J. Math. Anal. Appl. 416 (2014) 672–682

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Set partition asymptotics and a conjecture of Gould and Quaintance

Walaa Asakly [a], Aubrey Blecher [b], Charlotte Brennan [c], Arnold Knopfmacher [c,*], Toufik Mansour [a], Stephan Wagner [d]

[a] Department of Mathematics, University of Haifa, 3498838 Haifa, Israel
[b] School of Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa
[c] The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa
[d] Department of Mathematical Sciences, Stellenbosch University, Matieland 7602, South Africa

ARTICLE INFO

Article history:
Received 25 November 2013
Available online 6 March 2014
Submitted by U. Stadtmueller

Keywords:
Set partitions
Generating functions
Asymptotics
Bell numbers

ABSTRACT

The main result of this paper is the generalization and proof of a conjecture by Gould and Quaintance on the asymptotic behavior of certain sequences related to the Bell numbers. Thereafter we show some applications of the main theorem to statistics of partitions of a finite set S, i.e., collections $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ of non-empty disjoint subsets of S such that $\bigcup_{i=1}^k \mathcal{B}_i = S$, as well as to certain classes of partitions of $[n]$.

1. Introduction

A partition π of a set S is a collection $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ of non-empty disjoint subsets of S such that $\bigcup_{i=1}^k \mathcal{B}_i = S$ (see for example [6]). The $\mathcal{B}_i$'s are called *blocks*, and the size $|\mathcal{B}|$ of a block $\mathcal{B}$ is the number of elements in $\mathcal{B}$. We assume that $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ are listed in increasing order of their minimal elements, that is, $\min \mathcal{B}_1 < \min \mathcal{B}_2 < \cdots < \min \mathcal{B}_k$. This is known as the canonical representation. The collection of all set partitions of S is denoted by $\mathcal{P}(S)$. We define $[n]$ to be the set $\{1, 2, \ldots, n\}$. For example, the canonical representations of the five partitions of $[3]$ are

$$\{1,2,3\}; \quad \{1,2\}, \{3\}; \quad \{1,3\}, \{2\}; \quad \{1\}, \{2,3\} \quad \text{and} \quad \{1\}, \{2\}, \{3\}.$$

* Corresponding author.
E-mail addresses: walaa_asakly@hotmail.com (W. Asakly), Aubrey.Blecher@wits.ac.za (A. Blecher), Charlotte.Brennan@wits.ac.za (C. Brennan), Arnold.Knopfmacher@wits.ac.za (A. Knopfmacher), tmansour@univ.haifa.ac.il (T. Mansour), swagner@sun.ac.za (S. Wagner).

http://dx.doi.org/10.1016/j.jmaa.2014.02.061
0022-247X/

Let A be a (totally ordered) alphabet of k letters. A word w of size n over the alphabet A is an element of A^n. In the case $A = [k]$, an element of A^n is called k-ary word of size n. For example, the 2-ary words of size 3 are 111, 112, 121, 122, 211, 212, 221, and 222. In the word form of the set partition canonical representation, we indicate for each integer the block in which it occurs. Thus a partition into k blocks would be represented by a word $\pi = \pi_1\pi_2\cdots\pi_n$, where for $1 \leqslant j \leqslant n, \pi_j \in [k]$ and $\bigcup_{i=1}^{n}\{\pi_i\} = [k]$, and π_j indicates that $j \in \mathcal{B}_{\pi_j}$. For example, the above set partitions of [3] in canonical representation are respectively 111, 112, 121, 122, and 123. We denote the set of all partitions of $[n]$ by $\mathcal{P}([n])$, and the number of all set partitions of $[n]$ by $B_n = |\mathcal{P}([n])|$, with $B_0 = 1$ for the empty set. The B_n are known as the Bell numbers. Their sequence starts with $1, 1, 2, 5, 15, 52, 203, 877, \ldots$ for $n = 0, 1, 2, \ldots$ (see [7, Section 1.4], A000110 in [8]).

It is a fact (see [6]) that the canonical representations of all set partitions of $[n]$ are precisely the words $\pi = \pi_1\pi_2\cdots\pi_n$ such that $\pi_1 = 1$, and if $i < j$ then the first occurrence of the letter i precedes the first occurrence of j. Such words are known as restricted growth functions.

Set partitions (or restricted growth functions) have been extensively studied in the literature, see [6] and references therein. The exponential generating function for set partitions is given by

$$e^{e^x-1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!},$$

and the Bell numbers B_n satisfy the binomial recurrence

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k.$$

Let us define a sequence A_n by the following exponential generating function, as in [4]:

$$\sum_{n=0}^{\infty} A_n \frac{x^n}{n!} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!} \int_0^x e^{1-e^t}\, dt.$$

The sequence A_n, whose first terms are $0, 1, 1, 3, 9, 31, 121, 523, \ldots$, occurs in various contexts (see entry A040027 in [8]). In their paper [4], Gould and Quaintance conjectured that

$$\lim_{n\to\infty} \frac{A_n}{B_n} = \int_0^\infty \frac{e^{-u}}{1+u}\, du = \int_0^\infty e^{1-e^x}\, dx \approx 0.5963473623\ldots.$$

The main result of this paper is a generalisation of their conjecture, which can be formulated as follows.

Theorem 1. *Let g be an entire function in the complex plane that satisfies the growth condition $g(z) = O(e^{e^{(1-\epsilon)|z|}})$ for some $\epsilon > 0$ as $|z| \to \infty$, and let*

$$F(x) = e^{e^x} \int_0^x e^{-e^t} g(t)\, dt$$

be the solution to the differential equation

$$F'(x) = e^x F(x) + g(x)$$

with $F(0)=0$. *If* $F(x)=\sum_{n=0}^{\infty}\frac{a_n}{n!}x^n$ *and* $g(x)=\sum_{n=0}^{\infty}\frac{b_n}{n!}x^n$ *are the series expansions of* $F(x)$ *and* $g(x)$ *respectively, then the coefficients* a_n *satisfy the recursion*

$$a_{n+1}=\sum_{k=0}^{n}\binom{n}{k}a_k+b_n.$$

Moreover, we have

$$a_n=B_n\big(C+O\big(e^{-\kappa n/\log^2 n}\big)\big),$$

where $C=\int_0^\infty e^{1-e^t}g(t)\,dt$ *and* κ *is a positive constant.*

The paper is organised as follows. In the next section we present the proof of the conjecture of Gould and Quaintance [4]. In Section 3, we provide some combinatorial applications of Theorem 1, in particular we determine the asymptotic number of permutations and set partitions of $[n]$ under varying conditions.

Remark 1. For the case $g(x)=1$, the constant

$$\int_0^\infty \frac{e^{-u}}{1+u}\,du=\int_0^\infty e^{1-e^x}\,dx=-e\,\mathrm{Ei}(-1)\approx 0.5963473623\ldots$$

is known as the Euler–Gompertz constant [2, Section 6.2]; it will occur repeatedly in many of our other examples as well. This corresponds to the basic problem of Gould and Quaintance. We will henceforth denote this constant by G.

2. Proof of Theorem 1

The recursion for a_n follows directly by comparing coefficients in the differential equation. Now note that

$$\begin{aligned}F(x)&=e^{e^x}\int_0^x e^{-e^t}g(t)\,dt=e^{e^x-1}\int_0^\infty e^{1-e^t}g(t)\,dt-e^{e^x}\int_x^\infty e^{-e^t}g(t)\,dt\\&=C\sum_{n=0}^{\infty}\frac{B_n}{n!}x^n-e^{e^x}\int_x^\infty e^{-e^t}g(t)\,dt,\end{aligned}$$

so we only have to show that the coefficients of

$$H(x)=\sum_{n=0}^{\infty}\frac{h_n}{n!}x^n=e^{e^x}\int_x^\infty e^{-e^t}g(t)\,dt$$

satisfy $h_n=O(B_ne^{-\kappa n/\log^2 n})$. By our assumptions on g, the function $H(x)$ can be analytically continued to an entire function in the complex plane, so we can apply Cauchy's integral formula:

$$\frac{h_n}{n!}=\frac{1}{2\pi i}\int_{|z|=R} z^{-n-1}e^{e^z}\int_z^\infty e^{-e^r}g(r)\,dr\,dz. \tag{2.1}$$

Here, $R \sim \log n$ (defined implicitly by $Re^R = n+1$) is chosen as the saddle point of the function $\frac{e^{e^x-1}}{x^{n+1}}$ belonging to the generating function for the Bell numbers. A classical application of the saddle point method [3, Example VIII.6] yields the asymptotic formula

$$\frac{B_n}{n!} \sim \frac{R^{-n}e^{e^R-1}}{\sqrt{2\pi R(R+1)e^R}}. \tag{2.2}$$

Our goal will be to prove a uniform bound for

$$\left|\frac{e^{e^z}\int_z^\infty e^{-e^r}g(r)\,dr}{e^{e^R}}\right|$$

on the entire circle $|z| = R$, from which our theorem will follow immediately. Note first that

$$\left|\int_R^\infty e^{-e^r}g(r)\,dr\right| \leqslant \int_R^\infty e^{-e^r}\left|g(r)\right|dr = O\left(\int_R^\infty e^{e^{(1-\epsilon)r}-e^r}\,dr\right) = O\left(e^{e^{(1-\epsilon)R}-e^R}\right),$$

and that $|e^{e^z}| \leqslant e^{|e^z|} \leqslant e^{e^{|z|}} = e^{e^R}$ for $|z| = R$, which means that

$$e^{e^z}\int_R^\infty e^{-e^r}g(r)\,dr = O\left(e^{e^{(1-\epsilon)R}}\right).$$

So we consider

$$e^{e^z}\int_z^R e^{-e^r}g(r)\,dr$$

instead of extending the integral to ∞. Write $z = Re^{i\theta}$ with $|\theta| \leqslant \pi$, and assume without loss of generality that $\theta \geqslant 0$. We consider two different cases now:

- $\theta \leqslant \arcsin\frac{\pi}{3R}$: In this case, we note that for all $r = Re^{it}$ ($0 \leqslant t \leqslant \theta$) on the arc between R and z, the inequality

  $$\begin{aligned}\left|e^{-e^r}\right| &= \exp\left(-Re\left(e^r\right)\right) = \exp\left(-e^{R\cos t}\cos(R\sin t)\right)\\ &\leqslant \exp\left(-e^{R\cos t}\cos(R\sin\theta)\right) \leqslant \exp\left(-e^{R\cos t}/2\right)\end{aligned}$$

  holds. Moreover, since $\cos t \geqslant \cos\theta = \sqrt{1-\sin^2\theta} \geqslant \sqrt{1-\pi^2/(9R^2)} \geqslant 1-\pi^2/(9R^2)$, we end up with

  $$\left|e^{-e^r}\right| = O\left(e^{-c_1e^R}\right)$$

  for any positive constant $c_1 < 1/2$. This, the assumption that $g(r) = O(e^{e^{(1-\epsilon)|r|}})$ and the trivial estimate $|e^{e^z}| \leqslant e^{e^R}$ from before yield

  $$\left|\frac{e^{e^z}\int_z^\infty e^{-e^r}g(r)\,dr}{e^{e^R}}\right| = O\left(e^{e^{(1-\epsilon)R}-c_1e^R}\right)$$

  in this case.

- $\theta \geqslant \arcsin \frac{\pi}{3R}$: We now choose a different path of integration from R to z: the union of the line segments from R to 0 and from 0 to z. Of course,

$$\left|\int_0^R e^{-e^r} g(r)\,dr\right| \leqslant \int_0^\infty e^{-e^r}\left|g(r)\right|dr < \infty,$$

and since

$$\begin{aligned}\left|e^{e^z}\right| &\leqslant e^{|e^z|} = \exp\bigl(e^{R\cos\theta}\bigr)\\ &\leqslant \exp\bigl(e^{R\sqrt{1-\pi^2/(9R^2)}}\bigr) = \exp\left(e^R\left(1-\frac{\pi^2}{18R}+O\bigl(R^{-2}\bigr)\right)\right),\end{aligned}$$

we have

$$\left|\frac{e^{e^z}\int_0^R e^{-e^r} g(r)\,dr}{e^{e^R}}\right| = O\bigl(e^{-c_2e^R/R}\bigr)$$

for any constant $c_2 < \pi^2/18$. Thus we focus on the remaining integral from 0 to z, for which we have the estimate

$$\left|\int_0^z e^{-e^r} g(r)\,dr\right| = O\Bigl(Re^{e^{(1-\epsilon)R}}\max_{0\leqslant u\leqslant 1}\left|e^{-e^{uz}}\right|\Bigr)$$

by our assumptions on g. Let us consider

$$\begin{aligned}\left|e^{e^z}\right|\max_{0\leqslant u\leqslant 1}\left|e^{-e^{uz}}\right| &= \exp\Bigl(Re\bigl(e^z\bigr) - \min_{0\leqslant u\leqslant 1} Re\bigl(e^{uz}\bigr)\Bigr)\\ &= \exp\Bigl(e^{R\cos\theta}\cos(R\sin\theta) - \min_{0\leqslant u\leqslant 1} e^{uR\cos\theta}\cos(uR\sin\theta)\Bigr). \end{aligned} \tag{2.3}$$

Let the minimum be attained at $u = u_0$. If $\cos(u_0R\sin\theta) \geqslant 0$ or $\cos(R\sin\theta) - \cos(u_0R\sin\theta) \leqslant 1$, then we have

$$\left|e^{e^z}\right|\max_{0\leqslant u\leqslant 1}\left|e^{-e^{uz}}\right| \leqslant \exp\bigl(e^{R\cos\theta}\bigr) = O\bigl(\exp\bigl(e^R - c_2e^R/R\bigr)\bigr)$$

as before. Also, if $\cos\theta \leqslant 0$, then the entire expression in (2.3) is clearly $O(1)$. Thus we assume now that $\cos\theta \geqslant 0$ and

$$\cos(R\sin\theta) - \cos(u_0R\sin\theta) = 2\sin\left(\frac{u_0+1}{2}R\sin\theta\right)\sin\left(\frac{u_0-1}{2}R\sin\theta\right) \geqslant 1.$$

It follows that

$$\left|\sin\left(\frac{u_0-1}{2}R\sin\theta\right)\right| \geqslant \frac{1}{2}$$

and thus $u_0 \leqslant 1 - \pi/(3R\sin\theta)$. Now (2.3) gives us

$$\left|e^{e^z}\right|\max_{0\leqslant u\leqslant 1}\left|e^{-e^{uz}}\right| \leqslant \exp\bigl(e^{R\cos\theta} + e^{R\cos\theta - \pi/3\cot\theta}\bigr).$$

The maximum of $R\cos\theta - \pi/3\cot\theta$ (as a function of θ) is achieved when $\sin\theta = (\pi/(3R))^{1/3}$, and it is $R - (3\pi^2 R)^{1/3}/2 + O(1)$. Hence we end up with

$$\left|e^{e^z}\right| \max_{0\leqslant u\leqslant 1}\left|e^{-e^{uz}}\right| = O\bigl(\exp\bigl(e^R - c_2 e^R/R\bigr)\bigr)$$

once again.

Putting everything together, we find that

$$\left|\frac{e^{e^z}\int_z^\infty e^{-e^r} g(r)\,dr}{e^{e^R}}\right| = O\bigl(e^{-c_3 e^R/R}\bigr)$$

for some positive constant c_3, uniformly for all z with $|z| = R$. Applying this estimate to the integral in (2.1), we obtain

$$\frac{h_n}{n!} = O\bigl(R^{-n} e^{e^R - c_3 e^R/R}\bigr).$$

Combining this with (2.2) and the definition of R as the unique positive solution to the equation $Re^R = n+1$, we find that

$$h_n = O\bigl(B_n e^{-c_4 n/\log^2 n}\bigr)$$

for a positive constant c_4, which finally proves our claim.

3. Applications

In the following subsections, we make use of our main theorem, Theorem 1, to establish asymptotics (as $n\to\infty$) for the number of set partitions or classes of permutations of $[n]$ under varying conditions.

3.1. Restricted permutations

We say that a permutation $\pi = \pi_1\pi_2\cdots\pi_n$ avoids 1–23 if there are no $1 \leqslant i < j < n$ such that $\pi_i < \pi_j < \pi_{j+1}$ and we say that for $m \geqslant 2$, π begins with an m-long descent (or m-long rise, respectively) if $\pi_1 > \pi_2 > \cdots > \pi_m$ (or $\pi_1 < \pi_2 < \cdots < \pi_m$ respectively) and $n \geqslant m$. We denote the number of permutations of length n that avoid 1–23 and begin with an m-long descent by $a_{n,m}$. The exponential generating function for the sequence $a_{n,m}$ is given by $\sum_{n\geqslant m} a_{n,m}\frac{x^n}{n!} = \frac{1}{(m-1)!}e^{e^x}\int_0^x t^{m-1}e^{t-e^t}\,dt$, see Proposition 4 in [5] (for other examples see Proposition 26 in [5]). By Theorem 1 with $g(x) = \frac{1}{(m-1)!}x^{m-1}e^x$ we obtain the following corollary.

Corollary 1. *We have*

$$a_{n,m} = B_n\bigl(C_m + O\bigl(e^{-\kappa n/\log^2 n}\bigr)\bigr),$$

where $C_m = \frac{1}{(m-1)!}\int_0^\infty t^{m-1}e^{1+t-e^t}\,dt$ and κ is a positive constant. For $m = 2,\dots,5$, we have that $C_2 = G = 0.5963473623\dots$, $C_3 = 0.2659653850\dots$, $C_4 = 0.09678032514\dots$ and $C_5 = 0.03009381392\dots$.

Let A_n be the set of permutations of length n that avoid 1–23. Note that $|A_n| = B_n$, see [1]. Thus, the above corollary shows that approximately 59.63% of the permutations of A_n begin with a descent, but only approximately 26.60% of the permutations of A_n begin with a 3-long descent, when $n \to \infty$. In general, $\lim_{n\to\infty} \frac{a_{n,m}}{|A_n|} = C_m$, for all $m \geqslant 2$.

As another example, we denote the number of permutations of length n that avoid 1–23 and begin with a 2-long rise by b_n. It is known that the exponential generating function for the sequence b_n is given by $\sum_{n\geqslant 2} b_n \frac{x^n}{n!} = e^{e^x} \int_0^x e^{-e^t}(e^t-1)\,dt$ (see Proposition 24 in [5]). Thus, by Theorem 1 with $g(t) = e^t - 1$, we obtain the following result.

Corollary 2. *We have*

$$b_n = B_n\big(C + O\big(e^{-\kappa n/\log^2 n}\big)\big),$$

where $C = \int_0^\infty e^{1-e^t}(e^t-1)\,dt = 1 - G = 0.4036526378\ldots$ *and* κ *is a positive constant.*

The above corollary agrees with the previous one because each permutation of $n \geqslant 2$ begins with a 2-long descent or a 2-long rise.

3.2. The last block in a set partition

In this section we are interested in the average size of the last block in the canonical representation of a set partition. Equivalently, this is the number of maxima in the associated restricted growth function, which can be decomposed as

$$1\sigma^{(1)}2\sigma^{(2)}\cdots\big(k\sigma^{(k-1)}\big)^* k\sigma^{(k-1)},$$

where $*$ denotes a possibly empty sequence and where $\sigma^{(m)}$ denotes a word over the alphabet $[m]$.

This yields the generating function

$$R_k(x,u) = \left(\prod_{j=1}^{k-1} \frac{x}{1-jx}\right) \frac{xu\frac{1}{1-(k-1)x}}{1-\frac{xu}{1-(k-1)x}} = \prod_{j=1}^{k-1} \frac{x}{1-jx} \frac{xu}{1-(k-1)x-xu} \tag{3.1}$$

where x marks the size of the set partition and u marks the number of elements equal to k which are the maxima. This is equivalent to

$$\big(1-(k-1)x-xu\big)R_k(x,u) = xu\prod_{j=1}^{k-1}\frac{x}{1-jx}.$$

Now

$$\prod_{j=1}^{k-1}\frac{x}{1-jx} = \frac{x^{k-1}}{(1-x)(1-2x)\cdots(1-(k-1)x)} = \sum_{j=0}^{k-1}\frac{a_j}{1-jx},$$

where the coefficients in the partial fraction decomposition are given by

$$\begin{aligned} a_j &= \frac{1/j^{k-1}}{(1-1/j)(1-2/j)\cdots(1-(j-1)/j)(1-(j+1)/j)(1-(j+2)/j)\cdots(1-(k-1)/j)} \\ &= (-1)^{k-1-j}\frac{1}{(k-1)!}\binom{k-1}{j}. \end{aligned}$$

So

$$\frac{1}{x}R_k(x,u) - kR_k(x,u) - (u-1)R_k(x,u) = u\sum_{j=0}^{k-1}\frac{a_j}{1-jx}.$$

Now we translate this to exponential generating functions, in which each term of the form x^n is replaced by $\frac{x^n}{n!}$. Specifically, if $R_k(x,u) = \sum_{n=0}^{\infty} r_{k,n}(u)x^n$, we set $Q_k(x,u) = \sum_{n=0}^{\infty} r_{k,n}(u)x^n/n!$. Then we get

$$\frac{\partial}{\partial x}Q_k(x,u) - kQ_k(x,u) - (u-1)Q_k(x,u) = u\sum_{j=0}^{k-1} a_j e^{jx} = u\frac{(e^x-1)^{k-1}}{(k-1)!}.$$

Multiplying by v^k and summing over all $k \geqslant 1$ leads to a differential equation for $Q(x,u,v) = \sum_{k\geqslant 1} Q_k(x,u)v^k$:

$$\frac{\partial}{\partial x}Q(x,u,v) - v\frac{\partial}{\partial v}Q(x,u,v) - (u-1)Q(x,u,v) = \sum_{k\geqslant 1} uv^k\frac{(e^x-1)^{k-1}}{(k-1)!} = uve^{v(e^x-1)}$$

It is not hard to check that the solution of this partial differential equation (with initial condition $Q(0,u,v) = 0$) is

$$Q(x,u,v) = uv\int_0^x e^{v(e^x-e^t)+ut}\,dt.$$

To obtain the average size of the last block, we differentiate with respect to u and set $u = v = 1$, which gives us the generating function

$$\int_0^x (t+1)e^{e^x-e^t+t}\,dt = -(t+1)e^{e^x-e^t}\Big|_0^x + \int_0^x e^{e^x-e^t}\,dt = e^{e^x-1} - x - 1 + e^{e^x-1}\int_0^x e^{1-e^t}\,dt,$$

also making use of integration by parts. This can be expanded as

$$-1 - x + \sum_{n\geqslant 0}\frac{B_n x^n}{n!} + \sum_{n\geqslant 0}\sum_{j=0}^{n-1}\binom{n}{j}B_j c_{n-1-j}\frac{x^n}{n!},$$

where c_m are the m-th complementary Bell numbers (A000587 in [8]). So, for $n \geqslant 2$, we find that the sum of the sizes of the last blocks in all set partition of $[n]$ is

$$q_n = B_n + \sum_{j=0}^{n-1}\binom{n}{j}B_j c_{n-1-j},$$

and Theorem 1 immediately yields

Corollary 3. *The average number of elements in the last block of a set partition written in canonical form is*

$$1 + G + O\big(e^{-\kappa n/\log^2 n}\big),$$

where $1 + G = 1.5963473623\ldots$ *and where* κ *is a positive constant, when* $n \to \infty$.

Remark 2. Note that the generating function in (3.1) is also the generating function for the number of elements in the first block that are greater than the first maximum (smallest element of the last block). It follows immediately that the average number of elements in the first block that are greater than the minimal element of the last block is $G + O(e^{-\kappa n/\log^2 n})$. The same can be said for any other fixed block (second, third, ...).

3.3. Set partitions with m elements in the last block

The generating function method from the previous section could also be used to determine the asymptotic number of set partitions whose last block has precisely m elements (for some fixed m). Let us however discuss an alternative approach that gives a combinatorial interpretation for the recursive formula in Theorem 1 as well. We denote the number of set partitions with m elements in the last block by $a_m(n)$. Now, consider such a set partition of $[n+1]$, and suppose there are k elements, in addition to the mandatory number 1, in the first block. The $n-k$ elements not in this block form a set partition with the same property, i.e., there are $a_m(n-k)$ possibilities for them. This leads to the recursion

$$a_m(n+1) = \sum_{k=0}^{n} \binom{n}{k} a_m(k) = \sum_{k=0}^{n} \binom{n}{k} a_m(n-k),$$

with the initial conditions $a_m(0) = a_m(1) = \cdots = a_m(m-1) = 0$ and $a_m(m) = 1$ (where there is only one block).

Now we can apply Theorem 1 with $b_{m-1} = 1$ and $b_j = 0$ for $j \neq m-1$. From Theorem 1 the generating function for such set partitions is

$$F(x) = e^{e^x} \int_0^\infty e^{-e^t} g(t)\, dt,$$

where $g(t) = \frac{t^{m-1}}{(m-1)!}$. We obtain the following asymptotic result.

Corollary 4. *The probability that a set partition, when written in canonical form, has exactly m elements in the last block is*

$$\int_0^\infty e^{1-e^t} \frac{t^{m-1}}{(m-1)!}\, dt + O\big(e^{-\kappa n/\log^2 n}\big),$$

as $n \to \infty$. These probabilities are 0.596347, 0.265965, 0.0967803, 0.0300938, 0.0082299 *for* $m = 1, 2, 3, 4, 5$.

We remark that it would be possible to prove analogous results for the second-last block, etc.

3.4. The first element in the last block of a set partition

Let us now study the first maximum in a restricted growth function (or equivalently, the minimal element in the last block of the canonical representation of a set partition). It will be more convenient, however, to consider the number of elements greater or equal to the minimal element in the last block (if this element in a set partition of $[n]$ is m, then this number is $n-m+1$). For example the set partition $\{1, 6\}$, $\{2, 4, 8\}$, $\{3\}$, $\{5, 7\}$ of $[8]$ has $m = 5$ and $n - m + 1 = 4$.

From the decomposition

$$1\sigma^{(1)}2\sigma^{(2)}\cdots k\sigma^{(k)},$$

of restricted growth functions, we immediately get the generating function

$$R_k(x,u)=\left(\prod_{j=1}^{k-1}\frac{x}{1-jx}\right)\frac{xu}{1-kxu}$$

where x marks the size of the set partition and u marks the number of elements greater or equal to the first maximum. We can follow exactly the same steps as in Section 3.2 to obtain a trivariate exponential generating function, which is now

$$Q(x,u,v)=uv\int_0^x e^{ve^{ut}(e^{x-t}-1)+ut}\,dt.$$

Differentiating with respect to u and setting $u=v=1$ yields the generating function

$$\int_0^x \big(t\big(e^x-e^t+1\big)+1\big)e^{e^x-e^t+t}\,dt=te^{e^x-e^t+t}\big|_0^x+e^{e^x+x}\int_0^x te^{t-e^t}\,dt=xe^x+e^{e^x+x}\int_0^x te^{t-e^t}\,dt$$
$$=\frac{d}{dx}\left(e^{e^x}\int_0^x te^{t-e^t}\,dt\right).$$

Since differentiating an exponential generating function merely corresponds to a coefficient shift, we can apply Theorem 1 once again to obtain the following corollary:

Corollary 5. *The average number of elements greater or equal to the minimal element in the last block of a set partition of $[n]$ is*

$$\frac{GB_{n+1}}{B_n}+O\big(e^{-\kappa n/\log^2 n}\big)$$

for a positive constant κ. Here, the constant is $G=0.5963473623\ldots$ again. Consequently, the mean value of the minimal element in the last block is

$$n+1-\frac{GB_{n+1}}{B_n}+O\big(e^{-\kappa n/\log^2 n}\big). \tag{3.2}$$

Remark 3. Recall that $\frac{B_{n+1}}{B_n}-1\sim\frac{n}{\log n}$ is exactly the average number of blocks in a set partition of $[n]$.

Let us finally determine how many blocks on average contain an element that is greater than the smallest element of the last block. For our example $\{1,6\}$, $\{2,4,8\}$, $\{3\}$, $\{5,7\}$ the number of such blocks is 3. In terms of restricted growth functions, this is the number of distinct letters that occur after the first maximum. If k is the number of blocks (letters in the restricted growth function) and r the number of elements after the first maximum, then there are $k^r-(k-1)^r$ possibilities for these elements that contain a certain letter. Thus the generating function for the total number of distinct letters after the first maximum is

$$R_k(x)=\left(\prod_{j=1}^{k-1}\frac{x}{1-jx}\right)\sum_{r=0}^{\infty}k\big(k^r-(k-1)^r\big)x^{r+1}=\left(\prod_{j=1}^{k-1}\frac{x}{1-jx}\right)\left(\frac{kx}{1-kx}-\frac{kx}{1-(k-1)x}\right).$$

Passing on to the exponential generating function again, we end up with

$$Q(x,v) = v\int_0^x \left(e^t - 1\right)\left(v\left(e^x - e^t\right) + 1\right)e^{v(e^x - e^t)}\,dt,$$

with v marking the number of blocks. Setting $v = 1$ and simplifying yields

$$Q(x,1) = e^{e^x + x - 1} - \left(1 + \frac{d}{dx}\right)e^{e^x}\int_0^x e^{-e^t}\,dt,$$

which gives us the following corollary:

Corollary 6. *The average number of blocks with an element greater than the smallest element of the last block of a set partition of* $[n]$ *is*

$$\frac{(1-G)B_{n+1}}{B_n} - G + O\left(e^{-\kappa n/\log^2 n}\right)$$

for a positive constant κ.

Combining the results of the last two corollaries, we can say that the average number of elements following the last maximum is about 0.6 times the average number of blocks, while the average number of distinct blocks these elements belong to is about 0.4 times the average total number of blocks.

References

[1] A. Claesson, Generalised pattern avoidance, European J. Combin. 22 (2001) 961–971.
[2] S.R. Finch, Mathematical Constants, Encyclopedia of Mathematics and Its Applications, vol. 94, Cambridge University Press, 2003.
[3] P. Flajolet, R. Sedgewick, Analytic Combinatorics, Cambridge University Press, 2009.
[4] H. Gould, J. Quaintance, A linear binomial recurrence and the Bell numbers and polynomials, Appl. Anal. Discrete Math. 1 (2007) 371–385.
[5] S. Kitaev, T. Mansour, Simultaneous avoidance of generalized patterns, Ars Combin. 75 (2005) 267–288.
[6] T. Mansour, Combinatorics of Set Partitions, Discrete Mathematics and Its Applications, CRC Press, Chapman and Hall, Taylor and Francis Group, 2013.
[7] R.P. Stanley, Enumerative Combinatorics, vol. I, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, 1999.
[8] The on-line encyclopedia of integer sequences, published electronically at http://oeis.org, 2013.