UNIVERSITÉ DE BOURGOGNE DIJON	DÉPARTEMENT DE MATHÉMATIQUES	INSTITUT DE MATHÉMATIQUES DE BOURGOGNE U.M.R 5584

Habilitation à diriger des recherches

Présentée par

VÉRONIQUE MAUME-DESCHAMPS

Un point de vue statistique sur les systèmes dynamiques ; application en analyse d'algorithmes.

Soutenue publiquement le 14 septembre 2005 devant le Jury composé de

Viviane	BALADI
Jean-Pierre	CONZE
Paul	DOUKHAN
Shizan	FANG
Philippe	ROBERT
Bernard	SCHMITT
Brigitte	VALLÉE

RAPPORTEURS :
Jean-Pierre CONZE
Mariusz URBANSKI
Brigitte VALLÉE.

Jean-Pierre Conze est l'un des "pères" de la théorie ergodique en France. Je suis très touchée de l'intérêt qu'il a manifesté pour mes travaux et je le remercie d'avoir accepté la tâche de rapporteur.

Mariusz Urbanski a accepté avec enthousiasme d'être rapporteur, je lui suis très reconnaissante de l'attention qu'il a portée à mon travail.

C'est grâce à Brigitte Vallée que j'ai découvert l'utilisation des systèmes dynamiques en analyse d'algorithmes. Elle m'a communiqué son enthousiasme pour ce domaine de la recherche à l'interface entre mathématique et informatique. J'ai beaucoup profité de son savoir scientifique et apprécié ses qualités humaines. Je la remercie vivement d'avoir établi un rapport sur mon mémoire d'habilitation à diriger des recherches.

Viviane Baladi s'est toujours intéressée à mon travail, je suis heureuse d'avoir pu profiter de son dynamisme scientifique et je la remercie chaleureusement de faire partie de mon jury.

L'intérêt de Paul Doukhan pour mon travail m'a beaucoup touchée. Je lui suis très reconnaissante de faire partie de ce jury.

Je remercie sincèrement Shizan Fang et Philippe Robert de l'honneur qu'ils me font en participant à ce jury.

Depuis que je connais Bernard Schmitt, j'apprécie son ouverture d'esprit et son enthousiasme scientifique. Il a dirigé ma thèse de doctorat puis nous avons continué de fructueux échanges. Je suis heureuse de lui exprimer ma profonde gratitude pour la confiance et le soutien qu'il m'a témoignés et je le remercie pour sa participation à ce jury.

C'est souvent au contact de collaborateurs que la recherche scientifique progresse. Je suis heureuse d'avoir pu nouer de nombreuses relations scientifiques qui m'ont beaucoup apporté.
J'ai eu beaucoup de plaisir à travailler avec mes co-auteurs : V. Baladi, M. Benedicks, C. Bose, X. Bressaud, J. Buzzi, F. Chazal, P. Collet, B. Daireaux, A. Jebrane, A. Kondah, C. Labruère, C. Liverani, S. Martinez, B. Schmitt, S. Shin, M. Urbanski, B. Vallée, A. Zdunick. Je tiens à les remercier chaleureusement.
Je remercie aussi les membres des communautés ALEA et AofA pour leur accueil chaleureux. Les rencontres qu'ils organisent autour de l'analyse d'algorithmes et des structures aléatoires sont toujours très riches scientifiquement.

C'est très appréciable au quotidien de profiter d'un environnement dynamique et chaleureux. Je remercie les membres de l'équipe Applications des Mathématiques de l'IMB de maintenir cette ambiance propice au travail. C'est grâce au dynamisme et à l'ouverture d'esprit de Jean-Marc Gambaudo que l'équipe a pu voir le jour. Merci de nous avoir fait confiance !
Un grand merci au personnel non chercheur de l'IMB et du département de mathématiques pour leur aide quotidienne, leur efficacité et leur gentillesse.

Enfin, mes pensées vont à toute ma famille et mes proches pour leur soutien et leurs encouragements.
Je dédie ce mémoire à Corto et Romanée dont les dessins égaient mon bureau et à Sébastien qui accompagne mes projets de sa compréhension.

Ce texte est une introduction aux travaux suivants :

(1) B. Daireaux, V. Maume-Deschamps, B. Vallée *The Lyapunov Tortoise and the dyadic Hare* 2005 international Conference on Analysis of Algorithms, Discrete Matematics and Theoritical Computer Science, proc. **AD**, (2005), 71–94.

(2) V. Maume-Deschamps *Concentration inequalities and estimation of conditional probabilities* accepté pour publication à Lect. notes in Stat., proceedings of StatDep 2005.

(3) J. Buzzi, V. Maume-Deschamps *Decay of correlations on towers for potentials with summable variation*, Discrete and Continuous Dynamical Systems, 12, (2005), no 4, 639-656.

(4) F. Chazal, V. Maume-Deschamps, *General Markov Dynamical sources : applications to information theory*, Discrete Mathematics and Theoretical Computer Science, 6, (2004), no 2, 283-314.

(5) F. Chazal, V. Maume-Deschamps, B. Vallée *Erratum to "Dynamical sources in Information Theory : Fundamentals Intervals and Word Prefixes" by B. Vallée.* Algorithmica 38 (2004), no. 4, 591-596.

(6) V. Maume-Deschamps, B. Schmitt, M. Urbanski, A. Zdunik *Pressure and recurrence.* Fund. Math. 178 (2003), no. 2, 129-141.

(7) C. Liverani, V. Maume-Deschamps *Lasota-Yorke maps with holes : conditionally invariant probability measures and invariant probability measures on the survivor set*, Ann. IHP, 39, no 3, (2003), 385-412.

(8) C. Bose, V. Maume-Deschamps, B. Schmitt, S. Shin *Invariant measures for piecewise convex transformations of an interval.* Studia Mathematica, 152, (2002), no 3, 263-297.

(9) J. Buzzi, V. Maume-Deschamps *Decay of correlations for piecewise invertible maps in higher dimensions*, Israel J. Math. 131 (2002), 203-220.

(10) V. Baladi, M. Benedicks, V. Maume-Deschamps *Almost sure rates of mixing for i.i.d. unimodal maps* Ann. Sci. École Norm. Sup. (4) 35 (2002), no. 1, 77-126.
V. Baladi, M. Benedicks, V. Maume-Deschamps, *Corrigendum : "Almost sure rates of mixing for i.i.d. unimodal maps" [Ann. Sci. École Norm. Sup. (4) 35 (2002), no. 1, 77–126 ; MR1886006 (2003d :37027)].* Ann. Sci. École Norm. Sup. (4) 36 (2003), no. 2, 319–322.

(11) V. Maume-Deschamps *Projective metrics and mixing properties on towers*, Trans. A.M.S. **353**, (2001), 8, 3371-3389.

(12) V. Maume-Deschamps *Correlation decay for Markov maps on a countable state space.* Erg. Th. and Dyn. Sys., **21**, (2001), 165-196.

(13) P. Collet, S. Martinez, V. Maume-Deschamps *Conditionally Invariant Probability Measures in Dynamical Systems.* Nonlinearity, **13**, (2000), 1263-1274.
P. Collet, S. Martinez, V. Maume-Deschamps *Corrigendum : "On the existence of conditionally invariant probability measures in dynamical systems"* Nonlinearity , **17**, (2004), 1985-1987.

(14) V. Maume-Deschamps *Equilibrium states for non hölderian random dynamical systems.* Random and Computational Dynamics (1997) **5** (4) 319-335.

(15) A. Kondah, V. Maume, B. Schmitt *Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non höldériennes.* Ann. Inst. Poincaré Sec. Prob. Stat. (1997) **33** (6) 675-695.

(16) A. Kondah, V. Maume, B. Schmitt *Dynamique symbolique : vitesse de convergence vers l'état d'équilibre.* C. R. Acad. Sci. Paris (1996), **323**, 393-396.

Table des matières

Introduction

Ce document présente une synthèse de mes travaux sur les propriétés statistiques de processus faiblement dépendants, plus particulièrement des systèmes dynamiques, et de leurs applications. Ces travaux se répartissent selon trois thèmes : mesures invariantes absolument continues et décroissance des corrélations pour les systèmes dynamiques, propriétés statistiques plus fines (mesures conditionnellement invariantes, récurrence, autres théorèmes limites), utilisation des systèmes dynamiques en analyse d'algorithmes.

Plus précisément, étant donnée une application (ou système dynamique discret) $T : M \longrightarrow M$, en général au moins différentiable par morceaux, la théorie ergodique cherche à décrire le comportement statistique asymptotique des orbites : x, $T(x)$, ..., $T^n(x)$, ... pour $x \in M$. Ce comportement est vu au travers d'une *mesure invariante*. On considère alors les orbites comme un processus stationnaire dont on étudie les propriétés asymptotiques : théorème de la limite centrale, propriétés de récurrence, inégalités exponentielles,... De telles propriétés sont satisfaites pour des processus indépendants ; les processus provenant des systèmes dynamiques ne sont pas en général indépendants mais possèdent souvent des comportements asymptotiques proches des processus indépendants.
La première étape permettant d'obtenir ces propriétés asymptotiques est "l'oubli rapide du passé" autrement dit, une décroissance des corrélations rapide (sommable). Les coefficients de décroissance des corrélations sont étroitement liés aux propriétés spectrales de *l'opérateur de transfert*. Dans un premier temps, la démarche est alors la suivante : étant donné un système dynamique discret, on recherche une mesure invariante ayant un sens "physique" : mesure absolument continue ou de "Sinaï-Ruelle-Bowen", on estime la décroissance des corrélations en étudiant, par exemple, l'opérateur de transfert.
On est alors dans le cadre des processus faiblement dépendants. L'étude de leurs propriétés asymptotiques est un vaste champ de recherche actuel. Mes travaux sur le comportement asymptotique des systèmes dynamiques et/ou des processus faiblement dépendants portent sur les propriétés de récurrence et les inégalités de concentration. Ma motivation principale étant d'exhiber des estimateurs paramétriques et non paramétriques pour les systèmes dynamiques : par exemple des estimateurs de la pression, de la fonction potentiel,... dans le but d'utiliser les systèmes dynamiques en modélisation et en simulation.
Plus récemment, les systèmes dynamiques sont apparus en théorie de l'information et en analyse d'algorithmes. Dans le cadre de la théorie de l'information, les systèmes dynamiques permettent de considérer des sources plus générales que celles habituellement étudiées (indépendantes

ou markoviennes) ; on cherche alors le comportement asymptotique de quantités comme le nombre $B(\rho)$ de mots de masse supérieure à ρ, $\rho \sim 0$ ... En analyse d'algorithmes, on souhaite évaluer, par exemple, le nombre moyen d'itérations, la distribution asymptotique du nombre d'itérations,... On peut considérer l'algorithme et l'ensemble de ses données comme un système dynamique sur un espace discret que l'on plonge dans un système dynamique sur un espace continu. Dans les deux cadres, on relie les quantités étudiées (ou leurs séries génératrices) à l'opérateur de transfert. Le comportement asymptotique est alors guidé par les propriétés spectrales de l'opérateur de transfert. Ces techniques sont particulièrement efficaces pour étudier les algorithmes arithmétiques.

Dans un premier temps, nous donnerons quelques résultats d'estimation de décroissance des corrélations pour des systèmes dynamiques uniformément dilatants en dimension 1 ou symboliques, vérifiant diverses hypothèses de régularité. Nous montrerons ensuite que les "tours de Young" constituent un outil puissant pour obtenir de telles estimations pour des systèmes non uniformément dilatants et/ou en dimension plus grande.
Dans une deuxième partie, nous étudierons des propriétés statistiques plus fines : récurrence et inégalités de concentration. Le but ultime de cette étude est d'obtenir suffisamment d'informations quantitatives sur les systèmes dynamiques pour les utiliser en modélisation et en simulation.
Enfin, nous expliciterons les problématiques de la théorie de l'information et de l'analyse d'algorithmes et verrons comment les systèmes dynamiques et les produits aléatoires de matrices permettent d'obtenir des propriétés statistiques fines des mots générés par des sources dynamiques et de certains algorithmes arithmétiques.

Avant de détailler ces résultats et les techniques utilisées, rappelons quelques définitions et résultats élémentaires sur les systèmes dynamiques discrets.

Quelques définitions

1. Systèmes dynamiques inversibles par morceaux

Nous allons considérer des systèmes dynamiques sur M, un espace métrique, par exemple un espace symbolique $M \subset A^{\mathbb{N}}$, A au plus dénombrable ou M un ouvert de $\mathbb{R}^n$ ou encore M une variété riemannienne. Dans le cadre différentiable, on supposera que $T : M \to M$ est différentiable par morceaux. C'est-à-dire qu'il existe des ouverts U_i disjoints tels que la réunion des $\overline{U_i}$ soit M (on dira que les U_i forment une *partition topologique de* M) et $T : U_i \to T(U_i)$ est un C^2 difféomorphisme, continu sur $\overline{U_i}$.
Dans le cadre symbolique, T est le décalage : $T(x_0, \ldots, x_n, \ldots) = (x_1, \ldots, x_{n-1}, \ldots)$. Pour $0 < \theta < 1$, on définit une distance sur $A^{\mathbb{N}}$ par $d(x, y) = \theta^n$ si $x_i = y_i$ pour $i = 0, \ldots, n-1$ et $x_n \neq y_n$.
Nous considérerons que T est *inversible par morceaux* : il existe une partition topologique $\mathcal{P}$, au plus dénombrable, telle que la restriction à chaque élément P de $\mathcal{P}$ est inversible de P dans $T(P)$. Une telle partition est *génératrice* si $\bigvee_{i=0}^{\infty} T^{-i}\mathcal{P}$ est la partition en points. Cette définition de partition génératrice est équivalente à : $\mathrm{diam}(C) \overset{n\to\infty}{\longrightarrow} 0$, pour $C \in \mathcal{P}^n$ où $\mathcal{P}^n$ est la partition d'inversibilité de T^n ; les éléments de $\mathcal{P}^n$ s'écrivent $P_0 \cap T^{-1}P_1 \cap \cdots T^{-n-1}P_{n-1}$, $P_i \in \mathcal{P}$.
L'application est dite *non singulière* par rapport à une mesure m si la mesure image $T_\star m$ de m par T est absolument continue par rapport à m.
Une mesure de probabilité μ sur M est *invariante* par T si pour tout Borélien A, $\mu(T^{-1}(A)) = \mu(A)$, une mesure de probabilité μ invariante est dite *ergodique* si $T^{-1}A = A$ implique $\mu(A) = 0$ ou $\mu(A) = 1$. Toutes les mesures invariantes ne sont pas pertinentes : par exemple une mesure dont le support est une orbite périodique finie ne donnera pas beaucoup d'information. Nous nous intéresserons aux mesures ayant un sens "physique" : dans le cadre géométrique, on peut considérer des mesures de "Sinaï-Ruelle-Bowen". Une mesure de probabilité invariante est dite *mesure de Sinaï-Ruelle-Bowen (mesure SRB)* s'il existe un borélien A de mesure de Lebesgue positive, tel que pour $x \in A$, pour toute fonction f continue,

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \longrightarrow \int f d\mu.$$

En particulier, si μ est ergodique et absolument continue par rapport à la mesure de Lebesgue, le théorème ergodique de Birkoff montre que μ est une mesure SRB. Le plus souvent dans ce travail, nous considérerons des mesures invariantes absolument continues par rapport à la mesure

de Lebesgue.
Dans le cadre symbolique, nous considérerons des mesures de probabilité absolument continues par rapport à une mesure de probabilité ν dite *conforme* pour une fonction continue g appelée potentiel. Pour $c \in \mathbb{R}_\star^+$, une mesure de probabilité ν est c-conforme pour g si

$$\frac{d\mu \circ T}{d\mu} = (cg)^{-1}.$$

Les réels $c > 0$ pour lesquels une mesure conforme existe sont intimement liés au formalisme thermodynamique. Rappelons que la *pression topologique* [DenGS] d'un sous-ensemble S de M est

$$P(S,T) = \limsup_{n\to\infty} \frac{1}{n} \log \sum_{\substack{A\in\mathcal{P}^n \\ \overline{A}\cap S\neq\varnothing}} g^{(n)}(A),$$

avec $g^{(n)}(A) = \sup_{x\in A} g(x)g(Tx)\dots g(T^{n-1}x)$, on notera $P_{top}(T,g) = P(M,T)$ la pression topologique de tout le système. Il est facile de voir que $\log c \le P_{top}(M,T)$. Dans les cas que nous considérerons, on a en fait l'égalité $\log c = P_{top}(T,g) = P(M,T)$ (voir [BuPaS]). Si $g \equiv 1$, la pression topologique $P(M,g)$ est égale à l'entropie topologique $h_{top}(T)$.
On voit facilement que la mesure de Lebesgue est 1-conforme pour $g = JT^{-1}$ le déterminant de la différentielle de T^{-1}.
Enfin, un système dynamique inversible par morceaux est dit *markovien* si pour i, $T(U_i)$ est une union de U_j. Les systèmes dynamiques markoviens sont souvent beaucoup plus faciles á étudier, en particulier parce que l'opérateur de transfert agit sur des espaces de fonctions au moins continues par morceaux (et souvent C^1 par morceaux). Une stratégie (voir [Ho]) pour étudier des systèmes dynamiques non markoviens est de se ramener à un système markovien par une "extension markovienne".

2. Opérateur de transfert et propriétés ergodiques

L'outil essentiel que nous allons utiliser pour étudier le comportement asymptotique des systèmes dynamiques inversibles par morceaux est l'opérateur de transfert ou opérateur de Perron-Frobenius. C'est le dual - pour une mesure conforme - de l'opérateur $f \mapsto f \circ T$. Dans le cadre géométrique, nous considérerons pour ν la mesure de Lebesgue, dans le cadre symbolique, nous supposerons qu'une mesure 1-conforme pour un potentiel g existe (voir [BuPaS] pour des critères d'existence d'une telle mesure).
Plus précisément, pour $\varphi \in L^1(\nu)$, $\psi \in L^\infty(\nu)$ l'opérateur de transfert $\mathcal{L}$ est défini par :

$$\int (\varphi \circ T)\psi d\nu = \int \varphi \mathcal{L}(\psi) d\nu.$$

On a aussi l'expression suivante pour $\mathcal{L}$:

$$\mathcal{L}(\psi)(x) = \sum_{T(y)=x} g(y)\psi(y),$$

cette expression permet d'étendre l'opérateur $\mathcal{L}$ à $L^1(\nu)$. Rappelons encore qu'une mesure de probabilité invariante μ est mélangeante si et seulement si pour toutes fonctions φ et ψ dans $L^2(\mu)$,

$$\int (\varphi \circ T^n)\psi d\mu \longrightarrow \int \varphi d\mu \int \psi d\mu.$$

Une mesure mélangeante est ergodique.
Les coefficients de décroissance des corrélations sont les covariances entre $\varphi \circ T^n$ et ψ :

$$C_n(\varphi, \psi) = \left| \int (\varphi \circ T^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \right|.$$

Les propriétés ergodiques d'une mesure μ invariante, absolument continue par rapport à ν sont fortement reliées à l'opérateur de transfert :

- $\mu = h\nu$ est invariante si et seulement si $\mathcal{L}(h) = h$,
- si $\mu = h\nu$ est invariante, elle est ergodique si et seulement si 1 est simple comme valeur propre de $\mathcal{L}$,
- si $\mu = h\nu$ est invariante et mélangeante alors 1 est la seule valeur propre de module 1 de $\mathcal{L}$ (la réciproque est vraie sous certaines hypothèses supplémentaires).
- les coefficients de mélange sont liés au comportement asymptotique de l'opérateur de transfert par :

$$C_n(\varphi, \psi) = \left| \int \varphi[\mathcal{L}^n(\psi h) - \psi h] d\nu \right|.$$

3. Différents types de mélange

Si on considère $X_i = T^i$ comme des variables aléatoires de loi μ, on voit le mélange comme de l'indépendance asymptotique entre la tribu engendrée par les X_{i+n}, ..., $X_{i+n+\ell}$ et la tribu engendrée par les X_0, ..., X_i. On s'attend à ce que certaines propriétés asymptotiques des processus indépendants restent vraies pour des systèmes dynamiques mélangeants. Sous quelques conditions techniques supplémentaires, C. Liverani ([Li2]) montre que si $C_n(f, f)$ est sommable alors f vérifie un théorème de la limite centrale dans le sens suivant : il existe $\sigma \geq 0$ tel que

$$\frac{1}{\sqrt{n}} \left[S_n f - n \int_X f d\mu \right] \xrightarrow{\text{loi}} \mathcal{N}(0, \sigma),$$

où $\mathcal{N}(0,\sigma)$ désigne la loi normale centrée d'écart-type σ et
$S_nf = \frac{1}{n}\sum_{i=0}^{n-1} f\circ T^i$.
Ce résultat étend ceux de Gordin et Ibragimov-Linnik ([Gor, IL]) pour des processus Φ-mélangeants.
Si on ne fait aucune hypothèse de régularité sur φ et ψ, pour un système mélangeant, la décroissance vers 0 des coefficients $C_n(\varphi,\psi)$ peut être arbitrairement lente. Pour obtenir d'autres résultats asymptotiques, on a besoin d'une certaine uniformité en φ et ψ. En probabilité, on fait souvent l'hypothèse de Φ-mélange : un processus stationnaire X_0, ..., X_n, ... de loi μ est Φ-mélangeant s'il existe une suite $(\Phi(n))_{n\in\mathbb{N}}$ décroissant vers 0 telle que, pour tout i,j,n, $A\in\mathcal{M}_i^j$ et $B\in\mathcal{M}_{j+n}^\infty$,

$$|\mu(A\cap B)-\mu(A)\mu(B)|\leq \Phi(n)\mu(A)\mu(B),$$

avec $\mathcal{M}_i^j$ la tribu engendrée par les X_i, ..., X_j.
Même des systèmes dynamiques très simples comme $T(x)=2x \bmod 1$ sur $I=[0,1]$ ne satisfont pas le Φ-mélange (voir par exemple [DedPr]). Par contre, en restreignant les coefficients de mélange à certaines classes de fonctions, on peut obtenir, pour les systèmes dynamiques, un contrôle uniforme de la décroissance des corrélations. Étant donné un espace de Banach $\mathcal{C}$ de fonctions bornées sur M, si la norme sur $\mathcal{C}$ se décompose en : $\|f\|_{\mathcal{C}} = C(f)+\|f\|$ avec C une semi-norme et $\|\ \|$ une norme sur $\mathcal{C}$, dans [2], nous définissons le $\Phi_{\mathcal{C}}$-mélange pour un processus X_1, ..., X_n, ..., inspiré des coefficients de mélange de Dedeker-Prieur ([DedPr]). Soit :

$$\begin{aligned}\Phi_{\mathcal{C}}(n) = \sup\{&|\mathbb{E}(Yf(X_{i+n}))-\mathbb{E}(Y)\mathbb{E}(f(X_{i+n}))|\ i\in\mathbb{N},\ Y \text{ est}\\ &\mathcal{M}_i-\text{mesurable et } \|Y\|_1\leq 1,\ f\in\mathcal{C}_1\},\end{aligned} \tag{*}$$

où $\mathcal{M}_i$ est la tribu engendrée par les X_0, ..., X_i. Nous dirons que le processus est $\Phi_{\mathcal{C}}$ mélangeant si $(\Phi_{\mathcal{C}}(n))_{n\in\mathbb{N}}$ tend vers 0. Nous dirons que le processus est $\Phi_{\mathcal{C}}$-faiblement dépendant si $(\Phi_{\mathcal{C}}(n))_{n\in\mathbb{N}}$ est sommable. Nous verrons plus en détail ces coefficients dans la section II.2.2. Notons dès à présent que si $X_i=T^i$, le processus n'est pas $\Phi_{\mathcal{C}}$-faiblement dépendant mais on peut lui associer un processus "à temps inversé" qui est $\Phi_{\mathcal{C}}$-faiblement dépendant sous de bonnes hypothèses pour T.
Les exemples d'espaces $\mathcal{C}$ que nous considérerons sont : l'espace des fonctions à variation bornée, l'espace des fonctions Lipschitziennes, l'espace des fonctions C^1, l'espace des fonctions Hölder, ...
Si M est un espace métrique, pour $0<\alpha\leq 1$, une fonction f sur M est *α-Hölder* s'il existe $K>0$ tel que pour tout $x,y\in M$,

$$|f(x)-f(y)|\leq Kd(x,y)^\alpha,$$

l'infimum $K(f)$ des $K>0$ tel que l'inégalité ci-dessus est satisfaite est la constante de Hölder de f, si $\alpha=1$, on dit que f est *Lipschitz* et on

parle alors de constante de Lipschitz.
Si M est ordonné, une fonction f sur M est *à variation bornée* si

$$\sup\left\{\sum_{i=0}^{p-1}|f(a_i)-f(a_{i+1})|\right\} := \bigvee f < \infty,$$

le supremum ci-dessus étant pris sur les familles finies de $a_i \in M$, tels que $a_0 < a_1 < \cdots < a_p$. La quantité $\bigvee f$ s'appelle la variation totale de f. Les fonctions à variation bornée sont bien adaptées à la dimension 1 ou aux espaces symboliques mais sont beaucoup plus difficiles à manier en dimension supérieure.

Première partie I. Vitesse de mélange pour les systèmes dynamiques

Pour un système dynamique qui admet une mesure invariante ayant un sens physique (par exemple une mesure absolument continue ou une mesure SRB), le contrôle de la vitesse de décroissance des corrélations (ou vitesse de mélange) permet d'obtenir des propriétés statistiques du système, par exemple un théorème de la limite centrale. Ma thèse portait sur l'estimation de la vitesse de mélange pour des systèmes dynamiques markoviens ([16, 15, 14, 12]) avec une régularité éventuellement non höldérienne et éventuellement une partition infinie dénombrable.
À la suite de ces travaux de thèse j'ai considéré des systèmes dynamiques non markoviens et avec différentes hypothèses de régularité. Dans [8] nous avons montré que des hypothèses de convexité suffisent pour obtenir une mesure invariante absolument continue et un contrôle de la décroissance des corrélations.
Une autre technique, lorsque le système n'est pas markovien, est de construire une extension qui elle est markovienne et qui possède les mêmes propriétés statistiques que le système de départ ; on construit ainsi un modèle de système dynamique permettant de traiter de nombreux cas. Ce point de vue a été développé par L.S. Young ([Y1, Y2]) : on appellera *tours de Young* les extensions markoviennes ainsi construites. On cherche donc, d'une part à obtenir des propriétés statistiques pour les tours de Young ([3, 11]), d'autre part à conjuguer les systèmes à des tours. Dans [9] nous obtenons ainsi un contrôle de la vitesse de mélange pour des applications dilatantes en dimension plus grande que 1. Le point de vue des tours de Young permet aussi de traiter le cas de compositions aléatoires de systèmes dynamiques et plus particulièrement d'applications unimodales ([10]).

1. Systèmes symboliques et systèmes dynamiques dilatants en dimension 1

Le cas des systèmes symboliques sur un alphabet fini, associé à un potentiel Höldérien a été largement étudié notamment par Ruelle ([Ru]) dans le cadre du formalisme thermodynamique et de la mécanique statistique et par Bowen ([Bow]) en liaison avec l'étude des propriétés ergodiques de difféomorphismes de type Axiom A. Les applications C^2 uniformément dilatantes (en dimension quelconque) ou C^2 par morceaux (sur un nombre fini de morceaux) et markoviennes peuvent être codées sur un tel système symbolique. Les deux sections suivantes présentent une extension des résultats de Ruelle et Bowen à des potentiels non Höldériens (pour un système géométrique, cela correspond à une application dérivable dont la dérivée n'est pas Hölder) ou à des alphabets infinis.

Le cas des systèmes dilatants par morceaux, non markoviens, en dimension 1 a été aussi largement étudié (notamment par Lasota-Yorke ([LaY1])) dans le cas d'applications suffisamment régulières (essentiellement à dérivée à variation bornée). Dans la section 1.3 nous montrons que des hypothèses de convexité sont suffisantes pour obtenir une décroissance exponentielle des corrélations.

Pour une fonction potentiel $g : M \to \mathbb{R}^+$ nous contrôlerons son module de continuité par la suite :

$$w_n = \sup_{C \in \mathcal{P}^n} \sup_{x,y \in C} \log \frac{g(x)}{g(y)}, \tag{1.1}$$

cette suite est le module de continuité de $\log g$.
Les systèmes dynamiques symboliques que nous considérons sont des *sous décalages de type finis*, c'est à dire qu'il existe B une $A \times A$ matrice de 0 et 1 telle que

$$M = \{x \in A^{\mathbb{N}} \ / \ B_{x_i, x_{i+1}} = 1 \text{ pour tout } i \in \mathbb{N}\},$$

autrement dit la matrice B décrit les suites admissibles. Le décalage σ agit sur M, est inversible par morceaux pour la partition

$$\mathcal{P} = \{[a] \ a \in A\}, \quad [a] = \{x \in M \ / \ x_0 = a\}.$$

Cette partition est clairement markovienne et génératrice. Les éléments de $\mathcal{P}^n$ décrivent les suites de n-uplet admissibles et sont appelés *n-cylindres.*
Si la matrice B est *irréductible* - c'est à dire que pour tout couple $(i, j) \in A^2$ il existe $n \in \mathbb{N}$ tel que $(B^n)_{i,j} > 0$ - on dit que le sous décalage est irréductible. Cela exprime que pour tout couple $(i, j) \in A^2$, le système peut passer de i à j.
Si la matrice B est *apériodique* - c'est à dire que pour tout couple $(i, j) \in A^2$ il existe $N \in \mathbb{N}$ tel que pour $n > N$, $(B^n)_{i,j} > 0$ - on dit que le sous décalage est apériodique.

1.1. **Systèmes symboliques sur un alphabet fini associé à un potentiel non Höldérien.** Nous supposerons que la suite w_n est sommable ; lorsqu'elle est géométrique, on dit que g est Höldérienne. Remarquons que dans ce cas, il existe $0 < \theta_0 < 1$ tel que g est Hölder pour la métrique d_θ quelque soit $\theta_0 < \theta < 1$; de plus, si le sous décalage de type fini provient du codage d'une application markovienne dilatante par morceaux de l'intervalle (ou d'un difféomorphisme de type Axiom A), si g est l'inverse de $|T'|$, alors $|T'|$ est Hölder (pour la métrique usuelle) si et seulement si la suite w_n associée à g est géométrique.
On considère alors une métrique sur M qui rend g Lipschitz :

$$d(x, y) = \sum_{j > s(x,y)} w_j,$$

où $s(x,y)$ est le temps de séparation entre x et y c'est à dire le plus grand entier n tel que pour $0 \leq p < n$, $\sigma^p(x)$ et $\sigma^p(y)$ appartiennent au même élément de la partition $\mathcal{P}$. Soit L est l'espace des fonctions sur M Lipschitz pour cette distance d, on considère la norme $\|f\| = \max(\|f\|_\infty, L(f))$. Si $w_n = O(\theta^n)$ alors la métrique d est équivalente à la métrique d_θ.
Soit $c = \exp(P_{top}(\sigma, g))$.
La première partie du résultat suivant (existence et unicité d'une mesure conforme et d'une mesure invariante) a été montrée par P. Walters), l'estimation de la décroissance des corrélations a été obtenue en collaboration avec A. Kondah et B. Schmitt.

Théorème I.1. *(*[W],[15],[21]*) Il existe une unique mesure de probabilité ν conforme pour le potentiel g, il existe une unique mesure de probabilité μ invariante par σ et absolument continue par rapport à ν, sa densité h appartient à L. De plus, il existe $0 < \delta < 1$, $C > 0$ et $(\ell(n))_{n\in\mathbb{N}}$ une suite d'entiers tendant vers l'infini tels que, pour $f \in L$,*

$$\left\| \frac{\mathcal{L}^n f}{c^n} - h\nu(f) \right\|_\infty \leq C\ \delta^{\ell(n)}\ \|f\|. \tag{1.2}$$

Enfin,
- *si $w_n = O(n^{-\alpha-1})$, $\alpha > 0$ alors $\delta^{\ell(n)} = O(n^{-\alpha})$,*
- *si $w_n = O(\theta^n)$, $0 < \theta < 1$ alors il existe $0 < \gamma < 1$ tel que $\delta^{\ell(n)} = O(\gamma^n)$,*
- *si $w_n = O(\theta^{(\log n)^\alpha})$, $0 < \theta < 1$, $\alpha > 1$ alors, pour tout $\varepsilon > 0$, $\delta^{\ell(n)} = O(\theta^{(\log n)^{\alpha-\varepsilon}})$.*

Ce résultat implique que le processus à temps inversé $(Y_n)_{n\in\mathbb{N}}$ défini par $(Y_0, \ldots, Y_n) \stackrel{\text{Loi}}{=} (X_n, \ldots, X_0)$ est Φ_L mélangeant.

Corollaire I.2. *(voir* [BGR, DedPr, 2] *pour plus de détails)*
Il existe $C > 0$ tel que pour $\psi \in L$ et $\varphi \in L^1(\nu)$, pour tout $n \in \mathbb{N}$,

$$C_n(\varphi, \psi) \leq C\delta^{\ell(n)} L(\varphi)\|\psi\|_{L^1}.$$

Si $w_n = O(\frac{1}{n^\alpha})$, $\alpha > 1$, le processus à temps inversé $(Y_n)_{n\in\mathbb{N}}$ est Φ_L-mélangeant, il est Φ_L-faiblement dépendant si $w_n = O(\frac{1}{n^\alpha})$, $\alpha > 2$.

Remarque. Stratégie de la preuve
Dans la cas où la suite w_n est géométrique, l'estimation de la vitesse de convergence des itérés de l'opérateur de transfert vers la projection sur l'espace propre engendré par h peut être obtenue en montrant que $\mathcal{L}$ est quasi-compact. Dans le cas où la suite n'est pas géométrique, l'opérateur n'est plus quasi-compact (on montre [15] que tous les points du disque $D(0,c)$ sont des valeurs propres de multiplicité infinie). La stratégie de la preuve du Théorème I.1 consiste à exhiber une suite de cônes C_k de L et une suite d'entiers n_k tels que $\mathcal{L}^{n_k}(C_{k-1}) \subset C_k$ en contrôlant le diamètre hyperbolique Δ_k de $\mathcal{L}^{n_k}(C_{k-1})$ dans C_k. La

théorie de Birkhoff ([Bi1, Bi2]) permet alors de montrer que - projectivement - les itérés $\mathcal{L}^{n_1+\cdots+n_k}$ convergent vers un opérateur de rang 1 avec vitesse : $\prod_{i=1}^{k}\tanh(\Delta_i)$. Nous avons adapté la stratégie mise au point par Ferrero et Schmitt ([FeSc]) pour obtenir des décroissances exponentielles des corrélations de certaines compositions aléatoires d'opérateurs à un cadre "non quasi-compact".

La stratégie de preuve décrite ci-dessus peut être "aléatoirisé" pour considérer des compositions aléatoires d'opérateurs. Soit $(\Omega,\mathcal{F},\mathbb{P},S)$ un système dynamique inversible et ergodique. Soit $G\ :\Omega\to\ C(M)^+$, $\omega\mapsto g_\omega$ une application mesurable et $(w_n)_{n\in\mathbb{N}}$ une suite positive, sommable. On note $w_n(\omega)$ le module de continuité de $\log g_\omega$ défini comme dans (1.1). On considère $\mathcal{L}_\omega$ l'opérateur de transfert associé au potentiel g_ω qui agit sur les fonctions continues sur M :

$$\mathcal{L}_\omega(f)(x)=\sum_{\sigma(y)=x} g_\omega(y)f(y).$$

Théorème I.3. [14, 21] *Si l'application G vérifie :*
– $\int \|\log g_\omega\|_\infty dP<\infty$,
– $\sup_{p\in\mathbb{N}}\dfrac{\sum_{i=p+1}^{\infty} w_i(S^{-i}\omega)}{w_p}=C(\omega)$, $C(\omega)\in L^1(\Omega)$.
alors, il existe trois applications mesurables :

$$H:\ \begin{array}{l}\Omega\longrightarrow L\\ \omega\longrightarrow h_\omega\end{array}\quad C:\ \begin{array}{l}\Omega\longrightarrow\mathbb{R}^+/\{0\}\\ \omega\longrightarrow c_\omega\end{array}\quad \bar\nu:\ \begin{array}{l}\Omega\longrightarrow\mathcal{P}(X)\\ \omega\longrightarrow \nu_\omega\end{array}$$

telles que :
(1) $\forall\omega\in\Omega$, $\mathcal{L}_\omega h_\omega=c_\omega h_{S\omega}$ *et* $\mathcal{L}^*_\omega\nu_{S\omega}=c_\omega\nu_\omega$, $\nu_\omega(h_\omega)=1$,
(2) $\log c_\omega$ *et* $\|\log h_\omega\|$ *sont intégrables,*

Le triplet $(H,\bar\nu,C)$ est uniquement déterminé par (1). *De plus, il existe $0<\gamma<1$ et $\ell(n)$ une suite strictement croissante d'entiers qui ne dépend que de la suite w_n tels que pour $\psi\in L$ et φ telle que*

$$\sup_{s\in\Omega}(|\int\varphi d\nu_s|)<\infty,$$

alors

$$C_n^\omega(\varphi,\psi)\le Cte\ \gamma^{\ell(n)}\ \sup_{s\in\Omega}(|\int\varphi d\nu_s|)\|\psi\|\ \text{et}$$
$$C_n^{S^{-n}\omega}(\varphi,\psi)\le Cte\ \gamma^{\ell(n)}\ \sup_{s\in\Omega}(|\int\varphi d\nu_s|)\|\psi\|,$$

$C_n^\omega(\varphi,\psi)$ et $C_n^{S^{-n}\omega}(\varphi,\psi)$ sont les corrélations aléatoires "futures" et "passée" :

$$C_n^\omega(\varphi,\psi)=\left|\int\psi(\varphi\circ\sigma^n)h_\omega d\nu_\omega-\int\psi h_{S^n\omega}d\nu_{S^n\omega}\int\varphi h_\omega d\nu_\omega\right|.$$

Enfin, si g est un potentiel vérifiant les hypothèses du Théorème I.1 et si les g_ω sont des perturbés de g tels que :

$$\|\log g_\omega - \log g\|_\infty < \varepsilon \ \forall \omega \in \Omega,$$

et pour tout $\omega \in \Omega$ et $n \in \mathbb{N}$,

$$\widetilde{w}_n(\omega) \leq \varepsilon w_n \text{ avec } \widetilde{w}_n(\omega) \text{ le module de continuité de } \log g - \log g_\omega,$$

alors on a les résultats de stabilité forte suivants :

$$\lim_{\varepsilon \to 0} \sup_{\omega \in \Omega} \|h - h_\omega\|_\infty = 0,$$

$$\lim_{\varepsilon \to 0} \sup_{\omega \in \Omega} \frac{c}{c_\omega} = 1,$$

$$\forall f \in C(M), \ \lim_{\varepsilon \to 0} \sup_{\omega \in \Omega} |\nu(f) - \nu_\omega(f)| = 0.$$

Ce résultat s'applique en particulier à des compositions aléatoires de petites perturbations d'un difféomorphisme Axiom A.

1.2. **Systèmes symboliques sur un alphabet infini.** Lorsque l'alphabet A n'est plus fini, le sous décalage M n'est plus compact, ni même en général localement compact. Dans ([12]), nous avons considéré un sous décalage de type fini sur un alphabet dénombrable sur lequel il existe une mesure 1-conforme m de potentiel g Lipschitz et borné. Sous des hypothèses de "faible contribution de l'infini" qui expriment essentiellement que la contribution à l'opérateur de transfert des cylindres $[k] = \{x \in \Sigma \ / \ x_0 = k\}$ avec k dans le complémentaire d'une partie finie de A, est plus petite que 1, nous obtenons l'existence d'une mesure invariante absolument continue par rapport à m, elle est unique si le sous décalage est irréductible. Si, de plus, le sous décalage est apériodique et que le contrôle de la contribution des cylindres hors d'une partie finie de A est plus précis, nous obtenons un contrôle de la vitesse de convergence des itérés de l'opérateur de transfert : il existe $k \in \mathbb{N}^\star$ et une suite $(\alpha_j(N))_{j \in \mathbb{N}}$, $\alpha_j(N) \to 0$ tels que $\forall f \in L$, $\forall j \in \mathbb{N}$,

$$\|\mathcal{L}^{kj} f - h \int f\|_N \leq \alpha_j(N) \ \|f\| + m([\![0, N]\!]^c) \sup f,$$

où $[\![0, N]\!]$ désigne l'ensemble des x de M tels que $x_0 \leq N$ et $\| \ \|_N$ la norme uniforme sur $[\![0, N]\!]$. La suite $(\alpha_j)_{j \in \mathbb{N}}$ peut être déterminée explicitement et dépend de la contribution à l'opérateur $\mathcal{L}^k \mathbf{1}$ du complémentaire d'un nombre fini de cylindres.
En choisissant convenablement N par rapport j, on obtient ainsi une estimation de la vitesse de convergence sur chaque compact de M et de la décroissance des corrélations : pour $\varphi \in L^\infty$ et $\psi \in L$, pour tout $n \in \mathbb{N}$,

$$C_n(\varphi, \psi) \leq \text{Const} \widetilde{\alpha}_n \|\varphi\|_\infty \|\psi\|,$$

où $\widetilde{\alpha}_n$ est une suite tendant vers 0 qui s'exprime en fonction de $\alpha_{j(n)}(n)$ et $m([\![0, n]\!]^c)$.

Des exemples de systèmes vérifiant nos conditions sont, par exemple, des applications a “petites branches” ne vérifiant pas les hypothèses de [Bre], [LiSV] ou [MU]. De telles applications sont un modèle dynamique de chaînes de markov de type “naissance et mort”. On estime aussi la décroissance des corrélations pour des applications non uniformément dilatantes de l'intervalle, de type Gaspard-Wang ([GasW]) sur un espace d'observables contenant les fonctions lipschtziennes.

Les idées développées dans les deux sections précédentes ont été appliquée ensuite pour étudier des systèmes dynamiques type “tours de Young”, c'est l'objet du paragraphe 2. Avant de considérer ces modèles, nous donnons une dernière généralisation de résultats “classiques”. Il s'agit de systèmes dynamiques en dimension 1 vérifiant une propriété de convexité.

1.3. **Systèmes dynamiques convexes par morceaux.** Les applications inversibles par morceaux de l'intervalle on été largement étudiés depuis l'article de Lasota-Yorke ([LaY1]). Les hypothèses classiques de travail dans ce cadre sont que T est inversible sur un nombre fini de morceaux, C^2 par morceaux et uniformément dilatante. Dans ce cas, il existe des mesures invariantes absolument continues par rapport à la mesure de Lebesgue, celles qui sont ergodiques sont, à une période près, mélangeantes et le mélange pour des fonctions à variation bornée est exponentiel. De nombreuses généralisations ont été envisagées : partition d'inversibilité dénombrable, mesure conforme pour d'autres potentiels que $\frac{1}{|T'|}$ ([BaKe, Broi, LiSV]). On reste néanmoins dans un cadre uniformément dilatant avec un potentiel g au moins à variation bornée. L'idée de Lasota-Yorke était d'utiliser les systèmes inversibles par morceaux dans la modélisation du forage pétrolier, les hypothèses “classiques” peuvent s'avérer difficiles à vérifier dans ce cadre. Dans [LaY2], ils ont montré que des hypothèses de convexité sur T, associées à des conditions géométriques suffisent pour obtenir les propriétés de mélange. En collaboration avec C. Bose, B. Schmitt et S. Shin, nous avons donné des conditions de convexité à la fois moins fortes que celles de Lasota-Yorke et plus naturelles qui demandent moins de conditions géométriques, pas de dilatation uniforme et garantissent les propriétés de mélange.

Nous considérons une application T de l'intervalle $[0,1]$, non singulière par rapport à la mesure de Lebesgue, inversible et continue par morceaux (sur un nombre fini de morceaux) : P_1, ..., P_n. Sur chaque P_i, $T_i = T_{|P_i}$ est supposée croissante, on note F_i une version continue à droite de $\dfrac{dm \circ T_i^{-1}}{dm}$, on la prolonge par 0 hors de $T(P_i)$ afin de la considérer comme une fonction de $[0,1]$. Nos hypothèses sur T sont les suivantes :

Condition C

– pour chaque $k = 1, \ldots, n$, la fonction $F_1 + \cdots + F_n$ est décroissante,

– $F_1(0) < 1$.

Dans [LaY2], Lasota-Yorke supposent que chaque F_i est décroissante et que pour tout $P_i = [a_{i-1}, a_i]$, $T(a_i) = 0$. Cette dernière hypothèse implique en particulier que les branches T_i sont croissantes mais est clairement plus forte. L'hypothèse $F_1(0) < 1$ implique que la première branche de T doit être dilatante mais aucune hypothèse de dilatation n'est faite sur les autres branches.

Remarquons que le fait que T soit croissante par morceaux et que F_1 soit décroissante implique que $F_1(0) > 0$ (sinon F_1 est identiquement nulle est donc la branche T_1 n'existe pas). Ainsi, $T(0) = 0$ car si $0 \notin T(P_1)$ alors on a convenu que $F_1(0) = 0$.

Théorème I.4. [8] *Soit T une application croissante et continue par morceaux, satisfaisant la condition de convexité* **C** *ci-dessus. Alors T admet une mesure invariante μ absolument continue par rapport à la mesure de Lebesgue, de densité g, décroissante. De plus, il existe $\beta \in [0, 1]$ tel que $\cup_{k=0}^{\infty} T^k(P_1) = [0, \beta]$. La restriction μ_β de μ à $[0, \beta]$ est mélangeante (en fait elle est exacte) et si $T'(\beta) > 1$ alors la décroissance des corrélations est exponentielle pour des fonctions de $[0, \beta]$ à variation bornée : il existe $C > 0$ et $0 < \kappa < 1$ tels que pour $\psi \in BV([0, \beta])$, pour $\varphi \in L^1(\mu_\beta)$,*

$$C_n(\varphi, \psi) \leq C\kappa^n \|\varphi\|_{L^1} \|\psi\|_{BV}.$$

La preuve consiste, dans un premier temps à montrer que l'ensemble $\mathcal{J}$ des fonctions décroissantes est invariant par l'opérateur de transfert (la condition **C** est en fait nécessaire et suffisante pour obtenir cette invariance), on montre ensuite une inégalité de "Lasota-Yorke" pour les fonctions de $\mathcal{J}$ qui permet d'obtenir l'existence d'une mesure invariante absolument continue. L'exactitude de μ_β et l'estimation de la décroissance des corrélations se fait "à la main" en adaptant la preuve de Bowen ([Bow]).

Corollaire I.5. *Il existe $C > 0$ et $0 < \kappa < 1$ tels que pour $\psi \in BV([0, \beta])$, pour $\varphi \in L^1(\mu_\beta)$,*

$$C_n(\varphi, \psi) \leq C\kappa^n \|\varphi\|_{L^1} \bigvee \psi.$$

Le processus $(Y_n)_{n\in\mathbb{N}}$ à temps inversé est Φ_{BV}-faiblement dépendant.

2. Les tours "de Young"

La notion de tours en théorie ergodique remonte à Rohlin, l'idée essentielle est d'induire le système dynamique, c'est à dire de considérer des itérés $T^{R(x)}(x)$ ou R est une application à valeur entière dite application de retour. Les tours constituent un modèle d'extension markovienne auquel se ramène une très large catégorie de systèmes dynamiques. À la suite des travaux de Hofbauer et Keller ([HoKe, Ke]), L-S Young ([Y1, Y2]) a largement développé cette stratégie, le but est de considérer de "bons retours" de manière à ce que l'application "tour" vérifie de bonnes propriétés.

2.1. **Tours de Young : définition et propriétés élémentaires.** Formellement, une tour est donnée par :

- un espace de probabilité (Δ_0, m_0),
- une application $T\ : \Delta_0 \to \Delta_0$ inversible par morceaux, non singulière. On note $\mathcal{P}_0 = \{\Delta_{0,j},\ j \in \mathbb{N}\}$ la partition (modulo m_0) d'inversibilité,
- et une application de retour $R\ : \Delta_0 \to \mathbb{N}$, m_0 intégrable.

On "déploie" alors le système dynamique en un nouveau système dynamique F sur la tour

$$\Delta = \{(x,\ell) \in \Delta_0 \times \mathbb{N}\ /\ R(x) > \ell\} = \bigcup_{\ell\in\mathbb{N}} \Delta_\ell$$

par $F(x,\ell) = (x,\ell+1)$ si $R(x) > \ell+1$ et $F(x,\ell) = (T^{R(x)}(x), 0)$ sinon. Les Δ_ℓ sont les étages de la tour.
La tour Δ est munie de la mesure m définie par $m(A \times \{\ell\}) = C \cdot m_0(A)$ si $A \times \{\ell\} \subset \Delta$, la constante $C > 0$ est choisie pour que m soit une mesure de probabilité, une telle constante existe car R est m_0-intégrable.
Pour x et y dans Δ_0, le *temps de séparation* $s_0(x,y)$ est le plus grand entier p tel que pour $0 \le j \le p$, $(F^R)^j(x)$ et $(F^R)^j(y)$ appartiennent au même élément $\Delta_{0,j}$ de $\mathcal{P}_0$, pour (x,ℓ) et $(y,\ell) \in \Delta$, on définit le temps de séparation par $s((x,\ell),(y,\ell)) = s_0(x,y)$. Si x et $y \in \Delta$ ne sont pas sur le même étage, on pose $s(x,y) = 0$. La tour est alors munie d'une métrique dynamique : soit $0 < \beta < 1$, $d(x,y) = \beta^{s(x,y)}$.
On déploie aussi la partition $\mathcal{P}_0$ en une partition $\mathcal{P}$ sur Δ dont les atomes sont $\Delta_{\ell,j} = \{(x,\ell) \in \Delta\ /\ x \in \Delta_{0,\ell}\}$. Une telle tour sera appelée "tour de Young" si elle vérifie les propriétés suivantes :

(1) **Partition génératrice** La partition $\mathcal{P}$ est génératrice.

(2) **Propriété de Markov** L'application F est markovienne (par rapport à partition $\mathcal{P}$).

(3) **Propriété de grandes images**

$$\inf\{m_0(T(\Delta_{0,j}))\ /\ j \in \mathbb{N}\} > 0.$$

(4) **Régularité du Jacobien** On note JF^R le jacobien - par rapport à la mesure m_0 - de F^R. Il existe $C > 0$ tel que pour tout

$j \in \mathbb{N}$, pour tout $x, y \in \Delta_{0,j}$,

$$\left| \frac{JF^R(x)}{JF^R(y)} - 1 \right| \leq Cd(F^R(x), F^R(y)).$$

(5) **Mélange topologique** pour tout i, j, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$,

$$m_0((F^R)^{-n}\Delta_{0,j} \cap \Delta_{0,i}) > 0.$$

Remarque. Le mélange topologique, la propriété de grandes images et celle de Markov sont garanties par les deux propriétés suivantes : pour tout élément de la partition $\Delta_{j,0}$, $F^{R(x)}(\Delta_{j,0}) = \Delta_0$ et $pgcd\{R(x)\ x \in \Delta_0\} = 1$. Ce sont ces conditions qu'utilise L-S. Young.

Nous allons travailler avec les fonctions f bornées sur la tour, Lipschitz sur chaque Δ_ℓ avec une constante de Lipschitz uniformément bornée (en ℓ), on notera L l'ensemble de ces fonctions. Elles vérifient : il existe $K > 0$ tel que pour tout ℓ, pour x et $y \in \Delta_\ell$,

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

La constante de Lipschitz $L(f)$ est l'infimum des $K > 0$ tels que l'inégalité ci-dessus est vérifiée pour tout ℓ. L'espace L est muni de la norme $\|f\| = \max(\|f\|_\infty, L(f))$ qui en fait un espace de Banach.
Dans ce cadre, L-S. Young a montré l'existence et l'unicité d'une mesure invariante par F, absolument continue par rapport à m, mélangeante et a montré que la décroissance des corrélations pour des fonctions de L est relié à $m_0(R > n)$. Si $m_0(R > n)$ décroît exponentiellement, il est va de même pour les coefficients de corrélation. Si $m_0(R > n)$ décroît en $\frac{1}{n^\alpha}$, il en va de même pour les coefficients de corrélation.
Nous allons présenter deux résultats qui précisent ceux de L.S. Young : d'une part nous obtenons une estimation de la décroissance des corrélations pour des fonctions Lipschitz sur les tours par une autre méthode que celle de Young : cette méthode présente l'avantage de conduire à un mélange proche du Φ_L-mélange pour un processus à temps inversé. D'autre part, nous montrons, et c'est un résultat obtenu avec J. Buzzi, que l'on peut estimer la décroissance des corrélations sous des conditions de régularité moins fortes.
Nous donnerons ensuite deux exemples de systèmes conjugués à des tours de Young : certains systèmes uniformément dilatants en dimension supérieure (travail en collaboration avec J. Buzzi) et certaines compositions aléatoires d'applications unimodales (travail en collaboration avec V. Baladi et M. Benedicks).

2.2. **Décroissance des corrélations sur les tours.** Considérons donc une tour de Young. L'opérateur de transfert associé

$$\mathcal{L}(f)(x) = \sum_{F(y)=x} JF(y)^{-1} f(y)$$

agit sur L. Les techniques de cônes utilisées dans [12] peuvent être adaptées aux tours de Young et permettent de montrer le résultat suivant.

Théorème I.6. [11] *L'opérateur $\mathcal{L}$ admet un unique point fixe non nul $h \in L$, la mesure $\mu = hm$ est ergodique et mélangeante. De plus, $\mathcal{L}^n f$ converge vers $hm(f)$ uniformément sur chaque sous ensemble compact de Δ et dans $L^1(m)$ pour toute fonction f uniformément continue sur Δ.*
Pour toute suite croissante $(v_n)_{n\in\mathbb{N}}$ telle que $\sum_{\ell\in\mathbb{N}} m_0(\Delta_\ell)\cdot v_\ell < \infty$, pour toute fonction $\psi \in L$ et $\varphi \in L^1(v\cdot m)$, il existe $0 < \gamma < 1$ tel que

$$C_n(\varphi,\psi) \leq Ct\ \max[(v_n)^{-1},\ \gamma^n]\|\psi\|\ \|\varphi v\|_1.$$

En particulier
- *s'il existe des constantes $0 < e_1$ et $0 < \theta < 1$ telles que $\ell \in \mathbb{N}$, $m_0(R > \ell) \leq e_1\theta^\ell$ alors il existe $0 < \gamma < 1$ tel que pour tout $\theta' < \theta$, et $\varphi \in L^1(v\cdot m)$ $(v_n = \theta'^{-n})$,*
$$C_n(\varphi,\psi) \leq\ Ct\ (\theta')\ \gamma^n\ \|\psi\|\ \|\varphi\cdot v\|_1\,,$$
- *s'il existe des constantes $0 < e_1$, $0 < \theta < 1$ et $0 < \beta < 1$ telles que pour tout $\ell \in \mathbb{N}$, $m_0(R > \ell) \leq e_1\theta^{\ell^\beta}$ alors pour tout $\beta' < \beta$, soit $v_\ell = \theta^{-\ell^{\beta'}}$,*
$$C_n(\varphi,\psi) \leq Ct\ (\beta')\theta^{n^{\beta'}}\ \|\psi\|\ \|\varphi\cdot v\|_1,$$
- *s'il existe des constantes $0 < e_1$, $\beta > 1$ telles que pour tout $\ell \in \mathbb{N}$, $m_0(R > \ell) \leq e_1\ell^{-\beta}$, alors pour tout $\gamma > 1$, soit $v_\ell = \left[\frac{(\ln\ell)^\gamma}{\ell^\beta}\right]^{-1}$,*
$$C_n(\varphi,\psi) \leq Ct\ (\gamma)\frac{(\ln n)^\gamma}{n^\beta}\ \|\psi\|\ \|\varphi\cdot v\|_1.$$

Remarque. Les résultats de Young ([Y2]) sont très proches des estimations ci-dessus. La méthode utilisée est complètement différente. Notons que la dépendance des coefficients de corrélation $C_n(\varphi,\psi)$ se fait dans notre cas par une norme L^1 alors que c'est une norme L^∞ pour Young. Ceci n'est pas tout à fait suffisant pour obtenir du Φ_L-mélange mais donne néanmoins une notion de mélange plus forte que celle obtenue par Young. Remarquons enfin que nos vitesses ne sont pas tout à fait optimales dans le cas polynômial : Young obtient $O(\frac{1}{n^\beta})$ là où nous obtenons $O(\frac{(\ln n)^\gamma}{n^\beta})$ pour tout $\gamma > 1$.

L'association des techniques de cônes utilisées dans [15] et [12] permet d'affaiblir la condition de régularité sur JF. Si on note $\mathcal{P}^n$ la partition d'inversibilité de F^n. Le module de continuité de JF est contrôlé par la suite :

$$w_n = \sup_{C\in\mathcal{P}^n}\sup_{x,y\in C}\log\frac{JF(x)}{JF(y)},$$

l'hypothèse de régularité sur JF est que la suite $(w_n)_{n\in\mathbb{N}}$ est sommable. Dans ce paragraphe, le temps de séparation $s'(x,y)$ entre deux éléments x et y de Δ sera le plus grand entier p tel que pour $0 \leq j \leq p$, $F^j(x)$ et $F^j(y)$ appartiennent au même atome de la partition $\mathcal{P}$. La distance d' est construite pour que la famille de fonctions $\log(JF^n)$ soit uniformément Lipschitz :

$$d'(x,y) = \sum_{j>s'(x,y)} w_j.$$

En collaboration avec Jérôme Buzzi, nous avons obtenu le résultat suivant.

Théorème I.7. [3] *Il existe une unique mesure μ de probabilité invariante par F et absolument continue par rapport à m. Elle est ergodique et mélangeante. Pour toute fonction $\psi \in L$ et $\varphi \in L^\infty(\Delta)$,*

$$C_n(\varphi,\psi) \leq C \cdot \|\psi\|\|\varphi\|_{L^1(\mu)} \cdot u_n \qquad \text{pour tout } n \geq 0$$

pour une constante $C < \infty$ et une suite $u = (u_n)_{n=0}^\infty$ qui converge vers zéro et qui peut être spécifiée :
- *si $w_n = O(\rho^n)$ pour un $0 < \rho < 1$et $m_0(R > n) = O(\alpha^n)$ pour un $0 < \alpha < 1$ alors $u_n = \kappa^n$ pour un $0 < \kappa < 1$,*
- *si $w_n = O(n^{-\alpha})$ pour $\alpha > 1$ et $m(R > n) = O(n^{-\beta})$ avec $\beta > 1$ alors $u_n = n^{-\min(\alpha-1,\beta-\varepsilon)}$ pour tout $\varepsilon > 0$.*
- *si $w_n = O(e^{-n^\alpha})$ et $m(R > n) = O(e^{-n^\beta})$ avec $0 < \alpha, \beta < 1$, alors $u_n = e^{-n^{\min(\alpha,\beta)-\varepsilon}}$ for tout $\varepsilon > 0$.*

Remarque. La métrique utilisée est plus petite que la métrique d du paragraphe précédent. L'espace des fonctions Lipschitz pour d' est donc plus petit mais le type de mélange obtenu est plus fort : le théorème montre - là encore - que le processus à temps inversé est Φ_L-mélangeant.

Nous avons donc vu que pour les systèmes dynamiques du type “tours de Young”, on peut estimer la vitesse de décroissance des corrélations. De nombreux systèmes dynamiques peuvent être conjugués à une telle tour : applications unimodales et de type Hénon, certains billards, des applications de type “Lasota-Yorke”, des applications de l'intervalle à point fixes neutres, des difféomorphismes partiellement hyperboliques, ... ([BeY], [Che], [Y1, Y2],[Dol], [Ca]). Dans [9], nous avons obtenu, en collaboration avec Jérôme Buzzi une estimation de la décroissance des corrélations pour des applications dilatantes par morceaux en dimension > 1 en conjuguant le système à une tour de Young. La présentation de ce résultat est l'objet du paragraphe suivant.

2.3. **Systèmes dynamiques dilatants en dimension supérieure.** Le but est de généraliser les propriétés ergodiques et statistiques bien connues en dimension 1 ([LaY1], [HoKe], [Sc], [Li3] ...) à un cadre multidimensionnel. Dans ce cadre, diverses questions ont été considérées

(existence et caractérisation de mesures d'équilibre [Bu3], de mesures conformes [BuPaS], de mesures absolument continues [Bu1, Bu2, Bu4, Cow, GoBo, Sau]).
Nous considérons une application T inversible par morceaux, sur un espace métrique compact M, $\mathcal{P}$ est une famille finie d'ouverts disjoints et bornés tels que $M = \overline{\bigcup_{P\in\mathcal{P}} P}$. Les hypothèses générales sur T sont les suivantes :

(1) T est inversible sur chaque $P \in \mathcal{P}$ et continue sur $\overline{P}$.
(2) $g : M \to \mathbb{R}$ est la fonction potentiel, g est γ-Hölder et admet une mesure conforme ν (dans ce cas, elle est nécessairement $e^{P_{top}(T)}$-conforme, voir [BuPaS]), il existe $n \in \mathbb{N}$ avec $T^n(M) \subset \mathrm{supp}(\nu)$.
(3) T est uniformément dilatante sur chaque $P \in \mathcal{P}$.
(4) La partition $\mathcal{P}$ est génératrice.
(5) Notre condition essentielle porte sur la frontière $\partial\mathcal{P}$ de $\mathcal{P}$: $\partial\mathcal{P} = \bigcup_{P\in\mathcal{P}} \partial P$. On suppose :

$$P(\partial\mathcal{P}, T) < P_{top}(T, g).$$

Remarque. La condition sur la pression du bord de $\mathcal{P}$ est satisfaite dans de nombreux cas. Si T est uniformément dilatante, pour le potentiel $g = |T'|^{-1}$ ou une petite perturbation de ce potentiel, elle est satisfaite : 1) en dimension 1, 2) en dimension 2 pour T analytique réelle, 3) en toute dimension pour T affine par morceaux. Il existe des contre-exemples C^∞ en dimension 2.

Théorème I.8. [9] *Soit $(M, T, \mathcal{P}, g)$ vérifiant les conditions (1-5) ci-dessus. Il existe un nombre fini de mesures de probabilités invariantes et ergodiques absolument continues par rapport à la mesure conforme ν. Soit μ une telle mesure. Alors (quitte à remplacer T par un itéré), la mesure μ est mélangeante et il existe $C > 0$ et $0 < \kappa < 1$ tels que pour $\varphi \in L^1(\mu)$, ψ γ-Hölder :*

$$C_n(\varphi, \psi) \le C\|\varphi\|_{L^1(\mu)} \cdot K(\psi) \cdot \kappa^n,$$

Remarque. On peut affaiblir la condition de régularité du potentiel et éviter la condition de dilatation uniforme : il suffit que T ne soit pas contractante (pour tout $x, y \in M$, $d(T(x), T(y)) \ge d(x, y)$) et que le module de continuité de g est sommable. On utilise alors le Théorème I.7 pour estimer la décroissance des corrélations. Elle est essentiellement équivalente au reste de la série $\sum_{k>n} w_k$ pour ψ suffisamment régulière.

La preuve de ce théorème repose sur la construction d'une tour de Young à laquelle le système initial est conjugué. Cette extension markovienne permet d'éviter d'avoir à traiter avec des généralisations multidimensionnelles de fonctions à variations bornées. Sans trop entrer dans les détails techniques, signalons que cette extension est fabriquée à partir de l'extension markovienne à la Hofbauer ([Ho]) : la base Δ_0 de la tour est un itéré de

$$\widehat{X_0} := \{(x,P) \ / \ x \in P, \ P \in \mathcal{P}\}.$$

L'application T s'étend en une application $\widehat{T}$ sur $\widehat{X} := \bigcup\{D \times \{D\} \ / \ D = T^n P, \ P \in \mathcal{P}^n\}$. Cette application est naturellement markovienne mais ne vérifie pas toutes les conditions de la section précédente. Nous devons en particulier estimer le temps de retour. L'ingrédient essentiel de la construction est la proposition suivante.

Proposition I.9. *Pour tout $\delta > 0$, il existe $N_\star$ tel que si $\widehat{X}_\star = \widehat{T}^{N_\star}\widehat{X_0}$, le temps de retour sur $\widehat{X}_\star$, $\widehat{R}(x,D) = \inf\{n \geq 1 \ / \ \widehat{T}^n(x,D) \in \widehat{X}_\star\}$ vérifie pour tout $D \in \mathcal{P}_\star$:*

$$\nu(x \in D \ / \ \widehat{R}(x,D) > n) \leq Const \exp[-n(P_{top}(T,g) - P(\partial P,T) - \delta)].$$

Ce contrôle du temps de retour en fonction de la quantité $P_{top}(T,g) - P(\partial P,T)$ illustre le rôle central de l'inégalité $P(\partial P,T) < P_{top}(T,g)$.

Nous concluons cette partie en appliquant les techniques de tours “à la Young” à des compositions aléatoires d'applications unimodales. Ce travail a été réalisé en collaboration avec V. Baladi et M. Benedicks.

2.4. **Compositions aléatoires d'applications unimodales.** Les applications unimodales sont un premier exemple d'applications non uniformément hyperboliques. Elles ont étés largement étudiées ([ColE], [BeY] pour les premières approches du point de vue théorie ergodique ...). On connaît des conditions (satisfaites pour un ensemble de mesure positive du paramètre dans la famille $x \mapsto a - x^2$) qui garantissent l'existence et l'unicité d'une mesure invariante absolument continue et la décroissance exponentielle des corrélations ([Y1]). Nous nous intéressons ici aux compositions aléatoires de telles applications et plus précisément aux compositions aléatoires de petites perturbations d'une application unimodale. La motivation est double :

(1) les compositions aléatoires d'applications unidimensionnelles peuvent servir de modèles simplifiés de systèmes non uniformément hyperboliques multi-dimensionnels, il s'agit en quelque sorte de systèmes de dimension “$1 + \varepsilon$”,
(2) les compositions aléatoires de perturbations d'une application sont une modélisation plus proche de la réalité que l'itération d'une application et permettent d'obtenir des résultats de stabilité.

Avant de définir précisément les applications auxquelles nous allons nous intéresser, précisons le cadre et la problématique globale.
Étant donné une application $T : I \to I$ de l'intervalle, pour $\varepsilon > 0$

fixé, nous considérons $T_\omega(x) = T(x) + \omega_0$ avec $\omega \in [-\varepsilon, \varepsilon]^{\mathbb{Z}} = \Omega$ et par induction $T^n_\omega(x) = T^{n-1}_{\sigma\omega} \circ T_\omega(x)$ où $\sigma\ : [-\varepsilon, \varepsilon]^{\mathbb{Z}} \to [-\varepsilon, \varepsilon]^{\mathbb{Z}}$ est le décalage $(\sigma(\omega))_k = \omega_{k+1}$. Si on considère une mesure ν_ε sur $[-\varepsilon, \varepsilon]$ et la mesure produit sur $[-\varepsilon, \varepsilon]^{\mathbb{Z}}$, on peut étudier la chaîne de Markov :

$$\mathbb{P}(X_{n+1} \in E \mid X_n = x) = \int\limits_{[-\varepsilon,\varepsilon]} \mathbf{1}_E(T_\omega)(x) d\nu_\varepsilon^{\otimes\mathbb{Z}}(\omega).$$

L'existence d'une unique mesure stationnaire μ_ε pour cette chaîne est garantie sous des hypothèses assez faible sur T. Se posent alors les questions du mélange, de l'estimation de la vitesse de mélange et de la stabilité (i.e. la convergence de μ_ε vers une mesure μ invariante par T). Ce point de vue "intégré" : on intègre sur $[-\varepsilon, \varepsilon]^{\mathbb{Z}}$ les orbites x, $T_\omega(x)$, ..., $T^n_\omega(x)$ ..., a été étudié pour T une application unimodale par Benedick-Young ([BeY]), Baladi-Young ([BaY]), Baladi-Viana ([BaVi]).

Le point de vue physique ou "expérimental" consiste à étudier - $\nu_\varepsilon^{\otimes\mathbb{Z}}$ presque toutes - les orbites x, $T_\omega(x)$, ..., $T^n_\omega(x)$ ..., $\omega \in [-\varepsilon, \varepsilon]^{\mathbb{Z}}$. Les *mesures stationnaires échantillonnées absolument continues* constituent un outil naturel pour cet étude. Il s'agit de familles de mesures sur I absolument continues, $\mu_\omega = h_\omega \text{Leb}$ telles que $(T_\omega)_\star \mu_\omega = \mu_{\sigma\omega}$, les h_ω seront appelé *densités quasi-invariantes*. Ces mesures μ_ω peuvent être obtenues est désintégrant $\mu_\varepsilon \times (\nu_\varepsilon)^{\otimes\mathbb{Z}}$. Se posent alors les questions de mélange aléatoires de ces mesures, d'estimation de la vitesse de mélange et de stabilité. Plus précisément, on recherche une estimation des quantités :

$$C^{(f)}_{\omega,\varphi,\psi}(n) = \left| \int (\varphi \circ f^n_\omega)\psi \, d\text{Leb} - \int \varphi \, d\mu_{\sigma^n\omega} \int \psi \, d\text{Leb} \right|,$$

$$C^{(p)}_{\omega,\varphi,\psi}(n) = \left| \int (\varphi \circ f^n_{\sigma^{-n}\omega})\psi \, d\text{Leb} - \int \varphi \, d\mu_\omega \int \psi \, d\text{Leb} \right|.$$

Le but est d'obtenir, pour $\mathbb{P} := \nu_\varepsilon^{\otimes\mathbb{Z}}$-presque tout ω, des bornes supérieures du type $C_\omega C_{\varphi,\psi} \cdot \rho(n)$, avec $\rho(n) \to 0$, pour certaines classes de fonctions φ et ψ, ainsi que des bornes sur $\mathbb{P}(\omega \in \Omega \ / \ C_\omega > n)$.

Précisons maintenant quelles sont les applications considérées et les résultats.

2.4.1. *Applications unimodales.* Soit $I = [L, R]$ un intervalle compact contenant 0 dans son intérieur et $T : I \to I$ une application C^2 unimodale (i.e., T est croissante sur $[L, 0]$, décroissante sur $[0, R]$) satisfaisant de plus : $T''(0) \neq 0$, $\sup_I |T'| < 8$ et

(H1) Il existe $\alpha > 0$ et $1 < \lambda \leq 4$ avec $200\alpha < (\log \lambda)$ tels que

(1) $|(T^n)'(T(0))| \geq \lambda^n$ pour tout $n \in \mathbb{N}_+$.

(2) $|T^n(0)| \geq e^{-\alpha n}$ pour tout $n \in \mathbb{N}_+$.

(H2) Pour tout $\delta > 0$ suffisamment petit, il existe $M = M(\delta) \in \mathbb{N}_+$ pour lequel :

(1) si $x, \ldots, T^{M-1}(x) \notin (-\delta, \delta)$ alors $|(T^M)'(x)| \geq \lambda^M$;

(2) pour tout n, si $x, \ldots, T^{n-1}(x) \notin (-\delta, \delta)$ et $T^n(x) \in (-\delta, \delta)$, alors $|(T^n)'(x)| \geq \lambda^n$.

(H3) $T(I)$ est inclus dans l'intérieur de I.
(H4) T est topologiquement mélangeant sur $[T^2(0), T(0)]$.
Quelques commentaires : Les hypothèses **(H1)**, **(H2)**, **(H4)** sont standards dans l'étude des applications unimodales. L'hypothèse **(H1)** garantit que l'image $f(0)$ du point critique est hyperbolique et que l'orbite du point critique 0 ne revient pas trop près de 0. L'hypothèse **(H2)** exprime l'hyperbolicité des morceaux d'orbites à l'extérieur d'un voisinage de 0. Ce sont des hypothèses de type Benedicks-Carleson ([BeC]). On sait qu'elles sont satisfaites pour un ensemble de mesure de Lebesgue positive de paramètres a, pour des applications $x \mapsto a-x^2$. L'hypothèse **(H3)** permet des perturbations de la forme $T+\omega_0$, $\omega_0 \in [-\varepsilon, \varepsilon]$ pour ε assez petit.
Fixons ε tel que $\forall x \in I$, $T(x) \pm \varepsilon \in I$. On considère une mesure ν_ε sur $[-\varepsilon, \varepsilon]$ telle que : il existe une constante $C > 0$ telle que pour tout sous intervalle $J \subset I$,

$$\nu_\varepsilon(J) \leq C \frac{\mathrm{Leb}(J)}{\varepsilon}.$$

Remarques. On pourrait affaiblir cette condition en remplaçant $\mathrm{Leb}(J)$ par $\mathrm{Leb}(J)^u$ pour un $0 < u \leq 1$ fixé, une telle hypothèse influe sur la décroissance des corrélations mais ne modifie pas fondamentalement la preuve.
On ne peut se passer d'une telle hypothèse ; en effet, dans le cas de la famille $x \mapsto a - x^2$ au voisinage de a telle que l'application vérifie les hypothèses **H1-4**, il existe des paramètres tels que la dynamique admet une orbite périodique attractante. Il nous faut donc une hypothèse qui garantisse que ν_ε n'est pas une mesure de Dirac sur un tel point.
L'hypothèse sur ν_ε est vérifiée si ν_ε est absolument continue par rapport à la mesure de Lebesgue, de densité bornée.
Enfin nous ne supposons pas que 0 est dans le support de ν_ε.

2.4.2. *Résultats.* Sous les hypothèses de la section 2.4.1 nous obtenons le résultat suivant.

Théorème I.10. [10] *Pour ε suffisamment petit, pour $(\nu_\varepsilon)^{\otimes \mathbb{Z}}$ presque tout $\omega \in [-\varepsilon, \varepsilon]^{\mathbb{Z}}$, il existe une densité quasi-invariante $h_\omega \in L^1(dLeb)$. Il existe $C(\varepsilon) \geq 1$ et pour presque tout $\omega \in [-\varepsilon, \varepsilon]^{\mathbb{Z}}$, il existe $C_\omega^{(1)} > 0$, $C_\omega^{(2)} > 0$ (qui dépendent de ε) tels que pour toute fonction Lipschitz $\psi : I \to \mathcal{C}$ et toute fonction bornée $\varphi : I \to \mathcal{C}$, les corrélation "passées" et "futures" vérifient pour $n \geq 1$*

$$\left| \int \varphi \circ f^n_{\sigma^{-n}\omega} \, \psi \, dLeb - \int \varphi h_\omega \, dLeb \int \psi \, dLeb \right| \leq C_\omega^{(1)} \sup |\varphi| \, L(\psi) \, e^{-(n^{1/16}/C(\varepsilon))} ,$$

et

$$\left| \int \varphi \circ f_\omega^n \, \psi \, dLeb - \int \varphi h_{\sigma^n \omega} \, dLeb \int \psi \, dLeb \right| \leq C_\omega^{(2)} \sup |\varphi| \, L(\psi) e^{-(n^{1/16}/C(\varepsilon))} .$$

Il existe $C(\varepsilon)$ *et* $v > 1$ *tel* $C_\omega = \max(C_\omega^{(1)}, C_\omega^{(2)}, C_\omega^{(3)})$ *vérifie*

$$\mathbb{P}(\{\omega \in \Omega_\varepsilon \mid C_\omega > n\}) \leq \frac{C(\varepsilon)}{n^v} .$$

2.4.3. *Stratégie de la preuve.* Pour obtenir ce résultat, nous construisons une famille $(\Delta_\omega, F_\omega, R_\omega)$ de tours telles que $F_\omega \; : \Delta_\omega \to \Delta_{\sigma\omega}$ et qui vérifient une version "aléatoirisée" des propriétés des tours de Young. Nous adaptons la preuve de Young par couplage à notre modèle aléatoire et revenons au problème initial par une projection $\pi_\omega \; : \Delta_\omega \to I$ telle que $\pi_{\sigma\omega} F_\omega = T_\omega \circ \pi_\omega$.

Deuxième partie II. D'autres propriétés statistiques

Pour les systèmes dynamiques dont on sait estimer la vitesse de décroissance des corrélations, il est naturel de rechercher d'autres propriétés statistiques. Mes travaux dans ce domaine s'orientent dans deux directions : existence et propriétés de mesures conditionnellement invariantes; inégalités de concentration et estimation pour des systèmes dynamiques.

D'une part, les mesures conditionnellement invariantes permettent de mesurer le taux d'échappement ou d'absorption du système. Plus précisément, étant donné un système dynamique $(M, \mathcal{B}, T)$ et $Y \subset M$: le "trou", M_n est l'ensemble des points dont l'orbite par T ne rencontre pas Y avant le temps n. Une mesure de probabilité ν est conditionnellement invariante si pour tout n et $A \in \mathcal{B}$, $\nu(T^{-n}A \cap M_n) = \nu(A)\nu(M_n)$. Si τ est le temps d'absorption, la loi de τ sous ν est une loi géométrique. Un système à trous modélise des systèmes ouverts (type "billards"), τ est alors vu comme le temps d'échappement.

Dans [13], nous avons donné des conditions générales d'existence d'une mesure conditionnellement invariante et d'une mesure conditionnellement invariante absolument continue par rapport à une mesure de référence. Ces conditions sont vérifiées pour des systèmes dynamiques Φ-mélangeants, Axiom A, avec Y un trou markovien. Dans [7], nous avons donné des conditions d'existence et d'unicité d'une mesure conditionnellement invariante absolument continue par rapport à la mesure de Lebesgue pour des applications de type "Lasota-Yorke". Le comportement asymptotique des systèmes dynamiques à trou est aussi lié aux propriétés de récurrence et à la pression du système. Dans [6], nous montrons une généralisation d'un théorème d'Orstein et Weiss qui relie le temps de récurrence à la pression du système.

D'autre part, dans le but d'utiliser les systèmes dynamiques pour la modélisation et la simulation de systèmes complexes, il faut estimer des paramètres et des fonctions afin de reconstruire le système dynamique. Cette problématique est en développement. [2] est un premier pas dans cette direction : on montre des inégalités exponentielles pour des systèmes faiblement dépendants; appliquées aux systèmes dynamiques, ces inégalités permettent d'estimer la fonction potentiel g du système ($g^{-1} = \frac{d\mu \circ T}{d\mu}$, si μ est une mesure T-invariante).

1. Temps d'absorption, récurrence et pression

Ce paragraphe présente trois articles autour des thèmes de récurrence et pression. Les deux premiers concernent les systèmes dynamiques "à trous" et les mesures conditionnellement invariantes, le troisième relie la fonction de retour à la pression du système, généralisant ainsi un résultat de Orstein et Weiss ([OW]).

1.1. **Systèmes à trous et mesures conditionnellement invariantes.** La notion de mesure de probabilité conditionnellement invariante a été introduite pour les chaînes de Markov avec un état absorbant ([VeJ]). Le problème de l'existence de ces mesures pour des systèmes dynamiques a été abordé par Pianigiani et Yorke ([PiY]) puis par Collet-Martinez-Schmitt ([ColMSc2, ColMSc3, ColMSc4]) dans le cadre des applications dilatantes et des chaînes de Markov topologiques. Ces questions ont aussi été étudiées pour des difféomorphismes d'Anosov par Chernov-Markarian-Troubeskoy et Urbanski ([CheMT], [U]).
Dans [13], en collaboration avec Pierre Collet et Servet Martinez, nous donnons des conditions générales d'existence d'une mesure de probabilité conditionnellement invariante, éventuellement absolument continue par rapport à une mesure de référence. Ces conditions sont satisfaites, par exemple, pour des difféomorphismes de type Axiom A avec un trou markovien. Dans [7], en collaboration avec Carlangelo Liverani, nous abordons ces questions pour des applications de type Lasota-Yorke en dimension 1.
Avant de détailler ces résultats, nous rappelons quelques définitions.

1.1.1. *Définitions.* Nous considérons M un espace métrique, $\mathcal{B}$ la tribu des boréliens et $T : M \to M$ une application mesurable. Soit $Y \subset M$ un ensemble mesurable non trivial (appelé le trou), et $M_0 = M \backslash Y$. Pour $x \in M$, le temps d'entrée dans Y est $\tau(x) = \inf\{n \geq 0 \ / \ T^n(x) \in Y\}$ et $M_n = \{x \in M \ / \ \tau(x) > n\}$, les points qui n'ont pas été absorbé au temps n. Une mesure de probabilité ν portée par M_0 est une mesure *conditionnellement invariante* si : pour tout $n \geq 0$, pour tout $A \in \mathcal{B}$,

$$\nu(T^{-n}A \cap M_n) = \nu(A)\nu(M_n).$$

On remarque que pour une telle mesure, il existe $\theta \in]0,1]$ tel que pour tout $n \geq 0$, $\nu(M_n) = \theta^n$. Si $\theta \neq 1$, nous dirons que ν est non triviale.
Étant donnée une mesure de probabilité μ portée par M_0, on considère F^μ la fonction de répartition de τ sous μ : $F^\mu(n) = \mathbb{P}_\mu(\tau \leq n) = 1 - \mu(M_n)$. Nous allons donner des conditions suffisantes d'existence d'une mesure de probabilité conditionnellement invariante portant sur F^μ.
Dans un deuxième temps, nous nous intéressons à l'existence de mesures conditionnellement invariantes absolument continues par rapport à une mesure de référence m. Si T est non singulière (par rapport à cette mesure), on peut considérer l'opérateur de transfert $\mathcal{L}_0$ défini par :

$$\int_M (\mathcal{L}_0 f)\varphi dm = \int_M f \cdot \varphi \circ T dm,$$

définissons alors l'opérateur $\mathcal{L}$ par $\mathcal{L}(f) = \mathcal{L}_0(f\mathbf{1}_{M_0})$. On vérifie facilement qu'il existe une mesure de probabilité conditionnellement invariante, absolument continue si et seulement si il existe $h \neq 0$, $h \in L^1_+(m)$ et $\alpha \in]0,1]$ tels que $\mathbf{1}_{M_0}\mathcal{L}(h) = \alpha h \mathbf{1}_{M_0}$.

1.1.2. *Conditions générales d'existence.* Nous nous plaçons sous les hypothèses du paragraphe précédent.

Théorème II.1. [13] *Supposons que $T : M \longrightarrow M$ est continue, que les ensembles M_n, $n \geq 0$ sont soit tous ouverts soit tous fermés et que pour n assez grand, $M \setminus M_n$ est inclus dans un ensemble compact. S'il existe une mesure μ portée par M_0 telles que les trajectoires (sous μ) sont absorbées géométriquement i.e. il existe $C < \infty$, $\theta_0 \in]0,1[$ tel que pour tout $n \geq 0$,*

$$1 - F^\mu(n) \leq C\theta_0^n,$$

si, de plus, $\mu(\bigcup_{n\geq 0} \partial(T^{-n}Y)) = 0$ alors il existe une mesure de probabilité conditionnellement invariante non triviale.

Les conditions topologiques sur les M_n sont satisfaites dès que Y est soit ouvert soit fermé et T continue. L'hypothèse de continuité de T n'est pas tout à fait nécessaire et pourrait être remplacée par de la continuité par morceaux. La condition sur la mesure du bord de Y est satisfaite dès lors que $\mu(\partial Y) = 0$ et T est non-singulière par rapport à μ. La condition sur F^μ est satisfaite sous certaines hypothèses de mélange, c'est l'objet de la proposition suivante.

Proposition II.2. [13] *Soit $M_0 \in \mathcal{B}$ et ν une mesure de probabilité T-invariante telle que $\nu(M_0) < 1$. Supposons qu'il existe une suite $\varepsilon_n = \varepsilon_n(M_0)$ avec $\varepsilon_n \to 0$ quand $n \to \infty$, telle que*

$$\forall B \in \Gamma(M_0) = \left\{ \bigcap_{\ell=0}^{q} T^{-\ell r} M_0 : q \in \mathbb{N}, r \geq 0 \right\}$$

on a :

$$|\nu(X_0 \cap T^{-n}B) - \nu(X_0)\nu(B)| \leq \nu(B)\varepsilon_n \quad \forall n \geq 0.$$

Alors il existe $\theta > 0$ et $C > 0$ tels que :

$$1 - F^\nu(t) \leq Ce^{-\theta t} \quad \forall t \geq 0.$$

L'hypothèse de mélange ci-dessus est satisfaite pour la plupart des applications considérées dans la partie 1 à condition que $\mathbf{1}_{X_0}$ soit suffisamment régulière (par exemple à variation bornée dans le cas unidimensionnel ou Lipschitz dans le cas symbolique).
Comme pour les mesures invariantes, toutes les mesures conditionnellement invariantes ne sont pas pertinentes. Un moyen de sélectionner des mesures conditionnellement invariantes intéressantes est de considérer des mesures conditionnellement invariantes absolument continues par

rapport à une mesure de référence. On suppose que T est non singulière par rapport à une mesure de référence m (par exemple la mesure de Lebesgue ou une mesure conforme ou une mesure SRB).

Définition 1. Une suite $\big(\rho_n\big)_{n\in\mathbb{N}}$ de mesures de probabilité est *asymptotiquement uniformément absolument continue* par rapport à m si pour tout $\gamma > 0$, il existe un nombre $\delta = \delta(\gamma) > 0$ et pour tout ensemble U ouvert tel que $\lambda(U) < \delta$, un entier $N(U, \delta)$ tel que pour tout entier $n > N(U, \delta)$ on a $\rho_n(U) \le \gamma$.

On vérifie facilement (voir [13]) qu'une mesure conditionnellement invariante absolument continue par rapport à m est asymptotiquement uniformément absolument continue par rapport à m. Le résultat suivant donne des conditions suffisantes pour obtenir une mesure conditionnellement invariante absolument continue.

Théorème II.3. [13] *Soit m une mesure de probabilité telle que pour tout n, $m\big(M_n\big) > 0$. Soit*

$$\mu_n(A) = \frac{m\big(T^{-n}(A) \cap M_n\big)}{m\big(M_n\big)} \ .$$

Supposons que les trois conditions suivantes sont satisfaites :

(1) la suite $\big(\mu_n\big)_{n\in\mathbb{N}}$ est asymptotiquement uniformément absolument continue par rapport à m.

(2) $\vartheta = \max\Big\{x \ : \ \sum_{j=0}^{\infty} x^j m\big(M_j\big) < \infty\Big\} < \infty$.

(3) M_0 est compact.

Alors il existe une mesure conditionnellement invariante absolument continue par rapport à m.

La condition 2 ci-dessus est - par exemple - satisfaite si zéro n'est pas valeur d'adhérence de la suite $\dfrac{m(M_{n+1})}{m(M_n)}$, c'est le cas pour les applications markoviennes avec m une mesure conforme et Y un 1-cylindre tel que pour tout 1-cylindre C, $TC \not\subset Y$. Pour que la condition 1 ci-dessus soit satisfaite, il suffit que pour tout U ouvert, pour tout n assez grand,

$$m(T^{-n}(U) \cap M_n) \le \mathrm{Const} m(U) m(M_n).$$

Dans [13], on utilise le théorème ci-dessus pour montrer que pour un difféomorphisme Axiom A mélangeant, pour un trou markovien, il existe une unique mesure conditionnellement invariante dont la désintégration le long des feuilles instables est absolument continue par rapport à la désintégration le long des feuilles instables de la mesure SRB.

Le but du paragraphe suivant est l'étude de mesures conditionnellement invariantes pour des systèmes qui ne sont pas markovien (ou qui sont markoviens avec des trous non markoviens).

1.1.3. *Applications de Lasota-Yorke à trous.* On appellera *application de Lasota-Yorke* une application de l'intervalle $I = [0, 1]$, C^1 et monotone par morceaux. On notera $\mathcal{P}$ la partition de monotonie de T et $\mathcal{P}^{(n)}$ la partition de monotonie de T^n.
Soient $T : I \longrightarrow I$ une application de type Lasota-Yorke sur l'intervalle I, Y un sous intervalle non trivial et $g^0 : I \longrightarrow \mathbb{R}^+$ un potentiel strictement positif qui admet une mesure conforme m. Nous donnons des conditions constructives sur Y qui assurent l'existence d'une mesure de probabilité absolument continue (par rapport à m), invariante conditionnellement à la non absorption dans Y. Ces conditions impliquent aussi l'existence d'une mesure de probabilité, invariante par T et supportée par l'ensemble M_∞ des points qui ne tombent pas dans le trou. Nos conditions autorisent des trous relativement gros.
Notons $P_{top}(T, g^0)$ la pression topologique du système et $\Theta(g^0)$ l'exposant de g^0 définit par $\log \Theta(g^0) := \lim_{n\to\infty} \frac{1}{n} \log \sup_I g_n^0$, avec $g_n^0(x) = g^0(x) \cdot g^0(Tx) \dots g^0(T^{n-1}x)$. Nos hypothèses sur g^0 sont les suivantes :

Hypothèse 1.1.
- $\inf g^0 > 0$,
- le potentiel g^0 est *contractant* i.e., $\Theta(g^0) < e^{P_{top}(T,g^0)}$,
- le potentiel g^0 appartient à l'espace BV des fonctions à variations bornées,
- g^0 admet une mesure de probabilité conforme m.

Soit $Y \subset I$ un sous intervalle non trivial. Les sous ensembles M_n sont définis comme précédemment et décrivent les points qui n'ont pas été absorbés (par Y) au temps n. Soit g le potentiel du système à trou : $g = \mathbf{1}_{M_0} g^0$ et $\Theta = \Theta(g)$ l'exposant de g. Rappelons que si $\mathcal{L}_0$ est l'opérateur de transfert du système, l'opérateur $\mathcal{L}$ du système à trou est défini par $\mathcal{L}(f) = \mathcal{L}_0(f \mathbf{1}_{M_0})$, c'est l'opérateur de transfert associé au potentiel g :

$$\mathcal{L}(f)(x) = \sum_{T(y)=x} g(y) f(y).$$

Nous imposons deux conditions supplémentaires sur le système à trou. L'une assure que tous les points ne tombent pas dans le trou, l'autre est plus géométrique et exprime que les élément de $\mathcal{P}^{(n)}$ qui sont inclus dans M_n ne doivent pas être “trop contigus”.

Hypothèse 1.2. Soit $D_n := \{x \in I \mid \mathcal{L}^n 1(x) \neq 0\}$. Nous supposerons que

$$D_\infty := \bigcap_{n\in\mathbb{N}} D_n \neq \varnothing.$$

Notons $\mathcal{P}_\star^{(n)}$ les éléments de $\mathcal{P}^{(n)}$ qui sont inclus dans M_n.

Définition 2. La définition et la condition données ci-dessous ne sont pas tout à fait celle de [7], ce sont des versions “simplifiées” des conditions utilisées.
Nous dirons que deux éléments de $\mathcal{P}_\star^{(n)}$ sont *contigus* soit s'ils sont contigus dans le sens usuel, soit s'ils sont séparés par une composante connexe de $Y_n := \bigcup_{i=0}^{n-1} T^{-i}Y$.

Proposition II.4. *Si les Hypothèses 1.1 et 1.2 sont satisfaites, la suite*

$$\rho_n = \inf_{x \in D_n} \frac{\mathcal{L}^{n+1}\mathbf{1}(x)}{\mathcal{L}^n\mathbf{1}(x)}$$

converge vers $\rho > 0$.

Nous allons montrer que ρ est le rayon spectral de l'opérateur $\mathcal{L}$. Soit $c := e^{P_{top}(T,g^0)}$, c est aussi le rayon spectral de $\mathcal{L}_0$.

Hypothèse 1.3. Nous supposerons que le système vérifie :
Il existe des constantes $K \geq 0$ et $\xi \geq 1$, telles que pour tout $n \in \mathbb{N}$ il existe au plus $K\xi^n$ éléments de $\mathcal{P}_\star^{(n)}$ qui sont contigus. De plus, $\xi\Theta < \rho$.
Remarquons que cela implique, en particulier $\Theta < \rho$.

Théorème II.5. [7] *Supposons que les hypothèses 1.1, 1.2 et 1.3 sont satisfaites. Alors il existe une unique mesure de probabilité conditionnellement invariante* $\nu = hm$ *et absolument continue par rapport à* m. *Il existe une unique mesure de probabilité* μ *supportée par* $M_\infty := \bigcap_{n \in \mathbb{N}} M_n$ *telle que* $\mu(\mathcal{L}f) = \rho\mu(f)$, *avec* $\rho \leq c$, *pour toute fonction bornée* f. *La mesure* $\lambda = h\mu$ *est la seule mesure invariante par* T *supportée par* M_∞ *et absolument continue par rapport à* μ. *De plus, il existe* $\kappa < 1$ *tel que pour toute* $f \in BV$ *et tout* $A \in \mathcal{B}$:

$$\left\| \frac{\mathcal{L}^n f}{\rho^n} - h\mu(f) \right\|_\infty \leq Ct\kappa^n \|f\|_{BV},$$

$$|m(T^{-n}A|M_{n-1}) - \nu(A)| \leq Ct\kappa^n$$

$$\text{et } |\nu(A|M_{n-1}) - \lambda(A)| \leq Ct\kappa^n.$$

Dans [7], nous donnons des exemples d'applications de Lasota-Yorke à pour lesquelles nos conditions sont satisfaites : applications markoviennes avec trous non markovien ; β-applications ...

Comme sous produit du Théorème II.5 nous obtenons le résultat suivant sur la dimension de Hausdorff de l'ensemble M_∞ des survivants.
Pour $0 \leq t \leq 1$, définissons

$$\mathcal{L}_t f(x) = \sum_{Ty=x} (g^0)^t(y)\mathbf{1}_{M_0}(y)f(y)$$

et par Θ_t, ρ_t et $P(t)$ les nombres correspondants à Θ, ρ, P dans le cas $t = 1$.
Nous dirons que g^0 Vérifie la propriété de distorsion bornée s'il existe $C > 1$ tel que pour tout $n \in \mathbb{N}$, $P \in \mathcal{P}^{(n)}$ et x, $y \in Z$,

$$\frac{g_n^0(x)}{g_n^0(y)} \leq C. \tag{1.1}$$

Nous dirons que T est à grandes images si

$$\inf_{n\in\mathbb{N}} \inf_{P\in\mathcal{P}^{(n)}} m(T^n P) > 0. \tag{1.2}$$

Nous dirons que T est à grandes images par rapport à Ysi pour tout $n \in \mathbb{N}$, pour tout $P \in \mathcal{P}^{(n)}$, $P \cap M_\infty \neq \varnothing$, $T^n(P \cap M_{n-1}) \supset M_\infty$.

Théorème II.6. [7] *Soit $g^0 = \frac{1}{T'}$. Supposons que pour $0 \leq t \leq 1$, les hypothèses 1.1, 1.2 et 1.3 pour $g_t^0 = \left(\frac{1}{T'}\right)^t$ sont vérifiées. Alors, il existe un unique $0 < t_0 \leq 1$ tel que pour $0 \leq t < t_0$, $\rho_t > 1$ et pour $1 \geq t > t_0$, $\rho_t < 1$. Si T est à grandes images et à grandes images par rapport à Y alors la dimension de Hausdorff de M_∞ vérifie $HD(M_\infty) = t_0$.*

Ce résultat sur la dimension de Hausdorff de l'ensemble de Cantor M_∞ s'apparente aux résultats de Saussol et Wu ([SauW]) sur le spectre multifractal du temps de retour.

Remarquons que nos conditions 1.1, 1.2 et 1.3 permettent des trous relativement gros, nous n'utilisons pas d'arguments de perturbation. Néanmoins, une approche perturbative est possible et donne le résultat suivant.

Théorème II.7. [7] *Supposons que g^0vérifie l'hypothèse 1.1. Si l'application de Lasota-Yorke $T : I \to I$ admet une unique mesure de probabilité invariante μ_0 absolument continue par rapport à la mesure conforme m et que le système(I, T, μ_0) est mélangeant, alors il existe $\varepsilon > 0$ tel que pour tout trou Y, $m(Y) \leq \varepsilon$, les conclusions du Théorème II.5 sont vérifiées.*

Récemment, Demers ([Dem1, Dem2]) a considéré des "tours de Young à trous" et donné des hypothèses suffisantes sur ces tours qui garantissent l'existence de mesures conditionnellement invariantes absolument continues par rapport à la mesure de référence m. Ces techniques s'appliquent, par exemple, aux applications unimodales à trous.

1.2. **Récurrence et pression.** Dans le paragraphe précédent, nous avons étudié le temps d'entrée (ou le temps d'absorption) dans un ensemble fixé Y. Le théorème de récurrence de Poincaré est essentiel en théorie ergodique. Il implique que, sous une mesure de probabilité invariante, presque tous les points reviennent dans tout voisinage (infiniment souvent). Les premiers résultats quantitatifs dans ce domaine

arrivent avec Ornstein et Weiss ([OW]) et relient la récurrence à l'entropie du système.
Étant donné un système dynamique $(M, \mathcal{B}, T)$ inversible par morceaux, $\mathcal{P}$ la partition d'inversibilité de T et $\mathcal{P}^{(n)}$ la partition d'inversibilité de T^n. Nous supposerons que la partition $\mathcal{P}$ est génératrice et noterons $P_n(x)$ l'élément de $\mathcal{P}^{(n)}$ qui contient x. Le temps de retour d'ordre n est défini par :

$$\tau_n(x) = \min\{j \geq 1 : T^j(x) \in P_n(x)\}.$$

Si μ est une mesure de probabilité T invariante, Ornstein et Weiss ([OW]) on montré que pour μ presque tout x :

$$\lim_{n\to\infty} \frac{1}{n} \log \tau_n(x) = h_\mu,$$

où h_μ est l'entropie de la mesure μ définie dans notre cadre par :

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{A \in \mathcal{P}^{(n)}} \mu(A) \log(\mu(A)).$$

Les quantités :

$$\sum_{j=0}^{\tau_n(x)} \mu(P_n(T^j(x)))$$

donnent la masse des voisinages d'ordre n le long de l'orbite de x avant son retour dans $P_n(x)$. De telles quantités peuvent intervenir en théorie de l'information, notamment dans la compression de données. La question est d'obtenir le comportement asymptotique de ces quantités. En collaboration avec Mariusz Urbanski, Anna Zdunik et Bernard Schmitt, nous avons relié le comportement asymptotique de ces sommes à la pression du système dans le cadre des mesures de Gibbs ([6]).

Nous considérons M un sous décalage de type fini sur un alphabet fini, apériodique, g un potentiel Hölder et μ une mesure d'équilibre (i.e. μ est la mesure de probabilité invariante absolument continue par rapport à la mesure conforme ν). Rappelons que la pression topologique de système est donnée par :

$$P(\sigma, g) = \lim_{n\to\infty} \frac{1}{n} \log \left(\sum_{P \in \mathcal{P}^{(n)}} g^{(n)}(P) \right).$$

Dans le cas où $g \equiv 1$, on retrouve l'entropie topologique du système.

Théorème II.8. [6] *Soit $\sigma : M \to M$ un sous décalage de type fini, apériodique et $g : M \to \mathbb{R}^+$ un potentiel Hölder, μ l'unique mesure d'équilibre et $\mathcal{P}$ la partition de M en 1-cylindres. Alors pour μ presque* *tout* $x \in M$,

$$\lim_{n\to\infty} \frac{1}{n} \log \sum_{j=0}^{\tau_n(x)} g^{(n)} \circ T^j(x) = h_\mu + P_{top}(\sigma, g^2) - P_{top}(\sigma, g). \tag{1.3}$$

$$\lim_{n\to\infty} \frac{1}{n} \log \sum_{j=0}^{\tau_n(x)} \mu(P_n(T^j(x))) = h_\mu + P_{top}(\sigma, g^2) - 2P_{top}(\sigma, g). \tag{1.4}$$

Dans le cas où $g \equiv 1$, (1.3) est exactement le théorème de Orstein et Weiss. La preuve de ce théorème repose d'une part sur des inégalités simples comme Bienaimé-Chebitchev et d'autre part sur des résultats de grandes déviations pour la fonction "énergie libre" $c_{g,\mu}(t)$ de la mécanique statistique. Elle est définie par :

$$c_{g,\mu}(t) = \limsup_{n\to\infty} \frac{1}{n} \log \int g_n^t(x) d\mu(x).$$

Dans notre cas, l'énergie libre est liée à la pression par la relation : $c_{g,\mu}(t) = P_{top}(\sigma, g^{(1+t)}) - P_{top}(\sigma, g)$.

Dans le but d'estimer "empiriquement" (i.e. à partir de la seule donnée d'une orbite finie) la pression du système, il pourrait être intéressant de remplacer dans le Théorème $\mu(P_n(T^j(x)))$ par un estimateur. Ce genre de questions rentre tout à fait dans le cadre de la section suivante.

2. Inégalités exponentielles pour les systèmes dynamiques

L'objet de cette section est de présenter une problématique récente et en développement actuellement : les inégalités exponentielles pour les systèmes dynamiques. Ce type de questions est aussi étudié en statistiques pour les processus faiblement dépendants. Nous allons voir que les systèmes dynamiques peuvent être vus comme un cas particulier de processus faiblement dépendants. Notre intérêt dans cette thématique est d'utiliser ces inégalités pour estimer certaines fonctionnelles liées aux systèmes dynamiques.

Informellement, pour un processus $(X_0, \ldots, X_n, \ldots)$ et φ une fonction de $\mathbb{R}^n \to \mathbb{R}$, les inégalités de concentration mesurent l'écart de $\varphi(X_1, \ldots, X_n)$ à son espérance. Si la borne est exponentielle, on parle d'inégalités exponentielles :

$$\mathbb{P}(|\varphi(X_1, \ldots, X_n) - \mathbb{E}(\varphi(X_1, \ldots, X_n))| > t) \le \exp(-t^2 r(n) K(\varphi)).$$

2.1. **Inégalités exponentielles pour les processus faiblement dépendants.** Il existe diverses définitions de mélange faible pour un processus, notamment introduites par Doukhan-Louhichi ([DouL], voir aussi [Dou]). Ces définitions reposent sur les quantités :

$$c(n) = \sup\{|\mathbb{E}(Yf(X_{i+n})) - \mathbb{E}(Y)\mathbb{E}(f(X_{i+n}))| \ i \in \mathbb{N}, \ Y \text{ est } \mathcal{M}_i - \text{mesurable et } \|Y\|_1 \leq 1\},$$

avec f plus ou moins régulière. Les notions de dépendance faible supposent en général que les coefficients de mélange $c(n)$ ci-dessus sont sommable.

Dans la partie “Quelques définitions”, section 3, nous avons défini une notion de mélange par rapport à un espace de Banach de fonctions bornées $\mathcal{C}$ dont la norme est de la forme $\|f\|_{\mathcal{C}} = C(f) + \| \ \|$, où $C(\ )$ est une semi norme sur $\mathcal{C}$ et $\| \ \|$ est une norme sur $\mathcal{C}$; $\mathcal{C}_1$ est l'ensemble des fonctions de $\mathcal{C}$ telles que $C(f) \leq 1$. Les coefficients de $\Phi_{\mathcal{C}}$-mélange sont les coefficients $c(n)$ ci-dessus avec $f \in \mathcal{C}_1$ (voir page 8). Dans la suite, nous considérerons que $\mathcal{C}$ est un espace de fonctions bornées.

Suivant les travaux de Dedecker et Prieur, nous obtenons l'inégalité exponentielle - de type Hoeffding - suivante pour les processus $\Phi_{\mathcal{C}}$ faiblement dépendants.

Proposition II.9. *(*[DedPr], [2] *pour la généralisation à un espace de Banach $\mathcal{C}$). Soit $(X_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires satisfaisant le $\Phi_{\mathcal{C}}$-mélange. Pour $\varphi \in \mathcal{C}$, soit $S_n(\varphi) = \sum_{i=1}^{n} \varphi(X_i)$. On a l'inégalité :*

$$\mathbb{P}\left(|S_n(\varphi) - \mathbb{E}(S_n(\varphi))| > t\right) \leq e^{\frac{1}{e}} \exp\left(\frac{-t^2}{2e(C(\varphi))^2 \sum_{k=0}^{n-1}(n-k)\Phi_{\mathcal{C}}(k)}\right).$$

On peut obtenir des résultats de ce type pour $\varphi \ : \mathbb{R}^n \to \mathbb{R}$ (à la place de $S_n(\varphi)$) une fonction Lipschitzienne sous diverses hypothèses de mélange faible. De tels résultat ont été obtenus notamment par Rio ([Rio]), Dedecker-Prieur ([DedPr]). D'autres types d'inégalité (de type Bernstein) qui font intervenir la variance de $S_n(\varphi)$, φ fonction Lipschtizienne, ont été démontrées récemment par Kallabis-Neumann ([KaNe]) et Doukhan-Neumann ([DouNe]).

Dans le cadre des systèmes dynamiques, on peut bien sûr considérer le processus $X_i = T^i$. Comme nous l'avons déjà remarqué dans la partie I, il s'avère qu'en général ce processus n'est pas $\Phi_{\mathcal{C}}$ mélangeant mais que c'est le “processus à temps inversé” $(Y_i)_{i\in\mathbb{N}}$, défini par

$$(Y_0, \ldots, Y_n) \overset{\text{Loi}}{=} (X_n, \ldots, X_0)$$

qui est susceptible de l'être. Dans la section 2.3, nous appliquerons le résultat ci-dessus au processus à temps inversé. Une autre stratégie est possible : on peut travailler directement sur le système dynamique et obtenir des inégalités exponentielles pour des fonctions $\varphi \ : \mathbb{R}^n \to \mathbb{R}$, Lipschitziennes, pour des systèmes dynamiques dilatants par morceaux ([ColMSc1]) et pour certains systèmes dynamiques non uniformément hyperboliques ([ChaColSc]).

2.2. **Typicité et typicité conditionnelle.** L'inégalité de concentration ci-dessus permet d'aborder les questions de "typicité" et "typicité conditionnelles" pour des processus faiblement dépendants. Nous suivons ici les travaux de Flajolet-Kirschenhofer-Tichy ([FlKT] dans le cadre indépendant) et Csiszár-Shield ([Cs, CsS] pour des processus à valeur dans un alphabet fini et Φ mélangeants).
La problématique est la suivante. Étant donné X_0, ..., X_n, ... un processus aléatoire à valeur dans un espace complet M, $(\mathcal{P}_k)_{k\in\mathbb{N}}$ une suite de partitions de M, on recherche des estimations empiriques des probabilités $\mathbb{P}(X_j \in P)$, $P \in \mathcal{P}_k$ et des probabilités conditionnelles $\mathbb{P}(X_j \in P \mid X_{j-1} \in \widetilde{P})$, $P \in \mathcal{P}_k$ (avec $X_j \in P \overset{\text{p.s.}}{\Rightarrow} X_{j-1} \in \widetilde{P}$). Il est naturel de considérer les estimateurs suivants :

$$N_i^\ell(P) = \sum_{j=i}^{\ell} \mathbf{1}_P(X_j), \text{ et } \hat{g}_n(P) = \frac{N_1^{n+1}(P)}{N_0^{n-1}(\widetilde{P})}\frac{n}{n+1}.$$

On recherche des informations quantitatives sur la proximité entre les probabilités empiriques $\frac{N_i^{i+n}(P)}{n}$ et l'espérance $\mathbb{E}_i^{i+n}(P) := \mathbb{E}\left(\frac{N_i^{i+n}(P)}{n}\right)$ et entre $\hat{g}_n(P)$ et $\mathbb{E}_n(P|\widetilde{P}) := \dfrac{\mathbb{E}_1^{n+1}(P)}{\mathbb{E}_0^{n-1}(\widetilde{P})}$.
Remarquons que, dans le cas stationnaire, $\mathbb{E}_i^{i+n}(P) = \mathbb{P}(X_j \in P)$ et $\mathbb{E}_n(P|\widetilde{P}) = \mathbb{P}(X_j \in P \mid X_{j-1} \in \widetilde{P})$.
Nous portons une attention particulière sur la "typicité de grande échelle", c'est à dire que k croît vers l'infini avec n.
Nous dirons que la suite $(\mathcal{P}_k)_{k\in\mathbb{N}}$ de partitions dénombrables de Σ est *emboîtée* si pour tout $j, k \in \mathbb{N}$,

$$\forall P \in \mathcal{P}_k\ \exists!\ \widetilde{P} \in \mathcal{P}_{k-1}\ ,\ X_j \in P\ \Rightarrow X_{j-1} \in \widetilde{P}.$$

Théorème II.10 (Typicité et typicité conditionnelle)**.** [2] *Soit* $(X_p)_{p\in\mathbb{N}}$ *une suite de variables aléatoires satisfaisant le* $\Phi_{\mathcal{C}}$*-mélange, si* $(\mathcal{P}_k)_{k\in\mathbb{N}}$ *est une suite de partitions emboîtées telle que pour tout* $P \in \mathcal{P}_k$, $\mathbf{1}_P \in \mathcal{C}$. *Il existe* $C > 0$ *tel que, pour tout* $P \in \mathcal{P}_k$ *on a :*

$$\mathbb{P}\left(\left|\frac{N_i^{n+i}(P)}{n} - \mathbb{E}_i^{n+i}(P)\right| > t\right) \le e^{\frac{1}{e}} e^{\left(-\frac{Ct^2 n}{C(\mathbf{1}_P)^2}\right)}.$$

si, de plus, pour $0 < \varepsilon < 1$ *on a* $\dfrac{\mathbb{E}_0^{n-1}(\widetilde{P})}{C(\mathbf{1}_P)}$ *et* $\dfrac{\mathbb{E}_0^{n-1}(\widetilde{P})}{C(\mathbf{1}_{\widetilde{P}})} \ge n^{-\frac{\varepsilon}{2}}$, *alors il existe* $K > 0$ *tel que :*

$$\mathbb{P}\left(\left|\hat{g}_n(P) - \mathbb{E}_n(P|\widetilde{P})\right| > t\right) \le 4e^{-Kt^2 n^{1-\varepsilon}} + 2e^{-Kn^{1-\varepsilon}},$$

si la suite est stationnaire :

$$\mathbb{P}\left(\left|\hat{g}_n(P) - \mathbb{P}(X_1 \in P \mid X_0 \in \widetilde{P})\right| > t\right) \le 4e^{-Kt^2 n^{1-\varepsilon}} + 2e^{-Kn^{1-\varepsilon}}.$$

Ce résultat de typicité conditionnelle va nous permettre de construire un estimateur empirique de la fonction de pression g de certains systèmes dynamiques.

2.3. **Reconstruire les systèmes dynamiques ?** Dans ce paragraphe, nous montrons comment des estimations statistiques fines peuvent permettre d'estimer certaines fonctionnelles liées aux systèmes dynamiques. Le but ultime serait, étant donné un morceau d'orbite finie provenant d'un système dynamique, de reconstruire le système dynamique, ceci dans un but de modélisation et simulation.
Considérons un système dynamique (Σ, T, μ). Σ est un espace complet, $T : \Sigma \to \Sigma$ est une application mesurable, μ est une mesure de probabilité sur Σ, T-invariante. Soit $\mathcal{C}$ un espace de Banach de fonctions sur Σ. On suppose que le système dynamique vérifie la propriété de mélange suivante : pour tout $\varphi \in L^1(\mu)$, $\psi \in \mathcal{C}$,

$$\left| \int_\Sigma \psi \cdot \varphi \circ T^n d\mu - \int_\Sigma \psi d\mu \int_\Sigma \varphi d\mu \right| \leq \Phi(n) \|\varphi\|_1 \|\psi\|_{\mathcal{C}}, \tag{2.1}$$

avec $\Phi(n)$ sommable.
Soit $X_j = T^j$, $j \in \mathbb{N}$. On peut réécrire la condition de mélange : pour $\varphi \in L^1(\mu)$, $\psi \in \mathcal{C}$,

$$|\mathrm{Cov}(\psi(X_0), \varphi(X_n))| \leq \Phi(n) \|\varphi\|_1 \|\psi\|_{\mathcal{C}}.$$

Si la norme sur $\mathcal{C}$ est telle que pour tout $\psi \in \mathcal{C}$, il existe un nombre réel $R(\psi)$ tel que $\|\psi + R(\psi)\| \leq C(\psi)$, alors, en utilisant la stationarité de la suite (X_j), on a pour tout $i \in \mathbb{N}$, pour $\psi \in \mathcal{C}_1$, $\varphi \in L^1$, $\|\varphi\|_1 \leq 1$,

$$|\mathrm{Cov}(\psi(X_i), \varphi(X_{n+i}))| \leq 2\Phi(n).$$

On se donne une partition dénombrable de Σ. Soit $\mathcal{P}_k$ la partition dénombrable de Σ dont les atomes sont : pour $i_0, \ldots, i_{k-1}$,

$$A_{i_0}^{i_{k-1}} = \{x \in \Sigma \ / \ \text{for } j = 0, \ldots, k-1, \ T^j(x) \in A_{i_j}.\}$$

Évidement, si $X_j \in A_{i_0}^{i_{k-1}}$ alors $X_{j+1} \in A_{i_1}^{i_{k-1}}$.
On supposera que pour tout k-uplets $i_0, \ldots, i_{k-1}$, $f = \mathbf{1}_{A_{i_0}^{i_{k-1}}} \in \mathcal{C}$ et on notera $C(i_{i_0}, \ldots, i_{k-1}) = C(f)$.
Ainsi, pour un processus à temps renversé, la condition de $\Phi_{\mathcal{C}}$-mélange est satisfaite et la suite de partitions $\mathcal{P}_k$ est emboîtée.
Plus précisément, si $(Y_n)_{n\in\mathbb{N}}$ est tel que $(Y_n, \cdots, Y_0)$ à la même loi que $(X_0, \cdots, X_n)$ pour tout n, alors $\mathrm{Cov}(\psi(X_i), \varphi(X_{n+i})) = \mathrm{Cov}(\psi(Y_{i+n}), \varphi(Y_i))$ et le processus $(Y_n)_{n\in\mathbb{N}}$ est $\Phi_{\mathcal{C}}$-mélangeant, de plus, $Y_j \in A_{i_0}^{i_{k-1}}$ implique $Y_{j-1} \in A_{i_1}^{i_{k-1}}$ presque sûrement.

Le résultat sur la typicité conditionnelle s'applique à $(Y_n)_{n\in\mathbb{N}}$; en utilisant :

$$\sum_{j=0}^{n-2} \mathbf{1}_{\widetilde{P}}(Y_j) \overset{\text{en loi}}{=} \sum_{j=0}^{n-2} \mathbf{1}_{\widetilde{P}}(X_j),$$

on obtient :

Théorème II.11. [2] *Soit* $\hat{g}_n(A_{i_0}^{i_{k-1}}) = \frac{N_0^n(A_{i_0}^{i_{k-1}})}{N_0^{n-1}(A_{i_1}^{i_{k-1}})}\frac{n-1}{n}$*, il existe* $K > 0$ *tel que pour tout* $\varepsilon < 1$*, si*

$$\frac{\mu(A_{i_1}^{i_{k-1}})}{C(i_0,\dots,i_{k-1})} \ \textit{et}\ \frac{\mu(A_{i_1}^{i_{k-1}})}{C(i_1,\dots,i_{k-1})} \geq n^{-\frac{\varepsilon}{2}},$$

alors :

$$\mathbb{P}\left(\left|\hat{g}_n(A_{i_0}^{i_{k-1}}) - \mathbb{P}(X_0 \in A_{i_0}^{i_k} | X_1 \in A_{i_1}^{i_{k-1}})\right| > t\right) \leq 4e^{-Kt^2n^{1-\varepsilon}} + 2e^{-Kn^{1-\varepsilon}}.$$

Remarque. Si $\sup C(\mathbf{1}_{A_j}) < \infty$ alors il existe $\gamma < 1$ tel que pour tout $P \in \mathcal{P}_k$, $\mu(P) \leq \gamma^k$ (voir [Pac, 2]).
Ainsi, si k tend vers l'infini avec n, pour que la condition sur $\mu(P)$ soit satisfaite, il est nécessaire que $k = O(\ln n)$.

Nous allons spécifier les résultats obtenus pour des applications dilatantes de l'intervalle. C'est à dire, T est une application monotone par morceaux de l'intervalle, on note $A_1,\dots,A_\ell$ la partition finie sur laquelle T est monotone. Soit m la mesure de Lebesgue sur l'intervalle.

Hypothèse 2.1.
(1) la restriction de T à chaque $\overline{A_j}$ est une bijection C^2 de $\overline{A_j}$ dans $T(\overline{A_j}) =: B_j$.
(2) T est dilatante, c'est à dire qu'il existe $1 < \theta^{-1}$ tel que pour tout $x \in I$, $\theta^{-1} \leq |T'(x)|$.
(3) si T est une application markovienne (c'est à dire que pour tout j, B_j est une union de certains A_i), on suppose que l'application est apériodique : il existe $N \in \mathbb{N}$ tel que pour tout $i, j = 1,\dots,\ell$, pour tout $n \geq N$,
$$T^{-n}A_i \cap A_j \neq \varnothing.$$
(4) si T n'est pas nécessairement markovienne, on suppose que l'application est "couvrante" : pour tout $k \in \mathbb{N}$, il existe $N(k)$ tel que pour tout $P \in \mathcal{P}_k$,
$$T^{N(k)}P = [0,1].$$

Remarque. Les trois conditions (1-2)-(3 ou 4) sont suffisantes pour assurer l'existence et l'unicité d'une mesure invariante absolument continue par rapport à la mesure de Lebesgue : $\mu = hm$; ainsi qu'une estimation de la décroissance des corrélations pour des fonctions à variations bornées (voir [Broi, Col, Li3] ...).

Comme précédemment, on considère le processus à temps inversé $(Y_i)_{i\in\mathbb{N}}$ associé à $(X_i)_{i\in\mathbb{N}}$. Le processus $(Y_i)_{i\in\mathbb{N}}$ est une chaîne de Markov de noyau $\mathcal{L}$ ([BGR]) :

$$\mathcal{L}(f)(x) = \sum_{y/T(y)=x} g(y)f(y),$$

où g est le potentiel du système : $g = \frac{h}{|T'|\cdot h\circ T}$,
g^{-1} est aussi la dérivée de Radon-Nicodym de $\mu\circ T$ par rapport à μ.
$\mathcal{L}$ est aussi l'opérateur "transformateur de densité" normalisé ($\mathcal{L}(\mathbf{1}) = \mathbf{1}$). Les résultats précédents permettent d'estimer empiriquement la fonction g. L'espace $\mathcal{C}$ est l'espace des fonctions à variations bornées, les partitions $\mathcal{P}_k$ sont engendrées par la partition $A_1, \ldots, A_\ell$.
Soit $A_k(x)$l'élément $A_{i_0}^{i_{k-1}}$ de $\mathcal{P}_k$ qui contient x, soit $\hat{g}_{n,k}(x) = \hat{g}_n(A_k(x))$.

Théorème II.12. *Pour tout $\kappa > \gamma$ et $\kappa \geq \theta$, il existe D_k et E_k des unions finies d'éléments de $\mathcal{P}_k$ tels que $\mu(D_k) \leq 2\ell\gamma^k + a\left(\frac{\gamma}{\kappa}\right)^k$, $\mu(E_k) \leq \gamma^k$ et il existe $L > 0$ tel que si*
- *$x \notin D_k \cup E_k$,*
- *$\dfrac{\ln(\frac{t}{2K})}{\ln(\kappa)} \leq k \leq \dfrac{\varepsilon}{2}\dfrac{\ln 2n}{\ln(\frac{\ell}{\gamma})}$*

alors $\mathbb{P}(|\hat{g}_{n,k}(x) - g(x)| > t) \leq 4e^{-Lt^2n^{1-\varepsilon}} + 2e^{-Ln^{1-\varepsilon}}$.

Idée de la preuve. Les résultats précédents appliqués aux applications dilatantes de l'intervalle donnent :

$$\mu\left(\left|\hat{g}_{n,k}(x) - \frac{\mu(A_k(x))}{\mu(A_{k-1}(x))}\right| > t\right) \leq 4e^{-Kt^2n^{1-\varepsilon}} + 2e^{-Kn^{1-\varepsilon}},$$

à condition que $\mu(A_k(x))$ ne soit pas trop petit (remarquons que $A_k(x)$ est un intervalle et donc sa variation vaut 2). Maintenant, comme le montrent les deux lemmes suivants, $\frac{\mu(A_k(x))}{\mu(A_{k-1}(x))}$ est "presque" $g(x)$.

Lemme II.13. *Pour T une application dilatante et markovienne de l'intervalle, il existe $K > 0$ tel que pour tout $k \in \mathbb{N}$, pour tout $x \in I$,*

$$(1 - K\theta^k)g(x) \leq \frac{\mu(A_k(x))}{\mu(A_{k-1}(Tx))} \leq (1 + K\theta^k)g(x).$$

Dans le cas où l'application n'est pas markovienne, l'encadrement n'est pas aussi bon mais suffit pour obtenir le résultat.

Lemme II.14. *Pour T une application dilatante de l'intervalle (pas nécessairement markovienne), il existe $K > 0$ tel que pour tout $k \in \mathbb{N}$, pour tout $\kappa > \gamma$ et $\kappa \geq \theta$,*

$$\mu\left\{x \in I \;/\; (1 - K\kappa^k)g(x) \leq \frac{\mu(A_k(x))}{\mu(A_{k-1}(Tx))} \leq (1 + K\kappa^k)g(x)\right\}$$
$$\geq 1 - (2\ell\gamma^k + a\left(\frac{\gamma}{\kappa}\right)^k).$$

L'ensemble D_k est donné par le lemme ci-dessus. L'ensemble E_k est l'ensemble des x pour lesquels $A_k(x)$ est trop petit. $\square$

Un corollaire du théorème précédent est la convergence dans L^p, $\forall p \geq 1$ de $\hat{g}_n = \hat{g}_{n,k(n)}$ vers g, où $k(n)$ est une suite croissante d'entiers telle que

$$\alpha \frac{\ln n}{\ln \frac{1}{\kappa}} \leq k(n) \leq \frac{\varepsilon}{2} \frac{\ln 2n}{\ln(\frac{\ell}{\gamma})},$$

avec $\kappa > \gamma$ et $\kappa \geq \theta$ fixés et $\alpha < \dfrac{\varepsilon}{2} \dfrac{\ln \frac{1}{\kappa}}{\ln \frac{\ell}{\gamma}}$.

Corollaire II.15. *Soit $\hat{g}_n = \hat{g}_{n,k}$, alors pour tout $p \geq 1$, $\hat{g}_n$ converge vers g dans $L^p(\mu)$.*

Remarquons, pour conclure cette section, que pour obtenir le Théorème II.12 et le Corollaire II.15, outre la propriété de mélange (2.1), l'ingrédient essentiel est le Lemme II.14. Celui-ci est vérifié dans un cadre beaucoup plus large que les applications dilatantes de l'intervalle. Cette généralisation est l'objet d'un travail en cours [17].

Troisième partie III. Systèmes dynamiques et analyse d'algorithmes

L'utilisation des systèmes dynamiques en algorithmique a été largement développée par B. Vallée (voir [19] pour une synthèse).
D'une part, les systèmes dynamiques fournissent un cadre général à l'étude des sources rencontrées en théorie de l'information qui englobe les sources "classiques" (indépendantes ou markoviennes). On entend par *source* un mécanisme qui génère des suites (infinies) aléatoires, appelées *mots*, d'éléments d'un alphabet dénombrable ; les systèmes dynamiques de l'intervalle monotones par morceaux fournissent un tel mécanisme. Pour les sources définies par certains systèmes dynamiques analytiques de l'intervalle, B. Vallée ([Va1]) a relié les propriétés statistiques des mots émis aux propriétés spectrales de l'opérateur de transfert associé. En collaboration avec Frédéric Chazal et Brigitte Vallée, nous avons généralisé ces résultats à des classes plus larges de systèmes dynamiques non analytiques ([5, 4]). De tels systèms dynamiques peuvent aider à la modélisation dans des cas où les sources "classiques" ne reflètent pas une réalité concrète.
D'autre part, c'est une idée naturelle de considérer un algorithme et l'ensemble de ses données comme un système dynamique (discret) que l'on plonge dans un système dynamique continu. Les opérateurs de transfert associés au système dynamique permettent alors l'analyse des séries génératrices utilisées classiquement pour l'analyse en moyenne d'algorithmes : on cherche à déterminer un comportement asymptotique moyen de l'algorithme (nombre moyen d'itérations, coût moyen ...). Ces techniques permettent aussi, dans certains cas, d'obtenir l'analyse en distribution de l'algorithme : distribution asymptotique du nombre d'itérations par exemple. Ce point de vue d'analyse dynamique s'est révélé très efficace dans l'analyse en moyenne et en distribution de certains algorithmes arithmétiques (essentiellement d'algorithmes euclidiens), voir notamment les articles de Brigitte Vallée ([Va2, Va3, Va4]) et Brigitte Vallée et Viviane Baladi ([BaVa]). En collaboration avec Benoit Daireaux et Brigitte Vallée, nous exploitons ces techniques pour l'analyse en moyenne d'algorithmes de division sur les bits de poids faibles ([1]).

1. Outils et techniques

L'analyse en moyenne d'algorithmes vise à déterminer le comportement "moyen" des algorithmes. Cette analyse, souvent plus réaliste que l'analyse dans le pire des cas, est aussi plus difficile à mettre en œuvre. L'outil classique pour ce type d'analyse est celui des séries génératrices. À l'aide d'outils d'analyse complexe comme les théorèmes taubériens, on relie les singularités des séries génératrices au comportement asymptotique moyen de l'algorithme. Cette méthodologie est décrite par exemple dans les livres de Flajolet et Sedgewick ([Fl1, FlSe]). Des théorèmes probabilistes comme le théorème des quasi-puissances de Hwang ([Hw]) permettent alors d'obtenir aussi des informations sur la distribution, par exemple, du nombre d'itérations de l'algorithme.

Lorsque les algorithmes sont trop "corrélés" l'analyse des singularités des séries génératrices peut s'avérer impossible directement. L'idée est alors de plonger l'algorithme dans un système dynamique et d'utiliser l'opérateur de transfert associé à ce système dynamique pour analyser les séries génératrices. Cette stratégie a été formalisée par Brigitte Vallée ([Va1, Va2, Va3, Va4]) et s'est avérée particulièrement efficace dans deux domaines de l'algorithmique : l'algorithmique du texte et l'analyse d'algorithmes arithmétiques. Avant de détailler nos contributions dans ces deux domaines, rappelons les définitions et résultats classiques relatifs aux séries génératrices.

1.1. **Séries génératrices.** Une des techniques les plus standard en analyse d'algorithmes et en combinatoire analytique est, pour étudier une famille d'objets $\mathcal{A}$ munie d'une taille $\ell(a)$, $a \in \mathcal{A}$, d'utiliser les séries génératrices. Nous allons considérer des séries génératrices entières et de Dirichlet. Les séries entières les plus simples permettent essentiellement d'étudier la taille des éléments de $\mathcal{A}$:

$$A(z) = \sum_{a \in \mathcal{A}} z^{\ell(a)} = \sum_{n \geq 1} a_n z^n,$$

si a_n est le nombre d'éléments de taille n. Si on s'intéresse à un paramètre particulier défini sur $\mathcal{A}$, par exemple si un coût $c(a)$ est associé à chaque élément $a \in \mathcal{A}$, on considère la série entière bivariée :

$$A(z, u) = \sum_{a \in \mathcal{A}} z^{\ell(a)} u^{c(a)},$$

ou encore

$$\widetilde{A}(z, w) = \sum_{a \in \mathcal{A}} z^{\ell(a)} \exp(wc(a)).$$

Les séries de Dirichlet sont aussi très utiles dans le but d'une analyse probabiliste. Ces séries sont définies par :

$$F(s, w) = \sum_{a \in \mathcal{A}} \frac{\exp(wc(a))}{\ell(a)^s} = \sum_{n \geq 1} \frac{f_n^c(w)}{n^s},$$

où $f_n^c(w)$ est la somme cumulée

$$f_n^c(w) = \sum_{\substack{a \in \mathcal{A} \\ \ell(a) = n}} \exp(wc(a)).$$

Si on considère c comme une variable aléatoire sur le sous ensemble $\mathcal{A}_n$ supposé fini des éléments de $\mathcal{A}$ de taille inférieure où égale à n muni de l'équiprobabilité, l'espérance de $\exp(wc)$ est :

$$\mathbb{E}_n[\exp(wc)] = \frac{\sum_{p=1}^{n} f_p^c(w)}{|\mathcal{A}_n|},$$

ou en terme de série entière :

$$\mathbb{E}_n[\exp(wc)] = \frac{[z^n]\widetilde{A}(z,w)}{|\mathcal{A}_n|},$$

où $[z^n]\widetilde{A}(z,w)$ désigne le $n$ième coefficient de la série.
Lorsque l'on connaît bien les moments exponentiels d'une variable aléatoire, on peut obtenir un théorème de la limite centrale avec une expression de la moyenne et de la variance asymptotique pour la variable c sur $\mathcal{A}_n$ (voir le théorème de Hwang dans la section suivante).
Les séries bivariées ci-dessus qui permettent une analyse en distribution sont parfois difficiles à analyser. Dans ce cas, on peut considérer d'autres séries univariées, plus simples qui permettent l'analyse en moyenne d'une fonction coût :

$$T_c(s) = \sum_{a \in \mathcal{A}} \frac{c(a)}{\ell(a)^s} = \sum_{n \geq 1} \frac{t_n^c}{n^s},$$

où t_n^c est la somme cumulée :

$$t_n^c = \sum_{\substack{a \in \mathcal{A} \\ \ell(a)=n}} c(a).$$

L'espérance de c sur $\mathcal{A}_n$ est alors donnée par :

$$\mathbb{E}_n(c) = \frac{\sum_{p=1}^{n} t_n^c}{\sum_{p=1}^{n} t_n^1},$$

si t_n^1 est associé au coût constant égal à 1.
De même pour les séries entières :

$$A_c(z) = \sum_{a \in \mathcal{A}} z^{\ell(a)} c(a) = \sum_{n \geq 1} t_n^c z^n,$$

et

$$\mathbb{E}_n(c) = \frac{[z^n]A_c(z)}{[z^n]A_1(z)}.$$

L'analyse en moyenne du coût c vise à donner un équivalent de $\mathbb{E}_n(c)$.
Ainsi, l'analyse en moyenne ou en distribution de la fonction de coût c se ramène à l'extraction des coefficients des séries de Dirichlet ou des séries entières.
Dans le paragraphe suivant, nous rappelons les outils d'analyse complexe qui permettent cette extraction. Nous énonçons aussi le théorème de Hwang.

1.2. **Outils d'analyse complexe et probabilistes.** Nous donnons deux formes d'un théorème taubérien dû à Delange ([Del]). L'une des formes relie le comportement asymptotique des coefficients d'une série de Dirichlet aux singularités de la série. L'autre relie le comportement asymptotique d'une fonction positive croissante aux singularités de sa transformée de Laplace.

Théorème A. [Del] *Soit $F(s)$ une série de Dirichlet à coefficients positifs $(a_n)_{n\in\mathbb{N}^\star}$:*

$$F(s) = \sum_{n\geq 1} \frac{a_n}{n^s}.$$

On suppose que

(1) $F(s)$ converge dans le demi-plan $\Re(s) > \sigma > 0$ et est analytique pour $\Re(s) = \sigma$, $s \neq \sigma$,

(2) il existe $\gamma \geq 0$ tel que $F(s) = A(s)(s-\sigma)^{-\gamma-1} + C(s)$ où A et C sont analytiques en σ et $A(\sigma) \neq 0$.

Alors, lorsque $N \to \infty$,

$$\sum_{1\leq n\leq N} a_n = \frac{A(\sigma)}{\sigma\Gamma(\gamma+1)} N^\sigma \log^\gamma N[1+\varepsilon(N)],\ \ \varepsilon(N) \to 0.$$

Théorème B. [Del] *Soit $V(s)$ une fonction qui admet dans le demi-plan $\Re(s) > \sigma > 0$ la représentation intégrale*

$$V(s) = s \int_0^\infty A(y) e^{-sy} dy$$

où A est positive et croissante. Supposons que $V(s)$ est analytique sur $\Re(s) = \sigma$, $s \neq \sigma$ et que pour $\gamma \geq 0$,

$$V(s) = \frac{g(s)}{(s-\sigma)^{\gamma+1}} + l(s),$$

où g, l sont analytiques en σ et $g(\sigma) \neq 0$. Alors

$$A(x) = \frac{g(\sigma)}{\sigma\Gamma(\gamma+1)} e^{x\sigma} x^\gamma (1+\varepsilon(x)), \quad \textit{quand} \ \ x \to \infty.$$

Le cas des séries entières est beaucoup plus simple puisqu'il suffit d'appliquer le théorème de Cauchy pour extraire les coefficients. L'utilisation des séries entières est aussi plus fructueuse puisqu'elle s'applique aussi aux série bivariées.

Théorème C (Théorème des quasi-puissances)**.** [Hw] *Soit R_n une suite de variables aléatoires dont la série génératrice des moments $\mathbb{E}[\exp(wR_n)]$ est analytique dans un voisinage complexe $\mathcal{W}$ de $w = 0$, et satisfait :*

$$\mathbb{E}[\exp(wR_n)] = \exp[\beta_n U(w) + V(w)] \left(1 + O(\kappa_n^{-1})\right), \tag{1.1}$$

avec β_n, $\kappa_n \to \infty$ quand $n \to \infty$, $U(w)$, $V(w)$ sont analytiques sur $\mathcal{W}$ et le terme en O est uniforme sur $\mathcal{W}$. Alors l'espérance et la variance de R_n vérifient :

$$\mathbb{E}[R_n] = U'(0)\cdot\beta_n + V'(0) + O(\kappa_n^{-1}), \quad Var[R_n] = U''(0)\cdot\beta_n + V''(0) + O(\kappa_n^{-1}).$$

De plus, si $U''(0) \neq 0$, la distribution de R_n est asymptotiquement Gaussienne avec vitesse de convergence : $O(\kappa_n^{-1} + \beta_n^{-1/2})$:

$$\mathbb{P}\left[x \;\middle|\; \frac{R_n(x) - U'(0)n}{\sqrt{U''(0)n}} \leq Y\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y} e^{-y^2/2}\, dy + O(\kappa_n^{-1} + \beta_n^{-1/2}).$$

1.3. **Opérateurs de transfert généralisés.** Dans les deux sections suivantes, nous allons appliquer la stratégie d'analyse décrite sommairement en introduction ci-dessus. Nous allons exprimer les séries génératrices et/ou les transformées de Laplace des paramètres de type "coût moyen" d'un algorithme en terme d'opérateurs de transfert généralisés associés à un système dynamique inversible par morceaux.

Considérons un système dynamique (M, T) inversible et différentiable par morceaux, notons $\mathcal{P}$ la partition (au plus dénombrable) d'inversibilité. Soit m une mesure de référence par rapport à laquelle T est non-singulière - par la suite m sera soit la mesure de Lebesgue, soit la mesure de Haar sur un groupe compact, dans ces deux cas, m est une mesure 1-conforme pour le potentiel $|T'|^{-1}$. L'opérateur de transfert s'écrit :

$$\mathcal{L}f(x) = \sum_{T(y)=x} |T'(y)|^{-1} f(y)$$

ou encore, si on note $\mathcal{H}$ l'ensemble des branches inverses de T : $\mathcal{H} = \{h_P,\ P \in \mathcal{P}\}$, $h_P\ : Y_{h_P} \to\ P$, avec $Y_{h_P} = T(P)$ et h_P est l'inverse de $T_{|P}$,

$$\mathcal{L}f(x) = \sum_{h\in\mathcal{H}} |h'(x)| f \circ h(x) \mathbf{1}_{Y_h}(x).$$

Comme nous l'avons déjà remarqué, les propriétés spectrales de cet opérateur sont liées aux propriétés ergodiques du système dynamique. En particulier les densités invariantes sont données par les fonctions propres (pour la valeur propre 1), la vitesse de décroissance des corrélations dépend de la vitesse de convergence des itérés de $\mathcal{L}$ vers la projection sur l'espace propre associé à 1.

Dans le but d'obtenir les propriétés asymptotiques et d'analyticité nécessaires à l'application des Théorèmes A, B et C, nous allons considérer des familles à paramètres (complexes) d'opérateurs de transfert généralisés. Il s'agit en fait d'une généralisation des opérateurs de transfert considérés par Ruelle [Ru] dans le cadre du formalisme thermodynamique.

Les opérateurs qui permettront d'engendrer les séries de Dirichlet les plus simples seront :

$$\mathcal{L}_s f(x) = \sum_{h\in\mathcal{H}} |h'(x)|^s f\circ h(x)\mathbf{1}_{Y_h}(x).$$

Si on associe à chaque branche inverse du système un coût $c(h)$, $h\in\mathcal{H}$, les opérateurs permettant de générer les séries génératrices des coûts seront :

$$\mathcal{L}_{s,w} f(x) = \sum_{h\in\mathcal{H}} |h'(x)|^s\cdot\exp(wc(h))\cdot f\circ h(x)\mathbf{1}_{Y_h}(x).$$

Enfin, nous utiliserons aussi les opérateurs de transfert "sécants" qui agissent sur des fonctions sur $M\times M$:

$$\mathbf{L}_s[f](u,v) := \sum_{h\in\mathcal{H}} H_h^s(u,v)f(h(u),h(v))\mathbf{1}_{Y_h\times Y_h}(u,v),$$

où les H_h sont les sécantes des branches inverses :

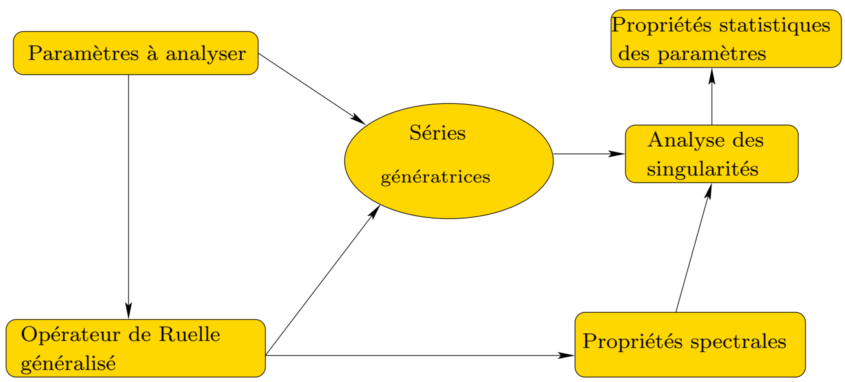
$$H_h(u,v) := \left|\frac{h(u)-h(v)}{u-v}\right|,$$

évidemment, $\mathbf{L}_s[f](u,u) = \mathcal{L}_s(\tilde{f})(u)$, avec $\tilde{f}(u) = f(u,u)$.

Remarque. Les divers opérateurs définis ci-dessus rentrent dans le cadre général des opérateurs de transfert associé à un potentiel g. Nous nous intéresserons ici à leurs propriétés spectrales et plus encore à leurs propriétés d'analyticité en s, w.

Rappelons pour terminer ce paragraphe, qu'un opérateur $\mathcal{L}$ agissant sur un espace de Banach de rayon spectral $r(\mathcal{L})$ est *quasi-compact* s'il existe $r < r(\mathcal{L})$ tel que si λ est une valeur spectrale avec $r < \lambda \le r(\mathcal{L})$ alors λ est une valeur propre de multiplicité finie. L'infimum des $r > 0$ vérifiant cette propriété est le rayon spectral essentiel noté $r_{ess}(\mathcal{L})$ (voir [Nu] pour plus de détails).

Nous allons maintenant mettre en œuvre la stratégie d'analyse dite *d'analyse dynamique* sur deux exemples. Cette stratégie peut se résumer par le graphique suivant.



2. Systèmes dynamiques et théorie de l'information

En algorithmique du texte ou en théorie de l'information, on considère des mots sur un alphabet $\mathcal{M}$ (au plus dénombrable). Les propriétés statistiques de ces mots et des algorithmes qui leurs sont liés (algorithmes de recherche, de compression, ...) dépendent de la façon dont sont générés ces mots. Le concept de *source dynamique* introduit par Brigitte Vallée ([Va1]) permet de généraliser les sources classiques (indépendantes ou markoviennes) : les symboles émis peuvent dépendre de tout le passé.
Étant donné un mot w fini sur $\mathcal{M}$, on notera p_w la probabilité d'émission de ce mot. Les grandeurs qui nous intéressent : *probabilité de coïncidence*, *équirépartition des mots de longueur* k et *nombre de préfixes les plus probables*, s'expriment en terme de séries de Dirichlet des probabilités p_w.

2.1. **Sources dynamiques et paramètres à analyser.** Une source dynamique sur un alphabet $\mathcal{M}$ est définie par un système dynamique de l'intervalle I, inversible par morceaux dont la partition d'inversibilité $\mathcal{P}$ est en bijection avec $\mathcal{M}$. Pour $m \in \mathcal{M}$, on notera $h_m \in \mathcal{H}$ la branche inverse définie sur $T(P_m) = Y_m$ où P_m est l'élément de $\mathcal{P}$ correspondant à m. On notera σ l'application qui à $x \in I$ associe $\sigma(x) = m$ si $x \in P_m$.
À une orbite $(x, \ldots, T^n(x), \ldots)$ correspond un mot infini : $M(x) = (\sigma(x), \ldots, \sigma(T^n x), \ldots)$; si la partition $\mathcal{P}$ est génératrice, cette application est injective. On notera $\mathcal{M}_a^\star$ l'ensemble de tous les mots finis admissibles, c'est à dire pouvant être produits par le procédé ci-dessus, et $\mathcal{M}_a^k$ l'ensemble des mots admissibles de longueur k.
Nous dirons que la source est *complète* si pour tout $m \in \mathcal{M}$, $T(P_m) = I$. Dans ce cas tous les mots finis formés sur les lettres de $\mathcal{M}$ sont admissibles (autrement dit, $\mathcal{M}_a^\star = \mathcal{M}^\star$).
Nous dirons que la source est markovienne si T est markovienne pour la partition $\mathcal{P}$. Dans ce cas, l'ensemble des mots infinis admissibles est un sous-ensemble de $\mathcal{M}^{\mathbb{N}}$ qui est un sous décalage de type fini.
On considère une loi μ sur I absolument continue par rapport à la mesure de Lebesgue, $x \in I$ est émis avec cette loi. Étant donné un mot fini $w = (w_0, \ldots, w_{n-1}) \in \mathcal{M}_a^\star$, la probabilité p_w de ce mot est la mesure $\mu(P_{h_w})$ où $h_w = h_{w_0} \circ \cdots \circ h_{w_{n-1}}$ et P_{h_w} est l'atome de la partition $\mathcal{P}^{(n)}$ correspondant à la branche inverse h_w de T^n.
Parmi les quantités qui interviennent en algorithmique du texte nous allons étudier les suivantes que nous appellerons par la suite *paramètres intrinsèques de la source* :

(1) *la coïncidence* $C(x, y)$ entre deux mots $M(x)$ et $M(y)$ pour deux éléments x et y de I tirés indépendamment selon la même loi est la longueur du plus long préfixe commun ; la probabilité pour que $M(x)$ et $M(y)$ aient le même préfixe commun de longueur

k est donc $\mathbb{P}(C(x,y) \geq k)$. Cet évènement se produit si et seulement si x et y appartiennent au même élément de $\mathcal{P}^{(k)}$. On a donc :

$$\mathbb{P}(C(x,y) \geq k) = \sum_{w \in \mathcal{M}_a^k} p_w^2.$$

La probabilité de coïncidence $c(T)$ est

$$c(T) = \lim_{k \to \infty} \left(\sum_{w \in \mathcal{M}_a^k} p_w^2 \right)^{\frac{1}{k}}.$$

(2) *équirépartition des mots de longueur* k. On cherche à décrire les probabilités des mots de longueur k. On définit sur $\mathcal{M}_a^k$ la variable aléatoire $\ell_k(w) = \log p_w$. On étudiera sa distribution asymptotique à l'aide de sa série génératrice des moments :

$$\mathbb{E}(e^{s\ell_k}) = \sum_{w \in \mathcal{M}_a^k} p_w p_w^s = \sum_{w \in \mathcal{M}_a^k} p_w^{s+1}.$$

(3) *préfixes les plus probables*. La quantité $B(\rho)$ désigne le nombre de préfixes w de probabilités supérieure où égale à ρ. On cherche son comportement asymptotique au voisinage de $\rho = 0$. Nous étudierons ce comportement à l'aide des transformées de Laplace et Mellin :

$$L(A)(s) = \int_0^\infty A(y) e^{-sy} dy \quad M(B)(s) = \int_0^\infty B(x) x^{s-1} dx,$$

$L(A)$ est la transformée de Laplace de A, $M(B)$ est la transformée de Mellin de B. Remarquons que si $A(y) = B(e^{-y})$ alors $L(A)(s) = M(B)(s)$ et

$$\sum_{w \in \mathcal{M}_a^\star} p_w^s = sM(B)(s).$$

On est donc amené naturellement à considérer les séries de Dirichlet des probabilités p_w :

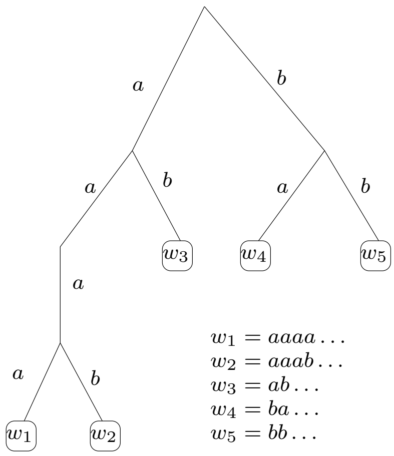
$$\Lambda_k(s) = \sum_{w \in \mathcal{M}_a^k} p_w^s \quad \text{et} \quad \Lambda(s) = \sum_{w \in \mathcal{M}_a^\star} p_w^s.$$

2.1.1. *Arbres digitaux (tries).* Une structure essentielle en algorithmique du texte est celle d'*arbre digital* (ou trie). Elle permet notamment d'implémenter des algorithmes de recherche dans des dictionnaires.
Étant donnés n mots générés indépendamment par la source dynamique : $w_1, \ldots, w_n$, on construit un arbre digital de la façon suivante.

Chaque feuille de l'arbre contient l'un des w_i. Si $\mathcal{M} = \{a_1, \dots, a_r\}$. L'arbre digital associé aux n mots $X = (w_1, \cdots, w_n)$ est défini récursivement par :

trie$(X) = \langle$trie$(X \backslash a_1), \dots,$ trie$(X \backslash a_r)\rangle$,

où $X \setminus a_i$ est contient les mots de X commençant par a_i et dont la première lettre a_i a été supprimée. La récurrence s'arrête dès que X contient au plus 1 élément.



Remarque. Dans la définition ci-dessus, on a supposé que l'alphabet $\mathcal{M}$ est fini. En fait, on peut définir l'arbre digital de la même manière si l'alphabet est infini : comme on considère un nombre fini de mots, à chaque étape de la construction, le nombre de lettres qui interviennent dans la définition récursive est fini.

On s'intéresse à la *taille* (nombre de nœuds internes) $S^{[n]}$, la *longueur de cheminement* (somme des longueurs de tous les chemins de la racine aux feuilles) $P^{[n]}$ et la *hauteur* (longueur maximale d'un chemin de la racine à une feuille) $H^{[n]}$ d'un arbre digital construit à partir de n mots indépendants. Pour estimer les espérances de ces quantités, on passe par le modèle de Poisson : on tire un entier n avec une loi de Poisson de paramètre z, puis on tire indépendamment n mots suivant le modèle de la source dynamique. On note S_z, P_z et H_z les variables aléatoires taille, longueur de cheminement et hauteur correspondantes. Clément, Flajolet et Vallée ([ClFlVa]) ont montré que les transformées de Mellin des espérances de S_z, P_z et H_z s'expriment aussi en terme de séries de Dirichlet des p_w. Par un procédé de "dépoissonisation", on déduit du comportement asymptotique des espérances "poissonisées" le comportement asymptotique des espérances à n fixé.

2.1.2. *Lien avec les opérateurs.* Suivant notre stratégie, nous devons maintenant relier les séries de Dirichlet aux opérateurs de Ruelle généralisés.
Rappelons que les probabilités p_w sont relatives à une mesure μ absolument continue. Notons f la densité de μ et F sa fonction de répartition. Supposons qu'il existe $\gamma < 1$ tel que pour $\Re(s) > \gamma$, les séries qui définissent les opérateurs $\mathcal{L}_s$ et $\mathbf{L}_s[f]$ convergent uniformément. Alors (voir [Va1, 4]), pour tout $\Re(s) > \gamma$, $k \geq 0$,

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

avec $\varphi_F(x,x') = \frac{|F(x)-F(x')|}{|x-x'|}$ et $P_m =]a_m, b_m[$.

Remarque. Les expressions ci-dessus permettent effectivement de donner le comportement asymptotique des paramètres intrinsèques de la source. Par contre pour les paramètres des arbres digitaux associés, les relations avec les séries de Dirichlet sont plus compliquées et font intervenir d'autres généralisations de l'opérateur de Ruelle qui agissent sur des espaces de fonctions à quatre variables.

2.2. **Résultats.** Les résultats que nous allons donner dans ce paragraphe ont étés obtenus par Brigitte Vallée ([Va1]) pour les paramètres intrinsèques des sources dynamiques et par Brigitte Vallée, Philippe Flajolet et Julien Clément ([ClFlVa]) pour les paramètres des arbres digitaux, dans le cas de sources dynamiques complètes, uniformément dilatantes et analytiques par morceaux. En collaboration avec Frédéric Chazal, nous avons généralisé ces résultats à une classe plus large de sources dynamiques, non nécessairement complètes et non analytiques. Nos hypothèses sur le système dynamique sont les suivantes :

- (d1) **Contraction.** Il existe $0 < \eta_m \leq \delta_m < \delta < 1$ tels que $\eta_m \leq |h'_m(x)| \leq \delta_m$ pour $x \in J_m$.
- (d2) Il existe $\gamma < 1$ tel que pour $\mathcal{R}(s) > \gamma$, la série $\displaystyle\sum_{\substack{m\in\mathcal{M}\\ x\in Y_m}} \delta_m^s$ converge uniformément pour $x \in I$ et $\displaystyle\sum_{m\in\mathcal{M}} |P_m|^s$ converge.
- (d3) **Distorsion bornée.** Il existe une constante $A < +\infty$ telle que pour tout $m \in \mathcal{M}$ et tous $x, y \in P_m$, $|h''_m(x)/h'_m(y)| < A$.
- (d4) **Markov.** L'application T est markovienne.
- (d5) **Positivité.** Condition technique mais plus ou moins équivalente à l'aperiodicité pour les chaînes de Markov classiques.

Remarque. Dans [4], nous donnons des exemples de systèmes dynamiques de l'intervalle vérifiant ces propriétés. Par exemple, (d5) est satisfaite pour des applications markoviennes à grandes branches et fortement apériodiques. Nous donnons aussi un exemple d'application markovienne qui vérifie (d5) mais n'est pas à grandes branches.

Théorème III.1. *(*[Va1] *dans le cas analytique complet,* [4, 5]*)*
Supposons que T vérifie les hypothèses (d1-5) ci-dessus et que la densité f de la mesure μ soit bornée, Lipschitz sur chaque P_m et de constante de Lipschitz uniformément bornée. Alors, il existe $\lambda(s) > 0$, $\Phi(s) > 0$ et $0 \leq \rho(s) < 1$ trois fonctions analytiques sur un voisinage complexe de la demi-droite $\{s \in \mathbb{R} \ / \ s > \gamma\}$ telles que pour tout $k \geq 1$,

$$\Lambda_k(s) = \lambda^k(s)\left(\Phi(s) + O(\rho^k(s))\right).$$

$\lambda(s)$ est la valeur propre dominante de $\mathcal{L}_s$ sur l'espace des fonctions bornées et Lipschitz sur chaque P_m et de constante de Lipschitz uniformément bornée.

$\Lambda(s)$ est analytique sur $\Re(s) > 1$ et a un pôle simple en $s = 1$.
En appliquant les Théorèmes B et C, on obtient alors :

- *La variable C est asymptotiquement géométrique :*
$$\mathbb{P}(C \geq k) = A_F \cdot \lambda(2)^k (1 + O(\kappa^k))$$
où $A_F > 0$ ne dépend que de la distribution F et $0 < \kappa < 1$.
- *Si $\lambda''(1) - \lambda'(1)^2 \neq 0$ alors la variable $\log \ell_k$ est asymptotiquement normale. De plus, $\lambda''(1) - \lambda'(1)^2 = 0$ si et seulement si T est conjugué à une application linéaire par morceaux, de même pentes, la conjugaison est C^{1+Lip} sur chaque P_m.*
- *Soit 1 est le seul pôle de $\Lambda(s)$ sur $\Re(s) = 1$, dans ce cas*
$$B(x) \simeq \frac{-1}{\lambda'(1)x},$$
ou bien, T est conjuguée à une application affine par morceaux, de pentes α^k, $\alpha > 1$, $k \in \mathbb{Z}$, la conjugaison est C^{1+Lip} sur chaque P_m. Dans ce cas, il existe A, B,
$$\frac{A}{x} \leq B(x) \leq \frac{B}{x}.$$

Remarque. Le résultat sur ℓ_k existe dans le “folklore” des systèmes dynamiques (voir [ParPo]) et de la théorie de l'information dans le cas où μ est une mesure invariante vérifiant une propriété de Gibbs. Il s'agit d'un raffinement du théorème d'équipartition de Shannon-MacMillan-Breiman.

La preuve du théorème réside essentiellement dans les propriétés spectrales de opérateurs $\mathcal{L}_s$ et $\mathbf{L}_s$ agissant sur les espaces $L_{pw}(I)$ des fonctions sur I, bornées, Lipschitz sur chaque P_m, avec une constante de Lipschitz bornée et $\widetilde{L_{pw}}$ des fonctions définies sur $\bigcup_{m \in \mathcal{M}} P_m \times P_m$, bornées, Lipschitz sur chaque $P_m \times P_m$, avec une constante de Lipschitz bornée. On a le résultat suivant.

Théorème III.2. [4] *Supposons que T vérifie les hypothèses (d1-5) ci-dessus et que la densité f de la mesure μ appartienne à $L_{pw}(I)$. Pour tout réel $s > \gamma$, l'opérateur $\mathbf{L}_s$ agissant sur $\widetilde{L_{pw}}$ est quasi-compact et admet une unique valeur propre dominante $\lambda(s)$. En particulier, pour tout $k \in \mathbb{N}$, $\varphi \in \widetilde{L_{pw}}$, on a la décomposition*
$$\mathbf{L}_s^k \varphi = \lambda^k(s) \left(\Pi_s(\varphi) + S_s^k \varphi \right),$$
où Π_s est la projection spectrale sur le sous-espace propre dominant et S_s est un opérateur sur $\widetilde{L_{pw}}$ dont le rayon spectral est strictement plus petit que $\lambda(s)$ et $S_s \circ \Pi_s = \Pi_s \circ S_s = 0$.
Soit s un nombre complexe tel que $\Re(s) > \gamma$. On a

(1) *La fonction $s \to \lambda(s)$ est strictement décroissante le long de l'axe réel $s > \gamma$.*
(2) *Sur chaque droite verticale $\Re(s) = \sigma$, on a $r(s) \leq \lambda(\sigma)$.*
(3) *Si $r(s) = \lambda(\sigma)$ pour $s = \sigma + it$ alors $\mathbf{L}_s$ a une valeur propre $\lambda = e^{ia}\lambda(\sigma)$, $a \in \mathcal{R}$ qui appartient au spectre de $\mathcal{L}_s$.*

Ces propriétés spectrales peuvent être montrées en utilisant le théorème d'Hennion ([HenHer]) puis en utilisant des techniques classiques sur les opérateurs quasi-compacts (afin de montrer qu'il existe une unique valeur propre dominante) ou directement par des méthodes de cônes. Remarquons que même pour l'opérateur $\mathcal{L}_s$ le résultat ci-dessus est original (les systèmes dynamiques que nous considérons ne vérifient pas les hypothèses "classiques" des systèmes dynamiques dilatants par morceaux avec une infinité de branches [Broi, BaKe, LiSV]).

Les mêmes techniques appliquées à des opérateurs "multi-sécants" agissant sur des fonctions de quatre variables donnent le résultat suivant.

Théorème III.3. *(*[ClFlVa] *dans le cas analytique complet,* [4]*)* *Supposons que T vérifie les hypothèses (d1-5) ci-dessus et que la densité f de la mesure μ appartienne à $L_{pw}(I)$. Notons $S^{[n]}, P^{[n]}, H^{[n]}$ la taille, la longueur de cheminement et la hauteur d'un arbre digital construit avec n mots générés indépendamment. Le comportement asymptotique de l'espérance (quand $n \to \infty$) de ces paramètres est donné par*

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

où $h(T) = \lambda'(1)$ est l'entropie de la source et $c(T) = \lambda(2)$ la probabilité de coïncidence.

3. Analyse d'algorithmes arithmétiques

Cette section est consacrée à la mise en œuvre de la startégie d'analyse dynamique pour étudier des paramètres comme *nombre d'itérations* ou la *complexité en bit* d'algorithmes arithmétiques et plus particulièrement d'un algorithme de calcul de pgcd sur les bits de poids faibles. L'intérêt de l'étude des algorithmes de pgcd est d'abord historique : l'algorithme d'Euclide est d'après Knuth ([Kn]) le "grand-père de tous les algorithmes". Néanmoins la compréhension de son comportement reste partielle : récemment, Viviane Baladi et Brigitte Vallée ont établi que le nombre d'itérations suit asymptotiquement une loi gaussienne ([BaVa]). De nombreux paramètres restent encore à étudier.
La principale motivation de telles analyses est, néanmoins, que le calcul de pgcd est une opération très fréquente dans de nombreuses applications informatiques (calcul fractionnaire, cryptographie à clé publique, calcul formel ...). L'optimisation de ces algorithmes est donc essentielle. Nous allons, dans un premier temps, expliquer la problématique générale

de l'analyse en moyenne d'algorithmes de calcul de pgcd. Nous détaillerons ensuite l'algorithme LSB de calcul de pgcd sur les bits de poids faibles. Nous énoncerons pour terminer les résultats obtenus pour cet algorithme en collaboration avec Benoit Daireaux et Brigitte Vallée.

3.1. **Problématique.** Commençons par décrire le formalisme lié aux algorithmes euclidiens (voir [Da] pour une présentation complète).
Les algorithmes de calcul de pgcd type "Euclide" reposent sur une division euclidienne de la forme : pour u et v deux entiers,

$$u = vq + r,$$

si on impose $0 \leq r < v$, on a la division euclidienne standard. Pour l'algorithme sur les bits de poids faibles qui nous concerne, on impose, une condition sur le reste r qui fait intervenir la norme 2-adique des entiers (voir section suivante).
Un algorithme euclidien est alors une succession de divisions et d'échanges qui se termine avec un reste nul (ou égal à un dans certains cas). L'avant-dernier reste est le pgcd de u et v. Le but de l'analyse probabiliste de tels algorithmes est de déterminer le comportement asymptotique moyen ou en distribution de divers paramètres comme : le nombre d'itérations, la complexité en bit, la longueur des quotients, des restes ...
Si $\widetilde{\Omega}$ est l'ensemble des couples d'entiers (u, v) qui forment une entrée valide pour l'algorithme, on définit une taille $\ell(u, v)$ sur $\widetilde{\Omega}$ comme étant la longueur binaire de la norme $\|(u, v)\|$, cette norme diffère suivant les algorithmes. Nous étudierons le comportement des différents paramètres sur les sous-ensembles de $\widetilde{\Omega}$:

$$\Omega = \{(u, v) \in \widetilde{\Omega} \;\; / \;\; pgcd(u, v) = 1\}, \;\; \Omega_n = \{(u, v) \in \Omega \;\; / \;\; \ell(u, v) = n\},$$
$$\widetilde{\Omega_n} = \{(u, v) \in \widetilde{\Omega} \;\; / \;\; \ell(u, v) = n\}.$$

Ces deux derniers ensembles Ω_n et $\widetilde{\Omega_n}$ sont munis de l'équiprobabilité, on notera $\mathbb{P}_n$ et $\widetilde{\mathbb{P}}_n$ les probabilités sur Ω_n et $\widetilde{\Omega_n}$ respectivement. En fait nous travaillons sur Ω_n et montrons que les comportements asymptotiques sur Ω_n et $\widetilde{\Omega_n}$ sont les mêmes.
Si on associe un coût $c(q)$ à chaque quotient q, étant donnée une entrée (u, v) de l'algorithme, telle que l'exécution de l'algorithme requière p divisions, on note q_1, ..., q_p les quotients successifs, le coût $C(u, v)$ est défini par

$$C(u, v) = \sum_{i=1}^{p} c(q_i).$$

Des coûts particulièrement intéressants sont $c \equiv 1$ qui génère le nombre d'itérations, $c =$ la taille binaire ...
De tels coûts ont été étudiés pour diverses variantes de l'algorithme euclidien standard (voir [AVa, Va2, Va3, Va4, BaVa]). Ces études ont notamment permis de classer les différentes variantes en fonction de

leur efficacité : pour les algorithmes dits *rapides*, le nombre moyen d'itérations est linéaire en n la taille des entrées - ce résultat s'étend à une catégorie de coûts dits à croissance modérée -, le nombre moyen d'itérations est quadratique pour les algorithmes dits *lents*. La complexité en bit des algorithmes rapides est quadratique. Enfin, pour les algorithmes rapides et les coûts à croissance modérée, le coût total est asymptotiquement normal.
Nous allons décrire la méthode et les résultats pour une variante de l'algorithme d'Euclide qui repose sur la division sur les bits de poids faibles.

3.2. **L'algorithme euclidien sur les bits de poids faibles.** L'algorithme LSB ("least signifiquant bits") repose sur la division sur les bits de poids faibles, c'est à dire, si u et v sont des entiers, la division de u par v renvoie un reste r qui est divisible par une puissance de 2 plus grande que la plus grande puissance de 2 qui divise v. Autrement dit, on cherche à avoir plus de 0 à droite dans le développement en base 2. La division standard produit des 0 à gauche. Cet algorithme est intéressant parce qu'il est plus stable et plus facile à implanter que les autres algorithmes de calcul de pgcd. L'algorithme de calcul de pgcd LSB a été utilisé pour mettre au point un algorithme de calcul de pgcd de type "Diviser pour régner" (algorithme récursif) par Stehlé et Zimmermann ([StZ]). Cet algorithme s'avère (expérimentalement) compétitif par rapport à d'autres algorithmes "Diviser pour régner" dont la brique de base est l'algorithme d'Euclide standard (algorithme de Lehmer-Euclide ou de Shönhage, voir [DaVa] pour une analyse en moyenne de l'algorithme de Lehmer-Euclide). L'algorithme LSB a été introduit et étudié par Stehlé et Zimmermann ([StZ]), en collaboration avec Benoit Daireaux et Brigitte Vallée, nous prouvons leurs observations expérimentales.

3.2.1. *La division LSB.* Rappelons que la valuation 2-adique $\nu(u)$ d'un entier u est la plus grande puissance de 2 qui divise u. Cette valuation s'étend sur $\mathbb{Q}$ par $\nu(\frac{u}{v}) = \nu(u) - \nu(v)$ et permet de définir la norme 2-adique sur $\mathbb{Q}$:

$$|x|_2 = 2^{-\nu(x)}.$$

Si u est un entier impair et v un entier pair, la division centrée sur les bits de poids faibles est décrite par :

$$u = vq + r \quad \text{avec} \quad |q| < 1 \quad \text{et} \quad 0 \leq |r|_2 < |v|_2.$$

La division entre u et v requière un certain nombre de décalages et de soustractions. Par exemple, la division entre $u = 29 = 11101_2$ et $v = 12 = 1100_2$ est obtenue en décalant v de deux crans, en soustrayant 11_2 à 11101_2, on soustrait ensuite le résultat obtenu à 110_2 et on recommence ... : $11101_2 - 11_2 = 11010_2$, $11010_2 - 110_2 = 10100_2$,

$10100_2 - 1100_2 = 1000_2$, on a alors,

$$11101_2 = 1100_2 \times \frac{1_2 + 10_2 + 100_2}{100_2} + 1000_2,$$

on centre alors le quotient (de façon à obtenir $|q| < 1$) en effectuant une nouvelle soustraction : $1000_2 - 111_2 = 1_2$ et finalement on obtient :

$$11101_2 = 1100_2 \times \frac{-1_2}{100_2} + 100000_2,\ 29 = \frac{-1}{4}12 + 32.$$

On a alors $\nu(r) > \nu(v) > 0$ et la paire $(v', r') := (2^{-\nu(v)}v, 2^{-\nu(v)}r)$ est telle que v' est impair et r' pair. L'algorithme de calcul de pgcd consiste alors à itérer le processus : on divise v' par r' et on enlève les zéros à droite de r' et du nouveau reste ... Jusqu'à obtenir un reste nul.

Remarque. On peut se demander pourquoi centrer le quotient, la division non centrée peut paraître plus naturelle. En fait si on utilise la division non centrée, l'algorithme de calcul de pgcd décrit ci-dessous ne s'arrête pas toujours. Par contre, il termine, pour toute entrée (u, v) valide, si on utilise l'algorithme centré (voir [StZ]).

3.2.2. *L'agorithme LSB.* L'ensemble $\widetilde{\Omega}$ des entrées valides de l'algorithme est :

$$\widetilde{\Omega} = \{(u, v) \in \mathbb{Z}^2 \ /\ u \text{ est impair},\ v \text{ est pair}\}.$$

Sur l'ensemble des couples (u, v) de $\widetilde{\Omega}$, l'algorithme LSB réalise une succession de divisions et de décalages. Il est décrit comme suit : $(u_0 := u, u_1 := v)$

$$\left\{\begin{array}{ccc} u_0 = q_1 u_1 + r_1, & u_2 := 2^{-\nu(u_1)} \cdot r_1, & u_1 := 2^{-\nu(u_1)} \cdot u_1, \\ u_1 = q_2 u_2 + r_2, & u_3 := 2^{-\nu(u_2)} \cdot r_2, & u_2 := 2^{-\nu(u_2)} \cdot u_2, \\ \cdots & \cdots & \cdots \\ u_{i-1} = q_i u_i + r_i, & u_{i+1} := 2^{-\nu(u_i)} \cdot r_i, & u_i := 2^{-\nu(u_i)} \cdot u_i \\ \cdots & \cdots & \cdots \end{array}\right\}.$$

L'algorithme s'arrête après p itérations avec $u_{p+1} = 0$. Le pgcd de u et v vaut alors $2^k u_p$ avec $k = k_1 + \cdots + k_p$ le nombre de décalages effectués. La Figure 1 décrit la suite de quotients et de restes sur un exemple. Remarquons que la suite u_i n'est pas décroissante (comme c'est le cas pour l'algorithme d'Euclide classique).

On obtient aussi un développement en fractions continues du rationnel v/u,

$$\frac{v}{u} = \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots + \cfrac{1}{q_p + 0}}}} \tag{3.1}$$

Entrée : $(u, v) = (72001, 2011176)$
En base 2 $(u, v) = (10001100101000001_2, 111101011000000101000_2)$.

i	u_i [base 2]	u_i[base 10]	quotient $a_i/2^{k_i}$
0	**10001100101000001**	72001	
1	**111101011000000101**000	2011176	-3 / 8
2	**1100100110110101**0000	826192	1 / 2
3	**11000011000101**0000000	1598080	1 / 8
4	**100110001111**00000000	626432	-1 / 2
5	**111010010101**000000000	1911296	-1 / 2
6	**11000001001**0000000000	1582080	1 / 2
7	**1000100011**00000000000	1120256	-1 / 2
8	**1000001011**000000000000	2142208	1 / 2
9	**11**00000000000000	49152	1 / 4
10	**1000001**000000000000000	2129920	-1 / 2
11	**10001**0000000000000000	1114112	1 / 2
12	**11**0000000000000000000	1572864	-5 / 8
13	**1**000000000000000000000	2097152	3 / 4

FIG. 1. Une exécution de l'algorithme LSB

L'ensemble des quotients possibles est

$$\mathcal{Q} = \left\{ \frac{a}{2^k} \ / \ k \geq 1, \ a \text{ impair}, \ a \in [-2^k + 1, 2^k - 1] \right\}.$$

Pour $q \in \mathcal{Q}$, on considère les matrices :

$$M_{[q]} = \begin{pmatrix} 0 & 2^k \\ 2^k & a \end{pmatrix}, \qquad N_{[q]} = \begin{pmatrix} 0 & 1 \\ 1 & q \end{pmatrix} = 2^{-k} M_{[q]},$$

étant donnée une entrée valide $(u, v) \in \widetilde{\Omega}$, la paire intermédiaire (v, r) et la nouvelle paire (v', r') sont données par :

$$\begin{pmatrix} v \\ u \end{pmatrix} = N_{[q]} \begin{pmatrix} r \\ v \end{pmatrix}, \qquad \begin{pmatrix} v \\ u \end{pmatrix} = M_{[q]} \begin{pmatrix} r' \\ v' \end{pmatrix} .$$

Pour une entrée (u, v) de pgcd égal à d, l'exécution de l'algorithme est décrite par un produit de matrices :

$$\begin{pmatrix} v \\ u \end{pmatrix} = M \begin{pmatrix} 0 \\ d \end{pmatrix} = N \begin{pmatrix} 0 \\ 2^k d \end{pmatrix}, \text{ avec}$$

$$M := M_{[q_1]} \cdot M_{[q_2]} \cdot \ldots \cdot M_{[q_p]}, \quad N := \frac{1}{2^k} M, \quad k = k_1 + \cdots k_p.$$

Si $h_{[q]}(x)$ désigne l'homographie associée à la matrice $M_{[q]}$ [ou $N_{[q]}$], elle est définie par

$$h_{[q]}(x) := \frac{1}{q + x} = \frac{2^k}{a + 2^k x},$$

la décomposition en fraction continue s'écrit alors :

$$\frac{v}{u} = h_{[q_1]} \circ h_{[q_2]} \circ \cdots \circ h_{[q_p]}(0) = h(0).$$

On note $\mathcal{M}$ et $\mathcal{N}$ les ensembles de matrices $M_{[q]}$ et $N_{[q]}$. Ces matrices vont bien sûr jouer un rôle fondamental dans la suite. Remarquons que :

$$\sum_{q \in \mathcal{Q}} \frac{1}{|q|_2^2} = 1 \quad (|q|_2 = |\frac{a}{2^k}|_2 = 2^k).$$

Nous allons considérer des produits aléatoires des matrices $M_{[q]} \in \mathcal{M}$ ou $N_{[q]} \in \mathcal{N}$, choisies avec la probabilité $\delta_q = \frac{1}{|q|_2^2}$.

3.2.3. *Comportement asymptotique de l'algorithme LSB.* Nous définissons la taille d'une entrée valide (u, v) par :

$$\ell(u, v) = \frac{1}{2}\text{longueur binaire de } (u^2 + v^2).$$

Autrement dit, c'est la taille liée à la norme euclidienne. Nous allons étudier différents coûts sur les ensembles Ω_n et $\widetilde{\Omega}_n$ associés à cette taille. La classe de coûts pour laquelle nous savons mener l'analyse en moyenne est celle des coûts *à croissance modérée*, ils vérifient :

$$\mu(c) := \sum_{q \in \mathcal{Q}} \frac{1}{|q|_2^2} \cdot c(q) < \infty. \tag{3.2}$$

Dans le cas où $c \equiv 1$, le coût total est alors le nombre d'itérations de l'algorithme noté P. Si $c(\frac{a}{2^k}) = k(q) = k$, alors le coût total est K le nombre de décalages effectués au cours de l'exécution de l'algorithme. Si $c(\frac{a}{2^k}) = s(q)$ est le nombre de 1 consécutifs à gauche dans la représentation binaire de a si $a > 0$ ou 1 plus le nombre de 1 consécutifs à gauche dans la représentation binaire de $a + 2^{k+1}$ si $a < 0$, alors le coût total est S le nombre de soustractions effectuées. La complexité en bit B de l'algorithme est donnée par le coût $b(u, v) = \ell(u) \cdot (k(q) + s(q))$ où $\ell(u)$ est la longueur binaire de u. Ce coût n'est pas dans la classe des coûts considérés ci-dessus : il dépend à la fois du quotient q et de l'entrée u. On peut néanmoins estimer le comportement moyen de la complexité en bit pour une entrée (u, v) :

$$B(u, v) = \sum_{i=1}^{p} b(u_i, u_{i+1}).$$

Comme nous l'avons remarqué, la suite des u_i n'est pas décroissante - mais il semble néanmoins que la taille moyenne augmente faiblement - par contre à chaque étape le nombre de 0 à droite dans le développement binaire augmente. Nous voyons alors l'algorithme comme une course entre un lièvre (binaire) et une tortue (réelle). Contrairement à la fable, le lièvre rattrape la tortue et l'algorithme termine.

3.3. **Le lièvre et la tortue.** Le comportement asymptotique moyen des coûts s'exprime en fonction de γ_0 l'exposant de Lyapunov binaire des compositions aléatoires de matrices $N \in \mathcal{N}$:

$$\gamma_0 := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}(\log_2 \|N_1 N_2 \cdots N_n\|).$$

Cet exposant décrit la croissance moyenne des u_i. Des simulations numériques faites par Philippe Flajolet ([Fl2]) montrent que $\gamma_0 \sim 0,0497$. Le nombre de 0 à droite du développement en base 2 augmente lui - en moyenne - de 2. Le théorème suivant indique que le comportement asymptotique des coûts additifs relève bien d'une course entre le lièvre et la tortue. Par contre, comme on le voit dans le Théorème III.5, pour la taille des continuants, le lièvre et la tortue ne sont plus en compétition mais marchent ensemble et la taille de continuant pour des entrées de tailles n croît en $(2+\gamma_0)\cdot n$.

Théorème III.4. [1] *Soit C un coût additif associé à c vérifiant (3.2). Sur les ensembles $\Omega_n, \widetilde{\Omega}_n$, munis de l'équiprobabilité, le comportement asymptotique moyen de C sur les entrées de taille n est donné par*

$$\mathbb{E}_n[C] \sim \tilde{\mathbb{E}}_n[C] \sim \frac{1}{2-\gamma_0}\cdot \mu[c]\cdot n,$$

Pour P (nombre d'itérations), K (nombre de décalages), S (nombre de soustractions) on a

$$\mu[p] = 1, \qquad \mu[k] = 2, \qquad \mu[s] = \frac{5}{2}.$$

De plus, le comportement asymptotique de la complexité en bits est donné par :

$$\mathbb{E}_n[B] \sim \tilde{\mathbb{E}}_n[B] \sim \frac{1}{2-\gamma_0}\cdot \mu[k+s]\cdot \frac{n^2}{2}.$$

Le second type de paramètres auxquels nous nous intéressons est la taille des “continuants”. C'est la taille binaire des quotients successifs du développement en fractions continues. Le $n$ième continuant est le vecteur $Q_n(x) = (p_n(x), r_n(x))$ définit par la relation :

$$\begin{pmatrix} p_n(x) \\ r_n(x) \end{pmatrix} = M_{[q_1]}\cdot M_{[q_2]}\cdot \ldots \cdot M_{[q_n]} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.3}$$

le rationnel $p_n(x)/r_n(x) = h_{[q_1]} \circ h_{[q_2]} \circ \cdots \circ h_{[q_n]}(0)$ approxime le nombre 2-adique x [par rapport à la norme 2−adique]. La suite des q_i forme le développement en fractions continues - centré - du nombre 2-adique x. Ce développement a été étudié par Browkin ([Brow1, Brow2]).
Rappelons que l'ensemble $\mathbb{Q}_2$ des nombres 2-adiques est le complété de $\mathbb{Q}$ pour la norme $|\ |_2$

Théorème III.5. [1] *Soit x un élément de la boule unité $\mathcal{B}$ de $\mathbb{Q}_2$, soit $Q_n(x) := (p_n(x), r_n(x))$ le vecteur de $\mathbb{Z}^2$ définit par (3.3), soit $||.||$*

la norme Euclidienne. On considère $\mathcal{B}$ muni de la densité uniforme par rapport à la mesure de Haar η. La variable aléatoire $\log ||Q_n||$ *est asymptotiquement normale et la vitesse de convergence est optimale (i.e. en $O(1/\sqrt{n})$). De plus, la variance et la moyenne vérifient :*

$$\mathbb{E}[\log_2 ||Q_n||] = (2+\gamma_0)\cdot n+a+O(\tau^{-n}), \quad Var[\log_2 ||Q_n||] = b\cdot n+c+O(\tau^{-n})$$

a, b, c, τ sont des constantes telles que $b > 0, \tau < 1$.

La preuve de ces deux théorèmes consiste à relier les séries génératrices associées aux différents coûts à des opérateurs de transfert généralisés dont on montre qu'ils sont quasi-compacts et qu'ils possèdent de bonnes propriétés d'analyticité, le Théorème A donne alors le Théorème III.4 et le Théorème C donne le Théorème III.5. Les opérateurs de transfert que nous allons considérer font intervenir la norme 2-adique des quotients q et les matrices $M_{[q]}$. Ces matrices agissent sur la droite projective, on peut aussi considérer leur action sur le tore $J = \mathbb{R}/\pi\mathbb{Z}$ vu comme un ensemble d'angles, on la note $\ell_{[q]} : J \to J$. Les opérateurs de transfert $\mathbf{G}_{t,z}$ définis par (3.4) sont relatifs à ce système de fonctions itérées.

Nous allons considérer les séries univariées :

$$T_C(s) := \sum_{(u,v)\in\Omega} \frac{C(u,v)}{(u^2+v^2)^s} = \sum_{n\geq 1} \frac{t_n^C}{n^s},$$

où t_n^c est le coût cumulé C sur l'ensemble des paires (u, v) telles que $(u^2 + v^2)$ vaut n :

$$t_n^C = \sum_{\substack{(u,v)\in\Omega \\ (u^2+v^2)=n}} C(u,v).$$

L'espérance de C sur Ω_N s'écrit :

$$\mathbb{E}_N[C] := \frac{\sum_{n=2^{2N-1}}^{2^{2N}} t_n^c}{\sum_{n=2^{2N-1}}^{2^{2N}} |\Omega_n|}.$$

Nous allons relier les séries de Dirichlet aux opérateurs :

$$\mathbf{G}_{t,z}^{[c]}[f](y) := \sum_{q\in\mathcal{Q}} \delta_q^t \cdot c(q) \cdot |\ell_q'(y)|^z \cdot f \circ \ell_q(y), \tag{3.4}$$

et plus particulièrement aux opérateurs : $\mathbf{G}_{s,s}^{[1]} =: \mathbf{G}_s$ et $\mathbf{G}_{s,s}^{[c]} =: \mathbf{G}_s^{[c]}$, notons aussi $\mathbf{G}_{t,z}$ pour $\mathbf{G}_{t,z}^{[1]}$.

Proposition III.6. [Da, 1] *Les séries univariées vérifient :*

$$T_C(s) = (I - \mathbf{G}_s)^{-1} \circ \mathbf{G}_s^{[c]} \circ (I - \mathbf{G}_s)^{-1}[1](0). \tag{3.5}$$

Pour la complexité en bit B on a la relation suivante :

$$T_B(s) = (I-\mathbf{G}_s)^{-1}\circ\mathbf{G}_s^{[k+s]}\circ(I-\mathbf{G}_s)^{-1}\circ\frac{d}{ds}\mathbf{G}_s\circ(I-\mathbf{G}_s)^{-1}[1](0). \tag{3.6}$$

Enfin, la série génératrice des moments de la norme des continuants $\|Q_n(x)\|$ *vérifie :*

$$\mathbb{E}[\exp(2w \log ||Q_n||)] = \mathbf{G}^n_{1-w,-w}[1](0).$$

Le théorème ci-dessous permet alors d'appliquer les Théorèmes A (au voisinage de $\sigma = 1$) et C.

Théorème III.7. [1] *Soient* $\mathcal{A} := \{(t,z) \in \mathcal{C}^2; \Re t > 1/2\}$, *et* $\mathcal{D}_0$ *un voisinage complexe (suffisamment petit) de* $(1,0)$ *et* $\mathcal{D}_1$ *un voisinage complexe (suffisamment petit) de* $(1,1)$. *Les opérateurs de transfert généralisés vérifient :*
(i) *Pour* $(t,z) \in \mathcal{A}$, *les opérateurs* $\mathbf{G}_{t,z}, \mathbf{G}^{[c]}_{t,z}$ *agissent sur* $\mathcal{C}^1(J)$. *De plus, l'application* $(t,z) \mapsto \mathbf{G}_{t,z}$ *est analytique.*
(ii) *Pour* $(t,z) \in \mathcal{D}_0 \cup \mathcal{D}_1$, *l'opérateur* $\mathbf{G}_{t,z}$, *agissant sur* $\mathcal{C}^1(J)$ *est quasi-compact et admet une unique valeur propre* $\lambda(t,z)$ *dominante.*
(iii) *La valeur propre dominante* $\lambda(t,z)$ *vérifie les relations suivantes :*

$$\lambda(t,0) = \frac{2^{1-2t}}{1-2^{1-2t}}, \qquad \lambda(t,z) = \lambda(t,1-z), \qquad \lambda(1,0) = 1 = \lambda(1,1),$$

et l'exposant de Lyapunov γ *associé au produit aléatoire des matrices de* $\mathcal{N}$ *vérifie*

$$2\gamma = \lambda'_z(1,1) = -\lambda'_z(1,0).$$

(iv) *L'application* $w \mapsto \log \lambda(1-w,-w)$ *a une dérivée seconde non nulle en* $w = 0$.
(v) *Sur* $\Re s \geq 1, s \neq 1$, *le rayon spectral* $r(s)$ *de* $\mathbf{G}_s$ *est strictement plus petit que* 1.

La preuve de ce dernier théorème utilise les résultats sur les produits aléatoires de matrices de Furstenberg, Guivarc'h-Raugi et Le Page que l'on peut trouver dans le livre de Bougerol-Lacroix ([BouL]). On obtient ainsi le comportement des opérateurs $\mathbf{G}_{t,z}$ au voisinage de $(1,0)$ sur l'espace des fonctions α-Höldériennes (pour α suffisamment petit). Nous montrons alors que le comportement est le même sur l'espace des fonctions C^1. Une relation intégrale entre $\mathbf{G}_{t,z}$ et $\mathbf{G}_{t,1-z}$ et le théorème de Hennion ([HenHer]) nous donne le comportement de $\mathbf{G}_{t,z}$ au voisinage de $(1,1)$ sur l'espace des fonctions continues.

Quatrième partie IV. Perspectives

Pour conclure ce document de synthèse, donnons quelques perspectives de recherche.
Mes projets de recherche s'orientent autour de deux thèmes principaux : estimation dans les systèmes dynamiques dans le but de les utiliser en modélisation et simulation ; méthodes dynamiques en analyse d'algorithmes notamment pour les algorithmes "auto-organisés" et les algorithmes arithmétiques.

1. Simulation, modélisation par des systèmes dynamiques

Suivant les idées développées dans la section II.2, dans le but d'utiliser les systèmes dynamiques pour la modélisation et la simulation de systèmes complexes, nous souhaitons estimer des paramètres et des fonctions afin de reconstruire le système dynamique. L'idée est d'utiliser des inégalités de concentration comme Hoeffding ou Bernstein pour obtenir des estimateurs fonctionnels convergents.

1.1. **Encore d'autres propriétés statistiques.** Dans la section II.2.3, nous avons utilisé des inégalités exponentielles de type Hoefding pour obtenir un estimateur convergent de la fonction potentiel du système. Ces inégalités font intervenir des coefficients de mélange faible. Ces coefficients sont difficiles à estimer. Les inégalités exponentielles de type Bernstein, par contre, font intervenir la variance qui, elle, est plus facilement estimable. Paul Doukhan et Michael H. Neumann ([DouNe]) ont montré de telles inégalités pour des sommes de variables aléatoires vérifiant certaines conditions de dépendance faible. On pourrait, d'une part, essayer d'utiliser ces inégalités pour obtenir la "typicité" et la "typicité conditionelle" dans les systèmes dynamiques. D'autre part, on peut s'intéresser à obtenir de telles inégalités pour des fonctions $f_n(X_1, \dots, X_n)$, plus générales que la somme le long de la suite de variables aléatoires (étendant ainsi les résultats de [DedPr, ColMSc1, ChaColSc]).

1.2. **Simulation, modélisation.** Dans la section II.1.2 nous avons relié une somme mobile dépendant du temps de retour à la pression du système. Pour être utilisée en estimation, il faudrait remplacer la mesure des cylindres par un estimateur de celle-ci. Les inégalités de typicité obtenues dans la section II.2 devraient permettre d'obtenir ainsi un estimateur exploitable.
Dans la section II.2.3, nous avons présenté un estimateur de la fonction potentiel du système. Cet estimateur présente deux inconvénients : d'une part il est très irrégulier, on souhaiterait donc disposer d'une version "lissée" ; d'autre part, il suppose la connaissance de la partition d'inversibilité de T^k, ce qui n'est pas réaliste si l'on veut faire de la modélisation. Pour parer à cet inconvénient, il faudrait utiliser des

voisinages des points x qui ne dépendent pas de T ou qui en dépendent mais peuvent être construit à partir de la connaissance d'une orbite finie.

2. Analyse d'algorithmes

Dans la partie III, nous avons développé la problématique de l'analyse d'algorithmes et expliqué les méthodes d'analyse dynamique. À la suite des travaux exposés dans cette partie, mes projets dans ce domaine portent sur l'analyse en distribution d'algorithmes sur les "bits de poids faibles" (en collaboration avec Brigitte Vallée et Benoit Daireaux) et des algorithmes "auto-organisés" type "move-to-front" (en collaboration avec Servet Martìnez et Frédéric Chazal).

2.1. **Analyse dynamique des algorithmes de "cache" ; recherche "auto-organisée".** Il s'agit d'étudier d'un point de vue dynamique et probabiliste divers problèmes d'allocation de ressources. Les algorithmes étudiés sont de deux types. Tout d'abord, les algorithmes de "cache" ont pour but de maintenir un accès rapide à un grand nombre de données en gardant un petit "cache" de données auquel on peut accéder très rapidement. Les requêtes de données sont adressées suivant une distribution généralement fixée et le contenu du cache est (éventuellement) modifié en fonction de ces requêtes. Un problème classique consiste alors à étudier la distribution de probabilité de l'événement "coûteux" : "la donnée requise n'est pas dans le cache". Ensuite, les algorithmes de recherche "auto-organisée" opèrent sur une liste de fichiers qui est mise à jour de la manière suivante : à chaque unité de temps, un fichier est requis et déplacé en tête de liste (règle Move-To-Front). Le coût de la requête est la position du fichier requis dans la liste. Le problème est alors d'étudier ce coût en fonction de la façon dont sont générées les requêtes.
Ces deux types d'algorithmes ont été largement étudiés dans le cas où les requêtes sont générées indépendamment les unes des autres suivant une distribution de probabilité pré-définie (voir [FlGaT] par exemple). En pratique ce n'est généralement pas le cas. Par exemple, lors d'une navigation sur le web, un utilisateur ne consulte pas les pages web indépendamment les unes des autres. Il est donc important d'analyser ces algorithmes dans un cadre plus général. Nous nous intéressons au cas où les requêtes sont générées par une source dynamique ou par un système dynamique plus général. Nous nous plaçons ainsi dans un cadre très général englobant les requêtes produites par une source de Bernoulli ou une chaîne de Markov. Le comportement des algorithmes peut alors être modélisé par des systèmes dynamiques dits "produits croisés". Nous avons pour objectif (en collaboration avec F. Chazal et S. Martìnez) de généraliser à notre contexte les résultats déjà connus dans le cas indépendant grâce à l'étude des propriétés de ces systèmes.

2.2. **Analyse en distribution.** La section III.3.2 était consacrée à l'analyse en moyenne de coûts “à croissance modérée” pour l'algorithme de calcul de pgcd sur les bits de poids faible. Nous avons aussi obtenu la distribution asymptotique de la norme des continuants. En collaboration avec B. Daireaux et B. Vallée nous poursuivons ce travail dans deux directions : l'analyse en distribution des coûts à croissance modérée et l'analyse (en moyenne puis en distribution) de coûts plus généraux. Lorsque les coûts ne sont plus à croissance modérée, on suspecte que la distribution asymptotique n'est plus normale mais α-stable.

Références

[AVa] A. Akhavi, B. Vallée, *m Average bit-complexity of Euclidean algorithms. Automata, languages and programming* (Geneva, 2000), 373–387, Lecture Notes in Comput. Sci., 1853, Springer, Berlin, 2000.

[BaKe] V. Baladi & G. Keller. *Zeta functions and transfer operator for Piecewise monotone transformations.* Commun. Math. Phys. (1990), **127**, 459-477.

[BaVa] V. Baladi, B. Vallée *Euclidean algorithms are Gaussian.* J. Number Theory, **110**, (2005), no. 2, 331–386.

[BaVi] V. Baladi, M. Viana *Strong stochastic stability and rate of mixing for unimodal maps.* Ann. Sci. École Norm. Sup. (1996) **29**, 4, 483-517.

[BaY] V.Baladi, L.-S. Young *On the spectra of randomly perturbed expanding maps.* Comm. Math. Phys. (1993) **156**, 2, 355-385.

[BGR] A.D. Barbour, R. Gerrard, G. Reinert, *Iterates of expanding maps.* Probab. Theory Related Fields 116 (2000), no. 2, 151–180.

[BRY] A. Barron, J. Rissanen, B. Yu, *The minimum description length principle in coding and modeling.* Information theory : 1948–1998. IEEE Trans. Inform. Theory 44 (1998), no. 6, 2743–2760.

[BeY] M. Benedicks & L.-S. Young *Decay of correlations for certain Henon maps.* (1996) preprint.

[BeC] M. Benedicks, L. Carleson *On iterations of* $1 - ax^2$ *on (-1,1)* Ann. of Math. (2), 122, (1985), 1–25

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

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

[BouL] P. Bougerol, J. Lacroix, *Products of random matrices with applications to Schrödinger operators.* Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985.

[Bow] R. Bowen *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Lect. Notes in Math. (1975), **470** Springer Verlag.

[Bre] X. Bressaud, *Subshifts on an infinite alphabet.* Ergodic Theory Dynam. Systems **19** (1999), no. 5, 1175–1200.

[BreGaF] X Bressaud, R. Fernàndez, A. Galves, *Decay of correlations for non-Hölderian dynamics. A coupling approach.* Electron. J. Probab. 4 (1999), no. 3, 19 pp.

[Broi] A. Broise, *Transformations dilatantes de l'intervalle et théorèmes limites.* Études spectrales d'opérateurs de transfert et applications. Astérisque 1996, no. 238, 1–109.

[Brow1] Browkin, J. *Continued fractions in local fields. II,* Math. Comp. 70 (2001), no. 235, pp.1281-1292.

[Brow2] Browkin, J. *Continued fractions in local fields. I,* Demonstratio Math. 11 (1978), no. 1, pp.67-82.

[Bu1] J. Buzzi Intrinsic ergodicity of affine maps in $[0,1]^d$. *Monatsh. Math.* 124 (1997), 97–118.

[Bu2] J. Buzzi *Markov extensions for multi-dimensional dynamical systems.* Israel J. Math. 112 (1999), 357–380.

[Bu3] J. Buzzi *Thermodynamical formalism for picewise invertible maps : absolutely continuous invariant measures as equilibrium states* in Smooth Ergodic Theory and its Applications (Seattle, WA 1999), Proceedings of Symposia in Pure Mathematics **69**, AMS, RI, (2001), 749–783.

[Bu4] J. Buzzi Absolutely continuous invariant probability measures for arbitrary expanding piecewise **R**-analytic mappings of the plane. *Ergodic Theory Dynam. Systems* **20** (2000), 697–708.

[BuPaS] J. Buzzi, F. Paccaut, B. Schmitt, *Conformal measures for multidimensional piecewise invertible maps.* Ergodic Theory Dynam. Systems 21 (2001), no. 4, 1035–1049.

[Ca] A.A. Castro *Backward inducing and exponential decay of correlations for partially hyperbolic attractors with mostly contracting central direction.* PhD Thesis (1998).

[ChaColSc] J-R. Chazottes, P. Collet, B. Schmitt *Statistical consequences of Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems* à paraître à Nonlinearity

[Che] N. Chernov *Statistical properties of piecewise smooth hyperbolic systems in high dimensions.* Discrete Contin. Dynam. Systems (1999), **5**, 2, 425-448.

[CheMT] N. Chernov, R. Markarian & S. Troubetzkoy. *Conditionally invariant measures for Anosov maps with small holes.*

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

[Col] P. Collet, *Some ergodic properties of maps of the interval.* Dynamical systems (Temuco, 1991/1992), 55–91, Travaux en Cours, 52, Hermann, Paris, 1996.

[ColE] P. Collet, J.-P. Eckmann *On the abundance of aperiodic behaviour for maps on the interval.* Comm. Math. Phys. 73 (1980), no. 2, 115–160.

[ColGaSc] P. Collet, A. Galves, B. Schmitt *Unpredictability of the occurrence time of a long laminar period in a model of temporal intermitency.* Ann. Inst. H. Poincaré Phys. Théor. (1992), **57**, 3, 319-331.

[ColMSc1] P. Collet,S. Martinez, B. Schmitt, *Exponential inequalities for dynamical measures of expanding maps of the interval.* Probab. Theory Related Fields 123 (2002), no. 3, 301–322.

[ColMSc2] P. Collet, S. Martínez & B. Schmitt. *The Pianigiani-Yorke measure for topological Markov chains.* Israel Journal of Math. **97**, 61-70 (1997).

[ColMSc3] P. Collet, S. Martínez & B. Schmitt. *The Pianigiani-Yorke measure and the asymptotic law on the limit Cantor set of expanding systems.* Nonlinearity **7**, 1437-1443 (1994).

[ColMSc4] P. Collet, S. Martínez & B. Schmitt. *Quasi-stationary distribution and Gibbs measure of expanding systems.* Instabilities and nonequilibrium structures, 205-219, (1996).

[Cow] W. Cowieson, *Piecewise smooth expanding maps in* $\mathbb{R}^d$*,* Ph.D. Thesis, U.C.L.A., Berkeley (see `http ://math.usc.edu/˜cowieson`).

[Cs] I. Csiszár, *Large-scale typicality of Markov sample paths and consistency of MDL order estimators.* Special issue on Shannon theory : perspective, trends, and applications. IEEE Trans. Inform. Theory 48 (2002), no. 6, 1616–1628.

[CsS] I. Csiszár, P. C. Shields, *The consistency of the BIC Markov order estimator.* Ann. Statist. 28 (2000), no. 6, 1601–1619.

[Da] B. Daireaaux Thèse de l'université de Caen.

[DaVa] B. Daireaux, B. Vallée *Dynamical Analysis of the Parametrized Lehmer-Euclid Algorithm* Combinatorics, Probability and Computing , 13, (2004), 499-536.

[DedDou] J. Dedecker, P. Doukhan, *A new covariance inequality and applications.* Stochastic Process. Appl. 106 (2003), no. 1, 63–80.

[DedPr] J. Dedecker, C. Prieur, *New dependence coefficients. Examples and applications to statistics.* To appear in Probab. Theory and Relat. Fields.

[Del] H. Delange, *Généralisation du Théorème d'Ikehara,* Ann. Sc. ENS, (1954), **71**, pp 213–242.

[Dem1] M. Demers *Markov extensions for dynamical systems with holes : an application to expanding maps of the interval* , to appear in Israel J. of Math.

[Dem2] M. Demers *Markov extensions and conditionally invariant measures for certain logistic maps with small holes*, to appear in Ergod. Th. and Dynam. Sys.

[DenGS] M. Denker, C. Grillenberger, K. Sygmund *Ergodic theory on compact spaces.* Lecture notes in mathematics **527**, Springer, Berlin, 1976.

[Dol] D. Dolgopyat *On dynamics of mostly contracting diffeomorphisms.* Comm. Math. Phys. 213 (2000), no. 1, 181–201.

[Dou] P. Doukhan *Mixing. Properties and examples.* Lecture Notes in Statistics, 85. Springer-Verlag, New York, (1994).

[DouL] P. Doukhan, S. Louhichi, *m A new weak dependence condition and applications to moment inequalities.* Stochastic Process. Appl. 84 (1999), no. 2, 313–342.

[DouNe] P. Doukhan, M. Neumann. *A Bernstein type inequality for times series* prépublication.

[FeSc] P. Ferrero, B. Schmitt *Ruelle Perron Frobenius theorems and projective metrics.* Colloque Math. Soc. J. Bolyai Random Fields. Estergom (Hungary) (1979).

[Fl1] P. Flajolet, *m Analytic analysis of algorithms.* Automata, languages and programming (Vienna, 1992), 186–210, Lecture Notes in Comput. Sci., **623**, Springer, Berlin, 1992.

[Fl2] P. Flajolet, Personal communication.

[FlGaT] P. Flajolet, D. Gardy, L. Thimonier. Birthday paradox, coupon collectors, caching algorithms and self-organizing search, *Discrete Applied Mathematics 39 (1992) 207-229, North Holland.*

[FlKT] P. Flajolet, P. Kirschenhofer, R.F. Tichy, *Deviations from uniformity in random strings.* Probab. Theory Related Fields 80 (1988), no. 1, 139–150.

[FlSe] P. Flajolet, R. Sedgewick *Analytic Combinatoric* livre en préparation, chapitre disponibles à `http ://algo.inria.fr/flajolet/Publications/books.html`.

[GoBo] P. Góra, A. Boyarsky Absolutely continuous invariant measures for piecewise expanding C^2 transformation in R^N. *Israel J. Math.* **67** (1989), 272–286.

[GalS] A. Galves & B. Schmitt *Inequalities for hitting time in mixing dynamical systems.* Random and Computational Dynamics (1997), **5**, 4, 337-347.

[GasW] P. Gaspard & X.J. Wang . Proc. Math. Acad. Sci. USA (1988) **85** 4591.

[Gor] M.I. Gordin *The central limit theorem for stationary processes.* Soviet. Math. Dokl. (1969), **10**, (5), 1174-1176.

[HenHer] H. Hennion, L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness.* Lecture Notes in Mathematics, 1766. Springer-Verlag, Berlin, 2001.

[HoKe] F. Hofbauer, G. Keller *Ergodic properties of invariant measures for piecewise monotonic transformations* Math. Z. **180**, (1982), 119–140.

[Ho] F. Hofbauer *Piecewise invertible dynamical systems.* Probab. Theory Relat. Fields **72** (1986), no. 3, 359–386.

[Hw] H.-K. Hwang *On convergence rates in the central limit theorem for combinatorial structures*, European Journal of Combinatorics, **19**, (1998), pp 329–343.

[IL] I. M. Ibragimov, Y. Linnik *Independant and stationary sequences of random variables* Walters-Noardhoff Pub. Groningen (1971)

[KaNe] R. Kallabis, M. Neumann *A Bernstein inegality under weak dependance* prépublication.

[Ke] G. Keller *Markov extensions, zeta-functions, and Fredholm theory for piecewise invertible dynamical systems* Trans. Amer. Math. Soc. 314(1989), 433-497.

[KeNo] G. Keller, T. Nowicki *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps.* Comm. Math. Phys. (1992), **149**, 1, 31-69.

[Kn] D .E. Knuth, *Theart of computer programming. Vol. 2. Seminumerical algorithms.* Third edition. Addison-Wesley Series in Computer Science and Information Processing. Addison-Wesley Publishing Co., Reading, Mass., 1998.

[LaY1] A. Lasota, J. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **186** (1973), 481–488.

[LaY2] A. Lasota, J. Yorke, *Exact dynamical systems and the Frobenius-Perron operator.* Trans. Amer. Math. Soc. **273**, (1982), no. 1, 375–384.

[Li1] C. Liverani *Decay of correlations.* Ann. of Math. (1995), **142** (2), 239-301.

[Li2] C. Liverani *Central limit theorem for deterministic systems.* Proceedings of the International Congress on Dynamical Systems, Montevideo 95, Research Notes in Mathematics series, Pittman, (1997).

[Li3] C. Liverani, *Decay of correlations for piecewise expanding maps.* J. Statist. Phys. 78 (1995), no. 3-4, 1111–1129.

[LiSV] C. Liverani, B. Saussol, S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems.* Ergodic Theory Dynam. Systems 18 (1998), no. 6, 1399–1420.

[MU] R.D. Mauldin, M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet.* Israel J. Math. **125**, (2001), 93–130.

[Nu] R. Nussbaum, *The radius of the essential spectrum.* Duke Math. J., **37**, (1970), 473–478.

[OW] D. Ornstein, B. Weiss, Entropy and data compression schemes, IEEE Trans. Inf. Theory 39 (1993), 78-83.

[Sar] O. Sarig *Thermodynamic Formalism for Countable Markov Shifts.* (1997).

[SauW] B.Saussol, J.Wu *Recurrence spectrum in smooth dynamical system* Nonlinearity, **16**, (2003) 1991-2001.

[Pac] F. Paccaut *Propriétés statistiques de systèmes dynamiques non markoviens* Ph D Thesis Université de Bourgogne (2000). Available at `http ://www.lamfa.u-picardie.fr/paccaut/publi.html`

[ParPo] W. Parry, M. Pollicott *Zeta functions and the periodic orbit structure of hyperbolic dynamics.* Asterisque (1990), **187-188**

[PiY] G. Pianigiani & J.A. Yorke. *Expanding maps on sets which are almost invariant : decay and chaos.* Trans. Amer. Math. Society **252**, 433-497 (1989).

[Po] M. Policott *Rates of mixing for potentials of summable variation.* Trans. Amer. Math. Soc. 352 (2000), no. 2, 843–853.

[Ris] J. Rissanen, non k-Markov points.*Stochastic complexity in statistical inquiry.* World Scientific Series in Computer Science, 15. World Scientific Publishing Co., Inc., Teaneck, NJ, 1989. vi+178 pp.

[Rio] E. Rio, *Inégalités de Hoeffding pour les fonctions lipschitziennes de suites dépendantes.*

[Ru] D. Ruelle *Thermodynamic formalism.* Addison Wesley New-York (1978).

[Sau] B. Saussol Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116** (2000), 223–248.

[Sc] B. Schmitt, *Ergodic theory and thermodynamic of one-dimensional Markov expanding endomorphisms.* Dynamical systems (Temuco, 1991/1992), 93–123, Travaux en Cours, 52, Hermann, Paris, 1996.

[StZ] D. Stehlé, P. Zimmermann, *A Binary Recursive Gcd Algorithm.* ANTS 2004, Lect. Notes in Comput. Science **3076**, (2204), 411-425.

[U] M. Urbański, *m The Hausdorff dimension of the set of points with non-dense orbit under a hyperbolic dynamical system.* Nonlinearity, **4**, (1991), no. 2, 385–397

[Va1] B. Vallée *Dynamical sources in information theory : fundamental intervals and word prefixes* Algorithmica, **29**, 262-306, (2001).

[Va2] B. Vallée *Dynamics of the binary Euclidean algorithm : functional analysis and operators. Average-case analysis of algorithms.* Algorithmica, **22**, (1998), no. 4, 660–685.

[Va3] B. Vallée *Digits and continuants in Euclidean algorithms. Ergodic versus Tauberian theorems.* J. Théor. Nombres Bordeaux 12 (2000), no. 2, 531–570.

[Va4] B. Vallée *m Dynamical analysis of a class of Euclidean algorithms.* Theoret. Comput. Sci. **297**, (2003), no. 1-3, 447–486.

[VeJ] D. Vere-Jones. *Geometric ergodicity in denumerable Markov chains.* Quart. J. Math. **13**, 2, 2, 7-28 (1962).

[W] P. Walters *Invariant measures and equilibrium states for somme mappings which expand distances.* Trans. Amer. Math. Soc. (1978), **236**, 121-153.

[Y1] L.-S. Young *Statistical properties of dynamical systems with some hyperbolicity.* Ann. of Math. (2) (1998), **147**, 3, 585-650.

[Y2] L.-S. Young *Recurrence times and rates of mixing.* Israel J. Math. 110 (1999), 153–188.

Liste des travaux

(les articles portant une * sont relatifs à la thèse)

Articles parus ou acceptés (avec procédure de referee)

[1] B. Daireaux, V. Maume-Deschamps, B. Vallée *The Lyapunov Tortoise and the dyadic Hare*, 2005 international Conference on Analysis of Algorithms, Discrete Matematics and Theoritical Computer Science, proc. **AD**, (2005), 71–94.

[2] V. Maume-Deschamps *Concentration inequalities and estimation of conditional probabilities* accepté pour publication à Lect. notes in Stat., proceedings of StatDep 2005.

[3] J. Buzzi, V. Maume-Deschamps *Decay of correlations on towers for potentials with summable variation*, Discrete and Continuous Dynamical Systems, 12, (2005), no 4, 639-656.

[4] F. Chazal, V. Maume-Deschamps, *General Markov Dynamical sources : applications to information theory*, Discrete Mathematics and Theoretical Computer Science, 6, (2004), no 2, 283-314.

[5] F. Chazal, V. Maume-Deschamps, B. Vallée *Erratum to "Dynamical sources in Information Theory : Fundamentals Intervals and Word Prefixes" by B. Vallée.* Algorithmica 38 (2004), no. 4, 591-596.

[6] V. Maume-Deschamps, B. Schmitt, M. Urbanski, A. Zdunik *Pressure and recurrence.* Fund. Math. 178 (2003), no. 2, 129-141.

[7] C. Liverani, V. Maume-Deschamps *Lasota-Yorke maps with holes : conditionally invariant probability measures and invariant probability measures on the survivor set*, Ann. IHP, 39, no 3, (2003), 385-412.

[8] C. Bose, V. Maume-Deschamps, B. Schmitt, S. Shin *Invariant measures for piecewise convex transformations of an interval.* Studia Mathematica, 152, (2002), no 3, 263-297.

[9] J. Buzzi, V. Maume-Deschamps *Decay of correlations for piecewise invertible maps in higher dimensions*, Israel J. Math. 131 (2002), 203-220.

[10] V. Baladi, M. Benedicks, V. Maume-Deschamps *Almost sure rates of mixing for i.i.d. unimodal maps* Ann. Sci. École Norm. Sup. (4) 35 (2002), no. 1, 77-126.
V. Baladi, M. Benedicks, V. Maume-Deschamps, *Corrigendum : "Almost sure rates of mixing for i.i.d. unimodal maps" [Ann. Sci. École Norm. Sup. (4) 35 (2002), no. 1, 77–126 ; MR1886006 (2003d :37027)].* Ann. Sci. École Norm. Sup. (4) 36 (2003), no. 2, 319–322.

[11] V. Maume-Deschamps *Projective metrics and mixing properties on towers*, Trans. A.M.S. **353**, (2001), 8, 3371-3389.

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

[13] P. Collet, S. Martinez, V. Maume-Deschamps *Conditionally Invariant Probability Measures in Dynamical Systems.* Nonlinearity, **13**, (2000), 1263-1274.
P. Collet, S. Martinez, V. Maume-Deschamps *Corrigendum : "On the existence of conditionally invariant probability measures in dynamical systems"* Nonlinearity , **17**, (2004), 1985-1987.

[14] * V. Maume-Deschamps *Equilibrium states for non hölderian random dynamical systems.* Random and Computational Dynamics (1997) **5** (4) 319-335.

[15] * A. Kondah, V. Maume, B. Schmitt *Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non höldériennes.* Ann. Inst. Poincaré Sec. Prob. Stat. (1997) **33** (6) 675-695.

[16] * A. Kondah, V. Maume, B. Schmitt *Dynamique symbolique : vitesse de convergence vers l'état d'équilibre.* C. R. Acad. Sci. Paris (1996), **323**, 393-396.

Prépublications, actes.

[17] V. Maume-Deschamps *Estimation of the pressure function in dynamical systems* en préparation.

[18] X. Bressaud, V. Maume-Deschamps *Ergodic theorems for group (or semi-group) actions.* Prépublication de l'Institut de Mathématiques de Bourgogne.

[19] F. Chazal, V. Maume-Deschamps, B. Vallée *Notes du cours de V. Baladi et B. Vallée, ALEA 2002 : Systèmes dynamiques et algorithmique*, Algorithms Seminar 2001-2002, F. Chyzak Ed., INRIA Res. Report 5003, November 2003, pp. 121–150.

[20] V. Baladi, M. Benedicks, V. Maume-Deschamps *Decay of random correlation functions for unimodal maps*, Proceedings 31st Symposium on Mathematical Physics, May 18-21 1999, Torun, Reports on Mathematical Physics, 46, 15-26 (2000).

[21] * V. Maume-Deschamps *Propriétés de mélange pour des sytèmes dynamiques markoviens*, thèse de l'université de Bourgogne, (1998).