \frac{T^*(0)}{\ln 2}$.
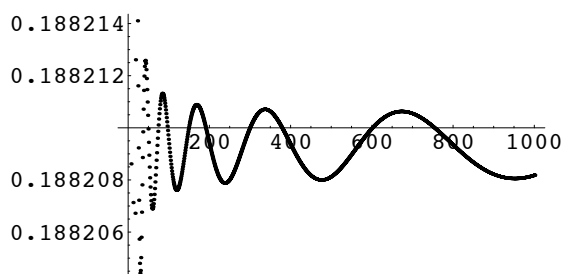
Fig. 3: *Plot of p_n for $b = 1$ and $1 \leq n \leq 1000$.*
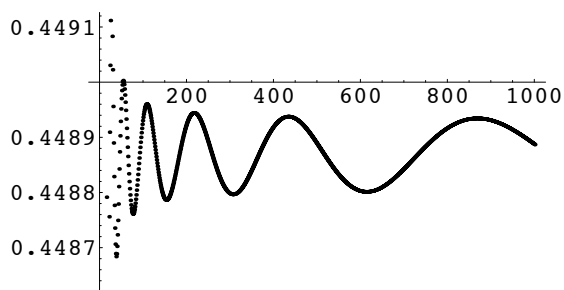
Fig. 4: *Plot of p_n for $b = 2$ and $1 \leq n \leq 1000$.*

Acknowledgements

References

[1] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley, 2000.

[2] G. E. Andrews, The Theory of Partitions, Addison – Wesley, Reading, MA, 1976.

[3] L. Devroye, A limit theory for random skip lists, *Ann. Appl. Probab.* **2** (1992) 597–609.

[4] P. Flajolet, X. Gourdon and P. Dumas, Mellin Transforms and Asymptotics: Harmonic Sums, *Theoret. Comput. Sci.* **144** (1995) 3–58.

[5] P. Flajolet and G.N. Martin, Probabilistic Counting Algorithms for Data Base Applications, *J. Comput. Syst. Sci.* **31** (1985) 182–209.

[6] P. Hitczenko and A. Knopfmacher, Gap-free compositions and gap-free samples of geometric random variables, *Discrete Math.* **294** (2005) 225–239.

[7] P. Hitczenko and G. Louchard, Distinctness of compositions of an integer: a probabilistic analysis, *Random Struct. Alg.* **19** (2001) 407–437.

[8] P. Hitczenko and C. D. Savage, On the multiplicity of parts in a random composition of a large integer, *SIAM J. Discrete Math.* **18** (2004) 418–435.

[9] P. Jacquet and W. Szpankowski, Analytical de–Poissonization and its applications, *Theoret. Comput. Sci.* **201** (1998) 1–62.

[10] S. Janson and W. Szpankowski, Analysis of the asymmetric leader election algorithm, *Electr. J. Combin.* **4** (1997) #R17.

[11] P. Kirschenhofer and H. Prodinger, On the Analysis of Probabilistic Counting, *Lecture Notes in Mathematics* **1452** (1990) 117–120.

[12] G. Louchard and H. Prodinger, The number of gaps in sequences of geometrically distributed random variables, *Discrete Math.* **308** (2008) 1538–1562.

[13] H. Prodinger, Combinatorics of geometrically distributed random variables: Left-to-right maxima, *Discrete Math.* **153** (1996) 253–270.

[14] W. Pugh, Skip lists: a probabilistic alternative to balanced trees, *Comm. ACM* **33** (1990) 668–676.

[15] W. Szpankowski, Average Case Analysis of Algorithms on Sequences, John Wiley and Sons, New York, 2001.