

Vers une compréhension du principe de maximisation de production d'entropie

Martin Mihelich

▶ **To cite this version:**

Martin Mihelich. Vers une compréhension du principe de maximisation de production d'entropie. Physique mathématique [math-ph]. Université Paris Saclay (COmUE), 2015. Français. NNT : 2015SACLS038 . tel-01308677

HAL Id: tel-01308677

https://theses.hal.science/tel-01308677

Submitted on 28 Apr 2016

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

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

NNT : 2015SACLS038

THESE DE DOCTORAT DE L'UNIVERSITE PARIS-SACLAY PREPAREE A L'Université Paris-Sud

ECOLE DOCTORALE N° 564 « Physique de l'Ile-de-France »

En Physique

Par

Mr Martin Mihelich

Vers une compréhension du principe de maximisation de production d'entropie.

Thèse présentée et soutenue à Saclay, le 26/10/2015

Composition du Jury :

Mr Trizac Emmanuel	Professeur (LPTMS)	Président
Mr Bertin Eric	Chercheur CNRS (IPHT)	Rapporteur
Mr Lucarini Valerio	Professeur (CEN,Hamburg)	Rapporteur
Mr Luck Jean-Marc	Chercheur CEA (IPHT)	Examinateur
Mme, Dubrulle, Bérengère	Directrice de recherche CNRS	Directeur de thèse
Mr Paillard Didier	Chercheur CEA (LSCE)	Co-directeur de thèse

A mon père,

1 Remerciements

1.1 Remerciements

Cette tentative de compréhension du principe de maximisation de production d'entropie n'a été possible qu'avec un maximum de bienveillance de la part de ma directrice de thèse Bérengère Dubrulle. Je ne saurais trop comment l'en remercier assez et me souviendrai toujours des multiples discussions autour d'une tasse d'eau chaude. Je remercie aussi bien évidemment Didier Paillard pour nos discussions fructueuses sur ce tableau blanc où les jours où nous trouvions du premier coup un feutre qui marchait était probablement de mesure nulle...Comment ne pas penser à Davide Faranda, chercheur talentueux toujours plein d'optimisme et d'idées inventives ! Merci aussi à Denis qui a partagé mon bureau (et mes histoires). Une pensée à Corentin Herbert qui m'a beaucoup aidé notamment au début de ma thèse, sans oublier les discussions en salle café avec Hervé, Vincent, François D, Cécile, François L et tout le SPHYNX évidemment !

Je remercie aussi Eric Bertin et Valerio Lucarini qui ont accepté d'être mes rapporteurs de thèse. Mais aussi Emmanuel Trizac, Françoise Cornu ainsi que Jean-Marc Luck qui ont eu la gentillesse d'être membre de mon jury.

Une pensée va tout naturellement à mes anciens professeurs de prépa qui ont su me transmettre leur passion pour les sciences.

Réfléchir, écrire ma thèse n'a également été possible qu'avec l'aide précieuse de mes amis : Quentin et nos appels quasi quotidiens, Manu et nos soirées, Maxime et nos souvenirs, Steve et ses défaites tennistiques, Ludo et nos appels Skype manqués, Cyril et ses cafés sans oublier Geff, Louis et Marion, Isabelle aussi et surtout, ainsi que mes amis du tennis Warren et Franck, Fabrice, Fahrid...

Merci enfin à ma famille de France et du Canada : mes grands parents, mes cousins, mes cousines, mes oncles et tantes qui m'ont toujours supporté et su m'aider aux bons moments.

Enfin, je ne remercierais jamais assez mon frère et ma mère sans oublier mon père.

Résumé

Dans cette thèse nous essayons de comprendre pourquoi le **Principe de Maximisation de Production d'Entropie** (MEP) donne de très bons résultats dans de nombreux domaines de la physique hors équilibre et notamment en climatologie. Pour ce faire nous étudions ce principe sur des systèmes jouets de la physique statistique qui reproduisent les comportements des modèles climatiques. Nous avons notamment travaillé sur l'Asymmetric Simple Exclusion Process (ASEP) et le Zero Range Process (ZRP). Ceci nous a permis tout d'abord de relier MEP à un autre principe qui est le principe de maximisation d'entropie de Kolmogorov-Sinai (MKS). De plus, l'application de MEP à ces systèmes jouets donne des résultats physiquement cohérents. Nous avons ensuite voulu étendre le lien entre MEP et MKS dans des systèmes dynamiques plus compliqués avant de montrer que, pour les chaines de Markov, maximiser l'entropie de KS revenait à minimiser le temps que le système prend pour atteindre son état stationnaire (mixing time). Enfin nous avons appliqué MEP à la convection atmosphérique.

Mots clefs : **Entropie, Physique statistique, chaîne de Markov, théorie des graphes.**

Abstract

In this thesis we try to understand why the maximum entropy production principle gives really good results in a wide range of Physics fields and notably in climatology. thus we study this principle on classical toy models which mimic the behaviour of climat models. In particular we worked on the Asymmetric Simple Exclusion Process (ASEP) and on the Zero Range Process (ZRP). This enabled us first to connect MEP to an other principle which is the maximum Kolmogorov-Sinaï entropy principle (MKS). Moreover the application of MEP on these systems gives results that are physically coherent. We then wanted to extend this link between MEP and MKS in more complicated systems, before showing that, for Markov Chains, maximise the KS entropy is the same as minimise the time the system takes to reach its stationnary state (mixing time). Thus, we applied MEP to the atmospheric convection.

Keywords : Entropy, statistical Physics, Markov chain, graphs theory.

Table des matières

1 Remerciements . . . iii
1.1 Remerciements . . . iii
Résumé . . . v
Abstract . . . vi
Table des matières . . . vii
2 Introduction générale . . . 1

1 Principe de maximisation et de minimisation de la production d'entropie. 5
1 La Physique et ses Principes. . . . 5
1.1 Différents Principes. . . . 5
1.2 Le Principe de Maximisation de Production d'entropie (MEP). 5
2 MEP en climatologie. . . . 8
2.1 Le modèle à deux boites. . . . 8
2.2 Le modèle de Paltridge . . . 9
2.3 Vers une amélioration du modèle de Paltridge. . . . 11
3 Prigogine vs MEP : Maximisation ou Minimisation de la production d'entropie ? . . . 12
3.1 Minimisation de la production d'entropie. . . . 12
4 Le principe de maximisation de l'entropie de KS. . . . 15
4.1 definition formelle de l'entropie de Kolmogorov-Sinaï . . . 15
4.2 Entropie de KS pour un décalage de Bernouilli : . . . 17
4.3 Chaine de Markov et entropie de KS. . . . 17
4.4 Production d'entropie dynamique . . . 18
4.5 Principe de Maximisation de l'entropie de KS . . . 19

2 Principe de maximisation de production d'entropie et modèle ASEP. 21
1 Chaines de Markov. . . . 21
1.1 Définition et outils associés. . . . 21
1.2 Convergence des chaines de Markov. . . . 22
1.3 Production d'entropie microscopique . . . 23
2 Une chaine de Markov particulière : le modèle SSEP et ASEP. . . . 24
2.1 Motivations et lien avec le climat. . . . 24
2.2 Explication des modèles SSEP et ASEP. . . . 24
2.3 Modèle CASEP et Langmuir-ASEP. . . . 25
3 Article. . . . 26
3.1 Discussion sur les résultats trouvés. . . . 37

3 Principe de maximisation de production d'entropie et Zero Range Process. 39
1 Introduction . . . 39
2 ZRP . . . 39
2.1 Description . . . 39
2.2 Les différentes applications du ZRP. . . . 40
3 ZRP et calculs analytiques . . . 41
3.1 Formalisme quantique . . . 41
3.2 Calcul de l'expression de la fugacité, du courant de particules... 43
3.3 Quelques cas particuliers . . . 44
3.4 Production d'entropie et entropie de KS . . . 44
4 Article . . . 47
5 Principaux résultats . . . 58
5.1 travaux futurs. . . . 58

4 Maximisation de la production d'entropie et coarse graining. 59
1 Introduction. . . . 59
2 Introduction aux systèmes dynamiques. . . . 59
2.1 Systèmes dynamiques : Définition et quelques exemples. . . . 60
2.2 Chaos, Exposants de Lyapunov et théorème de Pesin. . . . 61
3 Lien entre production d'entropie et entropie de KS. . . . 62
3.1 Entropie coarse grainée. . . . 62
3.2 Existe-t-il un véritable lien dans les systèmes dynamiques entre la production d'entropie et l'entropie de KS ? . . . 65
3.3 Limitations dues au coarse graining. . . . 65
4 Une nouvelle méthode pour coarse grainer. . . . 66
4.1 Méthodes classiques de coarse graining. . . . 66
4.2 Nouvelle méthode : La convolution. . . . 67
5 Production d'entropie microscopique. . . . 68
5.1 Expression analytique de l'entropie microscopique. . . . 69
5.2 $dS/dt<0$? . . . 70
6 Coarse grained par rapport au temps ou à l'espace. . . . 70
6.1 Un autre coarse graining. . . . 70
6.2 Coarse grained par rapport au temps. . . . 71
6.3 Coarse grained par rapport à l'espace. . . . 71
7 Conclusion. . . . 74

5 Maximisation de l'entropie de KS et minimisation du mixing time. 75
1 Introduction. . . . 75
2 Chaine de Markov et temps de mélange : . . . 76
2.1 Définition du temps de mélange. . . . 76
2.2 Lien avec les valeurs propres. . . . 76
2.3 Quelques inégalités sur le temps de mélange. . . . 77
3 A la recherche de la dynamique qui minimise le temps de mélange. . . 78
3.1 Quelques exemples élémentaires de chaines qui mélangent rapidement . . . 79
3.2 Algorithmes qui recherchent la dynamique qui mélange le plus rapidement. . . . 79

4 A la recherche de la dynamique qui maximise l'EKS. 81
4.1 Théorème de Parry 81
4.2 Méthode analytique pour trouver la dynamique qui maximise l'EKS. 81
5 Lien entre Minimisation du temps de mélange et maximisation de l'EKS. 82
5.1 Article 83
6 Pour aller plus loin : Maximisation de l'EKS et Minimisation du temps de mélange dans les systèmes dynamiques. 88
6.1 Entropie topologique et entropie métrique. 88
6.2 Volume growth (L.S Young) 89

6 Application de MEP à la convection **91**
1 Introduction, position du problème. 91
2 Outils nécessaires. 91
2.1 Calcul sur un graphe 92
2.2 Résolution en Température. 93
2.3 Résolution en Flux. 94
2.4 Quelques cas particuliers. 94
3 Méthode KKT pour la résolution. 95
3.1 Température potentielle équivalente. 95
3.2 Problème explicite une colonne. 96
3.3 Problème explicite 2 colonnes et n colonnes. 98
3.4 Principe permettant de trouver une unique solution parmi toutes les solutions du domaine admissible. 98
4 Résultats. 99
4.1 Modèle 1 colonne. 99
4.2 Modèle 2 colonnes. 100
5 Discussion 100
6 Conclusion 101
7 Conclusion générale 102
7.1 Conclusion 102
7.2 Perspectives et travaux futurs. 103
1 Annexes 104

Bibliographie **113**

Bibliographie **113**

2 Introduction générale

Un **Principe physique** est une loi physique (souvent non démontrée) qui permet de décrire et de comprendre un ensemble de phénomènes. L'avantage des Principes réside dans le caractère général qui permet de ne plus étudier les systèmes au cas par cas. C'est pourquoi les physiciens ont cherché et cherchent à établir des **principes physique** comme le principe d'inertie en mécanique classique, le principe de moindre action qui peut être appliqué dans de nombreux domaines comme en électromagnétisme ou en relativité générale, le principe d'exclusion de Pauli en physique quantique...

De plus, bien que la physique statistique à l'équilibre soit relativement bien comprise aujourd'hui, il en est tout autrement de la physique hors équilibre où bien souvent chaque phénomène doit être étudié séparément. Ainsi, une manière de décrire la physique hors équilibre serait de trouver des principes et de nombreux chercheurs s'y attèlent. Parmi ces principes il en existe un en particulier (il sera l'objet central de cette thèse) qui est appelé le **Principe de Maximisation de Production d'entropie (MEP)**.

Le principe de maximisation de production d'entropie a été introduit par Filyukov et Karpov en 1967 [1]. Dans cet article ils expliquent comment le principe de maximisation de production d'entropie permet de décrire les processus stationnaires. Ensuite, en 1975 ce principe a été utilisé pour la première fois en climatologie par Paltridge [2] pour décrire la température ainsi que les flux de chaleur terrestre, donnant des résultats surprenamment proches de la réalité. Il est aussi utilisé en océanographie [3] pour décrire la circulation générale des océans et les échanges de flux de chaleurs dans les océans. De façon plus surprenante, Zupanic et al [4] ont déduit de MEP les lois de Kirchoff ou encore Ivanchenko [5] a analysé le comportement électrodynamique d'un réseau périodique de jonctions Josephson en utilisant MEP pour trouver la valeur du paramètre de synchronisation. On pourrait encore citer de nombreuses applications de MEP en chimie, biologie, physique des solides...

MEP stipule qu'un système hors équilibre pouvant admettre plusieurs états stationnaire possibles admet comme état final celui qui maximise la production d'entropie. Le but de cette thèse est d'essayer de donner une meilleure compréhension de MEP, c'est à dire essayer de comprendre son domaine de validité, de le lier à d'autres principes et de comprendre son sens physique. La difficulté majeure est que ce principe est utilisé dans des domaines extrêmement variés. Qui plus est, MEP s'applique souvent pour des systèmes extrêmement complexes avec un très grand nombre de degrés de liberté comme le climat, l'océanographie...Il est par conséquent difficile dans des systèmes physique aussi compliqués de comprendre pourquoi MEP donne de bons résultats. Nous avons donc opté pour une nouvelle approche qui consiste à étudier MEP sur des systèmes relativement simples. Plus exactement, nous avons regardé l'application de MEP sur des modèles jouets de la physique statistique qui reproduisent les comportements de systèmes beaucoup plus complexes.

Cette nouvelle approche nous a permis de lier MEP au principe de maximisation d'entropie de Kolmogorov-Sinai, tout d'abord de façon numérique dans le modèle

ASEP (Asymmetric Simple Exclusion Process) puis de façon analytique dans le ZRP (Zero Range Process). Nous avons aussi démontré que les résultats donnés par MEP sur ces systèmes étaient physiquement cohérent dans le sens ou proche de l'équilibre l'état choisit par MEP est diffusif alors que plus le système est mis hors équilibre plus l'état devient convectif. De plus cette analyse nous a permis de définir une taille de coarse graining optimale. Enfin nous avons montré que maximiser l'entropie de Kolmogorov-Sinai correspondait à minimiser le temps que le système met pour atteindre son état stationnaire, cette connexion nous permet finalement de comprendre plus clairement MEP.

De manière plus précise nous avons divisé le manuscrit en 6 chapitres :

1. Dans le chapitre 1 nous commençons par expliquer le principe de maximisation de production d'entropie et notamment comment il est appliqué en climatologie avec le modèle en boite utilisé par Paltridge. Nous parlons ensuite du principe de minimisation de production d'entropie (Prigogine) et nous donnons une nouvelle vision qui permet de rallier ces deux principes à priori contradictoires. Enfin nous introduisons l'entropie de Kolmogorov-Sinai (KS) et donnons son expression analytique dans certains cas particuliers.

2. Dans le chapitre 2 nous commençons par introduire les quantités utiles et rappeler les théorèmes fondamentaux sur les chaines de Markov. Nous décrivons ensuite le modèle ASEP qui est un modèle jouet imitant le modèle climatique en boites de Paltridge. Nous expliquons ensuite comment nous avons appliqué MEP au modèle ASEP. Puis nous commentons les résultats obtenus notamment le lien numérique entre MEP et la maximisation de l'entropie de KS ainsi que le fait que plus le système est mis hors équilibre plus l'état choisit pas MEP est convectif.

3. Dans le chapitre 3 nous nous intéressons au Zero Range Process (ZRP). Après avoir décrit le ZRP nous calculons des quantités caractéristiques (fugacité, courant de particules...) qui nous permettent de donner une expression analytique de la production d'entropie et de l'entropie de KS. Nous établissons ensuite analytiquement le lien entre MEP et la maximisation de l'entropie de KS. Ceci nous permet de définir un coarse graining optimal.

4. Dans le chapitre 4 nous essayons de voir si cette connexion entre ces deux principes peut être étendue à des systèmes dynamiques plus généraux. Nous remarquons qu'en utilisant la méthode naturelle pour coarse grainer l'espace des phases une telle connection existe mais qu'avec cette méthode il existe des limitations. Nous introduisons ensuite une nouvelle manière de coarse grainer qui donne des résultats plus satisfaisants. Enfin nous nous intéressons à l'expression de la production d'entropie microscopique en fonction des exposants de Lyapunov ainsi qu'au comportement de ces derniers en fonction du coarse graining.

5. Dans le chapitre 5 nous montrons que maximiser l'entropie de KS revient à minimiser le temps de mélange (qui est le temps que le système prend pour

atteindre son état stationnaire dans les systèmes Markoviens). Nous commençons par définir proprement le temps de mélange puis nous décrivons des algorithmes qui recherchent la dynamique qui minimise le temps de mélange. Nous expliquons ensuite comment on peut trouver une expression analytique de la dynamique qui maximise l'entropie de KS (théorème de Parry) et nous comparons ces deux dynamiques. Nous étudions aussi ce lien sur des graphes diffusifs. Enfin nous étudions la possibilité d'étendre cette relation aux systèmes dynamiques plus généraux en introduisant notamment la notion de volume growth.

6. Le chapitre 6 est un chapitre plus appliqué que les précédents dans lequel nous appliquons MEP à la convection atmosphérique. Après quelques rappels sur les graphes nous montrons que la question revient à résoudre un problème d'optimisation avec contraintes d'égalités et d'inégalités (KKT). Nous résolvons ce problème et discutons les résultats trouvés qui sont caractérisés notamment par la présence d'updraft (mouvement d'air ascendant) et de downdraft (mouvement d'air descendant).

Chapitre 1

Principe de maximisation et de minimisation de la production d'entropie.

1 La Physique et ses Principes.

1.1 Différents Principes.

Un Principe physique est une loi physique qui ne peut ou qui n'a pas encore été démontrée mais qui est en accord avec l'expérience. Les Principes sont en quelque sorte les axiomes de la physique sur lesquelles elle repose. Il n'y a donc pas de physique sans Principe. C'est ainsi que la mécanique analytique classique trouve sa source dans le principe de d'Alembert [6] puis plus généralement dans le principe de moindre action largement développé par Lagrange. La thermodynamique prend son origine dans les trois principes de la thermodynamique développés entre autre par Boltzmann, Carnot et Ernst. La relativité quant à elle s'appuie sur le principe de la constante de la lumière. Le principe d'exclusion de Pauli ou encore une fois le principe de moindre action sont à la base de la physique quantique.

Les Principes physiques sont donc primordiaux et les grands scientifiques laissent d'ailleurs souvent leurs noms à un Principe. Cependant un Principe physique n'est pas universel. Il donne des résultats justes sous certaines conditions et moins exacts sous d'autres (par exemple le principe de conservation de la masse valide en mécanique classique mais non valide en relativité restreinte). Le comprendre passe donc par la connaissance de ses domaines d'application et de ses limitations.

1.2 Le Principe de Maximisation de Production d'entropie (MEP).

Origine de MEP

La physique statistique à l'équilibre, largement étudiée par Kelvin, Maxwell, Gibbs et Boltzmann, est bien décrite. La physique statistique hors équilibre est quant à elle très mal connue et un des principaux enjeux du 21 ème siècle est sa compréhension. La principale difficulté résulte dans le fait qu'il existe de multiples

façons d'être hors équilibre et chaque situation doit aujourd'hui être étudiée au cas par cas. Pour donner une certaine généralité aux phénomènes hors équilibres et essayer de les décrire de nombreux chercheurs ont cherché des Principes. Par exemple Evans [7] décrit les états stationnaires hors équilibre en maximisant l'entropie de Jaynes (entropie sur les chemins) ou Monthus [8] maximise la dérivée de l'entropie de Jaynes par rapport au temps pour décrire les systèmes hors équilibres. Malkus maximise quand à lui le transport de chaleur...

Un autre principe qui permet de décrire les systèmes hors équilibres dans un état stationnaire est le **Principe de Maximisation de production d'entropie**. Ce principe est l'objet central de cette thèse. Bien que controversé ce Principe est largement utilisé dans de nombreux domaines de la physique comme en physique des solides [9], en électromagnétisme [10] ou encore en physique quantique. Ce principe est aussi utilisé en chimie ainsi qu'en biologie [11] et en climatologie [2]. Ce Principe stipule qu'un système hors équilibre dans un état stationnaire choisi son état final de telle sorte à maximiser la production d'entropie. On trouve aussi une énonciation non stationnaire (transitoire) du principe qui dit que la dynamique choisit par le système est celle qui maximise la production d'entropie à tout instant. Il est essentiel de bien le distinguer du second principe de la thermodynamique qui s'intéresse à la maximisation de l'entropie et non la production d'entropie ! Afin de mieux le comprendre commençons par définir la production d'entropie et donnons une formulation la plus générale de ce principe.

Formulation de MEP

Considérons un système Σ pouvant ne pas être à l'équilibre. La variation de la densité d'entropie à l'intérieur du système s prend la forme [12, 11] :

$$\frac{\partial s}{\partial t} = \sigma - \operatorname{div} j_s, \tag{1.1}$$

où σ est la production de densité d'entropie et j_s est le flux de densité d'entropie. La production d'entropie totale du système σ_{tot} prend alors la forme naturelle :

$$\sigma_{tot} = \int_\Sigma \sigma. \tag{1.2}$$

De plus, σ peut être reliée aux densités de flux thermodynamiques J_i (qui peuvent être des flux de particules, des flux de chaleur...) et aux forces thermodynamiques X_i associées par la relation :

$$\sigma = \sum_i J_i X_i. \tag{1.3}$$

Voici la forme la plus générale de la production d'entropie dont l'écriture n'est pas remise en cause. En outre, c'est la formulation du Principe de Maximisation de production d'entropie (MEP) qui peut prendre plusieurs formes.

L'énonciation la plus générale de MEP est la suivante : **Un système hors équilibre choisit à tout instant ses flux thermodynamiques J_i et ses forces**

thermodynamiques X_i de telle sorte à maximiser sa production d'entropie totale σ_{tot} étant donnés les contraintes.

Il existe d'autres variantes qui sont en réalité des restrictions de l'énonciation générale précédente comme par exemple celle que fait Ziegler [13] qui dit que si les forces thermodynamiques sont fixées alors les flux thermodynamiques s'organisent afin de maximiser la production d'entropie. Ou encore que seuls les flux et les forces dues à la dissipation turbulente doivent être prise en compte [14], ou encore que les flux extérieurs doivent être fixés ou enfin que ceci n'est valide que pour les systèmes hors équilibres dans un état stationnaire...

Si le système est isolé alors l'équation (1.1) intégrée sur Σ devient $\frac{\partial s_{tot}}{\partial t} = \sigma_{tot}$ et donc le système choisit le chemin qui maximise la dérivée de s_{tot} par rapport au temps. Ainsi le système choisit le chemin le plus rapide en terme de s_{tot}. C'est à dire qu' il choisit le chemin qui maximise la vitesse par rapport à la distance (ou divergence) de Kullback-Leibler de la densité de probabilité du système à l'instant t notée $P(x,t)$ par rapport à la densité de probabilité finale $P_f(x)$. On rappelle que la distance de Kullback-Leibler est définie par :

$$D_{KL}(P,Q) = \int_{x \in Etats} P(x) \log(\frac{P(x)}{Q(x)}). \tag{1.4}$$

Schématiquement ceci est très bien représenté dans [11] :

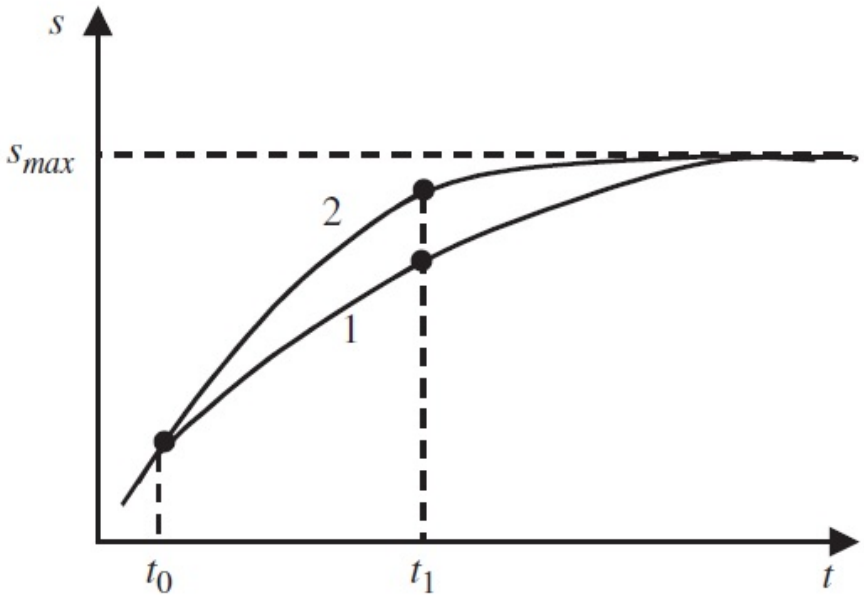


FIGURE 1.1 – Deux trajectoires possibles pour Σ. MEP prétend que la trajectoire privilégiée sera la numéro 2. Cette figure est extraite de [11].

Ainsi, à la différence du second principe qui ne parle que de l'état final, MEP est lié à la dynamique. Nous reviendrons sur ce point plus tard lorsque nous introduirons l'entropie de Kolmogorov-Sinaï.

2 MEP en climatologie.

MEP est utilisé en climatologie comme en témoignent les travaux de Paltridge, Herbert, Paillard... [2, 15]. Afin de comprendre comment MEP est utilisé commençons par le modèle le plus simple : le modèle à deux boites.

2.1 Le modèle à deux boites.

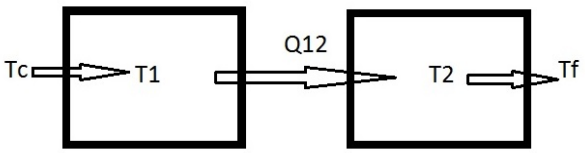


FIGURE 1.2 – Shéma du modèle à deux boites utilisé en climatologie.

Considérons deux boites 1 et 2 reliées à un réservoir de température chaude maintenu à la température T_c et un réservoir de température froide maintenu à la température T_f comme indiqué sur la figure Fig.(1.2). Notons $Q_{c\to 1}$ les flux de chaleur entre le réservoir T_c et la boite 1 et de manière analogue notons $Q_{1\to 2}$ et $Q_{2\to f}$ les flux de chaleur entre les boites 1 et 2 ainsi que la boite 2 et le réservoir à la température T_f. Nous supposons que les températures des réservoirs sont fixés mais que les températures des boites T_1 et T_2 ainsi que le flux $Q_{1\to 2}$ sont des inconnues. Nous supposons enfin que le système est dans un état stationnaire c'est à dire que ces grandeurs ne dépendent pas du temps. En utilisant 1.3, sachant que la force thermodynamique associée à un flux de chaleur dû à un gradient de température vaut grad$(\frac{1}{T})$, on trouve que la production d'entropie interne prend la forme :

$$\sigma(Q_{1\to 2}, T_1, T_2) = Q_{1\to 2}(\frac{1}{T_2} - \frac{1}{T_1}). \tag{1.5}$$

De plus, supposons que $Q_{c\to 1}$ soit une fonction de T_c et T_1 : $Q_{c\to 1} = G_1(T_c, T_1) = G_1(T_1)$ car T_c est fixée et de même $Q_{2\to f} = G_2(T_2, T_f) = G_2(T_2)$. Enfin, supposons que G_1 et G_2 soient inversibles.

Comme le système est supposé dans un état stationnaire nous obtenons :

$$Q_{c\to 1} = Q_{1\to 2} = Q_{2\to f}, \tag{1.6}$$

On a alors :

$$\sigma(Q_{1\to 2}) = Q_{1\to 2}(\frac{1}{G_1^{-1}(Q_{1\to 2})} - \frac{1}{G_2^{-1}(Q_{1\to 2})}). \tag{1.7}$$

σ n'est donc plus qu'une fonction de $Q_{1\to 2}$. De plus $Q_{1\to 2}$ caractérise l'état final du système. Les différentes valeurs prises par $Q_{1\to 2}$ correspondent aux différents états stationnaires pouvant être pris par le système. Sans condition supplémentaire toutes les valeurs prises par $Q_{1\to 2}$ peuvent être envisagées. MEP va nous permettre de déterminer une valeur précise de $Q_{1\to 2}$: **le flux choisit par le système maximise** σ. Prenons un exemple concret en se fixant G_1 et G_2. De manière naturelle nous pouvons écrire :

$$\begin{cases} G_1(T_c, T_1) = k_1(T_c - T_1). \\ G_2(T_2, T_f) = k_2(T_2 - T_f). \end{cases} \tag{1.8}$$

En utilisant cette dernière relation et l'équation eq.(1.7) nous obtenons :

$$\sigma(Q_{1\to2}) = Q_{1\to2}\left(\frac{1}{T_f + \frac{Q_{1\to2}}{k2}} - \frac{1}{T_c - \frac{Q_{1\to2}}{k1}}\right). \tag{1.9}$$

Lorsque $Q_{1\to2} = 0$ alors $\sigma(Q_{1\to2}) = 0$ et lorsque $Q_{1\to2}$ est assez grand $T_1 = T_2$ et donc $\sigma(Q_{1\to2}) = 0$. Schématiquement σ a l'allure suivante Fig(1.3) :

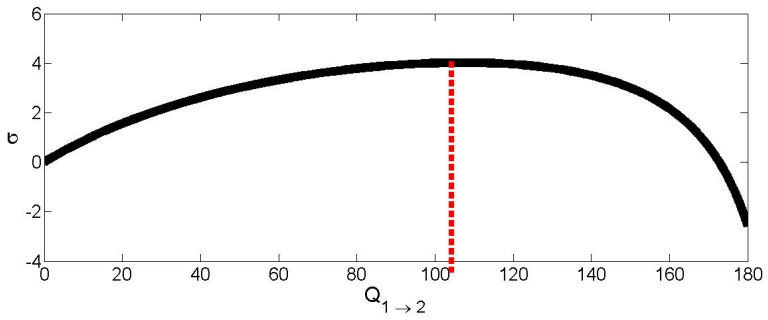


FIGURE 1.3 – σ fonction de $Q_{1\to2}$ pour $k_1 = 1$ $k_2 = 10$ $T_c = 200$ et $T_f = 10$. On voit que σ admet un maximum en un certain $Q^*_{1\to2}$.

σ admet donc un maximum en un certain $Q^*_{1\to2}$. MEP affirme que le système s'organise pour que le flux dans l'état stationnaire soit égal à $Q^*_{1\to2}$.

2.2 Le modèle de Paltridge

L'équation générale de convection diffusion de la température est donnée par :

$$\frac{\partial T}{\partial t} + \mathbf{v}.\nabla T = \kappa \Delta T + Q, \tag{1.10}$$

où T est la température, $\mathbf{v}$ représente la vitesse du fluide, κ la diffusivité moléculaire et Q le terme dit de chauffage diabatique qui comprend le bilan radiatif ainsi que la chaleur latente pour une atmosphère humide. Cette équation est souvent impossible à résoudre même numériquement car elle demande la connaissance de la vitesse du fluide en tout point.

On l'approxime alors par l'équation :

$$\frac{\partial T}{\partial t} = Q(T) + \gamma \tag{1.11}$$

où γ représente la convergence du flux d'énergie dû à tous les processus autres que le rayonnement. Les états stationnaires sont caractérisés par le fait que T ne dépende plus du temps et donc :

$$Q(T) + \gamma = 0 \tag{1.12}$$

De plus il existe une contrainte liée à la conservation de l'énergie :

$$\int_V \gamma dV = 0. \tag{1.13}$$

Cependant, il existe une infinité de distributions de température T qui vérifient les équations eq.(1.12) et eq.(1.15). L'idée de Paltridge est de choisir celle qui maximise la production d'entropie σ définie par :

$$\sigma = \int_V \frac{\gamma}{T} dV. \tag{1.14}$$

Paltridge est donc amené à résoudre le problème variationnel suivant :

$$\max_T \{\sigma(T)\,, \int_V Q(T)dV = 0\}. \tag{1.15}$$

Plus précisément, l'idée de Paltridge est de modéliser la surface terrestre (une zone latitudinaire 1D) du pôle nord vers le pôle sud par une chaine de 10 boites pouvant échanger des flux de chaleurs.

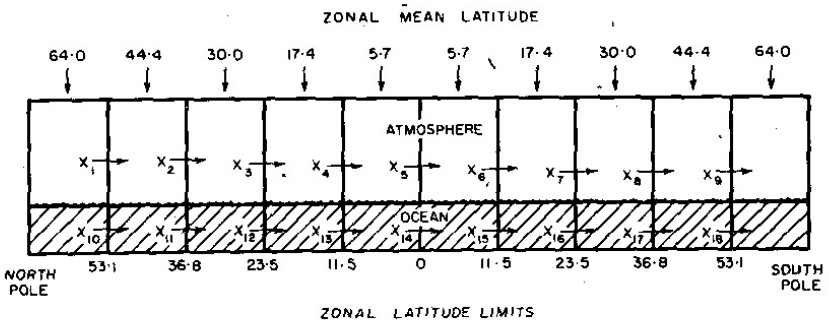


FIGURE 1.4 – Représentation du modèle de Paltridge constitué de 10 boites. Les boites peuvent échanger entre elles des flux de chaleurs représentés par les flèches. Cette figure est extraite de [2].

Plus précisément, comme le montre la figure Fig.(1.4) il existent des inconnus qui sont : les 18 flux de chaleur entre les boites, les températures dans les boites, la valeur des couvertures nuageuses, ainsi que les flux absorbés par le soleil et les flux radiatifs renvoyés par les boites. Néanmoins Paltridge fait des hypothèses qui lui permettent de réduire les inconnus aux 18 flux entre les boites. La connaissance de ces flux nous donne la valeur des autres variables. Pour trouver la valeur de ces flux, Paltridge utilise MEP.

En utilisant MEP Paltridge trouve alors les résultats observés sur la figure Fig.(1.5).

Les résultats de Paltridge sont donc proches de l'expérience. Néanmoins il existe dans l'article de Paltridge quelques points qui restent flous :

i) Dans sa modélisation Paltridge fixe les flux de surfaces et n'utilise MEP que pour déterminer les flux méridionaux alors que l'on aurait pu utiliser MEP pour déterminer les deux.

ii) les coefficients radiatifs sont supposés constants et déterminé de manière totalement empirique.

iii) L'utilisation de ce modèle pour les paléoclimats ou les climats d'autres planètes est difficile au vue de la paramétrisation assez contraignante.

La thèse de Corentin Herbert, mon prédécesseur qui fût aussi le doctorant de Bérengère Dubrulle et Didier Paillard a permis d'améliorer ce modèle.

2.3 Vers une amélioration du modèle de Paltridge.

C.Herbert [15] a construit un modèle dans lequel les flux de surface et les flux méridionaux sont traités par MEP. De plus il n'utilise dans son modèle aucun paramètres ajustables ce qui permet de faire un modèle beaucoup plus robuste que celui de Paltridge. Il applique son modèle aux climats présents et aux climats glaciers et trouve des résultats en accord avec les simulations de l'IPSL. L'avantage incontestable de MEP est qu'il ne s'intéresse pas à résoudre la dynamique du problème. Ainsi MEP est moins couteux en calcul que les modèles classiques utilisés en climatologie. Pour avoir un ordre de grandeur MEP donne un profil de température en quelques secondes alors que les modèles de l'IPSL prennent des jours, voir parfois des mois.

C.Herbert a aussi utilisé MEP pour décrire les rétroactions climatiques et les

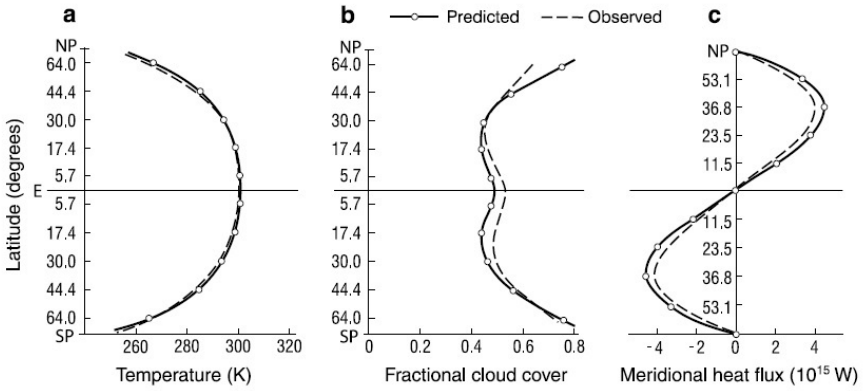


FIGURE 1.5 – Les résultats du modèle de Paltridge en traits pleins comparés aux observations. Cette figure est extraite de [2].

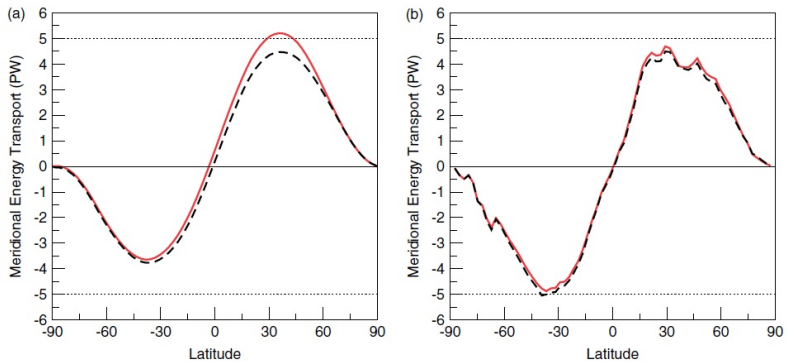


FIGURE 1.6 – Energie méridionale totale prédit par MEP (en traits pleins) et pour une simulation de l'IPSL (en pointillés). A gauche c'est une simulation avec les conditions du Last Glacial Maximum et à droite avec les conditions pre-industrielle. Cette figure est extraite de [15].

équilibres multiples en tenant compte du cycle saisonnier [16]

MEP est un principe empirique qui n'a pas de justifications rigoureuses. Dewar a essayé de donner une preuve de MEP [17, 18] qui s'avère être extrêmement intéressante mais pas parfaitement rigoureuse. De plus on connait assez mal les domaines de validité de MEP. Enfin, les chercheurs ne croient pas en ce principe car il existe un autre principe plus connu : le principe de minimisation de production d'entropie de Prigogine. Ces deux principes semblent contradictoires. Nous allons donner une explication nouvelle qui concilie MEP et le principe de minimisation de Prigogine.

3 Prigogine vs MEP : Maximisation ou Minimisation de la production d'entropie ?

La maximisation et la minimisation de la production d'entropie sont deux principes totalement différents qu'il faut absolument savoir distinguer pour pouvoir correctement les utiliser. De plus, ces deux principes semblent contradictoires car d'une part on minimise la production d'entropie et de l'autre on la maximise. Ceci est à la base d'une confusion dans la communauté scientifique. Martuyshev et Seleznev [11] ont donné une explication de la différence entre ces deux principes. Pour eux, lorsque les forces (flux) sont fixés mais que les flux (forces) associés sont libres alors MEP permet de les déterminer. En outre, lorsque le système est faiblement hors équilibre et dans un état stationnaire, et qu'il y a des forces libres alors ils vont s'ajuster (et ajuster les flux associés) afin de minimiser la production d'entropie. A temps court et forces thermodynamiques fixées le système **maximise** la production d'entropie alors que à temps long et à forces libres le système **minimise** sa production d'entropie.

Ma recherche m'a permise de donner une explication nouvelle quant à la différence entre ces deux principes. Elle permet aussi des les conciliers car comme nous allons le voir MEP apparait être un MaxMin.

Dans la section suivante nous rappelerons les hypothèses très strictes pour pouvoir appliquer le principe de minimisation de production d'entropie (appelé aussi principe de Prigogine) et nous expliquerons la différence avec le principe de maximisation de production d'entropie. Pour illustrer cette différence nous prendrons l'exemple d'un processus aléatoire Markovien.

3.1 Minimisation de la production d'entropie.

En 1945 Prigogine [19] explique son principe en prenant un système composé de deux récipients remplis d'un même fluide portés à deux températures différentes communicant par l'intermédiaire d'un petit orifice. Ce système est caractérisé par deux forces thermodynamiques X_m et X_{th} et deux flux J_m et J_{th} associés au transport de matière et au transport thermique. La production d'entropie s'écrit alors :

$$\sigma = J_m.X_m + J_{th}.X_{th}$$

On impose X_{th} constant, X_m peut varier, et l'on montre expérimentalement que très proche de l'équilibre le système admet un état stationnaire caractérisé par $J_m =$

0 et $J_{th} = Cst.X_{th}$. Théoriquement, on impose au système d'être suffisamment proche de l'équilibre pour que les relations d'Onsager entre les flux et les forces thermodynamiques soient vérifiées et que les coefficients de linéarité soient constants, c'est à dire :

$$\begin{cases} J_{th} = L_{11}.X_{th} + L_{12}.X_m \\ J_m = L_{21}.X_{th} + L_{22}X_m \end{cases}$$

avec $L_{12} = L_{21}$

Sous ces conditions très restrictives on a :

$$\sigma = L_{11}.X_{th}^2 + 2L_{12}X_{th}.X_m + L_{22}X_m^2$$

et donc :

$$\frac{\partial \sigma}{\partial X_m} = 0 \Leftrightarrow J_m = 0$$

Ainsi, σ étant une forme quadratique positive, l'état stationnaire minimise la production d'entropie par rapport à X_m. Et grâce aux relations d'Onsager l'état stationnaire minimise la production d'entropie. On remarquera que le principe de Prigogine n'est vrai que très proche de l'équilibre car lorsque l'écart de température entre les deux boites est suffisant il y a un transport de matière J_m non nul.

De même S.R. de Groot et P. Mazur [20] montent assez simplement que pour un système de conduction thermique (tout comme un système de diffusion de particules), vérifiant toujours les relations d'Onsager avec des coefficients constants, caractérisé par un flux J on a

$$\text{si } \sigma \text{ admet un minimum alors en ce minimum } \mathrm{div}(J) = 0$$

Ainsi, si σ est minimale alors on est dans un état stationnaire. Cependant **la réciproque est fausse** et est trop souvent employée sous le nom de principe de Prigogine ! C'est à dire que ce n'est pas parce que l'on est dans l'état stationnaire choisi par le système que la production d'entropie est minimale. En effet, s'il y a plusieurs états stationnaires possibles considérés, la réciproque ne peut-être déduite du sens direct. Par exemple dans le cas de la diffusion thermique on a $div(J) = 0$ mais il y a une infinité de J qui vérifient cette relation (c'est à dire une infinité d'états stationnaires possibles si aucune autre contrainte n'est spécifiée) et rien nous dit que l'état final choisi par le système sera celui qui minimisera la production d'entropie parmi les J possibles.

Nous pouvons penser de manière formelle à un système dont la production d'entropie en fonction de ses états à la forme représentée sur la figure Fig.(1.7).

Dans un tel cas, la minimisation de la production d'entropie nous donne potentiellement trois états stationnaires. Parmi ces trois états stationnaires nous ne pouvons donc savoir celui qui sera choisi en réalité par le système et nous verrons par la suite que ce ne sera pas l'état n°1 comme le prétendrait la réciproque du principe de Prigogine.

Un cas particulier où l'on peut dire que l'état stationnaire final sera celui qui minimisera la production d'entropie est lorsque la minimisation de celle-ci nous donne qu'un unique état stationnaire possible.

Pour résumer, le principe de Prigogine qui dit que l'état stationnaire choisi par le système est celui qui minimise la production d'entropie n'est vrai que si les conditions très restrictives suivantes sont vérifiées :

i) Le système se trouve suffisamment poche de l'équilibre pour que les relations d'Onsager soient vérifiées avec des coefficients de linéarité constants.

ii) La minimisation de la production d'entropie ne nous donne qu'un unique état stationnaire possible.

Ainsi ces hypothèses sont indispensables et trop d'articles critiquent le principe de Prigogine sans les avoir vérifiées [21].

Le principe de Maximisation d'entropie : quand l'hypothèse ii) n'est pas vérifiée.

Revenons tout d'abord sur le sens de l'hypothèse ii). Pour décrire un système hors équilibre dans un état stationnaire il est souvent nécessaire de considérer plusieurs états stationnaires possibles. Par exemple dans un modèle de convection thermique (ou de diffusion particulaire) d'un barreau 1D soumis à deux températures différentes T_A et T_B aux extrémités les différents états stationnaires sont donnés par $J = Cst$; il y en a ici une infinité. Il reste maintenant à savoir quel J va effectivement décrire le système. **Le principe de Maximisation de production d'entropie stipule que parmi tous les états stationnaires considérés respectant les contraintes, celui qui va être choisi par le système est celui qui maximise la production d'entropie**. Ce principe n'est pas démontré mais comme nous l'avons mentionné en introduction est largement utilisé par la communauté scientifique dans différents

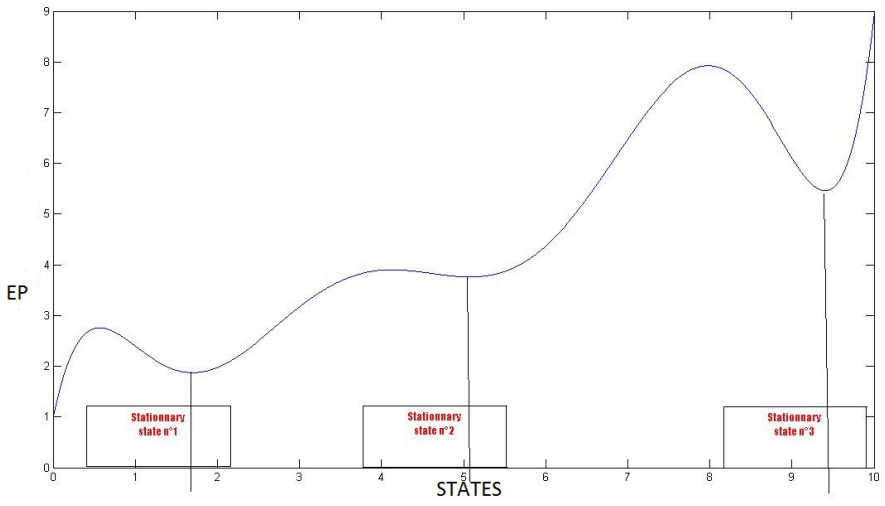


FIGURE 1.7 – Entropy Production vs States

domaines de la physique.

Ce principe est donc totalement différent du principe de Prigogine et est beaucoup plus général dans la mesure ou il ne suppose rien quant à la proximité du système par rapport à l'équilibre. MEP apparait donc ici comme un MaxMin : On minimise d'abord la production d'entropie Fig.(1.7) pour connaitre les états stationnaires puis, s'il en existe plusieurs, on maximise pour savoir lequel le système choisit.

4 Le principe de maximisation de l'entropie de KS.

L'entropie de Kolmogorov-Sinai (KS) ou entropie métrique a été initialement introduite par Kolmogorov en 1958 [22] comme un outil permettant de voir que deux systèmes dynamiques ne sont pas conjugués (isomorphes). Cette notion a été reprise et développée par son élève du moment : Y.Sinai (Prix Abel 2014) . Cette entropie est encore aujourd'hui très étudiée (et notamment en physique) car elle est intimement reliée au chaos. En effet, une transformation chaotique est une transformation d'entropie de KS non nulle. De fait, empiriquement, plus l'entropie de Kolmogorov-Sinai est grande plus le système est chaotique.

Parallèlement, en 1967 Fyliokov et Karpov propose un principe régissant les sytème hors équilibres dans un état stationnaire consistant à maximiser la production de l'entropie dynamique c'est à dire la dérivée de l'entropie sur les chemins. Ce principe a été repris par Evans et une explication est donnée dans [8]. De plus cette production de l'entropie dynamique est fortement reliée à l'entropie de KS, voir égale pour les chaines de Markov.

Nous commençons par donner une définition la plus rigoureuse possible de l'entropie de KS en s'appuyant sur le livre de Billingsley [23], de Katok et Hasselblatt [24] ainsi que le cours de F.Paulin et Y.Benoist [25] afin d'en comprendre sa signification. Puis nous discuterons cette entropie dans le cas particulier des chaines de Markov : nous démontrerons en particulier que l'entropie de KS sur les chaines de Markov et la production d'entropie dynamique sont égales. Nous introduirons ensuite la production d'entropie dynamique d'un système physique plus général et discuterons le principe de maximisation de l'entropie de KS.

4.1 definition formelle de l'entropie de Kolmogorov-Sinaï

Soit $(X, \mathcal{B}, \mu)$ un espace de probabilité. Soient A_1,A_2...A_n des éléments non vides de $\mathcal{B}$ tels que $\forall i, j\ i \neq j\ A_i \cap A_j = \emptyset$ et $\cup_i A_i = X$ alors $\alpha = A_1, ..., A_n$ est appelé une partition mesurable de X. Associé à cette partition on peut définir une entropie.

Définition : l'entropie de la partition α vaut

$$H(\alpha) = -\sum_{A \in \alpha} \mu(A) \log(\mu(A)). \tag{1.16}$$

Etant données deux partitions α et β on définit la partition Fig.(1.8) :

$$\alpha \vee \beta = \{A \cap B, A \in \alpha, B \in \beta \text{ et } A \cap B \neq \emptyset.\}$$

De manière récursive étant données une famille de partitions $\alpha_1 ... \alpha_n$ on définit alors :

$$\bigvee_i \alpha_i = \alpha_1 \vee \bigvee_{i \neq 1} \alpha_i.$$

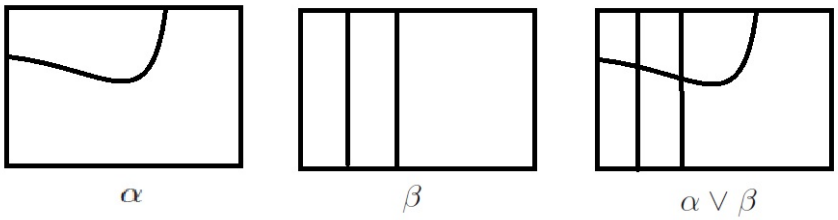


FIGURE 1.8 – Shéma de la partition $\alpha \vee \beta$ à partir des partitions α et β.

Proposition : Soient α et β deux partitions mesurables, on a :

$$H(\alpha \vee \beta) \leq H(\alpha) + H(\beta). \tag{1.17}$$

Définition : Un système dynamique probabilisé est la donnée d'un espace probabilisé $(X, \mathcal{B}, \mu)$ ainsi qu'une application f mesurable telle que $\forall B \in \mathcal{B}\ \mu(f^{-1}(B)) = \mu(B)$. On dit aussi que f préserve la mesure et on note naturellement le système dynamique probabilisé $(X, \mathcal{B}, \mu, f)$. De plus le système est dit inversible si f bijective et si f^{-1} conserve μ.

Définition : On note $f^{-1}\alpha = \{f^{-1}(A), A \in \alpha\}$. Il est aisé de vérifier vu la bijectivité de f que $f^{-1}\alpha$ est encore une partition et que $H(f^{-1}\alpha) = H(\alpha)$ vu que f conserve la mesure.

Définition/Proposition : Soit α une partition mesurable de X. La suite $\frac{1}{n} H(\bigvee_{i=1}^{n-1} f^{-1}\alpha)$ converge. On note sa limite $h(f, \alpha) = h_\mu(f, \alpha)$. On définit finalement l'entropie de Kolmogorov-Sinai comme le nombre :

$$h_\mu(f) = sup_\alpha h_\mu(f, \alpha). \tag{1.18}$$

Une bonne image est de supposer qu'à l'instant initial on ne distingue pas tous les états et on les regroupes donc par paquets ce qui constitue la partition α. Ensuite $\bigvee_{i=0}^{n-1} f^{-1}\alpha$ représente notre connaissance du système (des paquets) au bout de n itérations. $h_\mu(f, \alpha)$ apparait donc comme le gain d'information moyen à chaque itération.

Définition : la partition α est dite f-génératrice si la suite des partitions $\alpha_n = \bigvee_{i=0}^{n-1} f^{-1}\alpha$ engendre la σ-algèbre $\mathcal{B}$ ou si f est inversible et si la suite de partitions $\alpha_n = \bigvee_{i=-n}^{n} f^{-1}\alpha$ engendre la σ-algèbre $\mathcal{B}$

Proposition Si σ est f-génératrice alors $h_\mu(f) = h_\mu(f, \alpha)$

Cette dernière proposition est très utile car elle permet souvent le calcul de $h_\mu(f)$ comme dans l'exemple ci-après.

4.2 Entropie de KS pour un décalage de Bernouilli :

Considérons un alphabet de N lettres $\Sigma = 1, ..., N$ muni de la tribu discrète. Soit $(p_1, ..., p_N)$ N réels positifs de somme égale à 1. Définissons sur Σ une probabilité ν par $\nu(i) = p_i$. Soit alors l'espace des mots infinis sur cet alphabet : $\Sigma_N = 1, ..., N^{\mathbb{Z}}$. Définissons sur Σ_N la distance ultra-métrique :

$$d(x, y) = 2^{-a(x,y)} \text{ ou } a(x, y) = \inf |n|, x_n \neq y_n. \tag{1.19}$$

Prenons alors sur Σ_N la tribu de Borel associée à d et munissons Σ_N de la mesure produit : $\mu = \nu^{\mathbb{Z}}$. Considérons le décalage de Bernoulli :

$$\begin{aligned} \sigma : \Sigma_N &\to \Sigma_N \\ \sigma((x_i)_{i\in\mathbb{Z}}) &= (x_{i+1})_{i\in\mathbb{Z}}, \end{aligned} \tag{1.20}$$

σ est mesurable et la tribu de Borel sur Σ_N est engendrée par les cylindres :

pour $j_1, ..., j_k \in \Sigma$ et $i_1, ..., i_k \in \mathbb{Z}$

$$C\binom{j_1, ..., j_N}{i_1, ..., i_N} = \{x \in \Sigma_N, x_{i_1} = j_1 ... x_{i_k} = j_k\}. \tag{1.21}$$

De plus,

$$\mu(C\binom{j_1, ..., j_N}{i_1, ..., i_N}) = p_{j_1} ... p_{j_N}. \tag{1.22}$$

Enfin on montre facilement que σ conserve μ
Soit la partition $\alpha = \{C\binom{j}{0}, j \in \Sigma\}$. Elle est σ - génératrice et

$$\begin{aligned} H(\bigvee_{i=0}^{n-1} \sigma^{-1}\alpha) &= -\sum_{j_0,...,j_{n-1}} \mu(C\binom{j_0, ..., j_{n-1}}{0, ..., n-1}) \log(\mu(C\binom{j_0, ..., j_{n-1}}{0, ..., n-1})) \\ &= -\sum_{j_0,...,j_{n-1}} p_{j_0} ... pj_{n-1} \log(p_{j_0} ... pj_{n-1}) \\ &= -n \sum_j p_j \log(p_j). \end{aligned} \tag{1.23}$$

Donc des équations eq.(1.23) ainsi que eq.(1.18) on en déduit :

$$h_\mu(\sigma) = -\sum_{j\in\Sigma} p_j \log(p_j). \tag{1.24}$$

4.3 Chaine de Markov et entropie de KS.

Sur l'alphabet Σ définissons maintenant un graphe orienté $\overrightarrow{\Sigma}$ dont les sommets sont les éléments de Σ. $\overrightarrow{\Sigma}$ peut être représenté par une matrice carrée A de taille $N \times N$ définie par $A(i, j) = 1$ s'il existe une arrête entre i et j et $A(i, j) = 0$ sinon. De plus, nous supposerons le graphe connexe c'est à dire que pour tout i et j il existe un chemin allant de i à j. Soit l'ensemble

$$\Sigma_A = \{x \in \Sigma_N \, , \forall i \in \mathbb{Z}(x_i, x_{i+1}) \in \overrightarrow{\Sigma}\}, \tag{1.25}$$

Σ_A est stable par σ et notons σ_A sa restriction.

Soit $P = (p_{ij})$ une matrice carrée de taille N (appelée matrice de transition) telle que :

$$\begin{aligned} \forall (i,j) \in \Sigma^2 \quad & p_{ij} \geq 0. \\ \forall (i,j) \in \Sigma^2 \quad & P_{ij} > 0 \Leftrightarrow A(i,j) = 1 \text{ où A est la matrice d'adjacence.} \\ \forall i \in \Sigma \quad & \sum_{j=1...N} p_{ij} = 1. \end{aligned} \tag{1.26}$$

En appliquant le théorème de Perron-Frobenius à P il existe un unique vecteur $(p_1^{stat}, ..., p_N^{stat})$ tel que :

$$\begin{aligned} \forall i \in \Sigma \quad & p_i^{stat} \geq 0. \\ \forall i \in \Sigma \quad & \sum_{i=1}^{N} p_i^{stat} = 1. \\ \forall i \in \Sigma \quad & \sum_{i=1}^{N} p_i^{stat} p_{ij} = p_j^{stat}. \end{aligned} \tag{1.27}$$

Définissons alors sur Σ_A une mesure μ_A naturelle telle que σ_A préserve μ_A :

$$\mu_A(C\begin{pmatrix} j_0, ..., j_k \\ i, ..., i+k \end{pmatrix}) = p_{j_0}^{stat} p_{j_0 j_1} p j_1 j_2 ... p_{j_{k-1} j_k}. \tag{1.28}$$

De la même manière que dans la partie précédente on trouve aisément que l'entropie de KS prend la forme :

$$h_{\mu_A}(\sigma_A) = - \sum_{(i,j) \in \Sigma^2} p_i^{stat} p_{ij} \log(p_{ij}). \tag{1.29}$$

4.4 Production d'entropie dynamique

Considérons un système stochastique évoluant durant un intervalle de temps $[0, t]$ et notons $\Omega_{[0,t]}$ l'ensemble de ces trajectoires. L'entropie de Jaynes est alors définie par :

$$S_{dyn}(t) = - \sum_{\gamma \in \Omega_{[0,t]}} p(\gamma) \log(p(\gamma)), \tag{1.30}$$

où $p(\gamma)$ est la probabilité d'obtenir la trajectoire γ.

La production d'entropie dynamique vaut alors, si la limite existe :

$$\dot{S}_{dyn} = \lim_{t \to +\infty} \frac{S_{dyn}(t)}{t}. \tag{1.31}$$

Prenons le cas d'une chaine de Markov ayant N états comme dans la section précédente. Durant l'état stationnaire, la probabilité de la trajectoire $\gamma = (i_0, ..., i_k)$

vaut $p_{i_0}^{stat} p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{k-1} i_k}$. Ainsi, le calcul de $\dot{S_{dyn}}$ est totalement similaire à celui effectué pour trouver les équations eq.(1.23) et eq.(1.28) et donc on retrouve l'expression :

$$\dot{S}_{dyn} = h_{KS} = - \sum_{(i,j) \in \{1...n\}^2} p_i^{stat} p_{ij} \log(p_{ij}). \tag{1.32}$$

Ainsi, l'entropie de KS d'un point de vue mathématique et la production d'entropie dynamique coïncident pour une chaine de Markov.

4.5 Principe de Maximisation de l'entropie de KS

Fyliokov et Karpov [1] (1967) puis Monthus ont montré que maximiser l'entropie dynamique pour prédire l'état stationnaire des systèmes faiblement hors équilibre donnait de bons résultats, cohérents avec l'expérience et avec la physique à l'équilibre. De même, Luck et al [26] maximisent l'entropie de KS sur un réseau non régulier et observent que l'état obtenu reproduit un phénomène de localisation. De nombreux autres articles maximisent l'entropie de KS [27]...Grâce à la construction de Kolmogorov on comprend que plus le système est chaotique plus la partition α se déformera vite et donc plus l'entropie de KS sera grande. Néanmoins ceci n'est que qualitatif. Nous essaierons de donner des arguments plus quantitatifs quant à la maximisation de l'entropie de KS dans le cinquième chapitre.

Chapitre 2

Principe de maximisation de production d'entropie et modèle ASEP.

1 Chaines de Markov.

1.1 Définition et outils associés.

Considérons un processus aléatoire discret (X_n) modélisant une variable de l'état du système à l'instant n, comme par exemple sa taille si le système est une population ou encore la position d'un marcheur aléatoire... Souvent, pour un système physique, les valeurs de la variable aléatoire X_{n+1} dépendent du système aux temps passés $(t \leq n)$. Un système Markovien est caractérisé par le fait que la connaissance du système au temps $t = n + 1$ ne dépend que de sa connaissance au temps $t = n$ et non pas des temps précédents. Un processus de Markov est donc un processus aléatoire assez simple "sans mémoire". Comme son nom l'indique il a été introduit par Andrei Markov. Né en 1856 à Ryazan, Andrei Markov étudia à l' Université d' État de Saint-Pétersbourg en 1874 sous la tutelle de Tchebychev et en 1886, il devint membre de l'académie des Sciences de Saint-Pétersbourg. Comme exemples simples de systèmes Markovien on peut penser au célèbre jeu des serpents et des échelles ou encore au marcheur aléatoire ivre...

Les variables aléatoires $(X_n)_{n\in\mathbb{N}}$ sont définies sur un espace de probabilité $(\Omega, \mathcal{A}, \mu)$ et sont à valeurs dans l'espace des états $\mathcal{E}$ qui peut être soit fini soit dénombrable.

De manière rigoureuse, un processus aléatoire discret $(X_n)_{n\in\mathbb{N}}$ est Markovien ssi les probabilités conditionnelles vérifient :

$$\begin{aligned} \forall\,(i_0, i_1, ..., i_{n+1}) \in \mathcal{E}^{n+2} \text{ tel que } P(X_{n+1} = i_{n+1}, X_n = i_n, ..., X_0 = i_0) > 0. \\ P(X_{n+1} = i_{n+1}|X_n = i_n, ..., X_0 = i_0) = P(X_{n+1} = i_{n+1}|X_n = i_n). \end{aligned} \tag{2.1}$$

Une chaine de Markov est dite homogène si $\forall\,(i, j) \in \mathcal{E}^2$ la probabilité d'être dans l'état i à l'instant n et de passer dans l'état j à l'instant $n + 1$ ne dépend pas du temps n. On peut dès lors définir les probabilités de transitions de l'état i vers l'état j par

$$P(X_{n+1} = j|X_n = i) = P(i, j). \tag{2.2}$$

De plus, si l'espace des états est fini de cardinal N, on peut dès lors définir la matrice de transition P telle que $P = (P(i,j)) = (p_{ij})_{(i,j)\in\mathbb{N}^2}$. Cette matrice est stochastique et on vérifie aisément que $\forall\,(k,n) \in \mathbb{N}\ (P^k)_{ij} = P(X_{n+k} = j | X_n = i)$. On a donc, si μ_n désigne le vecteur représentant le loi de la v.a X_n :

$$\mu_{n+1} = P^t \mu_n \text{ ou encore } \mu_n = (P^t)^n \mu_0. \tag{2.3}$$

Une chaine de Markov peut-être vue comme un graphe où les noeuds sont indicés par les états et les liens correspondent aux transitions de i vers j : Il existe un lien de i vers j si et seulement si $p_{ij} > 0$. De plus, l'évolution de la chaine de Markov est alors semblable à l'évolution d'un marcheur aléatoire sur ce graphe qui, s'il est au point i à l'instant n, a la probabilité p_{ij} d'aller au point j à l'instant $n+1$.

Un outil très utile est la matrice d'adjacence A définie par :

$$\begin{aligned} A(i,j) = 1 \text{ si } p_{ij} > 0, \\ A(i,j) = 0 \text{ sinon .} \end{aligned} \tag{2.4}$$

On vérifie dès lors que le terme $(A^n)_{ij}$ correspond aux nombres de chemins de longueurs n allant de i à j.

1.2 Convergence des chaines de Markov.

On dit que la chaine de Markov est **irréductible** si le graphe associé est connexe, c'est à dire si quelque soit les états on peut trouver un chemin allant d'un état à un autre. De manière plus formelle, le graphe est connexe ssi :

$$\forall\ (i,j) \in [1,N]^2\ \exists n \in \mathbb{N}, P^n(i,j) > 0. \tag{2.5}$$

Lorsque la chaine de Markov est irréductible on a l'existence et l'unicité d'une mesure stationnaire μ_{stat} telle que :

$$\mu_{stat} = P^t \mu_{stat}. \tag{2.6}$$



FIGURE 2.1 – Andrei Markov

De plus, on peut définir la période d'un état i par :

$$T(i) = \text{pgcd}\{n \geq 1\,, P^n(i,j) > 0\}. \tag{2.7}$$

On dit alors qu'un chaine de Markov est **apériodique** si ses états ont pour période 1.

Si une chaine de Markov est irréductible et apériodique alors on montre que :

$$\exists M \in \mathbb{N}\,, \forall n \geq M\, \forall (i,j) \in [1,N]^2\, P^n(i,j) > 0, \tag{2.8}$$

Ceci permet de démontrer que si une chaine de Markov est irréductible et apériodique alors quel que soit la mesure initiale μ_0, la suite μ_n converge vers μ_{stat}. Nous considèrerons dans la suite des chaines de Markov irréductibles et apériodiques.

1.3 Production d'entropie microscopique

Il existe de nombreuses définitions de la production d'entropie qui prennent d'ailleurs parfois des formes différentes suivant l'approche considérée [28]. Néanmoins, pour un processus Markovien, nous pouvons définir une production d'entropie microscopique qui caractérise **l'irréversibilité** du système [29]. Considérons une trajectoire sur l'espace des états $\omega = (i_0, ..., i_T)$ et sa trajectoire renversée en temps $\theta\omega = (i_T, ..., i_0)$. De plus, notons par P_μ la probabilité sur l'espace des trajectoires du processus de Markov commençant avec la distribution de probabilité μ. Et de même, notons $P_{\mu T}$ la probabilité sur la chaine de Markov réciproque partant de μT, la distribution de probabilité qui est l'évolution de μ au temps T.

On démontre alors [30] qu'il existe une densité entre ces deux probabilités. Dès lors, la dérivée de Radon Nikodym est appelée la production d'entropie :

$$S_\mu^T(\omega) = \log(\frac{dP_\mu}{dP_{\mu T}})(\omega). \tag{2.9}$$

La production d'entropie moyenne $\sigma(\mu)$ est alors définie par :

$$\sigma(\mu) = \lim_{T \to 0} \frac{1}{T} < S_\mu^T >_\mu . \tag{2.10}$$

La production d'entropie microscopique représente donc de combien le système n'est pas réversible : plus le système est irréversible, plus la production d'entropie $\sigma(\mu)$ est grande.

Avec les notations introduites avant l'équation 2.10 devient :

$$\sigma(\mu) = \sum_{(i,j)} (\mu_i p_{ij} - \mu_j p_{ji}) \log(\frac{\mu_i p_{ij}}{\mu_j p_{ji}}). \tag{2.11}$$

Cette expression qui a été trouvée ici à partir d'une vision mathématique fondée sur des arguments d'irréversibilité a aussi une signification thermodynamique [31].

En effet, à partir de l'expression générale de l'entropie de Gibbs pour une chaine de Markov :

$$S_t = -\sum_i \mu_i(t) \log(\mu_i(t)), \tag{2.12}$$

et de 2.3 on déduit :

$$S_{t+1} - S_t = \Delta_e S + \Delta_i S, \tag{2.13}$$

où $\Delta_e S$ est le flux d'entropie et vaut :

$$\Delta_e S = -\sum_{(i,j)} \mu_i(t) p_{ij} \log(\frac{\mu_j(t+1)p_{ij}}{\mu_j(t)p_{ji}}), \tag{2.14}$$

et $\Delta_i S$ est la production d'entropie irréversible et vaut $\sigma(\mu(t))$.

2 Une chaine de Markov particulière : le modèle SSEP et ASEP.

2.1 Motivations et lien avec le climat.

Nous rappelons que le principe de maximisation de production d'entropie utilisé en climatologie prend la forme suivante : Un système hors équilibre, pouvant avoir différents états stationnaires suivant les valeurs des flux qui ne sont pas fixés, choisit son état stationnaire final de telle sorte à maximiser la production d'entropie. Afin de mieux comprendre ce principe j'ai voulu trouver un système simple qui reproduisait le mieux possible le modèle utilisé par Paltridge.
Il devait donc répondre aux critères suivants :

i) Système en boite.

ii) Système qui peut admettre différents états stationnaires possibles.

iii) Système dans lequel il existe des flux qui ne sont pas fixés.

iv) Système couplé avec un réservoir extérieur tout comme le système terrestre est couplé avec le soleil...

J'ai alors pensé au Symmetric Simple Exclusion Process (SSEP) et à l' Asymmetric Simple Exclusion Process (ASEP).

2.2 Explication des modèles SSEP et ASEP.

Le système SSEP consiste en un réseau 1D couplé avec deux réservoirs extérieurs sur lequel des particules peuvent se déplacer. Plus exactement il est constitué de L cases pouvant contenir une ou zéro particule. A chaque pas de temps de $dt << 1$ les particules qui se trouvent à l'intérieur du réseau (c'est à dire sur les cases notés de 2 à $L-1$) peuvent sauter à droite ou a gauche avec une probabilité pdt si celles-ci sont vides, où p est un réel compris entre 0 et 1. En ce qui concerne les cases du bord, elles sont couplées avec des réservoirs de la manière suivante : si la case de droite est vide une particule peut entrer avec une probabilité αdt et si elle est pleine elle peut sortir avec une probabilité γdt. A gauche la particule peut entrer avec une probabilité δdt et sortir avec une probabilité βdt. α, γ, β, δ sont ici quatre réels.

Le système ASEP ressemble au système SSEP à la différence qu'il n'y a plus la symétrie pour les particules de sauter à droite et à gauche. Soient p et q deux réels compris entre 0 et 1, les particules à l'intèrieur du réseau peuvent alors sauter à droite avec un probabilité pdt et à gauche avec une probabilité qdt.

Les modèles ASEP et SSEP sont des systèmes simples de diffusion sur réseaux. Ils sont utilisés aujourd'hui pour modéliser des systèmes physiques comme le trafic routier [32], les comportements sociaux, la propagation d'une maladie [33]...Ils sont aussi étudié pour leurs "beautés" dans la mesure ou ils peuvent se résoudre analytiquement [34, 35].

Néanmoins, lorsque p et q sont fixés le système admet un unique état stationnaire et nous ne pouvons donc pas appliquer le principe de maximisation de production d'entropie. L'idée va être **de faire varier p et q afin d'obtenir une infinité d'états stationnaires, de voir s'il existe $p*$ et $q*$ qui maximisent la production d'entropie, et enfin d'étudier à quoi correspondent ces $p*$ et $q*$.**

2.3 Modèle CASEP et Langmuir-ASEP.

Commençons par décrire les deux modèles que nous étudierons. Le premier est le modèle CASEP (pour changing ASEP). Il consiste à prendre $q = 1 - p$ et $dt = \frac{1}{L}$. Ceci équivaut à choisir à chaque pas de temps une case au hasard et si elle contient une particule et que la case de droite est libre de la faire sauter à droite avec une probabilité p ou, de même, à gauche avec une probabilité $1 - p$. En ce qui concerne le modèle Langmuir-ASEP il permet de plus aux particules de n'importe quelle case de rentrer avec une probabilité w_a et de sortir avec une probabilité w_b. Ceci permet de représenter de façon plus réaliste le modèle de Paltridge dans lequel les flux ne sont pas uniquement entre les boites voisines mais aussi vers l'extérieur.

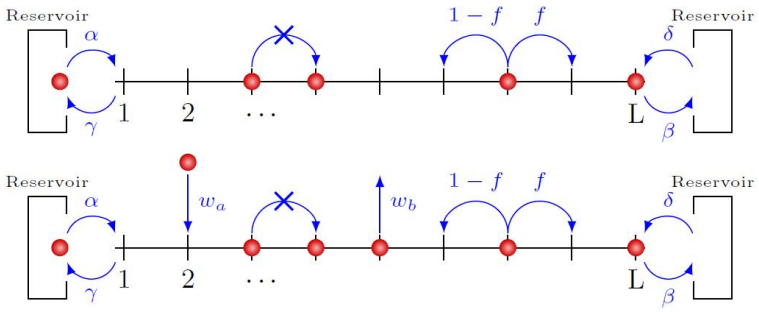


FIGURE 2.2 – Représentation des modèles CASEP en haut et LK-ASEP en bas.

Ces deux modèles sont des systèmes Markovien à 2^L états. En effet chaque état pouvant être représenté par une succesion d'un nombre L de 0 ou 1 (0110...111 par exemple) où 0 représente une case vide et 1 une case pleine. On peut dès lors s'intéresser, entre autre, à trois quantités :

i) La production d'entropie microscopique eq.(2.11)

ii) La production d'entropie macroscopique eq.(1.3)

iii) L'entropie de Kolmogorov-Sinaï eq.(1.31)

Ces trois quantités dépendent des paramètres p α β γ δ w_a et w_b. Cependant, si nous fixons les contraintes extérieurs, ces paramètres ne dépendent plus que de p. De plus p est fortement relié au flux de particules. On comprend que plus p est élevé, plus le flux de particules de la gauche vers la droite est important. Ainsi, en faisant varier p on fait varier le flux de particule et par conséquent l'état stationnaire de système. On remarque numériquement que la production d'entropie microscopique eq.(2.11) est une fonction croissante de p et n'admet donc pas de maximum. Par contre, la production d'entropie macroscopique ainsi que l'entropie de KS admettent toutes deux un maximum. Nous allons voir que l'endroit ($p*$) où est atteint ce maximum est le même pour ces deux quantités. Nous établissons ainsi un lien numérique entre deux principes à priori différents : Le principe de Maximisation de production d'entropie ainsi que le principe de maximisation de l'entropie de KS. C'est ce que nous détaillons dans l'article ci-après.

3 Article.

Entropy **2014**, *16*, 1037-1046; doi:10.3390/e16021037

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Maximum Entropy Production *vs*. Kolmogorov-Sinai Entropy in a Constrained ASEP Model

Martin Mihelich [1,*], Bérengère Dubrulle [1], Didier Paillard [2] and Corentin Herbert [3]

[1] Laboratoire SPHYNX, CEA/IRAMIS/SPEC, CNRS URA 2464, F-91191 Gif-sur-Yvette, France; E-Mail: berengere.dubrulle@cea.fr

[2] Laboratoire des Sciences du Climat et de l'Environnement, IPSL, Orme des Merisiers, Gif-sur-Yvette 91191, France; E-Mail: didier.paillard@lsce.ipsl.fr

[3] National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA; E-Mail: cherbert@ucar.edu

* Author to whom correspondence should be addressed; E-Mail: martin.mihelich@cea.fr; Tel.: +33-1-6908-7072.

Received: 25 November 2013; in revised form: 6 January 2014 / Accepted: 10 February 2014 / Published: 19 February 2014

Abstract: The asymmetric simple exclusion process (ASEP) has become a paradigmatic toy-model of a non-equilibrium system, and much effort has been made in the past decades to compute exactly its statistics for given dynamical rules. Here, a different approach is developed; analogously to the equilibrium situation, we consider that the dynamical rules are not exactly known. Allowing for the transition rate to vary, we show that the dynamical rules that maximize the entropy production and those that maximise the rate of variation of the dynamical entropy, known as the Kolmogorov-Sinai entropy coincide with good accuracy. We study the dependence of this agreement on the size of the system and the couplings with the reservoirs, for the original ASEP and a variant with Langmuir kinetics.

Keywords: maximum entropy production; Kolmogorov-Sinai Entropy; ASEP model

1. Introduction

The theory of dynamical systems at statistical equilibrium consists of well identified principles lying on firm mathematical foundations [1–3]. Although the representative point of the system in phase space (the *microstate*) can have a complicated dynamics, the statistics of any function on phase space (an

observable) are well described, at least for large enough systems (in the *thermodynamic limit*) by simple invariant measures depending only on the invariants of the system, and not on the details of the dynamical equations. On the contrary, general theories describing the statistics of non-equilibrium systems are still in their infancy; although there have been attempts at formulating general principles governing non-equilibrium systems [4,5], they have not reached the level of mathematical rigor of their equilibrium counterpart, partly because there may not exist a universal way of being out-of-equilibrium, and remain valid essentially close to equilibrium.

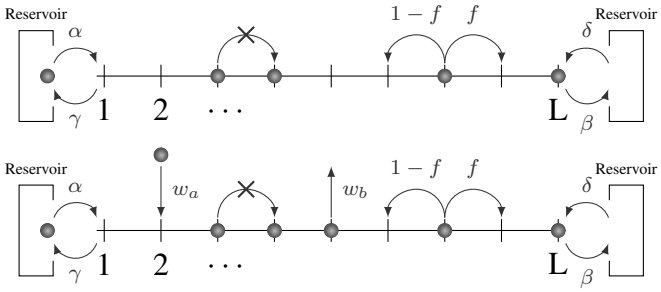
Nevertheless, some progress have been achieved recently, in particular in the framework of "toy models". Similarly to the Ising model which played a central role in the theory of critical phenomena, the so-called *asymmetric simple exclusion process* (ASEP) [6,7], which describes the transport of particles between two reservoirs, have become the dominant paradigm for non-equilibrium systems. Over the past few years, this system and others have benefited from the theory of large deviations, developed by Donsker and Varadhan [8–11]. This theory has introduced the large deviation function, which characterizes the statistics of the system in an asymptotic regime (large system, large time, low noise,...) and is a natural candidate for playing the out-of-equilibrium analogous role to the equilibrium free energy. In the long-time limit, the ASEP reaches a steady-state, and a macroscopic current is established between the two reservoirs. The average steady-state current, and the corresponding density profile, have been computed exactly [6,12,13], and more recently, the full large deviation function, which also describes the atypical fluctuations around these averages, has also been solved exactly [14], using a generalization of the *matrix product ansatz*.

Powerful as these techniques may be, they remain far from being applicable to more realistic systems. Indeed, if many systems in nature share some conceptual properties with the ASEP (establishment of a steady-state with fluctuations dominating the average behaviors and the resulting anomalous transport properties), not all of them are easily written as stochastic processes driven by a master equation. For instance, it has become clear over the years that deterministic systems with a very large number of degrees of freedom such as turbulent fluids, and in particular the atmosphere and the ocean of the Earth, should be thought of in terms of non-equilibrium statistical mechanics. However, up to now, attempts to apply such tools in the climate system have been restricted to phenomenological applications of the "Maximum Entropy Production" variational principle [15–17]. The success of these approaches points at a deeper connection with the general properties of non-equilibrium systems. A justification has been attempted [18], and dismissed [19,20], following the ideas of Jaynes that non-equilibrium systems should be characterized by a probability distribution on the trajectories in phase space, instead of just the points in phase space at equilibrium. In the context of Markov chains, this idea has proved to provide a natural generalization of equilibrium statistical mechanics [21], by considering the Kolmogorov-Sinai (KS) entropy. In this Letter, we establish a bridge between the phenomenological principle of maximum entropy production on the one hand, and the maximization of KS entropy on the other hand, in the context of the ASEP. We also consider a variant of the ASEP with Langmuir kinetics [22], which mimics the phenomenological climate model designed by Paltridge [15]. In both cases, rather than solving exactly the system for given dynamical rules, we show that the maxima location of the thermodynamic entropy production and the KS entropy as functions of the rate of transition coincide, and study the validity of this agreement when the system size and reservoir couplings vary.

2. Model Description

In this paper, we have considered a modified version to the ASEP model-hereafter named the CASEP model (for Changing ASEP model) or the LK-CASEP model when Langmuir dynamics is added- so as to be able to change the particle flux. A class of toy models for out-of-equilibrium systems consists of variants of the *exclusion process* (with open boundaries): let us consider a one dimensional lattice gas model composed of L sites between two reservoirs. Each site contains at most one particle. The particle undergoes a random walk on the lattice with the following rules: at each time step a site is chosen randomly. If this site belongs to the bulk and is occupied by a particle, the particle can jump right with a probability p and left with a probability q provided the neighboring sites are empty. At the boundaries, the system is coupled with the two reservoirs: at the left (resp. right) boundary, particles from the reservoir can enter the lattice with probability α and particles from the lattice can exit with probability γ (resp. δ and β). This coupling is equivalent to imposing the respective densities $\rho_a = \frac{\alpha}{\alpha+\gamma}$ and $\rho_b = \frac{\delta}{\delta+\beta}$. When $\rho_a = \rho_b$, the system reaches an equilibrium state, while it becomes increasingly out-of-equilibrium as the difference between ρ_a and ρ_b increases. The numerous variants of the exclusion process depend on choices of constraints on the hopping rates p and q: the *symmetric simple exclusion process* (SSEP) imposes $p = q$. Imposing a preferred direction for particle transport (e.g., $p > q$) yields the *asymmetric simple exclusion model* (ASEP). An extreme case is the *totally asymmetric simple exclusion model* (TASEP), for which $q = 0$. Here, we shall consider a particular class of ASEP, by imposing that the particle has to hop to a neighboring site—provided it is empty; it cannot remain in place unless both the neighboring sites are occupied (see Figure 1).

Figure 1. The dynamical rules of the asymmetric simple exclusion process (ASEP), without (top) and with (bottom, LK-ASEP) Langmuir kinetics.



This amounts to setting $p = f$ and $q = 1 - f$. We shall call this variant the *Changing asymmetric simple exclusion process* (CASEP).

The exclusion processes can be seen as microscopic models of transport. We are interested here in the connection with heat transport in turbulent flows at a macroscopic scale, like the atmosphere and the ocean. For such systems, phenomenological variational principles were suggested [15] to compute the energy transport associated with the establishment of a temperature gradient, outside of the diffusive regime. In the atmosphere, radiative exchanges on the vertical coexist with this meridional transport. To mimic this effect, we shall also consider a variant of the CASEP which includes Langmuir Kinetic dynamics (LK-CASEP) [22]: at each time step, particles can appear in an empty site with probability w_a and disappear from an occupied site with probability w_b.

The exclusion processes, and in particular the CASEP, are special cases of discrete Markov processes with 2^L states. It is therefore difficult to study numerically this system for L larger than 10. These processes are characterized by their transition matrix $P = (p_{ij})$ which is irreducible. Thus, the probability measure on the states converges to the stationary probability measure $\mu^{stat} = (\mu_1^{stat}, ..., \mu_{2^L}^{stat})$ which satisfies:

$$\mu_i^{stat} = \sum_{j=1}^{2^L} \mu_j^{stat}.p_{ji} \; \forall i \in [\![1, 2^L]\!] \tag{1}$$

The states are labeled using the base 2. The p_{ij} are equal to $f.dt$ (resp. $(1-f).dt$) for a transition corresponding to a right jump of particule (resp. left) inside the bulk. For a transition corresponding to a jump at the boundary of the lattice, the p_{ij} are equal to $\alpha.dt$, $\beta.dt$, $\gamma.dt$, $\delta.dt$. In the LK-CASEP Model, for a transition corresponding to the appearance or disappearance of a particle, $p_{ij} = w_a.dt$ or $p_{ij} = w_b.dt$. Once the transition matrix has been defined through these simple rules, the dynamics can then be obtained by moving particules at each time steps, until a stationary states has been reached. The density of particules on the lattice $\rho = (\rho_1, ..., \rho_L)$ in such a stationary states can then be computed numerically.

Most of the studies of the ASEP model consider a fixed dynamics (given value for p and q) and try to solve exactly the resulting statistics for the steady-state current [14]. Here, we rather consider that the dynamics is not known exactly, and vary the parameter f in order to obtain different stationary states and calculate their entropy production σ function of f. Indeed, changing the parameter f is equivalent to change the flux of particles J. At fixed f, we can also use the dynamical properties of the CASEP model to compute the Kolmogorov-Sinai Entropy h_{KS} as a function of f and compare it to $\sigma(f)$.

In the $f = \frac{1}{2}$ case, the system reaches a steady state that is reminiscent of a "conductive state", with a linear density behaviour over the lattice, going from $\rho_a = \frac{\alpha}{\alpha+\gamma}$ on the left side to $\rho_b = \frac{\delta}{\delta+\beta}$ on the right side [6]. When $f \neq \frac{1}{2}$, the system reaches another stationary state, ressembling a "convective state", in which the density is mainly constant in the bulk, with steep transitions near the edges towards the left and the right side densities ρ_a and ρ_b.

3. Dynamical and Thermodynamical Entropy

We now define both the Macroscopic Entropy production and the Kolmogorov-Sinai Entropy for the CASEP model in order to compare them.

3.1. Komogorov-Sinai Entropy

There are many ways to estimate the Kolmogorov-Sinai entropy associated with a Markov chain [21,23]. Here we compute Kolmogorov-Sinai Entropy as the time derivative of the Jaynes Entropy. To characterize the dynamics of the system during the time interval $[0, t]$, one considers the possible dynamical trajectories $\Gamma_{[0,t]}$ and the associated probabilities $p_{\Gamma_{[0,t]}}$. The dynamical trajectories entropy—the Jaynes entropy—reads:

$$S_{Jaynes}(t) = -\sum_{\Gamma_{[0,t]}} p_{\Gamma_{[0,t]}}.\log(p_{\Gamma_{[0,t]}}) \tag{2}$$

For a Markov chain we find that:

$$S_{Jaynes}(t) - S_{Jaynes}(t-1) = -\sum_{(i,j)} \mu_{i_{stat}} p_{ij} \log(p_{ij}) \tag{3}$$

Thus, the Kolmogorov-Sinai Entropy for the Markov chain is:

$$h_{KS} = -\sum_{(i,j)} \mu_{i_{stat}} p_{ij} \log(p_{ij}), \tag{4}$$

where μ_i is the stationary measure and p_{ij} the transition matrix.

3.2. Entropy Production

For a macroscopic system subject to thermodynamic forces X_i and fluxes J_i, the thermodynamics entropy production is given by [5,24]:

$$\sigma_{thermo} = \sum_i J_i X_i \tag{5}$$

The fluxes to consider for a diffusive particules model are fluxes of particules and the thermodynamics forces can be written $X = \Delta(-\frac{\nu}{T})$ where T is the temperature and ν the chemical potential proportionnal to $\log(\rho)$ for an ideal gas [24]. So, as the temperature is here fixed, the Thermodynamics Entropy production of a given stationary state takes the form:

$$\sigma_{thermo} \propto \sum_{i=1}^{L} J.(\log(\rho_i) - \log(\rho_{i+1})) = J.(\log(\rho_1) - \log(\rho_L)) \tag{6}$$

where ρ is the stationary density distribution and J the particle fluxes, that are both (nonlinear) function of f. Note that we can find this result considering a continuous system. Indeed, the statistical entropy for an ideal gas confined between the walls A_+ and B_- is:

$$S(t) = -\int_{A_+}^{B_-} \rho(x,t) \log(\rho(x,t)) dx \tag{7}$$

Taking the times derivative of Equation (7) with the continuity equation $div(J) + \frac{\partial \rho}{\partial t} = 0$ yields:

$$\frac{\partial S}{\partial t} = \int_{A_+}^{B_-} \frac{\partial J(x,t)}{\partial x} (1 + \log(\rho(x,t))) dx \tag{8}$$

We can write Equation (8) to make entropy production appears:

$$\frac{\partial S}{\partial t} - \int_{A_+}^{B_-} div(J(x,t)(1 + \log(\rho(x,t)))) = -\int_{A_+}^{B_-} J(x,t) \frac{\partial \log(\rho(x,t))}{\partial x} dx \tag{9}$$

Thus, the entropy production reads:

$$\sigma_{thermo} = -\int_{A_+}^{B_-} J(x,t) \frac{\partial \log(\rho(x,t))}{\partial x} dx \tag{10}$$

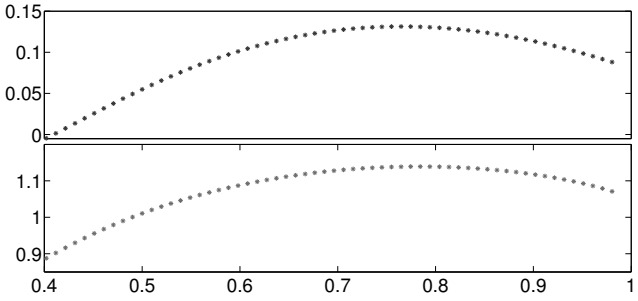
This is nothing but the continuous limit of Equation (6).

4. Numerical Results

4.1. Comparison between KS Entropy and Macroscopic Entropy Production

Given a value of L, ρ_a and ρ_b, we have computed the stationary states of the CASEP model as a function of f and used them to compute numerically the entropy production in Equation (6) and the Kolmogorov-Sinai Entropy in Equation (4). An example is provided in Figure 2 for $\rho_a = 0.75$, $\rho_b = 0.1$ and $L = 10$. We see that both function are of approximate parabolic shape, reaching a maximum for $f \geq 1/2$. In the sequel, we note $f_{max_{ep}}$ (resp. $f_{max_{ks}}$ the value of f at which the maximum of the entropy production (resp. of the Kolmogorov-Sinai entropy) is reached. The existence of a maximum of entropy production can be simply understood by considering the case $L = 2$, *i.e.*, a model with only two boxes, where we assume for simplicity that $\rho_a > \rho_b$. Noting ρ_1 and ρ_2 the density of the box number 1 and number 2 respectively, it is easy to see that for $J = 0$, $\sigma_{thermo} = 0$ whereas for J large enough $\rho_1 = \rho_2$ and $\sigma_{thermo} = 0$. Thus, in between these two values of J, σ_{thermo} has at least one maximum. We do not have such heuristic explanation for the existence of the maximum of the Kolmogorov-Sinai entropy. We note that although the entropy production and the Kolmogorov-Sinai entropy have similar parabolic shape, they do not coincide exactly: the entropy production and the Kolmogorov-Sinai entropy differ. For $\rho_a = \rho_b$ (*i.e.*, at equilibrium), their maxima take a common value $f_{max_{ep}} = f_{max_{ks}} = 1/2$, corresponding to the maximum of entropy: at equilibrium, all functionals are maximum for the same stationary state, corresponding to the "conductive state". As the difference between ρ_a and ρ_b increases, the two maxima $f_{max_{ep}}$ and $f_{max_{ks}}$ deviates from the equilibrium value $\frac{1}{2}$. This is the signature of a "convective dynamics". We would like now to explore the behaviour of the two entropy around this conductive state, by studying the behaviour of the maxima difference, Δf_{max}=$|f_{max_{ep}} - f_{max_{ks}}|$.

Figure 2. Macroscopic entropy production (top) and Kolmogorov-Sinai entropy (bottom) production function of f for $\rho_a = 0.75$ and $\rho_b = 0.1$ and $L = 10$. For these parameter values, we have limited the range of f from 0.4 to 1. Indeed $f \leq 0.4$ corresponds to a state where the flux of particules is in the same direction as the gradient of ρ which is physically impossible.

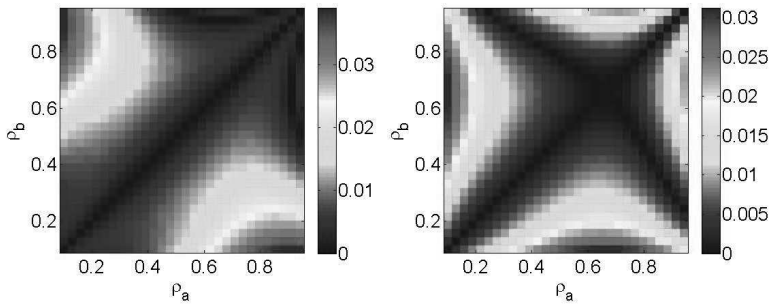


4.2. Behaviour of the Maxima Difference

Let us first fix the size of the box, L, and study the variation of Δf_{max}=$|f_{max_{ep}} - f_{max_{ks}}|$ as function of ρ_b and ρ_a. This is shown in Figure 3 (Left). It can be seen that Δf_{max} is zero for at perfect equilibrium, and remains small near equilibrium, for $\rho_a \approx \rho_b$. The difference then increases as the system deviates from equilibrium, reaching a maximal value of about 4% at $L = 10$. These general features are conserved

when including the Langmuir-Dynamics as can be seen in Figure 3 (Right): for $L = 8$ the difference between the maximum of entropy production and the maximum of the Kolmogorov-Sinai entropy also of the order of 4%. This is still valid varying w_a and w_b in $[0, \frac{1}{L}]$. These shows that the coincidence between the maxima of the KS entropy and the entropy production is robust and independent of inclusion of vertical fluxes.

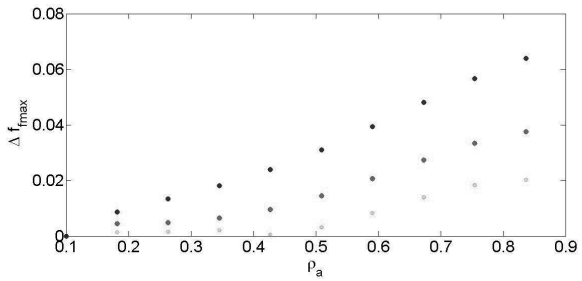
Figure 3. $\Delta f_{max} = |f_{max_{ep}} - f_{max_{ks}}|$ as function of ρ_a and ρ_b for fixed box size L: left at $L = 1$) in the CASEP model. Right: at $L = 8$ for the LK- CASEP model with $w_a = 0.06$ and $w_b = 0.03$.



4.3. Influence of Box Size

The previous results have been obtained for a number of boxes comparable with the number of boxes used in climate models. However, it is interesting to investigate possible finite size effects by varying L and observe the orresponding variations of Δf_{max}. Due to limited computational power, we were not able to simulate systems with L larger than $L = 10$. We could however decrease L and observe the result. This is done in Figure 4, where both $f_{max_{ep}}$ and $f_{max_{ks}}$ are plotted as a function of ρ_a and L for fixed ρ_b. One sees that for fixed ρ_a, the difference between the two values Δf_{max} decreases as L increases. So it is likely that for $L \to \infty$, this quantity converges to 0 for any fixed ρ_a and ρ_b.

Figure 4. Maxima difference Δf_{max} as function of ρ_a at fixed $\rho_b = 0.1$, for different system sizes L: blue: $L = 6$; red: $L = 8$; green: $L = 10$.



5. Discussion and Conclusions

In the present paper, we have built a bridge between stationary states that are maxima of Macroscopic Entropy production and maxima of Kolmogorov Sinai entropy, within a simple model of out-of-equilibrium physics. Both maxima coincide at equilibrium, and their difference increases with

increasing out-of equilibrium. For fixed out-of-equilibrium conditions, however, the difference decreases with increasing system size. We may then postulate that for any fixed out-of-equilibrium conditions, both the Macroscopic Entropy production and the Kolmogorov Sinai entropy peak at the same stationary state, that thereby acquire a priviledged status. Whether such a stationary state is the natural one selected by the dynamics is actually an open question. Many studies point towards the relevence of the Kolmogorov Sinai entropy for out-of-equilibrium systems. For example, Latora and Baranger [25] established a link between the Kolmogorov-Sinai entropy, which is a microscopic quantity, with the derivative of the coarse grained entropy, which is a macroscopic quantity, in an out of equilibrium dynamical system. In the present paper, we have evidenced another connection between the Kolmogorov-Sinai entropy and a macroscopic quantity. This link, one hand, strengthens the hypothesis that Kolmogorov-Sinai entropy may play a key role in the study of out of equilibrium systems. On the other hand, its link with the Maximum Entropy Production may explain its rather successful role played in in many area of physics [26] such as solid physics [27], electromagnetism, quantum physics or climate science [15,17]. At present time, the connection has only been established for a 1D discrete special system. Generalizing our result to continuous ASEP model is straightforward, using the expression of the Kolmogorov-Sinai entropy found by [28]. However, it would be very interesting to generalize the link between maxima of Maximum Entropy Production and maxima of Kolmogorov-Sinai entropy to more complicated systems, such as chaotic dynamical systems, lattice Bolzmann or turbulent flows.

Acknowledgments

Martin Mihelich thanks IDEEX Paris-Saclay for financial support. We thank The National Center for Atmospheric Research which is sponsored by the National Science Foundation. The authors would like to thank the two anonymous reviewers whose comments help to clarify the manuscript.

Author Contributions

Each author contributed equally to this work.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Khinchin, A. *The Mathematical Foundations of Statistical Mechanics*; Dover Publications: New York, NY, USA, 1949.
2. Ruelle, D. *Statistical Mechanics: Rigorous Results*; Benjamin: Amsterdam, The Netherlands, 1969.
3. Ellis, R.S. *Entropy, Large Deviations, and Statistical Mechanics*; Springer: New York, NY, USA, 1985.
4. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes*; Interscience: New York, NY, USA, 1967.

5. De Groot, S.; Mazur, P. *Non-Equilibrium Thermodynamics*; Dover Publications: New York, NY, USA, 2011.
6. Derrida, B. Non-equilibrium steady states: Fluctuations and large deviations of the density and of the current. *J. Stat. Mech.* **2007**, doi:10.1088/1742-5468/2007/07/P07023.
7. Chou, T.; Mallick, K.; Zia, R.K.P. Non-equilibrium statistical mechanics: From a paradigmatic model to biological transport. *Rep. Prog. Phys.* **2011**, *74*, 116601.
8. Donsker, M.D.; Varadhan, S.R.S. Asymptotic evaluation of certain Markov process expectations for large time, I. *Commun. Pure Appl. Math.* **1975**, *28*, 1–47.
9. Donsker, M.D.; Varadhan, S.R.S. Asymptotic evaluation of certain Markov process expectations for large time, II. *Commun. Pure Appl. Math.* **1975**, *28*, 279–301.
10. Donsker, M.D.; Varadhan, S.R.S. Asymptotic evaluation of certain Markov process expectations for large time—III. *Commun. Pure Appl. Math.* **1976**, *29*, 389–461.
11. Donsker, M.D.; Varadhan, S.R.S. Asymptotic evaluation of certain Markov process expectations for large time. IV. *Commun. Pure Appl. Math.* **1983**, *36*, 183–212.
12. Schütz, G.M. *Phase Transitions and Critical Phenomena*; Domb, C., Lebowitz, J.L., Eds.; Academic: New York, NY, USA, 2001; Volume 19.
13. Blythe, R.A.; Evans, M.R. Nonequilibrium steady sstates of matrix product form: A solver's guide *J. Phys. A* **2007**, *40*, 333–441.
14. Gorissen, M.; Lazarescu, A.; Mallick, K.; Vanderzande, C. Exact current statistics of the asymmetric simple exclusion process with open boundaries. *Phys. Rev. Lett.* **2012**, *109*, 170601.
15. Paltridge, G.W. Global dynamics and climate—a system of minimum entropy exchange. *Q. J. R. Meteorol. Soc.* **1975**, *101*, 475–484.
16. Ozawa, H.; Ohmura, A.; Lorenz, R.; Pujol, T. The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle. *Rev. Geophys.* **2003**, *41*, doi:10.1029/2002RG000113.
17. Herbert, C.; Paillard, D.; Kageyama, M.; Dubrulle, B. Present and Last Glacial Maximum climates as states of maximum entropy production. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 1059–1069.
18. Dewar, R.C. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J. Phys. A* **2003**, *36*, 631–641.
19. Grinstein, G.; Linsker, R. Comments on a derivation and application of the 'maximum entropy production' principle. *J. Phys. A* **2007**, *40*, 9717–9720.
20. Bruers, S. A discussion on maximum entropy production and information theory. *J. Phys. A* **2007**, *40*, 7441–7450.
21. Monthus, C. Non-equilibrium steady states: Maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints. *J. Stat. Mech.* **2011**, doi:10.1088/1742-5468/2011/03/P03008.
22. Parmeggiani, A.; Franosch, T.; Frey, E. Phase coexistence in driven one-dimensional transport. *Phys. Rev. Lett.* **2003**, *90*, 086601.
23. Billingsley, P. *Ergodic Theory and Information*; Wiley: Weinheim, Germany, 1965.

24. Balian, R. *Physique Statistique et Themodynamique Hors équilibre* (in French); Ecole Polytechnique: Palaiseau, France, 1992.
25. Latora, V.; Baranger, M.; Rapisarda, A.; Tsallis, C. The rate of entropy increase at the edge of chaos. *Phys. Lett. A* **2000**, *273*, 97–103.
26. Martyushev, L.M.; Seleznev, V.D. Maximum entropy production principle in physics, chemistry and biology. *Phys. Rep.* **2006**, *426*, 1–45.
27. Kirkaldy, J.S. Entropy criteria applied to pattern selection in systems with free boundaries. *Metall. Mater. Trans. A* **1985**, *16*, 1781–1797.
28. Lecomte, V.; Appert-Rolland, C.; van Wijland, F. Thermodynamic formalism for systems with Markov dynamics. *J. Stat. Phys.* **2007**, *127*, 51–106.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).

3.1 Discussion sur les résultats trouvés.

Revenons sur les principaux résultats trouvés dans cet article :

Premièrement nous établissons un lien numérique entre MEP et MKS. Ces deux principes sont à priori différents. En effet, l'entropie de Kolmogorov-Sinaï est reliée à la dynamique comme le montre les équations eq.(1.30) et eq.(1.31) alors que la production d'entropie macroscopique est liée à l'état du système eq.(1.3). De plus, l'entropie de Kolmogorov-Sinai est une quantité microscopique alors que la production d'entropie macroscopique (comme son nom l'indique) est macroscopique. De plus, l'article de [8] montre que MKS redonne les résultats classiques de la physique à l'équilibre et étend, de façon cohérente, ces résultats à la physique faiblement hors équilibre. Ce lien permet de mieux comprendre pourquoi l'application de MEP à la physique hors équilibre donne des résultats proches de ceux observés dans la réalité.

De plus, nous remarquons que l'application de MEP au modèle CASEP donne des résultats physiquement cohérents. En effet, lorsque le système est proche de l'équilibre, c'est à dire lorsque $\rho_a = \frac{\alpha}{\alpha+\gamma} \approx \rho_b = \frac{\delta}{\delta+\beta}$, l'état choisit par MEP est diffusif ($p* \approx \frac{1}{2}$) alors que plus le système est mis hors équilibre, plus l'état choisit par MEP devient convectif. Ceci reflète donc ce qui se passe dans la réalité : lorsqu'un faible gradient de température est imposé entre deux parties d'un fluide la diffusion impose une loi de Fick puis, lorsque le gradient devient plus important, des cellules convectives apparaissent laissant place à un régime diffuso-convectif.

Cependant, ces résultats sont numériques et non analytiques. La plus grande limitation du numérique étant la taille du système (le nombre de boites) qui ne peut guère excéder 20. En effet, 20 boites donne donc $2^{20} \approx 10^6$ états. La matrice de transition est alors une matrice de taille 10^6 ce qui commence à prendre beaucoup de temps de calcul... Il existe néanmoins de nombreux résultats analytiques sur le modèle ASEP [35] utilisant pour la plus part l'Ansatz de Bethe. On peut notamment trouver une expression analytique pour la densité moyenne de particules ou encore pour la densité stationnaire. Néanmoins, il est difficile de donner une expression simple de l'entropie de KS et de trouver où celle-ci est maximale...

Chapitre 3

Principe de maximisation de production d'entropie et Zero Range Process.

1 Introduction

Le lien établit entre MEP et MKS est un lien numérique. Nous avons cherché ce lien de façon analytique dans le modèle CASEP mais nous n'y sommes pas arrivés. Nous avons donc cherché un système similaire dans lequel les calculs seraient plus simple. Nous nous sommes aperçu que la difficulté des calculs dans le modèle ASEP (ou CASEP) résidait dans le phénomène d'exclusion. En effet, la fait qu'une particule puisse sauter dans la case voisine seulement si celle-ci est vide rend les calculs complexes. Nous avons donc opté pour une système Markovien analogue au système ASEP mais sans principe d'exclusion : Le Zero-Range Process (ZRP).

Dans une première partie nous détaillerons le ZRP et nous donnerons des exemples d'applications. Puis, nous introduirons le formalisme quantique qui nous permettra de donner une expression analytique de la distribution stationnaire mais aussi de la production d'entropie macroscopique ainsi que de l'entropie de Kolmogorov-Sinai. Nous verrons ensuite que ces deux quantités admettent des maxima et que ceux ci ont le même développement au premier ordre en la différence à l'équilibre. Enfin, nous verrons que pour une bonne valeur du coarse-graining l'entropie de KS et la production d'entropie admettent le même développement au second ordre. Nous discuterons enfin ces résultats.

2 ZRP

2.1 Description

Le ZRP, qui a été introduit initialement par Spitzer [36], est un système 1D constitué de N cases sur lequel des particules peuvent se déplacer. Tout comme pour le modèle ASEP il existe différents types de ZRP : le ZRP avec des conditions périodiques, le ZRP à bords ouverts... Dans le ZRP à bords ouverts, le système est

couplé avec deux réservoirs de particules : au niveau du bord gauche les particules peuvent entrer avec une probabilité α **quelque soit le nombre de particules dans la première case (il n'y a plus de principe d'exclusion)** et sortir avec une probabilité γw_n où w_n dépend du nombre de particules dans la première case. De même au niveau du bord droit les particules peuvent entrer avec une probabilité δ et sortir avec une probabilité βw_n. Pour les particules qui se trouvent à l'intérieur du réseau, si la case contient n particules, elles peuvent sauter à droite avec une probabilité pw_n et à gauche avec une probabilité qw_n. Les cases peuvent donc contenir un nombre quelconque de particules à la différence du modèle ASEP. La plupart des études ont étudié le ZRP avec des conditions périodiques [37, 38]. Cependant, nous voulons reproduire le modèle de Paltridge et avoir un système couplé avec des réservoirs. Nous nous intéresserons donc au ZRP à bords ouverts [39].

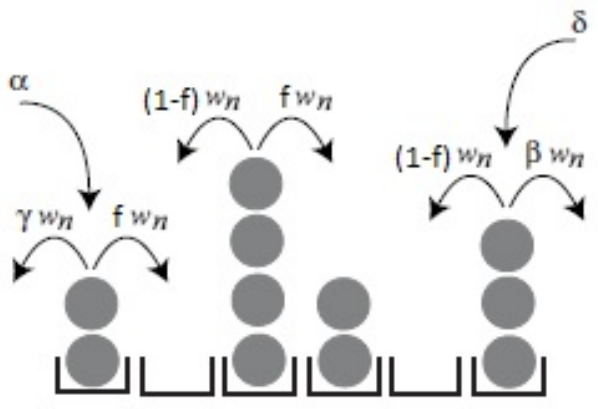


FIGURE 3.1 – Schéma représentant le modèle ZRP à bords ouverts. Cette figure est extraite de [39].

Avant une étude plus théorique du ZRP et du lien entre MEP et MKS commençons par faire un bref état de l'art sur le ZRP ainsi que ses différentes applications.

2.2 Les différentes applications du ZRP.

Le modèle ZRP représente de nombreux comportements liés au transport des particules. Il peut par exemple représenter la dynamique d'avalanche [40] mais aussi un modèle de transport plus récent : le modèle "des granulaires vibrés". Le système consiste en un récipient divisé en N compartiments contenant des granulaires. Les parois séparant les compartiments sont troués à la hauteur h ce qui permet aux particules de se déplacer d'un compartiment à l'autre.

Le modèle ZRP peut aussi représenter le trafic routier ainsi que la dynamique de l'évolution des interfaces, la dynamique des polymères et des verres...

Le modèle ZRP a aussi largement été étudié pour le phénomène de condensation [41]. En effet, en ajustant les paramètres du système (p q et w_n) certains sites vont accumuler un nombre de particules importants : il s'agit du phénomène de condensation. Cette condensation ressemble à la coalescence dans les milieux granulaires

ou encore à la condensation de Bose Einstein dans les réseaux [42].

3 ZRP et calculs analytiques

3.1 Formalisme quantique

Le nombre de particules dans chaque case n'étant pas borné, le ZRP est un système de Markov avec une infinité d'états. Chaque état peut être caractérisé par un N-uplet $n = (n_1, n_2, ..., n_N)$ où $\forall i \in 1, ..., N \; n_i \in \mathbb{N}$. Afin de calculer la probabilité stationnaire nous allons utiliser le formalisme quantique bien décrit dans l'article de [39]. Nous allons donc nous placer dans l'espace de Fock $(\mathbb{C}^\infty)^{\otimes L}$ dans lequel à chaque état n est associé un vecteur de la base $|n\rangle$. On définit alors sur cet espace le vecteur probabilité $|P\rangle$ définit par :

$$|P\rangle = \sum_n P_n |n\rangle . \tag{3.1}$$

Tout comme en physique quantique l'évolution de $|P\rangle$ est donnée par :

$$\frac{d}{dt} |P(t)\rangle = -H |P(t)\rangle , \tag{3.2}$$

où H est l'hamiltonien quantique. Pour $n \neq n'$ $H_{nn'}$ correspond à la probabilité de transition de l'état n vers l'état n'. Les termes diagonaux de H sont ajustés afin que la somme des termes sur une ligne vaille 1.

On peut alors écrire :

$$H = h_1 + h_L + \sum_{k=1}^{L-1} h_{k,k+1}, \tag{3.3}$$

où h_1 et h_L sont les opérateurs agissant sur les bords et $h_{k,k+1}$ est l'opérateur agissant entre la case k et la case $k+1$.

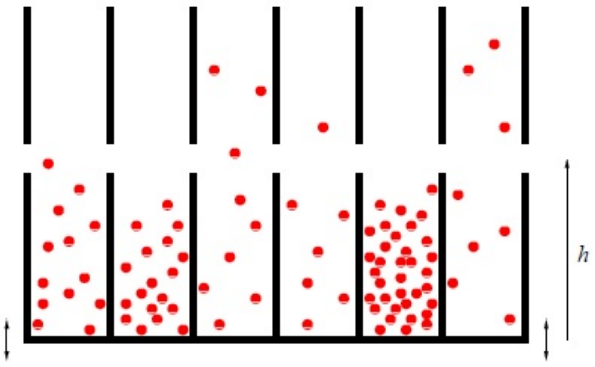


FIGURE 3.2 – Shéma représentant le modèle des "granulaires vibrés". On retrouve ici 6 compartiments ainsi que les parois trouées à la hauteur h. Cette figure est extraite de [37].

Plus exactement, définissons les opérateurs de création et d'annihilation a^+ et a^- qui peuvent être représentés par les matrices infinies :

$$a^+ = \begin{pmatrix} 0 & 0 & 0 & 0 & ... \\ 1 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & ... \end{pmatrix} \tag{3.4}$$

et

$$a^- = \begin{pmatrix} 0 & w_1 & 0 & 0 & ... \\ 0 & 0 & w_2 & 0 & ... \\ 0 & 0 & 0 & w_3 & ... \\ 0 & 0 & 0 & 0 & ... \end{pmatrix} \tag{3.5}$$

Nous définissons aussi la matrice infinie diagonale d par :

$$d = \begin{pmatrix} 0 & 0 & 0 & 0 & ... \\ 0 & w_1 & 0 & 0 & ... \\ 0 & 0 & w_2 & 0 & ... \\ 0 & 0 & 0 & w_3 & ... \end{pmatrix} \tag{3.6}$$

Enfin, nous définissons naturellement a_k^+ et a_k^- comme les opérateurs de création et d'annihilation liés au site k ainsi que d_k comme la matrice diagonale liée au site k.

Nous pouvons alors montrer que :

$$h_{k,k+1} = -p(a_k^- a_{k+1}^+ - d_k) - q(a_k^+ a_{k+1}^- - d_{k+1}), \tag{3.7}$$

$$h_1 = \alpha(Id - a_1^+) + \gamma(d_1 - a_1^-), \tag{3.8}$$

et

$$h_L = \delta(Id - a_L^+) + \beta(d_L - a_L^-). \tag{3.9}$$

Les équations eq.(3.7), eq.(3.8) et eq.(3.9) ainsi que l'équation eq.(3.3) nous donne l'expression du Hamiltonien.

Définissons alors le vecteur $|P_k^*)$ tel que $|P_k^*)_n = P^*(n_k = n)$ soit la probabilité de trouver n particules dans le site k. En suivant la démarche effectuée dans [39] nous définissons la fugacité z_k dans chaque case k comme une quantité liée à la densité de particules ρ_k (nous reviendrons sur ce point plus tard) et nous cherchons $|P_k^*)$ sous la forme :

$$P^*(n_k = n) = \frac{z_k^n}{Z_k} \prod_{i=1}^{n} w_i^{-1}, \tag{3.10}$$

où

$$Z_k = \sum_{n=0}^{\infty} z_k^n \prod_{i=1}^{n} w_i^{-1}. \tag{3.11}$$

Enfin, nous rechercherons la mesure stationnaire $|P^*\rangle$ comme le produit tensoriel des $|P_k^*)$:

$$|P^*\rangle = |P_1^*) \otimes |P_2^*) \otimes ... \otimes |P_N^*). \tag{3.12}$$

On vérifie aisément que :

$$a^+|P_k^*) = \frac{1}{z_k} d|P_k^*) \text{ et } a^-|P_k^*) = z_k|P_k^*). \tag{3.13}$$

Enfin, en utilisant 3.3 on obtient :

$$-H\,|P^*\rangle = \left(\sum_{k=1}^{L-1}(pz_k - qz_{k+1})(\frac{1}{z_{k+1}}d_{k+1} - \frac{1}{z_k}d_k) + (\alpha - \gamma z_1)(\frac{1}{z1}d_1 - Id) + (\delta - \beta z_L)(\frac{1}{z_L}d_L - Id)\right. \tag{3.14}$$

Nous verrons comment ce formalisme nous permet de calculer plutôt aisément les grandeurs caractéristiques du système ainsi que la mesure stationnaire.

3.2 Calcul de l'expression de la fugacité, du courant de particules...

On sait que $|P^*\rangle$ est une mesure stationnaire ssi $\frac{d}{dt}\,|P^*(t)\rangle = -H\,|P^*(t)\rangle = 0$ ce qui est en particulier vérifié si :

$$\forall\, k \in [|2, n-2|]\ pz_k - qz_{k+1} = \alpha - \gamma z_1 = \beta z_L - \delta = c, \tag{3.15}$$

où c est le courant de particules dans l'état stationnaire.
On peut réecrire 3.15

$$z_{k+1} = \frac{p}{q}z_k - \frac{c}{q}, \tag{3.16}$$

et donc (z_k) vérifie une relation arithmético-géométrique. On trouve donc, en posant $a = \frac{c}{q-p}$ que :

$$\forall k \in [|1, N|]\ z_k = (\frac{p}{q})^{k-1}(z_1 - a) + a. \tag{3.17}$$

Grâce à 3.15 et 3.17 on obtient alors :

$$c = (p-q)\frac{-\gamma\delta + \alpha\beta(\frac{p}{q})^{L-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{L-1}}. \tag{3.18}$$

Enfin, en réutilisant 3.17 on obtient que $\forall k \in [|1, N|]$

$$z_k = \frac{((\alpha+\delta)(p-q) - \alpha\beta + \gamma\delta)(\frac{p}{q})^{k-1} - \gamma\delta + \alpha\beta(\frac{p}{q})^{L-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{L-1}}. \tag{3.19}$$

Ainsi, connaissant z_k, on obtient une expression analytique de la probabilité stationnaire $|P*\rangle$ en fonction de α,β,γ,δ,p,q et L.

De plus, comme on l'avait mentionné au début de cette partie, la fugacité z_k est reliée à la densité de particules ρ_k par :

$$\rho_k = z_k \frac{\partial}{\partial z_k} \log(Z_k). \tag{3.20}$$

3.3 Quelques cas particuliers

Grâce à l'équation 3.18 on remarque que si :

$$\gamma\delta = \alpha\beta(\frac{p}{q})^{L-1}, \tag{3.21}$$

alors $c = 0$ et le système est à l'équilibre.

De plus, si $p = q$ c'est à dire si l'on se place dans le cas symétrique, alors l'expression de c ainsi que celle des z_k sont prolongeables par continuité et, par exemple, dans le cas $p = q = 1$ la fugacité est linéaire et vaut :

$$z_k = \frac{\alpha + \delta + \alpha\beta(L-1) - (\alpha\beta - \gamma\delta)(k-1)}{\beta + \gamma + \beta\gamma(L-1)}, \tag{3.22}$$

et le courant de particules vaut :

$$c = \frac{\alpha\beta - \gamma\delta}{\beta + \gamma + \beta\gamma(L-1)}. \tag{3.23}$$

Dans ce cas on retrouve ainsi une loi de Fick et un régime purement diffusif. Ceci est en parfait accord avec les résultats trouvés dans le chapitre précédent sur le modèle CASEP.

En outre, plus p devient différent de q plus le système devient convectif comme nous le verrons dans la partie suivante. Ainsi, pour mesurer l'écart au régime diffusif, et de manière analogue au travail effectué dans le modèle CASEP, nous pouvons poser $f = p - q$. Le but de la partie suivante va être d'exprimer la production d'entropie macroscopique σ ainsi que l'entropie de KS en fonction de f et des paramètres internes du système.

3.4 Production d'entropie et entropie de KS

Expression de la production d'entropie macroscopique

Dans cette partie nous allons considérer le ZRP sans interaction, c'est à dire que $w_i = w$ est constant. De plus, nous supposerons $p+q = 1$ et $p \leq q$ ce qui correspond à un flux de la gauche vers la droite. Tout comme dans l'équation eq.(1.3) la production d'entropie prend la forme :

$$\sigma = c(\log(\rho_1) - \log(\rho_N)). \tag{3.24}$$

En effet, la force thermodynamique entre la boite i et la boite $i+1$ s'écrit $X_i = \log(\rho_i) - \log(\rho_{i+1})$. Cette expression de la force thermodynamique provient d'une hypothèse de gaz parfait et est expliquée dans [12].

De plus, grâce aux équations 3.11 et 3.20 on trouve :

$$\rho_k = z_k \frac{\partial}{\partial z_k} \log(Z_k) = \frac{z_k}{1 - z_k}. \tag{3.25}$$

Et donc en utilisant 3.19 on obtient :

$$\sigma(f) = \frac{\alpha f}{f+\gamma}(\log(\frac{\alpha}{f+\gamma-\alpha}) - \log(\frac{(\alpha+\delta)f+\gamma\delta}{f(\beta-\alpha-\delta)+\beta\gamma-\gamma\delta})). \tag{3.26}$$

L'expression de $\sigma(f)$ est relativement simple. On remarque que puisque $f \geq 0$ alors σ est positive ssi $\rho_1 \geq \rho_N$ ce qui est physiquement cohérent car ceci correspond à un flux dans le sens contraire du gradient de densité de particules. De plus, lorsque $f = 0$ alors $\sigma(f) = 0$ et lorsque f devient assez grand alors $\rho_1(f) = \rho_N(f)$ et $\sigma(f) = 0$. On peut donc dire qu'entre ces deux valeurs σ admet un maximum noté $f_{max_{EP}}$

Expression de l'entropie de Kolmogorov-Sinai

On rappelle que pour une chaine de Markov l'entropie de KS h_{KS} prend la forme :

$$h_{KS} = -\sum_{(i,j)} \mu_{i_{stat}} p_{ij} \log(p_{ij}). \tag{3.27}$$

Ainsi, pour le ZRP nous pouvons écrire :

$$\begin{aligned} h_{KS} &= -\sum_i \mu_{i_{stat}} \sum_j p_{ij} \log(p_{ij}) \\ &= -\sum_{m_1=0}^{+\infty} ... \sum_{m_L=0}^{+\infty} \\ &\quad P(m_1, m_2, ..., m_L) \sum_j p_{(m_1,...,m_L)\to j} \log(p_{(m_1,...,m_L)\to j}) \\ &= -\sum_{m_1=0}^{+\infty} P(m_1) ... \sum_{m_L=0}^{+\infty} P(m_L) \\ &\quad \sum_j p_{(m_1,...,m_L)\to j} \log(p_{(m_1,...,m_L)\to j}). \end{aligned} \tag{3.28}$$

Nous devons donc calculer $\sum_j p_{(m_1,...,m_L)\to j} \log(p_{(m_1,...,m_L)\to j})$ que nous noterons $(*)$. Par simplicité de calcul nous prendrons $p+q = \alpha+\delta = \beta+\gamma = 1$ and $dt = \frac{1}{L}$ afin de négliger la probabilité de rester sur un même site comparé à celle de sauter sur un autre site. Nous devons considérer cinq cas différents :

1. if $\forall i\ m_i \geq 1$ les transitions possibles sont :
   $(m_1, m_2, ..., m_L) \to (m_1 \pm 1, m_2, ..., m_L)$ avec probabilité respectivement α et δ
   $(m_1, m_2, ..., m_L) \to (m_1, m_2, ..., m_L \pm 1)$ avec probabilité respectivement γ et β
   et $(m_1, ..., m_k, ..., m_L) \to (m_1, ..., m_k \pm 1, ..., m_L)$ avec probabilité respectivement p et q

   Donc,

$$(*) = \alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (L-1)(p \log(p) + q \log(q)). \quad (3.29)$$

2. si $m_1 \geq 1$ et $m_L \geq 1$ et si i désigne le nombre de m_i compris entre 2 et $L-1$ égal à 0. Avec le même argument que précédemment nous obtenons :

$$(*) = \alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (L-1-i)(p \log(p) + q \log(q)) \quad (3.30)$$

3. si $m_1 = 0$ et $m_L \geq 1$ et si i désigne le nombre de m_i entre 2 et $L-1$ égal à 0 nous obtenons :

$$(*) = \alpha \log \alpha + \delta \log \delta + \beta \log \beta + (L-2-i)p \log(p) + (L-1-i)q \log(q). \quad (3.31)$$

4. En appliquant le même raisonnement si $m_1 \geq 1$ et $m_L = 0$ et si i désigne le nombre de m_i compris entre 2 et $L-1$ égal à 0 nous obtenons :

$$(*) = \alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + (L-1-i)p \log(p) + (L-2-i)q \log(q). \quad (3.32)$$

5. finalement, si $m_1 = 0$ et si $m_L = 0$ et si i désigne le nombre de m_i entre 2 et $L-1$ égal à 0 nous obtenons :

$$(*) = \alpha \log \alpha + \delta \log \delta + (L-2-i)(p \log(p) + q \log(q). \quad (3.33)$$

De plus, en utilisant les équations 3.10 et 3.11 on obtient :

$$p_k(n_k = m) = (1 - z_k) z_k^m. \quad (3.34)$$

Et donc d'après 3.34 nous trouvons que : $P(m_k = 0) = 1 - z_k$ et $\sum_{i=1}^{+\infty} P(m_k = i) = z_k$.

Nous obtenons donc h_{KS} est la somme de cinq termes :

$$\begin{aligned} h_{KS} = &-(\alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (L-1)(p \log(p) + q \log(q)) \\ &+ (p \log(p) + q \log(q))(\sum_{r=0}^{L} r \sum_{i_1 \ldots i_L} \prod_{i=i_1,\ldots i_r} (1 - z_i) \prod_{i \neq i_1 \ldots i_r} z_i) \\ &+ (\gamma \log(\gamma) + p \log(p)) z_L (1 - z_1)(\sum_{i_2 \ldots i_{L-1}} \prod_{i=i_2,\ldots i_r} (1 - z_i) \prod_{i \neq i_2 \ldots i_r} z_i) \\ &+ (\beta \log(\beta) + q \log q) z_1 (1 - z_L)(\sum_{i_2 \ldots i_{L-1}} \prod_{i=i_2,\ldots i_r} (1 - z_i) \prod_{i \neq i_2 \ldots i_r} z_i) \\ &+ (\beta \log(\beta) + \gamma \log \gamma + p \log p + q \log q)(\sum_{i_2 \ldots i_{L-1}} \prod_{i=i_2,\ldots i_r} (1 - z_i) \prod_{i \neq i_2 \ldots i_r} z_i) \quad (3.35) \end{aligned}$$

Cette expression qui parait compliquée à première vue peut être simplifiée. En effet, en considérant la fonction $F(a) = \prod_1^L (z_k + a(1 - z_k))$ et en la dérivant par rapport à a nous obtenons la relation suivante :

$$\sum_{r=0}^{L} r \sum_{i_1 \dots i_L} \prod_{i=i_1,\dots i_r} (1 - z_i) \prod_{i \neq i_1 \dots i_r} z_i = \sum_{i=1}^{L} (1 - z_i). \tag{3.36}$$

Ainsi, nous pouvons simplifier l'expression de h_{KS} et nous obtenons :

$$\begin{aligned} h_{KS} = -(&\alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (L-1)(p \log(p) + q \log(q)) \\ &+ (p \log(p) + q \log(q)) \sum_{i=1}^{L} (1 - z_i) \\ &+ (\gamma \log(\gamma) + p \log(p))(1 - z_1) \\ &+ (\beta \log(\beta) + q \log(q))(1 - z_L). \end{aligned} \tag{3.37}$$

Ainsi grâce à l'expression des z_i trouvée 3.22 nous pouvons exprimer h_{KS} relativement assez simplement en fonction des paramètres du système.

Nous verrons dans l'article que h_{KS} admet un maximum en fonction de f noté $f_{max_{KS}}$ et que celui-ci, sous certaines conditions, est proche de $f_{max_{EP}}$.

4 Article

Nonlin. Processes Geophys., 22, 187–196, 2015
www.nonlin-processes-geophys.net/22/187/2015/
doi:10.5194/npg-22-187-2015
© Author(s) 2015. CC Attribution 3.0 License.

Nonlinear Processes in Geophysics
Open Access

Statistical optimization for passive scalar transport: maximum entropy production versus maximum Kolmogorov–Sinai entropy

M. Mihelich, D. Faranda, B. Dubrulle, and D. Paillard

Laboratoire SPHYNX, CEA/IRAMIS/SPEC, CNRS URA 2464, 91191 Gif-sur-Yvette, France

Correspondence to: B. Dubrulle (berengere.dubrulle@cea.fr)

Received: 18 September 2014 – Published in Nonlin. Processes Geophys. Discuss.: 18 November 2014
Revised: 20 February 2015 – Accepted: 28 February 2015 – Published: 25 March 2015

Abstract. We derive rigorous results on the link between the principle of maximum entropy production and the principle of maximum Kolmogorov–Sinai entropy for a Markov model of the passive scalar diffusion called the Zero Range Process. We show analytically that both the entropy production and the Kolmogorov–Sinai entropy, seen as functions of a parameter f connected to the jump probability, admit a unique maximum denoted $f_{\max_{EP}}$ and $f_{\max_{KS}}$. The behaviour of these two maxima is explored as a function of the system disequilibrium and the system resolution N. The main result of this paper is that $f_{\max_{EP}}$ and $f_{\max_{KS}}$ have the same Taylor expansion at first order in the deviation from equilibrium. We find that $f_{\max_{EP}}$ hardly depends on N whereas $f_{\max_{KS}}$ depends strongly on N. In particular, for a fixed difference of potential between the reservoirs, $f_{\max_{EP}}(N)$ tends towards a non-zero value, while $f_{\max_{KS}}(N)$ tends to 0 when N goes to infinity. For values of N typical of those adopted by Paltridge and climatologists working on maximum entropy production ($N \approx 10$–100), we show that $f_{\max_{EP}}$ and $f_{\max_{KS}}$ coincide even far from equilibrium. Finally, we show that one can find an optimal resolution N_* such that $f_{\max_{EP}}$ and $f_{\max_{KS}}$ coincide, at least up to a second-order parameter proportional to the non-equilibrium fluxes imposed to the boundaries. We find that the optimal resolution N^* depends on the non-equilibrium fluxes, so that deeper convection should be represented on finer grids. This result points to the inadequacy of using a single grid for representing convection in climate and weather models. Moreover, the application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system.

1 Introduction

A major difficulty in the modelling of nonlinear geophysical or astrophysical processes is the taking into account of all the relevant degrees of freedom. For example, fluid motions obeying Navier–Stokes equations usually require of the order of $N = Re^{9/4}$ modes to faithfully describe all scales between the injection scale and the dissipative scale (Frisch, 1995). In atmosphere, or ocean, where the Reynolds number exceeds 10^9, this amounts to $N = 10^{20}$, a number too large to be handled by any existing computers (Wallace and Hobbs, 2006). The problem is even more vivid in complex systems such as planetary climate, where the coupling of litho- bio- and cryosphere with ocean and atmosphere increases the number of degrees of freedom beyond any practical figure. This justifies the long historical tradition of parametrization and statistical model reduction, to map the exact equations describing the system onto a set of simpler equations involving few degrees of freedom. The price to pay is the introduction of *free parameters*, describing the action of discarded degrees of freedom, that need to be prescribed.

When the number of free parameters is small, their prescription can be successfully done empirically through calibrating experiments or by a posteriori tuning (Rotstayn, 2000). When the number of parameters is large, such as in climate models where it reaches several hundreds (Murphy et al., 2004), such empirical procedure is inapplicable, because it is impossible to explore the whole parameter space. In that respect, it is of great interest to explore an alternative road to parametrization via application of a statistical optimization principle, such as minimizing or maximizing of a suitable cost functional. As discussed by Turkington (2013) and Pascale et al. (2012), this strategy usually leads to closed reduced equations with adjustable parameters in the closure

appearing as weights in the cost functional that can be computed explicitly. A renowned example in climate is given by a principle of maximum entropy production (MEP) that allowed (Paltridge, 1975) to derive the distribution of heat and clouds at the Earth surface with reasonable accuracy, without any parameters and with a model of a dozen of degrees of freedom (boxes). Since then, refinements of the Paltridge model have been suggested to increase its generality and range of prediction (Herbert et al., 2011). MEP states that a stationary non-equilibrium system chooses its final state in order to maximize the entropy production as is explained in Martyushev and Seleznev (2006). Rigorous justifications of its application have been searched using e.g. information theory (Dewar and Maritan, 2014) without convincing success. More recently, we have used the analogy of the climate box model of Paltridge with the asymmetric exclusion Markov process (ASEP) to establish numerically a link between the MEP and the principle of maximum Kolmogorov–Sinai entropy (MKS) (Mihelich et al., 2014). The MKS principle is a relatively new concept which extends the classical results of equilibrium physics (Monthus, 2011). This principle applied to Markov chains provides an approximation of the optimal diffusion coefficient in transport phenomena (Gómez-Gardeñes and Latora, 2008) or simulates random walk on irregular lattices (Burda et al., 2009). It is therefore a good candidate for a physically relevant cost functional in passive scalar modelling.

The goal of the present paper is to derive rigorous results on the link between MEP and MKS using a Markov model of the passive scalar diffusion called the Zero Range Process (Andjel, 1982). We find that there exists an optimal resolution N_* such that both maxima coincide to second order in the distance from equilibrium. The application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system. This suggests that the MEP and MKS principle may be unified when the Kolmogorov–Sinai entropy is defined on opportunely coarse-grained partitions.

2 From passive scalar equation to ZRP model

The equation describing the transport of a passive scalar like temperature in a given velocity field $u(x,t)$ reads as

$$\partial_t T + u\partial_x T = \kappa \partial_x^2 T, \tag{1}$$

with appropriate boundary conditions, or equivalently, in non-dimensional form:

$$\partial_t T + u\partial_x T = \frac{1}{RePr}\partial_x^2 T, \tag{2}$$

where κ, Re and Pr are respectively the molecular diffusivity, the Reynolds and the Prandtl number. To solve this equation, one must know both the velocity field and the boundary conditions, and use as many number of modes as necessary to describe the range of scales up to the scales at which molecular diffusivity takes place, i.e. roughly $(RePr)^{3/2}$ modes, where Re is the Reynolds number of the convective flow, and Pr is its Prandtl number. In geophysical flows, this number is too large to be handled even numerically (Troen and Mahrt, 1986). Moreover, in typical climate studies, the velocity flow is basically unknown as it must obey a complicated equation involving the influence of all the relevant climate components. In order to solve the equation, one must necessarily prescribe the heat flux $f = -uT + \kappa\nabla T$. The idea of Paltridge was then to discretize the passive scalar equation in boxes and prescribe the heat flux $f_{i(i+1)}$ between boxes i and $i+1$ by maximizing the associated thermodynamic entropy production $\dot{S} = \sum_i f_{i(i+1)}(\frac{1}{T_{i+1}} - \frac{1}{T_i})$. Here, we slightly modify the Paltridge discretization approximation to make it amenable to rigorous mathematical results on Markov chains. For simplicity, we stick to a one-dimensional case (corresponding to boxes varying only in latitude) and impose the boundary conditions through two reservoirs located at each end of the chain (mimicking the solar heat flux at pole and equator). We consider a set of N boxes that can contain an arbitrary number $n \in \mathbb{N}$ of particles. We then allow transfer of particles in between two adjacent boxes via decorrelated jumps (to the right or to the left) following a 1-D Markov dynamics governed by a coupling with the two reservoirs imposing a difference of chemical potential at the ends. The resulting process is called the Zero Range Process (ZRP) (Andjel, 1982). The different jumps are described as follows. At each time step a particle can jump right with probability pw_n or jump left with probability qw_n where w_n is a parameter depending of the number of particles inside the box. Physically it represents the interactions between particles. At the edges of the lattice the probability rules are different: at the left edge a particle can enter with probability α and exit with probability γw_n whereas at the right edge a particle can exit with probability βw_n and enter with probability δ. Choices of different w_n give radically different behaviours. For example $w_n = 1 + b/n$ where $b \geq 0$ described condensation phenomena (Großkinsky et al., 2003) whereas $w_1 = w$ and $w_n = 1$ if $n \geq 2$ has been used to modelled road traffic. We consider in this paper the particular case where $w = 1$ by convenience of calculation. Moreover, without loss of generality we will take $p \geq q$ which corresponds to a particle flow from the left to the right and denote $f = p - q$. After a sufficiently long time the system reaches a non-equilibrium steady state. The interest of this toy model is that it is simple enough so that exact computations are analytically tractable.

Taking the continuous limit of this process, it may be checked that the fugacity z, which is a quantity related to the average particle density (see Eq. 9 below) of stationary solutions of a system consisting of boxes of size $\frac{1}{N}$, follows

the continuous equation (Levine et al., 2005) :

$$f\frac{\partial z}{\partial x} - \frac{1}{2N}\frac{\partial^2 z}{\partial x^2} = 0, \tag{3}$$

corresponding to a stationary solution of a non-dimensional passive scalar equation with non-dimensional velocity f and a non-dimensional diffusivity $\frac{1}{RePr} = \frac{1}{2N}$. Therefore, the fugacity of the Zero Range Process is a passive scalar obeying a convection–diffusion equation, with advection velocity controlled by the probability to jump to the right or to the left, and diffusivity controlled by the number of boxes: the larger the number of boxes (the finer the resolution) the smaller the diffusivity. This observation illuminates the well-known observation that the numerical diffusion of a discrete model of diffusion is inversely proportional to the resolution. The parameter f controls the regime: $f = 0$ corresponds to a purely conductive regime whereas the larger f the more convective the regime. In the sequel, we calculate the entropy production and the Kolmogorov–Sinai entropy function of f. These two quantities reach a maximum denoted respectively $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$. The MEP principle (resp. the MKS principle) states that the system will choose $f = f_{\max_{\mathrm{EP}}}$ (resp $f = f_{\max_{\mathrm{KS}}}$).

We will first show that numerically $f_{\max_{\mathrm{EP}}} \approx f_{\max_{\mathrm{KS}}}$ even far from equilibrium for a number of boxes N roughly corresponding to the resolution taken by Paltridge (1975) in his climate model. This result is similar to what we found for the ASEP model (Mihelich et al., 2014) and thus gives another example of a system in which the two principles are equivalent. Moreover we will see analytically that $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$ have the same behaviour in first order in the difference of the chemical potentials between the two reservoirs for N large enough. These results provide a better understanding of the relationship between the MEP and the MKS principles.

3 Notation and useful preliminary results

This Markovian process is a stochastic process with an infinite number of states in bijection with $\mathbb{N}^N$. In fact, each state can be written $n = (n_1, n_2, \ldots, n_N)$ where n_i is the number of particle lying in site i. We call P_n the stationary probability to be in state n. In order to calculate this probability it is easier to use a quantum formalism than the Markovian formalism (Domb, 2000; Levine et al., 2005).

The probability to find m particles in the site k is equal to $p_k(n_k = m) = \frac{z_k^m}{Z_k}$ where Z_k is the analogue of the grand canonical repartition function and z_k is the fugacity between 0 and 1. Moreover $Z_k = \sum_{i=0}^{\infty} z_k^i = \frac{1}{1-z_k}$. So, finally

$$p_k(n_k = m) = (1 - z_k)z_k^m. \tag{4}$$

We can show that the probability P over the states is the tensorial product of the probability p_k over the boxes:

$$P = p_1 \otimes p_2 \otimes \cdots \otimes p_N.$$

Thus events $(n_k = m)$ and $(n'_k = m')$ for $k \neq k'$ are independent and so

$$P(m_1, m_2, \ldots, m_N) = p_1(n_1 = m_1) * \cdots * p_N(n_N = m_N). \tag{5}$$

So finally

$$P(m_1, m_2, \ldots, m_N) = \prod_{k=1}^{N} (1 - z_k) z_k^{m_k}. \tag{6}$$

Moreover, with the Hamiltonian equation found from the quantum formalism we can find the exact values of z_k function of the system parameters:

$$z_k = \frac{(\frac{p}{q})^{k-1}[(\alpha+\delta)(p-q) - \alpha\beta + \gamma\delta] - \gamma\delta + \alpha\beta(\frac{p}{q})^{N-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{N-1}}, \tag{7}$$

and the flux of particles c:

$$c = (p-q)\frac{-\gamma\delta + \alpha\beta(\frac{p}{q})^{N-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{N-1}}. \tag{8}$$

Finally, the stationary density is related to the fugacity by the relation

$$\rho_k = z_k \frac{\partial \log Z_k}{\partial z_k} = \frac{z_k}{1 - z_k}. \tag{9}$$

3.1 Entropy production

For a system subject to internal forces X_i and associated fluxes J_i the macroscopic entropy production is well known (Onsager, 1931) and takes the form

$$\sigma = \sum_i J_i \cdot X_i.$$

The physical meaning of this quantity is a measure of irreversibility: the larger σ the more irreversible the system.

In the case of the ZRP, irreversibility is created by the fact that $p \neq q$. We will parametrize this irreversibility by the parameter $f = p - q$ and we will take $p + q = 1$. In the remainder of the paper, we take, without loss of generality, $p \leq q$ which corresponds to a flow from left to right. Moreover, the only flux to be considered here is the flux of particles c and the associated force is due to the gradient of the density of particles $\rho : X = \nabla \log \rho$ (Balian, 1992).

Thus, when the stationary state is reached, i.e. when c is constant,

$$\sigma = \sum_{i=1}^{N-1} c.(\log(\rho_i) - \log(\rho_{i+1})) = c \cdot (\log(\rho_1) - \log(\rho_N)). \tag{10}$$

Thus, according to Eqs. (7), (8), (9) and (10) when N tends to $+\infty$ we obtain:

$$\sigma(f) = \frac{\alpha f}{f+\gamma}\left(\log\left(\frac{\alpha}{f+\gamma-\alpha}\right) - \log\left(\frac{(\alpha+\delta)f + \gamma\delta}{f(\beta-\alpha-\delta) + \beta\gamma - \gamma\delta}\right)\right). \tag{11}$$

Because $f \geq 0$ the entropy production is positive if and only if $\rho_1 \geq \rho_N$ and iff $z_1 \geq z_N$. This is physically coherent be-

cause fluxes are in the opposite direction of the gradient. We remark that if $f = 0$ then $\sigma(f) = 0$. Moreover, when f increases $\rho_1(f)$ decreases and $\rho_2(f)$ increases until they take the same value. Thus there exists f, large enough, for which $\sigma(f) = 0$. Between these two values of f the entropy production has at least one maximum.

3.2 Kolmogorov–Sinai entropy

There are several ways to introduce the Kolmogorov–Sinai entropy which is a mathematical quantity introduced by Kolmogorov and developed by the renowned mathematicians Sinai and Billingsley (Billingsley, 1965). Nevertheless, for a Markov process we can give it a simple physical interpretation: the Kolmogorov–Sinai entropy is the time derivative of the Jaynes entropy (entropy over the path):

$$S_{\text{Jaynes}}(t) = -\sum_{\Gamma_{[0,t]}} p_{\Gamma_{[0,t]}} \log(p_{\Gamma_{[0,t]}}), \tag{12}$$

For a Markov Chain we thus have

$$S_{\text{Jaynes}}(t) - S_{\text{Jaynes}}(t-1) = -\sum_{(i,j)} \mu_{i_{\text{stat}}} p_{ij} \log(p_{ij}), \tag{13}$$

where $\mu_{\text{stat}} = \mu_{i_{\text{stat}}} i = 1, \ldots, N$ is the stationary measure and where the p_{ij} are the transition probabilities.

Thus the Kolmogorov–Sinai entropy takes the following form:

$$h_{\text{KS}} = -\sum_{(i,j)} \mu_{i_{\text{stat}}} p_{ij} \log(p_{ij}). \tag{14}$$

For the ZRP, we show in the Appendix that it can be written as

$$\begin{aligned} h_{\text{KS}} = &-(\alpha \log\alpha + \delta \log\delta + \gamma \log\gamma + \beta \log\beta \\ &+ (N-1)(p\log(p) + q\log(q))) + (p\log(p) + q\log(q)) \\ &\sum_{i=1}^{N}(1 - z_i) + (\gamma\log(\gamma) + p\log(p))(1 - z_1) \\ &+ (\beta\log(\beta) + q\log(q))(1 - z_N). \end{aligned} \tag{15}$$

4 Results

We start by pointing to some interesting properties of $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$, then present numerical experiments on the ZRP model and conclude with some analytical computations.

Note that for N, α, β, γ and δ fixed the entropy production as well as the Kolmogorov–Sinai entropy seen as functions of f both admit a unique maximum. When N tends to infinity and $f = 0$, using Eq. (7) (i.e. the symmetric case), we find that $z_1 = \frac{\alpha}{\gamma}$ and $z_N = \frac{\delta}{\beta}$. Thus, the system is coupled with two reservoirs with respective chemical potential $\frac{\alpha}{\gamma}$(left) and $\frac{\delta}{\beta}$ (right). For $\frac{\alpha}{\gamma} \neq \frac{\delta}{\beta}$ the system is out of equilibrium. We assume, without loss of generality, $z_1 \geq z_N$, which corresponds to a flow from left to right. As a measure of deviation from equilibrium we take $s = z_1 - z_N$: the larger s, the more density fluxes we expect into the system.

First we remark that $f_{\max_{\text{EP}}}$ hardly depends on N whereas $f_{\max_{\text{KS}}}$ depends strongly on N. This is easily understood because σ depends only on z_1 and z_N whereas h_{KS} depends on all the z_i. Moreover, the profile of the z_i depends strongly on N. In particular, for a fixed difference of potential between the reservoirs, $f_{\max_{\text{EP}}}(N)$ tends towards a non-zero value, while $f_{\max_{\text{KS}}}(N)$ tends to 0 when N goes to infinity.

Moreover, $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$ coincide even far from equilibrium for N corresponding to the choice of Paltridge (1975) $N \approx 10 \sim 100$. For N fixed, as large as one wants, and for all ϵ, as small as one wants, there exists ν such that for all $s \in [0; \nu]$ $|f_{\max_{\text{EP}}} - f_{\max_{\text{KS}}}| \leq \epsilon$.

These observations are confirmed by the results presented in Figs. 1 and 3 where EP and KS are calculated using Eqs. (7) and (15) for $s = 0.13$ and three different partitions: $N = 20$, $N = 100$ and $N = 1000$. The figure shows that $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$ coincide with good approximation for $N = 20$ and $N = 100$. But then when N increases $f_{\max_{\text{KS}}}(N)$ tends to 0 whereas $f_{\max_{\text{EP}}}(N)$ tends to a non-zero value.

In Fig. 2 we represent the entropy production (top) and KS entropy (bottom) as functions of f for $N = 1000$ and for three value of s: $s = 0.13$; $s = 0.2$; $s = 0.04$. This supports the claim that for N fixed, we can try different values of s such that $s \in [0; \nu]$ $|f_{\max_{\text{EP}}} - f_{\max_{\text{KS}}}| \leq \epsilon$. Figure 3 shows that $\Delta f_{\max}$ is minimum when the system is close to equilibrium whereas the further the system is from equilibrium (when s increases) the more $\Delta f_{\max}$ increases. Moreover, the optimal resolution where $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$ coincide is approximately 10–100. Then $\Delta f_{\max}$ is maximum at $N = 500$ and $s = 0.05$. $\Delta f_{\max}$ is obviously linear in s, for small values of s, but the behaviour with N is more complicated.

Such numerical investigations help to understand why $f_{\max_{\text{KS}}}(N)$ and $f_{\max_{\text{EP}}}(N)$ have different behaviour functions of N, and why for N large enough $f_{\max_{\text{KS}}}$ and $f_{\max_{\text{EP}}}$ have the same behaviour of first order in the deviation from equilibrium measured by the parameter s. We will see that we can get a precise answer to such questions by doing calculations and introducing a sort of hydrodynamics approximation.

4.1 Taylor expansion

From Eq. (15) it is apparent that $f_{\max_{\text{KS}}}$ depends on N whereas from Eq. (10) we get that $f_{\max_{\text{EP}}}$ hardly depends on N. Indeed there is a difference between $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$, i.e. a difference between the two principles for the ZRP. Nevertheless, we have seen numerically that there is a range of N, namely $N \approx$10–100 for which the maxima fairly coincide.

Using Eqs. (15), (7), and (11) we compute analytically the Taylor expansion of $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$ in s. We will show the main result: $f_{\max_{\text{EP}}}$ and $f_{\max_{\text{KS}}}$ have the same Taylor expansion in first order in s for N large enough. Their Taylor

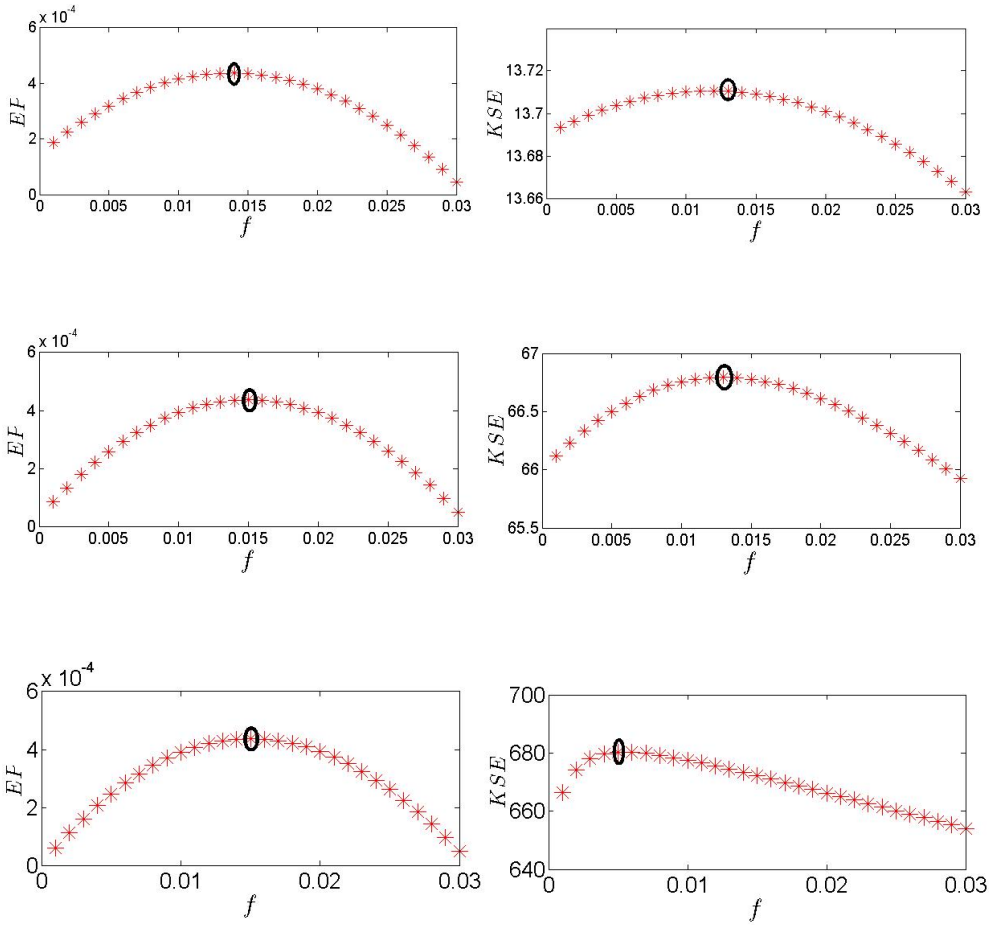


Figure 1. Entropy production calculated using Eq. (11) (left) and KS entropy calculated using Eqs. (7) and (15) (right) as functions of f for $s = 0.13$ and respectively $N = 20$, $N = 100$ and $N = 1000$.

expansions are different up to the second order in s but there exists an N, i.e. a resolution, such that $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$ coincident up to the second order.

Let us start by computing $f_{\max_{\mathrm{KS}}}$. It does not depend of the constant terms of h_{KS} in Eq. (15) and therefore we need only concern ourselves with

$$\begin{aligned}&-(p\log(p)+q\log(q))\left(\sum_{i=1}^{N}(z_i)-1\right)+(\gamma\log(\gamma)\\&+p\log(p))(1-z_1)+(\beta\log(\beta)+q\log(q))(1-z_N)\\&=N\cdot H(f,N,\alpha,\gamma,\beta,\delta).\end{aligned}\tag{16}$$

Using Eq. (7), the expression of $H(f,N,\alpha,\gamma,\beta,\delta)$ takes an easy form. To simplify the calculations, we restrict the space of parameter by assuming $\alpha+\gamma=1$ and $\beta+\delta=1$ and we parametrize the deviation from equilibrium by the parameter $\overline{s}=\alpha-\delta$. Moreover let us denote $a=\frac{1}{N}$. Thus, we have $H(f,N,\alpha,\gamma,\beta,\delta)=H(f,a,\alpha,\overline{s})$. In order to know the Taylor expansion to the first order in $\overline{s}$ of $f_{\max_{\mathrm{KS}}}$ we develop $H(f,a,\alpha,\overline{s})$ up to the second order in f; i.e. we have $H(f,a,\alpha,\overline{s})=C+Bf+Af^2+o(f^2)$. Then we find $f_{\max_{\mathrm{KS}}}=-B/2A$ which we will develop in the power of $\overline{s}$. This is consistent if we assume $f\ll a$.

After some tedious but straightforward calculations, we get at the first order in $\overline{s}$:

$$f_{\max_{\mathrm{KS}}}(\overline{s})=\frac{1}{4}\frac{(1-\alpha)-a(\alpha+2)}{\alpha(1-\alpha)+2a\alpha(\alpha-1)}\overline{s}+o(\overline{s})\tag{17}$$

and so

$$f_{\max_{\mathrm{KS}}}(\overline{s})=\frac{1}{4\alpha}\overline{s}+\frac{3a}{4(\alpha-1)}\overline{s}+o(\overline{s})+o(a\overline{s}).\tag{18}$$

We repeat the same procedure starting from Eq. (11) and we obtain

$$f_{\max_{\mathrm{EP}}}(\overline{s})=\frac{\overline{s}}{4\alpha}+o(\overline{s})+o(a).\tag{19}$$

Thus, since $a=\frac{1}{N}\ll 1$ the behaviour of $f_{\max_{\mathrm{KS}}}(\overline{s})$ and $f_{\max_{\mathrm{EP}}}(\overline{s})$ is the same for $\overline{s}$ small enough.

We remark that we can strictly find the same result by solving the hydrodynamics continuous approximation given by Eq. (3). This equation is a classical convection–diffusion

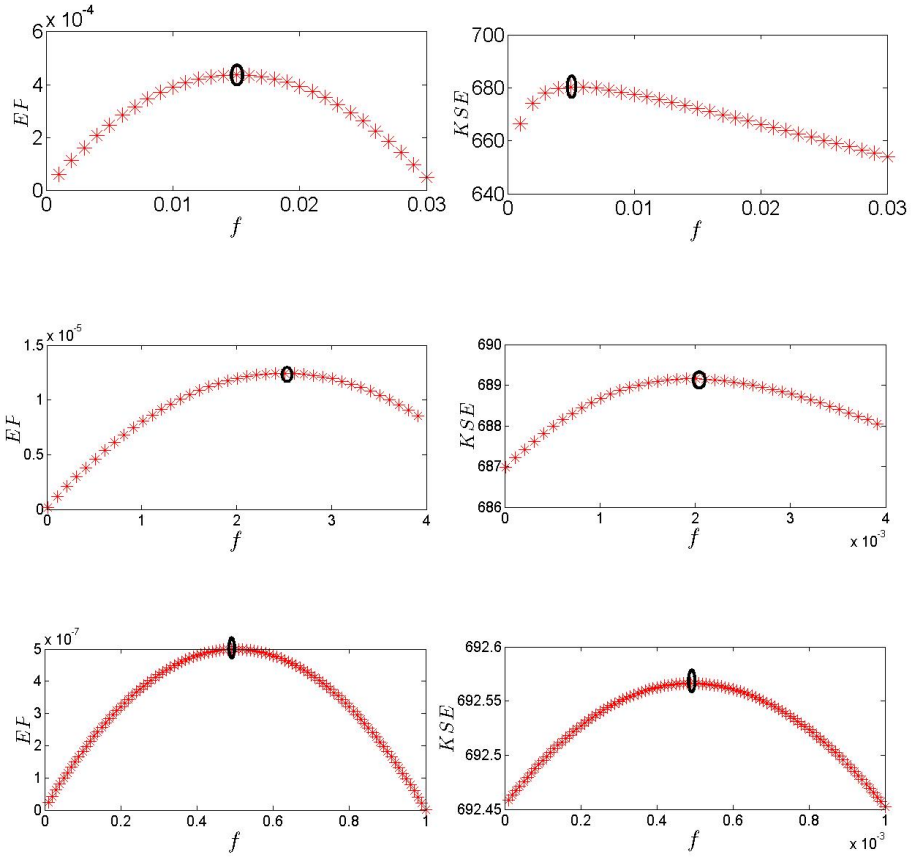


Figure 2. Entropy production (left) and KS entropy (right) as functions of f for $N = 1000$ and respectively $s = 0.13$, $s = 0.2$ and $s = 0.04$.

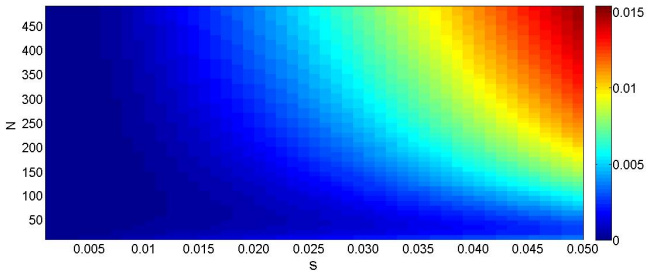


Figure 3. 2-D plot representing $\Delta f_{\max} = f_{\max_{\mathrm{EP}}} - f_{\max_{\mathrm{KS}}}$ in the (N, s) space.

equation. We remark that, by varying f, we change the convective behaviour: $f = 0$ corresponds to a purely diffusive regime whereas by increasing f we enhance the role of convection. If the system is near equilibrium then $f_{\max_{\mathrm{EP}}} \approx f_{\max_{\mathrm{KS}}} \approx 0$ and the system is purely diffusive. When the system is out of equilibrium $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$ are different from 0 and the system corresponds to an (optimal) trade-off between purely diffusive and convective behaviour.

One can verify this numerically: we first calculate the exact values of the entropy production function of f using Eq. (7) and the Kolmogorov–Sinai entropy function of f using Eqs. (7) and (15). Then we approximate these two curves with a cubic spline approximation in order to find $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$.

In order to find the optimal resolution N_* we can go one step further by expanding $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$ up to the second order in $\overline{s}$:

$$f_{\max_{\mathrm{EP}}}(\overline{s}) = \frac{\overline{s}}{4\alpha} + \frac{\overline{s}^2(\alpha+1)}{8\alpha^2(\alpha-1)} + o(\overline{s^2}) + o(a), \tag{20}$$

$$f_{\max_{\mathrm{KS}}}(\overline{s}) = \frac{1}{4}\frac{(1-\alpha)-a(\alpha+2)}{\alpha(1-\alpha)+2a\alpha(\alpha-1)}\overline{s} + \frac{(1-\alpha)^2+a(\alpha^2-2\alpha+1)}{8\alpha^2(\alpha-1)^2(1-2a)}\overline{s}^2 + o(\overline{s}^2). \tag{21}$$

Thus, $f_{\max_{\mathrm{EP}}}$ and $f_{\max_{\mathrm{KS}}}$ coincide in second order in $\overline{s}$ iff a satisfies the quadratic equation

$$(4\alpha - 6\alpha^2 + 6\alpha^3 - 4\overline{s} + 3\alpha^2\overline{s})a^2 - \frac{1}{2}(8\alpha - 8\overline{s} + 3\alpha^2\overline{s} - 6\alpha^2 + 6\alpha^3)a - (1-\alpha) = 0. \tag{22}$$

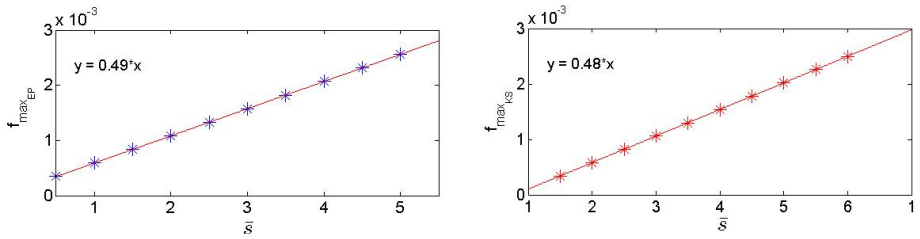


Figure 4. $f_{\max_{\mathrm{EP}}}$ (left) and $f_{\max_{\mathrm{KS}}}$ (right) as functions of $\overline{s}$ for $\alpha = 0.5$ and $N = 100$. Both $f_{\max_{\mathrm{KS}}}$ and $f_{\max_{\mathrm{EP}}}$ have a linear behaviour with slope respectively 0.48 and 0.49, close to $\frac{1}{4\alpha} = 0.5$.

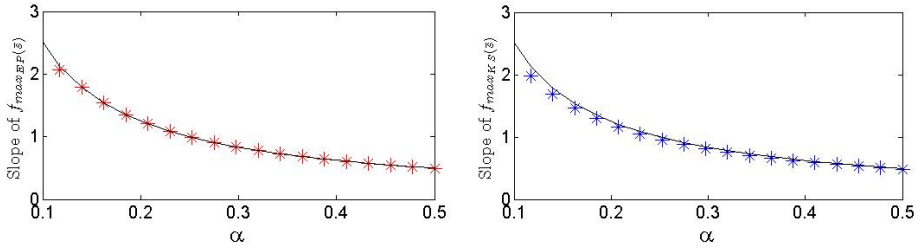


Figure 5. $f_{\max_{\mathrm{KS}}(\overline{s})}$ (left) and $f_{\max_{\mathrm{EP}}(\overline{s})}$ (right) as functions of α and in black the curve $f(\overline{s}) = \frac{1}{4\alpha}\overline{s}$. Note that the approximation $f_{\max_{\mathrm{KS}}}(\overline{s}) \approx f_{\max_{\mathrm{EP}}}(\overline{s}) \approx \frac{1}{4\alpha}\overline{s}$ is good.

This equation has a unique positive solution because the leading coefficient is positive for s small enough ($4\alpha - 6\alpha^2 + 6\alpha^3 - 4\overline{s} + 3\alpha^2\overline{s}) \geq 0$ and the constant term is negative $-(1-\alpha) \leq 0$. We remark that the optimal resolution $N_* = \frac{1}{a*}$ depends on the parameters of the system namely on the degree of non-equilibrium. This fact can be the explanation for two well-known issues in climate/weather modelling. First, it explains that, when downgrading or upgrading the resolution of convection models, the relevant parameters must be changed as they depend on the grid size. Second, it suggests that if the resolution is well tuned to represent a particular range of convective phenomena, it might fail in capturing the dynamics out of this range: since finer grids are needed to better represent deep convection phenomena, the deviations between model and observations observed in the distribution of extreme convective precipitation may be due to an inadequacy of the grid used.

5 Conclusions

We have shown how a simple 1-D Markov process, the Zero Range Process, can be used to obtain rigorous results on the problem of parametrization of the passive scalar transport problem, relevant to many geophysical applications including temperature distribution in climate modelling. Using this model, we have derived rigorous results on the link between a principle of maximum entropy production and the principle of maximum Kolmogorov–Sinai entropy using a Markov model of the passive scalar diffusion called the Zero Range Process. The Kolmogorov–Sinai entropy seen as a function of the convective velocity admits a unique maximum. We show analytically that both have the same Taylor expansion at the first order in the deviation from equilibrium. The behaviour of these two maxima is explored as a function of the resolution N (equivalent to the number of boxes, in the box approximation). We found that, for a fixed difference of potential between the reservoirs, the maximal convective velocity predicted by the maximum entropy production principle tends towards a non-zero value, while the maximum predicted using Kolmogorov–Sinai entropy tends to 0 when N goes to infinity. For values of N typical of those adopted by climatologists ($N \approx$10–100), we show that the two maxima nevertheless coincide even far from equilibrium. Finally, we show that there is an optimal resolution N_* such that the two maxima coincide to second order in $\overline{s}$, a parameter proportional to the non-equilibrium fluxes imposed to the boundaries. The fact that the optimal resolution depends on the intensity of the convective phenomena to be represented, points to new interesting research avenues, e.g. the introduction of convective models with adaptive grids optimized with maximum entropy principles on the basis of the convective phenomena to be represented.

The application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system. It would be interesting to apply it to a more realistic passive scalar transport problem, to see if it would yield a model that could be numerically handled (i.e. corresponding to a number of bows that is small enough to be handled by present computers). In view of applications to atmospheric convection, it

would be interesting to apply this procedure to the case of an active scalar, coupled with a Navier–Stokes equation for the velocity. In such a case, the role of f will be played by the turbulent subgrid Reynolds stresses. The heat fluxes and N^* will be fixed by the coarse-graining length, and the optimization procedure will in principle provide the optimum subgrid Reynolds stresses at a given resolution N. Moreover, by imposing coincidence of MKS and MEP, one could get both the Reynolds stresses, heat fluxes and the optimum resolution. Further, on the theoretical side, it will be interesting to study whether for general dynamical systems, there exists a smart way to coarse grain the Kolmogorov–Sinai entropy such that its properties coincide with the thermodynamic entropy production. This would eventually justify the use of the MEP principle and explain the deviations as well as the different representations of it due to the dependence of the dynamic (Kolmogorov–Smirnov–Tsallis–Jaynes) entropies on the kind of partition adopted.

Appendix A: Computation of the K-S entropy

In this appendix, we compute the Kolmogorov–Sinai entropy for the Zero Range Process, starting from its definition in Eq. (14). In the frame of our Zero Range Process, we use Eqs. (14) and (6) to write it as

$$h_{\mathrm{KS}} = -\sum_i \mu_{i_{\mathrm{stat}}} \sum_j p_{ij} \log(p_{ij}) = -\sum_{m_1=0}^{+\infty} \dots \sum_{m_N=0}^{+\infty} P(m_1, m_2, \dots, m_N) \sum_j p_{(m_1,\dots,m_N)\to j} \log(p_{(m_1,\dots,m_N)\to j})$$
$$= -\sum_{m_1=0}^{+\infty} P(m_1) \dots \sum_{m_N=0}^{+\infty} P(m_N) \sum_j p_{(m_1,\dots,m_N)\to j} \log(p_{(m_1,\dots,m_N)\to j}). \quad \text{(A1)}$$

We thus have to calculate $\sum_j p_{(m_1,\dots,m_N)\to j} \log(p_{(m_1,\dots,m_N)\to j})$ that we will refer to as $(*)$. We will take $p+q=\alpha+\delta=\beta+\gamma=1$ and $\mathrm{d}t=\frac{1}{N}$ in order to neglect the probabilities to stay in the same state compare to the probabilities of changing state. There are five different cases to consider:

1. if $\forall i\, m_i \geq 1$ so the possible transitions are:
   $(m_1, m_2, \dots, m_N) \to (m_1 \pm 1, m_2, \dots, m_N)$ with respective probabilities α and δ;
   $(m_1, m_2, \dots, m_N) \to (m_1, m_2, \dots, m_N \pm 1)$ with respective probabilities γ and β;
   and $(m_1, \dots, , m_k, \dots, m_N) \to (m_1, \dots, m_k \pm 1, \dots, m_N)$ with respective probabilities p and q;
   thus,
   $$(*) = \alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta + (N-1)(p\log(p)+q\log(q)); \quad \text{(A2)}$$
2. if $m_1 \geq 1$ and $m_N \geq 1$ and let i be the number of m_i between 2 and $N-1$ equal to 0. With the same argument as previously we have
   $$(*) = \alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta + (N-1-i)(p\log(p)+q\log(q)); \quad \text{(A3)}$$
3. if $m_1 = 0$ and $m_N \geq 1$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have
   $$(*) = \alpha\log\alpha + \delta\log\delta + \beta\log\beta + (N-2-i)p\log(p) + (N-1-i)q\log(q); \quad \text{(A4)}$$
4. the same applies if $m_1 \geq 1$ and $m_N = 0$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have
   $$(*) = \alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + (N-1-i)p\log(p) + (N-2-i)q\log(q); \quad \text{(A5)}$$
5. finally, if $m_1 = 0$ and $m_N = 0$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have
   $$(*) = \alpha\log\alpha + \delta\log\delta + (N-2-i)(p\log(p)+q\log(q)). \quad \text{(A6)}$$

Using Eq. (4) we find that $P(m_k=0) = 1-z_k$ and $\sum_{i=1}^{+\infty} P(m_k=i) = z_k$.
We thus obtain that

$$\begin{aligned} h_{\mathrm{KS}} = &-(\alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta) \\ &+ (N-1)(p\log(p)+q\log(q)) \\ &+ (p\log(p)+q\log(q)) \left(\sum_{r=0}^{N} r \sum_{i_1\dots i_N} \prod_{i=i_1,\dots i_r} (1-z_i) \prod_{i\neq i_1\dots i_r} z_i\right) \\ &+ (\gamma\log(\gamma)+p\log(p)) z_N (1-z_1) \left(\sum_{i_2\dots i_{N-1}} \prod_{i=i_2,\dots i_r} (1-z_i) \prod_{i\neq i_2\dots i_r} z_i\right) \\ &+ (\beta\log(\beta)+q\log q) z_1 (1-z_N) \left(\sum_{i_2\dots i_{N-1}} \prod_{i=i_2,\dots i_r} (1-z_i) \prod_{i\neq i_2\dots i_r} z_i\right) \\ &+ (\beta\log(\beta)+\gamma\log\gamma+p\log p+q\log q) \left(\sum_{i_2\dots i_{N-1}} \prod_{i=i_2,\dots i_r} (1-z_i) \prod_{i\neq i_2\dots i_r} z_i\right). \end{aligned} \quad \text{(A7)}$$

This expression, though complicated at first sight, can be simplified. Indeed, interested in the function $F(a) = \prod_1^N (z_k + a(1-z_k))$ and by deriving subject to a we show that

$$\sum_{r=0}^{N} r \sum_{i_1\dots i_N} \prod_{i=i_1,\dots i_r} (1-z_i) \prod_{i\neq i_1\dots i_r} z_i = \sum_{i=1}^{N} (1-z_i). \quad \text{(A8)}$$

Thus we can simplify the earlier equation to obtain

$$\begin{aligned} h_{\mathrm{KS}} = &-(\alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta \\ &+ (N-1)(p\log(p)+q\log(q))) \\ &+ (p\log(p)+q\log(q)) \sum_{i=1}^{N} (1-z_i) \\ &+ (\gamma\log(\gamma)+p\log(p))(1-z_1) \\ &+ (\beta\log(\beta)+q\log(q))(1-z_N). \end{aligned} \quad \text{(A9)}$$

Acknowledgements. M. Mihelich thanks IDEEX Paris Saclay for financial support.

Edited by: J. M. Redondo
Reviewed by: two anonymous referees

References

Andjel, E. D.: Invariant measures for the zero range process, Ann. Probabil., 525–547, 1982.

Balian, R.: Physique statistique et themodynamique hors équilibre, Ecole Polytechnique, 1992.

Billingsley, P.: Ergodic theory and information, Wiley, 1965.

Burda, Z., Duda, J., Luck, J. M., and Waclaw, B.: Localization of the maximal entropy random walk, Phys. Rev. Lett., 102, 160602, 2009.

Dewar, R. C. and Maritan, A.: A theoretical basis for maximum entropy production, in: Beyond the Second Law, 49–71, Springer, 2014.

Domb, C.: Phase transitions and critical phenomena, Vol. 19, Academic Press, 2000.

Frisch, U.: Turbulence: the legacy of AN Kolmogorov, Cambridge University Press, 1995.

Gómez-Gardeñes, J. and Latora, V.: Entropy rate of diffusion processes on complex networks, Phys. Rev. E, 78, 065102, 2008.

Großkinsky, S., Schütz, G. M., and Spohn, H.: Condensation in the zero range process: stationary and dynamical properties, J. Stat. Phys., 113, 389–410, 2003.

Herbert, C., Paillard, D., Kageyama, M., and Dubrulle, B.: Present and Last Glacial Maximum climates as states of maximum entropy production, Q. J. Roy. Meteorol. Soc., 137, 1059–1069, 2011.

Levine, E., Mukamel, D., and Schütz, G.: Zero-range process with open boundaries, J. Stat. Phys., 120, 759–778, 2005.

Martyushev, L. M. and Seleznev, V. D.: Maximum entropy production principle in physics, chemistry and biology, Phys. Rep., 426, 1–45, 2006.

Mihelich, M., Dubrulle, B., Paillard, D., and Herbert, C.: Maximum Entropy Production vs. Kolmogorov-Sinai Entropy in a Constrained ASEP Model, Entropy, 16, 1037–1046, 2014.

Monthus, C.: Non-equilibrium steady states: maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints, J. Stat. Mech., p. P03008, 2011.

Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Onsager, L.: Reciprocal relations in irreversible processes. I., Phys. Rev., 37, 2265–2279, 1931.

Paltridge, G. W.: Global dynamics and climate-a system of minimum entropy exchange, Q. J. Roy. Meteorol. Soc., 101, 475–484, 1975.

Pascale, S., Gregory, J. M., Ambaum, M. H., and Tailleux, R.: A parametric sensitivity study of entropy production and kinetic energy dissipation using the FAMOUS AOGCM, Clim. Dynam., 38, 1211–1227, 2012.

Rotstayn, L. D.: On the "tuning" of autoconversion parameterizations in climate models, J. Geophys. Res.-Atmso. (1984–2012), 105, 15495–15507, 2000.

Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, Boundary-Lay. Meteorol., 37, 129–148, 1986.

Turkington, B.: An optimization principle for deriving nonequilibrium statistical models of hamiltonian dynamics, J. Stat. Phys., 152, 569–597, 2013.

Wallace, J. M. and Hobbs, P. V.: Atmospheric science: an introductory survey, Vol. 92, Academic press, 2006.

5 Principaux résultats

Revenons quelques instants sur les résultats principaux démontrés dans cet article.

Le résultat majeur est le résultat analytique suivant : $f_{max_{EP}}$ et $f_{max_{KS}}$ ont le même développement limité au premier ordre en la différence à l'équilibre. Ainsi, dans le cas du modèle CASEP de la partie précédente nous avions démontré ce résultat de façon numérique alors que dans le cas du modèle ZRP le résultat est analytique. Le fait que $f_{max_{EP}}$ et $f_{max_{KS}}$ soient proches montre le lien entre MEP et MKS car les états finaux choisi par le système selon MEP et selon MKS seront alors proches eux aussi.

Le second résultat important est le fait que leurs développements au second ordre sont à priori différents. Cependant, il existe une certaine taille de résolution (un certain coarse-graining) pour lequel leurs développements coïncident. Ainsi, ceci peut permettre d'expliquer la difficulté de trouver un coarse graining adéquat, notamment en climatologie. En effet, nous pourrions penser naïvement que dans les modèles plus la résolution est fine et plus les résultats sont bons. Mais ceci n'est pas vrai []. Cet article permet donc de définir un coarse graining optimal.

5.1 travaux futurs.

On peut penser à plusieurs continuité de cet article :

- Nous avons établit un lien entre MEP et MKS dans un cas sans interaction du ZRP ($w_n = constante$). Que se passe t-il si w_n dépend de n ? C'est à dire si les probabilités de sauter à droite ou à gauche dépendent du nombre de particules contenu dans chaque case ? De plus, sous certaines conditions un phénomène de condensation apparait. On peut dès lors ce demander s'il existe un rapport quelconque entre ce phénomène et la maximisation de la production d'entropie ?
- La généralisation du procédé décrit dans cet article à un système quelconque régit par une équation de diffusion convection du type :

$$\partial_t T + u\partial_x T = \kappa \partial_x^2 T, \tag{3.38}$$

parait possible. Nous pouvons dès lors nous demander s'il serait possible de généraliser ce procédé à des systèmes plus généraux. Nous pouvons par exemple penser à un système régit par une équation de Navier Stokes du type :

$$(\vec{u}.\vec{\nabla})\vec{u} + \vec{\nabla}(p^*) = -\frac{1}{Re}\vec{\nabla}^2 u, \tag{3.39}$$

dans laquelle le nombre de Reynolds Re ne serait pas fixé (ou connu). Nous pourrions alors appliquer MEP, ou de façon quasi équivalente si nous sommes proches de l'équilibre MKS comme nous venons de le démontrer dans l'article, pour déterminer le nombre de Reynolds $Re_{max_{EP}}$ choisit par MEP. Si de plus la résolution N n'est pas fixée nous pouvons calculer la résolution dite optimale $N*$ telle que $Re_{max_{EP}}(N*) = Re_{max_{KS}}(N*)$.

Chapitre 4

Maximisation de la production d'entropie et coarse graining.

1 Introduction.

Nous avons observé et démontré dans les deux chapitres précédents qu'il existe un lien entre le principe de maximisation de production d'entropie et le principe de maximisation de l'entropie de Kolmogorov-Sinaï dans des systèmes Markoviens qui sont des modèles jouets classiques de la physique statistique.Il est dès lors naturel de se demander si ceci perdure dans des systèmes dynamiques plus généraux que les systèmes Markoviens ?
Après quelques rappels sur les systèmes dynamiques, nous allons voir qu'un lien existe entre la production d'entropie macroscopique et l'entropie de Kolmogorov-Sinaï pour des systèmes dynamiques qui ne sont pas forcément Markovien. Ce résultat a été établi par Latora et Baranger [43]. Cependant, nous verrons que ce lien est limité et que ceci est notamment dû au coarse graining. Nous introduirons par la suite une nouvelle façon de coarse grainer. Ensuite nous étudierons et nous donnerons l'expression générale de la production d'entropie microscopique pour un système dynamique et nous verrons, comme pour les chaines de Markov, que celle-ci diffère de l'entropie de Kolmogorov-Sinai. Puis nous étudierons la dépendance des exposants de Lyapunov en fonction de la taille du coarse graining en espace et en temps. Enfin nous nous intéresserons à l'évolution de la production d'entropie en fonction du coarse graining.

2 Introduction aux systèmes dynamiques.

Le terme système dynamique est apparu au début du XX ème siècle entre la publication du traité fondateur de Poincaré : "Les méthodes nouvelles de la mécanique céleste" en 1899 [44] et celle, en 1927, de la monographie de Birkhoff intitulée "Dynamical systems" [45]. L'idée prédominante à cette époque est de cesser de s'intéresser à la résolution explicite des équations différentielles mais plutôt au comportement qualitatif. On va donc chercher l'allure des solutions au voisinage d'un point, les quantités invariantes par le mouvement, les points d'équilibre, leur stabilité...Une des premières utilisations des systèmes dynamiques est la résolution du problème biologique proie-prédateur de Lotka Volterra [] qui a permis d'expliquer pourquoi

pendant la première guerre mondiale, alors que les pêcheurs n'allaient plus pêcher, le nombre de sardines a diminué.
De plus, les gens observent que deux états proches à un instant initial peuvent admettre des évolutions totalement différentes, c'est ce qui est à l'origine de la théorie du chaos et qui constitue sûrement la plus grande application de la théorie des systèmes dynamiques. Aujourd'hui les systèmes dynamiques sont utilisés dans quasiment tous les domaines de la physique.

Décrire les résultats majeurs concernant les sytèmes dynamiques constituerait bien plus qu'une (même 1000 !) thèse. C'est pourquoi j'ai choisi de ne pas m'attarder sur la théorie des systèmes dynamiques mais plutôt d'introduire les notions importantes qui seront utiles pour la suite.

2.1 Systèmes dynamiques : Définition et quelques exemples.

Un système dynamique est la donnée d'un triplet (T, E, S) où T est un ensemble inclus dans $\mathbb{R}$, E est un espace appelé espace des phases et S est une application de $T * E$ dans E telle que :

$$\begin{cases} \forall\, x \in E \; S(0,x) = x \\ \forall\, t \text{ et } s \text{ tel que } t+s \in I \; S(t, S(s,x)) = S(t+s,x). \end{cases} \tag{4.1}$$

Pour illuster la notion de systèmes dynamiques donnons quelques exemples élémentaires utiles pour la suite.

Soit X un élément de $\mathbb{R}^N$ et F une application de $\mathbb{R}^N$ dans $\mathbb{R}^N$ continue. Alors

$$\dot{X} = F(X), \tag{4.2}$$

constitue un système dynamique continu. Nous verrons de multiples exemples comme le système de Lorenz donné par les équations :

$$\dot{X} = \sigma(Y - X) \quad \dot{Y} = -XZ + rX - Y \quad \dot{Z} = XY - bZ, \tag{4.3}$$

où r et b sont des paramètres du système.

Mais aussi le système de Van Der Pol donné par l'équation différentielle :

$$\ddot{X} = \epsilon\omega_0(1 - X^2)\dot{X} - \omega_0^2 X, \tag{4.4}$$

où ϵ et ω_0 sont des paramètres du système.

Nous pouvons définir les systèmes dynamiques à temps discrets par une équation du type :

$$X_{n+1} = F(X_n), \tag{4.5}$$

où F est une application de E dans E. Dans les cas envisagés par la suite E sera souvent isomorphe à une partie de $\mathbb{R}^n$. Nous verrons de multiples exemples de systèmes discrets comme la Baker map, la Standard map...

2.2 Chaos, Exposants de Lyapunov et théorème de Pesin.

Les exposants de Lyapunov mesurent la divergence (ou la convergence) de deux trajectoires ayant des conditions initiales très proches. Commençons par donner la définition des exposants de Lyapunov dans le cas d'un système dynamique discret de $\mathbb{R}$ dans $\mathbb{R}$: $X_{n+1} = F(X_n)$ où F est dérivable. Partons d'un point X_0 de l'espace des phases ainsi que d'un point voisin $X_0 + \epsilon_0$. Au temps suivant le premier point se retrouvera en $X_1 = F(X_0)$ alors que le second se retrouvera en $X_1 + F'(X_0)\epsilon_0$. Ainsi l'erreur au temps suivant vaut $\epsilon_1 = F'(X_0)\epsilon_0$. De manière générale, si l'erreur au temps $n-1$ est ϵ_{n-1} alors l'erreur au temps n sera $\epsilon_n = F'(X_{n-1})\epsilon_{n-1}$ et donc on trouve que :

$$|\epsilon_n| = \frac{|\epsilon_n|}{|\epsilon_{n-1}|}\frac{|\epsilon_{n-1}|}{|\epsilon_{n-2}|}...\frac{|\epsilon_1|}{|\epsilon_0|}|\epsilon_0| = (\prod_0^{n-1}|F'(X_n)|)|\epsilon_0| \tag{4.6}$$

Donc si $\epsilon_n \sim e^{\lambda n}|\epsilon_0|$ on trouve finalement :

$$\lambda = \frac{1}{n}\sum_0^{n-1}\log(|F'(X_i)|). \tag{4.7}$$

Ainsi, si le système dynamique admet une mesure stationnaire μ on en déduit que l'exposant de Lyapunov vaut :

$$\lambda = \int_E \log(|F'(x)|)d\mu. \tag{4.8}$$

Pour un système dynamique continu, l'expression est identique.

Considérons maintenant un système dynamique discret de $\mathbb{R}^N$ dans $\mathbb{R}^N$: $X_{n+1} = F(X_n)$. Définissons la matrice jacobienne J :

$$J(X) = (\frac{\partial F_i}{\partial X_j}(X))_{i,j}, \tag{4.9}$$

De la même manière que précédemment, une erreur ϵ_n au temps n donne au temps $n+1$ une erreur $\epsilon_{n+1} = J(U_n)\epsilon_n$. Ainsi,

$$\epsilon_n = J(X_{n-1})J(X_{n-2})...J(X_0)\epsilon_0. \tag{4.10}$$

Définissons alors la matrice $M_n = J(X_{n-1})J(X_{n-2})...J(X_0)$

Comme dans le cas précédent, on peut penser que les valeurs propres de M_n pour les n grands se comportent comme $e^{\lambda_i n}$. Plus précisément, sous certaines hypothèses qui constituent le théorème d'Oseledec, les valeurs propres de la matrice :

$$(M_n^t M_n)^{\frac{1}{2n}}, \tag{4.11}$$

convergent. Les exposants de Lyapunov sont alors défini comme le logarithme de ces valeurs propres.

3 Lien entre production d'entropie et entropie de KS.

La définition de l'entropie de Gibbs ($\int \rho \log(\rho)$) [46] pose plusieurs difficultés. La première est qu'elle suppose connue la fonction de densité ρ sous sa forme exacte (par exemple analytique) ce qui n'est jamais le cas dans une expérience car il y a toujours une échelle liée aux mesures à laquelle s'ajoute des erreurs, du bruit...De plus il est fort probable que la nature ait introduit une échelle intrinsèque dans l'espace des phases comme le montre la relation d'incertitude d'Heisenberg. Pour tenir compte de ces échelles, les physiciens ont partitionné l'espace des phases en cellules (souvent cubique c'est à dire liées à la norme infinie) et ont calculé l'entropie en l'approximant par la somme de Riemann correspondante. Cette dernière entropie que nous appellerons entropie coarse grainée n'est pas simplement une approximation de l'entropie de Gibbs et se comporte différemment. En effet, celle-ci se rapproche de la formulation initiale de Boltzmann $S = -k_B \log(W)$ où les différentes cellules de notre partition représente les différents états. De plus, alors que l'entropie de Gibbs est constante pour un système Hamiltonien l'entropie coarse grainée croît ce qui est en accord avec le second principe.

Nous définirons l'entropie coarse grainée puis nous verrons que celle-ci croit avec une pente égale à l'entropie de KS.

3.1 Entropie coarse grainée.

Considérons un système dynamique du type $X_{n+1} = F(X_n)$ où F est une application de Γ dans Γ. Γ sera appelé l'espace des phases. Si Γ est un carré (cube, hypercube...) on peut le partitionner naturellement en petites boites ayant des faces parallèles aux faces du carrés :

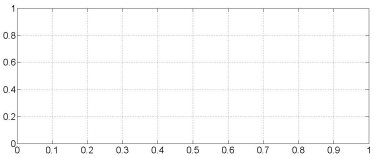


FIGURE 4.1 – Partition en carrés de l'espace des phases $[0,1]^2$.

Partons alors d'une configuration initiale dans laquelle toutes les particules (N) sont regroupées dans une même case et laissons le système évoluer selon la dynamique $X_{n+1} = F(X_n)$. Un exemple est donné en figure Fig.(4.2) avec la Baker Map où la fonction F est définie par :

$$F(x,y) = \begin{cases} (2x, y/2) & \text{for } 0 \leq x < \frac{1}{2} \\ (2-2x, 1-y/2) & \text{for } \frac{1}{2} \leq x < 1. \end{cases} \tag{4.12}$$

Numérotons les cases et notons N_i le nombre de particules qui se trouvent dans la case i. On peut dès lors définir l'entropie macroscopique $S(t)$ à l'instant t par :

$$S(t) = -\sum_i \frac{N_i(t)}{N} \log(\frac{N_i(t)}{N}), \tag{4.13}$$

Et calculer $\frac{dS}{dt}$ avant que le système atteigne un état stationnaire. Lorsque l'on fait ceci sur plusieurs configurations initiales on peut alors moyenner et on remarque [43] que l'on a une relation du type :

$$\frac{dS}{dt} \approx h_{KS}. \tag{4.14}$$

Nous avons codé l'exemple de la Baker Map. Ce système dynamique est largement chaotique avec deux exposants de Lyapunov $\lambda_x = \log(2)$ et $\lambda_y = \log(\frac{1}{2})$. Ainsi, ce système préserve le volume et $h_{KS} = \lambda_x = \log(2)$. Nous avons tracé numériquement $\frac{d\overline{S}_{\epsilon cg}}{dt}$ pour la Baker-Map et nous l'avons comparé à une fonction affine de pente log(2) Fig.(4.3)

Cette relation est aussi vérifiée pour d'autres systèmes dynamiques comme pour la standart map, la cat map, la generalized cat map [43]...Nous pouvons donc nous demander pourquoi une telle relation existe t-elle ?

Commençons par un exemple assez simples : Considérons une map 2D, agissant dans un espace des phases divisé en boites de taille ϵ, qui multiplie les longueurs par $\lambda_1 \geq 1$ dans une certaine direction et $\lambda_2 = \frac{1}{\lambda_1}$ dans une direction perpendiculaire de telle sorte à préserver l'espace des phases. Partons d'une configuration initiale où les particules (la distribution de probabilité) sont confinées dans une seule boite. Au bout d'un régime transitoire assez cours la tache de particules initiale se déforme en

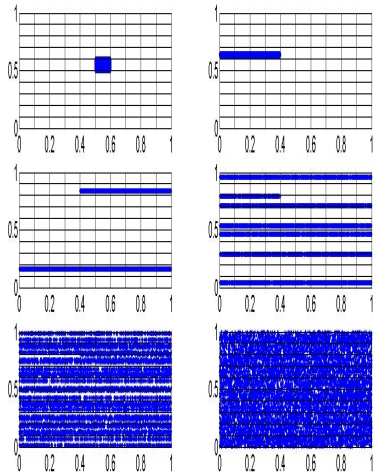


FIGURE 4.2 – Exemple de l'évolution de la Baker map.

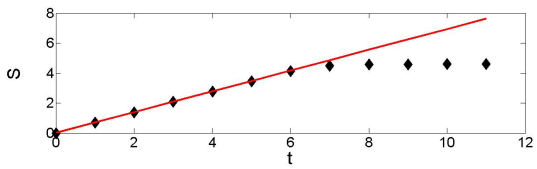


FIGURE 4.3 – S fonction de t pour la Baker-Map en trait pointillé bleu. En rouge une droite affine de pente log(2). Nous pouvons remarquer que la relation $\frac{dS}{dt} \approx h_{KS}$ est relativement bien vérifiée.

un rectangle de dimension $L \gg 1$ et $l \leq \epsilon \ll 1$. Ainsi, l'entropie définie dans l'équation eq.(4.13) vaut approximativement, si on suppose que les boites contiennent le même nombre de particules, $S \approx \log(L\epsilon)$. Au temps suivant le rectangle aura pour dimension $\lambda_1 L$ et $\lambda_2 l$. Ainsi l'entropie sera donc égale $S \approx \log(\lambda_1 L\epsilon)$. On obtient donc que la production d'entropie $\Delta S \approx \log(\lambda_1) = h_{KS}$. Dans cet exemple élémentaire on relie la production d'entropie coarse grainée par rapport aux boites à l'entropie de Kolmogorov-Sinaï.

De manière plus générale considérons un système dynamique de $\mathbb{R}^n$ ayant n exposants de Lyapunov λ_i $i \in [|1...n|]$. Nous supposerons $\lambda_i \geq 0$ $i \in [|1...r|]$ et $\lambda_i < 0$ $i \in [|r+1...n|]$. Pour une densité de probabilité initiale gaussienne :

$$f(x_1, ..., x_n) = \prod_1^n \frac{1}{\sqrt{(2\pi)}\sigma_0} exp(-\frac{x_k^2}{2\sigma_0^2}), \tag{4.15}$$

la densité coarse grainée au bout d'un temps t tel que $exp(t\lambda_i) \ll \epsilon$ $\forall i \in [|r+1...n|]$ prend la forme :

$$f_{cg}(x_1, ..., x_n) \approx \prod_1^r \frac{1}{\sqrt{(2\pi)}\sigma_i(t)} exp(-\frac{x_k^2}{2\sigma_i^2}) \prod_{r+1}^n \frac{1}{\sqrt{(2\pi)}\epsilon} exp(-\frac{x_k^2}{2\epsilon^2}), \tag{4.16}$$

où $\sigma_i(t) \approx exp(\lambda_i t)$. Ainsi on trouve que :

$$\frac{dS}{dt} \approx \sum_{i=1}^r \lambda_i = h_{KS}. \tag{4.17}$$

Ainsi, ces arguments qualitatifs simples nous permettent de comprendre que pour des systèmes dynamiques qui ne sont pas trop compliqués la production d'entropie macroscopique vaut approximativement l'entropie de KS. Cependant, dans notre raisonnement nous avons fait plusieurs hypothèses :

i) Nous avons supposé que la valeur des exposants de Lyapunov était constante dans l'espace des phases alors que d'une façon générale les exposants de Lyapunov dépendent de la position dans l'espace des phases.

ii) Nous avons supposé que les boites qui contenait des particules contenait le même nombre de particules ce qui revient à dire (grosso modo) que la mesure stationnaire est uniforme.

iii) Nous avons supposé que les directions d'expansion et de contraction étaient orthogonales et que les faces des boites étaient dans les mêmes directions que celles-ci.

Ainsi cette connexion entre la production d'entropie et l'entropie de KS parait limitée car contrainte à beaucoup d'hypothèses. Ce constat a déjà été fait dans l'article de Falcioni [47]. Nous reviendrons sur ces limitations et sur certains contre-exemples dans lesquels la relation n'est pas vérifiée. Nous pourrons alors essayer de trouver une méthode afin d'élargir le domaine de validité de cette relation.

3.2 Existe-t-il un véritable lien dans les systèmes dynamiques entre la production d'entropie et l'entropie de KS ?

Commençons par regarder comment varie la relation lorsque l'hypothèse 1 n'est plus vérifiée. Prenons, comme dans [47], un système dynamique 1D dans lequel l'exposant de Lyapunov dépend de la position dans l'espace des phases. Plus précisément considérons la tent map définie sur le segment $[0,1]$ par :

$$F(x) = \begin{cases} \frac{x}{p} & \text{pour } 0 \leq x \leq p \\ \frac{1-x}{1-p} & \text{pour } p < x \leq 1. \end{cases} \tag{4.18}$$

Où p est un paramètre compris entre 0 et 1. Si $p = \frac{1}{2}$ on retrouve la tent map habituelle. Sur $[0,p]$ F est dérivable et $F'(x) = 1/p$ donc sur $[0,p]$ l'exposant de Lyapunov est positif est vaut $\lambda_1 = \log(1/p) = -\log(p)$. De même, sur $[p,1]$ l'exposant de Lyapunov est positif est vaut $\lambda_2 = -\log(1-p)$. Ainsi, l'entropie de KS vaut pour ce système dynamique :

$$h_{KS} = -p\log(p) - (1-p)\log(1-p). \tag{4.19}$$

L'idée de Falcioni a été de calculer la production d'entropie coarse grainée (en divisant l'intervalle $[0,1]$ en boites de longueur ϵ selon la procédure expliquée dans la section précédente) et de la comparer à l'entropie de KS et ceci pour différents p Fig.([?])

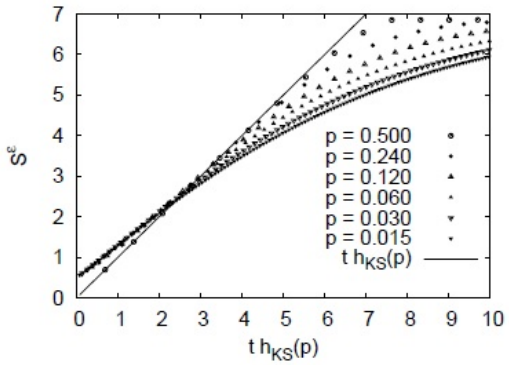


FIGURE 4.4 – Comparaison entre l'entropie coarse grainée et l'entropie de KS pour différentes valeurs de p. On remarque que pour $p = \frac{1}{2}$ la relation entre les deux entropies est bien vérifiée. Cependant, plus p est proche de 0 plus cette relation est mise à défaut. Cette figure est extraite de [47].

Ce contre exemple n'est pas isolé et il existe de nombreux autres systèmes dynamiques dans lesquels la relation n'est pas vérifiée.

3.3 Limitations dues au coarse graining.

Une autre raison importante qui limite la validité de cette relation est le fait que les axes d'expansion et de contraction ne sont souvent pas les mêmes que les axes des boites. Ainsi le choix de la direction des axes des boites est arbitraire et une possibilité serait de prendre des directions qui varieraient dans l'espace des phases en suivant les directions d'expansion et de contraction. Cependant ceci est assez difficile

à mettre en place. De plus, coarse grainer signifie que l'on ne peut pas voir en dessous d'une certaine échelle. Cependant, lorsque l'on effectue le coarse graining avec des boites de taille ϵ on impose une "disposition" particulière. C'est à dire que l'on peut très bien imaginer avec ce coarse graining deux particules qui sont à une distance plus petite que ϵ mais qui appartiennent à deux boites différentes et qui donc seront discernables. Ainsi, cette façon de coarse grainer avec des boites a des limites.

Nous allons dans la section suivante, après avoir fait un bref rappel du coarse graining de manière générale, introduire une nouvelle façon de coarse grainer plus adéquate.

4 Une nouvelle méthode pour coarse grainer.

Nous allons donner dans cette partie une nouvelle façon de coarse grainer qui apparait plus appropriée notamment pour le calcul de la production d'entropie macroscopique. Avant cela, nous allons commencer par faire un bref rappel sur les méthodes de coarse graining classiques. Nous n'entrerons pas dans les détails mais nous donnerons les références si le lecteur est intéressé.

4.1 Méthodes classiques de coarse graining.

La méthode du coarse graining a été introduite en 1911 par les époux Ehrenfest [48]. Elle consiste, de la même manière que le font Latora et Baranger, à diviser l'espace des phases en cellules de petites tailles et de considérer n'importe quelle fonction sur cet espace des phases comme une fonction constante par morceaux sur les cellules. Cette méthode est très bien expliquée dans l'article de N.Gorban [49] dans lequel il l'illustre par l'évolution d'une densité de probabilité dans l'espace des phases pour un système hamiltonien :

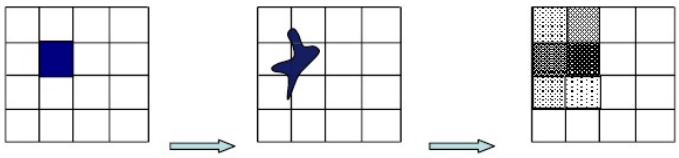


FIGURE 4.5 – De gauche à droite : Évolution d'une densité de probabilité dans l'espace des phases (le volume est donc conservé). Puis coarse graining. On voit bien que lors de l'évolution microscopique il n'y a pas de variation d'entropie alors que pendant l'étape de coarse graining il y a une variation d'entropie positive.

Une autre manière de coarse grainer résulte du constat suivant : Les systèmes physiques sont en général soumis aux fluctuations d'un environnement extérieur qui sont très difficilement prédictibles. Ainsi, on peu étudier l'évolution d'un système dynamique en présence d'un bruit stochastique. Plus précisément, à la place de considérer une évolution du type $\dot{X} = F(X)$ nous pouvons considérer une évolution du type $\dot{X} = F(X + \zeta_1) + \zeta_2$ où ζ_1 et ζ_2 sont deux bruits. Nous reviendrons sur ce type de coarse graining plus tard dans ce chapitre.

On peut aussi coarse grainer une équation aux dérivées partielles (ou ses solutions) en ne gardant que les modes qui paraissent importants [50].

Le coarse graining est aussi à la base de la théorie de renormalisation et notamment avec la transformation de Kadanoff pour les réseaux de spins. **Nous avons utilisé (Davide Faranda, Bérengère Dubrulle, moi et al) une méthode analogue pour l'étude d'un réseau de spin 2D mis hors équilibre. Cet article est présenté dans les annexes.**

Le coarse graining est aussi à la base du groupe de renormalisation de Wilson [51] ainsi que des nombreuses techniques de renormalisation récentes.

4.2 Nouvelle méthode : La convolution.

Afin d'éviter que l'entropie ne dépende de la partition choisie, on peut penser à convoler f plutôt que de la moyenner sur des cellules. La convolution comme moyen de coarse grainer est utiliser en turbulence dans les équations de Navier-Stockes [] (référence Bérengère)

Soit G_ϵ une fonction que nous appellerons fonction poids c'est à dire une fonction "piquée en 0" et telle que $\int_\Omega G_\epsilon = 1$.
Par exemple dans $\mathbb{R}^n$,

$$G_\epsilon(x_1, x_2, ..., x_n) = \frac{1}{(\epsilon\sqrt{(2\pi)})^n} \exp(-\frac{(x_1^2 + ... + x_n^2)}{2\epsilon^2}), \tag{4.20}$$

qui est une fonction gaussienne ou encore, si l'on note $B_2(0, \epsilon)$ la boule pour la norme euclidienne, on peut considérer la fonction plateau :

$$G_\epsilon(x_1, x_2, ..., x_n) = \frac{1}{Vol(B_2(0, \epsilon)} 1_{B_2(0,\epsilon)}. \tag{4.21}$$

On notera

$$\overline{f}_{\epsilon cg}(X) = f * G_\epsilon = \int_\Omega f(u) G_\epsilon(X - u) du$$

On définie alors l'entropie coarse grainée $\overline{S}_{\epsilon cg}$ par :

$$\overline{S}_{\epsilon cg} = \int_\Omega \overline{f}_{\epsilon cg} \log(\overline{f}_{\epsilon cg}) d\Omega. \tag{4.22}$$

On remarquera que par propriété élémentaire sur la convolution on a encore :

$$\int_\Omega \overline{f}_{\epsilon cg} d\Omega = 1.$$

La ϵ-convolution représente physiquement l'impossibilité de voir à une échelle plus fine que ϵ. Contrairement au coarse graining à l'aide de partition, elle ne donne pas d'axe privilégié. De plus, si $f \in L^1(\Omega)$ et G_ϵ est C^k alors
$\overline{f}_{\epsilon cg}$ est C^k . Ainsi si G_ϵ est régulière alors $\overline{f}_{\epsilon cg}$ le sera aussi.

Enfin, pour des G_ϵ usuelles (voir ceux cités en exemple) $\overline{f}_{\epsilon cg}$ converge uniformément vers f sur Ω et donc :

$$\lim_{\epsilon \to 0} \overline{S}_{\epsilon cg} = S,$$

ou S est l'entropie de Gibbs.

Nous allons voir les propriétés intéressantes de $\overline{S}_{\epsilon cg}$, notamment le fait que $\frac{d\overline{S}_{\epsilon cg}}{dt} \approx h_{KS}$ ou h_{KS} est l'entropie de Kolmogorov-Sinaï. Et ceci de manière plus satisfaisante que l'entropie utilisée par Latora et Baranger pour laquelle les cellules sont fixées.

En effet, nous avons calculé l'entropie coarse grainée pour le système dynamique introduit par Falcioni 4.18 et nous comparons cette entropie à l'entropie de KS. La figure obtenue Fig.(4.6) est donc à comparer avec la figure Fig.(4.4). Nous remarquons que sur la figure Fig.(4.6) les différentes courbes qui correspondent aux différentes valeurs de p coincident. Ceci montre qu'en effectuant un coarse graining, la relation entre la production d'entropie macroscopique et l'entropie de KS est vérifiée de façon plus satisfaisante. Néanmoins, coarse grainer grâce à la convolution demande beaucoup plus de temps de calcul.

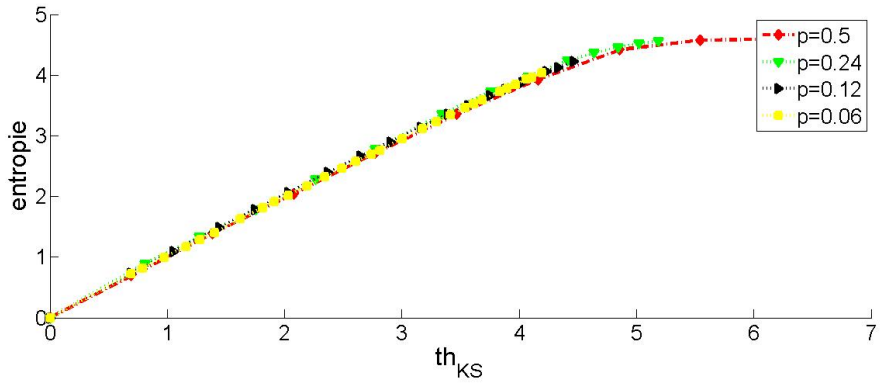


FIGURE 4.6 – Comparaison entre la nouvelle entropie coarse grainée avec la convolution et l'entropie de KS pour différentes valeurs de p.

5 Production d'entropie microscopique.

Comme nous l'avons mentionné auparavant, l'entropie calculée par un ordinateur ou celle observée par un expérimentateur est l'entropie macroscopique car il existe toujours une échelle limite en dessous de laquelle il est impossible de voir. Cependant, il est parfois utile de s'intéresser à l'entropie microscopique. L'entropie microscopique ou encore appelée entropie de Gibbs demande la connaissance de le densité de probabilité ρ en tout point X de l'espace des phases et son expression est bien connue :

$$S = -\int_{\Gamma} \rho(X) \log(\rho(X)) d\Gamma. \tag{4.23}$$

La question de cette partie est : Que vaut la production d'entropie microscopique ?
Tout comme l'a démontré Ruelle [52] nous allons commencer par donner l'expression

analytique exacte de la production d'entropie microscopique. Nous allons voir que l'on obtient le même comportement que pour les chaines de Markov : la production d'entropie macroscopique est proche de l'EKS alors que la production d'entropie microscopique diffère.

5.1 Expression analytique de l'entropie microscopique.

Soit un système dynamique défini par $\frac{dX}{dt} = G(X)$ et notons g^t le système dynamique correspondant, c'est à dire que $X(t) = g^t(X_0)$ où X_0 est la position à l'état initial.
Nous allons décrire les états non pas par des points dans l'espace des phases mais par des mesures de probabilité. Ceci est justifié par l'approche statistique due au très grand nombre de particules à considérer. Nous supposerons de plus que les mesures de probabilité sont assez régulières pour définir des densités de probabilité associées. Partant d'une densité de probabilité initiale ρ, la densité de probabilité à l'instant t ρ_t vérifie :

$$\rho_t(X) = \frac{\rho(g^{-t}(X))}{J_t(g^{-t}(X))}, \tag{4.24}$$

où J_t est le Jacobien (déterminant de la matrice Jacobienne) de g^t
Donc :

$$S(\rho_t) = -\int \rho_t(X)\log(\rho_t(X))dX = -\int \frac{\rho(g^{-t}(X))}{J_t(g^{-t}(X)}\log(\frac{\rho(g^{-t}(X))}{J_t(g^{-t}(X)})dX, \tag{4.25}$$

et après changement de variable :

$$S(\rho_t) = -\int \rho(X)\log(\frac{\rho(X)}{J_t(X)})dX = S(\rho) + \int \rho(X)\log(J_t(X))dX, \tag{4.26}$$

et donc :

$$\Delta S = S(\rho_t) - S(\rho) = \int \rho(X)\log(J_t(X))dX, \tag{4.27}$$

au premier ordre en dt nous pouvons réécrire $\frac{dX}{dt} = G(X)$ sous la forme $X_{t+dt} = X_t + dt.G(X_t) + O(dt^2)$ et donc $g^{dt}(X) = X + dt.G(X)$.

Ainsi, en utilisant par exemple la différentielle du déterminant en l'identité on obtient : $J_{dt}(X) = 1 + div(G(X))dt$

et donc :

$$\frac{dS}{dt} = \int \rho(X)div(G(X))dX. \tag{4.28}$$

Ainsi, on trouve l'expression générale :

$$\frac{dS}{dt} = \sum_i < \lambda_i >, \tag{4.29}$$

où les λ_i sont les exposants de Lyapunov et <> signifie la moyenne par rapport à ρ.

Ainsi pour un système Hamiltonien comme il y a conservation de l'espace des phases ($J_t = 1$) on trouve que $\frac{dS}{dt} = 0$.

5.2 dS/dt<0 ?

Pour un système gouverné par de la dissipation on trouve que $\sum_i < \lambda_i >< 0$ et donc $\frac{dS}{dt} < 0$ ce qui semble être contradictoire avec le second principe. Un moyen de comprendre ce paradoxe est de mettre en place la notion de thermostat [53] et de voir que la production d'entropie liée au thermostat doit être positive afin de compenser celle du système. Un autre moyen d'éviter ce paradoxe est d'utiliser la notion d'epsilon entropie (une autre forme de coarse grained entropie) pour trouver une production d'entropie positive [54, 55].

Dans la section suivante nous allons étudier le comportement des exposants de Lyapunov en fonction de la taille du coarse-graining et voir que nous retrouvons les résultats précédents sur l'entropie coarse grainée.

6 Coarse grained par rapport au temps ou à l'espace.

6.1 Un autre coarse graining.

Dans les coarses graining que nous avons considéré auparavant nous avons supposé que la dynamique F était parfaitement connue mais pas l'espace des phases Γ. Cependant, s'il existe une imprécision sur la position de la particule il parait naturel qu'il en existe une aussi sur la dynamique. Nous pouvons plus rigoureusement envisager plusieurs cas :

i) Dans le coarse graining précédent nous avons séparé l'espace des phases en boites. Dans une même boite les particules sont considérées indiscernables. Cependant la dynamique est appliquée à la position exacte de la particule (alors que l'on ne connait pas sa position exacte) ce qui peut paraitre contradictoire. On peut envisager le cas où chaque particule dans une même boite a une dynamique qui est celle d'un point aléatoire dans cette même boite. Ceci revient à prendre une dynamique du type :

$$\dot{X} = F(X + \zeta), \tag{4.30}$$

où ζ est un bruit (bruit blanc, gaussien, uniforme...)

ii) On peut aussi envisager un bruit sur la dynamique. C'est à dire considérer un système dynamique de la forme :

$$\dot{X} = F_\zeta(X). \tag{4.31}$$

Une manière particulière de faire ceci est d'envisager un coarse graining par rapport au temps en envisageant d'accroitre les erreurs dans la méthode d'Euler en augmentant le pas de temps dt.

iii) Enfin, nous pouvons imaginer une erreur à la fois sur la dynamique et sur la position de la particule.

6.2 Coarse grained par rapport au temps.

Nous allons voir dans cette partie que nous retrouvons le lien entre la production d'entropie et l'entropie de KS lorsque nous coarse grainons simplement par rapport au temps. En effet, lorsque nous augmentons le pas de temps dt et que nous calculons les exposants de Lyapunov le long de la trajectoire nous remarquons que les exposants de Lyapunov négatifs augmentent alors que les exposants de Lyapunov positifs restent constants. Lorsque nous avons assez augmenté le pas de temps pour que tous les exposants de Lyapunov négatifs soient devenu positifs ou nuls on remarque que la structure de l'attracteur est préservée même si les directions de contraction et de dilatation ne le sont plus. Nous avons testé cette procédure sur différents systèmes dynamiques discrets (Henon, Baker maps, Standard maps) ainsi que sur différents systèmes dynamiques continues (Van der Pol, Duffing, Rossler et L'attracteur de Lorenz). Les résultats sont équivalents. Ici nous allons présenter le cas de l'attracteur de Lorenz [56] donné par le système d'équations eq.(4.3)

Nous prenons les valeurs typiques pour les paramètres : $\sigma = 3$, $r = 26.5$ and $b = 1$. Nous considérons aussi deux map de Lorenz couplées qui représente une dynamique chaotique lente et rapide comme expliqué dans [50]. C'est à dire que nous considérons un système avec plusieurs échelles de temps. Les résultats trouvés sont décrit dans la figure Fig. 4.7 : En haut ce sont les résultats du système de Lorenz simple alors qu'en bas ce sont les résultats de la map couplée. La figure nous montre le comportement décrit précédemment : la valeur des exposants de Lyapunov négatifs augmente alors que celle des exposants positifs reste constante.

Le plus grand Δt considéré est le plus petit Δt pour lequel tous les exposants de Lyapunov deviennent positifs, ainsi pour ce Δt l'égalité Eq. (4.17) est vérifiée. De plus, pour voir l'effet du changement de Δt sur l'attracteur nous avons tracé l'attracteur de Lorenz pour trois différents $\Delta t = 6 \cdot 10^{-3}, 0.5, 20.1$. Les couleurs représentent le temps : au début la trajectoire est tracée en bleue puis en rouge. Enfin, nous avons aussi tracé la projection de la trajectoire suivant l'axe X. On remarque que bien que Δt soit grand la structure de l'attracteur est conservée.

6.3 Coarse grained par rapport à l'espace.

Nous envisageons ici un coarse graining en temps, c'est à dire un système dynamique modifié de la forme $\dot{X} = F(X + \epsilon)$ où ϵ est un bruit. Nous avons envisagé plusieurs bruits (bruit gaussien, bruit uniforme...). Les résultats sont sensiblement semblables au coarse graining en temps : plus le coarse graining augmente plus les exposants négatifs augmentent. Cependant, dans le coarse graining en temps les exposants de Lyapunov positifs sont constants alors qu'ici ils augmentent lorsque le coarse graining en espace augmente. Néanmoins, les exposants de Lyapunov positifs augmentent moins rapidement que les exposants négatifs. C'est ce que nous pouvons voir sur la figure Fig.4.9. Nous avons pris un bruit uniforme sur $[-r, r]$ et nous avons tracé les exposants de Lyapunov fonction de r.

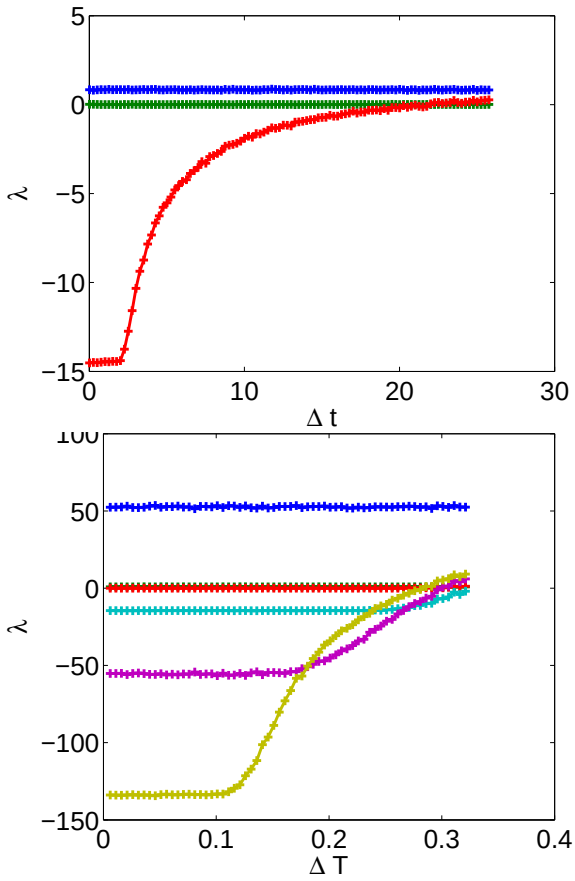


FIGURE 4.7 – Comportement des exposants de Lyapunov en fonction du pas de temps Δt pour l'attracteur de Lorenz classique (haut) et pour un modèle couplé de deux attracteurs de Lorenz (voir [50]).

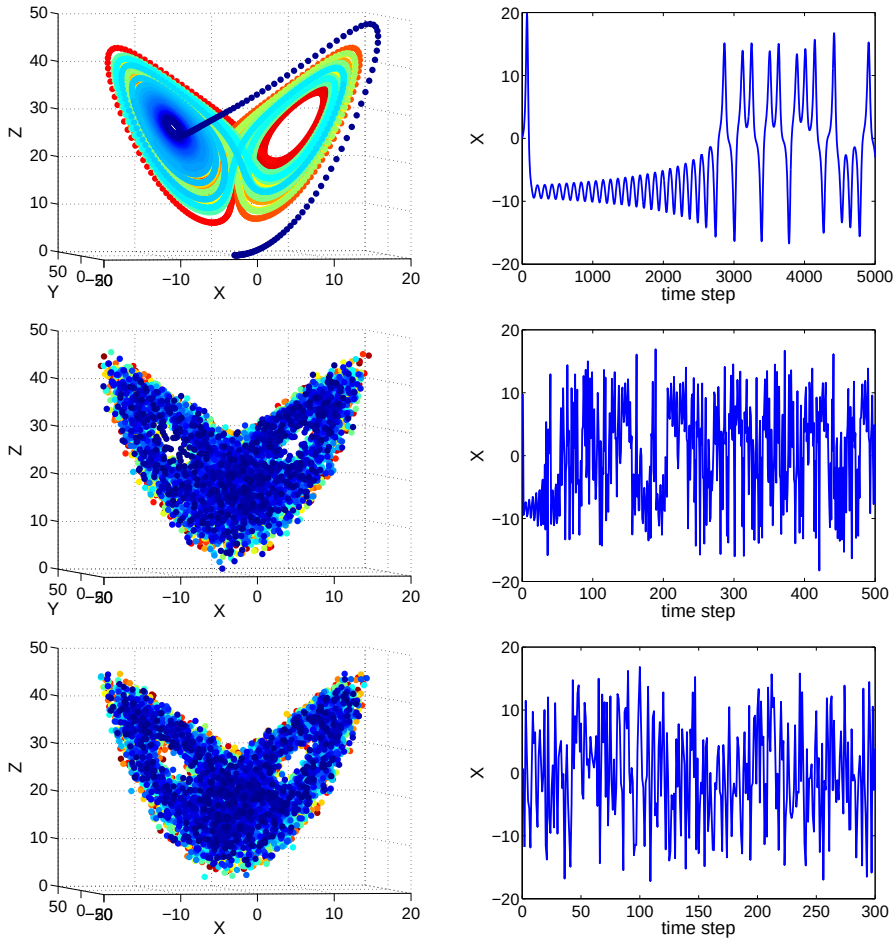


FIGURE 4.8 – Attracteurs de Lorenz (gauche) et son projeté sur la variable X seulement (droite) pour trois différents coarse graining : $\Delta t = 6 \cdot 10^{-3}, 0.5, 20.1$. L'échelle de couleur représente le temps : au début la trajectoire est bleue puis elle devient rouge.

7 Conclusion.

Dans cette partie nous avons vu que la relation entre la principe de maximisation de production d'entropie coarse grainée et le principe de maximisation d'entropie de KS pouvait être étendue à des systèmes dynamiques non nécessairement markovien. Néanmoins ce lien n'est pas universel et il existe des systèmes dynamiques dans lesquels il est mis en défaut. Une des raisons est dans la façon de coarse grainer. Nous avons donc introduit une manière de coarse grainer plus propre qui donne de meilleurs résultats. Puis, nous avons donné l'expression analytique de la production d'entropie microscopique associée à un système dynamique qui est égale à la somme de tous les exposants de Lyapunov. Enfin, nous avons montré que lorsque l'on coarse graine les exposants de Lyapunov négatif augmentent ce qui montre que l'on retrouve l'expression de la production d'entropie macroscopique avec celle de la production d'entropie microscopique.

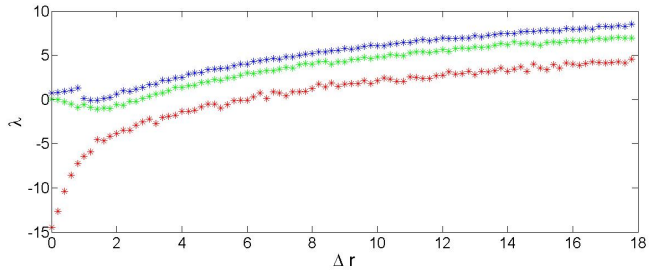


FIGURE 4.9 – Exposants de Lyapunov fonction de r pour la Lorenz map simple 4.3

Chapitre 5

Maximisation de l'entropie de KS et minimisation du mixing time.

1 Introduction.

La théorie des réseaux est une branche de la physique-mathématique qui est aujourd'hui en pleine essor vu le nombre croissant de ses applications : World Wide Web, réseaux sociaux, étude des neurones... Plus précisément, les chercheurs se sont intéressés aux marches aléatoires sur différents types de réseaux. Une quantité qui peut être utilisée pour l'étude de ces marches aléatoire est l'entropie de KS [26, 27]. La maximisation de l'entropie de KS permet d'obtenir une dynamique pour laquelle les chemins de même taille sont équiprobables et de retrouver des phénomènes naturels. Elle permet aussi de définir des dynamiques diffusives "bien mélangeantes" comme il l'est mentionné dans l'article de Roberta Sinatra [**?**]. Néanmoins, l'intuition selon laquelle la dynamique qui maximise l'entropie de KS "mélange bien" n'a pas de fondement théorique précis. Nous allons dans ce chapitre donner des arguments plus précis que ceux existant. En effet, nous allons établir un lien entre la maximisation de l'entropie de KS et la minimisation du temps de mélange, ce dernier étant défini comme le temps que le système prend pour atteindre son état stationnaire. La minimisation du temps de mélange est un problème très important notamment en informatique où l'on veut que les algorithmes arrivent le plus rapidement vers l'état stationnaire.

Dans ce chapitre nous allons commencer par donner la définition du temps de mélange (mixing time) ainsi que quelques propriétés qui lui sont associées. Nous allons ensuite expliquer quelles sont les méthodes numériques pour rechercher la dynamique qui minimise le temps de mélange. Nous allons ensuite montrer qu'il existe une méthode analytique (théorème de Parry) pour trouver la dynamique qui maximise l'entropie de KS. Nous allons enfin montrer le lien entre la minimisation du temps de mélange et la maximisation de l'entropie de KS, notamment à travers notre article. Enfin, nous allons décrire les perspectives et les travaux futures qui permettraient d'élargir ce lien aux systèmes dynamiques plus généraux.

2 Chaine de Markov et temps de mélange :

Dans toute la suite de ce chapitre (Ω, P, μ_{stat}) désignera une chaine de Markov sur un espace fini Ω avec une matrice de transition $P = (p_{ij})_{i,j}$ et admettant une unique mesure stationnaire μ_{stat}. On rappelle que l'on a :

$$\forall\, i,j\; p_{ij} \geq 0\,; \forall\, i\; \sum_j p_{ij} = 1\,; \forall\, j\; \sum_i p_{i,j}\mu_{stat_i} = \mu_{stat_j}. \tag{5.1}$$

Nous supposerons de plus que P est irréductible c'est à dire que le graphe associé est connexe. Enfin, nous supposerons que $\forall\ \mu$ mesure quelconque $(P^t)^n\mu$ converge vers μ_{stat}. Une condition suffisante pour obtenir cette propriété est que P soit apériodique.

2.1 Définition du temps de mélange.

Soient μ et π deux mesures sur $P(\Omega)$ l'ensemble des parties de Ω. On définit la distance usuelle par, si p est un nombre positif :

$$d_p(\mu,\pi) = (\sum_{x\in\Omega} |\mu(x) - \pi(x)|^p)^{\frac{1}{p}}. \tag{5.2}$$

Pour $p = 2$ on retrouve la norme euclidienne alors que pour $p = 1$ on peut facilement démontrer que l'on retrouve (à un facteur $1/2$ près) la "Total Variation" définie par :

$$d_{TV}(\mu,\pi) = \max_{A\in P(\Omega)} |\mu(A) - \pi(A)|, \tag{5.3}$$

μ étant un vecteur de $\mathbb{R}^n$ où n est le cardinal de Ω.

Etant donnée une distance d on défini le temps de mélange $t(\epsilon)$ dépendant d'un paramètre ϵ assez petit par :

$$t(\epsilon) = \min\, n :\ \forall x \in \Omega\ d(P(x,.), \mu_{stat}) \leq \epsilon. \tag{5.4}$$

Le temps de mélange représente donc le temps minimal à partir duquel, quelque soit la configuration initiale, le système est à une distance plus petite que ϵ de son état stationnaire.

2.2 Lien avec les valeurs propres.

Nous allons établir dans cette partie un lien entre le temps de mélange et la seconde amplitude maximale des valeurs propres de la matrice de transition P. Cette relation est montrée par exemple dans l'article de Diaconis [57].

Nous supposerons ici P diagonalisable dans $\mathbb{C}$. Ceci est évidemment vérifié lorsque P est symétrique car diagonalisable dans $\mathbb{R}$. De plus, lorsque l'on prend une matrice de transition aléatoire elle est λ presque sûrement diagonalisable où λ est la mesure de Lebesgue sur $\mathbb{R}^{n^2-n}$ (la puissance $n^2 - n$ provient du fait qu'une matrice a n^2 coefficients et que nous avons n relations du fait que la matrice doit être stochastique). Donc cette hypothèse n'est en faite pas trop restrictive.

Notons donc les valeurs propres par ordre décroissant :

$$1 = \lambda_1 > |\lambda_2| \geq .... \geq |\lambda_m| \geq 0. \tag{5.5}$$

D'après le théorème de Perron Frobenius le plus grand module est 1 associé à la valeur propre 1 et le sous espace propre associé est de dimension 1 engendré par μ_{stat}.

Notons $\mu_{stat}, X_2...X_m$ des vecteurs propres associés aux valeurs propres λ_1,λ_2,...,λ_m. Partons d'une configuration initiale μ_0 que nous décomposons sur la base des vecteurs propres :

$$\mu_0 = \alpha_1\mu_{stat} + \alpha_2 X_2 + ...\alpha_m X_m, \tag{5.6}$$

où α_1, α_2,...,α_m sont m complexes.
Après n itérations (au temps n) nous obtenons :

$$\mu_n = (P^t)^n\mu_0 = \alpha_1\mu_{stat} + \sum_2^m \alpha_k\lambda_k^n X_k = \alpha_1\mu_{stat} + \lambda_2^n \sum_2^m \alpha_k \frac{\lambda_k}{\lambda_2}^n X_k. \tag{5.7}$$

Et donc on voit bien que si l'on note :

$$\lambda(P) = \max_{i=2...m} |\lambda_i| = |\lambda_2|, \tag{5.8}$$

On obtient d'après eq.(5.7) :

$$||(P^t)^n\mu_0 - \mu_{\text{stat}}|| \propto (\lambda(P))^n. \tag{5.9}$$

En utilisant les équations Eqs. (5.4) et (5.9) on trouve que $\lambda(P)^{t(\epsilon)} \propto \epsilon$. Donc :

$$\lambda(P) \propto \epsilon^{1/t(\epsilon)}. \tag{5.10}$$

Ainsi, plus $\lambda(P)$ est petit plus le temps de mélange est lui aussi petit. Nous verrons que les relations précédentes sont effectivement vérifiées dans l'article présenté à la fin du chapitre.

2.3 Quelques inégalités sur le temps de mélange.

Comme nous l'avons mentionné, la recherche d'un temps de mélange minimal est un enjeu majeur notamment en informatique. Il existe donc de nombreuses inégalités pour majorer et minorer le temps de mélange. Une synthèse d'une grande partie de ces inégalités est très bien décrite dans [58]. Donnons quelques inégalités importantes et commençons par une majoration assez simple du temps de mélange. Pour cela nous considèrerons une chaine de Markov P irréversible et apériodique (voir définition et propriétés dans le chapitre 2). On a donc l'existence et l'unicité de la mesure stationnaire μ_{stat} ainsi que la convergence, quelque soit la mesure initiale, vers μ_{stat}. Nous allons montrer, comme le font D-A.Levin et Y.Peres [59] qu'il existe α, appartenant à $]0,1[$, tel que :

$$\max_{x\in\Omega} d_{TV}(P^n(x,.),\mu_{stat}) \leq \alpha^n, \tag{5.11}$$

Où $P^n(x,.)$ représente la densité de probabilité en itérant la matrice P n fois à partir de la configuration initiale x.

Pour cela, comme P est apériodique :

$$\exists r \in \mathbb{N}, \forall (i,j)(P^r)_{ij} > 0. \tag{5.12}$$

Notons Π la matrice de rang 1 ayant toutes ses lignes égales à μ_{stat}. Pour δ assez petit nous obtenons que $P^r(i,j) \geq \delta\mu_{stat}(j)$ et ceci pour quelque soit (i,j). Ainsi, nous pouvons écrire

$$P^r = \delta\Pi + (1-\delta)Q, \tag{5.13}$$

où Q est une matrice stochastique. En remarquant que quelle que soit la matrice stochastique M on a $M\Pi = \Pi$ et que quelle que soit la matrice vérifiant $\mu_{stat}M = \mu_{stat}$ on a $\Pi M = \Pi$, on montre par récurrence que :

$$\forall k \in \mathbb{N} P^{rk} = (1-(1-\delta)^k)\Pi + (1-\delta)^k Q^k. \tag{5.14}$$

Ainsi, en multipliant par P^j pour j entier et en servant du fait que $\Pi P^j = \Pi$ on obtient :

$$P^{rk+j} - \Pi = (1-\delta)^k(Q^kP^j - \Pi), \tag{5.15}$$

En prenant la norme L^1 des deux cotés, puis en divisant par 2 pour faire apparaitre la distance d_{TV} et enfin en prenant le max nous obtenons :

$$\max_{x\in\Omega} d_{TV}(P^{rk+j}(x,.),\mu_{stat}) = (1-\delta)^k \max_{x\in\Omega} d_{TV}(Q^kP^j(x,.),\mu_{stat}). \tag{5.16}$$

En sachant que $\max_{x\in\Omega} d_{TV}(Q^kP^j(x,.),\mu_{stat}) \leq 1$, alors $\alpha = (1-\delta)^{\frac{1}{2r}}$ convient pour vérifier l'équation eq.(5.11).

3 A la recherche de la dynamique qui minimise le temps de mélange.

Le problème est ici le suivant : On se donne un graphe fixé (c'est à dire une matrice d'adjacence A fixée) et on cherche la matrice de transition P_{mix} (c'est à dire la dynamique) vérifiant les conditions du graphe ($(P_{mix})_{ij} = 0$ dès que $(A)_{ij} = 0$) , qui minimise le temps de mélange. Mathématiquement, grâce notamment à l'équation eq.(5.10), ce problème prend la forme suivante :

$$\begin{cases} \min\limits_{P\in M_n} \lambda(P), \\ P(i,j) \geq 0, P * \mathbf{1} = \mathbf{1}, \\ A(i,j) = 0 \Rightarrow P(i,j) = 0. \end{cases} \tag{5.17}$$

Ce problème ne connait pas de solution analytique connue.

Après avoir donné quelques exemples de graphes pour lesquels le temps de mélange est petit nous verrons les méthodes numériques pour résoudre ce problème.

3.1 Quelques exemples élémentaires de chaines qui mélangent rapidement

Nous allons donner ici des exemples de dynamiques classiques pour lesquelles le temps de mélange est assez faible [57].

On rappelle pour la suite que le degrés d'un nud i noté d_i est le nombre de nuds j qui sont connectés avec i. Ainsi on peut écrire $d_i = \sum_j A(i,j)$. L'exemple le plus classique et le plus naturel de dynamique qui respecte le graphe $\mathcal{E}$ est l'exemple de la marche aléatoire sur le graphe pour lequel la probabilité d'aller du nud i vers le nud j ne dépend pas de j et vaut $\frac{1}{d_i}$ si $A(i,j) = 1$ et 0 sinon. Plus précisément :

$$\begin{cases} P_{ma}(i,j) = \frac{1}{d_i} \, si \, A(i,j) = 1, \\ P_{ma}(i,j) = 0 \text{ sinon.} \end{cases} \tag{5.18}$$

Il s'avère que cette dynamique n'est pas réversible. Néanmoins on peut construire une dynamique réversible (c'est à dire qui vérifie le bilan détaillé) qui admet comme mesure stationnaire la même mesure stationnaire que la dynamique préccédente de la marche aléatoire. Cette méthode s'appelle la méthode de Metropolis-Hastings. Pour cela notons $(\mu_1, \mu_2, ..., \mu_m)$ la mesure stationnaire de la dynamique correspondante à la marche aléatoire ainsi que $R_{ij} = \frac{\mu_j P^{ma}(j,i)}{\mu_i P^{ma}(i,j)}$. Nous pouvons alors définir la matrice de transition de Metropolis-Hastings par :

$$\begin{cases} P_{MH}(i,j) = P_{ma}(i,j) \min\{1, R_{ij}\} \, si \, A(i,j) = 1 \text{ et si } i \neq j, \\ P_{MH}(i,j) = P_{ma}(i,j) + \sum_{k:A(i,k)=1} P_{ma}(i,k)(1 - \min\{1, P_{ma}(i,k)\}) \text{ si } i = j, \\ P_{MH}(i,j) = 0 \text{ sinon.} \end{cases} \tag{5.19}$$

La chaine de Markov associée à cette matrice est réversible et le temps de mélange est assez faible.

Nous pouvons définir une autre dynamique classique qui est rapidement mélangeante : le graphe de degré maximal. Pour cela notons $d_{max} = \max_i\{d_i\}$. Le graphe de degré maximal est alors défini par :

$$\begin{cases} P_{dm}(i,j) = \frac{1}{d_{max}} \text{ si } i \neq j \text{ et } A(i,j) = 1, \\ P_{dm}(i,j) = 1 - \frac{d_i}{d_{max}} \text{ si } i = j, \\ P_{dm}(i,j) = 0 \text{ sinon.} \end{cases} \tag{5.20}$$

Ces deux exemples ont des temps de mélange assez faibles mais ne minimisent pas (sauf cas particuliers) le temps de mélange.

3.2 Algorithmes qui recherchent la dynamique qui mélange le plus rapidement.

Il n'existe pas (on ne connait pas) aujourd'hui d'expression analytique pour la matrice P_{mix} qui minimise le temps de mélange. Néanmoins, il existe des algorithmes (souvent assez lourd en temps de calcul) qui permettent d'approximer P_{mix}. Donnons en quelques exemples.

Nous allons dans le début de cette partie commencer par rechercher la matrice qui minimise le temps de mélange non pas parmi toutes les matrices de transition possibles mais parmi les matrices de transition qui sont symétriques. Nous cherchons alors à résoudre le problème suivant :

$$\begin{cases} \min\limits_{P\in M_n} \lambda(P), \\ P(i,j)\geq 0,\ P*\mathbf{1}=\mathbf{1},\ P^t=P, \\ A(i,j)=0\Rightarrow P(i,j)=0. \end{cases} \tag{5.21}$$

Nous savons que 1 est valeur propre et qu'un vecteur propre associé est $\mathbf{1}$. $\mathbf{1}$ étant le vecteur de $\mathbb{R}^n$ ayant que des 1 comme composantes. De plus, P est symétrique donc ses sous espaces propres sont orthogonaux. Ainsi, la valeur propre de second module maximal est à chercher dans l'orthogonal de la droite engendrée par $\mathbf{1}$. Comme la projection orthogonale sur $Vect(\mathbf{1})^{ortho}$ s'écrit $p=\mathbf{I}-\frac{1}{n}\mathbf{1}\mathbf{1}^t$ et que le module de la plus grande valeur propre d'un matrice symétrique$A$ est égale à la norme spectrale de celle-ci notée $|||A|||_2$, on peut donc réécrire le problème d'optimisation sous la forme :

$$\begin{cases} \min\limits_{P\in M_n} |||P(\mathbf{I}-\frac{1}{n}\mathbf{1}\mathbf{1}^t)|||_2 = |||P-\frac{1}{n}\mathbf{1}\mathbf{1}^t|||_2, \\ P(i,j)\geq 0,\ P*\mathbf{1}=\mathbf{1},\ P^t=P, \\ A(i,j)=0\Rightarrow P(i,j)=0. \end{cases} \tag{5.22}$$

Ce problème 5.22 est un problème d'optimisation sous contraintes d'égalités et d'inégalités. Nous pouvons donc penser en première approche à le résoudre avec une méthode de type Karush-Kuhn-Tucker (KKT). C'est ce que nous avons mis en place avec Matlab pour obtenir les résultats figurant dans l'article.

La restriction de la recherche de la dynamique optimale au sous espace des matrices symétriques est néanmoins assez restrictive. Nous allons voir ici comment nous pouvons résoudre un problème plus général : Nous nous fixons une densité stationnaire $\mu_{stat}=(\mu_{stat_1},\mu_{stat_2},...,\mu_{stat_m})$ et nous recherchons parmi toutes les matrices réversibles celle(s) qui minimise(nt) $\lambda(P)$.

Si nous notons $\Pi=diag(\mu)$. Alors P est réversible si et seulement $\Pi P=\Pi^t P$. Et donc en particulier $\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}$ est symétrique et a les mêmes valeurs propres que Π. Nous sommes donc quasiment ramené au problème précédent. Enfin, $p=(\sqrt{\mu_{stat_1}},...,\sqrt{\mu_{stat_m}})$ est un vecteur propre de $\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}$ associé à la valeur propre 1. Donc, le problème de minimisation prend la forme suivante :

$$\begin{cases} \min\limits_{P} |||(I_d-\frac{1}{n}\mathbf{q}\mathbf{q}^t)\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}(I_d-\frac{1}{n}\mathbf{q}\mathbf{q}^t)|||_2 \\ \qquad = |||\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}-\frac{1}{n}\mathbf{q}\mathbf{q}^t|||, \\ P(i,j)\geq 0, P*\mathbf{1}=\mathbf{1}, \Pi P=\Pi^t P, \\ A(i,j)=0\Rightarrow P(i,j)=0. \end{cases} \tag{5.23}$$

Nous avons résolu le problème d'optimisation précédent (5.23) qui est un problème d'optimisation sous contraintes sous Matlab. Cet algorithme n'est pas optimal en terme de temps de calcul bien que polynomiale. Il donne des résultats assez rapidement pour des graphes comportant $\approx 10^2$, 10^3 nuds.
Il existe de nombreux autres algorithmes qui calculent la matrice P_{mix} pour un nombre plus important de nuds [57], le maximum étant $\approx 10^5$. Néanmoins ces autres

algorithmes comportent d'autres désavantages comme la non convergence quel que soit la condition initiale, la difficulté d'estimer si l'on se trouve, ou non, près de la solution...

4 A la recherche de la dynamique qui maximise l'EKS.

Le problème est ici le suivant : On se donne un graphe fixé (c'est à dire une matrice d'adjacence A fixée) et on cherche la matrice de transition P_{KS} (c'est à dire la dynamique) vérifiant les conditions du graphe ($(P_{KS})_{ij} = 0$ dès que $(A)_{ij} = 0$) , qui maximise l'entropie de KS. Nous verrons dans une première partie que ce problème admet une solution analytique qui a été trouvée par William Parry (1934-2006), mathématicien anglais qui a beaucoup travaillé sur les systèmes dynamiques et la théorie ergodique. Ensuite, nous donnerons l'expression exacte de la matrice P_{KS}.

4.1 Théorème de Parry

Plutôt que de réécrire plus ou moins l'article de Parry [60] qui s'appuie sur un formalisme mathématique assez lourd, nous allons, dans cette partie, donner l'idée et la méthode qui ont conduit Parry à découvrir P_{KS}.

4.2 Méthode analytique pour trouver la dynamique qui maximise l'EKS.

Ici A est fixée et supposée symétrique. Notons λ la plus grande valeur propre de A et Ψ le vecteur propre normalisé associé c'est à dire $A\Psi = \lambda\Psi$ et $\sum_i \Psi_i^2 = 1$. Définissons P_{KS} telle que

$$P_{KS}(i,j) = \frac{A(i,j)}{\lambda}\frac{\Psi_j}{\Psi_i}. \tag{5.24}$$

On a dès lors bien entendu :

$$\forall\, i \sum_j P_{KS}(i,j) = 1, \tag{5.25}$$

et de plus :

$$\sum_j P_{KS}(j,i)\Psi_j^2 = \sum_j \frac{A(j,i)\Psi_i\Psi_j}{\lambda} = \Psi_i, \tag{5.26}$$

car $A(i,j) = A(j,i)$. Donc $P_{KS}^t\Psi^2 = \Psi^2$ et donc la densité stationnaire de P_{KS} est $\pi_{stat} = \Psi^2$.

On a donc d'après l'expression de l'entropie de KS pour les chaines de Markov :

$$h_{KS} = -\sum_{(i,j)} \mu_{stat_i} P_{KS}(i,j)\log(P_{KS}(i,j)), \tag{5.27}$$

et en remplaçant avec les expressions précédentes on trouve :

$$h_{KS} = -\frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(\frac{A(i,j)}{\lambda}\frac{\Psi_i}{\Psi_j}), \tag{5.28}$$

que l'on peut séparer en deux termes :

$$h_{KS} = -\frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(\lambda) - \frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(A(i,j)\frac{\Psi_j}{\Psi_i}), \tag{5.29}$$

Le terme de gauche vaut $\log(\lambda)$ car Ψ vecteur propre de A et le terme de droite vaut 0 par symétrie de A.

Donc, pour cette matrice P_{KS}, l'entropie de KS vaut :

$$h_{KS} = \log(\lambda). \tag{5.30}$$

De plus, nous avons vu dans le chapitre 2 que si A désigne la matrice d'adjacence, le nombre de trajectoires de longueurs n allant de i à j vaut $A^n(i,j)$. Ainsi, le nombre de trajectoires total de longueur n vaut :

$$N_n = \sum_{(i,j)} (A^n)(i,j). \tag{5.31}$$

En se souvenant que, pour les chaines de Markov, l'entropie de KS est aussi égale à la limite lorsque n tend vers ∞ de la dérivée de l'entropie sur les chemins par rapport au temps (voir fin du chapitre 1), on peut dès lors majorer l'entropie de KS par la valeur qu'elle aurait si tout les chemins étaient équiprobables. On trouve donc pour un temps asymptotiquement long :

$$h_{KS_{max}} = \frac{\log(N_n)}{n} \to \log(\lambda), \tag{5.32}$$

en diagonalisant la matrice A.

Enfin P_{KS} vérifie

$$\pi_{stat_i} P_{KS}(i,j) = \pi_{stat_j} P_{KS}(j,i) \ \forall \ i,j, \tag{5.33}$$

et donc P_{KS} est réversible.

5 Lien entre Minimisation du temps de mélange et maximisation de l'EKS.

L'idée générale, basée notamment sur le théorème de Pesin selon laquelle plus le système est chaotique plus son entropie de KS est grande nous donne l'intuition que plus que le système a une entropie de KS élevée, plus celui-ci atteint rapidement son état stationnaire. Néanmoins ceci n'est qu'une idée qualitative. De plus, comme nous l'avons déjà mentionné dans l'introduction, il existe de nombreux articles qui maximisent l'entropie de KS pour obtenir soit des comportements caractéristiques, soit des comportements que l'on retrouve dans la réalité, mais sans véritablement

comprendre pourquoi. Le lien que nous allons établir dans l'article suivant entre la maximisation de l'entropie de KS et la minimisation du temps de mélange va nous permettre de donner un sens plus précis à la raison pour laquelle maximiser l'entropie de KS donne des résultats cohérents. De plus, nous pouvons rapprocher ce résultat de celui obtenu dans le second et le troisième chapitre où nous avions établi un lien entre MEP et la maximisation de l'entropie de KS. En effet, il est physiquement cohérent qu'un système ayant plusieurs états stationnaires possibles supposés équiprobables ait un état stationnaire final plus proche de celui dont la dynamique va le plus vite possible. Pour illustrer cela nous pouvons penser à un enfant qui joue dans des escaliers et qui monte deux marches à la fois et n'en descend qu'une. Il a donc deux dynamiques possibles : celle qui consiste à monter et celle qui consiste à descendre. La première admet comme état stationnaire l'enfant en haut avec probabilité 1 et la seconde l'enfant en bas. Comme il monte les marches deux par deux il ira plus vite en haut de l'escalier qu'en bas et donc la dynamique montante est plus rapide. Si maintenant nous supposons que les deux dynamiques peuvent agir de manière équiprobable, il en résultera un état stationnaire qui sera plus proche de la dynamique montante que celle descendante.

5.1 Article

Maximum Kolmogorov-Sinai entropy vs minimum mixing time in Markov chains

M. Mihelich,[1,*] B. Dubrulle,[1] D. Paillard,[2] Q. Kral,[3,4] and D. Faranda[2]

[1] *Laboratoire SPHYNX, Service de Physique de l'Etat Condensé, DSM, CEA Saclay, CNRS UMR 3680, 91191 Gif-sur-Yvette, France*
[2] *Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA-CNRS-UVSQ, UMR 8212, 91191 Gif-sur-Yvette, France*
[3] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[4] *LESIA-Observatoire de Paris, UPMC Univ. Paris 06, Univ. Paris-Diderot, 92195 Meudon Cedex, France*

(Writing in progress: February 24, 2015)

We establish a link between the maximization of Kolmogorov Sinai entropy (KSE) and the minimization of the mixing time for general Markov chains. Since the maximisation of KSE is analytical and easier to compute in general than mixing time, this link provides a new faster method to approximate the minimum mixing time dynamics. It could be interesting in computer sciences and statistical physics, for computations that use random walks on graphs that can be represented as Markov chains.

PACS numbers:

Many modern techniques of physics, such as computation of path integrals, now rely on random walks on graphs that can be represented as Markov chains. Techniques to estimate the number of steps in the chain to reach the stationary distribution (the so-called "mixing time"), are of great importance in obtaining estimates of running times of such sampling algorithms [1] (for a review of existing techniques, see e.g. [2]). On the other hand, studies of the link between the topology of the graph and the diffusion properties of the random walk on this graph are often based on the entropy rate, computed using the Kolmogorov-Sinai entropy (KSE)[3]. For example, one can investigate dynamics on a network maximizing the KSE to study optimal diffusion [3], or obtain an algorithm to produce equiprobable paths on non-regular graphs [4].

In this letter, we establish a link between these two notions by showing that for a system that can be represented by Markov chains, **a non trivial relation exists between the maximization of KSE and the minimization of the mixing time**. Since KSE are easier to compute in general than mixing time, this link provides a new faster method to approximate the minimum mixing time that could be interesting in computer sciences and statistical physics and gives a physical meaning to the KSE. We first show that on average, the greater the KSE, the smaller the mixing time, and we correlated this result to its link with the transition matrix eigenvalues. Then, we show that the dynamics that maximises KSE is close to the one minimizing the mixing time, both in the sense of the optimal diffusion coefficient and the transition matrix.

Consider a network with m nodes, on which a particle jumps randomly. This process can be described by a finite Markov chain defined by its adjacency matrix A and its transition matrix P. $A(i,j) = 1$ if and only if there is a link between the nodes i and j and 0 otherwise. $P = (p_{ij})$ where p_{ij} is the probability for a particle in i to hop on the j node. Let us introduce the probability density at time n $\mu_n = (\mu_n^i)_{i=1...m}$ where μ_n^i is the probability that a particle is at node i at time n. Starting with a probability density μ_0, the evolution of the probability density writes: $\mu_{n+1} = P^t \mu_n$ where P^t is the transpose matrix of P.
Within this paper, we assume that the Markov chain is irreducible and thus has a unique stationary state.

Let us define:

$$d(n) = max||(P^t)^n \mu - \mu_{stat}|| \ \forall \ \mu, \tag{1}$$

where $||.||$ is a norm on $\mathbb{R}^n$. For $\epsilon > 0$, the mixing time, which corresponds to the time such that the system is within a distance ϵ from its stationary state is defined as follows:

$$t(\epsilon) = \min n, \ d(n) \leq \epsilon. \tag{2}$$

For a Markov chain the KSE takes the analytical form [5]:

$$h_{KS} = -\sum_{ij} \mu_{stat_i} p_{ij} \log(p_{ij}). \tag{3}$$

Random m size Markov matrices are generated by assigning to each p_{ij} $(i \neq j)$ a random number between 0 and $\frac{1}{m}$ and $p_{ii} = 1 - \sum_{j \neq i} p_{ij}$. The mean KSE is plotted versus the mixing time (Fig. 1) by working out h_{KS} and $t(\epsilon)$ for each random matrix. (Fig. 1) shows that KSE is on average a decreasing function of the mixing time.

We stress the fact that this relation is only true on average. We can indeed find two special Markov chains $P1$ and $P2$ such that $h_{KS}(P1) \leq h_{KS}(P2)$ and $t_1(\epsilon) \leq t_2(\epsilon)$. We illustrate this point further.

The link between the mixing time and the KSE can be understood via their dependence as a function of the transition matrix eigenvalues. A general irreducible transition matrix P is not necessarily diagonalizable on $\mathbb{R}$. However, since P is chosen randomly, it is almost everywhere diagonalizable on $\mathbb{C}$. According to Perron Frobenius theorem, the largest eigenvalue is 1 and the associated eigen-space is one-dimensional and equal to the vectorial space generated by μ_{stat}. Without loss of generality, we can label the eigenvalues in decreasing order of their module:

$$1 = \lambda_1 > |\lambda_2| \geq .... \geq |\lambda_m| \geq 0$$

The convergence speed toward μ_{stat} is given by the second maximum module of the eigenvalues of P [6], [7]:

$$\lambda(P) = \max_{i=2...m} |\lambda_i| = |\lambda_2|$$

The eigenvalues $\lambda_1 = 1, ..., \lambda_m$ of P and P^t being equal, let us denote their associated eigenvectors $\mu_1 = \mu_{\text{stat}}, ..., \mu_m$. For any initial probability density μ_0, we find:

$$||(P^t)^n \mu_0 - \mu_{\text{stat}}|| \propto (\lambda(P))^n. \quad (4)$$

According to Eqs. (1) and (2), $\lambda(P)^{t(\epsilon)} \propto \epsilon$, i.e. $\lambda(P) \propto \epsilon^{1/t(\epsilon)}$. Hence, the smaller $\lambda(P)$ the shorter the mixing time (Fig. 1). h_{KS} being a decreasing function of $t(\epsilon)$ and $\lambda(P)$ being an increasing function of $t(\epsilon)$, we deduce that h_{KS} is a decreasing function of $\lambda(P)$.

This link between maximum KSE and minimum mixing time actually also extends naturally to optimal diffusion coefficients. Such a notion has been introduced by Gomez-Gardenes and Latora [3] in networks represented by a Markov chain depending on a diffusion coefficient. Based on the observation that in such networks, KSE has a maximum as a function of the diffusion coefficient, they define an optimal diffusion coefficient as the value of the diffusion corresponding to this maximum. In the same spirit, one could compute an optimal diffusion coefficient with respect to the mixing time, corresponding to the value of the diffusion coefficient which minimizes the mixing time -or equivalently the smallest second largest eigenvalue $\lambda(P)$. This would roughly correspond to the diffusion model reaching the stationary time in the fastest time. To define such an optimal diffusion coefficient, we follow Gomez and Latora and vary the transition probability depending on the degree of the graph nodes. More accurately, if $k_i = \sum_j A(i,j)$ denotes the degree of node i, we set:

$$p_{ij} = \frac{A_{ij} k_j^\alpha}{\sum_j A_{ij} k_j^\alpha}. \quad (5)$$

If $\alpha < 0$ we favor transitions towards low degrees nodes, if $\alpha = 0$ we find the typical random walk on network and if $\alpha > 0$ we favor transitions towards high degrees nodes. We assume here that A is symmetric. It may then be checked that the stationary probability density is equal to:

$$\pi_{stat_i} = \frac{c_i k_i^\alpha}{\sum_j c_j k_j^\alpha}, \quad (6)$$

where $c_i = \sum_j A_{ij} k_j^\alpha$,

Using Eqs. (5) and (6), we check that the transition matrix is reversible and then has m real eigenvalues. From this stationary probability density, we can thus compute both the KSE and the second largest eigenvalue $\lambda(P)$ as a function of α. The result is provided in (Fig. 2).

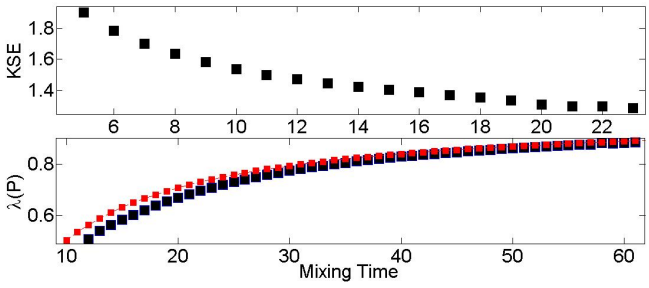


FIG. 1: Averaged KSE versus mixing time (top) for 10^6 random $m = 10$ size matrices and averaged $\lambda(P)$ versus mixing time (bottom) for 10^6 random $m = 10$ size matrices in curve blue and $f(t) = \epsilon^{1/t}$ in red. $\epsilon = 10^{-3}$ and the norm is chosen to be the euclidian one.

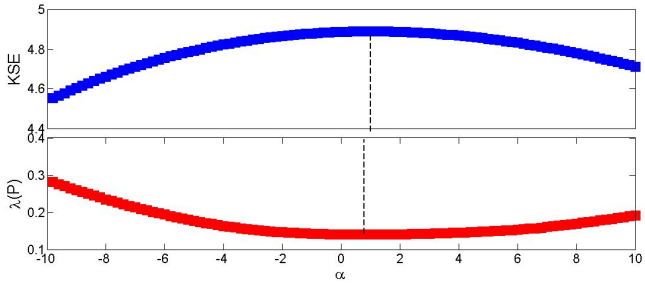


FIG. 2: KSE (top) and $\lambda(P)$ (bottom) function of α for a network of size $m = 400$ with a proportion of 0 in A equal to 1/3.

We observe in (Fig. 2) that the KS entropy has a maximum at a value that we denote α_{KS}, in agreement with the findings of [3]. Likewise, $\lambda(P)$ (i.e. the mixing time) presents a minimum for $\alpha = \alpha_{mix}$. Moreover, α_{KS} and α_{mix} are close. This means that the two optimal diffusion coefficients are close to each other. Furthermore, looking at the ends of the two curves, we can find two special Markov chains $P1$ and $P2$ such that $h_{KS}(P1) \leq h_{KS}(P2)$ and $t_1(\epsilon) \leq t_2(\epsilon)$, illustrating that

the link between KSE and the minimum mixing time is only true in a general statistical sense.

We have thus shown that, for a given transition matrix P (or equivalently for given jump rules) the greater the KSE, the smaller the mixing time. We now investigate whether a similar property holds for dynamics, i.e. whether transition rules that maximise KSE are close to the ones minimizing the mixing time. For a given network, i.e. for a fixed A, there is a well known procedure to compute the transition matrix P_{KS} which maximizes the KSE with the constraints $A(i,j) = 0 \Rightarrow P_{KS}(i,j) = 0$ [4]. It proceeds as follow: let us note λ the greatest eigenvalue of A and Ψ the normalized eigenvector associated i.e $A\Psi = \lambda\Psi$ and $\sum_i \Psi_i^2 = 1$. We define P_{KS} such that:

$$P_{KS}(i,j) = \frac{A(i,j)}{\lambda}\frac{\Psi_j}{\Psi_i}. \tag{7}$$

We have $\forall i \sum_j P_{KS}(i,j) = 1$. Moreover, using the fact that A is symmetric we find:

$$\sum_j P_{KS}(j,i)\Psi_j^2 = \sum_j \frac{A(j,i)\Psi_i\Psi_j}{\lambda} = \Psi_i^2. \tag{8}$$

Hence, $P_{KS}^t\Psi^2 = \Psi^2$ and the stationary density of P_{KS} is $\pi_{stat} = \Psi^2$.

Using Eqs. (3) and (7), we have:

$$h_{KS} = -\frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(\frac{A(i,j)}{\lambda}\frac{\Psi_i}{\Psi_j}). \tag{9}$$

Eq. (9) can be split in two terms:

$$\begin{aligned} h_{KS} &= \frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(\lambda) \\ &- \frac{1}{\lambda}\sum_{(i,j)} A(i,j)\Psi_i\Psi_j \log(A(i,j)\frac{\Psi_j}{\Psi_i}). \end{aligned} \tag{10}$$

The first term is equal to $\log(\lambda)$ because Ψ is an eigenvector of A and the second term is equal to 0 due to the symmetry of A. Thus:

$$h_{KS} = \log(\lambda). \tag{11}$$

Moreover, for a Markov chain the number of trajectories of length n is equal to $N_n = \sum_{(i,j)}(A^n)(i,j)$. For a Markov chain the KSE can be seen as the time derivative of the path entropy leading that KSE is maximal when the paths are equiprobable. For an asymptotic long time the maximal KSE is:

$$h_{KS_{max}} = \frac{\log(N_n)}{n} \to \log(\lambda), \tag{12}$$

by diagonalizing A. Using Eqs. (11) and (12) we find that P_{KS} defined as in Eq. (7) maximises the KSE. Finally P_{KS} verifies $\pi_{stat_i}P_{KS}(i,j) = \pi_{stat_j}P_{KS}(j,i)\ \forall\ (i,j)$ and thus P_{KS} is reversible.

In a similar way, we can search for a transition matrix P_{mix} which minimizes the mixing time -or, equivalently the transition matrix minimizing its second eigenvalue $\lambda(P)$. This problem is much more difficult to solve than the first one, given that the eigenvalues of P_{mix} can be complex. Nevertheless, two cases where the matrix P_{mix} is diagonalizable on $\mathbb{R}$ can be solved [6]: the case where P_{mix} is symmetric and the case where P_{mix} is reversible for a given fixed stationary distribution. Let us first consider the case where P is symmetric. The minimisation problem takes the following form:

$$\begin{cases} \min\limits_{P\in S_n} \lambda(P) \\ P(i,j) \geq 0, P * \mathbf{1} = \mathbf{1} \\ A(i,j) = 0 \Rightarrow P(i,j) = 0 \end{cases} \tag{13}$$

given the strict convexity of λ and the compactness of the stochastic matrices, this problem admits an unique solution.

P is symmetric thus $\mathbf{1}$ is an eigenvector associated with the largest eigenvalue of P. Then the eigenvectors associated to $\lambda(P)$ are in the orthogonal of $\mathbf{1}$.The orthogonal projection on $\mathbf{1}^\perp$ writes: $I_d - \frac{1}{n}\mathbf{1}\mathbf{1}^t$

Moreover, if we take the matrix norm associated with the euclidiean norm i.e. for M any matrix $|||M||| = \max \frac{||MX||_2}{||X||_2}\ X \in \mathbb{R}^n\ X \neq 0$ it is equal to the square root of the largest eigenvalue of MM^t and then if M is symmetric it is equal to $\lambda(M)$.

Then the minimization problem can be rewritten:

$$\begin{cases} \min\limits_{P\in S_n} |||(I_d - \frac{1}{n}\mathbf{1}\mathbf{1}^t)P(I_d - \frac{1}{n}\mathbf{1}\mathbf{1}^t)||| = |||P - \frac{1}{n}\mathbf{1}\mathbf{1}^t||| \\ P(i,j) \geq 0, P * \mathbf{1} = \mathbf{1} \\ A(i,j) = 0 \Rightarrow P(i,j) = 0 \end{cases} \tag{14}$$

We solve this constrained optimization problem (Karush-Kuhn-Tucker) with Matlab and we denote P_{mix} the matrix which minimizes this system.

We remark that the mixing time of P_{KS} is smaller than the mixing time of P_{mix}. This is coherent because in order to calculate P_{KS} we take the minimum on all the matrix space whereas to calculate P_{mix} we restrict us to the symmetric matrix space. Nevertheless, we can go a step further and calculate, the stationary distribution being fixed, the reversible matrix which minimizes the mixing time. If we note π the stationary measure and $\Pi = diag(\pi)$. Then P is reversible if and only if $\Pi P = \Pi^t P$. Then in particular $\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}$ is symmetric and has the same eigenvalues as Π. Finally, $p = (\sqrt{\pi_1}, ..., \sqrt{\pi_n})$ is an eigenvector of $\Pi^{\frac{1}{2}}P\Pi^{-\frac{1}{2}}$ associated to the eigenvalue

1. Then the minimization problem can be written as the following system:

$$\begin{cases} \min\limits_{P} |||(I_d - \frac{1}{n}\mathbf{q}\mathbf{q}^t)\Pi^{\frac{1}{2}} P \Pi^{-\frac{1}{2}}(I_d - \frac{1}{n}\mathbf{q}\mathbf{q}^t)||| \\ \qquad\qquad = |||\Pi^{\frac{1}{2}} P \Pi^{-\frac{1}{2}} - \frac{1}{n}\mathbf{q}\mathbf{q}^t||| \\ P(i,j) \geq 0, P * \mathbf{1} = \mathbf{1}, \Pi P = \Pi^t P \\ A(i,j) = 0 \Rightarrow P(i,j) = 0 \end{cases} \quad (15)$$

When we implement this problem in Matlab with $\pi = \pi_{KS}$ we find a matrix P_{mix} such that naturally $\lambda(P_{mix}) \leq \lambda(P_{KS})$. Moreover we can compare both dynamics by evaluating $|||P_{KS} - P_{mix}|||$ compared to $|||P_{KS}|||$ which is approximatively equal to $|||P_{mix}|||$. We remark that $|||P_{KS} - P_{mix}|||$ depends on the density ρ of 0 in the matrix A. For a density equal to 0 the matrices P_{KS} and P_{mix} are equal and the quantity $|||P_{KS} - P_{mix}|||$ will increase continuously when ρ increases. This is shown in (Fig. 3).

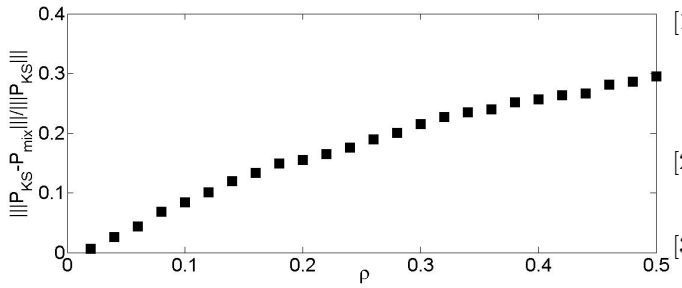


FIG. 3: $|||P_{KS} - P_{mix}|||/|||P_{KS}|||$ as a function of the density ρ of 0 present in A.

From this, we conclude that the rules which maximize the KSE are close to those which minimize the mixing time. This becomes increasingly accurate as the fraction of removed links in A is weaker. Since the calculation of P_{mix} quickly becomes tedious for quite large values of m, we offer here a much cheaper alternative by computing P_{KS} instead of P_{mix}.

Moreover, maximizing the KSE appears today as a method to describe out of equilibrium complex systems [8], to find natural behaviors [4] or to define optimal diffusion coefficients in diffusion networks. This general observation however provides a possible rationale for selection of stationary states in out-of-equilibrium physics: it seems reasonable that in a physical system with two simultaneous equiprobable possible dynamics, the final stationary state will be closer to the stationary state corresponding to the fastest dynamics (smallest mixing time). Through the link found in this letter, this state will correspond to a state of maximal KSE. If this is true, this would provide a more satisfying rule for selecting stationary states in complex systems such as climate than the maximization of the entropy production, as already suggested in [9].

Acknowledgments Martin Mihelich thanks IDEEX Paris-Saclay for financial support. Quentin Kral was supported by the French National Research Agency (ANR) through contract ANR-2010 BLAN-0505-01 (EXOZODI).

* Electronic address: martin.mihelich@cea.fr

[1] Prateek Bhakta, Sarah Miracle, Dana Randall, and Amanda Pascoe Streib. Mixing times of markov chains for self-organizing lists and biased permutations. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1–15. SIAM, 2013.

[2] Venkatesan Guruswami. Rapidly mixing markov chains: A comparison of techniques. *Available: cs. washington. edu/homes/venkat/pubs/papers. html*, 2000.

[3] Jesús Gómez-Gardeñes and Vito Latora. Entropy rate of diffusion processes on complex networks. *Physical Review E*, 78(6):065102, 2008.

[4] Z Burda, J Duda, J M Luck, and B Waclaw. Localization of the maximal entropy random walk. *Phys. Rev. Lett.*, 102(16):160602, 2009.

[5] P. Billingsley. *Ergodic theory and information*. Wiley, 1965.

[6] Stephen Boyd, Persi Diaconis, and Lin Xiao. Fastest mixing markov chain on a graph. *SIAM review*, 46(4):667–689, 2004.

[7] Pierre Bremaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. springer, 1999.

[8] C Monthus. Non-equilibrium steady states: maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints. *J. Stat. Mech.*, page P03008, 2011.

[9] Martin Mihelich, Bérengère Dubrulle, Didier Paillard, and Corentin Herbert. Maximum entropy production vs. kolmogorov-sinai entropy in a constrained asep model. *Entropy*, 16(2):1037–1046, 2014.

6 Pour aller plus loin : Maximisation de l'EKS et Minimisation du temps de mélange dans les systèmes dynamiques.

Nous avons relié le temps qu'une chaine de Markov prend pour atteindre son état stationnaire à l'entropie de KS. Il apparait donc naturel de se demander si une généralisation est possible pour des systèmes dynamiques plus généraux. Nous allons voir que le temps de mélange qui est bien défini pour les chaines de Markov ne trouve pas son analogue dans les systèmes dynamiques généraux. Ceci se comprend assez facilement pour deux raisons principales je pense : la première est que les systèmes dynamiques généraux admettent souvent plusieurs mesures stationnaires. De plus, même si le système admet une unique mesure stationnaire μ_{stat}, on peut très souvent montrer qu'il existe ϵ aussi petit que l'on veut tel que quel que soit M aussi grand que l'on veut on puisse trouver une mesure initiale μ_0 telle que $\forall\, n \leq M\, d(\mu_n, \mu_{stat}) \geq \epsilon$. Ainsi, le temps de mélange serait infini. Nous allons voir qu'il existe des notions assez similaires au temps de mélange pour les systèmes dynamiques généraux et notamment le "volume growth" [61]. Avant d'introduire cette quantité nous allons définir l'entropie topologique et la comparer à l'entropie métrique déjà étudiée dans le chapitre premier.

6.1 Entropie topologique et entropie métrique.

L'entropie topologique a été introduite après l'entropie métrique par Adler, Konheim et McAndrew. Néanmoins la définition suivante provient des travaux de Bowen et Dinaburg.

Soit $f : X \leftarrow X$ un système dynamique continu sur un espace compact. Nous supposerons que l'espace est métrique (muni d'une distance d) même si cette hypothèse n'est pas obligatoire.

Pour $\epsilon > 0$ et $n \in \mathbb{N}$, E est appelée une partie (n, ϵ)-séparée si :

$$\forall (x, y) \in E^2\, \exists i \in [|0, n|]\, , d(f^i(x), f^i(y)) \geq \epsilon. \tag{5.34}$$

Autrement dit, si nous nous fixons un coarse graining qui consiste à ne pas voir à une échelle plus petite que ϵ, une partie E est (n, ϵ)-séparée si toutes les orbites partant de points de E sont distinguables.

De plus, on note :

$$N(n, \epsilon) = \max_{E\ (n,\epsilon)-separee} \text{Cardinal} E. \tag{5.35}$$

On peut alors définir l'entropie topologique par :

$$h_{top}(f) = \lim_{\epsilon \to 0} (\limsup_{n \to \infty} \frac{1}{n} \log(N(n, \epsilon))). \tag{5.36}$$

Pour revenir à l'idée du coarse graining, si $N(n, \epsilon)$ représente le nombre d'orbite distinguable, alors il augmente comme $e^{n h_{top}(\epsilon)}$

L'entropie topologique et l'entropie métrique sont reliées par un principe variationnel.

Soit $f : X \leftarrow X$ qui est toujours un système dynamique continu sur un espace compact, on dit qu'une mesure μ est f-invariante si et seulement si $\forall A$ élément de la tribu mise sur X $\mu(f^{-1}(A)) = \mu(A)$. On a alors :

$$h_{top}(f) = \sup_{\mu \in \Gamma} h_\mu(f), \tag{5.37}$$

où $h_\mu(f)$ est l'entropie métrique et Γ est l'ensemble des mesures boréliennes f invariantes.

6.2 Volume growth (L.S Young)

Soit M une variété riemannienne C^∞ de dimension m et $f : M \leftarrow M$ un système dynamique de classe C^1. Pour l et k deux entiers strictement positifs, on note $\Sigma(k, l)$ l'ensemble des applications σ de classe C^k du cube unité de dimension l noté Q^l vers M. Notons de plus $w(\sigma)$ le volume (en dimension l) de $\sigma(Q^l)$. σ n'est pas forcément injective et le volume est donc calculé avec multiplicité c'est à dire que si deux parties de l'ensemble de départ donnent le même ensemble d'arrivé, on compte le volume de ce dernier deux fois. Pour $l \leq m$ définissons :

$$V_{l,k}(f) = \sup_{\sigma \in \Sigma(k,l)} \limsup_{n \to \infty} \frac{1}{n} \log w(f^n \sigma). \tag{5.38}$$

Notons enfin :

$$V(f) = \max_l V_{l,\infty}(f). \tag{5.39}$$

Ainsi $V(f)$ représente le taux de grossissement maximal des volumes sous l'action de f.

Ce qui est assez impressionnant c'est que l'on peut relier $V(f)$ à l'entropie topologique $h_{top}(f)$. En effet, un premier théorème dû à S. Newhouse [62] montre que si f est de classe $C^{1+\epsilon}$ avec $\epsilon > 0$ alors on a :

$$h(f) \leq V(f). \tag{5.40}$$

Un second théorème dû à Y. Yomdin [63] montre que si l'on note :

$$R(f) = \lim_{n \to \infty} \frac{1}{n} \log \max_{x \in M} ||Df^n(x)||, \tag{5.41}$$

où D est la différentielle, alors si f est de classe C^k on obtient :

$$V_{l,k}(f) \leq h_{top}(f) + \frac{2l}{k} R(f). \tag{5.42}$$

Et donc, si f est C^∞, grâce aux équations eq.(5.40) et eq.(**??**) on trouve :

$$h_{top}(f) = V(f). \tag{5.43}$$

Cette dernière équation nous montre bien le lien entre l'entropie topologique et la vitesse de grossissement des volumes sous l'action de f. Ainsi, qualitativement, plus l'entropie topologique est élevée, plus les volumes grossissent rapidement dans l'espace des phases et donc ceci nous amène à penser que plus le système atteint

rapidement son état stationnaire (ce qui est en accord avec les résultats trouvés dans l'article). Enfin, l'entropie métrique étant liée à l'entropie topologique par le principe variationnel eq.(5.37) ceci est aussi valable pour l'entropie métrique. Ceci nous amène à étudier les mesures μ qui maximisent $h_\mu(f)$.

Chapitre 6

Application de MEP à la convection

Les chapitres antérieurs étudient de façon théorique MEP. L'approche adoptée dans ce chapitre est différente, elle est beaucoup plus appliquée. En effet nous étudions ici l'application de MEP à la convection atmosphérique sans nous demander pourquoi MEP fonctionne. Cette partie de ma thèse a été faite au LSCE avec et grâce à Didier Paillard.

1 Introduction, position du problème.

La convection est le phénomène de mise en mouvement d'un fluide. Cette mise en mouvement peut être obtenue lorsqu'une différence de température est imposée ou encore un gradient de densité. Elle peut aussi être forcée par l'action de turbine par exemple. La question générale que l'on se pose dans cette partie est : **Est-ce que le Principe de Maximisation de Production d'entropie peut prédire la convection atmosphérique ?** La convection atmosphérique joue un très grand rôle dans la circulation atmosphérique. Elle est présente à petite échelle avec la formation de nuage de type cumulus ou cumulonimbus mais aussi à moyenne échelle avec la formation des orages, des moussons et des tornades. Elle est aussi présente à grande échelle sous le nom de cellules de Hadley. Les cellules de Hadley sont de gigantesques cellules convectives qui permettent la redistribution de l'énergie de l'équateur vers les pôles.

2 Outils nécessaires.

Comme il est habituel de le faire pour appliquer MEP en climatologie nous allons utiliser un modèle en boite [2, 64]. Nous allons donc modéliser une colonne atmosphérique par une succession de N boites qui peuvent échanger des flux de chaleur ainsi que des flux de matières Fig.(6.1). Il sera utile de modéliser ces boites par un graphe sur lequel les boites seront les nuds. Les échanges de flux entre les boites seront possibles que s'il existe un lien (côté du graphe) entre celles-ci. Nous allons introduire quelques outils sur les graphes qui vont nous permettre de rendre les calculs plus aisés. Je tiens à spécifier que nous ne nous intéressons pas ici à la dynamique et donc il n'y a pas de rapport direct avec les chaines de Markov.

2.1 Calcul sur un graphe

Soit un graphe $G = (V, E)$ où V est l'ensemble des sommets ($V = \{1, ..., N\}$) et E est l'ensemble des liens du graphe. On supposera le graphe orienté et donc E est un sous ensemble de $\{(i, j) \in [|1, N|]^2\}$, on notera M son cardinal et on numérotera les éléments de E. On peut alors définir la matrice S (start) de taille $M \times N$ par :

$$\begin{cases} S(i,j) = 1 \text{ si le côté } i \text{ a pour point de départ le sommet } j \\ S(i,j) = 0 \text{ sinon .} \end{cases} \tag{6.1}$$

De même, on peut définir la matrice T (target) de taille $M \times N$ par :

$$\begin{cases} T(i,j) = 1 \text{ si le côté } i \text{ a pour point d'arrivée le sommet } j \\ T(i,j) = 0 \text{ sinon .} \end{cases} \tag{6.2}$$

Ces deux matrices nous permettent de définir la matrice gradient ∇ de taille $M \times N$ définie par :

$$\nabla = T - S. \tag{6.3}$$

Ainsi, si $x = (x_i)$ $i \in [|1, N|]$ défini un champ sur V alors ∇x défini un champ sur E et correspond aux $x_j - x_i$. On retrouve ainsi le gradient habituel.

Si on regarde la transposée de cette matrice ∇^t elle correspond à la divergence. En effet, si $u = u_{i,j}$ est un champs sur E alors $\nabla^t u$ a pour i-ème composante $\sum_j u_{ji} - \sum_i u_{ij}$. C'est à dire la somme des flux entrant moins la somme des flux sortant.

On définit aussi la matrice Laplacien Δ de taille $M \times M$ par :

$$\Delta = \nabla\nabla^t. \tag{6.4}$$

On peut définir une advection-diffusion sur ce graphe à l'aide de deux champs vectoriels positifs sur E notés u^+ et u^-. u^+ est le flux de masse dans le sens positif,

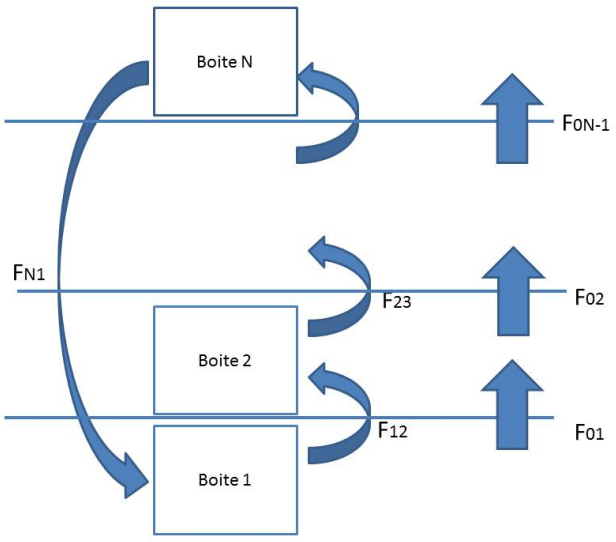


FIGURE 6.1 – Modèle en boite et les flux associés.

c'est à dire de i vers j avec $i < j$ et u^- est le flux de masse dans le sens négatif. Nous pouvons dès lors définir l'opérateur de transport par :

$$Tr_{(u^+,u^-)}x = Diag(Sx)u^+ - Diag(Tx)u^- = (Diag(u^+)S - Diag(u^-)T)x. \quad (6.5)$$

Le transport net de masse est défini par $u = (u_e)_{e\in E} = (u_e^+ - u_e^-)$.

On peut le décomposer en une partie diffusive $f_e = \min(u_e^+, u_e^-)$ et une partie advective $a_e = u_e^+ - f_e$.

Ainsi, lorsque le système est purement diffusif on a $u^+ = u^-$. Le but de ce chapitre est de trouver les valeurs des flux diffusifs et convectifs.

2.2 Résolution en Température.

Nous allons voir dans cette partie comment nous pouvons trouver la distribution de température $(T_1, T_2, ..., T_N)$ dans les boites. Ce travail a été en partie développé par C.Herbert [16]. Ici nous ne nous intéressons pas aux flux. Commençons par noter :

$$x_i = \frac{1}{T_i} \ \forall i \in [|1, N|]. \quad (6.6)$$

Une approximation linéaire du bilan local d'énergie se met sous la forme :

$$R(x) = R^0 + R.x, \quad (6.7)$$

où R est une matrice carrée [16]. On désigne ici par bilan local d'énergie la somme de toutes les énergies radiatives arrivant de l'extérieur. La production d'entropie est alors donnée par la forme quadratique suivante :

$$\sigma = x^t R(x) = x^t(R^0 + R.x). \quad (6.8)$$

Pour des valeurs de R et R^0 trouvées par C.Herbert la forme quadratique est définie négative. De plus, on veut maximiser σ sachant une conservation globale d'énergie : $\mathbf{1}^t(R^0 + R.x) = 0$. Ce problème peut être résolu facilement. On trouve alors le profil de température $T = 1/x$ en fonction de la boite Fig.(6.2).

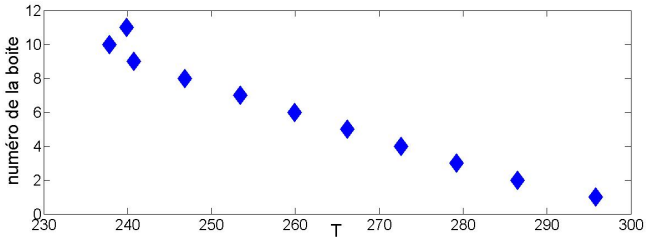


FIGURE 6.2 – Profil de la température en fonction de la boite.

2.3 Résolution en Flux.

Il peut être plus facile de résoudre le problème en flux plutôt qu'en température. En effet, si nous notons $F^0 = (F_1^0, F_2^0, ..., F_{N-1}^0)$ les flux de chaleur non pas entre les boites mais entre les étages Fig.(6.1), nous pouvons dès lors écrire :

$$R^0 + R.x = \nabla^t F^0, \tag{6.9}$$

où la matrice ∇ est une matrice de taille $(N-1) \times N$ définie par :

$$\begin{cases} \forall i \in [|1, N-1|] \nabla(i,i) = 1 \\ \forall i \in [|1, N-1|] \nabla(i,i+1) = -1. \end{cases} \tag{6.10}$$

La matrice ∇ correspond au graphe sur lequel il y a des liens entre i et $i+1$ et $i+1$ et i pour i allant de 1 à $N-1$.

L'avantage est que la condition $\mathbf{1}^t(R^0 + R.x)$ est directement vérifiée. Sachant que R est inversible on peut exprimer x en fonction de F^0 : $x = R^{-1}(\nabla^t F^0 - R^0)$. Puis la production d'entropie σ fonction de F^0 :

$$\sigma = ((F^0)^t \nabla - (R^0)^t)(R^t)^{-1} \nabla^t F^0, \tag{6.11}$$

que l'on peut réécrire :

$$\sigma = (F^0)^t A F^0 - b^t F^0, \tag{6.12}$$

où $A = \nabla(R^t)^{-1}\nabla^t$ et $b = (R^0)^t(R^t)^{-1}\nabla^t$.

La production d'entropie σ est encore une forme quadratique définie négative. Un calcul simple de la différentielle à partir de eq.(6.12) nous montre qu'elle admet un unique maximum en $(F^0)^*$ tel que :

$$(F^0)^* = (A + A^t)^{-1} b. \tag{6.13}$$

Sachant que $(A + A^t)$ est bien inversible. On peut calculer $(F^0)^*$.

2.4 Quelques cas particuliers.

Nous avons pris précédemment un schéma le plus simplifié possible où la boite numérotée i ne peut avoir d'échange qu'avec la boite $i+1$ et $i-1$, sauf les deux boites aux extrémités qui ne peuvent correspondre qu'avec une seule boite. Cependant nous pouvons envisager des modèles plus complexes. Nous pouvons par exemple envisager un modèle dans lequel, en plus des liens du modèle précédent, la boite numérotée 1 peut correspondre avec la boite numérotée N (Fig.(6.1)). Dans ce cas, la matrice ∇ est de taille $N \times N$ et prend la forme :

$$\begin{cases} \forall i \in [|1, N|] \, \nabla(i,i) = 1 \\ \forall i \in [|1, N-1|] \nabla(i,i+1) = -1 \text{ et } \nabla(N,1) = -1. \end{cases} \tag{6.14}$$

Ici, la matrice divergence ∇^t a un noyau non réduit à 0. On montre facilement que $\ker(\nabla^t) = vect((1,1,...,1)^t)$. Ainsi, il n'existe plus une unique solution au problème de maximisation de la production d'entropie σ (eq.(6.12)) mais l'ensemble des

solutions qui maximisent σ est un espace affine de dimension 1 qui est de la forme : $(F_0, 0)^t + \ker(\nabla^t)$.

De plus, nous pouvons remarquer une propriété importante qui sera très utile par la suite et qui ne dépend pas de la complexité du réseau choisi : tous les flux qui maximisent σ vérifient que la somme de ces flux entre les boites i et $i+1$ vaut F_i^0.

Ainsi, la méthode générale de recherche des flux entre les boites va être la suivante :

1. Calculer tout d'abord les flux (à travers les étages) F_0 à partir de l'équation (eq.(6.13)).
2. Choisir un réseau entre les boites, c'est à dire choisir quelles boites sont connectées entre elles.
3. Implémenter un programme de type recherche d'extrema sous contraintes d'égalités et d'inégalités (KKT) afin de déterminer les flux entre chaque boite.

3 Méthode KKT pour la résolution.

Dans cette section nous commencerons par définir la notion de **température potentielle équivalente** qui va être une notion fondamentale pour la suite. Ensuite nous expliciterons le problème d'optimisation avec ses contraintes et la méthode de résolution.

3.1 Température potentielle équivalente.

La température potentielle équivalente notée θ_e a été introduite pour pouvoir déterminer si une masse d'air à une certaine altitude est stable ou instable [65]. Elle est la température d'une masse d'air à laquelle on aurait enlevé adiabatiquement toute sa masse d'eau et que l'on aurait ramenée à la pression standard de référence P_0 ($P_0 = 100$ kilo Pascal).

Son expression analytique prend la forme :

$$\theta_e = T_e(\frac{P_0}{P})^\kappa, \tag{6.15}$$

où T_e est la température équivalente (température qu'aurait la masse d'air si on lui avait enlevé adiabatiquement sa masse d'eau) et $\kappa = \frac{R}{C_p} \approx 0.286$. De plus, la température équivalente prend la forme suivante :

$$T_e \approx T + \frac{L_v}{C_p} r, \tag{6.16}$$

où L_v est la chaleur latente d'évaporation, C_p est la chaleur spécifique de l'air sec à pression constante et r le rapport de la masse d'eau sur la masse d'air sec.

Ainsi, θ_e vaut :

$$\theta_e \approx (T + \frac{L_v}{C_p} r)(\frac{P_0}{P})^\kappa. \tag{6.17}$$

Lorsque θ_e augmente avec l'altitude ($\frac{d\theta_e}{dz} > 0$) la masse d'air est stable alors que lorsque θ_e diminue avec l'altitude ($\frac{d\theta_e}{dz} < 0$) la masse d'air est instable.

On remarque que θ_e est d'abord une fonction décroissante puis croissante de l'altitude. De plus, par un argument énergétique une masse d'air partant de la boite 1 ne pourra pas atteindre une boite dont la température potentielle θ_e a une valeur plus grande que $(\theta_e)_1$. On va donc considérer les échanges entre les boites $1, 2, ..., 8$. C'est à dire que l'on va considérer ici $N = 8$.

3.2 Problème explicite une colonne.

Nous avons commencé par envisager des schémas de graphe simples comme celui représenté sur la figure Fig.(6.1) mais il n'existe pas de solution cohérente, c'est à dire de solution vérifiant toutes les contraintes, sur de tels schémas. Nous avons donc considéré le problème le plus général possible, c'est à dire que chaque boite peut échanger des flux avec toutes les autres boites. Ainsi, si le système contient N boites alors le graphe associé aura N sommets et $N(N-1)$ arêtes orientées (on distingue l'arrête allant de i vers j et celle allant de j vers i et on considère qu'il n'y a pas d'arête entre i et i). Nous considèrerons tout d'abord un modèle constitué d'une colonne.

Le flux surfacique F_{ij} entre la boite i et la boite j prend la forme :

$$F_{ij} = u_{ij} C_p (\theta_e)_i, \tag{6.18}$$

où u_{ij} représente le flux surfacique de masse d'air allant de la boite i vers la boite j.

Comme nous l'avons mentionné dans le paragraphe précédent eq.(6.13) nous connaissons $F^0 = (F_1^0, F_2^0, ..., F_N^0)$. On rappelle que F^0 a été trouvé en utilisant MEP. Comme F_p^0 est la somme des flux passant entre le niveau de la boite p et de la boite $p+1$ on en déduit une première relation sur les flux qui lie F^0 aux F_{ij} :

$$\forall p \in [|1, N-1|] \; F_p^0 = \sum_{i \leq p \text{ et } j > p} F_{ij} - \sum_{l \leq p \text{ et } k > p} F_{kl}, \tag{6.19}$$

Et donc si l'on réécrit cette dernière équation en terme de flux de masse u_{ij} :

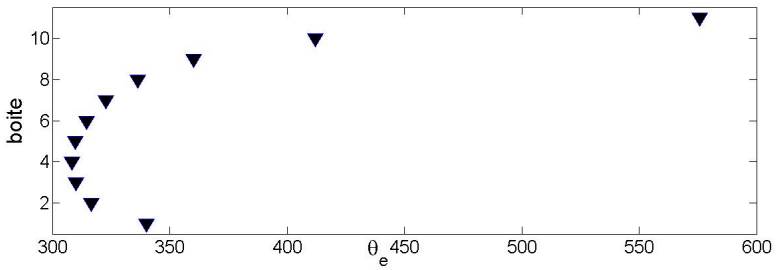


FIGURE 6.3 – Sur cette figure est tracée θ_e en abscisses et le numéro de la boite correspondante en ordonnée.

$$\forall p \in [|1, N-1|]\ F_p^0 = C_p \sum_{i \leq p \text{ et } j > p} u_{ij}(\theta_e)_i - C_p \sum_{l \leq p \text{ et } k > p} u_{kl}(\theta_e)_k, \tag{6.20}$$

Ces dernières relations nous donnent donc $N-1$ contraintes d'égalités linéaires.

De plus, la somme des flux de masse dans chaque boite doit être nul car nous sommes en régime stationnaire. Cette contrainte se traduit par, si $u = (u_{ij})$:

$$\nabla^t u = 0. \tag{6.21}$$

Ceci constitue N contraintes d'égalités linéaires supplémentaires.

De plus, les flux de masses doivent être positifs ce qui correspond aux N^2 contraintes d'inégalités :

$$\forall (i,j)\, u_{ij} > 0, \tag{6.22}$$

Enfin, les précipitations dans chaque boite i pour $i \geq 2$ (c'est à dire la somme de la masse d'eau qui entre moins la somme de celle qui part) doivent être positives. En effet, des précipitations positives correspondent à la possibilité d'avoir de la pluie. De plus, ceci n'est vrai qu'à partir de la seconde boite car la première boite est en contact avec la mer (ou le sol) qui peut donc lui donner un flux d'eau positif.

Ceci se réécrit comme $N-1$ contraintes d'inégalités linéaires :

$$\forall p \in [|2, N|]\ Pr(p) = \sum_{k \neq p} u_{kp} r(k) - \sum_{k \neq p} u_{pk} r(p) \geq 0, \tag{6.23}$$

où $r(k)$ est le rapport de la masse d'eau sur la masse d'air dans la boite k. Nous supposerons $r(k)$ connu. Plus exactement, nous prendrons comme expression de la pression de vapeur saturante :

$$es(T) = 6,112 e^{(17,67T/(T+243,5))},$$

et pour expression du rapport de mélange :

$$r(p,T) = es(T)/(p - es(T)).$$

Le but est donc de savoir si le problème défini par les équations eq.(6.19), eq.(6.20), eq.(6.21), eq.(6.22) et eq.(6.23) admet un domaine admissible. C'est à dire s'il existe u qui vérifie toutes ces équations. Notons (E_1) ce problème.
Nous verrons dans la partie résultats que ces équations admettent en effet un domaine admissible et que celui-ci est relativement restreint, c'est à dire que la structure des solutions est la même quelques soient les solutions du domaine. Ainsi, comme le domaine admissible n'est pas réduit à un point, il n'y a pas unicité de la solution mais comme le domaine est petit les solutions sont semblables. Néanmoins, nous verrons un critère qui nous permettra de ne retenir qu'une unique solution.

Nous avons envisagé ici un système unidimensionnel. Nous verrons dans la partie suivante que nous pouvons aussi envisager des systèmes 2D.

3.3 Problème explicite 2 colonnes et n colonnes.

Commençons par le système 2D le plus simple possible : 2 colonnes reliées entre elles par des liaisons aux extrémités Fig.(6.4). Chaque colonne est traitée comme dans la section précédente, c'est à dire que chaque boite de la première colonne peut interagir avec toutes les autres boites de cette même colonne. De même pour la seconde colonne. De plus, la boite 1 de la première colonne est liée à la boite 1 de la seconde et la boite N de la première colonne est liée à la boite N de la seconde. On dénombre donc en tout $2N(N-1)+4$ liaisons possibles.

On peut dès lors facilement ajuster les équations eq.(6.19), eq.(6.20), eq.(6.21), eq.(6.22) et eq.(6.23) en tenant compte du fait que les boites numérotées 1 et N des colonnes 1 et 2 sont reliées entre elles.

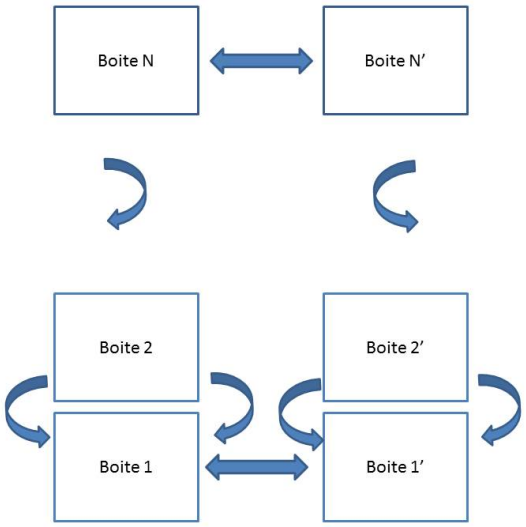


FIGURE 6.4 – Modèle en boite des deux colonnes.

De la même manière on peut envisager de mettre "en parallèle" n colonnes.

3.4 Principe permettant de trouver une unique solution parmi toutes les solutions du domaine admissible.

Le problème (E_1) n'admet pas (sauf cas exceptionnel) une unique solution. Il nous faut donc un critère permettant de n'en retenir qu'une seule. Pour un flux F de $\mathbb{R}^{N(N-1)}$ associé à u solution de E_1, nous pouvons décomposer F sous la forme suivante :

$$F = F_{//} + F_{\perp}, \tag{6.24}$$

où $F_{//}$ appartient à $\ker(\nabla)$ et $F_{\perp}$ appartient à l'orthogonal de $\ker(\nabla)$. $F_{//}$ représente la partie de F qui tourne.

Ainsi il parait judicieux de choisir comme critère :

$$u \text{ tel que } ||F_{//}|| \text{ soit minimal,} \tag{6.25}$$

où $||.||$ désigne la norme euclidienne. Nous noterons ce problème le problème (E).

Si (e_i) est une base orthonormée de $\ker(\nabla)$ alors on a :

$$||F_{//}||^2 = \sum_i < F, e_i >^2, \tag{6.26}$$

et si E_i représente e_i dans $\mathbb{R}^N$ on obtient :

$$||F_{//}||^2 = F^t(\sum_i E_i E_i^t)F. \tag{6.27}$$

De plus, si l'on note $D = C_p Diag(\theta_e)$ on a dès lors $F = C_p Diag(\theta_e)u$ et par conséquent en injectant cette expression dans l'équation eq.(6.28) on obtient :

$$||F_{//}||^2 = u^t D(\sum_i E_i E_i^t)Du. \tag{6.28}$$

Le problème (E) revient donc à résoudre un problème de minimisation d'une fonction quadratique sachant des conditions d'égalités et d'inégalités linéaires. Ceci est donc un problème d'optimisation quadratique. Nous résolvons ce problème avec Matlab.

4 Résultats.

Nous présentons ici les résultats obtenus, c'est à dire la solution vérifiant le problème (E) pour le modèle 1 colonne puis pour le modèle 2 colonnes.

4.1 Modèle 1 colonne.

Nous nous intéressons ici à une colonne vertical située au niveau de l'Équateur.

Nous présentons les résultats des flux F_{ij} (en Watt) dans la matrice $M = (m_{ij} = F_{ij})$ pour i et j allant de 1 à 8.

$$M = \begin{pmatrix} 0 & 14 & 5 & 0 & 0 & 0 & 0 & 2110 \\ 0 & 0 & 13 & 11 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 18 & 15 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 22 & 31 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 26 & 48 & 0 \\ 50 & 0 & 0 & 0 & 0 & 0 & 10 & 0 \\ 460 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1540 & 11 & 18 & 27 & 40 & 92 & 324 & 0 \end{pmatrix} \tag{6.29}$$

Nous remarquons dans la matrice (eq.(6.31)) que les flux prédominants sont les flux F_{18} et F_{81}. F_{18} correspond à un updraft c'est à dire un courant ascendant d'air directement de la boite 1 vers la boite 8 et F_{81} correspond lui à un downdraft c'est à dire un courant d'air descendant directement de la boite 8 vers la boite 1. Il existe

aussi d'autres flux non négligeables comme F_{71} et F_{87}.

Pour ce qui est des précipitations, on trouve que toutes les précipitations proviennent de la 8ème et dernière boite. De plus, en utilisant u solution du problème (E) on trouve que les précipitations correspondent à 3.2 mètres d'eau par année ce qui n'est pas si loin de la valeur réelle de la hauteur des précipitations en Équateur qui vaut 2.3 mètres.

4.2 Modèle 2 colonnes.

Pour le modèle 2 colonnes nous avons pris une colonne qui se situe à l'équateur et l'autre au pôle.

Nous présentons les résultats des flux F_{ij} (en Watt) de la colonne 1 dans la matrice $M_1 = (m_{ij} = F_{ij})$ pour i et j allant de 1 à 8. Puis nous présentons les flux de la colonne 2 dans la matrice M_2.

$$M_1 = \begin{pmatrix} 0 & 12 & 4 & 0 & 0 & 0 & 0 & 25335 \\ 0 & 0 & 11 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 19 & 11 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 29 & 22 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 51 & 25 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 160 & 0 \\ 511 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 24758 & 9 & 16 & 25 & 39 & 91 & 334 & 0 \end{pmatrix} \tag{6.30}$$

$$M_2 = \begin{pmatrix} 0 & 7 & 12 & 14 & 46 & 385 & 5 & 41 \\ 10 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 43 & 0 & 0 & 0 & 0 & 0 & 0 & 83 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 133 \\ 133 & 5 & 14 & 26 & 0 & 0 & 0 & 0 \\ 178 & 0 & 13 & 48 & 138 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 210 \\ 0 & 0 & 0 & 0 & 0 & 0 & 215 & 0 \end{pmatrix} \tag{6.31}$$

Enfin, le flux $F_{1^1 \to 1^2} = 62$; $F_{1^2 \to 1^1} = 112$;$F_{8^1 \to 8^2} = 70$;$F_{8^2 \to 8^1} = 16$

Nous pouvons remarquer que les flux à l'équateur sont plus importants que les flux qui se situent dans la colonne au pole ce qui est cohérent. Néanmoins les valeurs des flux $F_{1 \to 8}$ et $F_{8 \to 1}$ dans la colonne 1 sont élevés par rapport aux résultats 1 colonne.

5 Discussion

Nous avons choisi dans ce chapitre d'utiliser la démarche en deux étapes suivante :

- Appliquer MEP à un modèle en boites pour déduire la distribution de température ainsi que la divergence des flux dans les boites et la somme des flux à travers chaque niveau.
- Grâce à une méthode d'optimisation sous contraintes nous déterminons les flux entre les boites. Ceci nous donne en particulier les valeurs des précipitations.

Néanmoins on pourrait penser à une méthode alternative qui consisterait à maximiser MEP sous les contraintes voulues. Cette méthode serait différente car nous avons implicitement considéré ici 8 boites, ce nombre de boites correspond à la dernière boite pour laquelle $(\theta_e)_8 \leq (\theta_e)_1$ comme nous pouvons le voir sur la figure Fig.(6.3). Cependant nous aurions pu considérer la totalité des boites, c'est à dire dans notre cas 11 boites. Mais alors le domaine admissible aurait été réduit à l'ensemble vide.

Ainsi, la prochaine étape de ce travail serait d'utiliser cette nouvelle méthode et de la comparer aux résultats précédents.

6 Conclusion

Dans ce chapitre nous avons essayé de voir si MEP pouvait d'écrire la convection atmosphérique. L'avantage de MEP est qu'il ne requiert pas de paramètre ajustable et ne s'intéresse pas à la dynamique qui est dans ce type de problème extrêmement complexe.

Nous avons démontré que MEP prédisait la présence d'updraft et de downdraft ce qui constitue une "surprise" dans la mesure où la dynamique n'est pas prise en compte. De plus les flux ainsi que les précipitations prédits dans le modèle 1 colonne sont proches des valeurs réelles.

7 Conclusion générale

7.1 Conclusion

Le premier jour lorsque je suis arrivé au laboratoire mes directeurs de thèse m'ont posé la question suivante : Est-il possible de mieux comprendre le principe de maximisation de production d'entropie (MEP). "Mieux comprendre" signifie ici essayer de donner une justification à ce principe, principe qui est par ailleurs souvent contesté. "Mieux comprendre" ce principe c'est aussi délimiter son domaine de validité, lui donner un sens physique, le relier à d'autres principes existant... En effet, MEP est utilisé sans qu'aucune justification rigoureuse n'en soit donnée. Quelques chercheurs ont essayé de comprendre de façon théorique ce principe mais les preuves n'étaient pas exactes et des contres exemples ont été trouvé. De plus, comme le dit R. Goody dans [66] ainsi que A. Kleidon dans [67], pour pouvoir l'appliquer au climat dans de multiples situations il faudrait le comprendre théoriquement de façon plus satisfaisante. Ceci constitue l'objet principal de cette thèse : essayer de comprendre MEP. Je ne peux pas affirmer que nous avons résolu totalement l'énigme MEP mais nous avons apporté dans cette thèse quelques éléments de réponse ; je pense.

Beaucoup de chercheurs contestent le principe de maximisation de production d'entropie car ils voient une contradiction entre ce dernier et le principe de Prigogine qui minimise la production d'entropie. Le premier résultat de cette thèse (voir chapitre 1) a été d'interpréter MEP comme un prolongement du principe de Prigogine. En effet, nous expliquons comment MEP apparait non plus comme un maximum mais comme un MinMax.

Par ailleurs, MEP est souvent utilisé dans des systèmes extrèmement complexes (climat, océan, sociaux...), notre démarche (qui est nouvelle dans un certain sens) a été de l'étudier sur des systèmes assez simples de la physique statistique pour mieux le comprendre. C'est ainsi que dans l'Asymmetric Simple Exclusion Process (ASEP) nous avons relié numériquement MEP à un autre principe qui est le principe de maximisation de l'entropie de Kolmogorov-Sinaï (chapitre 2). De plus, nous avons vu que MEP donnait des résultats physiquement cohérent dans la mesure ou proche de l'équilibre l'état choisit par MEP est purement diffusif alors que plus le système est hors équilibre plus l'état devient convectif.

Nous avons ensuite démontré analytiquement le lien entre MEP et le principe de maximisation de l'entropie de Kolmogorov Sinai (MKS) dans un système jouet où les calculs se révèlent être assez commodes : le Zero Range Process (chapitre 3). Nous discutons ce résultat en terme de coarse graining optimal.

Puis nous avons démontré que maximiser la production d'entropie revenait à minimiser le mixing time c'est à dire le temps que le système met pour atteindre son état stationnaire (chapitre 5). Ceci constitue un autre résultat important de cette thèse et donne un sens physique à MEP.

Enfin nous avons travaillé avec une approche plus appliquée et moins théorique pour étudier l'application de MEP à la convection atmosphérique (chapitre 6).

7.2 Perspectives et travaux futurs.

Ce travail ouvre sur de nombreuses perspectives :

Une des plus importantes est la généralisation du lien entre MEP et MKS dans des systèmes dynamiques plus généraux que seuls les systèmes markoviens. Une question naturelle est donc de savoir si l'on peut démontrer analytiquement le lien entre MEP et MKS pour des systèmes dynamiques quelconques ?

Une autre ouverture est de comprendre les intéractions entre la production d'entropie et le coarse graining. Plus précisément il serait intéressant de poursuivre l'étude faite dans le chapitre 4 sur la nouvelle manière de coarse grainer.

Il serait aussi intéressant d'étendre la relation trouvée entre MEP et la minimisation du mixing time. Pour cela il faudrait tout d'abord définir une notion de mixing time pour des systèmes dynamiques généraux ce qui n'a pas encore été fait. Une perspective intéressante serait de comparer ces résultats purement théoriques aux simulations. On pourrait penser à utiliser un code modélisant le système climatique terrestre, arrêter les forçages extérieurs puis les réappliquer et regarder si effectivement plus la production d'entropie est élevée plus le système atteint rapidement son état stationnaire.

Enfin il reste du travail concernant l'application de MEP à la convection atmosphérique. Notamment la généralisation à n colonnes et l'optimisation sous contraintes directement à partir de l'expression de la production d'entropie.

En espérant que des travaux futures nous donne les solutions à ces énigmes...

1 Annexes

Nous présentons ici l'article évoqué dans le chapitre 4 sur le modèle d'Ising hors équilibre.

A non equilibrium Ising model of turbulence

Davide Faranda,* Simon Thalabard, Martin Mihelich, Brice Saint-Michel, and Bérengère Dubrulle
Laboratoire SPHYNX, Service de Physique de l'Etat Condensé, DSM,
CEA Saclay, CNRS URA 2464, 91191 Gif-sur-Yvette, France

In this paper we introduce a model of interacting lattices at different resolutions driven by the 2-D Ising dynamics with a nearest-neighbor interaction. We study this model both with tools borrowed from equilibrium statistical mechanics as well as non-equilibrium thermodynamics. We find that this model keeps the signature of the equilibrium phase transition. Moreover the critical temperature of the equilibrium models correspond to the state maximizing the entropy and delimitates two out of equilibrium regimes, one satisfying the Onsager relations for systems close to equilibrium and one resembling convecting turbulent states. Since the model preserves the fluxes of entropy and energy in the scale space, it seems a good candidate for parametric studies of out of equilibrium turbulent systems.

I. INTRODUCTION

Equilibrium systems can often be described through well founded statistical mechanics methods, built up from the classical laws of thermodynamics. In this sense, they represent an exception with respect to most systems found in nature, that are subject to flux of matter and energy to and from other systems and/or to chemical reactions. As such, the latter are intrinsically out-of-equilibrium and their description remains beyond the scope of present statistical mechanics [1]. Among all out-of-equilibrium systems, we may single out those leading to (out-of-equilibrium) steady states, the so-called NESS. These states have on average constant energy, and a natural question is how much they differ from the equilibrium states of the corresponding ideal system (without say, forcing and dissipation). For example, there is now strong evidence that the large scales of the steady states of 2D or quasi-2D turbulent system can be described as equilibria states of the corresponding Euler equation [2, 3], while there is a non-zero energy or enstrophy flux towards smaller scale, characteristic of out-of-equilibrium dynamics. On the other hand, there are several examples of steady states of out-of-equilibrium systems that do not correspond to any equilibrium states, like for an Ising model driven by a temperature gradient [4], or quantum open system in contact with two reservoir [5]. One way to get some insights about why and how NESS approach their equilibrium counterpart is via the study of simple toy model, that are well controlled and analytically tractable. The most studied case is probably the ASEP or the zero range process on a lattice [**?**], mimicking diffusion processes. This case is however not appropriate to understand the turbulent NESS, because it does not include an essential ingredient that is the development of an energy cascade over a wide range of scales. This was recognized by [6], who devised a dissipative forced zero range process on a Cawley tree mimicking the hierarchy of scale. Varying a parameter characterizing the transfer scale-to-scale dynamics, they observed a transition between a quasi-equilibrated regime and a far-from equilibrium one, with non-zero energy flux. However, this example does not reproduce the observed turbulent case, where both the non-zero energy flux and the equilibrium state coexist.

In the present letter, we explore a new toy model of "turbulent systems" based on the Ising model [7, 8]. It approximates magnetic dipole moments of atomic spins through discrete variables with Boolean-like distribution (+1 or 1), interacting with their nearest neighbors onto a lattice. The properties of the model at the thermal equilibrium are well established and allow for the identification of second order phase transitions in 2D [9, 10]. The idea of exploring the existence of non-equilibrium steady states (NESS) of the Ising model as a paradigm for other phenomena hails from the work of Glauber [11], who devised one of the simplest non-equilibriumn models mimicking the dynamics of a thermostatted system. Other modifications of the Ising model in the same spirit have been suggested in [4, 12–16] where the addition of a temperature gradient within the lattice sets the system out of equilibrium. However, none of these modification of the ising model so far contain the dynamics of energy or enstrophy flux through scales. In order to take into account this feature, we propose a model of two interacting lattices A and B at different temperatures. Lattice A is the reference one, its dynamics is iterated at the temperature T_A, then the configuration is copied to the lattice B and an iteration is performed at T_B. Finally the first lattice is updated with the changes maded on the lattice B and currents are computed as the number of spin changed within

*Electronic address: davide.faranda@cea.fr

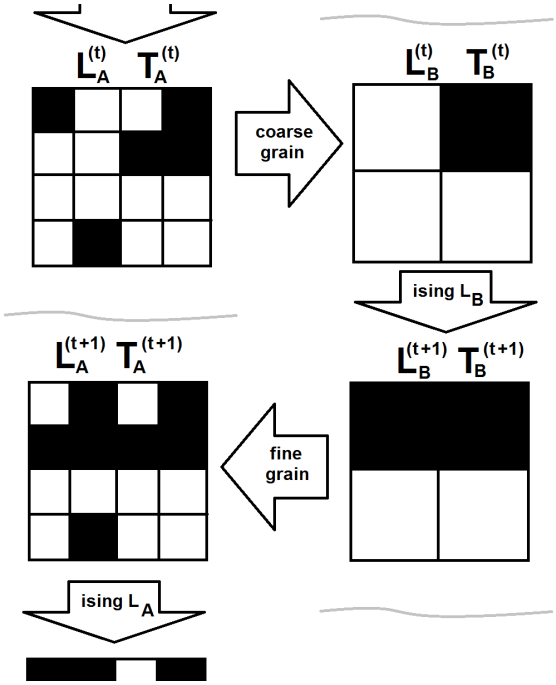


FIG. 1: Schematic representation of the two lattices model. In this representation the size of L_B is smaller than the one of L_A. Vertical steps are dynamical and related to the iteration of the Ising dynamics at temperature T_A or T_B. The horizontal steps describe the way the operations of coarse grain and fine grain are performed.

these dynamical steps. Such a procedure allows also for checking the role of coarse graining. This is simply realized by downgrading the resolution of lattice B of a certain factor.

In this model, we perform two different analysis. First, by using the equilibrium tools of classical statistical mechanics we explore whether the out of equlibrium dynamics keeps trace of the phase transitions observed in the classical equilibrium ising dynamics. Then we use no-equilibrium tools to study the thermodynamics of the system. We will show that coarse graining effects alone introduce non equilibrium dynamics on the model and that such dynamics depend both on the temperature gradient between the two lattices A and B but also on the temperature T_A of the fine grained grid.

The paper is organized as follows: in Section II we introduce the two lattices model. Then, in section III we study the systems with tools of equilibrium statistical mechanics. Section **??** contains the out of equilbirum analysis. Finally, in Section V we discuss the possible physical implications of our findings.

II. MODEL OF TWO INTERACTING LATTICES AT DIFFERENT TEMPERATURES.

We introduce a two lattices model A and B, each of them obeying the classical 2-D Ising dynamics with a nearest-neighbor interaction on a square lattice. The two lattices are at different temperature T_A and T_B and may be at two different resolutions L_A and L_B. This will be useful to analyse the non-equilibrium effects linked to a coarse grained dynamics. A schematic representation of the dynamics of the model is reported in Figure 1. The figure shows that first the lattice A is updated with a Monte Carlo step at temperature A, than, if $L_A \neq L_B$, the coarse grain operation is performed: for each spin of the lattice at lower resolution we count the number of corresponding spins in at full resolution and, if the majority is positive (black in the figure) we assign a positive spin to the corresponding coarse grain position. We then perform a Monte Carlo step at the temperature T_B on the lattice L_B. Eventually we update the grid L_A by reversing all the spins which correspond to a change in the lattice L_B.
In order to track non equilibrium effects, we have to introduce a suitable definition of a current $j(t)$. Although there is not an unique way to define currents, the most natural one is to compute them as the number of spins changed on the grid A as an effect of the operations (coarse grain and Monte Carlo steps) performed on the grid B at temperature T_B.

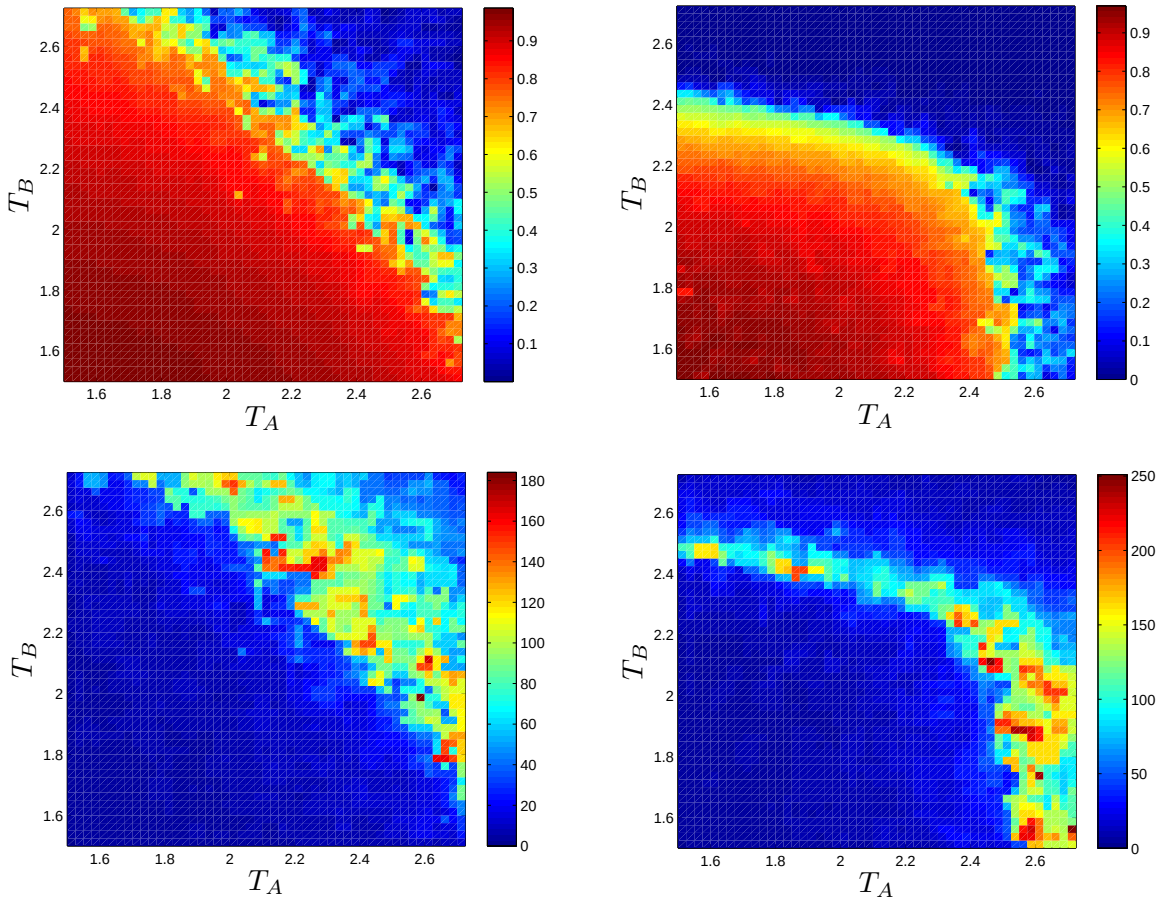


FIG. 2: Average magnetization $\langle \bar{M} \rangle$ (upper panels) and first zero of the autocorrelation function $C(\tau)$ as a function of the temperatures T_A and T_B. Left panel: lattice without coarse grain $L_A = L_B = 64$. Right panel: coarse grain with $L_A = 64, L_B = 8$

In the following we will describe the properties of such system, the role of non-equilibrium effects and of the phase transitions. We will see that there is no trivial way to completely describe the non-equilibrium effects by using the statistical tools devised for equilibrium dynamics, but relevant quantities of equilibrium dynamics still play an important role when the system is set out of equilibrium.

III. EQUILIBRIUM ANALYSIS AND PHASE TRANSITIONS

We explored the behavior of the systems for different temperature T_A and T_B and different sizes L_A and L_B from $L = 8$ to $L = 256$. On this range we have found that results do not qualitatively depend on the resolution. Therefore, in the following, we will consider two typical experiments: one where the two grids are at the same resolution but at two different temperatures $L_A = L_B = 64$ and another one where the two grids are at different temperatures and resolutions $L_A = 64$ and $L_B = 8$. For each experiments we compute the relevant quantities of the Ising dynamics on 1000 steps as described in Fig. 1. The temperature range analyzed is $(T_A, T_B) \in [T_c - 1, T_c + 1]$.
The first thing we observe is that the system sets very quickly in a stationary state where the magnetization and the energy fluctuate around a well defined value. It is interesting to see whether the system still displays a phase transition at a certain temperature and how such transition depends on the coarse grain. In order to do so, for each experiment we have computed the average magnetization $\langle \bar{M} \rangle$, and the first zero of the autocorrelation function $C(\tau)$ of the time series of magnetization. Both these quantities are represented in Fig. 2 for the non coarse grained experiment (left plots) and the coarse grained experiments (right).

We first remark that the behavior of the absolute value of $\langle \bar{M} \rangle$ (upper panels of Fig. 2) reproduces the one observed

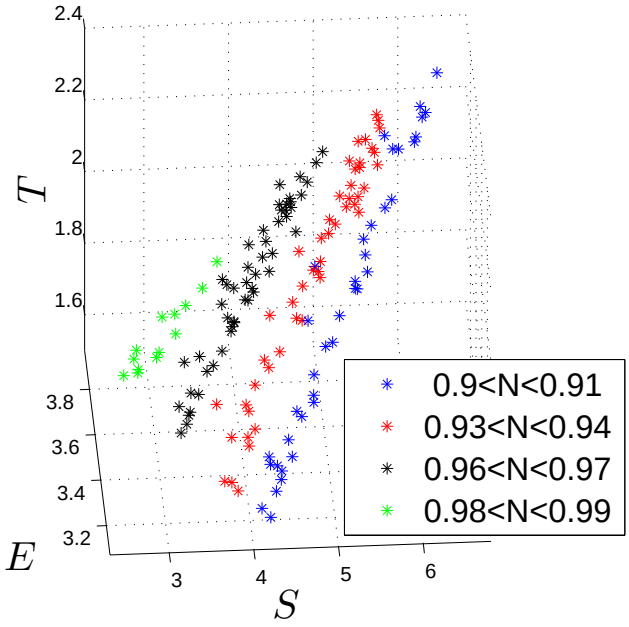


FIG. 3: Entropy of the configuration S vs Energy E for the coarse graining with $L_A = 64, L_B = 8$. The color scale represents the currents j (upper panel) and the temperature T_A (lower panel) . See text for an explanation.

for the classical Ising dynamics but with a higher pseudo-critical temperature $T_c(T_A, T_B)$. When no coarse grain is applied to the grid L_B the picture looks symmetric with respect T_A and T_B. In this sense, one could define $T_c(T_A, T_B)$ by following the curve on the T_A, T_B plane where $\langle \bar{M} \rangle$ changes abruptly. On the other hand, when the resolution of L_B is decreased (top right panel of the same figure), this curve does not look symmetric with respect to T_A and T_B. The transition looks sharper when the temperature of the coarse grained lattice L_B is varied. Our model also respects the symmetry of the magnetization with respect to zero: we obtain with equal probability positive or negative magnetizations.
In order to describe the phase transition we also verified that the change of the magnetization corresponds to an effective increase of correlation among the spins by computing the first zero of $C(\tau)$, reported in the lower panels of Fig. 2. For both the cases we observe an increase of such a quantity in a region of the (T_A, T_B) plane which can be directly used to identify T_c. Such region corresponds to the one identified by tracking the sharpest variation of the average magnetization $\langle \bar{M} \rangle$.

Now we turn to analyze the behavior of the temperature as a function of the entropy S and the energy E. In fact, as predicted by Onsager, if the system remains close to equilibrium, one must find: $dE/dS|_N = T_A$. The entropy used here is the usual Shannon Entropy computed in information theory which measures the *disorder* of the lattice L_A at the end of the simulation:

$$S = -\sum_{i,j} p(i,j) log_2(p(i,j))$$

with $p(i,j)$ the probability of observing a spin up or down at the position (i,j). The validity of such a relation can be effectively observed in Fig. 3 where the quantities E, S, T_A are plotted for three different values of N.

In the next section by considering thermodynamic quantities such as the entropy and the currents, we will perform a basic non equilibrium analysis and understand whether such a phase transition is still relevant when the system is set out of equilibrium.

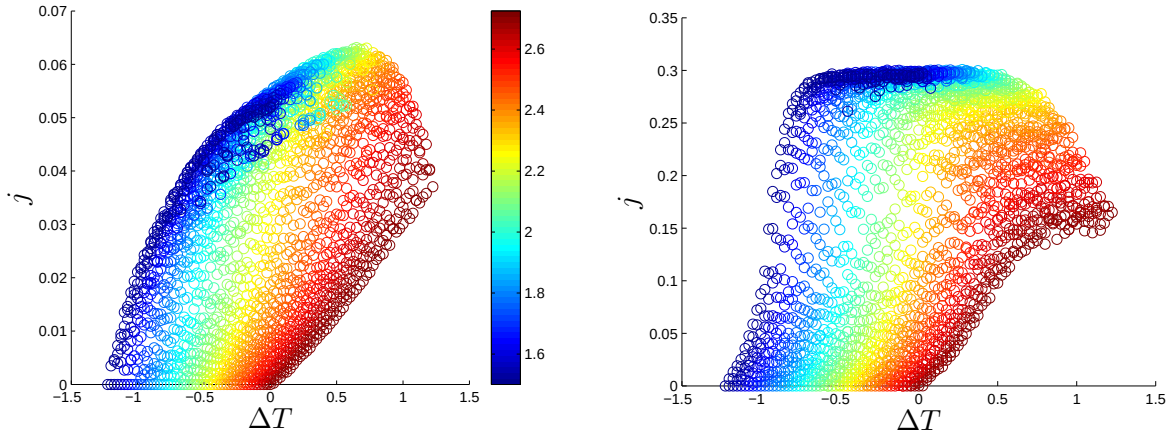


FIG. 4: Current j between the grid B and A, as a function of the temperature gradient $\Delta_T = T_B - T_A$. The colorscale represents temperature T_B. Left panel: lattice without coarse grain $L_A = L_B = 64$. Right panel: coarse grain with $L_A = 64, L_B = 8$

IV. NON-EQUILIBRIUM ANALYSIS

We start the analysis by verifying that the model proposed in this section possesses a genuine non equilibrium dynamics and then comment on some interesting thermodynamic properties of the system. We will use the definition of currents already described above, i.e. we define j as the number of spins changed on the lattice A after the steps performed on lattice B. We renormalize j to include the effects of the temperature gradient only.
In Fig. 4 we present the behavior of j for the two experiments detailed before: no coarse grain on left panel and coarse grain on the right. Colors represent the temperature T_B. For the non coarse grained dynamics we immediately see that, for each temperature - each color - T_B the current j increases proportionally to Δ_T. This is the signature of the Fourier law when the system is driven out of equilibrium and the *heat transfer* can be considered conductive. However, for higher temperature gradients and depending on T_B this regime is broken and currents saturate at a specific value. This happens also for the coarse grained dynamics (right panel) where such behavior is amplified and the saturation is indeed observed at all temperatures T_B.

Another important test is the existence of a maximum of the entropy production and the relation of such a maximum with the temperatures, the currents and the conductive/convective regimes. In fact, as pointed out in [17], for other out of equilibrium models such as the zero range process, a maximum of entropy production is clearly recognizable and falls between this two regimes. In our model we will define the macroscopic Entropy Production σ as the product of the fluxes by the thermodynamics forces. It is important to point out that this entropy is different from the microscopic entropy defined above. In terms of currents and temperatures the Entropy Production is written as:

$$\sigma = j\left(\frac{1}{T_B} - \frac{1}{T_A}\right), \tag{1}$$

where j is the current. Figure 5 represents σ as a function of the currents j and it is colored according to T_A (upper panels) and T_B (lower panels). We analyse first the non coarse grained dynamics (on the left panels) where maxima of σ are observed for each curve of T_B. Moreover, each curve has the maximum located at almost the same temperature T_A which roughly correspond to the critical temperature of the equilibrium Ising model T_c. Moreover, the maxima always fall in between the *conductive* and *convective* regimes. This fact is even more evident when analyzing the plots for the coarse grained dynamics (right panels) as the separation between these two regimes correspond to a sharp change of σ especially at lower temperature T_B. Also in this case, the location of the maximum corresponds to the critical temperature T_c.

This latter fact is indeed relevant for two reasons: first, it points to the fact that equilibrium properties and relevant quantities of equilibrium dynamics, e.g the critical temperature T_C, still play an important role when the system is set out of equilibrium. Second it points to a sort of thermodynamic explanation of the phenomenon of self organized criticality widely invoked to explain the fact that many physical systems set into non equilibrium states close to their critical temperature [18, 19]. Here the states obtained following a maximum entropy production principle coincide

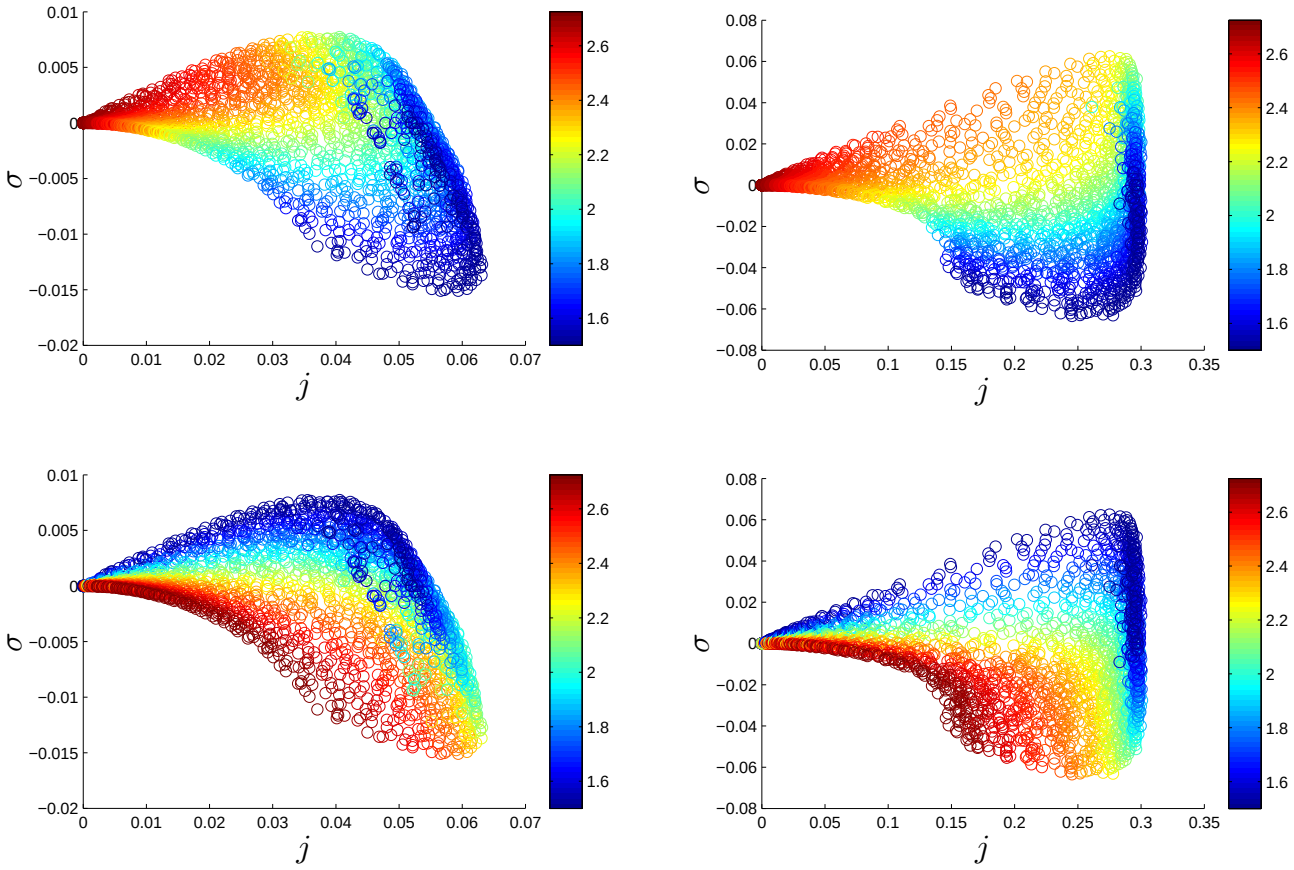


FIG. 5: Entropy production σ as a function of the currents j . Left panel: attice without coarse grain $L_A = L_B = 64$. Right panel: coarse grain with $L_A = 64, L_B = 8$. The color scale represents T_A (upper panels) and T_B (lower panels).

with the ones obtained invoking the self organized criticality.
Finally, it is interesting to study the distribution of σ in the plane (T_A, T_B) as shown in Fig. 6. Note that the scales of σ for the non coarse grained dynamics (left panel) is different from the one of the coarse grain simulations (right panels). The latter shows a sharper gradient of σ with respect to the non coarse grained case. This is consistent with what observed in Fig. 5.
The previous analysis seems to suggest that the coarse graining introduces a non smooth change of the thermodynamic properties of the system. In order to analyse more in details such an effect we introduce a cascade of coarse grain operations and build a model of interacting lattices in the scale space. In the numerical experiment, we start from a full resolution of $L_A = 64$ and halve several times the resolution. At each scale we perform a dynamic Ising step. The intermediate resolutions considered are $L_B = 32$, $L_C = 16$ and $L_D = 8$. We update the lattice L_A after the dynamical steps performed on L_D.

We first verify that the entropy production is constant through the scale space. In order to do so we compare several values of σ obtained at different change in resolutions. For example, we can compare $\sigma_{A\to B}$ with $\sigma_{B\to C}$ and $\sigma_{C\to D}$ where the resolution has changed by half, $\sigma_{A\to C}$ with $\sigma_{B\to D}$ and compare all of them with $\sigma_{A\to D}$ where the resolution changed by a factor 8. Results are shown in Fig. 7. One can observe that there is no substantial difference among the entropy productions defined. This is property is therefore reminiscent of the constant flux of energy observed in turbulence and necessary to explain the famous $k^{-5/3}$ energy spectrum. However if one looks at the shape of such curves and compare them with the ones shown in Fig. 5 several differences emerge. In the previous analysis we have performed an abrupt coarse graining shrinking the resolution by a factor 8 in one go. Here we have gone trough several operations where the resolution has been halved. As a result of this smoother procedure, the curves of entropy

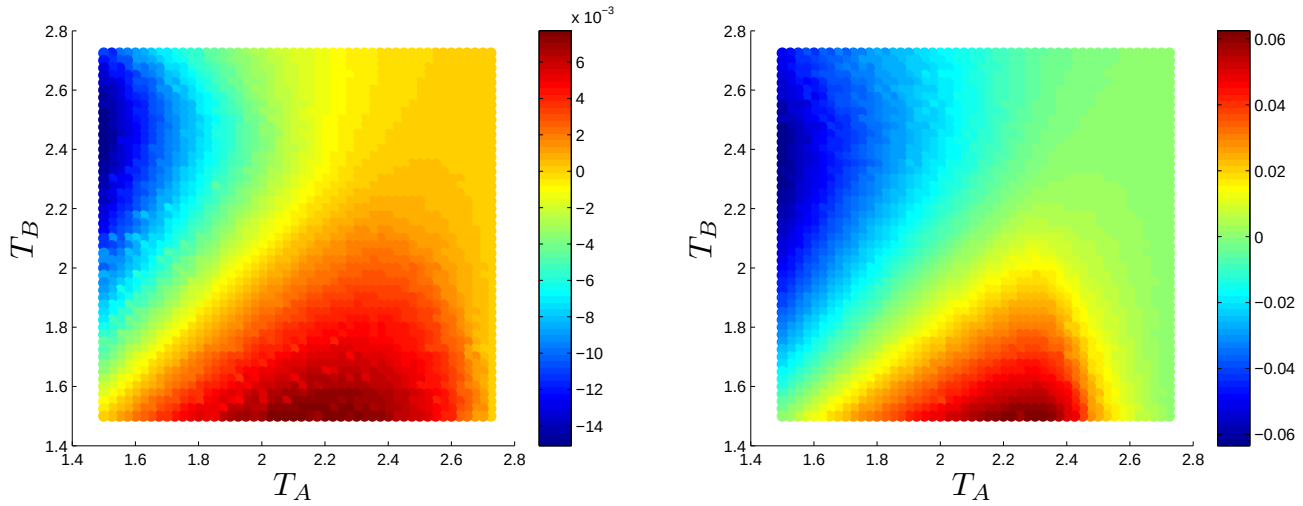


FIG. 6: Entropy production σ (in color scale) as a function of the temperatures T_A and T_B. Left panel: lattice without coarse grain $L_A = L_B = 64$. Right panel: coarse grain with $L_A = 64, L_B = 8$.

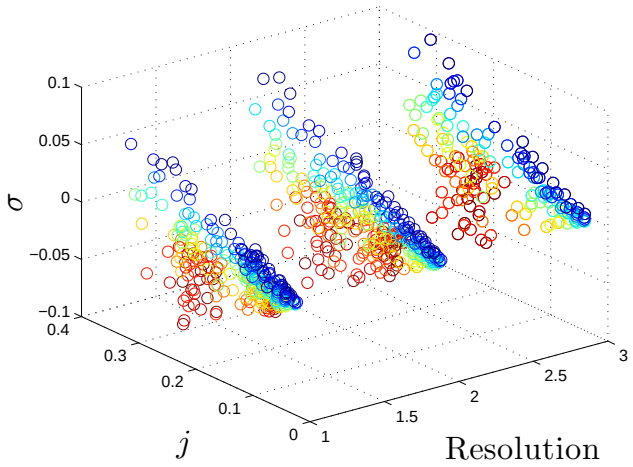


FIG. 7: Entropy production σ (in color scale) as a function of the currents and the change in the resolution. The color scale represents the temperature of the lattice at the lower resolution.

production shown in figure 7 do not show a clear maximum neither the switching between convective and conductive regimes. In order to observe the convective regime it seems therefore indispensable to change abruptly the resolution of the system.

V. DISCUSSION

In this paper we have obtained an effective way to introduce a genuine non equilibrium Ising dynamics under a temperature gradient field with an interaction in the scale space which reproduce some properties of turbulent systems. Previously, other authors have investigated the possibility of building such a model by using only a one lattice dynamics and introducing the temperature gradient within the site of the lattice. Although such a system possesses a rich dynamics with an interesting variety of metastable and stable states, they do not allow for interaction among different scales and presents discontinuities in the temperature fields which make difficult to test Onsager relations as well as out of equilibrium conjectures like the existence of maxima of entropy production.
Our model satisfies such constraints by composing the dynamics of two interacting lattice at different temperatures. This model not only respects the Fourier law $dE/dS|_N = T$, but also features the possibility of having the second

lattice at a coarse grained resolution, as it happens in several physical phenomena. Moreover, when comparing the coarse grained dynamics with the one of two lattices having the same spatial resolutions, evident differences appear. For example, the coarse grained dynamics enhance the breaking of the Fourier law (and therefore of conductive heat transfer) introducing, at high temperature gradients, a convective behavior confirmed by the appearance of a negative temperature in the relation . Moreover, the states in between conductive and convective behavior are the ones maximising the entropy production and correspond to the critical temperature of the equilibrium Ising dynamics. The convective states are a new features for the Ising dynamics and cannot be observed e.g. by adding an external magnetic field to the system. In addition, the introduction of the temperature gradient between the two different lattices preserves the symmetry spin up/down which is broken with the addition of magnetic fields.

Despite its relative simplicity, our model displays coexistence of quasi-equilibria states at the largest scale, with non-zero flux at smaller scale. It therefore shows that turbulent systems may not be so exceptional, and rather be representative of a wide class of oh-of-equilibrium systems with both spatial and scale dynamics.

An interesting outcome of our model could be the development of more realistic models of natural phenomena. For example, several models of atmospheric convection are based on the equilibrium Ising dynamics [20, 21]. We believe that a better representation of the convective atmospheric phenomena can be obtained by introducing the non equilibrium phenomena described in our model through the possible modeling both conductive and convective heat transfer. On the other hand, one can imagine that the second lattice is just some coarse grained properties of the first one and that temperature T_B maps some properties in the scale space.

Finally, the fact that the maxima of entropy production correspond to states at critical temperature for the lattice suggests that there is a link between the maximum entropy production principle and the self organized criticality. The latter is often invoked to explain why out of equilibrium systems settle into states close to some critical properties of the system but without a solid physical underlying motivation. We believe that such a connection might represent a step in solving this issue, but it needs further investigations.

[1] G. Lebon, D. Jou, and J. Casas-Vázquez, *Understanding non-equilibrium thermodynamics: foundations, applications, frontiers* (Springer, 2008).
[2] F. Bouchet and A. Venaille, Physics reports **515**, 227 (2012).
[3] S. Thalabard, B. Saint-Michel, É. Herbert, F. Daviaud, and B. Dubrulle (????).
[4] M. Pleimling, B. Schmittmann, and R. Zia, EPL (Europhysics Letters) **89**, 50001 (2010).
[5] X. Huang, B. Cui, and X. Yi, Modern Physics Letters B **25**, 175 (2011).
[6] E. Bertin and O. Dauchot, Physical review letters **102**, 160601 (2009).
[7] E. Ising, Zeitschrift für Physik A Hadrons and Nuclei **31**, 253 (1925).
[8] B. M. McCoy and T. T. Wu (1973).
[9] L. Onsager, Physical Review **65**, 117 (1944).
[10] L. Onsager, Suppl **2**, 249 (1949).
[11] R. GLAUBER, Journal of Mathematical Physics **4**, 294 (1963).
[12] P. Garrido and J. Marro, Physica A: Statistical Mechanics and its Applications **144**, 585 (1987).
[13] P. L. Garrido, J. L. Lebowitz, C. Maes, and H. Spohn, Physical Review A **42**, 1954 (1990).
[14] Z. Cheng, P. Garrido, J. Lebowitz, and J. Vallés, EPL (Europhysics Letters) **14**, 507 (1991).
[15] E. L. Præstgaard, B. Schmittmann, and R. Zia, The European Physical Journal B-Condensed Matter and Complex Systems **18**, 675 (2000).
[16] C. Maes and F. Redig, Journal de Physique I **1**, 669 (1991).
[17] M. Mihelich, B. Dubrulle, D. Paillard, and D. Faranda, Preprint (2014).
[18] P. Bak, C. Tang, and K. Wiesenfeld, Physical review A **38**, 364 (1988).
[19] P. Bak, Physica A: Statistical Mechanics and its Applications **163**, 403 (1990).
[20] A. J. Majda and B. Khouider, Proceedings of the National Academy of Sciences **99**, 1123 (2002).
[21] B. Khouider and A. J. Majda, Journal of the atmospheric sciences **63**, 1308 (2006).

Bibliographie

[1] AA Filyukov and V Ya Karpov. Description of steady transport processes by the method of the most probable path of evolution. *Journal of engineering physics*, 13(5) :326–329, 1967.

[2] G W Paltridge. Global dynamics and climate-a system of minimum entropy exchange. *Q. J. R. Meteorol. Soc.*, 101(429) :475–484, 1975.

[3] Shinya Shimokawa and Hisasi Ozawa. On the thermodynamics of the oceanic general circulation : entropy increase rate of an open dissipative system and its surroundings. *Tellus A*, 53(2) :266–277, 2001.

[4] Paško Županović, Davor Juretić, and Srećko Botrić. Kirchhoffs loop law and the maximum entropy production principle. *Physical Review E*, 70(5) :056108, 2004.

[5] Yu M Ivanchenko. High-efficiency frequency generation in a periodic array of josephson junctions. *Physical Review B*, 54(18) :13247, 1996.

[6] Jean Le Rond d'Alembert. *Traité de dynamique*. 1743.

[7] R M L Evans. Detailed balance has a counterpart in non-equilibrium steady states. *J. Phys. A : Math. Gen.*, 38(2) :293, 2004.

[8] C Monthus. Non-equilibrium steady states : maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints. *J. Stat. Mech.*, page P03008, 2011.

[9] J S Kirkaldy. Entropy criteria applied to pattern selection in systems with free boundaries. *Metallurgical and Materials Transactions A*, 16(10) :1781–1797, 1985.

[10] Thomas Christen. Application of the maximum entropy production principle to electrical systems. *J. Phys. D : Appl. Phys.*, 39(20) :4497, 2006.

[11] L M Martyushev and V D Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Phys. Rep.*, 426(1) :1–45, 2006.

[12] R Balian. *Physique statistique et themodynamique hors équilibre*. Ecole Polytechnique, 1992.

[13] H Ziegler, IN Sneddon, and R Hill. Progress in solid mechanics 4, 1963, 93.

[14] H Ozawa, A Ohmura, Ralph Lorenz, and T Pujol. The second law of thermodynamics and the global climate system : A review of the maximum entropy production principle. *Rev. Geophys.*, 41 :1018, 2003.

[15] C Herbert, D Paillard, M Kageyama, and B Dubrulle. Present and Last Glacial Maximum climates as states of maximum entropy production. *Q. J. R. Meteorol. Soc.*, 137(657) :1059–1069, 2011.

[16] Corentin Herbert, Didier Paillard, and Bérengère Dubrulle. Entropy production and multiple equilibria : the case of the ice-albedo feedback. *arXiv preprint arXiv :1103.0722*, 2011.

[17] Roderick C Dewar. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J. Phys. A : Math. Gen.*, 36(3) :631, 2003.

[18] Roderick C Dewar. Maximum entropy production and the fluctuation theorem. *J. Phys. A : Math. Gen.*, 38(21) :L371, 2005.

[19] Ilya Prigogine. Introduction à la thermodynamique des processus irréversibles. 1968.

[20] S.R. De Groot and P. Mazur. *Non-equilibrium thermodynamics*. Dover publications, 2011.

[21] V. Bertola and E. Cafaro. A critical analysis of the minimum entropy production theorem and its application to heat and fluid flow. *Int. J. Heat Mass Transfer*, 51(7) :1907–1912, 2008.

[22] A.N.Kolmogorov. New metric invariant of transitive dynamical systems and endomorphisms of lebesgue spaces. *Doklady of Russian Academy of Sciences*, 119(5) :861–864, 1958.

[23] P. Billingsley. *Ergodic theory and information*. Wiley, 1965.

[24] Anatole Katok and Boris Hasselblatt. Introduction to the modern theory of dynamical systems. *Cambridge, Cambridge*, 1995.

[25] Yves Benoist and Frédéric Paulin. Systèmes dynamiques élémentaires. *Cours de Magistère ENS*, 3, 2002.

[26] Z Burda, J Duda, J M Luck, and B Waclaw. Localization of the maximal entropy random walk. *Phys. Rev. Lett.*, 102(16) :160602, 2009.

[27] Jesús Gómez-Gardeñes and Vito Latora. Entropy rate of diffusion processes on complex networks. *Physical Review E*, 78(6) :065102, 2008.

[28] Christian Maes, Frank Redig, and Annelies Van Moffaert. On the definition of entropy production, via examples. *Journal of mathematical physics*, 41(3) :1528–1554, 2000.

[29] Christian Maes, Karel Netočnỳ, and Bram Wynants. On and beyond entropy production : the case of markov jump processes. *arXiv preprint arXiv :0709.4327*, 2007.

[30] M P Qian, M Qian, and G L Gong. The reversibility and the entropy production of Markov processes. *Contemp. Math.*, 118 :255–261, 1991.

[31] P. Gaspard. Time-reversed dynamical entropy and irreversibility in Markovian random processes. *J. Stat. Phys.*, 117(3) :599–615, 2004.

[32] Debashish Chowdhury, Ludger Santen, and Andreas Schadschneider. Statistical physics of vehicular traffic and some related systems. *Physics Reports*, 329(4) :199–329, 2000.

[33] Edy Soewono and Asep K Supriatna. A two-dimensional model for the transmission of dengue fever disease. *Bull. Malay. Math. Sci. Soc*, 24 :49–57, 2001.

[34] Bernard Derrida, Eytan Domany, and David Mukamel. An exact solution of a one-dimensional asymmetric exclusion model with open boundaries. *Journal of Statistical Physics*, 69(3-4) :667–687, 1992.

[35] M Gorissen, A Lazarescu, K Mallick, and C Vanderzande. Exact Current Statistics of the Asymmetric Simple Exclusion Process with Open Boundaries. *Phys. Rev. Lett.*, 109(17) :170601, 2012.

[36] Frank Spitzer. Interaction of markov processes. *Advances in Mathematics*, 5(2) :246–290, 1970.

[37] Martin R Evans and Tom Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *Journal of Physics A : Mathematical and General*, 38(19) :R195, 2005.

[38] Stefan Großkinsky, Gunter M Schütz, and Herbert Spohn. Condensation in the zero range process : stationary and dynamical properties. *Journal of statistical physics*, 113(3-4) :389–410, 2003.

[39] E Levine, D Mukamel, and GM Schütz. Zero-range process with open boundaries. *Journal of statistical physics*, 120(5-6) :759–778, 2005.

[40] JM Carlson, ER Grannan, and GH Swindle. Self-organizing systems at finite driving rates. *Physical Review E*, 47(1) :93, 1993.

[41] MR Evans. Bose-einstein condensation in disordered exclusion models and relation to traffic flow. *EPL (Europhysics Letters)*, 36(1) :13, 1996.

[42] Ginestra Bianconi and Albert-László Barabási. Bose-einstein condensation in complex networks. *Physical Review Letters*, 86(24) :5632, 2001.

[43] Vito Latora, Michel Baranger, Andrea Rapisarda, and Constantino Tsallis. The rate of entropy increase at the edge of chaos. *Phys. Lett. A*, 273(1) :97–103, 2000.

[44] Henri Poincaré and R Magini. Les méthodes nouvelles de la mécanique céleste. *Il Nuovo Cimento (1895-1900)*, 10(1) :128–130, 1899.

[45] George D Birkhoff. On the periodic motions of dynamical systems. *Acta Mathematica*, 50(1) :359–379, 1927.

[46] Josiah Willard Gibbs. Elementary principles in statistical physics, 1902.

[47] Massimo Falcioni, Luigi Palatella, and Angelo Vulpiani. Production rate of the coarse-grained gibbs entropy and the kolmogorov-sinai entropy : A real connection ? *Physical Review E*, 71(1) :016118, 2005.

[48] P Ehrenfest and T Ehrenfest. Encyklopädie der math. *Wissensch., IV*, 1911.

[49] Alexander N Gorban. Basic types of coarse-graining. In *Model Reduction and Coarse-Graining Approaches for Multiscale Phenomena*, pages 117–176. Springer, 2006.

[50] M Peña and E Kalnay. Separating fast and slow modes in coupled chaotic systems. *Nonlinear Processes in Geophysics*, 11(3) :319–327, 2004.

[51] Kenneth G Wilson and John Kogut. The renormalization group and the expansion. *Physics Reports*, 12(2) :75–199, 1974.

[52] David Ruelle. *Statistical Mechanics : Rigorous Results*. Amsterdam : Benjamin, 1969.

[53] David Ruelle. Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *Journal of Statistical Physics*, 95(1-2) :393–468, 1999.

[54] Pierre Gaspard. Entropy production in open volume-preserving systems. *Journal of statistical physics*, 88(5-6) :1215–1240, 1997.

[55] Thomas Gilbert and JR Dorfman. Entropy production : from open volume-preserving to dissipative systems. *Journal of statistical physics*, 96(1-2) :225–269, 1999.

[56] Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2) :130–141, 1963.

[57] Stephen Boyd, Persi Diaconis, and Lin Xiao. Fastest mixing markov chain on a graph. *SIAM review*, 46(4) :667–689, 2004.

[58] Ravi R Montenegro and Prasad Tetali. *Mathematical aspects of mixing times in Markov chains*, volume 3. Now Publishers Inc, 2006.

[59] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.

[60] William Parry. Intrinsic markov chains. *Transactions of the American Mathematical Society*, pages 55–66, 1964.

[61] Lai-Sang Young. Entropy in dynamical systems. In *Entropy*, pages 313–327. Princeton University Press Princeton, NJ, 2003.

[62] Sheldon E Newhouse. Entropy and volume. *Ergodic Theory and Dynamical Systems*, 8(8*) :283–299, 1988.

[63] Yosef Yomdin. Volume growth and entropy. *Israel Journal of Mathematics*, 57(3) :285–300, 1987.

[64] Corentin Herbert. *Application de la mécanique statistique à la modélisation du climat*. PhD thesis, Université Paris VI, 2012.

[65] David Bolton. The computation of equivalent potential temperature. *Monthly weather review*, 108(7) :1046–1053, 1980.

[66] Richard Goody. Maximum entropy production in climate theory. *Journal of the atmospheric sciences*, 64(7), 2007.

[67] Axel Kleidon and Ralph D Lorenz. *Non-equilibrium thermodynamics and the production of entropy : life, earth, and beyond*. Springer Science & Business Media, 2005.