Large Deviations Results for Spatially Extended Dynamical Systems

THÈSE

PRÉSENTÉE AU DÉPARTEMENT DE MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES
PAR

Jean-Baptiste BARDET

Mathématicien diplômé de l'Université de Lyon I
originaire de France

acceptée sur proposition du jury:

Prof. Anthony C. Davison, président du jury
Prof. Gérard Ben Arous, directeur de thèse
Prof. Viviane Baladi, rapporteur
Prof. Patrick Cattiaux, rapporteur
Prof. Charles-Edouard Pfister, rapporteur
Prof. Hans-Henrik Rugh, rapporteur

Lausanne, EPFL
2002

REMERCIEMENTS

C'est une grande chance d'avoir pu faire ma thèse sous la direction de Gérard Ben Arous. Dès les stages de maîtrise et de DEA, et plus encore pendant cette thèse, il m'a fait connaître le monde de la recherche en mathématiques comme un domaine riche, exigeant et fascinant. Travailler avec lui à Lausanne m'a permis de découvrir un vaste champ des mathématiques, des probabilités aux systèmes dynamiques. En plus de mon sujet de thèse proprement dit, je le dois aussi aux nombreuses activités de sa chaire (séminaires, groupes de travail, ...) et à son talent pour transmettre ses vastes connaissances. Je tiens aussi à le remercier pour son enthousiasme communicatif et la confiance qu'il a su me faire au cours de ce travail.

Je remercie chaleureusement Viviane Baladi, qui a accepté de faire partie de mon jury, après avoir joué un rôle important tout au long de ma thèse : comme référente en théorie ergodique, puisqu'elle a toujours accepté de répondre à mes questions sur ce domaine, mais aussi pour toutes les conférences où elle m'a proposé d'aller, en particulier l'Ecole d'Hiver de Sils-Maria, qu'elle organisait et qui a contribué à ma découverte du sujet.

Je suis très reconnaissant à Hans-Henrik Rugh d'avoir "écrit le bon article au bon moment", et surtout de m'avoir accordé beaucoup de temps et d'encouragements lorsque j'ai voulu lire et développer cet article pour obtenir une partie des résultats de cette thèse. Je le remercie aussi d'avoir accepté d'être dans mon jury.

Je remercie vivement Patrick Cattiaux de l'intérêt qu'il a montré pour ma recherche, dès mon premier passage à Nanterre et jusqu'à sa présence dans mon jury. Toutes ses remarques et propositions nourriront certainement ma recherche à venir.

Je remercie également Charles Pfister pour sa disponibilité lors de mes questions et sa présence dans mon jury, ainsi qu'Anthony Davison, qui a accepté de présider ce jury.

J'ai préparé l'essentiel de cette thèse à l'EPFL, où existait un très actif et sympathique groupe de thésards en probabilités. Je remercie chacun d'entre eux ainsi que tous les collègues du DMA. Anabela Querino s'est montrée particulièrement attentive et efficace pour m'aider à régler tous les problèmes administratifs et techniques liés à cette thèse, je l'en remercie vivement. Cette

thèse doit aussi beaucoup à Jirka et Sacha avec qui j'ai partagé beaucoup de balades à vélo ou en jazz, de bières et de maths, et qui m'ont souvent aidé à parer à mes insuffisances informatiques ou à mon éloignement de Lausanne.

En effet, j'ai passé la dernière année de cette thèse à l'Université de Nanterre, dans le laboratoire MODAL'X. J'apprécie énormément l'ambiance de cette équipe, où je me suis senti immédiatement intégré. Je remercie tous les Nanterrois pour leur accueil, et en particulier Sylvie Méléard pour son dynamisme permanent et contagieux.

J'ai aussi profité pendant cette thèse de plusieurs congrès, conférences et invitations, et tiens à remercier tous les organisateurs et les chercheurs que j'ai pu y rencontrer, en particulier X. Bressaud, J.R. Chazottes, J.D. Deuschel, G. Keller, Y. Kifer, C. Liverani, A. Le Ny et M. Viana. Beaucoup de ces rencontres étaient organisées et financées dans le cadre du programme PRODYN de l'ESF.

Cette thèse s'est développée en même temps que notre vie à deux avec Magali, ce qui donne à cette dernière un rôle bien particulier dans l'aboutissement de ce travail. Elle a supporté avec patience et m'a aidé à faire fructifier les moments de découragement quand ma thèse n'avançait pas comme les passages d'euphorie irraisonnée quand elle avançait. Je ne la remercierai jamais assez ici, mais compte bien me rattraper en la soutenant à mon tour dans son propre travail de thèse, et en vivant longtemps et heureux avec elle.

Il me reste à remercier tous ceux avec qui j'ai partagé des moments intenses de ma vie durant cette thèse, en particulier mes parents, Louise, Marie et Eric, Pierre-Luc, les parents de Magali et Nicolas, Thierry, Emilie, Marjo et Bruno, Marc, les copains du MRJC, et beaucoup d'autres...

ABSTRACT

This Ph-D Thesis is devoted to the study of limit theorems for coupled map lattices, which are models of discrete-time dynamical systems on lattices.
We consider a system constituted by expanding maps of the circle under a small coupling. We prove that the associated spatiotemporal measure satisfies under Lebesgue measure a Large Deviations Principle. The rate function is expressed in the setting of thermodynamic formalism and with a potential which appears in the main step of the proof - a Volume Lemma result - to describe sharp metric estimates of the system.
We study also the temporal asymptotics of the same coupled map lattices. Under stronger regularity assumptions, we prove that the temporal empirical mean of any regular enough observable satisfies a Central Limit Theorem and a Moderate Deviations Principle. We also establish a partial Large Deviations result, which implies in particular an exponential rate of convergence to equilibrium.

RÉSUMÉ

Nous étudions dans cette thèse des théorèmes limites pour des réseaux d'applications couplées, qui sont des modèles de systèmes dynamiques en temps discrets et sur réseau.
Nous considérons un réseau d'applications dilatantes du cercle avec faible couplage et prouvons que la mesure empirique spatiotemporelle associée satisfait, sous la mesure de Lebesgue, un Principe de Grandes Déviations. La fonction de taux s'exprime dans le cadre du formalisme thermodynamique, à l'aide d'un potentiel qui apparaît dans l'étape principale de la preuve - un résultat du type Lemme de Volume - pour décrire des estimées métriques fines sur le système.
Nous étudions aussi le comportement asymptotique en temps des mêmes réseaux d'applications couplées. Sous des hypothèses de régularité plus fortes, nous montrons que la moyenne empirique temporelle d'observables assez régulières satisfait sous la mesure de Lebesgue un Théorème Central Limite et un Principe de Déviations Modérées. Nous établissons aussi un résultat partiel de Grandes Déviations, qui implique en particulier la convergence exponentielle vers l'équilibre.

CONTENTS

INTRODUCTION

This thesis is devoted to the study of limit results, and in particular large deviations principles, for models of spatially extended dynamical systems, namely the coupled map lattices.
In this introduction, we briefly present the model and the results obtained. We insist on the links with existing results and on the perspectives offered by this work.

Coupled map lattices

Coupled map lattices are models of spatially extended dynamical systems, introduced by Kaneko in 1983. They consist of discrete time evolutions on an infinite product of spaces, indexed by a discrete lattice: the evolution at each step of time is the composition of local dynamics on each site and of coupling between sites.

Let us give a first example to make things more precise. It will also give us a simple setting for the presentation of our results. We work on $\mathcal{X} = (S^1)^{\mathbb{Z}}$, where we consider $S^1 = \mathbb{R}/\mathbb{Z}$ for the notations, and define on each site the local dynamics by an expanding map of the circle f,

$$|f'(x)| \geq \lambda > 1 \quad \forall x \in S^1.$$

A simple coupling is given by $G_\varepsilon : \mathcal{X} \to \mathcal{X}$ defined by

$$(G_\varepsilon(x))_i = (1-\varepsilon)x_i + \frac{\varepsilon}{2}x_{i-1} + \frac{\varepsilon}{2}x_{i+1}.$$

This is a diffusion operator with coupling strength ε (for $\varepsilon = 0$, $G_0 = \mathrm{Id}$). This example is used extensively for couplings between maps of the interval, in particular for numerical studies. However, in the context of circle maps, it is not well defined, and could for example be replaced by $(G_\varepsilon(x))_i = x_i + \frac{\varepsilon}{4\pi}\sin(2\pi(x_{i-1}-x_i)) + \frac{\varepsilon}{4\pi}\sin(2\pi(x_{i+1}-x_i))$, which gives a smooth adaptation of it.

The coupled map lattice is then $F = F_\varepsilon = G_\varepsilon \circ F_0$, where F_0 is the uncoupled map defined by $(F_0(x))_i = f(x_i)$. We want to study the iterations of this map F.

To see where the interest of this model is, we compare it to other evolution equations.

Consider for example a partial differential equation of the type

$$\frac{\partial u}{\partial t} = \varepsilon \frac{\partial^2 u}{\partial x^2} + \Phi(u),$$

with $\Phi : \mathbb{R} \to \mathbb{R}$. This is called a reaction-diffusion equation, where Φ governs the local reaction and $\frac{\partial^2 u}{\partial x^2}$ imposes diffusion.
Our coupled map can be seen as a discretized equivalent of a solution $u_t(x)$ of this equation: the reaction is localized on S^1 and expressed by the map f. The diffusion term becomes G_ε. These two effects are completely separated in two functions, and time is also discretized.

Our model can also be compared to a cellular automata, which is a map acting on $\{0, \ldots, N\}^{\mathbb{Z}}$, for example the well known example of the game of the life. For cellular automata, there are no interesting local dynamics, and in particular no local chaotic behavior, since the state space is finite. All the dynamics comes from the the couplings between sites, whereas for partial differential equations and coupled map lattices the interest lies in the competition between the disorder of the local dynamics and the organization due to the coupling.

If the initial project of using coupled map lattices as approximations of partial differential equations is far from having been completed yet, the intermediate complexity of coupled map lattices gives rise to many interesting phenomena. There occurs for instance at small coupling spatiotemporal chaos, with strong decorrelation in time and space. This gives a simpler version of turbulence observed in fluid dynamics.
For some choices of local maps, and for stronger coupling, very coherent structures appear with an intermediate regime of intermittency.

All these behaviors are briefly described in Chapter 1, with references to the rich literature in this field. We also give a review of existing results in the mathematical study of spatiotemporal chaos.

Since the aim of this thesis is to develop large deviations results for the asymptotic study of coupled map lattices, we present in Chapter 2 the existing results for the temporal behavior of dynamical systems on one site.

The remainder of the thesis is devoted to our results, which we present below.

Spatiotemporal description

We study in Chapter 3 the limit behavior of the spatiotemporal empirical measures associated to the coupled map F

$$R_T(x) = \frac{1}{T(2T+1)} \sum_{\substack{0 \le t < T \\ -T \le i \le T}} \delta_{S^i \circ F^t(x)} \in \mathcal{M}^1(\mathcal{X}),$$

where S denotes the spatial shift (defined by $(Sx)_i = x_{i+1}$) and $\mathcal{M}^1(\mathcal{X})$ the space of probability measures on $\mathcal{X}$.

Theorem 3.1.2 states that under the explicit assumptions of weak coupling (3.7), R_T satisfies under initial measure $\overline{m}$, the product of Lebesgue measures on the circles, a large deviations principle with rate function

$$I_{\mathrm{st}}(\mu) = \begin{cases} -h_{(F,S)}(\mu) - \int_{\mathcal{X}} \varphi \, d\mu & \text{if } \mu \text{ is invariant by } F \text{ and } S, \\ +\infty & \text{otherwise,} \end{cases}$$

with $h_{(F,S)}$ the metric entropy associated to the 2-dimensional dynamical system (F, S) and φ a potential associated to the dynamics: $-\varphi$ plays the role of the logarithm of Jacobian per site of F. This result is true for a larger class of systems (with more general couplings and in any dimension) and means, roughly, that

$$\overline{m}\{x \,:\, R_T(x) \sim \mu\} \sim \exp\Big(T(2T+1)\Big(h_{(F,S)}(\mu) + \int_{\mathcal{X}} \varphi \, d\mu\Big)\Big).$$

This implies in particular that R_T converges exponentially fast to the set of equilibrium measures associated to φ

$$\mathrm{EQ}(\varphi) = \{\nu \in \mathcal{M}^1(\mathcal{X}) \,:\, h_{(F,S)}(\nu) + \int_{\mathcal{X}} \varphi \, d\nu = 0\}.$$

This large deviations principle is related to the series of papers presented in Section 1.3.1: the potential φ is constructed in [53], where Jiang and Pesin prove also that under weak coupling assumptions, there is a unique equilibrium measure ν associated to φ, which is spatiotemporally mixing,

$$\int_{\mathcal{X}} g \circ F^t \circ S^n \cdot h \, d\nu \longrightarrow \int_{\mathcal{X}} g \, d\nu \int_{\mathcal{X}} h \, d\nu,$$

when t or $|n|$ tends to infinity.

In this range of weak coupling, our large deviations result is a refinement stating exponential convergence to the equilibrium measure ν.

However, our method does not use this result of uniqueness of equilibrium measure, neither the coding by a Gibbs measure on a shift system which is the key tool for all proofs of uniqueness of equilibrium measures. Our proof is developed directly in the thermodynamic formalism associated to the system (F, S). The main step of the proof is a "Volume Lemma" result, which is interesting by itself. It is stated as Theorem 3.1.1 and says that the potential φ describes the size under $\overline{m}$ of sets of points whose orbit follows a fixed orbit on given time and space.

Such a spatiotemporal large deviations principle is hence similar, concerning the result and the method of proof, to large deviations for Gibbs measures on shift systems, proven in [42, 31, 81]. These results are valid in great generality and are in particular known to occur in phase transition situations, i.e. where there are at least two Gibbs measures.

In our case, we must however work under restrictive assumptions: we need weak coupling to preserve expansivity of the map or even to construct the adequate potential φ. We don't know if our assumptions could cover a case with more than one equilibrium measure.

More generally, the situation of a stronger coupling needs to be clarified: we do not know if organization of the system into coherent structures is really linked to phase transitions in the sense of equilibrium measures or to other effects as, for example, bifurcations to stable periodic orbits.

We think that a step further in the comprehension of thermodynamic formalism for coupled map lattices can be done by developing a description of equilibrium measures as Gibbs measures. A possible starting point is the Gibbsian formalism introduced by Haydn and Ruelle [45, 94] for single maps satisfying expansiveness and specification (see a presentation of this formalism in Section 2.3.4). This is one of our aims for future research.

Temporal description

We adopt in Chapter 4 a different approach and study the behavior of the system under the temporal dynamics F only. We require that the local map is holomorphic in a neighborhood of the circle, and we need similar regularity assumptions on the coupling, which are satisfied by the example of this Introduction.

Most papers presented in Section 1.3.2 are concerned with this setting: Bricmont and Kupiainen [12], Baladi et al [4], and more recently Rugh [95] obtained the existence and uniqueness in a restricted set of measures of an invariant measure ν which is locally absolutely continuous with respect to Lebesgue measure and mixing.

We work under the same assumptions as in [95] and consider for any regular enough observable $u : \mathcal{X} \to \mathbb{R}$ the asymptotic behavior under various scalings of the empirical mean of u,

$$S_T u(x) = \sum_{t=0}^{T-1} u \circ F^t(x).$$

We write $m_u = \int_{\mathcal{X}} u \, d\nu$ and show in Theorem 4.2.3 that if u can not be written as $u = g - g \circ F$ with $g \in L^2(\nu)$, then there exists a positive constant σ_u^2 such that:

- Central Limit Theorem:

$$\left(\frac{S_T u - T m_u}{\sqrt{T}\sigma_u}\right)^* (\overline{m}) \xrightarrow{\text{Law}} \mathcal{N}(0,1);$$

- Moderate Deviations Principle: for any $1/2 < \alpha < 1$

$$\overline{m}\left\{x \,:\, \frac{S_T u(x) - T m_u}{T^\alpha} \sim a\right\} \sim \exp\Big(-T^{2\alpha-1}\frac{a^2}{2\sigma_u^2}\Big).$$

These two results describe the small and moderate deviations (of order T^α with $1/2 \leq \alpha < 1$) of the partial sum S_T around $T m_u$.

We also obtain, in Theorem 4.2.2, a partial large deviations result, which describes the fluctuations of order T: there is a convex rate function I_u with a unique zero at m_u and there are $a_u < m_u < b_u$ such that $S_T u/T$ satisfies under $\overline{m}$ a complete large deviations upper bound and a partial large deviations lower bound on the interval (a_u, b_u) with rate function I_u. This implies in particular that $S_T u/T$ converges exponentially fast to m_u: if A is such that $m_u \notin \bar{A}$, then

$$\limsup_{T\to\infty} \frac{1}{T} \log \overline{m} \left\{ z \;:\; \frac{S_T u(z)}{T} \in A \right\} < 0.$$

This is another refinement of the ergodic behavior of $S_T u/T$.

Proofs of these results rely on a modification of the arguments of [95] to construct perturbed transfer operators associated to observables U. A perturbation argument allows to preserve the spectral gap property proven in [95] to small observables. The rate function is then obtained around the mean m_u as the Legendre transform of the logarithm of the main eigenvalue of a perturbed operator. And the main limitation of our result is that we do not get an explicit form for this rate function. This limitation is strongly related to the temporal viewpoint adopted in this work: we can not express natural objects in terms of thermodynamic formalism since the infinite dimensional setup makes that the reference potential $-\log\det DF$ does not make sense and metric entropy is often infinite.

A comparison of both large deviations results on the trivial example of the uncoupled map F_0 clarifies the essential differences between them. In this case, one easily obtains a large deviations principle for the temporal empirical measure

$$L_T(x) = \frac{1}{T} \sum_{t=0}^{T-1} \delta_{F_0^t(x)}$$

by a projective limit method. $L_T(x)$ satisfies a large deviations principle with rate function

$$I_{\mathrm{temp}}(\mu) = \lim_{\Lambda\to\mathbb{Z}^d} I_\Lambda(\mu_\Lambda) = \lim_{\Lambda\to\mathbb{Z}^d} \left(-h_{F_{0,\Lambda}}(\mu_\Lambda) + \int \log\det DF_{0,\Lambda}\, d\mu_\Lambda \right)$$

if μ is F-invariant and $I_{\mathrm{temp}}(\mu)$ is infinite otherwise. For $\mu = \otimes_{i\in\mathbb{Z}^d}\mu_i$ a F-invariant product measure, we get the simpler expression

$$I_{\mathrm{temp}}(\mu) = \sum_{i\in\mathbb{Z}^d} \left[-h_f(\mu_i) + \int \log f'\, d\mu_i \right] = \sum_{i\in\mathbb{Z}^d} I_i(\mu_i),$$

whereas the rate function I_{st} governing large deviations of the spatiotemporal empirical measure simplifies in this particular case to

$$I_{\mathrm{st}}(\mu) = -h_{(F_0,S)}(\mu) + \int_{\mathcal{X}} \log f'\, d\mu_0 = \begin{cases} I_0(\mu_0) & \text{if } \mu_i = \mu_0 \text{ for all } i \in \mathbb{Z}, \\ \infty & \text{otherwise.} \end{cases}$$

This simple example shows that both large deviations results obtained in this thesis give two radically different images of the system: the spatiotemporal rate function sees (i.e. is finite on) only those measures which are invariant under the spatial shift whereas the only shift-invariant measure seen by the temporal rate function is the product of equilibrium measures.

Much remains to be done to better understand the temporal result and its associated variational principle: there exists nowadays no proof that the rate function in the coupled case remains the limit of the corresponding rate functions on finite lattices. This analysis would help to obtain an identification of the rate function or of the associated pressure with intrinsic quantities characterizing the system, or to define a new complete variational principle.

Other perspectives

We think it will be an important step for the comprehension of collective dynamics in coupled map lattices to derive the same kind of large deviations results in the mean field version of this model, the globally coupled map.

One can indeed hope that, as in Statistical Mechanics, mean field models are simpler to study than those with local interactions, because the geometry of the system is made simpler. We know also from Statistical Mechanics that large deviations can play a great role in the detection of phase transitions. Furthermore, in their recent preprint [8], Blank and Bunimovich claim that phase transitions situations are easier to obtain for globally coupled maps.

Another direction of research is to develop large deviations estimates for single site maps with a non-uniformly hyperbolic behavior.

The example of a map of the interval with an indifferent fixed point presented in Section 2.3.7 is promising since it exhibits non exponential decay for large deviations. A better comprehension of this phenomenon will be a first step to study coupled map lattices with such more general local maps.

From the physical literature we notice in particular that authors exhibit the so-called non-trivial collective behavior by a study of the evolution in time of the spatial average over the sites, see Section 1.2.3.

This phenomenon is completely different from the ones we studied and could present really interesting probabilistic properties. This constitutes another topic we want to develop.

1. COUPLED MAP LATTICES

This first Chapter gives a review of results concerning coupled map lattices since their introduction by Kaneko in 1983.

After a short description of the models, we present the results of physical literature: we restrict to a phenomenologic description of the various behaviors, and refer the interested reader to reference books and papers for more details on the analysis of theses phenomena.

We close this Chapter with a presentation in Section 1.3 of mathematical work on these models. Most of it is consacred to the case of weak coupling, where it is shown that various statistical properties of the uncoupled system are preserved.

1.1 Presentation of the different models

We find under the general terminology of "coupled map lattices" many models, all of which consist of deterministic interactions between local maps.
We distinguish between:

- coupled map lattices (CML), for which the interactions are local and the spatial structure of the lattices $\mathbb{Z}^d$ will be crucial;

- globally coupled maps (GCM), for which the interactions are global. This is a mean-field version of CML.

We follow for the choice of these terms Kaneko and Tsuda in [62].

Below, we present shortly both viewpoints, although our concern will be coupled map lattices, with local interactions.

1.1.1 Coupled map lattices

We consider the state space $\mathcal{X} = M^{\mathbb{Z}^d}$, with $d \geq 1$ and M the space on which the local maps act. It can be an interval of $\mathbb{R}$, the circle S^1 or more generally any compact Riemannian manifold.
At each step of time, the dynamics will consist in:

- the action of a local map $f_i : M \to M$ on each site $i \in \mathbb{Z}^d$;
- the action of a coupling map $G : \mathcal{X} \to \mathcal{X}$ on the whole space.

The dynamical system we study is the coupled map:

$$F = G \circ F_0 , \qquad \text{where } F_0 = \bigotimes_{i \in \mathbb{Z}^d} f_i$$

The aim is to understand the evolution of a given initial condition (i.e. an initial point $x_0 \in \mathcal{X}$, or an initial probability measure μ_0 on $\mathcal{X}$) under iterations of F.

In practical situations (e.g. for computer simulations), we must restrict the study to finite systems and consider, for Λ a finite subset of Z^d, a map F_Λ which acts on $\mathcal{X}_\Lambda = M^\Lambda$.

This finite approximation is also useful for mathematical studies, even to derive results for the infinite system.

More generally, the interest for the mathematical study of infinite systems stems from the fact that the behavior of an infinite system is considered, as in Statistical Mechanics, as a good approximation of the behavior of a large (but finite) system. For example, it is well known that phase transitions, although observable in laboratory, can only occur in the thermodynamic limit.

The local maps f_i and the coupling G will be specified in each case. We give some examples of the different choices that appear in the literature.
The most common choices for the local map f are:

- an expanding map of the circle $f : S^1 \to S^1$ $\mathcal{C}^{1+\alpha}$ and such that there is $\lambda > 1$ with:

  $$f'(x) \geq \lambda \qquad \forall\, x \in S^1$$

  This is the simplest example of a map with chaotic behavior, i.e. strong dependence on the initial conditions.

  Chaos is also characterized by the existence and uniqueness of an invariant probability measure absolutely continuous with respect to Lebesgue measure.

- an expanding map of the interval $f : [0,1] \to [0,1]$ such that there is an open partition $(I_l)_{1 \leq l \leq L}$ of $[0,1]$ with f $\mathcal{C}^{1+\alpha}$ and expanding on each I_l and $f(I_l) = (0,1)$.

  This is the interval map equivalent to previous one on the circle. Assumptions can be weakened to preserve the same statistical properties.

- a $\mathcal{C}^2$ diffeomorphism with hyperbolic attractor $f : M \to M$ on a Riemaniann manifold M containing a set Λ compact, f-invariant, attracting (i.e. there is a neighborhood U of Λ with $\overline{f(U)} \subset U$ and $\cap f^n(U) = \Lambda$) and hyperbolic: for all $x \in \Lambda$, the tangent space at x splits into the stable subspace E^s and the unstable subspace E^u, $T_xM = E^s(x) \oplus E^u(x)$, such that there is $\lambda < 1$ with:

$$\|Df^n v\| \leq \lambda^n \|v\| \quad \forall v \in E^s(x), n \geq 0$$
$$\|Df^{-n} v\| \leq \lambda^n \|v\| \quad \forall v \in E^u(x), n \geq 0$$

  This map satisfies also existence and uniqueness of an invariant probability measure absolutely continuous with respect to Riemannian measure along the unstable manifolds.

- a logistic map $f : [-1, 1] \to [-1, 1]$ defined by:

$$f(x) = 1 - ax^2 \tag{1.1}$$

  with $a \in [0, 2]$.

  The interest of this map is that it presents lot of different asymptotic regimes when the parameter a varies in the interval $[0, 2]$. We briefly describe this model, referring the reader to the main reference book [29] or the recent review paper [101] for details.

  The asymptotic behavior is described by the well known Feigenbaum diagram (see for example in [30]), where large iterates of an arbitrary initial point are drawn for various values of a.

  For low parameters, there is an increasing sequence $(a_n)_{n\geq 0}$, with $a_0 = 0$, tending to a limit parameter $a_\infty \sim 1.401\ldots$ and such that for any $a \in (a_n, a_{n+1})$ all orbits tend to a stable periodic orbit of period 2^n.

  Above this parameter a_∞, one can identify another sequence of $\bar{a}_n$ decreasing and tending to a_∞, with $\bar{a}_0 = 2$. For $a \in (a_\infty, \bar{a}_n)$, one can decompose the limit set in 2^n strips through which every orbit goes periodically. Some periodic orbits appear also above a_∞ (with periods all the integers which are not powers of 2).

  Concerning chaoticity of the dynamics, different behaviors appear:

  - on a completely disconnected set of parameters with null Lebesgue measure, the attractor is a Cantor set, with an ergodic but non-mixing invariant measure,

- on a completely disconnected set C of parameters with strictly positive Lebesgue measure, the attractor is an union of intervals,
- on a subset of C with same measure, there exists a mixing absolutely continuous invariant measure. This is for example the case for $a = 2$, where the invariant measure is in fact equivalent to Lebesgue measure.

The most interesting feature of this model (for high values of the parameter a) is the coexistence of areas of chaoticity, with high rate of dispersion (where the derivative of the logistic map is greater than one) and of a strongly laminar area, around the degenerate point 0, where the derivative vanishes.

The best criterion to ensure chaoticity is the Collet-Eckmann condition (stated in [30]): there exist $C > 0$ and $\lambda > 1$ such that:

$$|(f^n)'(1)| \geq C\lambda^n \qquad \forall n \geq 0$$

It states that the orbit of the critical point 0 is expanding, i.e. that any orbit does not lose too much time in the laminar area.

The simplest choice for the coupling is $G = G_\varepsilon$:

$$(G_\varepsilon(x))_i = (1-\varepsilon)x_i + \frac{\varepsilon}{2d}\sum_{j\sim i} x_j \tag{1.2}$$

where the parameter ε measures the strength of the coupling (note that if $\varepsilon = 0$,then $G_0 = \mathrm{Id}$) and the sum is over the nearest neighbors of i on the lattice.
This coupling is always used in physical literature. Mathematical results deal with general classes of couplings, but one can keep this one (or a smooth modification of it when one works on the circle) in mind for comprehension.

1.1.2 Globally coupled maps

In this model, we keep local maps, but the coupling does not depend on the geometry of the lattice: the action of the coupling on a site depends in the same way on all other sites.
In this sense, the state space of N sites can be taken as $\{1, \ldots, N\}$.

We formalize this by defining our map on $\mathcal{X}_N = M^{\{1,\ldots,N\}}$ by:

$$F_N(x) = G_N \circ F_0$$

where:

- $F_0 = \otimes_{i=1}^N f_i$ is the uncoupled product of local maps;
- the coupling G_N depends only on the value at site i and the empirical measure of all sites:

$$(G_N(x))_i) = g(x_i, R_N(x))$$

with $R_N(x) = \frac{1}{N}\sum_{j=1}^N \delta_{x_j} \in \mathcal{M}^1(\mathcal{X})$, the set of probability measures on $\mathcal{X}$, and $g : \mathcal{X} \times \mathcal{M}^1(\mathcal{X}) \to \mathcal{X}$.

We can then study the behavior of each finite dimensional dynamical system $(\mathcal{X}_N, F_N)$, and the effect of making N tend to infinity, although we can not define directly a corresponding infinite dimensional dynamical system.

The choice of local maps f_i will generally be the same as for CML. Thanks to the absence of spatial geometry, the analysis could get simpler, as in Statistical Mechanics, where for example the Curie-Weiss model is simpler to study than the Ising model, see [40].

1.2 Physical and numerical results

The literature about numerical results for coupled map lattices is wide and fast growing. Methods used to explain various behaviors are technical and hard to handle. We can not go into too much details and keep only the description of some significant observations. We refer the reader to the books [61, 62] and to the special issues of Physica D [22] and Chaos [60] for details on phenomenology, techniques and applications.

1.2.1 Phenomenologic study of locally coupled logistic maps

In his first papers on coupled map lattices [54, 55] (see also Chapter 3 of [62] for a detailed exposition), Kaneko studied computer simulations of the 1-dimensional CML formed on a finite interval $[-N, N]$ with the logistic maps (1.1) and the coupling $G = G_\varepsilon$ described in (1.2). This model depends on both parameters a, the nonlinearity parameter, and ε, the coupling strength. Kaneko obtains in [56] a detailed phase diagram.

The simplest way to see the link between both parameters is to say that a high nonlinearity parameter a tends to develop local chaotic behavior, as we saw for the single map, whereas a strong coupling strength favors spatial coherence between sites. These two effects compete to give the different behaviors of the phase diagram, which we present in this Section.

Coherent patterns

When the nonlinearity parameter is below a_∞, the periodic behavior of local maps is preserved by the coupling, giving domains where all sites oscillate periodically in phase. Between two such domains with same period 2^k, there appears a kink near unstable periodic points of period 2^{k-1}.

After this cascade of doublings, the system exhibits chaotic behavior but with a strong spatial structure organized in large domains on which sites are strongly correlated. The motion is periodic on some of these domains and chaotic on some others.

As nonlinearity is increased further, the large domains split into smaller domains, with shorter periods. The chaotic domains seem to disappear in this regime. According to Kaneko, there is no clear explanation for this suppression of chaos.

For strong coupling ($\varepsilon > 0.45$ according to Kaneko), the coherent structures described in this part can be moving and create traveling waves.

Spatiotemporal intermittency

The next feature appearing when nonlinearity a is increased is the spatiotemporal intermittency, in which any domain becomes unstable and each point oscillates irregularly between ordered states (called laminar regions) and regions with irregular motion (called burst regions).

Kaneko distinguishes between two types of spatiotemporal intermittency: in a first type, which occurs for small coupling, there is no spontaneous creation of burst areas. If a site and its neighbors are laminar, it remains laminar. This kind of intermittency has been studied further, see Section 1.2.2.
Another kind of intermittency, with spontaneous creation of burst regions, can occur for stronger coupling.

This occurrence of intermittency as a transition from regular structures to chaotic behavior has been observed in many examples, from numerical simulations of coupled map lattices or partial differential equations to various experiments.

Spatiotemporal chaos

For higher nonlinearity, any spatial structure disappears, giving fully developed spatiotemporal chaos, i.e. a motion with fast decorrelation between sites and in time.

Kaneko verified in [57] that the coupling destroys the windows of periodicity that are observed for a single logistic map at some $a > a_\infty$. He detects this by a computation of experimental entropy, which is increasing with a.

1.2.2 Spatiotemporal intermittency

To emphasize the appearance of intermittency, Chaté and Manneville propose in [25] a new model of coupled map lattices with the same coupling G_ε (see (1.2)) and with a simpler local map, for which the distinction between chaotic and laminar areas is clear. For $r > 2$, $f : [0, r/2] \to [0, r/2]$ is defined by:

$$f(x) = \begin{cases} rx & \text{if } 0 \le x \le 1/2, \\ r(1-x) & \text{if } 1/2 \le x \le 1, \\ x & \text{if } 1 < x \le r/2. \end{cases}$$

The dynamics of the single map is simple: Lebesgue-almost every orbit is moving on the uniformly expanding (with expansion rate r) part $[0, r/2]$ and then trapped at one time in the laminar part $[1, r/2]$ which is a complete interval of fixed points.

The same trivial behavior occurs for the coupled map lattices starting from a constant initial condition or under small coupling.

Nevertheless, simulations reveal a threshold ε_c (depending on r) above which spatiotemporal intermittency occurs. The sharp geometric structure of this intermittency depends strongly on the value of r, but in any case, laminar sites are spatially stable: a site can become chaotic only if some of its neighbors are chaotic.
In the same paper, following graphical similarities, Chaté and Manneville investigate the link between this model and a directed percolation model with adapted transition probabilities. However the calculation of characteristic exponents shows that the two models do not behave in the same way.

The model of Chaté and Manneville was further investigated by Grassberger and Schreiber in [44], in which it was shown that this model of coupled map lattices is in fact in the same universality class as directed percolation. To see it, one has to change the time scale in order to avoid the creation of spatial correlation, incompatible with percolation.
See also the PhD Thesis of Chaté [21] for more details.

1.2.3 Non-trivial collective behavior

Chaté and Manneville have emphasized in [26] another interesting feature of coupled map lattices: the non-trivial collective behavior. We give here a simple description of this phenomenon, and refer to [23, 75, 28] for studies with various methods.

We work in dimension 1 with the coupling (1.2) and a local logistic map with parameter $a \in (\bar{a}_2, \bar{a}_1)$ (see definitions in Section 1.1.1). In this case, the orbit of any initial point oscillates periodically between two strips I_0 and I_1. We denote for $i \in \{0,1\}$ by $\bar{p}_i$ the asymptotic distribution on each strip I_i ($\bar{p}_i = \lim \frac{1}{T}\sum_{t=0}^{T-1} \delta_{f^{2t}(x)}$ for Lebesgue-almost all $x \in I_i$) and explain how this periodic behavior is modified by the coupling.

1) For the uncoupled map formed by such logistic maps, each site evolves independently of the others, hence for N large the instantaneous empirical mean given by:

$$p_t = \frac{1}{2N+1}\sum_{i=-N}^{N} \delta_{f^t(x_i)} \in \mathcal{M}^1([-1,1])$$

is asymptotically a mixture of the limiting measures on each of the two strips:

$$p_t \sim c_t\bar{p}_0 + (1-c_t)\bar{p}_1$$

with $c_{t+1} = 1 - c_t$ and $c_{2t} = c_0$, because of the periodicity. The proportions of points in each strip are preserved (while exchanged) by the dynamics. In spite of local chaoticity, a global periodic behavior is preserved in time at this scale. For small coupling, the behavior remains the same.

2) On the other hand, when there is a strong coupling, a synchronization effect appears. If we denote again:

$$p_t \sim c_t \tilde{p}_0 + (1 - c_t)\tilde{p}_1$$

with $\tilde{p}_i$ concentrated on I_i, the system selects in this case a pure state in the sense that c_{2t} (or c_{2t+1}, depending on the initial conditions) tends to 1 as t tends to infinity.
Asymptotically, all points of the system are together in the same strip. This spatial synchronization coexists with local chaos in time.

In the already cited papers, strategies are developed to explain such emergence of collective behavior from chaos when the coupling strength increases, through collective bifurcation and increase of spatial correlation.
In [24], Chaté and Losson relate this non-trivial collective behavior to the asymptotic behavior of a transfer operator associated to the dynamics.

1.2.4 Globally coupled maps

Kaneko studied in [58] (see the presentation in Chapter 4 of [62]) the globally coupled version of previous model with local logistic maps. On $\mathcal{X}_N = [-1, 1]^{\{1,\dots,N\}}$, the coupling is defined by

$$G_N(x) = (1 - \varepsilon)x_i + (\varepsilon/N)\sum_{j=1}^{N} x_j$$

The competition between chaos induced by the nonlinearity parameter a and order imposed by the coupling gives rise to various behaviors (see [62] for the complete phase diagram):

- Coherent state with complete synchronization: when the coupling is strong, all sites take asymptotically the same value and follow a motion governed by a single logistic map.

- Completely desynchronized state: when the coupling is small, the system preserves the behavior of null coupling, where all sites take different values and evolve without synchronization.

- **Ordered state**: in an intermediary regime, there appear clusters, i.e. sets of sites taking the same values. The evolution is then governed by mean field equations defining interactions between clusters. If one has K clusters with sizes $(N_k)_{1\leq k\leq K}$ and x_t^k denotes the value of a site of the k^{th} cluster at time t, evolution follows:

$$x_{t+1}^k = (1-\varepsilon)f(x_t^k) + \sum_{j=1}^{K} \frac{\varepsilon N_j}{N} f(x_t^j)$$

  Kaneko makes a distinction between **ordered states**, with only few large clusters, and **partially ordered states**, with a greater number of clusters.

This appearance of clusters where motion is synchronized is experimentally observed to be independent of the size of the system N.

Kaneko has then developed many tools to study these behaviors. He introduced in particular a **split exponent** which measures the asymptotic splitting rate between two sites:

$$\begin{aligned}\lambda_{\text{split}}(i) &= \lim_{T\to\infty} \frac{1}{T} \log\Big(\prod_{n=t}^{T}(1-\varepsilon)f'(x_n(i))\Big)\\ &= \log(1-\varepsilon)+\lambda_0\end{aligned}$$

where λ_0 is the Lyapunov exponent of the single logistic map. A simple criterion to get complete synchronization is then $\lambda_{\text{split}} < 0$, which gives an explicit condition on the coupling strength ε.
A generalization of this simple derivation would allow to study occurrence and stability of clusters.

For the case of completely desynchronized state, one could hope that different sites behave independently from each other when the size of the system tends to infinity.
However, experiments [59] show that this is not always the case: Kaneko calls it "violation of the law of large numbers", although it is more a defect in the fluctuations: it remains asymptotically some correlation between sites, which contradicts the intuition of propagation of chaos in this mean field setup.
Shibata and Kaneko give recently [96, 97] some explanations of this phenomenon via hidden coherence, which is another appearance of collective behavior in this kind of system.

1.3 Mathematical results

1.3.1 Spatiotemporal viewpoint

The first mathematical result on coupled map lattices has been established by Bunimovich and Sinai in [19, 20]. They wanted to give a mathematical formulation of spatiotemporal chaos and used for this analogy with statistical mechanics for a model with weak coupling and local maps which are simpler than the coupled logistic maps.

For them, spatiotemporal chaos is characterized by the existence of a mixing invariant measure with finite box marginals absolutely continuous with respect to Lebesgue measure.
It is important to notice that they construct this measure for a 1-dimensional coupled map using a Gibbs measure on a shift space with 2 dimensions, one corresponding to space, the other to time. This approach makes no essential difference between temporal dynamics and spatial shifts.

They study a CML on $\mathcal{X} = [0,1]^{\mathbb{Z}}$ and take for local map an expanding map of the interval $f : [0,1] \to [0,1]$ such that there exists an open partition $(I_l)_{0 \leq l \leq L-1}$ of $I = [0,1]$ with $\bigcup \bar{I}_l = [0,1]$, $f(I_l) = (0,1)$ and f $\mathcal{C}^{1+\alpha}$ on each I_l with $f' \geq \lambda > 1$.

They choose small couplings of the form:

$$(G(x))_i = (1 - \alpha(x_i))x_i + \frac{\alpha(x_i)}{2}(x_{i-1} + x_{i+1})$$

with $\alpha \in \mathcal{C}^2([0,1])$ such that:

$$\begin{array}{rl} (i) & \alpha(y) = \varepsilon \quad \text{for } \delta \leq y \leq 1-\delta \\ (ii) & \alpha(0) = \alpha(1) = 0 \\ (iii) & \alpha'(y) \geq 0 \quad \text{for } 0 \leq y \leq \delta \quad \text{and } \alpha'(y) \leq 0 \quad \text{for } 1-\delta \leq y \leq 1 \end{array}$$

They get then:

Theorem 1.3.1. *For ε small enough, there exists a probability measure μ such that:*

1. *μ is invariant under $F = G \circ F_0$ and the spatial shift S;*

2. *The marginals of μ on any finite subspace are absolutely continuous to Lebesgue measure;*

3. *The dynamical system* $(\mathcal{X}, \mu, (F, S))$ *is mixing:*

$$\int_{\mathcal{X}} \varphi \circ F^t \circ S^n \cdot \psi \, d\mu \longrightarrow \int_{\mathcal{X}} \varphi \, d\mu \int_{\mathcal{X}} \psi \, d\mu$$

when t *or* $|n|$ *tends to infinity.*

It is well known (see for example [64]) that for expanding maps like f, one can construct a coding $\pi : \{0, \ldots, L-1\}^{\mathbb{N}} \to [0,1]$ by:

$$\pi(\omega) = \bigcap_{n \geq 0} \bar{I}_{\omega_0, \omega_1 \ldots, \omega_n} \quad \text{with} \quad I_{\omega_0, \omega_1 \ldots, \omega_n} = I_{\omega_0} \cap f^{-1} I_{\omega_1} \cap \cdots \cap f^{-n} I_{\omega_n}$$

It defines a semiconjugacy between the shift σ on $\{0, \ldots, L-1\}^{\mathbb{N}}$ (defined by $(\sigma\omega)_i = \omega_{i+1}$) and f:

$$\pi \circ \sigma = f \circ \pi$$

This coding has been used to construct interesting invariant measures for f, using the powerful theory of Gibbs measures constructed on shift spaces (see for references [11, 93]).
Indeed, the image by π of the Gibbs measure associated to the potential $\bar{\varphi}(\omega) = -\log f'(\pi(\omega))$ is known to be the unique mixing measure absolutely continuous with respect to Lebesgue measure. It is commonly called the SRB (Sinai-Ruelle-Bowen) measure.

Bunimovich and Sinai have defined the coupling G in an adequate manner to preserve the partitions described above. They use then these partitions to construct the limits of conditional probabilities of the measure μ, when spatial size and time tend to infinity.
These conditional probabilities give the adequate potential to define an associated Gibbs measure on $\Omega = \{0, \ldots, L-1\}^{\mathbb{N} \times \mathbb{Z}}$. Bunimovich and Sinai invoke then results of Dobrushin and Martirosian [35] to conclude[1] that the Gibbs measure is unique and mixing, hence its image by the coupling remains mixing.

This elaboration of a thermodynamic formalism, at least on the coding space, makes the authors of this seminal paper hope that the experimented appearance of coherent structures could be linked to the occurrence of phase transitions, i.e. cases with at least two Gibbs measures associated to the conditional probabilities (see Section 1.3.4 for advances on this idea).

[1] The assumptions of Dobrushin and Martirosian were in fact not satisfied by this system, for proofs of the uniqueness and mixing of these Gibbs systems, see [13, 14, 52].

This work has been soon generalized to the case of local uniformly hyperbolic maps by Pesin and Sinai [86].
They consider on a Riemannian manifold M a $\mathcal{C}^2$-diffeomorphism f possessing an attractor Λ (i.e., a set Λ invariant and compact with a neighborhood U such that $\overline{f(U)} \subset U$ and $\Lambda = \cap f^n(U)$) which is hyperbolic: for all $x \in \Lambda$, the tangent space at x splits into $T_xM = E^s(x) \oplus E^u(x)$, with:

$$\begin{aligned} \|Df^n v\| &\leq \lambda^n \|v\| \quad \forall v \in E^s(x), n \geq 0 \\ \|Df^{-n} v\| &\leq \lambda^n \|v\| \quad \forall v \in E^u(x), n \geq 0 \end{aligned}$$

for $\lambda < 1$. E^s is the stable subspace, E^u the unstable subspace.

They study 1-dimensional coupled maps associated to this local hyperbolic map with short range couplings, i.e. couplings for which strength of coupling between two sites decreases exponentially fast with the distance between these sites.
They obtain existence of a mixing invariant measure, which is locally absolutely continuous to Riemann measure along the unstable manifolds (which is the characterization of the SRB measure in this setup).

Since a hyperbolic map admits also a coding constructed on the so called Markov partitions (see Chapter 18 of [64]), Pesin and Sinai proceed by an adaptation to this setup of the strategy from [19].
The main difference is however that they use an abstract argument of structural stability to show that the coding obtained for the uncoupled map by tensorisation is preserved under small coupling.

The last improvements of this method have been given by Jiang and Pesin [53, 48, 49]. They study the same model as Pesin and Sinai previously: uniformly hyperbolic local maps with short range interactions. The method follows essentially the same lines.

However, they emphasize on the natural link between this analysis and the thermodynamic formalism which is directly defined for the coupled map and the spatial shifts. They identify explicitly the potential:

$$\varphi(x) = -\log \det D^u f(x_0) + \psi(x)$$

(see Section 3.2.2 for a complete derivation of this potential in a slightly different setup) associated to the spatiotemporal dynamics under Riemann measure, and show that the measure obtained by Pesin and Sinai is also the unique measure realizing the maximum in Gibbs variational principle:

$$P_{(F,S)}(\varphi) = \sup_{\nu \in \mathcal{M}^1_{\text{inv}}(\mathcal{X})} \left(h_{(F,S)}(\nu) + \int_{\mathcal{X}} \varphi \, d\nu \right)$$

It is called the **unique equilibrium measure** associated to potential φ and dynamics (F, S), see Section 3.8 for details.

They state hence explicitly uniqueness of the limit measure, among the measures which are invariant under F and spatial shifts. This is another characterization of spatiotemporal chaos.

They had also to rework the arguments giving uniqueness of the Gibbs measure, because their potential did not satisfy assumptions of the classical results. These generalizations were done in [52], see also [13, 14].

These results, and in particular their expression in the setup of thermodynamic formalism, have allowed further developments: Jiang proves [50] that the topological pressure of the potential φ is null. Dolgopyat [36] adopts the same viewpoint to analyze Lyapunov exponents associated to the system and prove the equivalent of Pesin formula, identifying entropy with the sum of positive Lyapunov exponents.
Jiang wrote also in [51] an adaptation of the proof of [53] to the setup of expanding maps of the circle.

1.3.2 Transfer operator approach

Another important tool associated to a regular expanding map f of the interval $I = [0,1]$ (or of the circle) is the transfer operator, acting on densities of measures as the dual of the composition by f: denoting m the Lebesgue measure, the transfer operator $\mathcal{L}_f$ is defined on $L^1(m)$ by either of the two equivalent definitions:

$$\int_I \varphi \circ f \cdot \psi \, dm = \int_I \varphi \cdot \mathcal{L}_f \psi \, dm \qquad \forall\, \varphi \in L^\infty(m),\ \psi \in L^1(m)$$

$$\mathcal{L}_f \varphi(x) = \sum_{y\,:\, fy = x} \frac{\varphi(y)}{|f'(y)|}$$

The interest of this transfer operator is that $\mu = hm$ is an invariant measure if and only if $\mathcal{L}_f h = h$, and asymptotic properties of μ are linked with spectral properties of $\mathcal{L}_f$.

A whole domain of modern ergodic theory consists in the study of these transfer operators, for which the hardest work is to find a good space on which they act and enjoy nice spectral properties. We refer the reader to the exhaustive presentation of this subject in [3].

The first attempt to use transfer operators for coupled map lattices is due to Keller and Küntzle [68, 65].
They construct adapted transfer operators on Banach spaces of functions with bounded variations for the case of a local map f expanding with same kind of conditions than in [19] and a coupling $G_\varepsilon(x) = x + A_\varepsilon(x)$ where the $\mathcal{C}^2$ norm of A_ϵ is controlled by ε (the strength of the coupling).

They get in the finite lattice case, i.e. with $\mathcal{X} = [0,1]^N$, the existence of an absolutely continuous invariant measure for ε small enough, and uniqueness for ε smaller ($\varepsilon < \varepsilon_U$). They notice that the limit parameter ε_U tends to 0 as N tends to infinity.

Indeed, for the infinite lattice case $\mathcal{X} = [0,1]^{\mathbb{Z}}$, they obtain only existence of a measure invariant by the coupled map and the spatial shift, with absolutely continuous marginals on finite sub-lattices.

In the finite lattice case, they define, with $\partial\mathcal{X}$ the boudary of $\mathcal{X} = [0,1]^N$ seen as a subset of $\mathbb{R}^N$:

$$\begin{aligned}
\tau &= \{\psi \in \mathcal{C}^1(\mathcal{X}) \,:\, \psi|_{\partial\mathcal{X}} = 0,\ |\psi| \le 1\} \\
\mathrm{var}(\varphi) &= \frac{1}{N}\sum_{j=1}^{N} \sup_{\psi\in\tau} \int_{\mathcal{X}} \varphi(x) D_j\psi(x)\, dx \\
\mathrm{BV}(\mathcal{X}) &= \{\varphi \in L^1(\mathcal{X}) \,:\, \mathrm{var}(\varphi) < \infty\} \quad \text{with } \|\varphi\|_{\mathrm{BV}} = \|\varphi\|_1 + \mathrm{var}(\varphi)
\end{aligned}$$

A weak formulation of the definition of bounded variation functions is necessary in this multidimensional setting.

They prove that the transfer operator $\mathcal{L}_F$ associated with the finite lattice coupled map preserves the Banach space $\mathrm{BV}(\mathcal{X})$ and enjoys on it nice spectral properties: in fact, they can use in this case, as for a single map, compactness arguments and the Theorem of Ionescu Tulcea-Marinescu. They can deduce the existence and uniqueness of invariant measure from this strong spectral result.

For the infinite lattice case, they manage to construct a transfer operator associated to F acting on a generalized BV space: they consider extended densities $\varphi = (\varphi_V)_{V\in\mathcal{F}}$, where $\mathcal{F}$ is the set of finite subsets of $\mathbb{Z}$ and $\varphi_V \in L^1([0,1]^V)$, then define:

$$\begin{aligned}
\mathrm{var}(\varphi) &= \sup(\mathrm{var}(\varphi_V) \,:\, V \in \mathcal{F}) \\
\mathrm{BV} &= \{\varphi = (\varphi_V)_{V\in\mathcal{F}} \,:\, \mathrm{var}(\varphi) < \infty\}
\end{aligned}$$

They get for this transfer operator weaker spectral properties, hence derive only the existence of the measure.

The only other result in this setup is the recent paper [69] where Keller and Zweimüller manage to establish uniqueness of the invariant absolutely continuous measure satisfying mixing conditions. They use for this really different methods and have to assume that the coupling is unidirectional. This is a strong assumption, which gives models also studied by physicists (see Section 3.5 in [62]).

All other works on transfer operators for coupled map lattices are done under stronger assumptions on the local maps and the coupling, requiring to work in a holomorphic context.

This approach has been initiated by Bricmont and Kupiainen in [12]. They work with a local map f expanding on the circle and holomorphic in a small ring around it and a coupling G with same regularity and with exponential decay of the coupling with the distance between sites.

They get hence:

Theorem 1.3.2. *There exists a probability measure μ on $\mathcal{X}$ such that:*

1. *μ is invariant under F and spatial shifts S. For any finite Λ, $\mu|_\Lambda$ is absolutely continuous with respect to Lebesgue measure.*
2. *μ is spatiotemporally mixing, there exist $\alpha > 0$ and $c < \infty$ such that*
$$\left| \int_{\mathcal{X}} \varphi \circ F^t \circ S^i \psi \, d\mu - \int_{\mathcal{X}} \varphi \, d\mu \int_{\mathcal{X}} \psi \, d\mu \right| \le e^{-\alpha(t+|i|)} e^{c|\Lambda|} \|F\| \|G\|$$
for F and G depending on the finite box Λ and holomorphic in a neighborhood of $(S^1)^\Lambda$.
3. *$F^n \overline{m}$ tends weakly to μ as n tends to infinity, with $\overline{m}$ the product of Lebesgue measures. μ is unique among the class of measures locally absolutely continuous with respect to Lebesgue measure with locally holomorphic densities.*

They use to prove this stronger properties of the finite lattice transfer operators on the space of real analytic densities: the uncoupled transfer operator is compact with 1 as simple maximal eigenvalue and the transfer operator of the coupling can be analyzed by complex analysis.

They obtain sharp estimates on the spectral gap of the coupled transfer operator by cluster expansions estimates. The main point is that these estimates are uniform in the size of the lattice, hence the results are preserved when the size of the system tends to infinity.

These authors have established other results on this subject [13, 14], where they generalize this viewpoint and make the link between the transfer operator approach and the work on coding space. They clarify in particular the arguments needed for uniqueness of the Gibbs measure.

Baladi et al develop this analysis in [4] to construct a transfer operator associated to the dynamics on the infinite lattice: they define in the same context as previous paper a Fréchet space (in the general case) and a Banach space (in dimension 1) on which a transfer operator acts. In the unidimensional case, cluster expansions techniques allow them to present the transfer operator associated to the coupled dynamics as a perturbation of the one associated to the uncoupled map.

They deduce from this and spectral analysis a localization of the coupled spectrum. This gives again spectral gap, hence existence and uniqueness among a subspace of invariant mixing locally absolutely continuous measure, and furthermore a description of the spectrum under this gap. However the study of this spectrum is limited by the absence of spatial homogeneity of the Banach space.

The next step in this study is due to Maes and Van Moffaert [78]. They study stochastic stability of the measure obtained in [12], and simplify in the course of the proof the cluster expansion argument.

This simplification is used by Fischer and Rugh in [41] (again in the context of holomorphic expanding maps of the circle), and associated to a new kernel representation of the transfer operator on a finite lattice to define in a simpler way a global transfer operator.
It acts on the Banach space $\mathcal{H}_\theta$ (for $\theta < 1$ well chosen) of families $\varphi = (\varphi_\Lambda)_{\Lambda\in\mathcal{F}}$ ($\mathcal{F}$ is the set of finite subsets of $\mathbb{Z}^d$) such that:

- $\varphi_{\Lambda_1} = \displaystyle\int_{(S^1)^{\Lambda^2\setminus\Lambda^1}} \varphi_{\Lambda_2}\, dm^{\Lambda^2\setminus\Lambda^1} \quad \text{when } \Lambda_1 \subset \Lambda_2$
- $\|\varphi\|_\theta = \displaystyle\sup_{\Lambda\in\mathcal{F}} \theta^{|\Lambda|}\|\varphi_\Lambda\| < \infty$

It means that one allows an exponential increase of the norm of the marginals with the size of the box.

This defines exactly the good Banach space to prove a spectral gap for some iterate of the transfer operator acting on $\mathcal{H}_\theta$, for θ and a coupling strength small enough.

They deduce from this spectral gap existence and uniqueness among $\mathcal{H}_\theta$ of an invariant measure with mixing properties.

Rugh gives more recently in [95] a simplified and generalized version of these results. In particular, he simplifies again the combinatorics and allows for the choice of couplings which decrease less than exponentially fast.

We do not emphasize these last results here, referring the reader to Chapter 4, where this method in exposed in details and used to get new limit theorems. We just remark that this last improvement allows also a study in more details of the spectrum under the spectral gap. This has been done in [5].

1.3.3 Globally coupled maps

We mention briefly the few results existing in the rigorous study of globally coupled maps. Järvenpää adapts in [47] the method of Bricmont and Kupiainen to prove existence of a mixing measure for each finite size system and weak convergence of these measures to a product measure as the size of the associated system tends to infinity.

Keller provides in [67] a general mathematical framework to study such models, using in particular the theory of exchangeable distributions to give a precise sense to what Kaneko calls "violation of the law of large numbers". He proves that this phenomenon does not occur for $\mathcal{C}^3$ expanding maps of the circle nor mixing tent maps with small couplings.

1.3.4 Phase transitions

Most articles on phase transition (see for instance [80, 9]) deal with numerical studies of such phenomena.

For a mathematical study, even the definition is not clear. If all authors agree on the fact that it must be linked to non-uniqueness of some specified invariant measures, the way to characterize these measures is in discussion.

For some of them [17, 18, 7, 8], the interesting measures are natural measures: for a dynamical system $(\mathcal{X}, F)$ with reference measure $\overline{m}$, a measure μ is a natural measure if there exists an open subset $U \subset \mathcal{X}$ (the basin of attraction of μ) such that for any measure $\nu \in \mathcal{M}^1(\mathcal{X})$ with support in U and locally absolutely continuous with respect to $\overline{m}$, we have:

$$\frac{1}{T}\sum_{t=0}^{T-1}(F^*)^t\nu \longrightarrow \mu$$

They construct some examples presenting non-uniqueness of natural measure, where this non-uniqueness is induced by some strong topological bifurcation, as two fixed points made stable by the coupling.

The strange fact in this definition is that it can be satisfied by finite systems, and even a simple map as:

$$f(x) = \begin{cases} -2-2x & \text{if } -1 \leq x \leq -1/2 \\ 2x & \text{if } -1/2 \leq x \leq 1/2 \\ 2-2x & \text{if } 1/2 \leq x \leq 1 \end{cases}$$

This is the reason why Gielis and MacKay [43] require for their study of phase transitions that the coupled map is indecomposable, which is a spatiotemporal specification property (extension of Definition 2.3.2).
In this setup, and assuming also that the coupled map admits a symbolic coding as described is Section 1.3.1, they say that phase transition occurs when there is more than one Gibbs measure for a convenient potential on the coding space.
They give an example: starting from a coding space with such a situation, they manage to construct the associated coupled map. This example has not the standard form of the composition of a coupling and a product of local maps, since the values at neighboring sites directly modify the local map.

The objection that can be done to this viewpoint is that it is more restrictive, asking for indecomposability and symbolic coding. For the second restriction, the authors invoke the use of thermodynamic formalism directly on the space. It could allow to call a phase transition a situation with several equilibrium measures associated to a given potential.

More generally, it seems to be unclear if one can hope to get interesting strong coupling cases which preserve global properties of the system as indecomposability or existence of an appropriate potential.

2. LARGE DEVIATIONS PRINCIPLES FOR DYNAMICAL SYSTEMS

We present in this Chapter some of the existing large deviations results for the temporal asymptotic study of **finite dimensional** dynamical systems.

Our presentation is inspired of lectures given by Gérard Ben Arous at the Winter School on Ergodic Theory, organized at Sils-Maria (Switzerland) in January 1999 by Viviane Baladi and Carlangelo Liverani.

As in these lectures, we have chosen to start with a short presentation of the large deviations formalism and some classical results in probability theory. It gives an introduction to main methods for proving large deviations principles, and these methods have been adapted to the setting of dynamical systems. For a more complete overview of the field of large deviations, see [34, 32].

The readers which are familiar with the basics of large deviations can skip this part and go directly to Section 2.3, where results for dynamical systems are presented.

As large deviations principles are natural results for Gibbs measures in statistical mechanics, the equivalent results for dynamical systems will be strongly related to the thermodynamic formalism, which is presented in Section 2.3.1.

2.1 Introduction to large deviations formalism

2.1.1 The example : Cramer's Theorem

As a first (and fundamental) example, we consider a sequence $(X_i)_{i\geq 0}$ of independent and identically distributed (i.i.d.) random variables in $\mathbb{R}$ with law $\mu \in \mathcal{M}_1(\mathbb{R})$ and the empirical mean of these variables :

$$M_n = \frac{1}{n}\sum_{i=0}^{n-1} X_i$$

We know by classical results of probability theory that :

- Law of Large Numbers: If $\bar{x} = \mathbb{E}X_0$ exists, $M_n \to \bar{x}$ almost everywhere;

- Central Limit Theorem: If in addition $\sigma^2 = \mathbb{E}(X_i - \bar{x})^2 < \infty$, then:

$$\frac{\sqrt{n}(M_n - \bar{x})}{\sigma} \longrightarrow \mathcal{N}(0,1) \quad \text{in law.}$$

These two results give the asymptotic behavior of the empirical mean and small fluctuations (of the order of $\frac{1}{\sqrt{n}}$) around this limit. With Large Deviations techniques, we want to estimate the speed of convergence to 0 of probabilities of greater fluctuations around the mean: events of the type of $\{M_n - \bar{x} > \alpha\}$ with $\alpha > 0$. For this we note

$$\begin{aligned} \Lambda(\lambda) &= \log \mathbb{E}(e^{\lambda X_0}) \quad \text{the logarithmic moment generating function,} \\ \mathcal{D}_\Lambda &= \{\lambda : \Lambda(\lambda) < \infty\} \quad \text{its domain,} \\ \Lambda^*(x) &= \sup_{\lambda \in \mathbb{R}} (\lambda x - \Lambda(\lambda)) \quad \text{its Legendre transform.} \end{aligned}$$

We have then, without any assumption on the function Λ, the following:

Theorem 2.1.1 (Cramer-Chernoff).

1. *For any* $F \subset \mathbb{R}$ *closed,* $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(M_n \in F) \leq -\inf_{x\in F} \Lambda^*(x)$
2. *For any* $G \subset \mathbb{R}$ *open,* $\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(M_n \in G) \geq -\inf_{x\in G} \Lambda^*(x)$
3. $\Lambda^* : \mathbb{R} \to [0,\infty]$ *is a convex lower semi-continuous function and satisfies* $\inf_{x\in\mathbb{R}} \Lambda^*(x) = 0$. *When* $\bar{x} = \mathbb{E}X_i$ *exists, then* $\Lambda^*(\bar{x}) = 0$.
4. *If* $0 \in \overset{\circ}{\mathcal{D}}_\Lambda$, *the level sets* $\{x : \Lambda^*(x) \leq c\}$ *are compact. If* $\mathcal{D}_\Lambda = \mathbb{R}$, *then* $\lim_{|x|\to\infty} \frac{\Lambda^*(x)}{|x|} = \infty$.

Examples :

$$\text{If } X_i \sim \text{ Bernoulli}(p),\ \Lambda^*(x) = \begin{cases} x \log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{1-x}{1-p}\right) & \text{if } 0 \leq x \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

$$\text{If } X_i \sim \mathcal{N}(0,\sigma^2),\ \Lambda^*(x) = \frac{x^2}{2\sigma^2}$$

Proof. We give the main steps of the proof of this first simple result (referring the reader to Th. 2.2.3 of [32] for the details) because the main tools that are used in Large Deviations techniques already appear in it: the Upper Bound relies on a well optimized exponential Chebyshev inequality, the Lower Bound on a change of measure argument.

Upper bound:

We treat the case where $\bar{x}$ exists and is finite. For $F = [x, \infty)$ with $\bar{x} < x$ (otherwise $\inf_{x\in F} \Lambda^*(x) = 0$ and there is nothing to prove), we get by the Chebyshev inequality for all $\lambda > 0$:

$$\begin{aligned}\mathbb{P}\left(M_n \in F\right) &= \mathbb{P}\Big(\sum_{i=0}^{n-1} X_i \geq nx\Big) \\ &\leq \mathbb{E}\left(e^{\lambda \sum_{i=0}^{n-1} X_i}\right) e^{-n\lambda x} = \exp\left[-n\left(\lambda x - \Lambda(\lambda)\right)\right]\end{aligned}$$

It is easy to verify that in this case the supremum of $(\lambda x - \Lambda(\lambda))$ is realized for a positive λ, hence:

$$\mathbb{P}\left(M_n \in F\right) \leq \exp(-n\Lambda^*(x))$$

In the same way, when $x < \bar{x}$, $\mathbb{P}\left(M_n \in (-\infty, x]\right) \leq \exp(-n\Lambda^*(x))$.

For a general closed set F such that $\inf_{x\in F} \Lambda^*(x) > 0$, then $\bar{x} \notin F$. In the simpler case, we have $F \subset (-\infty, a] \cup [b, \infty)$ with $a < \bar{x} < b$, hence:

$$\mathbb{P}\left(M_n \in F\right) \leq \exp(-n\Lambda^*(a)) + \exp(-n\Lambda^*(b)) \leq 2\exp\left(-n \inf_{x\in F} \Lambda^*(x)\right)$$

which gives the upper bound (and in fact an estimate available for all n). We can treat other cases in the same way.

Lower bound :

The lower bound is a local property. We need only to show that for all $x \in \mathbb{R}$ and $\delta > 0$:

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(M_n \in (x-\delta, x+\delta)\right) \geq -\Lambda^*(x)$$

and by change of variable, it is sufficient to prove it for $x = 0$:

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(M_n \in (-\delta, \delta)\right) \geq -\Lambda^*(0) = \inf_{x\in\mathbb{R}} \Lambda(x)$$

We will assume that μ has compact support, $\mu(-\infty, 0) > 0$ and $\mu(0, \infty) > 0$ (all other cases can be treated by approximation). Then Λ is finite and differentiable everywhere, and there exists η such that $\Lambda(\eta) = \inf_{x\in\mathbb{R}} \Lambda(x)$ characterized by $\Lambda'(\eta) = 0$.

We proceed then to a change of measure to define a new probability measure $\tilde{\mu}$ by:

$$\frac{d\tilde{\mu}}{d\mu}(x) = e^{\eta x - \Lambda(\eta)}$$

and $(\tilde{X}_i)_{i\geq 0}$ i.i.d. random variables with law $\tilde{\mu}$. We denote $\tilde{M}_n$ their empirical mean and get for $\varepsilon > 0$:

$$\begin{aligned}\mathbb{P}\left(|M_n| < \varepsilon\right) &= \int_{\left|\sum_{i=0}^{n-1} x_i\right| < n\varepsilon} \mu^{\otimes n}(dx_0, \dots, dx_{n-1}) \\ &\geq e^{-n\varepsilon|\eta|} \int_{\left|\sum_{i=0}^{n-1} x_i\right| < n\varepsilon} e^{\eta\left(\sum_{i=0}^{n-1} x_i\right)} \mu^{\otimes n}(dx_0, \dots, dx_{n-1}) \\ &\geq \exp(-n\varepsilon|\eta| + n\Lambda(\eta))\, \mathbb{P}(|\tilde{M}_n| < \varepsilon)\end{aligned}$$

We have just adapted the change of measures to get:

$$\mathbb{E}\tilde{X}_0 = \int_{\mathbb{R}} x e^{\eta x - \Lambda(\eta)}\, \mu(dx) = \Lambda'(\eta) = 0$$

hence $\lim_{n\to\infty} \mathbb{P}(|\tilde{M}_n| < \varepsilon) = 1$ by the law of large numbers. This gives:

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(|M_n| < \delta\right) &\geq \Lambda(\eta) - \varepsilon|\eta| && \forall\, 0 < \varepsilon < \delta \\ &\geq \inf_{x\in\mathbb{R}} \Lambda(x) = -\Lambda^*(0) && \text{making } \varepsilon \to 0\end{aligned}$$

□

2.1.2 Basic definitions of large deviations

Let $\mathcal{X}$ be a topological space and $\mathcal{B}$ its Borel σ-algebra. Consider $(\mu_n)_{n\geq 0}$ a family of probability measures on $(\mathcal{X}, \mathcal{B})$.

Definition 2.1.1. *$I : \mathcal{X} \to [0, \infty]$ is a* rate function *if it is lower semi-continuous, i.e. its level sets $\{x \in \mathcal{X} : I(x) \leq \alpha\}$ are closed for all $\alpha \in \mathbb{R}$.*
I is a good rate function *if the level sets are compact.*

Definition 2.1.2. *(μ_n) satisfies a* Large Deviations Principle (LDP) *with rate function I if:*

$$\text{Upper Bound}: \quad \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} I(x) \qquad \forall \text{ closed } F \subset \mathcal{X}$$

$$\text{Lower Bound}: \quad \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq -\inf_{x\in O} I(x) \qquad \forall \text{ open } O \subset \mathcal{X}$$

Examples : The Cramer Theorem is a Large Deviations Principle with convex rate function $I = \Lambda^*$. The simplest example to keep in mind is $d\mu_n(x) = \mathcal{Z}_n^{-1} e^{-nI(x)} dx$. In this case the rate function appears naturally as the density of μ_n.

Remarks : 1. If $\mathcal{X}$ is a regular space (i.e. for all closed F and $x \notin F$, there exist open sets G_1, G_2 such that $x \in G_1$, $F \subset G_2$, and $G_1 \cap G_2 = \emptyset$) then a family (μ_n) can have at most one rate function satisfying a Large Deviations Principle.
2. A Large Deviations result implies both measurable and topological structures, and depends heavily on this topology: the finer this one is, the stronger the Large Deviations result. The difficulty will often be to find the good topology: rough enough to make the Large Deviations Principle hold and fine enough to give a consistent result.

In a general setup, we have often to prove a Large Deviations Principle in two steps, proving first a weak version of it, then exponential tightness of the sequence of measures (a type of compactness property). We may already notice that weak and full Large Deviations Principles are equivalent if the space $\mathcal{X}$ is compact, which will generally be the case in the context of Dynamical Systems. We give however below the detailed Definitions and links between both properties.

Definition 2.1.3. $(\mu_n)_{n\geq 0}$ *satisfies a* weak Large Deviations Principle (WLDP) *with rate function I if :*

- *Upper bound is valid for any compact F,*
- *Lower bound is valid for any open G.*

Definition 2.1.4. (μ_n) *is* exponentially tight *if for every $L > 0$, there exists $K(L)$ compact such that :*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n \left(K(L)^C \right) \leq -L$$

Proposition 2.1.1. *If (μ_n) is exponentially tight then :*

1. *Upper Bound for all compact $F \Rightarrow$ Upper Bound for all closed F;*
2. *Lower Bound $\Rightarrow$ I is a good rate function;*
3. *WLDP $\Rightarrow$ LDP with a good rate function.*

Reciprocally, if (μ_n) satisfies a Large Deviations Principle with a good rate function on $\mathcal{X}$ Polish, then (μ_n) is exponentially tight.

2.1.3 Transformations of large deviations principles

We give two useful ways to derive new Large Deviations Principle from known ones. We give the simplest assumptions, although more general results exist, and can be found for example in Chapter 4 of [32].

Proposition 2.1.2 (Contraction Principle). *If $(\mu_n)_{n\geq 0}$ satisfies a Large Deviations Principle with good rate function I on $\mathcal{X}$ and $\Phi : \mathcal{X} \to \mathcal{Y}$ is continuous, then $\nu_n = \mu_n \circ \Phi^{-1}$ satisfies a Large Deviations Principle on Y with good rate function :*

$$J(y) = \inf\{I(x) \,:\, x \in \mathcal{X} \; s.t. \; \Phi(x) = y\}$$

Proposition 2.1.3 (Laplace-Varadhan Lemma). *Let F be bounded and continuous on $\mathcal{X}$ and $(\mu_n)_{n\geq 0}$ satisfy a Large Deviations Principle with good rate function I. Define $\nu_n = \mathcal{Z}_n^{-1} e^{-nF} d\mu_n$ with $\mathcal{Z}_n = \int_{\mathcal{X}} e^{-nF} d\mu_n$. Then:*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{Z}_n = -\inf_{x\in\mathcal{X}} (F(x) + I(x))$$

In addition, $(\nu_n)_{n\geq 0}$ satisfies a Large Deviations Principle with good rate function :

$$J(x) = F(x) + I(x) - \inf_{x\in\mathcal{X}} (F(x) + I(x))$$

This result is of great interest. If you manage to obtain a Large Deviations Principle for i.i.d. random variables on $\mathbb{R}$ (what Cramer Theorem gives), you get it when adding a potential e^{-nF}: this gives you directly Large Deviations Principle for Gibbs measures, and much more if the rate function behaves well.

2.1.4 Strategies to obtain Large Deviation Principles

As we said, there is no recipe available for all situations. We can however present the two main ways of getting Large Deviations results. A third one, relatively different will occur more specifically for Dynamical Systems (see Subsection 2.3.3 or Chapter 3).

WLDP from abstract nonsense and sub-additivity

Let $(\mu_n)_{n\geq 0}$ a sequence of probability measures on $\mathcal{X}$. Choosing $\mathcal{A}$ a basis for the topology of $\mathcal{X}$, we denote for $A \in \mathcal{A}$, $x \in \mathcal{X}$:

$$\underline{L}(A) = -\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(A) \qquad \underline{I}(x) = \sup_{x\in A, A\in\mathcal{A}} \underline{L}(A)$$

$$\overline{L}(A) = -\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(A) \qquad \overline{I}(x) = \sup_{x\in A, A\in\mathcal{A}} \overline{L}(A)$$

Theorem 2.1.2. *If $\mathcal{A}$ is such that $\underline{I}(x) = \overline{I}(x)$ $\forall x \in \mathcal{X}$, then (μ_n) satisfies a WLDP with rate function $I = \underline{I} = \overline{I}$.*

In particular:

Corollary 2.1.1. *If $\lim_{n\to\infty} \frac{1}{n} \log \mu_n(A)$ exists for all $A \in \mathcal{A}$, (μ_n) satisfies a WLDP with rate function I.*
Moreover, if $\mathcal{X}$ is a topological vector space and:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n \left(\frac{A+B}{2} \right) \geq -\frac{1}{2} \left(\underline{L}(A) + \underline{L}(B) \right) \quad \forall A, B \in \mathcal{A}$$

then I is convex.

We will usually take for $\mathcal{A}$ the set of open convex subsets of $\mathcal{X}$, and use a subadditivity argument to get the limit of $\frac{1}{n} \log \mu_n(A)$. The main interest of this method is that it can be applied to some non convex setup, i.e. with a non convex rate function. Its main limit is that it does not say much about the rate function, obtained only in an abstract sense.

WLDP from Cramer's idea: Gärtner-Ellis theorem

Assume $\mathcal{X}$ is a Hausdorff topological vector space and (Z_n) are random variables taking values in $\mathcal{X}$ with laws (μ_n).
We may define in this context the logarithmic moment generating function :

$$\Lambda_n(\lambda) = \log \mathbb{E} \left(e^{\langle \lambda, Z_n \rangle} \right) = \log \int_{\mathcal{X}} e^{\langle \lambda, x \rangle} \, \mu_n(dx) \quad \forall \, \lambda \in \mathcal{X}^*$$

and will suppose :

$$\lim_{n\to\infty} \frac{1}{n} \Lambda_n \left(n\lambda \right) = \Lambda(\lambda) \quad \text{exists in } \overline{\mathbb{R}} \qquad \forall \, \lambda \in \mathcal{X}^\star \tag{$\star$}$$

Theorem 2.1.3. *Under ($\star$):*

1. *Λ is convex and its Legendre transform*

   $$\Lambda^*(x) = \sup_{\lambda \in \mathcal{X}^*} \left(\langle \lambda, x \rangle - \Lambda(\lambda) \right)$$

   is a convex rate function.

2. *The LD Upper Bound is satisfied for every compact sets with rate function Λ^*.*

3. *If (ν_n) satisfies a Large Deviations Principle with good rate function I and $\Lambda(\lambda)$ is finite for every $\lambda \in \mathcal{X}^*$, then $\Lambda(\lambda) = \sup_{x\in\mathcal{X}}(\langle\lambda, x\rangle - I(x))$ and Λ^* is the affine regularization of I (i.e. $\Lambda^* \leq I$ and for all convex rate function $f \leq I$, then $f \leq \Lambda^*$)*

   In particular, if I is convex, $\Lambda^* = I$.

Remarks : 1. The two first points in the theorem remain true if we take the limsup instead of the limit in the definition of Λ given in $\star$. It gives a really general upper bound for compact sets.
2. We cannot have the lower bound in the same generality, in particular because this method relies on convexity properties, hence systematically gives a convex rate function. It may happen that this is not the good one to describe the real behavior of the process. We need an additional assumption to ensure we are (at least locally) in a convex setting: this gives Gärtner-Ellis theorem, stated below.

Definition 2.1.5. *$x \in \mathcal{X}$ is an* exposed point *for Λ^* if there exists $\lambda \in \mathcal{X}^*$,* exposing hyperplane *for Λ and x, such that:*

$$\langle\lambda, x\rangle - \Lambda^*(x) > \langle\lambda, z\rangle - \Lambda^*(z) \quad \forall z \neq x$$

We denote $\mathcal{F}$ the set of exposed points of Λ^* with an exposing hyperplane λ for which there exists $\gamma > 1$ such that $\Lambda(\gamma\lambda) < \infty$.

Theorem 2.1.4 (Gärtner-Ellis-Baldi). *Let $(\mu_n)_{n\geq 0}$ be exponentially tight and verifying ($\star$), then:*

$$\text{For any closed set } F \subset \mathcal{X} \quad \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \Lambda^*(x)$$
$$\text{For any open set } G \subset \mathcal{X} \quad \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x\in G\cap\mathcal{F}} \Lambda^*(x)$$

Corollary 2.1.2. *On $\mathcal{X}$ a Hausdorff locally convex topological vector space, let $(\mu_n)_{n\geq 0}$ exponentially tight and satisfying ($\star$). If Λ is finite and Gâteaux-differentiable everywhere, then $(\mu_n)_{n\geq 0}$ satisfies a Large Deviations Principle with the convex good rate function Λ^*.*

2.2 Some large deviation principles

We present in this Section different types (called levels) of Large Deviations results for i.i.d. random variables and Markov chains. We may apply in these cases the methods we have seen in previous Section, to obtain in some particular cases WLDP's and exponential tightness.

2.2.1 Level 1 large deviations

We have already seen the Cramer Theorem, for variables taking value in $\mathbb{R}$. This result can be generalized to a wider setup:

Theorem 2.2.1. *If (X_i) are i.i.d.r.v. with values in a separable Banach space $\mathcal{X}$, the law of the empirical mean $M_n = \frac{1}{n}\sum_{i=0}^{n-1} X_i$ satisfies a WLDP with the convex rate function :*

$$\Lambda^*(x) = \sup_{\lambda \in \mathcal{X}^*} \left(\langle \lambda, x\rangle - \Lambda(\lambda)\right) \quad \textit{where} \quad \Lambda(\lambda) = \log \int_{\mathcal{X}} e^{\langle \lambda, x\rangle}\, \mu(dx)$$

If in addition $\int_{\mathcal{X}} e^{t\|x\|}\, \mu(dx) < \infty$ for all t, the law of M_n satisfies the Large Deviations Principle with good rate function Λ^*.

First part of this Theorem is obtained by subadditivity for open convex sets and then identification of the rate function with Λ^* by convexity (see theorem 2.1.3 and remarks after).
Second part is obtained by Gärtner-Ellis Theorem. In the case where $\mathcal{X} = \mathbb{R}^d$ it is enough to suppose $0 \in \overset{\circ}{\mathcal{D}}_\Lambda$ to obtain the complete Large Deviations Principle.

2.2.2 Level 2 large deviations

We would like to get more general results for processes like $1/n \sum f(X_i)$ with any observable f. We will in fact state more abstract results for the empirical measure associated to the process $(X_i)_{i\geq 0}$, defined as:

$$L_n = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{X_i}$$

and considered as a random variable with value in $\mathcal{M}_1(\mathcal{X})$. Then:

Theorem 2.2.2 (Sanov). *For (X_i) i.i.d.r.v. with values in $\mathcal{X}$ a Polish space, the laws of L_n satisfy a Large Deviations Principle with good convex rate function:*

$$H(\nu|\mu) = \begin{cases} \int_{\mathcal{X}} f \log f \, d\mu & \textit{if } f = \frac{d\nu}{d\mu} \textit{ exists,} \\ \infty & \textit{otherwise.} \end{cases}$$

We call $H(\cdot|\mu)$ the relative entropy *w.r.t. μ, or the Kullback-Leibler information.*

This Theorem is in fact a particular case of the previous Cramer Theorem applied to the sequence of i.i.d. random variables (δ_{X_i}). We get hence an alternative expression of the relative entropy by the variational expression:

$$H(\nu|\mu) = \sup_{V \in C_b(\mathcal{X})} \left(\langle f, \nu \rangle - \log \int_{\mathcal{X}} e^{f(x)} \, \mu(dx) \right)$$

By the contraction principle (Proposition 2.1.2), this result contains all Large Deviations results of level 1 for empirical means $1/n \sum f(X_i)$, with f bounded and continuous.
These first results have already many applications, in particular to study mean-field models in statistical mechanics and the feature of propagation of chaos, see for example [6] for recent developments.

The case of Markov chains

Many results have also been developed to generalize such Large Deviations Principles to the class of Markov Chains (or Markov Processes) and more generally to mixing stationary processes. It is an important step to dynamical systems, since Markov chains can be seen as (the simplest) dynamical systems. We define (X_n) a Markov chain on a Polish space $\mathcal{X}$ by its transition kernel $\pi(x, dy) = \mathbb{P}(X_1 \in dy | X_0 = x)$, describing the probability of a transition from x to y. We define then $\mathbb{P}_x$ as the law of the chain starting from the point x:

$$\mathbb{P}_x((X_0, \ldots, X_n) \in A_0 \times \cdots \times A_n) = \int_{A_1} \cdots \int_{A_n} \pi(x_{n-1}, dx_n) \cdots \pi(x, dx_1) \delta_{A_0}(x)$$

and $\mathbb{P}_\mu = \int \mathbb{P}_x \, \mu(dx)$ the law of the chain starting with initial measure μ.

We want in this context again LD results for the associated empirical measure

$$L_n = \sum_{i=1}^{n} \delta_{X_i} \in \mathcal{M}_1(\mathcal{X})$$

The situation gives more choice : we may want results under $\mathbb{P}_\mu$, μ being the invariant measure of the chain, or under $\mathbb{P}_x$. Under $\mathbb{P}_x$, we must also specify if we want pointwise or uniform results. There are hence many different results under various assumptions. We will present only some of them. More details can be found in the reference books [32] and [34].

A first result has been obtained under a strong uniformity condition on the generalized transition probabilities $\pi^l(x,dy) = \mathbb{P}_x(X_l \in dy)$ We call it the (U) hypothesis:

(U) There exists $l \le N$ integers and $M \ge 1$ such that:

$$\pi^l(x,dz) \le \frac{M}{N}\sum_{m=1}^{N} \pi^m(y,dz) \quad \forall x,y \in \mathcal{X}$$

Theorem 2.2.3. *If the Markov Chain (X_n) satisfies Assumption (U) then:*

- *There exists a unique invariant measure μ.*
- *The limit $\Lambda(V) = \lim_{n\to\infty} \frac{1}{n}\log \mathbb{E}_x\left(e^{\sum_{i=0}^{n-1} V(x_i)}\right)$ exists for all $V \in C_b(\mathcal{X})$.*
- *$\Lambda^*(\nu) = \sup_{V\in C_b(\mathcal{X})}(\langle V,\nu\rangle - \Lambda(V))$ is a good convex rate function.*
- *The Large Deviations Principle is satisfied uniformly in x :*

$$-\inf_{\nu\in \mathring{A}} \Lambda^*(\nu) \le \liminf_{n\to\infty} \frac{1}{n}\log \inf_{x\in\mathcal{X}} \mathbb{P}_x(L_n \in A)$$
$$\le \limsup_{n\to\infty} \frac{1}{n}\log \sup_{x\in\mathcal{X}} \mathbb{P}_x(L_n \in A) \le -\inf_{\nu\in\bar{A}} \Lambda^*(\nu)$$

- *Finally, with $B(\mathcal{X})$ the set of bounded measurable functions on $\mathcal{X}$ and $\mu\pi = \int_{\mathcal{X}} \pi(x,\cdot)\mu(dx)$, we have:*

$$\Lambda^*(\nu) = \sup_{\substack{u\in C_b(\mathcal{X})\\ u\ge 1}} \left\{-\int_{\mathcal{X}} \log\frac{\pi u}{u}d\nu\right\} = \sup_{\substack{\mu\in\mathcal{M}_1(\mathcal{X})\\ \log\frac{d\mu\pi}{d\mu}\in B(\mathcal{X})}} \left\{-\int \log\frac{d\mu\pi}{d\mu}d\nu\right\}$$

This result is proven in [34] by the subadditivity method exposed in Theorem 2.1.2. It contains in fact Sanov Theorem as a particular case.

Assumption (U) is really too strong: for example the Ornstein-Uhlenbeck process, with generator $\frac{1}{2}\Delta + x\frac{\partial}{\partial x}$, doesn't satisfy this. Donsker and Varadhan have proposed in [37] another method which allows to relax this assumption.

Mixing processes

Another point of view is to consider a Markov Chain under its invariant measure as a particular case of a stationary process. In this general setup, Large Deviations Principle have been derived under some mixing assumptions. We can not hope to get any result uniformly in the initial condition in this way.

Hence, for $(X_k)_{k\geq 0}$ a stationary sequence of random variables, we define:

$$\mathcal{F}_a^b = \sigma\{X_i \,:\, a \leq i \leq b\}$$

and define the condition (S) of mixing, first introduced by Bryc and Dembo in [16]:

(S) For all $C > 0$, there exists a non-decreasing sequence $l(n)$ such that :

$$\sum_{n\geq 1} \frac{l(n)}{n(n+1)} < \infty$$

$$(S_-)\ \sup\Big\{P(A)P(B) - e^{l(n)}P(A\cap B)\ :\ A \in \mathcal{F}_0^{k_1},\, B \in \mathcal{F}_{k_1+l(n)}^{k_1+k_2+l(n)},\, k_1, k_2 \in \mathbb{N}\Big\} \leq e^{-Cn}$$

$$(S_+)\ \sup\Big\{P(A\cap B) - e^{l(n)}P(A)P(B)\ :\ A \in \mathcal{F}_0^{k_1},\, B \in \mathcal{F}_{k_1+l(n)}^{k_1+k_2+l(n)},\, k_1, k_2 \in \mathbb{N}\Big\} \leq e^{-Cn}$$

Theorem 2.2.4. *If the sequence of stationary random variables satisfies Assumption (S), then the law of L_n satisfies a Large Deviations Principle (for the τ-topology) in $\mathcal{M}_1(\mathcal{X})$ with rate function :*

$$\Lambda^*(\nu) = \sup_{V\in B(\mathcal{X})} (\langle V, \nu\rangle - \Lambda(V)) \quad \textit{where} \;\; \Lambda(V) = \lim_{n\to\infty} \frac{1}{n} \log E\left(e^{\sum_{i=1}^{n-1} V(X_i)}\right)$$

The τ-topology is the weakest topology on $\mathcal{M}_1(\mathcal{X})$ such that $\nu \mapsto \langle V, \nu\rangle$ is continuous for all $V \in B(\mathcal{X})$. It is finer than the weak topology, hence a Large Deviations Principle in the τ-topology is a stronger result than in the weak topology. In fact, Sanov Theorem is also available in the τ-topology (this is in particular implied by this last result).
The proof of this result relies again on a subadditivity argument, in an approximate version: if $f(n+m) \leq f(n) + f(m) + \delta(n+m)$ with $\delta(n)$ non-decreasing and $\sum_{k\geq 1} \frac{\delta(k)}{k(k+1)} < \infty$ then $\lim_{n\to\infty} \frac{f(n)}{n}$ exists and :

$$\lim_{n\to\infty} \frac{f(n)}{n} \leq \frac{f(m)}{m} - \frac{\delta(m)}{m} + 4 \sum_{k\geq 2m} \frac{\delta(k)}{k(k+1)} \qquad \forall m$$

Another mixing condition has been first introduced by Chiyonobu and Kusuoka in [27], and called the hypermixing condition. We give here a slightly different version, as stated in [16] or [32]:

$(H2)$ There exists $\beta(l) \in [1,\infty]$, $\gamma(l) \geq 0$ such that for all $k_1, k_2 \in \mathbb{N}$, $W \in L_\infty\left(\mathcal{F}_0^{k_1}\right)$, $Z \in L_\infty\left(\mathcal{F}_{k_1+l}^{k_1+k_2+l}\right)$:

$$|\mathbb{E}(W)\mathbb{E}(Z) - \mathbb{E}(WZ)| \leq \gamma(l)\left(\mathbb{E}|W|^{\beta(l)}\right)^{\frac{1}{\beta(l)}}\left(\mathbb{E}|Z|^{\beta(l)}\right)^{\frac{1}{\beta(l)}}$$

where $\lim_{l\to\infty}\gamma(l) = 0$ and $\lim_{l\to\infty}(\beta(l)-1)l(\log l)^{1+\delta}$ for some $\delta > 0$.

Bryc and Dembo have then proven that $(H2)$ implies the previous (S) Assumption, so we get the same level 2 Large Deviations Principle under this Assumption $(H2)$.

The importance of this hypermixing property comes from the fact that, for Markov Processes, hypermixing is implied by the hypercontractivity of the associated semi group, which is:

$$\exists\, T > 0 \;\text{ such that }\; \|P_T\|_{L_2 \mapsto L_4} = 1$$

And hypercontractivity is equivalent in a symmetric setting to the log-Sobolev inequality. We refer the reader to [34] for all these developments, and to the recent works of Stroock and Zegarlinski for applications of these concepts to the study of Glauber dynamics for Gibbs measures.

Many other mixing conditions exist. We do not detail them here, referring the reader to [16] for all their definitions and links between them.

2.2.3 Level 3 large deviations

We present now even more general Large Deviations results, concerning a measure associated to the whole history of the process $(X_i)_{i\geq 0}$. We define for this the associated empirical process:

$$R_n = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i(X)} \in \mathcal{M}^1\left(\mathcal{X}^{\mathbb{N}}\right) \tag{2.1}$$

where $X = (X_0, X_1, \ldots)$ is the entire process, and $T^i(X) = (X_i, X_{i+1}, \ldots)$ is the temporal shift by i.
Alternatively, we may also define

$$\tilde{R}_n = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i\tilde{X}_n} \in \mathcal{M}_1^s\left(\mathcal{X}^{\mathbb{Z}}\right)$$

with $\tilde{X}_n = (\ldots, X_0, \ldots, X_{n-1}, X_0, \ldots, X_{n-1}, \ldots)$ a periodic sequence of period n constructed from X. $\tilde{R}_n$ is the periodic transformation of R_n. Its principal interest is that it is with values in the set of stationary (for the shift T) probability measures on $\mathcal{X}^{\mathbb{Z}}$, which is noted $\mathcal{M}_S^1\left(\mathcal{X}^{\mathbb{Z}}\right)$. Large Deviations results for $\tilde{R}_n$ and for R_n will be equivalent.

Using again the Contraction Principle, we may note that a Level 3 result will contain the Level 2 one, by projection on the first marginal, but also lot of other estimates, for example for the behavior of an observable like $1/n \sum_{k=0}^{n-1} f(X_k, X_{k+1}, \ldots, X_{k+l})$. These estimates are of first importance for applications to Statistical Mechanics, for example when one wants to evaluate correlations between neighboring sites.

We present the way to obtain such results for a Markov Chain in the uniform case. We prove in fact first a result for the k-empirical measure, then go to the empirical process by a projective limit argument.
For two finite sequences $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_k)$, we define:

$$\pi_k(x, dy) = \pi(x_k, dy_k) \prod_{i=1}^{k-1} \delta_{x_{i+1}}(y_i)$$

Theorem 2.2.3 applies to the Markov chain $(X_i, \ldots, X_{i+k})_{i \geq 0}$ on $\mathcal{X}^k$, with transition kernel π_k, to give:

Theorem 2.2.5. *If the Markov Chain (X_n) satisfies Assumption (U), then the laws of the k-empirical measures $L_n^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i, \ldots, X_{i+k}} \in \mathcal{M}_1\left(\mathcal{X}^k\right)$ satisfy a Large Deviations Principle with good convex rate function :*

$$I_k(\nu) = \sup_{\substack{u \in C_b\left(\mathcal{X}^k\right) \\ u \geq 1}} \left\{ - \int_{\mathcal{X}} \log \frac{\pi_k u}{u} \, d\nu \right\}$$

We can obtain for $k \geq 2$ a better expression for I_k . If, for $\mu \in \mathcal{M}^1(\mathcal{X}^k)$, we note $p_i \mu \in \mathcal{M}^1(\mathcal{X}^i)$ its i first marginals, and for $\mu \in \mathcal{M}_1\left(\mathcal{X}^{k-1}\right)$:

$$(\mu \otimes_k \pi)(A) = \int_{\mathcal{X}^{k-1}} \left[\int_{\mathcal{X}} 1_{\{(x,y) \in A\}} \pi(x_{k-1}, dy) \right] \mu(dx)$$

then :

$$I_k(\nu) = \begin{cases} H(\nu | (p_{k-1}\nu) \otimes_k \pi) & \text{if } \nu \text{ is shift-invariant,} \\ \infty & \text{otherwise.} \end{cases}$$

From this result for the k-empirical measures we can deduce by a projective limit argument the level 3 Large Deviations Principle:

Theorem 2.2.6. *If the Markov Chain (X_n) satisfies Assumption (U), the law of $R_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{T^i X} \in \mathcal{M}_1\left(\mathcal{X}^{\mathbb{N}}\right)$ satisfies a Large Deviations Principle with good rate function:*

$$I_U(Q) = \begin{cases} \sup_{k\geq 2} H(p_k Q | p_{k-1}Q \otimes_k \pi) & \textit{if } Q \textit{ is shift-invariant,} \\ \infty & \textit{otherwise.} \end{cases}$$

We have in fact a simpler expression for the rate function :

$$I_U(Q) = \begin{cases} H(Q_1^* | Q_0^* \otimes_0 \pi) & \text{if } Q \text{ is shift-invariant,} \\ \infty & \text{otherwise.} \end{cases}$$

with $Q_0^* \in \mathcal{M}^1\left(\mathcal{X}^{\mathbb{Z}_-}\right)$ and $Q_1^* \in \mathcal{M}^1\left(\mathcal{X}^{\mathbb{Z}_-\cup\{1\}}\right))$ are defined by

$$\begin{aligned} & Q_0^*(x : (x_{1-k}, \dots, x_0) \in A) = (p_k Q)(A) \\ \text{and } & Q_1^*(x : (x_{2-k}, \dots, x_1) \in A) = (p_k Q)(A) \end{aligned}$$

For hypermixing Markov Chains, we have to add to Assumption $(H2)$ another one:

$(H1)$ There exists $l, \alpha < \infty$ such that for all $k, k_1, \dots, k_k$ and for all $(W_i)_{1\leq i\leq k}$ satisfying $W_i \in L_\infty\left(\mathcal{F}_{k_1+\cdots+k_{i-1}+(i-1)l}^{k_1+\cdots+k_i+(i-1)l}\right)$:

$$\mathbb{E}\Big(\Big|\prod_{i=1}^{k} W_i\Big|\Big) \leq \prod_{i=1}^{k} \mathbb{E}\big(|W_i|^\alpha\big)^{1/\alpha}$$

We can then proceed under these two conditions as in the uniform case to get:

Theorem 2.2.7. *If the Markov Chain (X_n), with stationary measure P satisfies Assumptions $(H1)$and $(H2)$, the law of R_n satisfies a Large Deviations Principle with good rate function:*

$$I(Q) = \begin{cases} I_U(Q) & \textit{if } p_k Q \ll p_k P \textit{ for all } k \geq 1, \\ \infty & \textit{otherwise.} \end{cases}$$

Chiyonobu and Kusuoka obtained (in [27], see also [34]) a Large Deviations Principle for R_n under the invariant measure of a stationary process satisfying slightly different hypermixing assumptions with the affine rate function $I_\infty(Q) = \lim_{k\to\infty} \frac{1}{k} H(p_k Q | p_k P_\mu)$ called the specific entropy.

Under condition (S), the level 3 Large Deviations Principle holds with the different rate function:

$$\overline{I}(Q) = \sup_{k\geq 2} \Lambda^*_{(k)}(p_k Q)$$

with $\Lambda^*_{(k)}$ the Legendre transform of:

$$\Lambda^{(k)}(V) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\Big(\exp\Big(\sum_{i=0}^{n-1} V(Y_i, \ldots, Y_{i+k-1})\Big)\Big)$$

We know for this rate function only that $\overline{I}(Q) = \infty$ if Q is not shift-invariant, and that $\overline{I} \leq I_\infty$, but this inequality can be strict.

All previous cases can be specialized to the of (X_i) i.i.d. of law μ. We get then for the rate function:

$$\begin{aligned} I(Q) &= \lim_{n\to\infty} \frac{1}{n} H(p_n Q|\mu^{\otimes n}) \\ &= \sup_{k\geq 2} H(p_k Q|p_{k-1}Q \otimes_k \mu) = H(Q_1^*|Q_0^* \otimes_0 \mu) \end{aligned}$$

And we can get in this setup a simpler expression of last term, giving the following:

Theorem 2.2.8. *If (X_n) are i.i.d.r.v. with law μ, then the laws of R_n satisfy a Large Deviations Principle with rate function:*

$$I(Q) = \int_{\mathcal{X}^{\mathbb{Z}}} H(Q^*_{1|\omega}|\mu)\, Q(d\omega)$$

where $Q^*_{1|\omega}$ is a regular version of the conditional probability $Q_1^*(\cdot|X_j, j\leq 0)(\omega)$.

2.3 Survey of large deviation results for dynamical systems

We present in this Section the existing large deviations results for dynamical systems. We consider a map $f : \mathcal{X} \to \mathcal{X}$ with $(\mathcal{X}, d)$ a compact metric space and study the law of the associated temporal empirical measure

$$R_T = \frac{1}{T} \sum_{t=0}^{T-1} \delta_{f^t(x)}$$

under a given initial measure.
If the context can look really different, due to the deterministic transition from x to $f(x)$, it is sufficient to look at the Definition of the empirical process associated to a stationary random process given in (2.1) to see that it could be understood as the action of the (temporal) shift on the space $\mathcal{X}^{\mathbb{N}}$. This is the viewpoint we will adopt in the sequel of this Chapter, and in the remaining of this thesis.
However, the context is different enough to make us express the rate functions in a more adequate formulations, using the setup of Thermodynamic formalism. This is the reason why we recall in first Subsection some facts of this theory. For a more general approach and all the proofs, we refer to the well-written introductory book of G. Keller [66].

Let us start by mentioning two essential properties that a dynamical system f on $(\mathcal{X}, d)$ may satisfy:

Definition 2.3.1. *f satisfies* expansiveness *with constant δ if, for $x, y \in \mathcal{X}$:*

$$d(f^t x, f^t y) < \delta \quad \forall\, t \geq 0 \quad \Rightarrow \quad x = y$$

Definition 2.3.2. *f satisfies* specification *if for all $\delta > 0$, there exists $p(\delta)$ such that for any $k \in \mathbb{N}^*$, $x_1, \cdots, x_k \in \mathcal{X}$, $T_1, \cdots, T_k \in \mathbb{N}$, $p_1, \cdots, p_{k-1} \geq p(\delta)$, there exists $x \in \mathcal{X}$ with:*

$$\begin{aligned}
d\left(f^t x, f^t x_1\right) &< \delta \quad \text{for } 0 \leq t < T_1 \\
d\left(f^{T_1+p_1+t} x, f^t x_2\right) &< \delta \quad \text{for } 0 \leq t < T_2 \\
&\vdots \\
d\left(f^{T_1+T_2+\cdots+T_{k-1}+p_1+\cdots+p_{k-1}+t} x, f^t x_k\right) &< \delta \quad \text{for } 0 \leq t < T_k
\end{aligned}$$

2.3.1 Thermodynamic Formalism

Ergodicity

Definition 2.3.3. *$\mathcal{M}^1_{inv}(\mathcal{X})$ denotes the set of probability measures which are* invariant *under f.*

An invariant measure ν is ergodic *if $\nu(A) = 0$ or 1 for any A invariant by f. We denote $\mathcal{M}^1_{erg}(\mathcal{X})$ the set of ergodic probabilities.*

The main result for ergodic measures is the ergodic theorem:

Theorem 2.3.1 (Ergodic Theorem). *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$, then for all $g \in L^1(\nu)$ and for ν-almost all x:*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} g \circ f^t(x) = \int_{\mathcal{X}} g \, d\nu$$

Entropy

For $\mathcal{A} = \{A_1, \ldots, A_K\}$ and $\mathcal{B} = \{B_1, \ldots, B_L\}$ finite partitions of $\mathcal{X}$, let

$$\mathcal{A} \vee \mathcal{B} = \{A_k \cap B_l \,:\, 1 \leq k \leq K, 1 \leq l \leq L\}$$

Then, for $\nu \in \mathcal{M}^1_{\text{inv}}(\mathcal{X})$ and $\mathcal{A}$ a partition of $\mathcal{X}$, we define:

- $h(\nu|\mathcal{A}) = -\sum_{A\in\mathcal{A}} \nu(A)\log(\nu(A))$ and $\mathcal{A}_T = \bigvee_{0\leq t<T} f^{-t}(\mathcal{A})$
- $h_f(\nu|\mathcal{A}) = \lim_{T\to\infty} \frac{1}{T} h(\nu|\mathcal{A}_T)$
- $h_f(\nu) = \sup\{h_f(\nu|\mathcal{A}) \,:\, \mathcal{A} \text{ finite partition of } \mathcal{X}\}$

This last quantity is the metric entropy of ν under f.

Proposition 2.3.1. *h_f is convex affine: if $\nu = \sum_{l=1}^L a_l \nu_l$ with $\sum_{l=1}^L a_l = 1$, then $h_f(\nu) = \sum_{l=1}^L a_l h_f(\nu_l)$.*

If furthermore f is expansive with constant δ, then:

1. *For $\nu \in \mathcal{M}^1_{inv}(\mathcal{X})$ and for any partition $\mathcal{A}$ such that $\nu(\partial\mathcal{A}) = 0$ and $\operatorname{diam}(\mathcal{A}) < \delta$, we have:*
$$h_f(\nu) = h_f(\nu|\mathcal{A})$$
2. *h_f is upper semi-continuous.*

A well known result about entropy is the Shannon-Mc Millan-Breiman theorem, which expresses that, for an ergodic measure, entropy describes the asymptotic size of elements of the partition:

Theorem 2.3.2 (Shannon-Mc Millan-Breiman). *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$ and $\mathcal{A}$ is a finite partition, then for ν-almost all x:*

$$-\frac{\log \nu(\mathcal{A}_T(x))}{T} \xrightarrow[T\to\infty]{} h_f(\nu|\mathcal{A})$$

where $\mathcal{A}_T(x)$ denotes the element of the partition $\mathcal{A}_T$ which contains x.

A metric equivalent of this theorem, which tells that, for an ergodic measure, the metric entropy describes the number of balls necessary to cover a significant set, will be useful in many proofs of lower bounds of large deviations.
For $T \geq 0$, $\delta > 0$ and $0 < b < 1$, we denote

$$b_x(T;\delta) = \{y \in \mathcal{X} \,:\, d(f^t(x), f^t(y)) \leq \delta \quad \forall\, 0 \leq t \leq T\} \tag{2.2}$$

$$N(T;\delta,b) = \min\Big\{\mathrm{Card}(Y) \,:\, \nu\Big(\bigcup_{x\in Y} b_x(T,\delta)\Big) > b\Big\} \tag{2.3}$$

(we call a set Y satisfying the condition in 2.3 a $(T;\delta,b)$-covering set for ν)

Theorem 2.3.3. *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$, then for all $0 < b < 1$:*

$$h_f(\nu) = \lim_{\delta\to 0}\liminf_{T\to\infty}\frac{1}{T}\log N(T;\delta,b) = \lim_{\delta\to 0}\limsup_{T\to\infty}\frac{1}{T}\log N(T;\delta,b)$$

This characterization of metric entropy as a measure of exponential rate of decreasing for dynamical balls is first due to Katok [63]. A proof of a multi-dimensional generalization is given in Theorem 3.8.3.

Topological pressure

A set $Y \subset \mathcal{X}$ is $(T;\delta)$-separated if

$$x, x' \in Y, x \neq x' \quad \Longrightarrow \quad x' \notin b_x(T;\delta)$$

We define for $V \in \mathcal{C}(\mathcal{X})$

$$Z_f(V,\delta,T) = \sup\left\{\sum_{x\in Y}\exp\Big(\sum_{t=0}^{T-1} V\circ f^t(x)\Big) \,:\, Y\ (T,\delta)\text{-separated set}\right\}.$$

Then:

$$P_f(V) = \lim_{\delta\to 0}\limsup_{T\to\infty}\frac{1}{T}\log Z_f(V,\delta,T)$$

is the topological pressure of V for the dynamic of f. The main result for this quantity is the Gibbs Variational Principle, which defines it as a variational expression of the entropy:

Theorem 2.3.4 (Gibbs Variational Principle). *For any $V \in \mathcal{C}(\mathcal{X})$:*

$$P_f(V) = \sup_{\nu\in\mathcal{M}^1_{inv}(\mathcal{X})}\left(h_f(\nu) + \int_{\mathcal{X}} V\,d\nu\right) \tag{2.4}$$

and, if h_f is upper semi-continuous, for any $\nu \in \mathcal{M}^1_{inv}(\mathcal{X})$:

$$h_f(\nu) = \inf_{V \in \mathcal{C}(\mathcal{X})} \left(P_f(V) - \int_{\mathcal{X}} V \, d\nu \right) \tag{2.5}$$

In the particular case of $V = 0$, the pressure is called the topological entropy of the system:

$$h_{\text{top}}(f) = P_f(0) = \sup_{\nu \in \mathcal{M}^1_{\text{inv}}(\mathcal{X})} h_f(\nu)$$

2.3.2 Via mixing conditions

Orey and Pelikan prove in [83] a uniform Large Deviations Principle for stationary processes under a mixing condition (called Ratio-Mixing) and a Feller-type hypothesis (Continuous Dependence) on a compact space, adapting results of Donsker and Varadhan.
They show that this applies to Gibbs measures for subshifts of finite type and adapt it in [84] to $\mathcal{C}^2$ Anosov diffeomorphisms, using Markov partitions. The rate function is in this case the "defect" in Pesin's formula.

Stationary processes

We work on $\Omega = \Gamma^{\mathbb{Z}}$ with Γ a compact metric space and will denote by σ the shift on Ω.
We set $\Omega_- = \Gamma^{\mathbb{Z}_-}$, $\Omega_+ = \Gamma^{\mathbb{N}}$ and $\mathcal{F}_{m,n} = \sigma\{\omega_i \, : \, m \leq i \leq n\}$.
We have already mentioned that this setup is the same as for Level 3 large deviations described in Section 2.2, taking $X_i(\omega) = \omega_i$. We have then seen large deviations results under various mixing conditions.

Orey and Pelikan introduce the Ratio-Mixing condition :

(RM) There exists a non-decreasing function $m(n)$ such that $0 < m(n) < n$, $\frac{m(n)}{n} \to 0$ and :

$$\lim_{n \to \infty} \sup \left\{ \log \frac{\mu^*_{\eta_-}(A)}{\mu^*_{\omega_-}(A)} \, : \, \eta_-, \, \omega_- \in \Omega_-, \, A \in \mathcal{F}_{m(n),n} \right\} = 0$$

where $\mu^*_{\omega_-}$ are regular conditional probabilities of μ given $\mathcal{F}_{-\infty,0}$.

Define :

$$R_n(\omega) = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k \omega} \in \mathcal{M}^1(\Omega)$$
$$Q_{n,\omega_-}(A) = \mu^*_{\omega_-}(R_n \in A) \quad \text{for } A \subset \mathcal{M}^1(\Omega)$$

Then:

Theorem 2.3.5. *Let $\mu \in \mathcal{M}^1_{inv}(\Omega)$ with μ^* satisfying (RM). The conditional measures $\mu^*_{\omega_-}$ satisfy then a uniform Large Deviations Principle with rate function I:*

$$\liminf_{n\to\infty} \frac{1}{n} \log \inf_{\omega_-\in\Omega_-} Q_{n,\omega_-}(A) \geq -\inf_{\nu\in A} I(\nu) \qquad \textit{for } A \textit{ open}$$
$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{\omega_-\in\Omega_-} Q_{n,\omega_-}(A) \leq -\inf_{\nu\in A} I(\nu) \qquad \textit{for } A \textit{ closed}$$

Proof. (Main steps) They just use sub-additivity as under condition (U), establishing directly a complete Large Deviations Principle since the state space Ω is compact.
They find a basis $\mathcal{A}$ of the topology of Ω such that for any $A \in \mathcal{A}$, $Q_n(A) = \inf_{\omega_-\in\Omega_-} Q_{n,\omega_-}(A)$, is super-multiplicative and then :

$$\begin{aligned} L(A) &= \lim_{n\to\infty} \frac{1}{n} \log Q_n(A) \qquad \text{exists } \forall\ A \in \mathcal{A}, \\ I(\nu) &= -\inf\{L(A)\ :\ \nu \in A,\ A \in \mathcal{A}\} \end{aligned}$$

is lower semi-continuous and convex. Hence, using (RM) to compare Q_n and $\sup_{\omega_-\in\Omega_-} Q_{n,\omega_-}$, we get the Large Deviations Principle with rate function I. $\square$

To identify this rate function, another assumption is needed:

(CD) For every $Y : \Omega \to \mathbb{R}$ continuous and $\mathcal{F}_{-\infty,1}$-measurable, $\omega \to \mu^*_{\omega_-}(Y)$ is continuous.

We have then:

Theorem 2.3.6. *For $\mu \in \mathcal{M}^1_{inv}(\Omega)$ with $\mu^*_{\omega_-}$ satisfying (RM) and (CD), the rate function for the Large Deviations Principle in Theorem 2.3.5 is :*

$$I(\nu) = \begin{cases} \int_{\Omega_-} H\big(\nu_{\omega_-|1} | \mu^*_{\omega_-|1}\big)\nu(d\omega_-) & \textit{if } \nu \in \mathcal{M}_{inv}(\Omega) \textit{ and } \nu_{\omega_-|1} \ll \mu^*_{\omega_-|1}\ \nu\textit{-as} \\ \infty & \textit{otherwise.} \end{cases}$$

Gibbs measures

We specialize here to the case where $\Gamma = \{0, 1, \cdots, N-1\}$ and, for A an $N \times N$ matrix of 0 and 1 assumed to be irreducible (i.e. such that there exists m with $(A^m)_{i,j} > 0$ for any i, j) we denote :

$$\Omega^A = \big\{\omega \in \Omega\ :\ A_{\omega_i,\omega_{i+1}} = 1 \quad \forall i \in \mathbb{Z}\big\}$$

(Ω^A, σ) is called a subshift of finite type. A sequence $(x_i)_{i\in I}$ is A-admissible if $A_{x_i,x_{i+1}} = 1$ for every $i \in \mathbb{Z}$.

For $B_i = \{\omega \in \Omega^A : \omega_1 = i\}$, (CD) reduces in this case to :

(CD') $\omega_- \to \mu^*_{\omega_-}(B_i)$ is continuous for $0 \leq i < N$.

And if we assume $\inf\{\mu^*_{\omega_-}(B_i) : \omega_- \in \Omega^A_-, 0 \leq i < N \text{ s.t. } A_{\omega_-(0),i} = 1\} > 0$, then $\mu^*_{\omega_-}$ satisfies (RM) with $m(n) = m$.

For φ a Hölder continuous function on Ω^A, we know that there exists a unique $\mu \in \mathcal{M}^1_{\text{inv}}(\Omega^A)$ which is a Gibbs measure, that is such that there exists C, $C' > 0$, $P \in \mathbb{R}$ with:

$$C \leq \frac{\mu(\eta : \eta_i = \omega_i \quad \forall\, 0 \leq i < n)}{\exp\left(-Pn + \sum_{i=0}^{n-1} \varphi\left(\sigma^k \omega\right)\right)} \leq C'$$

This result on uniqueness of Gibbs measures is proven in the Lecture Notes of Bowen [11], using the Ruelle-Perron-Frobenius Theorem.

We may apply Theorem 2.3.6 to this measure μ to get:

Theorem 2.3.7. *There exists a version μ^* satisfying (CD) and (RM), hence the Large Deviations Principle holds for the laws of R_n under any initial measure μ^*_ω with:*

$$I(\nu) = \begin{cases} -h_\sigma(\nu) - \int_{\Omega^A} \varphi \, d\nu & \text{if } \nu \text{ shift-invariant,} \\ \infty & \text{otherwise.} \end{cases}$$

Anosov case

This result applies also for uniformly hyperbolic dynamical systems. For $\mathcal{X}$ a Riemannian compact manifold, we take now $f : \mathcal{X} \to \mathcal{X}$ a transitive (i.e. such that there is a point with dense orbit) C^2 Anosov diffeomorphism, that is with a continuous splitting of the tangent bundle $T\mathcal{X} = E^s + E^u$ with $C > 0$ and $\lambda \in (0, 1)$ such that :

$$\begin{aligned} \|Df^n(x) \cdot v\| &\leq C\lambda^n \|v\| \quad \forall n \geq 0,\, v \in E^s_x \\ \|Df^{-n}(x) \cdot v\| &\leq C\lambda^n \|v\| \quad \forall n \geq 0,\, v \in E^u_x \end{aligned}$$

We say also in this case that f is uniformly hyperbolic on $\mathcal{X}$. Then $\varphi^u(x) = -\log\left|\det\left(Df(x)_{|E^u_x}\right)\right|$ is Hölder continuous and we can use Markov partitions construction to code this dynamical system.
It allows to construct a transition matrix A and the coding $\pi : \Omega^A \to \mathcal{X}$ (see Chapters 18 and 19 of [64] for details on this construction).

Let $\overline{\varphi} = \varphi^u \circ \pi$. $\overline{\varphi}$ is also Hőlder continuous, but now on Ω^A. We can then use previous result for the associated unique Gibbs measure $\overline{\mu}$. Gibbs measures are in fact also equilibrium measures, hence $\overline{\mu}$ is characterized as the unique minimizer of $\nu \to h_f(\nu) + \int_{\Omega^A} \overline{\varphi}\, d\nu$ among f-invariant probability measures. And its image $\mu = \pi(\overline{\mu})$ is the unique equilibrium measure on $(\mathcal{X}, f)$ for φ. It is also called the SRB (Sinai-Ruelle-Bowen) measure of the system.
We get then in this context, by an application of the Contraction Principle:

Theorem 2.3.8. *With m the normalized Lebesgue measure on $\mathcal{X}$, the Large Deviations Principle holds on $\mathcal{X}$ under m for $R_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k(x)}$ with rate function :*

$$I(\nu) = \begin{cases} -h_\nu(f) - \int_{\mathcal{X}} \varphi^u \, d\nu & \textit{if } \nu \in \mathcal{M}^1_{inv}(\mathcal{X}), \\ \infty & \textit{otherwise.} \end{cases}$$

During the proof, we obtain also the Large Deviations Principle under the SRB measure μ with the same rate function.
We can identify $I(\nu) = -h_\nu(f) + \sum_{\lambda_i > 0} \lambda_i$, with λ_i the ν-Lyapunov exponents for f. So $I(\nu) \geq 0$ means exactly $h_\nu(f) \leq \sum_{\lambda_i > 0} \lambda_i$ which is the Pesin formula: the rate function I measures the defect in Pesin formula.

ψ-mixing

Denker gives in [33] a proof of the Large Deviations Principle for the behavior of $1/n \sum g \circ \sigma^i$ under the Gibbs measure associated to a subshift of finite type, when g is Hőlder continuous. This result is an equivalent of level 1 LDP in Section 2.2.

He states for this that a Gibbs measure is ψ-mixing, that is (his definition is slightly different from that in [16]):

$$\psi(n) = \sup \left\{ \left| \frac{\mu(A \cap B)}{\mu(A)\mu(B)} - 1 \right| \; : \; A \in \mathcal{F}_{-\infty,0},\; B \in \mathcal{F}_{n,+\infty} \right\}$$

decreases to zero at an exponential rate.

2.3.3 Via volume estimates and Shannon Theorem

Seminal study of Takahashi

The first approach to large deviations theory for dynamical systems was in fact done by Takahashi in the early 80's. In particular, in [98], if he does not state explicitly any large deviations result, he precisely describes the natural objects linked to this theory.

He defines for $f : \mathcal{X} \to \mathcal{X}$ and m a probability measure:

$$
\begin{aligned}
P(U) &= \limsup_{n\to\infty} \frac{1}{n} \log \int_{\mathcal{X}} \exp\Big(- \sum_{i=0}^{n-1} U(f^i(x))\Big)\, m(dx) \\
h(\nu) &= \inf_{U\in\mathcal{C}(\mathcal{X})} \Big(P(U) + \int_{\mathcal{X}} U\, d\nu \Big) \\
Q(G) &= \limsup_{n\to\infty} \frac{1}{n} \log m\Big\{ x \in \mathcal{X} \,:\, \frac{1}{n}\sum_{i=0}^{n-1} \delta_{f^i(x)} \in G \Big\} \\
g(\nu) &= \inf\{ Q(G) \,:\, \nu \in G \text{ open in } \mathcal{M}^1(\mathcal{X})\}
\end{aligned}
$$

He proves that $g \le h$.

In the particular case of the shift σ on a finite space Ω with m having a positive continuous version j_m of Jacobian ($j_m(ax) = m(\sigma^{-1}(\omega_0) = a|x)$ m-a.e.), he gets that both functionals are equal and identifies them with quantities intrinsically linked to the system:

$$
h(\nu) = g(\nu) = h_\sigma(\nu) + \int_{\Omega^{\otimes\mathbb{N}}} \log j_m \, d\nu
$$

To prove that $g(\nu) \ge h_\sigma(\nu) + \int_{\Omega^{\otimes\mathbb{N}}} \log j_m\, d\nu$, he uses the fact that for ν ergodic, $\log j_m$ describes the exponential speed of decreasing of the measure under m of a cylinder, and $h_f(\nu)$ the exponential rate of the number of cylinders necessary to cover a ν significant set (as is described by Shannon-Mac Millan-Breiman Theorem, see 2.3.2).

This method is not far from a complete proof of large deviations. He would have done it taking liminf instead of limsup in the definition of Q. He does this remark and some other extensions in his later paper [99].

Another interesting point of this seminal paper [98] is the conjecture that the functional f is affine when the map f is structurally stable. He proves it in a particular setting.

The same kind of method has been generalized to many other situations. We give these developments here, and the generalization to the setup of maps satisfying expansiveness and specification in Section 2.3.4.

Gibbs systems

The complete development of this method to get a general large deviations result for Gibbs measures on multidimensional shift systems has been done by

Főllmer and Orey in [82, 42] (at the same time, different proofs of the same result where given in [81, 31]).

On $\mathcal{X} = \Gamma^{\mathbb{Z}^d}$ with Γ a finite set, they study the dynamics of the shifts $(\sigma^k)_{k\in\mathbb{Z}^d}$ defined by $(\sigma^k x)_i = x_{k+i}$ and a stationary interaction potential (U_V) indexed by all finite subsets V of $\mathbb{Z}^d$ such that:

$$\sum_{0\in V} \frac{\|U_V\|}{|V|} < \infty$$

Definition 2.3.4. *A measure $\mu \in \mathcal{M}^1_{inv}(\mathcal{X})$ is a Gibbs measure associated to U if, for any boundary condition η, the conditional measures satisfy:*

$$\mu_V(x_V|\eta) = \mathcal{Z}_V(\eta)^{-1} \exp\Big(- \sum_{A\cap V\neq\emptyset} U_A(x_V \vee \eta_{V^C})\Big)$$

Főllmer and Orey prove that the associated empirical process

$$R_n(x) = \frac{1}{n^d} \sum_{i\in V_n} \delta_{\sigma^i x}$$

(where $V_n = [0,n)^d$) satisfies under any Gibbs measure μ a Large Deviations Principle with rate function:

$$\begin{aligned} I(\nu) &= \lim_{n\to\infty} \frac{1}{n^d} H\left(\nu|_{V_n} \mid \mu|_{V_n}\right) \\ &= -h_\sigma(\nu) + \int_{\mathcal{X}} \sum_{0\in V} \frac{U_V}{|V|}\, d\nu + P_\sigma\Big(- \sum_{0\in V} \frac{U_V}{|V|}\Big) \end{aligned}$$

Their method is a complete development of the arguments proposed by Takahashi. It has been recently generalized to multidimensional subshifts of finite type by Eizenberg, Kifer and Weiss in [39].
The fact that we study the system under a Gibbs measure, for which the weight of cylinders is well known, plays a great role in these proofs.

Generalization to smooth systems

Young generalizes in [105] this method by noting that this Gibbs property can be replaced in some smooth contexts by a result of the type Volume Lemma. Specifically, she uses in an abstract setup local estimates of entropy and sharp description of the size of balls under the initial measure to get large deviations results.
She works with $f : \mathcal{X} \to \mathcal{X}$ continuous on a compact metric space $\mathcal{X}$, and m a probability measure. Denote $S_n\varphi = \sum_{i=0}^{n-1} \varphi \circ f^i$, and, as in Section 2.3.1:

$$b_x(T;\delta) = \{y \in \mathcal{X} \,:\, d(f^t(x), f^t(y) \le \delta \quad \forall\, 0 \le t \le T\}$$

She gets in this context an upper bound assuming the existence of a potential describing the size of balls under m:

Theorem 2.3.9. *If there exists $\xi \in \mathcal{C}(\mathcal{X})$ and constants $C, \varepsilon > 0$ such that:*

$$m(b_x(n,\varepsilon)) \le Ce^{-S_n\xi(x)} \quad \forall\, x \in \mathcal{X}, n \ge 0$$

then for every $\varphi \in \mathcal{C}(\mathcal{X})$ and $c \in \mathbb{R}$, we get :

$$\limsup_{n\to\infty} \frac{1}{n} \log m\left(\frac{1}{n}S_n\varphi \ge c\right) \le -\inf\Big\{ \int_{\mathcal{X}} \xi\, d\nu - h_f(\nu) \ : \\ \nu \in \mathcal{M}^1_{inv}(\mathcal{X}) \ \ s.t. \ \ \int_{\mathcal{X}} \varphi\, d\nu \ge c\Big\}$$

She gets also a lower bound of the same type under the additional assumption of specification (see Definition 2.3.2) which allows "gluing orbits" to go from ergodic measures to general ones:

Theorem 2.3.10. *If f satisfies specification and there exists $\xi \in \mathcal{C}(\mathcal{X})$ such that there is ε arbitrarily small and $C = C(\varepsilon) > 0$ with:*

$$m(b_x(n,\varepsilon)) \ge Ce^{-S_n\xi(x)} \quad \forall\, x \in \mathcal{X},\, n \ge 0$$

then for every $\varphi \in C(\mathcal{X})$ and $c \in \mathbb{R}$, we get:

$$\liminf_{n\to\infty} \frac{1}{n} \log m\left(\frac{1}{n}S_n\varphi > c\right) \ge -\inf\Big\{ \int_{\mathcal{X}} \xi\, d\nu - h_f(\nu) \ : \\ \nu \in \mathcal{M}^1_{inv}(\mathcal{X}) \ \ s.t. \ \ \int_{\mathcal{X}} \varphi\, d\nu > c\Big\}$$

Although she states these large deviations estimates in some particular way, the method can be extended to a general Large Deviations Principle for the empirical measure linked to the system, in particular because the estimates are available for any continuous observable φ.

From this abstract result, Young obtains various applications to concrete cases. For instance, if $f : M \to M$ is a $\mathcal{C}^2$ diffeomorphism on a $\mathcal{C}^\infty$ Riemannian manifold M, she gets:

Theorem 2.3.11. *If $\Lambda \subset U$ is an Axiom-A attractor in M, i.e.:*

- *Λ is compact invariant and uniformly hyperbolic for f,*
- *U is an open set such that $\overline{U}$ is compact, $f\overline{U} \subset U$ and $\cap_{n\in\mathbb{N}} f^n U = \Lambda$,*

- *$f_{|\Lambda}$ has a dense orbit,*

then for every $\varphi \in \mathcal{C}(\overline{U})$, $\frac{1}{n}S_n\varphi$ satisfies a Large Deviations Principle under the Riemannian measure with rate function:

$$K(s) = \inf\left\{-h_f(\nu) + \sum_{\lambda_i > 0} \lambda_i \, : \, \nu \in \mathcal{M}^1_{inv}(\mathcal{X}), \int_{\overline{U}} \varphi \, d\nu = s\right\}$$

As previously, the λ_i denote the Lyapunov exponents associated to the measure ν.

The abstract result of Theorems 2.3.9 and 2.3.10 being stated, the way to establish this result is to find a potential ξ that describes well the behavior of the size under initial measure m of the balls $b_x(n;\varepsilon)$.
This is generally called a Volume Lemma, and can be established in this setup with $\xi = \log|\det(Df|_{E^u})|$ where E^u is the unstable manifold.

2.3.4 For a map satisfying expansiveness and specification

Haydn and Ruelle have introduced in [45, 94] a general Gibbsian formalism adapting what existed for shift systems to dynamical systems satisfying expansiveness and specification. And it appears to be a good setup to obtain a Large Deviations Principle, although it was not clearly written in the litterature till recently, to our knowledge.
This is the reason why we will here present this Gibbsian formalism, as stated in [94], and state the Large Deviations Principle for Gibbs measures in this setup. It contains most results of the previous Section as consequences, since subshifts of finite type, Anosov diffeomorphisms or expanding maps of the circle satisfy expansiveness and specification. It applies moreover to new cases.

We work during the whole section with $(\mathcal{X}, d)$ a compact metric space and a continuous map f on $\mathcal{X}$ which is assumed to satisfy expansiveness with constant δ_0 and specification (see Definitions 2.3.1 and 2.3.2).

For any potential $\varphi \in \mathcal{C}(\mathcal{X})$, we define:

$$\begin{aligned}
b_x(n;\delta) &= \{y \in \mathcal{X} \, : \, d(f^t(x), f^t(y) \leq \delta \quad \forall\, 0 \leq t \leq n\} \\
K_\varphi(\delta, n) &= \sup\left\{\left|\sum_{k=0}^{n-1} \varphi(f^k x) - \varphi(f^k y)\right| \, : \, x, y \text{ s.t. } y \in b_x(n;\delta)\right\} \\
K_\varphi &= \sup_{n \geq 1} K_\varphi(\delta_0, n)
\end{aligned}$$

The set of potentials we will consider is then:

$$\mathcal{V} = \{\varphi \in \mathcal{C}(\mathcal{X}) \, : \, K_\varphi < \infty\}$$

Haydn and Ruelle do not require the continuity of potentials, and this is one of the interests of their approach, but we restrict here to continuous potentials, for simplicity.

We need other definitions to introduce the notion of Gibbs measures associated to a potential $\varphi \in \mathcal{V}$:

Definition 2.3.5. *x and y are* **conjugate** *if* $\lim_{t\to\infty} d(f^t x, f^t y) = 0$.
*They are T-***conjugate** *if* $f^T x = f^T y$.

Expansiveness property implies that:

Proposition 2.3.2. *x and y are conjugate iff there exists T such that they are T-conjugate.*

Definition 2.3.6. *(U, τ) is a* **conjugating homeomorphism** *if U is compact, $\tau : U \mapsto \tau U$ is a homeomorphism and there exists T such that for any $x \in U$ x and τx are T-conjugate.*

Gibbs measures for potential $\varphi \in \mathcal{V}$ are measures for which change by a conjugating homeomorphism can be expressed with this potential. This is a generalization of an unusual definition for shift systems:

Definition 2.3.7.
$\mu \in \mathcal{M}^1(\mathcal{X})$ is a **Gibbs state for** *$\varphi \in \mathcal{V}$ if for any conjugating homeomorphism (U, τ):*

$$\frac{d\tau\mu}{d\mu} = \exp \sum_{k=0}^{\infty} (\varphi \circ f^k \circ \tau^{-1} - \varphi \circ f^k)$$

$\mu \in \mathcal{M}^1(\mathcal{X})$ is a **quasi-Gibbs state for** *$\varphi \in \mathcal{V}$ if there is $C > 0$ such that for any conjugating homeomorphism (U, τ):*

$$\frac{d\tau\mu}{d\mu} \leq C \exp \sum_{k=0}^{\infty} (\varphi \circ f^k \circ \tau^{-1} - \varphi \circ f^k)$$

The main characterization of quasi-Gibbs states is given by the following equivalent conditions:

Theorem 2.3.12. *For a potential $\varphi \in \mathcal{V}$, the following conditions are equivalent:*

1. *μ is a quasi-Gibbs state*

2. *there exists $\tilde{c} > 0$ such that for all $x \in \mathcal{X}$, $T \geq 0$ and $\delta < \delta_0$:*

$$\exp\Big[\sum_{t=0}^{T-1} \varphi(f^t x) - TP_f(\varphi) - \tilde{c}\Big] \leq \mu(b_x(n;\delta)) \\ \leq \exp\Big[\sum_{t=0}^{T-1} \varphi(f^t x) - TP_f(\varphi) + \tilde{c}\Big]$$

3. *If ν is a quasi-Gibbs state, then μ and ν are equivalent, with $d\nu/d\mu \in L^\infty(\mu)$ and $d\mu/d\nu \in L^\infty(\nu)$.*

The second formulation will be crucial for a proof of Large Deviations: it expresses what we called in previous Section a Volume Lemma. More precisely a Volume Lemma states that the studied measure (Lebesgue or Riemann measure) is a quasi-Gibbs state for the associated potential.

Ruelle proves also that under these assumptions of expansiveness and specification, there is for each $\varphi \in \mathcal{V}$ a unique associated Gibbs measure, and a unique invariant quasi-Gibbs measure. He identifies also this last one with the unique equilibrium measure for φ and with quantities linked to the spectrum of the transfer operator.
But we do not need this property of uniqueness of Gibbs state to prove a general Large Deviations Principle. We get indeed under general assumptions:

Theorem 2.3.13. *Let $\varphi \in \mathcal{V}$ and μ_φ an associated quasi-Gibbs measure. Then the sequence of empirical means $R_T(x) = \frac{1}{T}\sum_{t=0}^{T-1} \delta_{f^t(x)}$ satisfies under μ_φ a Large Deviations Principle with rate function:*

$$I(\nu) = \begin{cases} P_f(\varphi) - h_f(\nu) - \int_{\mathcal{X}} \varphi \, d\nu & \text{if } \nu \text{ is } f\text{-invariant} \\ +\infty & \text{otherwise} \end{cases}$$

This result is not new, since the abstract Theorems 2.3.9 and 2.3.10 of Young apply to this general case. Maes and Verbitski remark this in their recent preprint [79].

They add that the result of Kifer (Theorem 2.3.14) applies also. It seems to us that we should add in this case the assumption that the set of potentials $\mathcal{V}$ is dense in $\mathcal{C}(\mathcal{X})$. It means that the method of Young is more efficient since it allows to avoid this additional assumption.

For other developments on maps satisfying expansiveness and specification, in particular multifractal formalism, see [102, 100].

2.3.5 Via spectral gap of the transfer operator

Lalley proposes in [73] another way of obtaining limit theorems for dynamical systems. He states results for subshifts of finite type $\sigma : \Omega^A \to \Omega^A$ (notations are the same as in Section 2.3.2). In this context, one can introduce transfer operators associated to any real valued Hőlder-continuous map f:

$$\mathcal{L}_f g(x) = \sum_{y\,:\,\sigma y = x} e^{f(y)} g(y)$$

Ruelle Perron-Frobenius Theorem states (see proofs in [11]) that under these assumptions, $\mathcal{L}_f$ satisfies a spectral gap property with a simple maximal eigenvalue $\lambda_f = e^{P_\sigma(f)} > 0$, associated eigenfunction $h_f > 0$ and eigenmeasure ν_f for the action of the dual operator. The Gibbs measure already presented in Section 2.3.2 is in fact $\mu_f = h_f \nu_f$.

We may remark that for ψ Hőlder-continuous:

$$\int_{\Omega^A} e^{S_n\psi} g \, d\mu_f = \lambda_f^{-n} \int_{\Omega^A} \mathcal{L}^n_{f+\psi}(g h_f) \, d\nu_f$$

which allows to identify the limit of log-Laplace transforms, with the use of the spectral gap property for $\mathcal{L}_{f+\psi}$ and $\mathcal{L}_f$ and a perturbation argument. We get hence:

$$\Lambda(z) = \lim_{n\to\infty} \frac{1}{n} \log \int_{\Omega^A} e^{zS_n\psi} \, d\nu_f = \log \frac{\lambda_{f+z\psi}}{\lambda_f} = P_\sigma(f + z\psi) - P_\sigma(f)$$

is a smooth function of z.

We can hence apply Gắrtner-Ellis Theorem (Theorem 2.1.4) to conclude that $\frac{1}{n} S_n\psi$ satisfies under μ_f a Large Deviations Principle (of the level 1 type, under the Hőlder-continuous observable ψ) with rate function Λ^*.

Lalley uses in fact a refined Ruelle-Perron-Frobenius Theorem, proven by Pollicott in [87] for complex-valued potentials, to get sharper results: local limit theorems and sharp large deviations estimates.

At the same time, Rousseau-Egele [92] develops a similar strategy for the class of expanding transformations of the unit interval for which the spectrum of the associated transfer operator is described by the Theorem of Ionescu-Tulcea and Marinescu. He gets in those cases also central and local limit theorems.

This has been recently generalized by Anne Broise in [15]. She gets the same kind of results and sharp large deviations estimates.

Another generalization of this method to Anosov flows is due to Waddington in [103].

It has to be noticed that all these results are established at the observable level, and under regular enough observables. For maps of the interval, the perturbation argument allows only the statement of a partial lower bound around the mean value, what makes harder the identification of the rate function.

Large deviations results for subshifts or Anosov flows are also linked with many results concerned with the study of closed geodesics on compact manifolds of negative curvature (see for example [74], or [2, 1] for recent and sharp results).

2.3.6 Variations on Gärtner-Ellis theorem

Kifer develops in [71] a way of going from large deviations for the empirical mean of regular observables to a measure level result.

He establishes a generalization of Gấrtner-Ellis Theorem when the state space $\mathcal{X}$ is compact, showing that the differentiability of the pressure on a dense separable set of potentials is sufficient to get the lower bound.
For (X_n) a stochastic process on $\mathcal{X}$, $V \in \mathcal{C}(\mathcal{X})$ and $\nu \in \mathcal{M}^1_{\text{inv}}(\mathcal{X})$ he defines:

$$\begin{aligned} \Lambda(V) &= \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\Big(e^{\sum_{i=0}^{n-1} V(X_i)}\Big) \\ I(\nu) &= \sup_{V\in\mathcal{C}(\mathcal{X})} \Big(\int_{\mathcal{X}} V\, d\nu - \Lambda(V)\Big) \end{aligned} \tag{2.6}$$

hence by duality:

$$\Lambda(V) = \sup_{\nu\in\mathcal{M}^1_{\text{inv}}(\mathcal{X})} \Big(\int_{\mathcal{X}} V\, d\nu - I(\nu)\Big) \tag{2.7}$$

He gets then:

Theorem 2.3.14.

1. *If the limit defining Λ in (2.6) exists for any $V \in \mathcal{C}(\mathcal{X})$, then $L_n = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{X_i}$ satisfies a LDP Upper Bound with rate function I.*

2. *If furthermore there exists a dense separable set $\mathcal{V} \subset \mathcal{C}(\mathcal{X})$ such that, for any $V \in \mathcal{V}$, there is a unique measure μ_V realizing the supremum in (2.7), then L_n satisfies also the Lower Bound with the same rate function I.*

This result applies then to the same kind of dynamical systems as previously (subshifts under Gibbs measures, expanding transformations, hyperbolic systems), in fact every one-dimensional map where a Volume Lemma and uniqueness of equilibrium measure for enough potentials can be established.

Kifer states this in the case of a non invariant subset Γ of $\mathcal{X}$, getting a slightly more general result where the escape rate from $\mathcal{X}$ occurs: he assumes that $\Gamma \subset \mathcal{X}$ is closed and puts:

$$\begin{aligned} \Gamma_n =& \{x \in \mathcal{X} : f^i x \in \Gamma \quad \forall\, 0 \leq i < n\} \\ R_n(x) =& \frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k x} \qquad \text{for } x \in \Gamma_n \\ b_{\Gamma,x}(n;\delta) =& \left\{y \in \Gamma_n \;:\; d\left(f^i x, f^i y\right) \leq \delta \quad \forall 0 \leq i \leq n\right\} \end{aligned}$$

and $P_{\Gamma,f}(V)$ the topological pressure of V restricted to Γ.
He assumes furthermore that m is a probability measure on $\mathcal{X}$ such that $\operatorname{supp}(m) = \Gamma$ and there exists $\varphi \in \mathcal{C}(\mathcal{X})$ with:

$$(A_\delta(n))^{-1} \leq m\left(U_\delta(x,n,\Gamma)\right) e^{-S_n\varphi(x)} \leq A_\delta(n)$$

for all $t, \delta > 0$ and $x \in \Gamma_n$, and with $\lim \frac{1}{n}\log A_\delta(n) = 0$.

He states that under these assumptions:

Theorem 2.3.15.

1. *If $h_\mu(f)$ is upper semi-continuous for every $\mu \in \mathcal{M}_{inv}(\Gamma)$, then for any closed $K \subset \mathcal{M}(\Gamma)$:*

$$\limsup_{n\to\infty} \frac{1}{n}\log m(x \in \Gamma_n \;:\; R_n(x) \in K) \leq -\inf\{I(\nu) \;:\; \nu \in K\} \leq P_{\Gamma,f}(\varphi)$$

2. *If in addition there exists a dense separable set $\mathcal{V} \subset \mathcal{C}(\mathcal{X})$ such that any $V \in \mathcal{V}$ verifies $P_{\Gamma,f}(\varphi+V) = \int_\Gamma V\, d\mu_V - I(\mu_V)$ for a unique $\mu_V \in \mathcal{M}(\Gamma)$, then for any open $G \subset \mathcal{M}(\Gamma)$:*

$$\liminf_{n\to\infty} \frac{1}{n}\log m(x \in \Gamma_n \;:\; R_n(x) \in G) \geq -\inf\{I(\nu) \;:\; \nu \in G\}$$

with:

$$I(\nu) = \begin{cases} -h_f(\nu) - \int_\Gamma \varphi \, d\nu & \text{if } \nu \in \mathcal{M}_{inv}(\Gamma), \\ \infty & \text{otherwise.} \end{cases}$$

In a later paper [72], Kifer uses the same abstract Theorem 2.3.14 in the general setup of a map satisfying expansiveness and specification with a dense set of regular functions (see Section 2.3.4 for definitions and remarks) to establish a large deviations result about the equidistribution of periodic orbits.

He denotes CO the set of periodic orbits and CO(t) the set of those with some period less than t. For γ a closed orbit, he denotes $\tau(\gamma)$ its smallest period and $\zeta_\gamma = \tau(\gamma)^{-1} \sum_{i=1}^{\tau(\gamma)} \delta_{f^i(x)}$ with x a point of γ.
He defines also ν_t the normalized counting measure on CO(t): for $\Gamma \subset$CO,

$$\nu_t(\Gamma) = \frac{1}{N_t} \mathrm{Card}(\Gamma \cap \mathrm{CO}(t)) \quad \text{where} \quad N_t = \mathrm{Card}(\mathrm{CO}(t))$$

Then the sequence of measures $\zeta_\cdot^\star(\nu_t)$ on $\mathcal{M}^1(\mathcal{X})$ satisfies a Large Deviations Principle with rate function:

$$I(\nu) = \begin{cases} h_{\mathrm{top}}(f) - h_f(\nu) & \text{if } \nu \in \mathcal{M}^1_{\mathrm{inv}}(\mathcal{X}), \\ +\infty & \text{otherwise.} \end{cases}$$

which gives an exponential rate of convergence of $N_t^{-1} \sum_{\gamma \in \mathrm{CO}(t)} \zeta_\gamma$ to the measure of maximal metric entropy (the equilibrium measure for the potential 0).

The results of Lalley [74], Babillot-Ledrappier [2] and Anantharaman [1] can be seen as refinements of this Large Deviations Principle.
A similar result weighted by any Hőlder-continuous potential has been obtained by Pollicott in [89].

2.3.7 Maps with an indifferent fixed point

Few results exist for dynamical systems which are not uniformly expanding or hyperbolic. The only progress has been done recently for maps of the interval with indifferent fixed points.

In [90], Pollicott, Sharp and Yuri study the map $f : [0,1] \to [0,1]$ defined by:

$$f(x) = x + x^{1+s} \mod 1$$

with $0 < s < 1$, called the Manneville-Pomeau map.

They give a new proof of the fact that there exists a finite invariant measure which is absolutely continuous with respect to Lebesgue measure, with unbounded density.
They prove that μ is not the unique equilibrium measure associated to the potential $\varphi = -\log f'$. The Dirac mass at 0 is another one and the set of equilibrium measures is:

$$\mathcal{A} = \{\alpha\mu + (1-\alpha)\delta_0 \,:\, 0 \leq \alpha \leq 1\}$$

They prove also a large deviations upper bound for the family of weighted averages of empirical means:

$$\Delta_n = \Big(\sum_{T^n y = x} \frac{1}{(T^n)'(y)}\Big)^{-1} \Big(\sum_{T^n y = x} \frac{1}{(T^n)'(y)} \big(\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k(y)}\big)\Big)$$

for any reference point x, with rate function $I(\nu) = -h_f(\nu) - \int_{[0,1]} \varphi \, d\nu$ which vanishes on the whole interval $\mathcal{A}$.

This result is in the spirit of the last presented result of Kifer in the previous Section, taking pre-images of a point instead of periodic orbits. It does not tell much about convergence since the rate function vanishes on a complete interval.

3. SPATIOTEMPORAL LARGE DEVIATIONS PRINCIPLE FOR COUPLED CIRCLE MAPS

We consider in this Chapter[1] a spatially invariant coupled map lattice between expanding maps of the circle under a coupling which is weak and with short range, i.e. such that the strength of the coupling between two sites decreases exponentially fast with the distance between the sites.

Our main result is a large deviations principle for the spatiotemporal empirical measure associated to the dynamics, stated in Theorem 3.1.2.
The spatiotemporal approach allows to relate the rate function with the thermodynamic formalism associated to the $(d+1)$ dynamical system of the coupled map and the spatial shifts (see Section 3.8 for a presentation of this theory).

This result is linked to previous papers of Jiang and Pesin [53, 51], where they prove for the same system under a slightly different weak coupling assumption the existence and uniqueness of the equilibrium measure associated to a potential φ which they construct.
This potential governs also our large deviations principle. We improve with this theorem the result of Jiang, giving an exponential rate of convergence to the equilibrium measure.

The main step in our proof is a Volume Lemma, see Theorem 3.1.1.
It makes our proof direct (i.e. without the use of coding by a Gibbs system) and independent of the uniqueness of the equilibrium measure.

We give our precise Assumptions and Results in Section 3.1. In Section 3.2 we recall the derivation of the potential we are interested in, done in [53] and [51]. We explain then in Section 3.3 how our assumptions give what we call the preserved expanding property, the key estimate for our proof. Section 3.4 is then devoted to the proof of the Volume Lemma and Sections 3.5 and 3.6 to the proof of the Large Deviations Principle.

[1] This Chapter corresponds to an article with same title, written with Gérard Ben Arous and submitted for publication.

3.1 Settings and results

3.1.1 The state space

We work on the state space $\mathcal{X} = (S^1)^{\mathbb{Z}^d}$ (with $d \geq 1$), equipped with the reference measure $\overline{m} = m^{\otimes \mathbb{Z}^d}$ where m is the Lebesgue measure on the circle. On the circle $S^1 = \mathbb{R}/\mathbb{Z}$, the distance is $d(x,y) = \min_{k\in\mathbb{Z}} |x+k-y| \leq 1/2$. We put on $\mathcal{X}$ two distances constructed from this one:

- $d(x,y) = \sup_{i\in\mathbb{Z}^d} d(x_i,y_i)$ which makes $\mathcal{X}$ an infinite dimensional manifold;
- $d_\rho(x,y) = \sup_{i\in\mathbb{Z}^d} \rho^{|i|} d(x_i,y_i)$ where we take for $i \in \mathbb{Z}^d$ the norm $|i| = \max_{1\leq k\leq d} |i_k|$ and $\rho < 1$ is a fixed parameter. The main interest of d_ρ is that $(\mathcal{X}, d_\rho)$ is a compact space, hence we can use the thermodynamic formalism to describe the system.

We denote by S^k the spatial shift of vector $k \in \mathbb{Z}^d$ on $\mathcal{X}$: if $x = (x_i)_{i\in\mathbb{Z}^d}$ then $(S^k x)_i = x_{i+k}$. For $N \in \mathbb{N}$, we write $\Lambda_N = [-N,N]^d \subset \mathbb{Z}^d$.

3.1.2 The coupled map

Let the uncoupled expanding map be $F_0 = \otimes_{i\in\mathbb{Z}^d} f_i$ where $f_i = f : S^1 \to S^1$ is $\mathcal{C}^{1+\alpha}$ and expanding, i.e. satisfies:

$$1 < \gamma \leq |f'(x)| \leq M \qquad \forall x \in S^1 \tag{3.1}$$

and f' hence $\log|f'|$ is α-Hőlder continuous:

$$|\log|f'(x)| - \log|f'(y)|| \leq C_1 d^\alpha(x,y) \qquad \forall x,y \in S^1 \tag{3.2}$$

We define also the coupling map $G : \mathcal{X} \to \mathcal{X}$ as a $\mathcal{C}^2$ map (for the distance d) commuting with all the spatial translations $(S^k)_{k\in\mathbb{Z}^d}$ and which satisfies the following estimates:

$$\left|\frac{\partial G_i}{\partial x_j} - \delta_{i,j}\right| \leq \mathcal{E}\,\theta^{2|i-j|} \qquad \forall i,j \in \mathbb{Z}^d \tag{3.3}$$

$$\left|\frac{\partial^2 G_i}{\partial x_j \partial x_k}\right| \leq \mathcal{E}\,\theta^{2\max(|i-j|,|i-k|)} \qquad \forall i,j,k \in \mathbb{Z}^d \tag{3.4}$$

with $\mathcal{E} > 0$ and $0 < \theta < 1$.
We denote $\mathcal{K} = \mathcal{E}\sum_{i\in\mathbb{Z}^d} \theta^{|i|}$ and $\mathcal{K}_2 = \mathcal{E}\sum_{i\in\mathbb{Z}^d} \theta^{2|i|}$.

The first derived estimates are:

$$d_i(G(x)-x, G(y)-y) \leq \mathcal{E} \sum_{k\in\mathbb{Z}^d} \theta^{2|i-k|} d_k(x,y) \qquad \forall i \in \mathbb{Z}^d,\ x,y \in \mathcal{X} \tag{3.5}$$

$$\left|\frac{\partial G_i}{\partial x_j}(x) - \frac{\partial G_i}{\partial x_j}(y)\right| \leq \mathcal{E} \sum_{k\in\mathbb{Z}^d} \theta^{2|i-k|} d_k(x,y) \quad \forall i,j \in \mathbb{Z}^d,\ x,y \in \mathcal{X} \tag{3.6}$$

The associated coupled map is then:

$$F = G \circ F_0$$

We say that F satisfies Assumption $(\mathcal{H})$ if it is the composition of two such maps whose parameters satisfy the two conditions:

$$\begin{cases} \theta < \rho & (H1) \\ \gamma - M\mathcal{K} > 1 & (H2) \end{cases} \tag{3.7}$$

The first assumption is essentially technical, to get functions regular enough for the distance d_ρ. It implies in particular with (3.5) that G is Lipschitz continuous for d_ρ, hence is α-Hőlder continuous.
$(H2)$ expresses exactly the preservation of the expanding property for the coupled map and implies two essential estimates:

$$\tilde{\gamma} = \gamma - M\mathcal{K}_2 > 1 \tag{3.8}$$

$$\mathcal{K} < 1 \tag{3.9}$$

Remark: These conditions for coupling are similar to those given in previous papers on this type of system (they are called short range maps in [53] or [51]).

3.1.3 Volume Lemma

We define for $T \in \mathbb{N}$ and E a finite subset of $\mathbb{Z}^d$

$$B_x(T,E;\delta) = \left\{y \,:\, d_\rho\left(S^i \circ F^t(x), S^i \circ F^t(y)\right) < \delta \quad \forall\, 0 \leq t \leq T,\, i \in E\right\} \tag{3.10}$$

the ball associated to a distance which describes the dynamics of F and the spatial shifts S. It contains the points whose orbit stays near a given orbit under fixed space and time translations. The Volume Lemma describes the measure of this ball in terms of local derivatives along the orbit of x:

Theorem 3.1.1. *If F satisfies Assumption $(\mathcal{H})$, then there exists a potential function $\varphi : \mathcal{X} \mapsto \mathbb{R}$ Hőlder continuous for the distance d_ρ, such that for any $x \in \mathcal{X}$, $0 < \delta < \frac{1}{2M}$, E a finite subset of $\mathbb{Z}^d$ and $T \geq 1$, we have:*

$$C_2(T,E,\delta,\rho)\exp\Big(\sum_{\substack{0\le t<T\\ i\in E}}\varphi\circ S^i\circ F^t(x)\Big) \le \overline{m}(B_x(T,E;\delta))$$
$$\le C_3(T,E,\delta)\exp\Big(\sum_{\substack{0\le t<T\\ i\in E}}\varphi\circ S^i\circ F^t(x)\Big) \quad (3.11)$$

with:

$$\lim_{\substack{T\to\infty\\ n\to\infty}}\frac{1}{T|E_n|}\log C_2(T,E_n,\delta,\rho) = \lim_{\substack{T\to\infty\\ n\to\infty}}\frac{1}{T|E_n|}\log C_3(T,E_n,\delta) = 0 \qquad (3.12)$$

for all $\delta < \frac{1}{2M}$, $\theta < \rho < 1$ and for any sequence E_n converging to $\mathbb{Z}^d$ in the sense of Van Hove (see Definition 3.7.1).

Remarks: 1. The potential function φ is defined in 3.26 in Section 3.2.2 following readily the construction given in [51] and [53]. From this Definition and the role it plays in the Volume Lemma (see for example [70] for an equivalent result in the case of a single map), φ can be called the "logarithm of Jacobian per site" of the map F.

2. The speeds of convergence in time and space are completely independent. We can even take one limit before the other, if we understand then the first limit as limsup and liminf.

3. This result is in fact true not only under Lebesgue measure but also for any probability measure μ which is locally absolutely continuous with respect to it, with a Radon-Nikodym derivative satisfying with $0 < A < B$:

$$A^{|E|} \le \left.\frac{d\mu}{d\overline{m}}\right|_E \le B^{|E|} \qquad \forall E \subset \mathbb{Z}^d$$

A direct consequence of this result, or of Proposition 3.5.1, concerns the topological pressure (see Subsection 3.8.3 for a definition) of the potential φ:

Corollary 3.1.1. *If F satisfies Assumption $(\mathcal{H})$, the topological pressure of the potential φ under the dynamical system (F,S) is null:*

$$P_{(F,S)}(\varphi) = 0$$

This was the main result of [50] in the context of Anosov maps and by the use of coding. In our context, it is an important result because it ensures with the Gibbs Variational Principle 3.8.4 that the rate function I (defined in (3.14) below) is non negative.

3.1.4 Large Deviations Principle

We can use the previous Volume Lemma to prove a spatio-temporal Large Deviations Principle for the empirical process

$$R_{T,E}(x) = \frac{1}{T|E|} \sum_{\substack{0 \le t < T \\ i \in E}} \delta_{S^i \circ F^t(x)} \in \mathcal{M}^1(\mathcal{X}) \tag{3.13}$$

under the initial measure $\overline{m}$ (and, more generally, under the same probability measures as for Volume Lemma, see Remark 3 after Theorem 3.1.1).

We introduce the function I defined on $\mathcal{M}^1(\mathcal{X})$ by:

$$I(\nu) = \begin{cases} -h_{(F,S)}(\nu) - \int_{\mathcal{X}} \varphi \, d\nu & \text{if } \nu \in \mathcal{M}^1_{\text{inv}}(\mathcal{X}) \\ +\infty & \text{otherwise} \end{cases} \tag{3.14}$$

(see Section 3.8 for the definitions and properties of $\mathcal{M}^1_{\text{inv}}(\mathcal{X})$ and the metric entropy $h_{(F,S)}$).
We have then:

Theorem 3.1.2. *Assume F satisfies Assumption $(\mathcal{H})$. Then I is a non negative, convex and lower semi-continuous function.*
For any map $s : \mathbb{N} \to \mathbb{N}$ non decreasing and such that $s(T)$ tends to infinity as T tends to infinity, the sequence $(R_{T,\Lambda_{s(T)}})^*(\overline{m})$ satisfies a Large Deviations Principle with rate function *I, i.e.:*

1. *For any K closed subset of $\mathcal{M}^1(\mathcal{X})$, we have:*

$$\limsup_{T\to\infty} \frac{1}{T|\Lambda_{s(T)}|} \log \overline{m}\{x \,:\, R_{T,\Lambda_{s(T)}}(x) \in K\} \le -\inf_{\nu \in K} I(\nu) \tag{UB}$$

2. *For any O open subset of $\mathcal{M}^1(\mathcal{X})$, we have:*

$$\liminf_{T\to\infty} \frac{1}{T|\Lambda_{s(T)}|} \log \overline{m}\{x \,:\, R_{T,\Lambda_{s(T)}}(x) \in O\} \ge -\inf_{\nu \in O} I(\nu) \tag{LB}$$

Remarks: 1. This result remains in fact true for more general sequences of sets: the upper bound is valid for any spatial sequence E_T converging to $\mathbb{Z}^d$ in the sense of Van Hove, the lower bound for any special averaging sequence (see Definition 3.7.2). Proofs are given in Sections 3.5 and 3.6 in this general setup.

2. Whatever, the main fact is that time and space averagings must tend together to infinity but at completely independent speeds.

3. The independence of speeds of convergence in time and space is not surprising because we know that for weak coupling there is a semi-conjugacy between (F, S) and shifts of a $(d+1)$ dimensional Gibbs system (see Theorem 2 in [51]). The time direction becomes then a spatial shift like others on the coding space.
This semi-conjugacy allows in fact to deduce a Large Deviations Principle for R_{T,E_T} from the same result for Gibbs systems (see [42], [81], [31] or [39]) by a contraction principle (Theorem 4.2.1 of [32]). We could not identify the rate function obtained in this way, hence preferred to develop a direct proof, without coding. It has however to be noticed that our analysis of inverse branches in Subsection 3.3.2 is not far from the construction of a Markov partition for the system.

4. Furthermore, the direct proof of large deviations result we give here leads to new questions: if we manage to avoid the restriction on the coupling due to the semi-conjugacy, we need to preserve in fact the expanding property. Could we hope such a Large Deviations Principle in a more general setup, with stronger coupling?

5. It is proved in [51] that for small enough coupling, there is a unique minimizing measure for I (called an equilibrium measure for the potential φ). But we do not know for which coupling a phase transition case (i.e. a case where there are at least two equilibrium measures) could occur. To analyze such a situation, it would be necessary to describe the equilibrium measures as Gibbs measures. As far as we know, such a characterization does not exist in this context.

3.2 Expansion of the derivative

In this section, we follow [53] to derive the potential φ by a sharp analysis of the derivative of the map F restricted to finite boxes. We give all the steps, referring the reader to Section 5 of [53] for the detailed computations.

3.2.1 Finite box maps

For Λ a finite subset of $\mathbb{Z}^d$ and $\eta \in \mathcal{X}$ a fixed boundary condition, we define

$$\begin{aligned} F_{\Lambda,\eta} \,:\, \mathcal{X}_\Lambda = \left(S^1\right)^\Lambda &\longrightarrow \mathcal{X}_\Lambda \\ x_\Lambda &\longmapsto F(x_\Lambda \vee \eta_{\Lambda^C})|_\Lambda \end{aligned}$$

with $w = x_\Lambda \vee \eta_{\Lambda^C}$ defined by $w_i = x_i$ when $i \in \Lambda$ and $w_i = \eta_i$ otherwise. In fact $F_{\Lambda,\eta} = G_{\Lambda,F_0(\eta)} \circ F_0$ with $G_{\Lambda,\eta} = G(x_\Lambda \vee \eta_{\Lambda^C})$.

$G_{\Lambda,\eta}$ is a $\mathcal{C}^2$ map and if we write $DG_{\Lambda,\eta} = Id_\Lambda + A_{\Lambda,\eta}$ with $A_{\Lambda,\eta} = \left(a_{i,j}\right)_{i,j\in\Lambda}$, we get from estimates (3.3) and (3.6) the following estimates for any $i, j \in \Lambda$, $x_\Lambda, y_\Lambda \in \mathcal{X}_\Lambda$:

$$|a_{i,j}(x_\Lambda)| \leq \mathcal{E}\theta^{2|i-j|} \tag{3.15}$$

$$|a_{i,j}(x_\Lambda) - a_{i,j}(y_\Lambda)| \leq \mathcal{E}\sum_{k\in\Lambda}\theta^{2|i-k|}d_k(x_\Lambda, y_\Lambda) \tag{3.16}$$

$$|a^{(\eta)}_{i,j}(x_\Lambda) - a^{(\eta')}_{i,j}(x_\Lambda)| \leq \frac{\mathcal{K}}{2}\,\theta^{d(i,\Lambda^C)} \tag{3.17}$$

$$|a^{(\Lambda)}_{i,j}(x_\Lambda) - a^{(\Lambda')}_{i,j}(y_{\Lambda'})| \leq \frac{\mathcal{K}}{2}\,\theta^{d(i,\Lambda'\setminus\Lambda)} \tag{3.18}$$

if $\Lambda \subset \Lambda'$ and $y_{\Lambda'}|_\Lambda = x_\Lambda$.

3.2.2 Expansion

Then, using (3.9):

$$\|A\|_\infty \leq \max_{i\in\Lambda}\left(\mathcal{E}\sum_{j\in\Lambda}\theta^{2|i-j|}\right) \leq \mathcal{K}_2 \leq \mathcal{K} < 1$$

hence $\log(Id + A)$ exists and we can write:

$$\begin{aligned}\log|\det DF_{\Lambda,\eta}(x_\Lambda)| &= \log|\det DF_0(x_\Lambda)\det DG_{\Lambda,F_0(\eta)}(F_0(x_\Lambda))| \\ &= \sum_{i\in\Lambda}\log|f'(x_i)| + \log|\det(\exp\log(Id+A)(F_0(x_\Lambda)))| \\ &= \sum_{i\in\Lambda}\log|f'(x_i)| + \log\exp(\operatorname{tr}\log(Id+A)(F_0(x_\Lambda))) \\ &= \sum_{i\in\Lambda}\log|f'(x_i)| + \operatorname{tr}\left(-\sum_{t\geq 1}\frac{(-1)^t}{t}A^t(F_0(x_\Lambda))\right) \\ &= \sum_{i\in\Lambda}(\log|f'(x_i)| - w_{\Lambda,\eta,i}(x_\Lambda))\end{aligned}$$

where $w_{\Lambda,\eta,i}(x_\Lambda) = \sum_{t\geq 1}\frac{(-1)^t}{t}a^{(t)}_{i,i}(F_0(x_\Lambda))$, denoting $A^t = \left(a^{(t)}_{i,j}\right)$.

Estimates (3.15) to (3.18) give analogous results for w under the same condition (3.9):

$$|w_{\Lambda,\eta,i}(x_\Lambda)| \leq \frac{\mathcal{E}}{1-\mathcal{K}} \tag{3.19}$$

$$|w_{\Lambda,\eta,i}(x_\Lambda) - w_{\Lambda,\eta,i}(y_\Lambda)| \leq \frac{M\mathcal{E}}{1-\mathcal{K}}\sum_{k\in\Lambda}\theta^{|i-k|}d_k(x_\Lambda, y_\Lambda) \tag{3.20}$$

$$|w_{\Lambda,\eta,i}(x_\Lambda) - w_{\Lambda,\eta',i}(x_\Lambda)| \leq \frac{1}{2(1-\mathcal{K})}\,\theta^{d(i,\Lambda^C)} \tag{3.21}$$

$$|w_{\Lambda,\eta,i}(x_\Lambda) - w_{\Lambda',\eta,i}(y_{\Lambda'})| \leq \frac{1}{2(1-\mathcal{K})}\,\theta^{d(i,\Lambda'\setminus\Lambda)} \tag{3.22}$$

if $\Lambda\subset\Lambda'$ and $y_{\Lambda'}|_\Lambda = x_\Lambda$.

Proof. To get (3.19), we start proving recursively on t that:

$$|a^{(t)}_{i,j}| \leq \mathcal{E}\mathcal{K}^{t-1}\theta^{|i-j|}$$

This is indeed true for $t=1$, as stated in (3.15). And if this is true for t, then:

$$\begin{aligned}|a^{(t+1)}_{i,j}| \leq \sum_{l\in\mathbb{Z}^d}|a^{(t)}_{i,l}||a_{l,j}| &\leq \sum_{l\in\mathbb{Z}^d}\mathcal{E}^2\mathcal{K}^{t-1}\theta^{|i-l|}\theta^{2|l-j|} \\ &\leq \mathcal{E}^2\mathcal{K}^{t-1}\theta^{|i-j|}\sum_{l\in\mathbb{Z}^d}\theta^{|l|} \leq \mathcal{E}\mathcal{K}^t\theta^{|i-j|}\end{aligned}$$

because $\theta^{|i-l|}\theta^{|l-j|} \leq \theta^{|i-j|}$ for all i, j, l.

From this and the definition of $w_{\Lambda,\eta,i}$, we get:

$$w_{\Lambda,\eta,i} \leq \sum_{t\geq 1} \frac{1}{t}|a_{i,i}^{(t)}| \leq \mathcal{E} \sum_{t\geq 1} \frac{\mathcal{K}^{t-1}}{t} \leq \frac{\mathcal{E}}{1-\mathcal{K}}$$

For (3.20), we proceed identically, proving that for all t, if $x_l = y_l$ for any $l \neq k$:

$$|a_{i,j}^{(t)}(x_\Lambda) - a_{i,j}^{(t)}(y_\Lambda)| \leq t\mathcal{E}\mathcal{K}^{t-1}\theta^{|i-k|}d_k(x_\Lambda, y_\Lambda)$$

(3.19) states it for $t = 1$ and if this is true for t, then, using also estimate (3.19):

$$\begin{aligned}
&|a_{i,j}{}^{(t+1)}(x_\Lambda) - a_{i,j}^{(t+1)}(y_\Lambda)| \\
&\quad\leq \sum_{l\in\Lambda} |a_{i,l}^{(t)}(x_\Lambda)||a_{l,j}(x_\Lambda) - a_{l,j}(y_\Lambda)| + |a_{l,j}(x_\Lambda)||a_{i,l}^{(t)}(x_\Lambda) - a_{i,l}^{(t)}(y_\Lambda)| \\
&\quad\leq \sum_{l\in\Lambda} \left(\mathcal{E}\mathcal{K}^{t-1}\theta^{|i-l|}\mathcal{E}\theta^{2|l-k|}d_k(x_\Lambda, y_\Lambda) + \mathcal{E}\theta^{2|l-j|}t\mathcal{E}\mathcal{K}^{t-1}\theta^{|i-k|}d_k(x_\lambda, y_\Lambda)\right) \\
&\quad\leq \mathcal{E}^2\mathcal{K}^{t-1}\theta^{|i-k|}d_k(x_\Lambda, y_\Lambda)\sum_{l\in\Lambda}(\theta^{|l-k|} + t\theta^{|l-j|}) \leq (t+1)\mathcal{E}\mathcal{K}^t\theta^{|i-k|}d_k(x_\Lambda, y_\Lambda)
\end{aligned}$$

This gives the desired estimate for $w_{\Lambda,\eta,i}$.

We proceed in the same way for the following estimates (3.21) and (3.22), with the intermediate results:

$$|a_{i,j}^{(\eta,t)}(x_\Lambda) - a_{i,j}^{(\eta',t)}(x_\Lambda)| \leq \frac{t\mathcal{K}^t}{2}\theta^{d(i,\Lambda^C)}$$

$$|a_{i,j}^{(\Lambda,t)}(x_\Lambda) - a_{i,j}^{(\Lambda',t)}(y_{\Lambda'})| \leq \frac{t\mathcal{K}^t}{2}\theta^{d(i,\Lambda'\setminus\Lambda)}$$

□

All these estimates imply that $\psi_i(x) = \lim_{N\to\infty} w_{\Lambda_N,\eta,i}(x|_{\Lambda_N})$ exists, is independent of the boundary conditions, shift invariant (i.e. $\psi_i = \psi_0 \circ S^i$ for all $i \in \mathbb{Z}^d$) and satisfies:

$$|\psi_0(x)| \leq \frac{\mathcal{E}}{1-\mathcal{K}} \tag{3.23}$$

$$|\psi_0(x) - \psi_0(y)| \leq \frac{M\mathcal{E}}{1-\mathcal{K}}\sum_{k\in\mathbb{Z}^d}\theta^{|i-k|}d_k(x,y) \tag{3.24}$$

$$|\psi_0(x) - w_{\Lambda,\eta,0}(x|_\Lambda)| \leq \frac{1}{2(1-\mathcal{K})}\,\theta^{d(i,\Lambda^C)} \tag{3.25}$$

We deduce from (3.24) that assumption $(H1)$ implies moreover that ψ_0 is Lipschitz continuous for the distance d_ρ.

We define hence

$$\varphi(x) = -\log|f'(x_0)| + \psi_0 \tag{3.26}$$

as the potential of interest to describe the dynamic of the system (F, S). φ is well-defined and α-Hőlder continuous for the distance d_ρ.

3.3 Conservation of the expanding property

We introduce $\emptyset \neq E \subset \Lambda$ two finite subsets of $\mathbb{Z}^d$, a time $T \in \mathbb{N}$ and $x \in \mathcal{X}$ a reference point.
We choose a finite box restriction of F^T to Λ, F^T_Λ with boundary conditions changing with time: $F^t_\Lambda = F_{\Lambda,F^{t-1}(x)} \circ \cdots \circ F_{\Lambda,F(x)} \circ F_{\Lambda,x}$. It implies in particular that:

$$F^t_\Lambda(x|_\Lambda) = F^t(x)\big|_\Lambda \qquad \forall\, 0 \leq t \leq T \tag{3.27}$$

This will essentially simplify the step from F_Λ to F in the proof of the Volume Lemma. We don't mention explicitly the dependence on the boundary conditions following the orbit of x: we have already seen in the previous section that the limit potential doesn't depend on it.

3.3.1 Bijectivity of the coupling map

First of all, our assumptions on the coupling map G are sufficient to get:

Proposition 3.3.1. *Under assumption $(H2)$, G_Λ is a $\mathcal{C}^1$ diffeomorphism.*

Proof. We get from estimate (3.5) and the triangle inequality that

$$d_i(G_\Lambda(x), G_\Lambda(y)) \geq d_i(x,y) - \mathcal{E} \sum_{k\in\Lambda} \theta^{2|i-k|} d_k(x,y) \qquad \forall i \in \Lambda$$

hence if $x \neq y$, let i_0 be such that $d_{i_0}(x,y) = \max_{i\in\Lambda} d_i(x,y) > 0$. Then:

$$d_{i_0}(G_\Lambda(x), G_\Lambda(y)) \geq d_{i_0}(x,y) \left(1 - \mathcal{E} \sum_{k\in\Lambda} \theta^{2|i-k|}\right) \geq (1-\mathcal{K}_2) d_{i_0}(x,y) > 0$$

because $\mathcal{K}_2 \leq \mathcal{K} < 1$ by (3.9). This proves that G_Λ is one-to-one.

We have already noticed that $\|A\|_\infty < 1$, which gives that DG_Λ is invertible, hence that G_Λ is everywhere a local diffeomorphism. The range of G_Λ is then open, and closed by compactness of $\mathcal{X}_\Lambda$, hence its range is the whole space $\mathcal{X}_\Lambda$ because it is connected.

G_Λ is then a bijection and a local diffeomorphism, then a diffeomorphism. $\square$

Remark: G is also a bijection (one-to-one in the same way, onto taking limit of pre-images on finite boxes). This was a specific assumption in most previous papers.

3.3.2 Inverse branches of F_Λ^T

The single site map $f : S^1 \to S^1$ is of degree $p = \int_{S^1} |f'(x)|\, dx$, an integer between γ and M, and has then locally p inverse branches around each point. We can in fact construct them globally except in one point (see Section 2.4 of [64]).

We will use this to construct inverse branches for F_0 around the orbit of x. Associated to the fact that G is a diffeomorphism, it will give us inverse branches for F_Λ^T.

We denote $\mathcal{C}[\Lambda] = \{0, \dots, p-1\}^\Lambda$ to enumerate the inverse branches of F_0. At each time $0 \le t < T$, we construct them around $F^t(x)$. We take

$$A_t = \{y \in \mathcal{X}_\Lambda \,:\, d_i(y, F_0 \circ F^t(x)) < 1/2 \quad \forall\, i \in \Lambda\}$$

(then $m^\Lambda(A_t) = 1$) and for any site $i \in \Lambda$ we denote $x_0^{(t,i)}$, $x_1^{(t,i)}$, …, $x_{p-1}^{(t,i)}$ (resp. $a_0^{(t,i)}$, $a_1^{(t,i)}$, …, $a_{p-1}^{(t,i)}$) the pre images by f of $(F_0 \circ F^t(x))_i$ (resp. $(F_0 \circ F^t(x))_i - 1/2$), indexed such that:

- $x_0^{(t,i)} = F_i^t(x)$
- $x_0^{(t,i)} < a_1^{(t,i)} < x_1^{(t,i)} < \cdots < a_0^{(t,i)} < x_0^{(t,i)}$

Then, for all $\beta \in \mathcal{C}[\Lambda]$, we define:

$$x_\beta^{(t)} = \left(x_{\beta(i)}^{(t,i)}\right)_{i\in\Lambda} \qquad \text{the pre images by } F_0 \text{ of } F_0 \circ F^t(x)$$
$$A_{\beta,t} = \prod_{i\in\Lambda} \left(a_{\beta_i}^{(t,i)}, a_{\beta(i)+1}^{(t,i)}\right)$$

satisfying the following straightforward properties:

- $x_0^{(t)} = F^t(x)$
- $x_\beta^{(t)} \in A_{\beta,t} \quad \forall\, \beta \in \mathcal{C}[\Lambda]$
- $m^\Lambda\Big(\bigcup_{\beta\in\mathcal{C}[\Lambda]} A_{\beta,t}\Big) = 1$
- F_0 is a bijection from $A_{\beta,t}$ onto A_t

We denote $F_{0,t,\beta}^{-1}$ its inverse characterized by $F_{0,t,\beta}^{-1}(y) = A_{\beta,t} \cap F_0^{-1}(y)$ for any $y \in A_t$. These inverse branches satisfy a contraction property, which has to be precisely described:

Lemma 3.3.1. *For all $y, z \in A_t$, there exists $\varphi_{y,z}$ permutation of $\mathcal{C}[\Lambda]$, with $y, z \mapsto \varphi_{y,z}$ measurable, such that $\forall\, \beta, \tilde\beta \in \mathcal{C}[\Lambda]$,$\forall\, i \in \Lambda$, if $\beta(i) = \tilde\beta(i)$, then:*

$$\frac{1}{M} d_i(y,z) \le d_i\left(F_{0,t,\tilde\beta}^{-1}(y), F_{0,t,\varphi_{y,z}(\beta)}^{-1}(z)\right) \le \frac{1}{\gamma} d_i(y,z) \tag{3.28}$$

If y or z equals $F_0 \circ F^t(x)$, then $\varphi_{y,z} = Id$.

Proof. The left inequality is obvious, because $d_i(F_0(\tilde{y}), F_0(\tilde{z})) \leq M d_i(\tilde{y}, \tilde{z})$ is always true.
For the contraction rate, we have to be careful because the partition is adapted to $F^t(x)$ but not to all other points. What has to be understood is how $d_i(y, z)$ is realized at each site $i \in \Lambda$:

- if the shortest arc from y_i to z_i (defining the distance) does not contain $(F_0 \circ F^t(x))_i - 1/2$ (case (i) of Figure 3.1), then $\varphi_{y,z}(\beta)(i) = \beta(i)$;
- otherwise, $\varphi_{y,z}(\beta)(i) = \beta(i) \pm 1$, depending on the order of the three points y, z and $(F_0 \circ F^t(x))_i - 1/2$ (cases (ii) and (iii) of the Figure)) but not on β.

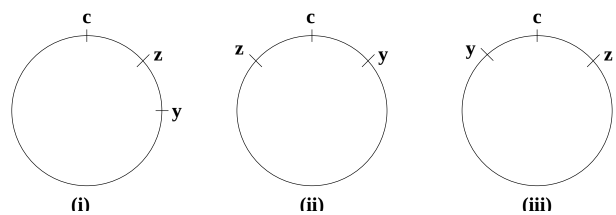

Fig. 3.1: The three cases, where $c = F_0 \circ F^t(x) - 1/2$. If f preserves the direction on the circle (i.e. $f' > 0$), (ii) corresponds to $\varphi_{y,z}(\beta)(i) = \beta(i) + 1$, (iii) to $\varphi_{y,z}(\beta)(i) = \beta(i) - 1$, and this is reversed otherwise.

This defines $\varphi_{y,z}$ as a one-to-one map, and if we are interested in site i, the inverse maps β and $\tilde{\beta}$ are indistinguishable, hence:

$$d_i\left(F^{-1}_{0,t,\tilde{\beta}}(y), F^{-1}_{0,t,\varphi_{y,z}(\beta)}(z)\right) = d_i\left(F^{-1}_{0,t,\beta}(y), F^{-1}_{0,t,\varphi_{y,z}(\beta)}(z)\right) \leq \frac{1}{\gamma} d_i(y, z)$$

If y or z is equal to $F_0 \circ F^t(x)$, we always are in the first case.
It is not hard to check that $\varphi_{y,z}$ depends on y and z only through the distance and the order of their coordinates in the open sets $S^1 \setminus \{(F_0 \circ F^t(x))_i - 1/2\}$, which are measurable maps of y and z. $\square$

We have also, from the left inequality of (3.28), that:

$$\left\{ y \,:\, d_i(F^t(x), y) < \frac{1}{2M} \right\} \subset \bigcup_{\substack{\beta \in \mathcal{C}[\Lambda] \\ \beta(i)=0}} A_{\beta,t} \tag{3.29}$$

We can then describe the inverse branches of F^T_Λ, with:

$$\mathcal{C}[T, \Lambda] = \{0, \ldots, p-1\}^{[1,\ldots,T] \times \Lambda}$$
$$\mathcal{C}[T, \Lambda, E] = \{\alpha \in \mathcal{C}[T, \Lambda] \,:\, \alpha_{t,i} = 0 \quad \forall\, 1 \leq t \leq T,\, i \in E\}$$

Then:

Proposition 3.3.2. *We associate in a unique way to each $\alpha \in \mathcal{C}[T,\Lambda]$ an open subset $\mathcal{A}_\alpha(x)$ of $\mathcal{X}_\Lambda$ such that:*

- $\mathcal{A}_\alpha(x) \cap \mathcal{A}_{\alpha'}(x) = \emptyset$ *if* $\alpha \neq \alpha'$;
- $m^\Lambda(\cup \mathcal{A}_\alpha(x)) = 1$;
- *There exists* $\mathcal{A} \subset \mathcal{X}_\Lambda$ *with* $m^\Lambda(\mathcal{A}) = 1$ *such that for all* $\alpha \in \mathcal{C}[T,\Lambda]$, F_Λ^T *is one-to-one from* $\mathcal{A}_\alpha(x)$ *onto* $\mathcal{A}$. *We denote* $F_{\Lambda,\alpha}^{-T}$ *its inverse.*

Moreover:

$$\left\{ y \in \mathcal{X}_\Lambda \,:\, d_i(F^t(x), F_\Lambda^t(y)) < \frac{1}{2M} \quad \forall\, 0 \le t < T,\, i \in E \right\} \subset \bigcup_{\alpha \in \mathcal{C}[T,\Lambda,E]} \mathcal{A}_\alpha(x) \tag{3.30}$$

Proof. We define:

$$\mathcal{A} = \bigcap_{t=0}^{T-1} F^{T-1-t} \circ G(A_t)$$

to avoid any problem of definition ($m^\Lambda(\mathcal{A}) = 1$ by preservation of total measure by F_0 and G, and by finite intersection) and

$$F_{\Lambda,\alpha}^{-T} = F_{0,0,\alpha(0,\cdot)}^{-1} \circ G^{-1} \circ F_{0,1,\alpha(1,\cdot)}^{-1} \circ G^{-1} \circ \cdots \circ F_{0,T-1,\alpha(T-1,\cdot)}^{-1} \circ G^{-1}$$

which is well defined on $\mathcal{A}$. All properties are then easily deduced from those of $F_{0,i,\beta}^{-1}$'s with:

$$\begin{aligned} \mathcal{A}_\alpha(x) &= F_{\Lambda,\alpha}^{-T}(\mathcal{A}) \\ &= \bigcap_{t=0}^{T-1} F^{-t}\left(A_{t,\alpha(t,\cdot)}\right) \bigcap F^{-T}(\mathcal{A}) \end{aligned}$$

□

Remark: 1. $\mathcal{A}_\alpha(x)$ can be really complicated sets, due to the perturbation term G and the non compatibility of inverse branches. But we avoid problems using the contraction property as described in Lemma 3.3.1.
2. In fact, this construction (except the inclusion 3.30) requires only the local Markov structure of expanding maps and the bijectivity of the coupling.

Notation: In the following, when $\alpha \in \mathcal{C}[T,\Lambda]$ and $0 < t < T$, the notation $F_{\Lambda,\alpha}^{-t}$ denotes in fact $F_\Lambda^{T-t} \circ F_{\Lambda,\alpha}^{-T}$, so that:

$$F_{\Lambda,\alpha}^{-t} = F_{0,T-t,\alpha(T-t,\cdot)}^{-1} \circ G^{-1} \circ F_{\Lambda,\alpha}^{-t+1} \tag{3.31}$$

3.3.3 Expanding property

We can then use the weak coupling assumptions and the inverse branch analysis of F_Λ to get a sharp form of the preservation of the expanding property when we replace F_0 by F_Λ:

Proposition 3.3.3. *Suppose that F satisfies Assumption $(H2)$, $y \in \mathcal{A}$ satisfies $d_i(F^T(x), y) \leq \delta$ for any $i \in E \subset \Lambda$, and that $\alpha \in \mathcal{C}[T, \Lambda, E]$. Then:*

$$d_i\left(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)\right) \leq \frac{\delta}{\tilde{\gamma}^t} + \lambda \cdot \theta^{d(i,E^C)} \qquad \forall\, 0 \leq t \leq T,\, i \in E \tag{3.32}$$

where $\lambda = \frac{M\mathcal{K}}{2(\gamma - M\mathcal{K} - 1)}$ and θ, M, $\mathcal{K}$ and $\tilde{\gamma} = \gamma - M\mathcal{K}_2$ are defined Section 3.1.2.

Remark: This Proposition gives a complete decoupling of temporal expanding property and spatial weak coupling, uniformly in time and space.

Proof. We know that G_Λ is invertible, and by the estimate (3.5) on the coupling and the triangle inequality, we have for $y, z \in \mathcal{X}_\Lambda$ and $i \in \Lambda$:

$$d_i(y, z) \leq d_i(G_\Lambda(y), G_\Lambda(z)) + \mathcal{E} \sum_{k \in \Lambda} \theta^{2|i-k|} d_k(x, y)$$

then for each $1 \leq t \leq T$ and $i \in \Lambda$:

$$\begin{aligned} d_i(G_\Lambda^{-1} \circ F^{T-t+1}(x), G_\Lambda^{-1} \circ F_{\Lambda,\alpha}^{-t+1}(y)) &\leq d_i(F^{T-t+1}(x), F_{\Lambda,\alpha}^{-t+1}(y)) \\ &+ \mathcal{E} \sum_{k \in \Lambda} \theta^{2|i-k|} d_k(G_\Lambda^{-1} \circ F^{T-t+1}(x), G_\Lambda^{-1} \circ F_{\Lambda,\alpha}^{-t+1}(y)) \end{aligned}$$

For the inverse of F_0, we can use Lemma 3.3.1, with the permutation $\varphi = \mathrm{Id}$ because one of the points is on the orbit of x, to get for all $i \in E$ (because $\alpha \in \mathcal{C}[T, \Lambda, E]$):

$$\begin{aligned} \frac{1}{M} d_i(G_\Lambda^{-1} \circ F^{T-t+1}(x), G_\Lambda^{-1} \circ F_{\Lambda,\alpha}^{-t+1}(y)) &\leq d_i(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)) \\ &\leq \frac{1}{\gamma} d_i(G_\Lambda^{-1} \circ F^{T-t+1}(x), G_\Lambda^{-1} \circ F_{\Lambda,\alpha}^{-t+1}(y)) \end{aligned}$$

Combining these two estimates gives for any $i \in E$ and $1 \leq t \leq T$:

$$\begin{aligned} d_i(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)) &\leq \frac{1}{\gamma} d_i(F^{T-t+1}(x), F_{\Lambda,\alpha}^{-t+1}(y)) \\ &+ \frac{M\mathcal{E}}{\gamma} \sum_{k \in E} \theta^{2|i-k|} d_k(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)) + \frac{M\mathcal{E}}{2\gamma} \sum_{k \in \Lambda \setminus E} \theta^{2|i-k|} \end{aligned} \tag{3.33}$$

We want now to go from this time to time estimate to a global one (in time and space). We will estimate this term from above by a double sequence which can be entirely solved by a generating function method.
We do this with an analyze of the behavior of all points at a given distance of E^C. With $E^{(i)}$ as defined in Section 3.7, we denote for $0 \leq t \leq T$ and $i \geq 0$:

$$v(i,t) = \sup_{j \in E^{(-i)}} d_j(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y))$$

(and $v(i,t) = 0$ if $E^{(-i)} = \emptyset$)
If $j \in E^{(-i)}$, for any $0 \leq k \leq i$, we have the inclusion $j + \Lambda_k \subset E^{(k-i)} \subset E$, then (3.33) becomes for $t \geq 1$:

$$\begin{aligned} & d_j(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)) \\ & \quad \leq \frac{1}{\gamma} d_j(F^{T-t+1}(x), F_{\Lambda,\alpha}^{-t+1}(y)) + \frac{M\mathcal{E}}{\gamma} \sum_{k=0}^{i} \sum_{|l|=k} \theta^{2|l|} d_{j+l}(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)) \\ & \qquad + \frac{M\mathcal{E}}{2\gamma} \sum_{k>i} \sum_{|l|=k} \theta^{2|l|} \\ & \quad \leq \frac{1}{\gamma} v(i,t-1) + \frac{M\mathcal{E}}{\gamma} \sum_{k=0}^{i} \sum_{|l|=k} \theta^{2|l|} v(i-k,t) + \frac{M\mathcal{E}}{2\gamma} \sum_{k>i} \sum_{|l|=k} \theta^{2|l|} \end{aligned}$$

Hence for $i \geq 0$ and $1 \leq t \leq T$:

$$v(i,t) \leq \frac{1}{\gamma} v(i,t-1) + \frac{1}{\gamma} \sum_{k=0}^{i} \alpha_k v(i-k,t) + \frac{1}{\gamma} \sum_{k>i} \frac{\alpha_k}{2} \tag{3.34}$$

with $\alpha_k = M\mathcal{E} c_k \theta^{2k}$ and $c_k = \mathrm{Card}(l \in \mathbb{Z}^d \, : \, |l| = k)$. We define then, for $\delta \geq 0$ the double sequence:

$$u(i,t) = \begin{cases} \frac{1}{2} & \text{if } i < 0 \\ \delta & \text{if } i \geq 0, t = 0 \\ \frac{1}{\gamma} u(i,t-1) + \frac{1}{\gamma} \sum_{k \geq 0} \alpha_k u(i-k,t) & \text{if } i \geq 0, t > 0 \end{cases}$$

We have the following upper bound for v:

Lemma 3.3.2. *If $v(i,t)$ satisfies recursive relation (3.34), $\sup_{i \geq 0} v(i,0) = v(0,0) \leq \delta$, and if $\alpha_0/\gamma < 1$, then:*

$$v(i,t) \leq u(i,t) \qquad \forall\, i \geq 0, t \geq 0 \tag{3.35}$$

Proof. By induction on t, then on i, because $1 - \alpha_0/\gamma > 0$ and:

$$\left(1 - \frac{\alpha_0}{\gamma}\right) v(i,t) \leq \frac{1}{\gamma} v(i,t-1) + \frac{1}{\gamma} \sum_{k=1}^{i} \alpha_k v(i-k,t) + \frac{1}{\gamma} \sum_{k>i} \alpha_k u(i-k,t)$$

□

The fact that $\alpha_0/\gamma < 1$ is a direct consequence of the assumption $(H2)$ because $\alpha_0 \leq \sum \alpha_k = M\mathcal{K}_2 \leq M\mathcal{K} < \gamma$. $(H2)$ implies also that assumptions of Proposition 3.9.1 are satisfied with α_k and $\tilde{\alpha}_k = M\mathcal{E}c_k\theta^k$. This Proposition and Lemma 3.3.2 imply:

$$v(i,t) \leq \frac{\delta}{(\gamma - M\mathcal{K}_2)^t} + \lambda \cdot \theta^{i+1}$$

Optimizing for any $j \in E$, since $j \in E^{(-d(i,E^C)+1)}$, we get the desired estimate (3.32). □

We can evaluate in the same way the effect of a change of finite box restriction on the inverse iterates of the map:

Proposition 3.3.4. *If F satisfies Assumption $(H2)$ then for any $y \in \mathcal{A}$, there is a bijection $\varphi_y : \mathcal{C}[T,\Lambda,E] \to \mathcal{C}[T,\Lambda \setminus E]$ such that $y \mapsto \varphi_y$ is measurable, and for all $\alpha \in \mathcal{C}[T,\Lambda,E]$:*

$$d_i\left(F_{\Lambda,\alpha}^{-t}(y), F_{\Lambda\setminus E,\varphi_y(\alpha)}^{-t}(y)\right) \leq \lambda \cdot \theta^{d(i,E)} \qquad \forall\, 0 \leq t \leq T,\ i \in \Lambda \setminus E \tag{3.36}$$

Proof. For the coupling, we have exactly the same type of estimate as in the context of Proposition 3.3.3 for any $i \in \Lambda \setminus E$:

$$d_i(G_\Lambda^{-1}(y), G_{\Lambda\setminus E}^{-1}(z)) \leq d_i(y,z) + \mathcal{E} \sum_{k \in \Lambda \setminus E} \theta^{2|i-k|} d_k(y,z) + \frac{\mathcal{E}}{2} \sum_{k \in E} \theta^{2|i-k|} \tag{3.37}$$

The inverse branches of F_0 are constructed in Subsection 3.3.2 independently on each site and around the orbit of x. Since $F_\Lambda^t(x) = F_{\Lambda\setminus E}^t(x) = F^t(x)$, these inverse branches are in fact independent of the finite box. We can then use the same method as in the proof of Lemma 3.3.1 to choose inverse branches such that the contraction property applies well to pre images of y.
At first step, we compare for $i \in \Lambda \setminus E$ the relative positions of the points $\left(G_\Lambda^{-1}(y)\right)_i$, $\left(G_{\Lambda\setminus E}^{-1}(y)\right)_i$ and $\left(F_0 \circ F^{T-1}(x)\right)_i - 1/2$ to define the action of φ_y at

time $T-1$ (see Figure 3.1 in the proof of Lemma 3.3.1) such that:

$$\begin{aligned}
&\frac{1}{M}d_i\left(G_\Lambda^{-1}(y), G_{\Lambda\setminus E}^{-1}(y)\right)\\
&\qquad\leq d_i\left(F^{-1}_{0,T-1,\alpha(T-1,\cdot)}\circ G_\Lambda^{-1}(y), F^{-1}_{0,T-1,\varphi_y(\alpha)(T-1,\cdot)}\circ G_{\Lambda\setminus E}^{-1}(y)\right)\\
&\qquad\qquad\leq \frac{1}{\gamma}d_i\left(G_\Lambda^{-1}(y), G_{\Lambda\setminus E}^{-1}(y)\right)
\end{aligned}$$

Then, if φ_y is well defined for times greater or equal to $T-t+1$, we compare at each $i\in\Lambda\setminus E$ the relative positions of $\left(G_\Lambda^{-1}\circ F_{\Lambda,\alpha}^{-t+1}(y)\right)_i$, $\left(G_{\Lambda\setminus E}^{-1}\circ F_{\Lambda\setminus E,\alpha}^{-t+1}(y)\right)_i$ and $\left(F_0\circ F^{T-t}(x)\right)_i - 1/2$ to define the action of φ_y at time $T-t$ such that for all $\alpha\in\mathcal{C}[T,\Lambda,E]$:

$$\begin{aligned}
&\frac{1}{M}d_i\left(G_\Lambda^{-1}\circ F_{\Lambda,\alpha}^{-t+1}(y), G_{\Lambda\setminus E}^{-1}\circ F^{-t+1}_{\Lambda\setminus E,\varphi_y(\alpha)}(y)\right)\\
&\leq d_i\left(F^{-1}_{0,T-t,\alpha(T-t,\cdot)}\circ G_\Lambda^{-1}\circ F_{\Lambda,\alpha}^{-t+1}(y), F^{-1}_{0,T-t,\varphi_y(\alpha)(T-t,\cdot)}\circ G_{\Lambda\setminus E}^{-1}\circ F^{-t+1}_{\Lambda\setminus E,\alpha}(y)\right)\\
&\qquad\leq \frac{1}{\gamma}d_i\left(G_\Lambda^{-1}\circ F_{\Lambda,\alpha}^{-t+1}(y), G_{\Lambda\setminus E}^{-1}\circ F^{-t+1}_{\Lambda\setminus E,\varphi_y(\alpha)}(y)\right)
\end{aligned}$$

We get in the same way as for Lemma 3.3.1 that φ_y is a measurable function of y.
This gives then, combined with (3.37), for any $i\in\Lambda\setminus E$:

$$\begin{aligned}
d_i(F^{-t}_{\Lambda,\alpha}(y), F^{-t}_{\Lambda\setminus E,\varphi_y(\alpha)}(y)) &\leq \frac{1}{\gamma}d_i(F^{-t+1}_{\Lambda,\alpha}(y), F^{-t+1}_{\Lambda\setminus E,\varphi_y(\alpha)}(y))\\
&+\frac{M\mathcal{E}}{\gamma}\sum_{k\in\Lambda\setminus E}\theta^{2|i-k|}d_k(F^{-t}_{\Lambda,\alpha}(x), F^{-t}_{\Lambda\setminus E,\varphi_y(\alpha)}(y)) + \frac{M\mathcal{E}}{2\gamma}\sum_{k\in E}\theta^{2|i-k|}
\end{aligned}\tag{3.38}$$

We can hence proceed as in the proof of Proposition 3.3.3, with:

$$v(i,t) = \sup_{j\in\Lambda\setminus\left(E^{(i)}\right)} d_j\left(F^{-t}_{\Lambda,\alpha}(y), F^{-t}_{\Lambda\setminus E,\varphi_y(\alpha)}(y)\right)$$

and $\delta = 0$. □

3.3.4 Expansiveness

A first consequence of the expanding property 3.3.3 is the expansiveness of the dynamical system (F,S):

Proposition 3.3.5. *If $d_\rho(S^i\circ F^t(x), S^i\circ F^t(y)) < \delta_0 = \frac{1}{2M}$ for all $i\in\mathbb{Z}^d$ and $t\in\mathbb{N}$, then:*

$$x = y$$

Proof. The inclusion (3.30) and the Proposition 3.3.3 can be combined to get that under assumption $(H2)$, if $d_i(F_\Lambda^t(x), F_\Lambda^t(y)) < \delta_0$ for all $0 \leq t \leq T$ and $i \in E$, then we have in fact the better estimate:

$$d_i\left(F_\Lambda^t(x), F_\Lambda^t(y)\right) \leq \frac{\delta}{\tilde{\gamma}^{T-t}} + \lambda \cdot \theta^{d(i,E^C)}$$

We can then take $\Lambda = \Lambda_N$ and N tends to infinity which gives the same property for the global map F. But the assumption done for this Proposition clearly implies that $d_i(F^t(x), F^t(y)) < \delta_0$ for all $i \in \mathbb{Z}^d$ and $t \in \mathbb{N}$, hence:

$$d_i\left(x, y\right) \leq \frac{\delta}{\tilde{\gamma}^{T}} + \lambda \cdot \theta^{d(i,E^C)}$$

for all $E \subset \mathbb{Z}^d$ and $T \in \mathbb{N}$. taking $E = \Lambda_n$ then T and n going to infinity, we can conclude that $x = y$. □

A classical and essential consequence of this property is that the metric entropy $h_{(F,S)}$ associated to the system is an upper semi-continuous function of the probability measures (see Proposition 3.8.1). This (and the continuity of the potential function φ) proves that the rate function I of the Large Deviations Principle defined in (3.14) is lower semi-continuous and allows to use the Gibbs variational principle for the proof of the Upper Bound.

3.4 Proof of the Volume Lemma

We begin by proving an intermediate Volume Lemma for the finite box map F_Λ with constraints on the orbit on the smaller box E, then use it to prove Theorem 3.1.1 for the global system (F, S).

Proposition 3.4.1. *Under assumption $(\mathcal{H})$, for x, $E \subset \Lambda$, T and $0 < \delta < \frac{1}{2M}$ as in Section 3.3 with Λ large enough, we have:*

$$\begin{aligned}
\exp\Big(\sum_{\substack{0\le t<T\\ i\in E}} \varphi \circ S^i \circ F^t(x) - T|E|\tilde{C}_2(T,E,\delta) - C_4(\Lambda,T,E)\Big) \\
\le m^\Lambda \left\{ y \,:\, d_i\left(F_\Lambda^t(x), F_\Lambda^t(y)\right) < \delta \quad \forall\, 0 \le t \le T,\, i \in E \right\} \\
\le \exp\Big(\sum_{\substack{0\le t<T\\ i\in E}} \varphi \circ S^i \circ F^t(x) + T|E|\tilde{C}_3(T,E,\delta) + C_5(\Lambda,T,E))\Big)
\end{aligned} \tag{3.39}$$

with:

$$\lim_{N\to\infty} C_4(\Lambda_N,T,E) = \lim_{N\to\infty} C_5(\Lambda_N,T,E) = 0 \qquad \forall T \ge 1,\ E \subset \mathbb{Z}^d \tag{3.40}$$

$$\lim_{\substack{T\to\infty\\ n\to\infty}} \tilde{C}_2(T,E_n,\delta) = \lim_{\substack{T\to\infty\\ n\to\infty}} \tilde{C}_3(T,E_n,\delta) = 0 \qquad \forall \delta < \frac{1}{2M} \tag{3.41}$$

for any sequence E_n tending to $\mathbb{Z}^d$ in the sense of Van Hove. Moreover $\tilde{C}_2$ and $\tilde{C}_3$ are continuous in δ.

The essential idea to prove this result is to do a change of variable by F_Λ^T. This must be done with some precautions to ensure we are on domains where this map is injective and to analyze all the terms.

3.4.1 Proof of the Upper Bound of Proposition 3.4.1

We decompose $\mathcal{X}_\Lambda$ in the subsets $(\mathcal{A}_\alpha(x))_{\alpha\in\mathcal{C}[T,\Lambda]}$, on each of which F_Λ^T is one-to-one. It has to be noticed that we do not lose anything because $m^\Lambda(\cup\mathcal{A}_\alpha(x)) = 1$ and that since $\delta < \frac{1}{2M}$ the intervals which appear are those corresponding to $\mathcal{C}[T,\Lambda,E]$ (see Proposition 3.3.2 for these properties):

$$\begin{aligned}
& m^\Lambda \left\{ y \in \mathcal{X}_\Lambda \,:\, d_i\left(F^t(x), F_\Lambda^t(y)\right) < \delta \quad \forall\, 0 \le t \le T,\, i \in E \right\} \\
&= \sum_{\alpha\in\mathcal{C}[T,\Lambda,E]} m^\Lambda \left\{ y \in \mathcal{A}_\alpha(x) \,:\, d_i\left(F^t(x), F_\Lambda^t(y)\right) < \delta \quad \forall\, 0 \le t \le T,\, i \in E \right\} \\
&= \sum_{\alpha\in\mathcal{C}[T,\Lambda,E]} \int_{\mathcal{X}_\Lambda} \prod_{\substack{0\le t\le T\\ i\in E}} \mathbb{1}_{\left\{d_i\left(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)\right)<\delta\right\}} \frac{1}{\left|DF_\Lambda^T(F_{\Lambda,\alpha}^{-T}(y))\right|}\, m^\Lambda(dy)
\end{aligned} \tag{3.42}$$

by a change of variables with F_Λ^T, bijection from $\mathcal{A}_\alpha(x)$ onto $\mathcal{A}$.
We apply then the results of Section 3.2.2 to get:

$$\begin{aligned}\frac{1}{\left|DF_\Lambda^T(F_{\Lambda,\alpha}^{-T}(y))\right|} &= \exp\Bigg(-\sum_{0\le t<T}\log\left|DF_{\Lambda,F^t(x)}\circ F_{\Lambda,\alpha}^{t-T}(y)\right|\Bigg)\\ &= \exp\Big(\sum_{\substack{0\le t<T\\ i\in\Lambda}}(-\log|f_i'|+w_{\Lambda,i})\circ F_{\Lambda,\alpha}^{t-T}(y)\Big)\end{aligned}$$

where we denote $w_{\Lambda,i}=w_{\Lambda,F^t(x),i}$ for any t: we don't mention the boundary conditions since all our estimates are uniform in them.
We treat differently the terms corresponding to $i\in E$ and to $i\in\Lambda\setminus E$. In the first case, we want to replace them by $\varphi\circ S^i\circ F^t(x)$ while in the second we want to reconstitute $D(F_{\Lambda\setminus E,\varphi_y(\alpha)}^{-T}(y))$ and integrate it to 1 by another change of variables on $\mathcal{X}_{\Lambda\setminus E}$.

Hence, if $i\in E$:

$$\begin{aligned}&|(-\log|f_i'|+w_{\Lambda,i})\circ F_{\Lambda,\alpha}^{t-T}(y)-\varphi\circ S^i\circ F^t(x)|\\ &\le|\log|f_i'|\circ F_{\Lambda,\alpha}^{t-T}(y)-\log|f_i'|\circ F^t(x)|+|w_{\Lambda,i}\circ F_{\Lambda,\alpha}^{t-T}(y)-w_{\Lambda,i}\circ F^t(x)|\\ &\qquad\qquad+|w_{\Lambda,i}\circ F^t(x)-\psi_i\circ F^t(x)|\end{aligned}$$

The third term is easily estimated by the speed of convergence of $w_{\Lambda,i}$ to ψ_i given in (3.25). Summing over all times and sites gives

$$\sum_{\substack{0\le t<T\\ i\in E}}|w_{\Lambda,i}\circ F^t(x)-\psi_i\circ F^t(x)|\le\frac{T}{2(1-\mathcal{K})}\sum_{i\in E}\theta^{d(i,\Lambda^C)}=C_5(\Lambda,T,E)\quad(3.43)$$

then we get $C_5(\Lambda_N,T,E)\to 0$ when N goes to infinity.
For the two other terms, we use the fact that $d_i\left(F^{T-t}(x),F_{\Lambda,\alpha}^{-t}(y)\right)<\delta$ for all $0\le t\le T$ and $i\in E$ which implies with Proposition 3.3.3 that:

$$d_i\left(F^{T-t}(x),F_{\Lambda,\alpha}^{-t}(y)\right)\le\frac{\delta}{\tilde\gamma^t}+\lambda\cdot\theta^{d(i,E^C)}\qquad\forall\,0\le t\le T,\,i\in E$$

This combined with the α-Hőlder property of $\log|f'|$ (see (3.2)) and the concavity of $x\to x^\alpha$ gives:

$$\begin{aligned}&\frac{1}{T|E|}\sum_{\substack{0\le t<T\\ i\in E}}|\log|f_i'|\circ F_{\Lambda,\alpha}^{t-T}(y)-\log|f_i'|\circ F^t(x)|\\ &\qquad\qquad\le C_1\Big(\frac{\delta}{T}\sum_{0\le t<T}\frac{1}{\tilde\gamma^{t-T}}+\frac{\lambda}{|E|}\sum_{i\in E}\theta^{d(i,E^C)}\Big)^\alpha\quad(3.44)\end{aligned}$$

which goes to 0 as T tends to infinity and E tends to $\mathbb{Z}^d$ in the sense of Van Hove, because $\tilde{\gamma} > 1$ and $1/|E| \sum_{i\in E} \theta^{d(i,E^C)}$ goes to 0 by Proposition 3.7.1. For $w_{\Lambda,i}$, we use estimate (3.20) and get, with $\mathcal{K}_{1/2} = \sum_{i\in\mathbb{Z}^d} \theta^{\frac{1}{2}|k|}$:

$$\begin{aligned}
&|w_{\Lambda,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda,i} \circ F^t(x)| \\
&\qquad \le \frac{M\mathcal{E}}{1-\mathcal{K}} \sum_{k\in\Lambda} \theta^{|i-k|} d_k(F^{t-T}_{\Lambda,\alpha}(y), F^t(x)) \\
&\qquad \le \frac{M\mathcal{K}}{1-\mathcal{K}} \frac{\delta}{\tilde{\gamma}^{t-T}} + \frac{\lambda M \mathcal{K}_{1/2}}{1-\mathcal{K}} \theta^{\frac{1}{2}d(i,E^C)} + \frac{M\mathcal{E}}{2(1-\mathcal{K})} \sum_{k\in E^C} \theta^{|i-k|}
\end{aligned}$$

$$\begin{aligned}
\text{Then :}\quad &\frac{1}{T|E|} \sum_{\substack{0\le t<T \\ i\in E}} |w_{\Lambda,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda,i} \circ F^t(x)| \\
&\le \frac{M\mathcal{K}}{1-\mathcal{K}} \frac{\delta}{T} \sum_{0\le t<T} \frac{1}{\tilde{\gamma}^{t-T}} + \frac{M\mathcal{K}_{1/2}}{1-\mathcal{K}} \left(\frac{1}{2} + \lambda\right) \frac{1}{|E|} \sum_{i\in E} \theta^{\frac{1}{2}d(i,E^C)} \qquad (3.45)
\end{aligned}$$

which goes also to 0 as $T \to \infty$ and $E \to \mathbb{Z}^d$.

In the same way, for $i \in \Lambda\backslash E$, we use the link between behaviors of $F^{t-T}_{\Lambda,\alpha}(y)$ and $F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)$ given in Proposition 3.3.4:

$$\begin{aligned}
&|(-\log|f_i'| + w_{\Lambda,i}) \circ F^{t-T}_{\Lambda,\alpha}(y) - (-\log|f_i'| + w_{\Lambda\backslash E,i}) \circ F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)| \\
&\le |\log|f_i'| \circ F^{t-T}_{\Lambda,\alpha}(y) - \log|f_i'| \circ F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)| + |w_{\Lambda,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda,\alpha}(y)| \\
&\qquad + |w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)|
\end{aligned}$$

and, using Proposition 3.3.4 instead of Proposition 3.3.3 and estimate (3.22) instead of (3.25):

$$\frac{1}{T|E|} \sum_{\substack{0\le t<T \\ i\in\Lambda\backslash E}} |\log|f_i'| \circ F^{t-T}_{\Lambda,\alpha}(y) - \log|f_i'| \circ F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)| \le C_1 \frac{\lambda^\alpha}{|E|} \sum_{i\in E^C} \theta^{\alpha d(i,E)} \qquad (3.46)$$

$$\frac{1}{T|E|} \sum_{\substack{0\le t<T \\ i\in\Lambda\backslash E}} |w_{\Lambda,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda,\alpha}(y)| \le \frac{1}{2(1+\mathcal{K})} \frac{1}{|E|} \sum_{i\in E^C} \theta^{d(i,E)} \qquad (3.47)$$

$$\begin{aligned}
&\frac{1}{T|E|} \sum_{\substack{0\le t<T \\ i\in\Lambda\backslash E}} |w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda,\alpha}(y) - w_{\Lambda\backslash E,i} \circ F^{t-T}_{\Lambda\backslash E,\varphi_y(\alpha)}(y)| \\
&\qquad\qquad \le \frac{\lambda M \mathcal{K}_{1/2}}{1-\mathcal{K}} \frac{1}{|E|} \sum_{i\in E^C} \theta^{\frac{1}{2}d(i,E)} \qquad (3.48)
\end{aligned}$$

all these terms tending to 0 when E tends to $\mathbb{Z}^d$ in the sense of Van Hove by estimate (3.58).

We take finally for $\bar{C}_3$ the sum of RHS in formulas (3.44), (3.45), (3.46), (3.47) and (3.48) and get the global estimate:

$$\frac{1}{|DF_\Lambda^T(F_{\Lambda,\alpha}^{-T}(y))|} \leq \frac{1}{|DF_{\Lambda\setminus E}^T(F_{\Lambda\setminus E,\varphi_y(\alpha)}^{-T}(y))|} \times \exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}} \varphi\circ S^i\circ F^t(x) + T|E|\bar{C}_3(T,E,\delta) + C_5(\Lambda,T,E)\Big)$$

On the other hand, we get an upper bound for the product of indicator functions in (3.42) by the terms corresponding to $t=0$, and use the identity:

$$\sum_{\alpha\in\mathcal{C}[T,\Lambda,E]} \frac{1}{DF^T\circ F_{\Lambda,\varphi_y(\alpha)}^{-T}} = \sum_{\alpha\in\mathcal{C}[T,\Lambda\setminus E]} \frac{1}{DF^T\circ F_{\Lambda,\alpha}^{-T}}$$

do to the bijectivity of φ_y from $\mathcal{C}[T,\Lambda,E]$ onto $\mathcal{C}[T,\Lambda\setminus E]$. We can then separate the terms in E and those in $\Lambda\setminus E$ and integrate the last ones by a change of variable:

$$\begin{aligned}
&m^\Lambda\left\{y\in\mathcal{X}_\Lambda\,:\,d_i\left(F_\Lambda^t(x),F_\Lambda^t(y)\right)<\delta\quad\forall\, 0\leq t\leq T,\, i\in E\right\}\\
&\leq \int_{\mathcal{X}_{\Lambda\setminus E}} \sum_{\alpha\in\mathcal{C}[T,\Lambda\setminus E]} \frac{1}{|DF_{\Lambda\setminus E}^T(F_{\Lambda\setminus E,\alpha}^{-T}(y))|} m^{\Lambda\setminus E}(dy)\\
&\qquad\times m^E\{y\,:\,d_i(F^T(x),y)<\delta\quad\forall i\in E\}\\
&\qquad\times \exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}} \varphi\circ S^i\circ F^t(x) + T|E|\bar{C}_3(T,E,\delta) + C_5(\Lambda,T,E)\Big)\\
&= m^{\Lambda\setminus E}\Big(\bigcup_{\alpha\in\mathcal{C}[T,\Lambda\setminus E]} \mathcal{A}_\alpha(x)\Big)(2\delta)^{|E|}\\
&\qquad\times \exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}} \varphi\circ S^i\circ F^t(x) + T|E|\bar{C}_3(T,E,\delta) + C_5(\Lambda,T,E)\Big)
\end{aligned}$$

Hence:

$$\begin{aligned}
&m^\Lambda\left\{y\in\mathcal{X}_\Lambda\,:\,d_i\left(F_\Lambda^t(x),F_\Lambda^t(y)\right)<\delta\quad\forall\, 0\leq t\leq T,\, i\in E\right\}\\
&\qquad\leq \exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}} \varphi\circ S^i\circ F^t(x) + T|E|\tilde{C}_3(T,E,\delta) + C_5(\Lambda,T,E)\Big)
\end{aligned}$$

where $\tilde{C}_3 = \bar{C}_3 + \frac{1}{T}\log(2\delta)$ satisfies the announced limit.

3.4.2 Proof of the Lower Bound of Proposition 3.4.1

For the lower bound, we use the same kind of estimates that for the upper bound, except for the term

$$\prod_{\substack{0\leq t\leq T \\ i\in E}} \mathbb{1}_{\left\{d_i\left(F_{\Lambda,0}^{-t}(F^T(x)),F_{\Lambda,\alpha}^{-t}(y)\right)<\delta\right\}}$$

Indeed, to insure this, we have to assume that $d_i(F^T(x), y) < \delta$ for i in a set larger than E: we choose L such that

$$\frac{\delta}{\tilde{\gamma}} + \lambda \cdot \theta^L \leq \delta$$

and assume that $E^{(L)} \subset \Lambda$ (this is the sense of Λ large enough in Proposition 3.4.1).
Then, if $d_i(F^T(x), y) < \delta$ for all $i \in E^{(L)}$, Proposition 3.3.3 implies that when $\alpha \in \mathcal{C}[T, \Lambda, E^{(L)}]$:

$$d_i\left(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)\right) \leq \frac{\delta}{\tilde{\gamma}^t} + \lambda \cdot \theta^{d(i,(E^{(L)})^C)} \qquad \forall\, 0 \leq t \leq T,\ i \in E^{(L)}$$

and in particular

$$d_i\left(F^{T-t}(x), F_{\Lambda,\alpha}^{-t}(y)\right) \leq \delta \qquad \forall\, 0 \leq t \leq T,\ i \in E$$

The assumption $\alpha \in \mathcal{C}[T, \Lambda, E^{(L)}]$ imposes then to restrict the sum in the decomposition of $\mathcal{X}_\Lambda$. This does not perturb the asymptotic estimates because $\frac{|E^{(L)}\setminus E|}{|E|} \to 0$ when E tends to $\mathbb{Z}^d$ in the sense of Van Hove. Then:

$$\begin{aligned}
&m^\Lambda \left\{y \in \mathcal{X}_\Lambda \,:\, d_i\left(F_\Lambda^t(x), F_\Lambda^t(y)\right) < \delta \quad \forall\, 0 \leq t \leq T,\ i \in E\right\} \\
&\quad \geq \sum_{\alpha\in\mathcal{C}[T,\Lambda,E^{(L)}]} \int_{\mathcal{X}_\Lambda} \prod_{i\in E^{(L)}} \mathbb{1}_{\{d_i(F^T(x),y)<\delta\}} \\
&\qquad\qquad \times \exp\Bigg(\sum_{\substack{0\leq t<T \\ i\in\Lambda}} (-\log f_i' + w_{\Lambda,i}) \circ f_{\Lambda,\alpha}^{t-T}(y)\Bigg)\, m^\Lambda(dy)
\end{aligned}$$

$$\geq m^{\Lambda\setminus E^{(L)}}\Big(\bigcup_{\alpha\in\mathcal{C}[T,\Lambda\setminus E^{(L)}]}\mathcal{A}_\alpha(x)\Big)m^{E^{(L)}}\{y\,:\,d_i(F^T(x),y)<\delta\quad\forall\,i\in E^{(L)}\}$$
$$\times\exp\Big(\sum_{\substack{0\leq t<T\\ i\in E^{(L)}}}\varphi\circ S^i\circ F^t(x)-T|E^{(L)}|\bar{C}_3(T,E^{(L)},\delta)-C_5(\Lambda,T,E^{(L)})\Big)$$
$$\geq\exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}}\varphi\circ S^i\circ F^t(x)-T|E|\tilde{C}_2(T,E,\delta)-C_4(\Lambda,T,E)\Big)$$

where
$$\tilde{C}_2(T,E,\delta)=\frac{|E^{(L)}|}{|E|}\tilde{C}_3(T,E^{(L)},\delta)+\frac{|E^{(L)}\setminus E|}{|E|}|\varphi|_\infty$$

tends to 0 as T goes to infinity and E tends to $\mathbb{Z}^d$ in the sense of Van Hove, and $C_4(\Lambda,T,E)=C_5(\Lambda,T,E^{(L)})$.

3.4.3 Proof of Theorem 3.1.1

We approximate F by F_{Λ_N} using convergence on a finite box for finite time: for any $0<\varepsilon<\frac{1}{2M}-\delta$, there exists N_0 such that for all $N\geq N_0$:

$$\begin{cases} d_i(F^t_{\Lambda_N}(y),F^t(y))\leq\varepsilon\quad\forall\,0\leq t\leq T,\ i\in E\text{ and }y\in\mathcal{X}\\ C_5(\Lambda_N,T,E)\leq\varepsilon\end{cases}$$

We deduce then from the upper bound of Proposition 3.4.1 applied to F_{Λ_N}:

$$\begin{aligned}&\overline{m}(B_x(T,E;\delta))\\ &\leq\overline{m}\left\{y\in\mathcal{X}\,:\,d_i\left(F^t(x),F^t(y)\right)<\delta\quad\forall\,0\leq t\leq T,\ i\in E\right\}\\ &\leq m^{\Lambda_N}\left\{y\in\mathcal{X}_{\Lambda_N}\,:\,d_i\left(F^t_{\Lambda_N}(x),F^t_{\Lambda_N}(y)\right)<\delta+\varepsilon\quad\forall\,0\leq t\leq T,\ i\in E\right\}\\ &\leq\exp\Big(\sum_{\substack{0\leq t<T\\ i\in E}}\varphi\circ S^i\circ F^t(x)+T|E|\tilde{C}_3(T,E,\delta+\varepsilon)+C_5(\Lambda_N,T,E)\Big)\end{aligned}$$

We take then $N\to\infty$, $\varepsilon\to0$ and use continuity of $\tilde{C}_3$ in δ to get the desired upper bound with $C_3=\exp(T|E|\tilde{C}_3)$.

In the same way, for the lower bound, let $\tilde{L}$ be such that $\frac{1}{2}\rho^{\tilde{L}+1}<\delta\leq\frac{1}{2}\rho^{\tilde{L}}$, and for any $0<\varepsilon<\delta$ let N_1 such that for all $N\geq N_1$:

$$\begin{cases} d_i(F^t_{\Lambda_N}(y),F^t(y))\leq\varepsilon\quad\forall\,0\leq t\leq T,\ i\in E^{(\tilde{L})}\text{ and }y\in\mathcal{X}\\ C_4(\Lambda_N,T,E^{(\tilde{L})})\leq\varepsilon\end{cases}$$

Then:

$$
\begin{aligned}
&\overline{m}(B_x(T,E;\delta)) \\
&= \overline{m}\left\{ y \;:\; \begin{array}{ll} d_i(F^t(x),F^t(y)) < \delta & \forall\, i \in E, \\ d_i(F^t(x),F^t(y)) < \delta\rho^{-1} & \forall\, i \in E^{(1)} \setminus E, \\ \vdots & \vdots \\ d_i(F^t(x),F^t(y)) < \delta\rho^{-\tilde{L}} & \forall\, i \in E^{(\tilde{L})} \setminus E^{(\tilde{L}-1)}, \end{array} \quad \forall\, 0 \le t \le T \right\} \\
&\ge \overline{m}\left\{ y \in \mathcal{X} \;:\; d_i\left(F^t(x),F^t(y)\right) < \delta \quad \forall\, 0 \le t \le T,\, i \in E^{(\tilde{L})} \right\} \\
&\ge m^{\Lambda_N}\left\{ y \in \mathcal{X}_{\Lambda_N} \;:\; d_i\left(F^t_{\Lambda_N}(x),F^t_{\Lambda_N}(y)\right) < \delta - \varepsilon \quad \forall\, 0 \le t \le T,\, i \in E^{(\tilde{L})} \right\} \\
&\ge \exp\Bigg(\sum_{\substack{0 \le t < T \\ i \in E}} \varphi \circ S^i \circ F^t(x) - T|E^{(\tilde{L})}|\tilde{C}_2(T,E^{(\tilde{L})},\delta-\varepsilon) - C_4(\Lambda_N,T,E^{(\tilde{L})}) \Bigg)
\end{aligned}
$$

We get the desired lower bound with $C_2 = \exp\left(T|E^{(\tilde{L})}|\tilde{C}_2(T,E^{(\tilde{L})},\delta)\right)$ when ε tends to 0. The only dependence of C_2 on the constant ρ defining the distance comes from the choice of $\tilde{L}$.

3.5 Large deviations upper bound

Our proof of the upper bound of the large deviations principle follows, at least for the main steps, the method of Kifer in [71]. It presents no particular difficulty since the space $\mathcal{M}^1(\mathcal{X})$ is compact for the weak-* topology, i.e. the weakest topology for which the evaluations $\nu \mapsto \int g \, d\nu$ are continuous for any $g \in \mathcal{C}(\mathcal{X})$. The Volume Lemma gives the identification of the log Laplace transforms.
For E_T a given sequence of subsets of $\mathbb{Z}^d$, we denote:

$$R_T(x) = R_{T,E_T}(x) = \frac{1}{T|E_T|} \sum_{\substack{0 \le t < T \\ i \in E_T}} \delta_{S^i \circ F^t(x)} \in \mathcal{M}^1(\mathcal{X})$$

the associated empirical process.

3.5.1 Identification of the pressure

The first step in this proof is the identification of the limit of the log-Laplace transforms of the empirical process R_T integrated against any continuous potential V with the topological pressure of $V + \varphi$ (see Section 3.8.3 for the definition):

Proposition 3.5.1. *Under assumption $(\mathcal{H})$, for any sequence $(E_T)_{T \ge 0}$ tending to $\mathbb{Z}^d$ in the sense of Van Hove and $V \in \mathcal{C}(\mathcal{X})$, we have:*

$$\limsup_{T \to \infty} \frac{1}{T|E_T|} \log \int_{\mathcal{X}} \exp \left(T|E_T| \int_{\mathcal{X}} V \, dR_T(x) \right) \overline{m}(dx) = P_{(F,S)}(V + \varphi) \tag{3.49}$$

Corollary 3.1.1 is immediately deduced from this Proposition, taking $V = 0$.

Proof. For $\delta > 0$ and $T \ge 0$, we take Y a maximal (T, δ)-separated set in $\mathcal{X}$, which means that:

$$x, x' \in Y \text{ and } x \ne x' \Longrightarrow x' \notin B_x(T, E_T; \delta)$$

and Y is maximal for this property.
Then $\cup_{x \in Y} B_x(T, E_T; \delta) = \mathcal{X}$ by maximality and if $x, x' \in Y$ are distinct then

$$B_x(T, E_T; \delta/2) \cap B_{x'}(T, E_T; \delta/2) = \emptyset$$

Hence, denoting $\gamma_V(\delta) = \sup\{|V(x) - V(y)| : d_\rho(x, y) < \delta\}$, a quantity which goes to 0 with δ by continuity, we decompose the integral in small balls and get:

$$\sum_{x\in Y}\exp\Big(\sum_{\substack{0\le t<T\\ i\in E_T}}(V\circ S^i\circ F^t(x)-\gamma_V(\delta/2))\Big)\overline{m}(B_x(T,E_T;\delta/2))$$
$$\le\int_{\mathcal{X}}\exp\Big(\sum_{\substack{0\le t<T\\ i\in E_T}}V\circ S^i\circ F^t(x)\Big)\,\overline{m}(dx)$$
$$\le\sum_{x\in Y}\exp\Big(\sum_{\substack{0\le t<T\\ i\in E_T}}(V\circ S^i\circ F^t(x)+\gamma_V(\delta))\Big)\overline{m}(B_x(T,E_T;\delta))$$

We use then the Volume Lemma to get:

$$C_2(T,E_T,\delta/2,\rho)\Big[\sum_{x\in Y}\exp\Big(\sum_{\substack{0\le t<T\\ i\in E_T}}(V+\varphi)\circ S^i\circ F^t(x)-\gamma_V(\delta/2)\Big)\Big]$$
$$\le\int_{\mathcal{X}}\exp\left(T|E_T|\int_{\mathcal{X}}V\,dR_T(x)\right)\,\overline{m}(dx)$$
$$\le C_3(T,E_T,\delta)\Big[\sum_{x\in Y}\exp\Big(\sum_{\substack{0\le t<T\\ i\in E_T}}(V+\varphi)\circ S^i\circ F^t(x)+\gamma_V(\delta)\Big)\Big]$$

We take know the logarithm of each term, divide by $T|E_T|$ and take successively the supremum on maximal (T,δ)-separated sets, the limsup when T goes to infinity (makes the terms C_2 and C_3 disappear) and the limit $\delta\to 0$. We get hence the desired result directly from the definition of topological pressure. □

3.5.2 Proof of the upper bound

For $\delta>0$ and $V\in\mathcal{C}(\mathcal{X})$ fixed, $\mathcal{M}^1(\mathcal{X})$ is compact and any closed subset F can be included in a finite union of balls of the type $\beta_\nu(V;\delta)=\{\mu\,:\,|\int V\,d\mu-\int V\,d\nu|<\delta\}$:

$$F\subset\bigcup_{l=1}^{d}\beta_{\nu_l}(V;\delta)\quad\text{with }\nu_l\in F\tag{3.50}$$

By the Chebychev inequality:

$$\overline{m}\,\{x\,:\,R_T(x)\in\beta_\nu(V;\delta)\}\le e^{T|E_T|(\delta-\int_{\mathcal{X}}V\,d\nu)}\int_{\mathcal{X}}e^{T|E_T|R_T(x)}\,\overline{m}(dx)$$

then, using Proposition 3.5.1, we have for such an open ball:

$$\limsup_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(R_T\in\beta_\nu(V;\delta))\le\delta-\int_{\mathcal{X}}V\,d\nu+P_{(F,S)}(V+\varphi)$$

The inclusion (3.50) implies now for F closed:

$$\begin{aligned}\limsup_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(R_T\in F) &\leq \max_{1\leq l\leq d}\left(\limsup_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(R_T\in\beta_{\nu_l}(V;\delta))\right)\\ &\leq \max_{\nu\in F}\left(\delta-\int_{\mathcal{X}}V\,d\nu+P_{(F,S)}(V+\varphi)\right)\end{aligned}$$

We can then make δ tend to 0, optimize on V continuous and use a minimax type result (available because F is compact) to get:

$$\begin{aligned}\limsup_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(R_T\in F) &\leq \max_{\nu\in F}\left(\inf_{V\in\mathcal{C}(\mathcal{X})}\left(P_{(F,S)}(V+\varphi)-\int_{\mathcal{X}}V\,d\nu\right)\right)\\ &= \sup_{\nu\in F\cap\mathcal{M}^1_{\mathrm{inv}}(\mathcal{X})}\left(h_{(F,S)}(\nu)-\int_{\mathcal{X}}\varphi\,d\nu\right)\\ &= -\inf_{\nu\in F}I(\nu)\end{aligned}$$

where we used the dual Gibbs variational principle (because h is upper semi-continuous, see Section 3.8) for invariant measures, and the fact that if $\nu\notin\mathcal{M}^1_{\mathrm{inv}}(\mathcal{X})$,

$$\inf_{V\in\mathcal{C}(\mathcal{X})}\left(P_{(F,S)}(V+\varphi)-\int_{\mathcal{X}}V\,d\nu\right)=-\infty=-I(\nu).$$

(We check it taking $V_\alpha=\alpha(g\circ F^t\circ S^i-g)-\varphi$ with g such that $\int_{\mathcal{X}}g\circ F^t\circ S^i\,d\nu\neq\int_{\mathcal{X}}g\,d\nu$.)

3.6 Large deviations lower bound

The large deviations lower bound is a local property in the sense that it is equivalent to prove for all open set O such that $\nu \in O$:

$$\liminf_{T\to\infty} \frac{1}{T|E_T|} \log \overline{m}\{x \,:\, R_T(x) \in O\} \geq -I(\nu) = h_{(F,S)}(\nu) + \int_{\mathcal{X}} \varphi \, d\nu$$

or, equivalently:

$$\liminf_{T\to\infty} \frac{1}{T|E_T|} \log \overline{m}\{x \,:\, R_T(x) \in \beta_\nu(V_1, \ldots, V_K; \delta)\} \geq h_{(F,S)}(\nu) + \int_{\mathcal{X}} \varphi \, d\nu$$

for all $\nu \in \mathcal{M}^1(\mathcal{X})$, $V_1, \ldots, V_K \in \mathcal{C}(\mathcal{X})$ and $\delta > 0$, denoting $\beta_\nu(V_1, \ldots, V_K; \delta) = \{\mu \,:\, |\int_{\mathcal{X}} V_k \, d\mu - \int_{\mathcal{X}} V_k \, d\nu| < \delta \quad \forall\, 1 \leq k \leq K\}$, because this gives a basis of the weak* topology on $\mathcal{M}^1(\mathcal{X})$.

The idea for the lower bound is a geometric estimate, which comes from [105] and is better expressed for an ergodic probability ν: we decompose the set $\{x \,:\, R_T(x) \in \beta_\nu(V_1, \ldots, V_K; \delta)\}$ in small balls $B_x(T, E_T; \delta)$. We need approximately $e^{(T|E_T|h_{(F,S)}(\nu))}$ of them (by Proposition 3.8.3) and each is approximately of size $e^{(T|E_T| \int_{\mathcal{X}} \varphi \, d\nu)}$ under $\overline{m}$ (by the Volume Lemma and the Ergodic Theorem 3.8.1).

We will write it directly for convex combinations of ergodic measures. We need for this a strong mixing result, the Specification Property. We obtain the general case by an approximation argument.

3.6.1 Specification property

This strong quantitative mixing property is again a consequence of the preservation of expanding property.

Proposition 3.6.1. *If F satisfies $(H2)$, then for all $\delta > 0$, there exists $p(\delta) \in \mathbb{N}$ such that for any $T_1, \ldots, T_L \in \mathbb{N}$, $x^1, \ldots, x^L \in \mathcal{X}$ and $p_1, \ldots, p_{L-1} \geq p(\delta)$, there exists $x \in \mathcal{X}$ such that:*

$$\begin{aligned}
d\left(F^t(x), F^t\left(x^1\right)\right) &< \delta & \forall\, 0 \leq t \leq T_1 \\
d\left(F^{t+T_1+p_1}(x), F^t\left(x^2\right)\right) &< \delta & \forall\, 0 \leq t \leq T_2 \\
&\vdots & \vdots \\
d\left(F^{t+\sum_{l=1}^{L-1}(T_l+p_l)}(x), F^t\left(x^L\right)\right) &< \delta & \forall\, 0 \leq t \leq T_L
\end{aligned}$$

Proof. We work in this proof with the global map F and the topology associated to the distance $d(x,y) = \sup_{i\in\mathbb{Z}^d} d_i(x,y)$. Let

$$V_x(T;\delta) = \{y \,:\, d(F^t(x), F^t(y)) < \delta \quad \forall\, 0 \le t \le T\}$$

be the dynamic neighborhood around the orbit of x. We want to show that:

$$V_{x^1}(T_1;\delta) \cap F^{-T_1-p_1}(V_{x^2}(T_2;\delta)) \cap \cdots \cap F^{-\sum_{l=1}^{L-1}(T_l+p_l)}(V_{x^L}(T_L;\delta)) \neq \emptyset$$

By a simple induction argument, it is sufficient to show that for all $x \in \mathcal{X}$, $T \ge 0$, $0 < \delta < \frac{1}{2M}$, $p \ge p(\delta)$ and A such that $\mathrm{Int}(A) \neq \emptyset$, we have:

$$V_x(T;\delta) \cap F^{-T-p}(\mathrm{Int}(A)) \neq \emptyset \Longleftrightarrow \mathrm{Int}(F^T(V_x(T;\delta)) \cap F^{-p}(A)) \neq \emptyset$$

We can proceed as in the proof of Proposition 3.3.2 in the infinite dimensional case to get that for any $\alpha \in \mathcal{C}[T,\mathbb{Z}^d] = \{0,\ldots,\underline{p-1}\}^{[1,\ldots,T]\times\mathbb{Z}^d}$, there exists $\mathcal{A}_\alpha(x)$ defining an infinite open partition of $\mathcal{X}$ ($\cup\overline{\mathcal{A}_\alpha(x)} = \mathcal{X}$) such that F^T is injective on $\mathcal{A}_\alpha(x)$ with inverse branch F_α^{-T}.
As in Subsection 3.3.2, if $\delta < \frac{1}{2M}$ then $V_x(T;\delta) \subset \mathcal{A}_0(x)$ and $F^T(V_x(T;\delta)) = \{y \,:\, d(F^T(x),y) < \delta\}$ is a product of intervals of size 2δ around $F^T(x)$.
In the same way, F_0^{-T} is a contraction around the orbit of x:

$$d(F^{T-t}(x), F_0^{-t}(y)) \le \frac{1}{\tilde{\gamma}^t} d(F^T(x), y)$$

Then, if we construct the inverse branches of F^p around the orbit of $F^T(x)$, we know that almost all points of $\mathcal{X}$ have a pre-image by F^p at distance less than $\frac{1}{2\tilde{\gamma}^p}$ of $F^T(x)$ (because F_0^{-p} is $\frac{1}{\tilde{\gamma}^p}$ contracting). We choose then $p(\delta)$ such that $\frac{1}{\tilde{\gamma}^{p(\delta)}} < 2\delta$ and get the Specification Property. □

3.6.2 Proof of the lower bound

If $\nu \notin \mathcal{M}^1_{inv}(\mathcal{X})$

In this case $I(\nu) = +\infty$, hence there is nothing to do.

If $\nu = \sum_{l=1}^{L} a_l \nu_l$ with $\nu_l \in \mathcal{M}^1_{erg}(\mathcal{X})$, $a_l \ge 0$ and $\sum_{l=1}^{L} a_l = 1$

For $\eta > 0$, $T \ge 1$ and any $1 \le l \le L$, we define

$$\hat{R}_T^l(x) = \frac{1}{\lceil a_l T\rceil |E_T|} \sum_{\substack{0\le t<\lceil a_l T\rceil \\ i\in E_T}} \delta_{S^i\circ F^t(x)}$$

$$\Gamma_T^l = \left\{x \,:\, \hat{R}_T^l(x) \in \beta_{\nu_l}(V_1,\ldots,V_K;\delta/4) \text{ and } \int_{\mathcal{X}} \varphi\, d\hat{R}_T^l(x) \ge \int_{\mathcal{X}} \varphi\, d\nu_l - \eta\right\}$$

Then by application of the Ergodic Theorem (Theorem 3.8.1), we know that $\nu_l(\Gamma_T^l)$ goes to 1 as T tends to infinity. Hence, for a fixed $0 < b < 1$, we choose T_0 such that for any $T \geq T_0$ and any $1 \leq l \leq L$:

$$\nu_l(\Gamma_T^l) \geq b \tag{3.51}$$

Using Proposition 3.8.3, we take ε_0 and T_1 such that for all $\varepsilon < \varepsilon_0$ and $T \geq T_1$, then for $1 \leq l \leq L$:

$$\frac{1}{\lceil a_l T \rceil |E_T|} \log N^l(\lceil a_l T \rceil, E_T, \varepsilon, b) \geq h_{(F,S)}(\nu_l) - \eta \tag{3.52}$$

where N^l denotes the number of balls necessary to cover a set of ν_l measure b (see (3.60) for the precise definition).
Let now $\varepsilon < \frac{\varepsilon_0}{4}$ and $T \geq \max(T_0, T_1)$. We can then choose for $1 \leq l \leq L$ a set $S_T^l \subset \Gamma_T^l$ which is maximal $(\lceil a_l T \rceil, E_T, 4\varepsilon)$-separated in Γ_T^l. Hence, by maximality, we have

$$\Gamma_T^l \subset \bigcup_{x \in S_T^l} B_x(\lceil a_l T \rceil, E_T; 4\varepsilon)$$

and this gives, combined with estimates (3.51) and (3.52):

$$\mathrm{Card}(S_T^l) \geq \exp\left(\lceil a_l T \rceil |E_T| (h_{(F,S)}(\nu_l) - \eta)\right)$$

We use now the Specification Property (Proposition 3.6.1) to construct from these sets S_T^l a set S_T of points which are typical for ν. Indeed, for any choice of $x^1 \in S_T^1$, $x^2 \in S_T^2$, ..., $x^L \in S_T^L$, there exists a point which ε-follows the orbits of each x^l during time $\lceil a_l T \rceil$, precisely:

$$d_\rho\left(S^i \circ F^{\sum_{m=0}^{l-1} \lceil a_m T \rceil + (l-1)p(\varepsilon) + t}(x), S^i \circ F^t(x^l)\right) < \varepsilon \quad \forall\, 0 \leq t \leq \lceil a_l T \rceil,\ i \in \mathbb{Z}^d$$

Let S_T be the set of all such constructed points: as S_T^l are $(\lceil a_l T \rceil, E_T, 4\varepsilon)$-separated, then all constructed points are distinct, hence:

$$\mathrm{Card}(S_T) = \prod_{l=1}^{L} \mathrm{Card}(S_T^l) \geq \exp\left(|E_T| \sum_{l=1}^{L} \lceil a_l T \rceil (h_{(F,S)}(\nu_l) - \eta)\right)$$

And S_T is $(\hat{T}, E_T, 2\varepsilon)$-separated, with $\hat{T} = \sum_{l=1}^{L} \lceil a_l T \rceil + (L-1)p(\varepsilon)$, what implies:

$$B_x(\hat{T}, E_T; \varepsilon) \cap B_y(\hat{T}, E_T; \varepsilon) = \emptyset \quad \forall\, x \neq y \text{ in } S_T \tag{3.53}$$

We choose then ε_1 such that $d_\rho(x,y) < \varepsilon_1$ implies that $|\varphi(x) - \varphi(y)| < \eta$ and $|V_k(x) - V_k(y)| < \frac{\delta}{4}$ for all $1 \le k \le K$. For $x \in S_T$ following the orbits of $x^1 \in S_T^1$, $x^2 \in S_T^2$, $\ldots$, $x^L \in S_T^L$, for $\varepsilon < \varepsilon_1$ and $1 \le k \le K$, then

$$\begin{aligned}
&\left| \int_{\mathcal{X}} V_k \, dR_T(x) - \int_{\mathcal{X}} V_k \, d\nu \right| = \left| \int_{\mathcal{X}} V_k \, dR_T(x) - \sum_{l=1}^{L} a_l \int_{\mathcal{X}} V_k \, d\nu_l \right| \\
&\quad \le \left| \int_{\mathcal{X}} V_k \, dR_T(x) - \int_{\mathcal{X}} V_k \, d\hat{R}_T(x) \right| \\
&\qquad + \left| \int_{\mathcal{X}} V_k \, d\hat{R}_T(x) - \sum_{l=1}^{L} a_l \int_{\mathcal{X}} V_k \, d\hat{R}_T^l(F^{\lceil a_1 T\rceil + \ldots + \lceil a_{l-1} T\rceil + (l-1)p(\epsilon)}(x)) \right| \\
&\qquad + \sum_{l=1}^{L} a_l \left(\left| \int_{\mathcal{X}} V_k \, d\hat{R}_n^l(F^{\lceil a_1 T\rceil + \ldots + \lceil a_{l-1} T\rceil + (l-1)p(\epsilon)}(x)) - \int_{\mathcal{X}} V_k \, d\hat{R}_n^l(x^l) \right| \right. \\
&\qquad\qquad \left. + \left| \int_{\mathcal{X}} V_k \, d\hat{R}_n^l(x^l) - \int_{\mathcal{X}} V_k \, d\nu_l \right| \right) \\
&\quad \le 2\left(1 - \frac{T}{\hat{T}}\right) \|V_k\| + \sum_{l=1}^{L} \left(\frac{\lceil a_l T \rceil}{\hat{T}} - a_l \right) \|V_k\| + \frac{(L-1)p(\epsilon)}{\hat{T}} \|V_k\| + \frac{\delta}{4} + \frac{\delta}{4}
\end{aligned}$$

and, by the same kind of computations,

$$\int_{\mathcal{X}} \varphi \, d\hat{R}_T(x) \ge \int_{\mathcal{X}} \varphi \, d\nu - \sum_{l=1}^{L} \left(\frac{\lceil a_l T \rceil}{\hat{T}} - a_l \right) \|\varphi\| - \frac{(L-1)p(\epsilon)}{\hat{T}} \|\varphi\| - 2\eta$$

These two expressions, and the observation that $\lceil a_l T \rceil / T \to a_l$ and $\hat{T}/T \to 1$ as T tends to infinity, imply that for T great enough,

$$\int_{\mathcal{X}} \varphi \, d\hat{R}_T(x) \ge \int_{\mathcal{X}} \varphi \, d\nu - 3\eta \quad \text{and} \quad \left| \int_{\mathcal{X}} V_k \, dR_T(x) - \int_{\mathcal{X}} V_k \, d\nu \right| \le \frac{3\delta}{4}$$

The last estimate implies that if $x \in S_T$ then $R_T(x) \in \beta_\nu(V_1, \ldots, V_K; \frac{3\delta}{4})$, and also, with previous estimate on V_k:

$$B_x(\hat{T}, E_T; \varepsilon) \subset B_x(T, E_T; \varepsilon) \subset \{ y \, : \, R_T(y) \in \beta_\nu(V_1, \ldots, V_K; \delta) \}$$

We associate this with disjunction of such balls stated in (3.53), the lower bound of the Volume Lemma and estimates for the cardinal of S_T to get:

$$\begin{aligned}
&\overline{m}\{y \,:\, R_T(y) \in \beta_\nu(V_1, \ldots, V_K; \delta)\} \\
&\quad \geq \sum_{x \in S_T} \overline{m}(B_x(\hat{T}, E_T; \varepsilon)) \\
&\quad \geq \sum_{x \in S_T} C_2(\hat{T}, E_T, \varepsilon, \rho) \exp\left(\hat{T}|E_T| \int_{\mathcal{X}} \varphi \, d\hat{R}_T(x)\right) \\
&\quad \geq C_2(\hat{T}, E_T, \varepsilon, \rho) \exp\left(|E_T| \sum_{l=1}^{L} \lceil a_l T \rceil (h_{(F,S)}(\nu_l) - \eta) + \hat{T}|E_T| \int_{\mathcal{X}} \varphi \, d\nu - 3\eta\right)
\end{aligned}$$

Then:

$$\liminf_{T \to \infty} \frac{1}{T|E_T|} \log \overline{m}\{y \,:\, R_T(y) \in \beta_\nu(V_1, \ldots, V_K; \delta)\} \geq h_{(F,S)}(\nu) + \int_{\mathcal{X}} \varphi \, d\nu - 4\eta$$

because $\frac{1}{T} \sum_{l=1}^{L} \lceil a_l T \rceil h_{(F,S)}(\nu_l)$ tends to $h_{(F,S)}(\nu)$ and $\frac{\hat{T}}{T}$ to 1 as T goes to infinity. It suffices then to make η go to zero.

If $\nu \in \mathcal{M}^1_{inv}(\mathcal{X})$

We want to approximate such a probability measure by $\bar{\nu} = \sum a_l \nu_l$ from the previous case with a good control on the entropy. For this we take $\eta > 0$ and fix ε such that:

$$\operatorname{dist}_{\mathcal{M}^1(\mathcal{X})}(\tau_1, \tau_2) < \varepsilon \Rightarrow \begin{cases} |\int_{\mathcal{X}} V_k \, d\tau_1 - \int_{\mathcal{X}} V_k \, d\tau_2| < \frac{\delta}{2} & \forall\, 1 \leq k \leq K \\ |\int_{\mathcal{X}} \varphi \, d\tau_1 - \int_{\mathcal{X}} \varphi \, d\tau_2| < \eta & \end{cases}$$

We choose then $\mathcal{P} = \{P_1, \ldots, P_L\}$ a partition of $\mathcal{M}^1(\mathcal{X})$ with diameter less than ε. We know by the ergodic decomposition theorem (Theorem 2.3.3 in [66]) that there exists a probability π on $\mathcal{M}^1(\mathcal{X})$ concentrated on $\mathcal{M}^1_{\text{erg}}(\mathcal{X})$ and such that $\nu = \int_{\mathcal{M}^1(\mathcal{X})} \tau \, \pi(d\tau)$. We take, for $1 \leq l \leq L$, $a_l = \pi(P_l)$ and $\nu_l \in P_l \in \mathcal{M}^1_{\text{erg}}(\mathcal{X})$ such that $h_{(F,S)}(\nu_j) \geq h_{(F,S)}(\tau) - \eta$ for π-almost all $\tau \in P_l$. Then, with $\bar{\nu} = \sum_{l=1}^{L} a_l \nu_l$, we have:

$$\begin{aligned}
h_{(F,S)}(\bar{\nu}) &\geq h_{(F,S)}(\nu) - \eta \\
\int_{\mathcal{X}} \varphi \, d\nu &\geq \int_{\mathcal{X}} \varphi \, d\nu - \eta \\
\beta_{\bar{\nu}}(V_1, \ldots, V_K; \delta/2) &\subset \beta_\nu(V_1, \ldots, V_K; \delta)
\end{aligned}$$

This implies

$$\begin{aligned}\liminf_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(y\ :\ R_T(y)\in\beta_\nu(V_1,\dots,V_K;\delta))&\\ \geq \liminf_{T\to\infty}\frac{1}{T|E_T|}\log\overline{m}(y\ :\ R_T(y)\in\beta_{\bar\nu}(V_1,\dots,V_K;\delta/2))&\\ \geq h_{(F,S)}(\bar\nu)+\int_{\mathcal{X}}\varphi\,d\bar\nu\geq h_{(F,S)}(\nu)+\int_{\mathcal{X}}\varphi\,d\nu-2\eta&\end{aligned}$$

and we conclude making ε then η tend to 0.

3.7 Convergence of subsequences of $\mathbb{Z}^d$

We introduce in this Section two different notions of convergence for subsets of $\mathbb{Z}^d$, and their main properties.

Definition 3.7.1. *A sequence $(E_n)_{n\geq 0}$ of finite subsets of $\mathbb{Z}^d$* tends to $\mathbb{Z}^d$ in the sense of Van Hove *if $\lim_{n\to\infty}|E_n| = \infty$ and*

$$\lim_{n\to\infty}\frac{|(E_n+l)\Delta E_n|}{|E_n|} = 0 \qquad \forall\, l \in \mathbb{Z}^d \tag{3.54}$$

(where Δ denotes the symmetric difference of sets, $A\Delta B = (A\setminus B)\cup(B\setminus A)$)

If E is a finite subset of $\mathbb{Z}^d$, we define enlarged and restricted sets in $\mathbb{Z}^d$ by:

$$E^{(i)} = \begin{cases}\{j \,:\, d(j,E)\leq i\} & \text{for } i\geq 0\\ \{j \,:\, d(j,E^C) > -i\} & \text{for } i<0\end{cases} \tag{3.55}$$

We have then two properties of sequences tending to $\mathbb{Z}^d$ in the sense of Van Hove:

Proposition 3.7.1. *If $(E_n)_{n\geq 0}$ tends to $\mathbb{Z}^d$ in the sense of Van Hove, then:*

1. *For all $i\in\mathbb{Z}$, $(E_n^{(i)})_{n\geq 0}$ tends to $\mathbb{Z}^d$ in the sense of Van Hove and*

$$\lim_{n\to\infty}\frac{|E_n^{(i)}|}{|E_n|} = 1 \tag{3.56}$$

2. *For all $\tau < 1$*

$$\lim_{n\to\infty}\frac{1}{|E_n|}\sum_{l\in E_n}\tau^{d(l,E_n^C)} = 0 \tag{3.57}$$

3. *For all $\tau < 1$*

$$\lim_{n\to\infty}\frac{1}{|E_n|}\sum_{l\in E_n^C}\tau^{d(l,E_n)} = 0 \tag{3.58}$$

Proof.
1. For $i\geq 1$, we have:

$$E_n \subset E_n^{(i)} = \bigcup_{l\in\Lambda_i}(E_n+l)$$

such that $E_n^{(i)}\setminus E_n = \cup_{l\in\Lambda_i}(E_n+l)\setminus E_n$, hence:

$$1 \leq \frac{|E_n^{(i)}|}{|E_n|} = 1 + \frac{|E_n^{(i)}\setminus E_n|}{|E_n|} \leq 1 + \sum_{l\in\Lambda_i}\frac{|(E_n+l)\setminus E_n|}{|E_n|} \xrightarrow[n\to\infty]{} 1$$

by definition of the convergence in the sense of Van Hove (see Definition 3.7.1). In the same way, $(E_n^{(i)}+k)\setminus E_n^{(i)} \subset \cup_{l\in\Lambda_i}(E_n+l+k)\setminus E_n$, then

$$\frac{|(E_n^{(i)}+k)\setminus E_n^{(i)}|}{|E_n^{(i)}|} \leq \frac{|E_n|}{|E_n^{(i)}|}\sum_{l\in\Lambda_i}\frac{|(E_n+l+k)\setminus E_n|}{|E_n|} \xrightarrow[n\to\infty]{} 0$$

We proceed similarly for $E_n^{(i)}\setminus(E_n^{(i)}+k) = k+(E_n^{(i)}-k)\setminus E_n^{(i)}$, and get that $E_n^{(i)}$ tends to $\mathbb{Z}^d$ in the sense of Van Hove.

For $i\leq -1$, we have the description:

$$E_n^{(i)} = \bigcap_{l\in\Lambda_{-i}}(E_n+l)\subset E_n$$

and computations are similar to those for $i\geq 1$.

2. For any $\varepsilon>0$, we choose $k\geq 0$ such that $\sum_{i\geq k}\tau^i\leq\varepsilon/2$ and write the sum in terms of the subsets $(E_n^{(i)})_{i\leq -1}$:

$$\begin{aligned}
\frac{1}{|E_n|}\sum_{l\in E_n}\tau^{d(l,E_n^C)} &= \sum_{i\geq 1}\frac{|E_n^{(1-i)}\setminus E_n^{(-i)}|}{|E_n|}\tau^i \\
&= \sum_{i=1}^{k-1}\frac{|E_n^{(1-i)}\setminus E_n^{(-i)}|}{|E_n|}\tau^i + \sum_{i\geq k}\frac{|E_n^{(1-i)}\setminus E_n^{(-i)}|}{|E_n|}\tau^i \\
&\leq \frac{|E_n\setminus E_n^{(1-k)}|}{|E_n|}+\frac{\varepsilon}{2}
\end{aligned}$$

where we used $\tau<1$ in the first term and $|E_n^{(1-i)}\setminus E_n^{(-i)}|\leq|E_n^{(1-i)}|\leq|E_n|$ in the second. By (3.56), the first term goes to 0, hence for n great enough:

$$\frac{1}{|E_n|}\sum_{l\in E_n}\tau^{d(l,E_n^C)}\leq\varepsilon$$

3. We use in this case the fact that $\sum_{i\geq 0}|\Lambda_i|\tau^i=\sum_{i\geq 0}(2i+1)^d\tau^i$ converges. Hence, for $\varepsilon>0$, we choose $k\geq 0$ such that $\sum_{i\geq k}|\Lambda_i|\tau^i\leq\varepsilon/2$ and decompose E^C in the subsets $(E^{(i)}\setminus E^{(i-1)})_{i\geq 1}$. Then

$$\begin{aligned}
\frac{1}{|E_n|}\sum_{l\in E_n^C}\tau^{d(l,E_n)} &= \sum_{i\geq 1}\frac{|E_n^{(i)}\setminus E_n^{(i-1)}|}{|E_n|}\tau^i \\
&\leq \frac{|E_n^{(k-1)}\setminus E_n|}{|E_n|}+\frac{\varepsilon}{2}
\end{aligned}$$

since $|E_n^{(i)}\setminus E_n^{(i-1)}|\leq|E_n^{(i)}|\leq|\Lambda_i||E_n|$. We conclude then as in 2. □

Convergence in the sense of Van Hove is too wide to use some existing results of ergodic theory, in particular the Ergodic Theorem and the Theorem of Shannon-Mac Millan-Breiman (see Section 3.8.1). We need to restrict the class of subsets to get the whole large deviations results:

Definition 3.7.2. *$(E_n)_{n\geq 0}$ is a* special averaging sequence *if it is increasing, it tends to $\mathbb{Z}^d$ in the sense of Van Hove and there exists $R>0$ such that*

$$|E_n - E_n| \leq R|E_n| \qquad \forall\, n \geq 0 \tag{3.59}$$

(with $A-B=\{a-b\,:\,a\in A,\, b\in B\}$)

We will use to apply results from ergodic theory, the following straightforward result:

Proposition 3.7.2. *If $(E_T)_{T\geq 1}$ is a special averaging sequence in $\mathbb{Z}^d$, then $([0,T-1]\times E_T)_{T\geq 1}$ is a special averaging sequence in $\mathbb{N}\times\mathbb{Z}^d$.*

Remark: We could use some recent results of Lindenstrauss to work with tempered sequences, a notion more general than special averaging sequences. He proves indeed in [76] and [77] that the ergodic results we use remain valid in this context.

3.8 Thermodynamic Formalism

We gave already a presentation of main objects and results of Thermodynamic Formalism in Section 2.3.1. The main modification here is that we work in a multidimensional setup, with the dynamical system formed by F and the spatial shifts. Generalization of metric entropy and topological pressure to this setup are classical. We recall them for completeness, and give formulations of Ergodic Theorem and Shannon-Mc Millan-Breiman Theorem (proofs can be found in [85]). We give also the proof of the metric formulation of Shannon-Mc Millan-Breiman Theorem we use, because it is not written in the usual literature.

3.8.1 Ergodicity

Definition 3.8.1. *$\mathcal{M}^1_{inv}(\mathcal{X})$ denotes the set of probability measures which are* invariant *under F and all spatial shifts $(S^k)_{k\in\mathbb{Z}^d}$.*
An invariant measure ν is ergodic *if $\nu(A) = 0$ or 1 for any A invariant by (F,S). We denote $\mathcal{M}^1_{erg}(\mathcal{X})$ the set of ergodic probabilities.*

The main result for ergodic measures is the ergodic theorem, valid under some restricted assumptions on the sequence of spatial sets:

Theorem 3.8.1 (Ergodic Theorem). *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$ and $(E_T)_{T\geq 0}$ is a special averaging sequence, then for all $g \in L^1(\nu)$ and for ν-almost all x:*

$$\lim_{T\to\infty} \frac{1}{T|E_T|} \sum_{\substack{0\leq t<T\\ i\in E_T}} g\circ S^i \circ F^t(x) = \int_{\mathcal{X}} g\, d\nu$$

3.8.2 Entropy

For $\mathcal{A} = \{A_1,\ldots,A_K\}$ and $\mathcal{B} = \{B_1,\ldots,B_L\}$ finite partitions of $\mathcal{X}$, let

$$\mathcal{A}\vee\mathcal{B} = \{A_k\cap B_l \,:\, 1\leq k\leq K, 1\leq l\leq L\}$$

Then, for $\nu\in\mathcal{M}^1_{\text{inv}}(\mathcal{X})$, E_T a sequence tending to $\mathbb{Z}^d$ in the sense of Van Hove, and $\mathcal{A}$ a partition of $\mathcal{X}$, we define:

- $h(\nu|\mathcal{A}) = -\sum_{A\in\mathcal{A}} \nu(A)\log(\nu(A))$ and $\mathcal{A}_T = \bigvee_{\substack{0\leq t<T\\ i\in E_T}} F^{-t}\circ S^{-i}(\mathcal{A})$
- $h_{(F,S)}(\nu|\mathcal{A}) = \lim_{T\to\infty} \frac{1}{T|E_T|} h(\nu|\mathcal{A}_T)$
- $h_{(F,S)}(\nu) = \sup\{h_{(F,S)}(\nu|\mathcal{A}) \,:\, \mathcal{A} \text{ finite partition of } \mathcal{X}\}$

This last quantity is the metric entropy of ν under (F,S), which does not depend on the choice of the sequence $(E_T)_{T\geq 0}$.

Proposition 3.8.1.

1. *$h_{(F,S)}$ is convex affine: if $\nu = \sum_{l=1}^{L} a_l \nu_l$ with $a_l \geq 0$ and $\sum_{l=1}^{L} a_l = 1$, then:*

$$h_{(F,S)}(\nu) = \sum_{l=1}^{L} a_l h_{(F,S)}(\nu_l)$$

2. *For $\nu \in \mathcal{M}^1_{inv}(\mathcal{X})$ and for any partition $\mathcal{A}$ such that $\nu(\partial\mathcal{A}) = 0$ and $\mathrm{diam}(\mathcal{A}) < \delta_0 = \frac{1}{2M}$, we have:*

$$h_{(F,S)}(\nu) = h_{(F,S)}(\nu|\mathcal{A})$$

3. *$h_{(F,S)}$ is upper semi-continuous.*

The two last properties are consequences of the expansiveness of the system stated in Proposition 3.3.5 (see Theorem 4.5.6 in [66] and its proof).

Theorem 3.8.2 (Shannon-Mc Millan-Breiman). *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$, $\mathcal{A}$ is a finite partition and $(E_T)_{T\geq 0}$ is a special averaging sequence, then for ν-almost all x:*

$$-\frac{\log \nu(\mathcal{A}_T(x))}{T|E_T|} \xrightarrow[T\to\infty]{} h_{(F,S)}(\nu|\mathcal{A})$$

where $\mathcal{A}_T(x)$ denotes the element of the partition $\mathcal{A}_T$ which contains x.

We use in our proof of the lower bound of Large Deviations a metric equivalent of this theorem, which tells that for an ergodic measure, the metric entropy describes the number of balls necessary to cover a significant set. For $T \geq 0$, $\delta > 0$, $0 < b < 1$ and $(E_T)_{T\geq 0}$ a special averaging sequence, we denote:

$$N(T, E_T; \delta, b) = \min\left\{\mathrm{Card}(Y) \,:\, \nu\Big(\bigcup_{x\in Y} B_x(T, E_T; \delta)\Big) > b\right\} \tag{3.60}$$

(see definition of $B_x(T, E_T; \delta)$ in formula (3.10). We call a set Y as in the definition a $(T, E_T; \delta, b)$-covering set for ν)

Theorem 3.8.3. *If $\nu \in \mathcal{M}^1_{erg}(\mathcal{X})$ and $(E_T)_{T\geq 0}$ is a special averaging sequence, then for all $0 < b < 1$:*

$$\begin{aligned} h_{(F,S)}(\nu) &= \lim_{\delta\to 0} \liminf_{T\to\infty} \frac{1}{T|E_T|} \log N(T, E_T; \delta, b) \\ &= \lim_{\delta\to 0} \limsup_{T\to\infty} \frac{1}{T|E_T|} \log N(T, E_T; \delta, b) \end{aligned}$$

This result in dimension 1 is due to Katok [63]. We adapt to our multidimensional context the proof from [88].

Proof. We prove this in two steps:

1. $\lim_{\varepsilon\to 0}\limsup_{\mathbf{T}\to\infty}\frac{\mathbf{1}}{\mathbf{T}|\mathbf{E_T}|}\log\mathbf{N}(\mathbf{T},\mathbf{E_T};\varepsilon,\mathbf{b})\le\mathbf{h_{(F,S)}}(\nu)$

For $\varepsilon>0$ and $\eta>0$, let $\mathcal{A}$ be a partition with diameter (of each element) less than ε. We denote then:

$$C_T=\left\{x\in\mathcal{X}\ :\ -\frac{\log\nu(\mathcal{A}_T(x))}{T|E_T|}\le h_{(F,S)}(\nu|\mathcal{A})+\eta\right\}$$

Then, by Theorem 3.8.2, $\nu(C_T)$ tends to 1 as T goes to infinity. We choose T_0 such that $\nu(C_T)\ge b$ for all $T\ge T_0$.
We can find also a finite covering of C_T by elements of the partition $\mathcal{A}_T$: there exists $Y_T=(x_l)_{1\le l\le L}$ such that $C_T=\cup\mathcal{A}_T(x_l)$. By the choice of the diameter of $\mathcal{A}$, each $\mathcal{A}_T(x_l)$ is included in the ball $B_x(T,E_T;\varepsilon)$, and since $x_l\in C_T$, we know that

$$\nu(\mathcal{A}_T(x_l))\ge\exp(-T|E_T|(h_{(F,S)}(\nu|\mathcal{A})+\eta))$$

This implies that Y_T $(T,E_T;\varepsilon)$-covers C_T, a set of ν-measure b, and allows to estimate its cardinal L by:

$$Le^{-T|E_T|(h_{(F,S)}(\nu|\mathcal{A})+\eta)}\le\sum_{l=1}^{L}\nu(\mathcal{A}_T(x_l))=\nu(\cup_{l=1}^{L}\mathcal{A}_T(x_l))\le 1$$

Hence

$$\frac{1}{T|E_T|}\log N(T,E_T;\varepsilon,b)\le\frac{1}{T|E_T|}\log\mathrm{Card}(Y_T)\le h_{(F,S)}(\nu)+\eta$$

We take then limsup as T goes to infinity, and limits as ε, then η go to 0.

2. $\lim_{\varepsilon\to 0}\liminf_{\mathbf{T}\to\infty}\frac{\mathbf{1}}{\mathbf{T}|\mathbf{E_T}|}\log\mathbf{N}(\mathbf{T},\mathbf{E_T};\varepsilon,\mathbf{b})\ge\mathbf{h_{(F,S)}}(\nu)$

For $\eta>0$, we take $\mathcal{A}=\{A_1,\dots,A_K\}$ a finite partition such that $h_{(F,S)}(\nu|\mathcal{A})\ge h_{(F,S)}(\nu)-\eta$. We choose then another partition $\mathcal{B}=\{B_0,B_1,\dots,B_K\}$ related to $\mathcal{A}$ by:

- For $1\le k\le K$, B_k is compact and included in A_k such that $\nu(A_k\setminus B_k)\le\frac{\eta}{K\log K}$;
- $B_0=\mathcal{X}\setminus\cup_{k=1}^{K}B_k$, hence $\nu(B_0)\le\frac{\eta}{\log K}$.

We can then compare entropies of ν related to both partitions, using Lemma 3.2.15 of [66]:

$$
\begin{aligned}
& h_{(F,S)}(\nu|\mathcal{A}) \leq h_{(F,S)}(\nu|\mathcal{B}) + H_\nu(\mathcal{A}|\mathcal{B}) \\
\text{with} \quad & H_\nu(\mathcal{A}|\mathcal{B}) = -\sum_{A\in\mathcal{A}, B\in\mathcal{B}} \nu(A\cap B)\log\frac{\nu(A\cap B)}{\nu(B)} \\
& \qquad = -\nu(B_0)\sum_{k=1}^{K}\frac{\nu(A_k\cap B_0)}{\nu(B_0)}\log\frac{\nu(A_k\cap B_0)}{\nu(B_0)} \leq \nu(B_0)\log K \leq \eta
\end{aligned}
$$

by convexity of $x \to x\log x$. This gives the estimate $h_{(F,S)}(\nu|\mathcal{B}) \geq h_{(F,S)}(\nu) - 2\eta$.

We define then:

$$
C_T = \left\{ x \in \mathcal{X} \;:\; -\frac{\log\nu(\mathcal{B}_T(x))}{T|E_T|} \geq h_{(F,S)}(\nu) - 3\eta \right\}
$$

and note that, by Theorem 3.8.2, $\nu(C_T)$ goes to 1 as T tends to infinity. We choose T_0 such that $\nu(C_T) \geq 1 - \frac{b}{2}$ for all $T \geq T_0$ and $\varepsilon < \mathrm{Min}\{d(B_k, B_l) \;:\; 1 \leq k < l \leq K\}$. Then each ball $B_x(T, E_T; \varepsilon)$ intersects at most $2^{T|E_T|}$ elements of the partition $\mathcal{B}_T$, and the size of each element of this partition is controlled if we are on C_T. Hence, for any $x \in \mathcal{X}$:

$$
\nu(B_x(T, E_T; \varepsilon) \cap C_T) \leq 2^{T|E_T|} e^{-T|E_T|(h_{(F,S)}(\nu) - 3\eta)}
$$

For $T \geq T_0$ and ε as above, let Y_T be a $(T, E_T; \varepsilon, b)$-covering set for ν and $Y'_T = \{x \in Y \;:\; B_x(T, E_T, \varepsilon) \cap C_T \neq \emptyset\}$. Then:

$$
\begin{aligned}
\mathrm{Card}(Y_T) &\geq \mathrm{Card}(Y'_T) \\
&\geq \sum_{x\in Y'_T} \nu(B_x(T, E_T; \varepsilon) \cap C_T) e^{T|E_T|(h_{(F,S)}(\nu) - 3\eta - \log 2)} \\
&\geq \nu\left(\bigcup_{x\in Y} B_x(T, E_T; \varepsilon) \cap C_T\right) e^{T|E_T|(h_{(F,S)}(\nu) - 3\eta - \log 2)} \\
&\geq \left(b + 1 - \frac{b}{2} - 1\right) \exp\left(T|E_T|(h_{(F,S)}(\nu) - 3\eta - \log 2)\right)
\end{aligned}
$$

We optimize then on the covering Y_T, and take limits as T tends to infinity and ε, then η go to 0, to get:

$$
h_{F,S)}(\nu) - \log 2 \leq \lim_{\varepsilon\to 0} \liminf_{T\to\infty} \frac{1}{T|E_T|} \log N(T, E_T; \varepsilon, b)
$$

The term $\log 2$ disappears by writing the same inequality for the dynamical system (F^r, S), getting:

$$rh_{F,S)}(\nu) - \log 2 \leq r \lim_{\varepsilon\to 0} \liminf_{T\to\infty} \frac{1}{T|E_T|} \log N(T, E_T; \varepsilon, b)$$

The desired result is then obtained as r goes to infinity. ☐

3.8.3 Topological pressure

A set $Y \subset \mathcal{X}$ is $(T, E; \delta)$-separated if

$$x, x' \in Y, x \neq x' \quad \Longrightarrow \quad x' \notin B_x(T, E; \delta)$$

It is separated maximal if it is maximal for this separation property.

We define then for $V \in \mathcal{C}(\mathcal{X})$, $(E_T)_{T\geq 0}$ a sequence tending to $\mathbb{Z}^d$ in the sense of Van Hove and $Y \subset \mathcal{X}$ finite:

$$P_{(F,S)}(V; T, Y) = \log \sum_{x\in Y} \exp\Big(\sum_{\substack{0\leq t<T \\ i\in E_T}} V \circ S^i \circ F^t(x) \Big)$$

Then

$$P_{(F,S)}(V) = \lim_{\delta\to 0} \limsup_{T\to\infty} \frac{1}{T|E_T|} \sup \{P_{(F,S)}(V; T, Y)\}$$

where the supremum is taken over all sets Y which are $(T, E_T; \delta)$-separated (or, equivalently, $(T, E_T; \delta)$-separated maximal), is the topological pressure of V for the dynamic of (F, S). This definition is independent of the choice of the sequence (E_T). The main result for this quantity is the Gibbs Variational Principle, which expresses it as a variational expression of the entropy:

Theorem 3.8.4 (Gibbs Variational Principle). *For any $V \in \mathcal{C}(\mathcal{X})$:*

$$P_{(F,S)}(V) = \sup_{\nu\in\mathcal{M}^1_{inv}(\mathcal{X})} \left(h_{(F,S)}(\nu) + \int_{\mathcal{X}} V\, d\nu \right) \tag{3.61}$$

and, since $h_{(F,S)}$ is convex affine and upper semi-continuous in our case, for any $\nu \in \mathcal{M}^1_{inv}(\mathcal{X})$:

$$h_{(F,S)}(\nu) = \inf_{V\in\mathcal{C}(\mathcal{X})} \left(P_{(F,S)}(V) - \int_{\mathcal{X}} V\, d\nu \right) \tag{3.62}$$

3.9 Generating function method for the iteration sequence

For $\delta > 0$, $\gamma > 1$ and (α_k) a sequence of non-negative reals, let $u(i,t)$ be defined for $i \in \mathbb{Z}$ and $t \in \mathbb{N}$ by:

$$u(i,t) = \begin{cases} \frac{1}{2} & \text{if } i < 0 \\ \delta & \text{if } i \geq 0, t = 0 \\ \frac{1}{\gamma}u(i,t-1) + \frac{1}{\gamma}\sum_{k\geq 0}\alpha_k u(i-k,t) & \text{if } i \geq 0, t > 0 \end{cases} \tag{3.63}$$

We have then for such a sequence:

Proposition 3.9.1. *Suppose there exists $\theta < 1$ such that for any $k \geq 0$, $\alpha_k = \theta^k \tilde{\alpha}_k$ and denote $S = \sum_{k\geq 0}\alpha_k$ and $\tilde{S} = \sum_{k\geq 0}\tilde{\alpha}_k$. Then, under the assumption*

$$\gamma - \tilde{S} > 1$$

we have for all $i \geq 0$ and $t \geq 0$:

$$u(i,t) \leq \frac{\delta}{(\gamma - S)^t} + \theta^{i+1}\frac{\tilde{S}}{2(\gamma - \tilde{S} - 1)} \tag{3.64}$$

Proof. We solve this equation by a generating function method (see [104] for a general introduction and many useful tools). Let $f(x,y)$ be the formal series defined by:

$$f(x,y) = \sum_{\substack{i\geq 0\\ t\geq 1}} u(i,t)x^i y^t$$

Then the inductive definition of $u(i,t)$ implies for f:

$$\begin{aligned} f(x,y) &= \sum_{\substack{i\geq 0\\ t\geq 1}} \Big(\frac{1}{\gamma}u(i,t-1) + \frac{1}{\gamma}\sum_{k\geq 0}\alpha_k u(i-k,t)\Big)x^i y^t \\ &= \frac{\delta y}{\gamma}\sum_{i\geq 0}x^i + \frac{y}{\gamma}\sum_{\substack{i\geq 0\\ t\geq 1}} u(i,t)x^i y^t + \frac{1}{\gamma}\sum_{\substack{i\geq 0\\ t\geq 1}}\Big(\sum_{k=0}^{i}\alpha_k u(i-k,t)x^i\Big)y^t \\ &\qquad + \frac{1}{2\gamma}\sum_{\substack{i\geq 0\\ t\geq 1}}\Big(\sum_{k>i}\alpha_k\Big)x^i y^t \\ &= \frac{\delta y}{\gamma}\sum_{i\geq 0}x^i + \frac{1}{2\gamma}\sum_{\substack{i\geq 0\\ t\geq 1}} R_i x^i y^t + \frac{1}{\gamma}\Big(y + \sum_{k\geq 0}\alpha_k x^k\Big)f(x,y) \\ &= \Big(\frac{\delta y}{\gamma}\sum_{i\geq 0}x^i + \frac{1}{2\gamma}\sum_{\substack{i\geq 0\\ t\geq 1}} R_i x^i y^t\Big)\Big(1 - \frac{1}{\gamma}\Big(y + \sum_{k\geq 0}\alpha_k x^k\Big)\Big)^{-1} \end{aligned}$$

where $R_i = \sum_{k>i} \alpha_k$. We invert formally this expression, using that:

$$\begin{aligned}
\Big(1 - \frac{1}{\gamma}\Big(y + \sum_{k\geq 0} \alpha_k x^k\Big)\Big)^{-1} &= \sum_{n\geq 0} \sum_{u=0}^{n} \binom{n}{u} \frac{1}{\gamma^n} y^u \Big(\sum_{k\geq 0} \alpha_k x^k\Big)^{n-u} \\
&= \sum_{\substack{u\geq 0\\ l\geq 0}} \binom{u+l}{u} \frac{1}{\gamma^{u+l}} y^u \sum_{k_1,\ldots,k_l\geq 0} \alpha_{k_1}\cdots\alpha_{k_l} x^{k_1+\cdots+k_l} \\
&= \sum_{\substack{n\geq 0\\ u\geq 0}} \Big(\sum_{l\geq 0} \binom{u+l}{u} \frac{1}{\gamma^{u+l}} \sum_{\substack{k_1,\ldots,k_l\geq 0\\ k_1+\cdots+k_l=n}} \alpha_{k_1}\cdots\alpha_{k_l}\Big) x^n y^u
\end{aligned}$$

Hence, using in the upper bound that $R_{i-n} \leq \theta^{i-n+1}\tilde{S}$, we get:

$$\begin{aligned}
u(i,t) &= \frac{\delta}{\gamma} \sum_{n=0}^{i} \Big(\sum_{l\geq 0} \binom{t-1+l}{t-1} \frac{1}{\gamma^{t-1+l}} \sum_{\substack{k_1,\ldots,k_l\geq 0\\ k_1+\cdots+k_l=n}} \alpha_{k_1}\cdots\alpha_{k_l}\Big) \\
&\qquad + \frac{1}{2\gamma} \sum_{\substack{0\leq n\leq i\\ 0\leq u<t}} R_{i-n}\Big(\sum_{l\geq 0} \binom{u+l}{u} \frac{1}{\gamma^{u+l}} \sum_{\substack{k_1,\ldots,k_l\geq 0\\ k_1+\cdots+k_l=n}} \alpha_{k_1}\cdots\alpha_{k_l}\Big) \\
&\leq \frac{\delta}{\gamma^t} \sum_{l\geq 0} \binom{t-1+l}{t-1} \Big(\frac{S}{\gamma}\Big)^l + \frac{\theta^{i+1}}{2\gamma} \sum_{u\geq 0} \frac{\tilde{S}}{\gamma^u} \sum_{l\geq 0} \binom{u+l}{u} \Big(\frac{\tilde{S}}{\gamma}\Big)^l \\
&= \frac{\delta}{(\gamma - S)^t} + \theta^{i+1} \frac{\tilde{S}}{2(\gamma - \tilde{S} - 1)}
\end{aligned}$$

□

Remark: We obtained in fact in the course of the proof an exact (but complicated) expression for the sequence $u_{(i,t)}$.

4. LIMIT THEOREMS FOR COUPLED ANALYTIC MAPS

We work in this Chapter[1] with coupled map lattices with local expanding maps on the circle and weak coupling. We need to assume strong regularity: the local maps (resp. the coupling) have to be holomorphic in a neighborhood of the circle (resp. the product of circles).

We derive in this setup new limit theorems for the asymptotic behavior of the temporal empirical measure associated to the system: for a large class of observables, we obtain in Theorem 4.2.3 a Central Limit Theorem and Moderate Deviations Principle.

We obtain also a partial Large Deviations Principle, see Theorem 4.2.2. This partial result implies in particular exponential convergence to equilibrium. Indeed, our results follow from methods of [95], where (in the same vein as Bricmont and Kupiainen [12] or Baladi et al [4]) Rugh proves the existence and uniqueness in a restricted class of an invariant locally absolutely continuous measure.

Rugh uses a sharp combinatorial analysis of the finite system operators to construct a transfer operator associated to the infinite system satisfying a spectral gap property on an adequate Banach space.
Our proof consists in an adaptation of this construction to perturbed transfer operators and a perturbation argument to preserve the spectral gap property.

We define the model in Section 4.1 and give the results in Section 4.2. We then prove the probabilistic results in Section 4.3. The method is similar to the proofs of [92] or [15]. Proof of an intermediate result on the existence of transfer operators and their spectral properties is given in Sections 4.4 and 4.5.

---

[1] This Chapter corresponds to an article with same title, to appear in Probability Theory and Related Fields.

4.1 Definitions

4.1.1 Expanding maps

We consider $S^1 = \mathbb{R}/\mathbb{Z}$ as a subset of the complex cylinder $\mathcal{C} = \mathbb{C}/\mathbb{Z}$. This allows us to work with functions not only real-analytic on the circle but holomorphic on a small annulus $A[\rho] = \{z \in \mathcal{C} : |\text{Im } z| \leq \rho\}$ for $\rho > 0$. For such functions, we are able to use complex analysis and this is the basis of the method introduced in [41] and [95] to construct transfer operators.

Thus, the single-site functions we will use are real-analytic expanding functions on the circle in the following sense:

Definition 4.1.1. *For $\rho > 0$ and $\lambda > 1$, we say that $f : A[\rho] \to \mathcal{C}$ is a* real analytic (ρ, λ)-expanding map *if f is continuous in $A[\rho]$, holomorphic in its interior, $f(S^1) = S^1$ and $f(\partial A[\rho]) \cap A[\lambda\rho] = \emptyset$.*
The set of all such functions is denoted $\mathcal{E}(\rho, \lambda)$.

Remark: Functions of $\mathcal{E}(\rho, \lambda)$ are also λ-expanding in the classical sense, i.e. they verify $|f'| \geq \lambda > 1$ on the circle (see Appendix A in [95]).

4.1.2 Configuration space

We take Ω an index set and define the configuration space of our dynamical system as the product of circles :

$$S_\Omega = \prod_{p \in \Omega} S^1 \subset A_\Omega = \prod_{p \in \Omega} A[\rho]$$

Ω can be quite general and could even be uncountable. But our main interest will be $\Omega = \mathbb{Z}^d$. For this case, some spatial behavior can be studied (see [95] or [5] for such applications).

4.1.3 Spaces of coupling and observables

Let $\mathcal{F}$ be the set of finite subsets of Ω, containing the empty set. For all $\Lambda \in \mathcal{F}$, we denote $S_\Lambda = \prod_{p \in \Lambda} S^1 \subset A_\Lambda = \prod_{p \in \Lambda} A[\rho]$. We call E_Λ the set of functions which are continuous in A_Λ and holomorphic in its interior.
For $K \subset \Lambda$, we denote $j_{\Lambda,K} : E_K \to E_\Lambda$ and $j_\Lambda : E_\Lambda \to \mathcal{C}(A_\Omega)$ the natural injections, then define $E(A_\Omega)$ as the closure of $\cup_{\Lambda \in \mathcal{F}} j_\Lambda(E_\Lambda)$. $E(A_\Omega)$ is in fact the space of weakly holomorphic continuous functions on A_Ω (see Appendix B of [95]).

We want to control the spatial expansion of the functions which will play the role of coupling and observables. For this, we choose a parameter $0 < \theta \leq 1$ and define:

$$H_\theta = \left\{ \phi \in E(A_\Omega) \, : \, \phi = \sum_{\Lambda \in \mathcal{F}} j_\Lambda \phi_\Lambda \text{ with } \phi_\Lambda \in E_\Lambda \text{ and } \sum_{\Lambda \in \mathcal{F}} \theta^{-|\Lambda|} |\phi_\Lambda| < \infty \right\}$$

with, for $\phi \in H_\theta$:

$$|\phi|_\theta = \inf \left\{ \sum_{\Lambda \in \mathcal{F}} \theta^{-|\Lambda|} |\phi_\Lambda| \, : \, (\phi_\Lambda)_{\Lambda \in \mathcal{F}} \text{ such that } \phi_\Lambda \in E_\Lambda \text{ and } \phi = \sum_{\Lambda \in \mathcal{F}} j_\Lambda \phi_\Lambda \right\}$$

Then $(H_\theta, |\cdot|_\theta)$ is a θ-penalized inductive limit of the spaces E_Λ. This defines a Banach algebra. If $\theta < 1$, functions of H_θ depend weakly of big sets Λ. For $\theta = 1$, $H_1 = E(A_\Omega)$ and $|\cdot|_\theta = |\cdot|_\infty$. We denote H_θ^r the set of real-analytic maps of H_θ.

4.1.4 Coupled maps

We can now define the class of dynamical systems we want to study:

Definition 4.1.2. *For $\rho > 0$, $\lambda > 1$, $0 < \theta \leq 1$ and $0 \leq \kappa < \infty$, we take $(f_p)_{p \in \Omega}$ expanding maps from $\mathcal{E}(\rho, \lambda)$, and $(g_p)_{p \in \Omega}$ coupling maps from H_θ^r such that $|g_p|_\theta < \kappa$.*
We define the associated coupled analytic map *as $F_\Omega = (F_p)_{p \in \Omega} \, : \, A_\Omega \to \mathcal{C}^\Omega$, where :*

$$F_p(z) = f_p(z_p) + g_p(z) \quad \forall p \in \Omega$$

We denote $CM[\rho, \lambda, \theta, \kappa]$ the space of all such coupled analytic maps.

4.2 Results

For all observable $b \in \mathcal{C}(S_\Omega)$ and all $T \geq 1$, we write:

$$S_T b = \sum_{t=0}^{T-1} b \circ F^t$$

In [95], an Ergodic Theorem for the random variables $S_T b$ under Lebesgue measure and decay of correlations for the limit measure are proved under the assumption that the coupling is weak enough:

Theorem 4.2.1 (Th. 2.1 of [95]). *For every $\rho > 0$, $\lambda > 1$, there exists $\theta_0(\rho, \lambda) \in (0, \frac{1}{3})$ such that for $\theta < \theta_0$ there is $\kappa > 0$ for which the following holds for all $F \in CM[\rho, \lambda, \theta, \kappa]$:*

1. *There exists a natural probability measure ν invariant under F, i.e. $F^*\nu = \nu$,*
2. *For all $b \in \mathcal{C}(S_\Omega)$, m^Ω-almost every x (with m the Lebesgue measure on the circle),*
$$\lim_{T\to\infty} \frac{1}{T} S_T b = \int_{S_\Omega} b \, d\nu \tag{4.1}$$
3. *There exists $\gamma > 1$ and $\theta < \vartheta < 1$ such that for all $b \in H_\theta$, $a \in H_\vartheta$ and $T \geq 1$,*
$$\left| \int_{S_\Omega} b \circ F^T \cdot a \, d\nu - \int_{S_\Omega} b \, d\nu \int_{S_\Omega} a \, d\nu \right| \leq 2 |b|_\theta |a|_\vartheta \gamma^{-T} \tag{4.2}$$

These properties are consequences of a more technical result, the fact that a transfer operator associated to F exists on a well chosen Banach space and has a spectral gap below 1, which is the simple maximal eigenvalue. They are really an infinite dimensional version of classical single site results.

Our method consists in generalizing the construction of this operator to its perturbations by potentials and then extending the spectral gap by perturbation theory (see Theorem 4.3.1 and its proof Section 4.4 for more details).

We improve the result of [95] with the following large deviations upper bound, and an associated partial lower bound (see Theorem 4.3.2 for a more precise statement):

Theorem 4.2.2. *Under the same conditions on the parameters as in Theorem 4.2.1, for all $u \in H_\theta^r$, there exists a lower semi-continuous, convex and non-negative function $I_u : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, with a unique zero at $\int_{S_\Omega} u \, d\nu$, and there are $a_u < \int_{S_\Omega} u \, d\nu < b_u$ such that:*

1. *For all closed $F \subset \mathbb{R}$:*
$$\limsup_{T\to\infty} \frac{1}{T} \log m^{\Omega}\left(z\ :\ \frac{S_T u(z)}{T} \in F\right) \leq -\inf_{x\in F} I_u(x) \quad \text{(Upper Bound)}$$
2. *For all $x \in (a_u, b_u)$ and $\delta > 0$:*
$$\liminf_{T\to\infty} \frac{1}{T} \log m^{\Omega}\left(z\ :\ \frac{S_T u(z)}{T} \in B(x,\delta)\right) \geq -I_u(x) \quad \text{(Lower Bound)}$$

This implies in particular that the convergence in (4.1) is exponential, which means that for all $A \in \mathbb{R}$ such that $\int_{S_\Omega} u\,d\nu \notin \bar{A}$:

$$\limsup_{T\to\infty} \frac{1}{T} \log m^{\Omega}\left\{z\ :\ \frac{S_T u(z)}{T} \in A\right\} < 0 \tag{4.3}$$

Moreover, we obtain new probabilistic results for the random variables $S_T u$ under Lebesgue measure, namely a Central Limit Theorem and a Moderate Deviations Principle:

Theorem 4.2.3. *Suppose the hypotheses of Theorem 4.2.1 are satisfied. For every $u \in H^r_\theta$, we write $m_u = \int_{S_\Omega} u\,d\nu$. Then the limit*

$$\lim_{T\to\infty} \int_{S_\Omega} \left(\frac{S_T u - T m_u}{\sqrt{T}}\right)^2 d\nu$$

exists and is non negative. We denote it σ_u^2 and have the following condition:

$$\sigma_u^2 = 0 \qquad \textit{iff} \qquad \exists\, v \in L^2(\nu) \ \textit{ such that } \ u = v - v\circ F \ \textit{ in } L^2(\nu) \tag{4.4}$$

For u such that $\sigma_u^2 > 0$, we have:

$$\left(\frac{S_T u - T m_u}{\sqrt{T}\sigma_u}\right)^{*} (m^{\Omega}) \xrightarrow{Law} \mathcal{N}(0,1) \tag{CLT}$$

and for all $\frac{1}{2} < \alpha < 1$, $A \subset \mathbb{R}$ Borel set:

$$\begin{aligned} -\inf_{x\in \mathring{A}} \frac{x^2}{2\sigma_u^2} &\leq \liminf_{T\to\infty} \frac{1}{T^{2\alpha-1}} \log m^{\Omega}\left(z\ :\ \frac{S_T u(z) - T m_u}{T^{\alpha}} \in A\right) \\ &\leq \limsup_{T\to\infty} \frac{1}{T^{2\alpha-1}} \log m^{\Omega}\left(z\ :\ \frac{S_T u(z) - T m_u}{T^{\alpha}} \in A\right) \leq -\inf_{x\in\bar{A}} \frac{x^2}{2\sigma_u^2} \end{aligned} \tag{MDP}$$

Remark. All results above are given with Lebesgue measure as initial probability. In fact, they remain true taking measures in the Banach space on which our operators act (exactly on the subset of this Banach space which contains probabilities, denoted $\mathcal{M}^p_\vartheta$, see Section 4.3.1). We will prove our results in this more general context. The same generalization for the ergodic theorem (4.1) is valid and the proof of [95] adapted in a simple way.

4.3 Use of the spectral gap

In this section, we will prove Theorems 4.2.2 and 4.2.3 given an intermediate result (Theorem 4.3.1) on the spectral gap for perturbed operators. We use in these proofs the same type of methods as in the papers of J. Rousseau-Egele [92] or A. Broise [15].

4.3.1 Space of densities

For $K \subset \Lambda$, let $\pi_{K,\Lambda} : E_\Lambda \to E_K$ be the projection defined by:

$$\pi_{K,\Lambda}\phi_\Lambda(z_K) = \int_{S_{\Lambda\setminus K}} \phi_\Lambda(z_\Lambda)\, m^{\Lambda\setminus K}(dz_{\Lambda\setminus K})$$

If $\Lambda = \Omega$, we will note $\pi_K = \pi_{K,\Omega}$.

Following [95], we define now the Banach space on which our operators work. We need to take it sufficiently large, and specifically not included in $L^1(dm^\Omega)$. Indeed, in the uncoupled case (when the couplings g_p are zero), we know that the natural measure will be the infinite product of the SRB measures $h_p dm$ for the single site functions f_p, which will generally not be absolutely continuous with respect to Lebesgue measure. To get a large enough space, we choose a parameter $0 < \theta \leq 1$ and define:

$$\mathcal{M}_\theta = \left\{ \phi = (\phi_\Lambda)_{\Lambda\in\mathcal{F}} \,:\, \pi_{\Lambda,\Lambda'}\phi_{\Lambda'} = \phi_\Lambda \;\forall \Lambda \subset \Lambda' \text{ and } \|\phi\|_\theta = \sup_{\Lambda\in\mathcal{F}} \theta^{|\Lambda|}|\phi_\Lambda| < \infty \right\}$$

$(\mathcal{M}_\theta, \|\cdot\|_\theta)$ is a Banach space and a H_θ-module: $g = \sum_{\Lambda'\in\mathcal{F}} g_{\Lambda'}$ element of H_θ acts on $\phi = (\phi_\Lambda)_{\Lambda\in\mathcal{F}}$ to get $g * \phi \in \mathcal{M}_\theta$ defined by:

$$(g*\phi)_\Lambda = \sum_{\Lambda'\in\mathcal{F}} \pi_{\Lambda,\Lambda\cup\Lambda'}\left(j_{\Lambda\cup\Lambda',\Lambda'}(g_{\Lambda'})\cdot\phi_{\Lambda\cup\Lambda'}\right)$$

and the following bound holds: $\|g*\phi\|_\theta \leq |g|_\theta \|\phi\|_\theta$.
As soon as $\theta^{-1} > \sup_{p\in\Omega} |h_p|_\infty$, $\mathcal{M}_\theta$ contains the uncoupled natural measure $\otimes_{p\in\Omega}(h_p dm)$. This measure is represented by $\phi = \left(\phi_\Lambda = \prod_{p\in\Lambda} h_p(z_p)\right)_{\Lambda\in\mathcal{F}}$, although it is not absolutely continuous with respect to Lebesgue measure. More generally, if we consider the following subset of $\mathcal{M}_\theta$:

$$\mathcal{M}_\theta^m = \left\{ \phi \in \mathcal{M}_\theta \,:\, \sup_{\Lambda\in\mathcal{F}} \int_{S_\Lambda} |\phi_\Lambda(z_\Lambda)|\, dz_\Lambda < \infty \right\}$$

then every $\phi \in \mathcal{M}_\theta^m$ can be seen as a measure on S_Ω defined by

$$\int_{S_\Omega} g\, d\phi = \phi(g) = \lim_{\Lambda\to\Omega} \int_{S_\Lambda} g_\Lambda \phi_\Lambda dm^\Lambda$$

for every $g \in \mathcal{C}(S_\Omega)$ and $g_\Lambda \in \mathcal{C}(S_\Lambda)$ such that $g_\Lambda \to g$.
All these measures have finite marginals on S_Λ which are absolutely continuous with respect to m^Λ, with density $\phi_\Lambda \in E(A_\Lambda)$. We will denote $\mathcal{M}_\theta^p$ the set of probability measures in $\mathcal{M}_\theta^m$.

4.3.2 Spectral gap for perturbed transfer operators

We state now the existence and the property of spectral gap for perturbed transfer operators:

Theorem 4.3.1. *For $F \in CM[\rho, \Lambda, \theta, \kappa]$, whose parameters satisfy conditions of Theorem 4.2.1 and with ϑ, γ and ν as in this result, there exists for all $T \geq 1$ an analytic functional:*

$$\begin{aligned} M^{(T)} \,:\, H_\theta &\longrightarrow L(\mathcal{M}_\vartheta, \mathcal{M}_\theta) \\ u &\longmapsto M_u^{(T)} \end{aligned} \tag{4.5}$$

satisfying:

- *There exists $T_0 \geq 1$ such that $M_u^{(T)} \in L(\mathcal{M}_\vartheta)$ if $T \geq T_0$*
- $\|M_u^{(T)}\| \leq e^{T|u|_\theta}$ • $\|M_u^{(T)} - M_0^{(T)}\| \leq e^{T|u|_\theta} - 1$ (4.6)
- $M_u^{(t)} \circ M_u^{(T)} = M_u^{(t+T)}$ *for $t \geq 1, T \geq T_0$* (4.7)
- $M_u^{(T)}\left(\mathcal{M}_\vartheta^m\right) \subset \mathcal{M}_\theta^m$ (4.8)
- $\displaystyle\int_{S_\Omega} b \circ F^T \exp\left(S_T u\right)\, d\phi = \int_{S_\Omega} b\, d\left(M_u^{(T)}\phi\right) \quad \forall\, b \in \mathcal{C}(S_\Omega),\ \phi \in \mathcal{M}_\vartheta^m$ (4.9)

Moreover, for all $\delta < \frac{1-\gamma^{-T_0}}{3}$, there exists $\rho > 0$ such that if $|u|_\theta < \rho$, we can write for $k \geq 1$:

$$M_u^{(kT_0)} = \lambda^{kT_0}(u)\, Q_u + R_u^k \tag{4.10}$$

with, for $D_\theta(0, \rho)$ the ball of radius ρ around 0 in H_θ:

- *$\lambda \,:\, u \in D_\theta(0,\rho) \longmapsto \lambda(u) \in \mathbb{C}$ is analytic and satisfies $\lambda^{T_0}(u) \in D(1,\delta)$ and $\lambda(0) = 1$,*
- *$Q \,:\, u \in D_\theta(0,\rho) \longmapsto Q_u \in L(\mathcal{M}_\vartheta)$ is analytic and satisfies $Q_u^2 = Q_u$, $Q_0 = \nu\pi_\emptyset$ and $\|Q_u - \nu\pi_\emptyset\| \leq \delta^2$,*
- *$R \,:\, u \in D_\theta(0,\rho) \longmapsto R_u \in L(\mathcal{M}_\vartheta)$ is analytic and satisfies $Sp(R_u) \subset D(0, \gamma^{-T_0} + \delta)$ and $\|R_u^k\| \leq (\gamma^{-T_0} + 2\delta)^k$.*

Remark. The important fact in these estimates is that they imply for such u:

$$\lim_{k\to\infty} \frac{\|R_u^k\|}{|\lambda^{kT_0}(u)|} = 0$$

so that $\lambda(u)$ will give the main term in asymptotic estimates.

4.3.3 Identification of the derivatives of $\lambda(u)$

Analyticity in the previous result is understood in the general sense given for example in Definition 3.17.2 of [46]: namely a map is analytic when it is expandable around each point as a convergent series of homogeneous terms with increasing degree. For λ, an analytic function of u on $D_\theta(0,\rho)$, we can write its expansion around 0:

$$\lambda(u) = \sum_{n\geq 0} \frac{1}{n!}\partial^n\lambda(0;u)$$

where in fact $\partial^0\lambda(0;u) = \lambda(0) = 1$ and $\partial^n\lambda(0;u) = \frac{\partial^n}{\partial z^n}\big|_{z=0}\lambda(zu)$.

The key of our probabilistic study is the identification of the first two derivatives of λ in real-analytic directions with statistical estimates of the system.

Proposition 4.3.1. *For every $u \in H^r_\theta$, we have the two following identities:*

$$\partial^1\lambda(0;u) = \int_{S_\Omega} u\, d\nu = m_u \qquad \partial^2\lambda(0;u) = \lim_{T\to\infty}\int_{S_\Omega}\left(\frac{S_T u - T m_u}{\sqrt{T}}\right)^2 d\nu = \sigma_u^2$$

Remark: The identifications of $\partial^1\lambda(0;u)$ and $\partial^2\lambda(0;u)$ with the mean and the asymptotic variance of u under the equilibrium state ν are natural results in view of classical thermodynamic formalism results (see [93]): $\lambda(u)$, in the domain where it is defined, really plays the role of a topological pressure.

Proof.

- *Identification of $\partial^1\lambda(0;u)$.* We will decompose each $T\geq 1$ as $T = kT_0 + \tilde{T}$, with $0 \leq \tilde{T} < T_0$, and write:

$$\int_{S_\Omega} \exp\left(\frac{1}{T}S_T u\right) d\nu = \int_{S_\Omega} \exp\left(S_{kT_0}\left(\frac{u}{T}\right) + \frac{1}{T}S_{\tilde{T}}\left(u\circ F^{kT_0}\right)\right) d\nu$$

We have then a uniform estimate for the term with $\tilde{T}$:

$$\exp\left(-\frac{T_0}{T}|u|_\infty\right) \leq \exp\left(\frac{1}{T}S_{\tilde{T}}\left(u\circ F^{kT_0}\right)\right) \leq \exp\left(\frac{T_0}{T}|u|_\infty\right) \tag{4.11}$$

For the remaining term, if $T > \frac{|u|_\theta}{\rho}$, we apply the identity (4.9) and the spectral decomposition (4.10) to $M^{(kT_0)}_{\frac{u}{T}}$ to get:

$$\begin{aligned}\int_{S_\Omega} \exp\left(\frac{1}{T}S_{kT_0}u\right) d\nu &= \pi_\emptyset\left(M^{(kT_0)}_{\frac{u}{T}}(\nu)\right)\\ &= \lambda^{kT_0}\left(\frac{u}{T}\right)\pi_\emptyset\left(Q_{\frac{u}{T}}(\nu)\right) + \pi_\emptyset\left(R^k_{\frac{u}{T}}(\nu)\right)\end{aligned} \tag{4.12}$$

We can now evaluate the limit as T tends to infinity of each term in this expression:

$$\lambda^{kT_0}\left(\frac{u}{T}\right) = \left(1 + \frac{1}{T}\partial^1\lambda(0;u) + o\left(\frac{1}{T}\right)\right)^{kT_0} \longrightarrow \exp\left(\partial^1\lambda(0;u)\right)$$

because the derivatives $\partial^n\lambda(0;u)$ are n-homogeneous and $\frac{kT_0}{T} \to 1$. It will be the main term in our estimation.

We control the two others:

$$\left|\pi_\emptyset\left(Q_{\frac{u}{T}}(\nu)\right) - 1\right| \leq \left\|Q_{\frac{u}{T}} - Q_0\right\| \|\nu\|_\vartheta \longrightarrow 0$$

by continuity of Q_u, and:

$$\left|\pi_\emptyset\left(R^k_{\frac{u}{T}}(\nu)\right)\right| \leq \left\|R_{\frac{u}{T}}\right\|^k \|\nu\|_\vartheta \leq \left(\gamma^{-T_0} + 2\delta\right)^k \|\nu\|_\vartheta \longrightarrow 0$$

We get, using estimate (4.11) and inserting previous limits in (4.12):

$$\lim_{T\to\infty}\int_{S_\Omega} \exp\left(\frac{1}{T}S_T u\right)\, d\nu = \lim_{T\to\infty} \pi_\emptyset\left(M^{(kT_0)}_{\frac{u}{T}}(\nu)\right) = \exp\left(\partial^1\lambda(0;u)\right) \tag{4.13}$$

On the other hand, (4.2) implies that ν is mixing, hence ergodic, which gives the limit, because u is bounded:

$$\lim_{T\to\infty}\int_{S_\Omega} \exp\left(\frac{1}{T}S_T u\right)\, d\nu = \exp\left(\int_{S_\Omega} u\, d\nu\right) \tag{4.14}$$

We can identify both RHS in (4.13) and (4.14) to get:

$$\partial^1\lambda(0;u) = \int_{S_\Omega} u\, d\nu$$

- *Identification of* $\partial^2\lambda(0;u)$. It is enough to show that for $u \in H^r_\theta$ such that $\partial^1\lambda(0;u) = \int_{S_\Omega} u\, d\nu = 0$, we have:

$$\lim_{T\to\infty}\int_{S_\Omega}\left(\frac{S_T u}{\sqrt{T}}\right)^2 d\nu = \partial^2\lambda(0;u)$$

For this, u being bounded, we know that we can write:

$$\int_{S_\Omega}\left(\frac{S_T u}{\sqrt{T}}\right)^2 d\nu = \left.\frac{\partial^2}{\partial t^2}\right|_{t=0}\int_{S_\Omega}\exp\left(\frac{t}{\sqrt{T}}S_T u\right)\, d\nu = \left.\frac{\partial^2}{\partial t^2}\right|_{t=0}\pi_\emptyset\left(M^{(T)}_{\frac{tu}{\sqrt{T}}}(\nu)\right) \tag{4.15}$$

For $T > \left(\frac{|t||u|_\theta}{\rho}\right)^2$, we write again $T = kT_0 + \tilde{T}$ with $0 \le \tilde{T} < T_0$ and use the composition rule (4.7) and the spectral decomposition (4.10) to get:

$$\pi_\emptyset\left(M^{(T)}_{\frac{tu}{\sqrt{T}}}(\nu)\right) = \lambda^{kT_0}\left(\frac{tu}{\sqrt{T}}\right)\pi_\emptyset\left(M^{(\tilde{T})}_{\frac{tu}{\sqrt{T}}} \circ Q_{\frac{tu}{\sqrt{T}}}(\nu)\right) + \pi_\emptyset\left(M^{(\tilde{T})}_{\frac{tu}{\sqrt{T}}} \circ R^k_{\frac{tu}{\sqrt{T}}}(\nu)\right)$$

We want to derive twice this last expression:

$$\left.\frac{\partial^2}{\partial t^2}\right|_{t=0}\left(\lambda^{kT_0}\left(\frac{tu}{\sqrt{T}}\right)\pi_\emptyset\left(M^{(\tilde{T})}_{\frac{tu}{\sqrt{T}}} \circ Q_{\frac{tu}{\sqrt{T}}}(\nu)\right)\right)$$
$$= (vw)''(0) = v''(0)w(0) + 2v'(0)w'(0) + v(0)w''(0)$$

with $v(t) = \lambda^{kT_0}\left(\frac{tu}{\sqrt{T}}\right)$, so that $v(0) = 1$, $v'(0) = 0$, and $v''(0) = \frac{kT_0}{T}\partial^2\lambda(0;u)$, and $w(t) = \pi_\emptyset\left(M^{(\tilde{T})}_{\frac{tu}{\sqrt{T}}} \circ Q_{\frac{tu}{\sqrt{T}}}(\nu)\right)$, so that $w(0) = 1$, and

$$w''(0) = \frac{1}{T}\pi_\emptyset\Big(M^{(\tilde{T})}_0 \circ \partial^2 Q(0;u) + 2\partial^1 M^{(\tilde{T})}(0;u) \circ \partial^1 Q(0;u)$$
$$+ \partial^2 M^{(\tilde{T})}(0;u) \circ Q_0\Big)(\nu)$$

which goes to zero when T goes to infinity.
In the same way

$$\left.\frac{\partial^2}{\partial t^2}\right|_{t=0}\left(\pi_\emptyset\left(M^{(\tilde{T})}_{\frac{tu}{\sqrt{T}}} \circ R^k_{\frac{tu}{\sqrt{T}}}(\nu)\right)\right) = \frac{1}{T}\pi_\emptyset\Big(M^{(\tilde{T})}_0 \circ \partial^2 R^k(0;u)$$
$$+ 2\partial^1 M^{(\tilde{T})}(0;u) \circ \partial^1 R^k(0;u) + \partial^2 M^{(\tilde{T})}(0;u) \circ R^k_0\Big)(\nu)$$

which goes to zero when T goes to infinity since $\lim_{k\to\infty} R^k = 0$.

Combining all these results, we get

$$\lim_{T\to\infty}\left.\frac{\partial^2}{\partial t^2}\right|_{t=0}\pi_\emptyset\left(M^{(T)}_{\frac{tu}{\sqrt{T}}}(\nu)\right) = \lim_{T\to\infty}\frac{kT_0}{T}\partial^2\lambda(0;u) = \partial^2\lambda(0;u)$$

This, together with equation (4.15) implies the desired equality:

$$\partial^2\lambda(0;u) = \lim_{T\to\infty}\int_{S_\Omega}\left(\frac{S_T u}{\sqrt{T}}\right)^2 d\nu = \sigma_u^2 \ge 0$$

and gives also the existence of the limit. □

4.3.4 Condition for positivity of σ_u^2

It is straightforward that $u = v - v \circ F$ implies $\sigma_u^2 = 0$ because in this case $S_T u = v - v \circ F^T$.

For the necessary condition in (4.4), we have to introduce the adjoint of the composition by F, $P : L^2(\nu) \to L^2(\nu)$ defined by

$$\int_{S_\Omega} \varphi \circ F \cdot \psi \, d\nu = \int_{S_\Omega} \varphi \cdot (P\psi) \, d\nu \qquad \forall \varphi, \psi \in L^2(\nu)$$

and we note that if $u \in \mathcal{C}_\vartheta(S_\Omega)$ and $g \in H_\theta$, then we can use the general formulation of the spectral gap property of M_0 (see Theorem 4.4.3) to get:

$$\begin{aligned}\int_{S_\Omega} u \cdot P^T g \, d\nu &= \int_{S_\Omega} u \, d\left(M_0^{(T)}(g \star \nu)\right) \\ &= \left(\int_{S_\Omega} u \, d\nu\right) \cdot \left(\int_{S_\Omega} g \, d\nu\right) + \int_{S_\Omega} u \, d\left(R_0^T(g \star \nu)\right)\end{aligned}$$

Hence, when $m_u = 0$:

$$\left|\int_{S_\Omega} u \cdot P^T g \, d\nu\right| \le |u|_\vartheta \gamma^{-T} |g|_\theta \|\nu\|_\vartheta \tag{4.16}$$

This estimate allows to give another expression for σ_u^2. We write:

$$\begin{aligned}\frac{1}{T}\int_{S_\Omega} (S_T u)^2 \, d\nu &= \int_{S_\Omega} u^2 \, d\nu + 2\sum_{k=1}^{T-1}\left(1 - \frac{k}{T}\right)\int_{S_\Omega} u \cdot P^k u \, d\nu \\ &= -\int_{S_\Omega} u^2 \, d\nu + 2\sum_{k=0}^{T-1}\left(1 - \frac{k}{T}\right)\int_{S_\Omega} u \cdot P^k u \, d\nu\end{aligned}$$

and (4.16) implies the existence of

$$\sigma_u^2 = \lim_{T\to\infty} \frac{1}{T}\int_{S_\Omega} (S_T u)^2 \, d\nu = -\int_{S_\Omega} u^2 \, d\nu + 2\sum_{k\ge 0}\int_{S_\Omega} u \cdot P^k u \, d\nu$$

If $\sigma_u^2 = 0$, then $\int_{S_\Omega} (S_T u)^2 \, d\nu = -2T \sum_{k \ge T} \int_{S_\Omega} u \cdot P^k u \, d\nu - 2\sum_{k=0}^{T-1} k \int_{S_\Omega} u \cdot P^k u \, d\nu$, hence $S_T u$ is bounded in $L^2(\nu)$ by estimate (4.16) : $\|S_T u\|_{L^2(\nu)} \le C$ for all $T \ge 0$.

Again with estimate (4.16), for $g \in H_\theta$:

$$l(g) = \lim_{T\to\infty} \int_{S_\Omega} S_T u \cdot g \, d\nu = \sum_{k\ge 0}\int_{S_\Omega} u \cdot P^k g \, d\nu$$

defines a $L^2(\nu)$-bounded linear functional on H_θ, because $|l(g)| \leq C\|g\|_{L^2(\nu)}$. This functional extends then to $L^2(\nu)$ by density of H_θ, and there is $S_u \in L^2(\nu)$ such that:

$$l(g) = \lim_{T\to\infty} \int_{S_\Omega} S_T u \cdot g \, d\nu = \int_{S_\Omega} S_u \cdot g \, d\nu$$

In the same way, $\lim_{T\to\infty} \int_{S_\Omega} u \circ F^T \cdot g \, d\nu = 0$ for every $g \in L^2(\nu)$.

We get then for every $g \in L^2(\nu)$:

$$\begin{aligned} \int_{S_\Omega} u \cdot g \, d\nu &= \int_{S_\Omega} S_T u \cdot g \, d\nu - \int_{S_\Omega} S_T u \circ F \cdot g \, d\nu + \int_{S_\Omega} u \circ F^T \cdot g \, d\nu \\ &= \int_{S_\Omega} S_T u \cdot g \, d\nu - \int_{S_\Omega} S_T u \cdot Pg \, d\nu + \int_{S_\Omega} u \circ F^T \cdot g \, d\nu \\ &\longrightarrow \int_{S_\Omega} S_u \cdot g \, d\nu - \int_{S_\Omega} S_u \cdot Pg \, d\nu = \int_{S_\Omega} (S_u - S_u \circ F) \cdot g \, d\nu \end{aligned}$$

as T goes to infinity. This proves the desired identity $u = S_u - S_u \circ F$ in $L^2(\nu)$.

4.3.5 Proof of the Central Limit Theorem

To show that $\frac{S_T u - T m_u}{\sqrt{T}\sigma_u}$ converges in law under any initial probability $\phi \in \mathcal{M}^p_\vartheta$ to the standard normal law, it is enough to show the convergence of its Laplace transform. We treat only the centered case and note that this result is valid even if $\sigma_u^2 = 0$.

Proposition 4.3.2. *For all $u \in H^r_\theta$ such that $m_u = 0$ and for all $\phi \in \mathcal{M}^p_\vartheta$, we have:*

$$\lim_{T\to\infty} \int_{S_\Omega} \exp\left(\frac{t}{\sqrt{T}} S_T u\right) d\phi = \exp\left(\frac{t^2}{2}\sigma_u^2\right) \qquad \forall t \in \mathbb{R} \tag{4.17}$$

If $\sigma_u^2 > 0$, this implies the central limit theorem (CLT) by Lévy's Theorem (see for example Theorem 2.5.1 in [10]). The case $\sigma_u^2 = 0$ corresponds to a faster convergence to the limit.

Proof. We proceed as in the proof of the first part of Proposition 4.3.1, but replace $\frac{1}{T}$ by $\frac{t}{\sqrt{T}}$. We can then use the decomposition (4.10) as soon as $T > \left(\frac{t|u|_\theta}{\rho}\right)^2$.

As we take $m_u = \partial^1\lambda(0;u) = 0$, the main term in $\lambda^{kT_0}\left(\frac{tu}{\sqrt{T}}\right)$ will be the second derivative:

$$\lambda^{kT_0}\left(\frac{tu}{\sqrt{T}}\right) = \left(1 + \frac{t^2}{2T}\partial^2\lambda(0;u) + o\left(\frac{1}{T}\right)\right)^{kT_0} \longrightarrow \exp\left(\frac{t^2}{2}\partial^2\lambda(0;u)\right)$$

when T goes to infinity. □

4.3.6 Proof of the Moderate Deviations Principle

A Moderate Deviations Principle with parameter $\frac{1}{2} < \alpha < 1$ is in fact a large deviations result for the laws of the random variables $\frac{S_T u}{T^\alpha}$. For these the exponential scale of probabilities is known to be of the order of $T^{2\alpha-1}$. This will then be the speed of the large deviations result (see Theorem 3.7.1 in [32]). It is hence sufficient to evaluate the appropriate log-Laplace transform:

$$\begin{aligned}\Lambda_\alpha(\beta) &= \lim_{T\to\infty} \frac{1}{T^{2\alpha-1}} \log \int_{S_\Omega} \exp\left(\beta T^{2\alpha-1} \frac{S_T u}{T^\alpha}\right) d\phi \\ &= \lim_{T\to\infty} \frac{1}{T^{2\alpha-1}} \log \int_{S_\Omega} \exp\left(\beta \frac{S_T u}{T^{1-\alpha}}\right) d\phi\end{aligned}$$

Proposition 4.3.3. *For all fixed $\frac{1}{2} < \alpha < 1$, for all $\phi \in \mathcal{M}_\vartheta^p$ and all $u \in H_\theta^r$ such that $m_u = 0$, we have:*

$$\Lambda_\alpha(\beta) = \frac{\beta^2}{2}\sigma_u^2 \tag{4.18}$$

The analyticity of $\Lambda_\alpha(\beta)$ allows to apply a general form of Gärtner-Ellis Theorem (see Theorem II.6.1 in [40], in fact a generalization to various speeds of convergence in the Corollary 2.1.2). The latter says that $\frac{S_T u}{T^\alpha}$ satisfies a Large Deviations Principle with speed $T^{2\alpha-1}$ and rate function given by the Legendre transform of Λ_α:

$$I_\alpha(x) = \Lambda_\alpha^*(x) = \sup_{\beta\in\mathbb{R}} (\beta x - \Lambda_\alpha(\beta)) = \frac{x^2}{2\sigma_u^2},$$

which is independent of α, if $\sigma_u^2 > 0$. This result is exactly the property (MDP) of Theorem 4.2.3.
If $\sigma_u^2 = 0$, then $I_\alpha(0) = 0$ and $I_\alpha(x) = +\infty$ for all $x \neq 0$, which corresponds to a trivial case.

Proof. We can proceed as for the Central Limit Theorem because $T^{1-\alpha} \to \infty$ as $T \to \infty$ so that, for T great enough, we are again in the domain where Theorem 4.3.1 can be applied. The main difference is that $\int_{S_\Omega} \exp\left(\beta \frac{S_T u}{T^{1-\alpha}}\right) d\phi$ diverges exponentially fast so that we need the factor $\frac{1}{T^{2\alpha-1}}$ to rescale it.
For $T = kT_0 + \tilde{T}$ with $T > \left(\frac{|\beta||u|_\theta}{\rho}\right)^{\frac{1}{1-\alpha}}$, we denote $u_T = \frac{\beta u}{T^{1-\alpha}}$. Then:

$$\exp\left(-\frac{\beta T_0}{T^{1-\alpha}}|u|_\infty\right) \leq \exp\left(S_{\tilde{T}} u_T \circ F^{kT_0}\right) \leq \exp\left(\frac{\beta T_0}{T^{1-\alpha}}|u|_\infty\right)$$

and:

$$\int_{S_\Omega} \exp\left(\beta \frac{S_{kT_0} u}{T^{1-\alpha}}\right) d\phi = \pi_\emptyset\left(M_{u_T}^{(kT_0)}\phi\right) = \lambda^{kT_0}(u_T)\,\pi_\emptyset\left(Q_{u_T}\phi\right) + \pi_\emptyset\left(R_{u_T}^k \phi\right)$$

with:

$$\frac{1}{T^{2\alpha-1}} \log\left(\lambda^{kT_0}(u_T)\right) = \frac{kT_0}{T} T^{2-2\alpha} \log\left(1 + \frac{\beta^2}{2T^{2-2\alpha}} \partial^2 \lambda(0;u) + o\left(\frac{1}{T^{2-2\alpha}}\right)\right)$$
$$\longrightarrow \frac{\beta^2}{2} \partial^2 \lambda(0;u) = \frac{\beta^2}{2} \sigma_u^2 \quad \text{as } T \to \infty$$

We have for the remaining term:

$$\frac{1}{T^{2\alpha-1}} \log\left(\pi_\emptyset\left(Q_{u_T}\phi\right) + \frac{\pi_\emptyset\left(R_{u_T}^k \phi\right)}{\lambda^{kT_0}(u_T)}\right)$$

which tends to zero when T goes to infinity since

$$\pi_\emptyset\left(Q_{u_T}\phi\right) \longrightarrow 1 \qquad \text{and} \qquad \left|\frac{\pi_\emptyset\left(R_{u_T}^k \phi\right)}{\lambda^{kT_0}(u_T)}\right| \leq \left(\frac{\gamma^{-T_0} + 2\delta}{1-\delta}\right)^k \cdot \|\phi\|_\vartheta \longrightarrow 0$$

We get in conclusion that

$$\Lambda_\alpha(\beta) = \lim_{T\to\infty} \frac{1}{T^{2\alpha-1}} \log \int_{S_\Omega} \exp\left(\frac{\beta}{T^{1-\alpha} S_T u}\right) d\phi = \frac{\beta^2}{2}\sigma_u^2$$

□

4.3.7 Proof of the large deviations result

We cannot prove a complete Large Deviations Principle because the existence of the spectral gap for $M_u^{(T_0)}$ in Theorem 4.3.1 is valid only for small u and the scaling taken to compute the log-Laplace transform is not the same as for moderate deviations (it corresponds to the case $\alpha = 1$). What we can obtain is for every $u \in H_\theta^r$ a complete upper bound and a partial lower bound controlled by a rate function with a unique minimum.

For $u \in H_\theta^r$ and $\phi \in \mathcal{M}_\vartheta^p$ such that $\|\phi\|_\vartheta < \delta^{-2}$ (this is a technical assumption which is not very important: it is satisfied by Lebesgue measure, and we can modify δ in Theorem 4.3.1 such that it is satisfied by any fixed ϕ), we write:

$$\Lambda_u(\beta) = \limsup_{T\to\infty} \frac{1}{T} \log \int_{S_\Omega} \exp(\beta S_T u)\, d\phi$$

the limit of log-Laplace transforms of $S_T u$.

Proposition 4.3.4. *For $|\beta| < \frac{\rho}{|u|_\theta}$, the map*

$$\beta \longmapsto \Lambda_u(\beta) = \lim_{T\to\infty} \frac{1}{T} \log \int_{S_\Omega} \exp(\beta S_T u)\, d\phi = \log(\lambda(\beta u)) \tag{4.19}$$

is analytic.

Proof. We proceed exactly as in the proof of Proposition 4.3.3, with $|\beta| < \frac{\rho}{|u|_\theta}$ such that:

$$\pi_\emptyset\left(M_{\beta u}^{(kT_0)}\right) = \lambda^{kT_0}(\beta u)\, \pi_\emptyset\left(Q_{\beta u}(\phi)\right) + \pi_\emptyset\left(R_{\beta u}^k(\phi)\right)$$

with

$$\frac{1}{T} \log \lambda^{kT_0}(\beta u) \longrightarrow \log \lambda(\beta u) \quad \text{when } T \to \infty$$

and

$$\frac{1}{T} \log \left(\pi_\emptyset\left(Q_{\beta u}(\phi)\right) + \frac{\pi_\emptyset\left(R_{\beta u}^k(\phi)\right)}{\lambda^{kT_0}(\beta u)} \right)$$

tends to zero when T goes to infinity, since

$$|\pi_\emptyset(Q_{\beta u}\phi) - 1| \leq \delta^2 \|\phi\|_\vartheta \qquad \text{and} \qquad \left| \frac{\pi_\emptyset\left(R_{\beta u}^k(\phi)\right)}{\lambda^{kT_0}(\beta u)} \right| \longrightarrow 0$$

□

This local differentiability implies the following partial large deviations result:

Theorem 4.3.2. *For all $u \in H_\theta^r$ and $\phi \in \mathcal{M}_\vartheta^p$ such that $\|\phi\|_\vartheta \leq \delta^{-2}$, we define:*

$$I_u(x) = \sup_{\beta\in\mathbb{R}} \left(\beta x - \Lambda_u(\beta)\right)$$

Then:

1. *I_u is convex and lower semi-continuous, $I_u(x) = +\infty$ if $|x| > |u|_\infty$, $I_u(x) \geq 0$ and:*

$$I_u(x) = 0 \quad \textit{if and only if } x = m_u \tag{4.20}$$

2. *For all closed $F \subset \mathbb{R}$:*

$$\limsup_{T\to\infty} \frac{1}{T} \log \phi\left(z \;:\; \frac{S_T u(z)}{T} \in F\right) \leq -\inf_{x\in F} I_u(x) \qquad \text{(Upper Bound)}$$

3. For all $x \in \Lambda'_u\left(-\frac{\rho}{|u|_\theta}, \frac{\rho}{|u|_\theta}\right)$ *and* $\delta > 0$*:*

$$\liminf_{T\to\infty} \frac{1}{T} \log \phi\left(z \,:\, \frac{S_T u(z)}{T} \in B(x,\delta)\right) \geq -I_u(x) \qquad \text{(Lower Bound)}$$

Proof. We are exactly in the context of Gärtner-Ellis Theorem stated in Theorem 2.1.4. I_u is the Legendre transform of Λ_u, hence convex and lower semi-continuous. Since $|\Lambda_u(\beta)| \leq |\beta u|_\infty$ for any β, I_u is infinite outside $[-|u|_\infty, |u|_\infty]$. When there is $\lambda \in \left(-\frac{\rho}{|u|_\theta}, \frac{\rho}{|u|_\theta}\right)$ such that $x = \Lambda'_u(\lambda)$, then, by convexity of Λ_u:

$$\begin{aligned}\Lambda_u(\lambda) &\leq \Lambda_u(\eta) - (\eta - \lambda)x \quad \forall \eta \in \mathbb{R}\\ &\leq \lambda x - I_u(x)\end{aligned}$$

and $\Lambda_u(\lambda) = \sup(\lambda y - I_u(y)) \geq \lambda x - I_u(x)$, hence $\Lambda - u(\lambda) = \lambda x - I_u(x)$. Moreover, if $\Lambda_u(\lambda) = \lambda y - I_u(y)$, then $\theta y \leq \Lambda_u(\lambda + \theta) - \Lambda_u(\lambda)$ for all θ which gives taking θ to 0:

$$y = \Lambda'(\lambda) = x$$

Then x is an exposed point for I_u with exposing hyperplane λ. We can hence apply Theorem 2.1.4 to obtain points 2. and 3. of this result.
The previous analysis gives also that $I_u(x) = 0$ if and only of $x = \Lambda'_u(0) = m_u$. □

4.4 Perturbed transfer operators and spectral gap

In this Section, we explain how to modify [95] to obtain perturbed transfer operators and preserve the spectral gap property. This analysis gives the proof of the intermediate Theorem 4.3.1. The precise construction of the operators is given in Section 4.5.

4.4.1 Finite box operators

The construction of the transfer operators is well understood by looking at restrictions of the coupled map to finite boxes: we fix some boundary condition $\xi \in S_\Omega$ and define for all $\Lambda \in \mathcal{F}$:

$$\begin{aligned} F_\Lambda \ : A_\Lambda &\to \mathcal{C}^\Lambda \\ z_\Lambda &\mapsto F(z_\Lambda \vee \xi_{\Lambda^C})|_\Lambda \end{aligned}$$

where $z_\Lambda \vee \xi_{\Lambda^C}$ denotes the point $w \in S_\Omega$ such that $w_i = z_i$ for all $i \in \Lambda$ and $w_i = \xi_i$ for all $i \in \Lambda^C$.
This function F_Λ is expanding as soon as $\kappa < (\lambda - 1)\rho$ and we can define the associated transfer operator $L_\Lambda \ : \ E_\Lambda \to E_\Lambda$ as follows:

$$\int_{S_\Lambda} \varphi \circ F_\Lambda \cdot \psi \, dm^\Lambda = \int_{S_\Lambda} \varphi \cdot L_\Lambda(\psi) \, dm^\Lambda \quad \forall \varphi, \psi \in E_\Lambda \tag{4.21}$$

This is a classical tool to study asymptotic properties of such dynamical systems (see [3] for an extended study of this domain).

In the same way, for $u \in E_\Lambda$, we can define a perturbed operator by:

$$\int_{S_\Lambda} \varphi \circ F_\Lambda \cdot e^u \cdot \psi \, dm^\Lambda = \int_{S_\Lambda} \varphi \cdot M_{\Lambda,u}(\psi) \, dm^\Lambda \quad \forall \varphi, \psi \in E_\Lambda$$

Or, equivalently:

$$M_{\Lambda,u}(\psi) = L_\Lambda \left(e^u \cdot \psi \right) \tag{4.22}$$

The interest of $M_{\Lambda,u}$ comes from the formula:

$$\int_{S_\Lambda} \exp\left(\sum_{t=0}^{T-1} u \circ F_\Lambda^t \right) \cdot \psi \, dm^\Lambda = \int_{S_\Lambda} M_{\Lambda,u}^T(\psi) \, dm^\Lambda$$

which identifies the Laplace transform of $\sum u \circ F_\Lambda^t$ with some spectral characteristic of $M_{\Lambda,u}$. We thus need an infinite dimensional equivalent of these operators, as described in Theorem 4.3.1.

The method to construct them is based on the following perturbative expansion, derived from Theorem 3.2 and Lemma 3.4 of [95]:

Theorem 4.4.1. *If $\kappa < (\lambda-1)\rho$, L_Λ has the integral representation:*

$$L_\Lambda(\psi)(\omega_\Lambda) = \pm \int_{\Gamma_\Lambda} \prod_{p\in\Lambda} \Big(k(\omega_p, f_p(z_p)) + \sum_{V\in\mathcal{F}} \beta_{p,V}(\omega_p, z_{V\cap\Lambda} \vee \xi_{V\cap\Lambda^C}) \Big) \psi(z_\Lambda)\, \mu^\Lambda(dz_\Lambda) \quad (4.23)$$

where:

- $\Gamma_\Lambda = \prod_{p\in\Lambda} \partial A[\rho]$ *and* μ^Λ *is the unique holomorphic differential form on* $\prod_{p\in\Lambda} \mathcal{C}$ *which extends* m_Λ.
- k *is the periodic Cauchy kernel:*

$$k(\omega, z) = \frac{1}{2i}\cot(\pi(z-\omega)) = \frac{1}{2\pi i}\sum_{n\in\mathbb{Z}} \frac{1}{z-\omega+n}$$

- $\beta(p,V)$ *are weakly holomorphic on* $D_{p,V} = A_p \times \Gamma_p \times \prod_{q\in V\setminus\{p\}} A_q$ *(i.e. continuous in all variables and holomorphic in* $w_p \in Int(A_p)$ *and* $z_{V\setminus\{p\}} \in Int(A_{V\setminus\{p\}})$*, see Appendix B of [95]) such that:*

$$\sum_{V\in\mathcal{F}} \theta^{-|V|}|\beta_{p,V}| \le C_\beta = \frac{e^{2\pi\kappa}}{e^{2\pi(\lambda-1)\rho} - e^{2\pi\kappa}} - \frac{1}{e^{2\pi(\lambda-1)\rho}-1}$$

and

$$\int_{S_p} \beta_{p,V}(\omega_p, z_{V\cup\{p\}})\, d\omega_p = 0 \quad \forall z_p \in \Gamma_p, z_{V\setminus\{p\}} \in A_{V\setminus\{p\}}$$

4.4.2 Existence of the operators

We can write a similar integral representation for L_Λ^T and expand it by interchange of products and sums. This gives naturally the kind of configurational operators we need to introduce and control to define an infinite-dimensional transfer operator. That is the method implemented by H.H. Rugh in Section 4 of [95].

To generalize his construction to perturbed operators, we cannot, for technical reasons, proceed as in (4.22) for the finite dimensional setting. We have to adapt the proof of [95] with some additional terms corresponding to the expansion of perturbations $U \in H_\theta$ and to control the new estimates. This allows to construct the general operators described in the following result:

Theorem 4.4.2. *Suppose the hypotheses of Theorem 4.2.1 are verified, then there is $\theta < \vartheta < 1$ such that for all $T \geq 1$ we have a multilinear functional:*

$$\begin{aligned} \mathcal{L}^{(T)} : \qquad\qquad H_\theta^T &\longrightarrow L(\mathcal{M}_\vartheta, \mathcal{M}_\theta) \\ (U_0, \ldots, U_{T-1}) &\longmapsto L^{(T)}_{[U_0,\ldots,U_{T-1}]} \end{aligned}$$

with the following properties:

- *There exists $T_0 \geq 1$ such that $L^{(T)}_{[U_0,\ldots,U_{T-1}]} \in L(\mathcal{M}_\vartheta) \quad \forall T \geq T_0$*
- $$\left\| L^{(T)}_{[U_0,\ldots,U_{T-1}]} \right\| \leq \prod_{t=0}^{T-1} |U_t|_\theta \tag{4.24}$$
- $$L^{(t+T)}_{[V_0,\ldots,V_{t-1},U_0,\ldots,U_{T-1}]} = L^{(t)}_{[V_0,\ldots,V_{t-1}]} \circ L^{(T)}_{[U_0,\ldots,U_{T-1}]} \qquad \text{if } T \geq T_0 \tag{4.25}$$
- $$L^{(T)}(\mathcal{M}_\vartheta^m) \subset \mathcal{M}_\theta^m \tag{4.26}$$
- $$\int_{S_\Omega} b \circ F^T \cdot \prod_{t=0}^{T-1} U_t \circ F^t \, d\phi = \int_{S_\Omega} b \, d\left(L^{(T)}_{[U_0,\ldots,U_{T-1}]} \phi \right) \quad \forall\, b \in \mathcal{C}(S_\Omega),\ \phi \in \mathcal{M}_\vartheta^m \tag{4.27}$$

Proof. As already mentioned, this proof essentially follows the construction of Part IV of [95]. We only have to add in configurations the terms corresponding to the perturbations U_i. We can do it and keep good estimates uniformly in $U \in H_\theta$. We detail the proof in Section 4.5. □

Remark: The operator constructed in [95] corresponds to $L^{(T)}_{[1,\ldots,1]}$.

Operators of Theorem 4.3.1 are a particular case of the general operators constructed in Theorem 4.4.2. For $u \in H_\theta$, we take

$$M_u^{(T)} = L^{(T)}_{[e^u,\ldots,e^u]}$$

We can already obtain some properties of these operators:

Theorem 4.3.1 - First part. Since it is the composition of the analytic function $u \mapsto e^u$ and of the multilinear map $\mathcal{L}^{(T)}$, $M_u^{(T)}$ is analytic.

We can write explicitly the series expansion of $M_u^{(T)}$ around a point u:

$$M_{u+h}^{(T)} = \sum_{n \geq 0} \frac{1}{n!} \partial^n M^{(T)}(u; h),$$

where

$$\partial^n M^{(T)}(u; h) = \sum_{\substack{n_0,\ldots,n_{T-1} \geq 0 \\ \sum_{t=0}^{T-1} n_t = n}} \frac{n!}{n_0! \cdots n_{T-1}!} L^T_{[h^{n_0} e^u, \ldots, h^{n_{T-1}} e^u]},$$

which is an element of $L(\mathcal{M}_\vartheta, \mathcal{M}_\theta)$ (or $L(\mathcal{M}_\vartheta)$ if $T \geq T_0$), is homogeneous of degree n and satisfies the bound

$$\left\|\partial^n M^{(T)}(u;h)\right\| \leq (T|h|_\theta)^n e^{T|u|_\theta}$$

hence we control the difference between two operators by:

$$\left\|M_{u+h}^{(T)} - M_u^{(T)}\right\| \leq \left(e^{T|h|_\theta} - 1\right) e^{T|u|_\theta}$$

Estimate (4.6) is the particular case of this inequality around $u = 0$. Formulas (4.7), (4.8) and (4.9) are easily deduced from (4.25), (4.26) and (4.27). □

4.4.3 Proof of the Spectral gap property

The central result in [95] is a spectral gap property for $M_0^{(T)} = L_{[1,\cdots,1]}^{(T)}$. It is a direct consequence of his Lemma 4.25 and can be stated as:

Theorem 4.4.3. *Under the same assumptions on the parameters as in Theorem 4.4.2, $M_0^{(T)}$ can be written for all $T \geq T_0$:*

$$M_0^{(T)}(\phi) = \left(\int_{S_\Omega} \phi \, dm^\Omega\right) \nu + R^T(\phi)$$

with $\nu \in \mathcal{M}_\vartheta$, $M_0^{(T)}(\nu) = \nu$ and $\|R^T(\phi)\| \leq \gamma^{-T}$, so that $Sp\,(R^T) \subset D(0, \gamma^{-T})$.

We cannot generalize this result to all our perturbed operators, but only extend it to small u as stated in the second part of Theorem 4.3.1. The proof uses an adaptation to our case of the Theorem of Kato-Rellich (see Theorem XII.8 in [91] or Theorem VII.6.9 of [38] for a more general result). We recall below the main steps of its proof and specify some estimates:

Theorem 4.3.1 - Second part. For $M \in L(\mathcal{M}_\vartheta)$, let $\mathrm{Res}(M) = \mathbb{C} \setminus \mathrm{Sp}(M)$ denote the resolvent set, and for $\lambda \in \mathrm{Res}(M)$

$$R(\lambda, M) = (\lambda \mathrm{Id} - M)^{-1}$$

the associated resolvent function.
We denote $M_u = M_u^{(T_0)} \in L(\mathcal{M}_\vartheta)$. Then, by Lemmas VII.6.3 and VII.6.4 of [38], we get that for any fixed $\delta < \frac{1-\gamma^{-T_0}}{3}$, there exists $\varepsilon > 0$ such that $\|M_u - M_0\| < \varepsilon$ implies:

1. $\{\lambda \,:\, d\,(\lambda, \mathrm{Sp}(M_0)) \geq \delta\} \subset \mathrm{Res}(M_u)$
2. $\|R\,(\lambda, M_u) - R\,(\lambda, M_0)\| < \delta$ $\qquad \forall\lambda$ s.t. $d\,(\lambda, \mathrm{Sp}(M_0)) \geq \delta$
3. $u \to R(\lambda, M_u)$ is analytic $\qquad \forall\lambda$ s.t. $d\,(\lambda, \mathrm{Sp}(M_0)) > \delta$

The last statement is a straightforward generalization of the proof of Lemma VII.6.4: the set of analytic functions in our sense is stable by the same operations as for classical analytical functions.

Then $\mathrm{Sp}(M_u) \subset D(1,\delta) \cup D\left(0, \gamma^{-T_0} + \delta\right)$ and if we denote

$$Q_u = -\frac{1}{2\pi i} \int_{|\lambda-1|=\delta} R(\lambda, M_u)\, d\lambda = -\frac{1}{2\pi i} \int_{|\lambda-1|=2\delta} R(\lambda, M_u)\, d\lambda$$

the projection associated to the spectrum of M_u included in $D(1,\delta)$, we get that Q_u is an analytic function of u and

$$\|Q_u - Q_0\| \leq \delta \int_0^1 \|R(1+\delta e^{i2\pi\theta}, M_u) - R(1+\delta e^{i2\pi\theta}, L^{(T_0)})\| d\theta \leq \delta^2 < 1 \quad (4.28)$$

This, with Lemma VII.6.7 of [38] and the fact that $\mathrm{Sp}(L(T_0)) \cap D(1,\delta) = \{1\}$ where 1 is a simple eigenvalue, implies that $\mathrm{Sp}(M_u) \cap D(1,\delta) = \{\lambda^{T_0}(u)\}$, where

$$\lambda^{T_0}(u) = \frac{M_u \circ Q_u(1)}{Q_u(1)}$$

is a simple eigenvalue and an analytic function of u.

Now, setting

$$R_u = M_u - \lambda^{T_0}(u)\, Q_u = M_u \circ \left(-\frac{1}{2\pi i} \int_{\{|\lambda| = \gamma^{-T_0} + \delta\}} R(\lambda, M_u)\, d\lambda \right),$$

which is the projection on the rest of the spectrum, we get:

$$M_u^{(kT_0)} = \lambda^{kT_0}(u)\, Q_u + R_u^k, \qquad \text{with } \mathrm{Sp}(R_u) \subset D\left(0, \gamma^{-T_0} + \delta\right)$$

and

$$\begin{aligned} \|R_u - R_0\| &\leq \|M_u\| \left\| -\frac{1}{2\pi i} \int (R(\lambda, M_u) - R(\lambda, M_0))\, d\lambda \right\| \\ &\qquad + \|M_u - L^{(T_0)}\| \left\| -\frac{1}{2\pi i} \int R(\lambda, M_0)\, d\lambda \right\| \\ &\leq (1+\varepsilon)\left(\gamma^{-T_0} + \delta\right)\delta + \varepsilon \\ &\leq 2\delta \qquad \text{taking } \varepsilon \text{ smaller if necessary} \end{aligned}$$

so that $\|R_u^k\| \leq \left(\gamma^{-T_0} + 2\delta\right)^k$ for every $k \geq 1, u \in D_\theta(0,\rho)$. □

4.5 Construction of the operators

This Section contains a sketch of the construction of the transfer operator done in [95] and presents also the main modifications which have to be done to extend it to perturbed operators.

4.5.1 Single site operators

For $f_p \in \mathcal{E}(\rho, \lambda)$ an expanding map on A_p, the associated transfer operator $L_{f_p} : E_p \longrightarrow E_p$ can be written (this is a particular case of identity (4.23)):

$$L_{f_p}\phi(\omega_p) = \int_{\Gamma_p} k(\omega_p, f_p(z_p))\phi(z_p)\, \mu^p(dz_p)$$

It satisfies $l_p \circ L_{f_p} = l_p$ with $l_p(\phi) = \int_{S_p} \phi(z_p)\, dz_p$ and enjoys a spectral gap property with the following estimates, uniformly in $f_p \in \mathcal{E}(\rho, \lambda)$:

$$\left\| L_{f_p}^T \right\| \le c_h \quad \text{and} \quad \left\| L_{f_p}^T \Big|_{\mathrm{Ker}\ l_p} \right\| \le c_r \eta^T \tag{4.29}$$

where $c_h \ge 1$, $c_r > 0$ and $\eta < 1$. A proof of these results can be found in Appendix A of [95].

4.5.2 Configurations

We define what a branching, the main element to define the configurations, is:

Definition 4.5.1. *A* branching pair (S, V) *is composed by a subset* $S \in \mathcal{F}$ *and a function* $V : S \to \mathcal{F}$. *We denote* $V[S] = S \cup (\cup_{p \in S} V(p))$.

Given $K \in \mathcal{F}$, (S, V) a branching pair such that $S \subset K$, $U \in H_\theta$ with fixed expansion $U = \sum_{W \in \mathcal{F}} U_W$ and $W \in \mathcal{F}$, we define $H = K \cup V[S] \cup W$ and the operator $L_{K,(S,V),(U,W)} : E_H \to E_K$ by:

$$\begin{aligned} L_{K,(S,V),(U,W)}(\varphi_H)(\omega_K) = \pm \int_{S_{H \setminus K}} m^{H \setminus K}(dz_{H \setminus K}) \int_{\Gamma_K} \prod_{p \in S} \beta_{p,V(p)}(\omega_p, z_{V(p) \cup p}) \\ \times \prod_{p \in K \setminus S} k(\omega_p, f_p(z_p)) U_W(z_W) \varphi_H(z_H)\, \mu^K(dz_K) \end{aligned}$$

We then have some compatibility properties for these operators:

Lemma 4.5.1. *We have:*

$$\pi_{K\setminus\{p\},K}L_{K,(S,V),(U,W)} = \begin{cases} 0 & \text{if } p \in S \\ L_{K\setminus\{p\},(S,V),(U,W)} & \text{if } p \in (V[S]\cup W)\setminus S \\ L_{K\setminus\{p\},(S,V),(U,W)} \circ \pi_{H\setminus\{p\},H} & \text{if } p \in K\setminus(V[S]\cup W) \end{cases} \tag{4.30}$$

and the sum

$$\sum_{W\in\mathcal{F}} L_{K,(S,V),(U,W)} \circ \pi_H(\phi) = L_{K,(S,V),(1,\emptyset)} \circ \pi_{K\cup V[S]}(U \star \phi)$$

is independent of the decomposition $U = \sum_{W\in\mathcal{F}} U_W$ *of* $U \in H_\theta$.

Proof. For the first part, we proceed as for Lemma 4.2 of [95]: the last equation of Theorem 4.4.1 states that $\beta_{p,V}$ is in the kernel of l_p, i.e. of $\pi_{K\setminus\{p\},K}$. Hence if $p \in S$, the term is null. When $p \notin S$, the action of $\pi_{K\setminus\{p\},K}$ removes the corresponding term $k(w_p, f(z_p))$ from the operator and replaces it by l_p (because $l_p \circ L_{f_p} = l_p$). If $p \in (V[S]\cup W)\setminus S$, this term was already in the operator, otherwise it is added by the term $\pi_{H\setminus\{p\},H}$.
For the second part, we commute sum and integral, obtaining

$$\sum_{W\in\mathcal{F}} L_{K,(S,V),(U,W)} \circ \pi_H(\phi) = L_{K,(S,V),(1,\emptyset)} \circ \left(\sum_{W\in\mathcal{F}} \pi_{K\cup V[S],K\cup V[S]\cup W}(U_W\phi_H)\right)$$

and use then the projectivity of ϕ and the definition of the module product on $\mathcal{M}_\theta$. $\square$

Given $T \geq 1$, $U_0, \ldots, U_{T-1} \in H_\theta$, we want to construct for any $K \in \mathcal{F}$ an operator $L^{(T)}_{K,[U_0,\ldots,U_{T-1}]} : \mathcal{M}_\vartheta \to E_K$ and control its norm.
We introduce configurations and associated configurational operators :

Definition 4.5.2. *A* configuration *on* $K \in \mathcal{F}$ *at time* $T \geq 1$ *is the choice of :*

- $W_{T-1}, \ldots W_0 \in \mathcal{F}$, *for the expansion of the perturbative terms* U,
- $(S_{T-1}, V_{T-1}), \ldots, (S_0, V_0)$, *branching pairs for the expansion of the* $\beta_{p,V}$,
- $I \in \mathcal{F}$ *an* initial state,

such that if K *is expanded by* $K_T = K$ *and* $K_t = K_{t+1} \cup V_t[S_t] \cup W_t$ *for* $0 \leq t < T$, *the following conditions are satisfied:*

$$S_t \subset K_{t+1} \ \text{ for } 0 \leq t < T \qquad \text{and} \qquad I \subset K_0$$

We denote $\mathcal{C}[K,T]$ *the set of all these configurations.*

To each configuration $C \in \mathcal{C}[K,T]$, we associate a configurational operator:

$$L_{K,[U_0,\dots,U_{T-1}]}[C] = L^{(T-1)} \circ \dots \circ L^{(0)} \circ Q_I^{K_0} \circ \pi_{K_0} \,:\, \mathcal{M}_\vartheta \to E_K$$

where

$$L^{(t)} = L_{K_{t+1},(S_t,V_t),(U_t,W_t)} \quad \text{and} \quad Q_I^{K_0} = \prod_{p\in I}(1-l_p) \prod_{p\in K_0\setminus I} l_p$$

The following equivalent of Proposition 4.3 of [95] remains valid and will be useful to construct the global operator on the projective Banach space $\mathcal{M}_\vartheta$. It is obtained applying recursively Lemma 4.5.1:

Proposition 4.5.1. *If $\alpha \subset K \in \mathcal{F}$, then:*

$$\pi_{\alpha,K} L_{K,[U_0,\dots,U_{T-1}]}[C] = \begin{cases} L_{\alpha,[U_0,\dots,U_{T-1}]}[C] & \textit{if } C \in \mathcal{C}[\alpha,T] \\ 0 & \textit{otherwise} \end{cases} \tag{4.31}$$

Remark : The initial set I, introduced in [95] to prove the spectral gap, is not necessary here. We keep it however to verify that in the case where $U_t = 1$ for all $0 \le t < T$, we really get the operator constructed in this paper.

For $C \in \mathcal{C}[K,T]$ a given configuration, with (S_t,V_t) the branching pairs, W_t the perturbative expansions and $K = K_T \subset K_{T-1} \subset \dots \subset K_0$ the expansion of K, we call the points $(q,t) \in \cup_{t=0}^{T} K_t \times \{t\}$ points of the configuration and classify them, calling (q,t):

- an inner point if $q \in W_t$,
- a vertex point if $q \in V_t[S_t] \setminus W_t$,
- an apex point if $t \ge 1$ and $q \in S_{t-1} \setminus (V_t[S_t] \cup W_t)$,
- a free point otherwise.

A chain is a maximal sequence of points of the configuration $\gamma = (q,t)_{t_1 \le t \le t_2}$ such that $q \notin S_{t_2-1}$ and $(q,t)_{t_1<t<t_2}$ are free points. t_1 is called the starting time of the chain and $|\gamma| = t_2 - t_1$ its length. Such a chain is called:

- an apex chain if (q,t_1) is an apex point,
- an initial chain if $t_1 = 0$, $(q,0)$ is a free point and $q \in I$,
- an end chain otherwise.

This analysis allows to separate the contributions of chains in $L_{K,[U_0,\dots,U_{T-1}]}[C]$. This is possible because the free points of a chain occur only in both uncoupled operators before and after it, hence can be separated by Fubini Theorem of all other integrands: If $\mathrm{ch}(t)$ denotes all the chains starting at time t, interchange of the terms in the integral gives:

$$L_{K,[U_0,\dots,U_{T-1}]}[C] = \tilde{L}^{(T-1)} \circ \tilde{U}^{(T-1)} \circ \tilde{L}^{(T-2)} \circ \cdots \tilde{U}^{(1)} \circ \tilde{L}^{(0)} \circ \tilde{U}^{(0)} \circ j_{K_0,I} \circ Q_I^{K_0} \circ \pi_{K_0}$$

where $\tilde{U}^{(t)} : E_{K_t} \to E_{K_t}$ is defined by $\tilde{U}^{(t)}(\phi) = (j_{K_t,W_t} U_{t,W_t})\phi$ and $\tilde{L}^{(t)} : E_{K_t} \to E_{K_{t+1}}$ by:

$$\tilde{L}^{(t)} = \left(\prod_{\gamma \in \mathrm{ch}(t)} L_\gamma \right) \pi_{K_{t+1},K_t} \prod_{p \in S_t} M_{K_t,\beta_{p,V_t(p)}}$$

with $L_\gamma = (L_{f_p})^{|\gamma|}$ and:

$$M_{K_t,\beta_{p,V_t(p)}} \phi(\omega_p, z_{K_t \setminus \{p\}}) = \pm \int_{\Gamma_p} \beta_{p,V_t(p)}(\omega_p, z_{V_t(p) \cup \{p\}}) \phi(z_{K_t})\, \mu^p(dz_p)$$

Using this last expression, we can bound the norm of each $L_{K,[U_0,\dots,U_{T-1}]}[C]$ by the product of the estimation for each term in its expression, using the following estimates:

$$\|Q_I^{K_0} \circ \pi_{K_0}\| \leq \left(1 + \frac{1}{\vartheta}\right)^{|I|} \tag{4.32}$$

since $Q_I^{K_0} \circ \pi_{K_0} = \sum_{J \subset I} (-1)^{|J|} j_{K_0,J} \pi_J$ and $\|\pi_J\| \leq \vartheta^{-|J|}$.

$$\|M_{\Lambda,\beta_{p,V}}\| \leq 2|\beta_{p,V}| \quad \text{and} \quad \|\tilde{U}^{(t)}\| \leq |U_{t,W_t}| \tag{4.33}$$

$$\|L_\gamma\| \leq c_h \quad \text{or} \quad \|L_\gamma\| \leq c_r \eta^{|\gamma|} \quad \text{for } \gamma \text{ initial or apex chain} \tag{4.34}$$

This last fundamental estimate comes from the spectral gap result for the single site operator L_{f_p} (see estimates (4.29)).

4.5.3 Tree structures

We will now associate to each configuration a tree structure in an injective way. This will allow us to bound the norms of configurational operators by some more computable estimates. The set of trees is exactly the same as in [95], but we will keep more of them to describe a configuration.

Definition 4.5.3. *For $T \geq 0$ and $p \in \Omega$, the collection of* trees $\mathcal{Y}[p,T]$ *is defined recursively on T:*

- *$\mathcal{Y}[p,0]$ contains two elements: an* initial leaf *and an* end leaf
- *for $t \geq 1$, $\mathcal{Y}[p,t]$ is constituted of the following trees:*
  - *an end leaf*
  - *an initial chain of length t followed by an initial leaf*
  - *an apex chain of length $0 \leq k < t$ followed by a branching over a set V; at each $q \in V \cup \{p\}$, we attach a tree $y_q^{t-k-1} \in \mathcal{Y}[q, t-k-1]$*

We associate now to each $C \in \mathcal{C}[K,T]$ a collection of trees, in fact one $y_{p,T} \in \mathcal{Y}[p,T]$ for each $(p,T)_{p\in K}$ and one $y_{p,t} \in \mathcal{Y}[p,t]$ for each (p,t) inner point, i.e. such that $0 \leq t < T$ and $p \in W_t$. We do this recursively on T, giving us a total ordering of Ω to go through the points associated to a given time (see Figure 4.1 for an illustration of this construction):
For $t = 0$, we associate to each $p \in K_0$ a tree $y_{p,0}$ which is an initial leaf if $p \in I$ and an end leaf otherwise.

Then, for $1 \leq t \leq T$:

- we go through the $p \in S_{t-1} \subset K_t$ and we take $y(p,t)$ a branching over the set $V_{t-1}(p)$, and we attach at each $q \in V_{t-1}(p) \cup \{p\}$:
  - $y(q,t-1)$ if $q \notin W_{t-1}$ and $y(q,t-1)$ has not yet been attached to another tree
  - an end leaf otherwise
- for $p \in K_t \cap [(V_{t-1}[S_{t-1}] \cup W_{t-1}) \setminus S_{t-1}]$, we forget the tree $y(p,t-1)$ (already attached to another tree or kept until the end) and take for $y(p,t)$ an end leaf
- for $p \in K_t \setminus (V_{t-1}[S_{t-1}] \cup W_{t-1})$:
  - $y(p,t)$ is an end leaf if $y(p,t-1)$ was already one (we forget the length of end chains because it is useless in the estimates)
  - otherwise, $y(p,t)$ is a chain of length 1 attached to $y(p,t-1)$

It should be noted that all $y(p,t)$ for (p,t) inner points are never attached to other ones: we keep them in our description of the configuration in term of trees. In fact, the terms U_{t,W_t}, for U_t chosen in H_θ, will exactly compensate the weights of these trees (see Proposition 4.5.4).
We define the weight of a tree as the product of the bounds of its components, and for those, we take the bounds (4.33) for the branchings and (4.34) for the chains. We estimate the other terms by:

$$\|\text{ end leaf }\| = c_h \quad \text{and} \quad \|\text{ initial leaf }\| = 1 + \frac{1}{\vartheta}$$

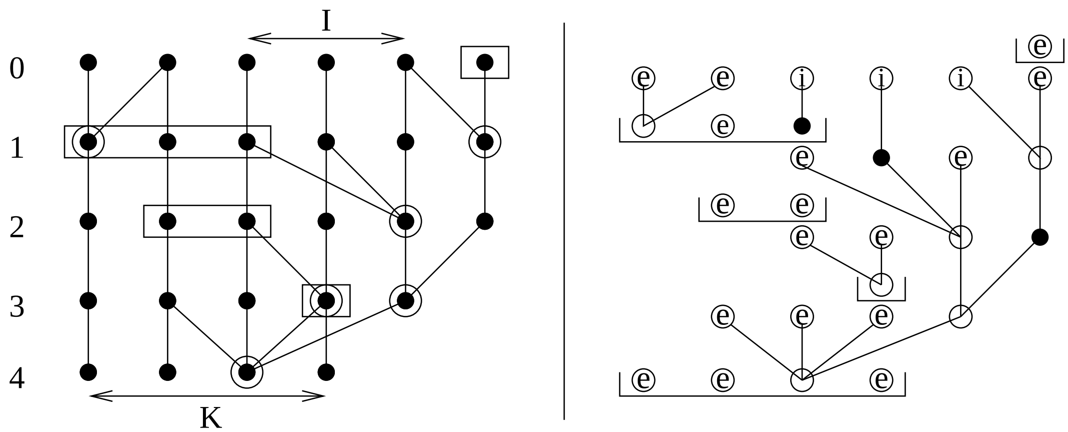

Fig. 4.1: An example of configuration (left) with its associated tree structure (right). Each circle represents the basis of a branching pair, each rectangle represents the perturbative term U_t. To each of these inner points and to the bottom points are associated independent trees.

Proposition 4.5.2. *The map which to every $C \in \mathcal{C}[K,T]$ associates the family of trees ($y(p,t)$ where $t = T$ and $p \in K$, or (p,t) is an inner point) is injective and we have the bound:*

$$\left\| L_{K,[U_0,\ldots,U_{T-1}]}[C] \right\|_{\mathcal{M}_\vartheta \to E_K} \leq \prod_{p\in K} \|y(p,T)\| \prod_{t=0}^{T-1} \left[|U_{t,W_t}| \prod_{p\in W_t} \|y(p,t)\| \right] \quad (4.35)$$

Proof. The tree structure contains in fact the whole information on the configuration defining the operator, hence this one can be constructed back from the tree.
Trees associated to a configuration contain exactly $|I|$ initial leaves, giving factor $(1+\frac{1}{\vartheta})^{|I|}$, as well as all branchings and chains, with length kept in case of initial or apex chains. Some additional end leaves may appear in the construction, each giving a weight $c_h \geq 1$. We get then the desired upper bound. □

Our trees are exactly the same as in [95]. We can then use its bounds for the weights of trees under the condition (TR) of [95]. We don't write this condition here, but just notice that for ρ and λ given, there exists $\theta_0(\rho,\lambda) \in (0,1/3)$ such that for all $\theta < \theta_0$ we can find $\gamma < \eta^{-1}$ (η is the gap of the simple site operator) and κ such that any $F \in CM[\rho,\lambda,\theta,\kappa]$ satisfies (TR) with this γ. We write below the results of Lemmas 4.20 and 4.21 of [95], which give these bounds:

Proposition 4.5.3. *The size of a tree is the sum of the length of its chains added to the number of its branchings. Define:*

$$u_p^T(s) = \sum_{y \in \mathcal{Y}[p,T]} \|y\| s^{size\ (y)}$$

If condition (TR) is satisfied with $\gamma \in (1, \eta^{-1})$, then there exists $\vartheta \in (\theta, 1)$ and $T_0 \geq 1$ such that:

$$\begin{aligned} &u_p^T(\gamma) \leq \theta^{-1} \\ &u_p^T(\gamma) \leq \vartheta^{-1} \qquad \text{if } T \geq T_0 \end{aligned} \tag{4.36}$$

4.5.4 Global estimates

We deduce from Propositions 4.5.2 and 4.5.3 above that:

Proposition 4.5.4.

$$\begin{aligned} \sum_{C \in \mathcal{C}[K,T]} \left\| L_{K,[U_0,\ldots,U_{T-1}]}[C] \right\|_{\mathcal{M}_\vartheta \to E_K} &\leq \theta^{-|K|} \prod_{t=0}^{T-1} |U_t|_\theta \\ &\leq \vartheta^{-|K|} \prod_{t=0}^{T-1} |U_t|_\theta \qquad \text{if } T \geq T_0 \end{aligned} \tag{4.37}$$

Proof. Because of the injectivity of the description by trees, we have:

$$\begin{aligned} &\sum_{C \in \mathcal{C}[K,T]} \left\| L_{K,[U_0,\ldots,U_{T-1}]}[C] \right\|_{\mathcal{M}_\vartheta \to E_K} \\ &\quad \leq \sum_{C \in \mathcal{C}[K,T]} \prod_{p \in K} \|y(p,T)\| \prod_{t=0}^{T-1} \left[|U_{t,W_t}| \prod_{p \in W_t} \|y(p,t)\| \right] \\ &\quad \leq \sum_{W_0,\ldots,W_{T-1}} \prod_{p \in K} u_p^T(1) \prod_{t=0}^{T-1} |U_{t,W_t}| \prod_{p \in W_t} u_p^t(1) \\ &\quad \leq \prod_{p \in K} u_p^T(1) \prod_{t=0}^{T-1} \left(\sum_{W \in \mathcal{F}} |U_{t,W}| \prod_{p \in W} u_p^t(1) \right) \end{aligned}$$

and we can conclude with estimates (4.36). □

These bounds, together with the first part of Proposition 4.5.1 (which assures compatibility of the operators constructed for different subsets K) make it possible to define:

$$L^{(T)}_{[U_0,\ldots,U_{T-1}]} = \left(\sum_{C \in \mathcal{C}[K,T]} L_{K,[U_0,\ldots,U_{T-1}]}[C] \right)_{K \in \mathcal{F}}$$

as an operator from $\mathcal{M}_\vartheta$ to $\mathcal{M}_\theta$, or to $\mathcal{M}_\vartheta$ when $T \geq T_0$, satisfying the announced bound (4.24).
We see also that this operator is independent of the decompositions $U_t = \sum_{W \in \mathcal{F}} U_{t,W}$ writing, with $\hat{L}^{(t)} = L_{K_{T+1},(S_t,V_t),(1,\emptyset)}$ and K_t the corresponding expansion:

$$\begin{aligned} &\left(\sum_{C \in \mathcal{C}[K,T]} L_{K,[U_0,\ldots,U_{T-1}]}[C] \right)(\phi) \\ &= \sum_{(S_{T-1},V_{T_1})} \hat{L}^{(T-1)} \left[\left(U_{T-1} \star \left[\sum_{(S_{T-2},V_{T_2})} \hat{L}^{(T-2)} \cdots \hat{L}^{(0)} \left[(U_0 \star \phi)_{K_0} \right] \right] \right)_{K_{T-1}} \right] \end{aligned} \tag{4.38}$$

For this, we use inductively the second part of Proposition 4.5.1 and the fact that any intermediate operator defines a projective family.

The multilinearity in the perturbation terms $U_0, \ldots, U_{T-1}$ is clear from this last expression. We also can remark that $L^{(T)}_{[1,\ldots,1]} = L^{(T)}$ is exactly the Perron Frobenius operator of [95], because in this case, all configurational operators with a $W_t \neq \emptyset$ are null.
We can then deduce the other properties of these operators by straightforward adaptations of equivalent results in [95]: for the composition rule (4.25), we just have to identify the configurations appearing in both sides. For the properties (4.26) and (4.27), we prove that the constructed operator $L^{(T)}$ is the limit of the restricted equivalent on finite boxes Λ when Λ goes to Ω, then deduce the properties going to the limit.

BIBLIOGRAPHY

[1] N. Anantharaman, Precise counting results for closed orbits of Anosov flows. *Ann. Sci. Ec. Norm. Super.*, **33** (1), 33–56, 2000.

[2] M. Babillot and F. Ledrappier, Lalley's theorem on periodic orbits of hyperbolic flows. *Ergodic Theory Dyn. Syst.*, **18** (1), 17–39, 1998.

[3] V. Baladi, *Positive transfer operators and decay of correlations*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[4] V. Baladi, M. Degli Esposito, S. Isola, E. Järvenpää and A. Kupiainen, The spectrum of weakly coupled map lattices. *J. Math. Pures Appl.*, **77** (6), 539–584, 1998.

[5] V. Baladi and H. Rugh, Floquet spectrum of weakly coupled map lattices. *Commun. Math. Phys.*, **220**, 561–582, 2001.

[6] G. Ben Arous and O. Zeitouni, Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.*, **35** (1), 85–102, 1999.

[7] M. Blank, Generalized phase transitions in finite coupled map lattices. *Phys. D*, **103**, 34–50, 1997.

[8] M. Blank and L. Bunimovich, Multicomponent dynamical systems: SRB measures and phase transitions, 2002. Preprint.

[9] C. Boldrighini, L. Bunimovich, G. Cosimi, S. Frigio and A. Pellegrinotti, Ising-type and other transitions in one-dimensional coupled map lattices with sign symmetry. *J. Statist. Phys.*, **102** (5-6), 1271–1283, 2001.

[10] V. S. Borkar, *Probability theory. An advanced course*. Springer-Verlag, New York, 1995.

[11] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, vol. 470 of *Lecture Notes in Math.*. Springer, 1975.

[12] J. Bricmont and A. Kupiainen, Coupled analytic maps. *Nonlinearity*, **8**, 379–396, 1995.

[13] J. Bricmont and A. Kupiainen, High temperature expansions and dynamical systems. *Commun. Math. Phys.*, **178** (3), 703–732, 1996.

[14] J. Bricmont and A. Kupiainen, Infinite dimensional SRB measures. *Phys. D*, **103** (1–4), 18–33, 1997.

[15] A. Broise, Transformations dilatantes de l'intervalle et théorèmes limites. In Etudes spectrales d'opérateurs de transfert et applications. *Astérisque*, **238**, 1–109, 1996.

[16] W. Bryc and A. Dembo, Large deviations and strong mixing. *Ann. IHP, Probab. Stat.*, **32** (4), 549–569, 1996.

[17] L. Bunimovich, Coupled map lattices: one step forward and two steps back. *Phys. D*, **86** (1-2), 248–255, 1995.

[18] L. Bunimovich and E. Carlen, On the problem of stability in lattice dynamical systems. *J. Diff. Eq.*, **123**, 213–219, 1995.

[19] L. Bunimovich and Y. Sinai, Spacetime chaos in coupled map lattices. *Nonlinearity*, **1**, 491–516, 1988.

[20] L. Bunimovich and Y. Sinai, Statistical mechanics of coupled map lattices. In K. Kaneko, editor, *Theory and Application of Coupled Map Lattices*, 169–189, J. Wiley & Sons, 1993.

[21] H. Chaté, *Transition vers la turbulence via intermittence spatio-temporelle*. Ph.D. thesis, Université Paris VI, 1989.

[22] H. Chaté and M. Courbage, editors, *Lattice dynamics*, 1997. Special issue of Phys. D, 103.

[23] H. Chaté, A. Lemaître, P. Marcq and P. Manneville, Non-trivial collective behavior in extensively-chaotic dynamical systems: an update. *Phys. A*, **224**, 447–457, 1996.

[24] H. Chaté and J. Losson, Non-trivial collective behavior in coupled map lattices: a transfer operator perspective. *Phys. D*, **103**, 51–72, 1997.

[25] H. Chaté and P. Manneville, Spatio-temporal intermittency in coupled map lattices. *Phys. D*, **32** (3), 409–422, 1988.

[26] H. Chaté and P. Manneville, Collective behaviors in spatially extended systems with local interactions and synchronous updating. *Prog. Theor. Phys.*, **87** (1), 1–60, 1992.

[27] T. Chiyonobu and S. Kusuoka, The large deviation principle for hypermixing processes. *Probab. Th. Rel. Fields*, **78**, 627–649, 1988.

[28] L. Cisneros, M. Cosenza, J. Jiménez and A. Parravano, Information transfer and nontrivial collective behavior in chaotic coupled map networks. *Phys. Rev. E*, **65**, 2002.

[29] P. Collet and J. Eckmann, *Iterated maps on the interval as dynamical systems*. Progress in Physics, Birkhauser, 1980.

[30] P. Collet and J. Eckmann, Positive Liapunov exponents and absolute continuity for maps of the interval. *Ergodic Theory Dynam. Systems*, 13–46, 1983.

[31] F. Comets, Grandes déviations pour des champs de Gibbs sur $\mathbb{Z}^d$. *C. R. Acad. Sci. Paris Sér. I Math.*, **303** (11), 511–513, 1986.

[32] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer, 1998.

[33] M. Denker, Large deviations and the pressure function. In *Information theory, statistical decision functions, Random Processes, Trans. 11th Prague Conf. (Prague, 1990)*, vol. A, 21–33, 1992.

[34] J. Deuschel and D. Stroock, *Large Deviations*. Academic Press, 1989.

[35] R. Dobrushin and M. Martirosyan, Nonfinite perturbations of Gibbs fields. *Theoret. and Math. Phys.*, **74** (1), 10–20, 1988.

[36] D. Dolgopyat, Entropy of coupled map lattices. *J. Statist. Phys.*, **86** (1-2), 377–389, 1997.

[37] M. Donsker and S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time. II. *Comm. Pure Appl. Math.*, **28**, 279–301, 1975.

[38] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*. Interscience Publishers, Inc., New York, 1958.

[39] A. Eizenberg, Y. Kifer and B. Weiss, Large deviations for $\mathbb{Z}^d$-actions. *Commun. Math. Phys.*, **164** (3), 433–454, 1994.

[40] R. S. Ellis, *Entropy, large deviations, and statistical mechanics*. Springer-Verlag, New York, 1985.

[41] T. Fischer and H. Rugh, Transfer operators for coupled analytic maps. *Ergodic Theory Dyn. Syst.*, **20** (1), 109–143, 2000.

[42] H. Föllmer and S. Orey, Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.*, **16** (3), 961–977, 1988.

[43] G. Gielis and R. MacKay, Coupled map lattices with phase transitions. *Nonlinearity*, **13**, 867–888, 2000.

[44] P. Grassberger and T. Schreiber, Phase transitions in coupled map lattices. *Phys. D*, **50** (2), 177–188, 1991.

[45] N. Haydn and D. Ruelle, Equivalence of Gibbs and equilibrium states for homeomorphisms satisfying expansiveness and specification. *Comm. Math. Phys.*, **148** (1), 155–167, 1992.

[46] E. Hille and R. S. Phillips, *Functional analysis and semi-groups*. American Mathematical Society, Providence, R. I., 1974.

[47] E. Järvenpää, A SRB-measure for globally coupled circle maps. *Nonlinearity*, **10**, 1435–1469, 1997.

[48] M. Jiang, Equilibrium states for lattice models of hyperbolic type. *Nonlinearity*, **8** (5), 631–659, 1995.

[49] M. Jiang, *Ergodic properties of coupled map lattices of hyperbolic type*. Ph.D. thesis, Pennsylvania State University, 1995.

[50] M. Jiang, The entropy formula for SRB-measures of lattice dynamical systems. *J. Statist. Phys.*, **95** (3-4), 791–803, 1999.

[51] M. Jiang, Sinai-Ruelle-Bowen measures for lattice dynamical systems, 2002. Preprint.

[52] M. Jiang and A. Mazel, Uniqueness and exponential decay of correlations for some two-dimensional spin lattice systems. *J. Statist. Phys.*, **82** (3-4), 797–821, 1996.

[53] M. Jiang and Y. B. Pesin, Equilibrium measures for coupled map lattices: existence, uniqueness and finite-dimensional approximations. *Commun. Math. Phys.*, **193** (3), 675–711, 1998.

[54] K. Kaneko, Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattices. Towards a prelude of a "field theory of chaos". *Prog. Theor. Phys.*, **72** (3), 480–486, 1984.

[55] K. Kaneko, Spatiotemporal intermittency in coupled map lattices. *Prog. Theor. Phys.*, **74** (5), 1033–1044, 1985.

[56] K. Kaneko, Pattern dynamics in spatiotemporal chaos. Pattern selection, diffusion of defect and pattern competition intermittency. *Phys. D*, **34** (1-2), 1–41, 1989.

[57] K. Kaneko, Towards thermodynamics of spatiotemporal chaos. *Prog. Theoret. Phys. Suppl.*, **99**, 263–287, 1989.

[58] K. Kaneko, Clustering, coding, switching, hierarchical ordering, and control in network of chaotic elements. *Phys. D*, **41**, 137, 1990.

[59] K. Kaneko, Globally coupled chaos violates law of large numbers. *Phys. Rev. Lett.*, **65**, 1–4, 1990.

[60] K. Kaneko, editor, *Focus issue on coupled map lattices*, 1992. Special issue of Chaos, 2.

[61] K. Kaneko, editor, *Theory and Application of Coupled Map Lattices*. J. Wiley & Sons, 1993.

[62] K. Kaneko and I. Tsuda, *Complex Systems: Chaos and Beyond. A Constructive Approach with Applications in Life Sciences*. Springer, 2000.

[63] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publ. Math. IHES*, **51**, 137–173, 1980.

[64] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.

[65] G. Keller, Coupled map lattice via transfer operators on functions of bounded variations. In *Stochastic and spatial structures of dynamical systems (Amsterdam, 1995)*, 71–80, North-Holland, 1996.

[66] G. Keller, *Equilibrium States in Ergodic Theory*. Cambridge University Press, 1998.

[67] G. Keller, An ergodic theoretic approach to mean field coupled maps. In *Fractal geometry and stochastics, II (Greifswald/Koserow, 1998)*, 183–208, Birkhäuser, 2000.

[68] G. Keller and M. Künztle, Transfer operators for coupled map lattices. *Ergodic Theory Dyn. Syst.*, **12**, 297–318, 1992.

[69] G. Keller and R. Zweimüller, Unidirectionnally coupled interval maps: between dynamics and statistical mechanics. *Nonlinearity*, **15** (1), 1–24, 2002.

[70] Y. Kifer, *Ergodic Theory of Random Transformations*. Birkhauser, 1986.

[71] Y. Kifer, Large deviations in dynamical systems and stochastic processes. *Trans. AMS*, **321** (2), 505–524, 1990.

[72] Y. Kifer, Large deviations, averaging and periodic orbits of dynamical systems. *Comm. Math. Phys.*, **162** (1), 33–46, 1994.

[73] S. Lalley, Ruelle's Perron-Frobenius theorem and a central limit theorem for additive functionals of one-dimensional Gibbs states. In *Adaptative statistical procedures and related topics (Upton, 1985)*, 428–446, 1985.

[74] S. Lalley, Closed geodesics in homology classes on surfaces of variable negative curvature. *Duke Math. J.*, **58** (3), 795–821, 1989.

[75] A. Lemaître, *Systèmes chaotiques couplés : comportements collectifs et universalité*. Ph.D. thesis, Ecole Polytechnique, 1998.

[76] E. Lindenstrauss, Pointwise theorems for amenable groups. *Electron. Res. Announc. Amer. Math. Soc.*, **5**, 82–90, 1999.

[77] E. Lindenstrauss, Pointwise theorems for amenable groups. *Invent. Math.*, **146** (2), 259–295, 2001.

[78] C. Maes and A. Van Moffaert, Stochastic stability of weakly coupled lattice maps. *Nonlinearity*, **10** (3), 715–730, 1997.

[79] C. Maes and E. Verbitski, Large deviations and a fluctuation symmetry for chaotic homeomorphisms, 2002. Preprint.

[80] J. Miller and D. Huse, Macroscopic equilibrium from microscopic irreversibility in a chaotic coupled map lattice. *Phys. Rev. E*, **48**, 1185–1205, 1993.

[81] S. Olla, Large deviations for Gibbs random fields. *Probab. Th. Rel. Fields*, **77**, 343–357, 1988.

[82] H. F. ollmer, On entropy and information gain in random fields. *Z. Wahrsch. Verw. Gebiete*, **26**, 207–217, 1973.

[83] S. Orey and S. Pelikan, Large deviation principles for stationary shifts. *Ann. Probab.*, **16** (4), 1481–1495, 1988.

[84] S. Orey and S. Pelikan, Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms. *Trans. A.M.S.*, **315** (2), 741–753, 1989.

[85] D. Ornstein and B. Weiss, The Shannon-McMillan-Breiman theorem for a class of amenable groups. *Israel J. Math.*, **44** (1), 53–60, 1983.

[86] Y. Pesin and Y. Sinai, Space-time chaos in chains of weakly interacting hyperbolic mappings. In *Dynamical systems and statistical mechanics (Moscow, 1991)*, 165–198, Amer. Math. Soc., 1991.

[87] M. Pollicott, A complex Ruelle-Perron-Frobenius theorem and two counterexamples. *Ergodic Theory Dyn. Syst.*, **4**, 135–146, 1984.

[88] M. Pollicott, *Lectures on Ergodic Theory and Pesin Theory on Compact Manifolds*. Cambridge University Press, 1993.

[89] M. Pollicott, Large deviations, Gibbs measures and closed orbits for hyperbolic flows. *Math. Z.*, **220** (2), 219–230, 1995.

[90] M. Pollicott, R. Sharp and M. Yuri, Large deviations for maps with indifferent fixed points. *Nonlinearity*, **11**, 1173–1184, 1998.

[91] M. Reed and B. Simon, *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.

[92] J. Rousseau-Egele, Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux. *Ann. Probab.*, **11** (3), 772–788, 1983.

[93] D. Ruelle, *Thermodynamic Formalism*. Addison-Wesley, 1978.

[94] D. Ruelle, Thermodynamic formalism for maps satisfying positive expansiveness and specification. *Nonlinearity*, **5** (6), 1223–1236, 1992.

[95] H. Rugh, Coupled maps and analytic function spaces, 2002. To appear in Ann. Sci. ENS. Available on http://mpej.unige.ch/mp_arc/index-01.html.

[96] T. Shibata and K. Kaneko, Collective chaos. *Phys. Rev. Lett.*, **81**, 4116, 1998.

[97] T. Shibata and K. Kaneko, Tongue-like bifurcation structures of the mean-field dynamics in a network of chaotic elements. *Phys. D*, **124**, 163, 1998.

[98] Y. Takahashi, Entropy functional (free energy) for dynamical systems and their random perturbations. In *Stochastic Analysis (Kataka/Kyoto, 1982)*, 437–467, North-Holland, 1984.

[99] Y. Takahashi, Two aspects of large deviation theory for large time. In *Probabilistic Methods in Mathematical Physics (Kataka/Kyoto, 1985)*, 363–384, Academic Press, 1987.

[100] F. Takens and E. Verbitski, Multifractal analysis of local entropies for expansive homeomorphisms with specification. *Comm. Math. Phys.*, **203** (3), 593–612, 1999.

[101] H. Thunberg, Periodicity versus chaos in one-dimensional dynamics. *SIAM Rev.*, **43** (1), 3–30, 2001.

[102] E. Verbitski, *Generalized Entropies in Dynamical Systems*. Ph.D. thesis, University of Groningen, 2000.

[103] S. Waddington, Large deviation asymptotics for Anosov flows. *Ann. IHP, Anal. non Lin.*, **13** (4), 445–484, 1996.

[104] H. Wilf, *Generatingfunctionology. 2nd ed.*. Academic Press, Boston, MA, 1994.

[105] L.-S. Young, Some large deviation results for dynamical systems. *Trans. A.M.S.*, **318** (2), 525–543, 1990.

CURRICULUM VITAE

Jean-Baptiste BARDET
Born on the 31st of March 1975
French nationality
Married

Education

1998–2002	Ph-D in Mathematics under the supervision of Professor Gérard Ben Arous, Ecole Polytechnique Fédérale de Lausanne.
1998	"Agrégation de Mathématiques" (rank: 32).
1998	Advanced Studies Diploma in Mathematics, University Lyon I.
1995–1998	Student at Ecole Normale Supérieure de Lyon. Academic studies in Mathematics.
1993–1995	Preparation for the French Engineering School and ENS, Lycée du Parc, Lyon
1993	"Baccalauréat C", Lycée Lalande (Bourg en Bresse)

Professional background

2001–2003	Research and Teaching Assistant (ATER), University Paris X-Nanterre.
1999–2001	Research and Teaching Assistant of Professor Gérard Ben Arous, EPF-Lausanne.

Languages

- French, mother-tongue,
- English, good knowledge,
- German, basic knowledge.