



**HAL**  
open science

## Elephant polynomials

Hélène Guérin, Lucile Laulin, Kilian Raschel

► **To cite this version:**

Hélène Guérin, Lucile Laulin, Kilian Raschel. Elephant polynomials. *Aequationes Mathematicae*, 2024, 10.1007/s00010-024-01095-9. hal-04386848v2

**HAL Id: hal-04386848**

**<https://hal.science/hal-04386848v2>**

Submitted on 1 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Elephant polynomials

Hélène Guérin\*, Lucile Laulin† and Kilian Raschel‡

January 19, 2024

## Abstract

In this note, we study a family of polynomials that appear naturally when analysing the characteristic functions of the one-dimensional elephant random walk. These polynomials depend on a memory parameter  $p$  attached to the model. For certain values of  $p$ , these polynomials specialise to classical polynomials, such as the Chebychev polynomials in the simplest case, or generating polynomials of various combinatorial triangular arrays (e.g. Eulerian numbers). Although these polynomials are generically non-orthogonal (except for  $p = \frac{1}{2}$  and  $p = 1$ ), they have interlacing roots. Finally, we relate some algebraic properties of these polynomials to the probabilistic behaviour of the elephant random walk. Our methods are reminiscent of classical orthogonal polynomial theory and are elementary.

**Keywords:** Elephant random walk; Explicit distribution; Orthogonal polynomials; Interlacing roots; Eulerian numbers

**AMS MSC 2020:** 60E05, 60E10, 60J10, 60G50, 05A10

## 1 Introduction and main results

**A one-parameter family of polynomials.** In this paper, our main objective is to study a family of polynomials defined as follows:  $R_1(x) = x$  and for  $n \geq 1$ ,

$$R_{n+1}(x) = xR_n(x) - \frac{a}{n}(1-x^2)R'_n(x), \quad (1)$$

where  $a \in \mathbb{R}$  is some parameter. Due to a strong connection with the elephant random walk (ERW) when  $a \in [-1, 1]$ , which we shall now present, we call them *elephant polynomials*. The first three elephant polynomials are given by

$$R_1(x) = x, \quad R_2(x) = (a+1)x^2 - a \quad \text{and} \quad R_3(x) = x((a+1)^2x^2 - a(a+2)). \quad (2)$$

---

\*Université du Québec à Montréal (UQAM), Département de Mathématiques, Canada. Funded by Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grant (RGPIN-2020-07239); [guerin.helene@uqam.ca](mailto:guerin.helene@uqam.ca)

†Université Paris Nanterre, Modal'X, UMR CNRS 9023, UPL, France, and FP2M, CNRS FR 2036; [lucile.laulin@math.cnrs.fr](mailto:lucile.laulin@math.cnrs.fr)

‡Université d'Angers, CNRS, Laboratoire Angevin de Recherche en Mathématiques, France. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 759702 and from Centre Henri Lebesgue, programme ANR-11-LABX-0020-01; [raschel@math.cnrs.fr](mailto:raschel@math.cnrs.fr)

**The elephant random walk.** The one-dimensional elephant random walk  $(S_n)_{n \geq 0}$  is defined as follows [15]. We denote by  $(X_n)_{n \geq 0}$  its successive steps. The elephant starts at the origin at time zero:  $S_0 = 0$ . For the first step  $X_1$ , the elephant moves one step to the right (+1) with probability  $q$  or one step to the left (-1) with probability  $1 - q$ , for some  $q$  in  $[0, 1]$ . The next steps are performed by choosing uniformly at random an integer  $k$  among the previous times. Then the elephant moves exactly in the same direction as at time  $k$  with probability  $p \in [0, 1]$ , or in the opposite direction with probability  $1 - p$ . In other words, defining for all  $n \geq 1$ ,

$$X_{n+1} = \begin{cases} +X_k & \text{with probability } p, \\ -X_k & \text{with probability } 1 - p, \end{cases}$$

with  $k \sim \mathcal{U}\{1, \dots, n\}$ , the position of the ERW at time  $n + 1$  is given by  $S_{n+1} = S_n + X_{n+1}$ . The probability  $q$  is called the first step parameter (in this paper we will take  $q = \frac{1}{2}$ ) and  $p$  the memory parameter of the ERW.

**The characteristic functions as trigonometric polynomials.** The characteristic function of the process at time  $n$  is defined by

$$\varphi_n(t) = \mathbb{E}[e^{itS_n}]. \quad (3)$$

Taking  $q = \frac{1}{2}$ , we have  $\varphi_1(t) = \cos t$  and will justify later on that for  $n \geq 1$ ,

$$\varphi_{n+1}(t) = \cos(t)\varphi_n(t) + \frac{a}{n} \sin(t)\varphi_n'(t), \quad (4)$$

with  $a = 2p - 1 \in [-1, 1]$ . The sequence  $\{\varphi_n(t)\}_{n \geq 1}$  naturally defines a sequence of polynomials  $\{R_n(x)\}_{n \geq 1}$  via the formula

$$\varphi_n(t) = R_n(\cos t). \quad (5)$$

Using (4), one immediately obtain the recurrence relation (1) mentioned at the beginning of the paper. Although the elephant random walk model is classical and well studied in the literature (see e.g. [1, 2, 3, 5, 6, 7, 8, 10, 11, 14, 15]), the properties of the polynomials we will prove in this paper have remained unnoticed.

Interestingly, specialising the parameter  $a$  to some particular values, we recover classical orthogonal or combinatorial polynomials. Our main results in this direction can be summarized in Table 1. The distribution of  $S_n$  in these specific cases is consistent with [3] on the number of returns to zero, and the fact that the random walk is positive recurrent for  $a < \frac{1}{2}$ .

The very special form of the polynomials  $R_n$  at  $a = 0$  and  $a = 1$  in Table 1 admits a direct probabilistic interpretation. For  $a = 0$ , the ERW's memory parameter  $p = \frac{1+a}{2} = \frac{1}{2}$ , and  $\varphi_n(t)$  is the characteristic function of the classical simple random walk on  $\mathbb{Z}$ , namely  $\cos(t)^n$ . It is also clear that for  $a = 1$ , one should get for  $\varphi_n(t)$  the Chebychev polynomial of the first kind, namely,  $\varphi_n(t) = \cos(nt)$ , as from a probabilistic point of view the ERW's memory parameter  $p = 1$ , this means that at time  $n$  the elephant is either at  $n$  or  $-n$ , with probability  $\frac{1}{2}$  each. On the other hand, we don't have a probabilistic interpretation in the other cases appearing in Table 1, for instance  $a = -1$  ( $p = 0$ ) and  $a = -\frac{1}{2}$  ( $p = \frac{1}{4}$ ). Note that from a probabilistic viewpoint, it does not make sense to take  $a > 1$  or  $a < -1$  (which corresponds to  $p > 1$  or  $p < 0$ ); however, the recursive definition (1) is well defined for any value of  $a$ .

In addition to the results presented in Table 1, which only concern a few values of  $a$ , we will prove structural results on the roots of  $R_n(x)$ . Our first result analyses the situation  $a > 0$ .

$a$	$-1$	$-\frac{1}{2}$	$0$	$1$	$\pm\infty$
$R_n(x)$	<a href="#">A101280</a> (Prop. 8)	<a href="#">A034839</a> (Prop. 6)	$x^n$ (Lem. 5)	Chebychev (Lem. 5)	<a href="#">A008293</a> (Prop. 10)
$\mathbb{P}(S_n = k)$	$\frac{A(n-1, \frac{n+k}{2})}{(n-1)!}$ (Prop. 9) <a href="#">[12, 4]</a>	$\frac{1}{2^{2n-1}} \binom{2n}{n+k}$ (Prop. 7)	$\frac{1}{2^n} \binom{n}{\frac{n+k}{2}}$ (simple RW)	$\frac{\delta_{k,+n} + \delta_{k,-n}}{2}$	n.a.

Table 1: For certain values of  $a$ , the polynomials  $R_n(x)$  specialise to classical polynomials or classical combinatorial sequences, and the exact distribution of the random walk is computed. In the last row, the quantities  $A(n, k)$  stand for the Eulerian numbers, see [A008292](#) in [\[16\]](#); by definition, they count increasing rooted trees with  $n + 1$  nodes and  $k$  leaves.

**Proposition 1.** *For  $a > 0$  and  $n \geq 1$ ,  $R_n$  admits  $n$  real roots, which are mutually distinct and on  $(-1, 1)$ . Moreover, the zeros of  $R_n$  and the zeros of  $R_{n+1}$  interlace.*

See [Figure 1](#) for an illustration of [Proposition 1](#). When  $a = 0$  we have  $R_n(x) = x^n$  (see

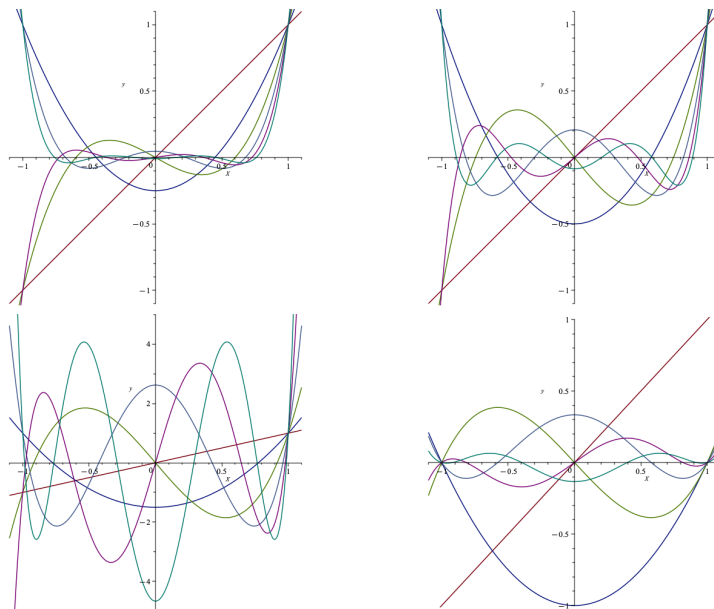


Figure 1: According to [Proposition 1](#), the first six polynomials  $R_1(x), \dots, R_6(x)$  have interlaced roots. Top left display:  $a = \frac{1}{4}$ ; top right:  $a = \frac{1}{2}$ , these are the Chebychev polynomials; bottom left:  $a = \frac{3}{2}$ ; bottom right:  $a = +\infty$ .

(1)), so that all roots collapse at 0. In the result below, we prove that the behavior of the roots dramatically changes when  $a < 0$ , as they all become purely imaginary. However, they still possess an interlacing property.

**Proposition 2.** *Define  $S_n(x) = (-i)^n R_n(ix)$ . For  $a < 0$  and  $n \geq 1$ ,  $S_n(x)$  is a real polynomial, which admits  $n$  real roots if  $a \neq -1$ , and  $|n - 2|$  real roots if  $a = -1$ , which are mutually distinct in all cases. Moreover, the zeros of  $S_n$  and the zeros of  $S_{n+1}$  interlace.*

See [Figure 2](#) for an illustration of [Proposition 2](#).

The interlacing property is most classical for orthogonal polynomials [\[17\]](#), so it is useful to notice that:

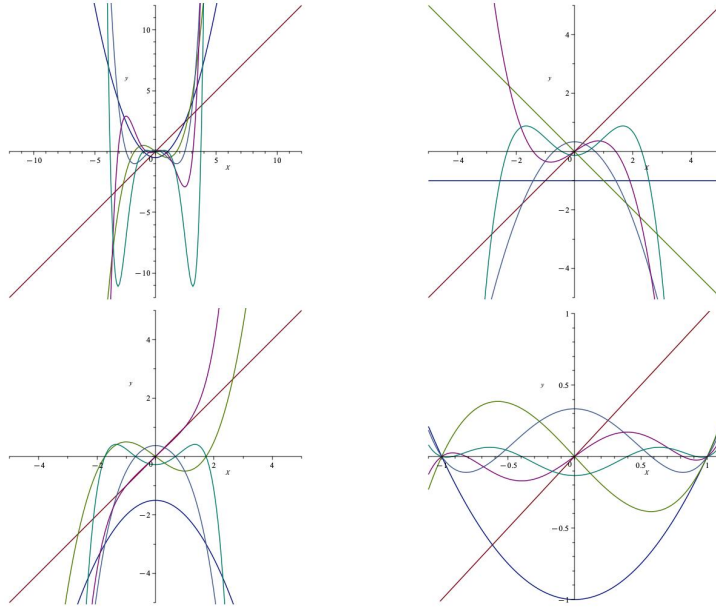


Figure 2: From top left to bottom right, the first six polynomials  $S_1(x), \dots, S_6(x)$  for  $a = -\frac{1}{2}, -1, -\frac{3}{2}, -\infty$  have interlaced roots, according to Proposition 2 (except  $S_2(x) = -1$  in the case  $a = -1$ )

**Proposition 3.** *Except for  $a = 0$  and  $a = 1$ , the  $R_n(x)$  and  $S_n(x)$  are not orthogonal.*

Finally, from a probabilistic point of view, it is natural to compute the asymptotics of  $R'_n(1)$  as  $n \rightarrow \infty$ , in order to guess limit theorems for the elephant random walk. Indeed, if  $a = 0$ , the model reduces to the classical simple random walk, for which the most classical central limit theorem yields

$$\varphi_n\left(\frac{t}{\sqrt{n}}\right) = R_n\left(\cos\left(\frac{t}{\sqrt{n}}\right)\right) \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}.$$

When  $n \rightarrow \infty$ ,

$$R_n\left(\cos\left(\frac{t}{\sqrt{n}}\right)\right) = R_n\left(1 - \frac{t^2}{2n} + O\left(\frac{1}{n^2}\right)\right) = 1 - R'_n(1)\frac{t^2}{2n} + (R'_n(1) + 3R''_n(1))\frac{t^4}{24n^2} + \dots,$$

so the asymptotics of the derivatives of  $R_n(x)$  at  $x = 1$  should reflect the correct scaling to have a central limit theorem.

**Proposition 4.** *For all  $n \geq 1$  and  $a \in [0, 1]$ ,*

$$R'_n(1) = \frac{\Gamma(n+2)\Gamma(2a) - \Gamma(n+1+2a)}{(1-2a)\Gamma(2a)\Gamma(n+1)}. \quad (6)$$

Moreover, we have the following asymptotics as  $n \rightarrow \infty$ :

$$R'_n(1) \sim \begin{cases} \frac{n}{1-2a} & \text{if } a < \frac{1}{2}, \\ n \ln(n) & \text{if } a = \frac{1}{2}, \\ \frac{n^{2a}}{(2a-1)\Gamma(2a)} & \text{if } a > \frac{1}{2}. \end{cases}$$

As  $R'_n(1)$  represents the second-order moment of the elephant random walk, the proposition above aligns with the initial (and by now well-established) observations regarding the asymptotic behavior of this moment, such as in [15]. The values of the higher-order derivatives  $R''_n(1)$ , etc., seems to be fairly more complicated.

## 2 Interlacing property

We start with stating elementary properties of the  $R_n(x)$ .

**Lemma 5.** *For all  $a \in \mathbb{R}$ ,  $R_n$  is odd (resp. even) for odd (resp. even) values of  $n$ , and one has  $R_n(1) = 1$  and  $R_n(-1) = (-1)^n$ . For  $a \in \mathbb{R} \setminus \{-1\}$ , the polynomial  $R_n$  has degree  $n$  and dominant coefficient  $(a+1)^{n-1}$ . Moreover, for  $a = 0$  one has  $R_n(x) = x^n$  and for  $a = 1$ ,  $R_n(x) = T_n(x)$ , the  $n$ th Chebychev polynomial of the first kind. For  $a = -1$ , one has  $R_1(x) = x$  and for  $n \geq 2$ ,  $R_n(x)$  has degree  $n - 2$ , with dominant coefficient  $\frac{2^{n-2}}{(n-1)!}$ . Furthermore, when  $a > 0$  the coefficients of the polynomial  $R_n$  have alternating signs, and when  $a \in [-1, 0]$  the coefficients are nonnegative. Viewed as a polynomial in  $a$ ,  $R_n$  has degree  $n - 1$ .*

*Proof.* The above lemma is obvious, with perhaps the exception of the connection with Chebychev polynomials. By definition  $T_n(\cos t) = \cos(nt)$ , and then it is clear that  $\cos(nt)$  satisfies (4) as well as the initial condition, when  $a = 1$ . The sign of the coefficients is easily obtained by induction.  $\square$

*Proof of the identity (4).* Remember that  $\varphi_n(t)$  is the characteristic function of the position  $S_n$  of the elephant random walk at time  $n$ , see (3). Let us denote by  $(\mathcal{F}_n)_{n \geq 0}$  the natural filtration of  $(S_n)_{n \geq 0}$ . It is known that by simple calculations we have  $\mathbb{P}(X_{n+1} = 1 | \mathcal{F}_n) = \frac{1}{2}(1 + a \frac{S_n}{n})$ . Then,

$$\begin{aligned} \varphi_{n+1}(t) &= \mathbb{E}[e^{itS_{n+1}}] \\ &= \mathbb{E}[e^{itS_n} \mathbb{E}[e^{itX_{n+1}} | \mathcal{F}_n]] \\ &= \mathbb{E}\left[ e^{itS_n} \left( \frac{1}{2} \left( 1 + a \frac{S_n}{n} \right) e^{it} + \frac{1}{2} \left( 1 - a \frac{S_n}{n} \right) e^{-it} \right) \right] \\ &= \frac{e^{it} + e^{-it}}{2} \mathbb{E}[e^{itS_n}] + a \frac{e^{it} - e^{-it}}{2n} \mathbb{E}[S_n e^{itS_n}] \\ &= \cos(t) \varphi_n(t) + \frac{a}{n} \sin(t) \varphi'_n(t). \end{aligned} \quad \square$$

*Proof of Proposition 1.* The techniques used are elementary and reminiscent of orthogonal polynomials theory, see [17]. Let us denote the  $n$  (a priori complex and non-distinct) roots of  $R_n$  by  $\alpha_1^{(n)}, \dots, \alpha_n^{(n)}$ . We will prove by induction that for all  $n \geq 1$ , the  $\alpha_k^{(n)}$  are real, are in  $(-1, 1)$  and satisfy  $\text{sign}(R'_n(\alpha_k^{(n)})) = (-1)^{k+n}$ . In particular, they should be simple roots.

This is easily verified for  $n = 1$ , with  $R_1(x) = x$  and  $\alpha_1^{(1)} = 0$ . Let us now assume that the previous assumption holds for some  $n$ . Using (1), one immediately obtain that

$$R_{n+1}(\alpha_k^{(n)}) = -\frac{a}{n} (1 - (\alpha_k^{(n)})^2) R'_n(\alpha_k^{(n)}).$$

In particular, since  $a > 0$  we deduce that

$$\text{sign}(R_{n+1}(\alpha_k^{(n)})) = (-1)^{k+n+1}. \quad (7)$$

Introduce  $\alpha_0^{(n)} = -1$  and  $\alpha_{n+1}^{(n)} = 1$ . Note that (7) is true for  $k = 0$  and  $k = n + 1$  as well, using  $R_{n+1}(1) = 1$  and  $R_{n+1}(-1) = (-1)^{n+1}$  (see Lemma 5). We now apply to the function  $R_{n+1}$  the intermediate value theorem on the interval  $[\alpha_k^{(n)}, \alpha_{k+1}^{(n)}]$ , for any  $k$  from 0 to  $n$ . Since by (7) the signs of  $R_{n+1}(\alpha_k^{(n)})$  and  $R_{n+1}(\alpha_{k+1}^{(n)})$  are opposite, we immediately obtain the existence of a zero, which we denote by  $\alpha_{k+1}^{(n+1)}$ . By construction we obtain  $n + 1$  points in  $(-1, 1)$ , which are mutually distinct. Finally, since these zeros must be simple, the derivatives  $R'_{n+1}(\alpha_k^{(n+1)})$  must be non-zero and of alternating sign.  $\square$

*Proof of Proposition 2.* It is very similar to the one of Proposition 1. Let us first assume that  $a \in (-1, 0)$ . Using (1) we obtain the recurrence relation

$$S_{n+1}(x) = xS_n(x) + \frac{a}{n}(1+x^2)S'_n(x), \quad (8)$$

which is valid for all  $n \geq 1$ . The dominant coefficient of  $S_n(x)$  is  $(a+1)^{n-1}x^n$  by Lemma 5, so that

$$S_n(+\infty) = +\infty \quad \text{and} \quad S_n(-\infty) = (-1)^n \infty. \quad (9)$$

Following the proof of the previous proposition, let us denote the  $n$  (a priori complex and non-distinct) roots of  $S_n$  by  $\beta_1^{(n)}, \dots, \beta_n^{(n)}$ . We will prove by induction that for all  $n \geq 1$ , the  $\beta_k^{(n)}$  are real and satisfy  $\text{sign}(S'_n(\beta_k^{(n)})) = (-1)^{k+n}$ . In particular, they should be simple roots. Using (8) we find

$$S_{n+1}(\beta_k^{(n)}) = \frac{a}{n}(1 + (\beta_k^{(n)})^2)S'_n(\beta_k^{(n)}).$$

In particular, since  $a < 0$  we deduce that

$$\text{sign}(S_{n+1}(\beta_k^{(n)})) = (-1)^{k+n+1}. \quad (10)$$

Introduce  $\beta_0^{(n)} = -\infty$  and  $\beta_{n+1}^{(n)} = +\infty$ . Note that (10) is true for  $k = 0$  and  $k = n + 1$  as well, using the limits  $S_{n+1}(\pm\infty)$  computed in (9). We now apply to the function  $S_{n+1}$  the intermediate value theorem on the interval  $[\beta_k^{(n)}, \beta_{k+1}^{(n)}]$ , for any  $k$  from 0 to  $n$ . Since by (10) the signs of  $S_{n+1}(\beta_k^{(n)})$  and  $S_{n+1}(\beta_{k+1}^{(n)})$  are opposite, we immediately obtain the existence of a zero, which we denote by  $\beta_{k+1}^{(n+1)}$ . By construction we obtain  $n + 1$  points in  $\mathbb{R}$ , which are mutually distinct. Finally, since these zeros must be simple, the derivatives  $S'_{n+1}(\beta_k^{(n+1)})$  must be non-zero and of alternating sign.

If now  $a < -1$ , we prove similarly by induction that for all  $n \geq 1$ , the  $\beta_k^{(n)}$  are real and satisfy  $\text{sign}(S'_n(\beta_k^{(n)})) = (-1)^k$ , using the following changes: (9) should be replaced by  $S_n(+\infty) = (-1)^{n-1}\infty$  and  $S_n(-\infty) = -\infty$ ; Equation (10) should be modified as follows:  $\text{sign}(S_{n+1}(\beta_k^{(n)})) = (-1)^{k+1}$ .

Finally, the case  $a = -1$  is very similar to the previous situation where  $a < -1$ , with the only difference that the degree of  $S_n$  is  $n - 2$ . More precisely, one has  $S_1(x) = x$ ,  $S_2(x) = -1$  and for  $n \geq 2$ , the dominant term of  $S_n(x)$  is  $-\frac{2^{n-2}}{(n-1)!}x^{n-2}$ . The proof continues similarly as in the case  $a < -1$ .  $\square$

*Proof of Proposition 3.* Let us do the proof in the case of  $R_n(x)$ ; one would conclude in the case of the sequence  $S_n(x)$  by similar arguments. If the sequence  $R_n(x)$  is orthogonal, it has to satisfy a recurrence relation of the form

$$R_{n+1}(x) = (\alpha_n x + \beta_n)R_n(x) + \gamma_n R_{n-1}(x). \quad (11)$$

Analysing the dominant coefficient (see Lemma 5), one should have  $\alpha_n = a + 1$ . Moreover, due to the parity of the  $R_n(x)$ , see again Lemma 5, one must take  $\beta_n = 0$ . We thus deduce  $\gamma_n = -a$  from  $R_n(1) = 1$  (see Lemma 5), so that (11) becomes

$$R_{n+1}(x) = (a + 1)xR_n(x) - aR_{n-1}(x). \quad (12)$$

Actually, setting  $R_0(x) = 1$ , the identity (12) holds true for  $n = 1$  and 2; compare with (2).

We now look at (12) in the case  $n = 3$ . Computing  $R_4(x) - (a + 1)xR_3(x)$  gives a second-degree polynomial, which if (12) were true, should be proportional to  $R_2(x)$ . However, an elementary computation shows that the resultant of the polynomials  $R_4(x) - (a + 1)xR_3(x)$  and  $R_2(x)$  is given by  $\frac{a^4(a-1)^2}{9}$ . In other words, except for  $a = 0$  and  $a = 1$ , the recurrence relation (12) (and thus (11) as well) is not satisfied for  $n = 3$ .  $\square$

### 3 Special cases of the memory parameter

#### 3.1 The case $a = -\frac{1}{2}$

**Proposition 6.** *Let  $a = -\frac{1}{2}$ . For all  $n \geq 1$ , we have*

$$R_n(x) = \frac{1}{2^{n-1}} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} x^{n-2k} = \left(\frac{x-1}{2}\right)^n + \left(\frac{x+1}{2}\right)^n.$$

And indeed, for  $a = -1/2$ , with (4) and (5) we find

$$R_1(x) = x, \quad R_2(x) = \frac{x^2 + 1}{2}, \quad R_3(x) = \frac{x^3 + 3x}{2^2}, \quad R_4(x) = \frac{x^4 + 6x^2 + 1}{2^3}, \quad \text{etc.}$$

The coefficients in the expansions of the above polynomials are thus connected with the sequence [A034839](#), corresponding to the triangular array formed by taking every other term of each row of Pascal's triangle.

*Proof of Proposition 6.* It is a direct consequence of the recurrence equation (1). Indeed, an obvious computation shows that

$$\begin{aligned} \left(\frac{x-1}{2}\right)^{n+1} + \left(\frac{x+1}{2}\right)^{n+1} &= x \left( \left(\frac{x-1}{2}\right)^n + \left(\frac{x+1}{2}\right)^n \right) \\ &\quad + \frac{1-x^2}{4} \left( \left(\frac{x-1}{2}\right)^{n-1} + \left(\frac{x+1}{2}\right)^{n-1} \right), \end{aligned}$$

which corresponds to (1) with  $a = -\frac{1}{2}$ . Therefore  $R_n(x)$  and  $\left(\frac{x-1}{2}\right)^n + \left(\frac{x+1}{2}\right)^n$  have the same initial term for  $n = 1$  and satisfy the same recurrence, hence are equal for all  $n$ .  $\square$

**Proposition 7.** *Let  $a = -\frac{1}{2}$  (equivalently  $p = \frac{1}{4}$ ). For all  $n \geq 1$ ,  $-n \leq k \leq n$ ,  $k$  and  $n$  having same parity, one has*

$$\mathbb{P}(S_n = k) = \frac{1}{2^{2n-1}} \binom{2n}{n+k}. \quad (13)$$

*In particular, for any  $n \geq 1$ , we have*

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{2^{4n-1}} \binom{4n}{2n} \underset{n \rightarrow \infty}{\sim} \sqrt{\frac{2}{\pi n}}.$$



It is interesting to compare (13) with the distribution of the classical symmetric simple random walk on  $\mathbb{Z}$  (corresponding to  $p = \frac{1}{2}$  in our model):  $\mathbb{P}(S_n = k) = \frac{1}{2^n} \binom{n}{\frac{n+k}{2}}$ , it is as if all elements of the binomial coefficient should be divided by 2.

*Proof.* For the sake of conciseness, we focus on the case  $k = 0$ . Due to the periodicity of the model and for any memory parameter, one has  $\mathbb{P}(S_{2n-1} = 0) = 0$ . Moreover, by Proposition 6 we have

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_{2n}(t) dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \left( \frac{\cos t - 1}{2} \right)^{2n} + \left( \frac{\cos t + 1}{2} \right)^{2n} \right) dt, \quad (14)$$

from which the value given in Proposition 7 follows from standard integral computations (or simply taking generating functions in (14)). We obtain similarly the complete distribution of  $S_n$ , using the relation

$$\mathbb{P}(S_n = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikt} \varphi_n(t) dt. \quad \square$$

### 3.2 The case $a = -1$

Define a triangle of integers  $T(n, k)$  for  $n \geq 1$  and  $0 \leq k \leq \lfloor \frac{n-1}{2} \rfloor$  as follows:

$$T(1, 0) = 1 \quad \text{and for } n \geq 2, \quad T(n, k) = (k+1)T(n-1, k) + (2n-4k)T(n-1, k-1). \quad (15)$$

Further, for all  $n \geq 1$ , define

$$U_n(y) = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} T(n, k) y^k. \quad (16)$$

The  $T(n, k)$  and  $U_n(y)$  for  $n$  up to 7 are reproduced below:

$$\left\{ \begin{array}{ll} 1, & U_1(y) = 1, \\ 1, & U_2(y) = 1, \\ 1, 2, & U_3(y) = 1 + 2y, \\ 1, 8, & U_4(y) = 1 + 8y, \\ 1, 22, 16, & U_5(y) = 1 + 22y + 16y^2, \\ 1, 52, 136, & U_6(y) = 1 + 52y + 136y^2, \\ 1, 114, 720, 272, & U_7(y) = 1 + 114y + 720y^2 + 272y^3. \end{array} \right.$$

See [A101280](#) in the OEIS for various properties and characterizations of these numbers.

**Proposition 8.** *Let  $a = -1$  and  $U_n$  be the polynomials defined in (16). We have for all  $n \geq 2$ ,*

$$R_n(x) = \frac{(2x)^{n-2}}{(n-1)!} U_{n-1} \left( \frac{1}{4x^2} \right). \quad (17)$$

*Proof.* Let us make a change of function as in (17). A straightforward computation from (1) (with  $a = -1$ ) shows that the polynomial sequence  $(U_n)_{n \geq 0}$  defined by (17) must satisfy the following recurrence relation:  $U_1(y) = 1$  and for  $n \geq 2$ ,

$$U_n(y) = (2ny + 1 - 4y)U_{n-1}(y) + y(1 - 4y)U'_{n-1}(y).$$

On the other hand, based on (15) it is clear that the polynomials defined in (16) satisfy the same recurrence equation, so we conclude by uniqueness.  $\square$

In the OEIS [16], it is mentioned that the generating function of the  $U_n(y)$  admits the following closed-form expression:

$$\sum_{n \geq 1} U_n(y) \frac{z^n}{n!} = \sum_{n \geq 1, k \geq 0} T(n, k) y^k \frac{z^n}{n!} = \frac{C(y)(2 - C(y))}{\exp(-z\sqrt{1-4y}) + 1 - C(y)} - C(y), \quad (18)$$

where  $C(y) = \frac{1 - \sqrt{1-4y}}{2y}$  is the generating function of Catalan numbers. Notice that this equation holds for  $(y, z)$  such that  $|y| < \frac{1}{4}$  and  $z \in \mathbb{C}$ .

We now turn to the distribution of the random walk when  $a = -1$ , and give a new proof of the result. Denote by  $A(n, k)$  the Eulerian numbers, defined for  $n \geq 1$  and  $k \in \{1, \dots, n\}$ . Besides their combinatorial interpretation in terms of increasing rooted trees with  $n + 1$  nodes and  $k$  leaves, the Eulerian numbers admit the following probabilistic interpretation: for  $k \in \{1, \dots, n\}$ ,  $\frac{A(n, k)}{n!}$  is the probability that a sum of  $n$  independent uniform random variables on  $[0, 1]$  lies between  $k - 1$  and  $k$ , see [18].

**Proposition 9** ([12, 4]). *Let  $a = -1$  (equivalently  $p = 0$ ). For any  $n \geq 2$  and  $-(n - 2) \leq k \leq n - 2$  with the same parity as  $n$ , we have*

$$\mathbb{P}(S_n = k) = \frac{A(n - 1, \frac{n+k}{2})}{(n - 1)!}.$$

In particular, for any  $n \geq 1$ ,  $\mathbb{P}(S_{2n-1} = 0) = 0$  and

$$\mathbb{P}(S_{2n} = 0) = \frac{A(2n - 1, n)}{(2n - 1)!} \underset{n \rightarrow \infty}{\sim} \sqrt{\frac{3}{\pi n}},$$

where  $A(2n - 1, n)$  denotes the *central Eulerian number*, equal to  $\sum_{k=0}^n (-1)^k (n - k)^{2n-1} \binom{2n}{k}$ .

It suggests that a bijective proof of Proposition 9 should exist, meaning to find a bijection between ERW of length  $n$ , ending at altitude  $k$ , and increasing rooted trees with  $n + 1$  nodes and  $k$  leaves, with the parity condition on  $n$  and  $k$  mentioned in Proposition 9. Using the connection between ERW and Pólya urn models, let's mention that this result was already known for  $p = 0$  (see [12, Sec. 7.2.2] and [4, Lem. 2.1]). It is also interesting to note that the Pólya urn model in this case is equivalent to a time-shifted Internal Diffusion Limited Aggregation model (also called growth model visited by an explorer in [9]), for which the distribution had also been obtained in [13, Thm 1]. In these earlier articles, other techniques were used to prove the result.

*Proof of Proposition 9.* We prove the statement about the return probability  $\mathbb{P}(S_{2n} = 0)$ , the exact same proof would allow us to describe the full distribution of the model. It follows from (17) that

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_{2n}(t) dt = \frac{1}{2\pi} \frac{2^{2n-2}}{(2n - 1)!} \int_{-\pi}^{\pi} (\cos t)^{2n-2} U_{2n-1} \left( \frac{1}{4 \cos^2 t} \right) dt.$$

We now use the following identity, stated in A101280:

$$(1 + x)^{2n-2} U_{2n-1} \left( \frac{x}{(1 + x)^2} \right) = \sum_{k=0}^{2n-2} A(2n - 1, k + 1) x^k = A_{2n-1}(x),$$

where the last identity is the definition of the  $(2n - 1)$ th Eulerian polynomial. Setting  $x = e^{it}$ , we deduce that

$$(2 \cos t)^{2n-2} U_{2n-1} \left( \frac{1}{4 \cos^2 t} \right) = \exp(-2i(n - 1)t) A_{2n-1}(\exp(2it)). \quad (19)$$

(Notice that the left-hand side of (19) is an even function of  $t$ ; so should be the right-hand side, which corresponds to the well-known fact that the  $A_n$  are reciprocal polynomials.) We thus have

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{(2n-1)!} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-2i(n-1)t) A_{2n-1}(\exp(2it)) dt = \frac{A(2n-1, n)}{(2n-1)!}. \quad \square$$

The above proof (mostly (19)) shows that for  $n \geq 1$ ,

$$\varphi_{2n}(t) = e^{-2i(n-1)t} A_{2n-1}(e^{2it}) = \sum_{k=-(n-1)}^{n-1} A(2n-1, k+n) e^{2ikt}.$$

### 3.3 The case $a = \pm\infty$

In this section, we identify the polynomials

$$T_n(x) = \lim_{a \rightarrow \pm\infty} \frac{R_n(x)}{a^{n-1}}, \quad (20)$$

whose existence is guaranteed by Lemma 5 (considered as polynomials in  $a$ , the  $R_n$  have degree  $n-1$ ). To that purpose, define for any  $n \geq 0$  the polynomial  $V_n(x)$  such that

$$\tanh^{(n)}(x) = V_n(\tanh x),$$

where  $\tanh^{(n)}$  stands for the  $n$ th derivative of  $\tanh$ . For instance,

$$V_0(x) = x, \quad V_1(x) = 1 - x^2 \quad \text{and} \quad V_2(x) = 2x^3 - 2x.$$

See [A008293](#) and [A101343](#) for related sequences in the OEIS.

**Proposition 10.** *For all  $n \geq 1$ , we have*

$$T_n(x) = (-1)^{n-1} \frac{V_{n-1}(x)}{(n-1)!}. \quad (21)$$

*Proof.* Using the recurrence relation (1), we immediately obtain that the polynomials  $T_n(x)$  defined by (20) satisfy the recurrence relation

$$T_{n+1}(x) = \frac{x^2 - 1}{n} T_n'(x)$$

for  $n \geq 1$ , with the initial value  $T_1(x) = x$ . Define now the polynomials  $V_n$  by the relation (21). Using the above recurrence relation for the  $T_n$ , we deduce that the  $V_n$  satisfy the recurrence

$$V_{n+1}(x) = (1 - x^2) V_n'(x).$$

This exactly corresponds to our interpretation of the polynomials  $V_n$ . If indeed  $\tanh^{(n)}(x) = V_n(\tanh x)$ , then we have

$$\tanh^{(n+1)}(x) = (1 - \tanh^2 x) V_n'(\tanh x) = V_{n+1}(\tanh x),$$

concluding the proof. □

## 4 Connection with the probabilistic behavior of the model

In this part, our main objective is to prove Proposition 4, or its generating-function version stated below:

**Proposition 11.** *We have*

$$\sum_{n \geq 1} R'_n(1)x^n = \begin{cases} \frac{x}{(1-x)^2} \frac{1-2a(1-x)^{1-2a}}{1-2a} & \text{if } a \neq \frac{1}{2}, \\ \frac{x}{(1-x)^2} (1 - \ln(1-x)) & \text{if } a = \frac{1}{2}. \end{cases} \quad (22)$$

*Proof.* First note that  $R'_1(1) = 1$  by (2). Then take the derivative of (1) and evaluate the new identity at  $x = 1$ . This way, we obtain for  $n \geq 1$

$$R'_{n+1}(1) = R_n(1) + R'_n(1) + \frac{2a}{n} R'_n(1) = 1 + R'_n(1) \left(1 + \frac{2a}{n}\right), \quad (23)$$

where we have simplified  $R_n(1) = 1$  by Lemma 5. To proceed, we can just check that the sequence

$$\frac{\Gamma(n+2)\Gamma(2a) - \Gamma(n+1+2a)}{(1-2a)\Gamma(2a)\Gamma(n+1)},$$

which appears in the right-hand side of (6), satisfies the same recurrence as (23) and same initial value for  $n = 1$ . So it should coincide with  $R'_n(1)$ .

A more constructive proof consists in multiplying (23) by  $n$ , and to deduce that the series  $y(x) = \sum_{n \geq 1} R'_n(1)x^n$  satisfies the differential equation

$$(1-x)y'(x) - \left(\frac{1}{x} + 2a\right)y(x) = \frac{x}{(1-x)^2}$$

with initial condition ( $y(0) = 0$  and)  $y'(0) = 1$ . It is standard to solve this order-one differential equation in closed-form, and to conclude to (22). The formula (6) follows from a Taylor expansion of the series (22).  $\square$

Note that the series  $\sum_{n \geq 1} R_n^{(k)}(1)x^n$  corresponding to higher-order derivatives ( $k \geq 2$ ) admit similar, but considerably more complicated, closed-form expressions.

**Acknowledgments.** We would like to thank Alin Bostan for interesting discussions at an early stage of the project, as well as Jean Bertoin and Philippe Nadeau for their valuable comments.

## References

- [1] E. Baur and J. Bertoin. Elephant random walks and their connection to Pólya-type urns. *Phys. Rev. E*, 94:052134, 2016.
- [2] B. Bercu and L. Laulin. On the multi-dimensional elephant random walk. *J. Stat. Phys.*, 175(6):1146–1163, 2019.
- [3] J. Bertoin. Counting the zeros of an elephant random walk. *Trans. Amer. Math. Soc.*, (375):5539–5560, 2022.
- [4] J. Bertoin. Counterbalancing steps at random in a random walk. *J. Eur. Math. Soc.*, (375):1–23, 2023.

- [5] C. F. Coletti, R. Gava, and G. M. Schütz. Central limit theorem and related results for the elephant random walk. *J. Math. Phys.*, 58(5):053303, 8, 2017.
- [6] C. F. Coletti and I. Papageorgiou. Asymptotic analysis of the elephant random walk. *J. Stat. Mech. Theory Exp.*, (1):Paper No. 013205, 12, 2021.
- [7] J. C. Cressoni, M. A. A. da Silva, and G. M. Viswanathan. Amnestically induced persistence in random walks. *Phys. Rev. Lett.*, 98(7):070603, 4, 2007.
- [8] M. A. A. Da Silva, J. C. Cressoni, G. M. Schütz, G. M. Viswanathan, and S. Trimper. Non-gaussian propagator for elephant random walks. *Phys. Rev. E*, 88:022115, 2013.
- [9] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Sem. Mat. Univ. Politec. Torino*, 49(1):95–119, 1991.
- [10] H. Guérin, L. Laulin, and K. Raschel. A fixed-point equation approach for the superdiffusive elephant random walk. *arXiv*, 2308.14630, 2023.
- [11] R. Kürsten. Random recursive trees and the elephant random walk. *Phys. Rev. E*, 93(3):032111, 11, 2016.
- [12] H. M. Mahmoud. *Pólya urn models*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, 2008.
- [13] K. Mittelstaedt. A stochastic approach to Eulerian numbers. *Amer. Math. Monthly*, 127(7):618–628, 2020.
- [14] S. Qin. Recurrence and transience of multidimensional elephant random walks. *arXiv*, 2309.09795, 2023.
- [15] G. M. Schütz and S. Trimper. Elephants can always remember: Exact long-range memory effects in a non-markovian random walk. *Phys. Rev. E*, 70:045101, 2004.
- [16] N. J. A. Sloane and T. O. F. Inc. The on-line encyclopedia of integer sequences, 2020.
- [17] G. Szegő. *Orthogonal polynomials.*, volume 23 of *Colloq. Publ., Am. Math. Soc.* American Mathematical Society (AMS), Providence, RI, 1939.
- [18] S. Tanny. A probabilistic interpretation of Eulerian numbers. *Duke Math. J.*, 40:717–722, 1973.