THE WATER CAPACITY OF INTEGER COMPOSITIONS

AUBREY BLECHER, CHARLOTTE BRENNAN[1], AND ARNOLD KNOPFMACHER[1]

Abstract. We introduce the notion of capacity (ability to contain water) for compositions. Initially the compositions are defined on a finite alphabet $[k]$ and thereafter on $\mathbb{N}$. We find a capacity generating function for all compositions, the average capacity generating function and an asymptotic expression for the average capacity as the size of the composition increases to infinity.

1. Introduction

A composition of n is a sum of positive integers $x_1 + x_2 + \cdots + x_k = n$ and it can be represented graphically using a column of height x_i to represent the i-th part. Given an arbitrary composition made up of letters from $[k] := \{1, 2, \cdots, k\}$, we mean by its *capacity*, the amount of water the composition would retain if it was considered as a full container with the water allowed to escape left and right, subject to the usual rules of water flow. The notion of capacity in compositions is, as the name suggests, a natural two-dimensional model for the extent to which various shapes may contain water. This is useful for the theoretical modelling of dam capacity and provides the initial reason for this study. The notion is general enough to include structures such as words on a finite or infinite alphabet, Dyck paths, permutations. These cases are dealt with in further articles [3, 4, 5]. Capacity may also be thought of as a measure of a particular composition's degree of "concavity", hence an interesting concept in itself. Recently, certain specialized concave compositions have been an object of study, for example, see Andrews [1, 2]. In these papers, the particular type of compositions were studied because they allowed the problem of non-intersecting spiral walks to be completely elucidated. However, the degree of concavity did not figure in these studies. Here we are suggesting that the notions of capacity are interesting in their own right and may well become the subject of further applications. In these studies, generating functions were obtained. As preliminary examples the compositions **3 2 4 5 1 3 2 1** and **3 2 4 1 3 2 1 5** of **21**, which differ only in the position of the largest part, would have capacities 3 and 10 respectively, as shown by the shading below.

Date: February 14, 2019.

1991 *Mathematics Subject Classification.* Primary: 05A15, 05A16; Secondary:

Key words and phrases. Compositions, generating functions, asymptotics, water capacity.

[1]This material is based upon work supported by the National Research Foundation under grant numbers 86329 and 81021 respectively.

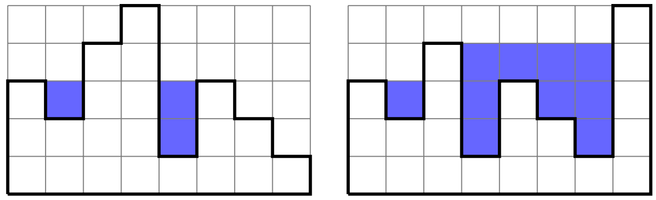

FIGURE 1. The water capacity of two compositions of $n = 21$

2. Definition

More precisely, for an arbitrary composition of n on $[k]$, say $\alpha_1\alpha_2\cdots\alpha_m$ where $n = \sum_{i=1}^{m}\alpha_i$, *the capacity contribution* of α_i, $1 \le i \le m$ is denoted $C(\alpha_i)$ and defined as

$$C(\alpha_1) = 0, \quad C(\alpha_m) = 0,$$

and for all i, $1 < i < m$,

$$C(\alpha_i) = 0, \quad \text{if } \alpha_i \ge \max\{\alpha_j; j < i\} \text{ or } \alpha_i \ge \max\{\alpha_j; j > i\}$$

and

$$C(\alpha_i) = \min\big\{\max\{\alpha_j; j < i\}, \max\{\alpha_j; j > i\}\big\} - \alpha_i, \quad \text{otherwise.}$$

The capacity of the composition of n constituted by $\alpha_1\alpha_2\ldots\alpha_m$ is defined to be $\sum_{i=1}^{m} C(\alpha_i)$.

Refer to examples in Figure 1.

3. The Main Capacity Generating Function

Our first aim is to provide a generating function for the capacity of compositions on an alphabet $[k]$, and thereafter to allow the alphabet to be $\mathbb{N}$.

Suppose, firstly, that there are at least two uses of the letter k in a particular composition. Consider the capacity of that part of the composition constituted by the first and last occurrences of k and everything in between.

We know that compositions on alphabet $[k]$ have generating function

$$\sum_{m\ge 0}\left(\sum_{a=1}^{k} z^a\right)^m = \frac{1}{1-\sum_{a=1}^{k} z^a}$$

where z marks the size of the composition, (see for example Theorem 3.13, page 71 [6]).

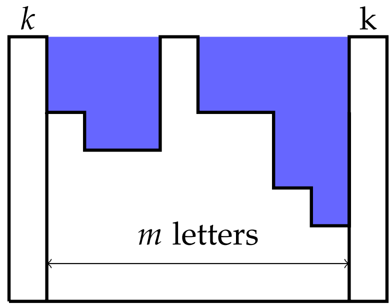

FIGURE 2. A composition with its capacity between the leftmost and rightmost occurrences of k

So, for such a composition having $m+2$ letters, the generating function for the capacity (i.e., area of shaded region in Figure 2) is therefore given by the exponent of x in

$$y^2 z^{2k} \sum_{m\geq 0} \left(\sum_{a=1}^{k} z^a x^{k-a} \right)^m y^m$$

where y marks the number of letters in the subword, x marks the capacity and z marks the size of the composition. Thus, assuming our composition has one or more occurrences of the letter k, the capacity for the sub-composition contained between the first and last occurrence of k (including these occurrences) is

$$y^2 z^{2k} \sum_{m\geq 0} \left(\sum_{a=1}^{k} z^a x^{k-a} \right)^m y^m + yz^k \tag{3.1}$$

where the last term is for the case where there is only one occurrence of k (and therefore zero capacity).

Since the main interest in compositions is with respect to their area marked by the variable x, we now set $y=1$ in (3.1) to obtain

$$\begin{aligned} & z^{2k} \sum_{m\geq 0} \left(\sum_{a=1}^{k} z^a x^{k-1} \right)^m + z^k \\ = \; & z^{2k} \frac{1}{1 - x\sum_{a=1}^{k} z^a x^{k-1}} + z^k \\ = \; & \frac{(x-z)z^{2k}}{(x-z) - z(x^k - z^k)} + z^k. \end{aligned} \tag{3.2}$$

Next, consider the occurrences of the letter $k-1$, either left or right of the subword above.

So, for example, when we consider the leftmost occurrence of $k-1$ to the left of k, we have this illustration.

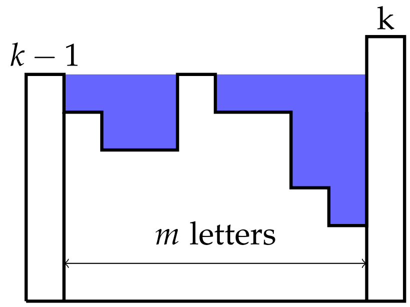

FIGURE 3. Leftmost occurrence of $k-1$ to the left of k

Arguing as above, the capacity generating function of this subword of our original composition is

$$\frac{(x-z)z^{k-1}}{(x-z)-z\left(x^{k-1}-z^{k-1}\right)}+1$$

where 1 is for the case where there is no $k-1$.

Iterating this process for smaller part sizes and noting that there is at least one occurrence of the largest part k in our word, we get the capacity generating function for compositions on the finite alphabet $[k]$ to be

$$C_k = z^k\left(1+\frac{(x-z)z^k}{(x-z)-z\left(x^k-z^k\right)}\right)\prod_{i=1}^{k-1}\left(1+\frac{(x-z)z^i}{(x-z)-z\left(x^i-z^i\right)}\right)^2.$$

Hence the capacity generating function for all compositions on the finite alphabet $[k]$ is given by

$$C(k) := \sum_{r=1}^{k} C_r = \sum_{r=1}^{k} z^r\left(1+\frac{(x-z)z^r}{(x-z)-z\left(x^r-z^r\right)}\right)\prod_{i=1}^{r-1}\left(1+\frac{(x-z)z^i}{(x-z)-z\left(x^i-z^i\right)}\right)^2$$

with

$$C := \sum_{r\geq 1} C_r = \sum_{r\geq 1} z^r\left(1+\frac{(x-z)z^r}{(x-z)-z\left(x^r-z^r\right)}\right)\prod_{i=1}^{r-1}\left(1+\frac{(x-z)z^i}{(x-z)-z\left(x^i-z^i\right)}\right)^2$$

for the infinite alphabet $\mathbb{N}$.

4. THE TOTAL CAPACITY GENERATING FUNCTION

Now let

$$C_r = z^r\underbrace{\left(1+\frac{(x-z)z^r}{(x-z)-z\left(x^r-z^r\right)}\right)}_{A}\cdot\underbrace{\prod_{i=1}^{r-1}\left(1+\frac{(x-z)z^i}{(x-z)-z\left(x^i-z^i\right)}\right)^2}_{B}.$$

So that $C(N) := \sum_{r=1}^{N} C_r$.

We need to calculate $\left.\frac{\partial C(N)}{\partial x}\right|_{x=1}$; so first compute $\left.\frac{\partial C_r}{\partial x}\right|_{x=1}$. Let

$$A = 1 + \frac{(x-z)z^r}{(x-z) - z(x^r - z^r)}.$$

$$B = \prod_{i=1}^{r-1} \left(1 + \frac{(x-z)z^i}{(x-z) - z(x^i - z^i)}\right)^2.$$

So

$$\left.\frac{\partial C_r}{\partial x}\right|_{x=1} = z^r (AB' + A'B)\big|_{x=1}.$$

Now

$$\begin{aligned}
A\big|_{x=1} &= 1 + \frac{(1-z)z^r}{(1-z) - z(1-z^r)} = \frac{(1-z) - z(1-z^{r-1})}{(1-z) - z(1-z^r)} = \frac{1-2z+z^r}{1-2z+z^{r+1}},\\
B\big|_{x=1} &= \prod_{i=1}^{r-1} \left[1 + \frac{(1-z)z^i}{(1-z) - z(1-z^i)}\right]^2 = \prod_{i=1}^{r-1} \left(\frac{(1-z) - z(1-z^{i-1})}{(1-z) - z(1-z^i)}\right)^2\\
&= \frac{(1-z)^2}{(1-2z+z^r)^2},\\
A'\big|_{x=1} &= \frac{\left[(1-z) - z(1-z^r)\right] z^r - (1-z)z^r(1-rz)}{\left((1-z) - z(1-z^r)\right)^2}\\
&= \frac{z^{r+1}(r-1) + z^{2r+1} - rz^{r+2}}{(1-2z+z^{r+1})^2} = \frac{z^{r+1}\left[(r-1) + z^r - rz\right]}{(1-2z+z^{r+1})^2},
\end{aligned}$$

and

$$\begin{aligned}
B'\big|_{x=1} &= \left(B\big|_{x=1}\right) \cdot \sum_{i=1}^{r-1} \left.\left(\frac{2}{1 + \frac{(x-z)z^i}{(x-z)-z(x^i-z^i)}}\right)\right|_{x=1} \left.\frac{\partial}{\partial x}\left(\frac{(x-z)z^i}{(x-z) - z(x^i - z^i)}\right)\right|_{x=1}\\
&= \frac{2(1-z)^2}{(1-2z+z^r)^2} \sum_{i=1}^{r-1} \frac{z^{i+1}\left[-1+i-iz+z^i\right]}{(1-2z+z^{i+1})(1-2z+z^i)}\\
&= \frac{2(1-z)^2}{(1-2z+z^r)^2} \sum_{i=1}^{r-1} \left(1 + \frac{2z-iz}{1-2z+z^i} - \frac{1-iz}{1-2z+z^{i+1}}\right)\\
&= \frac{2(1-z)^2}{(1-2z+z^r)^2} \left(r - 1 + \sum_{i=1}^{r-1} \frac{2z-iz}{1-2z+z^i} - \sum_{i=2}^{r} \frac{1-(i-1)z}{1-2z+z^i}\right)\\
&= \frac{2(1-z)^2}{(1-2z+z^r)^2} \left(r - 1 + \frac{z}{1-z} - \frac{1-(r-1)z}{1-2z+z^r} + \sum_{i=2}^{r-1} \frac{z-1}{1-2z+z^i}\right).
\end{aligned}$$

Therefore:

$$z^r A'B\big|_{x=1} = \frac{(1-z)^2 z^{2r+1}\big[(r-1)+z^r-rz\big]}{(1-2z+z^r)^2(1-2z+z^{r+1})^2}.$$

By splitting into partial fractions, we get

$$\begin{aligned}
&\frac{(1-z)^2 z^{2r+1}(r-1-rz+z^r)}{(1-2z+z^r)^2(1-2z+z^{r+1})^2}\\
&= \underbrace{\frac{-2z+rz+2z^2-rz^2}{(1-2z+z^r)^2}}_{1} + \underbrace{\frac{-5z+2rz+2z^2}{(-1+2z)(1-2z+z^r)}}_{2} + \underbrace{\frac{-1+z+rz-rz^2}{(1-2z+z^{1+r})^2}}_{3}
\end{aligned}$$

$$+ \underbrace{\frac{-1+4z-2rz^2}{(-1+2z)(1-2z+z^{1+r})}}_{4}. \tag{4.1}$$

Now consider:

$$\sum_{r=1}^{N} z^r A'B\Bigg|_{x=1}.$$

Restricting attention to the terms labelled 1 and 3 in (4.1), we obtain

$$\begin{aligned}
&\sum_{r=1}^{N}\frac{z(r-2)+z^2(2-r)}{(1-2z+z^r)^2} + \sum_{r=1}^{N}\frac{-1+z+rz-rz^2}{(1-2z+z^{r+1})^2}\\
&= \frac{-z+z^2}{(1-2z+z)^2} + \sum_{r=2}^{N}\frac{z(r-2)+z^2(2-r)-1+z+(r-1)z-(r-1)z^2}{(1-2z+z^r)^2}\\
&\quad + \frac{-1+z+Nz-Nz^2}{(1-2z+z^{N+1})^2}\\
&= -\frac{z}{1-z} + \sum_{r=2}^{N}\frac{(1-z)\big((2r-3)z-1\big)}{(1-2z+z^r)^2} + \frac{-1+z+Nz-Nz^2}{(1-2z+z^{N+1})^2}. \qquad (4.2)
\end{aligned}$$

And by restricting ourselves to terms labelled 2 and 4 in (4.1), we get

$$\begin{aligned}
&\sum_{r=1}^{N}\frac{-5z+2rz+2z^2}{(-1+2z)(1-2z+z^r)} + \sum_{r=1}^{N}\frac{-1+4z-2rz^2}{(-1+2z)(1-2z+z^{r+1})}\\
&= \frac{-5z+2z+2z^2}{(-1+2z)(1-2z+z)} + \sum_{r=2}^{N}\frac{-5z+2rz+2z^2-1+4z-2(r-1)z^2}{(-1+2z)(1-2z+z^r)}\\
&\quad + \frac{-1+4z-2Nz^2}{(-1+2z)(1+2z+z^{N+1})}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{z(-3+2z)}{(-1+2z)(1-z)} + \sum_{r=2}^{N} \frac{-1+(2r-1)z+z^2(4-2r)}{(-1+2z)(1-2z+z^r)} \\
&+ \frac{-1+4z-2Nz^2}{(-1+2z)(1+2z+z^{N+1})}.
\end{aligned} \tag{4.3}
$$

Next, consider

$$
z^r AB' = \frac{2z^r(1-z)^2}{(1-2z+z^r)(1-2z+z^{r+1})} \left(r-1+\frac{z}{1-z} - \frac{1-(r-1)z}{1-2z+z^r} + \sum_{i=2}^{r-1} \frac{z-1}{1-2z+z^i} \right). \tag{4.4}
$$

Again, by splitting into partial fractions, we have

$$
\frac{2(1-z)^2 z^r}{(1-2z+z^r)(1-2z+z^{r+1})} = \frac{2(-1+z)}{1-2z+z^r} - \frac{2(-1+z)}{1-2z+z^{r+1}}, \tag{4.5}
$$

and

$$
\begin{aligned}
&\frac{2(1-z)^2 z^r \big(1-(r-1)z\big)}{(1-2z+z^r)^2(1-2z+z^{r+1})} \\
&= -\frac{2(1-rz-z^2+rz^2)}{(1-2z+z^r)^2} + \frac{2(-1-z+rz)}{(-1+2z)(1-2z+z^r)} - \frac{2(-z-z^2+rz^2)}{(-1+2z)(1-2z+z^{r+1})}.
\end{aligned} \tag{4.6}
$$

Therefore:

$$
\begin{aligned}
z^r AB' &= -\frac{2z^r(1-z)^2\big(1-(r-1)z\big)}{(1-2z+z^r)^2(1-2z+z^{r+1})} \\
&- \left(r-1+\frac{z}{1-z} + \sum_{i=2}^{r-1} \frac{z-1}{1-2z+z^i} \right) 2(1-z) \left(\frac{1}{1-2z+z^r} - \frac{1}{1-2z+z^{r+1}} \right).
\end{aligned} \tag{4.7}
$$

Note that $z^r AB' = 0$ when $r=1$ so we can start the sum $\sum_r^N z^r AB'$ at $r=2$. Now

$$
\begin{aligned}
&\sum_{r=2}^{N} \frac{1}{1-2z+z^r} - \sum_{r=2}^{N} \frac{1}{1-2z+z^{r+1}} \\
&= \sum_{r=2}^{N} \frac{1}{1-2z+z^r} - \sum_{r=3}^{N+1} \frac{1}{1-2z+z^r} \\
&= \frac{1}{(1-z)^2} - \frac{1}{1-2z+z^{N+1}}.
\end{aligned}
$$

So in (4.7),

$$
\sum_{r=2}^{N} \left(1-\frac{z}{1-z}\right) 2(1-z) \left(\frac{1}{1-2z+z^r} - \frac{1}{1-2z+z^{r+1}} \right) = \frac{2(1-2z)}{(1-z)^2} - \frac{2(1-2z)}{1-2z+z^{N+1}}. \tag{4.8}
$$

Next, using (4.6),

$$\begin{aligned}
&\sum_{r=2}^{N} -\frac{2(1-z)^2 z^r\big(1-(r-1)z\big)}{(1-2z+z^r)^2(1-2z+z^{r+1})}\\
&= \sum_{r=2}^{N} \frac{2\big((1-rz)+(r-1)z^2\big)}{(1-2z+z^r)^2} - \frac{2}{(-1+2z)}\sum_{r=2}^{N}\frac{-1+(r-1)z}{1-2z+z^r}\\
&\quad + \frac{2}{(-1+2z)}\sum_{r=3}^{N+1}\frac{-z-z^2+(r-1)z^2}{1-2z+z^r}\\
&= 2\sum_{r=2}^{N}\frac{1-rz+(r-1)z^2}{(1-2z+z^r)^2} + \frac{2}{(-1+2z)}\frac{1}{(1-z)}\\
&\quad + \frac{2}{(-1+2z)}\sum_{r=3}^{N}\frac{1-rz+z^2(r-2)}{1-2z+z^r} + \frac{2z}{(1-2z)}\frac{(1+z-Nz)}{(1-2z+z^{N+1})}, \qquad (4.9)
\end{aligned}$$

and using (4.5),

$$\begin{aligned}
&\sum_{r=2}^{N}\frac{r2(1-z)^2z^r}{(1-2z+z^r)(1-2z+z^{r+1})}\\
&= 2(-1+z)\sum_{r=2}^{N} r\left(\frac{1}{1-2z+z^r}-\frac{1}{1-2z+z^{r+1}}\right)\\
&= 2(-1+z)\left(\sum_{r=2}^{N}\frac{r}{1-2z+z^r}-\sum_{r=3}^{N+1}\frac{r-1}{1-2z+z^r}\right)\\
&= \frac{-2^2}{1-z} - 2(1-z)\sum_{r=3}^{N}\frac{1}{1-2z+z^r} + \frac{2N(1-z)}{1-2z+z^{N+1}}. \qquad (4.10)
\end{aligned}$$

Finally

$$\begin{aligned}
&2(-1+z)\sum_{r=3}^{N}\left(\frac{1}{1-2z+z^r}-\frac{1}{1-2z+z^{r+1}}\right)\sum_{i=2}^{r-1}\frac{z-1}{1-2z+z^i}\\
&= 2(-1+z)\left[\sum_{r=3}^{N}\frac{1}{1-2z+z^r}\sum_{i=2}^{r-1}\frac{z-1}{1-2z+z^i} - \sum_{r=4}^{N+1}\frac{1}{1-2z+z^r}\sum_{i=2}^{r-2}\frac{z-1}{1-2z+z^i}\right]\\
&= 2(-1+z)\left[\sum_{r=3}^{N}\frac{1}{1-2z+z^r}\cdot\frac{z-1}{1-2z+z^{r-1}} - \frac{1}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{z-1}{1-2z+z^i}\right]\\
&= 2(1-z)^2\sum_{r=3}^{N}\frac{1}{(1-2z+z^r)(1-2z+z^{r-1})} - \frac{2}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{(z-1)^2}{1-2z+z^i}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2(1-z)}{1-2z}\left[\sum_{r=3}^{N}\frac{-z}{1-2z+z^r}+\sum_{r=3}^{N}\frac{1}{1-2z+z^{r-1}}\right]-\frac{2}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{(z-1)^2}{1-2z+z^i}\\
&= \frac{2(1-z)}{1-2z}\left[\sum_{r=3}^{N-1}\frac{-z+1}{1-2z+z^r}+\frac{1}{1-2z+z^2}-\frac{z}{1-2z+z^N}\right]\\
&\quad -\frac{2}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{(z-1)^2}{1-2z+z^i}\\
&= \frac{2}{(1-2z)(1-z)}+\frac{2(1-z)^2}{1-2z}\sum_{r=3}^{N-1}\frac{1}{1-2z+z^r}-\frac{2(1-z)z}{(1-2z)(1-2z+z^N)}\\
&\quad -\frac{2(1-z)^2}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{1}{1-2z+z^i}.
\end{aligned}
\tag{4.11}
$$

Together (4.8), (4.9), (4.10), and (4.11) take care of $\sum_{r=2}^{N} z^r AB'$. Now combining and simplifying all the above contributions to the generating function for the total capacity, namely (4.2), (4.3), (4.8), (4.9), (4.10), and (4.11), we eventually obtain

Theorem 1. *The total water capacity in compositions of n with parts in $[N]$ has the generating function*

$$
\begin{aligned}
V_N(z) = {} & \frac{1}{-1+z}+\sum_{r=2}^{N}\frac{(-1+z)^2}{(1-2z+z^r)^2}+\sum_{r=3}^{N}\frac{-1+z}{1-2z+z^r}\\
&+\frac{2(-1+z)}{(-1+2z)(-1+2z-z^N)}+\frac{-1+z+Nz-Nz^2}{(1-2z+z^{N+1})^2}\\
&-\frac{1-2z+2z^2}{(-1+2z)(1-2z+z^{1+N})}-\frac{2(1-N-2z+Nz)}{1-2z+z^{1+N}}\\
&-\frac{2(1-z)^2}{1-2z+z^{N+1}}\sum_{i=2}^{N-1}\frac{1}{1-2z+z^i}.
\end{aligned}
\tag{4.12}
$$

Now let us define a further function $v_N(z)$ which is obtained by setting all terms of the form z^N or z^{N+1} in $V_N(z)$ equal to zero. Further simplifications then occur in (4.12) and we find that

$$
\begin{aligned}
v_N(z) = {} & \frac{-2+2N+5z-5Nz-4z^2+3Nz^2}{(-1+2z)^2}+\frac{(1-z)(-3+4z)}{1-2z}\sum_{r=2}^{N}\frac{1}{1-2z+z^r}\\
&+\sum_{r=2}^{N}\frac{(1-z)^2}{(1-2z+z^r)^2}.
\end{aligned}
\tag{4.13}
$$

Note that we cannot let $N\to\infty$ in $V_N(z)$ or in $v_N(z)$, as the series are then divergent. Nevertheless, if we wish to study the total capacity of unrestricted compositions of any given n, it is sufficient to consider the case of $V_N(z)$, or the simpler $v_N(z)$, with $N \geq n$. In particular, we have that $[z^j]V_N(z) = [z^j]v_N(z)$ for $j \leq N$, where $[z^j]$ denotes the extraction operation.

For example, if we take $N = 10$ then the series expansion of $V_N(z)$ begins

$$z^5 + 6z^6 + 24z^7 + 79z^8 + 233z^9 + 640z^{10} + 1674z^{11} + 4224z^{12} + O\left(z^{13}\right).$$

Whereas, the series expansion of $v_N(z)$ begins

$$z^5 + 6z^6 + 24z^7 + 79z^8 + 233z^9 + 640z^{10} + 1673z^{11} + 4220z^{12} + O\left(z^{13}\right).$$

Both series correctly give the total capacity of unrestricted compositions of n at least up to $n = N = 10$. In fact, the sequence for the total capacity of compositions of n for $n = 1$ to 18 is given by

$$0, 0, 0, 0, 1, 6, 24, 79, 233, 640, 1674, 4224, 10370, 24912, 58800, 136767, 314201, 714209.$$

5. Asymptotic estimates for the average capacity

Theorem 2. *The mean water capacity W_n in compositions of n has the asymptotic expansion*

$$W_n = \frac{n \log_2 n}{2} + \frac{n}{2}\left(\frac{3\gamma - 4}{\log 2} - \frac{3}{2}\right) + \frac{n}{2\log 2}\delta_2\left(\log_2 n\right) + o(n),$$

where γ is Euler's constant and $\delta_2(x)$ is a periodic function of period 1, mean zero and small amplitude, which is given by the Fourier series

$$\delta_2(x) = \sum_{k \neq 0} (\chi_k + 3)\Gamma\left(-1 - \chi_k\right) e^{2k\pi i x}.$$

The complex numbers χ_k are given by $\chi_k = 2k\pi i / \log 2$.

Proof. Firstly we rewrite $v_N(z)$ in the form

$$\begin{aligned} v_N(z) &= \frac{3Nz^2 - 5Nz + 2N - 4z^2 + 5z - 2}{(1-2z)^2} \\ &\quad + (1-z)(-3+4z)\sum_{k=2}^{N}\left(\frac{z^{-k}}{1-2z} - \frac{z^{-k}}{1-2z+z^k}\right) + \sum_{k=2}^{N}\frac{(1-z)^2}{\left(1-2z+z^k\right)^2}. \end{aligned} \tag{5.1}$$

Then provided $N \geq n$, $W_n = \frac{1}{2^{n-1}}[z^n]V_N(z) = \frac{1}{2^{n-1}}[z^n]v_N(z)$ and satisfies from (5.1)

$$\begin{aligned} &W_n - 2^{-1}(-2(2+n) + N(5+n)) \\ &= 2^{1-n}\sum_{k=2}^{N}[z^n](1-z)(-3+4z)\left(\frac{z^{-k}}{1-2z} - \frac{z^{-k}}{1-2z+z^k}\right) + \sum_{k=2}^{N}[z^n]\frac{(1-z)^2}{\left(1-2z+z^k\right)^2}. \end{aligned} \tag{5.2}$$

Now taking for convenience $N = n$, the left hand side becomes $W_n - (\frac{n^2}{2} + \frac{3n}{2} - 2)$, while the right hand side is

$$2^{1-n}\sum_{k=2}^{n}[z^{k+n}](1-z)(-3+4z)\left(\frac{1}{1-2z}-\frac{1}{1-2z+z^k}\right)+2^{1-n}\sum_{k=2}^{n}[z^n]\frac{(1-z)^2}{\left(1-2z+z^k\right)^2}$$

(5.3)

$$:= \Sigma_1 + \Sigma_2.$$

To compute asymptotics we follow the approach used for the longest runs given in [7, Pages 308–311]).

For readability we do not give explicit O-estimates for the error terms in our estimates below, except where directly relevant. The notation $\sim$ will be used in all cases where the respective O-terms are majorized by explicit O-terms appearing elsewhere in the calculations, and always bearing in mind that we want an overall error in Theorem 2 that is $o(n)$ in magnitude.

We deal with the second sum first Σ_2 from (5.3). Let ρ_k be the smallest positive root of the denominator polynomial $1-2z+z^k$ that lies between $1/2$ and 1. An application of the principle of the argument or Rouche's Theorem shows such a root to exist with all other roots of modulus greater than $3/4$. By dominant pole analysis, for large n and fixed k,

$$\begin{aligned}
s_{n,k} &:= [z^n]\frac{(1-z)^2}{\left(z^k-2z+1\right)^2} \\
&\sim [z^n]\left(\frac{(1-\rho_k)^2}{(z-\rho_k)^2\left(2-k\rho_k^{k-1}\right)^2}-\frac{2\,(1-\rho_k)}{(z-\rho_k)\left(2-k\rho_k^{k-1}\right)^2}\right) \\
&\sim \frac{(1-\rho_k)^2}{\left(2-k\rho_k^{k-1}\right)^2}(n+1)\rho_k^{-n-2}+2\frac{(1-\rho_k)}{\left(2-k\rho_k^{k-1}\right)^2}\rho_k^{-n-1}.
\end{aligned}$$

The denominator polynomial $1-2z+z^k$ behaves like a perturbation of $1-2z$ near $z = 1/2$, so one expects ρ_k to be approximated by the root of $1-2z = 0$ as $k \to \infty$, namely $1/2$. By "bootstrapping", see [7, Page 309]), we find that

$$\rho_k = \frac{1}{2} + 2^{-k-1} + O(k2^{-2k}). \tag{5.4}$$

Consequently,

$$\rho_k^{-n} = \left(\frac{1}{2}\left(1+2^{-k}+O(k2^{-2k})\right)\right)^{-n} \sim 2^n \exp\left(-\frac{n}{2^k}\right).$$

The use of this approximation can be justified for a wide range of values of k with $2 \le k \le n$, and n sufficiently large, (see for example [8] or [9]).

Therefore

$$[z^n]\frac{(1-z)^2}{\left(1-2z+z^k\right)^2} \sim (n+1)2^{n-2}\exp\left(-\frac{n}{2^k}\right) + 2\cdot 2^{n-2}\exp\left(-\frac{n}{2^k}\right)$$

whence see (5.3)

$$\Sigma_2 = \frac{n+3}{2}\sum_{k=0}^{n}\exp\left(-\frac{n}{2^k}\right) + o(n).$$

Now

$$\frac{n+3}{2}\sum_{k=0}^{n}\exp\left(-\frac{n}{2^k}\right) \sim \frac{n+3}{2}\left(n+1-\sum_{k=0}^{\infty}\left(1-\exp\left(-\frac{n}{2^k}\right)\right)\right).$$

It remains to estimate

$$h(n) := \sum_{k\geq 0}\left(1-\exp\left(-\frac{n}{2^k}\right)\right),$$

as $n\to\infty$. For this we use Mellin transforms and find (see [7, Appendix B.7, equation (48)] and [7, Page 311])

$$h(n) = \log_2 n + \frac{1}{2} + \frac{\gamma}{L} - \delta\left(\log_2 n\right) + O(1/n). \tag{5.5}$$

Here $L = \log 2$; γ is Euler's constant and $\delta(x)$ is a periodic function of period 1 and mean 0 and small amplitude, which is given by the Fourier series

$$\delta(x) = \frac{1}{L}\sum_{k\neq 0}\Gamma\left(-\chi_k\right)e^{2k\pi i x}.$$

Hence by using (5.5),

$$\Sigma_2 = \frac{n^2}{2} - \frac{n\log_2 n}{2} + \frac{n}{2}\left(\frac{7}{2} - \frac{\gamma}{L} + \delta\left(\log_2 n\right)\right) + o(n). \tag{5.6}$$

Now for Σ_1 in (5.3), we must consider

$$\sum_{k=2}^{n}[z^{k+n}](1-z)(-3+4z)\left(\frac{1}{1-2z} - \frac{1}{1-2z+z^k}\right).$$

We need to estimate

$$s^*_{n,k} = [z^{n+k}]\frac{(1-z)(-3+4z)}{1-2z+z^k}.$$

As in the case of the estimation of Σ_2 above,

$$s^*_{n,k} \sim c^*_k\rho_k^{-n-k} \quad \text{with} \quad c^*_k = \frac{(1-\rho_k)(-3+4\rho_k)}{\rho_k(2-k\rho_k^{k-1})}.$$

Firstly $c^*_k \sim -\frac{1}{2}(1+(k-4)2^{-k} + O\left(\frac{k}{2^{2k}}\right))$ and

$$\rho_k^{-n-k} = \left(\frac{1}{2}\left(1+2^{-k}+O(k2^{-2k})\right)\right)^{-n-k}$$

$$= 2^{n+k} \exp\left(-\frac{n+k}{2^k}\right)\left(1+O\left(\frac{kn}{2^{2k}}\right)\right)$$

whence

$$s^*_{n,k} = -2^n \exp\left(-\frac{n+k}{2^k}\right)\left(2^k + (k-4) + O\left(\frac{kn}{2^k}\right)\right).$$

Thus

$$\Sigma_1 \sim -\sum_{k=0}^{n} 2^k \left(1-\exp\left(-\frac{n+k}{2^k}\right)\right) + \sum_{k=0}^{n}(k-4)\exp\left(-\frac{n+k}{2^k}\right) := \Sigma_{1,1} + \Sigma_{1,2}. \tag{5.7}$$

Now rewrite

$$\Sigma_{1,2} = \frac{1}{2}\left(n^2 - 7n - 8\right) - \sum_{k=0}^{n}(k-4)\left(1-\exp\left(-\frac{n+k}{2^k}\right)\right).$$

By considering the last sum for $k \le 3\log_2 n$ and $k > 3\log_2 n$ we get the bound

$$\Sigma_{1,2} = \frac{1}{2}\left(n^2 - 7n\right) + O(\log^2 n). \tag{5.8}$$

It remains to estimate

$$\begin{aligned} \Sigma_{1,1} &= -\sum_{k=0}^{n} 2^k \left(1-\exp\left(-\tfrac{n+k}{2^k}\right)\right) \\ &= -\tfrac{3}{2}n(n+1) - \sum_{k=0}^{n} 2^k\left(1-\exp\left(-\tfrac{n+k}{2^k}\right) - \tfrac{n+k}{2^k}\right). \end{aligned} \tag{5.9}$$

Now by splitting the sum into $k \le \log_2 n$, $\log_2 n < k < 3\log_2 n$ and $k \ge 3\log_2 n$ we can show that

$$\sum_{k=0}^{n} 2^k \left(1-\exp\left(-\frac{n+k}{2^k}\right) - \frac{n+k}{2^k}\right) \sim \sum_{k=0}^{\infty} 2^k\left(1-\exp\left(-\frac{n}{2^k}\right) - \frac{n}{2^k}\right) := \Sigma_{1,3}. \tag{5.10}$$

The latter sum can be estimated via Mellin transforms (see [10, 11]). Its Mellin transform is

$$-\frac{x^{-s}\Gamma(s)}{1-2^{s+1}}$$

which exists in the vertical strip $\langle -2, -1\rangle$. Using the Mellin inversion formula and evaluating the integral by residues

$$\Sigma_{1,3} \sim -n\log_2 n - \frac{(-2+2\gamma+L)n}{2L} - n\delta_3(\log_2 n) \tag{5.11}$$

with

$$\delta_3(x) = \frac{1}{L}\sum_{k\neq 0}\Gamma\left(-1-\chi_k\right)e^{2k\pi i x}.$$

Combining equations (5.7) through (5.11) yields

$$\Sigma_1 = -n^2 + n\log_2 n + n\left(\delta_3(\log_2 n) + \frac{-2+2\gamma+L}{2L}\right) - 5n + o(n).$$

This together with equations (5.2), (5.3) and (5.6) gives

$$W_n = \frac{n \log_2 n}{2} + \frac{n}{2}\left(\frac{3\gamma - 4}{\log 2} - \frac{3}{2}\right) + \frac{n}{2\log 2}\delta_2\left(\log_2 n\right) + o(n),$$

where $\delta_2(x) = L(-\delta(x) + 2\delta_3(x)) = \sum_{k\neq 0}(\chi_k + 3)\Gamma\left(-1 - \chi_k\right) e^{2k\pi i x}$, as required. $\square$

References

[1] G. Andrews, Concave compositions, *Electron. J. Combin.*, **18:2**, (2011), P6.

[2] G. Andrews, Concave and convex compositions, *Ramanujan J.*, **31**, (2013), 67–82.

[3] A. Blecher, C. Brennan and A. Knopfmacher, Capacity of words, *J. Combin. Math. Combin. Comput.*, **107**, (2018), 245–258.

[4] A. Blecher, C. Brennan, A. Knopfmacher and M. Shattuck, Water capacity of permutations, submitted.

[5] A. Blecher, C. Brennan and A. Knopfmacher, Catchment area for Dyck paths, in preparation.

[6] S. Heubach, and T. Mansour, Combinatorics of compositions and words, Discrete Mathematics and its applications, CRC press, Taylor and Francis group, 2010.

[7] P. Flajolet and R. Sedgewick, Analytic Combinatorics, Cambridge University Press 2008. (Web edition available at *http://algo.inria.fr/flajolet/Publications/books.html*.)

[8] X. Gourdon and H. Prodinger, A generating function approach to random subgraphs of the n-cycle, *Discr. Math.* **169** (1997) 227–232.

[9] D. E. Knuth, The average time for carry propagation, *Indagationes Math.* **40** (1978) 238–242.

[10] A. M. Odlyzko and B. Richmond, On the compositions of an integer, *Lecture Notes in Math.* **829** (1980) 199–210.

[11] W. Szpankowski, Average Case Analysis of Algorithms on Sequences, Wiley, 2001.

A. Blecher, 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

E-mail address: Aubrey.Blecher@wits.ac.za

C. Brennan, 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

E-mail address: Charlotte.Brennan@wits.ac.za

A. Knopfmacher, 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

E-mail address: Arnold.Knopfmacher@wits.ac.za