HAL open science

Pascal's formulas and vector fields

Philippe Chassaing, Jules Flin, Alexis Zevio

▶ **To cite this version:**

Philippe Chassaing, Jules Flin, Alexis Zevio. Pascal's formulas and vector fields. 2022. hal-03821769v3

HAL Id: hal-03821769

https://hal.science/hal-03821769v3

Preprint submitted on 14 Sep 2023

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.

Pascal's formulas and vector fields

Philippe Chassaing [*], Jules Flin [†], Alexis Zevio [‡]

September 14, 2023

Abstract

We consider four examples of combinatorial triangles $(T(n,k))_{0\leq k\leq n}$ (Pascal, Stirling of both types, Euler) : through saddle-point asymptotics, their *Pascal's formulas* define four vector fields, together with their field lines that turn out to be the conjectured limit of sample paths of four well known Markov chains. We prove this asymptotic behaviour in three of the four cases.
Keywords. Markov chain, combinatorial triangle, Pascal formula, hydrodynamic limit, vector field.

Contents

1 Introduction **2**
1.1 Pascal's formulas . . . 2
1.2 Transition probabilities . . . 2
1.3 Random walk, coupon collector, chinese restaurant, and internal DLA 3
1.3.1 Simple random walk . . . 3
1.3.2 Coupon collector's problem . . . 3
1.3.3 Chinese restaurant process . . . 3
1.3.4 One-dimensional Internal Diffusion Limited Aggregation . . . 4
1.4 Simulations . . . 4
1.5 Asymptotics of sample paths, and field lines of vector fields . . . 6
1.5.1 Description of φ . . . 7

2 Time reversal and Markov property for W **8**
2.1 Time reversal . . . 8
2.2 Simple random walk . . . 10
2.3 Coupon collector's problem . . . 11
2.4 Chinese restaurant process . . . 12
2.5 One-dimensional Internal Diffusion Limited Aggregation process . . . 13

3 The limit vector field : proof of Theorem 1 **15**

[*]Institut Élie Cartan, Université de Lorraine Email: chassaingph@gmail.com
[†]Institut Élie Cartan, Université de Lorraine Email: jules.flin8@etu.univ-lorraine.fr
[‡]Institut Élie Cartan, Université de Lorraine Email: alexis.zevio1@etu.univ-lorraine.fr

4 Sample path convergence **20**
4.1 Proof of Theorem 2 : Pascal's triangle. 20
4.2 Proof of Theorem 2 : Stirling numbers of the first kind. 22
4.3 Proof of Theorem 2 : Stirling numbers of the second kind. 23
4.4 Application to the enumeration of accessible complete deterministic automata with k letters and n vertices 25

1 Introduction

1.1 Pascal's formulas

Set $S = \{(n,k) \in \mathbb{N}^2,\ 0 \le k \le n\}$, $\mathring{S} = \{(n,k) \in \mathbb{N}^2,\ 0 < k < n\}$, and let $S^\star = S \setminus \{(0,0)\}$. Besides Pascal's triangle, other triangular arrays $(T(n,k))_{(n,k)\in S}$ of interest satisfy a recursion formula similar to Pascal's formula, i.e. of the following form, for $(n,k) \in S^\star$:

$$T(n,k) = a(n,k)T(n-1,k-1) + b(n,k)T(n-1,k), \tag{1}$$

with the convention that either $(n,k) \in S$ or $T(n,k) = 0$. For instance, relation (1) holds true for the following triangular arrays :

- for Pascal's triangle, if $(a,b)(n,k) = (1,1)$;
- for Stirling numbers of the second kind, if $(a,b)(n,k) = (1,k)$;
- for Stirling numbers of the first kind, if $(a,b)(n,k) = (1,n-1)$;
- for Euler's triangle, if $(a,b)(n,k) = (n-k,k+1)$.

1.2 Transition probabilities

In view of (1), for $(n,k) \in S^\star$, consider

$$(p_0(n,k), p_1(n,k)) = \left(\frac{b(n,k)T(n-1,k)}{T(n,k)}, \frac{a(n,k)T(n-1,k-1)}{T(n,k)} \right) \tag{2}$$

as some transition probabilities from (n,k) to $(n-1,k)$, resp. to $(n-1,k-1)$. For each of these four triangular arrays, the transition probabilities

$$(p_\varepsilon(n,k))_{(\varepsilon,(n,k))\in\{0,1\}\times S^\star},$$

together with the initial state (m,ℓ), define a Markov chain $W = (W_k)_{0\le k\le n}$ with terminal state $(0,0)$. These four Markov chains are closely related to the simple random walk, the coupon collector problem, the chinese restaurant process and the one-dimensional internal DLA, respectively : they are the time-reversed versions of these processes, once these processes are conditioned to be at level ℓ at time m, as explained in the next section.

1.3 Random walk, coupon collector, chinese restaurant, and internal DLA

Consider a random process defined by $X_0 = 0$, and, for $n \geqslant 0$, $X_{n+1} = X_n + Y_{n+1}$, in which the Y_i's are Bernoulli random variables. Set

$$W_n = (m-n, X_{m-n}) \in S, \quad 0 \leq n \leq m,$$
$$w_m(t) = m^{-1} X_{\lfloor mt \rfloor} \in S, \quad 0 \leq t \leq 1,$$

and note that, by definition, $W_n = (0,0)$ if and only if $n = m$.

1.3.1 Simple random walk

Assume that $(Y_i)_{i \geq 1}$ is a Bernoulli process, i.e. a sequence of i.i.d. Bernoulli random variables with parameter $p \in (0,1)$. Then

Proposition 1. *The stochastic process $W = (W_n)_{0 \leq n \leq m}$, conditioned to $W_0 = (m, \ell)$, or equivalently to $X_m = \ell$, is the Markov chain with transition probabilities $(p_\varepsilon(n,k))_{(\varepsilon,n,k) \in \{0,1\} \times S^\star}$ related to Pascal's triangle. Its distribution does not depend on p.*

This result goes back at least to Kennedy [Ken75], or even to the introduction of the concept of sufficiency by Fisher around 1920 [Sti73]. We recall its proof at Subsection 2. In the next cases, the Bernoulli random variables Y_i are not i.i.d. .

1.3.2 Coupon collector's problem

Consider the coupon collector's problem with N different items. Let X_n denote the number of different items in the collection after the nth step. Again :

Proposition 2. *The stochastic process $W = (W_n)_{0 \leq n \leq m}$, conditioned on $W_0 = (m, \ell)$, or equivalently on the number of different items in the collection after the mth step, X_m, to be equal to ℓ, is the Markov chain with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Stirling numbers of the second kind. Its distribution does not depend on N.*

1.3.3 Chinese restaurant process

In the Chinese restaurant process with $(0, \theta)$ seating plan, defined at Section 2 (see also, e.g., [Pit06, Ch. 3]), let X_n denote the number of occupied tables after the arrival of the nth customer.

Proposition 3. *The stochastic process $W = (W_n)_{0 \leq n \leq m}$, conditioned to $W_0 = (m, \ell)$, or equivalently to $X_m = \ell$, is the Markov chain with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Stirling numbers of the first kind. Its distribution does not depend on θ.*

Remark 1. As a consequence, in the three previous cases, given the data $(X_n)_{0 \leq n \leq m}$, X_m (or W_0) are sufficient statistics for the parameters p, N or θ, respectively.

1.3.4 One-dimensional Internal Diffusion Limited Aggregation

Finally, in the one-dimensional Internal Diffusion Limited Aggregation process (iDLA), let X_n denote the number of particles settled to the right of the origin after the release of the nth particle. Then

Proposition 4. *The stochastic process $W = (W_n)_{0 \leq n \leq m}$, conditioned to $W_0 = (m, \ell)$, is the Markov chain with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Euler's triangle.*

More precise definitions, and proofs, are to be found at Section 2.

1.4 Simulations

In the case of Pascal's triangle, the behaviour of this time-reversed Markov chain is well understood since forever. Quite recently, [AC19] gave a rather precise analysis of the analog time-reversed Markov chain related to Stirling numbers of the second kind, with combinatorial analysis of finite automata as a motivation. We hope to improve some of their results and proofs.

In this section, in order to surmise the behaviour of these time-reversed Markov chains, we present the result of some simulations. For each case, the figures below show sample paths starting at (m, mt) with $t \in \{0.05, ..., 0.95\}$ and $m = 500$:

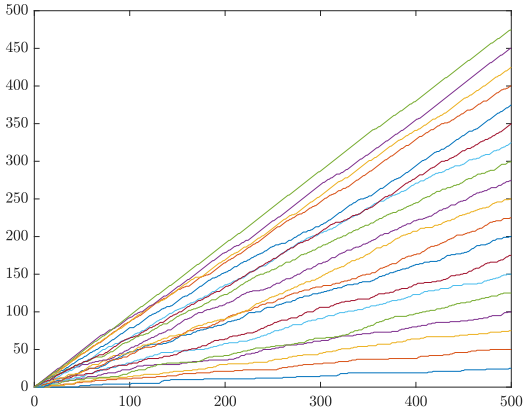

Figure 1: Pascal's triangle.

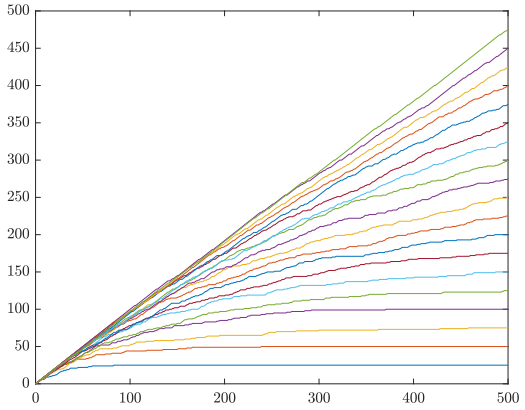

Figure 2: Stirling numbers of the second kind.

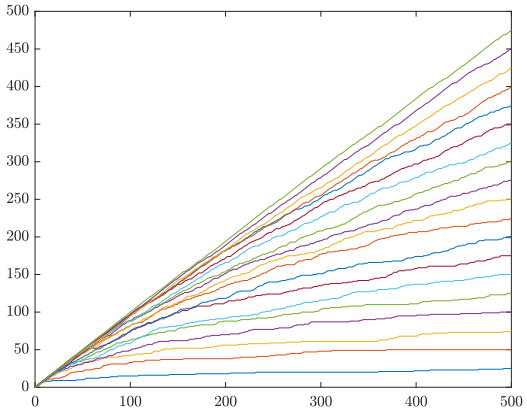

Figure 3: Stirling numbers of the first kind.

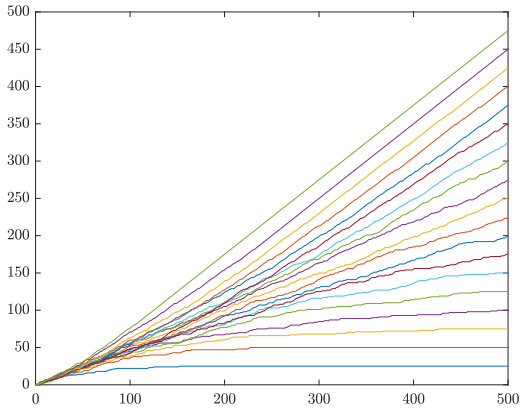

Figure 4: Eulerian numbers.

Now, in order to compare the four combinatorial triangles, we show the average of 100 sample paths for each triangle, for $m = 1000$ and $t \in \{0.05, ..., 0.95\}$:

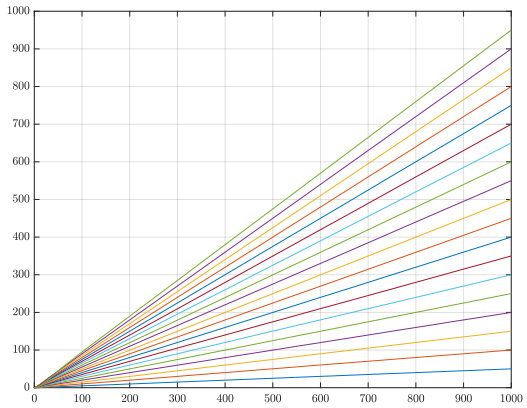

Figure 5: Pascal's triangle.

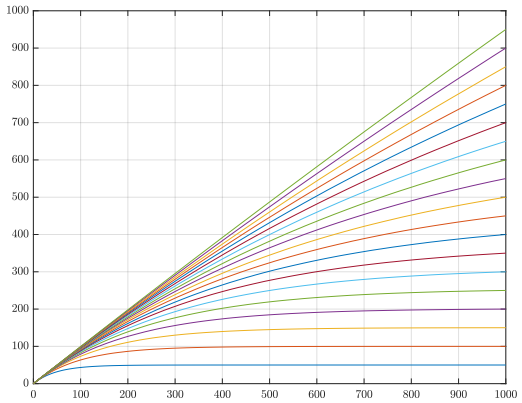

Figure 6: Stirling numbers of the second kind.

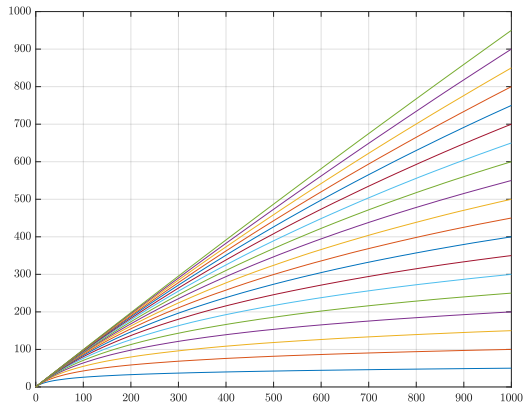

Figure 7: Stirling numbers of the first kind.

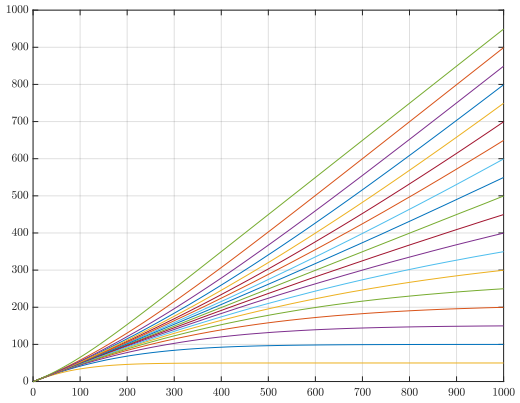

Figure 8: Eulerian numbers.

In the first two cases, the smooth nature of these averaged paths is not unexpected due to old, and more recent, fluid approximation results, see [AC19] or the next sections. This paper aims at a global explanation of the asymptotic behaviour of the four Markov chains exhibited by these simulations.

1.5 Asymptotics of sample paths, and field lines of vector fields

Combinatorial analysis, see [Goo61] or [Ben73], yields that,

Theorem 1. *In each of the four cases, there exists a function* $\varphi : (0, +\infty) \to [0, 1]$ *such that, for any positive number* λ_∞*, when* $(m, \ell) \to +\infty$ *and* $\lim m/\ell = 1 + \lambda_\infty$*,*

$$\lim p_1(m, \ell) = \varphi(\lambda_\infty).$$

At the end of this section, the function φ is described for each of the four cases.

We set, for $(m,\ell) \in \mathring{S}$,

$$\lambda = \lambda(m,\ell) = \frac{m-\ell}{\ell}.$$

As a direct consequence of Theorem 1, one expects a fluid approximation of the previous Markov chains by a special family of curves : let $\mathbb{P}_{(m,\ell)}$ denote the probability distribution of the Markov chain W starting from (m,ℓ), and let $(x,\gamma_\lambda(x))_{0\le x\le n}$ be the field line going through the point $(1,\ell/m) = (1,1/(1+\lambda))$ for the vector field $(1,\varphi(-1+x/y))$, or, equivalently, let γ_λ denote the solution of the ODE

$$y' = \varphi\left(\frac{x-y}{y}\right) \tag{3}$$

that satisfies $y(1) = 1/(1+\lambda)$. So far we have a complete proof of this approximation only in the first three cases :

Theorem 2. *In the first three cases, for any $\eta \in (0,1/2)$ and any $\lambda_\infty > 0$, when $(m,\ell) \to +\infty$ and $\lim \lambda(m,\ell) = \lambda_\infty$,*

$$\lim \mathbb{P}_{(m,\ell)}\left(\sup_{0\le t\le 1}\left(|w_m(t) - \gamma_{\lambda_\infty}(t)|\right) \ge m^{-\eta}\right) = 0.$$

Note that the special form of the ODE (3) entails that the set of field lines is invariant by positive homotheties.

This kind of statement seems to hold true for eulerian numbers, according to our simulations (see Section 1.4), but remains an open question. For Stirling numbers of the first kind, Theorem 2 seems to be new, as far as we know. For Stirling numbers of the second kind, Theorem 2 is a vastly improved version of a result that appeared in [AC19], in which the proof relies mainly on uniform bounds for

$$m\ \left|p_1(m,\ell) - \varphi\left(\lambda(m,\ell)\right)\right|,$$

on domains that approach $\mathring{S}$ as well as possible. These bounds follow from a careful asymptotic analysis of $T(m,\ell)$, that should have some interest in itself. However the proof given here is much simpler.

Our choice of combinatorial triangles may seem arbitrary, and we confess it is : for instance, Bell's triangle or Delannoy's triangle have also Pascal's formulas, but of a slightly different form. It remains to see if the approach of this paper still produces results for Bell's triangle or Delannoy's triangle, in spite of these slight differences.

1.5.1 Description of φ

- *Pascal's triangle.* It is well known that for all $(m,\ell) \in S^\star$,

  $$p_1(m,\ell) = \frac{\ell}{m},$$

  so that

  $$\varphi_1(\lambda) = \frac{1}{1+\lambda}.$$

  Relation (3) reduces to $y' = y/x$, with the linear functions as solutions, as expected.

- *Stirling numbers of the second kind.* For $\lambda > 0$, let φ_2 be defined, through $\zeta_2(\lambda)$, the unique positive solution of

$$\frac{\zeta_2}{1-e^{-\zeta_2}} = 1+\lambda, \quad \text{by } \varphi_2(\lambda) = e^{-\zeta_2}. \tag{4}$$

  Then, for $x \geq 0$,

$$\gamma_\lambda(x) = \frac{1-e^{-x\,\zeta_2(\lambda)}}{\zeta_2(\lambda)}. \tag{5}$$

- *Stirling numbers of the first kind.* For $\lambda > 0$, let φ_3 be defined, through $\zeta_3(\lambda)$, the unique solution, in $(0,1)$, of

$$\frac{\zeta_3}{(\zeta_3-1)\ln(1-\zeta_3)} = 1+\lambda, \quad \text{by } \varphi_3(\lambda) = 1-\zeta_3.$$

  Then, for $x \geq 0$,

$$\gamma_\lambda(x) = \frac{1-\zeta_3(\lambda)}{\zeta_3(\lambda)} \ln\left(\frac{1-\zeta_3(\lambda)+x\,\zeta_3(\lambda)}{1-\zeta_3(\lambda)}\right). \tag{6}$$

- *Eulerian numbers.* For $\lambda > 0$, let φ_4 be defined, through $\zeta_4(\lambda)$, the unique solution, in $\mathbb{R}$, of

$$\frac{1}{1+\lambda} = \frac{e^{\zeta_4}}{e^{\zeta_4}-1} - \frac{1}{\zeta_4}, \quad \text{by } \varphi_4(\lambda) = 1 - \frac{\zeta_4}{(1+\lambda)(e^{\zeta_4}-1)}.$$

  At the moment, we are unaware of any closed form formula for γ_λ in this case.

2 Time reversal and Markov property for W

In this section, we prove Propositions 1, 2, 3 and 4. The notations X_n, Y_n, W_n are defined at Section 1.3.

2.1 Time reversal

As already known at least since Kolmogorov, see [Kol35, (7)], a time-reversed Markov process is still a Markov process, but it is an inhomogeneous one. Let us recall the basic facts that we need here : if h_k denotes the probability distribution of X_k and if $X = (X_k)_{k\geq 0}$ is an inhomogeneous Markov chain with kernels $(Q_k)_{k\geq 0}$, i.e.

$$Q_{k,i,j} = \mathbb{P}\left(X_{k+1} = j \mid X_k = i\right),$$

then

Proposition 5. *$W = (W_n)_{0\leq n\leq m}$ is a Markov chain with state space S and with kernel P defined on $S^\star$ by*

$$P_{(n,i),(n-1,j)} = \frac{h_{n-1}(j)Q_{n-1,j,i}}{h_n(i)}.$$

Proof. First, since $(0,0)$ is only reached, eventually, at time m, there is no need to define $P_{(0,0),(.,.)}$. Also, P is a probability kernel due the Chapman-Kolmogorov equations for (h_n) and (Q_n). Then,

$$\mathbb{P}\left((X_k)_{0\le k\le m} = (x_k)_{0\le k\le m}\right) = h_0(x_0)\prod_{k=0}^{m-1} Q_{k,x_k,x_{k+1}},$$

thus, provided that $x_m = \ell$,

$$\begin{aligned}\mathbb{P}\left((X_k)_{0\le k\le m} = (x_k)_{0\le k\le m} \mid X_m = \ell\right) &= h_0(x_0)\prod_{k=0}^{m-1} Q_{k,x_k,x_{k+1}} \;/\; h_m(\ell),\\ &= \prod_{k=0}^{m-1} P_{(k+1,x_{k+1}),(k,x_k)}.\end{aligned}$$

That is,

$$\mathbb{P}\left((W_k)_{0\le k\le m} = (m-k, x_{m-k})_{0\le k\le m} \mid W_0 = (m,\ell)\right) = \prod_{k=0}^{m-1} P_{(k+1,x_{k+1}),(k,x_k)},$$

as expected. □

But for eulerian numbers, $h_n(k) = T(n,k)\theta^k/T_n(\theta)$, or $h_n(k) = T(n,k)\theta^{k\downarrow}/T_n(\theta)$, in which $T_n(\theta)$ is a normalizing constant :

$$T_n(\theta) = \sum_{k=0}^{n} T(n,k)\theta^k, \quad \text{or} \quad T_n(\theta) = \sum_{k=0}^{n} T(n,k)\theta^{k\downarrow}. \tag{7}$$

For eulerian numbers, $h_n(k) = T(n,k)/T_n(1) = T(n,k)/n!$.

Note that Q_n results from a natural growing mechanism with independent steps, that is, a Markovian growth process, obtained as follows :

- by addition of an $n+1$th letter, either $\mathtt{a}$ or $\mathtt{b}$, at the end of a random word of $\{\mathtt{a},\mathtt{b}\}^n$, in order to form an $n+1$-letters long word, for Pascal's triangle,
- by addition of the image of $n+1$ to a random mapping from $[\![n]\!]$ to $[\![N]\!]$, in order to form a random mapping from $[\![n+1]\!]$ to $[\![N]\!]$, for the the second Stirling triangle,
- by random insertion of $n+1$ in order to form a permutation on $[\![n+1]\!]$, starting from a permutation on $[\![n]\!]$, for the 2 other examples.

In each case, the added letter, or integer, is chosen independently of the previous history of the growth process, hence the Markovian character of these growth processes. For the sake of brevity, in this paper, we call the last growth process the *random permutation process*. For these three nonhomogeneous Markov growth processes, there exist well studied functionals that retain the Markov property, and whose one-dimensional distributions are given by the rows of the corresponding combinatorial triangle :

- the sequence of counts of letter a, in the sequence of words defined previously, forms one of the most studied Markov chain : the simple random walk, whose one-dimensional distributions h_n are binomial distributions, famously related to Pascal's triangle ;
- the sequence of sizes of images, derived from the sequence of random mappings, is a famous inhomogeneous Markov chain, related to the coupon collector problem : it is the sequence of successive sizes of the collection. Its one-dimensional distributions h_n have a simple expression in terms of the Stirling numbers of the second kind ;
- the sequence of number of cycles, derived from the random permutation process, is an inhomogeneous Markov chain, related to the chinese restaurant process. Its one-dimensional distributions h_n have a simple expression in terms of the Stirling numbers of the first kind ;
- the sequence of the number of descents, also derived from the random permutation process, is an inhomogeneous Markov chain, related to the internal diffusion limited aggregation process. Its one-dimensional distributions h_n have a simple expression in terms of eulerian numbers.

Chapman-Kolmogorov equations for these Markov chains are derived from Pascal's formulas for corresponding triangles through renormalization : in our settings, Q_n is defined by (a, b) as follows

$$Q_{n,x,y} = c_n(\theta)\ \big(b(n+1,y)\mathbb{1}_{y=x\in[\![n]\!]} + a(n+1,y)\ \theta\ \mathbb{1}_{y=x+1\in[\![n+1]\!]}\big)\,, \tag{8}$$

in which $c_n(\theta)$ denotes a normalizing factor $T_n(\theta)/T_{n+1}(\theta)$, and $\theta = \theta^{x+1}/\theta^x$ should be replaced, in the last factor of (8), with $\theta - x = \theta^{x+1\downarrow}/\theta^{x\downarrow}$ in the case of Stirling numbers of the second kind. For eulerian numbers, $\theta = 1$. Then, Pascal's formulas appear as special cases of the Chapman-Kolmogorov equation $h_{n-1}Q_{n-1} = h_n$, and relation (2) is just a special case of Proposition 5.

Here, $Q_{n,x_n,x_{n+1}} \neq 0$ only if $\varepsilon_{n+1} = x_{n+1} - x_n$ belongs to $\{0, 1\}$, thus $P_{(n,x),(n-1,y)} \neq 0$ only if $\varepsilon = x - y$ belongs to $\{0, 1\}$: in this paper, $P_{(n,x),(n-1,x-\varepsilon)}$ is abridged to $p_\varepsilon(n, x)$.

2.2 Simple random walk

Proof of Proposition 1. Here

$$T(n,k) = \binom{n}{k}, \quad \theta = \frac{p}{1-p}, \quad T_n(\theta) = (1+\theta)^n$$
$$h_n(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad c_n(\theta) = \frac{1}{1+\theta} = 1-p.$$

Then, for instance,

$$\begin{aligned} P_{(n,i),(n-1,i-1)} &= \frac{h_{n-1}(i-1)Q_{n-1,i-1,i}}{h_n(i)} \\ &= \frac{\binom{n-1}{i-1}p^{i-1}(1-p)^{n-i} \times p}{\binom{n}{i}p^i(1-p)^{n-i}} = \frac{\binom{n-1}{i-1}}{\binom{n}{i}} = p_1(n,i) \end{aligned}$$

as expected. $\square$

2.3 Coupon collector's problem

Let us recall the famous problem studied by Gauss and Laplace, among others : a collector wants to complete a collection of N different items (denoted $1, ..., N$). At each step, he receives a coupon chosen uniformly from $[\![1, N]\!]$. The average time to complete the collection is known to be NH_N, where

$$H_N = \sum_{k=1}^{N} \frac{1}{k}$$

is the Nth harmonic number. If X_n denotes the number of different items in the collection after the nth step, then we call the graph of $t \mapsto X_{\lfloor t \rfloor}$ the *completion curve*.

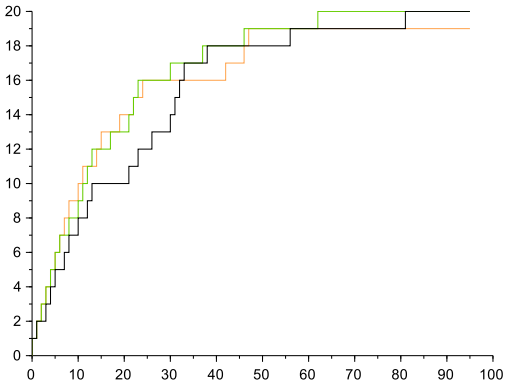

Figure 9: Three completion curves for a $n = 20$ items collection.

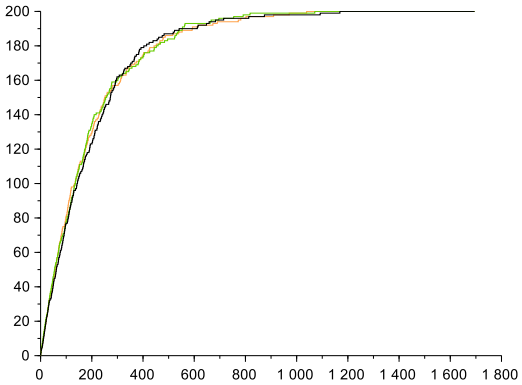

Figure 10: Three completion curves for a $n = 200$ items collection.

Proof of Proposition 2. See [AC19, Proposition 1], in which the proof is given for $m = N$. It fits with the frame given at Section 2.1 as follows : set

$$T(n, k) = \left\{ {n \atop k} \right\}, \quad \theta = N,$$

but consider a variant of T_n. Here :

$$T_n(\theta) = \sum_k T(n, k)\theta^{k\downarrow} = N^n$$

$$h_n(k) = \binom{N}{k} \left\{ {n \atop k} \right\} k! \, \frac{1}{N^n} = \left\{ {n \atop k} \right\} \frac{N^{k\downarrow}}{N^n}, \quad c_n(\theta) = \frac{1}{N}.$$

Then, for instance,

$$Q_{n,k,k+1} = a(n+1, k) \frac{\theta^{k+1\downarrow}}{\theta^{k\downarrow}} c_n(\theta) = \frac{N-k}{N}.$$

and

$$p_1(n,k) = P_{(n,k),(n-1,k-1)} = \frac{h_{n-1}(k-1)Q_{n-1,k-1,k}}{h_n(k)}$$
$$= \frac{\left\{ {n-1 \atop k-1} \right\} \frac{N^{k-1\downarrow}}{N^{n-1}} \times \frac{N-k+1}{N}}{\left\{ {n \atop k} \right\} \frac{N^{k\downarrow}}{N^n}} = \frac{\left\{ {n-1 \atop k-1} \right\}}{\left\{ {n \atop k} \right\}},$$

as expected. □

2.4 Chinese restaurant process

Set $\theta \in (0, +\infty)$. The chinese restaurant process, introduced in 1974 by Antoniak in [Ant74], is defined as follows : when entering a metaphoric chinese restaurant, the first customer seats at the first table. For $n > 1$, the nth customer seats at the kth (non-empty) table with probability $\frac{c_{n,k}}{n-1+\theta}$ (where $c_{n,k}$ is the number of customers seated at this table), or at an empty table with probability $\frac{\theta}{n-1+\theta}$. Let X_n denote the number of non-empty tables after the arrival of the nth customer. For exemple, let us sample the first 50 steps of the process, for $\theta = 1$:

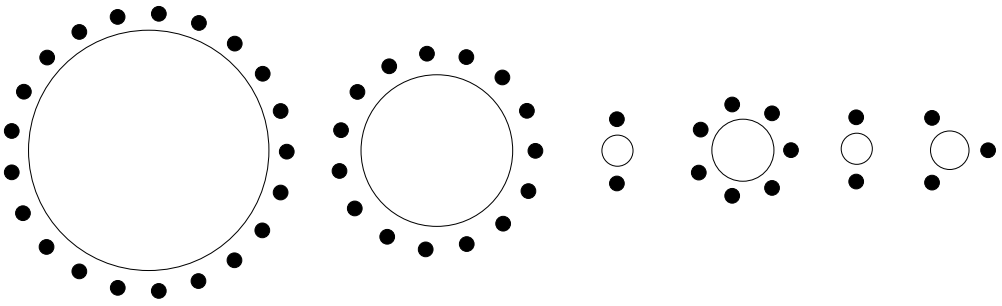

Figure 11: A realization of the chinese restaurant process (here $X_{50} = 6$).

Proof of Proposition 3. In this example, $(Y_i)_{i\geq 1}$ is a family of independent Bernoulli random variables with respective parameters $p_i = \theta/(i-1+\theta)$. We have :

$$T(n,k) = \begin{bmatrix} n \\ k \end{bmatrix}, \quad T_n(\theta) = \sum_k T(n,k)\theta^k = (\theta)^{\uparrow n}, \quad c_n(\theta) = \frac{1}{\theta + n}.$$

Thus the probability distribution of X_n is given, for $n \geq 1$, by:

$$h_n(\ell) = \mathbb{P}(X_n = \ell) = \frac{\theta^\ell}{(\theta)^{\uparrow n}} \begin{bmatrix} n \\ \ell \end{bmatrix} \mathbb{1}_{1\leq \ell \leq n},$$

see [Pit06, Section 3.1.3]. For instance,

$$Q_{n,k,k+1} = \frac{\theta}{n+\theta} = c_n(\theta)\ a(n+1,k)\ \theta.$$

and

$$p_1(n,k) = P_{(n,k),(n-1,k-1)} = \frac{h_{n-1}(k-1)Q_{n-1,k-1,k}}{h_n(k)}$$
$$= \frac{\frac{\theta^{k-1}}{(\theta)^{n-1\uparrow}}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix} \times \frac{\theta}{n-1+\theta}}{\begin{bmatrix} n \\ k \end{bmatrix}\frac{\theta^k}{(\theta)^{n\uparrow}}} = \frac{\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}}{\begin{bmatrix} n \\ k \end{bmatrix}},$$

as expected. □

2.5 One-dimensional Internal Diffusion Limited Aggregation process

Diaconis and Fulton [DF91] introduced the internal Diffusion Limited Aggregation process (iDLA). Lawler, Bramson and Griffeath [LBG92] coined the terminology *iDLA*, and obtained an asymptotic shape behaviour. In the iDLA process, an aggregate of particles on $\mathbb{Z}^d$ is built as follows:

i) the first particle settles at the origin;

ii) the next particles perform a symmetric random walk on $\mathbb{Z}^d$, starting from the origin, and settle at the first empty site they encounter.

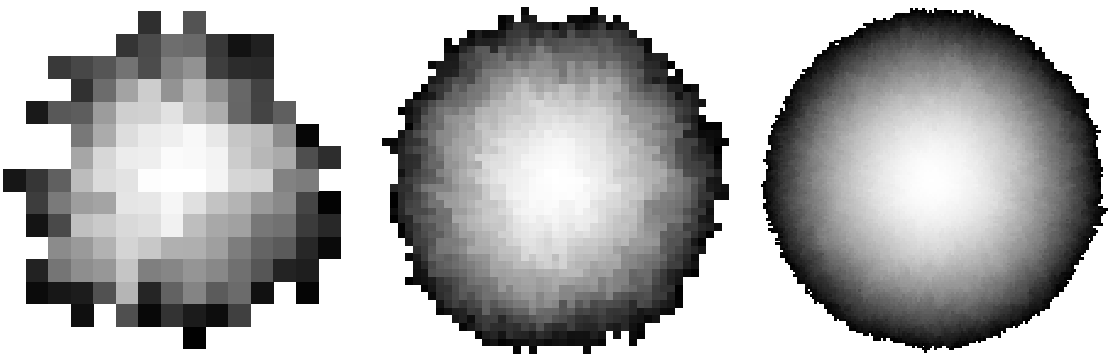

Figure 12: Normalized iDLA aggregates with 150, 1,500 and 15,000 particles on $\mathbb{Z}^2$.

When $d = 1$, let X_n denote the number of particles settled to the right of the origin after the nth step. Then, according to [Mit20], the process $(X_n)_n$ is an inhomogeneous Markov chain with the same distribution as the sequence of number of descents of the sequence of random permutations defined previously. Both processes have the one-dimensional distribution below

$$\mathbb{P}(X_n = k) = h_n(k) = \frac{\left\langle {n \atop k} \right\rangle}{n!}\mathbb{1}_{(n,k)\in S}.$$

In the case of the one-dimensional iDLA we can stack successive aggregates upon one another to form a space-time diagram. As with Figure 12, the longer it took to visit a cell, the darker we color it.

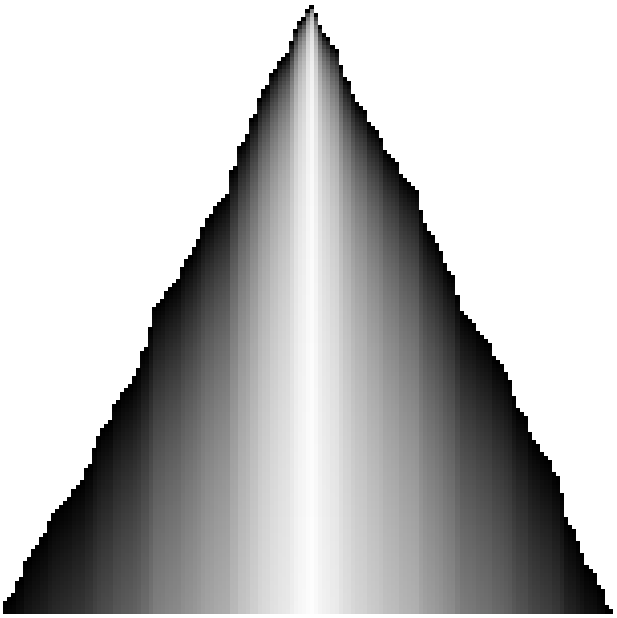

Figure 13: Space-time diagram of
a one-dimensional iDLA.

Proof of Proposition 4. In this example, we have :

$$T(n,k) = \left\langle {n \atop k} \right\rangle, \quad T_n(1) = \sum_k T(n,k) = n!, \quad c_n(1) = \frac{1}{n}.$$

Thus the probability distribution of X_n is given, for $n \geq 1$, by:

$$h_n(\ell) = \mathbb{P}(X_n = \ell) = \frac{1}{n!} \left\langle {n \atop \ell} \right\rangle \mathbb{1}_{1 \leq \ell \leq n},$$

see [Pit06, Section 3.1.3]. For instance,

$$Q_{n,k,k+1} = \frac{n-k}{n+1} = c_n(1)\ a(n+1,k).$$

and

$$\begin{aligned} p_1(n,k) = P_{(n,k),(n-1,k-1)} &= \frac{h_{n-1}(k-1)Q_{n-1,k-1,k}}{h_n(k)} \\ &= \frac{\frac{1}{n-1!}\left\langle {n-1 \atop k-1} \right\rangle \times \frac{n-k}{n}}{\left\langle {n \atop k} \right\rangle \frac{1}{n!}} = \frac{\left\langle {n-1 \atop k-1} \right\rangle (n-k)}{\left\langle {n \atop k} \right\rangle}, \end{aligned}$$

as expected. $\square$

3 The limit vector field : proof of Theorem 1

One can see the set v of average jumps $v(n,k)$, defined, for $(n,k) \in S$, by

$$\begin{aligned} v(n,k) &= p_0(n,k) \times (-1,0) + p_1(n,k) \times (-1,-1) \\ &= (-1, -p_1(n,k)), \end{aligned}$$

as a kind of discrete vector field v on S, with slope $p_1(n,k)$ at point (n,k). As a consequence, the convergence of the sample paths of the time-reversed Markov chains of Section 1.3 (see Theorem 2) requires a precise asymptotic analysis of

$$p_1(n,k) = \frac{a(n,k)T(n-1,k-1)}{T(n,k)},$$

and thus, of $T(n,k)$. Consider the generating functions V_k and H_n defined by

$$V_k(z) = \sum_{n=k}^{+\infty} \frac{1}{f_n} T(n,k) z^n, \quad H_n(w) = \sum_{k=0}^{n} T(n,k) w^k,$$

respectively. Here f_n is either 1 (for Pascal's, resp. Euler's, triangle) or $n!$, for the 2 Stirling's triangles : for the enumeration of sets of labelled structures, such as e.g. subsets or cycles, as for the 2 Stirling's triangles, cf. [FS09, Part A], the factor $n!$ is due to the use of EGFs, and since we consider *sets*, not sequences, of k objects, i.e. *unordered collections*, the generating function V_k contains a factor $1/k!$. In the first three cases, V_k exhibits a factorisation $A \times B^k$ suitable for the saddle-point method, while, for eulerian numbers, H_n is approximately of the form B^n, allowing the use of large deviations methods.

Due to these factorisations, the limit vector field depends only on the slope y/x, and the function φ depends on B alone, in the first 3 cases through the saddle-point equation

$$\frac{B'(\zeta)}{B(\zeta)} = \frac{1+\lambda}{\zeta}, \tag{9}$$

obtained by optimisation of the function $x \to \dfrac{B(x)}{x^{1+\lambda}}$ on $(0,+\infty)$, and, for eulerian numbers, through the Legendre transformation of $\ln B$, leading to the equation :

$$\frac{B'(\zeta)}{B(\zeta)} = \frac{1}{1+\lambda}. \tag{10}$$

Proof of Theorem 1. But for eulerian numbers, let $\zeta(\lambda)$ be defined implicitly by (9), i.e. let $\zeta(.)$ be the inverse function of :

$$x \longrightarrow \frac{xB'(x)}{B(x)} - 1.$$

The eulerian case is similar, but uses large deviations rather than saddle-point methods, and will be handled separately. In the remaining 3 cases, recall that $a(n,k) = 1$, and set :

$$1+\lambda = \frac{n}{k}, \quad \zeta = \zeta(\lambda), \quad 1+\tilde{\lambda} = \frac{n-1}{k-1}, \quad \tilde{\zeta} = \zeta(\tilde{\lambda}).$$

For these 3 cases, the saddle-point method, see [FS09, Part B, Chap. VIII], leads to

$$T(n,k) \sim \frac{f_n}{f_k} \left(\frac{B(\zeta)}{\zeta^{1+\lambda}}\right)^k g(n,k), \tag{11}$$

in which $g(.,.)$ is some factor such that $g(n,k) \sim g(n-1,k-1)$. The invariance by homothetie of the field lines results from the factorisation $V_k = A \times B^k$ and from the Cauchy formula, that leads to the key role of λ in the asymptotic behaviour (11), and is thus a consequence of the decomposability of the underlying combinatorial structures.

The factor f_n/f_k matters only for the 2 Stirling's triangles. As a consequence, for the two Stirling triangles, we have

$$\begin{aligned} p_1(n,k) &\sim \frac{a(n,k)\ k}{n} \frac{\tilde{\zeta}^{1+\tilde{\lambda}}}{B\left(\tilde{\zeta}\right)} \left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\tilde{\lambda}}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right)^k \\ &\sim \frac{1}{1+\lambda} \frac{\zeta^{1+\lambda}}{B(\zeta)} \left(\frac{\tilde{\zeta}^{1+\lambda}}{\tilde{\zeta}^{1+\tilde{\lambda}}}\right)^k \left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right)^k \\ &\sim \frac{1}{1+\lambda} \frac{\zeta^{1+\lambda}}{B(\zeta)} \zeta^{(\lambda-\tilde{\lambda})k}, \end{aligned}$$

the last step due to

$$\lim_k\ k\ \ln\left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right) = 0. \tag{12}$$

Actually, since ζ is solution of the saddle-point equation, the derivative of

$$x \to \ln\left(\frac{B(x)}{x^{1+\lambda}}\right)$$

vanishes at ζ, thus

$$\ln\left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right) = o\left(\tilde{\zeta} - \zeta\right) = o\left(\lambda - \tilde{\lambda}\right),$$

but

$$\lambda - \tilde{\lambda} = \frac{n}{k} - \frac{n-1}{k-1} \sim \frac{-\lambda}{k},$$

entailing (12). Thus

$$p_1(n,k) \sim \frac{1}{1+\lambda} \frac{\zeta}{B(\zeta)}.$$

Finally, for Stirling's triangles, the saddle-point equation (9) gives

$$p_1(n,k) \sim \frac{1}{B'(\zeta)},$$

For Pascal's triangle, $\binom{n}{k}$ enumerates words with n letters, k among them being a's and the $n-k$ others being b's, thus Pascal's triangle enumerates *sequences* (not sets) of *unlabelled* objects[1], for which one usually uses OGFs. As a consequence, $f_n = 1$, and, compared with the previous computation, we are rid of the factor $n!/k!$ in $T(n,k)$, and of the factor $k/n = 1/(1+\lambda)$ in $p_1(n,k)$, thus we obtain

$$p_1(n,k) \sim \frac{\zeta}{B(\zeta)}.$$

Before we turn to the case of eulerian numbers, let us derive φ for each of the 3 first cases :

- *Pascal's triangle :*

$$\begin{aligned} V_{k,1}(z) &= \sum_{n\geq k} \binom{n}{k} z^n = \frac{1}{1-z}\left(\frac{z}{1-z}\right)^k, \\ B_1(z) &= \frac{z}{1-z}, \\ p_1(n,k) &\sim \frac{\zeta}{B_1(\zeta)} = 1-\zeta. \end{aligned}$$

  Here (9) can be written

$$\frac{1}{1-\zeta} = 1+\lambda,$$

  thus $\varphi_1(k/n) = \frac{1}{1+\lambda} = k/n$, that is :

$$p_1(n,k) \sim \frac{k}{n},$$

  which is not a surprise, since it is well known that, actually, $p_1(n,k) = \frac{k}{n}$.

- *Stirling numbers of the second kind*

  Here :

$$\begin{aligned} V_{k,2}(z) &= \sum_{n\geq k} \left\{ {n \atop k} \right\} \frac{z^n}{n!} = \frac{1}{k!}\left(e^z-1\right)^k, \\ B_2(z) &= e^z - 1, \\ p_1(n,k) &\sim \frac{1}{B_2'(\zeta)} = e^{-\zeta}, \end{aligned}$$

[1] A word with n letters, k among them being a's and the $n-k$ others being b's, can be seen as a sequence of k words of the form $\mathtt{b}^m\mathtt{a}$ followed by a word of the form $\mathtt{b}^m$.

and (9) can be written

$$\frac{\zeta}{1-e^{-\zeta}} = 1+\lambda, \tag{13}$$

see [AC19]. Thus

$$\varphi_2(k/n) = e^{-\zeta\left(\frac{n}{k}-1\right)}.$$

Note that, according to Good [Goo61] and others, ζ is a smooth concave function of $\lambda > 0$, with positive values. Note also that (13) is the equation to be solved when one wants to tune the parameter ζ of a Poisson random variable *conditioned to be positive* in order to obtain the expectation $1+\lambda$.

- *Stirling numbers of the first kind (unsigned)*

$$\begin{aligned} V_{k,3}(z) &= \sum_{n\geq k} \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!} = \frac{1}{k!}\ \left(-\ln(1-z)\right)^k, \\ B_3(z) &= -\ln(1-z), \\ p_1(n,k) &\sim \frac{1}{B_3'(\zeta)} = 1-\zeta. \end{aligned}$$

Here (9) can be written

$$\frac{\zeta}{(\zeta-1)\ln(1-\zeta)} = 1+\lambda, \tag{14}$$

which defines ζ as smooth concave function of $\lambda > 0$, with values in $(0,1)$. Thus

$$\varphi_3(k/n) = 1-\zeta\left(\tfrac{n}{k}-1\right). \tag{15}$$

Note that (14) is the equation to be solved when one wants to tune the parameter ζ of a logarithmic probability distribution in order to obtain the expectation $1+\lambda$.

For *eulerian numbers*, though the computation of φ_4 has a similar flavour, it presents some notable differences. In order to sum up the asymptotic analysis of eulerian numbers, set, as done in [Ben73] :

$$t = \frac{k}{n} = \frac{1}{1+\lambda}.$$

In [Ben73, page 97], the main tool is the approximation of $H_n(e^s)$, the Laplace transform of h_n, by

$$B(s)^{n+1} = r(s)^{-n-1} = \left(\frac{e^s-1}{s}\right)^{n+1}.$$

In other terms, the key point in [Ben73] is that h_n is approximately the distribution of the sum of $n+1$ i.i.d. uniform random variables, with Laplace transform $B(s)$. This is reminiscent of Tanny's representation of eulerian numbers (cf. [Tan73]) :

$$h_n(k) = \frac{\left\langle {n \atop k} \right\rangle}{n!} = \mathbb{P}\left(\lfloor U_1 + U_2 + \cdots + U_n \rfloor = k\right). \tag{16}$$

Bender obtains the following asymptotic formula for $\left\langle {n \atop k} \right\rangle$ when (n,k) goes to infinity

$$\frac{\left\langle {n \atop k} \right\rangle}{n!} \sim \left(B(\zeta)e^{-\zeta t}\right)^n \; g(n,k)$$

in which $g(.,.)$ is some factor such that $g(n,k) \sim g(n-1,k-1)$, and in which ζ is the only real number such that

$$\begin{aligned} \frac{1}{1+\lambda} = t &= \frac{\partial}{\partial \zeta} \ln\left(\frac{e^\zeta - 1}{\zeta}\right) \\ &= \frac{e^\zeta}{e^\zeta - 1} - \frac{1}{\zeta}. \end{aligned} \tag{17}$$

One recognize in $\zeta(.)$ the derivative of the Legendre-Fenchel transformation of the cumulant-generating function of the uniform distribution, i.e. the unique solution of

$$\frac{\partial}{\partial \zeta} \ln\left(B(\zeta)e^{-\zeta t}\right) = \frac{B'(\zeta)}{B(\zeta)} - t = 0. \tag{18}$$

As a consequence, for eulerian numbers, we have

$$\begin{aligned} p_1(n,k) &\sim \frac{a(n,k)}{n} \left(\frac{B(\tilde{\zeta})^{n-1} e^{-\tilde{\zeta}(k-1)}}{B(\zeta)^n e^{-\zeta k}}\right) \\ &\sim \frac{a(n,k)}{n} \frac{e^{\tilde{\zeta}}}{B(\tilde{\zeta})} \left(\frac{B(\tilde{\zeta})^n e^{-\tilde{\zeta} k}}{B(\zeta)^n e^{-\zeta k}}\right) \\ &\sim \frac{(1-t)e^\zeta}{B(\zeta)} \left(\frac{B(\tilde{\zeta})e^{-\tilde{\zeta} t}}{B(\zeta)e^{-\zeta t}}\right)^n \\ &\sim \frac{(1-t)e^\zeta}{B(\zeta)} \end{aligned}$$

the last step due to

$$\lim_n \; n \; \ln\left(\frac{B(\tilde{\zeta})e^{-\tilde{\zeta} t}}{B(\zeta)e^{-\zeta t}}\right) = 0. \tag{19}$$

Actually, since ζ is solution of (18), the derivative of

$$x \to \ln\left(B(x)e^{-xt}\right)$$

vanishes at ζ, thus

$$\ln\left(\frac{B(\tilde{\zeta})e^{-\tilde{\zeta}t}}{B(\zeta)e^{-\zeta t}}\right) = o\left(\tilde{\zeta}-\zeta\right) = o\left(t-\tilde{t}\right),$$

but

$$t-\tilde{t} = \frac{k}{n} - \frac{k-1}{n-1} \sim \frac{1-t}{n},$$

entailing (19). Thus

$$p_1(n,k) \sim \frac{(1-t)\zeta e^{\zeta}}{e^{\zeta}-1} = \frac{\lambda\zeta}{(1+\lambda)(1-e^{-\zeta})} = \phi_4(\lambda).$$

Note that :

$$\zeta(1-t) = -\zeta(t), \quad \varphi_4(1-t) = 1-\varphi_4(t) = \frac{t\zeta}{e^{\zeta}-1},$$

as expected from the relation $\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n-k-1} \right\rangle$. □

4 Sample path convergence

This section is devoted to the proof of Theorem 2 for the first three triangles. For the sake of completeness, we first give the well known proof of Theorem 2 for Pascal's triangle. In the case of Stirling numbers of the second kind, a weaker form of Theorem 2 was obtained in [AC19] at the price of a tedious proof using Wormald method and saddle-point asymptotics. For Euler's triangle, we think that the same property holds true, but the proof is still a work in progress. In the case of Stirling triangles of both kind, we believe that the proofs given in the next sections are new.

4.1 Proof of Theorem 2 : Pascal's triangle.

Consider two probability distributions for the processes (W,X,Y) defined at section 1.3. Under $\mathbb{P}_{(m,\ell)}$, W is a Markov chain starting from (m,ℓ), with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Pascal's triangle, and the processes (X,Y) are distributed accordingly. On the other hand, under $\mathbb{P}_p$, $Y=(Y_k)_{1\leqslant k\leqslant m}$ is a sequence of i.i.d Bernoulli random variables with parameter p, and the processes (X,W) are distributed accordingly. According to Proposition 1, for any $p\in(0,1)$, and any set B in the relevant state space,

$$\begin{aligned}\mathbb{P}_{(m,\ell)}\left((W,X,Y)\in B\right) &= \mathbb{P}_p\left(\{(W,X,Y)\in B\}\cap\{X_m=\ell\}\right)/\mathbb{P}_p\left(X_m=\ell\right) \\ &= \mathbb{P}_p\left(\{(W,X,Y)\in B\}\cap\{X_m=\ell\}\right)/\mathbb{P}_p\left(W_0=(m,\ell)\right).\end{aligned} \tag{20}$$

By Hoeffding's inequality, for all $t>0$, and all $n\in[\![1,m]\!]$,

$$\begin{aligned}\mathbb{P}_p(|X_n-np|\geqslant t) &\leqslant 2\exp\left(-\frac{2t^2}{n}\right) \\ &\leqslant 2\exp\left(-\frac{2t^2}{m}\right).\end{aligned}$$

In particular, for any $\eta \in (0, 1/2)$ and for $t = m^{1-\eta}/\sqrt{2}$,

$$\mathbb{P}_p(|X_n - np| \geqslant m^{1-\eta}/\sqrt{2}) \leqslant 2\exp\left(-m^{1-2\eta}\right).$$

Thus

$$\begin{aligned}\mathbb{P}_p(\exists n \in [\![0, m]\!] \text{ s.t. } |X_n - np| \geqslant m^{1-\eta}/\sqrt{2}) &\leqslant \sum_{n=1}^{m} \mathbb{P}_p(|X_n - np| \geqslant m^{1-\eta}/\sqrt{2}) \\ &\leqslant 2m\exp\left(-m^{1-2\eta}\right).\end{aligned}$$

Set

$$A_m = \left\{\exists n \in [\![0, m]\!] \text{ s.t. } \|W_n - (m-n, (m-n)p)\|_1 \geqslant m^{1-\eta}/\sqrt{2}\right\}.$$

Then $\mathbb{P}_p(A_m \cap \{X_m = \ell\}) \leqslant 2m\exp\left(-m^{1-2\eta}\right)$ and, according to (20),

$$\mathbb{P}_{(m,\ell)}(A_m) = \frac{\mathbb{P}_p(A_m \cap \{X_m = \ell\})}{\mathbb{P}_p(X_m = \ell)} \leqslant \frac{2m\exp\left(-m^{1-2\eta}\right)}{\mathbb{P}_p(X_m = \ell)}$$

This is true for any $p \in (0, 1)$, thus for $p = \ell/m$ too, but, using Stirling formula, one finds

$$\begin{aligned}\mathbb{P}_{\ell/m}(X_m = \ell) &= \binom{m}{\ell}\left(\frac{\ell}{m}\right)^m\left(\frac{m-\ell}{m}\right)^{m-\ell} \\ &\sim \frac{1}{\sqrt{2\pi p(1-p)}}\,\frac{1}{\sqrt{m}}.\end{aligned}$$

Finally, $\mathbb{P}_{(m,\ell)}(A_m) = \mathcal{O}(m^{3/2}e^{-m^{1-2\eta}})$ and vanishes for $\eta \in (0, 1/2)$. For Pascal triangle, recall that $\gamma_{m,\ell}(t) = \ell t/m$, thus

$$A_m = \left\{\sup\left\{|w_m(t) - \gamma_{m,\ell}(t)|, mt \in [\![0, m]\!]\right\} \geq m^{-\eta}/\sqrt{2}\right\},$$

and

$$0 \leq \sup_{t\in[0,1]}\{|w_m(t) - \gamma_{m,\ell}(t)|\} - \sup_{mt\in[\![0,m]\!]}\{|w_m(t) - \gamma_{m,\ell}(t)|\} \leq \frac{\ell}{m^2} \leq \frac{1}{m},$$

so that, for m large enough,

$$\left\{\sup_{t\in[0,1]}|w_m(t) - \gamma_{m,\ell}(t)| \geq m^{-\eta}\right\} \subset A_m,$$

and, as expected,

$$\lim_m \mathbb{P}_{(m,\ell)}\left(\sup_{t\in[0,1]}|w_m(t) - \gamma_{m,\ell}(t)| \geq m^{-\eta}\right) = 0.$$

4.2 Proof of Theorem 2 : Stirling numbers of the first kind.

Consider two probability distributions for the processes (W, X, Y) defined at section 1.3. Under $\mathbb{P}_{(m,\ell)}$, W is a Markov chain starting from (m, ℓ), with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Stirling numbers *of the first kind*, and the processes (X, Y) are distributed accordingly. On the other hand, under $\mathbb{P}_\theta$, $(Y_i)_{i\geq 1}$ is a family of independent Bernoulli random variables with respective parameters $p_i = \theta/(i-1+\theta)$, and the processes (W, X, Y) are distributed accordingly : for instance, X_n can be seen as the number of non-empty tables after the arrival of the nth customer, as in Section 2.4. As before, according to Proposition 3, for any $\theta > 0$, and any set B in the relevant state space,

$$\begin{aligned}\mathbb{P}_{(m,\ell)}\left((W,X,Y)\in B\right) &= \mathbb{P}_\theta\left(\{(W,X,Y)\in B\}\cap\{X_m=\ell\}\right)/\mathbb{P}_\theta\left(X_m=\ell\right) \\ &= \mathbb{P}_\theta\left(\{(W,X,Y)\in B\}\cap\{X_m=\ell\}\right)/\mathbb{P}_\theta\left(W_0=(m,\ell)\right).\end{aligned} \tag{21}$$

Also, as in (14), recall that

$$\frac{\zeta}{(\zeta-1)\ln(1-\zeta)} = \frac{m}{\ell} = 1+\lambda,$$

and that, for $t \geq 0$,

$$\gamma_{m,\ell}(t) = \frac{1-\zeta}{\zeta}\ln\left(\frac{1-\zeta+t\,\zeta}{1-\zeta}\right).$$

Let $\mu = \mu_{\theta,n}$ denote the expectation of X_n under $\mathbb{P}_\theta$, that is

$$\mu = \mu_{\theta,n} = \sum_{k=1}^{n}\frac{\theta}{k-1+\theta},$$

and note that, for the choice $\theta_m = m(1-\zeta)/\zeta$,

$$|\mu_{\theta_m,n} - m\,\gamma_{m,\ell}(n/m)| \leq \frac{n\zeta}{m(1-\zeta)} \leq \frac{\zeta}{1-\zeta}. \tag{22}$$

According to Hoeffding's inequality, for all $t > 0$, and all $n \in [\![1, m]\!]$,

$$\begin{aligned}\mathbb{P}_\theta(|X_n-\mu_{\theta,n}|\geqslant t) &\leqslant 2\exp\left(-2t^2/n\right)\\ &\leqslant 2\exp\left(-2t^2/m\right).\end{aligned}$$

In particular, for all $\eta \in (0, 1/2)$ and for $t = m^{1-\eta}/\sqrt{2}$,

$$\mathbb{P}_\theta(|X_n-\mu_{\theta,n}|\geqslant m^{1-\eta}/\sqrt{2}) \leqslant 2\exp\left(-m^{1-2\eta}\right).$$

Set

$$A_m = \left\{\exists n\in[\![0,m]\!] \text{ s.t. } |X_n-\mu_{\theta,n}|\geqslant m^{1-\eta}/\sqrt{2}\right\}.$$

Thus

$$\begin{aligned}\mathbb{P}_\theta(A_m) &\leqslant \sum_{n=1}^{m}\mathbb{P}_\theta(|X_n-\mu_{\theta,n}|\geqslant m^{1-\eta}/\sqrt{2})\\ &\leqslant 2m\exp\left(-m^{1-2\eta}\right)\end{aligned}$$

Then $\mathbb{P}_\theta(A_m \cap \{X_m = \ell\}) \leqslant 2m \exp(-2m^{1-2\eta})$ and, according to (21),

$$\mathbb{P}_{(m,\ell)}(A_m) = \frac{\mathbb{P}_\theta(A_m \cap \{X_m = \ell\})}{\mathbb{P}_\theta(X_m = \ell)} \leqslant \frac{2m \exp(-m^{1-2\eta})}{\mathbb{P}_\theta(X_m = \ell)}$$

This is true for any $\theta > 0$, thus for $\theta_m = (1-\zeta)m/\zeta$ too, but, using relation (13) in [Goo61], one finds

$$\begin{aligned}\mathbb{P}_{\theta_m}(X_m = \ell) &= \frac{\theta_m^\ell}{(\theta_m)^{\uparrow m}} \begin{bmatrix} m \\ \ell \end{bmatrix} \mathbb{1}_{1 \leq \ell \leq m}, \\ &\sim \frac{1}{\sqrt{m}} \sqrt{\frac{\ln(1-\zeta)}{2\pi(1+\lambda)(\zeta + \ln(1-\zeta))}}.\end{aligned}$$

Finally, $\mathbb{P}_{(m,\ell)}(A_m) = \mathcal{O}(m^{3/2} e^{-m^{1-2\eta}})$ and vanishes for $\eta \in (0, 1/2)$. But, for m large enough,

$$B_m = \left\{ \sup_{t \in [0,1]} |w_m(t) - \gamma_{m,\ell}(t)| \geq m^{-\eta} \right\} \subset A_m,$$

and, as expected,

$$\lim_m \mathbb{P}_{(m,\ell)}(B_m) = 0.$$

Actually, due to (22), for $0 \leq n \leq m$,

$$\left| \frac{\mu_{N_0,n}}{m} - \gamma_{m,\ell}(n/m) \right| \leq \frac{\zeta}{m(1-\zeta)},$$

and

$$0 \leq \sup_{t \in [0,1]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} - \sup_{mt \in [\![0,m]\!]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} \leq \frac{1}{m},$$

thus $B_m \subset A_m$ provided that

$$\frac{m^{-\eta}}{\sqrt{2}} + \frac{\zeta}{m(1-\zeta)} + \frac{1}{m} \leq m^{-\eta}.$$

4.3 Proof of Theorem 2 : Stirling numbers of the second kind.

Consider two probability distributions for the processes (W, X, Y) defined at section 1.3. Under $\mathbb{P}_{(m,\ell)}$, W is a Markov chain starting from (m, ℓ), with transition probabilities $(p_\varepsilon(n,k))_{\varepsilon,n,k}$ related to Stirling numbers *of the second kind*, and the processes (X, Y) are distributed accordingly. On the other hand, under $\mathbb{P}_N$, X_n is the number of different coupons that have been collected after n draws with replacement in a collection of N available coupons, and the processes (X, Y, W) are distributed accordingly. According to Proposition 2, for any $N \geq \ell$, and any set B in the relevant state space,

$$\begin{aligned}\mathbb{P}_{(m,\ell)}((W,X,Y) \in B) &= \mathbb{P}_N(\{(W,X,Y) \in B\} \cap \{X_m = \ell\}) / \mathbb{P}_N(X_m = \ell) \qquad (23) \\ &= \mathbb{P}_N(\{(W,X,Y) \in B\} \cap \{X_m = \ell\}) / \mathbb{P}_N(W_0 = (m,\ell)).\end{aligned}$$

Let μ denote the expectation of X_n under $\mathbb{P}_N$, that is

$$\mu = \mu_{N,n} = N\left(1 - \left(1 - \frac{1}{N}\right)^n\right),$$

and note that

$$\left|\mu_{N,n} - N\left(1 - e^{-\frac{n}{N}}\right)\right| \leq \frac{n}{2N}. \tag{24}$$

Also, as in (13), set

$$\frac{\zeta}{1 - e^{-\zeta}} = 1 + \lambda = \frac{m}{\ell}.$$

According to [MR95, Ch. 4, Theorem 4.18], by Azuma-Hoeffding's inequality, for all $t > 0$, and all $n \in [\![1, m]\!]$,

$$\begin{aligned}\mathbb{P}_N(|X_n - \mu_{N,n}| \geqslant t) &\leqslant 2\exp\left(-t^2\,\frac{N - 1/2}{N^2 - \mu_{N,n}^2}\right)\\ &\leqslant 2\exp\left(-\frac{t^2}{2N}\right).\end{aligned}$$

In particular, for all $\eta \in (0, 1/2)$ and for $t = m^{1-\eta}/\sqrt{2}$,

$$\mathbb{P}_N(|X_n - \mu_{N,n}| \geqslant m^{1-\eta}/\sqrt{2}) \leqslant 2\exp\left(-m^{2-2\eta}/4N\right).$$

Set

$$A_m = \left\{\exists n \in [\![0, m]\!] \text{ s.t. } |X_n - \mu_{N,n}| \geqslant m^{1-\eta}/\sqrt{2}\right\}.$$

Thus

$$\begin{aligned}\mathbb{P}_N(A_m) &\leqslant \sum_{n=1}^{m} \mathbb{P}_N(|X_n - \mu_{N,n}| \geqslant m^{1-\eta}/\sqrt{2})\\ &\leqslant 2m\exp\left(-m^{2-2\eta}/4N\right)\end{aligned}$$

Then $\mathbb{P}_N(A_m \cap \{X_m = \ell\}) \leqslant 2m\exp\left(-m^{2-2\eta}/4N\right)$ and, according to (23),

$$\mathbb{P}_{(m,\ell)}(A_m) = \frac{\mathbb{P}_N(A_m \cap \{X_m = \ell\})}{\mathbb{P}_N(X_m = \ell)} \leqslant \frac{2m\exp\left(-m^{2-2\eta}/4N\right)}{\mathbb{P}_N(X_m = \ell)}$$

This is true for any $N \geq \ell$, thus for $N_0 = \lceil m/\zeta \rceil$ too, but, using relation (3) in [Goo61], one finds

$$\begin{aligned}\mathbb{P}_{\lceil m/\zeta \rceil}(X_m = \ell) &= \frac{N_0!\,N_0^{-m}}{N_0 - \ell!}\left\{{m \atop \ell}\right\}\\ &\sim \sqrt{\frac{\zeta\ e^{\zeta}}{2\pi(\zeta - \lambda)m}}.\end{aligned}$$

Finally, $\mathbb{P}_{(m,\ell)}(A_m) = \mathcal{O}(m^{3/2}e^{-\zeta m^{1-2\eta}})$ and vanishes for $\eta \in (0, 1/2)$. For Stirling numbers of the second kind, recall that $\gamma_{m,\ell}(t) = \left(1 - e^{-\zeta t}\right)/\zeta$, thus

$$B_m = \left\{\sup\left\{\left|w_m(t) - \gamma_{m,\ell}(t)\right|, mt \in [\![0, m]\!]\right\} \geq m^{-\eta}\right\} \subset A_m.$$

Actually, due to (24), for $0 \leq n \leq m$,

$$\begin{aligned}\left|\frac{\mu_{N_0,n}}{m} - \gamma_{m,\ell}(n/m)\right| &\leq \frac{n}{2N_0 m} + \left|\frac{N_0}{m}\left(1 - e^{-\frac{n}{N_0}}\right) - \gamma_{m,\ell}(n/m)\right| \\ &\leq \frac{n}{2N_0 m} + \left|\frac{1 - e^{-\tilde{\zeta} n/m}}{\tilde{\zeta}} - \frac{1 - e^{-\zeta n/m}}{\zeta}\right| \\ &\leq \frac{\tilde{\zeta}}{2m} + \frac{1 + \tilde{\zeta}}{m}, \end{aligned} \tag{25}$$

in which

$$\frac{m}{\tilde{\zeta}} = \left\lceil \frac{m}{\zeta} \right\rceil = N_0, \quad \text{thus} \quad 0 \leq \zeta - \tilde{\zeta} \leq \frac{\zeta\tilde{\zeta}}{m}.$$

Thus $B_m \subset A_m$ for m large enough, i.e. provided that

$$\frac{m^{-\eta}}{\sqrt{2}} + \frac{\tilde{\zeta}}{2m} + \frac{1 + \tilde{\zeta}}{m} \leq m^{-\eta}.$$

Finally

$$0 \leq \sup_{t \in [0,1]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} - \sup_{mt \in [\![0,m]\!]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} \leq \frac{1}{m},$$

so that, for m large enough,

$$\left\{\sup_{t \in [0,1]} |w_m(t) - \gamma_{m,\ell}(t)| \geq m^{-\eta}\right\} \subset A_m, \tag{26}$$

and, as expected,

$$\lim_m \mathbb{P}_{(m,\ell)}\left(\sup_{t \in [0,1]} |w_m(t) - \gamma_{m,\ell}(t)| \geq m^{-\eta}\right) = 0.$$

4.4 Application to the enumeration of accessible complete deterministic automata with k letters and n vertices

Let $a_{k,n}$ denote the number of accessible complete deterministic automata (ACDA) with k letters and n vertices (see [Nic00, AC19] for definitions). According to Koršunov [Kor78, Kor86], for any given $k \geq 2$,

$$a_{k,n} \sim c_k \left\{ {kn+1 \atop n} \right\} n!, \tag{27}$$

in which ζ_2 is defined by (4), and

$$c_k = 1 - k\, e^{-\zeta_2(k-1)}. \tag{28}$$

Following [AC19], this section gives a probabilistic interpretation of Koršunov's formula, that relies on Theorem 2 for Stirling numbers of the second kind : according to [Nic00], there exists a bijection between the set of ACDA with k letters and n vertices and a subset $\mathcal{A}_{k,n}$ of the set $\Omega_{kn+1,n}$ of surjections from $[\![kn+1]\!]$ to $[\![n]\!]$. Thus (27) states that the ratio $\#\mathcal{A}_{k,n}/\#\Omega_{kn+1,n}$ converges to c_k with n. But an element of $\Omega_{kn+1,n}$ can be seen as the sample path of a coupon collector process such that the collection of n items is complete at step $kn+1$. As a consequence, in the notations of Section 4.3,

$$\mathbb{P}_{(kn+1,n)}(\mathcal{A}_{k,n}) = \frac{\#\mathcal{A}_{k,n}}{\#\Omega_{kn+1,n}} = \frac{a_{k,n}}{\left\{ {kn+1 \atop n} \right\} n!}, \tag{29}$$

and Koršunov's formula can be rephrased as

$$\lim_n \mathbb{P}_{(kn+1,n)}(\mathcal{A}_{k,n}) = c_k. \tag{30}$$

Now, according to [Nic00], $\mathcal{A}_{k,n}$ is the set of elements $\omega \in \Omega_{kn+1,n}$ such that

$$\forall \ell \in [\![0, n-1]\!], \quad X_{\ell k+1}(\omega) \geq \ell + 1,$$

or, equivalently,

$$\forall \ell \in [\![0, kn]\!], \quad k\, X_\ell(\omega) \geq \ell. \tag{31}$$

Relation (31) is, as usual, required from a breadth first search walk to insure the connexity of the underlying graph.

We shall now sketch the argument, taken from [AC19], which, using Theorem 2, shows that $\Upsilon_n = \overline{\mathcal{A}_{k,n}}$ satisfies

$$\lim_n \mathbb{P}_{(kn+1,n)}(\Upsilon_n) = 1 - c_k = k\, e^{-\zeta_2(k-1)}.$$

Note that $\lim_n \lambda(kn+1, n) = k-1$, and that the corresponding concave limit field line,

$$\gamma_{m,\ell}(t) = \frac{1 - e^{-\zeta_2(k-1)t}}{\zeta_2(k-1)},$$

crosses the line $y = x/k$ only at its endpoints $(0,0)$ and $(k,1)$, so that, according to Theorem 2, but for an exponentially small probability, the sample path $\{(\ell, X_\ell), 0 \leq \ell \leq kn+1\}$ crosses the line $y = x/k$ only close to its endpoints. The probability that such a crossing occurs close to $(0,0)$ is very small too, see [AC19, Proposition 2]. As a consequence, $\mathbb{P}_{(kn+1,n)}(\Upsilon_n)$ has the same asymptotic behaviour than the probability that the sample path $\{(\ell, X_\ell), 0 \leq \ell \leq kn+1\}$ crosses the line $y = x/k$ close to its endpoint $(kn+1, n)$. Close to this endpoint, the sample path has approximately the same transition probabilities as a standard random walk with step distribution

$$\left(1 - e^{-\zeta_2(k-1)}\right) \delta_0 + e^{-\zeta_2(k-1)} \delta_{-1}.$$

The probability of a crossing of the line $y = x/k$ by such a standard random walk is $1 - c_k = k\, e^{-\zeta_2(k-1)}$, as follows for instance from the Pollaczek-Khinchine formula (cf. Corollary 6.6 of [Asm03], or Proposition 3 of [AC19], in which one can find a proof of Koršunov's formula along these lines).

Acknowledgements.

The authors are grateful to Antoine Lejay and Rémi Peyre for fruitful discussions, and to Antoine Lejay for pointing reference [Kol35].

References

[AC19] Anis Amri and Philippe Chassaing, *The impatient collector*, working paper or preprint, June 2019.

[Ant74] Charles E. Antoniak, *Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems*, Ann. Statist. **2** (1974), 1152–1174. MR 365969

[Asm03] Sören Asmussen, *Applied probability and queues*, second ed., Applications of Mathematics (New York), vol. 51, Springer-Verlag, New York, 2003, Stochastic Modelling and Applied Probability. MR 1978607

[Ben73] Edward A. Bender, *Central and local limit theorems applied to asymptotic enumeration*, J. Combinatorial Theory Ser. A **15** (1973), 91–111. MR 375433

[DF91] P. Diaconis and W. Fulton, *A growth model, a game, an algebra, Lagrange inversion, and characteristic classes*, vol. 49, 1991, Commutative algebra and algebraic geometry, II (Italian) (Turin, 1990), pp. 95–119 (1993). MR 1218674

[FS09] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR 2483235

[Goo61] I. J. Good, *An asymptotic formula for the differences of the powers at zero*, Ann. Math. Statist. **32** (1961), 249–256. MR 0120204

[Ken75] Douglas P. Kennedy, *The Galton-Watson process conditioned on the total progeny*, J. Appl. Probab. **12** (1975), 800–806 (English).

[Kol35] A. Kolmogoroff, *Zur Theorie der Markoffschen Ketten*, Math. Ann. **112** (1935), 155–160 (German).

[Kor78] A. D. Koršunov, *Enumeration of finite automata*, Problemy Kibernet. (1978), no. 34, 5–82, 272. MR 517814

[Kor86] ———, *On the number of nonisomorphic strongly connected finite automata*, Elektron. Informationsverarb. Kybernet. **22** (1986), no. 9, 459–462. MR 862029

[LBG92] Gregory F. Lawler, Maury Bramson, and David Griffeath, *Internal diffusion limited aggregation*, Ann. Probab. **20** (1992), no. 4, 2117–2140 (English).

[Mit20] Kiana Mittelstaedt, *A stochastic approach to Eulerian numbers*, Am. Math. Mon. **127** (2020), no. 7, 618–628 (English).

[MR95] Rajeev Motwani and Prabhakar Raghavan, *Randomized algorithms*, Cambridge: Cambridge Univ. Press, 1995 (English).

[Nic00] Cyril Nicaud, *Étude du comportement en moyenne des automates finis et des langages rationnels*, Ph.D. thesis, Paris VII, 2000.

[Pit06] J. Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR 2245368

[Sti73] Stephen M. Stigler, *Studies in the history of probability and statistics. XXXII: Laplace, Fisher, and the discovery of the concept of sufficiency*, Biometrika **60** (1973), 439–445 (English).

[Tan73] S. Tanny, *A probabilistic interpretation of Eulerian numbers*, Duke Math. J. **40** (1973), 717–722 (English).