Information theory : Sources, Dirichlet series, and realistic analyses of data structures.

Mathieu Roux
LMNO and GREYC (CNRS and University of Caen), France
mathieu.roux@unicaen.fr

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

Most of the text algorithms build data structures on words, mainly trees, as digital trees (tries) or binary search trees (bst). The mechanism which produces symbols of the words (one symbol at each unit time) is called a source, in information theory contexts. The probabilistic behaviour of the trees built on words emitted by the same source depends on two factors: the algorithmic properties of the tree, together with the information-theoretic properties of the source. Very often, these two factors are considered in a too simplified way: from the algorithmic point of view, the cost of the Bst is only measured in terms of the number of comparisons between words –from the information theoretic point of view, only simple sources (memoryless sources or Markov chains) are studied.
We wish to perform here a realistic analysis, and we choose to deal together with a general source and a realistic cost for data structures: we take into account comparisons between symbols, and we consider a general model of source, related to a dynamical system, which is called a dynamical source. Our methods are close to analytic combinatorics, and our main object of interest is the generating function of the source $\Lambda(s)$, which is here of Dirichlet type. Such an object transforms probabilistic properties of the source into analytic properties. The tameness of the source, which is defined through analytic properties of $\Lambda(s)$, appears to be central in the analysis, and is precisely studied for the class of dynamical sources. We focus here on arithmetical conditions, of diophantine type, which are sufficient to imply tameness on a domain with hyperbolic shape.

Plan of the paper. We first recall in Section 1 general facts on sources and trees, and define the probabilistic model chosen for the analysis. Then, we provide the statements of the main two theorems (Theorem 1 and 2) which establish the possible probabilistic behaviour of trees, provided that the source be tame. The tameness notions are defined in a general framework and then studied in the case of simple sources (memoryless sources and Markov chains). In Section 2, we focus on a general model of sources, the dynamical sources, that contains as a subclass the simple sources. We present sufficient conditions on these sources under which it is possible to prove tameness. We compare these tameness properties to those of simple sources, and exhibit both resemblances and differences between the two classes.

1 Probabilistic behaviour of trees built on general sources.

1.1. General sources. Throughout this paper, an ordered (possibly infinite denumerable) alphabet $\Sigma := \{a_1, a_2, \ldots, a_r\}$ is fixed.

A *probabilistic source*, which produces infinite words of $\Sigma^{\mathbb{N}}$, is specified by the set $\{p_w, w \in \Sigma^\star\}$ of *fundamental probabilities* p_w, where p_w is the probability that an infinite word begins with the finite prefix w. It is furthermore assumed that $\pi_k := \sup\{p_w : w \in \Sigma^k\}$ tends to 0, as $k \to \infty$.

P. Ambrož, Š. Holub and Z. Masáková (Eds.):
8th International Conference WORDS 2011
EPTCS 63, 2011, pp. 199–214, doi:10.4204/EPTCS.63.26

As it is usual in the domain of analytic combinatorics, well described in [11], our analyses involve the generating function of the source, here of Dirichlet type, first introduced in [25] and defined as

$$\Lambda(s) := \sum_{w\in\Sigma^\star} p_w^s, \qquad \Lambda_{(k)}(s) := \sum_{w\in\Sigma^k} p_w^s. \tag{1}$$

Since all the equalities $\Lambda_{(k)}(1) = 1$ hold, the series $\Lambda(s)$ is divergent at $s = 1$, and the probabilistic properties of the source can be expressed in terms of the regularity of Λ near $s = 1$, as it is known from previous works [25] and will be recalled later. For instance, the entropy $h(\mathscr{S})$ relative to a probabilistic source $\mathscr{S}$ is defined as the limit (if it exists) that involves the previous Dirichlet series

$$h(\mathscr{S}) := \lim_{k\to\infty} \frac{-1}{k} \sum_{w\in\Sigma^k} p_w \log p_w = \lim_{k\to\infty} \frac{-1}{k} \frac{d}{ds}\Lambda_{(k)}(s)_{|s=1}. \tag{2}$$

1.2. Simple sources: memoryless sources and Markov chains. A memoryless source, associated to the (possibly infinite) alphabet Σ, is defined by the set $(p_j)_{j\in\Sigma}$ of probabilities, and the Dirichlet series $\Lambda, \Lambda_{(k)}$ are expressed with

$$\lambda(s) = \sum_{i\in\Sigma} p_i^s, \qquad \text{under the form} \qquad \Lambda_{(k)}(s) = \lambda(s)^k, \qquad \Lambda(s) = \frac{1}{1-\lambda(s)}. \tag{3}$$

A Markov chain associated to the finite alphabet Σ, is defined by the vector R of initial probabilities $(r_i)_{i\in\Sigma}$ together with the transition matrix $P := [(p_{i|j})_{(i,j)\in\Sigma\times\Sigma}]$. We denote by $P(s)$ the matrix with general coefficient $p_{i|j}^s$, and by $R(s)$ the vector of components r_i^s. Then

$$\Lambda(s) = 1 + {}^t\mathbf{1} \cdot (I - P(s))^{-1} \cdot R(s). \tag{4}$$

If, moreover, the matrix P is irreducible and aperiodic, then, for any real s, the matrix $P(s)$ has a unique dominant eigenvalue $\lambda(s)$.

In both cases, the entropy satisfies $h(\mathscr{S}) = -\lambda'(1)$.

1.3. The first main data structure: the trie. A trie is a tree structure which is used as a dictionary in various applications, as partial match queries, text processing tasks or compression. This justifies considering the trie structure as one of the central general purpose data structures of Computer Science. See [13] or [23] for an algorithmic study of this structure.

The trie structure compares words via their prefixes: it is based on a splitting according to symbols encountered. If $\mathscr{X}$ is a set of (infinite) words over Σ, then the trie associated to $\mathscr{X}$ is defined recursively by the rule: $\mathtt{Trie}(\mathscr{X})$ is an internal node where are attached the tries $\mathtt{Trie}(\mathscr{X} \setminus a_1), \mathtt{Trie}(\mathscr{X} \setminus a_2), \ldots, \mathtt{Trie}(\mathscr{X} \setminus a_r)$. Here, the set $\mathscr{X} \setminus a$ denotes the subset of $\mathscr{X}$ consisting of strings that start with the symbol a stripped of their initial symbol a; recursion is halted as soon as $\mathscr{X}$ contains less than two elements: if $\mathscr{X}$ is empty, then $\mathtt{Trie}(\mathscr{X})$ is empty; if $\mathscr{X}$ has only one element X, then $\mathtt{Trie}(\mathscr{X})$ is a leaf labelled with X.

For $|\mathscr{X}| = n$, the trie $\mathtt{Trie}(\mathscr{X})$ has exactly n branches, and the depth of a branch is the number of (internal) nodes that it contains. The *path–length* equals the sum of the depth of all branches: this is the total number of symbols that need to be examined in order to distinguish all elements of $\mathscr{X}$. Divided by the number of elements, it is also by definition the cost of a positive search (i.e. searching for a word that is present in the trie). The *size* of the tree is the number of its internal nodes. Adding to the size, the cardinality of $\mathscr{X}$ gives the number of prefixes necessary to isolate all elements of $\mathscr{X}$. It gives also a precise estimate of the place needed in memory to store the trie in a real-life implementation. In this paper, we focus on two trie parameters: the size and the path-length.

1.4. The second main data structure: the binary search tree (`Bst`). We revisit here this well-known structure. Usually, this kind of tree contains keys and the path length of this tree measures the number of key comparisons that are needed to build the tree, and sort the keys by a method closely related to `QuickSort`. This usual cost –the number of key comparisons– is not realistic when the keys have a complex structure, in the context of data bases or natural languages, for instance. In this case, it is more convenient to view a key as a word, and now, the cost for comparing two words (in the lexicographic order) is closely related to the length of their largest common prefix, called the coincidence. The convenient cost of the bst is then the total number of symbol comparisons between words that are needed to build it; this is a kind of a weighted path length, called the *symbol path–length* of the `Bst`, also equal to the total symbol cost of `QuickSort`. For instance, for inserting the key F in the `Bst` of Figure 1, the number of key comparisons equals 3, whereas the number of symbol comparisons equals 18 (7 for comparing F to A, 8 for comparing F to B and 1 for comparing F to C). This is this symbol path length that is studied in the following.

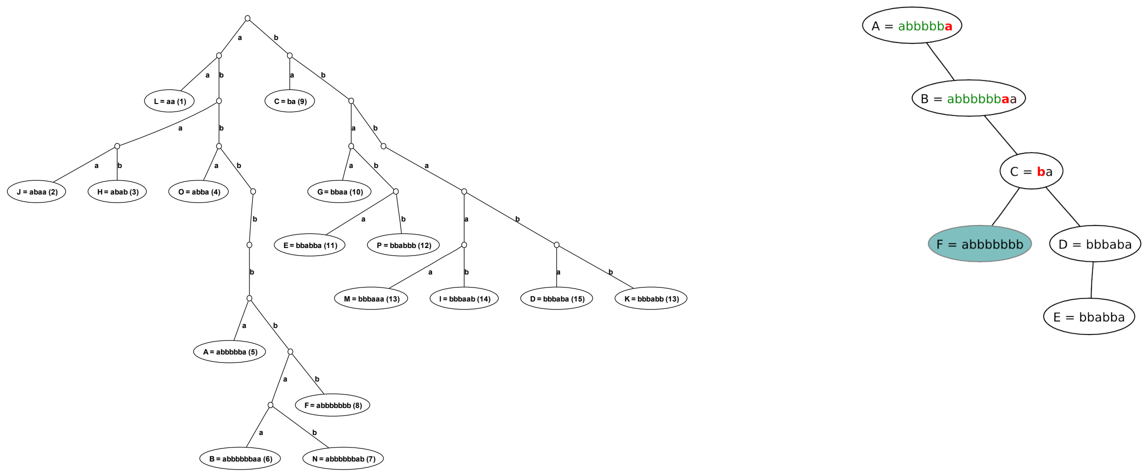

Figure 1: On the left, a trie built on sixteen words of $\{a,b\}^\star$. On the right, a binary search tree built on seven words of $\{a,b\}^\star$.

1.5. Average-case analysis: exact expressions of the three mean costs. The average–case analysis of structures (or algorithms) aims characterizing the mean value of their parameters under a well-defined probabilistic model that describes the initial distribution of its inputs. Here, we adopt the following quite general model: we work with a finite sequence $\mathscr{X}$ of infinite words independently produced by the same general source $\mathscr{S}$, and we wish to estimate the mean value of the parameters when the cardinality n of $\mathscr{X}$ becomes large. Here, in the paper, we focus on three main parameters, two for $\texttt{Trie}(\mathscr{X})$ and one for $\texttt{Bst}(\mathscr{X})$. When restricted to simple sources, there exist many works that study the trie parameters (see [9, 14, 15, 24]) or the symbol path length for `Bst` (see [8]). The same studies, in the case of a general source, are done in [4] for the `Trie` and in [26] for the `Bst`, and are summarized as follows:

Theorem 1 [Clément, Fill, Flajolet, Vallée]. *Let $\mathscr{S}$ be a general source. Consider a finite sequence $\mathscr{X}$ of n infinite words independently produced by $\mathscr{S}$. Then the expectations of the size R of* `Trie`$(\mathscr{X})$*, the path length C of* `Trie`$(\mathscr{X})$*, the symbol path–length B of the binary search tree* `Bst`$(\mathscr{X})$ *are all expressed under the form*

$$T(n) = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \varpi_T(k), \tag{5}$$

where the function $\varpi_T(s)$ is a Dirichlet series which depends on the parameter T and is closely related to the Dirichlet series $\Lambda(s)$ of the source $\mathscr{S}$, defined in (1)

$$\varpi_R(s) = (s-1)\,\Lambda(s), \qquad \varpi_C(s) = s\,\Lambda(s), \qquad \varpi_B(s) = \frac{2}{s(s-1)}\,\Lambda(s). \tag{6}$$

This result provides exact expressions for the mean values of parameters of interest, that are totally explicit for simple sources, due to formulae given in (3) or in (4). As we now wish to obtain an asymptotic form for these mean values, these nice exact expressions are not easy to deal with, due to the presence of the alternate sum. The Rice formula, described in [20, 21] and introduced by Flajolet and Sedgewick [10] into the analytic combinatorics domain, transforms an alternate sum into an integral of the complex plane, provided that the sequence of numerical values $\varpi(k)$ lifts into an analytic function $\varpi(s)$.

Let $T(n)$ be a numerical sequence which can be written as in (5), where the function $\varpi_T(s)$ is analytic in $\Re(s) > C$, with $1 < C < 2$, and is there of polynomial growth with order at most r. Then the sequence $T(n)$ admits a Nörlund–Rice representation*, for $n > r+1$ and any $C < d < 2$.*

$$T(n) = \frac{1}{2i\pi}\int_{-d-i\infty}^{-d+i\infty} \varpi_T(-s)\frac{n!}{s(s+1)\cdots(s+n)}\,ds \tag{7}$$

1.6. Importance of tameness of sources. The idea is now to push the contour of integration in (7) to the right, past -1. This is why we consider the possible behaviours for the function $\varpi_T(s)$ near $\Re s = 1$, more precisely on the left of the line $\Re s = 1$. Due to the close relations between the functions $\varpi_T(s)$ and the Dirichlet series $\Lambda(s)$ of the source given in (6), it is sufficient to consider possible behaviours for $\Lambda(s)$ itself. We will later show why the behaviours that are described in the following definition, already given in [26][1], and shown in Figure 2, arise in a natural way for a large class of sources.

Definition 1 *Let $\mathscr{R}$ be a region that contains the half-plane $\Re s \geq 1$.*
A source $\mathscr{S}$ is $\mathscr{R}$–entropic if $\Lambda(s)$ is meromorphic on $\mathscr{R}$ with a simple pole at $s = 1$, simple, whose residue involves the entropy $h(\mathscr{S})$ under the form $1/h(\mathscr{S})$.

A source is $\mathscr{R}$–tame if (i) it is $\mathscr{R}$–entropic, – (ii) $\Lambda(s)$ has no other pole than $s = 1$ in $\mathscr{R}$, – (iii) $\Lambda(s)$ is of polynomial growth in $\mathscr{R}$ as $|s| \to +\infty$.

A source is
(a) strongly–tame (S–tame in shorthand) of abscissa δ if there exists a vertical strip $\mathscr{R}$ of the form $\Re(s) > 1-\delta$, with $\delta > 0$, where $\Lambda(s)$ is $\mathscr{R}$– tame.
(b) hyperbolically tame (H–tame in shorthand) of exponent α if there exists a hyperbolic region $\mathscr{R}$, with $A, B, \alpha > 0$

$$\mathscr{R} := \left\{ s = \sigma + it;\ |t| \geq B,\ \sigma > 1 - \frac{A}{t^\alpha} \right\} \bigcup \left\{ s = \sigma + it;\ \sigma > 1 - \frac{A}{B^\alpha}, |t| \leq B \right\},$$

[1] There are slight differences between the two definitions but the "spirit" is the same.

where $\Lambda(s)$ is $\mathscr{R}$–tame.
A source $\mathscr{S}$ is periodic of abscissa δ, if there exists a vertical strip $\mathscr{R}$ of the form $\Re(s) > 1-\delta$, with $\delta > 0$, where $\Lambda(s)$ is $\mathscr{R}$ entropic and admits a singularity at a point $1+it_0$, for some real $t_0 > 0$ [2].

For an entropic source, the Dirichlet series $\varpi_T(s)$ has a pole of order 0 (for the `Trie` size, cost R), a pole of order 1 (for the `Trie` path length, cost C), a pole of order 2 (for the `Bst` symbol path length, cost B).

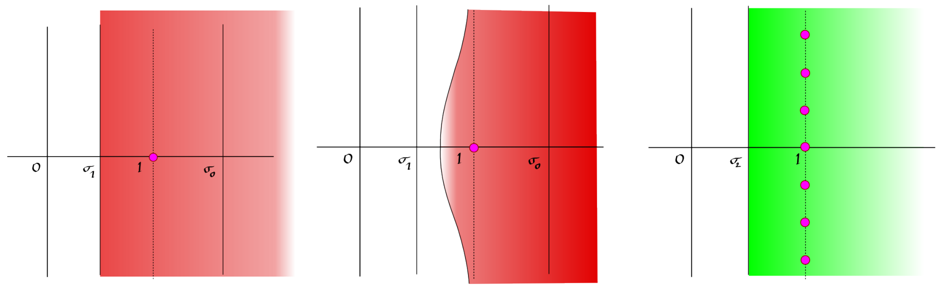

Figure 2: Three possible domains where the function $\varpi(s)$ is analytic and of polynomial growth.

1.7. Average-case analysis: asymptotic expressions of the three mean costs. Now, the following result shows that the shape of the tameness region (described by the order, the abscissa, the exponent) essentially determine the behaviour of the Rice integral in (7), and thus the asymptotic behaviour of our main parameters of interest: It provides a dictionary which transfers the tameness properties of the source into asymptotic properties of the sequence $T(n)$. The following theorem gathers and makes more precise results that are already obtained in [4] or [26] :

Theorem 2 [Clément, Fill, Flajolet, Vallée]. *The asymptotics of each cost $T(n)$ of interest, relative to a parameter of a tree built on a general source $\mathscr{S}$, and defined in Theorem 1, is of the general following form $T(n) = P_T(n) + E(n)$. The "principal term" $P_T(n)$ involves the entropy $h(\mathscr{S})$ under the form*

$$P_R(n) = \frac{1}{h(\mathscr{S})} n, \qquad P_C(n) = \frac{1}{h(\mathscr{S})} n\log n + an, \qquad P_B(n) = \frac{1}{h(\mathscr{S})} n\log^2 n + bn\log n + cn,$$

together with some other constants a,b,c. The "error term" $E(n)$ admits the possible following forms, depending on the tameness of the source

(a) If $\mathscr{S}$ is S–tame with abscissa δ_0, then $E(n) = O(n^{1-\delta})$, for any $\delta < \delta_0$.

(b) If $\mathscr{S}$ is H–tame with exponent α_0, then $E(n) = n \cdot O(\exp[-(\log n)^{\alpha}])$ for any $\alpha < 1/(\alpha_0+1)$.

(c) If $\mathscr{S}$ is periodic with abscissa δ_0, then $E(n) = n\cdot\Phi(\log n) + O(n^{1-\delta})$, for any $\delta < \delta_0$, where $n\cdot\Phi(\log n)$ is the part of the expansion brought by the family of the non real poles located on the vertical line $\Re s = 1$, and involves a periodic function Φ.

Note that the "error term" $E(n)$ is not always ... an actual error term: in the case of the trie size, for a periodic source, the fluctuation terms given by $E(n)$ arise in the main term. However, in all the other cases, the term $E(n)$ is indeed an error term. The main term of the principal term always involves a constant equal to $1/h(\mathscr{S})$, and the order of the main term depends on the tree parameter: it is always of the form $n\log^k n$, and the integer k equals the order of the pole $s=1$ for the Dirichlet series $\varpi_T(s)$: one

[2] This implies that $\varpi(s)$ admits singularities at all the points $1+ikt_0$ for any integer k, and is of polynomial growth on a family of horizontal lines $t = t_k$ with $t_k \to \infty$, and on vertical lines $\Re(s) = 1-\delta'$ with some $\delta' < \delta$

has $k = 0$ for the `Trie` size, $k = 1$ for the `Trie` path length, and $k = 2$ for the `Bst` symbol path length. This result proves that, with respect to the number of symbol comparisons, the `Bst` is much less efficient than the `Trie`.

1.8. Tameness of simple sources. We show that tameness properties that are described in Definition 1 arise in a natural way for simple sources. Even if S–tameness never occurs for simple sources, we will see later that it "often" occurs for most of more "complex" sources. We now focus on the memoryless case, defined by the probabilities $\mathfrak{P} = (p_1, p_2, \ldots, p_r)$, to which we associate the ratios $\alpha_{k,j} := \log p_j / \log p_k$. Then, tameness properties depend on arithmetic properties of the ratios $\alpha_{k,j}$.

Proposition 1 *Any simple source (memoryless source or irreducible aperiodic Markov chain) is entropic. A memoryless source is periodic if and only, for any fixed k, all the real numbers $\alpha_{k,j}$ are rationals with the same denominator.*

We now focus on non-periodic memoryless sources, where there exists, amongst all the reals $\alpha_{k,j}$, at least one real $\alpha_{k,j}$ which is irrational. In this case, there is no other pole of $\Lambda(s)$ than $s = 1$ on the vertical line $\Re s = 1$ but there exist poles of Λ which are arbitrary close to the vertical line $\Re s = 1$. This entails that a simple source is never strongly tame. The distribution of distances of the poles with respect to the vertical line $\Re s = 1$ depends on the degree of approximability of the family α by rationals, as it was first remarked in [7]. We recall some notions on diophantine approximations (see for instance [16]). The irrationality exponent of a real x is defined by

$$\omega(x) := \sup \left\{ \alpha, \left| x - \frac{p}{q} \right| \leq \frac{1}{q^{2+\alpha}} \quad \text{for an infinite number of pairs } (p,q) \in \mathbb{N}^2 \right\}.$$

A number x is diophantine if its irrationality exponent is finite. The following result provides a characterisation of H–tameness for simple sources. It can be found in a more precise form in [12], where the authors revisit previous results of [17].

Theorem 3 [Flajolet, Roux, Vallée]. *A memoryless source is H–tame if and only it is diophantine. Moreover, there is a relation between the exponent α of H–tameness and the irrationality exponent $\mu(\mathfrak{P})$: one can choose as α any real strictly greater than $2\mu(\mathfrak{P}) + 2$, and it is in a precise sense the best possible choice.*

With the general Theorem 2, together with Propositions 1 and 2, we can precisely describe the asymptotic probabilistic behaviour of two main tree data structures built on words produced by memoryless sources. Generally speaking, Theorem 2 can be applied to tree structures built on a general source as soon as its tameness may be studied. The following of the paper describes a general class of sources, which contains the simple sources, for which tameness properties can be precisely studied. We will see that tameness of these general sources may be quite different from tameness of simple sources.

2 Tameness of dynamical sources.

We first define the class of dynamical sources and explain their relation with simple sources. Then, we recall the expression of the Dirichlet series $\Lambda(s)$ as a function of the secant transfer operator of the underlying dynamical systems. Finally, we exhibit sufficient conditions on the underlying dynamical system under which it is possible to prove tameness properties [Theorem 4 for S–tameness, and Theorem 5 for H–tameness].

2.1. Definition of dynamical sources. A dynamical source, defined in [25] is closely related to a dynamical system on the interval.

Definition 2 *A dynamical system of the interval* $\mathscr{I} := [0,1]$ *is defined by a mapping* $T : \mathscr{I} \to \mathscr{I}$ *(called the shift) for which*

(a) *there exists a finite alphabet* Σ*, and a topological partition of* $\mathscr{I}$ *with disjoint open intervals* $(\mathscr{I}_m)_{m\in\Sigma}$*, i.e.* $\mathscr{I} = \bigcup_{m\in\Sigma} \overline{\mathscr{I}}_m$*.*

(b) *The restriction of* T *to each* $\mathscr{I}_m$ *is a* $\mathscr{C}^2$ *bijection from* $\mathscr{I}_m$ *to* $T(\mathscr{I}_m)$*.*

The system is complete when each restriction is surjective, i.e., $T(\mathscr{I}_m) = \mathscr{I}$*. The system is Markovian when each interval* $T(\mathscr{I}_m)$ *is a union of intervals* $\mathscr{I}_j$*.*

A dynamical system, together with a distribution G on the unit interval $\mathscr{I}$, defines a probabilistic source, which is called a dynamical source and is now described (See also Fig.1 at the end). The map T is used as a shift mapping, and the mapping τ whose restriction to each $\mathscr{I}_m$ is equal to m, is used for coding. The words are emitted as follows [see Figure 3]: To each real x, (except for a denumerable set), one associates the trajectory $\mathscr{T}(x) = (x, T(x), T^2(x), \dots T^j(x), \dots)$, which gives rise, via the mapping τ to the word $M(x) \in \Sigma^{\mathbb{N}}$,

$$M(x) = (m_1(x), m_2(x), \dots, m_n(x), \dots) \qquad \text{with} \quad m_j(x) = \tau(T^{j-1}(x)).$$

Given a prefix $w \in \Sigma^\star$, the set $\mathscr{I}_w$ of all reals x for which the word $M(x)$ begins with the prefix w is an interval, the fundamental interval associated to w, and the measure of this interval (with respect to distribution G), is the fundamental probability p_w of the source. In the case of a complete system, one denotes by $h_{[m]}$ the local inverse of T restricted to $\mathscr{I}_m$ and by $\mathscr{H}$ the set $\mathscr{H} := \{h_{[m]}, m \in \Sigma\}$ of all local inverses. Each local inverse of the k–th iterate T^k is then associated to a word $w = m_1 m_2 \dots m_k \in \Sigma^k$; it is of of the form $h_{[w]} := h_{[m_1]} \circ h_{[m_2]} \dots h_{[m_k]}$, and

$$\mathscr{I}_w = h_{[w]}(\mathscr{I}), \qquad p_w = |G(h_{[w]}(1)) - G(h_{[w]}(0))|. \tag{8}$$

The set of all the inverse branches of T^k is $\mathscr{H}^k = \{h_{[w]};\ w \in \Sigma^k\}$. For $h \in \mathscr{H}^k$, the number k is called the *depth* of h and it is denoted by $p(h)$. We denote by $\mathscr{H}^\star := \cup_{k\geq 0}\mathscr{H}^k$ the set of all inverse branches.

Such sources may possess a high degree of correlations, due to the *geometry* of the branches and also to the *shape* of branches.

The geometry of the branches is defined by the respective positions of "horizontal" intervals $\mathscr{I}_m$ with respect to "vertical" intervals $\mathscr{J}_\ell := T(\mathscr{I}_\ell)$ and allows to describe the set $\mathscr{S}_m$ formed with symbols which can be possibly emitted after symbol m. The geometry of the system then provides a first access to the correlation between successive symbols. In particular, in a *complete* system, any symbol of Σ can be emitted after any symbol m, and thus the equality $\mathscr{S}_m = \Sigma$ always holds.

The shape of the branches, and more precisely, the behavior of derivatives $|h'_m|$ has also a great influence on correlations between symbols. For a fixed geometry of the branches, a system with affine branches is "less correlated" than the other systems with the same geometry. The contraction properties of $\mathscr{H}$, (i.e., the fact that $|h'_m| < 1$) are also essential, since they give rise to chaotic behaviour of the trajectories.

2.2. Simple sources viewed as dynamical sources. All memoryless sources and all Markov chain sources belong to the general framework of dynamical sources and correspond to a piecewise linear shift, under this angle of dynamical sources. For instance, the standard binary system is obtained by $T(x) = \{2x\}$ ($\{\cdot\}$ is the fractional part). More precisely:

– A memoryless source is a complete dynamical source, with affine branches and a uniform initial distribution,
– A Markov chain is a Markovian dynamical source, with affine branches and a family of uniform initial distributions on each $\mathscr{I}_j$.

Figure 3 shows three instances of simple sources, viewed as dynamical sources.

However, as soon as the derivatives h' of the branches are not constant, there exist correlations between successive symbols, and the dynamical source is no longer simple. Dynamical sources with a non-linear shift allow for correlations that depend on the entire past. A main instance is the dynamical source relative to the Gauss map, represented in Figure 3, which underlies the Euclid Algorithm and is defined on the unit interval via the shift T

$$T(0) = 0, \qquad T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \quad (x \neq 0). \tag{9}$$

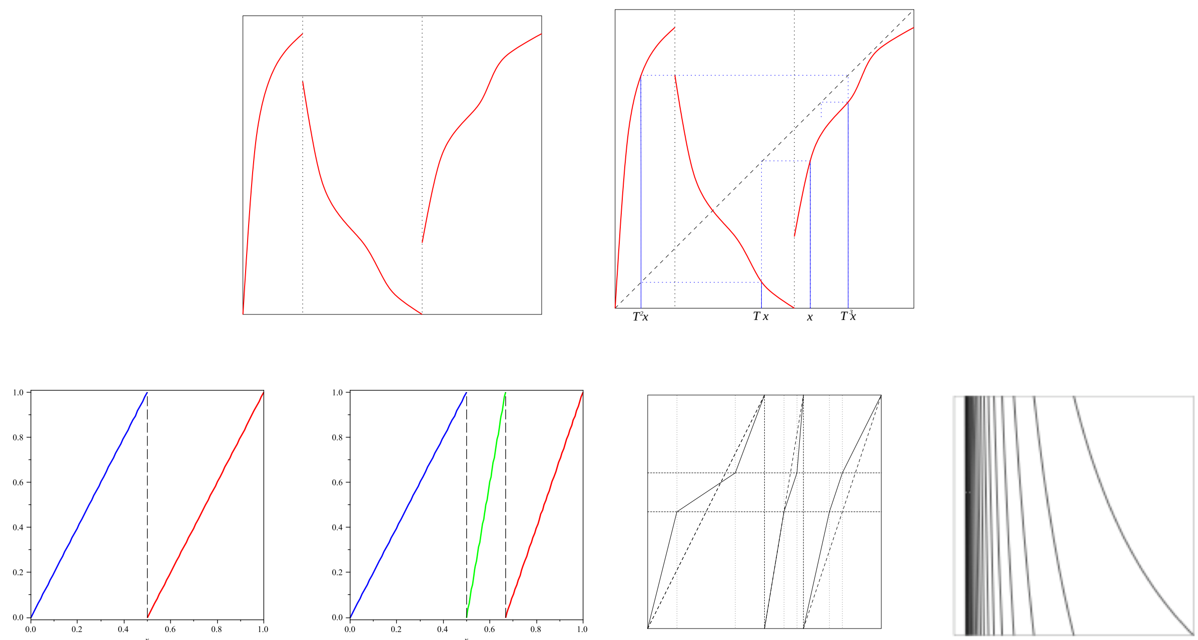

Figure 3: (Up) A dynamical system, with $\Sigma = \{a, b, c\}$ and a word $M(x) = (c, b, a, c \ldots)$ – (Down) Two memoryless sources and a Markov chain, viewed as dynamical sources. The continued fraction source.

2.3. Transfer operators. One of the main tools in dynamical system theory is the transfer operator introduced by Ruelle, denoted by $\mathbf{H}_s$. It generalizes the density transformer $\mathbf{H}$ that describes the evolution of the density.

We here consider the case of a complete dynamical system: if $f = f_0$ denotes the initial density on $\mathscr{I}$, and f_1 the density on $\mathscr{I}$ after one iteration of T, then f_1 can be written as $f_1 = \mathbf{H}[f_0]$, where $\mathbf{H}$ is defined by

$$\mathbf{H} := \sum_{h \in \mathscr{H}} \mathbf{H}_{(h)} \qquad \text{with} \quad \mathbf{H}_{(h)}[f](x) := |h'(x)|\, f \circ h(x).$$

The transfer operator extends the density transformer; it depends on a complex parameter s,

$$\mathbf{H}_s = \sum_{h \in \mathscr{H}} \mathbf{H}_{(h),s} \qquad \text{with} \quad \mathbf{H}_{(h),s}[f](x) := |h'(x)|^s \cdot f \circ h(x), \tag{10}$$

and coincides with **H** when $s = 1$. Here, we are interested by generating the fundamental probabilities, whose expression is provided in (8) in the case of a complete dynamical system. The main tool is a generalized version of the transfer operator –the secant transfer operator– introduced by Vallée in [25]. This operator involves the secant function of inverse branches (instead of their derivatives), it acts on functions F of two variables; for $s \in \mathbb{C}$, and $h \in \mathscr{H}$, we first define the component secant operator $\mathbb{H}_{(h),s}$ as

$$\mathbb{H}_{(h),s}[F](x,y) := \left|\frac{h(x)-h(y)}{x-y}\right|^s F(h(x),h(y)), \tag{11}$$

and the secant transfer operator is defined as

$$\mathbb{H}_s := \sum_{h \in \mathscr{H}} \mathbb{H}_{(h),s}, \tag{12}$$

Denote by $\operatorname{diag} F$ the function defined by $\operatorname{diag} F(x) := F(x,x)$. The equality $\mathbb{H}_s[F](x,x) = \mathbf{H}_s[\operatorname{diag} F](x)$ holds on the diagonal $x = y$ and shows that the secant operator is an extension of the plain transfer operator. Moreover, multiplicative properties of secants then entail the relation

$$\mathbb{H}_s^k = \sum_{h \in \mathscr{H}^k} \mathbb{H}_{(h),s} \qquad \text{so that} \quad \mathbb{H}_s^k[F](x,y) = \sum_{h \in \mathscr{H}^k} \left|\frac{h(x)-h(y)}{x-y}\right|^s F(h(x),h(y)).$$

Finally, the Dirichlet series can be expressed as a quasi–inverse of the secant operator: this is a nice extension of the expressions obtained for simple sources, in (3,4).

Proposition 2 [Vallée]. *For a complete dynamical source, relative to a shift T and a distribution G, the Dirichlet series of the source admits an alternative expression which involves the quasi–inverse of the secant operator, defined in (12) applied to the function L^s, where L is the secant of the distribution G,*

$$\Lambda(s) = (I - \mathbb{H}_s)^{-1}[L^s](0,1), \qquad \textit{with} \quad L(x,y) := \frac{G(x)-G(y)}{x-y}.$$

2.4. Tameness of dynamical sources. Here, we consider subclasses of dynamical sources, for which the quasi-inverse has nice spectral properties. This will entail, with Proposition 2, nice properties for the function $\Lambda(s)$, from which one deduces tameness properties. The main results are as follows: There exist natural instances of dynamical sources which are S–tame, or H–tame. A "random" dynamical source is "very often" S–tame: this happens as soon as its inverse branches have "not too often" the same "shape". A dynamical source can be periodic only if it "closely resembles" a memoryless source. A dynamical source is H–tame if, informally speaking, its arithmetical properties are the same as the arithmetical properties of a H-tame memoryless source. More precisely, we define three (large) subclasses of dynamical sources – the `Good` Class, the `UNI` Class, the `DIOP` Class– for which we can describe the tameness in an informal setting. The `UNI` Class has been already studied and described in previous works [5, 1, 2, 3]. The original part of our work is related to the `DIOP` Class, for which we revisit and extend previous results described in [6, 18, 19]. We first state the main tameness results for dynamical sources in an informal way:

Theorem. *All the sources of the* `Good-UNI` *Class are S–tame. All the sources of the* `Good-DIOP` *Class are H–tame. A source of the* `Good` *Class may be periodic only if it is conjugated to a source with affine branches.*

2.5. The `Good` Class. We first define the `Good` Class, for which the shift is expansive, and gives rise to a chaotic behaviour for the trajectories.

Definition 3 [`Good` Class]. *A dynamical system of the interval $(\mathscr{I}, T)$ belongs to the* `Good` *Class if it is complete, with a set $\mathscr{H}$ of inverse branches which satisfies the following:*

$(G1)$ *The set $\mathscr{H}$ is uniformly contracting, i.e., there exists a constant $\rho < 1$, for which*

$$\forall h \in \mathscr{H}, \quad \beta_h := \sup\{|h'(x)|;\ x \in \mathscr{I}\} \leq \rho.$$

$(G2)$ *There is a constant $A > 0$, so that every inverse branch $h \in \mathscr{H}$ satisfies $|h''| \leq A|h'|$.*

$(G3)$ *There exists $\sigma_0 < 1$ for which the series $\sum_{h\in\mathscr{H}} \beta_h^s$ converges on $\Re s > \sigma_0$.*

The essential condition is $(G1)$. The bounded distortion property $(G2)$ and the property $(G3)$ are technical conditions that always fulfilled for a finite alphabet Σ.

When the dynamical system belongs to the `Good` Class, the transfer operators (tangent and secant) act on spaces of functions of $\mathscr{C}^1$ class. They admit dominant spectral properties for s near the real axis, together with a spectral gap. This implies that, for s near 1, the function $\Lambda(s)$ is meromorphic for s with a small imaginary part, and admits a simple pôle at $s = 1$.

2.6. The `UNI` Condition. One first defines a probability $\Pr_n$ on each set $\mathscr{H}^n \times \mathscr{H}^n$, in a natural way, and lets $\Pr_n\{(h,k)\} := |h(\mathscr{I})| \cdot |k(\mathscr{I})|$, where $|\mathscr{J}|$ denotes the length of the interval $\mathscr{J}$. Furthermore, $\Delta(h,k)$ denotes the "distance" between two inverse branches h and k of same depth, defined as

$$\Delta(h,k) = \inf_{x\in\mathscr{I}} |\Psi'_{h,k}(x)| \qquad \text{with} \quad \Psi_{h,k}(x) = \log\left|\frac{h'(x)}{k'(x)}\right|. \tag{13}$$

The distance $\Delta(h,k)$ is a measure of the difference between the "shape" of the two branches h,k. The `UNI` Condition, stated as follows [5], is a geometric condition which expresses that the probability that two inverse branches have almost the same "shape" is very small:

Definition 4 [Condition `UNI`]. *A dynamical system $(\mathscr{I}, T)$ satisfies the* `UNI` *condition if its set $\mathscr{H}$ of inverse branches satisfies the following*

$(U1)$ *For any $a \in]0,1[$, and for any integer n, one has $\Pr_n[\,\Delta \leq \rho^{an}] << \rho^{an}$.*

$(U2)$ *Each $h \in \mathscr{H}$ is of class $\mathscr{C}^3$ and for any n, there exists B_n for which $|h'''| \leq B_n|h'|$ for any $h \in \mathscr{H}^n$.*

For a source with affine branches, the "distance" Δ is always zero, and the probabilities of Assertion $(U1)$ are all equal to 1. Such a source never satisfies the Condition `UNI`. Conversely, a dynamical source of the `Good-UNI` Class cannot be conjugated to a source with affine branches, as it is proven by Baladi and Vallée [1]. Then, the condition `UNI` excludes all the simple sources, which cannot be S–tame. The strength of the Condition `UNI` is due to the fact that this condition is sufficient to imply strong tameness :

Theorem 4 [Dolgopyat, Baladi–Vallée, Cesaratto–Vallée] *When the dynamical system of the* `Good` *Class satisfies the condition* `UNI`*, it gives rise to a S–tame source.*

There are natural instances of sources that belong to the `Good-UNI` Class, for instance the Euclidean dynamical system defined in (9), together with two other dynamical systems, of the Euclidean type.

2.7. The diophantine conditions. The `Good-UNI` Class gathers systems which are quite different from systems with affine branches. The `DIOP` Condition "copies" the behaviour of memoryless sources, when they are H–tame. In this case, we recall that there exists a ratio $\log p_i / \log p_k$ which is diophantine, i.e., whose irrationality exponent is finite.

The DIOP condition is an arithmetical condition, which extends this condition to a system of the Good Class. For an inverse branch h, one denotes by $h^\star$ its unique fixed point (such a point exists and is unique for a system of the Good Class), by $p(h)$ its depth, and one lets, for h,k,ℓ in $\mathscr{H}^\star$,

$$c(h) = \frac{\log|h'(h^\star)|}{p(h)}, \qquad c(h,k) = \frac{c(h)}{c(k)}, \qquad c(h,k,\ell) = \frac{c(h)-c(k)}{c(h)-c(\ell)}.$$

We can now state the definition of diophantine dynamical sources:

Definition 5 [DIO2 and DIOP3]. *A dynamical source is 2–diophantine (* [DIOP2] *in shorthand) if there exist two branches h et k of $\mathscr{H}^\star$ for which the ratio $c(h,k)$ is diophantine.*
A dynamical source is 3–diophantine ([DIOP3] *in shorthand) if there exist three branches h, k and ℓ of $\mathscr{H}^\star$ for which the ratio $c(h,k,\ell)$ is diophantine*

The following result proves that these conditions are sufficient to entail H–tameness of associated sources. This is the main contribution of Roux' PhD thesis [22]. The appendix contains hints on the proof, that will be detailed in the long version.

Theorem 5 [Dologopyat, Naud, Melbourne, Roux–Vallée]
(a) *A dynamical system of the* Good *Class, which is moreover* DIOP3*, gives rise to a H–tame source.*
(b) *A dynamical system of the Good Class, which is moreover* DIOP2*, gives rise to a H–tame source.*

2.8. A little piece of history. Dolgopyat, in two seminal papers [5, 6], introduces the Conditions UNI and DIOP2. He proves that, under these conditions, the quasi-inverse of the plain (tangent) transfer operator has nice properties in a region on the left of the line $\{\Re s = 1\}$: when the UNI Condition holds, the region is a vertical strip, and when the DIOP2 Condition holds, the region is of hyperbolic type. However, he does not consider the case of an infinite number of branches, and his results are extended to this case by Baladi and Vallée in [1, 2] for the UNI condition, and by Melbourne [18] in the case of the DIOP condition, who introduces the DIOP3 Condition. However, in order to deal with the Dirichlet series $\Lambda(s)$, one needs to extend the previous proofs to the secant operator. This have been done by Cesaratto and Vallée in [3] for the UNI Condition. Here, we deal with the DIOP conditions and we perform two extensions: we consider a possible infinite alphabet, we deal both with the DIOP3 (where we use a method due to Melbourne [18]) and the DIOP2 condition (where we use a method due to Naud [19]). We also extend these results to the secant operator.

Acknowledgements. This work takes place inside the ANR project MAGNUM [Méthodes Algorithmiques pour la Génération Non Uniforme: Modèles et Applications] [ANR 2010 BLAN 0204] .

References

[1] BALADI, V., AND VALLÉE, B. Euclidean algorithms are Gaussian. *Journal of Number Theory 110* (2005), 331–386. doi:10.1016/j.jnt.2004.08.008

[2] BALADI, V., AND VALLÉE, B. Exponential decay of correlations for surface semi-flows without finite Markov partitions. *Proceedings of the American Mathematical Society 133*, 3 (2005), 865–874 (electronic).

[3] CESARATTO, E., AND VALLÉE, B. Gaussian distribution of trie depth for dynamical sources. Manuscript,, 2010.

[4] CLÉMENT, J., FLAJOLET, P., AND VALLÉE, B. Dynamical sources in information theory: A general analysis of trie structures. *Algorithmica 29*, 1/2 (2001), 307–369. doi:10.1007/BF02679623

[5] DOLGOPYAT, D. On decay of correlations in Anosov flows, *Annals of Mathematics* 147 (1998) 357-390. doi:10.2307/121012

[6] DOLGOPYAT, D. Prevalence of rapid mixing in hyperboloic flows (I) *Ergodic Theory and Dynamical Systems* 18 (1998) 1097-1114. doi:10.1017/S0143385798117431

[7] FAYOLLE, G., FLAJOLET, P., AND HOFRI, M. On a functional equation arising in the analysis of a protocol for a multi-accessbroadcast channel. *Adv. Appl. Prob.*, 18 (1986), 441–472.

[8] FILL, J. A., AND JANSON, S. The number of bit comparisons used by Quicksort: An average-case analysis. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA04)* (2001), 293–300.

[9] FLAJOLET, P. The ubiquitous digital tree. In *STACS 2006* (2006), B. Durand and W. Thomas, Eds., vol. 3884 of *Lecture Notes in Computer Science*, pp. 1–22. Proceedings of 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, February 2006.

[10] FLAJOLET, P., AND SEDGEWICK, R. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science 144*, 1–2 (June 1995), 101–124. doi:10.1016/0304-3975(94)00281-M

[11] FLAJOLET, P., AND SEDGEWICK, R. *Analytic Combinatorics*. Cambridge University Press, 2009. Available electronically from the authors' home pages. doi:10.1017/CBO9780511801655

[12] FLAJOLET, P., ROUX, M., and VALLÉE, B. Digital Trees and Memoryless Sources: from Arithmetics to Analysis, *Proceedings of the AofA'10 conference*, DMTCS, proc AM, pp 231–258 (2010)

[13] GONNET, G. H., AND BAEZA-YATES, R. *Handbook of Algorithms and Data Structures: in Pascal and C*, second ed. Addison–Wesley, 1991.

[14] JACQUET, P., AND SZPANKOWSKI, W. Analysis of digital tries with Markovian dependency. *IEEE Transactions on Information Theory 37*, 5 (1991), 1470–1475. doi:10.1109/18.133271

[15] KNUTH, D. E. *The Art of Computer Programming*, 2nd ed., vol. 3: Sorting and Searching. Addison-Wesley, 1998.

[16] LAGARIAS, J. C. Best simultaneous Diophantine approximations I: Growth rates of best approximation denominators. *Transactions of the American Mathematical Society 272*, 2 (1982), 545–554. doi:10.2307/1998713

[17] LAPIDUS, M. L., AND VAN FRANKENHUIJSEN, M. *Fractal Geometry, Complex Dimensions and Zeta Functions: Geometry and Spectra of Fractal Strings*. Springer, 2006.

[18] MELBOURNE, I. Rapid decay of correlations for nonuniformly hyperbolic flows. *Trans. Amer. Soc. 359* (2007) 2421–2441.

[19] NAUD, F. Analytic continuation of a dynamical zeta function under a diophantine condition. *Nonlinearity 14* (2001) 995–1009. doi:10.1088/0951-7715/14/5/305

[20] NÖRLUND, N. E. Leçons sur les équations linéaires aux différences finies. In *Collection de monographies sur la théorie des fonctions*. Gauthier-Villars, Paris, 1929.

[21] NÖRLUND, N. E. *Vorlesungen über Differenzenrechnung*. Chelsea Publishing Company, New York, 1954.

[22] ROUX, M. Séries de Dirichlet, Théorie de l'information, et Analyse d'algorithmes, PhD thesis, in preparation

[23] SEDGEWICK, R. *Algorithms in C, Parts 1–4*, third ed. Addison–Wesley, Reading, Mass., 1998.

[24] SZPANKOWSKI, W. *Average-Case Analysis of Algorithms on Sequences*. John Wiley, 2001.

[25] VALLÉE, B. Dynamical sources in information theory: Fundamental intervals and word prefixes. *Algorithmica 29*, 1/2 (2001), 262–306. doi:10.1007/BF02679622

[26] VALLÉE, B., CLÉMENT, J., FILL, J. A., AND FLAJOLET, P. The number of symbol comparisons in QuickSort and QuickSelect. In *ICALP 2009, Part I* (2009), S. A. *et al.*, Ed., vol. 5555 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 750–763. Proceedings of the 36th International Colloquium on Automata, Languages and Programming.

3 Some hints on the proof of Theorem 5.

Since the Dirichlet series $\Lambda(s)$ is expressed with the quasi-inverse of the secant operator $\mathbb{H}_s$ we study the behaviour of this quasi-inverse on the vertical line $\Re s = 1$. It is closely related to the behaviour of the operators $\mathbf{M}_t, \mathbb{M}_t$ defined by

$$\mathbf{M}_t[f](x) := |T'(x)|^{it} f\circ T(x), \qquad \mathbb{M}_t[F[(x,y) := \left|\frac{T(x-T(y)}{x-y}\right|^{it} F(T(x),T(y)).$$

3.1. Various possibilities for the spectral radius of the operator $\mathbb{H}_s$ on $\Re s = 1$. The beginning point is the following proposition, that is classical for the tangent operator, and can be easily extended to the secant operator.

Proposition 3. *Consider a dynamical system of the* `Good` *Class and its secant transfer operator $\mathbb{H}_s$, acting on the space $\mathscr{C}^1(\mathscr{I}\times\mathscr{I})$ for a parameter s of the form $s = 1+it_0$, with $t_0 \neq 0$.*

(a) *For a complex number λ of modulus 1, the two conditions are equivalent:*

- $(a1)$ *The complex number λ belongs to the spectrum $\mathrm{Sp}\,\mathbb{H}_{1+it_0}$.*
- $(a2)$ *The complex number λ^{-1} is an eigenvalue of $\mathbf{M}_{t_0}$.*

(b) *Assume that there exists $t_0 \neq 0$ for which the condition $(a2)$ is satisfied. Then, there exist $a \neq 0$ and b for which the quantities $c(h)-b$ all belong to the $\mathbb{Z}$-module $\mathbb{Z}a$.*

- $(b1)$ *If λ is a root of unity, then all the ratios $c(h,k)$ are rationals.*
- $(b2)$ *If λ is any complex number of modulus 1, all the ratios $c(h,k,\ell)$ are rationals.*

(c) *If one of the two conditions is satisfied*

- $(c1)$ *there exists a ratio $c(h,k)$ which is not rational,*
- $(c2)$ *For any $t \neq 0$, the spectrum of the operator $\mathbb{M}_t$ does not contain $\lambda = 1$.*

then, the quasi-inverse $(I-\mathbb{H}_s)^{-1}$ is analytic on $\Re s = 1$ except at $s = 1$ where it has a simple pole.

(d) *If one of the two conditions is satisfied*

- $(d1)$ *there exists a ratio $c(h,k,\ell)$ which is not rational,*
- $(d2)$ *For any $t \neq 0$, the spectrum of the operator $\mathbb{M}_t$ does not contain any λ with $|\lambda| = 1$,*

then, the spectral radius of $\mathbb{H}_s$ is strictly less than 1 on $\{s; \Re s = 1, s \neq 1\}$ and, for any λ of modulus 1, the quasi-inverse $(I-\lambda\mathbb{H}_s)^{-1}$ is analytic on the line $\Re s = 1$ except at $s = 1$ where it admits a simple pole.

3.2. Reinforcement of conditions $(c),(d)$. The main question is now as follows: if one of the conditions $(c1)$ or $(d1)$ or $(c2)$ or $(d2)$ is replaced by a stronger condition, is it possible to obtain a conclusion about tameness, of the following kind:

(R) *There exists a region on the left of the vertical line $\Re s = 1$ on which the quasi-inverse $(I-\mathbb{H}_s)^{-1}$ is analytic except at $s = 1$ (where it admits a simple pole), and is of polynomial growth for $t = \Im s \to \infty$.*

We deal here with the Banach space $\mathscr{C}^1(\mathscr{I}\times\mathscr{I})$ formed with functions of class $\mathscr{C}^1$ on the unit square, endowed with the norm $||.||_1$ defined by $||u||_1 := \sup|u(x,y| + \sup||u'(x,y)||$, but we also use a norm $||.||_{(t)}$ which depends on the imaginary part t of s, defined by $||u||_{(t)} := \sup|u(x,y| + (1/|t|)\sup||u'(x,y)||$. Our main object of study is

$$\mathscr{R}(t) := \left\|(I-\mathbb{H}_{1+it})^{-1}\right\|_{(t)}. \tag{14}$$

A possible reinforcement DIOP3 of the condition $(d1)$ is "There exists a triple (h,k,ℓ) for which $c(h,k,\ell)$ is diophantine". A possible reinforcement $(d3)$ of the condition $(d2)$ is: "The operator $\mathbb{M}_t$ does not admit a system of almost eigenfunctions" for which a more formal statement will be provided later. We will also see that these two reinforcements are not independent since the implication DIOP3 $\Rightarrow (d3)$ holds

A possible reinforcement DIOP2 of the condition $(c1)$ is "There exists a pair (h,k) for which $c(h,k)$ is diophantine". A possible reinforcement $(c3)$ of the condition $(c2)$ is: "The operator $\mathbb{M}_t$ does not admit a system of almost invariant functions" for which a more formal statement will be provided later. We will also see that these two reinforcements are not independent since the implication DIOP2 $\Rightarrow (c3)$ holds

3.3. Precise statement of Theorem 5. There are two theorems, one for each condition DIOP2 or DIOP3.

Theorem 5(a)**.** [DIOP3] *Consider a dynamical source of the* Good *Class, with a possibly infinite denumerable alphabet, with a contraction ratio* $\rho < 1$. *If there exists a triple* $\{h,k,\ell\}$, *with* $\nu = \max\{c(h), c(k), c(\ell)\}$, *for which* $c(h,k,\ell)$ *is diophantine with exponent* μ, *then* $\mathscr{R}(t)$ *is of polynomial growth, with an exponent strictly larger than*

$$4\mu + 3 + \nu\frac{2\mu+4}{|\log\rho|}.$$

Theorem 5(b)**.** [DIOP2] *Consider a dynamical source of the* Good *Class, with a possibly infinite denumerable alphabet, with a contraction ratio* $\rho < 1$ *and a pression function* $s \mapsto L(s)$[3]. *Consider the real* ν_1 *defined from the pressure function by the two equations*

$$(1-\sigma_1)L'(\sigma_1) + L(\sigma_1) = \log\rho, \quad \nu_1 = -L'(\sigma_1).$$

If there exists a pair $\{h,k\}$, *with* $\nu_0 = \max\{c(h), c(k)\}$, *for which* $c(h,k)$ *is diophantine with exponent* μ, *then* $\mathscr{R}(t)$ *is of polynomial growth, with an exponent strictly larger than*

$$4\mu + 3 + \nu\frac{2\mu+4}{|\log\rho|} \qquad \textit{with } \ \nu = \max(\nu_0, \nu_1).$$

3.4. Main sets of interest. One considers triples $(\mathscr{T}, \mathscr{W}, \eta)$ formed with

(i) a subset $\mathscr{T}$ of the set $\{t \in \mathbb{R}, |t| \geq 1\}$,

(ii) a family $\mathscr{W}$ of functions, $\mathscr{W} := \{w_t,\ t \in \mathscr{T};\ w_t \in \mathscr{C}^1(\mathscr{I}\times\mathscr{I}),\ |w_t| = 1,\ ||w_t||_{(t)} \leq K\}$,

(iii) a family η of complex numbers, $\eta := \{\eta_t \in \mathbb{C},\ t \in \mathscr{T},\ |\eta_t| = 1\}$.

We consider properties which are satisfied only on subsets of the unit square $\mathscr{I}\times\mathscr{I}$ and only in an approximative way, and, for a given imaginary part t, these subsets, and the approximation will depend on t (in a polynomial way), and there are various parameters $(\alpha,\beta,\gamma,\delta)$ for the possible exponents.

One lets $n(\beta,t) := \lceil \beta \log|t| \rceil$, $n(\theta,t) := \lceil \theta \log|t| \rceil$ and considers the following subsets of $\mathscr{H}^\star$,

$$\mathscr{H}(t,\theta) := \mathscr{H}^{n(\theta,t)}, \qquad \mathscr{H}(t,\beta,\delta) := \left\{ h \in \mathscr{H}^{n(\beta,t)};\quad \min\{|h'(x)|, x \in \mathscr{I}\} \geq \frac{1}{t^\delta} \right\}.$$

The following subsets of $\mathscr{I}\times\mathscr{I}$ are called "fundamental unions"

$$\mathscr{I}(t,\beta,\delta) := \bigcup_{h\in\mathscr{H}(t,\beta,\delta)} h(\mathscr{I})\times h(\mathscr{I}), \qquad \mathscr{I}(t,\beta,\delta,\theta) := \bigcup_{\substack{h\in\mathscr{H}(t,\beta,\delta)\\ \ell\in\mathscr{H}(t,\theta)}} h\circ\ell(\mathscr{I})\times h\circ\ell(\mathscr{I}).$$

[3]the pression is the logarithm of the dominant eigenvalue $\lambda(s)$

In the proof, there are various subsets which intervene: Subsets $\mathscr{A}$, related to the notion of "almost eigenfunctions" – subsets $\mathscr{C}$ relared to the notion of "almost invariant functions" – subsets $\mathscr{E}$ which approximate subsets $\mathscr{A} \setminus \mathscr{C}$ – subsets $\mathscr{B}$ related to the behaviour of the iterate of the operator – Subsets $\mathscr{F}$ related to the growth of the quasi-inverse of the secant operator. The final subset of interest is the subset $\mathscr{F}$, and the other ones form a chain of subsets which will be compared to $\mathscr{F}$ in the proof. The first three ones involve the approximate subset $\mathscr{I}(t,\beta,\delta,\theta)$.

The set $\mathscr{A}(\alpha,\beta,\delta,\theta)$ gathers all the reals t for which there exists a pair (w_t,η_t) that satisfies,

$$|\mathbb{M}_t^{n(\beta,t)}[w_t](x,y) - \eta_t w_t(x,y)| \le \frac{1}{t^\alpha}, \qquad \text{for any } (x,y) \in \mathscr{I}(t,\beta,\delta,\theta). \tag{15}$$

The set $\mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ gathers all the reals t for which there exists a pair (w_t,η_t) that satisfies

$$|\eta_t^{k_0} - 1| \le \frac{1}{t^\gamma}. \qquad |\mathbb{M}_t^{n(\beta,t)}[w_t](x,y) - \eta_t w_t(x,y)| \le \frac{1}{t^\alpha}, \qquad \text{for any } (x,y) \in \mathscr{I}(t,\beta,\delta,\theta). \tag{16}$$

The set $\mathscr{E}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ gathers the reals t for which there exists a pair (w_t,η_t) that satisfies

$$|\eta_t^{k_0} - 1| > \frac{1}{t^\gamma}, \qquad |\mathbb{M}_t^{n(\beta,t)}[w_t](x,y) - \eta_t w_t(x,y)| \le \frac{1}{t^\alpha}, \qquad \text{for any } (x,y) \in \mathscr{I}(t,\beta,\delta,\theta). \tag{17}$$

The inclusion $\mathscr{A}(\alpha,\beta,\delta,\theta) \setminus \mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0) \subset \mathscr{E}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ holds.

Let ψ be the invariant function of $\mathbb{H}_1$. The set $\mathscr{B}(\alpha,\beta,\theta)$ gathers all the reals t for which there exists u_t, with $||u_t||_{(t)} \le 1$, that satisfies, for any $n \le 3n(\beta,t)$ and any $(x,y) \in \mathscr{I}(t,\theta)$,

$$|\mathbb{H}_{1+it}^n[\psi u_t](x,y)| \ge \psi(x,y)\left(1 - \frac{1}{t^\alpha}\right).$$

The set $\mathscr{F}(\alpha)$ gathers the reals t for which the quasi-inverse is of polynomial growth with exponent α

$$\mathscr{F}(\alpha) := \{t,\ \mathscr{R}(t) \le t^\alpha\}.$$

We wish to prove that there exists α for which $\mathscr{F}^c(\alpha)$ is bounded.

3.5. Relation between diophantine properties, subsets $\mathscr{A}$ and $\mathscr{C}$. There are two main results, described in Lemma 0 (with subset $\mathscr{A}$) and Lemma 1 (with subset $\mathscr{C}$).

Lemma 0. *Consider a triple $(h,k,\ell) \in \mathscr{H}^3$ and a real $\alpha > 1$. If there exists a triple (β,δ,θ) with*

$$\frac{\delta}{\beta} > \max\{c(h),c(k),c(\ell)\}, \tag{18}$$

for which $\mathscr{A}(\alpha,\beta,\delta,\theta)$ is unbounded, then $c(h,k,\ell)$ has an irrationality exponent at least equal to $\alpha - 1$. If $c(h,k,\ell)$ is diophantine with exponent μ, then for any 4-uple $(\alpha,\beta,\delta,\theta)$ avec $\alpha > \mu + 1$, et (β,δ) which satisfies (18), the subset $\mathscr{A}(\alpha,\beta,\delta,\theta)$ is bounded.

Lemma 1. *Consider a pair $(h,k) \in \mathscr{H}^2$ and a real $\mu > 1$. If there exists a 6–uple $(\alpha,\beta,\gamma,\delta,\theta,k_0)$ with $\min(\alpha,\gamma) = \mu$, and*

$$\frac{\delta}{\beta} > \max\{c(h),c(k)\}, \tag{19}$$

for which the subset $\mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ *is unbounded, then* $c(h,k)$ *has an irrationality exponent at least equal to* $\mu-1$.
If $c(h,k)$ *is diophantine with exponent* μ*, then, for each 6–uple* $(\alpha,\beta,\gamma,\delta,\theta,k_0)$ *with* $\min(\alpha,\gamma)>\mu+1$ *et* (β,δ) *that satisfies (19), the subset* $\mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ *is bounded.*

In the following of the proof, we use the notion of weak inclusion between two subsets $\mathscr{L}$ et $\mathscr{M}$ de $\mathbb{R}$. The subset $\mathscr{L}$ is said to be weakly included in $\mathscr{M}$ [this is denoted by $\mathscr{L}\,\tilde{\subset}\,\mathscr{M}$] if there exists $t_1\in\mathbb{R}$ for which $\mathscr{L}\cap[t_1,+\infty[\subset\mathscr{M}\cap[t_1,+\infty[$.

3.6. Relations between subsets $\mathscr{A}$ and $\mathscr{F}$. Lemma 2 compares subsets $\mathscr{B}$ and $\mathscr{A}$ whereas Lemma 3 compares subsets $\mathscr{B}$ and $\mathscr{F}$. Lemmas 2 and 3 are summarized in Lemma 4 which compares subsets $\mathscr{A}$ and $\mathscr{F}$. Lemmas 0 and 4 together prove Theorem 5(a).

Lemma 2. *For any 4-uple* $(\alpha,\beta,\delta,\theta)$ *that satisfies* $\beta|\log\rho|\geq\alpha+1$, *the weak inclusion* $\mathscr{B}(\alpha_1,\beta,\theta)\,\tilde{\subset}\,\mathscr{A}(\alpha,\beta,\delta,\theta)$ *holds, for any* $\alpha_1>2\alpha+\delta$.

Lemma 3. *For any triple* (α_1,β,θ) *that satisfies* $\theta|\log\rho|\geq\alpha_1+1$, *the weak inclusion* $\mathscr{B}^c(\alpha_1,\beta,\delta)\,\tilde{\subset}\,\mathscr{F}(\alpha_2)$ *holds for any* $\alpha_2>2\alpha_1+1$

Lemma 4. *For any 4-uple* $(\alpha,\beta,\delta,\theta)$ *that satisfies* $\beta|\log\rho|\geq\alpha+1,\quad\theta|\log\rho|\geq2\alpha+\delta+1$, *the weak inclusion* $\mathscr{F}^c(\alpha_2)\,\tilde{\subset}\,\mathscr{A}(\alpha,\beta,\delta,\theta)$ *holds, for any* $\alpha_2>4\alpha+2\delta+1$

3.7. Relation between subsets $\mathscr{E}$ and $\mathscr{F}$. This relation is described in Lemma 5.

Lemma 5. *One considers the logarithm* $L(s)$ *of the dominant eigenvalue* $\lambda(s)$ *of the operator* $\mathbb{H}_s$ *and the real* ν_1 *defined from the pressure function by the two equations*

$$(1-\sigma_1)L'(\sigma_1)+L(\sigma_1)=\log\rho,\;\;\nu_1=-L'(\sigma_1).$$

For any 5–uple $(\alpha,\beta,\gamma,\delta,\theta)$ *which satisfies the relations*

$$\alpha>\gamma,\qquad\beta\geq\frac{\alpha+1}{|\log\rho|},\qquad\frac{\delta}{\beta}>\nu_1,$$

there exists an integer k_0 *for which the weak inclusions*
$\mathscr{E}(\alpha,\beta,\gamma,\delta,\theta,k_0)\,\tilde{\subset}\,\mathscr{F}(2\alpha)$, *and thus* $\mathscr{F}^c(2\alpha)\,\tilde{\subset}\,\mathscr{A}^c(\alpha,\beta,\delta,\theta)\cup\mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0)$ *hold.*

3.8. Relation between subsets $\mathscr{C}$ and $\mathscr{F}$. One gathers the conclusions of Lemmas 4 and 5 in Lemma 6. Lemmas 1 et 6 together prove Theorem 5(b).

Lemma 6. *One considers the logarithm* $L(s)$ *of the dominant eigenvalue* $\lambda(s)$ *of the operator* $\mathbb{H}_s$ *and the real* ν_1 *defined from the pressure function by the two equations*

$$(1-\sigma_1)L'(\sigma_1)+L(\sigma_1)=\log\rho,\;\;\nu_1=-L'(\sigma_1).$$

For any 5–uple $(\alpha,\beta,\gamma,\delta,\theta)$ *which satisfies the relations*

$$\alpha>\gamma,\qquad\beta\geq\frac{\alpha+1}{|\log\rho|},\qquad\theta\geq\frac{2\alpha+\delta+1}{|\log\rho|},\qquad\frac{\delta}{\beta}>\nu_1,$$

there exists an integer k_0 *for which the weak inclusion*

$$\mathscr{F}^c(2\alpha+\delta)\,\tilde{\subset}\,\mathscr{C}(\alpha,\beta,\gamma,\delta,\theta,k_0)\text{ holds.}$$