

TRAVAIL D'ÉTUDE ET DE RECHERCHE

Introduction à la théorie des types et aux assistants de preuves.

Dorian Guillet

Encadré par Vincent Beffara.

Mai 2025.

Institut Fourier – M1 Mathématiques Générales

Table des matières

Introduction	1
I Théorie des Types	2
I.1 Introduction	2
I.2 Règles structurelles	3
I.2.1 Contexte	3
I.2.2 Égalité	3
I.3 Construction de types	4
I.3.1 Univers	4
I.3.2 Fonctions	4
I.3.3 Fonctions dépendantes (II-type)	6
I.3.4 Produits	7
I.3.5 Paires dépendantes (Σ -types)	8
I.3.6 Type vide	8
I.3.7 Type unité	9
I.3.8 Coproduit	9
I.3.9 Types inductifs : W-types	11
I.3.10 Entiers naturels	14
I.3.11 Listes	15
I.4 Correspondance preuves – programmes	18
II Formalisation du théorème de Bowen en Lean	20
II.1 Formalisation de l'énoncé du théorème	20
II.2 Preuve d'un résultat topologique	21
III Topologie des espaces ultramétriques	23
Bibliographie	26
Annexe	28

Introduction

L'objectif du TER est de découvrir un assistant de preuve en formalisant un résultat non trivial, puis de comprendre les fondements théoriques sur lesquels reposent les assistants de preuves actuels comme Lean ou Rocq. Pour ce faire, ce rapport est décomposé en trois parties distinctes.

La première partie est une présentation de la théorie des types (dépendants), dans la version introduite par Per Martin-Löf [Mar21]. Cette théorie introduit la notion fondamentale de *type* qui est un analogue des ensembles, et est largement utilisée par de nombreux langages de programmation. Elle permet de formaliser les mathématiques et de construire un lien entre informatique et mathématiques en étendant la correspondance de Curry-Howard entre les preuves et les programmes. On décrira ici seulement les règles structurant les relations entre les types et décrivant le système de preuve qu'il fournit, et enfin les liens entre les preuves et les programmes et par extension les assistants de preuve comme Lean.

La deuxième partie décrit la formalisation d'un théorème à l'aide de l'assistant de preuve Lean. Pour ce faire, on présente les outils qui ont été utilisés dans le cadre de la formalisation et les obstacles rencontrés. Initialement, le théorème choisi pour être formalisé établissait l'existence de mesure dite de Gibbs. Cependant la preuve due à Rufus Bowen [Bow75] était trop longue pour être complètement portée en Lean. Le choix final a été de montrer rigoureusement un de ses lemmes essentiels. On présente également dans cette section les reformulations des énoncés qui ont été nécessaires pour formaliser ces résultats en utilisant les types présents dans Lean.

La dernière partie est une preuve du lemme finalement choisi. Ce lemme est utile dans le théorème de Bowen pour obtenir des estimations sur la mesure des ouverts d'un espace de Bernoulli, nécessaire pour obtenir l'unicité des mesures de Gibbs. On en donne ici une version plus générale, permettant d'obtenir une partition des ouverts bornés par des boules en utilisant des espaces ultramétriques et en prouvant également des résultats élémentaires sur ses espaces métriques particuliers.

I.

Théorie des Types

I.1 Introduction

La théorie des types se veut être une alternative à la théorie des ensembles de Zermelo-Fraenkel (ZF). Les fondations de cette dernière s'appuient sur le système déductif de la logique du premier ordre, la théorie des ensembles étant formulée dans ce système. Cette théorie est donc composée de deux couches : la logique du premier ordre, puis la théorie des ensembles, formée de ses axiomes. On a donc deux objets fondamentaux dans cette approche : les propositions (qui se basent sur la logique du premier ordre) et les ensembles de ZF.

Afin d'éviter cette construction en deux couches, la théorie des types possède son propre système déductif. Un système déductif est défini par des jugements, décrivant quelles affirmations peuvent être prouvées, et des règles qui donnent la manière dont les jugements peuvent être dérivés les uns par rapport aux autres. Par conséquent, une fois les jugements et les règles du système déductif définies, il est prêt à être utilisé sans axiomes supplémentaires. Cette théorie ne possède qu'un objet fondamental : les **types**.

Pour pouvoir formuler des théorèmes, on a besoin de l'équivalent des propositions, qui sont ici des types, dont la construction suit des règles qui seront précisées dans les sections suivantes. Dans cette théorie, prouver un théorème est donc équivalent à construire un objet dont le type correspond au théorème (ici l'objet en question est une preuve).

On peut aussi voir les types d'un point de vue plus proche de la théorie des ensembles, en interprétant le fait qu'un élément a soit de type A comme l'affirmation $a \in A$. Cependant, l'affirmation $a \in A$ est une proposition alors que dire " a est de type A " (que l'on notera dorénavant $a : A$) est un jugement. En effet, dans la théorie des types un élément possède toujours un type déterminé.

Le système déductif de cette théorie est composé de trois jugements :

1. **Jugement de typage** $a : A$, qui affirme que a est de type A .
2. **Jugement d'égalité** $a \equiv b : A$, qui affirme que a et b sont égaux par définition dans le type A .
3. **Jugement de contexte** $\Gamma \text{ ctx}$, exprimant le fait que Γ est un contexte bien formé.

A noter que le symbole " \equiv " est différent de " $=$ ". En effet, si $a, b : A$, alors on a le type $a =_A b$ qui correspond à une égalité que l'on peut prouver, c'est l'égalité propositionnelle. Pour l'égalité "par définition" du système déductif, la prouver ou la supposer n'a pas réellement de sens étant donné qu'elle est vraie par définition (ou par construction).

Cette distinction entre proposition et jugement est fondamentale. Un jugement est l'affirmation qu'une proposition est vraie dans le système déductif de la théorie, alors qu'une proposition est un type, pouvant être non vide et se situe donc dans la théorie elle-même.

I.2 Règles structurelles

Pour prouver une proposition dans la théorie des types, on utilise des règles d'inférences, ou plus simplement règles, de la forme

$$\frac{\mathcal{I}_1 \quad \dots \quad \mathcal{I}_k}{\mathcal{I}} \text{NOM}$$

où \mathcal{I} est la conclusion de la règle tandis que les $\mathcal{I}_1, \dots, \mathcal{I}_k$ sont les prémisses ou hypothèses de la règle.

Avec de telles règles, on peut prouver une proposition en construisant un arbre de preuve dont la racine est le jugement à prouver et les liens entre les noeuds respectent les règles d'inférences de la théorie.

I.2.1 Contexte

Les règles concernant le jugement de contexte sont au nombre de deux et permettent de vérifier qu'un contexte Γ est bien formé, c'est-à-dire que toutes les variables et jugements apparaissant dans Γ sont dans un ordre cohérent. Pour ce faire on a ces règles :

$$\frac{}{\cdot \text{ ctx}} (\emptyset \text{ ctx}) \qquad \frac{a_1 : A_1, \dots, a_{n-1} : A_{n-1} \vdash A_n \text{ type}}{(a_1 : A_1, \dots, a_n : A_n) \text{ ctx}} (\text{ctx})$$

Le jugement " A_n type" signifie que A_n est un type et sera détaillé davantage dans la partie suivante. La première règle dit qu'un contexte vide est toujours bien formé, la seconde règle dit qu'un contexte est bien formé dès lors que la dernière variable est bien typé sous l'hypothèse que le contexte auquel on retire la dernière variable est lui-même bien formé.

I.2.2 Égalité

L'égalité par définition " \equiv " se comporte comme l'égalité classique, pour ce faire elle vérifie les mêmes propriétés élémentaires, à savoir la réflexivité, la symétrie et la transitivité que l'on exprime avec les règles suivantes.

$$\frac{\Gamma \vdash a : A}{\Gamma \vdash a \equiv a : A} (\text{refl}) \qquad \frac{\Gamma \vdash b \equiv a : A}{\Gamma \vdash a \equiv b : A} (\text{sym}) \qquad \frac{\Gamma \vdash a \equiv b : A \quad \Gamma \vdash b \equiv c : A}{\Gamma \vdash a \equiv c : A} (\text{trans})$$

On ajoute aussi deux règles supplémentaires qui permettent d'échanger le rôle de deux élément ou type sous l'hypothèse qu'ils sont égaux.

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash A \equiv B}{\Gamma \vdash a : B} (\equiv \text{ type}) \qquad \frac{\Gamma \vdash a \equiv b : A \quad \Gamma \vdash A \equiv B}{\Gamma \vdash a \equiv b : B} (\equiv \text{ eq})$$

Enfin, on ajoute une dernière règle permettant d'utiliser des jugements présent dans le contexte, qui s'exprime sous la forme suivante.

$$\frac{(a_1 : A_1, \dots, a_n : A_n) \text{ ctx}}{a_1 : A_1, \dots, a_n : A_n \vdash a_i : A_i} (\text{var})$$

I.3 Construction de types

Pour construire un nouveau type à partir d'autres, on donne généralement 3 règles :

- **Formation du type**, qui permet de construire un type à partir d'autres types ou famille de types,
- **Introduction**, qui explique comment sont construits les éléments de ce type,
- **Élimination**, décrivant comment utiliser ces éléments.

Partant de ces 3 règles, on peut construire un type et éventuellement ajouter des règles supplémentaires concernant leur comportement par rapport à l'égalité par définition.

I.3.1 Univers

Plus tôt, on a utilisé un jugement de la forme " $\Gamma \vdash A$ type" pour signifier que A est un type. Cependant ce jugement n'est pas un jugement de la théorie, on se ramène alors à un autre jugement déjà présent : le jugement de typage. Pour ce faire, on doit introduire un type dont les éléments sont des types, un type univers \mathcal{U} . Pour la même raison, on voudrait pouvoir dire que \mathcal{U} est lui-même un élément d'un type plus grand et ainsi de suite. Pour résoudre ce problème, on postule qu'il existe une hiérarchie de types

$$\mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_i, \dots$$

telle que $\mathcal{U}_i : \mathcal{U}_{i+1}$.

Pour ces types, on a seulement une règle d'introduction et une autre décrivant la hiérarchie entre ces types :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathcal{U}_i : \mathcal{U}_{i+1}} (\mathcal{U} \text{ i}) \qquad \frac{\Gamma \vdash A : \mathcal{U}_i}{\Gamma \vdash A : \mathcal{U}_{i+1}} (\mathcal{U} \text{ cumul})$$

Par conséquent, la règle sur les contextes devient la suivante :

$$\frac{(a_1 : A_1, \dots, a_{n-1} : A_{n-1}) \vdash A_n : \mathcal{U}_k}{(a_1 : A_1, \dots, a_n : A_n) \text{ ctx}} (\text{ctx})$$

I.3.2 Fonctions

Étant donné deux types $A, B : \mathcal{U}_i$, on peut former le type des fonctions de A dans B noté $A \rightarrow B$ grâce à la règle de formation suivante :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash B : \mathcal{U}_i}{\Gamma \vdash A \rightarrow B : \mathcal{U}_i} (\rightarrow \text{ f})$$

Une fonction est donc un élément f de type $A \rightarrow B$. Si on se donne $a : A$, on peut évaluer f en a ce qui donne un élément de type B que l'on note $f \ a$ ou bien $f(a)$ qui est appelé valeur

$$f(x, y) = (\lambda(a : A).f(a, y))x = ((\lambda(a : A).(\lambda(b : B).f(a, b))) x) y \quad (1)$$

$$= (\lambda(a : A).\lambda(b : B).f(a, b)) x y. \quad (2)$$

On a ainsi obtenu une fonction de type $A \rightarrow B \rightarrow C$ qui prend les mêmes valeurs que f grâce à ce processus de *curryfication*.

Exemple. Considérons deux types $A, B : \mathcal{U}_i$ et la fonction K définie par

$$K \equiv \lambda(x : A).(\lambda(y : B).x)$$

Grâce à un arbre similaire au précédent, on trouve que le type de K est $A \rightarrow B \rightarrow A$ (on omet les parenthèses en prenant comme convention que $A \rightarrow B \rightarrow C$ est égal à $A \rightarrow (B \rightarrow C)$).

I.3.3 Fonctions dépendantes (Π -type)

Une manière de généraliser le type des fonctions est de définir un type de fonctions dépendantes (ou Π -type), qui est un type de fonction dont le codomaine dépend du point d'application de la fonction.

Pour former ce nouveau type, on considère $A : \mathcal{U}_i$ un type et $B : A \rightarrow \mathcal{U}_i$ une famille de type indexée sur A . On peut alors former le type $\prod_{a:A} B(a)$ grâce à la règle de formation suivante :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma, a : A \vdash B(a) : \mathcal{U}_i}{\Gamma \vdash \prod_{a:A} B(a) : \mathcal{U}_i} (\Pi f)$$

On construit des éléments de type $\prod_{a:A} B(a)$ de manière analogue aux fonctions en utilisant les λ -abstractions. En effet, $\lambda(a : A).\Phi$ est de type $\prod_{a:A} B(a)$ si lorsque $a : A$, alors $\Phi : B(a)$. Cette règle énoncée plus formellement est la règle d'introduction de ce nouveau type :

$$\frac{\Gamma, a : A \vdash \Phi : B(a)}{\Gamma \vdash \lambda(a : A).\Phi : \prod_{a:A} B(a)} (\Pi i)$$

Ensuite, la règle d'élimination est similaire à celle des types de fonctions. Si on considère $f : \prod_{a:A} B(a)$ et $a : A$, alors $f a$ est de type $B(a)$, ce que l'on peut traduire avec la règle suivante :

$$\frac{\Gamma \vdash f : \prod_{a:A} B(a) \quad \Gamma \vdash a : A}{\Gamma \vdash f a : B(a)} (\Pi e)$$

Également, la règle β et le principe d'unicité des fonctions se généralisent au Π -type :

$$\frac{\Gamma, x : A \vdash \Phi : B(x) \quad \Gamma \vdash a : A}{\Gamma \vdash (\lambda(x : A).\Phi) a \equiv \Phi[a/x]} (\beta)$$

$$\frac{\Gamma \vdash f : \prod_{a:A} B(a)}{\Gamma \vdash f \equiv \lambda(a : A). f \ a : \prod_{a:A} B(a)} (\rightarrow \text{u})$$

L'utilité des types de fonctions dépendantes est notamment de pouvoir définir des fonctions de façon polymorphes, comme le montre les exemples suivants.

Exemple. La fonction identité du paragraphe précédent était définie sur un type $A : \mathcal{U}_i$ quelconque. On peut alors généraliser sa définition de la manière suivante :

$$I \equiv \lambda(A : \mathcal{U}_i). \lambda(a : A). a : \prod_{A:\mathcal{U}_i} A \rightarrow A.$$

Ainsi définie, la fonction I ne dépend pas d'un type préexistant.

On peut également redéfinir la fonction constante K d'une manière analogue :

$$K \equiv \lambda(A : \mathcal{U}_i). \lambda(B : \mathcal{U}_i). \lambda(x : A). \lambda(y : B). x : \prod_{A:\mathcal{U}_i} \prod_{B:\mathcal{U}_i} A \rightarrow B \rightarrow A.$$

Pour ces types et définitions polymorphes, on peut également avoir la notation I_A à la place de I A et $K_{A,B}$ pour K A B pour plus de clarté.

I.3.4 Produits

Etant donné deux types $A, B : \mathcal{U}_i$, comme en théorie des ensembles on veut construire le type $A \times B$ des paires d'éléments de A et de B . Pour construire ce type on a la règle de formation suivante :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash B : \mathcal{U}_i}{\Gamma \vdash A \times B : \mathcal{U}_i} (\times \text{ f})$$

Les éléments de ce type sont des paires d'éléments et sont donc de la forme $(a, b) : A \times B$ où $a : A$ et $b : B$, on a donc la règle d'introduction pour décrire les paires :

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B}{\Gamma \vdash (a, b) : A \times B} (\times \text{ i})$$

Ensuite pour les règles d'éliminations et de calculs, on introduit deux fonctions $\pi_1 : A \times B \rightarrow A$ et $\pi_2 : A \times B \rightarrow B$ qui projettent un élément $(a, b) : A \times B$ sur A et sur B .

$$\frac{\Gamma \vdash e : A \times B}{\Gamma \vdash \pi_1 \ e : A} (\times_1 \text{ e}) \qquad \frac{\Gamma \vdash e : A \times B}{\Gamma \vdash \pi_2 \ e : B} (\times_2 \text{ e})$$

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B}{\Gamma \vdash \pi_1 \ (a, b) \equiv a : A} (\times_1 \equiv) \quad \frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B}{\Gamma \vdash \pi_2 \ (a, b) \equiv b : B} (\times_2 \equiv)$$

Enfin, pour les types produits on a également une règle d'unicité reliant un élément à ses

projections :

$$\frac{\Gamma \vdash e : A \times B}{\Gamma \vdash e \equiv (\pi_1 e, \pi_2 e) : A \times B} (\times \text{ u})$$

On peut également voir les types produits comme un cas particulier des types de fonctions dépendantes dont le domaine est le type fini **2** à deux éléments, qui sera décrit dans I.3.8.

I.3.5 Paires dépendantes (Σ -types)

Si on considère un type $A : \mathcal{U}_i$ et une famille de type $B : A \rightarrow \mathcal{U}_i$ indexée sur A , on veut pouvoir construire l'union disjointe des $B(a)$ pour $a : A$. Pour ce faire, on introduit les types de paires dépendantes (ou Σ -type) qui se forment grâce à la règle :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma, a : A \vdash B(a) : \mathcal{U}_i}{\Gamma \vdash \Sigma_{a:A} B(a) : \mathcal{U}_i} (\Sigma \text{ f})$$

Un élément de ce type est une paire (a, b) où $a : A$ et $b : B(a)$ (d'où le nom de paire dépendante), ce que la règle suivante décrit :

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B(a)}{\Gamma \vdash (a, b) : \Sigma_{a:A} B(a)} (\Sigma \text{ i})$$

Ensuite, comme les paires dépendantes sont des paires, on a de la même manière que pour les types produits les projections $\pi_1 : \Sigma_{a:A} B(a) \rightarrow A$ et $\pi_2 : \prod_{e:\Sigma_{a:A} B(a)} B(\pi_1 e)$ sur chacune des coordonnées d'un élément de $\Sigma_{a:A} B(a)$. Ces projections se comportent de manière analogue à celle des produits, ce qu'on remarque grâce aux règles d'éliminations et de calculs des Σ -types :

$$\begin{array}{ll} \frac{\Gamma \vdash e : \Sigma_{a:A} B(a)}{\Gamma \vdash \pi_1 e : A} (\Sigma_1 \text{ e}) & \frac{\Gamma \vdash e : \Sigma_{a:A} B(a)}{\Gamma \vdash \pi_2 e : B(\pi_1 e)} (\Sigma_2 \text{ e}) \\ \frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B(a)}{\Gamma \vdash \pi_1 (a, b) \equiv a : A} (\Sigma_1 \equiv) & \frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B(a)}{\Gamma \vdash \pi_2 (a, b) \equiv b : B(a)} (\Sigma_2 \equiv) \end{array}$$

Enfin, on a également un équivalent de la règle d'unicité des types produits pour les Σ -types :

$$\frac{\Gamma \vdash e : \Sigma_{a:A} B(a)}{\Gamma \vdash e \equiv (\pi_1 e, \pi_2 e) : \Sigma_{a:A} B(a)} (\Sigma \text{ u})$$

Exemple (Cas où B est constant). Si on considère deux types $A, B : \mathcal{U}_i$, on peut poser $C \equiv \Sigma_{a:A} B$ qui est en tout point similaire au produit $A \times B$. En effet, si $(a, b) : C$ alors $a : A$ et $b : B(a) \equiv B$ et donc $(a, b) : A \times B$. De plus les règles étant très similaires, lorsque la famille B est constante on retrouve les règles des types produits.

I.3.6 Type vide

On introduit le type vide **0**, qui ne contient aucun élément et donc n'a pas de règle d'introduction mais seulement une règle de formation et d'élimination :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{0} : \mathcal{U}_i} (\mathbf{0} \text{ f}) \qquad \frac{\Gamma \vdash C : \mathcal{U}_i \quad \Gamma \vdash a : \mathbf{0}}{\Gamma \vdash \text{rec}_0(C, a) : C} (\mathbf{0} \text{ e})$$

Ainsi, on a construit une fonction rec_0 de type $\prod_{C:\mathcal{U}_i} \mathbf{0} \rightarrow C$ qui à un type $C : \mathcal{U}_i$ associe une fonction qui ne prend aucune valeur. Par conséquent, on ne devrait jamais calculer cette fonction, c'est la raison pour laquelle il n'y a pas de règle de calcul pour ce type. La règle d'élimination agit comme le principe *ex falso quodlibet*, en effet si on construit un élément du type vide (*ie.* une preuve de \perp), alors on peut construire des éléments de n'importe quel type.

I.3.7 Type unité

Ensuite, dans l'objectif de construire tous les types finis, on commence par construire le type unité $\mathbf{1}$, contenant seulement un élément. Pour ce faire on a la règle de formation analogue à celle du type $\mathbf{0}$ avec une règle d'introduction simple car le type $\mathbf{1}$ ne possède qu'un seul élément :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} (\mathbf{1} \text{ f}) \qquad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \star : \mathbf{1}} (\mathbf{1} \text{ i})$$

La règle d'élimination de $\mathbf{1}$ permet de construire des fonctions de $\mathbf{1}$ vers un autre type C , qui sont donc des fonctions constantes et c'est ce qu'affirme la règle de calcul du type unité.

$$\frac{\Gamma, x : \mathbf{1} \vdash C(x) : \mathcal{U}_i \quad \Gamma \vdash c : C(x) \quad \Gamma \vdash x : \mathbf{1}}{\Gamma \vdash \text{rec}_1(C, c, x) : C(x)} (\mathbf{1} \text{ e})$$

$$\frac{\Gamma, x : \mathbf{1} \vdash C(x) : \mathcal{U}_i \quad \Gamma \vdash c : C(\star)}{\Gamma \vdash \text{rec}_1(C, c, \star) \equiv c : C(\star)} (\mathbf{1} \equiv)$$

où la fonction rec_1 est de type $\prod_{C:\mathbf{1} \rightarrow \mathcal{U}_i} C(\star) \rightarrow \prod_{x:\mathbf{1}} C(x)$ De plus pour ce type, on a une règle d'unicité car le seul élément de $\mathbf{1}$ est \star :

$$\frac{\Gamma \vdash p : \mathbf{1}}{\Gamma \vdash p \equiv \star : \mathbf{1}} (\mathbf{1} \text{ u})$$

I.3.8 Coproduit

Le dernier élément manquant pour construire tous les types finis sont les coproduits, qui permettent de construire un équivalent de l'union disjointe de la théorie des ensembles à partir de deux types comme l'exprime la règle de formation suivante :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash B : \mathcal{U}_i}{\Gamma \vdash A + B : \mathcal{U}_i} (+ \text{ f})$$

Étant donné deux types $A, B : \mathcal{U}_i$, un élément e du type $A + B$ à deux manières d'être construit :

- soit e est issu d'un élément $a : A$ et est donc de la forme $\text{inl}(a)$,
- soit e est issu d'un élément $b : B$ et est donc de la forme $\text{inr}(b)$.

On a donc deux règles d'introduction pour les coproduits :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash B : \mathcal{U}_i \quad \Gamma \vdash a : A}{\Gamma \vdash \text{inl}(a) : A + B} (+ \text{inl}) \quad \frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash B : \mathcal{U}_i \quad \Gamma \vdash b : B}{\Gamma \vdash \text{inr}(b) : A + B} (+ \text{inr})$$

Pour prouver des propositions sur des coproduits, une manière naturelle de faire est de raisonner par cas : ou bien on a un élément de A ou bien il s'agit d'un élément de B . Pour décrire cette méthode de raisonnement, on a la règle d'élimination des coproduits pour deux types $A, B : \mathcal{U}_i$:

$$\frac{\Gamma, a : A \vdash g_0 a : C(\text{inl } a) \quad \Gamma, x : A + B \vdash C(x) : \mathcal{U}_i \quad \Gamma, b : B \vdash g_1 b : C(\text{inr } b) \quad \Gamma \vdash e : A + B}{\Gamma \vdash \text{rec}_{A+B}(C, g_0, g_1, e) : C(e)} (+ e)$$

Ainsi, la fonction rec_{A+B} est de type

$$\prod_{C:A+B \rightarrow \mathcal{U}_i} \left(\prod_{a:A} C(\text{inl } a) \right) \rightarrow \left(\prod_{b:B} C(\text{inr } b) \right) \rightarrow \left(\prod_{e:A+B} C(e) \right)$$

et permet de faire la disjonction de cas. On a de plus deux règles décrivant le comportement de la fonction rec_{A+B} en fonction de son dernier argument :

$$\frac{\Gamma, a : A \vdash g_0 a : C(\text{inl } a) \quad \Gamma, x : A + B \vdash C(x) : \mathcal{U}_i \quad \Gamma, b : B \vdash g_1 b : C(\text{inr } b) \quad \Gamma \vdash x : A}{\Gamma \vdash \text{rec}_{A+B}(C, g_0, g_1, \text{inl}(x)) \equiv g_0 x : C(\text{inl } x)} (+ \equiv_1)$$

$$\frac{\Gamma, a : A \vdash g_0 a : C(\text{inl } a) \quad \Gamma, x : A + B \vdash C(x) : \mathcal{U}_i \quad \Gamma, b : B \vdash g_1 b : C(\text{inr } b) \quad \Gamma \vdash y : B}{\Gamma \vdash \text{rec}_{A+B}(C, g_0, g_1, \text{inr}(y)) \equiv g_1 y : C(\text{inr } y)} (+ \equiv_2)$$

Exemple (Le type **2**). Avec le coproduit et le type **1** on peut désormais construire le type **2** à deux éléments. En effet, on définit **2** par $\mathbf{2} \equiv \mathbf{1} + \mathbf{1}$, ce qui donne la règle de formation

$$\frac{\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i}}{\Gamma \vdash \mathbf{1} + \mathbf{1} : \mathcal{U}_i}$$

que l'on peut résumer à :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{2} : \mathcal{U}_i} (\mathbf{2} \text{ f})$$

Ensuite, comme pour le coproduit on a deux règles d'introduction :

$$\frac{\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \star : \mathbf{1}}}{\Gamma \vdash \text{inl}(\star) : \mathbf{2}} \quad \frac{\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{1} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \star : \mathbf{1}}}{\Gamma \vdash \text{inr}(\star) : \mathbf{2}}$$

En posant $0_2 \equiv \perp \equiv \text{inl}(\star)$ et $1_2 \equiv \top \equiv \text{inr}(\star)$, on peut réécrire les deux règles d'introduction plus simplement :

$$\frac{\Gamma \text{ctx}}{\perp : \mathbf{2}} (\perp \text{ i}) \qquad \frac{\Gamma \text{ctx}}{\top : \mathbf{2}} (\top \text{ i})$$

On obtient alors la fonction rec_2 de type

$$\prod_{C: \mathbf{2} \rightarrow \mathcal{U}_i} \left(\prod_{x: \mathbf{1}} C(\text{inl } x) \right) \rightarrow \left(\prod_{y: \mathbf{1}} C(\text{inr } y) \right) \rightarrow \prod_{b: \mathbf{2}} C(b),$$

que l'on peut également voir comme le type $\prod_{C: \mathbf{2} \rightarrow \mathcal{U}} C(\perp) \rightarrow C(\top) \rightarrow \prod_{b: \mathbf{2}} C(b)$ (car les fonctions sur $\mathbf{1}$ sont constantes et égales à la valeurs qu'elles prennent sur \star). Cette fonction vérifie les règles d'élimination et de calcul suivantes :

$$\frac{\Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma, p : \mathbf{1} \vdash g_0 \ p : C(\text{inl } p) \quad \Gamma, p : \mathbf{1} \vdash g_1 \ p : C(\text{inr } p) \quad \Gamma \vdash e : \mathbf{2}}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, e) : C(e)}$$

$$\frac{\Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma, p : \mathbf{1} \vdash g_0 \ p : C(\text{inl } p) \quad \Gamma, p : \mathbf{1} \vdash g_1 \ p : C(\text{inr } p)}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, \perp) \equiv g_0 \ \perp : C(\perp)}$$

$$\frac{\Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma, p : \mathbf{1} \vdash g_0 \ p : C(\text{inl } p) \quad \Gamma, p : \mathbf{1} \vdash g_1 \ p : C(\text{inr } p)}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, \top) \equiv g_0 \ \top : C(\top)}$$

On peut les réécrire en modifiant les types de rec_2 , puisque les fonctions dont le domaine est $\mathbf{1}$ sont des constantes :

$$\frac{\Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma \vdash g_0 : C(\perp) \quad \Gamma \vdash g_1 : C(\top) \quad \Gamma \vdash e : \mathbf{2}}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, e) : C(e)} (\mathbf{2} \text{ e})$$

$$\frac{\Gamma \vdash g_0 : C(\perp) \quad \Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma \vdash g_1 : C(\top)}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, \perp) \equiv g_0 : C(\perp)} (\mathbf{2} \equiv_1) \quad \frac{\Gamma \vdash g_0 : C(\perp) \quad \Gamma, b : \mathbf{2} \vdash C(b) : \mathcal{U}_i \quad \Gamma \vdash g_1 : C(\top)}{\Gamma \vdash \text{rec}_2(C, g_0, g_1, \top) \equiv g_1 : C(\top)} (\mathbf{2} \equiv_2)$$

On remarque alors que la fonction rec_2 agit comme le "si .. alors .. sinon .." des booléens, ce qui renforce la comparaison entre le type $\mathbf{2}$ et les booléens.

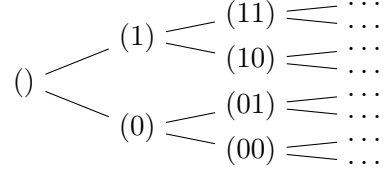
Plus généralement on peut construire tout les types finis \mathbf{T}_n comme étant n fois le coproduit du type $\mathbf{1}$, ce qui construit des types à n éléments.

I.3.9 Types inductifs : W-types

Pour construire le type des entiers \mathbf{N} , les constructeurs précédents ne suffisent pas car ils ne permettent pas de décrire le caractère inductif des entiers. Habituellement, on définit le type \mathbf{N} comme étant un élément $0 : \mathbf{N}$ et une fonction successeur $s : \mathbf{N} \rightarrow \mathbf{N}$ et un entier est de la forme 0 ou bien $s \circ s \circ \dots \circ s \ 0$. Une autre manière de le voir est de décrire \mathbf{N} comme un arbre dont 0 est une feuille et chacun des noeuds est d'arité un (*ie.* possède au plus un fils).

$$0 \text{ — } 1 \text{ — } 2 \text{ — } \dots$$

Un autre exemple de type défini de manière inductive est le type $L(A)$ des listes sur un type $A : \mathcal{U}_i$. La seule feuille est la liste vide notée $()$ et la fonction jouant le rôle de constructeur inductif est la fonction $c : A \rightarrow L(A) \rightarrow L(A)$ qui a un élément $a : A$ et une liste $l : L(A)$ renvoie la liste dont le premier élément est a et le reste est l . Ainsi, toutes les listes sont de la forme $() : L(A)$ ou bien $c\ a_1\ (c\ a_2\ (\dots (c\ a_n\ ()))) : L(A)$.



Pour généraliser les types inductifs, on introduit alors les W-types (*well-founded trees*) de Martin-Löf. Pour ce faire, on utilise un type $A : \mathcal{U}_i$ qui correspond aux *étiquettes* et un type $B : A \rightarrow \mathcal{U}$ qui permet de coder l'arité des noeuds : un noeud dont l'étiquette est $a : A$ aura une arité de $B(a)$. Le type résultant de ces deux types est $W_{a:A}B(a)$ que l'on construit via la règle de formation suivante :

$$\frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma, a : A \vdash B(a) : \mathcal{U}_i}{\Gamma \vdash W_{a:A}B(a) : \mathcal{U}_i} \text{ (W f)}$$

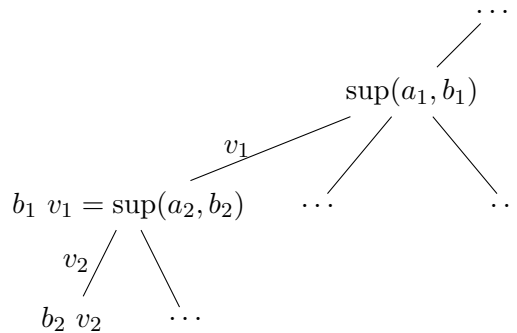
Les éléments d'un type $W_{a:A}B(a)$ sont donc des noeuds d'un arbre auxquels on associe la fonction donnant les noeuds précédents dans l'arbre. On construit ces éléments à l'aide de la règle d'introduction pour les W-types :

$$\frac{\Gamma \vdash a : A \quad \Gamma, a : A \vdash b : B(a) \rightarrow W_{x:A}B(x)}{\Gamma \vdash \text{sup}(a, b) : W_{x:A}B(x)} \text{ (W i)}$$

La fonction sup utilisée dans cette règle est donc de type

$$\text{sup} : \prod_{a:A} (B(a) \rightarrow W_{x:A}B(x)) \rightarrow W_{a:A}B(a).$$

Chaque noeud $c \equiv \text{sup}(a, b) : W_{x:A}B(x)$ est donc relié aux autres par la fonction $b : B(a) \rightarrow W_{x:A}B(x)$, en effet si on considère un élément $v : B(a)$ alors $b\ v$ est un prédécesseur du noeud c dans l'arbre $W_{x:A}B(x)$.



Reprenons l'exemple du type \mathbf{N} des entiers. On peut le décrire en terme de W-type en prenant

$\mathbf{N} \equiv W_{x:\mathbf{2}}B(x)$ où $B(0_2) \equiv \mathbf{0}$ et $B(1_2) \equiv \mathbf{1}$. En effet, il y a deux manières de construire un entier : ou bien c'est 0 ou bien il est de la forme $s\ k$ avec $k : \mathbf{N}$. Pour 0, son étiquette dans l'arbre est donc 0_2 tandis que les éléments de la forme $s\ k$ ont l'étiquette 1_2 . On peut donc écrire 0 avec la fonction sup :

$$0 \equiv \text{sup}(0_2, \text{rec}_{\mathbf{0}, \mathbf{N}}).$$

De même la fonction successeur $s : \mathbf{N} \rightarrow \mathbf{N}$ peut être définie de la manière suivante :

$$s \equiv \lambda n. \text{sup}(1_2, \lambda(x : \mathbf{1}). n) : \mathbf{N} \rightarrow \mathbf{N}.$$

Et donc on a $1 \equiv s\ 0 \equiv \text{sup}(1_2, \lambda x. 0)$ etc.

La règle d'élimination décrit le principe d'induction sur les W-types. Pour prouver une proposition sur un type inductif, de manière analogue au principe de récurrence des entiers, on montre que la proposition est héréditaire, dans le cas présent cela revient à montrer que si la proposition est vraie sur tout les sous-arbres, alors elle est vraie sur le noeud reliant ces sous-arbres, c'est ce que dit la dernière prémisse de la règle d'élimination, qui s'énonce de la manière suivante :

$$\frac{\begin{array}{c} \Gamma, a : A, b : B(a) \rightarrow W_{x:A}B(x), \\ \Gamma \vdash t : W_{x:A}B(x) \quad c : \prod_{v:B(a)} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(a, b)) \end{array}}{\Gamma \vdash \text{rec}_{W_{x:A}B(x)}(E, e, t) : E(t)} \text{ (W e)}$$

La fonction e présente dans la règle est de type

$$e : \prod_{a:A} \prod_{b:(B(a) \rightarrow W_{x:A}B(x))} \left(\prod_{v:B(a)} E(b\ v) \right) \rightarrow E(\text{sup}(a, b)).$$

et pour des éléments $a : A$, $b : B(a) \rightarrow W_{x:A}B(x)$ et une preuve c de $\prod_{v:B(a)} E(b\ v)$ renvoie une preuve de $E(\text{sup}(a, b))$. La fonction $\text{rec}_{W_{x:A}B(x)}$ présente dans la conclusion de la règle précédente est une fonction qui à un prédicat (ou famille de type) E sur le type $W_{x:A}B(x)$, une fonction e comme décrit juste avant et un élément $t : W_{x:A}B(x)$ retourne une preuve (ou un élément) de $E(t)$. Ainsi, la fonction e et la fonction $\text{rec}_{W_{x:A}B(x)}$ renvoie des preuves de même nature, on peut donc imaginer une relation liant les deux : c'est la règle d'égalité (ou règle de calcul de $\text{rec}_{W_{x:A}B(x)}$).

$$\frac{\begin{array}{c} \Gamma \vdash a : A \\ \Gamma \vdash b : B(a) \rightarrow W_{x:A}B(x) \\ \Gamma, x : A, y : B(x) \rightarrow W_{x:A}B(x), z : \prod_{v:B(x)} E(y\ v) \vdash e(x, y, z) : E(\text{sup}(x, y)) \end{array}}{\Gamma \vdash \text{rec}_{W_{x:A}B(x)}(E, e, \text{sup}(a, b)) \equiv e(a, b, \lambda v. \text{rec}_{W_{x:A}B(x)}(E, e, b\ v)) : E(\text{sup}(a, b))} \text{ (W } \equiv \text{)}$$

En effet, pour $a : A$ et $b : B(a) \rightarrow W_{x:A}B(x)$ la fonction $\lambda(v : B(a)). \text{rec}_{W_{x:A}B(x)}(E, e, b\ v)$

retourne une preuve (un élément) de $\prod_{v:B(a)} E(b\ v)$, ce qui correspond au dernier paramètre de la fonction e et donc le membre de droite de l'égalité est bien une preuve (un élément) de $E(\text{sup}(a, b))$ comme voulue. Les prémisses de cette règle sont de simple vérifications de typage de a , b et e .

I.3.10 Entiers naturels

Avec les W-types, nous pouvons désormais construire les entiers naturels. Pour ce faire, on note $\mathbf{N} \equiv W_{x:\mathbf{2}} B(x)$ où $B(0_{\mathbf{2}}) \equiv \mathbf{0}$ et $B(1_{\mathbf{2}}) \equiv \mathbf{1}$ (ie. $B \equiv \text{rec}_{\mathbf{2}}(\mathcal{U}_i, \mathbf{0}, \mathbf{1}) : \mathbf{2} \rightarrow \mathcal{U}_i$). On a déjà vu dans le paragraphe précédent que tout élément de \mathbf{N} s'exprime sous la forme 0 ou $s\ k$ avec $k : \mathbf{N}$ et de plus on avait les expressions de 0 et de $s : \mathbf{N} \rightarrow \mathbf{N}$ suivantes :

$$\begin{cases} 0 & \equiv \text{sup}(0_{\mathbf{2}}, \text{rec}_{\mathbf{0}, \mathbf{N}}) & : \mathbf{N} \\ s & \equiv \lambda(n : \mathbf{N}). \text{sup}(1_{\mathbf{2}}, \lambda(x : \mathbf{1}). n) & : \mathbf{N} \rightarrow \mathbf{N} \end{cases}$$

On peut également exprimer les règles de formation, d'introduction, d'élimination et d'égalité dans le cas de \mathbf{N} . Pour la règle de formation, on obtient :

$$\frac{\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{2} : \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma, a : \mathbf{2} \vdash B : \mathbf{2} \rightarrow \mathcal{U}_i} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{2} : \mathcal{U}_i} \quad \frac{\Gamma \vdash \mathbf{2} : \mathcal{U}_i \quad \Gamma, a : \mathbf{2} \vdash B(a) : \mathcal{U}_i}{\Gamma, a : \mathbf{2} \vdash B(a) : \mathcal{U}_i}}{\Gamma \vdash \mathbf{N} : \mathcal{U}_i}$$

On a donc la règle simplifiée :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash \mathbf{N} : \mathcal{U}_i} \mathbf{N\ i}$$

Ensuite, concernant la règle d'introduction on distingue deux cas puisqu'il y a deux manières de construire un élément $n : \mathbf{N}$, soit $n \equiv 0$, soit $n \equiv s\ k$ avec $k : \mathbf{N}$.

$$\frac{\frac{\Gamma \text{ ctx}}{\Gamma \vdash 0_{\mathbf{2}} : \mathbf{2}} \quad \frac{\Gamma, 0_{\mathbf{2}} : \mathbf{2} \vdash \text{rec}_{\mathbf{0}, \mathbf{N}} : \mathbf{0} \rightarrow \mathbf{N}}{\Gamma, 0_{\mathbf{2}} : \mathbf{2} \vdash \text{rec}_{\mathbf{0}, \mathbf{N}} : B(0_{\mathbf{2}}) \rightarrow \mathbf{N}}}{\frac{\Gamma \vdash \text{sup}(0_{\mathbf{2}}, \text{rec}_{\mathbf{0}, \mathbf{N}}) : \mathbf{N}}{\Gamma \vdash 0 : \mathbf{N}}}}$$

$$\frac{\frac{\Gamma \vdash s : \mathbf{N} \rightarrow \mathbf{N}}{\Gamma \vdash s : \mathbf{N} \rightarrow \mathbf{N}} \quad \Gamma \vdash k : \mathbf{N}}{\Gamma \vdash s\ k : \mathbf{N}}$$

Ce qui donne les deux règles d'introduction de \mathbf{N} :

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash 0 : \mathbf{N}} (0\ \text{i}) \quad \frac{\Gamma \vdash k : \mathbf{N}}{\Gamma \vdash s\ k : \mathbf{N}} (s\ \text{i})$$

Pour la règle d'élimination sur le type \mathbf{N} , on s'attend à retrouver le principe de récurrence

usuel sur les entiers, à savoir que si un prédicat $E : \mathbf{N} \rightarrow \mathcal{U}_i$ est vrai sur 0 et que si $E(k)$ est vrai alors $E(s\ k)$ est vrai, alors $E(k)$ est vrai sur tout $k : \mathbf{N}$.

$$\frac{\Gamma \vdash k : \mathbf{N} \quad \frac{\text{Lemme 1} \quad \text{Lemme 2}}{\Gamma, a : \mathbf{2}, b : B(a) \rightarrow \mathbf{N}, c : \prod_{v:B(a)} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(a, b))}}{\Gamma \vdash \text{rec}_{\mathbf{N}}(E, e, k) : E(k)}$$

Avec

$$\frac{\Gamma, a : \mathbf{1}, b : \mathbf{0} \rightarrow \mathbf{N}, c : \prod_{v:\mathbf{0}} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(\mathbf{0}\mathbf{2}, b))}{\Gamma, a : \mathbf{1}, b : B(\text{inl}(a)) \rightarrow \mathbf{N}, c : \prod_{v:B(\text{inl}(a))} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(a, b))} \text{Lemme 1}$$

et

$$\frac{\Gamma, a : \mathbf{1}, b : \mathbf{1} \rightarrow \mathbf{N}, c : \prod_{v:\mathbf{1}} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(\mathbf{1}\mathbf{2}, b))}{\Gamma, a : \mathbf{1}, b : B(\text{inr}(a)) \rightarrow \mathbf{N}, c : \prod_{v:B(\text{inr}(a))} E(b\ v) \vdash e(a, b, c) : E(\text{sup}(\text{inr}(a), b))} \text{Lemme 2}$$

On peut simplifier les prémisses du Lemme 1 en remarquant que ses hypothèses sont triviales et que $\text{sup}(\mathbf{0}\mathbf{2}, b) \equiv 0$ pour toute fonction $b : \mathbf{0} \rightarrow \mathbf{N}$, ce qui donne après simplification :

$$\Gamma \vdash e_0 : E(0).$$

De même, pour le Lemme 2 en remarquant qu'une fonction $b : \mathbf{1} \rightarrow \mathbf{N}$ est équivalente à choisir un $k : \mathbf{N}$ et que $c : \prod_{v:\mathbf{1}} E(b\ v)$ est similaire à $c : E(k)$, on peut réécrire la prémisse en

$$\Gamma, k : \mathbf{N}, p : E(k) \vdash e_s(k, p) : E(s\ k).$$

Finalement, on peut décomposer la fonction e du récursur : $e_0 : E(0)$ et $e_s : \prod_{k:\mathbf{N}} E(k) \rightarrow E(s\ k)$, ce qui permet de réécrire la règle d'élimination sous la forme plus simple :

$$\frac{\Gamma \vdash k : \mathbf{N} \quad \Gamma \vdash e_0 : E(0) \quad \Gamma, k : \mathbf{N}, p : E(k) \vdash e_s(k, p) : E(s\ k)}{\Gamma \vdash \text{rec}_{\mathbf{N}}(E, e_0, e_s, k) : E(k)} \mathbf{N\ e}$$

On retrouve bien le principe de récurrence usuel, en effet pour prouver un prédicat sur les entiers \mathbf{N} il suffit de le prouver sur 0 puis qu'il est héréditaire *ie.* s'il est vérifié sur un entier $k : \mathbf{N}$ alors il est vérifié sur $s\ k$ l'entier suivant.

I.3.11 Listes

Un autre exemple de type inductif est le type des listes sur un type $A : \mathcal{U}_i$, que l'on note $L(A)$. Pour la construire on utilise alors le W-type $W_{x:\mathbf{1}+A}B(x)$ où $B(\text{inl}(\star)) \equiv \mathbf{0}$ et $B(\text{inr}(a)) \equiv \mathbf{1}$,

autrement dit on pose

$$L(A) \equiv W_{x:\mathbf{1}+A} \text{rec}_{\mathbf{1}+A}(\mathcal{U}_i, \lambda(x:\mathbf{1}). \mathbf{0}, \lambda(a:A). \mathbf{1}, x).$$

De là, on déduit la règle de formation de $L(A)$:

$$\frac{\frac{\Gamma \text{ctx}}{\Gamma \vdash A : \mathcal{U}_i} \quad \frac{\Gamma \vdash \mathbf{1} : \mathcal{U}_i}{\Gamma \vdash \mathbf{1} + A : \mathcal{U}_i}}{\Gamma \vdash L(A) : \mathcal{U}_i} \quad \frac{\Gamma, x : \mathbf{1} + A \vdash \text{rec}_{\mathbf{1}+A}(\mathcal{U}_i, \lambda(y:\mathbf{1}). \mathbf{0}, \lambda(y:A). \mathbf{1}, x) : \mathcal{U}_i}{\Gamma, x : \mathbf{1} + A \vdash B(x) : \mathcal{U}_i}$$

Plus simplement,

$$\frac{\Gamma \vdash A : \mathcal{U}_i}{\Gamma \vdash L(A) : \mathcal{U}_i} L \text{ f}$$

Ensuite pour construire un élément on a deux constructeurs :

- La liste vide, notée $()$ que l'on définit comme $() \equiv \text{sup}(\text{inl}(\star), \text{rec}_{\mathbf{0}}(L(A))) : L(A)$,
- L'ajout d'un élément en tête de liste via la fonction $c : A \rightarrow L(A) \rightarrow L(A)$ qui est défini de la manière suivante :

$$c \equiv \lambda(a:A). \lambda(l:L(A)). \text{sup}(a, \lambda(x:\mathbf{1}). l).$$

que l'on peut également noter $::$ de manière infixé.

Ainsi, une liste $l \equiv (a_1, \dots, a_k) : L(A)$ est définie par

$$l \equiv c \ a_1 \ (a_2, \dots, a_k) \equiv a_1 :: (a_2, \dots, a_k) \equiv \text{sup}(a_1, \lambda(x:\mathbf{1}). (a_2, \dots, a_k)).$$

On remarque alors une structure d'arbre dans le type des listes, dont la racine est $()$ et la fonction donnant les successeurs d'un noeud l est la fonction $\lambda(a:A). c \ a \ l : A \rightarrow L(A)$.

On a donc deux règles d'introduction correspondant à ces deux cas.

$$\frac{\frac{\Gamma \vdash \mathbf{1} : \mathcal{U}_i \quad \Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash \star : \mathbf{1}}{\Gamma \vdash \text{inl}(\star) : \mathbf{1} + A} \quad \frac{\Gamma \vdash \text{rec}_{\mathbf{0}}(L(A)) : \mathbf{0} \rightarrow L(A)}{\Gamma \vdash \text{rec}_{\mathbf{0}}(L(A)) : B(\text{inl}(\star)) \rightarrow L(A)}}{\Gamma \vdash \text{sup}(\text{inl}(\star), \text{rec}_{\mathbf{0}}(L(A))) : L(A)} \quad \Gamma \vdash () : L(A)$$

$$\frac{\frac{\Gamma \vdash \mathbf{1} : \mathcal{U}_i \quad \Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash a : A}{\Gamma \vdash \text{inr}(a) : \mathbf{1} + A} \quad \frac{\frac{\Gamma \vdash l : L(A)}{\Gamma, x : \mathbf{1} \vdash l : L(A)} \quad \Gamma \vdash \lambda(x:\mathbf{1}). l : \mathbf{1} \rightarrow L(A)}{\Gamma \vdash \lambda(x:\mathbf{1}). l : \text{rec}_{\mathbf{1}+A}(\mathcal{U}_i, \mathbf{0}, \mathbf{1}, \text{inr}(a)) \rightarrow L(A)}}{\Gamma \vdash \text{sup}(\text{inr}(a), \lambda(x:\mathbf{1}). l) : L(A)} \quad \Gamma \vdash a :: l : L(A)$$

Plus simplement,

$$\frac{\Gamma \vdash A : \mathcal{U}_i}{\Gamma \vdash () : L(A)} \text{ () i} \quad \frac{\Gamma \vdash A : \mathcal{U}_i \quad \Gamma \vdash a : A \quad \Gamma \vdash l : L(A)}{\Gamma \vdash a :: l : L(A)} c \text{ i}$$

A partir de la règle d'élimination des W-types, on peut déduire la règle d'élimination du type $L(A)$:

$$\frac{\begin{array}{c} \text{Lemme 1} \quad \text{Lemme 2} \\ \Gamma \vdash l : L(A) \quad \Gamma, a : \mathbf{1} + A, b : B(a) \rightarrow L(A), c : \prod_{v:B(a)} E(b \ v) \vdash e(a, b, c) : E(\text{sup}(a, b)) \end{array}}{\Gamma \vdash \text{rec}_{L(A)}(E, e, l) : E(l)}$$

$$\frac{\Gamma, a : \mathbf{1}, b : \mathbf{0} \rightarrow L(A), c : \prod_{v:\mathbf{0}} E(b \ v) \vdash e(a, b, c) : E(())}{\Gamma, a : \mathbf{1}, b : B(\text{inl}(a)) \rightarrow L(A), c : \prod_{v:B(\text{inl}(a))} E(b \ v) \vdash e(a, b, c) : E(\text{sup}(\text{inl}(a), b))} \text{Lemme 1}$$

$$\frac{\Gamma, a : A, b : \mathbf{1} \rightarrow L(A), c : \prod_{v:\mathbf{1}} E(b \ v) \vdash e(a, b, c) : E(a :: b)}{\Gamma, a : A, b : B(\text{inr}(a)) \rightarrow L(A), c : \prod_{v:B(\text{inr}(a))} E(b \ v) \vdash e(a, b, c) : E(\text{sup}(\text{inr}(a), b))} \text{Lemme 2}$$

Le lemme 1 correspond à une preuve de E sur la liste vide $()$ dont les hypothèses sont triviales. On peut donc réécrire sa prémisse sous une forme plus simple

$$\Gamma \vdash e_0 : E(())$$

De même, on peut simplifier le lemme 2 en remarquant que $b : \mathbf{1} \rightarrow L(A)$ est équivalent à une liste $l \equiv b \star$, et que $c : \prod_{v:\mathbf{1}} E(b \ v)$ est une preuve de $E(l)$:

$$\Gamma, a : A, l : L(A), p : E(l) \vdash e_s(a, l, p) : E(a :: l).$$

On peut aussi réécrire e en le décomposant en $e_0 : E(())$ et $e_s : \prod_{a:A} \prod_{l:L(A)} E(l)$ pour correspondre à ces deux cas, ce qui donne finalement la règle

$$\frac{\Gamma \vdash l : L(A) \quad \Gamma \vdash e_0 : E(()) \quad \Gamma, a : A, l : L(A), p : E(l) \vdash e_s(a, l, p) : E(a :: l)}{\Gamma \vdash \text{rec}_{L(A)}(E, e_0, e_s, l) : E(l)} L \text{ e}$$

Ce qui correspond à la preuve par induction sur les listes, à savoir que si un prédicat E sur $L(A)$ est vraie sur la liste vide $()$ et qu'il est héréditaire *ie.* si E est vraie sur une liste l et si $a : A$ alors E est vérifié sur $a :: l$, alors E est vérifié sur toutes les listes.

I.4 Correspondance preuves – programmes

Dans la théorie des types, les propositions sont des types et les preuves de ces propositions sont des éléments du type correspondant. Ainsi, prouver un théorème correspond à construire un élément (ou un terme) du type correspondant à l'énoncé du théorème. On a décrit comment construire des types de manière générale et ces constructions peuvent être utilisés pour construire des propositions, comme décrit dans le tableau suivant.

Logique	Type
$A \implies B$	$A \rightarrow B$
$A \wedge B$	$A \times B$
$A \vee B$	$A + B$
\perp	$\mathbf{0}$
\top	$\mathbf{1}$
$\neg A$	$A \rightarrow \mathbf{0}$
$\forall x : A, P(x)$	$\prod_{x:A} P(x)$
$\exists x : A, P(x)$	$\sum_{x:A} P(x)$

On peut donc voir les règles d'inférences comme des règles de construction des preuves dans la théorie des types. Par exemple, pour prouver une implication $A \implies B$, en théorie des types il faut construire une fonction $f : A \rightarrow B$. Or pour construire une fonction f du type $A \rightarrow B$, on doit donner une manière de construire une preuve de B à partir d'une preuve de A , c'est-à-dire que si $a : A$ est une preuve de A , $f a$ est une preuve de B et ce pour toute preuve a de A . On peut alors voir la fonction f comme un algorithme qui donne une preuve de B à partir de n'importe quelle preuve de A .

Ainsi, les règles d'introduction des types donnent une manière de construire des éléments du type en question, et donc si ce type représente une proposition alors on peut voir la règle d'introduction comme une règle de construction de preuve. De la même manière les règles d'élimination donnent la manière d'utiliser une hypothèse.

Enfin, les deux derniers types du tableau sont les quantificateurs universels et existentiels et leurs équivalents : les types dépendants. Le caractère dépendant de ces types se situe dans le fait qu'ils dépendent d'un paramètre : pour les fonctions dépendantes le codomaine dépend du point d'application, pour les paires dépendantes c'est le type de la deuxième coordonnées qui dépend de la première coordonnée. Le caractère dépendant de ces types permet notamment la construction de polymorphismes, comme détaillé dans la section sur les fonctions dépendantes.

Prouver une proposition en théorie des types est équivalent à construire un élément d'un certain type. Pour ce faire, les assistants de preuves permettent de simplifier la création de terme (ou d'éléments). Par exemple, la proposition

$$(\neg(p \wedge q) \implies r) \implies (\neg p \implies r) \wedge (\neg q \implies r).$$

correspond au type

$$\begin{aligned} & (\neg(p \times q) \rightarrow r) \rightarrow (\neg p \rightarrow r) \times (\neg q \rightarrow r) \\ \equiv & (((p \times q) \rightarrow \mathbf{0}) \rightarrow r) \rightarrow ((p \rightarrow \mathbf{0}) \rightarrow r) \times ((q \rightarrow \mathbf{0}) \rightarrow r). \end{aligned}$$

dont un terme est

$$\lambda h. (\lambda x. h (\lambda a. x (\pi_1 a)), \lambda y. h (\lambda a. y (\pi_2 a)))$$

En effet si on considère $h : ((p \times q) \rightarrow \mathbf{0}) \rightarrow r$, pour construire un terme de type $(p \rightarrow \mathbf{0}) \rightarrow r$ (*resp.* $(q \rightarrow \mathbf{0}) \rightarrow r$), on se donne $x : p \rightarrow \mathbf{0}$ (*resp.* $y : q \rightarrow \mathbf{0}$), puis on cherche à construire un terme de type r dans les deux cas. Pour ce faire, on applique la fonction h à un terme du type $(p \times q) \rightarrow \mathbf{0}$ que l'on doit construire. Or, $\pi_1 a$ (*resp.* $\pi_2 a$) est de type p (*resp.* q) si a est de type $p \times q$, donc $x (\pi_1 a)$ (*resp.* $y (\pi_2 a)$) est de type $\mathbf{0}$. Les fonctions $\lambda a. x (\pi_1 a)$ et $\lambda a. y (\pi_2 a)$ sont de type $p \times q \rightarrow \mathbf{0}$, on peut alors appliquer h à chacune de ses deux fonctions pour obtenir un terme de type r comme voulu.

La construction d'une telle fonction est parfois compliquée, c'est pourquoi les assistants de preuves possèdent un système de construction de termes à partir de commande plus simple ayant un sens qui se rapproche des règles habituelles de raisonnement. Reprenons l'exemple précédent. En Lean, le code suivant qui permet de construire un terme de ce type :

```
example : (¬(p ∧ q) → r) → (¬p → r) ∧ (¬q → r) := by
  intro h                -- Introduit un terme de type ¬(p ∧ q) → r
  constructor            -- Sépare l'objectif en deux sous-objectifs
  <;> intro hn            -- Introduit un terme de type ¬p (resp. ¬q)
                          -- dans le premier (resp. second) sous-objectif
  <;> apply h             -- Applique h aux deux sous-objectifs
  <;> rintro <hp, hq>     -- Introduit des termes de type p et q
  . exact hn hp
  . exact hn hq          -- Construit des termes de type 0
```

Le terme créé par Lean à partir de ce code est :

```
fun h =>
  <fun hn => h fun a => And.casesOn a fun hp hq => hn hp,
  fun hn => h fun a => And.casesOn a fun hp hq => hn hq>
```

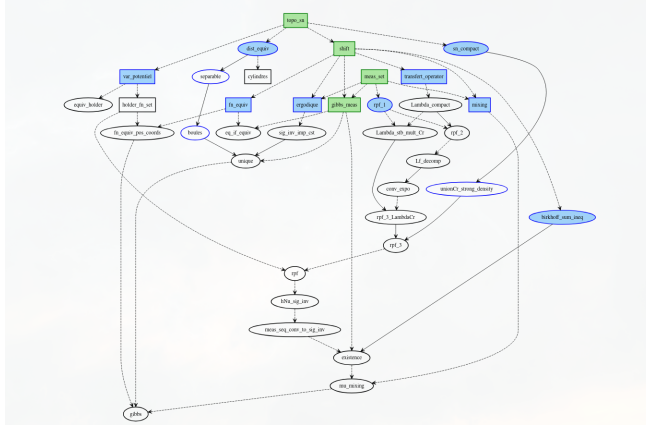
II.

Formalisation du théorème de Bowen en Lean

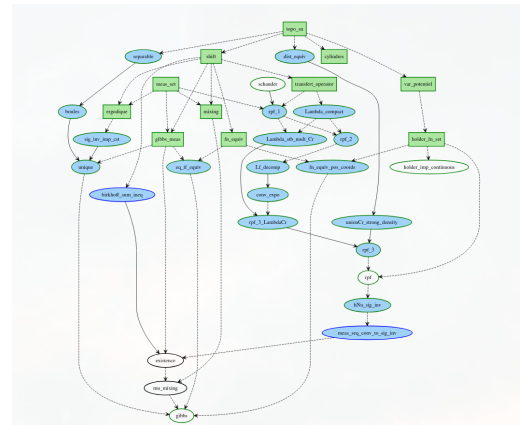
Cette section est un retour d'expérience sur l'utilisation de Lean en tant qu'outil de formalisation pour les mathématiques.

Durant ce TER, j'ai utilisé Lean pour formaliser le théorème de Bowen, que j'ai eu l'occasion d'étudier durant un stage et dont une preuve figure dans mon rapport de stage en annexe (preuve initialement due à R. Bowen). Ce théorème établit l'existence et l'unicité d'une certaine mesure de probabilité dite de Gibbs sur un espace métrique particulier, et donc ce théorème fait intervenir notamment de la topologie et de la théorie de la mesure.

Pour formaliser ce théorème, j'ai eu recours à des outils extérieurs à Lean comme `leanblueprint` [Mas] qui permet à partir d'un texte mathématiques d'obtenir un graphe décrivant les liens entre les différents lemmes, théorèmes et définitions intermédiaires et d'en donner l'état d'avancement dans le code Lean associé.



(a) État initial du graphe



(b) État final du graphe

Les cases rectangulaires sont les définitions, les ellipses sont les lemmes, théorèmes, corollaires et propositions. Les couleurs en fond indiquent l'état d'avancement des preuves : le vert foncé pour les lemmes prouvés et dont les parents sont également prouvés, le vert clair pour les lemmes prouvés et le bleu pour les lemmes prêt à être prouvés (*ie.* toutes les définitions nécessaires sont formalisées). Enfin les couleurs des bordures indiquent l'état d'avancement des énoncés des lemmes et des définitions : le vert pour les cases dont l'énoncé est formalisé et le bleu pour les cases dont l'énoncé n'est pas formalisé mais prêt à l'être.

II.1 Formalisation de l'énoncé du théorème

Dans un premier temps, j'ai donc converti chaque définition et résultat nécessaire pour le théorème de Bowen en code Lean afin de pouvoir écrire l'énoncé du théorème. Pour ce faire on a besoin notamment de définir la notion de mesure de Gibbs puis on peut alors énoncé le théorème avec le code suivant :

```

class IsGibbs (ϕ : Bernoulli Z n → R) (μ : Measure (Bernoulli Z n))
extends InvariantProb μ where
  gibbs_prop : Exists P : R, Exists c c' : NNReal, c > 0 ∧ c' > 0 ∧
    Forall x : Bernoulli Z n, Forall m : N,
      μ (cylinder x m) / nnexp (- P * m + ∑ k ∈ Ico 0 m, ϕ (shift^[k] x))
        ∈ Icc (c : ENNReal) (c' : ENNReal)

theorem gibbs (ϕ : Bernoulli Z n → R) [HolderLike ϕ] :
  ExistsUnique μ : Measure (Bernoulli Z n), IsGibbs ϕ μ := by sorry
    
```

Cependant, la preuve de ce résultat nécessite un certain nombre de lemmes intermédiaires qu'il faut également prouver et donc la preuve du théorème était trop longue par rapport au temps imparti.

On a donc décidé de se concentrer sur la formalisation des énoncés et sur la preuve d'un lemme important de topologie des espaces ultramétriques, pour expérimenter tout les aspects d'un assistant de preuves.

Pour la formalisation des énoncés il a donc fallu dans un premier temps adapter le rapport de stage au format demandé par l'outil lean-blueprint, en ajoutant des détails et en décomposant certains énoncés pour correspondre davantage à la structure du code Lean. L'étape suivante était de traduire chaque définition et lemme (sans les preuves) en un code le plus clair possible pour faciliter le travail d'écriture des preuves suivant l'ordre préconisé par le graphe de dépendance. Enfin, j'ai pu commencer l'écriture des morceaux de preuves pour les lemmes les plus simples. Finalement, au vu du nombre de preuve à convertir en code, il a fallu choisir un résultat dont la preuve était moins conséquente pour s'y consacrer.

II.2 Preuve d'un résultat topologique

Le choix qui a été fait était de prouver un théorème de topologie permettant de décomposer les ouverts d'un espace ultramétrique compact en une union disjointe de boules. Ce résultat était nécessaire pour réécrire la mesure des ouverts en somme de mesure de boules pour lesquelles on possède une estimation dans le cas des mesures de Gibbs.

Afin de prouver ce résultat, la première étape était de réécrire de manière la plus détaillée possible sa preuve et d'en donner les étapes intermédiaires pour ensuite les convertir en lemme dans Lean pour pouvoir écrire les preuves de ces lemmes. L'étape finale a donc été d'écrire la preuve du théorème en utilisant tout les lemmes précédents.

Cependant, la traduction des preuves en Lean ne se fait pas littéralement, il faut souvent adapter les énoncés pour les faire correspondre aux lemmes déjà existants et prouvés dans Mathlib (la bibliothèque de Lean contenant une large variété de théorèmes). Ainsi, nous avons donc transformés l'énoncé mathématique standard en un énoncé utilisant à la fois la théorie des types et des propriétés déjà présentes dans Mathlib.

Théorème. *Soit \mathcal{O} un ouvert borné de E , alors il existe une partie $R \subseteq \mathcal{O}$ et une fonction $r : R \rightarrow \mathbf{R}_+^*$ telle que*

$$\mathcal{O} = \bigsqcup_{x \in R} B(x, r(x)).$$

Cet énoncé transformé nécessite l'introduction d'un type $S_{\mathcal{O}}$ construit comme l'ensemble quotient de l'ensemble $B_{\mathcal{O}}$ des boules incluses dans \mathcal{O} par une relation d'équivalence, qui sera décrite dans la section suivante, et d'une fonction $\Phi : S_{\mathcal{O}} \rightarrow B$ qui à chaque classe d'équivalence associe l'union de ses éléments.

Théorème. *Soit \mathcal{O} un ouvert borné de E . Alors la fonction Φ vérifie les propriétés suivantes :*

1. $\bigcup_{s:S_{\mathcal{O}}} \Phi(s) = \mathcal{O}$
2. Si $s_1, s_2 : S_{\mathcal{O}}$ sont distincts alors $\Phi(s_1)$ et $\Phi(s_2)$ sont disjoints.

On peut alors remarquer que les deux énoncés sont en fait équivalents, il s'agit simplement d'une reformulation plus maniable en Lean d'un même théorème, car elle fait intervenir des types et des fonctions largement utilisées dans Mathlib et qui sont donc plus faciles d'utilisation. Une fois converti en Lean, il peut s'écrire avec le code suivant :

```
theorem partition (hO : IsOpen O)
  (O_bdd : Exists (x : X), Exists r > 0, O ⊆ ball x r) :
  let Φ := iBall (Obdd O_bdd)
  ⋃ s, Φ s = O ∧ Pairwise (onFun Disjoint (fun s => (Φ s).1)) := by sorry
```

Une preuve de ce théorème se trouve dans la section suivante et une formalisation de ce théorème est disponible sur [Github](#).

III.

Topologie des espaces ultramétriques

Pour montrer l'unicité dans le théorème de Bowen, on a besoin d'un résultat de topologie de certains espaces, appelés espaces ultramétriques, qui sont des espaces métriques où l'inégalité triangulaire habituelle est remplacée par une inégalité dite ultramétrique :

$$\forall x, y, z \in E, d(x, z) \leq \max(d(x, y), d(y, z)).$$

Cette inégalité donne de nombreux résultats topologiques très différents des espaces métriques, notamment sur les triangles et les boules.

Le théorème permettant de prouver l'unicité du théorème de Bowen est le suivant, et nous allons en donner une preuve dans cette section, puis une formalisation possible en Lean.

Théorème III.1. *Soit (E, d) un espace ultramétrique et \mathcal{O} un ouvert borné de E , alors il existe une partie $R \subseteq \mathcal{O}$ et une application $r: R \rightarrow \mathbf{R}_+^*$ tels que*

$$\mathcal{O} = \bigsqcup_{x \in R} B(x, r(x)).$$

Dans la suite on fixe (E, d) un espace ultramétrique. Commençons par quelques lemmes décrivant les relations entre les boules dans ces espaces.

Lemme III.2. *Soit $x, y \in E$ et un réel $r > 0$, si $y \in B(x, r)$, alors $B(x, r) = B(y, r)$*

Démonstration. Supposons que $y \in B(x, r)$. Soit $z \in B(x, r)$, alors

$$d(y, z) \leq \max(d(y, x), d(x, z)) < \max(r, r) = r,$$

car $y, z \in B(x, r)$. Donc on a une inclusion : $B(x, r) \subseteq B(y, r)$.

Réciproquement comme $y \in B(x, r)$ on a $x \in B(y, r)$ par symétrie de la distance et donc avec le même raisonnement on a l'autre inclusion. Finalement, $B(x, r) = B(y, r)$. \square

Lemme III.3. *Soit $x, y \in E$ et deux réels $r, s > 0$, alors il y a trois possibilités :*

1. $B(x, r) \subseteq B(y, s)$,
2. $B(y, s) \subseteq B(x, r)$,
3. $B(x, r) \cap B(y, s) = \emptyset$.

Démonstration. Supposons que $B(x, r) \cap B(y, s) \neq \emptyset$, alors soit $z \in B(x, r) \cap B(y, s)$. On peut alors écrire $B(x, r) = B(z, r)$ et $B(y, s) = B(z, s)$ grâce au lemme III.2, et ainsi on a bien le résultat voulu en fonction de l'ordre de r et s . \square

Lemme III.4. Soit $(x_i)_{i \in I}$ une famille de points de E et $(r_i)_{i \in I}$ une famille de réels strictement positifs. De plus, supposons qu'il existe $x \in E$ tel que $x \in B(x_i, r_i)$ pour tout $i \in I$ et que $R = \sup_{i \in I} r_i < \infty$. Alors

$$\bigcup_{i \in I} B(x_i, r_i) = B(x, R).$$

Démonstration. Commençons par réécrire chaque boule en changeant son centre grâce au lemme III.2 et notons $R = \sup_{i \in I} r_i < \infty$, on a alors l'inclusion

$$\bigcup_{i \in I} B(x_i, r_i) = \bigcup_{i \in I} B(z, r_i) \subseteq B(z, R).$$

Réciproquement si $y \in B(z, R)$, alors il existe $i_0 \in I$ tel que $d(y, z) < r_{i_0} \leq R$ et donc $y \in B(z, r_{i_0}) \subseteq \bigcup_{i \in I} B(z, r_i)$. Finalement, l'union de ces boules est encore une boule. \square

On notera dans la suite $\mathcal{O} \subseteq E$ un ouvert borné, et donc on a pour chaque $x \in \mathcal{O}$ un réel $r_x > 0$ tel que $B(x, r_x) \subseteq \mathcal{O}$. On a alors

$$\mathcal{O} = \bigcup_{x \in \mathcal{O}} B(x, r_x).$$

Remarque. On peut voir les $(r_x)_{x \in \mathcal{O}}$ comme une application $r: \mathcal{O} \rightarrow \mathbf{R}_+^*$.

Définition III.1. On note $B(\mathcal{O})$ l'ensemble des boules incluses dans \mathcal{O} , en particulier :

$$B(\mathcal{O}) := \{B(x, r) \mid x \in \mathcal{O} \wedge r > 0 \wedge B(x, r) \subseteq \mathcal{O}\} \subseteq \mathcal{P}(E).$$

Soit $\mathcal{U} \subseteq B(\mathcal{O})$, on définit une relation $\sim_{\mathcal{U}}$ (ou plus simplement \sim s'il n'y a pas d'ambiguïté) sur \mathcal{U} par :

$$\forall u, v \in \mathcal{U}, \quad u \sim_{\mathcal{U}} v \iff \exists w \in \mathcal{U}, u \cup v \subseteq w.$$

Proposition III.5. Pour tout $\mathcal{U} \subseteq B(\mathcal{O})$, la relation $\sim_{\mathcal{U}}$ est une relation d'équivalence.

Démonstration. Soit $u_1, u_2, u_3 \in \mathcal{U}$ trois boules.

Pour la réflexivité, on considère $w = u_1 \in \mathcal{U}$ et on a bien $u_1 \sim u_1$.

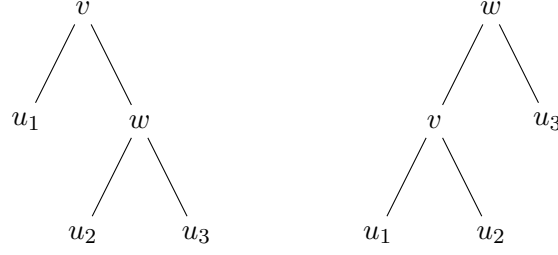
Pour la symétrie, la définition de la relation est clairement symétrique.

Enfin pour la transitivité, on suppose que $u_1 \sim u_2$ et $u_2 \sim u_3$. On a alors $v, w \in \mathcal{U}$ tels que

$$u_1 \subseteq v, u_2 \subseteq v, u_2 \subseteq w \text{ et } u_3 \subseteq w.$$

Ainsi, $\emptyset \neq u_2 \subseteq v \cap w$, or v et w sont des boules d'un espace ultramétrique. On a alors deux cas :

- $v \subseteq w$ dans ce cas on a $u_1 \overset{v}{\sim} u_3$
- $w \subseteq v$ et alors on a $u_1 \overset{w}{\sim} u_3$



Donc $\sim_{\mathcal{U}}$ est bien une relation d'équivalence sur \mathcal{U} . \square

Pour la suite de la section on fixe $\mathcal{U} \subseteq B(\mathcal{O})$ un sous-ensemble de l'ensemble des boules incluses dans \mathcal{O} .

Définition III.2. Soit $u \in \mathcal{U}$, la classe d'équivalence de u pour la relation \sim est noté \bar{u} . On introduit également $\mathcal{B}_u \in B(\mathcal{O})$ l'union des éléments de la classe d'équivalence de u :

$$\mathcal{B}_u := \bigcup_{u \sim u'} u' \subseteq \mathcal{O}.$$

Remarquons que \mathcal{B}_u est une boule dès lors que cette union est bornée (d'après le lemme III.4), or cette $\mathcal{B}_u \subseteq \mathcal{O}$ et \mathcal{O} est borné donc \mathcal{B}_u est toujours une boule.

Définition III.3. Soit R un système complet de représentant pour la relation \sim . On pose alors

$$\mathcal{U}' := \mathcal{U} \cup \{\mathcal{B}_u \mid u \in R\} \subseteq B(\mathcal{O}).$$

car $\mathcal{B}_u \in B(\mathcal{O})$ pour chaque $u \in R$.

On considère désormais $\sim' = \sim_{\mathcal{U}'}$ qui est encore une relation d'équivalence, et soit R' un système complet de représentant pour la relation \sim' .

Théorème III.6. Soit \mathcal{O} un ouvert borné de E . Avec les notations introduites précédemment on a

$$\mathcal{O} = \bigsqcup_{u \in R'} \mathcal{B}_u.$$

Démonstration. On commence par montrer que cette union est bien égale à \mathcal{O} . En considérant $\mathcal{U} = \{B(x, r_x) \mid x \in \mathcal{O}\}$,

$$\mathcal{O} = \bigcup_{u \in \mathcal{U}} u = \bigcup_{u \in \mathcal{U}'} u = \bigcup_{u \in R'} \left(\bigcup_{v \sim' u} v \right) = \bigcup_{u \in R'} \mathcal{B}_u$$

Il reste encore à montrer que cette union est disjointe. Supposons que $\mathcal{B}_u \cap \mathcal{B}_v \neq \emptyset$. Alors on a $\mathcal{B}_u \subseteq \mathcal{B}_v$ ou $\mathcal{B}_v \subseteq \mathcal{B}_u$ par le lemme III.3. Dans les deux cas on a $u \sim' v$ avec $w = \mathcal{B}_v \in \mathcal{U}'$ dans le premier cas et $w = \mathcal{B}_u \in \mathcal{U}'$ dans le second. Or R est un système complet de représentant pour \sim' et donc nécessairement $u = v$. Finalement les $(\mathcal{B}_u)_{u \in R'}$ sont deux à deux disjoints et leur union est égale à \mathcal{O} , ce qui conclut la preuve du théorème III.1. \square

Bibliographie

- [24] *Martin-Löf dependent type theory in nLab*. 2024. URL : <https://ncatlab.org/nlab/show/Martin-L%C3%B6f+dependent+type+theory>.
- [Bow75] Rufus BOWEN. « Equilibrium states and the ergodic theory of Anosov diffeomorphisms ». In : *Lecture notes in mathematics* 470 (1975), p. 11-25.
- [Car19] Mario CARNEIRO. « The Type Theory of Lean ». In : (2019). URL : <https://github.com/digama0/lean-type-theory/releases>.
- [Mar21] Per MARTIN-LÖF. « Intuitionistic Type Theory : Notes by Giovanni Sambin of a series of lectures given in Padova, June 1980 ». In : (2021).
- [Mas] Patrick MASSOT. *Lean blueprints*. URL : <https://github.com/PatrickMassot/leanblueprint>.
- [Uni13] The UNIVALENT FOUNDATIONS PROGRAM. *Homotopy Type Theory : Univalent Foundations of Mathematics*. Institute for Advanced Study : <https://homotopytypetheory.org/book>, 2013.

Annexe

Existence et unicité de mesures de Gibbs pour des dynamiques hyperboliques.

Stage de L3 encadré par Christophe Dupont à l'IRMAR (Rennes).

Dorian Guillet, étudiant au Magistère de Grenoble.

Juin 2024.

TABLE DES MATIÈRES

I	INTRODUCTION	1
II	TOPOLOGIE DU DÉCALAGE DE BERNOULLI	3
II.1	Définition de l'espace métrique Σ_n	3
II.2	Propriétés topologiques de Σ_n	4
III	MESURES DE GIBBS SUR Σ_n	6
III.1	Unicité de la mesure de Gibbs	6
III.2	Réduction de Σ_n à Σ_n^+	7
III.3	Théorème de Perron-Frobenius de Ruelle	10
III.4	Construction d'une mesure de Gibbs	15
III.5	Généralisation au sous-décalage de type fini	18
IV	DYNAMIQUE HYPERBOLIQUE	20
IV.1	Automorphismes hyperboliques du tore	20
IV.2	Dynamique symbolique	21
IV.3	Existence des partitions de Markov	24
IV.4	Généralisation aux difféomorphismes AXIOME A	28
IV.5	Exemple d'un codage sur \mathbf{T}^2	29
V	APPENDICE	31
V.1	Théorème de Riesz	31
V.2	Topologie des espaces de probabilités	31
	BIBLIOGRAPHIE	34

I.

INTRODUCTION

L'objectif de ce rapport de stage est de présenter une construction de mesure de Gibbs sur des espaces de Bernoulli et la manière dont on peut encoder la dynamique d'une application hyperbolique dans un espace de Bernoulli Σ_n , avec le décalage σ .

Définition (Mesure de Gibbs). Soit μ une mesure de probabilité σ -invariante sur Σ_n . On dit μ est une mesure de Gibbs pour un potentiel $\phi: \Sigma_n \rightarrow \mathbf{R}$ s'il existe $P \in \mathbf{R}$ et $c_1, c_2 > 0$ tels que pour tout $x \in \Sigma_n$ et $m \in \mathbf{N}$ on ait

$$c_1 \leq \frac{\mu\{y \in \Sigma_n \mid \forall i \in \llbracket 0, m-1 \rrbracket, x_i = y_i\}}{\exp(-Pm + \sum_{k=0}^{m-1} \phi(\sigma^k x))} \leq c_2.$$

Dans une première partie on se concentre sur la construction d'une mesure de Gibbs ergodique μ sur l'espace métrique Σ_n , afin de montrer le théorème principal, dont la preuve initiale se trouve dans [Bow75].

Théorème ([Bow75]). Soit ϕ une fonction de potentiel hölderienne. Alors il existe une unique mesure de Gibbs pour cette fonction ϕ .

Pour ce faire, on se ramène au cas où la fonction de potentiel ϕ ne dépend plus des coordonnées négatives. Ensuite, on considère l'opérateur de transfert \mathcal{L} défini par

$$\forall f \in \mathcal{C}(\Sigma_n^+), \forall x \in \Sigma_n^+, \quad \mathcal{L}f(x) = \sum_{y \in \sigma^{-1}x} f(y)e^{\phi(y)},$$

où Σ_n^+ est l'ensemble des suites à valeurs dans $\llbracket 1, n \rrbracket$ et indexées sur \mathbf{N} .

Le théorème suivant établit que cet opérateur admet une mesure propre et une fonction propre.

Théorème (Ruelle-Perron-Frobenius). Soit ϕ un potentiel et \mathcal{L} l'opérateur de transfert. Alors il existe $\lambda > 0, \nu \in \mathcal{M}(\Sigma_n^+)$ et $h \in \mathcal{C}(\Sigma_n^+), h > 0$ tels que :

1. ν vérifie $\mathcal{L}^*\nu = \lambda\nu$,
2. h vérifie $\mathcal{L}h = \lambda h$ et $\nu(h) = 1$,
3. et pour toute fonction $g \in \mathcal{C}(\Sigma_n^+)$, $\lim_{m \rightarrow \infty} \left\| \frac{1}{\lambda^m} \mathcal{L}^m g - \nu(g)h \right\| = 0$.

Pour prouver ce théorème, on utilisera le théorème de Schauder-Tychonoff afin de construire la mesure propre μ et la fonction propre h comme des points fixes de certains opérateurs, pour cela nous devrons d'abord établir la compacité de Σ_n et d'un certain ensemble de fonctions notamment grâce au théorème d'Ascoli. Puis pour établir la limite nous aurons besoin de la densité des fonctions en escaliers dans $\mathcal{C}(\Sigma_n^+)$ et des propriétés de l'opérateur de transfert.

Grâce à cette mesure propre ν et cette fonction propre h , on peut construire $\mu = h \cdot \nu$. Cette dernière mesure sur Σ_n^+ est alors σ -invariante, ce qui se montre grâce aux propriétés algébriques

de l'opérateur de transfert et permettra de construire une forme linéaire G sur Σ_n qui s'identifiera grâce au théorème de Riesz en une mesure $\tilde{\mu}$ sur Σ_n , qui sera la mesure de Gibbs pour le potentiel höldérien ϕ . Une fois $\tilde{\mu}$ construite, on montrera qu'elle est ergodique (et même mélangeante), ce qui permettra d'établir l'unicité.

Dans la partie suivante, l'objectif est de montrer que si $f: \mathbf{T}^n \rightarrow \mathbf{T}^n$ est hyperbolique (ou dans un cadre plus large, un difféomorphisme AXIOME A sur une variété riemannienne, dont on donnera un exemple sur \mathbf{C}^2 grâce à [FS92 ; BS91]), on peut coder sa dynamique dans un espace de Bernoulli grâce aux théorèmes suivants, le premier assurant l'existence d'un recouvrement particulier dit partition de Markov et le second donnant le codage associé à un tel recouvrement.

Les preuves, davantage exigeantes, des ces théorèmes s'inspirent de [Bow75], puis de [BP03], [KH95], et se basent sur le lemme de pistage et sur l'existence de variétés stables et instables permettant de munir l'espace d'un produit local.

Théorème ([Bow75]). *Soit $f: \mathbf{T}^n \rightarrow \mathbf{T}^n$ un automorphisme hyperbolique. Alors il existe une partition de Markov pour f de diamètre arbitrairement petit.*

La construction d'une telle partition passe d'abord par la construction d'un codage plus élémentaire à partir d'un recouvrement de \mathbf{T}^n par des boules, qui vérifie certaines propriétés similaires à celle des partitions de Markov, mais dont les pièces s'intersectent sur des parties trop grandes, générant un défaut d'injectivité du codage. Pour y remédier, on modifie le recouvrement précédent en un dernier recouvrement donc les pièces ne se chevauchent plus, donnant ainsi une partition de Markov.

Théorème ([BP03]). *Soit $\mathcal{R} = (R_i)_{1 \leq i \leq m}$ une partition de Markov et (Σ_A, σ) l'espace de Bernoulli associé à la matrice d'incidence A de la partition \mathcal{R} . Alors,*

1. *pour $\omega \in \Sigma_A$, l'intersection $\bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_i})$ est un singleton et on note $\pi(\omega)$ cet unique élément,*
2. *l'application $\pi: \Sigma_A \rightarrow \mathbf{T}^n$ est continue, surjective et $f \circ \pi = \pi \circ \sigma$,*
3. *si $\mu \in \mathcal{M}_\sigma(\Sigma_A)$ est ergodique de support Σ_A , alors*

$$\mu \{ \omega \in \Sigma_A \mid \text{Card } \pi^{-1}(\pi(\omega)) > 1 \} = 0.$$

De cette manière, on peut considérer que π est injective quitte à retirer un ensemble de mesure nulle pour certaines mesures (en particulier la mesure de Gibbs). La dynamique de f sur le tore peut alors être codée par un sous-décalage de Σ_m , permettant ainsi une étude plus simple de cette dynamique, et notamment de munir ces systèmes de mesures de probabilités. En effet, si μ est la mesure de Gibbs sur Σ_A alors la mesure $\pi_*\mu$ vérifie des propriétés sur le tore semblable à celle vérifiée sur Σ_A .

Finalement, on montre un exemple de construction d'un tel codage sur le tore \mathbf{T}^2 à deux dimensions, pour l'automorphisme du tore associé à la matrice

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

pour laquelle on peut construire une partition de Markov composée de trois morceaux, et dont la méthode de construction reprend celle utilisée dans [Nij17].

II.

TOPOLOGIE DU DÉCALAGE DE BERNOULLI

II.1 Définition de l'espace métrique Σ_n

Définition II.1.1. Soit $n \in \mathbf{N}$. L'ensemble des suites bi-infinies à valeurs dans $\llbracket 1, n \rrbracket$ est noté $\Sigma_n = \prod_{\mathbf{Z}} \llbracket 1, n \rrbracket$. Chaque $\llbracket 1, n \rrbracket$ est muni de la topologie discrète, et on munit alors Σ_n de la distance produit d donnée par

$$\forall x, y \in \Sigma_n, \quad d(x, y) = \sum_{i \in \mathbf{Z}} 2^{-|i|} \mathbf{1}_{x_i \neq y_i} < +\infty.$$

On notera $B_d(x, r)$ la boule centrée en $x \in \Sigma_n$ et de rayon $r \geq 0$ pour d .

Proposition II.1.1. *L'espace métrique (Σ_n, d) est compact.*

Preuve. Chacun des $\llbracket 1, n \rrbracket$ est compact pour sa topologie. En tant que produit Σ_n est donc également compact, grâce au théorème de Tychonoff. \square

Proposition II.1.2. *Pour $x, y \in \Sigma_n$ distincts, notons $N = \min \{i \in \mathbf{N} \mid x_i \neq y_i \text{ ou } x_{-i} \neq y_{-i}\}$. Alors $d_\beta(x, y) = \beta^N$ et $d_\beta(x, x) = 0$ définit une distance sur Σ_n et est équivalente à la distance d , et ce pour tout $\beta \in]0, 1[$.*

Preuve. Soit $\beta \in]0, 1[$. Alors d_β est distance car d_β est clairement symétrique, $d(x, y) = 0 \iff x = y$ pour tout $x, y \in \Sigma_n$ et d_β vérifie l'inégalité triangulaire.

De plus, d_β est équivalente à d . En effet pour $r > 0$, montrons qu'il existe $r_1, r_2 > 0$ tel que pour tout $x \in \Sigma_n$ on ait

$$B_\beta(x, r_1) \subseteq B_d(x, r) \subseteq B_\beta(x, r_2).$$

Soit $y \in \Sigma_n$. Supposons que $y \in B_d(x, r)$, alors $d(x, y) < r$, en particulier

$$\forall i \in \mathbf{Z}, \quad 2^{-|i|} \mathbf{1}_{x_i \neq y_i} \leq r$$

Soit $i \in \mathbf{Z}$, si $2^{-|i|} > r$, alors nécessairement $x_i = y_i$ et $|i| < -\frac{r}{\log 2}$, alors en posant $r_2 = \beta^{-\frac{r}{\log 2}}$ on a $y \in B_\beta(x, r_2)$, ce qui prouve une des deux inclusions.

Supposons désormais que $x \in B_\beta(x, r)$, c'est-à-dire $d_\beta(x, y) < r$, donc x et y coïncident sur les $m_r = \frac{r}{\log \beta}$ premières coordonnées. Ainsi,

$$d(x, y) = \sum_{i \in \mathbf{Z}} 2^{-|i|} \mathbf{1}_{x_i \neq y_i} \leq 2 \sum_{i \geq m_r + 1} 2^i = 2^{1-m_r} = r_1.$$

et finalement, $y \in B_d(x, r_1)$. Donc les distances sont équivalentes et les topologies associées à ces distances sont les mêmes. \square

On notera donc dans la suite, $B_\beta(x, r)$ les boules de centre $x \in \Sigma_n$ et de rayon $r > 0$ pour la distance d_β , et $B(x, r) = B_{\frac{1}{2}}(x, r)$. On peut alors noter que si $y \in B_\beta(x, r)$, si on note $m_r = \frac{r}{\log \beta}$, alors x et y coïncident sur les coordonnées entre $-m_r$ et m_r :

$$\forall i \in \llbracket -m_r, m_r \rrbracket, \quad x_i = y_i.$$

II.2 Propriétés topologiques de Σ_n

Définition II.2.1. Soit $\phi \in \mathcal{C}(\Sigma_n)$. La fonction de variation d'ordre k est donné par :

$$\text{var}_k(\phi) = \sup \{ |\phi(x) - \phi(y)| : \forall i \in \llbracket -k, k \rrbracket, x_i = y_i \}.$$

Proposition II.2.1. Soit $\beta \in]0, 1[$. Soit $\phi \in \mathcal{C}(\Sigma_n)$, si pour tout $k \in \mathbf{N}$ on $\text{var}_k(\phi) \leq b\alpha^k$ pour certaines constantes $b > 0$ et $\alpha \in]0, 1[$, alors il existe $\beta \in]0, 1[$ tel que ϕ soit hölderienne pour d_β .

Définition II.2.2. On notera $\mathcal{H}(\Sigma_n)$ l'ensemble des fonctions $\phi \in \mathcal{C}(\Sigma_n)$ pour lesquels il existe $\alpha \in]0, 1[$ et $b > 0$ vérifiant

$$\forall k \in \mathbf{N}, \quad \text{var}_k(\phi) \leq b\alpha^k.$$

Proposition II.2.2. Pour tout $m \in \mathbf{N}$, soit $T_m = \{x \in \Sigma_n \mid \forall |i| > m, x_i = 1\}$ et $T = \bigcup_{m \geq 0} T_m$. L'ensemble T est dense dans Σ_n et est dénombrable. Ainsi, Σ_n est séparable.

Preuve. Soit O un ouvert non vide et $x \in O$, alors il existe $r = 2^{-m+1} > 0$ tel que $B(x, r) \subseteq O$. Si on prend $y \in \Sigma_n$ tel que si $i \in \llbracket -m, m \rrbracket$ alors $y_i = x_i$ et sinon $y_i = 1$, alors $y \in T_m$ et $y \in B(x, r)$, donc $O \cap T \neq \emptyset$. Ainsi T est dense dans Σ_n .

Pour tout $m \in \mathbf{N}$, l'ensemble T_m contient exactement n^{2m+1} éléments, ie. T_m est fini. Donc par union dénombrable, T est dénombrable. \square

Proposition II.2.3. Soit O un ouvert de Σ_n . Pour tout $x \in O$, on note $r_x > 0$ tel que $B(x, r_x) \subseteq O$. Alors il existe $\mathcal{R} \subseteq \Sigma_n$ dénombrable tel que

$$O = \bigsqcup_{x \in \mathcal{R}} B(x, r_x),$$

où \bigsqcup dénote l'union disjointe.

Preuve. Soit O un ouvert de Σ_n , alors $O = \bigcup_{x \in O} B(x, r_x)$. Cette union n'est pas disjointe et indexée sur un ensemble fini. On introduit alors la relation :

$$\forall x, y \in O, x \sim y \iff B(x, r_x) \cap B(y, r_y) \neq \emptyset.$$

Cette relation est clairement une relation d'équivalence. De plus on a pour $x, y \in O$,

$$x \sim y \iff B(x, r_x) \subseteq B(y, r_y) \text{ ou } B(y, r_y) \subseteq B(x, r_x).$$

En effet, supposons que $B(x, r_x) \cap B(y, r_y) \neq \emptyset$, alors soit z dans l'intersection, on note $m_x = -\frac{r_x}{\log 2}$ et $m_y = -\frac{r_y}{\log 2}$. Sans perte de généralité, on peut supposer que $m_y \leq m_x$. On a alors

$$\begin{cases} \forall i \in \llbracket -m_x, m_x \rrbracket, x_i = z_i, \\ \forall i \in \llbracket -m_y, m_y \rrbracket, y_i = z_i. \end{cases}$$

Alors, pour tout $i \in \llbracket -m_y, m_y \rrbracket, x_i = y_i$ et donc $B(x, r_x) \subseteq B(y, r_y)$. Ainsi, il existe un système de représentant $\mathcal{R} \subseteq O$ pour la relation \sim tel que

$$O = \bigcup_{x \in O} B(x, r_x) = \bigsqcup_{x \in \mathcal{R}} B(x, r_x).$$

Il reste à montrer que \mathcal{R} est au plus dénombrable. Pour tout $x \in \mathcal{R}$, il existe $h(x) \in T$ tel que $h(x) \in B(x, r_x)$. L'application $h: \mathcal{R} \longrightarrow T$ ainsi définie est donc injective et T est dénombrable, ainsi \mathcal{R} est dénombrable. \square

Définition II.2.3. L'application de décalage (*shift* en anglais) σ sur Σ_n est donnée par

$$(\sigma x)_i = x_{i+1}$$

pour tout $x \in \Sigma_n$ et $i \in \mathbf{Z}$. Cette application est alors un homéomorphisme de Σ_n .

Définition II.2.4. Soit $x \in \Sigma_n$. On définit les cylindres de centre x et de rayon $m \in \mathbf{N}$ par

$$C(x, m) = \{y \in \Sigma_n \mid \forall i \in \llbracket 0, m-1 \rrbracket, x_i = y_i\}.$$

de cette manière on peut les relier aux boules pour la distance d_β aux cylindres :

$$C(x, 2m+1) = \sigma^{-m}(B_\beta(\sigma^m x, \beta^m)).$$

III.

 MESURES DE GIBBS SUR Σ_n

III.1 Unicité de la mesure de Gibbs

Théorème III.1.1 ([Bow75]). Soit $\phi \in \mathcal{H}(\Sigma_n)$, alors il existe une unique mesure $\mu_\phi = \mu \in \mathcal{M}_\sigma(\Sigma_n)$ tel qu'il existe $P(\phi) = P \in \mathbf{R}$, $c_1, c_2 > 0$ vérifiant

$$\frac{\mu(C(x, m))}{\exp(-Pm + S_m\phi(x))} \in [c_1, c_2]$$

pour tout $x \in \Sigma_n$ et $m \in \mathbf{N}$, où $S_m\phi(x) = \sum_{k=0}^{m-1} (\phi(\sigma^k x))$

Dans les sections suivantes, on construira une mesure de Gibbs mélangeante notamment grâce au théorème de Ruelle. C'est l'ergodicité de cette mesure qui entrainera l'unicité de la mesure de Gibbs.

Définition III.1.1. Soit $\mu \in \mathcal{M}(\Sigma_n)$. On dit que

- μ est ergodique (par rapport à σ) si pour tout borélien $E \in \mathcal{B}(\Sigma_n)$ tels que $\sigma^{-1}E = E$, on a

$$\mu(E) = 0 \text{ ou } \mu(E^c) = 0,$$

- μ est mélangeante (par rapport à σ) si pour tout borélien $E, F \in \mathcal{B}(\Sigma_n)$, on a

$$\mu(E \cap \sigma^{-n}F) \xrightarrow[n \rightarrow \infty]{} \mu(E)\mu(F).$$

Remarque. On a l'implication "mélangeante" \implies "ergodique".

Lemme III.1.2. Soit $f: \Sigma_n \rightarrow \mathbf{R}$ intégrable par rapport à une mesure $\mu \in \mathcal{M}_\sigma(\Sigma_n)$ ergodique. Supposons que $f \circ \sigma = f$ μ -presque partout. Alors f est constante μ -presque partout.

Preuve. Pour montrer ce lemme on considère les ensembles $E_c = f^{-1}(\{c\})$ pour tout $c \in \mathbf{R}$. Comme $f \circ \sigma = f$ μ -p.p., on a $\sigma^{-1}E_c = E_c$, et donc par ergodicité de μ ,

$$\mu(E_c) = 0 \text{ ou } \mu(E_c) = 1.$$

Il est donc clair qu'il existe au plus un $c \in \mathbf{R}$ tel que $f = c$ pour μ -presque tout $x \in \Sigma_n$. \square

Proposition III.1.3. Supposons qu'il existe une mesure de Gibbs μ ergodique pour $\phi \in \mathcal{C}(\Sigma_n)$. Alors cette mesure est l'unique mesure de Gibbs associé à ϕ .

Preuve. Soit $\mu, \mu' \in \mathcal{M}(\Sigma_n^+)$ deux mesures de Gibbs avec μ ergodique, $c_1, c'_1, c_2, c'_2 > 0$ et $P, P' \in \mathbf{R}$ des constantes telles que pour tout $m \in \mathbf{N}$ et $x \in \Sigma_n$ on ait

$$\begin{cases} c_1 \leq \frac{\mu(C(x, m))}{\exp(-Pm + S_m\phi(x))} \leq c_2 \\ c'_1 \leq \frac{\mu'(C(x, m))}{\exp(-P'm + S_m\phi(x))} \leq c'_2 \end{cases}$$

D'abord, montrons que $P = P'$. Soit $m \in \mathbf{N}$ et $T_m = \{x \in \Sigma_n \mid x_i = 1 \text{ si } i \notin \llbracket 0, m-1 \rrbracket\}$, de telle sorte que $\Sigma_n = \bigsqcup_{x \in T_m} C(x, m)$, et T_m est fini. On a alors

$$c'_1 e^{-P'm} \sum_{x \in T_m} e^{S_m \phi(x)} \leq \sum_{x \in T_m} \mu'(C(x, m)) = 1 = \sum_{x \in T_m} \mu'(C(x, m)) \leq c'_2 e^{-P'm} \sum_{x \in T_m} e^{S_m \phi(x)},$$

et donc

$$P - \frac{1}{m} \log c'_2 \leq \frac{1}{m} \log \sum_{x \in T_m} e^{S_m \phi(x)} \leq P - \frac{1}{m} \log c'_1.$$

Par le théorème des gendarmes, on trouve donc $P' = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{x \in T_m} e^{S_m \phi(x)}$. Grâce au même raisonnement, on en conclut que $P = P'$ car ils sont tout deux égaux à la même limite.

Grâce aux estimations sur μ et μ' sur les cylindres, on a alors pour tout $x \in \Sigma_n$ et $m \in \mathbf{N}$,

$$\mu'(C(x, m)) \leq \frac{c'_2}{c_1} \mu(C(x, m)).$$

Comme μ et μ' sont invariantes par σ , on peut étendre le résultat sur les ensembles de la forme $\{y \in \Sigma_n \mid \forall i \in \llbracket -m, m \rrbracket, x_i = y_i\}$ pour tout $x \in \Sigma_n$ et $m \in \mathbf{N}$, c'est-à-dire sur une base de la topologie de Σ_n . De plus, comme tout ouvert de Σ_n est une union disjointe dénombrable de ces éléments d'après la proposition II.2.3, on peut alors étendre cette inégalité aux ouverts de Σ_n . Par régularité extérieure par rapport aux ouverts, d'après la proposition V.2.4, on a pour tout borélien $B \in \mathcal{B}(\Sigma_n)$,

$$\mu'(B) = \inf \{ \mu'(O) \mid O \text{ ouvert}, B \subseteq O \}.$$

De là, on a pour tout ouvert $B \subseteq O$,

$$\mu'(B) \leq \mu'(O) \leq \frac{c'_2}{c_1} \mu(O),$$

et donc $\mu'(B) \leq \frac{c'_2}{c_1} \inf \{ \mu(O) \mid O \text{ ouvert}, B \subseteq O \} = \frac{c'_2}{c_1} \mu(B)$. Donc pour tout borélien $B \in \mathcal{B}(\Sigma_n)$ on a $\mu'(B) \leq \frac{c'_2}{c_1} \mu(B)$, ainsi μ' est absolument continue par rapport à μ . D'après le théorème de Radon-Nikodym, μ' admet une densité f par rapport à μ . En appliquant σ , on obtient

$$\mu' = \sigma_* \mu' = (f \circ \sigma^{-1}) \cdot \sigma_* \mu = (f \circ \sigma^{-1}) \cdot \mu.$$

Par unicité de la dérivée de Radon-Nikodym, $f = f \circ \sigma^{-1}$ μ -presque partout. Or comme μ est ergodique par hypothèse, alors il existe une constante $c \in \mathbf{R}$ tel que $f = c$ μ -presque partout. Finalement,

$$1 = \mu'(\Sigma_n) = \int_{\Sigma_n} c d\mu = c.$$

Donc $\mu = \mu'$, ce qui prouve l'unicité. □

III.2 Réduction de Σ_n à Σ_n^+

Pour prouver l'existence de telles mesures, on commence par montrer que peu importe la fonction de potentiel ϕ dans $\mathcal{H}(\Sigma_n)$, on peut commencer par trouver une fonction de potentiel $\psi \in \mathcal{H}(\Sigma_n^+)$ ne dépendant que des coordonnées positives ayant la même mesure de Gibbs que ϕ .

Définition III.2.1. Soit $\phi, \psi \in \mathcal{C}(\Sigma_n)$. On dit que ϕ et ψ sont équivalentes et on note $\phi \sim \psi$ dès lors qu'il existe $u \in \mathcal{C}(\Sigma_n)$ vérifiant :

$$\phi = \psi - u + u \circ \sigma.$$

Le prochain lemme justifie l'introduction de cette relation :

Lemme III.2.1. Soit $\phi \sim \psi \in \mathcal{H}(\Sigma_n)$. Dans ce cas, $\mu_\phi = \mu_\psi$ et $P(\phi) = P(\psi)$.

Preuve. Soit $u \in \mathcal{C}(\Sigma_n)$ tel que $\phi - \psi = u \circ \sigma - u$. Alors

$$\begin{aligned} |S_m \phi(x) - S_m \psi(x)| &= \left| \sum_{k=0}^{m-1} (\phi - \psi)(\sigma^k x) \right| = \left| \sum_{k=0}^{m-1} (u(\sigma^{k+1} x) - u(\sigma^k x)) \right| \\ &= |u(\sigma^m x) - u(x)| \leq 2\|u\| \end{aligned}$$

Ainsi pour $m \in \mathbf{N}$ et $x \in \Sigma_n$, on a

$$c_1 e^{-2\|u\|} \leq \frac{\mu_\phi(C(x, m))}{e^{-P(\phi)m + S_m \phi(x) + 2\|u\|}} \leq \frac{\mu_\phi(C(x, m))}{e^{-P(\phi)m + S_m \psi(x)}} \leq \frac{\mu_\phi(C(x, m))}{e^{-P(\phi)m + S_m \phi(x) - 2\|u\|}} \leq c_2 e^{2\|u\|}$$

Donc μ_ϕ et $P(\phi)$ conviennent aussi pour ψ . \square

Remarque. Soit $\phi \stackrel{u}{\sim} \psi \in \mathcal{C}(\Sigma_n)$. Si $\sum_k u \circ \sigma^k$ converge, alors

$$u = \sum_{k=0}^{\infty} (\phi - \psi) \circ \sigma^k$$

Cette remarque permet de mieux comprendre la forme de la fonction u dans un cadre favorable, et de donner l'intuition de ce qu'elle pourrait être pour donner une fonction équivalente à ϕ dépendant uniquement des coordonnées positives.

Preuve de la remarque précédente. La convergence de $\sum_{k \geq 0} (u - u \circ \sigma) \circ \sigma^k = \sum_k (\phi - \psi) \circ \sigma^k$ est assuré par celle de $\sum_k u \circ \sigma^k$. Ainsi par télescopage, $u(x) = \sum_{k=0}^{\infty} (u - u \circ \sigma)(\sigma^k x) = \sum_{k=0}^{\infty} (\phi - \psi)(\sigma^k x)$. \square

Avant le prochain lemme on introduit la fonction $r: \Sigma_n \rightarrow \Sigma_n$ définie par

$$\forall i \in \mathbf{Z}, \forall x \in \Sigma_n, (r(x))_i = \begin{cases} x_i & \text{si } i \geq 0, \\ 1 & \text{si } i < 0. \end{cases} = 111 \dots x_{|\geq 0}$$

De sorte que si $x, y \in \Sigma_n$ ont les mêmes coordonnées positives (ie. $\forall i \geq 0, x_i = y_i$), alors $r(x) = r(y)$.

De plus, lorsqu'on compose r avec σ , on garde l'indépendance vis-à-vis des coordonnées négatives : si $x = \dots x_{-2} x_{-1} \textcolor{teal}{x_0} x_1 x_2 x_3 \dots \in \Sigma_n$, alors on a :

$$\begin{aligned} x &= \dots x_{-2} x_{-1} \textcolor{teal}{x_0} x_1 x_2 \dots \\ \sigma x &= \dots x_{-1} x_0 \textcolor{teal}{x_1} x_2 x_3 \dots \\ r(\sigma x) &= \dots 111 \textcolor{teal}{x_1} x_2 x_3 \dots \\ r(x) &= \dots 111 \textcolor{teal}{x_0} x_1 x_2 \dots \\ \sigma r(x) &= \dots 11 x_0 \textcolor{teal}{x_1} x_2 x_3 \dots \end{aligned}$$

Donc $\sigma \circ r$ et $r \circ \sigma$ diffèrent seulement à la coordonnées -1 et ne dépendent que des coordonnées positives de x .

Lemme III.2.2. Soit $\phi \in \mathcal{H}(\Sigma_n)$, et

$$\psi := \phi \circ r + \left(\sum_{k=0}^{\infty} \phi \circ \sigma^k \right) \circ (\sigma \circ r - r \circ \sigma).$$

Alors,

1. $\psi \in \mathcal{H}(\Sigma_n)$ et ne dépend que des coordonnées positives,
2. Si on note $u = \sum_{k \geq 0} \phi \circ \sigma^k - \phi \circ (\sigma^k r)$, alors $u \in \mathcal{C}(\Sigma_n)$,
3. De plus, $\phi \stackrel{u}{\sim} \psi$.

Preuve. Soit $b \in \mathbf{R}$ et $\alpha \in]0, 1[$ tels que $\text{var}_k \phi \leq b\alpha^k$. Pour $k \geq 0$, $\sigma^k x$ et $\sigma^k r(x)$ coïncident de $-k$ à $+\infty$ donc $|\phi(\sigma^k x) - \phi(\sigma^k r(x))| \leq b\alpha^k$ et ainsi comme $|\alpha| < 1$, la série converge normalement pour tout $x \in \Sigma_n$. Donc u est bien définie et est continue.

Ensuite vérifions que $\psi = \phi - u + u \circ \sigma$. Pour $N \in \mathbf{N}$,

$$\begin{aligned} & \phi(x) - \sum_{k=0}^N \left(\phi(\sigma^k x) - \phi(\sigma^k r(x)) \right) + \sum_{k=0}^N \left(\phi(\sigma^{k+1} x) - \phi(\sigma^k r(\sigma x)) \right) \\ &= \phi(r(x)) - \sum_{k=0}^{N-1} \left(\phi(\sigma^{k+1} x) - \phi(\sigma^{k+1} r(x)) \right) + \sum_{k=0}^N \left(\phi(\sigma^{k+1} x) - \phi(\sigma^k r(\sigma x)) \right) \\ &= \phi(r(x)) + \underbrace{\phi(\sigma^{N+1} x) - \phi(\sigma^N r(\sigma x))}_{\leq b\alpha^N} + \sum_{k=0}^{N-1} \left(\phi(\sigma^{k+1} r(x)) - \phi(\sigma^k r(\sigma x)) \right) \\ &\longrightarrow \psi(x) \\ &\longrightarrow \phi(x) - u(x) + u(\sigma x) \end{aligned}$$

lorsque $N \rightarrow \infty$. Ainsi $\psi = \phi - u + u \circ \sigma$ ie. $\phi \stackrel{u}{\sim} \psi$.

Soit $x \in \Sigma_n$ et $y \in C(x, m)$. Soit $k \in \mathbf{N}$, distinguons plusieurs cas :

— Comme $\sigma^k x$ et $\sigma^k r(x)$ coïncident jusqu'à la k -ième coordonnée,

$$\left| \phi(\sigma^k x) - \phi(\sigma^k r(x)) - (\phi(\sigma^k y) - \phi(\sigma^k r(y))) \right| \leq 2b\alpha^k$$

— Si $k \leq \lfloor \frac{m}{2} \rfloor$, on peut faire mieux :

$$\begin{aligned} & \left| \phi(\sigma^k x) - \phi(\sigma^k y) + \phi(\sigma^k r(x)) - \phi(\sigma^k r(y)) \right| \\ & \leq \underbrace{\left| \phi(\sigma^k x) - \phi(\sigma^k y) \right|}_{\leq \text{var}_{m-k} \phi \leq b\alpha^{m-k}} + \underbrace{\left| \phi(\sigma^k r(x)) - \phi(\sigma^k r(y)) \right|}_{\leq b\alpha^{m-k}} \\ & \leq 2b\alpha^{m-k} \end{aligned}$$

car $\sigma^k x$ et $\sigma^k y$ coïncident sur $[-m-k, m-k]$

Donc en sommant chacun des termes,

$$\begin{aligned} |u(x) - u(y)| &\leq 2b \left(\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} \alpha^{m-k} + \sum_{k > \lfloor \frac{m}{2} \rfloor} \alpha^k \right) \leq 4b \sum_{k > \lfloor \frac{m}{2} \rfloor} \alpha^k = \frac{4b}{1-\alpha} \alpha^{\lfloor \frac{m}{2} \rfloor} \\ &\leq \frac{4b\alpha}{1-\alpha} (\sqrt{\alpha})^m \end{aligned}$$

Ainsi $u \in \mathcal{H}(\Sigma_n)$ et donc $\psi \in \mathcal{H}(\Sigma_n)$ également. \square

Remarque. On remarque que la construction de ψ dépend entièrement de la fonction r , ce qui signifie qu'il existe d'autres fonctions $\psi_r \in \mathcal{H}(\Sigma_n)$ ne dépendant que des coordonnées positives et équivalentes à ϕ .

III.3 Théorème de Perron-Frobenius de Ruelle

Définition III.3.1. Soit $\phi \in \mathcal{H}(\Sigma_n^+) = \mathcal{C}(\Sigma_n^+) \cap \mathcal{H}(\Sigma_n)$. On appelle opérateur de transfert ou opérateur de Ruelle l'application $\mathcal{L}_\phi: \mathcal{C}(\Sigma_n^+) \longrightarrow \mathcal{C}(\Sigma_n^+)$ définie par :

$$\forall f \in \mathcal{C}(\Sigma_n^+), \forall x \in \Sigma_n, \mathcal{L}_\phi f(x) = \sum_{\sigma y = x} e^{\phi(y)} f(y).$$

Remarque. A noter que le passage aux fonctions définies sur $\mathcal{C}(\Sigma_n^+)$ rend σ non injective, et donc \mathcal{L} a toujours n termes dans la somme qui le définit. C'est grâce à cette restriction que l'opérateur de transfert \mathcal{L} devient intéressant.

Pour la suite on fixe $\phi \in \mathcal{H}(\Sigma_n^+)$ et des constantes $b > 0, \alpha \in]0, 1[$ vérifiant $\text{var}_k \phi \leq b\alpha^k$, pour tout $k \in \mathbb{N}$.

Théorème III.3.1 (Théorème de Perron-Frobenius de Ruelle). *Soit $\phi \in \mathcal{H}(\Sigma_n^+)$ et $\mathcal{L} = \mathcal{L}_\phi$. Alors il existe $\lambda > 0, \nu \in \mathcal{M}(\Sigma_n^+)$ et $h \in \mathcal{C}(\Sigma_n^+), h > 0$ tels que :*

1. ν vérifie $\mathcal{L}^* \nu = \lambda \nu$,
2. h vérifie $\mathcal{L} h = \lambda h$ et $\nu(h) = 1$,
3. et pour toute fonction $g \in \mathcal{C}(\Sigma_n^+)$, on a

$$\left\| \frac{1}{\lambda^m} \mathcal{L}^m g - \nu(g) h \right\| \xrightarrow{m \rightarrow \infty} 0.$$

Pour prouver ce théorème, nous allons avoir besoin de quelques lemmes.

Proposition III.3.2. *Il existe $\nu \in \mathcal{M}(\Sigma_n^+)$ tel que $\mathcal{L}^* \nu = \lambda \nu$.*

Preuve. Comme $\mathcal{L}1(x) = \sum_{\sigma y = x} e^{\phi(y)} > 0$ pour tout $x \in \Sigma_n$, on a nécessairement $\mathcal{L}^* \mu(1) > 0$ pour tout $\mu \in \mathcal{M}(\Sigma_n^+)$. On peut alors poser $G(\mu) = \frac{1}{\mathcal{L}^* \mu(1)} \mathcal{L}^* \mu \in \mathcal{M}(\Sigma_n^+)$ pour toute mesure $\mu \in \mathcal{M}(\Sigma_n^+)$, ce qui donne une application $G: \mathcal{M}(\Sigma_n^+) \longrightarrow \mathcal{M}(\Sigma_n^+)$, (on peut se permettre $\mathcal{C}(\Sigma_n^+)^* = \mathcal{M}(\Sigma_n^+)$ grâce au théorème de Riesz). Ainsi définie, G est continue de $\mathcal{M}(\Sigma_n^+)$ dans lui-même. Or $\mathcal{M}(\Sigma_n^+)$ est un compact convexe. Par le théorème de Schauder-Tychonoff, on en déduit que G admet un point fixe $\nu \in \mathcal{M}(\Sigma_n^+)$. On pose alors $\lambda = \mathcal{L}^* \nu(1)$ et on a la relation voulue : $\mathcal{L} \nu = \lambda \nu$. \square

Pour la suite on notera pour tout $m \in \mathbf{N}$, les constantes

$$B_m = \exp \left(\sum_{k \geq m+1} 2b\alpha^k \right) \quad \text{et} \quad K = \lambda B_0 e^{\|\phi\|}.$$

On cherche à construire h comme un point fixe vérifiant (2), pour ce faire on considère l'ensemble $\Lambda \subseteq \mathcal{C}(\Sigma_n^+)$ définie par :

$$\forall f \in \mathcal{C}(\Sigma_n^+), f \in \Lambda \iff \begin{cases} f \geq 0, \\ \nu(f) = 1, \\ \forall m \in \mathbf{N}, \forall x \in \Sigma_n, x' \in C(x, m), f(x) \leq B_m f(x'). \end{cases}$$

On remarque que $1 \in \Lambda$, ce qui assure que $\Lambda \neq \emptyset$.

Lemme III.3.3. *L'ensemble Λ est compact. De plus si $f \in \mathcal{C}(\Sigma_n^+)$, alors $\inf \frac{1}{\lambda} \mathcal{L}f \geq K^{-1}$.*

Preuve. Pour ce faire on va utiliser le théorème d'Ascoli. Dans un premier temps, on va montrer que $\Lambda(x) \subseteq [0, K]$. Soit $f \in \Lambda$ et $x \in \Sigma_n^+$. Remarquons que pour $z \in \Sigma_n^+$ et $x_0 \in \llbracket 1, n \rrbracket$ arbitraire, on a

$$\begin{aligned} \frac{1}{\lambda} \mathcal{L}f(x) &= \sum_{\sigma y = x} e^{\phi(y)} f(y) \geq \lambda^{-1} e^{-\|\phi\|} f(x_0 x) \\ &\geq \lambda^{-1} e^{-\|\phi\|} B_0^{-1} f(z) = \frac{1}{K} f(z). \end{aligned}$$

Donc comme $1 = \nu(f) = \nu(\frac{1}{\lambda} \mathcal{L}f) \geq \frac{1}{K} f(z)$, on en déduit que $\|f\| \leq K$. De plus, comme $\nu(f) = 1$, il existe $z \in \Sigma_n^+$ tel que $f(z) \geq 1$. En appliquant l'inégalité précédente à un tel z , on obtient $\inf \frac{1}{\lambda} \mathcal{L}f \geq \frac{1}{K}$.

Soit $f \in \Lambda$ et $x, x' \in \Sigma_n^+$ tels que $x' \in C(x, m)$ pour un certain $m \in \mathbf{N}$, alors $f(x) \leq B_m f(x')$ et $f(x') \leq B_m f(x)$. Donc

$$|f(x) - f(x')| \leq (B_m - 1) \|f\| \leq (B_m - 1) K \xrightarrow{m \rightarrow \infty} 0.$$

Ainsi comme la majoration est indépendante de f , Λ est équicontinue. \square

Proposition III.3.4. *Il existe $h \in \Lambda$ tel que $h > 0$ et vérifie (2) (ie. $\nu(h) = 1$ et $\mathcal{L}h = \lambda h$).*

Preuve. On pose $F: \mathcal{C}(\Sigma_n^+) \rightarrow \mathcal{C}(\Sigma_n^+)$ définie par

$$\forall f \in \mathcal{C}(\Sigma_n^+), F(f) = \frac{1}{\lambda} \mathcal{L}f.$$

On vérifie que Λ est stable par F . Soit $f \in \Lambda$, il est clair que $\frac{1}{\lambda} \mathcal{L}f$ est positif car f l'est. Aussi, $\nu(\frac{1}{\lambda} \mathcal{L}f) = \frac{1}{\lambda} \mathcal{L}^* \nu(f) = \nu(f) = 1$. Enfin, pour $x, x' \in \Sigma_n^+$ tels que $x' \in C(x, m)$ pour un certain $m \in \mathbf{N}$, on a pour tout $j \in \llbracket 1, n \rrbracket$

$$e^{\phi(jx)} f(jx) \leq e^{\phi(jx') + b\alpha^{m+1}} B_{m+1} f(jx') \leq B_m e^{\phi(jx')} f(jx').$$

En sommant tous ces termes, on obtient l'inégalité voulue : $\mathcal{L}f(x) \leq B_m \mathcal{L}f(x')$. Par ailleurs, F est continue car $ne^{\|\phi\|}$ -lipschitzienne : pour tout $f, g \in \mathcal{C}(\Sigma_n^+)$ et $x \in \Sigma_n^+$, on a

$$|\mathcal{L}f(x) - \mathcal{L}g(x)| \leq \sum_{\sigma y = x} e^{\phi(y)} \|f - g\| \leq ne^{\|\phi\|} \|f - g\|,$$

et donc $\|\mathcal{L}f - \mathcal{L}g\| \leq ne^{\|\phi\|}\|f - g\|$.

En appliquant le théorème de Schauder-Tychonoff à F sur Λ , on obtient un point fixe $h \in \Lambda$ qui vérifie alors $\nu(h) = 1$ et $\mathcal{L}h = \lambda h$. De plus, par le lemme précédent, $0 < K^{-1} \leq \inf \lambda^{-1}\mathcal{L}h = \inf h$, d'où $h > 0$. \square

Lemme III.3.5. *Il existe $\eta \in]0, 1[$ tel que pour toute fonction $f \in \Lambda$, il existe $f' \in \Lambda$ vérifiant*

$$\frac{1}{\lambda}\mathcal{L}f = \eta h + (1 - \eta)f'.$$

De plus, $\eta \leq \min(\frac{u_1}{u_2} \frac{1-\alpha}{4\|h\|K}, \frac{1}{K\|h\|})$.

Preuve. Soit $\eta \in]0, 1[$ comme ci-dessus. Soit $f \in \Lambda$ et on pose la fonction $g = \frac{1}{\lambda}\mathcal{L}f - \eta h$. Alors si $\eta \leq \min(\frac{u_1}{u_2} \frac{1-\alpha}{4\|h\|K}, \frac{1}{K\|h\|})$, on a $\frac{1}{1-\eta}g \in \Lambda$. En effet, l'intégrale de g vaut $\nu(g) = \nu(\lambda^{-1}\mathcal{L}f) - \eta\nu(h) = 1 - \eta$, car f et h sont dans Λ .

Aussi $g \geq 0$ dès lors que $\eta\|h\| \leq \frac{1}{K}$ car

$$g = \lambda^{-1}\mathcal{L}f - \eta h \geq \inf \lambda^{-1}\mathcal{L}f - \eta\|h\| \geq 0,$$

d'après le lemme III.3.3.

Finalement, pour que $\frac{1}{1-\eta}g \in \Lambda$ on doit avoir $g(x) \leq B_m g(x')$ pour $x \in \Sigma_n^+$ et $x' \in C(x, m)$ pour $m \geq 0$, ce qui équivaut à :

$$\eta(B_m h(x') - h(x)) \leq B_m \lambda^{-1}\mathcal{L}f(x') - \lambda^{-1}\mathcal{L}f(x).$$

Or, $\mathcal{L}f(x) \leq B_{m+1}e^{b\alpha^{m+1}}\mathcal{L}f(x')$ (comme vu dans la preuve du lemme III.3.3). Une condition suffisante est

$$\eta(B_m - B_m^{-1})\|h\| \leq (B_m - B_{m+1}e^{b\alpha^{m+1}})K^{-1},$$

car,

$$\begin{cases} \eta(B_m h(x') - h(x)) \leq \eta(B_m h(x') - B_m^{-1}h(x')) \leq \eta(B_m - B_m^{-1})\|h\|, \\ (B_m - B_{m+1}e^{b\alpha^{m+1}})K^{-1} \leq (B_m - B_{m+1}e^{b\alpha^{m+1}})\lambda^{-1}\mathcal{L}f(x') \leq \lambda^{-1}(B_m \mathcal{L}f(x') - \mathcal{L}f(x)). \end{cases}$$

On rappelle que pour tout compact $K \subseteq \mathbf{R}^2$ et tout $x, y \in K$, on a $u_1(x-y) \leq e^x - e^y \leq u_2(x-y)$, ce qu'on applique ici avec le compact $K = [-L, L]^2$ de sorte que $\pm \log B_m, \log B_m e^{b\alpha^m}$ soient dans $[-L, L]$.

Ainsi il suffit de vérifier

$$\eta u_2\|h\|(\log B_m - \log B_m^{-1}) \leq K^{-1}u_1(\log B_m - \log B_{m+1} - b\alpha^{m+1}),$$

qui est équivalent à

$$\eta\|h\|u_2 \left(\frac{4b\alpha^{m+1}}{1-\alpha} \right) \leq K^{-1}u_1 b\alpha^{m+1},$$

ou encore

$$\eta \leq \frac{u_1}{u_2} \frac{1-\alpha}{4\|h\|K},$$

afin d'avoir $\frac{1}{1-\eta}g \in \Lambda$. \square

Le lemme suivant est le cas particulier du théorème III.3.1 pour les fonctions dans Λ , pour lesquelles en plus de la limite on a une convergence exponentielle.

Lemme III.3.6. *Pour tout $n \in \mathbf{N}$ et tout $f \in \Lambda$, on a $\|\lambda^{-n}\mathcal{L}^n f - h\| \leq (\|h\| + K)(1-\eta)^n = A\beta^n$ avec $0 < \beta < 1$ et $A > 0$.*

Preuve. Remarquons qu'en itérant le lemme précédent,

$$\begin{aligned}\lambda^{-1}\mathcal{L}f &= \eta h + (1-\eta)f'_1 = (1-(1-\eta)^1)h + (1-\eta)f'_1, \\ \lambda^{-2}\mathcal{L}^2 f &= \lambda^{-1}\mathcal{L}((1-(1-\eta))h + (1-\eta)f'_1) = (1-(1-\eta))h + \lambda^{-1}(1-\eta)\mathcal{L}f'_1 \\ &= (1-(1-\eta) + \eta(1-\eta))h + (1-\eta)^2 f'_2 \\ &= (1-(1-\eta)^2)h + (1-\eta)^2 f'_2,\end{aligned}$$

Ainsi, par récurrence et grâce au lemme précédent, on a pour tout $n \in \mathbf{N}$ et $f \in \Lambda$,

$$\frac{1}{\lambda^n}\mathcal{L}^n f = (1-(1-\eta)^n)h + (1-\eta)^n f_n,$$

où $f_n \in \Lambda$. Comme $\|f_n\| \leq K$, on obtient

$$\left\| \frac{1}{\lambda^n}\mathcal{L}^n f - h \right\| = (1-\eta)^n \|h + f_n\| \leq (1-\eta)^n (K + \|h\|).$$

□

Avant d'étendre le résultat à toutes les fonctions de $\mathcal{C}(\Sigma_n^+)$, on l'étend d'abord à un sous-ensemble dense constitué des fonctions "en escaliers". On pose alors

$$\mathcal{C}_r = \{f \in \mathcal{C}(\Sigma_n^+) \mid \text{var}_r(f) = 0\} \text{ et } \mathcal{C} = \bigcup_{r \geq 0} \mathcal{C}_r.$$

On établira la densité de \mathcal{C} dans un prochain lemme.

Lemme III.3.7. *Soit $F \in \Lambda$ et $f \in \mathcal{C}_r$ tels que $fF \neq 0$ et $f \geq 0$. Alors, $\frac{1}{\nu(fF)\lambda}\mathcal{L}^r(fF) \in \Lambda$.*

Preuve. La positivité de $g = \lambda^{-r}\mathcal{L}^r(fF)$ découle de la positivité de f et de F . Ensuite, soit $x, x' \in \Sigma_n^+$ et $m \in \mathbf{N}$ tels que $x' \in C(x, m)$. Remarquons d'abord que, grâce à une récurrence, on a

$$\mathcal{L}^r(fF)(x) = \sum_{j_1, \dots, j_r \in \llbracket 1, n \rrbracket} \exp \left(\sum_{k=0}^{r-1} \phi(\sigma^k(j_1 \dots j_r x)) \right) f(j_1 \dots j_r x) F(j_1 \dots j_r x).$$

Soit $j_1, \dots, j_r \in \llbracket 1, n \rrbracket$, alors comme $f \in \mathcal{C}_r$ on a $f(j_1 \dots j_r x) = f(j_1 \dots j_r x')$.

Aussi, $F(j_1 \dots j_r x) \leq B_{m+r} F(j_1 \dots j_r x')$ car les deux suites coïncident sur les $m+r$ premières coordonnées. Enfin,

$$\begin{aligned}B_{m+r} \exp \left(\sum_{k=0}^{r-1} \phi(\sigma^k(j_1 \dots j_r x)) \right) &\leq B_{m+r} \exp \left(\sum_{k=0}^{r-1} \left(\text{var}_{m+r-k}(\phi) + \phi(\sigma^k(j_1 \dots j_r x')) \right) \right) \\ &\leq \left(B_{m+r} \exp \left(\sum_{k=m+1}^{m+r} b\alpha^k \right) \right) \exp \left(\sum_{k=0}^{r-1} \phi(\sigma^k(j_1 \dots j_r x')) \right) \\ &\leq B_m \exp \left(\sum_{k=0}^{r-1} \phi(\sigma^k(j_1 \dots j_r x')) \right).\end{aligned}$$

Ainsi, chacun des termes de $\mathcal{L}^r(fF)(x)$ est majoré par B_m fois le terme correspondant dans $\mathcal{L}^r(fF)(x')$, d'où $\mathcal{L}^r(fF)(x) \leq B_m \mathcal{L}^r(fF)(x')$.

Finalement, il reste à vérifier que $\nu(fF) > 0$. Pour ce faire, si $x, z \in \Sigma_n^+$ alors,

$$\frac{1}{\lambda} \mathcal{L}(\mathcal{L}^r(fF))(x) = \lambda^{-1} \sum_{\sigma y=x} e^{\phi(y)} \mathcal{L}^r(fF)(y) \geq \lambda^{-1} e^{-\|\phi\|} B_0^{-1} \mathcal{L}^r(fF)(z) = K^{-1} \mathcal{L}^r(fF)(z).$$

Or comme $fF \neq 0$, il existe $z \in \Sigma_n^+$ tel que $(fF)(z) > 0$, donc $\mathcal{L}^r(fF)(\sigma^r z) > 0$. Ainsi,

$$\nu(fF) = \frac{1}{\lambda^r} \nu(\lambda^{-1} \mathcal{L}(\mathcal{L}^r(fF))) \geq \frac{1}{K \lambda^r} \mathcal{L}^r(fF)(\sigma^r z) > 0.$$

Et enfin, on a bien $\nu(\nu(fF)^{-1} \lambda^{-r} \mathcal{L}^r(fF)) = 1$, ce qui conclut ce lemme. \square

Lemme III.3.8. Soit $f \in \mathcal{C}_r$, $F \in \Lambda$ et $n \in \mathbf{N}$. Alors,

$$\left\| \frac{1}{\lambda^{n+r}} \mathcal{L}^{n+r}(fF) - \nu(fF)h \right\| \leq A\nu(|fF|)\beta^n.$$

Remarque. La fonction $F \in \Lambda$ n'a aucune utilité dans la preuve du théorème III.3.1 (pour le montrer on prendra $F = 1$), elle sert plus tard pour montrer que $h \cdot \nu$ est la mesure de Gibbs du théorème III.1.1.

Preuve. Décomposons f en sa partie positive f^+ et sa partie négative f^- toute deux positives et vérifiant $f = f^+ - f^-$. Si $f^\pm F \neq 0$, alors grâce aux lemmes III.3.7 et III.3.6, on a

$$\left\| \frac{1}{\lambda^{n+r}} \mathcal{L}^{n+r}(f^\pm F) - \nu(f^\pm F)h \right\| = \nu(f^\pm F) \left\| \frac{1}{\lambda^{n+r} \nu(f^\pm F)} \mathcal{L}^{n+r}(f^\pm F) - h \right\| \leq A\nu(f^\pm F)\beta^n.$$

Dans le cas où $f^\pm F = 0$ l'inégalité est triviale. Ainsi,

$$\left\| \frac{1}{\lambda^{n+r}} \mathcal{L}^{n+r}(fF) - \nu(fF)h \right\| \leq A(\nu(f^+ F) + \nu(f^- F))\beta^n = A\nu(|fF|)\beta^n.$$

\square

Lemme III.3.9. L'ensemble $\mathcal{C} = \bigcup_{r \geq 0} \mathcal{C}_r$ vérifie la propriété suivante :

$$\forall f \in \mathcal{C}(\Sigma_n^+), \forall \varepsilon > 0, \exists g_1, g_2 \in \mathcal{C}, \|g_1 - g_2\| \leq \varepsilon \text{ et } g_1 \leq f \leq g_2,$$

et donc \mathcal{C} est dense dans $\mathcal{C}(\Sigma_n^+)$.

Preuve. Soit $f \in \mathcal{C}(\Sigma_n^+)$ et $\varepsilon > 0$. Alors il existe $r \in \mathbf{N}$ tel que pour tout $x \in \Sigma_n^+$ et $x' \in C(x, r)$ on ait $|f(x) - f(x')| \leq \varepsilon$, car f est continue sur le compact Σ_n^+ donc uniformément continue par le théorème de Heine. On pose alors pour $x \in \Sigma_n^+$,

$$\begin{cases} g_1(x) = \inf_{z \in \Sigma_n^+} f(x_1 \cdots x_r z), \\ g_2(x) = \sup_{z \in \Sigma_n^+} f(x_1 \cdots x_r z), \end{cases}$$

de sorte que $g_1 \leq f \leq g_2$ et $g_1, g_2 \in \mathcal{C}_r \subseteq \mathcal{C}$. De plus, si on note $z, z' \in \Sigma_n^+$ tel que $g_1(x) = f(x_1 \cdots x_r z)$ et $g_2(x) = f(x_1 \cdots x_r z')$, alors

$$|g_1(x) - g_2(x)| \leq |f(x_1 \cdots x_r z) - f(x_1 \cdots x_r z')| \leq \varepsilon.$$

\square

Enfin, on arrive au dernier lemme permettant d'établir le (3) du théorème III.3.1, pour lequel il ne reste plus que la limite sans la convergence exponentielle.

Proposition III.3.10. *Soit $f \in \mathcal{C}(\Sigma_n^+)$, alors*

$$\left\| \frac{1}{\lambda^m} \mathcal{L}^m f - \nu(f)h \right\| \xrightarrow{m \rightarrow \infty} 0.$$

Preuve. Soit $\varepsilon > 0$, $r \in \mathbf{N}$ et $g_1, g_2 \in \mathcal{C}_r$ comme dans le précédent lemme. On applique le lemme III.3.8 avec $F = 1$, ce qui donne pour m assez grand,

$$\begin{aligned} \left\| \frac{1}{\lambda^m} \mathcal{L}^m g_i - \nu(f)h \right\| &\leq \left\| \lambda^{-m} \mathcal{L}^m g_i - \nu(g_i)h \right\| + |\nu(g_i) - \nu(f)| \|h\| \\ &\leq \varepsilon(1 + \|h\|). \end{aligned}$$

De plus, on a $\lambda^{-m} \mathcal{L}^m g_1 \leq \lambda^{-m} \mathcal{L}^m f \leq \lambda^{-m} \mathcal{L}^m g_2$ et donc pour m assez grand,

$$-\varepsilon(1 + \|h\|) \leq \lambda^{-m} \mathcal{L}^m g_1 - \nu(f)h \leq \lambda^{-m} \mathcal{L}^m f - \nu(f)h \leq \lambda^{-m} \mathcal{L}^m g_2 - \nu(f)h \leq \varepsilon(1 + \|h\|).$$

Finalement, on a bien $\lim_{m \rightarrow \infty} \|\lambda^{-m} \mathcal{L}^m f - \nu(f)h\| = 0$, pour tout $f \in \mathcal{C}(\Sigma_n^+)$. \square

Remarque. La densité de \mathcal{C} dans $\mathcal{C}(\Sigma_n^+)$, à savoir

$$\forall f \in \mathcal{C}(\Sigma_n^+), \forall \varepsilon > 0, \exists r \in \mathbf{N}, \exists f_r \in \mathcal{C}_r, \|f - f_r\| \leq \varepsilon,$$

ne suffit pas pour conclure. En effet, si on a une telle fonction $f_r \in \mathcal{C}_r$ pour un certain $r \in \mathbf{N}$,

$$\begin{aligned} \left\| \frac{1}{\lambda^m} \mathcal{L}^m f - \nu(f)h \right\| &\leq \frac{1}{\lambda^m} \|\mathcal{L}^m f - \mathcal{L}^m f_r\| + \left\| \frac{1}{\lambda^m} \mathcal{L}^m f_r - \nu(f_r)h \right\| + |\nu(f_r) - \nu(f)| \|h\| \\ &\leq \frac{1}{\lambda^m} \|\mathcal{L}^m f - \mathcal{L}^m f_r\| + A\nu(|f_r|)\beta^{m-r} + \varepsilon \|h\|. \end{aligned}$$

Or, on n'arrive pas à majorer le premier terme, notamment car on ne possède aucune estimation de λ .

III.4 Construction d'une mesure de Gibbs

Soit λ, ν et h comme dans le théorème 3.1. On pose $\mu = h \cdot \nu \in \mathcal{M}(\Sigma_n^+)$.

Lemme III.4.1. *La mesure de probabilité $\mu = h \cdot \nu$ est invariante par σ :*

$$\sigma_* \mu = \mu.$$

Preuve. Soit $f, g \in \mathcal{C}(\Sigma_n^+)$, on remarque que $(\mathcal{L}f)g = \mathcal{L}(f \cdot (g \circ \sigma))$. En effet, pour $x \in \Sigma_n$, on a

$$\begin{aligned} ((\mathcal{L}f)g)(x) &= \sum_{y \in \sigma^{-1}x} e^{\phi(y)} f(y)g(x) = \sum_{y \in \sigma^{-1}x} e^{\phi(y)} f(y)g(\sigma y) \\ &= \mathcal{L}(f \cdot (g \circ \sigma))(x). \end{aligned}$$

Donc pour $f \in \mathcal{C}(\Sigma_n^+)$,

$$\begin{aligned} \mu(f) &= \nu(h \cdot f) = \nu(\lambda^{-1}(\mathcal{L}h) \cdot f) = \lambda^{-1} \nu(\mathcal{L}(h \cdot (f \circ \sigma))) \\ &= \lambda^{-1} \mathcal{L}^* \nu(h \cdot (f \circ \sigma)) = \nu(h \cdot (f \circ \sigma)) = \mu(f \circ \sigma). \end{aligned}$$

\square

A chaque fonction $f \in \mathcal{C}(\Sigma_n)$, on associe une nouvelle fonction $[f] \in \mathcal{C}(\Sigma_n^+)$ définie par :

$$\forall x \in \Sigma_n^+, [f](x) = \min \{ f(y) \mid y \in \Sigma_n, \forall i \geq 0, x_i = y_i \}.$$

De cette façon, on retire la dépendance de f par rapport à ses coordonnées négatives. Par ailleurs, si $f \in \mathcal{C}(\Sigma_n^+)$, alors $[f] = f$.

Lemme III.4.2. *Pour tout $f \in \mathcal{C}(\Sigma_n)$, la suite $(\mu([f \circ \sigma^m]))_{m \geq 0}$ admet une limite, que l'on note $G(f)$. De plus G vérifie $G(1) = 1$ et pour toute fonction $f \in \mathcal{C}(\Sigma_n)$, $f \geq 0$, on a $G(f) \geq 0$.*

Par ailleurs,

$$G(f \circ \sigma) = G(f).$$

Preuve. Soit $f \in \mathcal{C}(\Sigma_n)$, on va montrer que $(\mu([f \circ \sigma^n]))$ est une suite de Cauchy. Soit $m, k \in \mathbf{N}$, alors pour tout $x \in \Sigma_n^+$, il existe $y, y' \in \Sigma_n$ tel que

$$\begin{aligned} [f \circ \sigma^m](\sigma^k x) &= f(\sigma^m(\cdots y_{-2} y_{-1} \mathbf{x}_k x_{k+1} \cdots)) = f(\cdots y_{-1} x_k \cdots \mathbf{x}_{m+k} x_{m+k+1} \cdots), \\ [f \circ \sigma^{m+k}](x) &= f(\sigma^{m+k}(\cdots y'_{-2} y'_{-1} \mathbf{x}_0 x_1 \cdots)) = f(\cdots y'_{-1} x_0 \cdots \mathbf{x}_{m+k} x_{m+k+1} \cdots). \end{aligned}$$

Ainsi,

$$\left\| [f \circ \sigma^m] \circ \sigma^k - [f \circ \sigma^{m+k}] \right\| \leq \text{var}_k(f),$$

et donc,

$$\left| \mu([f \circ \sigma^m]) - \mu([f \circ \sigma^{m+k}]) \right| = \left| \mu \left([f \circ \sigma^m] \circ \sigma^k - [f \circ \sigma^{m+k}] \right) \right| \leq \text{var}_k(f) \longrightarrow 0,$$

lorsque $k \rightarrow \infty$, car f est continue. On a alors prouvé que $(\mu([f \circ \sigma^m]))_m$ est une suite de Cauchy, ce qui assure sa convergence vers un réel $G(f)$.

Comme $1 \in \mathcal{C}(\Sigma_n^+)$ et pour tout $m \in \mathbf{N}$ on a $1 \circ \sigma^m = 1$, d'où

$$\mu([1 \circ \sigma^m]) = \mu(1) = 1.$$

Enfin, si $f \in \mathcal{C}(\Sigma_n)$ est positive, alors $[f] \geq 0$ et donc pour tout $m \in \mathbf{N}$, $[f \circ \sigma^m] \geq 0$, d'où $G(f) \geq 0$. \square

Grâce au lemme précédent, on a construit une forme linéaire G et par l'identification rendue possible par le théorème de Riesz, on peut considérer la mesure $\tilde{\mu} \in \mathcal{M}_\sigma(\Sigma_n)$ associée à G .

Théorème III.4.3. *La mesure $\tilde{\mu}$ est une mesure de Gibbs pour $\phi \in \mathcal{H}(\Sigma_n^+)$.*

Pour prouver ce résultat, nous allons avoir besoin d'un lemme supplémentaire. Pour ce faire, on définit la constante a donnée par :

$$a = \sum_{k=0}^{\infty} \text{var}_k(\phi) < \infty.$$

Lemme III.4.4. *Soit $x, y \in \Sigma_n$ et $m \in \mathbf{N}$ tels que $y \in C(x, m)$. Alors,*

$$|S_m \phi(x) - S_m \phi(y)| \leq a.$$

Preuve. Pour $y \in \Sigma_n$, on définit $y' \in \Sigma_n$ par

$$y'_i = \begin{cases} x_i & \text{si } i < 0, \\ y_i & \text{sinon.} \end{cases}$$

Pour $k \geq 0$, on a alors $\phi(\sigma^k y) = \phi(\sigma^k y')$ car $\phi \in \mathcal{C}(\Sigma_n^+)$. Ainsi,

$$\begin{aligned} |S_m \phi(x) - S_m \phi(y)| &= |S_m \phi(x) - S_m \phi(y')| \leq \sum_{k=0}^{m-1} |\phi(\sigma^k x) - \phi(\sigma^k y')| \\ &\leq \sum_{k=0}^{m-1} \text{var}_{m-1-k}(\phi) \leq a. \end{aligned}$$

□

Preuve du théorème III.4.3. Soit $E = C(x, m) = \{y \in \Sigma_n \mid \forall i \in \llbracket 0, m-1 \rrbracket, x_i = y_i\}$, et on veut montrer que

$$c_1 e^{-Pm + S_m \phi(x)} \leq \tilde{\mu}(E) \leq c_2 e^{-Pm + S_m \phi(x)},$$

pour certaines constantes $c_1, c_2 > 0$ et $P \in \mathbf{R}$.

Pour tout $z \in \Sigma_n^+$, il existe un unique $y' = x_0 x_1 \cdots x_{m-1} z_0 z_1 \cdots \in \Sigma_n^+$ tel que $\sigma^m y' = z$ et $y'_i = x_i$ pour $0 \leq i \leq m-1$. Ainsi,

$$\mathcal{L}^m(h \mathbf{1}_E)(z) = \sum_{\sigma^m y = z} e^{S_m \phi(y)} h(y) \mathbf{1}_E(y) \leq e^{S_m \phi(y')} h(y') \leq e^{S_m \phi(x)} e^a \|h\|,$$

et donc, comme $\mathbf{1}_E \in \mathcal{C}(\Sigma_n^+)$, on a $\tilde{\mu}(E) = \mu(E)$, d'où

$$\tilde{\mu}(E) = \nu(h \mathbf{1}_E) = \lambda^{-m} \nu(\mathcal{L}^m(h \mathbf{1}_E)) \leq \lambda^{-m} e^{S_m \phi(x)} \|h\| e^a.$$

On peut alors poser $c_2 = e^a \|h\| > 0$. Pour l'autre inégalité, remarquons que pour $z \in \Sigma_n^+$, il existe au moins un $y' = x_0 x_1 \cdots x_{m-1} z_0 z_1 \cdots \in \Sigma_n^+$ tel que $\sigma^{m+1} y' = z$ et $x_i = y'_i$ pour $i \in \llbracket 0, m-1 \rrbracket$, ce qui donne

$$\begin{aligned} \mathcal{L}^{m+1}(h \mathbf{1}_E)(z) &= \sum_{\sigma^{m+1} y = z} e^{S_m \phi(y) + \phi(\sigma^m y)} h(y) \mathbf{1}_E(y) \\ &\geq e^{S_m \phi(y') - \|\phi\|} h(y') \\ &\geq (\inf h) e^{-\|\phi\| - a} e^{S_m \phi(x)}, \end{aligned}$$

et donc

$$\tilde{\mu}(E) = \nu(h \mathbf{1}_E) = \lambda^{-(m+1)} \nu(\mathcal{L}^{m+1}(h \mathbf{1}_E)) \geq \lambda^{-m} e^{S_m \phi(x)} \underbrace{(\lambda(\inf h) e^{-\|\phi\| - a})}_{c_2}.$$

On a les inégalités souhaitées avec $P = \log \lambda$, ce qui montre que $\tilde{\mu}$ est bien une mesure de Gibbs. □

Proposition III.4.5. La mesure $\tilde{\mu}$ est mélangeante par rapport à σ : pour tout boréliens $E, F \in \mathcal{B}(\Sigma_n)$ on a

$$\tilde{\mu}(E \cap \sigma^{-m} F) \xrightarrow{m \rightarrow \infty} \tilde{\mu}(E) \tilde{\mu}(F).$$

Preuve. Soit $x \in \Sigma_n$, et $f \in \mathcal{C}(\Sigma_n^+)$. On prouve par récurrence que

$$\mathcal{L}^m f(x) = \sum_{\sigma^m y = x} e^{S_m \phi(y)} f(y).$$

Ainsi, si $g \in \mathcal{C}(\Sigma_n^+)$, on a

$$((\mathcal{L}^m f) \cdot g)(x) = \sum_{\sigma^m y = x} e^{S_m \phi(y)} f(y) g(\sigma^m y) = \mathcal{L}^m (f \cdot (g \circ \sigma^m)).$$

Pour prouver que $\tilde{\mu}$ est mélangeante, il suffit de le montrer pour E, F dans la base de la topologie de Σ_n . Soit alors $a, b \in \Sigma_n$ et $E = \{y \in \Sigma_n \mid y_i = a_i, r \leq i \leq s\}$ et $F = \{y \in \Sigma_n \mid y_i = b_i, r \leq i \leq s\}$. Comme $\tilde{\mu}$ est invariante par σ , on peut supposer $r, u \geq 0$, de telle sorte que $\mathbf{1}_E, \mathbf{1}_F \in \mathcal{C}(\Sigma_n^+)$.

$$\begin{aligned} \tilde{\mu}(E \cap \sigma^{-m} F) &= \tilde{\mu}(\mathbf{1}_E \mathbf{1}_{\sigma^{-m} F}) = \tilde{\mu}(\mathbf{1}_E (\mathbf{1}_F \circ \sigma^m)) \\ &= \nu(h \mathbf{1}_E (\mathbf{1}_F \circ \sigma^m)) \\ &= \lambda^{-m} \mathcal{L}^{*m} \nu(h \mathbf{1}_E (\mathbf{1}_F \circ \sigma^m)) \\ &= \nu(\lambda^{-m} \mathcal{L}^m (h \mathbf{1}_E) \mathbf{1}_F). \end{aligned}$$

De là, on obtient grâce au lemme III.3.8 et au fait que $\mathbf{1}_E \in \mathcal{C}_s$

$$\begin{aligned} |\tilde{\mu}(E \cap \sigma^{-m} F) - \tilde{\mu}(E) \tilde{\mu}(F)| &= |\mu(E \cap \sigma^{-m} F) - \nu(h \mathbf{1}_E) \nu(h \mathbf{1}_F)| \\ &= |\nu[(\lambda^{-m} \mathcal{L}^m (h \mathbf{1}_E) - \nu(h \mathbf{1}_E) h) \mathbf{1}_F]| \\ &\leq \|\lambda^{-m} \mathcal{L}^m (h \mathbf{1}_E) - \nu(h \mathbf{1}_E) h\| \nu(F) \\ &\leq A \nu(h \mathbf{1}_E) \nu(F) \beta^{m-s} \longrightarrow 0, \end{aligned}$$

lorsque $m \rightarrow \infty$, ce qui permet de conclure. \square

Finalement, on a construit une mesure de Gibbs mélangeante donc ergodique. En vertu de la proposition III.1.3, cette mesure est l'unique mesure de Gibbs pour cette fonction de potentiel ϕ .

III.5 Généralisation au sous-décalage de type fini

Considérons une matrice A de taille $n \times n$ dont chacun des coefficients est soit 1 soit 0. On dira qu'une telle matrice est une *matrice d'incidence*. On peut la voir comme la matrice associée à un graphe orienté $\mathcal{G} = (\mathcal{S}, \mathcal{A})$ dès lors que $\mathcal{S} = \llbracket 1, n \rrbracket$ et

$$A_{i,j} = 1 \iff (i, j) \in \mathcal{A}.$$

Ainsi, l'ensemble Σ_A des suites de sommets valide pour ce graphe est inclus dans Σ_n , plus précisément

$$\Sigma_A = \{x \in \Sigma_n \mid \forall i \in \mathbf{Z}, A_{x_i, x_{i+1}} = 1\}.$$

Dans la suite on supposera de plus que pour tout entier k compris entre 1 et n , il existe toujours $x \in \Sigma_A$ tel que $x_0 = k$.

Définition III.5.1. On appelle sous-décalage de type fini (de Σ_n) le couple (Σ_A, σ_A) tel que pour tout $x \in \Sigma_A$ on ait

$$\forall i \in \mathbf{Z}, (\sigma_A x)_i = x_{i+1},$$

pour une matrice d'incidence A .

Dans ce nouveau cadre, le théorème d'existence des mesures de Gibbs est encore valide sous une hypothèse supplémentaire :

Théorème III.5.1 (Existence des mesures de Gibbs). *Supposons qu'il existe $M \in \mathbf{N}$ tel que tout les coefficients de A^M soit strictement positifs, et $\phi \in \mathcal{H}(\Sigma_A)$.*

Alors il existe une unique mesure $\mu = \mu_\phi \in \mathcal{M}_\sigma(\Sigma_A)$ et des constantes $c_1, c_2 > 0$ et $P = P(\phi) \in \mathbf{R}$ telles que

$$\forall x \in \Sigma_A, \forall m \in \mathbf{N}, \frac{\mu(C_A(x, m))}{\exp(-Pm + S_m \phi(x))} \in [c_1, c_2],$$

où $C_A(x, m) = C(x, m) \cap \Sigma_A$.

Remarque. La preuve dans ce cadre reprend les mêmes idées que dans le cas où $A = (1)_{1 \leq i, j \leq n}$, l'essentiel à adapter étant de prendre des puissances de \mathcal{L} multiples de M (où M vérifie $A^M > 0$) pour avoir des chemins valides entre toutes les paires de points du graphe A . On peut retrouver la preuve du cas plus général dans [Bow75].

IV.

DYNAMIQUE HYPERBOLIQUE

IV.1 Automorphismes hyperboliques du tore

Soit f un endomorphisme linéaire bijectif de $E = \mathbf{R}^n$. On pose $\text{Sp}_- f = \{\lambda \in \text{Sp } f \mid |\lambda| < 1\}$ et $\text{Sp}_+ = \{\lambda \in \text{Sp } f \mid |\lambda| > 1\}$. On peut maintenant définir le sous-espace stable E_s et le sous-espace instable E_u de f par

$$\begin{cases} E_s &= \bigoplus_{\lambda \in \text{Sp}_- f} E_\lambda(f), \\ E_u &= \bigoplus_{\lambda \in \text{Sp}_+ f} E_\lambda(f), \end{cases}$$

On dira qu'une norme $\|\cdot\|$ est adaptée à f si pour tout $v_s \in E_s$ et $v_u \in E_u$ on a

$$\|v_s + v_u\| = \max\{\|v_s\|, \|v_u\|\}.$$

On introduit maintenant les endomorphismes linéaires du tore $\mathbf{T}^n = \mathbf{R}^n / \mathbf{Z}^n$. Pour ce faire, dans toute la suite on notera $p: \mathbf{R}^n \rightarrow \mathbf{T}^n$ la projection de \mathbf{R}^n sur le tore. De plus, si $\|\cdot\|$ est une norme sur \mathbf{R}^n , on définit la distance quotient d sur le tore donnée par

$$\forall x, y \in \mathbf{T}^n, d(x, y) = \inf\{\|u - v\| \mid u, v \in \mathbf{R}^n, p(u) = x, p(v) = y\}.$$

Proposition IV.1.1. *Soit M une matrice de taille $n \times n$ et $f = f_M$ l'endomorphisme associé à M . Si M est à coefficients entiers, alors f se factorise en un endomorphisme du tore \mathbf{T}^n .*

Preuve. Si M est à coefficients entiers, si on considère $x \in \mathbf{R}^n, y \in \mathbf{Z}^n$ alors $M(x + y) = Mx + My \in Mx + \mathbf{Z}^n$, et donc $M(x + y) \equiv Mx$ dans \mathbf{T}^n . Ainsi f se factorise sur le tore en $\tilde{f}(x) = Mx \pmod{1}$, de sorte que $p \circ f = \tilde{f} \circ p$. \square

Proposition IV.1.2. *Soit $M \in M_n(\mathbf{Z})$. Alors M est inversible dans $M_n(\mathbf{Z})$ si et seulement si $\det M = \pm 1$.*

Définition IV.1.1. Soit M une matrice à coefficients entiers. On dit qu'une matrice inversible M est hyperbolique si elle possède n valeurs propres (comptées avec leur multiplicité) de module différent de 1 ie. $\text{Sp } f \cap \mathbf{S}^1 = \emptyset$ et que $E = E_s \oplus E_u$.

On dit qu'un automorphisme $f = f_M$ du tore \mathbf{T}^n est hyperbolique si la matrice M est hyperbolique et de déterminant ± 1 .

Remarque. D'un point de vue géométrique, un automorphisme hyperbolique dilate l'espace E_u et contracte E_s .

Désormais, dans toute la suite on notera f un automorphisme hyperbolique du tore \mathbf{T}^n associé à la matrice M et $\|\cdot\|$ une norme adaptée à la décomposition de cette matrice en somme de sous-espaces stable et instable.

IV.2 Dynamique symbolique

Définition IV.2.1. Soit (x_i) une suite de points du tore.

— On dit que (x_i) est une η -pseudo-orbite si

$$\forall i \in \mathbf{Z}, d(f(x_i), x_{i+1}) \leq \eta.$$

— On dit que (x_i) est ε -pistée par l'orbite du point $x \in \mathbf{T}^n$ si

$$\forall i \in \mathbf{Z}, d(f^i(x), x_i) \leq \varepsilon.$$

Lemme IV.2.1 (Lemme de pistage). *Pour $\varepsilon > 0$, il existe un $\eta > 0$ tel que si (x_i) est une η -pseudo-orbite, alors il existe un unique point $x \in \mathbf{T}^n$ tel que (x_i) est ε -pistée par l'orbite de x .*

On peut trouver une preuve de ce lemme dans [Bow75]

Définition IV.2.2. Soit $\varepsilon > 0$ et $x \in \mathbf{T}^n$. La variété stable locale de f en x , notée $W_\varepsilon^s(x)$, est définie par

$$W_\varepsilon^s(x) = \{y \in \mathbf{T}^n \mid \forall n \geq 0, d(f^n(x), f^n(y)) \leq \varepsilon\},$$

et la variété instable locale de f en x , notée $W_\varepsilon^u(x)$, donnée par

$$W_\varepsilon^u(x) = \{y \in \mathbf{T}^n \mid \forall n \leq 0, d(f^n(x), f^n(y)) \leq \varepsilon\}.$$

Remarque. La variété stable de f en un point x donne l'ensemble des points du tore qui ont le même "futur" que x pour la dynamique donnée par f , et la variété instable de f donne les points qui ont le même "passé" que x pour f . On peut alors remarquer que pour que deux points aient le même futur, il est nécessaire que ces deux points soient dans le même sous-espace affine dirigé par E_s . Pour que deux points aient le même passé, il faut qu'ils soient sur le même sous-espace affine dirigé par E_u .

Proposition IV.2.2. *Soit $\varepsilon > 0$ et deux points x, y du tore et $u, v \in \mathbf{R}^n$ tels que $p(u) = x, p(v) = y$. Alors $W_\varepsilon^s(x) = p(B(u, \varepsilon) \cap (u + E_s))$ et $W_\varepsilon^u(x) = p(B(u, \varepsilon) \cap (u + E_u))$.*

Proposition IV.2.3. *Soit $\varepsilon > 0$ et $x, y \in \mathbf{T}^n$. Alors*

1. $f(W_\varepsilon^s(x)) \subseteq W_\varepsilon^s(f(x))$ et $f(W_\varepsilon^u(x)) \supseteq W_\varepsilon^u(f(x))$,
2. si $d(x, y) \leq \varepsilon$, alors $W_\varepsilon^s(x) \cap W_\varepsilon^u(y)$ est un singleton, et on note $[x, y]$ son unique élément,
3. l'application $(x, y) \mapsto [x, y]$ est continue et on l'appelle produit local.

Preuve. Soit $y \in W_\varepsilon^s(x)$, alors pour tout $n \geq 0$, on a $d(f^n(x), f^n(y)) \leq \varepsilon$. En particulier, pour tout $n \geq 0, d(f^n(f(x)), f^n(f(y))) \leq \varepsilon$, d'où $f(y) \in W_\varepsilon^s(f(x))$. De même pour l'inclusion pour les variétés instables, ce qui prouve le premier point.

Supposons que $d(x, y) \leq \varepsilon$, on considère alors $u, v \in \mathbf{R}^n$ comme dans la proposition précédente et vérifiant $\|u - v\| \leq \varepsilon$. Alors $\{w\} = (u + E_s) \cap (v + E_u)$ est un singleton, car $E_s \oplus E_u = \mathbf{R}^n$ et donc $E_s \cap E_u = \{0\}$. De plus $w \in B(u, \varepsilon) \cap B(v, \varepsilon) \neq \emptyset$ car $u - w \in E_s$ et $v - w \in E_u$

$$\|u - w\| \leq \max\{\|u - w\|, \|v - w\|\} = \|u - w + w - v\| \leq \varepsilon.$$

Ainsi, $p(w) \in W_\varepsilon^s(x) \cap W_\varepsilon^u(y)$ et c'est le seul élément dans cette intersection.

Pour la continuité du produit local, remarquons pour que $[x, y] \in B(z, \delta)$ alors $x \in B(z, r)$ et $y \in B(z, r)$ où $r = \min\{\varepsilon, \delta\}$, et $d(x, y) \leq \varepsilon$. \square

Définition IV.2.3. Soit $\mathcal{R} \subseteq \mathbf{T}^n$. On dit que \mathcal{R} est un rectangle dès lors que

$$\forall x, y \in \mathcal{R}, [x, y] \in \mathcal{R}.$$

On dira que \mathcal{R} est un rectangle propre si c'est un rectangle et que $\mathcal{R} = \overline{\mathcal{R}}$.

De plus, quand \mathcal{R} est un rectangle, on notera

$$W_{\mathcal{R}}^s(x) = W_{\varepsilon}^s(x) \cap \mathcal{R} \quad \text{et} \quad W_{\mathcal{R}}^u(x) = W_{\varepsilon}^u(x) \cap \mathcal{R}.$$

Proposition IV.2.4. Soit R un rectangle de \mathbf{T}^n . Alors en identifiant par rapport aux sous-espaces stables et instables, on peut décomposer le bord de R sous la forme $\partial R = \partial^s R \cup \partial^u R$, avec $\partial^s R = \{x \in R \mid W_{\varepsilon}^s(x) \cap \text{Int } R = \emptyset\}$ et $\partial^u R = \{x \in R \mid W_{\varepsilon}^u(x) \cap \text{Int } R = \emptyset\}$.

On peut enfin introduire la notion de partition de Markov, qui permet de coder la dynamique de f dans un espace de Bernoulli.

Définition IV.2.4. Une partition de Markov de \mathbf{T}^n est un recouvrement fini $\mathcal{R} = (R_i)$ de \mathbf{T}^n par des rectangles propres vérifiant :

1. pour tout $i \neq j$, on a $\overset{\circ}{R}_i \cap \overset{\circ}{R}_j = \emptyset$,
2. si $x \in \overset{\circ}{R}_i$ et $f(x) \in \overset{\circ}{R}_j$, alors

$$\begin{cases} f(W_{\overset{\circ}{R}_i}^s(x)) \subseteq W_{\overset{\circ}{R}_j}^s(f(x)), \\ f(W_{\overset{\circ}{R}_i}^u(x)) \supseteq W_{\overset{\circ}{R}_j}^u(f(x)). \end{cases}$$

De plus, la matrice d'incidence A (dont les coefficients sont dans $\{0, 1\}$) associé à la partition de Markov \mathcal{R} est donnée par

$$A_{i,j} = 1 \iff f(\overset{\circ}{R}_i) \cap \overset{\circ}{R}_j \neq \emptyset.$$

Théorème IV.2.5. Soit $\mathcal{R} = (R_i)_{1 \leq i \leq m}$ une partition de Markov et (Σ_A, σ) l'espace de Bernoulli associé à la matrice d'incidence A de la partition \mathcal{R} . Alors,

1. pour $\omega \in \Sigma_A$, l'intersection $\bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_i})$ est un singleton et on note $\pi(\omega)$ cet unique élément,
2. l'application $\pi: \Sigma_A \longrightarrow \mathbf{T}^n$ est continue, surjective et $f \circ \pi = \pi \circ \sigma$,
3. si $\mu \in \mathcal{M}_{\sigma}(\Sigma_A)$ est ergodique de support Σ_A , alors

$$\mu \{ \omega \in \Sigma_A \mid \text{Card } \pi^{-1}(\pi(\omega)) > 1 \} = 0.$$

De cette manière, on peut considérer que π est injective quitte à retirer un ensemble de mesure nulle pour certaines mesures (en particulier la mesure de Gibbs de la section précédente). La dynamique de f sur le tore peut alors être codée par un sous-décalage de Σ_m , permettant ainsi une étude plus simple de cette dynamique, et notamment de munir ces systèmes de mesures de probabilités. En effet si μ est la mesure de Gibbs sur Σ_A , alors la mesure $\pi_*\mu$ vérifie des propriétés sur le tore semblable à celle vérifiée sur Σ_A .

Preuve du théorème IV.2.5. Soit $\omega \in \Sigma_A$. Posons $K_n(\omega) = \bigcap_{i=-n}^n f^{-i}(R_{\omega_i})$ qui est un compact non vide pour tout $n \geq 1$. De plus la suite $(K_n(\omega))_{n \geq 1}$ est décroissante. Ainsi, en tant qu'intersection décroissante de compact non vide,

$$K = \bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_i}) = \bigcap_{i \geq 1} K_i(\omega)$$

est un compact non vide de \mathbf{T}^n . Reste à vérifier que K contient au plus un élément. Supposons par l'absurde que $x, y \in K$, alors pour tout $i \in \mathbf{Z}$ on a, en supposant que les rectangles de la partition de Markov sont de diamètre au plus ε ,

$$d(f^i(x), f^i(y)) \leq \varepsilon,$$

donc les deux points ont des orbites que se ε -pistent, et par le lemme de pistage il ne peut y en avoir qu'un. D'où $x = y$, et finalement K est un singleton et $\pi(\omega)$ est son unique élément.

Ensuite, π vérifie la relation de semi-conjugaison car

$$K(\sigma\omega) = \bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_{i+1}}) = f\left(\bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_i})\right) = f(K(\omega)),$$

et donc $\pi \circ \sigma = f \circ \pi$.

Concernant la surjectivité de π , soit $x \in \mathbf{T}^n$. On pose alors $\omega \in \Sigma_A$ de sorte que $f^i(x) \in R_{\omega_i}$, ce qui est possible car \mathcal{R} est un recouvrement de \mathbf{T}^n et un tel ω est bien dans Σ_A par construction de Σ_A . Ainsi $x \in \bigcap_{i \in \mathbf{Z}} f^{-i}(R_{\omega_i}) = \{\pi(\omega)\}$.

Pour la continuité de π , si on considère une boule B centrée en x et de rayon $r > 0$, alors, il existe $N \in \mathbf{N}$ tel que

$$\text{diam} \left(\bigcap_{-N \leq i \leq N} f^{-i}(R_{\omega_i}) \right) \leq r.$$

Ce dernier ensemble est bien un ouvert de Σ_A donc un voisinage de ω et $\pi(K_N(\omega)) \subseteq B$.

Il reste encore le point (3) à démontrer. Soit μ une mesure de probabilité σ -invariante, ergodique et de support Σ_A . Alors $\nu = \pi_*\mu$ la mesure image de μ par π est aussi ergodique, f -invariante et de support \mathbf{T}^n . Si on note $Z = \{x \in \mathbf{T}^n \mid \text{Card } \pi^{-1}(x) > 1\}$, alors le point (3) est équivalent à $\nu(Z) = 0$. Nécessairement, $Z \subseteq \bigcup_{i \in \mathbf{Z}} f^i(\partial\mathcal{R})$, il suffit donc de montrer que ce dernier est de mesure nulle pour ν . On note $\partial\mathcal{R} = \partial^u\mathcal{R} \cup \partial^s\mathcal{R}$, où $\partial^s\mathcal{R} = \bigcup_{R \in \mathcal{R}} \partial^s R$ et de même pour $\partial^u\mathcal{R}$. Or par la propriété (2) des partitions de Markov, $f(\partial^s\mathcal{R}) \subseteq \partial^s\mathcal{R}$, et donc la suite $(f^i(\partial^s\mathcal{R}))_i$ est décroissante, d'où par continuité décroissante et invariance de ν par rapport à f :

$$\nu \left(\bigcap_{i \geq 0} f^i(\partial^s\mathcal{R}) \right) = \lim_{i \rightarrow \infty} \nu(f^i(\partial^s\mathcal{R})) = \nu(\partial^s\mathcal{R}).$$

Or l'ensemble $F = \bigcap_{i \geq 0} f^i(\partial^s\mathcal{R})$ vérifie $f^{-1}(F) = F$, donc par ergodicité de ν , il est ou bien de mesure nulle ou égale à 1, cette dernière possibilité est exclue car ν est de support \mathbf{T}^n et F est strictement inclus dans le tore. Donc $\nu(F) = 0$, c'est-à-dire que $\nu(\partial^s\mathcal{R}) = 0$. On fait de même pour $\partial^u\mathcal{R}$ et on en conclut que $0 = \nu(\partial\mathcal{R}) \geq \nu(Z)$. Finalement, on a bien $\nu(Z) = 0$. \square

IV.3 Existence des partitions de Markov

L'objectif de ce paragraphe est de prouver le théorème suivant donnant l'existence de partition de Markov pour f un automorphisme hyperbolique du tore \mathbf{T}^n .

Théorème IV.3.1. *Soit f un automorphisme hyperbolique du tore \mathbf{T}^n , et $\varepsilon > 0$. Alors il existe une partition de Markov pour f dont ses rectangles sont de diamètre au plus ε .*

Considérons $\eta > 0$ et $(x_i)_{i \leq I}$ une suite η -dense dans \mathbf{T}^n (ie. tout point de \mathbf{T}^n est à distance au plus η d'un point x_i). On fixera plus tard des contraintes sur η pour faire en sorte que les rectangles de la partition de Markov soient de diamètre inférieur à ε . Soit A une matrice à coefficients dans $\{0, 1\}$ définie par

$$\forall i, j \leq I, \quad A_{i,j} = 1 \iff f(x_i) \in B(x_j, \eta(1 + \|M\|)),$$

et (Σ_A, σ) le sous-décalage associé à la matrice d'incidence A .

Proposition IV.3.2. *Il existe une application $\theta: \Sigma_A \longrightarrow \mathbf{T}^n$ vérifiant :*

1. θ est surjective et continue,
2. on a la relation de semi-conjugaison : $\theta \circ \sigma = f \circ \theta$,
3. pour tout $\omega \in \Sigma_A$ et pour tout $i \in \mathbf{Z}$, on a $d(f^i(\theta(\omega)), x_{\omega_i}) \leq \frac{\varepsilon}{2}$,

Preuve. Soit $\omega \in \Sigma_A$, alors la suite (x_{ω_i}) est une η' -pseudo-orbite où $\eta' = \eta(1 + \|M\|)$ car pour tout $i \in \mathbf{Z}$,

$$d(f(x_{\omega_i}), x_{\omega_{i+1}}) \leq \eta(1 + \|M\|).$$

Donc il existe un $\eta > 0$ tel qu'il existe un point $\theta(\omega) \in \mathbf{T}^n$ dont l'orbite $\frac{\varepsilon}{2}$ -piste la suite x_{ω_i} , par le lemme de pistage, et ainsi le point (3) est vérifié par construction de θ .

L'application θ ainsi définie est continue, pour le montrer on procède de la même manière que dans la preuve du théorème IV.2.5.

Elle est de plus surjective car si $x \in \mathbf{T}^n$, on prend ω tel que $f^i(x) \in B(x_{\omega_i}, \eta)$ pour tout $i \in \mathbf{Z}$. Une telle suite ω appartient bien à Σ_A car

$$\begin{aligned} d(f(x_{\omega_i}), x_{\omega_{i+1}}) &\leq d(f(x_{\omega_i}), f(f^i(x))) + d(f^{i+1}(x), x_{\omega_{i+1}}) \\ &\leq \|M\| d(x_{\omega_i}, f^i(x)) + d(x_{\omega_{i+1}}, f^{i+1}(x)) \\ &\leq \eta(1 + \|M\|) = \eta'. \end{aligned}$$

Ce qui permet de conclure le premier point de cette proposition.

Finalement, si $x = \theta(\omega)$ alors par la construction précédente, on remarque que $\theta(\sigma\omega) = f(x)$, et donc on a bien la relation de semi-conjugaison : $\theta \circ \sigma = f \circ \theta$. \square

Soit $\mathcal{U}_i = \{\omega \in \Sigma_A \mid \omega_0 = i\}$. On pose alors $T_i = \theta(\mathcal{U}_i)$ de sorte que $\mathcal{T} = (T_i)_{i \leq I}$ soit un recouvrement fini de \mathbf{T}^n par des fermés (car θ est continue et \mathcal{U}_i fermé). Cette étape fournit alors un codage intermédiaire de la dynamique de f .

Proposition IV.3.3. *Pour tout i , T_i est un rectangle de diamètre inférieur à ε .*

Pour démontrer cette proposition, nous allons avoir besoin d'introduire une structure de produit local sur Σ_A qui respecte le produit local sur \mathbf{T}^n et le codage θ .

Soit $\omega, \omega' \in \Sigma_A$ tels que $\omega_0 = \omega'_0$, alors on définit le produit local $[\omega, \omega']$ par

$$\forall i \in \mathbf{Z}, [\omega, \omega']_i = \begin{cases} \omega_i & \text{si } i \geq 0, \\ \omega'_i & \text{si } i \leq 0. \end{cases}$$

Lemme IV.3.4. *Soit $\omega, \omega' \in T_i$. Alors $[\theta(\omega), \theta(\omega')] = \theta([\omega, \omega'])$.*

Preuve. On note $x = \theta(\omega)$, $y = \theta(\omega')$ et $z = \theta([\omega, \omega'])$. Pour montrer ce résultat, il s'agit de montrer d'abord que $z \in W_\varepsilon^s(x)$ puis de manière identique que $z \in W_\varepsilon^u(y)$. Soit $i \geq 0$,

$$\begin{aligned} d(f^i(z), f^i(x)) &= d(f^i(\theta([\omega, \omega'])), f^i(\theta(\omega))) \\ &\leq d(f^i(\theta([\omega, \omega'])), x_{\omega_i}) + d(x_{\omega_i}, f^i(\theta(\omega))) \\ &\leq \varepsilon \end{aligned}$$

D'où, $z \in W_\varepsilon^s(x)$ et donc $z = [x, y]$. □

Preuve de la proposition IV.3.3. D'abord, on montre que T_i est un rectangle. Soit $x, y \in T_i$, alors il existe $\omega, \omega' \in \mathcal{U}_i$ tels que $x = \theta(\omega)$ et $y = \theta(\omega')$. Alors par le lemme précédent,

$$[x, y] = [\theta(\omega), \theta(\omega')] = \theta([\omega, \omega']) \in \theta(\mathcal{U}_i) = T_i.$$

Donc T_i est un rectangle.

De plus comme $x \in T_i$ et $\omega_0 = i$, alors grâce au point (3) de la proposition précédente

$$x = \theta(\omega) \in B\left(x_i, \frac{\varepsilon}{2}\right),$$

et de même pour y , donc par l'inégalité triangulaire, $d(x, y) \leq \varepsilon$. Ce qui prouve que le diamètre de T_i est majoré par ε . □

Ce recouvrement, bien qu'il ne soit pas encore une partition de Markov, vérifie une propriété semblable à la seconde condition des partitions de Markov.

Lemme IV.3.5. *Soit $\omega \in \Sigma_A$. On note $i = \omega_0, j = \omega_1$ et $x = \theta(\omega)$. Alors,*

$$\begin{cases} f(W_{T_i}^s(x)) \subseteq W_{T_j}^s(f(x)), \\ f(W_{T_i}^u(x)) \supseteq W_{T_j}^u(f(x)). \end{cases}$$

Preuve. On prouve ici seulement la première inclusion, la seconde se traitant de la même manière.

Soit $y \in W_{T_i}^s(x)$, donc déjà $f(y) \in W_\varepsilon^s(f(x))$. Il reste alors à montrer que $f(y) \in T_j$. Comme $y \in T_i$, il existe $\omega' \in \mathcal{U}_i$ tel que $\theta(\omega') = y$. Alors $f(y) = \theta([\omega, \omega'])$. En effet, $y = [x, y]$ car y est dans la variété stable de x , ainsi

$$f(y) = f([x, y]) = f(\theta[\omega, \omega']) = \theta(\sigma[\omega, \omega']) \in \theta(\mathcal{U}_j) = T_j,$$

car $\omega_1 = j$. Ce qui prouve l'inclusion $f(W_{T_i}^s(x)) \subseteq W_{T_j}^s(f(x))$. □

Lemme IV.3.6. *Soit $\omega \in \mathcal{U}_i$ et $x = \theta(\omega)$. Alors, $W_{T_i}^s(x) = \theta(\{\omega' \in \Sigma_A \mid \forall i \geq 0, \omega_i = \omega'_i\})$.*

Preuve. C'est le même argument que celui utilisé dans la preuve du lemme précédent. \square

Ce qui empêche ce recouvrement d'être une partition de Markov, c'est le fait que les rectangles qui le composent peuvent s'intersecter. Pour remédier à ce problème, on note pour chaque point $x \in \mathbf{T}^n$,

$$\mathcal{T}(x) = \{T_i \in \mathcal{T} \mid x \in T_i\} \text{ et } \mathcal{T}^*(x) = \{T_k \in \mathcal{T} \mid \exists T_j \in \mathcal{T}(x), T_j \cap T_k \neq \emptyset\}.$$

De plus, pour des rectangles T_i, T_j qui s'intersectent, on découpe T_i en 4 morceaux donnés par :

$$\begin{aligned} T(i, j, su) &= \{x \in T_i \mid W_{T_i}^s(x) \cap T_j \neq \emptyset, W_{T_i}^u(x) \cap T_j \neq \emptyset\} = T_i \cap T_j, \\ T(i, j, s0) &= \{x \in T_i \mid W_{T_i}^s(x) \cap T_j \neq \emptyset, W_{T_i}^u(x) \cap T_j = \emptyset\}, \\ T(i, j, 0u) &= \{x \in T_i \mid W_{T_i}^s(x) \cap T_j = \emptyset, W_{T_i}^u(x) \cap T_j \neq \emptyset\}, \\ T(i, j, 00) &= \{x \in T_i \mid W_{T_i}^s(x) \cap T_j = \emptyset, W_{T_i}^u(x) \cap T_j = \emptyset\}. \end{aligned}$$

Lemme IV.3.7. *Si T_i et T_j s'intersectent, alors pour tout $q \in \{su, s0, 0u, 00\}$, $T(i, j, q)$ est un rectangle.*

Preuve. Soit $x, y \in T(i, j, q)$. Alors comme $W_{T_i}^s(x) = W_{T_i}^s([x, y])$ et $W_{T_i}^u(y) = W_{T_i}^u([x, y])$, on en déduit que $[x, y] \in T(i, j, q)$ ie. $T(i, j, q)$ est un rectangle. \square

On a ainsi construit des nouveaux rectangles en nombre fini. Pour éviter les points proches des bords de ces rectangles, on définit Z^* par :

$$Z^* = \{x \in \mathbf{T}^n \mid W_\varepsilon^s(x) \cap \partial^s T_j = \emptyset, W_\varepsilon^u(x) \cap \partial^u T_j = \emptyset \text{ pour } T_j \in \mathcal{T}^*(x)\}.$$

qui est un ouvert dense de \mathbf{T}^n .

Pour chaque point $x \in \mathbf{T}^n$ on définit

$$R(x) = \bigcap \{\text{Int } T(i, j, q) \mid x \in T_i, T_i \cap T_j \neq \emptyset, x \in T(i, j, q)\},$$

qui est alors un rectangle en tant qu'intersection de rectangles. De plus, les rectangles de la forme $T(i, j, q)$ sont en nombre fini donc on ne peut en faire qu'un nombre fini d'intersection. Notons alors $\mathcal{R} = \{\overline{R(x)} \mid x \in \mathbf{T}^n\} = \{R_1, \dots, R_m\}$ ces nouveaux rectangles.

Proposition IV.3.8. *Pour $x, y \in \mathbf{T}^n$, si $y \in R(x)$ alors $R(x) = R(y)$. Ainsi, les intérieurs des rectangles de \mathcal{R} ne s'intersectent pas, et vérifie donc la propriété (1) des partitions de Markov.*

Preuve. Soit $y \in R(x)$. Alors $\mathcal{T}(x) = \mathcal{T}(y)$, en effet comme $y \in R(x)$ on a $\mathcal{T}(x) \subseteq \mathcal{T}(y)$ et si $x \notin T_i$, alors on a $R(x) \cap T_i = \emptyset$, ce qui donne l'inclusion réciproque : $\mathcal{T}(y) \subseteq \mathcal{T}(x)$.

Si $x \in T_i$ et $T_i \cap T_k \neq \emptyset$, alors il existe q tel que $x \in T(i, k, q)$. Or $y \in R(x) \subseteq T(i, k, q)$, donc finalement on a bien $R(x) = R(y)$.

Ainsi, si $R(x) \cap R(y) \neq \emptyset$, on considère z dans cette intersection, donc $R(x) = R(z) = R(y)$ par ce qui précède, ce qui donne la propriété (1) voulue. \square

Il reste maintenant à vérifier que le recouvrement \mathcal{R} vérifie la propriété (2) des partitions de Markov. Pour ce faire, on démontre d'abord un résultat intermédiaire, sur les points x tels que $W_\varepsilon^u(x)$ et $W_\varepsilon^u(f(x))$ ne rencontrent pas $\partial^u \mathcal{T} = \bigcup_{T_i \in \mathcal{T}} \partial^u T_i$

Lemme IV.3.9. *Soit $x, y \in \mathbf{T}^n$ tels que les variétés stables locales en $x, y, f(x)$ et $f(y)$ ne rencontrent $\partial^u \mathcal{T}$. Supposons de plus que $R(x) = R(y)$ et que $y \in W_\varepsilon^s(x)$. Alors $R(f(x)) = R(f(y))$.*

Preuve. Dans un premier temps, on montre que $\mathcal{T}(f(x)) = \mathcal{T}(f(y))$. Si $f(x) \in T_j$, considérons $\omega \in \Sigma_A$ tel que $\omega_0 = i$ et $\omega_1 = j$ et $f(x) = \theta(\sigma\omega)$. Alors $x \in T_i$, et donc $f(y) \in f(W_{T_i}^s(x)) \subseteq W_{T_j}^s(f(x)) \subseteq T_j$. Donc $\mathcal{T}(f(x)) \subseteq \mathcal{T}(f(y))$. Réciproquement, comme $W_{T_i}^s(x) = W_{T_i}^s(y)$, grâce au lemme IV.3.6, on a bien $\mathcal{T}(f(y)) \subseteq \mathcal{T}(f(x))$.

Ensuite, supposons que $f(x), f(y) \in T_j$ et qu'il existe T_k tel que $T_j \cap T_k \neq \emptyset$. On va montrer que $f(x)$ et $f(y)$ sont dans le même $T(j, k, q)$, ce qui permettra de conclure. Comme $f(y) \in W_\varepsilon^s(f(x))$, alors $W_{T_j}^s(f(y)) = W_{T_j}^s(f(x))$. Ainsi, si $f(x) \in T(j, k, q)$ pour $q \in \{su, s0\}$, alors $f(y) \in T(j, k, su) \cup T(j, k, s0)$, et de même lorsque $q \in \{0u, 00\}$.

On va désormais montrer l'implication suivante :

$$W_{T_j}^u(f(x)) \cap T_k \neq \emptyset \implies W_{T_j}^u(f(y)) \cap T_k \neq \emptyset.$$

Soit alors $f(z) \in W_{T_j}^u(f(y)) \cap T_k$. Notons $\omega \in \Sigma_A$ tel que $f(x) = \theta(\sigma\omega)$ avec $\omega_0 = s$ et $\omega_1 = j$. Alors par le lemme IV.3.5, on a

$$f(z) \in W_{T_j}^u(f(x)) \subseteq f(W_{T_s}^u(x)).$$

On peut alors prendre $z \in W_{T_s}^u(x)$. On note alors $\omega' \in \Sigma_A$ tel que $\theta(\sigma\omega') = f(z)$ avec $\omega'_0 = t$ et $\omega'_1 = k$. On a alors $T_s \in \mathcal{T}(x) = \mathcal{T}(y)$ et $T_t \cap T_s \neq \emptyset$. Comme $z \in T_t$, on a $z \in W_{T_s}^u(x) \cap T_t$. Or x, y sont dans le même $T(s, t, q)$ (qui est un rectangle), donc il existe $z' \in W_{T_s}^u(y) \cap T_t$. Ainsi,

$$z'' = [z, y] = [z, z'] \in \begin{cases} W_\varepsilon^s(z) \cap W_\varepsilon^u(y) & \text{par de définition du produit local,} \\ T_t & \text{car } z \text{ et } z' \text{ sont dans } T_t \text{ qui est un rectangle,} \\ T_s & \text{car } z \text{ et } z' \text{ sont dans } T_s \text{ qui est un rectangle.} \end{cases}$$

Donc $z'' \in W_{T_t}^s(z) \cap W_{T_s}^u(y)$. Or $f(z'') = [f(z), f(y)]$ qui est donc dans $W_\varepsilon^s(f(z)) \cap W_\varepsilon^u(f(y))$ par définition, mais aussi dans les rectangles T_j et T_k par hypothèse. On a ainsi construit un élément $f(z'') \in W_{T_j}^s(f(z)) \cap W_{T_k}^u(f(y)) \subseteq T_k \cap W_{T_j}^u(f(y))$ qui est donc non vide. On montre de la même manière l'implication réciproque, et finalement on en conclut que $R(f(x)) = R(f(y))$. \square

Proposition IV.3.10. *Le recouvrement \mathcal{R} vérifie la propriété (2) des partitions de Markov, à savoir : si $x \in \text{Int } R_i$ et $f(x) \in \text{Int } R_j$, alors*

$$f(W_{R_i}^s(x)) \subseteq W_{R_j}^s(f(x)) \text{ et } f(W_{R_i}^u(x)) \supseteq W_{R_j}^u(f(x)).$$

Preuve. Soit $x \in \text{Int } R_i \cap f^{-1}(\text{Int } R_j)$, alors les variétés stables locales de x et $f(x)$ ne rencontrent pas $\partial^u \mathcal{T}$. Si $y \in W_{R_i}^s(x)$, alors il en va de même, on peut alors appliquer le lemme précédent : $R_j = \overline{R(f(x))} = \overline{R(f(y))}$. Ainsi, $f(y) \in R_j \cap W_\varepsilon^s(f(x)) = W_{R_j}^s(f(x))$. L'autre inclusion se traite de la même manière. \square

Ce qui termine la preuve de ce théorème, c'est-à-dire que $\mathcal{R} = (R_i)_{1 \leq i \leq m}$ est une partition de Markov dont le diamètre de chacun de ses rectangles est inférieur à ε .

On peut alors étendre le cadre d'utilisation de ce théorème à une certaine classe de difféomorphisme, c'est l'objet du paragraphe suivant.

IV.4 Généralisation aux difféomorphismes AXIOME A

Désormais, on considère une variété riemannienne M et $f: M \rightarrow M$ un difféomorphisme.

Définition IV.4.1. On dit qu'un ensemble $\Lambda \subseteq M$ est hyperbolique si $f(\Lambda) = \Lambda$ et qu'il existe pour tout $x \in \Lambda$ deux sous-espaces E_x^s, E_x^u de $T_x M$ vérifiant :

1. $T_x M = E_x^s \oplus E_x^u$,
2. $df(E_x^s) = E_{f(x)}^s$ et $df(E_x^u) = E_{f(x)}^u$,
3. il existe des constantes $c > 0$ et $\lambda \in]0, 1[$ telles que pour tout $i \in \mathbf{N}$ on ait

$$\forall v \in E_x^s, \|df^i(v)\| \leq c\lambda^i \|v\|, \forall v \in E_x^u, \|df^{-i}(v)\| \leq c\lambda^i \|v\|.$$

4. E_x^s et E_x^u varient continûment par rapport à x .

Définition IV.4.2. Un point $x \in M$ est non-errant si pour tout voisinage U de x on a

$$U \cap \bigcup_{i>0} f^i(U) \neq \emptyset.$$

On note alors $\Omega(f)$ l'ensemble de tout les points non-errants pour f . Remarquons que les points ayant une orbite périodiques sont clairement non-errants.

On peut enfin définir ce qu'est un difféomorphisme axiome A :

Définition IV.4.3. On dit que f est un difféomorphisme AXIOME A si

1. $\Omega(f)$ est hyperbolique et compact,
2. l'ensemble des points d'orbite périodique est dense dans $\Omega(f)$.

En gardant les mêmes définitions de variété stables et instables locales dans ce cadre, on peut aussi définir ce qu'est une partition de Markov. Le théorème suivant, dû à R. Bowen, et dont la preuve garde la même structure et arguments que celle dans le cadre plus simple du tore \mathbf{T}^n , donne l'existence de partition de Markov pour ces difféomorphismes.

Théorème IV.4.1. *L'ensemble $\Omega(f)$ admet une partition de Markov de diamètre arbitrairement petit.*

Pour trouver un exemple de tels difféomorphismes, on peut considérer les difféomorphismes de \mathbf{C}^2 de la forme

$$f : \begin{cases} \mathbf{C}^2 & \longrightarrow \mathbf{C}^2 \\ (z, w) & \longmapsto (z^2 + c + aw, az) \end{cases}.$$

avec $a, c \in \mathbf{C}$, que l'on appelle applications de Hénon. Sous certaines conditions sur les constantes a et c , on peut montrer que f est un difféomorphisme AXIOME A.

Pour ce faire, posons

$$K^\pm = \{p \in \mathbf{C}^2 \mid (f^{\pm n}(p))_n \text{ est bornée}\} \text{ et } K = K^+ \cap K^-,$$

On note l'ensemble de Mandelbrot \mathcal{M} défini comme étant l'ensemble des points $c \in \mathbf{C}$ tel que la suite $(z_n)_n$ est bornée, où $z_0 = 0$ et $z_{n+1} = z_n^2 + c$ pour tout $n \in \mathbf{N}$.

Alors d'après le théorème 3.1 de [FS92] et le corollaire 6.14 de [BS91], à savoir

Théorème IV.4.2 ([FS92]). *Si $c \notin \mathcal{M}$, alors il existe $a \in \mathbf{C}$ tel que f soit hyperbolique sur K qui est alors un ensemble de Cantor.*

Théorème IV.4.3 ([BS91]). *Si f est hyperbolique, alors f vérifie l'Axiome A.*

On aboutit à ce résultat, donnant un exemple de difféomorphisme AXIOME A.

Corollaire IV.4.4. *Supposons que $c \notin \mathcal{M}$, alors il existe $a \in \mathbf{C}$ tel que l'application f donnée par $f(z, w) = (z^2 + c + aw, az)$ satisfait l'Axiome A.*

IV.5 Exemple d'un codage sur \mathbf{T}^2

Ce paragraphe est consacré à un exemple : celui d'une partition de Markov composée de trois rectangles pour un automorphisme hyperbolique du tore \mathbf{T}^2 en deux dimensions. On considère alors la matrice

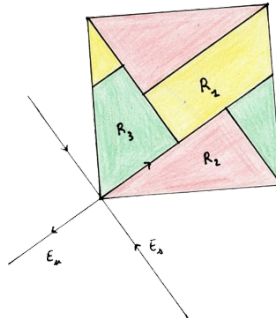
$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

et f l'automorphisme du tore \mathbf{T}^2 associé. Le polynôme caractéristique associé à cette matrice est $\chi(X) = X^2 - 3X + 1$ et donc son spectre est

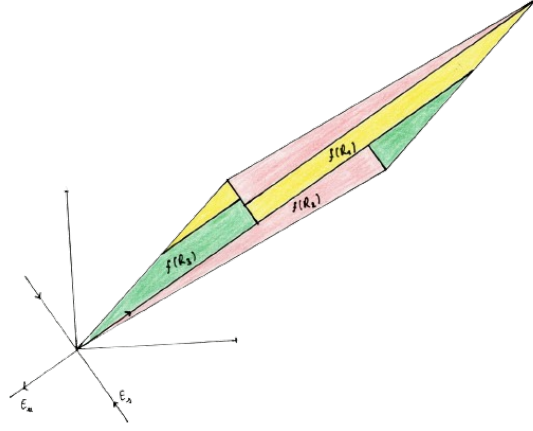
$$\text{Sp } f = \left\{ \frac{3 \pm \sqrt{5}}{2} \right\},$$

qui n'intersecte donc pas le cercle \mathbf{S}^1 . Un vecteur propre de cette matrice associé à la valeur propre de module plus grand que 1 est $v_u = \left(1 \quad \frac{\sqrt{5}-1}{2}\right)^\top$. Grâce au théorème spectral, on sait que les vecteurs propres pour la valeur propre de module plus petit que 1 sont orthogonaux à v_u , ainsi $v_s = \left(\frac{1-\sqrt{5}}{2} \quad 1\right)^\top$ en est un. On note alors E_s (resp. E_u) l'espace propre associé à v_s (resp. v_u) qui est le sous-espace stable (resp. instable) associé à f , qui sont supplémentaires et donc f est hyperbolique.

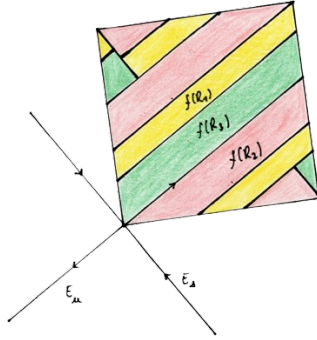
Pour construire une partition de Markov pour cet automorphisme, on commence par tracer le segment partant de l'origine dans la direction de v_u jusqu'à la droite d'équation $x = 1$. Ensuite, on trace le segment partant du point $(0, 1)$ dirigé par v_s jusqu'au segment précédent et on fait de même en partant du point $(1, 0)$. On prolonge alors le premier segment jusqu'à intersecter le second segment, comme ci-dessous (on identifie \mathbf{T}^2 au carré $[0, 1]^2$).



Pour vérifier qu'il s'agit bien d'une partition de Markov et pour obtenir la matrice d'incidence associé, on calcule les images des rectangles de la partition :



Enfin, on projette sur le tore \mathbf{T}^2 .



On en déduit donc que

$$\begin{array}{lll} \mathring{R}_1 \cap f(\mathring{R}_1) \neq \emptyset, & \mathring{R}_2 \cap f(\mathring{R}_1) \neq \emptyset, & \mathring{R}_3 \cap f(\mathring{R}_1) \neq \emptyset, \\ \mathring{R}_1 \cap f(\mathring{R}_2) = \emptyset, & \mathring{R}_2 \cap f(\mathring{R}_2) \neq \emptyset, & \mathring{R}_3 \cap f(\mathring{R}_2) \neq \emptyset, \\ \mathring{R}_1 \cap f(\mathring{R}_3) \neq \emptyset, & \mathring{R}_2 \cap f(\mathring{R}_3) = \emptyset, & \mathring{R}_3 \cap f(\mathring{R}_3) \neq \emptyset. \end{array}$$

Ce qui donne finalement la matrice d'incidence :

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Ainsi, comme $A^2 > 0$ par le théorème III.1.1, pour tout potentiel $\phi \in \mathcal{H}(\Sigma_A)$ on a une unique mesure de Gibbs μ_ϕ . De plus, par le théorème IV.2.5, la dynamique de f est conjuguée à celle de (Σ_A, σ_A) , et si on pose $\pi: \Sigma_A \rightarrow \mathbf{T}^2$ comme dans le théorème IV.2.5, alors π est injective presque partout pour μ_ϕ , ce qui conclut la construction de cette partition de Markov.

V.

APPENDICE

Considérons un espace métrique (X, d) compact. On note $\mathcal{B}(X)$ la tribu borélienne associée aux ouverts de X . Si $T: X \rightarrow X$ est un homéomorphisme de (X, d) , et μ une mesure de probabilité sur $(X, \mathcal{B}(X))$, on dira que T est un automorphisme dès lors que $T_*\mu = \mu$, où $T_*\mu$ est la mesure image de μ par T . On notera $\mathcal{M}(X)$ l'ensemble des mesures de probabilité sur $(X, \mathcal{B}(X))$ et $\mathcal{M}_T(X)$ le sous ensemble de $\mathcal{M}(X)$ constitué des mesures invariantes par T .

L'ensemble des fonctions continues de X à valeurs réelles, noté $\mathcal{C}(X)$, muni de la norme $\|f\| = \max_X |f| < \infty$ pour $f \in \mathcal{C}(X)$, est un espace de Banach.

V.1 Théorème de Riesz

On peut alors faire un lien entre $\mathcal{M}(X)$ et $\mathcal{C}(X)^*$ l'ensemble des formes linéaires de $\mathcal{C}(X)$, grâce au théorème de représentation de Riesz :

Théorème V.1.1 (Riesz). *Pour toute mesure $\mu \in \mathcal{M}(X)$, on pose $\alpha_\mu \in \mathcal{C}(X)^*$ définie par*

$$\forall f \in \mathcal{C}(X), \alpha_\mu(f) = \int_X f d\mu.$$

Alors l'application $\mu \mapsto \alpha_\mu$ est une bijection entre $\mathcal{M}(X)$ et

$$\{\alpha \in \mathcal{C}(X)^* \mid \alpha(1) = 1, \alpha(f) \geq 0 \text{ si } f \geq 0\}.$$

Dans toute la suite on notera plus simplement μ à la place de α_μ .

Par ailleurs, le fait que X soit un espace métrique compact assure le fait que $\mathcal{C}(X)$ soit séparable, *ie.* il existe une partie dénombrable dense dans $\mathcal{C}(X)$. Soit alors $(f_n)_n$ une suite de fonctions continues non nulles sur X dense dans $\mathcal{C}(X)$.

V.2 Topologie des espaces de probabilités

Proposition V.2.1. *Pour $\mu, \mu' \in \mathcal{M}(X)$, on pose*

$$d(\mu, \mu') = \sum_{n \in \mathbf{N}} 2^{-n} \|f_n\|^{-1} \left| \int_X f_n d\mu - \int_X f_n d\mu' \right|.$$

Alors l'application $d: \mathcal{M}(X) \times \mathcal{M}(X) \rightarrow \mathbf{R}$ définit une distance.

De plus, si $(\mu_n)_n$ est une suite de probabilité et $\mu \in \mathcal{M}(X)$, alors on a

$$d(\mu_n, \mu) \xrightarrow{n \rightarrow \infty} 0 \iff \forall f \in \mathcal{C}(X), \mu_n(f) \xrightarrow{n \rightarrow \infty} \mu(f).$$

Proposition V.2.2. *L'ensemble $\mathcal{M}(X)$ est convexe et compact pour la distance définit précédemment.*

Preuve. On va montrer que $\mathcal{M}(X)$ vérifie la propriété de Bolzano-Weierstrass. Soit (μ_n) une suite de probabilité et (f_n) une suite de fonctions dense dans $\mathcal{C}(X)$. On montre aisément par récurrence que pour tout $n \in \mathbf{N}$ il existe $\varphi_1, \dots, \varphi_k$ des extractrices et $G(f_1), \dots, G(f_k) \in \mathbf{R}$ tels que

$$\forall i \in \llbracket 1, k \rrbracket, \quad \mu_{\varphi_1 \circ \dots \circ \varphi_k(n)}(f_i) \xrightarrow{n \rightarrow \infty} G(f_i).$$

En notant $\psi_k = \varphi_1 \circ \dots \circ \varphi_k$, on a toujours $\mu_{\psi_k(n)}(f_i) \rightarrow G(f_i)$ lorsque $n \rightarrow \infty$ pour tout $i \leq k$. On note finalement $\psi(n) = \psi_n(n)$ et ainsi pour tout $i \in \mathbf{N}$, la suite $\mu_{\psi(n)}(f_i)$ converge vers $G(f_i)$ lorsque $n \rightarrow \infty$.

L'application $G: \{f_n \mid n \in \mathbf{N}\} \rightarrow \mathbf{R}$ est uniformément continue, en effet soit $\varepsilon > 0$ et $f, g \in \{f_n \mid n \in \mathbf{N}\}$ tels que $\|f - g\| \leq \varepsilon$. Alors,

$$|G(f) - G(g)| \leq |G(f) - \mu_{\psi(n)}(f)| + |\mu_{\psi(n)}(f) - \mu_{\psi(n)}(g)| + |\mu_{\psi(n)}(g) - G(g)| \leq 3\varepsilon,$$

pour n assez grand. Donc G est uniformément continue à valeurs dans un espace complet, on peut alors la prolonger de manière unique en une application $\tilde{G}: \mathcal{C}(X)^* \rightarrow \mathbf{R}$ continue et vérifiant $\tilde{G}(1) = 1$ et pour toute fonction $f \geq 0$, $\tilde{G}(f) \geq 0$. Par le théorème de Riesz, il existe une unique probabilité μ telle que $\tilde{G} = \alpha_\mu$, de plus $\mu_n \rightarrow \mu$, lorsque $n \rightarrow \infty$. Finalement, $\mathcal{M}(X)$ est compact. \square

Proposition V.2.3. *L'ensemble $\mathcal{M}_T(X)$ est un fermé non vide de $\mathcal{M}(X)$.*

Preuve. L'application $G: \mu \in (X) \mapsto T_*\mu - \mu$ est continue car si $(\mu_n)_n$ est une suite d'éléments de $\mathcal{M}(X)$ convergeant vers μ , alors par le théorème de transfert,

$$\int_X f dT_*\mu_n - \int_X f d\mu_n = \int_X (f \circ T - f) d\mu_n \rightarrow \int_X f \circ T - f d\mu = \int_X f dT_*\mu - \int_X f d\mu,$$

lorsque $n \rightarrow \infty$ et pour tout $f \in \mathcal{C}(X)$. Ainsi G est continue et donc $\mathcal{M}_T(X)$ est un fermé.

Soit $\mu \in \mathcal{M}(X)$, on considère $\mu_n = \frac{1}{n}(\sum_{i=0}^{n-1} (T^i)_*\mu)$ et une extractrice φ telle que $(\mu_{\varphi(n)})_n$ converge (possible car $\mathcal{M}(X)$ est compact) vers μ' . Alors $\mu' \in \mathcal{M}_T(X)$, car $T_*\mu_n - \mu_n = \frac{1}{n}(T_*^n\mu - \mu) \xrightarrow{n \rightarrow \infty} 0$. (car pour toute fonction f continue sur X , f est bornée). \square

Proposition V.2.4. *Soit $\mu \in \mathcal{M}(X)$ une mesure de probabilité borélienne. Alors μ est régulière :*

$$\begin{aligned} \forall E \in \mathcal{B}(X), \mu(E) &= \inf \{ \mu(O) \mid O \text{ ouvert}, E \subseteq O \} \\ &= \sup \{ \mu(F) \mid F \text{ fermé}, F \subseteq E \}. \end{aligned}$$

C'est à dire que pour tout borélien $E \in \mathcal{B}(X)$ et $\varepsilon > 0$, il existe un ouvert O contenant E et un fermé F inclus dans E tels que

$$\mu(E) - \varepsilon \leq \mu(F) \leq \mu(E) \leq \mu(O) \leq \mu(E) + \varepsilon. \quad (*)$$

Preuve. Pour ce faire, on considère l'ensemble

$$\tau = \{ E \in \mathcal{B}(X) \mid \forall \varepsilon > 0, \exists O \text{ ouvert}, \exists F \text{ fermé}, F \subseteq E \subseteq O \text{ et } (*) \},$$

et on montre qu'il s'agit d'une tribu, qui contient les fermés.

Soit $E \in \mathcal{B}(X)$, et $\varepsilon > 0$. On prend $F \subseteq E \subseteq O$ avec F fermé et O ouvert vérifiant (*). Alors en posant $F' = O^c$ qui est fermé et $O' = F^c$ qui est ouvert de sorte que $F' \subseteq E^c \subseteq O'$ et on a toujours (*). Donc $E^c \in \tau$.

Pour la stabilité par union dénombrable, on considère $(E_n)_{n \geq 1}$ une suite de boréliens et

$$E = \bigcup_{n \geq 1} E_n = \bigcup_{n \geq 1} \left(\bigcup_{k \leq n} E_k \right),$$

où la suite $(\bigcup_{k \leq n} E_k)_n$ est croissante, et donc $\lim_{n \rightarrow \infty} \mu(\bigcup_{k \leq n} E_k) = \mu(E)$. Soit $\varepsilon > 0$, alors il existe $N \in \mathbb{N}$ tel que

$$0 \leq \mu(E) - \mu\left(\bigcup_{k \leq N} E_k\right) \leq \frac{\varepsilon}{2}.$$

Pour chacun des E_k pour $k \in \llbracket 1, N \rrbracket$, il existe F_k un fermé inclus dans E_k et tel que $\mu(E_k) - \mu(F_k) \leq \frac{\varepsilon}{2N}$. Or, on a l'inclusion

$$\left(\bigcup_{k \leq N} E_k \right) \setminus \left(\bigcup_{k \leq N} F_k \right) \subseteq \bigcup_{k \leq N} (E_k \setminus F_k),$$

et donc

$$\mu\left(\bigcup_{k \leq N} E_k\right) - \mu\left(\bigcup_{k \leq N} F_k\right) \leq \sum_{k \leq N} \mu(E_k) - \mu(F_k) \leq \frac{\varepsilon}{2}.$$

On pose alors $F = \bigcup_{k \leq N} F_k$ qui est un fermé inclus dans E et

$$0 \leq \mu(E) - \mu(F) = \mu(E) - \mu\left(\bigcup_{k \leq N} E_k\right) + \mu\left(\bigcup_{k \leq N} E_k\right) - \mu(F) \leq \varepsilon.$$

On procède de la même manière pour construire un ouvert O contenant E de tel sorte qu'on ait l'inégalité (*). Ainsi $E \in \tau$, et finalement τ est une tribu.

Il reste enfin à montrer que τ contient les fermés. Soit F un fermé, alors en posant $O_n = \{x \in X \mid d(x, F) < \frac{1}{n}\}$ qui est ouvert par continuité de d et car F est fermé. De plus, la suite $(O_n)_{n \geq 1}$ est décroissante et $F = \bigcap_{n \geq 1} O_n$, d'où

$$\lim_{n \rightarrow \infty} \mu\left(\bigcap_{k \leq n} O_k\right) = \mu(F),$$

ce qui permet de conclure que $F \in \tau$.

Finalement, τ est une tribu qui contient les ouverts, donc $\tau = \mathcal{B}(X)$, et on a bien la propriété voulue. \square

BIBLIOGRAPHIE

- [Bow75] Rufus BOWEN. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*. Lecture Notes in Computer Science. Springer-Verlag, 1975. ISBN : 9780387071879.
- [BP03] Yves BENOIST et Frédéric PAULIN. *Systèmes dynamiques élémentaires*. 2003. URL : https://www.imo.universite-paris-saclay.fr/~frederic.paulin/notescours/cours_sysdyn.pdf.
- [BS91] Eric BEDFORD et John SMILLIE. « Polynomial diffeomorphisms of \mathbf{C}^2 : currents, equilibrium measure and hyperbolicity. » In : *Inventiones mathematicae* 103.1 (1991), p. 69-100. URL : <http://eudml.org/doc/143853>.
- [FS92] John Erik FORNÆSS et Nessim SIBONY. « Complex Hénon mappings on \mathbf{C}^2 and Fatou-Bieberbach domains ». In : *Duke Math Journal* (1992).
- [KH95] Anatole KATOK et Boris HASSELBLATT. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.
- [Nij17] Nadia NIJIMBERE. *Markov partition and symbolic dynamics of hyperbolic toral automorphisms*. 2017. URL : https://fse.studenttheses.ub.rug.nl/15162/1/Nadia_Nijimbere__2017_WM.pdf.