THE NECKLACE PROCESS: A GENERATING FUNCTION APPROACH

BENJAMIN HACKL AND HELMUT PRODINGER

ABSTRACT. The "necklace process", a procedure constructing necklaces of black and white beads by randomly choosing positions to insert new beads (whose color is uniquely determined based on the chosen location), is revisited. This article illustrates how, after deriving the corresponding bivariate probability generating function, the characterization of the asymptotic limiting distribution of the number of beads of a given color follows as a straightforward consequence within the analytic combinatorics framework.

1. INTRODUCTION

We consider the following process (illustrated in Figure 1) for constructing necklaces with two-colored beads:

– We start with , the necklace with one black and one white bead.
– New beads can be added between any two adjacent beads. The color of the new bead is determined by the color of those two beads: The new bead is white if and only if its two neighbors are black.

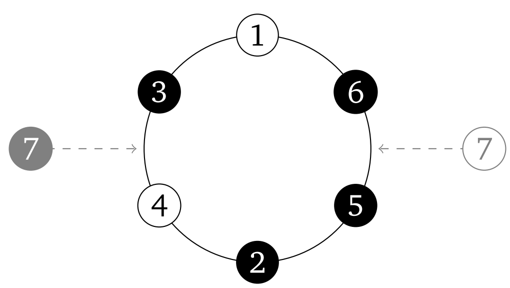

FIGURE 1. Construction of a necklace by the necklace process. The numbers on the beads correspond to the order in which they were inserted into the necklace.

Motivated by a simple network communication model, this "necklace process" was analyzed in [5]. Further variants of this process were discussed in [8], where an elegant approach

2010 *Mathematics Subject Classification.* 60C05; 05A15, 05A16.

Key words and phrases. Necklace process; bivariate generating function; quasi-power theorem.

B. Hackl is supported by the Austrian Science Fund (FWF): P 28466-N35 and by the Karl Popper Kolleg "Modeling-Simulation-Optimization" funded by the Alpen-Adria-Universität Klagenfurt and by the Carinthian Economic Promotion Fund (KWF). This paper has been written while he was a visitor at Stellenbosch University.

using Pólya urns to model the edges connecting the beads was pursued. The parameters investigated in these articles are the number of white beads in a random necklace of size n (i.e., consisting of n beads), as well as the number of runs of black beads of given length.

The purpose of this brief note is to provide an alternative access to the analysis of the number of beads of a given color in the necklace process by focusing on an appropriate generating function and using tools from analytic combinatorics. A similar approach has been successfully employed in, for example, [6, 7].

In Section 2 we briefly discuss the combinatorial structure surrounding the necklace process and elaborate how many different necklaces of given size (i.e., consisting of a given amount of beads) can be constructed. Then, in Section 3 we analyze the number of white and black beads in a random necklace of given size. Our main result is given in Theorem 1, which is an explicit formula for the bivariate probability generating function with respect to the number of white beads—which has a surprisingly nice closed form. Apart from some additional remarks on the structure of this generating function, we then show in Corollaries 3.1 and 3.2 how the qualitative results concerning the number of black and white beads obtained in [5] (expectation, variance, limiting distribution) are a straightforward consequence of the explicitly known bivariate generating function.

2. Number of Necklaces

Before diving straight into the analysis of the number of beads of a given color, for the sake of completeness, we briefly discuss the combinatorial structure of the objects we are constructing.

While it is rather easy to see that there are $(n-1)!$ possible necklace constructions[1] for a necklace of size n, many of those constructions yield the same necklace. Note that the number of different necklaces of size n is enumerated by sequence A000358 in [9]. We use the analytic combinatorics framework in order to analyze this quantity in detail.

Proposition 2.1. *Let $\mathcal{N}$ be the combinatorial class containing all different necklaces constructed by the necklace process. The corresponding ordinary generating function $N(z)$ enumerating these necklaces with respect to size is given by*

$$N(z) = \sum_{k\geq 1} \frac{\varphi(k)}{k} \log\left(\frac{1-z^k}{1-z^k-z^{2k}}\right), \tag{1}$$

where $\varphi(k)$ is Euler's totient function. Asymptotically, the number of necklaces of size n is given by

$$[z^n]N(z) = \left(\frac{1+\sqrt{5}}{2}\right)^n n^{-1} + O\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n/2} n^{-1}\right). \tag{2}$$

[1] Starting with , the necklace of size 2, there are 2 possible positions for a new bead. In the new necklace there are now 3 positions to choose from. Inductively, this proves that there are possible $(n-1)!$ construction processes for necklaces with n beads.

Proof. The generating function (1) can directly be obtained by means of the machinery provided by the symbolic method (see Chapter I and in particular Theorem I.1 in [2]). In fact, we can construct the combinatorial class $\mathcal{N}$ as

$$\mathcal{N} = \mathrm{Cyc}(\bigcirc \times \bullet^{+}),$$

where $\bigcirc$ and $\bullet^{+}$ represent the combinatorial classes for a single white and a non-empty sequence of black beads, respectively. Translating the construction of the combinatorial class $\mathcal{N}$ into the language of generating functions then immediately yields (1).

In order to obtain the asymptotic growth of the coefficients of $N(z)$, we use singularity analysis (see [1], [2, Chapter VI]), which requires us to identify the dominant singularities of $N(z)$, i.e., the singularities with minimal modulus.

In fact, by observing that all $\zeta \in \mathbb{C}$ satisfying $\zeta^k = \frac{-1\pm\sqrt{5}}{2}$ are roots of $1 - z^k - z^{2k} = 0$, it is easy to see that $N(z)$ has a unique dominant singularity located at $z = \frac{-1+\sqrt{5}}{2}$ which comes from the first summand of $N(z)$ in (1). Extracting the coefficient growth provided by the first summand and observing that the singularity with the next-larger modulus is located at $z = \sqrt{(-1+\sqrt{5})/2}$ (and comes from the second summand), we obtain (2). □

3. Beads of Equal Color

Let $n \geq 2$ and let W_n and B_n denote the random variables modeling the number of white and black beads in a necklace of size n that is constructed uniformly at random, respectively.

The fact that $W_n + B_n = n$ allows us to concentrate our investigation on W_n. Results from the characterization of W_n can be translated directly to B_n. Let $W(z,u)$ denote the shifted bivariate probability generating function corresponding to W_n, that is

$$W(z,u) = \sum_{n,k\geq 1} \mathbb{P}(W_n = k) z^{n-1} u^k.$$

In contrast to previous works on the necklace process, we give an explicit formula for the bivariate probability generating function $W(z,u)$.

Theorem 1. *The shifted bivariate probability generating function $W(z,u)$ corresponding to the random variables W_n modeling the number of white beads in a random necklace of size n is given by*

$$W(z,u) = \frac{u}{\sqrt{1-u}\coth(z\sqrt{1-u}\,) - 1}, \tag{3}$$

or, equivalently, by

$$W(z,u) = u\frac{\exp(z\alpha) - \exp(-z\alpha)}{\exp(z\alpha)(\alpha-1) + \exp(-z\alpha)(\alpha+1)} = u\frac{\exp(2z\alpha) - 1}{\exp(2z\alpha)(\alpha-1) + \alpha + 1}, \tag{4}$$

where $\alpha := \sqrt{1-u}$.

Proof. Analogously to the approach in [5] we also see the number of white beads in the context of a Markov chain with properties

$$\mathbb{P}(W_{n+1} = k \mid W_n = k) = \frac{2k}{n}, \qquad \mathbb{P}(W_{n+1} = k+1 \mid W_n = k) = 1 - \frac{2k}{n}. \tag{5}$$

This is because when choosing the position for the new bead uniformly at random among all possible n positions, the color of the bead to be inserted is white (which would increase the total number of white beads) if and only if we chose one of the positions not adjacent to a white bead. Given that there are k white beads among the n beads, there are $n - 2k$ such positions.

Let $w_{n,k} := \mathbb{P}(W_n = k)$ and let $w_n(u)$ denote the probability generating function for W_n. With the help of (5) and the law of total probability we find

$$\begin{aligned} w_{n+1,k} &= \mathbb{P}(W_{n+1} = k \mid W_n = k) w_{n,k} + \mathbb{P}(W_{n+1} = k \mid W_n = k-1) w_{n,k-1} \\ &= \frac{2}{n} k w_{n,k} - \frac{2}{n}(k-1) w_{n,k-1} + w_{n,k-1}. \end{aligned} \tag{6}$$

It is interesting to note that structurally, these probabilities $w_{n,k}$ are strongly connected to Eulerian numbers (see [4, Section 5.1.3.]): By setting $e_{n,k} := w_{n,k}/(n-1)!$ in (6) we obtain

$$e_{n,k} = 2k e_{n-1,k} + (n+1-2k) e_{n-1,k-1},$$

which strongly resembles the recurrence for Eulerian numbers as given in [4, 5.1.13.(2)].

By multiplication of (6) with u^k and summing over k this translates to

$$w_{n+1}(u) = \frac{2u(1-u)}{n} w_n'(u) + u w_n(u) \tag{7}$$

for $n \geq 2$ with initial value $w_2(u) = u$. Note that the bivariate probability generating function $W(z,u)$ can be expressed by the $w_n(u)$ by means of $W(z,u) = \sum_{n\geq 2} w_n(u) z^{n-1}$. After multiplying (7) with z^{n-1} and summation over $n \geq 2$ we find that $W(z,u)$ satisfies the first order linear partial differential equation

$$\partial_z W(z,u)(1 - zu) = 2u(1-u)\partial_u W(z,u) + uW(z,u) + u, \tag{8}$$

with the condition $W(0,u) = 0$. Solving this PDE (e.g., by means of the method of characteristics, or with the help of a computer algebra system) yields (3). The alternate form (4) follows from rewriting $\coth(z) = \frac{\exp(z)+\exp(-z)}{\exp(z)-\exp(-z)}$. □

Because of the particularly nice shape of the bivariate probability generating function $W(z,u)$ we are able to use the machinery around Hwang's quasi-power theorem (see [3], [2, Section IX.7]) in order to find the characterization of W_n from [5] as an immediate corollary.

Corollary 3.1 ([5, Section 3]). *The expected number of white beads in a necklace of size n constructed uniformly at random and the corresponding variance are given by*

$$\mathbb{E}W_n = \frac{n}{3} \qquad \text{and} \qquad \mathbb{V}W_n = \frac{2n}{45} \tag{9}$$

for $n \geq 6$. Furthermore, W_n is asymptotically normally distributed in the sense that for all $x \in \mathbb{R}$ we have

$$\mathbb{P}\left(\frac{W_n - n/3}{\sqrt{2n/45}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt + O(n^{-1/2}). \tag{10}$$

Proof. This explicit form of the bivariate probability generating function allows us to use standard techniques from analytic combinatorics in order to obtain the expected value $\mathbb{E}W_n$ and the variance $\mathbb{V}W_n$. By construction, $\mathbb{E}W_n$ is the coefficient of z^{n-1} in $\partial_u W(z,1)$. We find

$$\partial_u W(z,1) = \frac{(z^2 - 3z - 3)z}{3(1-z)^2} = z + \sum_{n\geq 3} \frac{n}{3} z^{n-1},$$

which proves $\mathbb{E}W_n = n/3$ for $n \geq 3$. Similarly, we can use the second partial derivative of $W(z,u)$ with respect to u in order to extract the second factorial moment, $\mathbb{E}(W_n(W_n - 1))$. Together with the well-known identity

$$\mathbb{V}W_n = \mathbb{E}(W_n(W_n - 1)) + \mathbb{E}W_n - (\mathbb{E}W_n)^2$$

this allows to verify that $\mathbb{V}W_n = 2n/45$ for $n \geq 6$.

Finally, the asymptotically normal limiting distribution is obtained immediately by using the explicit formula for $F(z,u) := zW(z,u)$ and applying [2, Theorem IX.12]. The corresponding necessary conditions are all checked easily:

- *Analytic perturbation.* We have

$$A(z,u) = 0, \quad B(z,u) = z^2 u, \quad C(z,u) = z(\sqrt{1-u}\coth(z\sqrt{1-u}) - 1)$$

  and $\alpha = 1$. $A(z,u)$ and $B(z,u)$ are obviously entire functions, and $C(z,u)$ is analytic in the domain $\{z \in \mathbb{C} : |z| < 2\pi\} \times \{u \in \mathbb{C} : |u - 1| < \varepsilon\}$ for some $\varepsilon > 0$. To see this, observe that $\coth(z)$ is an odd function with Laurent expansion

$$\coth(z) = \frac{1}{z} + \frac{z}{3} - \frac{z^3}{45} + \frac{2z^5}{945} + O(z^7),$$

  which means that only even powers of $\sqrt{1-u}$ occur. Hence, there is no branching point of the square root at $u = 1$. Also, $C(z,1) = 1 - z$ has a unique simple root at $\rho = 1$ with $B(\rho,1) = 1 \neq 0$.
- *Non-degeneracy.* It is straightforward to check $\partial_z C(\rho,1) \cdot \partial_u C(\rho,1) = 1/3 \neq 0$.
- *Variability.* By explicitly computing the variance above we already computed that the linear term does not vanish.

This proves (10) and concludes this proof. □

As mentioned above, the characterization W_n can immediately be carried over to B_n. Rewriting $B_n = n - W_n$ proves the following result.

Corollary 3.2. *The expected number of black beads in a necklace of size n constructed uniformly at random and the corresponding variance are given by*

$$\mathbb{E}B_n = \frac{2n}{3}, \qquad \text{and} \qquad \mathbb{V}B_n = \frac{2n}{45} \tag{11}$$

for $n \geq 6$. Furthermore, B_n is asymptotically normally distributed.

As a side effect of Theorem 1 we are also able to extract more information on the probability generating functions $w_n(u)$.

Corollary 3.3. *Let $\alpha = \sqrt{1-u}$. The shifted probability generating functions $w_n(u)/u$ can be expressed by means of even polynomials $r_n(\alpha)$ satisfying the recurrence relation*

$$r_n(\alpha) = \frac{(2\alpha)^{n-2}}{(n-1)!} + (1-\alpha)\sum_{k=0}^{n-2} \frac{(2\alpha)^k}{(k+1)!} r_{n-1-k}(\alpha), \tag{12}$$

for $n \geq 2$ with initial value $r_1(\alpha) = 0$.

Proof. From (4), and by the definition of $W(z,u)$, we find

$$\frac{\exp(2z\alpha)-1}{\exp(2z\alpha)(\alpha-1)+\alpha+1} = \sum_{n\geq 2} z^{n-1} w_n(u)/u = \sum_{n\geq 2} z^{n-1} r_n(\alpha).$$

After multiplying with the denominator and extracting the coefficient of z^{n-1} on both sides, we obtain

$$\frac{(2\alpha)^{n-1}}{(n-1)!} = (\alpha+1) r_n(\alpha) + (\alpha-1)\sum_{k=0}^{n-1} \frac{(2\alpha)^k}{k!} r_{n-k}(\alpha),$$

valid for $n \geq 2$. Rearranging this equation then yields (12). □

References

[1] Philippe Flajolet and Andrew Odlyzko, *Singularity analysis of generating functions*, SIAM J. Discrete Math. **3** (1990), 216–240.

[2] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.

[3] Hsien-Kuei Hwang, *On convergence rates in the central limit theorems for combinatorial structures*, European J. Combin. **19** (1998), 329–343.

[4] Donald E. Knuth, *The art of computer programming. Vol. 3: Sorting and searching*, second ed., Addison-Wesley, Reading, MA, 1998.

[5] Colin Mallows and Larry Shepp, *The necklace process*, J. Appl. Probab. **45** (2008), no. 1, 271–278.

[6] Basile Morcrette, *Fully analyzing an algebraic Pólya urn model*, LATIN 2012: theoretical informatics, Lecture Notes in Comput. Sci., vol. 7256, Springer, Heidelberg, 2012, pp. 568–581.

[7] Basile Morcrette and Hosam M. Mahmoud, *Exactly solvable balanced tenable urns with random entries via the analytic methodology*, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12), Discrete Math. Theor. Comput. Sci. Proc., AQ, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012, pp. 219–232.

[8] Toshio Nakata, *Necklace processes via Pólya urns*, J. Appl. Probab. **46** (2009), no. 1, 284–295.

[9] *The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org`, 2015.

(Benjamin Hackl) Institut für Mathematik, Alpen-Adria-Universität Klagenfurt, Universitätsstrasse 65–67, 9020 Klagenfurt, Austria
E-mail address: benjamin.hackl@aau.at

(Helmut Prodinger) Department of Mathematical Sciences, Stellenbosch University, 7602 Stellenbosch, South Africa
E-mail address: hproding@sun.ac.za