The Depoissonisation Quintet: Rice–Poisson–Mellin–Newton–Laplace

Brigitte Vallée
GREYC Laboratory, CNRS and University of Caen, France
brigitte.vallee@unicaen.fr

Abstract

This paper is devoted to the Depoissonisation process which is central in various analyses of the AofA domain. We first recall in Section 1 the two possible paths that may be used in this process, namely the Depoissonisation path and the Rice path. The two paths are rarely described for themselves in the literature, and general methodological results are often difficult to isolate amongst particular results that are more directed towards various applications. The main results for the Depoissonisation path are scattered in at least five papers, with a chronological order which does not correspond to the logical order of the method. The Rice path is also almost always presented with a strong focus towards possible applications. It is often very easy to apply, but it needs a tameness condition, which appears *a priori* to be quite restrictive, and is not deeply studied in the literature. This explains why the Rice path is very often undervalued.

Second, the two paths are not precisely compared, and the situation creates various "feelings": some people see the tools that are used in the two paths as quite different, and strongly prefer one of the two paths; some others think the two paths are almost the same, with just a change of vocabulary. It is thus useful to compare the two paths and the tools they use. This will be done in Sections 2 and 3. We also "follow" this comparison on a precise problem, related to the analysis of tries, introduced in Section 1.7.

The paper also exhibits in Section 4 a new framework, of practical use, where the tameness condition of Rice path is proven to hold. This approach, perhaps of independent interest, deals with the shifting of sequences and then the inverse Laplace transform, which does not seem of classical use in this context. It performs very simple computations. This adds a new method to the Depoissonisation context and explains the title of the paper. We then conclude that the Rice path is both of easy and practical use: even though (much?) less general than the Depoissonisation path, it is easier to apply.

2012 ACM Subject Classification Mathematics of computing

Keywords and phrases Analysis of Algorithms, Poisson model, Mellin transform, Rice integral, Laplace transform, Newton interpolation, Depoissonisation, sources, trie structure, algorithms on words

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.35

Related Version Complete version at https://arxiv.org/pdf/1802.04988.

Funding ANR Project MetaConc (ANR-15-CE40-0017-01) between Taiwan and France: Méthodes analytiques non conventionnelles en Combinatoire

Acknowledgements I wish to warmly thank Hsien-Kuei Hwang for many discussions about the Depoissonisation process. He notably introduced to me the Poisson Charlier expansion. I also wish to thank one of the referees, for precise reading of the manuscript, and many helpful remarks.

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 35; pp. 35:1–35:20
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 General framework

This section first recalls the two probabilistic models, the Bernoulli model and the Poisson model, together with the two main objects attached to a sequence f: the classical Poisson transform P_f, and another sequence, denoted as $\Pi[f]$ and called here the Poisson sequence. We insist on the involutive characteristic of the mapping Π and introduce two new notions, shift and canonical sequences. After the description of the two paths of interest, we present analyses on tries which strongly motivate the work, and will be performed within each path.

1.1 Probabilistic settings

Many algorithms deal with inputs that are finite sequences of data. We give some examples: (a) for text algorithms, data are words, and inputs are finite sequences of words; (b) for geometric algorithms, data are points, and inputs are finite sequences of points; (c) for a source, data are symbols, and inputs are finite sequences of symbols, namely finite words. The cardinality of the input sequence is often chosen as the input size, and, as usual, one is interested in the asymptotic behaviour of the algorithm for large input size.

The probabilistic framework is as follows: Each data (word or point) is produced along a distribution, and the set of data is thus a probabilistic space $(\mathcal{X}, \mathbb{P})$. Very often, the data are independently chosen with the same distribution and the set $(\mathcal{X}^n, \mathbb{P}_{[n]})$ is the product of order n of the space $(\mathcal{X}, \mathbb{P})$. The space of all the inputs is thus the set $\mathcal{X}^\star := \sum_{n\geq 0} \mathcal{X}^n$ of finite sequences $\boldsymbol{x}$ of elements of $\mathcal{X}$, and there are two main probabilistic models:

(i) The Bernoulli model $\mathcal{B}_n$ (more natural in algorithmics), where the cardinality $N(\boldsymbol{x})$ of $\boldsymbol{x}$ is fixed and equal to n (then tends to ∞);

(ii) The Poisson model $\mathcal{P}_z$ of parameter z, where the cardinality $N(\boldsymbol{x})$ is a random variable that follows a Poisson law of parameter z, where the fixed parameter z tends also to ∞,

$$\mathbb{P}[N(\boldsymbol{x}) = n] = e^{-z}\frac{z^n}{n!}\,.$$

This model has very nice probabilistic properties, notably properties of independence.

1.2 The Poisson transform and the Poisson sequence

There is a variable (or a cost) $R : \mathcal{X}^\star \to \mathbb{N}$ which describes the behaviour of the algorithm on the input; for instance, for $\boldsymbol{x} \in \mathcal{X}^\star$, $R(\boldsymbol{x})$ is the path length of a tree (trie or dst) built on the sequence $\boldsymbol{x} := (x_1, \ldots, x_n)$ of words x_i. Our final aim is the analysis of R in the Bernoulli model $\mathcal{B}_n$, i.e., the asymptotic study of the sequence $f : n \mapsto f(n)$, where $f(n) := \mathbb{E}_{[n]}[R]$ is the expectation of R in the model $\mathcal{B}_n$. We begin with the easier Poisson model $\mathcal{P}_z$, and study the expectation $\mathbb{E}_z[R]$ in the model $\mathcal{P}_z$ that satisfies

$$\mathbb{E}_z[R] = \sum_{n\geq 0} \mathbb{E}_z[R\,|N=n]\,\mathbb{P}_z[N=n] = \sum_{n\geq 0} \mathbb{E}_{[n]}[R]\,\mathbb{P}_z[N=n] = e^{-z}\sum_{n\geq 0} f(n)\frac{z^n}{n!}\,.$$

This leads us to the Poisson transform P_f of the sequence $f : n \mapsto f(n)$, that is written as an exponential generating function (with "signs")[1] and thus defines another sequence p,

$$P_f(z) := e^{-z}\sum_{n\geq 0} f(n)\frac{z^n}{n!} = \sum_{k\geq 0}(-1)^k\frac{z^k}{k!}p(k)\,, \quad \text{with} \quad p(k) := (-1)^k k![z^k]P_f(z)\,. \qquad (1)$$

[1] The Poisson transform is often called the Poisson generating function. The signs are added in order to get an involutive formula in (2).

▶ **Definition 1.** Consider a sequence $f : n \mapsto f(n)$. Then,

(a) the series P_f defined in (1) is the *Poisson transform* of the sequence f;

(b) the sequence $p : k \mapsto p(k)$ defined in (1) is the *Poisson sequence* of the sequence f.

The following holds:

▶ **Lemma 2.** *Consider a cost R defined on $\mathcal{X}^\star$, its expectation $f(n)$ in the model $\mathcal{B}_n$. Then,*

(a) *its expectation in the model $\mathcal{P}_z$ is the Poisson transform $P_f(z)$;*

(b) *there are binomial relations between the sequences f and p, namely*

$$p(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(k), \quad \text{and} \quad f(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} p(k)\,; \tag{2}$$

(c) *the map Π which associates with the sequence f the sequence p is involutive.*

1.3 Description of the two paths.

We only deal here with a sequence f of polynomial growth, for which the Poisson transform $z \mapsto P_f(z)$ is entire. The sequence f is often given in an implicit way, and we assume that we have some knowledge on $P_f(z)$, which may be of two different types

(a) about the Poisson transform $P_f(z)$ itself,

(b) about its coefficients, namely the sequence $\Pi[f]$.

The main question is now: *Is it possible to return to the initial sequence f and obtain some knowledge about its asymptotics?* There are two main return paths, one for each framework: the Depoissonisation path for (a), and the Rice path for (b). We first describe in Section 1.4 the classical toolbox, then, in Section 1.5, a new useful tool. As we aim to provide a precise comparison between the two paths, we perform a kind of "test" on a particular instance which arises when analysing the *trie* structure and is introduced in Sections 1.6 and 1.7.

1.4 Toolbox and main definitions

This section gathers various definitions about domains of the plane, behaviours of functions. Itthen presents the Mellin transform.

Cones and vertical strips. There are two important types of domains of the complex plane we deal with.

(i) The cones built on the real line $\mathbb{R}^+$, with two possible definitions,

$$\mathcal{C}(a, \theta) := \{z \mid |\arg(z - a)| < \theta\} \quad \text{for } \theta < \pi$$
$$\widehat{\mathcal{C}}(a, \gamma) = \{z \mid \Re(z - a) > \gamma |z - a|\} \text{ for } |\gamma| \le 1\,,$$

related by the relation $\widehat{\mathcal{C}}(a, \cos\theta) = \mathcal{C}(a, \theta)$. When $a = 0$, it is omitted.

(ii) The vertical strips, or halfplanes: $\mathcal{S}(a, b) := \{z \mid a < \Re z < b\}$, $\mathcal{S}(a) := \{z \mid \Re z > a\}$.

Polynomial growth. This notion plays a fundamental role: A function $s \mapsto \varpi(s)$ defined in an unbounded domain $\Omega \subset \mathbb{C}$ is said to be of polynomial growth if there exists r for which the estimate $|\varpi(s)| = O(|s|^r)$ holds as $s \to \infty$ on Ω. When $\Omega \subset \mathcal{S}(a, b)$, this means: $|\varpi(s)| = O(|\Im s|^r)$; when $\Omega \subset \mathcal{C}(\theta)$ with $\theta < \pi/2$, this means: $|\varpi(s)| = O(|\Re s|^r)$;

Tameness. A function $s \mapsto \varpi(s)$ is tame on $\Re s > c$ when it is analytic and of polynomial growth there. This notion is extended when $\varpi(s)$ stops being analytic on $\Re s = c$. We will say that ϖ is tame at $s = c$ if it is meromorphic and of polynomial growth in a larger region $\mathcal{R}$ on the left of the line $\Re s = c$ delimited by a frontier curve $\mathcal{F}$. (see [4] and the Annex).

Mellin transform. The Mellin transform of a function Q defined in $[0, +\infty]$ is defined as

$$Q^\star(s) := \int_0^{+\infty} Q(u)\, u^{s-1} du\,.$$

The Mellin transform plays a central role in each of the two paths (see its main properties in the survey paper [7]). In particular, the transform has a nice behaviour on harmonic sums:

$$Q(z) = \sum_k g(\mu_k z) \quad \Longrightarrow \quad Q^\star(s) = \left(\sum_k \mu_k^{-s}\right) g^\star(s)\,. \tag{3}$$

Moreover, the following lemma[2] proves that the function $\Gamma(s)$ and its derivatives $\Gamma^{(m)}(s)$ are exponentially small along vertical lines (when $|\Im(s)| \to \infty$).

▶ **Lemma 3** (Exponential Smallness Lemma, [7]). *If, inside the closure of the cone $\mathcal{C}(\theta)$ with $\theta > 0$, one has $Q(z) = O(|z|^{-\alpha})$ as $z \to 0$ and $Q(z) = O(|z|^{-\beta}$ as $|z| \to \infty$, then the estimate $Q^*(s) = O(\exp[-\theta|\Im(s)|])$ uniformly holds in the vertical strip $\mathcal{S}(\alpha, \beta)$.*

1.5 A first new tool: Shift and canonical sequences

The notions that are presented here are not introduced in this way in the literature, and, in particular, the notion of canonical sequence appears to be new (and useful), notably in Section 4.

▶ **Definition 4.** Consider a non zero real sequence $n \mapsto f(n)$.

(a) Its degree $\deg(f)$ and its valuation $\mathrm{val}(f)$ are defined as

$$\texttt{deg}(f) := \inf\{c \mid f(k) = O(k^c)\} \qquad \texttt{val}(f) := \min\{k \mid f(k) \neq 0\}\,.$$

A sequence f with finite degree is said to be of polynomial growth.

(b) A sequence $n \mapsto f(n)$ satisfies the Valuation-Degree Condition (VD), iff $\texttt{val}(f) > \texttt{deg}(f) + 1$.

(c) It is reduced if it satisfies $\texttt{val}(f) = 0$ and $\texttt{deg}(f) < -1$.

The VD Condition is essential in the Rice path. As we are (only) interested in the asymptotics of the sequence f, the VD condition is easy to ensure, as we now show: With a sequence F of polynomial growth, we associate the integer

$$\sigma(F) := 1 \quad (\text{if } \texttt{deg}(F) < 0), \quad \sigma(F) := 2 + \lfloor \texttt{deg}(F) \rfloor \quad (\text{if } \texttt{deg}(F) \geq 0)\,, \tag{4}$$

that satisfies the inequality $\sigma(F) > \texttt{deg}(F) + 1$; we only modify the first terms of the sequence F : we put zeroes for indices $k < \sigma(F)$ together with one 1 for $k = \sigma(F)$, and obtain a new sequence $\tau(F)$ of valuation $\sigma(F)$ that keeps the same asymptotics as the initial sequence F and now satisfies the VD condition.

We will need here to deal with the stronger notion of reduced sequences, and we now explain how to associate with a sequence F, and in a canonical way, a reduced sequence.

---

[2] It is called the Exponential Smallness Lemma in the paper [13], and we keep the same terminology.

▶ **Lemma 5.** *Consider the shifting map T which associates with a sequence f the sequence $T(f)$ defined, for any $n \geq 0$ as*

$$T[f](n) = \frac{f(n+1)}{n+1} \qquad \text{and thus, for } m \geq 1, \text{ as} \qquad T^m[f](n) = \frac{f(n+m)}{(n+1)\dots(n+m)}\,.$$

For $m \geq 1$, the inverse mapping T^{-m} associates with a sequence g the sequence f defined as $f(n) = n(n-1)\dots(n-m+1)\,g(n-m), \quad$ for $n \geq m$.

(a) *The shifting T^m anti-commutes with the involution Π, namely $T^m \circ \Pi = (-1)^m \Pi \circ T^m$.*

(b) *The sequence $\rho(F) := T^{\sigma(F)}(\tau(F))$ associated with F is reduced, with a degree equal to $\texttt{deg}(F) - \sigma(F)$. It is called the canonical sequence associated to F.*

Proof. Start with the sequence f with valuation ℓ. Then the Poisson transform $P_f(z)$ has itself valuation ℓ and is written as

$$P_f(z) = z^\ell Q(z) \quad \text{with} \quad Q(z) = e^{-z} \sum_{k\geq 0} g(k)\,\frac{z^k}{k!} = \sum_{k\geq 0} (-1)^k \frac{z^k}{k!} q(k)\,. \tag{5}$$

Then, the two sequences g and $q := \Pi[g]$ associated with f via Eqn (5) are expressed with the iterate of T of order ℓ, namely $g = T^\ell[f], \qquad q := \Pi[g] = (-1)^\ell T^\ell[\Pi[f]]$. ◀

In the sequel, it will be then *sufficient* to deal with the *canonical* sequence $\rho(F)$, and its Poisson sequence $\Pi(\rho(F)) = (-1)^{\sigma(F)} \rho(\Pi(F))$. Then, the results on the asymptotics on $\rho(F)$ will be easily transfered on the initial Poisson pair of F with Properties (a) and (b).

Example. In Section 1.6, we will deal with the following sequences F_0, F_1, F_2, all of valuation 2, which satisfy moreover $F_0(k) = 1, \quad F_1(k) = k, \quad F_2(k) = k \log k, \qquad$ for $k \geq 2$. Their canonical sequences are defined for $k \geq 0$, as

$$f_0(k) = f_1(k) = \frac{1}{(k+1)(k+2)}, \qquad f_2(k) = \frac{\log(k+3)}{(k+1)(k+2)}\,.$$

1.6 An instance of the context. Probabilistic analysis of tries

A source $\mathcal{S}$ is a probabilistic process which produces infinite words on the (finite) alphabet $\Sigma := [0..r-1]$. A trie is a tree structure, used as a dictionary, which compares words via their prefixes. Given a finite sequence $\boldsymbol{x}$ of (infinite) words emitted by the source $\mathcal{S}$, the trie $\mathcal{T}(\boldsymbol{x})$ built on the sequence[3] $\boldsymbol{x}$ is defined recursively by the following three rules which involve the cardinality $N(\boldsymbol{x})$ of the sequence $\boldsymbol{x}$:

(a) If $N(\boldsymbol{x}) = 0$, then $\mathcal{T}(\boldsymbol{x}) = \emptyset$

(b) If $N(\boldsymbol{x}) = 1$, with $\boldsymbol{x} = (x)$, then $\mathcal{T}(\boldsymbol{x})$ is a leaf labeled by x.

(c) If $N(\boldsymbol{x}) \geq 2$, then $\mathcal{T}(\boldsymbol{x})$ is formed with an internal node and r subtries equal to

$$\mathcal{T}(\boldsymbol{x}_{\langle 0\rangle}), \dots, \mathcal{T}(\boldsymbol{x}_{\langle r-1\rangle})\,,$$

where $\boldsymbol{x}_{\langle\sigma\rangle}$ denotes the sequence consisting of words of $\boldsymbol{x}$ which begin with symbol σ, stripped of their initial symbol σ. If the set $\boldsymbol{x}_{\langle\sigma\rangle}$ is non empty, the edge which links the subtrie $\mathcal{T}(\boldsymbol{x}_{\langle\sigma\rangle})$ to the internal node is labelled with the symbol σ.

[3] The trie depends only on the underlying set $\{x_1, x_2, \dots, x_n\}$.

Iterating the process, we consider, for a finite prefix $\boldsymbol{w}$, the sequence $\boldsymbol{x}_{\langle \boldsymbol{w} \rangle}$ consisting of words of $\boldsymbol{x}$ which begin with prefix $\boldsymbol{w}$, stripped of their initial prefix $\boldsymbol{w}$, and denote by $N_{\boldsymbol{w}}(\boldsymbol{x}) := N(\boldsymbol{x}_{\langle \boldsymbol{w} \rangle})$ the cardinality of such a sequence. Then, the internal nodes are used for directing the search: they are labelled by prefixes $\boldsymbol{w}$ with $N_{\boldsymbol{w}}(\boldsymbol{x}) \geq 2$. The leaves contain suffixes of $\boldsymbol{x}$, and there are as many leaves as words in $\boldsymbol{x}$.

Trie analysis aims at describing the average shape of a trie. We focus here on *additive* parameters, whose (recursive) definition exactly copies the (recursive) definition of the trie. With a sequence $f : \mathbb{N} \to \mathbb{R}$ – called a *toll* – which satisfies $f(0) = f(1) = 0$ and $f(k) \geq 0$ for $k \geq 2$, we associate a random variable R defined on the set $\mathcal{X}^\star$ as follows:

(ab) If $N(\boldsymbol{x}) \leq 1$, then $\quad R(\boldsymbol{x}) = 0$;

(c) if $N(\boldsymbol{x}) \geq 2$, then $\quad R(\boldsymbol{x}) = f(N(\boldsymbol{x})) + \sum_{\sigma \in \Sigma} R(\boldsymbol{x}_{\langle \sigma \rangle})$.

Iterating the recursion leads to the expression

$$R(\boldsymbol{x}) := \sum_{\boldsymbol{w} \in \Sigma^\star} f(N_{\boldsymbol{w}}(\boldsymbol{x}))\,. \tag{6}$$

The probabilistic properties of R will depend both on the toll f and the source $\mathcal{S}$:

– The probabilistic properties of the source $\mathcal{S}$ are encapsulated in the Dirichlet series $\Lambda(s)$ of the source, introduced in [19], and defined with the fundamental probabilities $\pi_{\boldsymbol{w}}$,

$$\Lambda(s) := \sum_{\boldsymbol{w} \in \Sigma^\star} \pi_{\boldsymbol{w}}^s\,, \qquad \text{with} \quad \pi_{\boldsymbol{w}} := \mathbb{P}[\text{a word emitted by } \mathcal{S} \text{ begins with the prefix } \boldsymbol{w}]\,. \tag{7}$$

The series $\Lambda(s)$ mainly intervenes via its behaviour near $s = 1$. We consider here a *tame* source, for which $s \mapsto \Lambda(s)$ is tame at $s = 1$, with a simple pole at $s = 1$ whose residue equals $1/h(\mathcal{S})$ where $h(\mathcal{S})$ is the entropy of the source. (See [4] about tameness of sources.)

– Here are some instances of natural tolls : the size is associated to the toll $f(k) = 1$ (for $k \geq 2$) and the path length to the toll $f(k) = k$ (for $k \geq 2$). A version of the `QuickSort` algorithm on words [4] leads to the *sorting toll* $f(k) = k \log k$ (for $k \geq 2$).

We focus here on this last toll, and are interested in the analysis of the associated cost R. The analysis was already performed in [4] with Depoissonisation path (a). We would have wished there to use the Rice path (b) (as we got used in our previous analyses) but we did not succeed. This failure was a strong motivation for the present study, and we now present here two proofs for the following result, each of them using one path.

▶ **Theorem 6.** *Consider a trie built on n words emitted by a* tame *source $\mathcal{S}$. Then the mean value of parameter R associated with the sorting toll f satisfies in the Bernoulli model $\mathcal{B}_n$*

$$r(n) \sim \frac{1}{2\,h(\mathcal{S})}\, n \log^2 n \qquad (n \to \infty)\,.$$

1.7 Main principles of trie analysis

We begin to deal with the Poisson model, that presents the following advantage: In the model $\mathcal{P}_z$, the cardinality $N_{\boldsymbol{w}}$ which appears in Eqn (6) follows a Poisson law of rate $z\,\pi_{\boldsymbol{w}}$ that involves the fundamental probability $\pi_{\boldsymbol{w}}$ defined in (7). We then adapt the general framework defined in Subsection 1.3, both for the initial sequence f and for the sequence r, and consider the two paths:

– Path (a) deals with the Poisson transforms $P_r(z)$ and $P_f(z)$; averaging Relation (6) in the model $\mathcal{P}_z$ entails a relation between the two Poisson transforms

$$P_r(z) = \sum_{\boldsymbol{w} \in \Sigma^\star} \mathbb{E}_z[f(N_{\boldsymbol{w}})] = \sum_{\boldsymbol{w} \in \Sigma^\star} P_f(z\,\pi_{\boldsymbol{w}}). \tag{8}$$

Then, the function $P_r(z)$ writes as a harmonic sum, and, with (3), its Mellin transform $P_r^*(s)$ factorises and involves the Λ function defined in (7), namely $P_r^*(s) = \Lambda(-s) \cdot P_f^*(s)$.

– Path (b) deals with the Poisson sequences $q = \Pi[r]$ and $p = \Pi[f]$. Then, Relation (8) entails the equality which also involves the Λ function, namely: $q(n) = \Lambda(n)\, p(n)$ for $n \geq 2$.

2 The Depoissonization path

We first provide a general description of the path; then, we apply it to the analysis of tries and obtain a first proof of Theorem 6.

2.1 General description

We first describe the main steps of the path in an informal way.

Main steps. The Depoissonization path deals with the Poisson transform $P_f(z)$:

(a) It compares $f(n)$ and $P_f(n)$ via the Poisson–Charlier expansion.
(b) It uses the tameness of the Mellin transform $P_f^\star(s)$ for the asymptotics of $P_f(n)$.
(c) Under Conditions $(\mathcal{JS})$ on the Poisson transform $P_f(z)$, the Poisson-Charlier expansion may be truncated and provides the asymptotic of $f(n)$ with a good remainder.
(d) Moreover, there exists a Condition $(\mathcal{DP})$ on the input sequence f under which the Conditions $(\mathcal{JS})$ hold.

We then describe more precisely the main objects that are involved.

The Poisson-Charlier expansion. Using the Taylor expansion of $P_f(z)$ at $z = n$, the term $f(n)$ admits an (infinite) expansion,

$$f(n) := n![z^n]\left(e^z P(z)\right) = \sum_{j \geq 0} \frac{P^{(j)}(n)}{j!} \tau_j(n),$$

where the coefficient $\tau_j(n) := n![z^n]\left((z-n)^j e^z\right)$ is a polynomial in n of degree $\lfloor j/2 \rfloor$, closely related to the (classical) Charlier polynomial.

Conditions $(\mathcal{JS})$**.** *An entire function $P(z)$ satisfies the Conditions $\mathcal{JS}(\alpha, \beta)$ if there exist $\theta \in]0, \pi/2[$, and $\delta < 1$ for which one has, for $z \to \infty$:*

(I) Inside cone $\mathcal{C}(\theta)$, one has $|P(z)| = O\left(|z|^\alpha \log^\beta(1 + |z|)\right)$.

(O) Outside cone $\mathcal{C}(\theta)$, one has $|P(z)e^z| = O\left(e^{\delta|z|}\right)$.

Condition $(\mathcal{DP})$**.** *There exists an analytic lifting φ for the sequence f which is of polynomial growth inside horizontal cones.*

We now state the two main results of the Depoissonisation path.

▶ **Theorem 7** ([15, 13])**.** *If the Poisson transform $P_f(z)$ satisfies the $\mathcal{JS}(\alpha, \beta)$ conditions, then the first terms of the Poisson-Charlier expansion provide the beginning of the asymptotic expansion of $f(n)$. More precisely, for any $k > 0$, one has:*

$$f(n) = \sum_{0 \leq j < 2k} P^{(j)}(n) \frac{\tau_j(n)}{j!} + O(n^{\alpha - k} \log^\beta n).$$

▶ **Theorem 8** ([16, 14]). *The two conditions are equivalent :*
(i) *the sequence f satisfies the Condition $(\mathcal{DP})$;*
(ii) *the Poisson transform P_f satisfies the Conditions $(\mathcal{JS})$.*

Bibliographic references. The Depoissonisation path is based on five main contributions, that are scattered in the literature. The path, together with its name, was systemized in 1998 by Jacquet and Szpankowski in [15]. They compare the asymptotics of the two sequences, the sequence $f(n)$ and the sequence $P_f(n)$. There were previous results of the same vein, notably a paper due to Hayman [11] in 1956, but they were not known by the AofA community. Jacquet and Szpankowski did not use the Poisson-Charlier expansion which was later introduced in 2010 into the AofA domain by Hwang, Fuchs and Zacharovas in [13]. Jacquet and Szpankowski also introduced conditions on the Poisson transform that we call (following the proposal of [13]) the Conditions $(\mathcal{JS})$. In [15], the authors prove that, under Conditions $(\mathcal{JS})$, it is possible to compare the two sequences $P_f(n)$ and $f(n)$. Later on, in 2010, using the Poisson Charlier expansion, the authors of [13] obtain a direct and natural proof of this comparison, with a more explicit remainder term. Finally, in two other papers, Jacquet and Szpankowski show that the two conditions – Condition $(\mathcal{DP})$ on the sequence f and Conditions $(\mathcal{JS})$ on P_f – are equivalent. The paper [16] deals with the necessary condition whereas the very recent paper [14] deals with the sufficient condition.

2.2 Application to the sorting toll in tries. First proof of Theorem 6.

This section ends with an example of the Depoissonisation path in the study of trie parameters. The Mellin transform of $P_f(z)$ satisfies,

$$P_f^*(s) = \sum_{k\geq 2} \frac{f(k)}{k!} \int_0^\infty e^{-z} z^k z^{s-1} dz = \sum_{k\geq 2} \frac{f(k)}{k!} \Gamma(k+s) = \sum_{k\geq 2} \frac{f(k)}{k} \frac{\Gamma(k+s)}{\Gamma(k)}.$$

The ratio of Gamma Functions can be estimated with the Stirling Formula,

$$\frac{\Gamma(k+s)}{\Gamma(k)} = \frac{(k+s)^{k+s}}{k^k} \frac{e^{-k-s}}{e^{-k}} \sqrt{\frac{k+s}{k}} \left[1 + O\left(\frac{1}{k}\right)\right] = k^s \left[1 + O\left(\frac{|s|}{k}\right)\right], \tag{9}$$

with a O-term uniform in k. Then, the Mellin transform of P_f satisfies, for $f(k) = k \log k$,

$$P_f^*(s) = \sum_{k\geq 2} k^s \log k \left[1 + O\left(\frac{|s|}{k}\right)\right] = -\zeta'(-s) + H_1(s), \quad H_1(s) \text{ analytic on } \Re s < 0. \tag{10}$$

Then $P_f^*(s)$ has a pole at $s = -1$ of order 2, and, together with the tameness of $\Lambda(s)$ at $s = 1$, this entails the following singular expressions for $P_f^*(s)$ and $P_r^*(s)$ at $s = -1$,

$$P_f^*(s) \asymp \frac{1}{(s+1)^2}, \qquad P_r^*(s) \asymp \frac{1}{h(\mathcal{S})} \frac{1}{(s+1)^3}.$$

The tamenesses of $P_f^*(s)$ and $\Lambda(-s)$ at $s = -1$ are enough to deduce, using standard Mellin inverse transform [7], the estimates, for $z \to \infty$,

$$P_f(z) = z \log z \,(1 + o(1)), \qquad P_r(z) = \frac{1}{2h(\mathcal{S})} z \log^2 z \,(1 + o(1)). \tag{11}$$

We now return to the Bernoulli model; we prove that $P_r(z)$ satisfies the Conditions $(\mathcal{JS})$. This will entail the estimate $r(n) \sim P_r(n)$ and end the proof. Assertion (I) is deduced from (11) in some cone $\mathcal{C}(\theta_1)$. For Assertion (O), we write $P_f(z)$ as

$$P_f(z) = z^2 e^{-z} G(z) \quad \text{with} \quad G(z) = \sum_{k=0} \frac{z^k}{k!} g(k) \quad \text{and} \quad g(k) := \frac{1}{k+1} \log(k+2). \tag{12}$$

As the sequence $g := T^2[f]$ satisfies Condition $(\mathcal{DP})$, Theorem 8 entails good behaviour for $G(z)$ outside horizontal cones. Namely, for some θ_2, and for all linear cones $\mathcal{C}(\theta)$ with $\theta < \theta_2$, there exist $\delta < 1$ and $A > 0$ such that the exponential generating function $G(z)$ of g satisfies

$$z \notin \mathcal{C}(\theta) \quad \Longrightarrow \quad (\forall w \in \Sigma^\star), \quad |G(p_w z)| \le A \exp(\delta |p_w z|) \tag{13}$$

We now consider, for $\gamma < 1$, a cone $\widehat{\mathcal{C}}(\gamma)$ defined in Section 1.4, with γ large enough to ensure the inclusions $\widehat{\mathcal{C}}(\gamma) \subset \mathcal{C}(\theta_1)$ (with θ_1 relative to Assertion (I) for $P_r(z)$) and $\widehat{\mathcal{C}}(\gamma) \subset \mathcal{C}(\theta_2)$ (with θ_2 relative to Eqn (13) for $G(z)$). With (8) and (12), and $\alpha := \max(\delta, \gamma)$, one has

$$\begin{aligned} \text{for } z \notin \widehat{\mathcal{C}}(\gamma), \quad |G(p_w z) \exp(z - p_w z)| &\le A \exp\left[\delta p_w |z| + \Re(z)(1 - p_w)\right] \\ &\le A \exp[|z|(\delta p_w + \gamma(1 - p_w))] \ \le A \exp(\alpha |z|). \end{aligned}$$

We then transfer the bounds on

$$P_r(z) e^z = e^z \sum_{w \in \Sigma^\star} P_f(z p_w) = z^2 \sum_{w \in \Sigma^\star} p_w^2\, G(p_w z)\, \exp(z - p_w z).$$

and obtain, with $B := A\Lambda(2)$, and for $|z|$ large enough

$$z \notin \widehat{\mathcal{C}}(\gamma) \Longrightarrow |P_r(z) e^z| \le B |z|^2 \exp(\alpha |z|) \ \le C \exp(\alpha' |z|)$$

with $\alpha' \in]\alpha, 1[$ and a given constant C. Finally, Assertion (O) of Condition $(\mathcal{JS})$ holds for $P_r(z)$ and this ends the proof.

3 The Rice path

In the Rice path, we deal with the Poisson sequence $\Pi[f]$. We assume the following condition, denoted as Condition $\mathcal{RM}$ [Rice-Mellin], to hold on the sequence $\Pi[f]$

Condition $(\mathcal{RM})$**:** *There is an analytic lifting $\psi(s)$ for the sequence $\Pi[f]$ which is tame.*

Then the binomial recurrence (2) is transfered into a relation which expresses the term $f(n)$ as an integral along a vertical line which involves the analytic lifting $\psi(s)$. With tameness of ψ, we obtain the asymptotics of the sequence f.

3.1 The three steps of the Rice path

The Rice path performs three steps. It deals with a sequence f which satisfies the `(VD)` conditions, but we describe it in the stronger case when f is reduced. The complete proofs are in the Annex.

Step 1. It proves the existence of an analytical lifting ψ of the sequence $\Pi[f]$, on a halfplane $\Re s > c$ (for some c). It uses the (direct) Mellin transform and the Newton interpolation, without any other condition on the sequence f.

▶ **Proposition 9** (Nördlund-Rice)**.** *The sequence $\Pi[f]$ associated with a reduced sequence f of degree $c < -1$ admits as an analytic lifting on $\Re s > c$ a function ψ, which is also an analytic extension of $P_f^*(-s)/\Gamma(-s)$ there.*

Step 2. If moreover ψ is of *polynomial growth* "on the right", the binomial relation (2) is transfered into a Rice integral expression

▶ **Proposition 10.** *Assume that the analytic lifting ψ of $\Pi[f]$ is of polynomial growth on the halfplane $\Re s > c$, with $c < -1$. Then, for any $a \in]c, 0[$ and $n \geq n_0$, the sequence $f(n)$ admits an integral representation of the form*

$$f(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} p(k) = \frac{1}{2i\pi} \int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s) \, ds \,,$$

where the Rice kernel $\quad L_n(s) := \dfrac{(-1)^{n+1}\, n!}{s(s-1)(s-2)\ldots(s-n)} = \dfrac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)}$

involves the Beta Function B with the equality $L_n(s) = B(n+1, -s)$.

This integral representation is valid for any abscissa a which belongs to the interval $]c, 0[$. We now shift the vertical line $\Re s = a$ to the left, and thus use *tameness conditions* on ψ at $s = c$, as defined in Section 1.4.

Step 3. If moreover ψ is *tame* "on the left", the integral is *shifted* to the left; this provides the asymptotics of of the sequence f.

▶ **Proposition 11.** *Consider a reduced sequence $f : n \mapsto f(n)$ with $\texttt{deg}(f) = c < -1$. If the lifting ψ of $\Pi[f]$ is tame at $s = c$ with a region $\mathcal{R}$ of tameness and a left frontier $\mathcal{F}$, then*

$$f(n) = -\left[\sum_{k | s_k \in \mathcal{R}} \text{Res}\left[L_n(s) \cdot \psi(s); s = s_k\right] + \frac{1}{2i\pi} \int_{\mathcal{F}} L_n(s) \cdot \psi(s) \, ds \right],$$

where the sum is over the poles s_k of ψ inside $\mathcal{R}$.

3.2 The main question about the Rice method: Tameness of ψ

The main results are due to Nörlund [18, 17], then to Rice who popularized them. Later on, with the paper [10], Flajolet and Sedgewick brought this methodology into the AofA domain. The Rice-Mellin method is also well described in [6]. There exist many analyses of various data structures or algorithms that are based on the application of the method: tries ([9, 8, 3, 1]), digital trees ([9, 12]), or fine complexity analyses of sorting or searching algorithms on sources ([4, 2]).

The situation for applying the Rice method is not the same as in Section 2: previously, with Condition ($\mathcal{DP}$), we know exactly when the Depoissonisation method may be applied. This is not the case for the Rice method. Even though the literature well explains how to use this method in various cases of interest, the following question is never asked: *What are sufficient conditions on the sequence f that would entail tameness of ψ?*

As $\psi(s)$ is closely related to the Mellin transform $P_f^*(-s)$, meromorphy is often easy to prove, and the poles often easy to find. In many natural contexts, the polynomial growth and the tameness of the Mellin transform $P_f^*(s)$ generally hold, and are often used in the Depoissonisation approach [see Section 2.2]. But the main difference between the Rice method and the Depoissonisation method is the division by $\Gamma(s)$.

Sometimes, and this is often the case in classical tries problems, the factor $\Gamma(s)$ already appears in $P_f^*(s)$, and $\psi(s)$ has an explicit form, from which its polynomial growth may be easily proven. For instance, for the toll $f = f_1$ associated to the path length, then

$P^*_f(s) = \Gamma(s+1)$ and $\psi(-s)$ is explicit, and equal to s. This is also the case for polynomial tolls f of the form $f = T^{-m}[f_1]$ with $m \geq 1$.

But what about other sequences, for instance the sorting toll $f(k) = k \log k$, and more generally, the basic sequence $f(k) = k^d \log^b k$ (with $d \in \mathbb{R}$ and an integer $b \geq 1$)? In this case, the following expansion holds for $P^*_f(-s)$, that involves the b-th derivative of the Riemann ζ function and generalizes (10),

$$P^*_f(-s) = (-1)^b \zeta^{(b)}(s-(d-1)) + H_1(s)\,, \tag{14}$$

where $H_1(s)$ is analytic on $\Re s > d-1$. Then principles of Depoissonisation apply in this case, due to good properties of the Riemann function. Now, in the Rice method, the function ψ satisfies $\psi(s) = P^*_f(-s)/\Gamma(-s)$, and the function $1/\Gamma(-s)$, even though it is analytic on the half-plane $\Re(s) \geq 0$, is of exponential growth along vertical lines. The Stirling formula indeed entails the estimate

$$\frac{1}{\Gamma(x+iy)} \sim \frac{1}{\sqrt{2\pi}} e^{\pi|y|/2}\, |y|^{1/2-x}, \qquad \text{as } |y| \to \infty\,.$$

It is thus not clear whether $\psi(s)$ attached to the sorting toll is tame at $s=1$. Then, the Rice method seems to have a more restrictive use than the Depoissonisation method. As we wish to compare the power of the two paths [Depoissonisation path and Rice path], we ask the two (complementary) questions: *Is the Rice path only useful for very specific tolls, where the Mellin transform $P^*_f(s)$ of the Poisson transform $P_f(s)$ factorizes with the factor $\Gamma(s)$, or is it useful for more general tolls?*

This leads us to study sufficient conditions under which the analytic lifting ψ may be proven to be tame. We now propose to use the (inverse) Laplace transform. With this tool, we prove the tameness of ψ for basic sequences (see Theorem 13).

4 The Rice–Laplace approach.

As in the previous Section, we deal with the Poisson sequence $\Pi[F]$. Our main result proves the tameness of the analytic continuation $\Psi(s)$ of $\Pi[F]$, when F is a basic sequence.

▶ **Definition 12.** Consider a pair (d, b) with a real d and an *integer* $b \geq 0$.

- **(i)** A sequence F is *basic* with pair (d, b) if it writes as $F_{b,d}(k) = k^d \log^b k$ for any $k \geq 2$.
- **(ii)** A sequence F is *extended basic* with pair (d, b) if it has an analytic extension Φ on some halplane $\Re s > a$, of the form $\Phi(z) = F_{d,b}(z)\, W(1/z)$, with W analytic at 0, and $W(0) = 1$.

▶ Remark. In the proof, an integral exponent $b \geq 0$ is needed to relate $F_{d,b}$ to a b-th derivative.

Our main result is as follows:

▶ **Theorem 13.** *Consider a basic extended sequence F with pair (d, b). Then, for some $\sigma_0 > 0$, the analytic continuation $\Psi(s)$ of the $\Pi[F]$ sequence is of polynomial growth on any halfplane $\Re s \geq a > d$. Moreover, it writes in terms of the integer $\ell := \sigma(F)$ defined in* (4),

$$\Psi(s) = \left[s(s-1)\dots(s-\ell+1) \sum_{m=0}^{b} a_m\, \Gamma^{(m)}_\ell(s-d) \right] + B(s) \tag{15}$$

on the halfplane $\Re s > d - \sigma_0$*, for* $\sigma_0 \in]0,1[$*. Here,* $B(s)$ *is of polynomial growth, and* $\Gamma_\ell^{(m)}$ *is the* m*-th derivative of the twisted* Γ *function that is defined for* $\Re s > 0$*, integers* $m \geq 0$ *as*

$$\Gamma_\ell^{(m)}(s) := \int_0^\infty e^{-\ell u} u^{s-1} \log^m u \, du \,. \tag{16}$$

The coefficients a_m *involve the derivatives of order* $k \leq b$ *of* $s \mapsto 1/\Gamma(s)$ *at* $s = \ell - d$*.*

Remarks.

(a) With Lemma 3, the twisted function Γ_ℓ and its derivatives are of exponential decrease along the vertical lines. This entails the tameness of Ψ at $s = d$.

(b) We already know the singular part of Ψ at $s = d$ which is given by the expansion (14), and the singular expansion given in (15) is just an alternative and complicate expression. What is new is the tameness, not the singular expansion.

4.1 Plan of the proof

We first recall the principles of Section 1.5 : with a initial sequence F, we associate its canonical sequence $f := \rho(F)$, and now deal with this new sequence f; it is easy to return (later) to the initial sequence F with Lemma 5. If the initial F admits an analytic lifting of polynomial growth on $\Re s > 0$, then the sequence $f = \rho(F)$ is reduced and admits an analytic lifting φ on $\Re s > -1$ that satisfies $\varphi(s) = O(|s+1|^c)$ there, with $c < -1$.

The first step performed in Section 4.2 deals with any reduced sequence f which admits an analytic lifting φ on $\Re s > -1$ that satisfies $\varphi(s) = O(|s+1|^c)$ there, with $c < -1$. With a strong use of the involutive character of Π, we first exhibit a new expression of the analytical extension ψ of $\Pi[f]$ which deals with the inverse Laplace transform $\widehat{\varphi}$ of the extension φ of the sequence f. The proof is then applied to the canonical sequence $\rho[F]$ of the initial sequence F.

Then, the sequel of the present section focuses on (extended) basic sequences $F_{d,b}$. Here, in this Section, we only deal with exact basic sequences. The extension to *extended basic* sequences will be done in the Annex. We first obtain in Section 4.3 a precise expression of the inverse Laplace transform $\widehat{\varphi}$ of extension φ of the canonical sequence $f_{d,b} := \rho(F_{d,b})$, that is transfered into a precise estimate of $\Pi[f_{d,b}]$. This leads to the proof of Theorem 13.

4.2 A new general expression for ψ with the inverse Laplace transform

This section is of independent interest and provides a new expression of the extension of the sequence $\Pi[f]$ in the case when f is reduced.

▶ **Proposition 14.** *Consider a sequence* f *which admits an analytic lifting* φ *on* $\Re s > -1$*, with the estimate* $\varphi(s) = O(|s+1|^c)$ *with* $c < -1$*. Then:*

(i) *The function* φ *admits an inverse Laplace transform* $\widehat{\varphi}$ *whose restriction to the real line* $[0, +\infty[$ *is written as the Bromwich integral for* $a \in]-1, 0[$*,*

$$\widehat{\varphi}(u) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) \, e^{su} ds \,, \qquad \textit{and satisfies} \quad |\widehat{\varphi}(u)| \leq K e^{au} \,.$$

(ii) *There is an analytical lifting* ψ *of the sequence* $\Pi[f]$ *that admits an integral form*

$$\psi(s) = \mathcal{I}_s[\widehat{\varphi}] \qquad \textit{with} \quad \mathcal{I}_s[h] := \int_0^\infty h(u)(1 - e^{-u})^s du \qquad \textit{for } \Re s > -1 \,. \tag{17}$$

Sketch of the proof. The complete proof is in the Annex.

(i) In a general context, where the analytic lifting $\varphi(s)$ is only defined on $\Re s > 0$, the Bromwich integral is written as an integral on a vertical line $\Re s = a$ with $a > 0$. Here, the hypotheses on φ are stronger and the Bromwich integral may be shifted to the left with $a \in]-1, 0[$. Moreover, the Bromwich integral is absolutely convergent, and the exponential bound on $\widehat{\varphi}(u)$ holds.

(ii) As φ is polynomial growth, we use the involutive character of Π and apply Proposition 10 to the pair $(p := \Pi[f], f = \Pi^2[f])$. It transfers the binomial expression of $\Pi[f]$ in terms of $\Pi^2[f] = f$ into a Rice integral, with $a \in]-1, 0[$,

$$p(n) = \frac{1}{2i\pi}\int_{\Re s = a} \varphi(s) L_n(s) ds, \qquad L_n(s) = \frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)} = B(n+1, -s)\,.$$

We now use the integral expression of the Beta function, and "exchange" the two integrals. ◀

The integral representation (17) leads us to introduce the two functions, defined on $[0, +\infty]$,

$$N_s(u) := \left(\frac{1-e^{-u}}{u}\right)^s, \; M_s(u) := \left[\left(\frac{1-e^{-u}}{u}\right)^s - 1\right], \tag{18}$$

that satisfy the two estimates, with $\sigma := \Re s$,

$$N_s(u) = \Theta(1), \quad (u \to 0), \qquad N_s(u) = O(u^{-\sigma}) \quad (u \to \infty)\,,$$

$$M_s(u) = \Theta(u) \quad (u \to 0), \qquad M_s(u) = O(u^{-\sigma}) \quad (u \to \infty, \sigma > 0)\,.$$

Then, for "good" functions h, the integral $\mathcal{I}_s[h]$ may be compared to the Mellin transform $h^\star(s+1)$. We now apply this idea to the particular cases where the behaviour of $h = \widehat{\varphi}$ is well-known. Then, there are two steps which deal with a reduced sequence f, and aim at studying the tameness of the analytical extension ψ of $\Pi[f]$:

(a) transfer properties of φ into properties of its inverse Laplace transform $\widehat{\varphi}$;

(b) with properties of $\widehat{\varphi}$, study the tameness of ψ, via the representation (17).

We now perform these two steps. The (inverse) Laplace transform is not well studied, and we do not know a general transfer result of type (a). This is why we only perform the two steps for canonical sequences related to basic sequences. The proofs of the following section are in the Annex.

4.3 Dealing with basic sequences

Step (a). We obtain first an expression for φ, then an expression for $\widehat{\varphi}$.

▶ **Proposition 15.**

(i) *The sequence $f_{b,d}$ is extended in a function φ defined on $\Re s > -1$*

$$\varphi(s) = (s+\ell)^{d-\ell} \log^b(s+\ell)\, U\left(\frac{1}{s+\ell}\right) \;; \quad \ell := \sigma(d)$$

Here U satisfies $U(u) = 1$ for $d < 0$. For $d \ge 0$, it is defined as

$$U(u) = (1-u)^{-1}(1-2u)^{-1}\dots(1-(\ell-1)u)^{-1} \quad (\textit{with } \ell = 2 + \lfloor d \rfloor)$$

For $d \ge 0$, the coefficient $a_j := [u^j]U(u)$ satisfies $a_j = \Theta_d(\ell-1)^j$.

(ii) *The inverse Laplace transform $\widehat{\varphi}(u)$ is a linear combination of functions, for $m \in [0..b]$,*

$$e^{-\ell u}\, u^{-c-1} \log^m u\, \left[1 + V^{\langle m\rangle}(u)\right], \qquad \textit{with} \quad |V^{\langle m\rangle}(u)| \le A_{(d,b)}\, u\, e^{(\ell-1)u}\,.$$

Step (b). The previous expression of $\widehat{\varphi}$ together with the representation (17) entail a decomposition for ψ. Using the estimates of functions defined in (18), each term is compared to the twisted version of the Γ function and its m-th derivative, defined in (16). This provides the estimate for the initial function $\Psi := \Pi[f]$.

▶ **Proposition 16.**

(i) *The extension ψ of the sequence $\Pi[f_{b,d}]$ is a linear combination of functions, for $m \in [0..b]$, each term being the sum of a main term $A^{\langle m \rangle}(s)$ and a remainder term $O(B^{\langle m \rangle}(s))$, with $c := d - \sigma(d) < -1$ and*

$$A^{\langle m \rangle}(s) := \mathcal{I}_s\left[e^{-\ell u}\, u^{-c-1}\, \log^m u\right], \ B^{\langle m \rangle}(s) := \mathcal{I}_s\left[e^{-u}\, u^{-c}\, \log^m u\right] .$$

(ii) *For $\Re s \geq 0$, the two functions $A^{\langle m \rangle}(s)$ and $B^{\langle m \rangle}(s)$ are bounded on the halfplane $\Re s \geq 0$. For any integer $m \geq 0$ and any integer $\ell \geq 1$, the two functions*

$$A^{\langle m \rangle}(s) - \Gamma_\ell^{(m)}(s-c), \quad B^{\langle m \rangle}(s)$$

are analytic and of bounded growth on the vertical strip $\Re s > c - \sigma_0$, with $\sigma_0 \in]0, 1[$.

4.4 A second proof for Theorem 6.

Within the framework of Section 1.7, we deal with the sorting toll $F_{1,1}$. The singular part of the extension $\Psi := \Pi[F_{1,1}]$ at $s = 1$ is obtained in (14). The tameness of Ψ at $s = 1$ is proven in Theorem 13. Together with the tameness of the Dirichlet series Λ at $s = 1$, this entails the tameness of $\Pi[r]$ and gives a three-lines proof of Theorem 6. We prefer this proof!

5 Final comparison between the two paths.

The Annex describes a formal comparison between the two paths. From analytical properties, the Rice-Laplace path remains of more restrictive use than the Depoissonisation path:

(a) We need the analytic extension φ of f to hold on a halfplane, whereas the Depoissonisation path needs it only on a horizontal cone.

(b) The analytic extension φ of f involves a precise expansion in terms of an analytic series W, whereas the Depoissonisation path only needs a rough asymptotic estimate of φ.

(c) The exponent of the log term must be an *integer* b, whereas the Depoissonisation path deals with any real exponent. The need of an integer exponent b is related to the interpretation in terms of b-derivatives, and this is a restriction which is also inherent in the method used by Flajolet in [5] in a similar context.

These are strong restrictions... However, most of the Depoissonisation analyses (at least for mean values) deal with extended basic sequences, where the Rice-Laplace path may be used. We let the final conclusion to the reader !

References

1 Eda Cesaratto and Brigitte Vallée. Gaussian distribution of trie depth for strongly tame sources. *Combinatorics, Probability and Computing*, 24, issue 01:54–103, 2014.

2 Julien Clément, Jim Fill, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of the Quickselect algorithm. *Theory of Computing Systems*, 58(4):528–578, 2016.

3 Julien Clément, Philippe Flajolet, and Brigitte Vallée. Dynamical sources in information theory: a general analysis of trie structures. *Algorithmica*, 29, 1-2:307–369, 2001.

4 Julien Clément, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of some popular sorting algorithms. *Combinatorics, Probability and Computing*, 24 issue 01:104–144, 2014.
5 Philippe Flajolet. Singularity analysis and asymptotics of Bernoulli sums. *Theoret. Comput. Sci.*, 215 1-2:371–381, 1999.
6 Philippe Flajolet. A journey to the outskirts of Quickbits. *Personal Communication*, 2010.
7 Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144 1-2:3–58, 1995.
8 Philippe Flajolet, Mireille Régnier, and Wojtek Szpankowski. Some uses of the Mellin integral transform in the analysis of algorithms. *In Combinatorial Algorithms on Words*, 12:241–254, 1995.
9 Philippe Flajolet and Robert Sedgewick. Digital search trees revisited. *SIAM J. Comput.*, 15 (3):48–67, 1986.
10 Philippe Flajolet and Robert Sedgewick. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science*, 144 (1-2):101–124, 1995.
11 Walter Kurt Hayman. A generalisation of Stirling's formula. *J. Reine Angew. Math.*, 196:67–95, 1956.
12 Kanal Hun and Brigitte Vallée. Typical depth of a digital search tree built on a general source. In *Proceedings of ANALCO*, 2014.
13 Hsien-Kuei Hwang, Michael Fuchs, and Vytas Zacharovas. Asymptotic variance of random symmetric digital search trees. *Discrete Mathematics and Theoretical Computer Science*, 12(2):103–166, 2010.
14 Philippe Jacquet. Trie structure for graphs sequences. In *Proceedings of AofA 2014*, volume 950 of *DMTCS Proceedings*, pages 181–192, 2014.
15 Philippe Jacquet and Wojtek Szpankowski. An analytic approach to the asymptotic variance of trie statistics and related structures. *Theoretical Computer Science*, 201 1-2:1–62, 1998.
16 Philippe Jacquet and Wojtek Szpankowski. Entropy computations via Analytic depoissonization. *IEEE Transactions on Information Theory*, 45 (4):1072–1081, 1999.
17 Niels Erik Nörlund. *Vorlesungen über Differenzenrechnung*. Chelsea Publishing Company, New-York, 1954.
18 Nils Erik Nörlund. Leçons sur les équations linéaires aux différences finies. In *Collection de monographies sur la théorie des fonctions*. Gauthier-Villars, Paris, 1929.
19 Brigitte Vallée. Dynamical sources in information theory: Fundamental intervals and word prefixes. *Algorithmica*, 29 (1/2):262–306, 2001.

A Annex

A.1 More on tameness

▶ **Definition 17** (Tameness). A function ϖ analytic and of polynomial growth on $\Re s > c$ is tame at $s = c$ if one of the three following properties holds:

(a) [S-shape] (shorthand for Strip shape) there exists a vertical strip $\Re(s) > c - \delta$ for some $\delta > 0$ where $\varpi(s)$ is meromorphic, has a sole pole (of order $b + 1 \geq 1$) at $s = c$ and is of polynomial growth as $|\Im s| \to +\infty$.

(b) [H-shape] (shorthand for Hyperbolic shape) there exists an hyperbolic region $\mathcal{R}$, defined as, for some $A, B, \rho > 0$

$$\mathcal{R} := \{s = \sigma + it; \quad |t| \geq B, \quad \sigma > c - \frac{A}{|t|^\rho}\} \bigcup \{s = \sigma + it; \quad \sigma > c - \frac{A}{B^\rho}, |t| \leq B\},$$

where $\varpi(s)$ is meromorphic, with a sole pole (of order $b+1$) at $s = c$ and is of polynomial growth in $\mathcal{R}$ as $|\Im s| \to +\infty$.

(c) [P-shape] (shorthand for Periodic shape) there exists a vertical strip $\Re(s) > c - \delta$ for some $\delta > 0$ where $\varpi(s)$ is meromorphic, has only a pole (of order $b + 1 \geq 1$) at $s = c$ and a family (s_k) (for $k \in \mathbb{Z} \setminus \{0\}$) of simple poles at points $s_k = c + 2ki\pi t$ with $t \neq 0$, and is of polynomial growth as $|\Im s| \to +\infty$[4].

A.2 Proofs of the Rice path

Proof of Proposition 9. In the strip $\mathcal{S}(0, -c)$, the Mellin transform $P_f^*(s)$ of $P_f(z)$ exists and satisfies

$$\frac{P_f^*(s)}{\Gamma(s)} = \frac{1}{\Gamma(s)} \sum_{k\geq 0} \frac{f(k)}{k!} \int_0^\infty e^{-z} z^k z^{s-1} dz = \sum_{k\geq 0} \frac{f(k)}{k!} \frac{\Gamma(k+s)}{\Gamma(s)}$$

where the exchange of integration and summation is justified by the estimates given in (9). On the strip $\mathcal{S}(c, 0)$, the series is a Newton interpolation series,

$$\psi(s) := \frac{P_f^*(-s)}{\Gamma(-s)} = \sum_{k\geq 0} (-1)^k \frac{f(k)}{k!} s(s-1)\dots(s-k+1)\,. \tag{19}$$

Such series converge in right halfplanes and thus the previous series converges on $\Re s > c$. Moreover, Relation (19), together the binomial relation (2), entails the equality

$$\psi(n) = \sum_{k=0}^{n} (-1)^k \frac{f(k)}{k!} n(n-1)\dots(n-k+1) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(k) = \Pi[f](n)\,.$$

This proves that ψ provides an analytic lifting of the sequence $\Pi[f]$ on $\Re s > c$ which is also an analytic extension of $P_f^*(-s)/\Gamma(-s)$. ◀

Proof of Proposition 10. (Sketch) Use the Residue Theorem and the polynomial growth of $\psi(s)$ "on the right". First, we consider the rectangle $\mathcal{A}_M$ delimited by the contour τ_M defined by the two vertical lines $\Re s = a$ and $\Re s = n + M$ and the two horizontal lines $\Im s = \pm M$. If the contour τ_M is taken counterclockwise, then the Residue Theorem applies,

$$\frac{1}{2i\pi} \int_{\tau_M} L_n(s) \cdot \psi(s)\, ds = \sum_{k=0}^{n} \mathrm{Res}[L_n(s) \cdot \psi(s); s = k] = -\sum_{k=0}^{n} (-1)^k \binom{n}{k} \Pi[f](k) = -f(n).$$

Next, the integral on the curve τ_M is the sum of four integrals. Let now M tend to ∞. The integrals on the right, top and bottom lines tend to 0, due to the polynomial growth of the function $\psi(s)$. The integral on the left becomes

$$-\int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s)\, ds,$$

and we have proven the result. For details on the proof, we may refer to papers [18, 17, 10]. ◀

---

[4] More precisely, this means that $\varpi(s)$ is of polynomial growth on a family of horizontal lines $t = t_k$ with $t_k \to \infty$, and on vertical lines $\Re(s) = \sigma_0 - \delta'$ with some $\delta' < \delta$.

Proof of Proposition 11. (Sketch) The proof is similar to the previous proof. With the tameness of $\psi(s)$ at $s = c$ with a tameness domain $\mathcal{R}$, we now deal with the residues of ψ; we consider the domains

$$\widehat{\mathcal{R}} := \mathcal{R} \cap \{\Re s < a\} \text{ and } \mathcal{R}_M := \widehat{\mathcal{R}} \cap \{|\Im s| \leq M\},$$

and denote $\mathcal{L}_M$ the curve (taken counterclockwise) which borders the region $\mathcal{R}_M$. As $\psi(s)$ is meromorphic in $\mathcal{R}_M$ and $L_n(s)$ analytic there, we apply the Residue Theorem to the function $L_n(s) \cdot \psi(s)$ inside $\mathcal{R}_M$, and obtain

$$\frac{1}{2i\pi} \int_{\mathcal{L}_M} L_n(s) \cdot \psi(s)\, ds = \sum_{s_k \in \mathcal{R}_M} \text{Res}\,[L_n(s) \cdot \psi(s); s = s_k]$$

where the sum is taken over the poles s_k of $\psi(s)$ inside $\mathcal{R}$. Now, when $M \to \infty$, the integrals on the two horizontal segments tend to 0, since $\psi(s)$ is of polynomial growth, and

$$\lim_{M\to\infty} \int_{\mathcal{R}_M} L_n(s) \cdot \psi(s)\, ds = \int_{a-i\infty}^{a+i\infty} L_n(s) \cdot \psi(s)\, ds - \int_{\mathcal{F}} L_n(s) \cdot \psi(s) ds$$
$$= 2i\pi \sum_{s_k \in \mathcal{R}} \text{Res}\,[L_n(s) \cdot \psi(s); s = s_k]\ ,$$

where the sum is taken over the poles s_k of $\psi(s)$ inside the domain $\mathcal{R}$. ◀

A.3 Proofs of the Rice-Laplace path

Proof of Proposition 14.

(i) In a general context, where the analytic lifting $\varphi(s)$ is only defined on $\Re s > 0$, the Bromwich integral is written as

$$\widehat{\varphi}(u) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s)\, e^{su} ds, \qquad (\text{with } a > 0)\,.$$

Here, the hypotheses on φ are stronger: we can shift the integral on the left and choose $a \in]-1, 0[$. Moreover, the Bromwich integral is ansolutely convergent, and the exponential bound on $\widehat{\varphi}(u)$ holds.

(ii) We use the involutive character of Π and apply Proposition 10 to the pair ($p := \Pi[f], f = \Pi^2[f]$). In the classical Rice path, it is applied to the pair $(f, \Pi[f])$, when $\Pi[f]$ is of polynomial growth, and it transfers the binomial expression of f in terms of $\Pi[f]$ into an integral expression. Here, due to the polynomial growth of $f = \Pi^2[f]$ on $\Re s > -1$, it transfers the binomial expression of $\Pi[f]$ in terms of $\Pi^2[f] = f$ into a Rice integral, with $a \in]-1, 0[$,

$$p(n) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) L_n(s) ds, \qquad L_n(s) = \frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)}\,.$$

We now deal with the Beta function

$$B(t+1, -s) = \frac{\Gamma(t+1)\Gamma(-s)}{\Gamma(t+1-s)}\,,$$

that is well defined for $\Re t > -1$ and $\Re s < 0$, and admits an integral expression

$$B(t+1, -s) = \int_0^\infty e^{su}(1 - e^{-u})^t du, \qquad (\text{for } \Re t > -1,\ \Re s < 0)\,.$$

Together with the equality $L_n(s) = B(n+1, -s)$, this entails an analytic extension ψ of the sequence $\Pi[f]$ on the halfplane $\Re t > -1$,

$$\psi(t) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) B(t+1, -s) ds, \quad (b < 0)$$

with an integral expression,

$$\psi(t) = \frac{1}{2i\pi} \int_{\Re s = a} \varphi(s) \left[\int_0^\infty e^{su} (1 - e^{-u})^t du \right] ds\,.$$

With properties of φ, it is possible to exchange the integrals: then, the equality holds

$$\psi(t) = \int_0^\infty (1 - e^{-u})^t \left[\frac{1}{2i\pi} \int_{\Re s = a} \varphi(s)\, e^{su} ds \right] du,$$

and the second integral is the inverse Laplace transform $\widehat{\varphi}$ of φ. This ends the proof. ◀

Proof of Proposition 15. For (i), letting $\ell := \sigma(d)$ with σ defined in (4), the canonical sequence f associated with F can be extended to a function φ defined on $]-1, +\infty[$ as

$$\varphi(x) = \log^b(x+\ell) \frac{(x+\ell)^d}{(x+1)(x+2)\dots(x+\ell)} = \log^b(x+\ell)\,(x+\ell)^{d-\ell}\, U\left(\frac{1}{x+\ell}\right),$$

and involves a function U defined as $U(u) = 1$ for $d < 0$ and, for $d \geq 0$ as

$$U(u) = (1-u)^{-1}(1-2u)^{-1} \dots (1-(\ell-1)u)^{-1} \quad (\text{with } \ell = 2 + \lfloor d \rfloor)\,. \tag{20}$$

Then, for $d \geq 0$, the coefficient $a_j := [u^i]U(u)$ satisfies $a_j = \Theta(\ell-1)^j$.

For (ii), there are three main steps, according to the type of the basic sequence.

Step 1. We begin with the particular case when $\varphi(s)$ is of the form $\varphi(s) = (s+\ell)^c$ (with $c < -1$). Its inverse Laplace transform $\widehat{\varphi}$ is then

$$\widehat{\varphi}(u) = \frac{1}{\Gamma(-c)} e^{-\ell u}\, u^{-c-1}\,.$$

Step 2. We now consider a function (without logarithmic factor) of the form

$$\varphi(s) = \varphi_c(s) = (s+\ell)^c U\left(\frac{1}{s+\ell}\right) = \sum_{j \geq 0} a_j\, (s+\ell)^{c-j}\,. \tag{21}$$

Then φ is a linear combination of functions of Step 1 and the inverse Laplace transform $\widehat{\varphi}$ of φ is written as

$$\widehat{\varphi}_c(u) = e^{-\ell u}\, \frac{u^{-c-1}}{\Gamma(-c)} [1 + V_c(u)], \quad \text{with } V_c(u) := \sum_{j \geq 1} a_j u^j G_j(c)\,, \tag{22}$$

where the function G_j is the rational fraction which associates with c the ratio

$$G_j(c) := \frac{\Gamma(-c)}{\Gamma(j-c)} = \frac{1}{-c(1-c)\dots(j-1-c)}\,. \tag{23}$$

As $c < -1$, the inequality $G_j(c) \leq (1/j!)$ holds and this entails the inequality $|V_c(u)| \leq A\, u\, e^{(\ell-1)u}$, where the constant A only depends on d.

Step 3. We add finally a logarithmic factor and consider a function of the form

$$\varphi(s) = (s+\ell)^c \log^b(s+\ell)\, U\left(\frac{1}{s+\ell}\right) \tag{24}$$

which is written as a b-th derivative. Indeed, the equality holds

$$U\left(\frac{1}{s+\ell}\right)(s+\ell)^c \log^b(s+\ell) = \frac{\partial^b}{\partial t^b}\left[(s+\ell)^{c+t} U\left(\frac{1}{s+\ell}\right)\right]\Bigg|_{t=0},$$

and we can take the derivative "under the Laplace integral": we then deduce that the inverse Laplace transform $\widehat{\varphi}$ of the function φ defined in (24) is equal to

$$\frac{\partial^b}{\partial t^b}\widehat{\varphi}_{c+t}(u)\Bigg|_{t=0} = e^{-\ell u}\frac{\partial^b}{\partial c^b}\left[\frac{u^{-c-1}}{\Gamma(-c)}(1+V_c(u))\right].$$

The coefficient of u^j in the k-th derivative of $c \mapsto V_c(u)$ involves the k-th derivative of the function $c \mapsto G_j(c)$, defined in (23) which satisfies the inequality

$$|G_j^{(k)}(c)| \le A_k \log^k(j+c)\, G_j(c) \quad \text{for some constant } A_k.$$

Then, the inequality holds,

$$\left|\frac{\partial^k}{\partial c^k} V_c(u)\right| \le A_{(d,b)}\, u\, e^{(\ell-1)u},$$

and involves a constant $A_{(d,b)}$ which depends on the pair (d,b). On the other hand, the following m-th derivative is a linear combination of the form

$$\frac{\partial^m}{\partial c^m}\left[\frac{u^{-c-1}}{\Gamma(-c)}\right] = u^{-c-1}\left[(-1)^m \sum_{a=0}^{m}\binom{m}{a}(\log^a u)\, H^{(m-a)}(c)\right],$$

where H is the function defined as $H(c) = 1/\Gamma(-c)$. This ends the proof. ◀

Proof of Proposition 16.

(a) is clear : For $\Re s \ge 0$, the result follows from the inequalities $(1-e^{-u})^\sigma \le 1$, $c < -1$, together with the integrability of the function $u \mapsto e^{-\ell u} u^{-c-1} \log^m u$ on the interval $[0, +\infty]$.

(b) The difference $A^{\langle m \rangle}(s) - \Gamma_\ell^{(m)}(s-c)$ is expressed with M_s, whereas $B^{\langle m \rangle}(s)$ is expressed with N_s, both defined in (18). Together with their estimates, this leads to the following bounds, for any $\rho > 0$,

$$A^{\langle m \rangle}(s) - \Gamma_\ell^{(m)}(s-c) = O_\rho\Big(\Gamma_\ell(\sigma - c + 1 - \rho)\Big), \quad B^{\langle m \rangle}(s) = O_\rho\Big(\Gamma(\sigma - c + 1 - \rho)\Big)$$

and also to the analyticity of the functions of interest on the vertical strip $\Re s > c - \sigma_0$, with $\sigma_0 \in]0,1[$. ◀

Extension to extended basic sequences. It is easy to extend the proof of Theorem 13 to this more general case: We denote by r the convergence radius of W, and we thus choose a shift T^ℓ with an integer which now satisfies

$$\ell \ge \max\Big[2 + \lfloor d \rfloor, a+1, (1/r)+1\Big],$$

and deal with the sequence $f := T^\ell[F]$. We replace the previous series U defined in Proposition 15 by the series $U \cdot W$ which has now a convergence radius $\tilde{r} := \min(r, 1/(\ell-1))$ for which the bound $1/\tilde{r} < \ell$ holds. We choose $\widehat{r} \in]1/\tilde{r}, \ell[$, and the new series V_c defined in (22) satisfies $|V_c(u)| \le Au\, e^{\widehat{r}u}$ and indeed gives rise to a remainder term. ◀

A.4 Description of a formal comparison between the two paths

As it is observed in the paper [13], there are formal manipulations which allow us to compare the two paths.

In the Depoissonisation path, the asymptotics of $f(n)$ is manipulated in two steps: first use the Cauchy integral formula

$$f(n) = \frac{n!}{2i\pi} \int_{|z|=r} P_f(z)\, e^z \frac{1}{z^{n+1}} dz\,. \tag{25}$$

then derive asymptotics of $P_f(z)$ for large $|z|$ by the inverse Mellin integral

$$P_f(z) = \frac{1}{2i\pi} \int_{\uparrow} P_f^*(s) z^{-s} ds = \frac{1}{2i\pi} \int_{\uparrow} P_f^*(-s) z^s ds\,, \tag{26}$$

where the integration path is some vertical line. This two-stage Mellin-Cauchy formula is the beginning point of the Depoissonization path.

We now compare the formula obtained by this two stage approach with the N'ordlund-Rice formula. First remark that, as the function $P_f(z)e^z$ is entire, we can replace the contour $\{|z| = r\}$ in (25) by a Hankel contour $\mathcal{H}$ starting at $-\infty$ in the upper halfplane, winding clockwise around the origin and proceeding towards $-\infty$ in the lower halfplane. Then (25) becomes

$$f(n) = \frac{n!}{2i\pi} \int_{\mathcal{H}} P_f(z)\, e^z \frac{1}{z^{n+1}} dz \tag{27}$$

Now, if we *formally substitute* (26) into (27), *interchange* the order of integration and use the equality

$$\frac{1}{\Gamma(n+1-s)} = \frac{1}{2i\pi} \int_{\mathcal{H}} e^z \frac{z^s}{z^{n+1}} dz\,,$$

we obtain the representation

$$f(n) = \frac{n!}{2i\pi} \int_{\uparrow} P_f^*(-s) \frac{1}{\Gamma(n+1-s)} ds\,, \tag{28}$$

and we recognize in (28) the Rice integral

$$f(n) = \frac{n!}{2i\pi} \int_{\uparrow} \frac{P_f^*(-s)}{\Gamma(-s)} \frac{\Gamma(-s)}{\Gamma(n+1-s)} ds = \frac{1}{2i\pi} \int_{\uparrow} \psi(s) \frac{(-1)^{n+1}\, n!}{s(s-1)\dots(s-n)} ds\,.$$

This exhibits a formal comparison between the two paths. However, this comparison is *only formal* because the previous manipulations may be meaningless due to the divergence of the integrals.