Discrete Mathematics and Theoretical Computer Science **6**, 2004, 283–314

Statistical properties of Markov dynamical sources: applications to information theory

F. Chazal and V. Maume-Deschamps

Université de Bourgogne B.P. 47870 21078 Dijon Cedex FRANCE
{fchazal, vmaume}@u-bourgogne.fr

received Jul. 1, 2003, *revised Nov. 28, 2003*, *accepted May. 6, 2004*.

In (V1), the author studies statistical properties of words generated by dynamical sources. This is done using generalized Ruelle operators. The aim of this article is to generalize the notion of sources for which the results hold. First, we avoid the use of Grothendieck theory and Fredholm determinants, this allows dynamical sources that cannot be extended to a complex disk or that are not analytic. Second, we consider Markov sources: the language generated by the source over an alphabet $\mathcal{M}$ is not necessarily $\mathcal{M}^*$.

Keywords: dynamical sources, information theory, transfer operator, markov sources

1 Introduction

Statistical properties of words describe the asymptotic behavior (or laws) of parameters such as "most probable prefixes," "coincidence probability" etc. These analyses have many applications in analysis of algorithms, pattern matching, study of tries, optimization of algorithms... Of course, statistical properties of words heavily depend on the way the words are produced.

In information theory contexts, a source is a mechanism which emits symbols from an alphabet $\mathcal{M}$ (finite or infinite countable) to produce (infinite) words. The two "classical" simpler models are memoryless sources where each symbol is emitted independently of the previous ones and Markov chains where the probability for a symbol to be emitted depends on a bounded part of the past. Sources encountered in practical situations are usually complex mechanisms, and one needs general models to study the statistical properties of emitted words (e.g. the distribution of the prefixes of the same fixed length) and the parameters of the sources (e.g. entropy). In (V1), B. Vallée introduces a model of *probabilistic dynamical source* which is based upon dynamical systems theory. It covers classical sources models (that is memoryless, some Markov chains) and some other processes with unbounded dependency on past history. A probabilistic dynamical source consists of two parts: a dynamical system on the unit interval $[0,1]$ representing the mechanism which produces words and a probability measure. More precisely, a dynamical source is defined by:

(a) A finite or infinite countable alphabet $\mathcal{M}$.

(b) A topological partition of $I := [0,1]$ into disjoint open intervals I_m, $m \in \mathcal{M}$, i.e. $\overline{I} = \bigcup_{m\in\mathcal{M}} \overline{I}_m$.

(c) A mapping σ which is constant and equal to m on each I_m.

(d) A mapping T whose restriction to each $\overline{I_m}$ is a C^2 bijection from $\overline{I_m}$ to $T(\overline{I_m}) = J_m$.
Let f be a probability density on I. Words on the alphabet $\mathcal{M}$ are produced in the following way: first, $x \in I$ is chosen at random with respect to the probability of density f, second, the infinite word $M(x) = (\sigma(x), \sigma(Tx), \cdots, \sigma(T^k x), \cdots)$ is associated to x.
The main tool in the analysis of such sources is a "generating operator," *the generalized Ruelle operator* depending on a complex parameter s and acting on a suitable Banach space. To derive results about the source, this operator must have a simple dominant eigenvalue $\lambda(s)$ defined for s in a neighborhood of the real axis. Thus some additional hypotheses on the mapping T are needed. For example, in the context of (V1), branches $T_{|I_m}$ need to be real analytic with a holomorphic extension to a complex neighborhood of $[0,1]$, complete (i.e. $T(I_m) = I$) and they need to satisfy a bounded distortion property (see (C,M,V)). Such sources produce the set $\mathcal{M}^*$ of all the words on the alphabet $\mathcal{M}$. The analyticity of T allows to use the powerful Grothendieck theory and Fredholm theory on operators on spaces of holomorphic functions. The aim of this work is to prove that the hypothesis of analyticity and completeness may be relaxed. We extend the results of (V1) to a larger class of *P-Markov sources* (see Definition 1). Our class contain various classes of examples of interest such as Markov sources on a finite alphabet, Markov sources with finitely many images or Markov sources with large images (see Section 2.1 and Figure 1).

The dominant eigenvalue function $s \to \lambda(s)$ is involved in all the results of the paper. First of all,

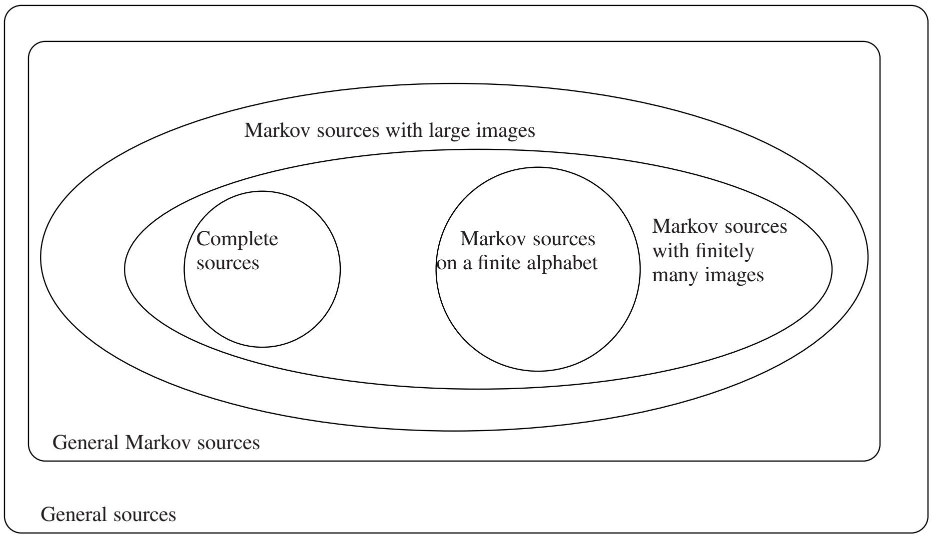

Fig. 1: Geometry of sources

parameters of the source $\mathcal{S}$ like entropy $h(\mathcal{S})$ or coincidence probability $c_b(\mathcal{S})$ depend on this function:

$$h(\mathcal{S}) = -\lambda'(1) \quad \text{and} \quad c_b(\mathcal{S}) = \lambda(b).$$

Secondly, statistical properties of word emitted by the source depend on $\lambda(s)$:

- the number $B(x)$ of finite words whose probability is at least x, satisfies

$$B(x) \simeq \frac{1}{\lambda'(1)x}$$

  if the source is not conjugate to some source with affine branches.

- let $\ell_k(x)$ be the probability of words having the same prefix of length k as x. This random variable follows asymptotically a log-normal law provided that the function $s \to \log \lambda(s)$ is not affine.

- the random variable $C(x,y)$ which is the length of the longest common prefix of the two words associated to $x, y \in [0,1]$ follow asymptotically a geometric law with ratio equal to $\lambda(2)$ if the x and y are drawn independently.

These results, proven by B. Vallée for holomorphic dynamical sources, remain valid in our setting and are explicitly stated in the following main theorem. Before stating the main theorem, let us recall that two dynamical systems T, $\widetilde{T} : I \to I$ are conjugate if there exists an homeomorphism g of I such that $\widetilde{T} = g \circ T \circ g^{-1}$. Roughly speaking, from a measurable dynamical point of view, if g is piecewise C^1 the systems are the same.

Theorem *Consider a P-Markov source and f a density of probability, which is bounded, Lipschitz on each I_m with uniformly bounded Lipschitz constant. Then there exists an analytic function $s \to \lambda(s)$ on a complex neighborhood of $\mathcal{R}(s) \geq 1$ such that:*

- *Either there exist $\alpha > 1$ and a sequence of integers $(k_m)_{m\in\mathcal{M}}$ such that the map T is conjugated to a piecewise affine map with slopes α^{k_m} on I_m, with the conjugacy C^{1+Lip} on each I_m. In that case, there exists A, B such that*

$$\frac{A}{x} \leq B(x) \leq \frac{B}{x},$$

  or

$$B(x) \simeq \frac{-1}{\lambda'(1)x}.$$

- *If $\lambda''(1) - \lambda'(1)^2 \neq 0$ then the variable $\log \ell_k$ follows asymptotically a normal law. Moreover $\lambda''(1) - \lambda'(1)^2 = 0$ if and only if the map T is conjugated to a piecewise affine map with slopes all equal, the conjugacy is C^{1+Lip} on each I_m.*

- *The variable C follows asymptotically a geometric law with ratio equal to $\lambda(2)$ if the x and y are drawn independently.*

As an immediate corollary we can give an answer to Conjecture 2 of (V1).

Corollary *Exceptional sources are those for which there exist $\alpha > 1$ and a sequence of integers $(k_m)_{m\in\mathcal{M}}$ such that the map T is conjugated to a piecewise affine map (not necessarily complete) with slopes α^{k_m} on I_m, with the conjugacy C^{1+Lip} on each I_m.*

As a consequence of the proof of main theorem, we solve Conjecture 1 of (V1) (see Remark 4).
Let us quickly present the strategy underlying the proof of the previous theorem. Important objects involved in the analysis of the sources are *fundamental intervals*: given a prefix h of length $k \in \mathbb{N}$, the set of

words starting with this prefix is an interval in $[0,1]$, the fundamental interval associated to h. Its measure (with respect to the probability density f) is denoted by u_h. It is not difficult to prove that all the studied quantities can be expressed in terms of the Dirichlet series of the fundamental measures:

$$\Lambda_k(F,s) = \sum_{h\in\mathcal{L}_k} u_h^s \text{ and } \Lambda(F,s) = \sum_{k\geq 0} \Lambda_k(F,s)$$

where $\mathcal{L}_k$ is the set of prefixes of length k (lemma 2.1). For P- Markov sources, these series define holomorphic functions of the variable s which admit a meromorphic extension to a half plane. Next we prove that these series can be expressed in terms of the generalized Ruelle operator. A careful study of spectral properties of Ruelle operators is then used to describe the singularities of Dirichlet series. Finally, parameters of the source are derived by mean of "classical" techniques: Tauberian theorem and Mellin transforms. This last part being exactly the same as in (V1), is not done in this paper. The reader is referred to B. Vallée's paper.
Let us mention that previous strategy initially developed by B. Vallée also has various important applications in the area of analysis of algorithms (especially for arithmetic algorithms), see (V2), (V3), (V4) for example.
At last, an important application of the asymptotic behavior of the parameters of P-Markov sources is the

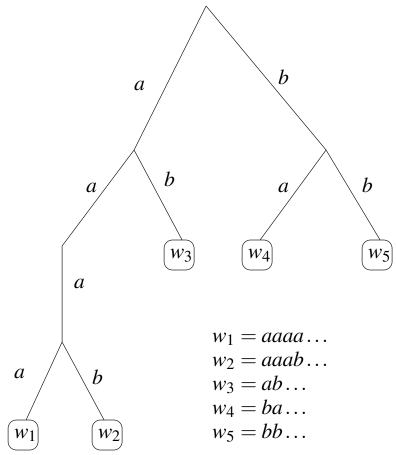

Fig. 2: An example of trie

analysis of trie (or digital tree) structure. Tries are tree data structures widely used in order to implement a search in a dictionary. They are constructed from a finite set $X = \{w_1,\cdots,w_n\}$ of words independently generated by a source. The nodes of the trie are used to manage the search in the dictionary, and each leaf contains a single word of the dictionary.
Formally, given a finite alphabet $\mathcal{M} = \{a_1,\ldots,a_r\}$, the trie associated to X is defined recursively by

$$\text{trie}(X) = \langle \text{trie}(X\setminus a_1),\ldots,\text{trie}(X\setminus a_r)\rangle,$$

where $X\setminus a_i$ is the subset of X consisting of words which begin with a_i with their first symbol a_i removed. The recursion is halted as soon as X contains less than 2 elements (see figure 2). We are concerned with

the standard parameters of trees: for example, size, path length, height.
The structure of tries have been intensively studied in the setting of independent sources (see (Sz) for example). The analysis of trie structures has been done recently in the setting of complete holomorphic sources by J. Clément, P. Flajolet and B. Vallée (C,F,V), (C) : roughly speaking, the expected values of size, path-length and height of tries can be expressed in terms of fundamental measures of the source and of Dirichlet series of fundamental measures. Thus the asymptotic behavior of these parameters is deduced from the spectral properties of some generalized Ruelle operators related to the source: some of these operators are defined over Banach spaces of functions of 4 variables. The definitions and spectral properties of these operators immediately extend to our setting.

Theorem *Let $\mathcal{S}$ be a P-Markov source. Denote by $S^{[n]}, P^{[n]}, H^{[n]}$ the size, the path-length and the height of a trie constructed over n independently drawn words of $\mathcal{S}$. The asymptotic expected value (when $n \to \infty$) of these parameters is given by*

$$\mathbf{E}[S^{[n]}] \simeq \frac{n}{h(\mathcal{S})} \qquad \mathbf{E}[P^{[n]}] \simeq \frac{n \log n}{h(\mathcal{S})} \qquad \mathbf{E}[H^{[n]}] \simeq \frac{\log n}{2|\log c(\mathcal{S})|}$$

where $h(\mathcal{S})$ is the entropy of the source and $c(\mathcal{S})$ is the coincidence probability of $\mathcal{S}$.

The paper is organized as follows. In section 2, we give precise definitions and statement of results. In section 3, we analyze the parameters of the source assuming some spectral properties of generalized Ruelle operators associated to our sources. In section 2.1 we consider some general classes of systems that satisfy our hypothesis and give some specific examples (in particular we exhibit a source that satisfy our hypothesis but that does not admit a complex extension). Finally, section 4 contains the proof of the spectral properties.

Acknowledgments: We are grateful to B. Vallée, P. Flajolet and J. Clément for interesting us in the theory of dynamical sources and for fruitful discussions. Many of these discussions were made possible thanks to a partial financial support of ALEA project.

2 Dynamical sources, intrinsic parameters and transfer operators

The following definition of dynamical sources extend B. Vallée's one. We try to give the minimal conditions ensuring that the generalized Ruelle operator associated to such a source is quasi compact on a "natural" Banach space. We call these sources *P-Markov dynamical sources* (for positive Markov dynamical sources).

Definition 1 *A dynamical P-Markov source is defined by the four following elements :*

(a) *An alphabet $\mathcal{M}$, finite or infinite countable.*

(b) *A topological partition of $I := [0,1]$ with disjoint open intervals I_m, $m \in \mathcal{M}$, i.e. $\overline{I} = \bigcup_{m \in \mathcal{M}} \overline{I}_m$, $I_m =]a_m, b_m[$.*

(c) *A mapping σ which is constant and equal to m on each I_m.*

(d) *A mapping T whose restriction to each $\overline{I_m}$ is a C^2 bijection from $\overline{I_m}$ to $T(\overline{I_m}) = J_m$. Let $h_m : J_m \to I_m$ be the local inverse of T restricted to $\overline{I_m}$. The mappings h_m satisfy the following conditions:*

(d1) **Contracting.** *There exist $0 < \eta_m \leq \delta_m < 1$ for which $\eta_m \leq |h'_m(x)| \leq \delta_m$ for $x \in J_m$.*

(d2) *There exists* $\gamma < 1$ *such that for* $\mathcal{R}(s) > \gamma$, *the series* $\sum_{m\in\mathcal{M}} \mathbf{1}_{J_m}(x)\delta_m^s$ *converge uniformly for* $x \in I$ *and* $\sum_{m\in\mathcal{M}} |I_m|^s$ *converges.*

(d3) **Bounded distortion.** *There exists a constant* $A < +\infty$ *such that for all* $m \in \mathcal{M}$ *and all* $x, y \in J_m$, $|h_m''(x)/h_m'(y)| < A$.

(d4) **Markov property.** *Each interval* J_m *is union of some of the* I_k*'s.*

(d5) **Positivity.** *See Condition 1 below.*

Remark 1 *(see the definition of operators* G_s *in section 2.2) The first part of condition (d2) is sufficient to have that the sum defining* G_s_Id *converges uniformly. Because the source is not necessarily complete, it does not imply the second part of condition (d2).*
Condition (d5) is a bit stronger than (d2), it implies that for all $m \in \mathcal{M}$, *there exists* $N \in \mathbb{N}$ *such that:*

$$\inf_{x\in I} \mathbf{G}_s^N[\mathbf{1}_{I_m}](x) > 0.$$

Remark 2 *If the alphabet* $\mathcal{M}$ *is infinite then Condition (d2) is equivalent to:*

$$\lim_{\substack{|Q|\to\infty \\ Q_{finite}}} \sup_{Q\subset\mathcal{M}} \sup_{x\in I} \sum_{\substack{m\in\mathcal{M} \\ m\notin Q}} \delta_m^s \mathbf{1}_{J_m}(x) = 0. \tag{2.1}$$

If the alphabet is finite then Condition (d2) is always satisfied.

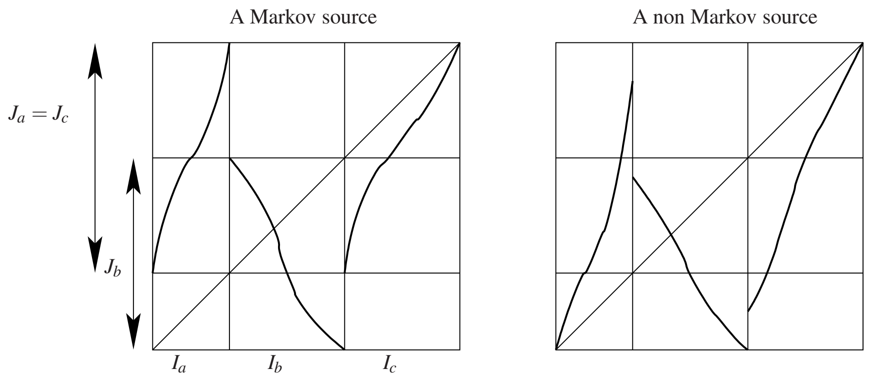

Fig. 3: Markov and non Markov sources

Such a source produces words on the alphabet $\mathcal{M}$, to each $x \in I$ we associate the infinite word

$$M(x) = (\sigma(x), \sigma(Tx), \cdots, \sigma(T^k x), \cdots).$$

For $k \in \mathbb{N}$, the kth prefix of $M(x)$ is

$$P_k(x) = (\sigma(x), \sigma(Tx), \cdots, \sigma(T^{k-1}x)).$$

We denote by $\mathcal{L}_k$ the subset of $\mathcal{M}^k$ of prefixes of length k that may be produced by the dynamical source. Remark that in our setting, $\mathcal{L}_k$ may be a strict subset of $\mathcal{M}^k$. For example in Figure 3, the word bc does not belong to $\mathcal{L}_2$. In the following, each element of $\mathcal{L}_k$ will be identified with an inverse branch of T^k of the form $h = h_{m_1} \circ \cdots \circ h_{m_k}$, $m_i \in \mathcal{M}$. Let J_h be the definition interval of $h \in \mathcal{L}_k$ and $I_h = h(J_h) = [a_h, b_h]$ *the fundamental interval of* h. If $h = h_{i_k} \circ \cdots \circ h_{i_1} \in \mathcal{L}_k$ then because of the Markov property, $J_h = J_{i_1}$. Define also $\eta_h = \inf_{x \in J_h} |h'(x)|$, and $\delta_h = \sup_{x \in J_h} |h'(x)|$.
Remark that P- Markov sources are a generalization of memoryless and classical Markov sources. Indeed, if the inverse branches h_m are affine (or equivalently if h'_m is constant) and complete (i.e. $J_m = I$) then the symbols emitted $\sigma(x)$, $\sigma(Tx)$ ... are independent (i.e. the source is memoryless). If the inverse branches are affine but not complete then the symbols emitted $\sigma(x)$, $\sigma(Tx)$ ... form a Markov chain (see Figure 4).

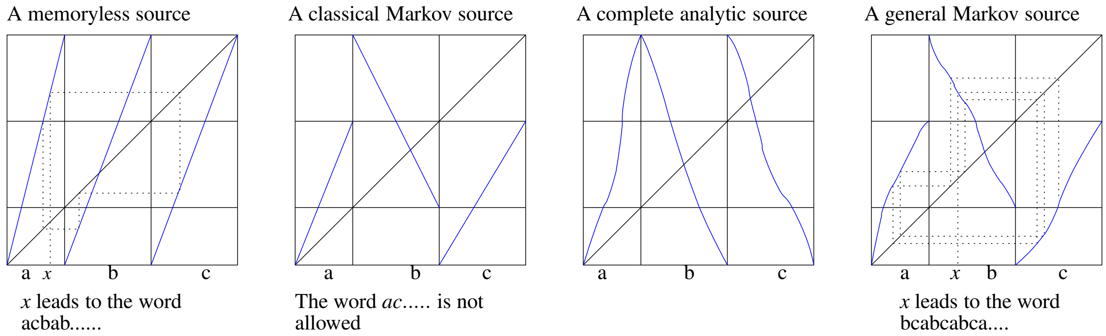

Fig. 4: Memoryless and classical Markov sources

We are now in position to express the positivity condition (d5).

Condition 1 *For all* $m \in \mathcal{M}$*, for all* $s > \gamma$*, there exists* $N \in \mathbb{N}$ *such that*

$$\inf_{x \in I} \sum_{h \in \mathcal{L}_N} \eta_h^s \, \mathbf{1}_{I_m}(h(x)) \mathbf{1}_{J_h}(x) > 0. \tag{2.2}$$

This condition is related to the aperiodicity condition of classical Markov chains. Indeed in the context of Markov (infinite) chains on an alphabet $\mathcal{M}$, let P be the (infinite) transition matrix. Then for $s = 1$, Condition (2.2) is equivalent to the following:
For all $m \in \mathcal{M}$, there exists $N \in \mathbb{N}$ such that the infimum of the coefficients of the mth column of the matrix P^N is strictly positive. If the alphabet $\mathcal{M}$ is finite this is equivalent to: there exists $N \in \mathbb{N}$ such that all the coefficients of the matrix P^N are strictly positive (i.e. the Markov chain is aperiodic). This point of view is developed in section 2.1.1 below.
Let us give some examples of sources satisfying our hypothesis.

2.1 Examples of P-Markov sources.

It is straightforward that complete holomorphic sources with bounded distortion ((V1), (C,M,V)) are P-Markov dynamical sources.

2.1.1 Some examples.

Let us give some large classes of sources satisfying our hypothesis. The simplest class is given by finite aperiodic Markov maps. Let us recall that a Markov map (i.e. a dynamical system satisfying (d4) is strongly aperiodic if there exists $M \in \mathbb{N}$ such that for any $i, j \in \mathcal{M}$, for any $n \geq M$,

$$T^{-n}I_j \cap I_i \neq \emptyset.$$

The strong aperiodicity condition is natural in the context of Markov maps (in some sense it means that the systems is not decomposable). It may be rewritten in terms of inverse branches as: there exists $M \in \mathbb{N}$ such that for $n \geq M$, for any $i, j \in \mathcal{M}$, there exists $h \in \mathcal{L}_n$ with $I_i \subset J_h$ and $h(J_h) \subset I_j$. Let us show that it suffices to ensure (d5) if the alphabet is finite, if the number of images is finite or if the system has large branches.

Example 1 *If $\mathcal{M}$ is finite and the system is strongly aperiodic then it defines a P-Markov.*
Indeed, the only point to verify is (d5). The aperiodicity condition implies that for all $n \geq M$, all $x \in I$ and $m \in \mathcal{M}$, there exists $h \in \mathcal{L}_n$ with $x \in J_h$ and $I_h \subset I_m$. Thus we have: for $m \in \mathcal{M}$, $x \in I$, $n \geq M$,

$$\sum_{\substack{h\in\mathcal{L}_n\\ x\in J_h}} \eta_h^s \mathbf{1}_{I_m}(hx) \geq \inf_{h\in\mathcal{L}_n} \eta_h^s.$$

Remark that Markov chains on a finite alphabet may always be obtained from an affine dynamical source. Thus, aperiodic Markov chains are P-Markov sources.

Example 2 *If the set $\{J_m \ / \ m \in \mathcal{M}\}$ is finite and the system is strongly aperiodic then it defines a P-Markov source provided (d2) and (d3) are satisfied.*
Indeed, let J_{i_1}, ..., J_{i_k} be the images of the system. The strong aperiodicity condition implies that for all $n \geq M$, all $m \in \mathcal{M}$ and all $j = 1, \ldots, k$, there exits $h_{i_j} \in \mathcal{L}_n$ such that $h_{i_j}(J_{i_j}) \subset I_m$. Now,

$$\sum_{\substack{h\in\mathcal{L}_n\\ x\in J_h}} \eta_h^s \mathbf{1}_{I_m}(hx) \geq \inf_{j=1,\ldots,k} \eta_{h_{i_j}}^s.$$

We would say that a source has large images if

$$\inf_{m\in\mathcal{M}} \{|J_m|\} > 0.$$

Example 3 *If the source has large images and is strongly aperiodic then it defines a P-Markov source provided (d2) and (d3) are satisfied.*
It suffices to remark that if the source has large branches and is strongly aperiodic then there exists finitely many J_m whose union is I. Then the same argument has above shows that (d5) is satisfied.

2.1.2 A P-Markov source with small branches.

For $0 < \theta < 1$, let $C = \frac{1}{1-\theta}$. Consider a partition of I into intervals I_m with $|I_m| = C\theta^m$, $m \geq 0$. Consider the piecewise affine map T such that $T(I_{2m}) = I_m$, $m \geq 1$, and $T(I_{2m+1}) = I$, $m \geq 0$. We have for all m,

$\eta_m = \delta_m$ and $\delta_{2m} = \theta^m$, $\delta_{2m+1} = \theta^{2m+1}$. Condition (d2) is satisfied. Let us show that (d5) is also satisfied. If $m = 2^k(2p+1)$, $k \geq 0$ then for all $x \in I$,

$$\sum_{\substack{h\in\mathcal{L}_{k+1}\\ x\in J_h}} \eta_h^s \mathbf{1}_{I_m}(h(x)) \geq \theta^{2p+1}\theta^{2(2p+1)}\cdots\theta^{2^k(2p+1)}.$$

This source is represented in Figure 5.

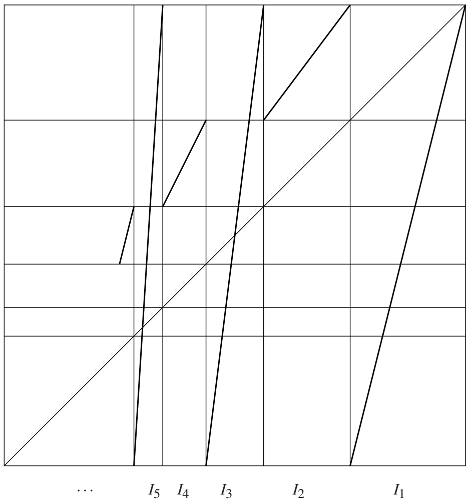

Fig. 5: A source with small branches

From now on, we have emphasized that our hypothesis allow various geometric behavior of the branches, let us now give an example showing that relaxing the holomorphic extension hypothesis of (V1) is a substantial gain.

2.1.3 A P-Markov source with no extension on a complex neighborhood.

Consider the source whose alphabet is $\mathbb{N}^*$ and inverse branches are given by

$$h_n(x) = \frac{1}{n+1} + C_n(f_n(x) - f_n(0))$$

where $f_n : [0,1] \to \mathbb{R}$ is defined by

$$f_n(x) = x - \frac{1}{x\sqrt{n}+1} - \frac{1}{\sqrt{n}(x\sqrt{n}+1)} - 2\frac{\log(x\sqrt{n}+1)}{\sqrt{n}}$$

and C_n is a constant defined by

$$C_n = (\frac{1}{n} - \frac{1}{n+1})\frac{1}{f_n(1)-f_n(0)} = \frac{1}{2(n+1)\sqrt{n}(\sqrt{n}-\log(\sqrt{n}+1))}.$$

For all $n \in \mathbb{N}$, the branch h_n maps $I = [0,1]$ onto interval $[\frac{1}{n+1}, \frac{1}{n}]$. The derivative of h_n is

$$h'_n(x) = C_n f'_n(x) = C_n \frac{x^2 + \frac{1}{\sqrt{n}}}{(x + \frac{1}{\sqrt{n}})^2}$$

and $h'_n(x) = 0$ if and only if $x = i/\sqrt[4]{n}$ or $x = -i/\sqrt[4]{n}$. Hence this source does not satisfy condition (d2) in (V1): there is no complex neighborhood of $[0,1]$ on which all the h'_n extend to a non vanishing function. Note that for any n,

$$\frac{C_n}{4n} < h'_n(x) < \delta_n \quad \text{for all } x \in [0,1]$$

with $\delta_n = C_n\sqrt{n}$ and for n sufficiently large,

$$\delta_n < \frac{1}{4n\sqrt{n}} \quad \text{for all } x \in [0,1].$$

It follows that there exists $\gamma < 1$ such that the series $\sum_{n\in\mathbb{N}^*} \delta_n^s$ converges on $\mathcal{R}(s) > \gamma$.
Now for any $y \in [0,1]$,

$$|h''_n(y)| = \frac{C_n}{\sqrt{n}} \frac{2|y-1|}{(y + \frac{1}{\sqrt{n}})^3} \leq \frac{2C_n}{\sqrt{n}} \frac{1}{\sqrt{n}^3} \frac{2C_n}{n^2}.$$

From previous inequalities, it results that for any $x, y \in [0,1]$,

$$\frac{|h''_n(y)|}{|h'_n(x)|} \leq \frac{8}{n^2\sqrt{n}} \leq 8$$

so that the source is a P-Markov dynamical source.

2.2 Intrinsic parameters and transfer operators.

Recall that a function f on a metric space X is *Lipschitz* if there exists $L \geq 0$ such that for all $x, y \in X$,

$$|f(x) - f(y)| \leq Ld(x,y).$$

The smallest constant L satisfying this property is called *the Lipschitz constant of f*.
The following definition introduces the notion of fundamental measures and the main parameters of the source ((V1)).

Definition 2 Fundamental measures and parameters of the source
Let $f > 0$ be a bounded, Lipschitz on each I_m with bounded Lipschitz constants, probability density on I and F its associated distribution function. The fundamental measures are:

$$u_h = |F(a_h) - F(b_h)|, \; h \in \bigcup_{k\in\mathbb{N}^*} \mathcal{L}_k.$$

For $b > \gamma$, denote by $c_b(F)$ the b-coincidence probability:

$$c_b(F) = \lim_{k\to\infty} \left(\sum_{h\in\mathcal{L}_k} u_h^b \right)^{\frac{1}{k}}.$$

Let $B(x)$ be the number of fundamental intervals whose measure is at least equal to x.
For $k \in \mathbb{N}^*$, ℓ_k is the random variable defined by $\ell_k(x) = u_h$ if $x \in I_h$, $h \in \mathcal{L}_k$.
Finally, C is the random variable on $I \times I$, defined by

$$C(x,y) = \max\{k \in \mathbb{N} \;/\; P_k(x) = P_k(y)\}.$$

The Dirichlet series of fundamental measures are:

$$\Lambda_k(F,s) = \sum_{h \in \mathcal{L}_k} u_h^s \text{ and } \Lambda(F,s) = \sum_{k \geq 0} \Lambda_k(F,s).$$

Lemma 2.1 *(V1) The parameters of the source may be expressed in terms of Dirichlet series of fundamental measures:*

$$c_b(F) = \lim_{k \to \infty} (\Lambda_k(F,b))^{\frac{1}{k}}.$$

$$\Lambda(F,s) = s \int_0^\infty B(x) x^{s-1} dx.$$

$$\mathbb{E}(\ell_k^s) = \Lambda_k(F,s+1).$$

$$\mathbb{P}(C \geq k) = \Lambda_k(F,2) \;\; \textit{and} \;\; \mathbb{E}(C) = \Lambda(F,2).$$

In (V1), the asymptotic behavior of Dirichlet series is obtained from spectral properties of generalized Ruelle operators associated to some analytic sources satisfying $\mathcal{L}_k = \mathcal{M}^k$ for all k. In this paper, we prove that generalized Ruelle operators associated to P-Markov sources have the same dominant spectral properties. We relate Dirichlet series to these operators in our setting. So the analysis on the parameters of the source remain valid.
Generalized Ruelle operators $\mathbf{G}_s$ involve secants of inverse branches

$$H_m(u,v) := \left| \frac{h_m(u) - h_m(v)}{u - v} \right|$$

and are defined by

$$\mathbf{G}_s[\Phi](u,v) := \sum_{m \in \mathcal{M}} H_m^s(u,v) \Phi(h_m(u), h_m(v)) \mathbf{1}_{J_m \times J_m}(u,v).$$

We are going to prove that these operators are quasi compact with unique and simple dominant eigenvalue $\lambda(s)$ that coincide with the dominant eigenvalue of the "classical" Ruelle operator:

$$G_s \phi(u) := \mathbf{G}_s[\Phi](u,u) \text{ with } \Phi(u,v) = \phi(u).$$

Recall that the spectrum $Sp(P)$ of a linear operator P acting on a Banach space B is the set of complex numbers λ such that $Id - \lambda P$ in not invertible. Such a spectral value λ may be either an eigenvalue (i.e. $Id - \lambda P$ is not injective) or $Id - \lambda P$ is not surjective.The spectral radius $R(P)$ is the largest modulus of an element of $Sp(P)$. An operator P is compact if the elements of $Sp(P) \setminus \{0\}$ are eigenvalues of finite multiplicity. An operator P is quasi-compact if there exists $0 < \varepsilon < R(P)$ such that the elements of $Sp(P) \setminus B(0,\varepsilon)$ are eigenvalues of finite multiplicity. The smallest such ε is called essential spectral radius

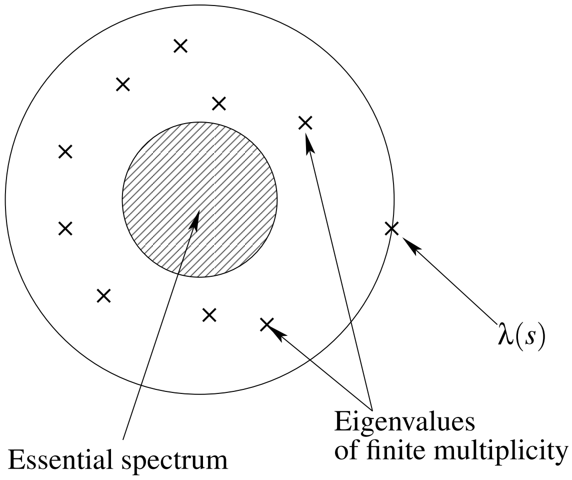

Fig. 6: Spectrum of a quasi compact operator

and $Sp(P) \cap B(0,\varepsilon)$ is called essential spectrum.

Remark that condition (d2) ensures that the operator G_s is well defined for $\mathcal{R}(s) > \gamma$ on bounded functions. Condition (d2) together with Taylor formula ensure that operators $\mathbf{G}_s$ are well defined for $\mathcal{R}(s) > \gamma$ on bounded functions.
Also, it is easy to see that:

$$\mathbf{G}_s^k \Phi(x,x') = \sum_{h \in \mathcal{L}_k} H_h^s(x,x') \Phi(h(x),h(x')) \mathbf{1}_{J_h \times J_h}(x,x'),$$

where H_h is the secant function associated to h. In our setting, the relation between Dirichlet series and Ruelle operators is given by the following proposition.

Proposition 2.2 *For all $\mathcal{R}(s) > \gamma$, $k \geq 0$,*

$$\Lambda_{k+1}(F,s) = \sum_{m \in \mathcal{M}} |a_m - b_m|^s \mathbf{G}_s^k L^s(a_m,b_m),$$

$$\textit{with} \;\; L(x,x') = \frac{|F(x) - F(x')|}{|x - x'|}.$$

Proof. – For any $m \in \mathcal{M}$, we have:

$$\begin{aligned}
\mathbf{G}_s^k L^s(a_m, b_m) &= \sum_{\substack{h \in \mathcal{L}_k \\ (a_m, b_m) \in J_h \times J_h}} H_h^s(a_m, b_m) L^s(h(a_m), h(b_m)) \\
&= \sum_{\substack{h \in \mathcal{L}_k \\ a_m \in J_h}} \frac{|F(h(a_m)) - F(h(b_m))|^s}{|a_m - b_m|^s} \\
&= \sum_{\substack{h \in \mathcal{L}_k \\ a_m \in J_h}} \frac{u_{h \circ h_m}^s}{|a_m - b_m|^s}
\end{aligned}$$

(remark that $\{h(a_m), h(b_m)\} = \{a_{h \circ h_m}, b_{h \circ h_m}\}$). Now, any $\tilde{h} \in \mathcal{L}_{k+1}$ may be uniquely written as $\tilde{h} = h \circ h_m$ for some $h \in \mathcal{L}_k$ and $m \in \mathcal{M}$. □

Our main theorem extends B. Vallée results to P-Markov dynamical sources.

Theorem 2.3 *Consider a dynamical P-Markov source. There exist $\lambda(s) > 0$, $\Phi(s) > 0$ and $0 \le \rho(s) < 1$ three analytic functions on a complex neighborhood of the half line $\{s \in \mathbb{R} \ / \ s > \gamma\}$ such that for any $k \ge 1$,*

$$\Lambda_k(F, s) = \lambda^k(s) \left(\Phi(s) + O(\rho^k(s)) \right). \tag{2.3}$$

$\lambda(s)$ is the dominant eigenvalue of G_s on a suitable functional space.
$\Lambda(F, s)$ is analytic on $\mathcal{R}(s) > 1$ and has a simple pole at $s = 1$.
The variable C follows asymptotically a geometric law.
If $\lambda''(1) - \lambda'(1)^2 \neq 0$ then the variable $\log \ell_k$ follows asymptotically a normal law. Moreover, $\lambda''(1) - \lambda'(1)^2 = 0$ if and only if the map T is conjugated to a piecewise affine map with equal slopes, the conjugacy is C^{1+Lip} on each I_m.
Either 1 is the only pole of $\Lambda(F, s)$ on $\mathcal{R}(s) = 1$, in that case

$$B(x) \simeq \frac{-1}{\lambda'(1) x},$$

or the map T is conjugated to a piecewise affine map with slopes of the form α^k, $\alpha > 1$, $k \in \mathbb{Z}$, with conjugacy C^{1+Lip} on each I_m. In that case, there exist A, B,

$$\frac{A}{x} \le B(x) \le \frac{B}{x}.$$

Theorem 2.3 is derived from dominant spectral properties of generalized real Ruelle operators. We will prove that these operators admit a unique maximal eigenvalue. To this aim, we use Birkhoff cones and projective metrics ((Bi1), (Bi2)). These techniques have been introduced in dynamical systems by P. Ferrero and B. Schmitt ((F,S)) and have been widely used by dynamicians to study Ruelle operators in many different situations. Here, we will use these techniques to prove that both operators $\mathbf{G}_s$ and G_s are quasi-compact and have a unique and simple dominant eigenvalue, for real $s > \gamma$. We will give the proofs for $\mathbf{G}_s$, the proofs for G_s may be obtained in the same way. Even for the operators G_s, our setting is not covered by previous works (see for example (Bre), (M), (Sa)).

Of course the spectral properties of the operators G_s and $\mathbf{G}_s$ depend on the space on which they act.

Because the system is not assumed to be complete (i.e. we do not assume $J_m = I$ for all $m \in \mathcal{M}$), the operators G_s and $\mathbf{G}_s$ do not act on continuous functions.
A function f is Lipschitz continuous on I_m if there exists a constant $K_m > 0$ such that for all $x, y \in I_m$,

$$|f(x) - f(y)| \le K_m |x - y|.$$

The smallest number K_m such that the above is satisfied is called the *Lipchitz constant* of f on I_m. Let $L_{pw}(I)$ be the space of functions that are bounded and Lipschitz continuous on each I_m, with the supremum of the Lipschitz constants on the I_m's finite. Denote by $\mathcal{J} \subset I \times I$ the union of all sets $I_m \times I_m$ and let $L_{pw}(\mathcal{J})$ the space of functions on $\mathcal{J}$, that are bounded and Lipschitz continuous on each $I_m \times I_m$, with bounded Lipschitz constant. In both cases, $\mathrm{Lip}(f)$ will denote the sup of the Lipschitz constants on the I_m's or on the $I_m \times I_m$'s. These spaces are endowed with the norm:

$$\|f\| = \|f\|_\infty + \mathrm{Lip}(f).$$

It is easy to see (and will in fact follow from Lemma 3.2) that G_s (resp. $\mathbf{G}_s$) acts on $L_{pw}(I)$ (resp. $L_{pw}(\mathcal{J})$).

Theorem 2.4 *For real $s > \gamma$, the operators $\mathbf{G}_s$ (resp. G_s) act on $L_{pw}(\mathcal{J})$ (resp. $L_{pw}(I)$), they are quasi compact and have a simple dominant eigenvalue. This dominant eigenvalue $\lambda(s)$ is the same for $\mathbf{G}_s$ and G_s. The corresponding eigenvectors are strictly positive and belong to $L_{pw}(\mathcal{J})$ (resp. $L_{pw}(I)$).*

Remark 3 *If the source were complete (i.e. $J_m = I$ for all m) and the density function f is C^1 on I, then we could work with spaces of C^1 functions. In that case, $\mathbf{G}_s$ acts on the space $C^1(I \times I)$ of functions that are C^1 on $I \times I$ and G_s acts on the space $C^1(I)$ of functions that are C^1 on I, they are quasi compact and have a simple dominant eigenvalue. This dominant eigenvalue $\lambda(s)$ is the same for $\mathbf{G}_s$ and G_s. The corresponding eigenvectors are strictly positive and belong to $C^1(I \times I)$ (resp. $C^1(I)$). The only change in our proof would be in the definition of the cone in section 2.4 (see Remark 8).*

We postpone the proof of Theorem 2.4 to the end of the paper (see section 4). Let us show how to use it to get Theorem 2.3.

3 Analysis of the parameters of the source

3.1 Preliminary results

The following lemma is an easy application of the derivation chain rule, (d3) and the fact that all h_m, $m \in \mathcal{M}$ are δ contractions with $\delta = \sup_{m \in \mathcal{M}} \delta_m < 1$.

Lemma 3.1 *For all $k \in \mathbb{N}^*$, for all $h \in \mathcal{L}_k$, $x, y \in J_h$,*

$$\frac{h''(x)}{h'(y)} \le \frac{A(1+A)}{1-\delta} := B.$$

Applying the integral Taylor formula at order 1 to h, the Taylor formula at order 1 to h' and Lemma 3.1 gives: for all $k \in \mathbb{N}^*$, for all $h \in \mathcal{L}_k$, $X = (x, x')$, $Y = (y, y') \in J_h \times J_h$,

$$\frac{H_h(X)}{H_h(Y)} \le 1 + d(X,Y)B, \tag{3.1}$$

$$\text{where } d(X,Y) = |x-y| + |x'-y'|. \tag{3.2}$$

The following lemma proves that the operators $\mathbf{G}_s$, $\mathcal{R}(s) > \gamma$ satisfy a "Doblin-Fortet" or "Lasota-Yorke" inequality. We are going to use a result by H. Hénnion ((H)) to conclude that they are quasi-compact for some complex s, $\mathcal{R}(s) > \gamma$. We could also use it to conclude that $\mathbf{G}_s$ are quasi-compact for real $s > \gamma$ then it would remain to prove that the dominant eigenvalue is unique and simple. This can be done "by hand" but we have preferred to give a self contained argument proving in the same time the quasi compactness and the dominant spectral property (see section 4).

Lemma 3.2 *For all s, $\mathcal{R}(s) = \sigma > \gamma$, there exists $K > 0$ such that for all $f \in L_{pw}(\mathcal{I})$, for all $n \in \mathbb{N}$,*

$$\begin{aligned} \operatorname{Lip}(\mathbf{G}_s^n f) &\leq \delta^n \|\mathbf{G}_\sigma^n \mathbf{1}\|_\infty \operatorname{Lip}(f) + K\|\mathbf{G}_\sigma^n |f|\|_\infty \\ &\leq \delta^n \|\mathbf{G}_\sigma^n \mathbf{1}\|_\infty \operatorname{Lip}(f) + K\|\mathbf{G}_\sigma^n \mathbf{1}\|_\infty \|f\|_\infty. \end{aligned} \tag{3.3}$$

Proof. $-$ Let $X = (x,x')$, $Y = (y,y')$ belong to the same $I_m \times I_m$. In that case, the sets $\{h \,/\, |h| = n$ and $X \in J_h \times J_h\}$ and $\{h \,/\, |h| = n$ and $Y \in J_h \times J_h\}$ are the same. We compute:

$$\begin{aligned} |\mathbf{G}_s^n f(X) - \mathbf{G}_s^n f(Y)| &\leq \sum_{\substack{h\in\mathcal{L}_n \\ X\in J_h\times J_h}} |H_h(X)|^\sigma |f(h(x),h(x')) - f(h(y),h(y'))| \\ &\quad + \sum_{\substack{h\in\mathcal{L}_n \\ X\in J_h\times J_h}} |f(h(Y))| H_h(Y)^\sigma \left| \left(\frac{H_h(X)}{H_h(Y)}\right)^s - 1 \right| \\ &\leq \delta^n \operatorname{Lip}(f) d(X,Y) \mathbf{G}_\sigma^n(\mathbf{1})(X) \\ &\quad + \sigma B d(X,Y)(1+Be^\sigma)\mathbf{G}_\sigma^n(|f|)(Y), \end{aligned}$$

(we have used (3.1)).
This gives the result with $K = \sigma B(1+Be^\sigma)$. □
Let us state Hénnion's theorem and show that we can apply it.

Theorem 3.3 *((H)) Let $(B, \|\cdot\|)$ be a Banach space, let $|\cdot|$ be another norm on B and Q be an operator on $(B, \|\cdot\|)$, with spectral radius $R(Q)$. If Q satisfies:*

1. *Q is compact from $(B, \|\cdot\|)$ into $(B, |\cdot|)$,*
2. *for all $n \in \mathbb{N}$, there exist positive numbers R_n and r_n such that $r = \liminf (r_n)^{\frac{1}{n}} < R(Q)$ and for all $f \in B$,*
$$\|Q^n f\| \leq R_n |f| + r_n \|f\|$$

then Q is quasi-compact and the essential spectral radius is less than r.

We will use this theorem with $B = L_{pw}(\mathcal{I})$ and $|\cdot|$ the sup norm. According to Lemma 3.2, in order to apply Theorem 3.3, we have to prove that the operators $\mathbf{G}_s$ are compact from $(L_{pw}(\mathcal{I}), \|\cdot\|)$ into $(L_{pw}(\mathcal{I}), \|\cdot\|_\infty)$. In other words, consider a sequence $(f_n)_{n\in\mathbb{N}}$, $f_n \in L_{pw}(\mathcal{I})$ with $\|f_n\| \leq 1$, we have to prove that there exists a subsequence n_k such that the sequence $(\mathbf{G}_s f_{n_k})$ converges for the sup norm $\|\cdot\|_\infty$. This will follow from remark 2.

Lemma 3.4 *For all s such that $\mathcal{R}(s) > \gamma$, $\mathbf{G}_s$ is compact from $(L_{pw}(\mathcal{J}), \|\cdot\|)$ into $(L_{pw}(\mathcal{J}), \|\cdot\|_\infty)$.*

Proof. – Let $(f_n)_{n\in\mathbb{N}}$, $f_n \in L_{pw}(\mathcal{J})$ with $\|f_n\| \le 1$, restricted to each $I_m \times I_m$ the functions f_n are uniformly equicontinuous. We may apply Ascoli's theorem on each $I_m \times I_m$ and use a diagonal principle to find a subsequence n_k such that the sequence f_{n_k} converges to some function f. Let us prove that $\mathbf{G}_s f_{n_k}$ converges uniformly to $\mathbf{G}_s f$. Denote $s = \sigma + it$, let $X = (x, x') \in \mathcal{J}$ and $\mathcal{Q}$ a finite subset of $\mathcal{M}$,

$$\begin{aligned}
&|\mathbf{G}_s f_{n_k}(X) - \mathbf{G}_s f(X)| \\
&= \left| \sum_{\substack{m\in\mathcal{M} \\ X\in J_m\times J_m}} H_m^s(X)(f_{n_k}(h_m(x), h_m(x')) - f(h_m(x), h_m(x'))\right| \\
&\le \sum_{\substack{m\in\mathcal{Q} \\ x\in J_m}} H_m^\sigma(X)\left|f_{n_k}(h_m(x), h_m(x')) - f(h_m(x), h_m(x'))\right| \\
&+ \sum_{\substack{m\notin\mathcal{Q} \\ x\in J_m}} H_m^\sigma(X)\left|f_{n_k}(h_m(x), h_m(x')) - f(h_m(x), h_m(x'))\right| \\
&\le \|\mathbf{G}_\sigma \mathbf{1}\|_\infty \sup_{\substack{m\in\mathcal{Q} \\ X\in I_m\times I_m}} |f_{n_k}(X) - f(X)| + 2 \sum_{\substack{m\in\mathcal{M},\ x\in J_m \\ m\notin\mathcal{Q}}} \delta_m^\sigma
\end{aligned}$$

(we have used that $\|f_n\|_\infty \le 1$ and Taylor equality). Fix $\varepsilon > 0$, choose $\mathcal{Q} \subset \mathcal{M}$, $\mathcal{Q}$ finite, such that $\sum_{\substack{m\notin\mathcal{Q} \\ x\in J_m}} \delta_m^\sigma < \varepsilon$, now choose k_0 such that for $k > k_0$,

$$\sup_{\substack{m\in\mathcal{Q} \\ X\in I_m\times I_m}} |f_{n_k}(X) - f(X)| < \varepsilon$$

(this can be done because the convergence is uniform on each $I_m \times I_m$ and $\mathcal{Q}$ is finite). We have:

$$|\mathbf{G}_s f_{n_k}(X) - \mathbf{G}_s f(X)| \le \varepsilon(\|\mathbf{G}_\sigma \mathbf{1}\|_\infty + 2).$$

In other words, $\mathbf{G}_s f_{n_k}$ goes to $\mathbf{G}_s f$ uniformly. □

Now the following result is a simple consequence of Theorem 3.3. For any s, $R(s)$ denotes the spectral radius of $\mathbf{G}_s$.

Proposition 3.5 *Let $\mathcal{R}(s) = \sigma > \gamma$, either $R(s) \le \delta R(\sigma)$ or $\mathbf{G}_s$ is quasi-compact. In particular, for $\sigma > \gamma$, $\mathbf{G}_\sigma$ is quasi-compact.*

Proof. – We have that $R(\sigma) = \lim_{n\to\infty} \|\mathbf{G}_\sigma^n \mathbf{1}\|_\infty^{\frac{1}{n}}$. Thus

$$\liminf (\delta^n \|\mathbf{G}_\sigma^n \mathbf{1}\|_\infty)^{\frac{1}{n}} = \delta R(\sigma).$$

The result follows. □

To conclude the proof of Theorem 2.4, it remains to prove that for real $s > \gamma$, $\mathbf{G}_s$ admits a unique simple dominant eigenvalue. We postpone this proof to section 4. Let us use Theorem 2.4 and Proposition 3.5 to obtain spectral properties of $\mathbf{G}_s$ for complex parameters s.

3.2 Spectral properties for complex parameters s and properties of Dirichlet series

For real $s > \gamma$, by Theorem 2.4, we have that for any $k \in \mathbb{N}$, $f \in L_{pw}(\mathcal{I})$,

$$\mathbf{G}_s^k f = \lambda^k(s)\Pi_s(f) + S_s^k f,$$

where Π_s is the spectral projection on the maximal eigenvalue and S_s is an operator on $L_{pw}(\mathcal{I})$ whose spectral radius strictly less than $\lambda(s)$ and such that $S_s \circ \Pi_s = \Pi_s \circ S_s = 0$. Now Proposition 2.2 gives:

$$\begin{aligned}
\Lambda_{k+1}(F,s) &= \sum_{m\in\mathcal{M}} |a_m - b_m|^s \mathbf{G}_s^k L^s(a_m,b_m) \\
&= \sum_{m\in\mathcal{M}} |a_m - b_m|^s \left(\lambda^k(s)\Pi_s(L^s)(a_m,b_m) + S_s^k L^s(a_m,b_m)\right) \\
&= \lambda^{k+1}(s)\left(\frac{\Phi(s)}{\lambda(s)} + O(\rho^k(s))\right),
\end{aligned}$$

with $\Phi(s) = \sum_{m\in\mathcal{M}} |a_m - b_m|^s \Pi_s(L^s)(a_m,b_m)$ and $\rho(s)$ the spectral radius of S_s over $\lambda(s)$. Remark that we have used that

$$\sum_{m\in\mathcal{M}} |a_m - b_m|^s$$

converges which follows from (d2). Thus we have proved (2.3) of Theorem 2.3 for real s. The fact that it holds on a complex neighborhood of $s > \gamma$ follows from perturbation theory (see for example Kato (K)). We now prove Proposition 8, Proposition 9 and Proposition 10 of (V1) in our context. Remark that her proofs are based upon Fredholm determinant theory thus we have to use others arguments. Also, some changes are due to the fact that we work with functions f that are continuous on each I_m but not on I. In particular, in general there does not exist $x \in I$ such that $f(x) = \sup_I f$.

Proposition 3.6

1. *The function $s \to \lambda(s)$ is strictly decreasing along the real axis $s > \gamma$.*
2. *On each vertical line $\mathcal{R}(s) = \sigma$, we have $R(s) \leq \lambda(\sigma)$.*
3. *If $R(s) = \lambda(\sigma)$ for $s = \sigma + it$ then $\mathbf{G}_s$ has an eigenvalue $\lambda = e^{ia}\lambda(\sigma)$, $a \in \mathbb{R}$ that belongs to the spectrum of G_s.*

Proof. − From (2.3), we deduce that:

$$\lambda(s) = \lim_{k\to\infty} \Lambda_k(\mathbf{1},s)^{\frac{1}{k}}.$$

Since for all $m \in \mathcal{M}$, h_m is a δ-contraction, we deduce:

$$\begin{aligned}
\Lambda_k(\mathbf{1},s+u) &= \sum_{h\in\mathcal{L}_k} |I_h|^{s+u} \\
&\leq \sum_{h\in\mathcal{L}_k} \delta_h^u |I_h|^s \leq \delta^{ku} \sum_{h\in\mathcal{L}_k} |I_h|^s.
\end{aligned}$$

Thus, $\lambda(s+u) \leq \delta^u \lambda(s)$ and we have proved item 1.
To prove item 2, it suffices to remark that for $f \in L_{pw}(\mathcal{I})$, $\|\mathbf{G}_s^k f\|_\infty \leq \|\mathbf{G}_\sigma^k f\|_\infty$. This together with Lemma 3.2 gives $R(s) \leq R(\sigma) = \lambda(\sigma)$.
Finally, if $R(s) = \lambda(\sigma)$ then by Proposition 3.5, the operator $\mathbf{G}_s$ is quasi compact and thus admits a eigenvalue $\lambda = e^{ia}\lambda(\sigma)$ of modulus $\lambda(\sigma)$. Let Ψ_s be such that $\mathbf{G}_s\Psi_s = \lambda\Psi_s$ and $\psi_s(x) = \Psi_s(x,x)$. Then $G_s\psi_s = \lambda\psi_s$. □
Let us study the spectral properties of $\mathbf{G}_s$ for $\mathcal{R}(s) = 1$. Let us remark that for any distribution F, we have (see also Proposition 5 in (V1)),

$$\Lambda_k(F,1) = 1.$$

Thus $\lambda(1) = 1$. For further use, let us denote $\varphi_1 \in L_{pw}(I)$ the eigenfunction of G_1 corresponding to the maximal eigenvalue $\lambda(1) = 1$ and satisfying $m(\phi_1) = 1$. Then the measure $\nu = \phi_1 m$ is T invariant.

Proposition 3.7 *Let $\mathcal{R}(s) = 1$, the operator may behave in two different ways.*

1. *Either for all $s \neq 1$, $\mathcal{R}(s) = 1$, $R(s) < 1$ (the aperiodic case),*
2. *or the set of $t \in \mathbb{R}$ such that 1 belongs to the spectrum of $\mathbf{G}_{1+it}$ is of the form $t_0\mathbb{Z}$ for some t_0 (the periodic case). In that case, the map T is conjugated to a piecewise affine map with slopes of the form α^k, $\alpha > 1$, the conjugacy is C^{1+Lip} on each I_m. Moreover, there exists $\sigma_0 < 1$ such that on the strip $\{\sigma_0 < \mathcal{R}(s) < 1\}$ the operator $(I - \mathbf{G}_s)^{-1}$ has no pole.*

Proof. − Let $s = 1+it$ and assume that 1 belongs to the spectrum of $\mathbf{G}_{1+it}$. Then using Proposition 3.6 we have that there exists $f \in L_{pw}(I)$ such that $G_s f = f$. Let us prove that $|f|$ is an eigenfunction for G_1 with eigenvalue 1. We have

$$|f| = |G_s f| \leq G_1|f|. \tag{3.4}$$

Recall that the Lebesgue measure is invariant by G_1 so that

$$\int_I G_1|f|(x)dx = \int_I |f|dx.$$

As a consequence, inequality (3.4) must be an equality. Now, because of Theorem 2.4, 1 is simple as an eigenvalue of G_1. Thus, let $f_1 > 0$ be a dominant eigenfunction of G_1. Let $\mu(x) = \frac{f(x)}{f_1(x)}$, multiplying if necessary f_1 by some constant, we may assume that $|\mu| \equiv 1$. Following B. Vallée's proof of Proposition 9, we obtain that for all $m \in \mathcal{M}$, $x \in J_m$,

$$h_m'(x)^{it}\mu \circ h_m(x) = \mu(x). \tag{3.5}$$

Reciprocally, let t be such that there exists a function μ satisfying (3.5) for all $m \in \mathcal{M}$ then $f = \mu \cdot f_1$ satisfies $G_{1+it}f = f$.
In other words, we have proved that 1 belongs to the spectrum of $\mathbf{G}_{1+it}$ if and only if there exists a function μ satisfying (3.5) for all $m \in \mathcal{M}$. This implies that the set of real t such that 1 belongs to the spectrum of $\mathbf{G}_{1+it}$ is a subgroup of $\mathbb{R}$: if

$$h_m'(x)^{it}\mu_t \circ h_m(x) = \mu_t(x) \text{ and } h_m'(x)^{it'}\mu_{t'} \circ h_m(x) = \mu_{t'}(x)$$

then

$$h'_m(x)^{i(t-t')}\left(\frac{\mu_t}{\mu_{t'}}\right)\circ h_m(x)=\left(\frac{\mu_t}{\mu_{t'}}\right)(x).$$

It cannot accumulate 0 because of the analyticity of $s\to\lambda(s)$ near $s=1$. Thus it is of the form $t_0\mathbb{Z}$.
There exists a real function $\theta\in L_{pw}(I)$ such that $\mu=e^{i\theta}$ (recall that $\mu=\frac{f}{f_1}\in L_{pw}(I)$), take $\phi=\exp(\frac{\theta}{t})$ and $\alpha=\exp(\frac{2\pi}{t})$. Equation (3.5) becomes:

$$\log|T'(x)|=\frac{\theta(x)}{t}-\frac{\theta(Tx)}{t}+\frac{2k(x)\pi}{t}$$

where $k(x)\in\mathbb{Z}$ and is constant on each I_m, and finally, equation (3.5) may be rewritten as:

$$|T'(x)|=\frac{\phi(x)}{\phi\circ T(x)}\alpha^{k(x)},$$

Now, we may find constants c_m and d_m, $m\in\mathcal{M}$ such that the function

$$g(x)=c_m\int_{a_m}^{x}\phi(t)dt+d_m\ \ x\in I_m$$

is continuous, maps I into I, is invertible, is derivable on each I_m with Lipschitz derivative on each I_m. Derivating $\widetilde{T}=g\circ T\circ g^{-1}$ we obtain that $\widetilde{T}$ is piecewise affine with slopes α^k.
Let us prove the existence of a strip free of poles. There exists $\gamma<\sigma_1<1$ such that for any $\sigma\in]\sigma_1,1[$, the operator G_σ has no eigenvalue of modulus 1. Let $\sigma_1<\sigma_0<1$ being such that $\delta\lambda(\sigma)<1$ for all $\sigma>\sigma_0$. Let $\sigma\in]\sigma_0,1[$ and $s=\sigma+i\tau$. Proposition 3.5 implies that either $\mathbf{G}_s$ is quasi-compact or $R(s)<1$ (in this last case 1 does not belong to the spectrum of $\mathbf{G}_s$). So assume that $\mathbf{G}_s$ is quasi-compact. If 1 is in the spectrum of $\mathbf{G}_s$, then it is an eigenvalue of $\mathbf{G}_s$ (Theorem 3.3) and of G_s. There exists $f\in L_{pw}(I)$ such that $G_s(f)=f$. Using that $\alpha^{ikt_0}=1$ for any integer k, one deduces that 1 is an eigenvalue of the operators $G_{\sigma+i(\tau+kt_0)}$for any $k\in\mathbb{Z}$. It follows that if there is no strip free of poles, then some of the points of the line $\mathcal{R}(s)=1$ are accumulated by a sequence of poles of $\Lambda(F,s)$. This is a contradiction since $\Lambda(F,s)$ is a meromorphic function in a neighborhood of $\mathcal{R}(s)=1$. □
We now prove the log-convexity of $s\to\lambda(s)$. Such a property is necessary to study the random variable $\log\ell_k$.

Proposition 3.8 *The function $s\to\log\lambda(s)$ is convex. Either it is strictly convex or it is affine. In this last case, the map T is conjugated to a piecewise affine map with slopes all equal. The conjugacy is C^{1+Lip} on each I_m.*

Proof. − We have to prove that for $t\in[0,1]$ and $s>\gamma$, $s'>\gamma$,

$$\lambda(ts+(1-t)s')\le\lambda(s)^t\cdot\lambda(s')^{1-t}.\tag{3.6}$$

Consider the function

$$\psi=f_{ts+(1-t)s'}(f_s)^{-t}(f_{s'})^{-(1-t)}$$

where f_σ denote a dominant eigenfunction of G_σ. We may normalize ψ to have $\sup_I \psi = 1$. Consider a sequence $x_n \in I$ such that $\psi(x_n) \to 1$.

$$\lambda(ts+(1-t)s')f_{ts+(1-t)s'}(x_n) = \sum_{h\in\mathcal{M}} |h'|(x_n)^{ts+(1-t)s'} f_{ts+(1-t)s'}(h(x_n)) \tag{3.7}$$

$$\leq \sum_{h\in\mathcal{M}} |h'|(x_n)^{ts} f_s(hx_n)^t \cdot |h'|(x_n)^{(1-t)s'} f_{s'}(hx_n)^{1-t}$$

$$\leq \left(\sum_{h\in\mathcal{M}} |h'|(x_n)^s f_s(hx_n)\right)^t \cdot \left(\sum_{h\in\mathcal{M}} |h'|(x_n)^{s'} f_{s'}(hx_n)\right)^{1-t} \tag{3.8}$$

$$= \lambda(s)^t f_s(x_n)^t \cdot \lambda(s')^{1-t} f_{s'}(x_n)^{1-t}.$$

(3.8) follows from Hölder inequality. Taking the limit when $n \to \infty$ gives (3.6). λ being analytic, if equality holds in (3.6) for some s, s', t then $\log\lambda$ is affine. In this last case, it remains to prove that the map T is conjugated to a piecewise affine map with slopes all equal.
Assume that $\log\lambda$ is affine then there exists $a < 1$ such that $\lambda(s) = a^{s-1}$. Choose s, s', t such that $ts+(1-t)s' = 1$, let us show that $f_s^t \cdot f_{s'}^{1-t}$ is a dominant eigenfunction of G_1. Hölder inequality implies that

$$G_1(f_s^t \cdot f_{s'}^{1-t}) \leq f_s^t \cdot f_{s'}^{1-t}.$$

As in the proof of Proposition 3.7, we use that G_1 leaves Lebesgue measure invariant to conclude that $G_1(f_s^t \cdot f_{s'}^{1-t}) = f_s^t \cdot f_{s'}^{1-t}$. As a consequence, $\psi \equiv 1$ and equality holds in (3.7) for all $x \in I$. This implies that there exists a function $k : I \to \mathbb{R}^+$ such that for all $h \in \mathcal{M}$,

$$|h'(x)|^s f_s(hx) = k(x)|h'(x)|^{s'} f_{s'}(hx).$$

Summing over $h \in \mathcal{M}$ and noting $\phi(x) = \dfrac{f_s(x)}{f_{s'}(x)}$ we get that

$$\phi(x) = k(x)\frac{\lambda(s)}{\lambda(s')}$$

and then T satisfy a cocycle relation:

$$|h'(x)|^{s-s'} \phi \circ h(x) = \frac{\lambda(s)}{\lambda(s')}\phi(x) \text{ for all } h \in \mathcal{M}. \tag{3.9}$$

Following the end of the proof of Proposition 3.7, we conclude that T is conjugated to a piecewise affine map with slopes all equal to $\frac{1}{a}$. □

Remark 4 *By the way, the cocycle argument used in the proofs of Proposition 3.7 and 3.8 resolve Conjecture 1 of B. Vallée:*
A source is similar to a source with affine branches if and only if it is conjugated to a source with affine branches. The conjugacy is C^{1+Lip} on each I_m.

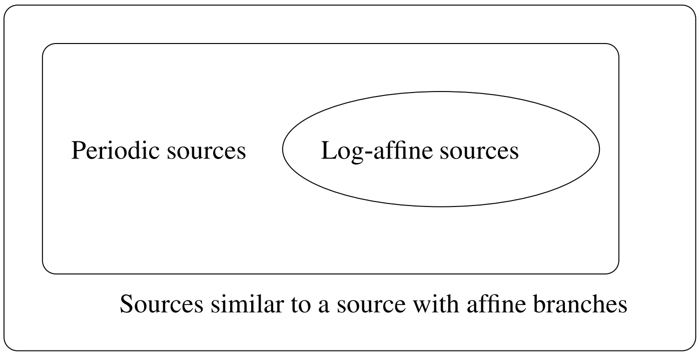

Fig. 7: Exceptional sources

Figure 7 shows relations between sources conjugated to piecewise affine sources.
Let $U(s) = \log \lambda(s)$. With propositions 3.6, 3.7, 3.8 the analysis of parameters of the source done in sections 7, 8, 9 of (V1) apply to our setting without any change. To conclude the proof Theorem 2.3, it remains to verify that if the source is not log-affine then $U''(1) \neq 0$. This is necessary to apply Hwang's quasi powers theorem and obtain the central limit theorem.

Proposition 3.9 *Let $U(s) = \log \lambda(s)$. The following assertions are equivalent :*

1. $U''(1) = 0$,
2. *there exist a C^1 function ϕ whose derivative belongs to the space $L_{pw}(I)$ such that for all $s > \gamma$, for all $h \in \mathcal{M}$, $|h'(x)|^s \phi \circ h(x) = \lambda(s)\phi(x)$.*

Proof. – We apply Proposition 6.10 and Theorem 6.8 of (Bro) with the function

$$f = -\log|T'| + \int_I \log|T'| d\nu$$

where ν is the probability measure which is T-invariant and absolutely continuous with respect to m. It is classical that

$$\int_I \log|T'| d\nu = -\lambda'(1).$$

Clearly, $f \in L_{pw}(I)$, also $L_{pw}(I)$ is dense in $L^1(\nu)$. With A. Broise's notations (the operators $\Phi_f(\theta)$ are defined in (Bro), pp 37), we have:

$$G_s(g) = e^{-(s-1)\lambda'(1)} \Phi_f(s-1)(g).$$

So that if $\tilde{\lambda}(s-1)$ is the dominant eigenvalue of $\Phi_f(s-1)$ then $\tilde{\lambda}(s-1) = e^{(s-1)\lambda'(1)}\lambda(s)$. Since

$$U''(1) = \frac{\lambda''(1) - \lambda'(1)^2}{\lambda(1)^2},$$

we deduce that $U''(1)=0$ if and only if $\tilde{\lambda}''(0)=0$ which is equivalent to item 1. by Proposition 6.10 and Theorem 6.8 of (Bro). $\square$

4 Spectral properties of real generalized Ruelle operators

The aim of this section is to prove Theorem 2.4. Let us recall definitions and properties of cones and projective metrics (see (L) or (L,S,V) for a complete presentation).

4.1 Cones and projective metrics

The theory of cones and projective metrics of G. Birkhoff (Bi1) is a powerful tool to study linear operators. P. Ferrero and B. Schmitt (F,S) applied it to estimate the correlation decay for random compositions of dynamical systems.

Definition 3 *Let $\mathcal{V}$ be a vector space. A subset $\mathcal{C}\subset\mathcal{V}$ which enjoys the following four properties*

(i) $\mathcal{C}\cap-\mathcal{C}=\emptyset$

(ii) $\forall\lambda>0\ \ \lambda\mathcal{C}=\mathcal{C}$

(iii) $\mathcal{C}$ *is a convex set*

(iv) $\forall f,g\in\mathcal{C}\,,\forall\alpha_n\in\mathbb{R}\,,\ (\alpha_n\to\alpha,\ \ g-\alpha_n f\in\mathcal{C})\Rightarrow(g-\alpha f\in\mathcal{C}\cup\{0\}).$

is called a convex cone.

We now define the Hilbert metric on $\mathcal{C}$:

Definition 4 *The distance $d_{\mathcal{C}}(f,g)$ between two points f,g in $\mathcal{C}$ is given by*

$$\begin{aligned} d_{\mathcal{C}}(f,g) &= \log\frac{\beta(f,g)}{\alpha(f,g)} \text{ where } \alpha \text{ and } \beta \text{ are defined as} \\ \alpha(f,g) &= \sup\{\lambda>0 | g-\lambda f\in\mathcal{C}\} \\ \beta(f,g) &= \inf\{\mu>0 | \mu f-g\in\mathcal{C}\} \end{aligned}$$

where we take $\alpha=0$ or $\beta=\infty$ when the corresponding sets are empty.

Remark 5 *In the sequel we will use that* $\beta(f,g)=\alpha(g,f)$.

The distance $d_{\mathcal{C}}$ is a pseudo-metric, because two elements can be at an infinite distance from each other, and it is a projective metric because any two proportional elements have a null distance.
Given two elements linearly independent f and $g\in\mathcal{C}$, consider the intersection of $\mathcal{C}$ with the two dimensional vector space spanned by f and g. Its boundary is the union of two half lines ℓ_1, ℓ_2. The distance $d_{\mathcal{C}}(f,g)$ is the log of the cross-ratio of the four half lines ℓ_1, ℓ_2, f, g (see figure 8).

Remark 6 *For example, if $\mathcal{V}$ is a space formed with real valued functions and $\mathcal{C}^+$ the cone of positive functions then an easy computation gives:*

$$d_{\mathcal{C}^+}(f,g)=\frac{\sup f}{\inf f}\frac{\sup g}{\inf g}.$$

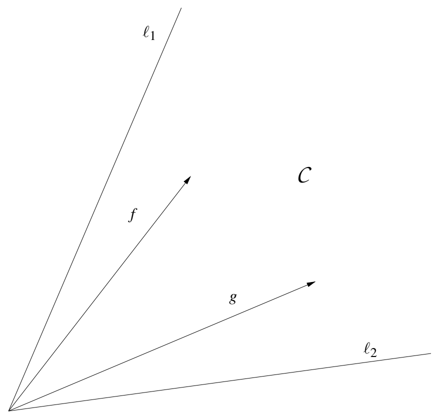

Fig. 8: Projective metric

Definition 5 *Let $\mathcal{V}$ be a vector space, $\mathcal{C} \subset \mathcal{V}$ a convex cone, a linear operator $L : \mathcal{V} \to \mathcal{V}$ is called a* positive operator *if $L\mathcal{C} \subset \mathcal{C}$.*

The next theorem, due to G. Birkhoff (Bi2), shows that every positive linear operator is a contraction, provided that the diameter of the image is finite.

Theorem 4.1 *Let $\mathcal{V}$ be a vector space, $\mathcal{C} \subset \mathcal{V}$ a convex cone (see definition above) and $L : \mathcal{V} \to \mathcal{V}$ a positive linear operator. Let $d_{\mathcal{C}}$ be the Hilbert metric associated to the cone $\mathcal{C}$. If we denote*

$$\Delta = \sup_{f,g \in L(\mathcal{C})} d_{\mathcal{C}}(f,g) \quad ,$$

then

$$d_{\mathcal{C}}(Lf, Lg) \le \tanh\left(\frac{\Delta}{4}\right) d_{\mathcal{C}}(f,g) \ \ \forall f,g \in \mathcal{C}$$

$(\tanh(\infty) = 1)$.

Remark 7 *Theorem 4.1 implies that if $\mathcal{C}_1$ and $\mathcal{C}_2$ are two convex cones such that $\mathcal{C}_1 \subset \mathcal{C}_2$ then for any $f,g \in \mathcal{C}_1$,*

$$d_{\mathcal{C}_2}(f,g) \le d_{\mathcal{C}_1}(f,g).$$

(apply Theorem 4.1 with $L = Id$). In particular, if $\mathcal{C} \subset \mathcal{C}^+$ then for $f,g \in \mathcal{C}$,

$$d_{\mathcal{C}}(f,g) \ge \frac{\sup f}{\inf f} \frac{\sup g}{\inf g}.$$

Theorem 4.1 alone is not completely satisfactory: given a cone $\mathcal{C}$ and its metric $d_{\mathcal{C}}$, we need to relate the distance $d_{\mathcal{C}}$ with a suitable norm on $\mathcal{V}$. The following lemma provides such a relation.

Lemma 4.2 *(L,S,V) Let $\|\cdot\|$ be a norm on $\mathcal{V}$ such that*

$$\forall f,g \in \mathcal{V} \ \ f-g, \ f+g \in \mathcal{C} \Rightarrow \|g\| \le \|f\|$$

and let $\ell : \mathcal{C} \to \mathbb{R}^+$ be a homogeneous and order preserving function, i.e.

$$\forall f \in \mathcal{C}, \forall \lambda \in \mathbb{R}^+ \qquad \ell(\lambda f) = \lambda \ell(f)$$
$$\forall f, g \in \mathcal{C} \qquad g - f \in \mathcal{C} \Rightarrow \ell(f) \leq \ell(g) ,$$

then

$$\forall f, g \in \mathcal{C} \;\; \ell(f) = \ell(g) > 0 \Rightarrow \|f - g\| \leq (\mathrm{e}^{d_{\mathcal{C}}(f,g)} - 1) \min(\|f\|, \|g\|)$$

4.2 Proof of Theorem 2.4

We are now going to use Theorem 4.1 and Lemma 4.2 to prove Theorem 2.4. Recall that we already know from Section 3 that the operators $\mathbf{G}_s$ are quasi-compact for real $s > \gamma$. It remains to prove that they have a unique dominant eigenvalue ; we prove it for $\mathbf{G}_s$ and leave to the reader the proof for G_s.
Let us sketch how to use cones to obtain the dominant spectral properties. To obtain a unique dominant eigenvalue, it is sufficient to find a cone $\mathcal{C}$ and an integer k such that $\mathbf{G}_s^k$ maps C into itself and the diameter Δ of $\mathbf{G}_s^k \mathcal{C}$ into $\mathcal{C}$ is finite. Indeed, the idea is that if $f \in \mathcal{C}$ then Theorem 4.1 applied $n-1$ times gives ($\kappa = \tanh \frac{\Delta}{4} < 1$):

$$d_{\mathcal{C}}(\mathbf{G}_s^{k+n} f, \mathbf{G}_s^n f) \leq \left(\tanh \frac{\Delta}{4}\right)^{n-1} \Delta = \kappa^{n-1} \Delta. \tag{4.1}$$

We deduce (using Lemma 4.2 with $\| \; \|_\infty$ as homogenous form) that the sequence of lines generated by $(\mathbf{G}_s^n f)_{n\in\mathbb{N}}$ is a Cauchy sequence and converges to a line generated by an eigenvector Ψ_s. This eigenvector corresponds to an eigenvalue $\lambda(s)$. On another hand, we construct an eigenvector ν_s for the dual operator. Then Lemma 4.2 (applied with ν_s as homogenous form) and equation (4.1) give that

$$\left\| \frac{\mathbf{G}_s^n f}{\lambda(s)^n} \; - \; \Psi_s \nu_s(f) \right\|_\infty \tag{4.2}$$

goes to zero exponentially fast for any $f \in \mathcal{C}$. Then we have to extend this result from the cone to the Banach space of piecewise Lipschitz functions (this is done using Lemma 4.9 below). The fact that (4.2) goes to zero exponentially fast implies that $\lambda(s)$ is the unique dominant eigenvalue of $\mathbf{G}_s$.

The following lemma proves the existence of a real positive eigenvalue for the dual operator of $\mathbf{G}_s$. The corresponding eigenvector is indeed a measure. Recall that is $\mathcal{V}$ is a topological Banach space, it topological dual $\mathcal{V}'$ is endowed with the weak topology that is: a sequence $(\nu_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{V}'$ converges to $\nu \in \mathcal{V}'$ if and only if for any $f \in \mathcal{V}$, the sequence $(\nu_n(f))_{n\in\mathbb{N}}$ converges to $\nu(f)$. Also, if L is a continuous linear operator on $\mathcal{V}$ then it defines a continuous linear operator L' on $\mathcal{V}'$ by: for $\nu \in \mathcal{V}'$, for any $f \in \mathcal{V}$,

$$L'(\nu)(f) = \nu(L(f)).$$

Lemma 4.3 *There exists a measure ν_s on $\mathcal{I}$ and a positive number $\lambda(s)$ such that for $f \in L_{pw}(\mathcal{I})$,*

$$\nu_s(\mathbf{G}_s f) = \lambda(s) \cdot \nu_s(f).$$

Proof. – Let $L_{pw}(\mathcal{J})^*$ be the topological dual of $L_{pw}(\mathcal{J})$ (i.e. $L_{pw}(\mathcal{J})^*$ is the set of continuous linear forms on $L_{pw}(\mathcal{J})$). Recall that the weak topology on $L_{pw}(\mathcal{J})^*$ is defined by: Γ_n weakly converges to $\Gamma \in L_{pw}(\mathcal{J})^*$ if and only if for all $f \in L_{pw}(\mathcal{J})$, $\Gamma_n(f)$ converges to $\Gamma(f)$.
Let $K \subset L_{pw}(\mathcal{J})^*$ be the positive forms Γ of $L_{pw}(\mathcal{J})^*$ such that $\Gamma(\mathbf{1}) = 1$. Define $\mathbf{P}_s$ which maps K into itself by:

$$\mathbf{P}_s\Gamma(f) = \frac{\Gamma(\mathbf{G}_s f)}{\Gamma(\mathbf{G}_s\mathbf{1})},$$

(remark that the positivity condition (d5) implies that $\inf(\mathbf{G}_s\mathbf{1}) > 0$). K is non empty, convex and weakly compact (see (Sc) Theorem $XIX.7.7$), $\mathbf{P}_s$ is continuous on it for the weak topology. Then the Schauder-Tychonoff theorem ((D,S)) implies that it admits a fixed point ν_s. We know that ν_s is additive, to prove that it is a measure, we need to prove that it is σ-additive.
Restricted to each $I_m \times I_m$, ν_s may be identified to a measure (by Riesz representation theorem on compact spaces), in particular, we may compute $\nu_s(\mathbf{1}_{I_m\times I_m})$. To conclude that it is a measure on the union $\mathcal{J}$ of the $I_m \times I_m$, it suffices to prove that:

$$\nu_s\Big(\sum_{m\in\mathcal{M}} \mathbf{1}_{I_m\times I_m}\Big) = \sum_{m\in\mathcal{M}} \nu_s(\mathbf{1}_{I_m\times I_m}).$$

This will follow from:

$$\lim_{\substack{|Q|\to\infty\ Q\subset\mathcal{M} \\ Q_{finite}}} \nu_s\Big(\sum_{m\notin Q} \mathbf{1}_{I_m\times I_m}\Big) = 0.$$

We have:

$$\nu_s\Big(\sum_{m\notin Q} \mathbf{1}_{I_m\times I_m}\Big) = \frac{\nu_s\left(\mathbf{G}_s(\sum_{m\notin Q}\mathbf{1}_{I_m\times I_m})\right)}{\lambda(s)}$$

and $\nu_s\left(\mathbf{G}_s(\sum_{m\notin Q}\mathbf{1}_{I_m\times I_m})\right) \le \sup \mathbf{G}_s(\sum_{m\notin Q}\mathbf{1}_{I_m\times I_m})$ (because $\nu_s \in K$ is a positive form). Now, for all $X = (x,x') \in \mathcal{J}$,

$$\begin{aligned}
\mathbf{G}_s\Big(\sum_{m\notin Q}\mathbf{1}_{I_m\times I_m}\Big)(X) &= \sum_{\substack{n\in\mathcal{M}\\ X\in J_n\times J_n}} H_n^s(X) \sum_{m\notin Q} \mathbf{1}_{I_m\times I_m}(h_n(x),h_n(x')) \\
&= \sum_{\substack{m\notin Q\\ X\in J_m\times J_m}} H_m^s(X) \\
&\le \sum_{\substack{m\notin Q\\ x\in J_m}} \delta_m^s.
\end{aligned}$$

The result follows from (2.1) which is equivalent to Condition (d2) (see Remark 2). □

The starting point of the construction of cones is usually a Lasota-Yorke inequality. Looking at Lemma 3.2, on sees that the operators $\mathbf{G}_s$ improve the Lipschitz constant of piecewise Lipschitz functions, this explains condition 3 in the definition of the cone below. If the alphabet $\mathcal{M}$ was finite, we could manage with this condition and the positivity condition 2 below. The idea is that condition 3 provides a good control of the functions on each $I_m \times I_m$ and if the alphabet is finite, we may obtain a uniform control on the union of the $I_m \times I_m$, iterating by $\mathbf{G}_s^k$ if necessary. Since we want to deal with infinite alphabet, we add

condition 4 below which controls the functions on the complementary of a well chosen finite part Q of $\mathcal{M}$.
Let Q be a finite subset of $\mathcal{M}$ such that:

$$\sup_{x\in I}\sum_{\substack{m\notin Q\\ x\in J_m}}\delta_m^s<\lambda(s)\delta.$$

The existence of such a subset Q follows from (2.1).
For $a>0$, $b>0$, let $C_{a,b}(s)$ be the set of functions f on $\mathcal{J}$ such that:

1. $f\in L_{pw}(\mathcal{J})$,
2. $\forall(u,v)\in\mathcal{J}$, $f(u,v)>0$,
3. $\forall m\in\mathcal{M}$, $\forall(u,u')=U,(v,v')=V\in I_m\times I_m$, $f(u,u')\leq e^{ad(U,V)}\cdot f(v,v')$ (the distance $d(U,V)$ has been defined in (3.1)),
4. for all $m\notin Q$, for $(u,v)\in I_m\times I_m$, $f(u,v)\leq b\nu_s(f)$.

Remark that $C_{a,b}(s)\subset L_{pw}(\mathcal{J})$ is a cone.

Remark 8 *As mentioned in Remark 3, if the source is complete, $\mathbf{G}_s$ acts on the space $C^1(I\times I)$ of functions that are C^1 on $I\times I$. To get the dominant eigenvalue result on this space, it suffices to replace item 1 in the definition of the cone above by "$f\in C^1(I\times I)$".*

Lemma 4.4

1. *For all $m\in\mathcal{M}$, $\nu_s(I_m\times I_m)>0$.*
2. *For $f\in C_{a,b}(s)$, $m\in\mathcal{M}$, $X\in I_m\times I_m$, $f(X)\leq e^{2a}\dfrac{\nu_s(f)}{\nu_s(I_m\times I_m)}$.*

Proof. – To prove Item 1, remark that (2.2) and Taylor equality imply that for all $m\in\mathcal{M}$, there exists $N\in\mathbb{N}$ such that

$$\inf_{X\in\mathcal{J}}\mathbf{G}_s^N\mathbf{1}_{I_m\times I_m}(X)>0.$$

Now, $\nu_s(\mathbf{1}_{I_m\times I_m})=\lambda(s)^{-N}\nu_s(\mathbf{G}_s^N\mathbf{1}_{I_m\times I_m})>0$.
Item 2. follows from the definition of the cone (condition 3.) by integrating with respect to ν_s on $I_m\times I_m$.
□

Lemma 4.5 *For any $s>\gamma$, for any $\delta<\xi<1$, there exists $a_0>0$, $b_0>0$ such that for all $a\geq a_0$, $b\geq b_0$ and for any $k\in\mathbb{N}^*$, $\mathbf{G}_s^k$ maps $C_{a,b}(s)$ into $C_{\xi a,\xi b}(s)$.*

Proof. – Let $f\in C_{a,b}(s)$. Because $C_{\xi a,\xi b}(s)\subset C_{a,b}(s)$, it suffices to proof the lemma for $k=1$. Let $f\in C_{a,b}(s)$, for any$m\in\mathcal{M}$, $X=(x,x')$, $Y=(y,y')\in I_m\times I_m$, we have to compare: $f(h_m(x),h_m(x'))H_m^s(x,x')$ with $f(h_m(y),h_m(y'))H_m^s(y,y')$. Because f belongs to $C_{a,b}(s)$, and for each $m\in\mathcal{M}$, h_m is a δ-contraction, we have:

$$f(h_m(x),h_m(x'))\leq e^{a\delta d(X,Y)}f(h_m(y),h_m(y')).$$

(3.1) implies that:

$$H^s_m(x,x') \le e^{sBd(X,Y)} H^s_m(y,y').$$

So, $\mathbf{G}_s f(X) \le e^{\xi a d(X,Y)} \mathbf{G}_s f(Y)$ provided $a \ge \dfrac{sB}{\xi - \delta}$.

Now, let $X \in I_m \times I_m$ with $m \notin Q$. Let $c := \inf_{m \in Q} \nu_s(I_m \times I_m)$, $c > 0$ because of Lemma 4.4 and the fact that Q is finite. We have:

$$\begin{aligned}
\mathbf{G}_s f(X) &= \sum_{\substack{m \in Q \\ X \in J_m \times J_m}} H^s_m(X) f(h_m(x), h_m(x')) \\
&\quad + \sum_{\substack{m \notin Q \\ X \in J_m \times J_m}} H^s_m(X) f(h_m(x), h_m(x')) \\
&\le \frac{e^{2a}}{c} \nu_s(f) \|\mathbf{G}_s \mathbf{1}\|_\infty + b \nu_s(f) \sup_{x \in I} \sum_{\substack{m \notin Q \\ x \in J_m}} \delta^s_m.
\end{aligned}$$

Now, we use that $\nu_s(f) = \dfrac{\nu_s(\mathbf{G}_s f)}{\lambda(s)}$ and since $\sup_{x \in I} \sum_{\substack{m \notin Q \\ x \in J_m}} \delta^s_m < \lambda(s)\delta$, we get:

$$\mathbf{G}_s f(X) \le \nu_s(\mathbf{G}_s f) \ (\frac{e^{2a} \|\mathbf{G}_s \mathbf{1}\|}{c \lambda(s)} + b\delta) \le b \xi \nu_s(\mathbf{G}_s f)$$

provided $b \ge \frac{e^{2a} \|\mathbf{G}_s \mathbf{1}\|}{\lambda(s) c (\xi - \delta)}$. □

Lemma 4.6 *Let $a \ge a_0$, $b \ge b_0$, there exists M such that for $k \ge M$, the projective diameter Δ of $\mathbf{G}^k_s C_{a,b}(s)$ into $C_{a,b}(s)$ is finite:*

$$\Delta = \sup_{f,g \ \in \ C_{a,b}(s)} d_{C_{a,b}(s)}(\mathbf{G}^k_s f, \mathbf{G}^k_s g) < \infty.$$

Proof. – Let $f, g \in C_{\xi a, \xi b}(s)$, let $\beta > 0$, we have that $\beta f - g \in C_{a,b}(s)$ if and only if:

1. $\beta > \dfrac{g(x,x')}{f(x,x')}$ for all $(x,x') \in \mathcal{I}$.
2. $\beta > \dfrac{e^{ad(X,Y)} g(y,y') - g(x,x')}{e^{ad(X,Y)} f(y,y') - f(x,x')} := u(X,Y)$ for all $(x,x'), (y,y') \in \mathcal{I}$.
3. $\beta > \dfrac{b \nu_s(g) - g(x,x')}{b \nu_s(f) - f(x,x')} := v(X)$ for all $(x,x') \in I_m \times I_m$, $m \notin Q$.

The quantity $u(X,Y)$ may be rewritten as:

$$u(X,Y) = \left(\frac{g(y,y')}{f(y,y')} \right) \left(\frac{e^{ad(X,Y)} - \frac{g(x,x')}{g(y,y')}}{e^{ad(X,Y)} - \frac{f(x,x')}{f(y,y')}} \right),$$

using that $f,g \in C_{\xi a,\xi b}(s)$, we get:

$$\begin{aligned} u(X,Y) &\leq \frac{g(y,y')}{f(y,y')} \frac{e^{ad(X,Y)} - e^{-a\xi d(X,Y)}}{e^{ad(X,Y)} - e^{a\xi d(X,Y)}} \\ &\leq \frac{g(y,y')}{f(y,y')} e^{a(1+\xi)} \frac{1+\xi}{1-\xi}. \end{aligned}$$

Moreover, because for all $X \in I_m \times I_m$, $m \notin Q$, $0 < f(X) \leq b\xi\nu_s(f)$, we have that $v(X) \leq \frac{\nu_s(g)}{\nu_s(f)(1-\xi)}$.
Remarking that

$$\sup_{(y,y')\in\mathcal{J}} \frac{g(y,y')}{f(y,y')} e^{a(1+\xi)} \frac{1+\xi}{1-\xi} \geq \sup_{(y,y')\in\mathcal{J}} \frac{g(y,y')}{f(y,y')},$$

we have proven that if $f,g \in C_{\xi a,\xi b}(s)$

$$\beta(f,g) \leq \max\left[\sup_{(y,y')\in\mathcal{J}} \frac{g(y,y')}{f(y,y')} e^{a(1+\xi)} \frac{1+\xi}{1-\xi}, \frac{\nu_s(g)}{\nu_s(f)(1-\xi)}\right]. \tag{4.3}$$

The same computation (recall that $\alpha(f,g) = \beta(g,f)$) gives:

$$\alpha(f,g) \geq \min\left[\inf_{(y,y')\in\mathcal{J}} \frac{g(y,y')}{f(y,y')} e^{-a(1+\xi)} \frac{1-\xi}{1+\xi}, \frac{\nu_s(g)(1-\xi)}{\nu_s(f)}\right]. \tag{4.4}$$

One sees that we have to control the quantities $\sup_{\mathcal{J}} f$, $\inf_{\mathcal{J}} f$ with respect to $\nu_s(f)$. This cannot be done for all functions in $C_{\xi a,\xi b}(s)$ but it can be done for $\mathbf{G}_s^k f$ for $f \in C_{a,b}(s)$ and $k \geq M$, M large enough. To this aim, we need two sublemmas. The first one (Sublemma 4.7) proves there exists a finite subset $\widetilde{Q} \subset \mathcal{M}$ such that a function in $C_{a,b}(s)$ is bounded from below at least on one $I_m \times I_m$, $m \in \widetilde{Q}$. Then Sublemma 4.8 proves that if k is large enough, for $f \in C_{a,b}(s)$, $\mathbf{G}_s^k f$ is bounded from below on all the $I_m \times I_m$.

Sublemma 4.7 *Let $\varepsilon = \frac{1}{2e^{2a}}$. Let $\widetilde{Q}$ be a finite subset of $\mathcal{M}$ which contains Q and satisfies:*

$$\sum_{m\notin\widetilde{Q}} \nu_s(I_m \times I_m) < \frac{1}{2b}.$$

For all $f \in C_{a,b}(s)$, there exists $m \in \widetilde{Q}$ such that for all $X \in I_m \times I_m$,

$$f(X) \geq \varepsilon\nu_s(f).$$

Proof. − If the sublemma was false then for all $m \in \widetilde{Q}$, there would exist $X_m \in I_m \times I_m$ such that $f(X_m) < \varepsilon\nu_s(f)$. Then, we would have:

$$\begin{aligned} \nu_s(f) &= \sum_{m\in\mathcal{M}} \nu_s(f\mathbf{1}_{I_m\times I_m}) \\ &= \sum_{m\in\widetilde{Q}} \nu_s(f\mathbf{1}_{I_m\times I_m}) + \sum_{m\notin\widetilde{Q}} \nu_s(f\mathbf{1}_{I_m\times I_m}) \\ &\leq \sum_{m\in\widetilde{Q}} f(X_m)e^{2a}\nu_s(\mathbf{1}_{I_m\times I_m}) + b\nu_s(f) \sum_{m\notin\widetilde{Q}} \nu_s(\mathbf{1}_{I_m\times I_m}) \\ &< e^{2a}\varepsilon\nu_s(f) + b\frac{\nu_s(f)}{2b} < \nu_s(f) \end{aligned}$$

a contradiction. □

Choose $M > 0$ such that,

$$\inf_{m\in\widetilde{Q}}\inf_{x\in I}\sum_{\substack{|h|=M\\ x\in J_h}}\mathbf{1}_{I_m}(h(x))\eta_h^s := D > 0 \tag{4.5}$$

such a M exists because if (2.2) is satisfied for $m\in\widetilde{Q}$ and N then (2.2) is also satisfied for m and kN for all $k\in\mathbb{N}^*$. Since $\widetilde{Q}$ is finite, we may take M a common multiple.

Sublemma 4.8 *Let M be given by (4.5). There exist constants K_1, K_2 such that for any $k\geq M$, for any $f\in C_{a,b}(s)$, for all $X\in\mathcal{J}$,*

$$K_2\nu_s(\mathbf{G}_s^k f)\leq \mathbf{G}_s^k f(X)\leq K_1\nu_s(\mathbf{G}_s^k f).$$

Proof. − From Lemma 4.5, we have that if $f\in C_{a,b}(s)$ then $\mathbf{G}_s^p f\in C_{a,b}(s)$ for all $p\in\mathbb{N}$. So, it suffices to prove the inequality for $k=M$.
Since $\mathbf{G}_s^M f\in C_{a,b}(s)$, we have for all $X\in\mathcal{J}$:

$$\mathbf{G}_s^M f(X)\leq \nu_s(\mathbf{G}_s^M f)\max[b,\frac{e^{2a}}{c}].$$

Now, using Sublemma 4.7, we find $m_0\in\widetilde{Q}$ such that for $X\in I_{m_0}\times I_{m_0}$, $f(X)\geq\varepsilon\nu_s(f)$. Now, for all $X\in\mathcal{J}$,

$$\mathbf{G}_s^M f(X)\geq\varepsilon\nu_s(f)\mathbf{G}_s^M(\mathbf{1}_{I_{m_0}\times I_{m_0}})(X)\geq\nu_s(f)\varepsilon D.$$

So the sublemma is proved with
$K_1=\max[b,\frac{e^{2a}}{c}]$ and $K_2=\frac{\varepsilon D}{\lambda(s)^M}$. □
Let us conclude the proof of Lemma 4.6. Equations (4.3) and (4.4) together with Sublemma 4.8 gives that for all $f,g\in C_{a,b}(s)$ and $k\geq M$,

$$\begin{aligned} d_{C_{a,b}(s)}(\mathbf{G}_s^k f,\mathbf{G}_s^k g) &\leq \log\left[\frac{\sup\mathbf{G}_s^k f}{\inf\mathbf{G}_s^k f}\cdot\frac{\sup\mathbf{G}_s^k g}{\inf\mathbf{G}_s^k g}\cdot e^{2(1+\xi)}\cdot\left(\frac{1+\xi}{1-\xi}\right)^2\right]\\ &\leq \log\left[\left(\frac{K_1}{K_2}\right)^2\cdot e^{2(1+\xi)}\cdot\left(\frac{1+\xi}{1-\xi}\right)^2\right].\end{aligned}$$

Let $K=\log\frac{K_1}{K_2}$, we conclude that the projective diameter Δ of $\mathbf{G}_s^k C_{a,b}(s)$ into $C_{a,b}(s)$ is finite :

$$\Delta\leq 2K+2a(1+\xi)+2\log\frac{1+\xi}{1-\xi}.$$

□

The following lemma shows that any function in $L_{pw}(\mathcal{J})$ may be pushed into the cone $C_{a,b}(s)$.

Lemma 4.9 *There exists $K_3>0$ satisfying:*
for any function $f\in L_{pw}(\mathcal{J})$, there exists $R(f)>0$ such that $R(f)+f\in C_{a,b}(s)$ and $R(f)\leq K_3\cdot\|f\|$.

Proof. – Take $R(f)$ satisfying:

$R(f) \geq \sup|f|$,

$R(f) \geq \dfrac{\mathrm{Lip}(f) - a\inf f}{a}$,

$R(f) \geq \dfrac{\sup f - b\inf f}{b-1}$. □

We are now in position to prove that $\mathbf{G}_s$ has a unique simple dominant eigenvalue. Let $\kappa = \tanh\frac{\Delta}{4}$. Lemmas 4.5 and 4.6 and Theorem 4.1 give: $\forall n, m \in \mathbb{N}^*$,

$$d_{C_{a,b}(s)}(\mathbf{G}_s^{n+m}\mathbf{1}, \mathbf{G}_s^n\mathbf{1}) \leq \kappa^{n-1} d_{C_{a,b}(s)}(\mathbf{G}_s^{m+1}\mathbf{1}, \mathbf{G}_s\mathbf{1}) \leq \Delta\kappa^{n-1}.$$

Now, apply Lemma 4.2 using $\| \ \|_\infty$ and ν_s to get:

$$\left\| \frac{\mathbf{G}_s^{n+m}\mathbf{1}}{\lambda(s)^{n+m}} - \frac{\mathbf{G}_s^n\mathbf{1}}{\lambda(s)^n} \right\|_\infty \leq \Delta\kappa^{n-1} \cdot \left\| \frac{\mathbf{G}_s^n\mathbf{1}}{\lambda(s)^n} \right\|_\infty.$$

Now use again Lemma 4.2 to prove that $\left\| \dfrac{\mathbf{G}_s^n\mathbf{1}}{\lambda(s)^n} \right\|_\infty$ is bounded:

$$\left\| \frac{\mathbf{G}_s^n\mathbf{1}}{\lambda(s)^n} - \frac{\mathbf{G}_s\mathbf{1}}{\lambda(s)} \right\|_\infty \leq \Delta \cdot \left\| \frac{\mathbf{G}_s\mathbf{1}}{\lambda(s)} \right\|_\infty.$$

So, the sequence $\left(\frac{\mathbf{G}_s^n\mathbf{1}}{\lambda(s)^n}\right)_{n\in\mathbb{N}}$ is a Cauchy sequence (for the sup norm), thus it converges to some function Ψ_s. This function satisfy $\mathbf{G}_s\Psi_s = \lambda(s)\Psi_s$ and for any $f \in C_{a,b}(s)$,

$$\left\| \frac{\mathbf{G}_s^n f}{\lambda(s)^n} - \Psi_s\nu_s(f) \right\|_\infty \leq \Delta\kappa^{n-1}\|\Psi_s\|_\infty\nu_s(f).$$

Lemma 4.9 implies that for any $f \in L_{pw}(\mathcal{I})$,

$$\left\| \frac{\mathbf{G}_s^n f}{\lambda(s)^n} - \Psi_s\nu_s(f) \right\|_\infty \leq \Delta\kappa^{n-1}\|\Psi_s\|_\infty(2K_3+1)\|f\|. \tag{4.6}$$

We already know that $\mathbf{G}_s$ is quasi compact (recall section 3). Now, equation (4.6) proves that $\lambda(s)$ is a maximal eigenvalue (i.e. $\lambda(s) = R(s)$ and any other eigenvalue has modulus strictly smaller than $\lambda(s)$. In other words, $\lambda(s)$ is simple and the unique eigenvalue of maximal modulus. The proof is the same for G_s. Let us denote $\widetilde{\lambda}(s)$ the dominant eigenvalue and ψ_s the dominant eigenvector of G_s. We have

$$\lambda(s) = \lim_{n\to\infty}(\mathbf{G}_s^n\mathbf{1}(0,0))^{\frac{1}{n}} \text{ and } \widetilde{\lambda}(s) = \lim_{n\to\infty}(G_s^n\mathbf{1}(0))^{\frac{1}{n}}$$

and the operators $\mathbf{G}_s$ and G_s coincide on the diagonal, we conclude that $\lambda(s) = \widetilde{\lambda}(s)$. □

References

[Bi1] G. Birkhoff *Extensions of Jentzch's theorem.* T.A.M.S. (1957), **85**, 219-227.

[Bi2] G. Birkhoff *Lattice theory (3rd edition).* Amer. Math. Soc. (1967).

[Bre] X. Bressaud *Opérateurs de transfert sur le décalage à alphabet dénombrable et applications.* Erg. Th Dyn. Syst. (1999), **19**, 5, 1175-1200.

[Bro] A. Broise *Transformations dilatantes de l'intervalle et théorèmes limites.* Études spectrales d'opérateurs de transfert et applications. Astérisque (1996), **238**, 1-109.

[C,M,V] F. Chazal, V. Maume-Deschamps, B. Vallée *Erratum to "Dynamical sources in Information Theory: Fundamentals Intervals and Word Prefixes" by B. Vallée* Algorithmica (2004), **38**, 4, 591-596.

[C] J. Clément *Arbres digitaux et sources dynamiques*, thèse de l'université de Caen (2000).

[C,F,V] J. Clément, P. Flajolet, B. Vallée *Dynamical sources in information theory: A general analysis of trie structures* Algorithmica, **29**(1/2), pp. 307-369, (2001).

[D,S] N. Dunford & J. T. Schwartz. *Linear operators Part I.* Interscience, New-York (1966).

[F,S] P. Ferrero, B. Schmitt *Ruelle Perron Frobenius theorems and projective metrics.* Col. Math. Soc. J. Bolyai Random Fields. Estergom (Hungary), Edited by J. Fritz, J.L. Lebowitz and D. Szasz (1979), 333 - 337.

[H] H. Hénnion *Sur un théorème spectral et son application aux noyaux lipschitziens.* Proc. AMS, **118**, 2, (1993), 627-634 .

[K] T. Kato *Perturbation theory for linear operators.* Springer-Verlag, New-York, (1980).

[L] C. Liverani *Decay of correlations.* Ann. of Math. (1995), **142** (2), 239-301

[L,S,V] C. Liverani, B. Saussol & S. Vaienti *Conformal measure and decay of correlations for covering weighted systems.* Ergodic Theory Dynam. Systems (1999), **19**, 3, 671-685.

[M] V. Maume-Deschamps *Correlation decay for Markov maps on a countable state space.* Erg. Th Dyn. Syst. (2001), **21**, 1, 165-196.

[Sa] O. Sarig *Thermodynamic Formalism for Countable Markov Shifts.* Erg. Th. Dyn. Syst. (1999), **19**, 6, 1565-1593.

[Sc] L. Schwartz *Topologie générale et analyse fonctionelle.*Hermann, (1970).

[Sz] W. Szpankowski *Average Case Analysis of Algorithms on sequences*, John Wiley and sons, New York, 2001.

[V1] B. Vallée *Dynamical sources in information theory: fundamental intervals and word prefixes* Algorithmica, **29**, 262-306, (2001).

[V2] B. Vallée *Digits and continuants in Euclidean algorithms. Ergodic versus Tauberian theorems* Journal de Théorie des Nombres de Bordeaux, **12**, 531-570, (2000).

[V3] B. Vallée *Dynamics of the binary Euclidean Algorithm: Functional Analysis and Operators* Algorithmica, **22**, 660-685, (1998).

[V4] B. Vallée *Dynamical analysis of a class of Euclidian algorithms.* Th. Comput. Science, **297**, (2003), 447-486.