

Asymptotic behaviour of iterated function systems and applications to variable length Markov chains

Blandine Dubarry

► **To cite this version:**

Blandine Dubarry. Asymptotic behaviour of iterated function systems and applications to variable length Markov chains. Probability [math.PR]. Université Rennes 1, 2017. English. NNT : 2017REN1S114 . tel-01777729

HAL Id: tel-01777729

<https://tel.archives-ouvertes.fr/tel-01777729>

Submitted on 25 Apr 2018

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

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

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Bretagne Loire

pour le grade de
DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Mention : Mathématiques et applications

Ecole doctorale MATISSE

Blandine Dubarry

Préparée à l'IRMAR – UMR CNRS 6625
Institut de recherche mathématique de Rennes
U.F.R de mathématiques

**Comportement
asymptotique des
systèmes de
fonctions itérées et
applications pour les
chaînes de Markov
d'ordre variable**

**Thèse soutenue à Rennes
le 14 juin 2017**

devant le jury composé de :

Eva LÖCHERBACH

Professeur à l'Université de Cergy-Pontoise /
rapporteur

Aihua FAN

Professeur à l'Université de Picardie Jules Verne /
rapporteur

Mihai GRADINARU

Professeur à l'Université de Rennes 1 / *examineur*

Dimitri PETRITIS

Professeur à l'Université de Rennes 1 / *directeur de
thèse*

Table des matières

1	Introduction	3
1.1	Applications des systèmes de fonctions itérées	4
1.1.1	Aspects combinatoires et dynamiques	5
1.1.2	Aspects probabilistes des IFS	6
1.1.3	Applications dans d'autres disciplines	9
1.2	Mesure invariante et IFS	11
1.3	Dimension de Hausdorff et IFS	15
1.4	Une application particulière des IFS : les chaînes de Markov d'ordre infini et les chaînes de Markov de longueur variable	17
2	Asymptotic stability of IFS and Hausdorff dimension (Article accepté dans Chaos, Solitons and Fractals sous réserve)	22
2.1	Introduction	23
2.2	Preliminaries and notation	25
2.3	Technical results	28
2.4	Invariant measure(s) of an IFS	32
2.5	Hausdorff dimension	41
2.6	Conclusion and open problems	48

3	A criterion of asymptotic stability for Variable Length Markov Chains (Article à soumettre)	49
3.1	Introduction	50
3.2	Preliminaries and notation	52
3.2.1	Variable Length Markov Chain and context tree	52
3.2.2	Asymptotic stability for VLMC related to a finite context tree . .	53
3.3	Asymptotic stability of a VLMC	55
3.4	Conclusion and open problems	59
4	Conclusion et problèmes ouverts	60
A	Dimensions fractales	62

Chapitre 1

Introduction

L'objet de cette thèse est l'étude de mesure(s) invariante(s) pour des systèmes de fonctions itérées. Commençons par définir ces éléments dans le cas général. On considère un espace métrique séparable complet X et un alphabet fini \mathcal{A} de cardinal $D \geq 2$ que nous noterons, sans perte de généralité, $\mathcal{A} = \{0, \dots, D-1\}$. Soit $(S_i)_{i \in \mathcal{A}}$ une famille de fonctions boréliennes de X dans X et $(p_i)_{i \in \mathcal{A}}$ une partition de l'unité sur X en des fonctions boréliennes positives. En d'autres termes, pour tout $i \in \mathcal{A}$, pour tout $x \in X$, on a $p_i(x) \geq 0$ et $\sum_{i \in \mathcal{A}} p_i(x) = 1$. Un ensemble $\mathcal{S} = \{(S_i, p_i), i \in \mathcal{A}\}$ est appelé système de fonctions itérées (IFS = iterated function system) avec probabilités dépendant de la position (=place-dependent probabilities). L'évolution d'un IFS se décrit comme suit : si à un temps $n \in \mathbb{N}$, le système se trouve dans un état $x \in X$, au temps $n+1$, il se trouvera dans l'état $S_i(x)$ avec une probabilité $p_i(x)$. La probabilité de transition $p_i(x)$ de l'état x vers l'état $S_i(x)$ peut donc dépendre de la position du système. On peut associer à l'IFS \mathcal{S} une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à valeurs dans X et de noyau de transition P défini par :

$$P(x, B) = \mathbb{P}(X_{n+1} \in B | X_n = x) = \sum_{i \in \mathcal{A}} p_i(x) \mathbb{1}_B(S_i(x)),$$

pour tout $x \in X$ et tout borélien $B \subset X$. En d'autres termes, si $(X_n)_n$ est une chaîne de Markov sur X de noyau de transition P , alors pour tout $x \in X$, tout $n \in \mathbb{N}$ et tout $i \in \mathcal{A}$,

$$\mathbb{P}(X_{n+1} = S_i(x) | X_n = x) = p_i(x).$$

L'opérateur de Markov P agit :

- de manière covariante sur l'espace des mesures boréliennes i.e. pour toute mesure borélienne μ , pour tout borélien $B \subset X$, on a :

$$\mu P(B) = \sum_{i \in \mathcal{A}} \int_{S_i^{-1}(B)} p_i(x) \mu(dx);$$

– de manière contravariante sur l'espace des fonctions f mesurables bornées :

$$Pf(x) = \sum_{i \in \mathcal{A}} p_i(x) f \circ S_i(x).$$

On peut aussi considérer le processus Markovien dilaté $(\xi_n)_{n \in \mathbb{N}}$ à valeurs dans $X \times \mathcal{A}$. L'alphabet \mathcal{A} étant fini, on peut le munir de sa tribu exhaustive. Par conséquent, il suffit de donner le noyau K de $\xi_n = (X_n, A_n)$. Pour tout $x \in X$, tout $B \subset X$, tous $(i, j) \in \mathcal{A}^2$,

$$K((x, i), B \times \{j\}) = \mathbb{P}(\xi_{n+1} \in B \times \{j\} | \xi_n = (x, i)) = p_j(x) \mathbb{1}_B(S_j(x)).$$

On constate alors que (X_k) est le processus obtenu de (ξ_n) marginalisé par intégration des observations relatives aux lettres (A_n) .

Ainsi, le comportement d'un IFS est déterminé par le couplage de trois types d'information. La première est une information *combinatoire* obtenue comme l'ensemble des trajectoires α infinies découlant d'un automate fini déterministe (cf. un exemple en figure 1.1). Il s'agit des mots $\alpha \in \mathcal{A}^{\mathbb{N}} = \{\alpha_1 \alpha_2 \cdots, \alpha_i \in \mathcal{A}\}$.

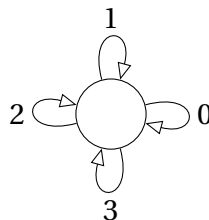


FIGURE 1.1 – Automate pour un alphabet à 4 éléments $\mathcal{A} = \{0, 1, 2, 3\}$

Cette information combinatoire induit une information *dynamique* $(S^{(n)}[\alpha])_n$ où $X \ni x \mapsto S^{(n)}[\alpha](x) := S_{\alpha_1} \circ \cdots \circ S_{\alpha_n}(x)$. La troisième information est une information *probabiliste* codée dans $(p_i(x))_{i \in \mathcal{A}, x \in X}$. Cette information induit une probabilisation des trajectoires α par $p_{\alpha_1}(x) p_{\alpha_2}(x) \cdots$ qui à son tour induit une probabilité \mathbb{P} sur $X^{\mathbb{N}}$.

1.1 Applications des systèmes de fonctions itérées

Avant de nous intéresser aux critères permettant de déterminer l'existence et l'unicité d'une mesure invariante pour les IFS, nous allons nous pencher sur différentes applications possibles des systèmes de fonctions itérées.

1.1.1 Aspects combinatoires et dynamiques

L'intérêt pour les IFS est apparu grâce à leur lien avec la théorie des fractals introduite par Mandelbrot [44] en 1975, pour désigner des ensembles qui sont généralement hautement irréguliers. De nombreuses définitions des fractals existent mais nous considérerons ici qu'un fractal est un ensemble qui possède une certaine forme d'auto-similarité, i.e. un ensemble composé de copies réduites de lui-même. Citons par exemple l'ensemble tri-adique de Cantor ou la courbe de von Koch.

Certains systèmes de fonctions itérées constituent un moyen de « créer » des fractals comme l'invariant d'une application F définie pour tout compact $A \subset X$ par

$$F(A) = \bigcup_{i \in \mathcal{A}} S_i(A).$$

L'application F est parfois appelée opérateur de Barnsley. En 1981, Hutchinson [30] a prouvé que si les fonctions S_i sont des contractions de X dans X , alors pour tout compact $A \subset X$, la limite de la suite $(F^n(A))_{n \in \mathbb{N}}$ où $F^n(A) = \bigcup_{\alpha \in \mathcal{A}^n} S^{(n)}[\alpha](A)$ existe, est indépendante du compact A et est un ensemble fractal (au sens large). La limite de ces suites, que nous noterons K_X , est appelée l'attracteur de l'IFS \mathcal{S} . On peut remarquer que l'ensemble fractal est défini uniquement à partir de la partie dynamique (déterministe) de l'IFS.

L'ensemble tri-adique de Cantor peut par exemple être obtenu comme l'invariant de l'IFS défini sur $X = [0, 1]$ par des homothéties de centres 0 et 1 et de rapport $\frac{1}{3}$. Autrement dit, l'IFS est constitué des transformations $S_0(x) = \frac{x}{3}$ et $S_1(x) = \frac{x+2}{3}$.

La courbe de von Koch peut quant à elle être obtenue de la manière suivante. On considère les points du plan $a_1(0, 0)$, $a_2\left(\frac{1}{3}, 0\right)$, $a_3\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$, $a_4\left(\frac{2}{3}, 0\right)$ et $a_5(1, 0)$. Ce sont les sommets obtenus à la suite de la première itération lorsqu'on veut construire la courbe de Koch. Soit $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$ l'IFS tel que $S_i : [0, 1]^2 \rightarrow [0, 1]^2$ est l'unique similitude directe qui envoie le vecteur $\overrightarrow{a_1 a_5}$ sur le vecteur $\overrightarrow{a_i a_{i+1}}$. L'attracteur de cet IFS est la courbe de von Koch.

Les attracteurs des IFS sont en général des ensembles hautement irréguliers qui ne peuvent pas être différenciés par leurs simples propriétés topologiques. Un arsenal de propriétés combinant des propriétés métriques avec des propriétés topologiques, connu sous le nom collectif de dimensions fractales (dimension de Hausdorff, dimension de boîtes, dimension d'entassement, etc.) ont été introduites afin de différencier des objets indiscernables par leurs propriétés topologiques.

1.1.2 Aspects probabilistes des IFS

Dans cette partie, nous allons voir que des IFS probabilistes peuvent apparaître dans l'étude d'autres objets mathématiques. Nous présentons ici quelques exemples : les marches aléatoires sur des graphes dirigés ou des pavages apériodiques, les systèmes dynamiques aléatoires. Un survol des applications des IFS peut se trouver dans [49]. Les chaînes de Markov d'ordre infini et d'ordre variable constituent deux autres exemples qui seront traités plus longuement par la suite.

Marches aléatoires sur des graphes dirigés ou non de \mathbb{Z}^d Dans ce paragraphe, nous nous intéressons à deux types de marches aléatoires : la marche aléatoire symétrique simple sur \mathbb{Z}^d et des marches aléatoires sur des graphes dirigés sur \mathbb{Z}^2 .

Les marches aléatoires symétriques simples sur \mathbb{Z}^d avec $d \geq 1$ peuvent être vues comme des systèmes de fonctions itérées. En effet, on peut considérer l'espace $X = \mathbb{Z}^d$ et l'alphabet $\mathcal{A} = \{\pm e_1, \dots, \pm e_d\}$, les e_j étant les vecteurs de la base canonique de \mathbb{Z}^d . Pour tout $i \in \mathcal{A}$, on définit l'application $X \ni x \mapsto S_i(x) := x + i$ ainsi que les probabilités de transition $p_i(x) := \frac{1}{2d}$ pour tout $x \in X$. Alors, l'IFS $\mathcal{S} = \{(S_i, p_i), i \in \mathcal{A}\}$ est équivalent à la marche aléatoire symétrique simple sur \mathbb{Z}^d introduite par Pólya en 1921 [50] et dont le comportement a été démontré dans ce papier séminal : elle est récurrente sur \mathbb{Z}^d pour $d \leq 2$ et est transiente pour $d \geq 3$. On peut noter que les marches aléatoires simples ont des applications physiques. En effet, tout phénomène décrit par une équation différentielle faisant intervenir un laplacien peut être représenté par une marche aléatoire.

De même, les marches aléatoires sur des graphes dirigés sur \mathbb{Z}^2 peuvent s'exprimer comme des IFS. En effet, on considère comme précédemment $X = \mathbb{Z}^2$, $\mathcal{A} = \{\pm e_1, \pm e_2\}$, les transformations $S_i(x) = x + i$ mais les probabilités de transition choisies p_i vont cette fois-ci dépendre de la position $x = (x_1, x_2) \in \mathbb{Z}^2$ du système. On peut alors s'intéresser à trois exemples de graphes dirigés (figure 1.2) avec une unique direction possible horizontalement qui ont été étudiés du point de vue mathématique pour la première fois dans [8].

Dans le premier graphe, les lignes horizontales sont orientées vers l'est ou l'ouest alternativement. Dans le deuxième graphe, les lignes de la partie supérieure à l'axe des abscisses sont orientées vers l'ouest et celles qui sont strictement inférieures sont orientées vers l'est. Dans le troisième graphe, les directions horizontales sont données de manière aléatoire en suivant une loi uniforme. σ_{x_2} est donc une variable aléatoire uniforme à valeur dans $\{-1, 1\}$.

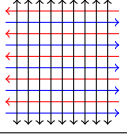
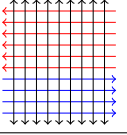
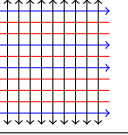
Graphe alterné horizontalement	Graphe à directions opposés sur les moitiés horizontales du plan	Graphe à direction horizontale aléatoire
		
$p_i(\mathbf{x}) = \begin{cases} 1/3 & \text{si } i = \pm e_2 \\ 1/3 & \text{si } i = e_1 \text{ et } x_2 \in 2\mathbb{Z} \\ 1/3 & \text{si } i = -e_1 \text{ et } x_2 \in 2\mathbb{Z} + 1 \\ 0 & \text{sinon.} \end{cases}$	$p_i(\mathbf{x}) = \begin{cases} 1/3 & \text{si } i = \pm e_2 \\ 1/3 & \text{si } i = e_1 \text{ et } x_2 < 0 \\ 1/3 & \text{si } i = -e_1 \text{ et } x_2 \geq 0 \\ 0 & \text{sinon.} \end{cases}$	$p_i(\mathbf{x}) = \begin{cases} 1/3 & \text{si } i = \pm e_2 \\ 1/3 & \text{si } i = \sigma_{x_2} e_1 \\ 0 & \text{sinon.} \end{cases}$

FIGURE 1.2 – Trois graphes ayant une unique direction horizontalement et les probabilités de transition p_i associées. Pour le troisième graphe, σ_{x_2} désigne un signe aléatoire.

L'étude de l'existence d'une mesure invariante pour ces graphes peut se trouver dans [8]. La marche aléatoire sur le premier graphe est récurrente alors que celle sur les graphes 2 et 3 est presque sûrement transiente.

On peut noter que, bien que les marches aléatoires sur \mathbb{Z}^d sont étudiées depuis une centaine d'années, les marches aléatoires sur les graphes dirigés ne sont quant à elles étudiées du point de vue mathématique que depuis 2003. Pourtant, ces modèles correspondent plus à la réalité si l'on souhaite modéliser des flux d'information sur internet ou un trafic routier dans un réseau urbain.

Marches aléatoires simples sur des pavages apériodiques Un autre exemple de marches aléatoires que l'on peut considérer comme un IFS sont les marches aléatoires sur des réseaux induits par des pavages apériodiques tels que celui de Penrose.

Nous allons tout d'abord nous pencher sur la construction de pavages et réseaux apériodiques. Pour cela, on considère une décomposition de l'espace euclidien \mathbb{R}^N en une somme directe de deux espaces orthogonaux E et E' tels que $\dim E = d$ et $\dim E' = N - d$. Nous nous intéresserons à deux exemples particuliers : celui avec $N = 2$ et $d = 1$ (exemple 1) et celui avec $N = 5$ et $d = 2$ (exemple 2). Notons C le cube unité de \mathbb{R}^N i.e. $C = \{y \in \mathbb{R}^N \mid y = \sum_{i=1}^N y_i e_i, y_i \in [0, 1]\}$ où les $(e_i)_{1 \leq i \leq N}$ sont les vecteurs de la base canonique de \mathbb{R}^N . Soit $\mathbf{B} = C + E$. Remarquons que pour l'exemple 1, \mathbf{B} est simplement la bande obtenue en traduisant le cube unité parallèlement à la droite E . On projette alors tous les points du réseau \mathbb{Z}^N inclus dans \mathbf{B} sur l'espace E . Si les directions principales de E sont incommensurables par rapport à \mathbb{Z}^N (i.e. pour l'exemple 1, E est une droite irrationnelle et pour l'exemple 2, E est un plan orienté irrationnellement), alors les points de $\mathbf{B} \cap \mathbb{Z}^N$ sont en bijection avec leur pro-

jection orthogonale sur E . Cette projection définit un pavage apériodique de E par un nombre fini de tuiles. Pour l'exemple 1, les tuiles sont des segments de deux longueurs différentes et pour l'exemple 2, ce sont deux types de losanges fins et épais présents selon cinq orientations différentes, ce qui représente dix tuiles différentes. Ce deuxième pavage est celui de Penrose. Les sommets des losanges forment un réseau sur \mathbb{R}^2 . Plus généralement, les sommets des tuiles forment un réseau sur \mathbb{R}^d .

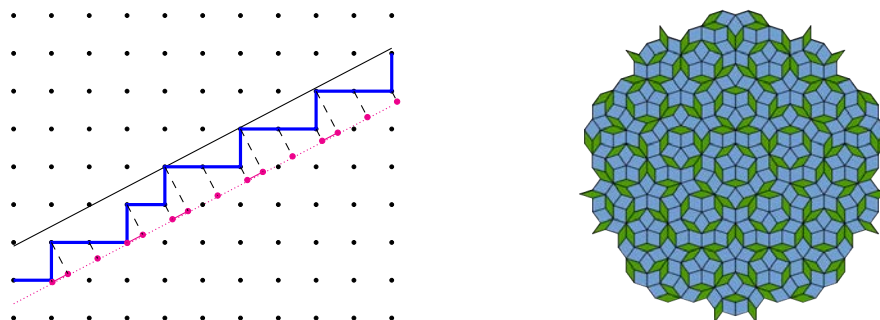


FIGURE 1.3 – Pavages apériodiques de la droite et du plan.

On considère alors la marche aléatoire sur les sommets au plus proche voisin avec probabilité uniforme. On peut alors montrer que la marche aléatoire est récurrente pour $d \leq 2$ et transiente pour $d \geq 3$ [12].

Cette marche aléatoire sur le réseau de Penrose est isomorphe à la marche aléatoire induite par l'IFS avec probabilités dépendant de la position en considérant l'espace $X = \mathbb{Z}^N$, l'alphabet $\mathcal{A} = \{\pm e_1, \dots, \pm e_N\}$ et

$$p_i(x) = \begin{cases} \frac{1}{d(x)} & \text{if } x + i \in \mathbf{B} \\ 0 & \text{otherwise,} \end{cases}$$

où pour $x \in \mathbf{B}$, on définit $d(x) := \#\{j \in \mathcal{A}, x + j \in \mathbf{B}\}$.

Systèmes dynamiques aléatoires Soit (X, d) un espace métrique et $\mathcal{B}(X)$ sa tribu borélienne. Soit Γ une famille de transformations $(S_i)_{i \in \mathcal{A}}$ de X dans X et soit Q un

noyau de probabilité sur $(\mathcal{A} \times X)$. Un système dynamique aléatoire est alors décrit par un triplet (X, Γ, Q) .

Soit $(S_{\alpha_n})_{n \in \mathbb{N}}$ une suite de fonctions aléatoires de Γ et X_0 une variable aléatoire à valeurs dans X , indépendante de la suite $(S_{\alpha_n})_{n \in \mathbb{N}}$. On définit par récurrence

$$X_1 = S_{\alpha_1}(X_0) \quad \text{et} \quad X_{n+1} = S_{\alpha_{n+1}}(X_n) = S_{\alpha_{n+1}} \circ \cdots \circ S_{\alpha_1}(X_0).$$

Alors X_n est un processus de Markov homogène dont la loi est donnée pour $x \in X$ et $B \in \mathcal{B}(X)$ par

$$P(x, B) = \sum_{i \in \mathcal{A}} \mathbb{1}_{S_i(B)}(x) Q(i, x) = Q((i, x) \in \mathcal{A} \times \{x\} | S_i(x) \in B).$$

En d'autres termes, le système dynamique aléatoire peut s'exprimer comme un IFS de la forme $\mathcal{S} = \{(S_i, Q(i, \cdot)) | i \in \mathcal{A}\}$. Notons aussi que tout processus de Markov peut être représenté comme un système dynamique aléatoire.

On s'intéresse maintenant à un exemple de système dynamique aléatoire : l'itération d'applications lipschitziennes aléatoires. Soient $(S_{\alpha_n})_n$ une suite d'applications lipschitziennes aléatoires de Γ i.i.d. Soit, pour $k \geq j$, L_j^k le coefficient de Lipschitz aléatoire de $S_{\alpha_k} \circ \cdots \circ S_{\alpha_j}$ i.e.

$$L_j^k(\alpha) := \sup \left\{ \frac{d(S_{\alpha_k} \circ \cdots \circ S_{\alpha_j}(x), S_{\alpha_k} \circ \cdots \circ S_{\alpha_j}(y))}{d(x, y)} \middle| x \neq y \right\}.$$

Soit $(X_n(x))_n$ le processus sur X défini par $X_n(x) = S_{\alpha_n} \circ \cdots \circ S_{\alpha_1}(x)$. Si l'espace (X, d) est compact et s'il existe un entier $r \geq 1$ tel que

$$-\infty \leq \mathbb{E}(\log L_1^r) < 0,$$

alors le processus de Markov $(X_n)_n$ admet une unique mesure de probabilité invariante. De plus, cette mesure est attractive. La démonstration de ce résultat figure dans [3].

1.1.3 Applications dans d'autres disciplines

Nous allons maintenant nous pencher sur des applications des IFS en biologie avec la classification de protéines et en physique avec les mesures quantiques répétées.

Classification de protéines Les systèmes de fonctions itérées et plus particulièrement les chaînes de Markov de longueur variable ont été utilisés ces dernières années sous le nom de Probabilistic Suffix Trees (PST) pour modéliser des données scientifiques en biologie. En effet, l'un des problèmes en bio-informatique est de déterminer la fonction biologique ou la conformation stéréo-chimique d'une protéine à partir de la séquence d'acides aminés qui la composent, dans le but de les classer. L'utilisation des PST est due à leur capacité à extraire les informations structurales des séquences. Dans un modèle PST pour une famille de protéines, l'alphabet \mathcal{A} représente un ensemble d'acides aminés et on s'intéresse aux séquences d'acides aminés de la protéine étudiée. Dans [41], Léonardi et Galves utilisent un modèle plus évolué – un SPST pour Sparse Probabilistic Suffix Tree – dans lequel certains acides aminés sont regroupés dans une classe d'équivalence : l'alphabet n'est alors plus constitué uniquement d'acides aminés mais de séquences d'acides aminés. Les auteurs présentent un algorithme qui identifie les équivalences entre les acides aminés dans différentes positions comme un motif en utilisant des tests de rapports de vraisemblance. Une fois la chaîne de Markov de longueur variable associée à la protéine déterminée, un test permet de décider son appartenance ou non à une famille donnée.

Mesures quantiques répétées L'état d'un système quantique est représenté par un opérateur ρ agissant sur un espace de Hilbert séparable H . L'opérateur ρ est borné, positif (i.e. $\text{Sp}(\rho) \subset \mathbb{R}_+$) – et donc auto-adjoint – de classe trace et a une trace normalisée. L'ensemble de ces opérateurs forme une partie convexe $\mathcal{D}(H)$ contenue dans l'espace de Banach $B(H)$ des opérateurs bornés sur H . La décomposition spectrale d'un opérateur $\rho \in \mathcal{D}(H)$, associant ρ avec sa mesure spectrale, établit l'analogie de $\mathcal{D}(H)$ avec les mesures de probabilités classiques. On se limitera au cas $H = \mathbb{C}^d$ dans la suite.

Une observable quantique X (l'analogie d'une variable aléatoire réelle en classique) est un opérateur auto-adjoint borné sur H . Autrement dit, c'est une matrice auto-adjointe quand $H = \mathbb{C}^d$. Sa décomposition spectrale $X = \sum_{i \in \mathcal{A}} x_i E_i$ établit une bijection entre X et la famille $(E_i)_{i \in \mathcal{A}}$ de ses projecteurs spectraux. La mécanique quantique étant intrinsèquement probabiliste, la mesure d'une observable est une expérience aléatoire. L'observable X peut prendre n'importe quelle valeur (propre) $x_i \in \text{Sp}(X)$ avec probabilité $\text{tr}(\rho E_i)$.

Les états extrémaux de $\mathcal{D}(H)$ sont ceux qui vérifient $\text{tr}(\rho^2) = \text{tr}(\rho) = 1$. Cette égalité impose en effet que la mesure spectrale de ρ se concentre sur une seule valeur spectrale de ρ ; ρ est alors un projecteur sur le sous-espace propre correspondant. Ceci est l'analogie quantique d'une mesure de probabilité discrète dont la combinaison

convexe des masses de Dirac dégénère en une unique masse de Dirac.

Une caractéristique totalement contre-intuitive de la mécanique quantique, découlant de son formalisme, est que l'observation des résultats d'une observable change irréversiblement l'état du système. Si $(E_i)_{i \in \mathcal{A}}$ est une décomposition de l'identité $I_{\mathbb{C}^d}$ en une somme finie de matrices positives (au sens que $E_i = A_i^* A_i$ pour tout $i \in \mathcal{A}$ avec $\sum_{i \in \mathcal{A}} A_i^* A_i = I_{\mathbb{C}^d}$) mais pas nécessairement à valeurs projectives (on peut avoir $E_i^2 < E_i$) alors $(E_i)_{i \in \mathcal{A}}$ correspond à une observation généralisée. Partant d'un état ρ , mesurer partiellement le système quantique équivaut à assigner au système l'état a posteriori $\rho_i = \varphi_i(\rho) := \frac{A_i \rho A_i^*}{\text{tr}(A_i \rho A_i^*)}$, la valeur i étant observée avec probabilité $p_i(\rho) = \text{tr}(\rho E_i) = \text{tr}(A_i \rho A_i^*)$.

Il s'ensuit que des mesures répétées sur un système quantique induisent une chaîne de Markov classique sur $\mathcal{D}(H)$ dont l'opérateur de Markov est donnée pour tout $D \subset \mathcal{B}(D(\mathcal{H}))$ par

$$\mathbb{P}(p_{n+1} \in D | p_n = x) = \sum_{i \in \mathcal{A}} p_i(x) \mathbb{1}_D(\varphi_i(x))$$

où $p_i(x) = \text{tr}(x E_i) = \text{tr}(A_i x A_i^*)$.

Autrement dit, les mesures répétées sur un système quantique induisent un IFS sur l'espace $\mathcal{D}(H)$.

On peut noter que cette chaîne de Markov classique a une origine purement quantique. En effet, en classique, $\varphi_i(\rho)$ est la probabilité conditionnelle après avoir observé i comme résultat de l'expérience aléatoire et $\sum_i p_i(\rho) \varphi_i(\rho) = \rho$ par la formule des probabilités totales. Ici, la non-commutativité empêche de récupérer ρ après la mesure. On peut se référer à [43] et pour [42] plus de détails.

1.2 Mesure invariante et IFS

Considérons un IFS avec les transformations S_i qui sont des contractions et des probabilités de transition p_i constantes. On peut montrer que si les $p_i \in]0, 1[$, alors l'IFS admet une unique mesure de probabilité invariante et son support est l'attracteur de l'IFS. Ce résultat est dû à Hutchinson [30] en 1981. Nous avons pu constater au travers des différentes applications présentées dans la section précédente que les systèmes de fonctions itérées avec probabilités constantes n'ont pas été les seuls à être étudiés. Le cas des IFS (et de leurs opérateurs de Markov associés) avec des probabilités de transition dépendant de la position a aussi été considéré. Plusieurs critères de stabilité asymptotique ont en effet été donnés en affaiblissant au fur et à mesure les hypothèses en termes d'espaces (espace compact, localement compact, σ -compact puis

complet), en termes de transformations (contractantes ou lipschitziennes contractantes en moyenne) et en termes de probabilités de transition (lipschitziennes, α -holdériennes puis Dini-continues). Différentes hypothèses ont aussi été formulées sur l'opérateur de Markov telles que la non-expansivité ou la concentration globale ou locale par exemple.

Le premier article concernant les probabilités dépendant de la position semble (sous cette appellation) être dû à Barnsley et al. [1] en 1988. Ces auteurs ont étudié l'existence et l'attractivité des mesures invariantes dans le cas où les probabilités de transition p_i n'étaient pas constantes, les transformations S_i n'étaient pas nécessairement des contractions et X n'était pas compact. Ils sont parvenus à montrer l'existence d'une unique mesure invariante attractive sous les hypothèses suivantes : X est localement compact, le module de continuité des p_i satisfait une condition de Dini et les transformations S_i sont lipschitziennes et satisfont une condition de contraction en moyenne entre deux points quelconques. Le point clé pour démontrer leur résultat est l'utilisation du théorème de représentation de Riesz.

Lasota, Yorke, Szarek et Stenflo ont eux aussi contribué à faire avancer l'étude de la stabilité asymptotique des opérateurs de Markov dans les années 1990–2000. Ils utilisent essentiellement les propriétés de contraction (au sens large) et de concentration d'un opérateur P .

Définition 1.2.1. *Un opérateur de Markov P est dit contractant (au sens large) si pour toutes mesures de probabilité μ_1 et μ_2 , on a $\|\mu_1 P - \mu_2 P\| \leq \|\mu_1 - \mu_2\|$ où la norme choisie est celle de Fortet-Mourier.*

Condition 1.2.2. *(Condition classique de concentration) Pour tout $\varepsilon > 0$, il existe un borélien $A \subset X$ de diamètre inférieur à ε et un réel $\alpha > 0$ tels que pour toute mesure de probabilité μ ,*

$$\liminf_{n \rightarrow \infty} \mu P^n(A) \geq \alpha.$$

Dans le cas localement compact et σ -compact, Lasota et Yorke [40] ont démontré en 1994 que si l'opérateur de Markov est contractant et vérifie la condition 1.2.2, alors il est asymptotiquement stable. Dans les années 2000, Szarek s'est intéressé aux mesures invariantes des IFS dans les espaces polonais. Il ne pouvait cependant pas étendre les méthodes utilisant le théorème de Riesz car dans un espace polonais, une fonctionnelle positive peut ne pas correspondre à une mesure. Il utilise alors la notion de tension et le théorème de Prokhorov et prouve ainsi dans [55] la stabilité asymptotique d'un opérateur de Markov vérifiant les mêmes hypothèses que celles données par Lasota et Yorke.

En 2002, Jaroszewska [35] cherche à étudier le comportement asymptotique des IFS

avec un nombre fini de contractions et des probabilités dépendant de la position continues (sans hypothèses plus fortes) et strictement positives. En 1953, Karlin [38] avait conjecturé qu'un IFS avec deux contractions affines et des probabilités continues et strictement positives était asymptotiquement stable. Or, en 2001, Stenflo [53] a montré que la conjecture de Karlin était erronée en construisant un exemple d'IFS avec des contractions affines qui admettait deux mesures de probabilités invariantes, mettant ainsi en exergue le défaut de la conjecture. Plus précisément, il a montré le théorème qui pourrait s'exprimer comme suit

Théorème 1.2.3. *Soient S_0 et S_1 des applications définies de $[0, 1]$ dans lui-même par*

$$S_0(x) = \frac{x}{3} \quad \text{et} \quad S_1(x) = \frac{x+2}{3}.$$

Alors il existe une fonction continue réelle $p : [0, 1] \rightarrow]0, 1[$ telle que l'IFS $\{(S_0, p(x)), (S_1, 1-p(x))\}$ admet plus d'une mesure invariante.

Jaroszewska [35] est parvenue à montrer que dans un espace métrique complet, un IFS dont les transformations S_i sont des contractions et les probabilités de transitions p_i sont continues strictement positives admet une mesure invariante qui peut ne pas être unique comme l'a montré par exemple Stenflo.

Une poursuite naturelle de tous ces travaux est naturellement l'étude de la stabilité asymptotique pour des IFS avec des probabilités de transition discontinues. Cependant, aucune des méthodes utilisées dans les précédentes études ne semble pouvoir s'étendre de part la perte de la continuité de l'opérateur de Markov. En effet, même si on se restreint à des espaces compacts, l'utilisation du théorème de Riesz n'est plus possible car si les probabilités de transition p_i sont discontinues, l'opérateur de Markov n'est plus Fellérien. D'autre part, si on veut utiliser le concept de tension et le théorème de Prokhorov pour des opérateurs de Markov sur des espaces compacts (même si Szarek les avait introduits pour traiter le cas des espaces polonais), on ne parvient pas à prouver l'existence d'une mesure invariante puisque l'opérateur de Markov ne sera pas contractant pour des probabilités de transitions discontinues. La contraction est une conséquence du caractère fellérien dont on ne dispose pas dans ce cas là (une preuve de ce résultat peut se trouver dans [40]).

Un premier pas a été effectué par Jaroszewska en 2013 dans [36] où elle prouve la stabilité asymptotique d'un IFS particulier. Ce résultat est résumé dans le

Théorème 1.2.4. *Soit $\mathcal{S} = \{(S_i, p_i), i \in \{0, 1\}\}$ un IFS sur $X = [0, 1]$ tel que*

$$S_0(x) = \frac{x}{2} \quad \text{et} \quad p_0(x) = \rho_{00} \mathbb{1}_{[0, \frac{1}{2})}(x) + \rho_{01} \mathbb{1}_{[\frac{1}{2}, 1]}(x),$$

$$S_1(x) = \frac{x+1}{2} \quad \text{et} \quad p_1(x) = \rho_{10} \mathbb{1}_{[0, \frac{1}{2})}(x) + \rho_{11} \mathbb{1}_{[\frac{1}{2}, 1]}(x),$$

avec $\rho_{00} + \rho_{01} = \rho_{10} + \rho_{11} = 1$ et $\rho_{ij} \geq 0$ pour $i, j \in \{0, 1\}$. Si tous les $\rho_{ij} \in]0, 1[$, alors l'IFS est asymptotiquement stable.

Elle se sert pour cela du théorème de Schauder-Tychonov mais la méthode qu'elle utilise semble être particulière à ce cas et ne pas pouvoir se généraliser.

L'un des objectifs de cette thèse a donc été d'étendre les résultats de Jaroszewska en donnant un critère d'existence et d'unicité d'une mesure invariante pour une classe d'IFS plus large. Pour cela, il a fallu considérer des transformations S_i qui sont des bijections contractantes sur des intervalles fermés de \mathbb{R} dont les images ne se chevauchent pas ou alors seulement en un point ainsi que des probabilités de transitions qui sont adaptées aux fonctions S_i . La notion d'adaptation signifie que les probabilités de transition p_i sont constantes sur les intervalles $S_{\alpha_1} \circ \dots \circ S_{\alpha_M}(X)$ où M est un entier fixé pour chaque IFS. Sur un tel intervalle, elle est notée $\rho_{i\alpha_1 \dots \alpha_M}$. Les critères de mesure invariante sont établis sous des conditions de non-nullité faible, à savoir essentiellement la stricte positivité d'au moins une des probabilités de transition. Cette condition sera précisée par la suite (voir (2.3.3)). Sans donner plus précisément les hypothèses, on énonce les critères d'existence et d'unicité de mesure invariante ainsi que celui de stabilité asymptotique.

Théorème 1.2.5. *Sous l'hypothèse (2.4.1) qui concerne la structure de l'IFS, l'IFS $\{(S_i, p_i), i \in \mathcal{A}\}$ admet une mesure de probabilité invariante μ qui est unique si on ajoute la condition de faible non-nullité (2.3.3). De plus, une mesure invariante peut être exprimée sur une famille d'ensembles qui engendrent la tribu borélienne sur K_X .*

La preuve de ce théorème repose principalement sur l'utilisation du système symbolique associé aux contractions, un alphabet étendu et des théorèmes usuels sur les chaînes de Markov. Elle est donnée dans le chapitre 2.

De plus, sous la même hypothèse et la condition de non-nullité qui garantit l'unicité de la mesure invariante, on a alors le résultat :

Théorème 1.2.6. *Sous la même hypothèse et la condition de faible non-nullité (2.3.3), l'IFS est asymptotiquement stable i.e. pour tout $\mu' \in \mathcal{M}_1(X)$, on a la convergence étroite*

$$\mu' P^n \xrightarrow[n \rightarrow \infty]{} \mu.$$

L'idée principale consiste à utiliser le théorème de Portmanteau et à montrer que nous avons la convergence de la mesure itérée $(\mu' P^n)_n$ pour tout ensemble borélien de μ -continuité.

1.3 Dimension de Hausdorff et IFS

Reprenons tout d'abord les travaux de Hutchinson. Dans [30], il a montré que si les transformations S_i sont des contractions et les probabilités de transitions associées sont constantes et vérifient $p_i \in]0, 1[$, alors le support de la mesure invariante est l'attracteur de l'IFS et donc un ensemble fractal. Nous allons donc nous intéresser à une dimension fractale de la mesure invariante des IFS. Comme nous l'avons évoqué précédemment, plusieurs dimensions fractales existent pour différencier des ensembles dont les propriétés topologiques ne le permettent pas. Dans toute la suite, nous nous concentrerons uniquement sur la dimension de Hausdorff.

Avant de donner les résultats sur la dimension de Hausdorff de la mesure invariante obtenue dans le théorème (1.2.5), nous allons nous pencher sur les résultats obtenus auparavant.

Une condition souvent utilisée pour parvenir à calculer la dimension de Hausdorff est appelée l'Open Set Condition. Cette condition technique a pour but de séparer les images des transformations sur un ouvert $O \subset X$ pour qu'elles ne se chevauchent pas.

Condition 1.3.1. (*Open Set Condition*) Un IFS $\mathcal{S} = \{(S_i, p_i), i \in \mathcal{A}\}$ satisfait l'Open Set Condition (OSC) s'il existe un ouvert O non-vide tel que

$$\bigcup_{i \in \mathcal{A}} S_i(O) \subset O \quad \text{et} \quad S_i(O) \cap S_j(O) = \emptyset \quad \text{pour } i \neq j.$$

Elle est par exemple utilisée dans le cas suivant. Si les S_i sont des similitudes de rapport r_i qui vérifient l'Open Set Condition et $p_i = r_i^s$ où s est l'exposant de similitude, alors la dimension de Hausdorff de l'attracteur (qui est le support de la mesure invariante) est égale à s_0 où s_0 est l'unique solution de l'équation $\sum_i r_i^{s_0} = 1$. Ce résultat est aussi dû à Hutchinson [30].

De même, Falconer [16] utilise cette condition pour calculer la dimension de Hausdorff de l'attracteur d'un IFS déterministe formé de contractions qui ne sont pas des similitudes. Il considère V un intervalle ouvert de \mathbb{R} et S_1, \dots, S_d des contractions sur \bar{V} deux fois différentiables sur V telles qu'il existe des constantes a et b vérifiant $0 < a < |S'_i(x)| < b < 1$ pour tout $x \in V$. Si les S_i satisfont l'OSC avec l'ouvert V , alors la limite

$$\lim_{k \rightarrow \infty} \left[\sum_{\mathcal{A}^k} \left| (S_{\alpha_1} \circ \dots \circ S_{\alpha_k})'(x) \right|^s \right]^{1/k} := g(s)$$

existe pour tout $s > 0$, est indépendante de x et est décroissante en s . Alors si on note K_X l'attracteur de l'IFS, on a $\dim_H K_X = s_0$ où s_0 est la solution de $g(s_0) = 1$.

Dans [19], Fan calcule la dimension de Hausdorff du support de la mesure invariante d'un IFS en considérant une hypothèse plus forte que l'OSC puisqu'il suppose que l'attracteur de l'IFS est totalement disconnecté. De plus, les transformations S_i sont comme pour Hutchinson des similitudes de rapport L_i . Par contre, les probabilités de transition associées p_i , bien que toujours constantes et dans l'intervalle $]0, 1[$, ne sont pas spécifiques. Fan montre que sous ces hypothèses, la dimension de Hausdorff est égale à

$$\dim_H \mu = \frac{\sum_{i \in \mathcal{A}} p_i \log p_i}{\sum_{i \in \mathcal{A}} p_i \log L_i}.$$

De manière plus générale, cela revient à dire que $\dim_H \mu = \frac{h_\mu}{\chi_\mu}$ où h_μ et χ_μ désignent respectivement l'entropie de μ et l'exposant de Lyapunov de l'IFS. Ces notions seront redéfinies dans le chapitre 2. Nous allons voir que, sous d'autres hypothèses, cette formule pour la dimension de Hausdorff est vérifiée ou à défaut qu'elle constitue une borne supérieure.

En effet, en 2002, Myjak et Szarek [45] établissent des hypothèses sur des IFS sous lesquelles le rapport entropie sur exposant de Lyapunov constitue une borne supérieure pour la dimension de Hausdorff du support de la mesure invariante. Ils considèrent X un espace polonais muni d'une distance ρ et des transformations S_i lipschitziennes contractantes en moyenne. La définition considérée n'étant pas la même chez tous les auteurs, nous précisons celle utilisée par Myjak et Szarek.

Définition 1.3.2. (*IFS contractant en moyenne*) Soient S_i des applications lipschitziennes de constantes de Lipschitz L_i . L'IFS est contractant en moyenne s'il existe une constante $\gamma \in]0, 1[$ telle que

$$\prod_{i \in \mathcal{A}} \rho(S_i(x), S_i(y))^{p_i(x)} \leq \gamma \rho(x, y) \quad \text{pour tous } x, y \in X.$$

La constante γ étant liée à l'exposant de Lyapunov, ce type de contraction est aussi appelé contraction de Lyapunov.

Sous une condition de forte non-nullité des probabilités de transition (2.5.2) et une condition de Dini sur leur module de continuité concave, ces auteurs établissent que l'IFS admet une mesure invariante μ dont la dimension de Hausdorff est majorée par $\frac{\ln \delta}{\ln \gamma}$ où $\delta = \inf_{x \in X} \prod_{i \in \mathcal{A}} p_i(x)^{p_i(x)} > 0$. On note que $-\ln \delta$ est une forme d'entropie et on retrouve bien le type de borne supérieure évoquée. Le calcul de cette borne supérieure se fait grâce à des outils d'analyse multifractale. Rappelons que le module de continuité $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ des p_i satisfait une condition de Dini si $\int_0^\sigma \frac{\omega(t)}{t} dt < \infty$ pour un $\sigma > 0$.

Jaroszewska et Rams [37] ont quant à eux déterminé une borne supérieure pour la dimension de Hausdorff égale à une généralisation du rapport h_μ/χ_μ pour des mesures invariantes ergodiques d'IFS sur un espace métrique séparable avec des probabilités de transition strictement positives. De plus, les hypothèses sur l'IFS qu'ils donnent correspondent à une forme faible d'une contraction en moyenne.

Dans [29], Heurteaux et Stos étudient des mesures sur $[0, 1]$ induites par une chaîne de Markov finie qui généralisent les produits de Bernoulli. Cela revient à étudier un IFS sur $X = [0, 1]$ avec D contractions $S_i(x) = \frac{x+i}{D}$ et des probabilités de transition p_i constantes par morceaux et adaptées. Grâce à des outils issus de l'analyse multifractale, ils parviennent à calculer la dimension de Hausdorff du support de l'unique mesure invariante (garantie par une hypothèse d'irréductibilité de la chaîne de Markov associée) et on peut montrer qu'elle est égale au rapport entropie sur exposant de Lyapunov.

Excepté pour les travaux de Heurteaux et Stos, les études précédentes concernent des IFS dont les probabilités de transition sont constantes ou au moins Dini-continues. Le but est donc de déterminer la dimension de Hausdorff de la mesure invariante quand il en existe une unique, pour la classe d'IFS avec probabilités de transition adaptées et dont les hypothèses précises seront données plus tard (2.4.1). Sous des conditions de forte non-nullité et de dérivabilité des transformations, on peut déterminer la dimension de Hausdorff du support de la mesure invariante.

Théorème 1.3.3. *Sous les hypothèses (1.2.5) et les conditions (2.5.1) et (2.5.2), la dimension de Hausdorff de la mesure invariante μ est égale à*

$$\dim_H \mu = \frac{h_\mu}{\chi_\mu}.$$

La preuve de ce théorème, donnée au chapitre 2, est assez technique mais repose principalement le choix de « bonnes » bornes supérieures et inférieures pour la mesure invariante.

1.4 Une application particulière des IFS : les chaînes de Markov d'ordre infini et les chaînes de Markov de longueur variable

Nous nous intéressons maintenant à deux types d'IFS particuliers : les chaînes d'ordre infini et les chaînes de Markov d'ordre variable.

Les chaînes de Markov d'ordre infini Une chaîne de Markov d'ordre infini est une suite de variables aléatoires $(A_n)_{n \in \mathbb{Z}}$ à valeurs dans un alphabet fini \mathcal{A} (on peut fixer $\mathcal{A} = \{0, \dots, D-1\}$) dont la loi est donnée par les probabilités de transitions $q_i(\cdot)$ définies pour tout $i \in \mathcal{A}$ et tout $\alpha \in \mathcal{A}^{\mathbb{N}}$ par :

$$\mathbb{P}(A_n = i | A_{n-1} = \alpha_1, A_{n-2} = \alpha_2, \dots) = q_i(\alpha).$$

Le futur de la chaîne dépend en général de son passé tout entier et c'est pourquoi ce processus est appelé chaîne d'ordre infini ou chaîne à liaisons complètes. Ce type de chaîne a pour la première fois été étudié par Onicescu et Mihoc en 1935 [47] sous le nom de chaînes à liaisons complètes, puis plus tard par Doeblin et Fortet en 1937 [14] et par Ionescu Tulcea et Marinescu en 1948 [31]. En 1955, dans [26], plutôt que de travailler avec des chaînes d'ordre infini, Harris a choisi de plonger la suite $(A_n)_{n \in \mathbb{Z}}$ dans l'intervalle unité. Il a alors obtenu une nouvelle suite $(X_n)_{n \in \mathbb{N}}$ et a étudié des conditions d'existence et d'unicité d'une mesure invariante pour ce nouveau processus. On peut définir X_n comme la variable aléatoire admettant $0, A_{n-1}A_{n-2} \dots$ comme possible représentation dans le système D -aire, i.e.

$$X_n = \sum_{j=1}^{\infty} \frac{A_{n-j}}{D^j}.$$

La loi de $(X_n)_n$ est alors donnée pour tout $x = 0, \alpha_1 \alpha_2 \dots \in [0, 1]$ par

$$\mathbb{P}\left(X_{n+1} = \frac{x+i}{D} \middle| X_n = x\right) = p_i(x),$$

où $p_i(0, \alpha_1 \alpha_2 \dots) := q_i(\alpha)$. Alors, le processus $(X_n)_n$ peut être vu comme la chaîne de l'IFS $\mathcal{S} = \{(S_i, p_i), i \in \mathcal{A}\}$ avec S_i et p_i comme définis précédemment.

La plupart des études sur les chaînes d'ordre infini se sont concentrées sur des hypothèses de continuité et de positivité stricte des probabilités de transition pour assurer l'existence et l'unicité d'une chaîne stationnaire. Dans [54], Stenflo rappelle les conditions supplémentaires données par Harris [26], Iosifescu et Spătaru [33], Comets, Fernández et Ferrari [11] avant d'en exprimer une nouvelle. Toutes ces conditions avaient pour but de contrôler les variations des probabilités de transition. Par exemple, celle de Harris était

$$\sum_{n=0}^{\infty} \prod_{m=0}^n \left(1 - \frac{D}{2} \sup_i \sup_{\alpha_1 \dots \alpha_m = \beta_1 \dots \beta_m} |q_i(\alpha) - q_i(\beta)|\right) = \infty$$

et celle de Stenflo était

$$\sum_{n=0}^{\infty} \prod_{m=0}^n \inf_{\alpha_1 \dots \alpha_m = \beta_1 \dots \beta_m} \sum_{i=0}^{D-1} \min(q_i(\alpha), q_i(\beta)) = \infty.$$

Les chaînes de Markov de longueur variable ou VLMC (= variable length Markov chains) ont été introduites par Rissanen [52] en 1983 comme un outil de compression de données pour des chaînes longues générées par des sources d'information non indépendantes. Plutôt que de considérer un modèle dans lequel la source est une chaîne de Markov d'ordre fixé, Rissanen a considéré que la longueur du passé dont on avait besoin afin de décider du symbole suivant dépendait du passé lui-même. Le modèle des VLMC a deux avantages principaux : il nécessite moins d'espace de stockage que les chaînes de Markov d'ordre fixé et est donc moins coûteux, et il met en évidence les dépendances structurelles entre les données. Les VLMC peuvent être utilisées pour modéliser des chaînes discrètes stochastiques prenant des valeurs dans un alphabet fini : c'est le cas en génomique pour classifier des séquences d'ADN ou des séquences de protéines. On peut trouver une étude générale sur les VLMC dans Galves et Löcherbach [25].

Afin de définir les VLMC, nous avons besoin d'introduire les arbres de contexte. Pour cela, on considère $\mathcal{A} = \{0, \dots, D-1\}$ un alphabet fini. Alors $\bigcup_{n \geq 0} \mathcal{A}^n$ est l'ensemble des mots finis sur \mathcal{A} et $\mathcal{A}^{\mathbb{N}}$ est l'ensemble des mots infinis sur \mathcal{A} . Soit \mathcal{T} un arbre sur \mathcal{A} tel que chaque noeud possède 0 ou D enfants et qui possède un ensemble fini ou dénombrable de feuilles (=contextes) noté $\mathcal{C}(\mathcal{T})$. À chaque contexte $c \in \mathcal{C}(\mathcal{T})$, on associe une mesure de probabilité $q_i(c)$ sur \mathcal{A} telle que $\sum_{i=0}^{D-1} q_i(c) = 1$. Le couple $(\mathcal{T}, (q_i(c))_{c \in \mathcal{C}(\mathcal{T})})$ est alors ce qu'on appelle un arbre de contexte.

On peut maintenant définir la fonction préfixe, notée pref . Si $w \in \mathcal{T}$, alors $\text{pref}(w) := w$. Si $w \notin \mathcal{T}$, $\text{pref}(w)$ est le plus long mot $w_1 \dots w_N \in \mathcal{C}(\mathcal{T})$ tel que w admette $w_1 \dots w_N$ comme préfixe. C'est cette fonction qui permet de saisir la partie pertinente du passé que l'on a besoin de regarder pour prédire l'élément suivant.

La chaîne de Markov de longueur variable associée à l'arbre de contexte est la chaîne $(Y_n)_{n \in \mathbb{N}}$, avec espace d'états $\mathcal{A}^{\mathbb{N}}$, définie par les probabilités de transitions suivantes : pour tout $n \in \mathbb{N}$, pour tout $i \in \mathcal{A}$ et pour tout $\alpha \in \mathcal{A}^{\mathbb{N}}$, on a

$$\mathbb{P}(Y_{n+1} = i \alpha | Y_n = \alpha) = q_i(\text{pref}(\alpha)).$$

On peut aussi étendre la fonction $q_i(\cdot)$ définie sur $\mathcal{C}(\mathcal{T})$ en une fonction $\tilde{q}_i(\cdot)$ sur $\mathcal{A}^{\mathbb{N}}$ en posant pour tout $\alpha \in \mathcal{A}^{\mathbb{N}}$, $\tilde{q}_i(\alpha) := q_i(\text{pref}(\alpha))$.

Alors la chaîne $(Y_n)_n$ peut être vue comme une chaîne de Markov associée à l'IFS $\mathcal{S} = \{(S_i, \tilde{q}_i), i \in \mathcal{A}\}$ sur $\mathcal{A}^{\mathbb{N}}$ où les S_i sont les applications $S_i(\alpha) = i\alpha$.

On s'intéresse maintenant aux études concernant l'existence et l'unicité d'une mesure invariante pour les VLMC. La plupart d'entre elles utilisent les deux mêmes hypothèses : la famille des probabilités de transition $(\tilde{q}_i)_{i \in \mathcal{A}}$ est faiblement positive (ce qui signifie qu'au moins l'une des fonctions \tilde{q}_i a un infimum strictement positif) et le

taux de continuité $(\beta_k)_{k \geq 1}$ de la famille $(\tilde{q}_i)_{i \in \mathcal{A}}$ est sommable. Le taux de continuité est défini par

$$\beta_k = \sup_i \sup_{\alpha_1 \dots \alpha_k = \beta_1 \dots \beta_k} |\tilde{q}_i(\alpha) - \tilde{q}_i(\beta)|.$$

La deuxième hypothèse implique que $\lim_{k \rightarrow +\infty} \beta_k = 0$; la famille de probabilités de transition est donc continue pour la distance ultramétrique. En 2011, Gallo ([23]) a réussi à se passer de l'hypothèse de continuité, en la remplaçant par une hypothèse sur la structure de l'arbre. Mais ce critère ne permettait pas d'étudier tous les types d'arbres. En 2012, Cénac et al. ([10]) ont entièrement déterminé pour deux exemples (dont le peigne infini représenté par la figure (3.1)) des conditions nécessaires et suffisantes pour obtenir une unique mesure invariante.

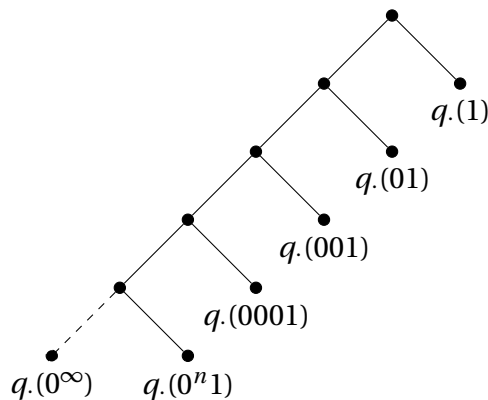


FIGURE 1.4 – Arbre de contexte : le peigne infini

Ensuite, Gallo et Paccaut ([24]) en 2013 se sont intéressés à des probabilités de transition discontinues. Leur critère d'existence se base sur des considérations sur l'ensemble des points de discontinuité. Pour avoir l'unicité, ils ajoutent des conditions sur l'ensemble des points de continuité qui permettent de contrôler les variations. Ils ont aussi besoin d'une condition de positivité stricte sur un sous-ensemble de $\mathcal{A}^{\mathbb{N}}$.

Pour faire suite à ces travaux, on peut énoncer un critère d'existence et d'unicité d'une mesure de probabilité invariante et même de stabilité asymptotique sous des hypothèses de continuité (pour la distance ultramétrique) et de non-nullité faible des probabilités de transitions. Aucune hypothèse de contrôle des variations n'est alors nécessaire.

Théorème 1.4.1. (*Stabilité asymptotique d'une VLMC*) Soit $(\mathcal{T}, (q.(c))_{c \in \mathcal{C}(\mathcal{T})})$ un arbre de contexte et $(Y_n)_n$ une VLMC associée à cet arbre de contexte. On appelle Q l'opérateur de Markov de la VLMC. Sous des conditions de non-nullité faible et de continuité (détaillées en (3.3.1) et en (3.3.2)), il existe une unique mesure de probabilité

$\mu \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$ telle que pour toute mesure $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$,

$$\mu' Q^k \xrightarrow[k \rightarrow \infty]{} \mu.$$

Le principe de la preuve consiste à approcher un arbre de contexte infini par une suite d'arbres de contexte fini et d'utiliser les résultats de stabilité asymptotique pour une VLMC associée à un arbre de contexte fini. La démonstration de la stabilité asymptotique pour un arbre de contexte fini se base sur celle de la stabilité asymptotique de la classe d'IFS donné dans [15]. En effet, le système de fonctions itérées sur l'espace $[0, 1]$ donné par les transformations $S_i(x) = \frac{x+i}{D}$ et des probabilités de transition adaptées fait partie de cette classe. Alors en plongeant le processus de la dernière lettre ajoutée la VLMC dans l'intervalle unité, nous avons presque une bijection avec le processus de Markov donné par l'IFS. Ainsi, les idées de la preuve sont très similaires, excepté que nous n'avons pas besoin de réduire l'ensemble $\mathcal{A}^{\mathbb{N}}$ puisque nous n'avons pas besoin de la bijection dans ce cas là. Enfin, c'est un critère de convergence dominée qui permettra de conclure.

Chapitre 2

Asymptotic stability of IFS and Hausdorff dimension (Article accepté dans Chaos, Solitons and Fractals sous réserve)

A CLASS OF ITERATED FUNCTION SYSTEMS WITH ADAPTED PIECEWISE
CONSTANT TRANSITION PROBABILITIES: ASYMPTOTIC STABILITY AND
HAUSDORFF DIMENSION OF THE INVARIANT MEASURE

Blandine DUBARRY

Abstract

A class of iterated function systems (IFS) with non-overlapping or just-touching contractions on closed real intervals and adapted piecewise constant transition probabilities are studied. We give criteria for the existence and the uniqueness of an invariant probability measure for the IFSs and for the asymptotic stability of the system in terms of bounds of transition probabilities. The proofs are mainly based on the symbolic system associated with the contractions, an extended alphabet and usual theorems for Markov chains. Additionally, in case there exists a unique invariant measure, we obtain its Hausdorff dimension as the ratio of the entropy over the Lyapunov exponent. This result extends the formula, established in the literature for continuous transition probabilities, to the case considered here of piecewise constant probabilities. The main idea of this proof consists in finding good lower and upper bounds for the invariant measure.

Keywords : Iterated function systems; Markov operators; invariant measure; asymptotic stability; Hausdorff dimension; fractals.

Mathematical Subject Classification : 60J05, 37B10, 37A30, 28A78, 28A80.

2.1 Introduction

Both for their rich mathematical structure and for their interest into several applied fields (dynamical systems, inhomogeneous random walks, random walks in random environment, quantum information, fractal geometry, etc.), deterministic and/or probabilistic iterated function systems (IFS) have been the object of many studies. Let us start by the deterministic point of view. If we consider a system of D contraction mappings $\mathcal{S} = \{S_0, \dots, S_{D-1}\}$ on a metric complete separable space X , the limit K_X of the sequence $(F^n(A))_n$, where $F(A) = \bigcup_i S_i(A)$ and A is a compact set, exists, is independent of the choice of the compact A , and is generically a fractal set. The limit set is called the attractor of the IFS \mathcal{S} . Now, if we choose a contraction S_i with constant probability $p_i \in [0, 1]$, we can define the Markov operator corresponding to this system which acts covariantly on the space of Borel measures :

$$\mu P = \sum_i p_i \mu \circ S_i^{-1}.$$

If the maps S_i are still contractions and the probabilities belong to $(0, 1)$ then the IFS admits a unique invariant measure μ and its support is the attractor, K_X , of the IFS and so a fractal set. These criteria of asymptotic stability were proved by Hutchinson ([30]) in 1981. In 1988, Barnsley ([2]) reuses them to improve his algorithm for computing pictures of fractals : the Deterministic Algorithm is replaced by the Random Iteration Algorithm, also known as the Chaos Game, and is still being developed.

Since then, some generalisations of the IFSs have been developed. The ones we will focus on concern the transition probabilities. Even if at the beginning they always were constant, the case of IFSs (and associated Markov operators) with continuous place-dependent probabilities has also been considered. Indeed, several criteria of asymptotic stability have been proved for more and more general spaces X (compact, locally compact, σ -compact, Polish) and for weaker and weaker conditions on transition probabilities (Lipschitz continuous, α -Hölder and then Dini-continuous)(see for instance [1, 40, 55]).

The natural continuation is the study of IFSs with discontinuous place-dependent probabilities. However, all the methods used in the previous studies break down because of the lack of continuity of the Markov operator. Indeed, on compact spaces,

the usual method is to use the Riesz theorem, which is not possible since in the discontinuous case, the Markov operator is not Fellerian. On the other hand, if we use the concept of tightness and the Prokhorov theorem for Markov operators on compact spaces (even if Szarek [55] developed the method to establish criteria on Polish spaces), we cannot prove the existence of an invariant measure because our operator will never be non expansive (which means that it can increase the Fortet-Mourier distance between two measures) in the discontinuous case. Consequently, other methods need to be found. A first step was done by Jaroszewska ([36]) in 2013 who proved the asymptotic stability of one particular IFS on $[0, 1]$ $\mathcal{S} = \{(S_i, p_i), i = 0, 1\}$ with $S_i(x) = (x + i)/2$ and adapted piecewise constant probabilities under the assumption of strict positivity of probabilities. The notion of adaptation will be explained later. In a first part of this paper, we will extend Jaroszewska's results by giving a criterion of existence and uniqueness of an invariant measure for a larger class of IFSs. Indeed, our IFSs consist of bijective contractions on closed intervals of \mathbb{R} which are non overlapping or just touching and with adapted transition probabilities. Our hope is that we can eventually extend the results to transition probabilities obtained as limits of piecewise constant functions.

Here we present the main results without expliciting the technical conditions (2.3.3, 2.5.1, 2.5.2) which will be given in detail later.

Theorem 2.1.1. *Under the global assumption on the structure of the IFS (see 2.4.1 below), the IFS admits an invariant probability measure μ ; this measure is unique if we add the weak non-nullness condition (see 2.3.3). Moreover, the measure can be expressed on a family of sets which generate the Borel measure on K_X .*

The main idea of the proof consists in using the symbolic system associated with the contractions, an extended alphabet and usual theorems for Markov chains. In addition, under the weak non-nullness condition, we can prove the following

Theorem 2.1.2. *Under the assumption (2.4.1) and the condition (2.3.3), the IFS is asymptotically stable i.e for any $\mu' \in \mathcal{M}_1(X)$, we have the following weak convergence:*

$$\mu' P^n \xrightarrow[n \rightarrow \infty]{} \mu.$$

The main idea consists in using the Portmanteau's theorem and showing that we have the convergence of the iterated measure for any Borel set of μ -continuity. It is enough to prove the convergence for all sets of a family that generate the Borel algebra on X . Even if it means adding well chosen fictitious transformations S_i so that $\bigcup_i S_i(X) = X$, the family $\{S^{(n)}[\alpha](X), \alpha \in \mathcal{A}^n, n \in \mathbb{N}\}$ is a good candidate.

In a second part, we will be interested in the Hausdorff dimension of the invariant measure (when it exists). This dimension generalises the idea of topological dimension, d , by a metric dimension enabling to classify subsets of \mathbb{R}^d whose d -dimensional volume vanishes. In 1981, Hutchinson calculated the Hausdorff dimension of an invariant measure for IFSs such that the S_i were similarities of ratio r_i which verified the open set condition (OSC) and such that the p_i were equal to r_i^s where s is the similarity exponent. In that case, the Hausdorff dimension is equal to s , which can be found by solving the equation $\sum_i r_i^s = 1$. In [19], Fan got more general results. If the contractions S_i are still similarities on \mathbb{R}^d with an attractor totally disconnected (in particular, the OSC is verified) and if the transition probabilities are constant but non specific this time, the Hausdorff dimension is equal to the entropy h_μ divided by the Lyapunov exponent χ_μ . Since then, studies with IFSs contracting in average, overlapping or that do not verify the OSC have been done and enabled to establish an upper bound for the Hausdorff dimension equal to the ratio h_μ/χ_μ (see for instance [37]). For these studies, the transition probabilities were constant or at least Dini-continuous. For our class of IFSs with piecewise constant probabilities, we are going to prove the following

Theorem 2.1.3. *Under the assumptions (2.4.1), (2.5.1) and (2.5.2) below, the Hausdorff dimension of the invariant measure μ is equal to*

$$\dim_H(\mu) = \frac{h_\mu}{\chi_\mu}.$$

The proof is quite technical but the main idea of the proof is to find good lower and upper bounds for our measure. We can notice that in 1957, Blackwell [5] had already defined an invariant measure of an IFS with piecewise constant probabilities to calculate the entropy of hidden Markov chains.

2.2 Preliminaries and notation

Let $D \geq 2$ be an integer and \mathcal{A} be the set $\{0, \dots, D-1\}$. Let $(S_i)_{i \in \mathcal{A}}$ be a collection of Borel measurable functions from a real closed interval X into X and $(p_i)_{i \in \mathcal{A}}$ be a non-negative Borel measurable partition of unity on X , i.e. for all $i \in \mathcal{A}$, for all $x \in X$, $p_i(x) \geq 0$ and $\sum_{i=0}^{D-1} p_i(x) = 1$. The functions $p_i, i \in \mathcal{A}$, will be called placement probabilities. A family of such transformations and probabilities $\{(S_i, p_i), i \in \mathcal{A}\}$ is called an iterated function system (IFS) with place-dependent probabilities. It can be proved ([30]) that for an IFS \mathcal{S} whose transformations are contracting on a complete

metric space, there exists a unique compact set K_X which is invariant with respect to \mathcal{S} . K_X is the attractor of the IFS and we have

$$K_X = \bigcup_{i \in \mathcal{A}} S_i(K_X) = \bigcap_{n \in \mathbb{N}} \left(\bigcup_{(\alpha_1, \dots, \alpha_n) \in \mathcal{A}^n} S_{\alpha_1} \circ \dots \circ S_{\alpha_n}(X) \right).$$

In order to study the IFS, we consider the following Markov kernel: for a given $x \in X$ and a Borel subset $B \subset X$, the transition probability from x to B is defined by

$$P(x, B) = \sum_{i=0}^{D-1} \mathbb{1}_{S_i^{-1}(B)}(x) p_i(x) = \sum_{i=0}^{D-1} \mathbb{1}_B(S_i(x)) p_i(x).$$

This kernel allows the definition of a discrete-time Markov process $(X_n)_{n \in \mathbb{N}}$ on X whose transition probabilities, defined above, yield

$$\mathbb{P}(X_{n+1} = S_i(x) | X_n = x) = p_i(x).$$

We denote by $\mathcal{M}_1(X)$ the space of all probability measures on the Borel algebra $\mathcal{B}(X)$ of X and by P the Markov operator of this chain; it acts on $\mathcal{M}_1(X)$ as follows:

$$\mu P(B) = \int_X P(y, B) \mu(dy) = \sum_{i=0}^{D-1} \int_{S_i^{-1}(B)} p_i(y) \mu(dy) \quad \text{for } B \in \mathcal{B}(X), \mu \in \mathcal{M}_1(X).$$

A measure $\mu \in \mathcal{M}_1(X)$ is said to be invariant for the chain $(X_n)_n$ if $\mu P = \mu$.

We can consider the set \mathcal{A} as an alphabet and associate a symbolic system with the Markov chain $(X_n)_n$. Indeed, let us denote by $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ the "last letter process" for $(X_n)_{n \in \mathbb{N}}$. The conditional law of the transition probabilities for the joint process $(A_n, X_n)_{n \in \mathbb{N}}$ is given by

$$\mathbb{P}(A_{n+1} = j, X_{n+1} \in B | A_n = i, X_n = x) = p_j(x) \delta_{S_j(x)}(B) \quad \text{for } B \in \mathcal{B}(X), i, j \in \mathcal{A}, x \in X.$$

The process $(A_n)_{n \in \mathbb{N}}$ is not, in general, a Markov chain but merely a hidden Markov chain. It becomes a Markov chain under the assumptions made later in the text. In that case, studying the existence of an invariant measure for $(X_n)_n$ will be equivalent to studying the existence of an invariant measure for $(A_n)_n$.

In order to calculate the Hausdorff dimension of the invariant measure, we need a few definitions and notation. For a non-empty subset U of X , we denote by $|U|$ its diameter.

$$|U| = \sup \{|x - y| : x, y \in U\}.$$

Definition 2.2.1. (*δ -cover of a subset*) Let F be a subset of X . A δ -cover of F is an at most countable collection $(U_i)_i$ of subsets of X such that $F \subset \bigcup_i U_i$ and $|U_i| < \delta$.

Let F be a subset of X and s a positive real. For any $\delta > 0$, we define

$$\mathcal{H}_\delta^s(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : (U_i)_i \text{ } \delta\text{-cover of } F \right\}.$$

It can be proved (see [16] for an easily accessible reference) that $\lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F) \in \overline{\mathbb{R}}_+$ and we can define

Definition 2.2.2. (*s-dimensional Hausdorff measure of a set*) Let F be a subset of X and s a positive real. The outer measure

$$\mathcal{H}^s(F) := \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F)$$

is the s -dimensional Hausdorff measure of the set F .

It can be shown that $\mathcal{H}^s(\cdot)$ is an outer measure for any positive s . On the other hand, for any subset F of X , we can prove that there exists a unique $s_0 > 0$ such that

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{if } s < s_0, \\ 0 & \text{if } s > s_0. \end{cases}$$

This s_0 is called the Hausdorff dimension of F and we have $0 \leq \mathcal{H}^{s_0}(F) \leq \infty$. Formally, we have

Definition 2.2.3. (*Hausdorff dimension of a set*) Let F be a subset of X . We define its Hausdorff dimension by

$$\dim_H(F) := \inf \{s : \mathcal{H}^s(F) = 0\} = \sup \{s : \mathcal{H}^s(F) = \infty\} \in \overline{\mathbb{R}}_+.$$

We can now define the Hausdorff dimension and the upper Hausdorff dimension following [17] and [18].

Definition 2.2.4. (*Hausdorff dimension of a Radon probability measure*) Let μ be a probability measure. Its Hausdorff dimension is defined by

$$\begin{aligned} \dim_H \mu &= \inf \{ \dim_H A : A \text{ Borel with } \mu(A) > 0 \} \\ &= \sup \left\{ s : \liminf_{r \rightarrow 0^+} \frac{\log \mu(\mathcal{B}(x, r))}{\log r} \geq s \text{ for } \mu\text{-almost all } x \right\}. \end{aligned}$$

Definition 2.2.5. (*Upper Hausdorff dimension of a Radon probability measure*) Let μ be a probability measure. Its upper Hausdorff dimension is defined by

$$\begin{aligned} \dim_H^* \mu &= \inf \{ \dim_H A : A \text{ Borel with } \mu(A) = 1 \} \\ &= \inf \left\{ s : \liminf_{r \rightarrow 0^+} \frac{\log \mu(\mathcal{B}(x, r))}{\log r} \leq s \text{ for } \mu\text{-almost all } x \right\}. \end{aligned}$$

Definition 2.2.6. A probability measure μ is exact dimensional if there exists a constant $C \geq 0$ such that

$$C = \liminf_{r \rightarrow 0^+} \frac{\log \mu(\mathcal{B}(x, r))}{\log r} = \limsup_{r \rightarrow 0^+} \frac{\log \mu(\mathcal{B}(x, r))}{\log r} \text{ for } \mu\text{-a.e. } x.$$

In that case, $\dim_H \mu = \dim_H^* \mu = C$.

Calculating the Hausdorff dimension from the definition is generally complicated and it can be easier to use the characterisation above.

We end by giving the definitions of the entropy and the Lyapunov exponent. Let $\mathcal{S} = \{(S_i, p_i), i \in \mathcal{A}\}$ an IFS which admits μ as an invariant measure. Then, the entropy h_μ , associated with μ , is

$$h_\mu := - \int_X \sum_{i \in \mathcal{A}} p_i(x) \log p_i(x) d\mu(x).$$

If $(S_i)_i$ are differentiable, the Lyapunov exponent χ_μ of the IFS with respect to μ is defined by

$$\chi_\mu := - \int_X \sum_{i \in \mathcal{A}} p_i(x) \log |S'_i(x)| d\mu(x).$$

2.3 Technical results

For any $x \in X$ and any Borel subset $B \subset X$, the form $P(x, B) = \sum_{i=0}^{D-1} \mathbb{1}_B(S_i(x)) p_i(x)$ implies that for any integer n ,

$$P^n(x, B) = \sum_{(a_1, \dots, a_n) \in \mathcal{A}^n} p_{a_1}(x) p_{a_2}(S_{a_1}(x)) \cdots p_{a_n}(S_{a_{n-1}} \circ \cdots \circ S_{a_1}(x)) \mathbb{1}_B(S_{a_n} \circ \cdots \circ S_{a_1}(x)).$$

In order to systematize, we introduce for any $n \in \mathbb{N}$, the notation

$$\mathcal{A}^n = \{\alpha = \alpha_1 \cdots \alpha_n \mid \alpha_i \in \mathcal{A}\} \quad \text{and} \quad \mathcal{A}^{\mathbb{N}} = \{\alpha = \alpha_1 \alpha_2 \alpha_3 \cdots \mid \alpha_i \in \mathcal{A}\}.$$

For $\alpha \in \mathcal{A}^n$, $n \in \mathbb{N} \cup \{+\infty\}$, and $k, l \leq n$, let $s_{[l:k]}(\alpha) = \alpha_l \cdots \alpha_k = \alpha_{[l:k]}$ be the segment of α between indices l and k and $S^{(k)}[\alpha] = S_{\alpha_1} \circ \cdots \circ S_{\alpha_k}$. We shall use indistinguishably the notation $s_{[l:k]}(\alpha)$ or $\alpha_{[l:k]}$.

Lemma 2.3.1. *Let*

1. $M \geq 1$ be a fixed integer,
2. π a probability vector on \mathcal{A}^M ,
3. $\rho : \mathcal{A}^{M+1} \rightarrow [0, 1]$ a map such that for all $\alpha \in \mathcal{A}^M$, $\sum_{i \in \mathcal{A}} \rho_{i\alpha} = 1$.

Then, there exists a measurable space (Ω, \mathcal{F}) , a unique probability $\mathbb{P}_\pi \in \mathcal{M}_1(\Omega, \mathcal{F})$, and a sequence $A = (A_n)_{n \in \mathbb{N}}$ of \mathcal{A} -valued random variables on (Ω, \mathcal{F}) such that for all $n \geq M$,

$$\mathbb{P}(A_n = i | \mathcal{F}_{n-1}) = \mathbb{P}(A_n = i | s_{[n-1:n-M]}(A)),$$

where \mathcal{F}_n is the σ -algebra generated by $(A_k)_{k \leq n}$ and $\mathbb{P}(A_n = i | s_{[n-1:n-M]}(A) = \alpha) = \rho_{i\alpha}$ for $\alpha \in \mathcal{A}^M$. Moreover, all finite-dimensional marginals are given for $\alpha \in \mathcal{A}^{\mathbb{N}}$ by

$$\mathbb{P}_\pi(A_{[n:1]} = \alpha_{[n:1]}) = \mathbb{P}_\pi(A_n = \alpha_n, \dots, A_1 = \alpha_1) = \left(\prod_{k=0}^{n-M-1} \rho_{\alpha_{[n-k:n-k-M]}} \right) \pi(\alpha_{[M:1]}).$$

Proof. This is a simple consequence of Ionescu-Tulcea theorem [32] (An easier accessible reference is [46]) since ρ defines a stochastic matrix $K : \mathcal{A}^M \times \mathcal{A}^M \rightarrow [0, 1]$ by

$$K(\alpha, \beta) = \begin{cases} \rho_{i\alpha} & \text{if } \beta = i\alpha_1 \cdots \alpha_{M-1} \\ 0 & \text{otherwise} \end{cases}$$

□

Remark 2.3.2. We can choose as minimal Ω , holding the whole sequence $(A_k)_k$, the trajectory space $\Omega = \mathcal{A}^{\mathbb{N}}$.

Condition 2.3.3. (*Weak non-nullness condition*) We can make the following assumptions :

1.

$$\sum_{i \in \mathcal{A}} \inf_{\alpha \in \mathcal{A}^M} \rho_{i\alpha} > 0$$

2.

$$\sup_{(i,\alpha) \in \mathcal{A} \times \mathcal{A}^M} \rho_{i\alpha} < 1.$$

Remarque 2.3.4. The first condition will be required to guarantee the uniqueness of invariant measure of different processes. The second one will be used to get a subspace of $\mathcal{A}^{\mathbb{N}}$ in bijection with K_X .

Corollary 2.3.5. Under the condition (2.3.3) and with $\Omega = \mathcal{A}^{\mathbb{N}}$, we have

1. $\mathbb{P}_\pi(\{\alpha\}) = 0$ for all $\alpha \in \mathcal{A}^{\mathbb{N}}$,
2. $\mathbb{P}_\pi(\Sigma_i) = 1$ where

$$\Sigma_i = \mathcal{A}^{\mathbb{N}} \setminus \{\alpha \in \mathcal{A}^{\mathbb{N}} : \exists N \in \mathbb{N}, \forall n \geq N, \alpha_n = i\}.$$

In the sequel, we can set $\Sigma = \Sigma_{D-1}$.

Remarque 2.3.6. The interest of defining Σ is to guarantee a bijection with K_X .

Proof.

1. Let $\alpha \in \mathcal{A}^{\mathbb{N}}$. We have

$$\begin{aligned} P_\pi(\{\alpha\}) &= \lim_{n \rightarrow \infty} \mathbb{P}(A_{[n:1]} = \alpha_{[n:1]}) \\ &= \lim_{n \rightarrow \infty} \left(\prod_{k=0}^{n-M} \rho_{\alpha_{[n-k:n-k-M-1]}} \right) \pi(\alpha_{[M:1]}) \\ &= 0. \end{aligned}$$

2. We have $\Sigma_i = \mathcal{A}^{\mathbb{N}} \setminus \bigcup_{n \geq 1} \mathcal{A}^n \{i\}^{\mathbb{N}}$. As $\bigcup_{n \geq 1} \mathcal{A}^n \{i\}^{\mathbb{N}}$ is a countable union of infinite words and each singleton has a measure 0, $\mathbb{P}_\pi(\Sigma_i) = 1$.

□

If $A = (A_n)_{n \in \mathbb{N}}$ is the previous process, define $B = (B_n)_{n \geq M+1}$ by

$$B_n = A_{[n:n-M+1]} = A_n \cdots A_{n-M+1}.$$

Lemma 2.3.7. $(B_n)_{n \in \mathbb{N}}$ is a \mathcal{A}^M -valued Markov chain of order 1 whose transition matrix Q is given for any $\alpha_1, \dots, \alpha_M, i \in \mathcal{A}$, by

$$Q(\alpha_1 \cdots \alpha_M, i \alpha_1 \dots \alpha_{M-1}) = \mathbb{P}_\pi(B_{n+1} = i \alpha_1 \dots \alpha_{M-1} | B_n = \alpha_1 \cdots \alpha_M) = \rho_{i \alpha_1 \dots \alpha_M}.$$

Proof. That is a simple substitution. □

Remarque 2.3.8. Notice that the chain $(B_n)_n$ may be not irreducible under the condition (2.3.3).

Proposition 2.3.9. (Asymptotic behaviour of the chain $(B_n)_n$) The \mathcal{A}^M -valued chain $(B_n)_n$ admits an invariant probability measure ν which is unique under the condition (2.3.3). In addition, under the same condition (2.3.3), we have for any probability measure $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$ and for any $\alpha \in \mathcal{A}^M$,

$$\nu' Q^k(\alpha) \xrightarrow[k \rightarrow \infty]{} \nu(\alpha).$$

Proof. As the state space is finite, there exists at least one communication class C , therefore, an invariant measure for the chain $(B_n)_n$.

Assume now that the condition (2.3.3) is verified and let i be one of the i 's whose $\inf_{\alpha \in \mathcal{A}^M} (\rho_{i,\alpha}) > 0$. Then, for any $\alpha_1 \cdots \alpha_M \in \mathcal{A}^M$, we have

$$Q^M(\alpha_1 \cdots \alpha_M, i \cdots i) = \prod_{k=1}^M \rho_{i^{M+1-k} \alpha_1 \cdots \alpha_k} > 0.$$

Then, the state $i^M \in \mathcal{A}^M$ is accessible from any state in \mathcal{A}^M , which implies that there is at most one communication class. Consequently, there is exactly one class C , which means this class is absorbing and supports a unique invariant probability measure ν .

Let $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$. As there is a unique communication class C , there exists $n \in \mathbb{N}$ such that for any $k \geq n$ the support of the measure $\nu' Q^k$ is included in C . Then, for any $\alpha \in \mathcal{A}^M \setminus C$, we have $\nu' Q^k(\alpha) \xrightarrow[k \rightarrow \infty]{} 0 = \nu(\alpha)$.

In addition, the measure $\nu' Q^k|_C$ which denotes the restriction of $\nu' Q^k$ on the set C is a probability measure on C for $k \geq n$. Similarly, we can denote by $\nu|_C$ the restriction of the invariant measure ν on the set C , which is also a probability measure on C .

The chain $(B_n)_n$ restricted on C is irreducible and also aperiodic since, considering the same state i as before, we have $Q(i^M, i^M) = \rho_{i^{M+1}} > 0$. Then, by the ergodic theorem, for any $\alpha \in C$, we have $\nu' Q^n|_C (Q|_C)^k(\alpha) \xrightarrow[k \rightarrow \infty]{} \nu|_C(\alpha)$ where $Q|_C$ denotes the submatrix of Q formed by the rows and columns corresponding to C . Therefore, for any $\alpha \in C$, we have $\nu' Q^k(\alpha) \xrightarrow[k \rightarrow \infty]{} \nu(\alpha)$.

In any case, notice that the measure(s) ν can be found by solving the linear system $\nu Q = \nu$. □

Remarque 2.3.10. An alternative proof of the proposition (2.3.9) can be given by use of the Dobrushing ergodicity coefficient ([13] and [34]).

Example 2.3.11. Let $\mathcal{A} = \{0, 1\}$ and $M = 1$. Then, the transition matrix Q of the chain $(B_n)_n$ is the following:

$$Q = \begin{pmatrix} \rho_{00} & \rho_{10} \\ \rho_{01} & \rho_{11} \end{pmatrix}$$

If ρ_{00} and $\rho_{01} \in (0, 1)$, then the chain $(B_n)_n$ admits a unique invariant probability measure given by :

$$\nu(0) = \frac{\rho_{01}}{\rho_{01} + \rho_{10}}, \quad \nu(1) = \frac{\rho_{10}}{\rho_{01} + \rho_{10}}.$$

2.4 Invariant measure(s) of an IFS

Let us denote by $S^{(n)}[\alpha]$ the function $S_{\alpha_1} \circ \dots \circ S_{\alpha_n}$ for any $\alpha = (\alpha_1, \alpha_2, \dots) \in \mathcal{A}^{\mathbb{N}}$ and by \mathbf{S} the set $\bigcup_{\alpha \in \mathcal{A}^M} S^{(M)}[\alpha](X)$.

Global assumption 2.4.1. (*Structure of the IFS*) In the following, we will consider a class of IFSs $\{(S_i, p_i), i \in \mathcal{A}\}$ on a closed interval $X \subset \mathbb{R}$ which satisfy the conditions :

1. For all $i \in \mathcal{A}$, $S_i(X)$ is a subinterval of X and $\bigcup_{i \in \mathcal{A}} S_i(X) \subseteq X$.
2. The images $S_i(X)$ are non-overlapping or just-touching (i.e for $i \neq j$, we have $S_i(\overset{\circ}{X}) \cap S_j(\overset{\circ}{X}) = \emptyset$).
3. The transformations S_i are ordered (i.e. for all $x \in X$, for all $i \leq i+1$, $S_i(x) \leq S_{i+1}(x)$).
4. For all $i \in \mathcal{A}$, there exists a constant $L_i < 1$ such that the map $S_i : X \rightarrow S_i(X)$ iuniques a bijective and L_i -contractive function.
5. The transition probabilities $(p_i)_{i \in \mathcal{A}}$ are adapted to the functions S_i i.e. there exist an integer M , a family $(\rho_{i\alpha})_{(i,\alpha) \in \mathcal{A}^{M+1}}$ and measurable functions $(q_i)_{i \in \mathcal{A}}$ on \mathbf{S}^c such that

$$p_i(x) = \sum_{\alpha \in \mathcal{A}^M} \rho_{i\alpha} \mathbb{1}_{S^{(M)}[\alpha](X)}(x) + q_i(x) \mathbb{1}_{\mathbf{S}^c}(x).$$

Remarque 2.4.2. In the case where two contractions S_i and S_j are just-touching, we can have $S_i(\sup X) = S_j(\inf X)$ or vice-versa and thus the transition probabilities in the assumption (2.4.1) are not well defined. To prevent this trivial ambiguity from occurring, we will consider, for any $(\alpha_2, \dots, \alpha_M) \in \mathcal{A}^{M-1}$, a left-closed right-open version of the interval $S_i \circ S_{\alpha_2} \circ \dots \circ S_{\alpha_M}(X)$.

Let us start with a few remarks on the assumptions. The set X is chosen to be complete in order to guarantee the existence and uniqueness of the attractive compact set K_X . The assumptions on separability of X and on bijectivity and contractivity of the functions S_i imply that the denumerable family of sets $\{S^{(n)}[\alpha](X) \cap K_X, \alpha \in \mathcal{A}^n, n \in \mathbb{N}\}$ generate the Borel algebra on K_X , and thus it is enough to express the invariant measure of these sets when there exists one. Adding the adapted form of the transition probabilities, the former assumptions enable to use a symbolic system and to deal with Markov chains on a finite state space. Finally, the last point of the assumption means that outside the intervals in the form of one of the $S^{(M)}[\alpha](X)$, the transition probabilities can take any values as long as the functions are measurable.

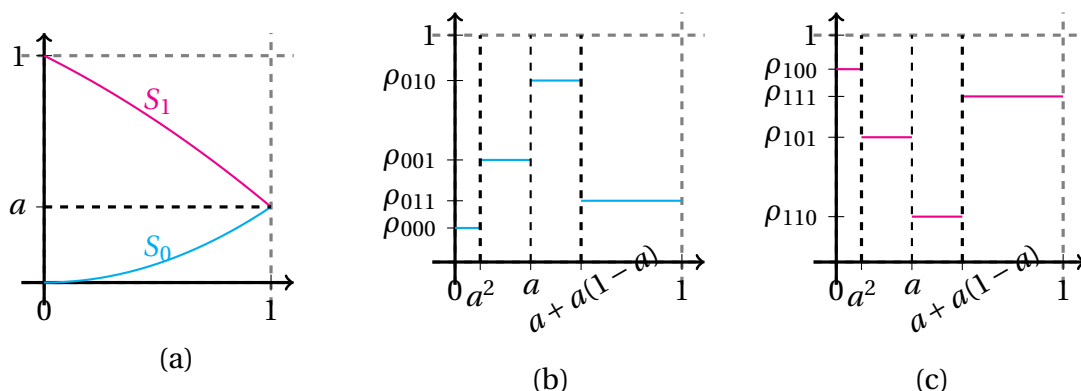


Figure 2.1: A first example of IFS verifying the assumption (2.4.1) with S_0 and S_1 depicted in (a) and adaptation of order $M = 2$ for transition probabilities depicted in (b) and (c). The number $a \in (0, 1)$ is defined so that $S_0(X) = [0, a]$ and $S_1(X) = [a, 1]$.

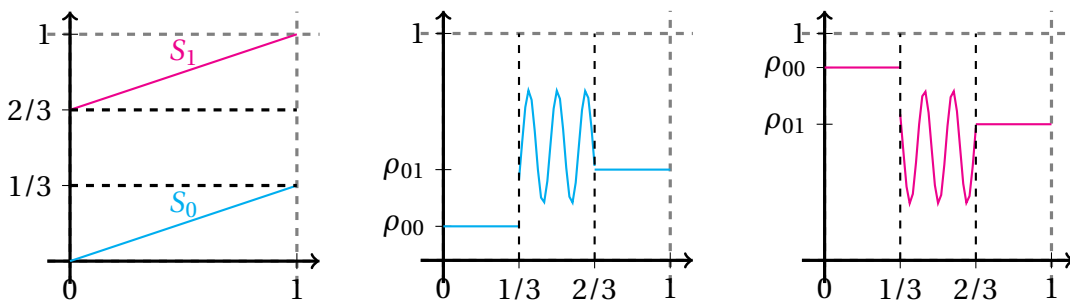


Figure 2.2: A second example of IFS verifying the assumption (2.4.1). The wiggly graph on $[\frac{1}{3}, \frac{2}{3}]$ is supposed to mean that there the functions p_i can be arbitrary.

Let us consider the mapping

$$\Sigma \ni \alpha \mapsto \phi(\alpha) := \lim_{n \rightarrow \infty} S^{(n)}[\alpha](x) \in K_X,$$

where x can be any element of X since the limit is independent of x .

Under the assumption (2.4.1), the function ϕ is well defined due to the contractivity of the functions S_i . In addition, the set Σ is chosen to guarantee the bijectivity of the function ϕ . Indeed, in the case where the contractions S_i are just-touching, if the domain of ϕ is $\mathcal{A}^{\mathbb{N}}$, some elements of K_X have two inverse images. The set Σ is chosen to avoid this problem.

Lemma 2.4.3. *If an IFS on X admits an invariant measure μ , its support is included in the attractor of the IFS K_X .*

Proof. Let μ be an invariant measure for the IFS. If $y \in \text{supp}\mu P$ and $\varepsilon > 0$, then $\mu P(\mathcal{B}(y, \varepsilon) \cap X) > 0$. Besides, as $\mu P(\mathcal{B}(y, \varepsilon) \cap X) = \int_X P(x, \mathcal{B}(y, \varepsilon) \cap X) \mu(dx)$, there exists an $x \in \text{supp}\mu$ such that $P(x, \mathcal{B}(y, \varepsilon) \cap X) > 0$. We deduce that

$$\mathcal{B}(y, \varepsilon) \cap \text{supp}P(x, \cdot) \neq \emptyset.$$

As ε is arbitrary, we deduce that there exists an $i \in \mathcal{A}$ such that $y = S_i(x)$. Consequently, each element of the support of μP (and thus each element of the support of μ) can be written as the image of an element of the support of μ by one of the maps S_i . In other words, for any $y \in \text{supp}(\mu)$, there exist $x_1 \in \text{supp}\mu$ and $\alpha_1 \in \mathcal{A}$ such that $y = S_{\alpha_1}(x_1)$. By induction, there exists $\alpha \in \Sigma$ such that for any $n \in \mathbb{N}$, there exists $x_n \in \text{supp}(\mu)$ verifying $y = S^{(n)}[\alpha](x_n)$. Then, $y \in \bigcap_{n \in \mathbb{N}} \bigcup_{\alpha \in \mathcal{A}^n} S^{(n)}[\alpha](X) = K_X$. Thus, $\text{supp}\mu \subset K_X$. \square

In the sequel, we will use several times the technical

Lemma 2.4.4. *For all $n \geq M + 1$, for all $\alpha \in \mathcal{A}^n$ and for all $\mu \in \mathcal{M}_1(X)$, we have*

$$\mu P(S^{(n)}[\alpha](X)) = \rho_{\alpha_1 \dots \alpha_{M+1}} \mu(S^{(n-1)}[\alpha_2 \dots \alpha_n](X)).$$

Proof. Let $n \geq M + 1$, $\alpha \in \mathcal{A}^n$ and $\mu \in \mathcal{M}_1(X)$.

$$\begin{aligned} \mu P(S^{(n)}[\alpha](X)) &= \sum_{i \in \mathcal{A}} \int_X p_i(x) \mathbb{1}_{S_i^{-1}(S^{(n)}[\alpha](X))}(x) \mu(dx) \\ &= \int_{S^{(n-1)}[\alpha_2 \dots \alpha_n](X)} p_{\alpha_1}(x) \mu(dx) \\ &= \rho_{\alpha_1 \dots \alpha_{M+1}} \mu(S^{(n-1)}[\alpha_2 \dots \alpha_n](X)). \end{aligned}$$

\square

Lemma 2.4.5. *Let $\mu \in \mathcal{M}_1(X)$ be such that $\text{supp}(\mu) \subset K_X$. We can associate with μ a probability measure $\nu \in \mathcal{M}_1(\mathcal{A}^M)$ defined by $\nu(\alpha) = \mu(S^{(M)}[\alpha](X))$. Then, for any $\alpha \in \mathcal{A}^M$ and for any $k \in \mathbb{N}$, we have the following relation*

$$\mu P^k(S^{(M)}[\alpha](X)) = \nu Q^k(\alpha).$$

Proof. We can prove the result by induction on $k \in \mathbb{N}$. The result is trivial for $k = 0$. Assuming now the property is true at rank $k-1$, we prove it remains at rank k . Indeed, for any $\alpha \in \mathcal{A}^M$,

$$\begin{aligned} \mu P^k(S^{(M)}[\alpha](X)) &= \int_{S^{(M-1)}[\alpha_2 \cdots \alpha_M](X)} p_{\alpha_1}(x) \mu P^{k-1}(dx) \\ &= \sum_{\alpha_{M+1} \in \mathcal{A}} \rho_{\alpha_{[1:M+1]}} \mu P^{k-1}(S^{(M-1)}[\alpha_2 \cdots \alpha_M] \circ S_{\alpha_{M+1}}(X)) \quad (*) \\ &= \sum_{\alpha_{M+1} \in \mathcal{A}} \rho_{\alpha_{[1:M+1]}} \nu Q^{k-1}(\alpha_2 \cdots \alpha_{M+1}) \\ &= \nu Q^k(\alpha_1 \cdots \alpha_M) \\ &= \nu Q^k(\alpha). \end{aligned}$$

The result (*) is true since $\text{supp}(\mu) \subset K_X$ and thus for any $k \in \mathbb{N}$, $\text{supp}(\mu P^k) \subset K_X$ due to the definition of P and the form of K_X . \square

Theorem 2.4.6. *(Existence and uniqueness of an invariant probability measure for the IFS) Under the assumption (2.4.1), the iterated function system $\{(S_i, p_i), i \in \mathcal{A}\}$ admits an invariant probability measure μ , which is unique under the condition (2.3.3). An invariant measure on X is determined by the formula*

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n), \quad (2.4.1)$$

for all $n \geq M+1$, for all $\alpha \in \mathcal{A}^n$ and where ν denotes an invariant measure for the related chain $(B_n)_n$.

Notation 2.4.7. *We will denote by $\bar{\nu}$ the law of the process $(A_n)_{n \in \mathbb{N}}$ when $A_M \cdots A_1$ is distributed according to ν . It is an invariant and ergodic measure for the shift map σ on $\mathcal{A}^{\mathbb{N}}$.*

Proof of the theorem 2.4.6. Let μ be a probability measure on $\mathcal{B}(X)$ whose support is included in K_X . As the functions S_i are injective and contractive, the family

$$\{S^{(n)}[\alpha](X) \cap K_X \mid \alpha \in \mathcal{A}^n, n \geq M\}$$

generate the Borel σ -algebra on K_X . Then, μ is invariant for $(X_n)_n$ if and only if for all $n \geq M$, for all $\alpha \in \Sigma$, we have

$$\mu P(S^{(n)}[\alpha](X) \cap K_X) = \mu(S^{(n)}[\alpha](X) \cap K_X).$$

But, as $\mu(U) = 0$ for all $U \subset (\text{supp}\mu)^c$ and thus for all $U \subset K_X^c$, we get that μ is invariant for $(X_n)_n$ if and only if for all $n \geq M$, for all $\alpha \in \Sigma$, we have

$$\mu P(S^{(n)}[\alpha](X)) = \mu(S^{(n)}[\alpha](X)).$$

On the other hand, by lemma (2.4.4), for any probability measure μ , we have for all $\alpha \in \Sigma$, for all $n \geq M$,

$$\mu P^{n-M}(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \mu(S^{(M)}[\alpha_{[n-M+1:n]}](X)).$$

Then, if μ is invariant, we have

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \mu(S^{(M)}[\alpha_{[n-M+1:n]}](X)).$$

Therefore, to define the measure μ on the Borel σ -algebra, it is enough to determine its value on the intervals $S^{(M)}[\alpha](X)$ for $\alpha \in \mathcal{A}^M$. By lemma (2.4.5), we can associate with μ a measure $\nu \in \mathcal{M}_1(\mathcal{A}^M)$ defined by $\nu(\alpha) = \mu(S^{(M)}[\alpha](X))$ which gives

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n).$$

Existence of an invariant probability measure By proposition (2.3.9), we know there exists an invariant measure for the process $(B_n)_n$. Then, choosing ν as an invariant measure, we have for $n \geq M + 1$,

$$\begin{aligned} \mu P(S^{(n)}[\alpha](X)) &= \rho_{\alpha_{[1:M+1]}} \mu(S^{(n)}[\alpha_2 \cdots \alpha_n](X)) \\ &= \rho_{\alpha_{[1:M+1]}} \prod_{i=2}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n) \\ &= \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n) \\ &= \mu(S^{(n)}[\alpha](X)). \end{aligned}$$

Consequently, there exists an invariant measure μ for the process $(X_n)_n$ defined by

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n),$$

where ν is an invariant measure on \mathcal{A}^M .

Uniqueness of the invariant measure under the condition (2.3.3) If μ is invariant, we have in particular for any $\alpha \in \mathcal{A}^{M+1}$,

$$\mu P(S^{(M+1)}[\alpha](X)) = \mu P^2(S^{(M+1)}[\alpha](X)),$$

i.e.

$$\rho_{\alpha_{[1:M+1]}} \mu(S^{(M)}[\alpha_2 \cdots \alpha_{M+1}](X)) = \rho_{\alpha_{[1:M+1]}} \mu P(S^{(M)}[\alpha_2 \cdots \alpha_{M+1}](X)).$$

The condition (2.3.3) implies there exists an $\alpha_i \in \mathcal{A}$ such that for any $\alpha_2 \cdots \alpha_{M+1}$, $\rho_{\alpha_1 \cdots \alpha_{M+1}} > 0$ which implies that for any $\alpha_2 \cdots \alpha_{M+1} \in \mathcal{A}^M$, we have necessarily $\mu(S^{(M)}[\alpha_2 \cdots \alpha_{M+1}](X)) = \mu P(S^{(M)}[\alpha_2 \cdots \alpha_{M+1}](X))$ if μ is invariant.

In others words, we have $\nu(\alpha) = \nu Q(\alpha)$ for all $\alpha \in \mathcal{A}^M$ by lemma (2.4.5). Thus, ν is the unique invariant measure for the chain $(B_n)_n$. This shows there exists a unique invariant measure μ for the chain $(X_n)_n$. \square

Remarque 2.4.8. The condition (2.3.3) implies that for all $i \in \mathcal{A}$, $\sup_{\alpha \in \mathcal{A}^M} \rho_{i\alpha} < 1$. Then, for any $\alpha \in \Sigma$, we have $\lim_n \mu(S^{(n)}[\alpha](X)) = 0$, which means that μ assigns 0 to single points, where μ denotes an invariant measure for $(X_n)_n$.

Exemple 2.4.9. Let $\mathcal{S} = \{(S_i, p_i), i \in \{0, 1\}\}$ be an IFS on $X = [0, 1]$ such that

$$S_0(x) = \frac{x}{2} \quad \text{and} \quad p_0(x) = \rho_{00} \mathbb{1}_{[0, \frac{1}{2}]}(x) + \rho_{01} \mathbb{1}_{[\frac{1}{2}, 1]}(x),$$

$$S_1(x) = \frac{x+1}{2} \quad \text{and} \quad p_1(x) = \rho_{10} \mathbb{1}_{[0, \frac{1}{2}]}(x) + \rho_{11} \mathbb{1}_{[\frac{1}{2}, 1]}(x).$$

We assume the condition (2.3.3) is satisfied i.e. $0 < \rho_{00} < 1$ and $0 < \rho_{01} < 1$. Then the IFS admits a unique invariant measure μ . The figures (2.3), (2.4), (2.5) and (2.6) depict the cumulative distribution function of μ for three different sets of parameters ρ_{00} and ρ_{01} .

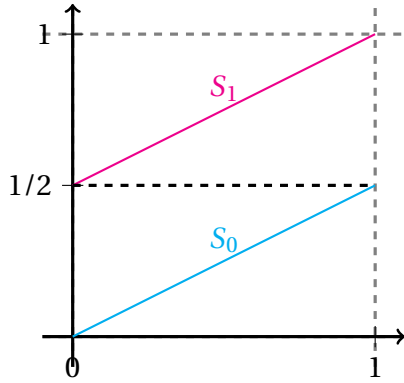


Figure 2.3: Contractions of the example (2.4.9)

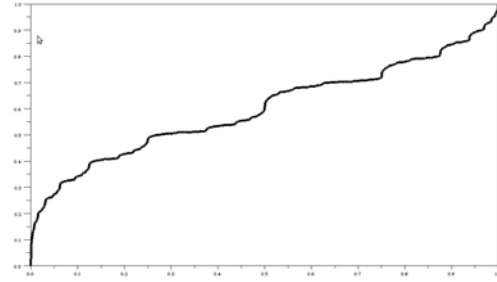


Figure 2.4: Cumulative distribution function of μ for $\rho_{00} = 0.8$ and $\rho_{01} = 0.3$

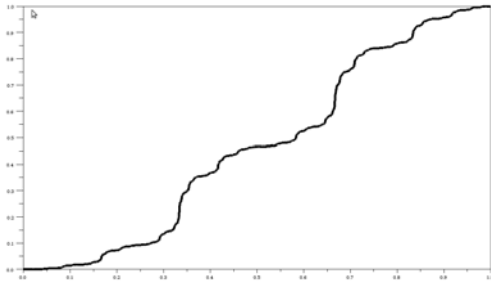


Figure 2.5: Cumulative distribution function of μ for $\rho_{00} = 0.2$ and $\rho_{01} = 0.7$

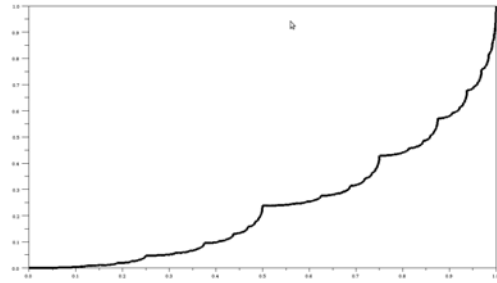


Figure 2.6: Cumulative distribution function of μ for $\rho_{00} = 0.2$ and $\rho_{01} = 0.25$

Theorem 2.4.10. Under the assumption (2.4.1) and the condition (2.3.3), the IFS is asymptotically stable i.e for any $\mu' \in \mathcal{M}_1(X)$, we have the following weak convergence

$$\mu' P^n \xrightarrow[n \rightarrow \infty]{} \mu,$$

where μ denotes the unique invariant measure of the IFS.

Proof.

Step 1 We start by proving that for any $\mu' \in \mathcal{M}_1(X)$, for any $n \in \mathbb{N}$, we have $\text{supp} \mu' P^n \subset \bigcup_{\alpha \in \mathcal{A}^n} S^{(n)}[\alpha](X)$ by induction on $n \in \mathbb{N}$.

Let $\mu' \in \mathcal{M}_1(X)$ and $n = 0$. We obviously have $\text{supp} \mu' \subset X$.

Then, assuming the property true at rank $n - 1$, we prove that it remains at rank n .

$$\begin{aligned}
\mu' P^n \left(\bigcup_{\alpha \in \mathcal{A}^n} S^{(n)}[\alpha](X) \right) &= \sum_{i \in \mathcal{A}} \int_{S_i^{-1}(\bigcup_{\alpha \in \mathcal{A}^{n-1}} S^{(n-1)}[\alpha](X))} p_i(x) \mu' P^{n-1}(dx) \\
&= \sum_{i \in \mathcal{A}} \int_{\bigcup_{\alpha \in \mathcal{A}^{n-1}} S^{(n-1)}[\alpha](X)} p_i(x) \mu' P^{n-1}(dx) \\
&= \mu' P^{n-1} \left(\bigcup_{\alpha \in \mathcal{A}^{n-1}} S^{(n-1)}[\alpha](X) \right) \\
&= 1 \quad \text{by induction on } n.
\end{aligned}$$

Step 2 With any $\mu' \in \mathcal{M}_1(X)$, we can associate a measure $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$ defined, for any $\alpha \in \mathcal{A}^M$ by

$$\nu'(\alpha) = \mu' P^M(S^{(M)}[\alpha](X)).$$

Unlike the measures considered in the lemma (2.4.5), the measures μ' can have their support in X and not only in K_X . Thus, we need to consider M iterations of the measure to have the support included in $\bigcup_{\alpha \in \mathcal{A}^M} S^{(M)}[\alpha](X)$ and thus the transition probabilities constant on each interval $S^{(M)}[\alpha](X)$. Then, we can prove that for any $\alpha \in \mathcal{A}^M$, we have by induction on $k \in \mathbb{N}$,

$$\nu' Q^k(\alpha) = \mu' P^{k+M}(S^{(M)}[\alpha](X)). \quad (2.4.2)$$

The result is trivial for $k = 0$. Assuming now the property is true at rank $k - 1$, we prove it remains at rank k . Indeed, for any $\alpha \in \mathcal{A}^M$,

$$\begin{aligned}
\mu' P^{M+k}(S^{(M)}[\alpha](X)) &= \int_{S^{(M-1)}[\alpha_2 \cdots \alpha_M](X)} p_{\alpha_1}(x) \mu' P^{M+k-1}(dx) \\
&= \sum_{\alpha_{M+1} \in \mathcal{A}} \rho_{\alpha_{[1:M+1]}} \mu' P^{M+k-1}(S^{(M-1)}[\alpha_2 \cdots \alpha_M] \circ S_{\alpha_{M+1}}(X)) \quad (*) \\
&= \sum_{\alpha_{M+1} \in \mathcal{A}} \rho_{\alpha_{[1:M+1]}} \nu' Q^{k-1}(\alpha_2 \cdots \alpha_{M+1}) \\
&= \nu' Q^k(\alpha_1 \cdots \alpha_M) \\
&= \nu' Q^k(\alpha).
\end{aligned}$$

The relation (*) is obtained since $\text{supp}(\mu P^{M+k-1}) \subset \bigcup_{\alpha \in \mathcal{A}^M} S^{(M)}[\alpha](X)$.

Step 3 It remains to prove that for any $\mu' \in \mathcal{M}_1(X)$, for any Borel set A of μ -continuity, $\mu' P^k(A) \xrightarrow[k \rightarrow \infty]{} \mu(A)$. By Portmanteau's theorem [4], we should then conclude that

$\mu' P^k \implies \mu$. It is enough to prove the convergence for all sets of a family that generate the Borel algebra on X .

If $\bigcup_{i \in \mathcal{A}} S_i(X) = X$ then the family of sets $\{S^{(n)}[\alpha](X), \alpha \in \mathcal{A}^n, n \in \mathbb{N}\}$ generate the Borel algebra on X . Otherwise, we extend the IFS \mathcal{S} by building fictitious transformations, i.e. transformations that will arise with probabilities 0. Indeed, we extend the alphabet \mathcal{A} into an alphabet \mathcal{A}' and we build $|\mathcal{A}' \setminus \mathcal{A}|$ affine transformations $(S_i)_{i \in \mathcal{A}' \setminus \mathcal{A}}$ whose images are intervals of the set $X \setminus \bigcup_{\alpha_1 \in \mathcal{A}} S_{\alpha_1}(X)$ so that $\bigcup_i S_{i \in \mathcal{A}'}(X) = X$. The transition probabilities p_i are equal to 0 for all $i \in \mathcal{A}' \setminus \mathcal{A}$. We denote by \mathcal{S}' this new IFS. Notice that the dynamics of the Markov chains generated by \mathcal{S} and \mathcal{S}' and the associated Markov operators are the same.

As we built a family of sets $\{S^{(n)}[\alpha](X), \alpha \in \mathcal{A}'^n, n \in \mathbb{N}\}$ that generate the Borel algebra on X , it is enough to prove that for any $\mu' \in \mathcal{M}_1(X)$, for any $\alpha \in \mathcal{A}'^n$, we have

$$\mu' P^k (S^{(n)}[\alpha](X)) \xrightarrow[k \rightarrow \infty]{} \mu(S^{(n)}[\alpha](X)).$$

Let $k \geq n \geq M$ and $\alpha = \alpha_1 \cdots \alpha_n \in \mathcal{A}'^n$.

Case 1. If one of the α_i belongs to $\mathcal{A}' \setminus \mathcal{A}$, then $S^{(n)}[\alpha](X) \subset (\bigcup_{\alpha \in \mathcal{A}^n} S^{(n)}[\alpha](X))^c$. By step 1, we deduce that

$$\mu' P^k (S^{(n)}[\alpha](X)) \xrightarrow[k \rightarrow \infty]{} 0 = \mu(S^{(n)}[\alpha](X)),$$

since $\text{supp}(\mu) \subset K_X$ and $S^{(n)}[\alpha](X) \subset (K_X)^c$.

Case 2. If all the α_i belong to \mathcal{A} , then use lemma (2.4.4) to iterate; using the step 2, we have

$$\begin{aligned} \mu' P^k (S^{(n)}[\alpha](X)) &= \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \mu' P^{k-n+M} (S^{(M)}[\alpha_{n-M+1} \cdots \alpha_n](X)) \\ &= \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu' Q^{k-n} (\alpha_{n-M+1} \cdots \alpha_n). \end{aligned}$$

But, for any $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$, for any $\alpha \in \mathcal{A}^M$, $\nu' Q^k(\alpha) \xrightarrow[k \rightarrow \infty]{} \nu(\alpha)$.

Thus,

$$\mu' P^k (S^{(n)}[\alpha](X)) \xrightarrow[k \rightarrow \infty]{} \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n) = \mu(S^{(n)}[\alpha](X)).$$

The asymptotic stability is thus proved. \square

2.5 Hausdorff dimension

In the case where there exists a unique invariant probability measure for the IFS, we will establish a formula giving the Hausdorff dimension of the invariant measure. We need to add the technical

Condition 2.5.1. *The contractions $(S_i)_{i \in \mathcal{A}}$ are differentiable and the derivatives are Hölder-continuous. In addition, we assume $\inf_{i,x} |S'_i(x)| > 0$.*

This condition will allow us to control the variations of the derivatives of the maps S_i and consider their logarithm.

Condition 2.5.2. *(Strong non-nullness condition) For all $i \in \mathcal{A}$, $\inf_{\alpha \in \mathcal{A}^M} \rho_{i\alpha} > 0$.*

Remarque 2.5.3.

1. All results obtained under the condition (2.3.3) remain true under the condition (2.5.2).
2. The support of ν is the set \mathcal{A}^{M+1} and for any $n \in \mathbb{N}$, for any $\alpha \in \Sigma$, we have

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-(M+1)} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M} \cdots \alpha_n) > \left(\inf_{\beta \in \mathcal{A}^{M+1}} \rho_\beta \right)^{n-(M+1)} \inf_{\beta \in \mathcal{A}^{M+1}} \nu(\beta) > 0.$$

We can therefore consider the logarithm of the invariant measure μ on any interval of the form $S^{(n)}[\alpha](X)$.

Theorem 2.5.4. *(Hausdorff dimension of the support of an invariant measure) Under the assumption (2.4.1) and the conditions (2.5.2) and (2.5.1), the Hausdorff dimension of the invariant measure μ of the considered IFS is given by*

$$\dim_H(\mu) = \frac{h_\mu}{\chi_\mu},$$

where h_μ and χ_μ denote respectively the entropy of μ and the Lyapunov exponent of the IFS given by

$$h_\mu = - \sum_{(i,\alpha) \in \mathcal{A} \times \mathcal{A}^M} \nu(\alpha) \rho_{i\alpha} \log \rho_{i\alpha} \quad \text{and} \quad \chi_\mu = - \sum_{(i,\alpha) \in \mathcal{A} \times \mathcal{A}^M} \rho_{i\alpha} \int_{S^{(M)}[\alpha](X)} \log |S'_i(x)| \mu(dx).$$

We aim at proving that μ is exact dimensional. More precisely, we want to show that for μ -a.e x , we have

$$\lim_{r \rightarrow 0^+} \frac{\log(\mu(\mathcal{B}(x, r)))}{\log r} = \frac{-h_\mu}{-\chi_\mu}.$$

The idea of the proof consists in finding N_{int} and N_{ext} which will depend on x (or equivalently on α and r) such that

$$\mu(S^{(N_{\text{int}})}[\alpha](X)) \leq \mu(\mathcal{B}(x, r)) \leq \mu(S^{(N_{\text{ext}})}[\alpha](X)) \text{ for all } r > 0, \quad (2.5.1)$$

$$\lim_{r \rightarrow 0^+} \frac{\log \mu(S^{(N_{\text{int}})}[\alpha](X))}{N_{\text{int}}} = \lim_{r \rightarrow 0^+} \frac{\log \mu(S^{(N_{\text{ext}})}[\alpha](X))}{N_{\text{ext}}} = -h_\mu \text{ for } \bar{\nu} - \text{a.e } \alpha, \quad (2.5.2)$$

and

$$\lim_{r \rightarrow 0^+} \frac{\log r}{N_{\text{int}}} = \lim_{r \rightarrow 0^+} \frac{\log r}{N_{\text{ext}}} = -\chi_\mu \text{ for } \bar{\nu} - \text{a.e } \alpha. \quad (2.5.3)$$

Remarque 2.5.5. The proof of the theorem is based on ideas borrowed from [19] and [20]. The main difference with these references is that in our case the attractor can be not disconnected. Therefore, we can not separate the compact sets and thus we cannot get a lower bound for the invariant measure of balls of radius r using their arguments. Consequently, we have to find an other way to get lower and upper bounds. After having found such bounds, we can calculate the Hausdorff dimension following [20].

Proof of the theorem 2.5.4.

Step 1 Let $\alpha \in \Sigma$ and $x \in K_X$ such that $x = \phi(\alpha)$. Let $r > 0$. We shall show that

$$\mathcal{B}\left(x, C^{-1}c^{d_{N(r)}(\alpha)}r\right) \subset S^{N(r)}[\alpha](X) \subset \mathcal{B}(x, Cr), \quad (2.5.4)$$

where the symbols C , c and $d_{N(r)}(\alpha)$ will be defined in the course of the proof. We start by defining

$$N(r) = \inf\{n \geq 0 : |S^{(n)}[\alpha]'(\phi(\sigma^n \alpha))| < r\} = \inf\left\{n \geq 0 : \prod_{i=1}^n \left|S'_{\alpha_i}\left(S^{(n-i)}[\sigma^i \alpha](\phi(\sigma^n \alpha))\right)\right| < r\right\}.$$

Let us show there exists a constant $C > 0$ such that for all $r > 0$,

$$C^{-1}r \leq |S^{N(r)}[\alpha](X)| \leq Cr. \quad (2.5.5)$$

By ([48] proposition 20.1(1)), there exists a constant K such that for all $n \geq 1$, all $\alpha \in \Sigma$ and all $z, y \in X$,

$$K^{-1} \leq \frac{|S^{(n)}[\alpha]'(z)|}{|S^{(n)}[\alpha]'(y)|} \leq K. \quad (2.5.6)$$

It follows that

$$|S^{N(r)}[\alpha](X)| \leq K|X| |S^{N(r)}[\alpha](\phi(\sigma^n \alpha))| < K|X|r.$$

The first inequality is obtained by choosing $n = N(r)$ and $y = \phi(\sigma^n \alpha)$ in the relation (2.5.6) and by integrating on the variable z . The second one is due to the definition of $N(r)$.

In addition, we have

$$\begin{aligned} |S^{N(r)}[\alpha](X)| &= |S^{N(r)-1}[\alpha] \circ S_{\alpha_{N(r)}}(X)| \\ &\geq \inf_x |S^{N(r)-1}[\alpha]'(X)| |S_{\alpha_{N(r)}}(X)| \\ &\geq K^{-1} r |X| \inf_{i,x} |S'_i(x)|. \end{aligned}$$

By choosing $C = \max\left(\left(|X| K^{-1} \inf_{i,x} |S'_i(x)|\right)^{-1}, K|X|\right)$, we get the relation (2.5.5).

As $x = \phi(\alpha) \in S^{N(r)}[\alpha](X)$, we deduce immediately the second inclusion in (2.5.4) i.e.

$$S^{N(r)}[\alpha](X) \subset \mathcal{B}(x, Cr).$$

Let us now define for $n \in \mathbb{N}$ and $\alpha \in \Sigma$,

$$d_n(\alpha) = \max(\inf\{i, \alpha_{n+i} \neq 0\}, \inf\{i, \alpha_{n+i} \neq D-1\}).$$

Then, we have $\alpha_{n+1} = \dots = \alpha_{n+d_n(\alpha)-1}$ and they are all equal to 0 or $D-1$.

Suppose they are all equal to 0. Then,

$$\begin{aligned} \left| \phi(\alpha) - \min_{x \in X} S^{N(r)}[\alpha](x) \right| &\geq \left| S_0^{d_{N(r)}(\alpha)}(S^{N(r)}[\alpha](X)) \right| \\ &\geq \left(\inf_x |S'_0(x)| \right)^{d_{N(r)}(\alpha)} C^{-1} r. \end{aligned}$$

Similarly, if $\alpha_{n+1} = \dots = \alpha_{n+d_n(\alpha)-1} = D-1$, we have

$$\begin{aligned} \left| \phi(\alpha) - \max_{x \in X} S^{N(r)}[\alpha](x) \right| &\geq \left| S_{D-1}^{d_{N(r)}(\alpha)}(S^{N(r)}[\alpha](X)) \right| \\ &\geq \left(\inf_x |S'_{D-1}(x)| \right)^{d_{N(r)}(\alpha)} C^{-1} r. \end{aligned}$$

Set $c = \min(\inf_x |S'_0(x)|, \inf_x |S'_{D-1}(x)|)$. Then, we get the first inclusion in (2.5.4), i.e.

$$\mathcal{B}(x, C^{-1} c^{d_{N(r)}(\alpha)} r) \subset S^{N(r)}[\alpha](X).$$

Step 2 From the previous step, we deduce that

$$S^{N(C^{-1}r)}[\alpha](X) \subset \mathcal{B}(x, r) \subset S^{N(Crc^{-d_{N(r)}(\alpha)})}[\alpha](X). \quad (2.5.7)$$

We have thus found $N_{\text{int}} = N(C^{-1}r)$ and $N_{\text{ext}} = N(Crc^{-d_{N(r)}(\alpha)})$ which satisfy the inequality (2.5.1). Notice that $N(r)$ tends to $+\infty$ when r goes to 0. Then, $N(C^{-1}r)$ tends to $+\infty$ when r goes to 0. On the other hand, $0 \leq Crc^{-d_{N(r)}(\alpha)} \leq Crc^{-1}$. As Crc^{-1} tends to 0 when r goes to 0, then $Crc^{-d_{N(r)}(\alpha)}$ also. Therefore, $N(Crc^{-d_{N(r)}(\alpha)})$ tends to $+\infty$ when r goes to 0.

Step 3 Let us show the equalities (2.5.2). It is enough to show that for almost all $\alpha \in \Sigma$, we have

$$\lim_{n \rightarrow +\infty} \frac{\log \mu(S^{(n)}[\alpha](X))}{n} = -h_\mu.$$

By the equality (2.4.1) and the remark (2.5.3), we have for any $n \geq M$, for any $\alpha \in \Sigma$,

$$\mu(S^{(n)}[\alpha](X)) = \prod_{i=1}^{n-M} \rho_{\alpha_{[i:i+M]}} \nu(\alpha_{n-M+1} \cdots \alpha_n) > 0.$$

Then, we have for all $n \geq M$,

$$\begin{aligned} \frac{\log \mu(S^{(n)}[\alpha](X))}{n} &= \frac{1}{n} \left(\sum_{i=1}^{n-M} \log \rho_{\alpha_{[i:i+M]}} + \log \nu(\alpha_{n-M+1} \cdots \alpha_n) \right) \\ &= \frac{1}{n} \left(\sum_{\beta \in \mathcal{A}^{M+1}} N_\beta^n[\alpha] \log \rho_\beta + \log \nu(\alpha_{n-M+1} \cdots \alpha_n) \right), \end{aligned}$$

where $N_\beta^n[\alpha]$ denotes the number of times where we can find the string $\beta \in \mathcal{A}^{M+1}$ in the initial segment of length $n \geq M+1$ of $\alpha \in \Sigma$.

But, by the ergodic theorem, we have for all $\beta \in \mathcal{A}^{M+1}$,

$$\frac{N_\beta^n[\alpha]}{n} \xrightarrow{n \rightarrow \infty} \rho_\beta \nu(\beta_2 \cdots \beta_{M+1}) \quad \text{for } \bar{\nu} \text{ almost all } \alpha \in \Sigma. \quad (2.5.8)$$

Then, we have

$$\frac{\log \mu(S^{(n)}[\alpha](X))}{n} \xrightarrow{n \rightarrow \infty} \sum_{i, \beta \in \mathcal{A} \times \mathcal{A}^M} \nu(\beta) \log \rho_{i\beta} \quad \text{for } \bar{\nu} \text{ almost all } \alpha \in \Sigma. \quad (2.5.9)$$

Indeed,

$$\begin{aligned} h_\mu &= - \int_X \sum_{i \in \mathcal{A}} p_i(x) \log p_i(x) d\mu(x) \\ &= - \int_{\bigcup_{\alpha \in \mathcal{A}^M} S^M[\alpha](X)} \sum_{i \in \mathcal{A}} p_i(x) \log p_i(x) d\mu(x) \quad \text{since } \text{supp } \mu \subset K_X \subset \bigcup_{\alpha \in \mathcal{A}^M} S^M[\alpha](X) \\ &= - \sum_{i \in \mathcal{A}} \sum_{\alpha \in \mathcal{A}^M} \rho_{i\alpha} \log(\rho_{i\alpha}) \mu(S^M[\alpha](X)) \\ &= - \sum_{(i, \alpha) \in \mathcal{A} \times \mathcal{A}^M} \log(\rho_{i\alpha}) \rho_{i\alpha} \nu(\alpha_1 \cdots \alpha_M) \end{aligned}$$

Therefore, we have

$$\lim_{r \rightarrow 0} \frac{\log \mu(S^{(N(C^{-1}r))}[\alpha](X))}{N(C^{-1}r)} = \lim_{r \rightarrow 0} \frac{\log \mu(S^{(N(Crc^{-d_{N(r)}(\alpha))})}[\alpha](X))}{N(Crc^{-d_{N(r)}(\alpha)})} = -h_\mu.$$

Step 4 It remains to prove the equality (2.5.3), i.e. that

$$\lim_{r \rightarrow 0} \frac{\log r}{N(C^{-1}r)} = \lim_{r \rightarrow 0} \frac{\log r}{N(Crc^{-d_{N(r)}(\alpha)})} = -\chi_\mu.$$

Let us start by proving that

$$\lim_{r \rightarrow 0} \frac{\log r}{N(r)} = -\chi_\mu.$$

Let $\Sigma \ni \omega \mapsto f(\omega) := \log |S'_{\omega_1}(\phi(\sigma\omega))| \in \mathbb{R}$. Notice that this function is well defined thanks to the condition (2.5.1).

Then for all $n \in \mathbb{N}^*$,

$$\frac{1}{n} \sum_{i=1}^n \log \left| S'_{\alpha_i} (S^{n-i} [\sigma^i \alpha] (\phi(\sigma^n \alpha))) \right| = \frac{1}{n} \sum_{i=1}^n f(\sigma^{i-1} \alpha) \xrightarrow{n \rightarrow \infty} \int_{\Sigma} f(\omega) \bar{\nu}(\mathrm{d}\omega)$$

by the ergodic theorem for $\bar{\nu}$ almost all $\alpha \in \Sigma$. Since N tends to infinity when r tends to 0, we have

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{\log r}{N(r)} &= \lim_{N \rightarrow \infty} \frac{\log (\prod_{i=1}^N |S'_{\alpha_i} (S^{n-i} [\sigma^i \alpha] (\phi(\sigma^n \alpha)))|)}{N} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \log \left| S'_{\alpha_i} (S^{n-i} [\sigma^i \alpha] (\phi(\sigma^n \alpha))) \right| \\ &= \int_{\Sigma} f(\omega) \mathrm{d}\bar{\nu}(\omega). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \int_{\Sigma} f(\omega) \mathrm{d}\bar{\nu}(\omega) &= \sum_{i \in \mathcal{A}} \int_{\{i\} \times \Sigma} \log |S'_i(\phi(\sigma\omega))| \mathrm{d}\bar{\nu}(\omega) \\ &= \sum_{(i, \alpha) \in \mathcal{A} \times \mathcal{A}^M} \int_{\{i\} \times \{\alpha\} \times \Sigma} \log |S'_i(\phi(\sigma\omega))| \rho_{i\alpha} \mathrm{d}\bar{\nu}(\sigma\omega) \\ &= \sum_{(i, \alpha) \in \mathcal{A} \times \mathcal{A}^M} \rho_{i\alpha} \int_{\{\alpha\} \times \Sigma} \log |S'_i(\phi(\omega))| \mathrm{d}\bar{\nu}(\omega) \\ &= \sum_{(i, \alpha) \in \mathcal{A} \times \mathcal{A}^M} \rho_{i\alpha} \int_{S^{(M)}[\alpha](X)} \log |S'_i(x)| \mathrm{d}\mu(x) = -\chi_{\mu}. \end{aligned}$$

Notice that we have

$$\lim_{r \rightarrow 0} \frac{\log r}{\log N(r)} = \lim_{r \rightarrow 0} \frac{\log C^{-1}r}{\log N(C^{-1}r)} = \lim_{r \rightarrow 0} \frac{\log r}{\log N(C^{-1}r)}.$$

On the other hand, we can show that $\frac{d_{N(r)}}{\log N(r)}$ tends to 0 when r goes to 0 for $\bar{\nu}$ almost all α in Σ . Indeed, if

$$\limsup_{n \rightarrow +\infty} \frac{\max(\inf\{i, \alpha_{n+i} \neq 0\}, \inf\{i, \alpha_{n+i} \neq D-1\})}{n} = \varepsilon > 0 \text{ for } \bar{\nu} \text{ almost all } \alpha,$$

then

$$\sigma^n \alpha \in s_{[\varepsilon n]:1}(0, 0, \dots) \cup s_{[\varepsilon n]:1}(D-1, D-1, \dots) \text{ infinitely often.}$$

But, $\sum_{n=1}^{\infty} \tilde{\nu}(\sigma^n \alpha \in s_{[[\varepsilon n]:1]}(0, 0, \dots) \cup s_{[[\varepsilon n]:1]}(D-1, D-1, \dots)) < \infty$ thanks to the second part of the condition (2.3.3), thus by Borel-Cantelli lemma, this cannot happen.

Then,

$$\lim_{r \rightarrow 0} \frac{d_{N(r)}}{\log r} = \lim_{r \rightarrow 0} \frac{d_{N(r)} N(r)}{N(r) \log r} = 0.$$

We deduce that

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{\log r}{\log N(r)} &= \lim_{r \rightarrow 0} \frac{\log(Crc^{-d_{N(r)}})}{\log N(Crc^{-d_{N(r)}})} = \lim_{r \rightarrow 0} \frac{\log(C) + \log r - d_{N(r)} \log c}{\log N(Crc^{-d_{N(r)}})} \\ &= \lim_{r \rightarrow 0} \frac{\log r \left(\frac{\log(C)}{\log r} + 1 - \frac{d_{N(r)} \log c}{\log r} \right)}{\log N(Crc^{-d_{N(r)}})} = \lim_{r \rightarrow 0} \frac{\log r}{\log N(Crc^{-d_{N(r)}})}. \end{aligned}$$

Therefore, the equality (2.5.3) is proved. □

In [29], the Hausdorff dimension of measures driven by Markov chains is studied. This study reduces to our case with specific contractions $S_i : x \mapsto \frac{x+i}{D}$ on the state space $X = [0, 1]$. The dimension turns out to be the same in both cases.

Example 2.5.6. We consider the same IFS as in the example (2.4.9) and we assume that the condition (2.3.3) is satisfied. Then, we can calculate the Hausdorff dimension of the support of μ . We have

$$h_\mu = - \sum_{\alpha_1, \alpha_2} \nu(\alpha_2) \rho_{\alpha_1 \alpha_2} \log \rho_{\alpha_1 \alpha_2} = - \frac{\rho_{01} (\rho_{00} \log \rho_{00} + \rho_{10} \log \rho_{10}) + \rho_{10} (\rho_{01} \log \rho_{01} + \rho_{11} \log \rho_{11})}{\rho_{01} + \rho_{10}},$$

and

$$\chi_\mu = - \sum_{\alpha_1, \alpha_2} \rho_{\alpha_1 \alpha_2} \times \log_2 \left(\frac{1}{2} \right) \mu(S_{\alpha_2}(X)) = \sum_{\alpha_2} \mu(S_{\alpha_2}(X)) \sum_{\alpha_1} \rho_{\alpha_1 \alpha_2} = \sum_{\alpha_2} \mu(S_{\alpha_2}(X)) = 1.$$

Thus, the Hausdorff dimension of μ is

$$\dim_H \mu = h_\mu.$$

Example 2.5.7. Let $\mathcal{S} = \{(S_i, p_i), i \in \{0, 1\}\}$ be an IFS on $X = [0, 1]$ such that

$$S_0(x) = \frac{x}{3} \quad \text{and} \quad p_0(x) = \rho_{00} \mathbb{1}_{[0, \frac{1}{3}]}(x) + q_0(x) \mathbb{1}_{[\frac{1}{3}, \frac{2}{3}]}(x) + \rho_{01} \mathbb{1}_{[\frac{2}{3}, 1]}(x),$$

$$S_1(x) = \frac{x+2}{3} \quad \text{and} \quad p_1(x) = \rho_{10} \mathbb{1}_{[0, \frac{1}{3})}(x) + q_1(x) \mathbb{1}_{[\frac{1}{3}, \frac{2}{3})}(x) + \rho_{11} \mathbb{1}_{[\frac{2}{3}, 1]}(x),$$

with q_0 and q_1 two positive measurable functions such that their sum is equal to 1. We assume that the condition (2.3.3) is satisfied. Then there exists a unique invariant measure μ and we can easily calculate its Hausdorff dimension. The entropy is the same as in the previous example and the Lyapunov exponent is

$$\begin{aligned} \chi_\mu &= - \sum_{\alpha_1, \alpha_2} \rho_{\alpha_1 \alpha_2} \times \log_2 \left(\frac{1}{3} \right) \mu(S_{\alpha_2}(X)) = \log_2(3) \sum_{\alpha_2} \mu(S_{\alpha_2}(X)) \sum_{\alpha_1} \rho_{\alpha_1 \alpha_2} \\ &= \log_2(3) \sum_{\alpha_2} \mu(S_{\alpha_2}(X)) = \log_2(3). \end{aligned}$$

Then, we have

$$\dim_H \mu = \frac{h_\mu}{\log_2(3)} = \log_3(2) h_\mu.$$

2.6 Conclusion and open problems

In this paper, we give criteria of existence and uniqueness of an invariant probability measure for iterated function systems with adapted piecewise constant transition probabilities. The natural continuation of this work is to extend these criteria to the same type of iterated function system but with general transition probabilities. This will be the object of a future paper. On the other hand, we show the Hausdorff dimension of the invariant measure is equal to the quotient of the entropy by the Lyapunov exponent. It could be interesting to verify whether the formula remains true with general transition probabilities.

Acknowledgements. I would like to thank both referees for their detailed reports and their numerous remarks and suggestions which have significantly improved this article. I also would like to thank my Ph.D advisor Dimitri Petritis for his precious help and his advice during the writing of this paper and the Centre Henri Lebesgue (programme “Investissements d’avenir” — ANR-11-LBX-0020-01) and the local government of Brittany for their financial support.

Chapitre 3

A criterion of asymptotic stability for Variable Length Markov Chains (Article à soumettre)

A CRITERION OF ASYMPTOTIC STABILITY FOR VARIABLE LENGTH MARKOV
CHAINS

Blandine DUBARRY

Abstract

The asymptotic stability of Variable length Markov chains is proved under conditions of weak non-nullness and continuity for the ultrametric distance of the transition probabilities.

Keywords : Variable length Markov chain; context tree; invariant measure; asymptotic stability; Markov operator.

Mathematical Subject Classification : 60J05; 60J10; 37B10; 37A30; 60K30; 94A45.

3.1 Introduction

The name variable length Markov chain (VLMC) was coined by Bühlmann and Wyner [7] in 1999 but chains with memory of variable length were introduced by Rissanen in 1983 as a tool for data compression of long strings generated by non independent information sources. Rather than considering a model in which the source is a fixed order Markov chain, Rissanen considered that the length of the past we need to know in order to decide the next symbol depended on the past itself. He called the relevant part of the past a context. The VLMC model has two main advantages : it requires less storage than a fixed order Markov chain and is so less expansive, and it highlights the structural dependencies present in the data. In [52], Rissanen gave the algorithm context which aimed at estimating the probabilistic context tree (see more formal definition below) generating the chain. Given a sample generated by a “finitely generated” source (as Rissanen called it), the algorithm builds inductively a context tree adapted to the sample reading one by one the symbols in the string. At each step, it associates with each context an estimated probability transition defined as the proportion of time the context appears in the sample followed by each one of the symbols in the alphabet. In addition, it associates a cost with each context and accepts a context only if its share in reducing entropy exceeds its cost. The algorithm stops when the estimated probabilities are stabilized and the tree remains unchanged. Several variants and improvements can be found in literature including the ones of Ferrari and Wyner ([22]) in 2003. To the best of our knowledge, these authors were the first to consider the case of unbounded probabilistic context trees. If the length of the contexts is not bounded, then chains with memory of variable length are chains of infinite order.

VLMCs have nowadays other fields of applications. Indeed, they can be used for modelling discrete general stochastic chains taking values in a finite alphabet. This is the case in genomics to classify DNA or protein sequences (see for instance [41]). Biologists call them Probabilistic Suffix Trees (PST). An overview on VLMC can be found in Galves and Löcherbach [25].

Let \mathcal{A} be a finite alphabet and $(\tilde{q}_i(\cdot))_{i \in \mathcal{A}}$ a family of transition probabilities on $\mathcal{A}^{\mathbb{N}}$. We consider a VLMC $(Y_n)_n$ on $\mathcal{A}^{\mathbb{N}}$ such that for any $i \in \mathcal{A}$, for any $\alpha \in \mathcal{A}^{\mathbb{N}}$, we have

$$\mathbb{P}(Y_{n+1} = i\alpha | Y_n = \alpha) = \tilde{q}_i(\alpha).$$

The existence and uniqueness of an invariant measure for VLMC have been the object of many studies. Most of the studies use the same two assumptions to guarantee the existence and uniqueness of an invariant measure for a VLMC : the family of transition probabilities $(\tilde{q}_i(\cdot))$ has to be weakly non-null (that means that at least one of

the infima on the \tilde{q}_i has to be positive) and the continuity rate $(\beta_k)_{k \geq 1}$ of the family $(\tilde{q}_i)_{i \in \mathcal{A}}$ has to be summable. The continuity rate is defined by

$$\beta_k = \sup_i \sup_{\alpha_1 \dots \alpha_k = \beta_1 \dots \beta_k} |\tilde{q}_i(\alpha) - \tilde{q}_i(\beta)|.$$

The second assumption implies that $\lim_{k \rightarrow +\infty} \beta_k = 0$ and consequently, the family of transition probabilities is continuous for the ultrametric distance. In 2011, Gallo ([23]) managed to get rid of this continuity assumption, replacing it by one over the structure of the context tree. But, this criterion did not allow him to study all types of context trees. In 2012, Cénac et al. ([10]) have entirely determined for two examples (including the infinite comb we will define later), necessary and sufficient conditions to have an invariant measure and to ensure its uniqueness. We can also mention the work of Gallo and Paccaut ([24]) in 2013 who were interested in discontinuous transition functions. Their criterion is mainly based on considerations on the set of discontinuity points to ensure existence and supplementary conditions to control the variations of the set of continuity points to provide uniqueness. They also use a condition of non-nullness of all transition probabilities on a subset of $\mathcal{A}^{\mathbb{N}}$.

In this paper, we prove the existence and uniqueness of an invariant probability measure and even its attractivity under continuity and non-nullness assumptions but without any variation control.

Theorem 3.1.1. *(Asymptotic stability of a VLMC) Let $(\mathcal{T}, (q.(c))_{c \in \mathcal{C}(\mathcal{T})})$ be a context tree and $(Y_n)_n$ a VLMC associated with this context tree (see definitions below). We denote by Q the Markov operator of the VLMC. Under conditions of weak non-nullness and continuity detailed below, there exists a unique invariant probability measure $\mu \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$ such that for any measure $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$,*

$$\mu' Q^k \xrightarrow[k \rightarrow \infty]{} \mu.$$

The main idea of the proof consists in approaching a general context tree by a sequence of finite context trees and use results of asymptotic stability for VLMC associated with finite context trees. The idea of approaching VLMC by Markov chains of finite order is not new. In [21], Fernández and Galves give an upper bound of Ornstein's \bar{d} -distance between a chain of infinite order and its canonical approximation at any order which is proportional to the continuity rate of the chain of infinite order.

3.2 Preliminaries and notation

3.2.1 Variable Length Markov Chain and context tree

Let $\mathcal{A} = \{0, \dots, D-1\}$ be a finite alphabet. Then $\bigcup_{n \geq 0} \mathcal{A}^n$ is the set of finite words on \mathcal{A} and $\mathcal{A}^{\mathbb{N}}$ is the set of right-infinite words on \mathcal{A} . Let \mathcal{T} be a context tree over \mathcal{A} , i.e. a tree whose nodes have either 0 or D children and which has a finite or denumerable set of leaves $\mathcal{C}(\mathcal{T})$, called contexts. With each context $c \in \mathcal{C}(\mathcal{T})$, we associate a probability measure $q.(c)$ over \mathcal{A} such that $\sum_{i=0}^{D-1} q_i(c) = 1$. The pair $(\mathcal{T}, (q.(c))_{c \in \mathcal{C}(\mathcal{T})})$ is called a context tree. Let us define the prefix function, noted pref . Let w a finite or infinite word on \mathcal{A} . If $w \in \mathcal{T}$, then $\text{pref}(w) := w$. If $w \notin \mathcal{T}$, $\text{pref}(w)$ is the largest word $w_1 \dots w_N \in \mathcal{C}(\mathcal{T})$ such that w has $w_1 \dots w_N$ as a prefix. Notice that the prefix function depends on \mathcal{T} . Not to burden notation, we omit this dependence in the sequel.

Example 3.2.1. Let us see how the function pref is defined on the infinite comb (see figure 3.1 for a description of the tree). For any word w , we have

$$\text{pref}(1w) = 1, \quad \text{pref}(0^n 1w) = 0^n 1 \quad \text{and} \quad \text{pref}(0^n) = 0^n.$$

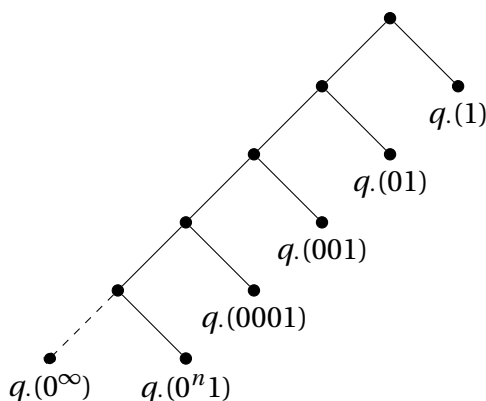


Figure 3.1: Infinite comb probabilistic context tree

The Variable Length Markov Chain (VLMC) associated with the context tree is the chain $(Y_n)_{n \geq 0}$ with state space $\mathcal{A}^{\mathbb{N}}$, defined by the transition probabilities for all $n \in \mathbb{N}$, for all $i \in \mathcal{A}$,

$$\mathbb{P}(Y_{n+1} = i | Y_n) = q_i(\text{pref}(Y_n)).$$

In addition, we define for all $i \in \mathcal{A}$ the extended transition functions $\tilde{q}_i : \mathcal{A}^{\mathbb{N}} \ni \alpha \mapsto \tilde{q}_i(\alpha) := q_i(\text{pref}(\alpha))$. If we denote by Q the Markov operator of the chain, we have for

any $i \in \mathcal{A}$, for any $\alpha \in \mathcal{A}^{\mathbb{N}}$,

$$Q(\alpha, i\alpha) = \tilde{q}_i(\alpha).$$

Denoting by $(A_n)_{n \in \mathbb{Z}}$ the last letter process of $(Y_n)_n$, we also have for any $\alpha = (\alpha_1, \alpha_2, \dots) \in \mathcal{A}^{\mathbb{N}}$, for any $i \in \mathcal{A}$,

$$\mathbb{P}(A_n = i | A_{n-1} = \alpha_1, A_{n-2} = \alpha_2, \dots) = \tilde{q}_i(\alpha) = q_i(\text{pref}(\alpha)).$$

We can notice that $(A_n)_{n \in \mathbb{Z}}$ becomes not Markovian as soon as the context tree has at least one infinite context. When the tree is finite, $(A_n)_{n \geq 0}$ is a Markov chain whose order is the length of its longest branch.

3.2.2 Asymptotic stability for VLMC related to a finite context tree

Before giving a criterion for the asymptotic stability of a VLMC associated with a general context tree, we start by giving one for a VLMC associated with a finite context tree.

Definition 3.2.2. (*Asymptotic stability*) A Markov chain (or its Markov operator R) is said asymptotically stable if it admits a unique invariant measure π which is attractive for the weak convergence i.e. for every measure π' , we have

$$\pi' R^k \xrightarrow[k \rightarrow \infty]{} \pi.$$

Proposition 3.2.3. We consider a finite context tree $(\mathcal{T}^M, (q_i^M(c))_{c \in \mathcal{C}(\mathcal{T}^M)})$ of length M . We assume that

$$\sum_{i \in \mathcal{A}} \inf_c q_i^M(c) > 0.$$

Then, the VLMC $(Y_n)_{n \in \mathbb{N}}$ associated with the context tree admits a unique invariant probability measure μ_M . In addition, if we denote by Q_M the Markov operator of the VLMC, it is asymptotically stable, i.e. for any $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$,

$$\mu' Q_M^k \xrightarrow[k \rightarrow \infty]{} \mu_M.$$

Remarque 3.2.4. The result is almost the same as in [15]. The author gives the same criterion of asymptotic stability for a class of iterated function systems. The iterated function system on the space $[0, 1]$ given by the D transformations $S_i(x) = \frac{x+i}{D}$ and adapted transition probabilities is part of the class. By mapping the process $(A_n)_n$ onto the unit interval, we almost have a bijection with the Markov process given by the iterated function system. Thus, the ideas of proof are the same except we do not reduce the set $\mathcal{A}^{\mathbb{N}}$ since we do not need the bijection. We give here the main ideas.

Main ideas of the proof. For the sake of simplicity, we assume that the tree \mathcal{T}^M is saturated i.e. each node has exactly D children. When a given context corresponds to a leaf c of the tree at some earlier generation $m < M$, the branch leading to c is extended up to generation M by assigning the same probability measure $q_i(c')$ to all c' in the umbra of c until generation M .

Step 1 We first consider an auxiliary chain $(B_n)_n$ on the extended alphabet \mathcal{A}^M and we prove that it is asymptotically stable.

If $(A_n)_n$ is the last letter process of $(Y_n)_n$, we define $(B_n)_n$ by $B_n := A_n \cdots A_{n-M+1}$. $(B_n)_n$ is a \mathcal{A}^M -valued Markov chain of order 1 whose transition matrix P_M is given for any $\alpha_1, \dots, \alpha_{M+1}$ and $i \in \mathcal{A}$, by

$$P_M(\alpha_1 \cdots \alpha_M, i \alpha_1 \dots \alpha_{M-1}) = q_i^M(\alpha_1 \dots \alpha_M).$$

Assuming $\sum_{i \in \mathcal{A}} \inf_c q_i^M(c) > 0$, the chain $(B_n)_n$ admits a unique invariant probability measure ν_M . In addition, for any $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$, for any $\alpha \in \mathcal{A}^{M+1}$, we have

$$\nu' P_M^k(\alpha) \xrightarrow[k \rightarrow \infty]{} \nu_M(\alpha).$$

Step 2 We can then prove that the VLMC $(Y_n)_n$ admits a unique invariant measure μ_M .

The family of cylinder sets $\{\alpha_1 \times \cdots \times \alpha_n \times \mathcal{A}^{\mathbb{N}} \mid \alpha_1, \dots, \alpha_n \in \mathcal{A}, n \in \mathbb{N}\}$ generates the σ -algebra of $\mathcal{A}^{\mathbb{N}}$. Then, a probability measure μ_M is invariant if and only if for $n \in \mathbb{N}$, for any $\alpha \in \mathcal{A}^{\mathbb{N}}$,

$$\mu_M Q_M(\{\alpha_1 \cdots \alpha_n \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}}) = \mu_M(\{\alpha_1 \cdots \alpha_n \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}}).$$

We can prove that we can express $\mu_M Q_M^n(\{\alpha_1 \cdots \alpha_n \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}})$ in terms of the family $(q_i(\alpha))_{(i, \alpha) \in \mathcal{A} \times \mathcal{A}^M}$ and $\mu_M(\{\alpha'_1 \cdots \alpha'_M \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}})$ with some $\alpha' \in \mathcal{A}^M$. Then, we can show that under the condition (3.3.1) μ_M is invariant for $(Y_n)_n$ if and only if ν_M is invariant for $(B_n)_n$ where ν_M is related to μ_M by $\nu_M(\alpha) = \mu(\{\alpha_1 \cdots \alpha_M \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}})$, for all $\alpha \in \mathcal{A}^M$. Then, we can conclude that there exists a unique invariant probability measure μ_M for the VLMC $(Y_n)_n$, defined for any $n \geq M$ and any $\alpha \in \mathcal{A}^n$, by

$$\mu_M(\{\alpha_1 \cdots \alpha_n \beta\} \mid \beta \in \mathbb{A}^{\mathbb{N}}) = \prod_{i=1}^{n-M} q_{\alpha_i}^M(\alpha_{i+1} \cdots \alpha_{i+M}) \nu_M(\alpha_{n-M+1} \cdots \alpha_n).$$

Step 3 We end the proof by the asymptotic stability of the Markov operator Q_M . It is enough to prove that for each element A of a family that generates the σ -algebra of $\mathcal{A}^{\mathbb{N}}$, we have for any measure $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$, $\mu' Q_M^k(A) \xrightarrow[k \rightarrow \infty]{} \mu_M(A)$. By Portmanteau's theorem [4], we should conclude that $\mu' Q_M^k \Rightarrow \mu_M$.

Let $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$. We can associate a measure $\nu' \in \mathcal{M}_1(\mathcal{A}^M)$ defined for any $\alpha \in \mathcal{A}^M$ by

$$\nu'(\alpha) = \mu'(\{\alpha_1 \cdots \alpha_M \beta\} | \beta \in \mathbb{A}^{\mathbb{N}}).$$

The key point is to use the convergence of $\nu' P_M^k$ (proved in step 1) to conclude. \square

3.3 Asymptotic stability of a VLMC

Throughout this section, we will consider a context tree $(\mathcal{T}, (q_i(c))_{c \in \mathcal{C}(\mathcal{T})})$ with extended transition functions $\tilde{q}_i(\alpha)$ for any $\alpha \in \mathcal{A}^{\mathbb{N}}$ and Markov operator Q .

The aim of this section is to prove the asymptotic stability for a VLMC associated with a general context tree under the following conditions.

Condition 3.3.1. (*Weak non-nullness condition*) We assume either of the two equivalent conditions

$$\sum_{i \in \mathcal{A}} \inf_{c \in \mathcal{C}(\mathcal{T})} q_i(c) > 0 \quad \text{or} \quad \sum_{i \in \mathcal{A}} \inf_{\alpha \in \mathcal{A}^{\mathbb{N}}} \tilde{q}_i(\alpha) > 0.$$

Condition 3.3.2. (*Continuity for the ultrametric distance*) The transition probabilities $\tilde{q}_i(\cdot)$ on $\mathcal{A}^{\mathbb{N}}$ are continuous for the ultrametric distance i.e. the distance defined for any $\alpha, \beta \in \mathcal{A}^{\mathbb{N}}$ by

$$d(\alpha, \beta) = 2^{-\sup_n (\alpha_1 \cdots \alpha_n = \beta_1 \cdots \beta_n)}.$$

We start by approaching the context tree by a sequence composed of the tree \mathcal{T} shortened at rank n and the transition probabilities $(\tilde{q}_i^n(\alpha))_n$ that converge towards the transition probabilities $\tilde{q}_i(\alpha)$ for any $i \in \mathcal{A}$ and $\alpha \in \mathcal{A}^{\mathbb{N}}$. The following lemma shows how to choose the $(\tilde{q}_i^n(\alpha))_n$ if the initial context tree verifies the condition (3.3.2).

Lemma 3.3.3. Under the condition (3.3.2), for any $i \in \mathcal{A}$, the sequence $(\tilde{q}_i^n(\cdot))_{n \in \mathbb{N}}$ defined for any $\alpha \in \mathcal{A}^{\mathbb{N}}$ by $\tilde{q}_i^n(\alpha) := \tilde{q}_i(\alpha_1 \cdots \alpha_n \mathbf{0})$ (where $\mathbf{0}$ denotes the right-infinite sequence of 0) converges towards the function $\tilde{q}_i(\cdot)$ i.e. for any $i \in \mathcal{A}$, for any $\alpha \in \mathcal{A}^{\mathbb{N}}$, we have

$$\tilde{q}_i^n(\alpha) \xrightarrow[n \rightarrow \infty]{} \tilde{q}_i(\alpha).$$

In addition, we notice that the functions $\tilde{q}_i^n(\alpha)$ are probability measures on \mathcal{A} for any $\alpha \in \mathcal{A}^{\mathbb{N}}$ and if the condition (3.3.2) is verified, then $\inf_{i \in \mathcal{A}, \alpha \in \mathcal{A}^{\mathbb{N}}} \tilde{q}_i^n(\alpha) > 0$.

Proof. Let $\varepsilon > 0$ and $\alpha \in \mathcal{A}^{\mathbb{N}}$. Let $i \in \mathcal{A}$. There exists $N \in \mathbb{N}$ such that for any $\beta \in \mathcal{A}^{\mathbb{N}}$ with $d(\alpha, \beta) \leq 2^{-N}$, we have $|\tilde{q}_i(\alpha) - \tilde{q}_i(\beta)| < \varepsilon$. In particular, for any $n \geq N$, we have $|\tilde{q}_i(\alpha) - \tilde{q}_i(\alpha_1 \cdots \alpha_n \mathbf{0})| < \varepsilon$. Therefore, for any $n \geq N$, we have $|\tilde{q}_i(\alpha) - \tilde{q}_i^n(\alpha)| < \varepsilon$ and hence the convergence. \square

We can now prove the asymptotic stability of the VLMC of a general context tree under the conditions (3.3.1) and (3.3.2).

Theorem 3.3.4. (*Asymptotic stability of a VLMC*) Let $(\mathcal{T}, (q_i(c))_{c \in \mathcal{C}(\mathcal{T})})$ be a context tree and $(Y_n)_n$ a VLMC associated with this context tree. We denote by Q the Markov operator of the VLMC. Under the conditions (3.3.1) and (3.3.2), there exists a unique invariant probability measure $\mu \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$ such that for any measure $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$,

$$\mu' Q^k \xrightarrow[k \rightarrow \infty]{} \mu.$$

Proof. We consider a sequence of finite context trees $(\mathcal{T}^n, (q_i^n(c))_{c \in \mathcal{C}(\mathcal{T}^n)})_{n \in \mathbb{N}}$ as in the lemma (3.3.3). We can denote by Q_n the Markov operators associated with the transition probabilities $(q_i^n(\cdot))_{i \in \mathcal{A}}$.

By the proposition (3.2.3), each of these context trees admits a unique invariant probability measure μ_n such that for any probability measure $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$, we have

$$\mu' Q_n^k \xrightarrow[k \rightarrow \infty]{} \mu_n.$$

As the space $\mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$ is a compact space, by Prokhorov theorem [51], there exists a subsequence $(\mu_{n_l})_{l \in \mathbb{N}}$ of $(\mu_n)_{n \in \mathbb{N}}$ and a probability measure $\mu \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$ such that

$$\mu_{n_l} \xrightarrow[l \rightarrow \infty]{} \mu.$$

We aim at proving that μ is invariant for the Markov operator Q and even attractive i.e. that for any $\mu' \in \mathcal{M}_1(\mathcal{A}^{\mathbb{N}})$, we have

$$\mu' Q^k \xrightarrow[k \rightarrow \infty]{} \mu.$$

We can write the difference $\mu' Q^k - \mu$ as a sum of three terms and show that each of them weakly converges towards 0.

$$\mu' Q^k - \mu = \mu' Q^k - \mu' Q_{n_l}^k + \mu' Q_{n_l}^k - \mu_{n_l} + \mu_{n_l} - \mu. \quad (3.3.1)$$

First term Let us show that $\mu'Q^k - \mu'Q_{n_l}^k \xrightarrow[k, l \rightarrow \infty]{} 0$.

Let f be a continuous bounded function on $\mathcal{A}^{\mathbb{N}}$.

$$\begin{aligned} & \int_{\mathcal{A}^{\mathbb{N}}} f(\beta) \mu'Q^k(d\beta) - \int_{\mathcal{A}^{\mathbb{N}}} f(\beta) \mu'Q_{n_l}^k(d\beta) \\ &= \int_{\mathcal{A}^{\mathbb{N}}} \sum_{\alpha_1 \in \mathcal{A}} \tilde{q}_{\alpha_1}(\beta) f(\alpha_1 \beta) \mu'Q^{k-1}(d\beta) - \int_{\mathcal{A}^{\mathbb{N}}} \sum_{\alpha_1 \in \mathcal{A}} \tilde{q}_{\alpha_1}^{n_l}(\beta) f(\alpha_1 \beta) \mu'Q_{n_l}^{k-1}(d\beta) \\ &= \dots \\ &= \int_{\mathcal{A}^{\mathbb{N}}} \sum_{\alpha_1, \dots, \alpha_k \in \mathcal{A}^k} \left[\prod_{i=1}^k \tilde{q}_{\alpha_i}(\beta) f(\alpha_{i-1} \dots \alpha_1 \beta) - \prod_{i=1}^k \tilde{q}_{\alpha_i}^{n_l}(\beta) f(\alpha_{i-1} \dots \alpha_1 \beta) \right] \mu'(d\beta). \end{aligned}$$

As the functions $\tilde{q}_i^{n_l}(\cdot)$ converge towards the functions $\tilde{q}_i(\cdot)$ for any $\alpha \in \mathcal{A}^{\mathbb{N}}$ by lemma (3.3.3), we conclude that the difference converges towards 0 by dominated convergence, which proves $\mu'P^k - \mu'Q_{n_l}^k \xrightarrow[l \rightarrow \infty]{} 0$ for any $k \in \mathbb{N}$. Then, we have $\mu'P^k - \mu'Q_{n_l}^k \xrightarrow[k, l \rightarrow \infty]{} 0$.

Second term The weak convergence of $\mu'Q_{n_l}^k - \mu_{n_l}$ towards 0 when k tends to infinity for any $l \in \mathbb{N}$ is ensured by the proposition (3.2.3). Then, we have $\mu'Q_{n_l}^k - \mu_{n_l} \xrightarrow[k, l \rightarrow \infty]{} 0$.

Third term Finally, we have $\mu_{n_l} - \mu \xrightarrow[l \rightarrow \infty]{} 0$ by construction of μ and thus $\mu_{n_l} - \mu \xrightarrow[k, l \rightarrow \infty]{} 0$.

Thus, the weak convergence is proved and the Markov operator P of the VLMC is asymptotically stable. \square

Example 3.3.5. Let $(\mathcal{T}_1, (q_0(c))_{c \in \mathcal{C}(\mathcal{T}_1)})$ be an infinite comb context tree. According to the theorem 3.3.4, if the infima and suprema of q_0 and q_1 are included in $(0, 1)$ and if the extended transition functions are continuous on $\mathcal{A}^{\mathbb{N}}$, then there exists a unique stationary VLMC with these transition probabilities.

Remarque 3.3.6. The previous result covers the results of [10]. Indeed, in the case when $q_0(0^\infty) \neq 1$, it is proved the existence of a unique stationary process if and only if the numerical series of general term $c_n := \prod_{k=0}^{n-1} q_0(0^k 1)$ converges. As for all n , c_n is majorised by $\sup_c q_0(c)$ which is strictly smaller than 1, we have convergence of the series.

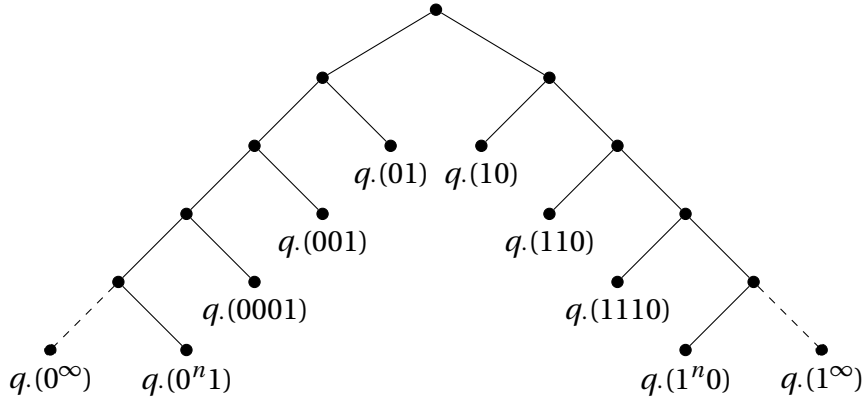


Figure 3.2: Double infinite comb probabilistic context tree

Then, in 2013, Cénac et al. ([9]) reuse results on infinite comb for persistent random walks and study a new example : the double infinite comb (see figure 3.2). A complete study on existence and uniqueness of invariant process is also given.

Example 3.3.7. Let $(\mathcal{T}_2, (q.(c))_{c \in \mathcal{C}(\mathcal{T})})$ be a double infinite comb probabilistic context tree. According to the theorem 3.3.4, if the infima and suprema of q_0 and q_1 are included in $(0, 1)$ and if the extended transition functions are continuous on $\mathcal{A}^{\mathbb{N}}$, then there exists a unique stationary VLMC with these transition probabilities.

Remarque 3.3.8. Again, the previous result covers the ones of [9]. Indeed, in the case when $q_0(0^\infty) \neq 1$ and $q_1(0^\infty) \neq 1$, it is proved the existence of a unique stationary process if both numerical series of general terms $c_n^0 := \prod_{k=0}^{n-1} q_0(0^k 1)$ and $c_n^1 := \prod_{k=0}^{n-1} q_1(0^k 1)$ converge. As previously, the convergences are ensured since the suprema are strictly smaller than 1.

We can notice that any continuity assumption was needed to ensure the existence and uniqueness of an invariant measure for the infinite comb and the double infinite comb. However, adding a continuity assumption, we add a new information by proving the asymptotic stability a the Markov operators associated with the VLMC of both trees. In addition, the two studies were very specific and the theorem 3.3.4 enables to give a criterion of asymptotic stability for any context tree.

3.4 Conclusion and open problems

In this paper, we give a criterion for the asymptotic stability of a VLMC under assumptions of positivity and continuity of the transition probabilities. The natural continuation of this work would be to relax the continuity assumption under a weaker one. However, without any control on transition probability, it seems to be hard to approach the transition probabilities.

Chapitre 4

Conclusion et problèmes ouverts

Dans cette thèse, nous avons pu étendre les critères de stabilité asymptotique pour une classe d'IFS avec des probabilités de transition adaptées aux transformations qui sont constantes par morceaux. Jusqu'à présent, la majeure partie des études nécessitaient des hypothèses de continuité sur les probabilités de transition dont nous avons pu nous passer ici. Cependant, il serait intéressant de parvenir à étendre les critères d'existence et d'unicité de mesure invariante à des IFS plus généraux avec des probabilités discontinues qui ne soient pas seulement constantes par morceaux. Une piste pour y parvenir pourrait être d'utiliser la notion de φ -irréductibilité [28]. Cependant, nous avons pu calculer la dimension de Hausdorff du support de la mesure dans le cas constant par morceaux et nous avons pu observer qu'elle était (à part cas particuliers) strictement inférieure à 1. Cela implique que la mesure invariante est singulière par rapport à la mesure de Lebesgue. Étant donné que les fonctions φ qui permettent d'obtenir la φ -irréductibilité sont absolument continues par rapport à la mesure invariante (et donc singulière à la mesure de Lebesgue), il semble difficile de parvenir à trouver une mesure φ adéquate. Une autre piste serait d'approcher les fonctions discontinues correspondant aux probabilités de placement par des suites de fonctions constantes par morceaux. La difficulté dans ce cas réside au fait que les intervalles d'adaptation deviennent de plus en plus petits.

D'autre part, nous avons étudié un autre type d'IFS : les chaînes de Markov de longueur variable. Nous avons pu établir pour ces chaînes un critère de stabilité asymptotique sous conditions de stricte positivité et de continuité pour la distance ultramétrique. Une poursuite de ce travail serait de remplacer la condition de continuité par une condition plus faible.

Enfin, nous avons seulement traité le cas de systèmes de fonctions itérées avec des

alphabets finis. Nous pourrions aussi regarder le cas d'un alphabet infini en utilisant par exemple le développement en fractions continues si nous avons besoin d'une bijection entre l'intervalle unité et les suites infinies.

Annexe A

Dimensions fractales

Les différentes dimensions fractales sont des quantités métriques qui quantifient la notion de « contenu » pour des ensembles ayant des dimensions topologiques triviales ou des mesures de Lebesgue nulles. La plus utilisée est la dimension de Hausdorff mais des dimensions de boîtes ou la dimension d'entassement, etc. peuvent parfois être utiles.

Dimension de Hausdorff Pour pouvoir définir la dimension de Hausdorff, nous avons besoin de quelques définitions et propriétés au préalable.

Tout d'abord, on note $|U|$ le diamètre d'un ensemble $U \subset X$. Si on considère F un sous-ensemble de X , un δ -recouvrement de F est une collection au plus dénombrable $(U_i)_i$ de sous-ensembles de X tels que $F \subset \bigcup_i U_i$ et $|U_i| < \delta$. Soit s un réel positif. Pour tout $\delta > 0$, on définit

$$\mathcal{H}_\delta^s(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : (U_i)_i \text{ } \delta\text{-cover of } F \right\}.$$

On peut montrer que $\lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F) \in \overline{\mathbb{R}}_+$ (une démonstration peut se trouver dans [16]) et on définit alors $\mathcal{H}^s(F) := \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F)$ comme la mesure s -dimensionnelle de Hausdorff de l'ensemble F . On peut montrer que $\mathcal{H}^s(\cdot)$ est une mesure extérieure pour tout réel $s > 0$. D'autre part, pour tout sous-ensemble $F \subset X$, on peut montrer qu'il existe un unique réel $s_0 > 0$ tel que

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{if } s < s_0, \\ 0 & \text{if } s > s_0. \end{cases}$$

Le réel s_0 est appelé la dimension de Hausdorff de F et nous avons $0 \leq \mathcal{H}^{s_0}(F) \leq \infty$.

Formellement, la dimension de Hausdorff d'un ensemble F est donnée par :

$$\dim_H(F) := \inf \{s : \mathcal{H}^s(F) = 0\} = \sup \{s : \mathcal{H}^s(F) = \infty\} \in \overline{\mathbb{R}}_+.$$

Si $0 < \mathcal{H}^{s_0}(F) < \infty$, on dit que le fractal a une dimension exacte de Hausdorff s_0 .

Ainsi, l'ensemble tri-adique de Cantor qui a une mesure de Lebesgue nulle et une dimension topologique égale à 0 (car il est totalement déconnecté) a pour dimension de Hausdorff $\ln(2)/\ln(3)$. La courbe de von Koch a, quant à elle, une mesure de Lebesgue 2-dimensionnelle nulle, une dimension topologique égale à 1 et une dimension de Hausdorff égale à $\ln(4)/\ln(3)$.

Dimension de boîtes Soit F un ensemble borné non vide tel que $F \subset \mathbb{R}^d$. Soit $N_\delta(F)$ le plus petit nombre d'ensembles de diamètre au plus δ qui recouvre F . On définit alors les dimensions de boîtes inférieures et supérieures respectivement par

$$\underline{\dim}_B(F) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad \text{et} \quad \overline{\dim}_B(F) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.$$

Si les dimensions inférieures et supérieures sont égales, alors la dimension de boîtes de F est définie par :

$$\dim_B(F) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.$$

Le nombre $N_\delta(F)$ peut être défini de plusieurs manières équivalentes : c'est le plus petit nombre de boules fermées de rayon δ qui recouvrent F , le plus petit nombre de cubes de côté δ qui recouvrent F ou encore le plus grand nombre de boules disjointes de rayon δ dont les centres sont dans F .

Dimension d'entassement (Packing dimension) Cette dimension est définie de manière similaire à la dimension de Hausdorff puisqu'elle est basée sur une mesure même si le type d'ensembles considérés correspond à celui de la dimension de boîtes. Soit $F \subset \mathbb{R}^d$ et

$$\mathcal{P}_\delta^s(F) = \sup \left\{ \sum_{i=1}^{\infty} |B_i|^s : \begin{array}{l} (B_i)_i \text{ ensemble de boules disjointes telles que} \\ |B_i| < \delta \text{ de centre dans } F \end{array} \right\}$$

Puisque $\mathcal{P}_\delta^s(F)$ décroît avec δ , la limite

$$\mathcal{P}_0^s(F) = \lim_{\delta \rightarrow 0} \mathcal{P}_\delta^s(F)$$

existe. Cependant, $\mathcal{P}_0^s(F)$ n'est pas une mesure si F est un ensemble dénombrable dense. On définit donc la mesure d'entassement s -dimensionnelle de $F \subset \mathbb{R}^d$ par

$$\mathcal{P}^s(F) = \inf \left\{ \sum_{i=1}^{\infty} \mathcal{P}_0^s(F_i) : F \subset \bigcup_{i=1}^{\infty} F_i \right\}.$$

La dimension d'entassement est alors définie par

$$\dim_P(F) = \sup \{s : \mathcal{P}^s(F) = \infty\} = \inf \{s : \mathcal{P}^s(F) = 0\}.$$

Relations entre les dimensions On a les relations suivantes entre les dimensions précédentes :

$$\dim_H(F) \leq \dim_P(F) \leq \overline{\dim}_B(F) \quad \text{et} \quad \dim_H(F) \leq \underline{\dim}_B(F) \leq \overline{\dim}_B(F).$$

On peut trouver ces résultats dans Harte [27] ou Falconer [16].

Bibliographie

- [1] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(3):367–394, 1988. [12](#), [23](#)
- [2] Michael Barnsley. *Fractals everywhere*. Academic Press, Inc., Boston, MA, 1988. [23](#)
- [3] Rabi Bhattacharya and Mukul Majumdar. *Random dynamical systems*. Cambridge University Press, Cambridge, 2007. Theory and applications. [9](#)
- [4] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication. [39](#), [55](#)
- [5] David Blackwell. The entropy of functions of finite-state Markov chains. In *Transactions of the first Prague conference on information theory, Statistical decision functions, random processes held at Liblice near Prague from November 28 to 30, 1956*, pages 13–20. Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957. [25](#)
- [6] Maury Bramson and Steven Kalikow. Nonuniqueness in g -functions. *Israel J. Math.*, 84(1-2):153–160, 1993.
- [7] Peter Bühlmann and Abraham J. Wyner. Variable length Markov chains. *Ann. Statist.*, 27(2):480–513, 1999. [50](#)
- [8] M. Campanino and D. Petritis. Random walks on randomly oriented lattices. *Markov Process. Related Fields*, 9(3):391–412, 2003. [6](#), [7](#)
- [9] P. Cénac, B. Chauvin, S. Herrmann, and P. Vallois. Persistent random walks, variable length Markov chains and piecewise deterministic Markov processes. *Markov Process. Related Fields*, 19(1):1–50, 2013. [58](#)
- [10] Peggy Cénac, Brigitte Chauvin, Frédéric Paccaut, and Nicolas Pouyanne. Context trees, variable length Markov chains and dynamical sources. In *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Math.*, pages 1–39. Springer, Heidelberg, 2012. [20](#), [51](#), [57](#)

- [11] Francis Comets, Roberto Fernández, and Pablo A. Ferrari. Processes with long memory: regenerative construction and perfect simulation. *Ann. Appl. Probab.*, 12(3):921–943, 2002. [18](#)
- [12] Basile De Loynes. Random walks on graphs induced by aperiodic tilings. Accepted for publication dans Markov Processes Related Fields, May 2016. [8](#)
- [13] R. L. Dobruš sin. Central limit theorem for nonstationary Markov chains. *Dokl. Akad. Nauk SSSR (N.S.)*, 102:5–8, 1955. [31](#)
- [14] Wolfgang Doeblin and Robert Fortet. Sur des chaînes à liaisons complètes. *Bull. Soc. Math. France*, 65:132–148, 1937. [18](#)
- [15] Blandine Dubarry. A class of iterated function systems with adapted piecewise constant transition probabilities: asymptotic stability and hausdorff dimension of the invariant measure. Submitted for publication, available at https://perso.univ-rennes1.fr/blandine.dubarry/articles/IFS_invariant_measure_Hausdorff_dimension.pdf. [21](#), [53](#)
- [16] Kenneth Falconer. *Fractal geometry*. John Wiley & Sons, Ltd., Chichester, 1990. Mathematical foundations and applications. [15](#), [27](#), [62](#), [64](#)
- [17] Kenneth Falconer. *Techniques in fractal geometry*. John Wiley & Sons, Ltd., Chichester, 1997. [27](#)
- [18] Ai Hua Fan. *Décompositions de mesures et recouvrements aléatoires*, volume 89 of *Publications Mathématiques d'Orsay [Mathematical Publications of Orsay]*. Université de Paris-Sud, Département de Mathématiques, Orsay, 1989. With an English summary. [27](#)
- [19] Ai Hua Fan. Ergodicity, unidimensionality and multifractality of self-similar measures. *Kyushu J. Math.*, 50(2):541–574, 1996. [16](#), [25](#), [42](#)
- [20] Ai Hua Fan and Ka-Sing Lau. Iterated function system and Ruelle operator. *J. Math. Anal. Appl.*, 231(2):319–344, 1999. [42](#)
- [21] R. Fernández and A. Galves. Markov approximations of chains of infinite order. *Bull. Braz. Math. Soc. (N.S.)*, 33(3):295–306, 2002. Fifth Brazilian School in Probability (Ubatuba, 2001). [51](#)
- [22] Fiorenzo Ferrari and Abraham Wyner. Estimation of general stationary processes by variable length Markov chains. *Scand. J. Statist.*, 30(3):459–480, 2003. [50](#)
- [23] Sandro Gallo. Chains with unbounded variable length memory: perfect simulation and a visible regeneration scheme. *Adv. in Appl. Probab.*, 43(3):735–759, 2011. [20](#), [51](#)
- [24] Sandro Gallo and Frédéric Paccaut. On non-regular g -measures. *Nonlinearity*, 26(3):763–776, 2013. [20](#), [51](#)

- [25] Antonio Galves and Eva Löcherbach. Stochastic chains with memory of variable length. *Festschrift in honour of the 75th birthday of Jorma Rissanen*, 2008. [19](#), [50](#)
- [26] T. E. Harris. On chains of infinite order. *Pacific J. Math.*, 5:707–724, 1955. [18](#)
- [27] David Harte. *Multifractals*. Chapman & Hall/CRC, Boca Raton, FL, 2001. Theory and applications. [64](#)
- [28] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Markov chains and invariant probabilities*, volume 211 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003. [60](#)
- [29] Yanick Heurteaux and Andrzej Stos. On measures driven by Markov chains. *J. Stat. Phys.*, 157(6):1046–1061, 2014. [17](#), [47](#)
- [30] John E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981. [5](#), [11](#), [15](#), [23](#), [25](#)
- [31] Tulcea Ionescu and G. Marinescu. Sur certaines chaînes à liaisons complètes. *C. R. Acad. Sci. Paris*, 227:667–669, 1948. [18](#)
- [32] C. T. Ionescu Tulcea. Mesures dans les espaces produits. *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 7:208–211 (1950), 1949. [29](#)
- [33] M. Iosifescu and A. Spătaru. On denumerable chains of infinite order. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:195–214, 1973. [18](#)
- [34] Marius Iosifescu. *Finite Markov processes and their applications*. John Wiley & Sons Ltd., Chichester, 1980. Wiley Series in Probability and Mathematical Statistics. [31](#)
- [35] Joanna Jaroszevska. Iterated function systems with continuous place dependent probabilities. *Univ. Iagel. Acta Math.*, (40):137–146, 2002. [12](#), [13](#)
- [36] Joanna Jaroszevska. A note on iterated function systems with discontinuous probabilities. *Chaos Solitons Fractals*, 49:28–31, 2013. [13](#), [24](#)
- [37] Joanna Jaroszevska and Michał Rams. On the Hausdorff dimension of invariant measures of weakly contracting on average measurable IFS. *J. Stat. Phys.*, 132(5):907–919, 2008. [17](#), [25](#)
- [38] Samuel Karlin. Some random walks arising in learning models. I. *Pacific J. Math.*, 3:725–756, 1953. [13](#)
- [39] Andrzej Lasota and Józef Myjak. Semifractals. *Bull. Polish Acad. Sci. Math.*, 44(1):5–21, 1996.
- [40] Andrzej Lasota and James A. Yorke. Lower bound technique for Markov operators and iterated function systems. *Random Comput. Dynam.*, 2(1):41–77, 1994. [12](#), [13](#), [23](#)

- [41] Florencia Leonardi and Antonio Galves. Sequence motif identification and protein family classification using probabilistic trees. In *Advances in Bioinformatics and Computational Biology*, pages 190–193. Springer, 2005. [10](#), [50](#)
- [42] Bunrith Jacques Lim. *Poisson boundaries of quantum operations and quantum trajectories*. Thèse, Université de Rennes 1, November 2010. <https://tel.archives-ouvertes.fr/tel-00637636>. [11](#)
- [43] Hans Maassen and Burkhard Kümmerner. Purification of quantum trajectories. In *Dynamics & stochastic*, volume 48 of *IMS Lecture Notes Monogr. Ser.*, pages 252–261. Inst. Math. Statist., Beachwood, OH, 2006. [11](#)
- [44] Benoit Mandelbrot. *Les objets fractals*. Flammarion, Editeur, Paris, 1975. Forme, hasard et dimension, Nouvelle Bibliothèque Scientifique. [5](#)
- [45] Józef Myjak and Tomasz Szarek. On Hausdorff dimension of invariant measures arising from non-contractive iterated function systems. *Ann. Mat. Pura Appl. (4)*, 181(2):223–237, 2002. [16](#)
- [46] Jacques Neveu. *Bases mathématiques du calcul des probabilités*. Préface de R. Fortet. Deuxième édition, revue et corrigée. Masson et Cie, Éditeurs, Paris, 1970. [29](#)
- [47] Octav Onicescu and Gheorghe Mihoc. Sur les chaines de variables statistiques. *Bull Sci., Math*, 59:174–92, 1935. [18](#)
- [48] Yakov B. Pesin. *Dimension theory in dynamical systems*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997. Contemporary views and applications. [43](#)
- [49] Dimitri Petritis. On the pertinence to physics of random walks induced by random dynamical systems: a survey. *Journal of Physics: Conference Series*, 738(1):012003, 2016. [6](#)
- [50] Georg Pólya. über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, 84(1-2):149–160, 1921. [6](#)
- [51] Yu. V. Prokhorov. Convergence of random processes and limit theorems in probability theory. 1(2):157–214, 1956. [56](#)
- [52] Jorma Rissanen. A universal data compression system. *IEEE Trans. Inform. Theory*, 29(5):656–664, 1983. [19](#), [50](#)
- [53] Örjan Stenflo. A note on a theorem of Karlin. *Statist. Probab. Lett.*, 54(2):183–187, 2001. [13](#)
- [54] Örjan Stenflo. Uniqueness in g -measures. *Nonlinearity*, 16(2):403–410, 2003. [18](#)
- [55] Tomasz Szarek. The stability of Markov operators on Polish spaces. *Studia Math.*, 143(2):145–152, 2000. [12](#), [23](#), [24](#)