



HAL
open science

Mécanique statistique et dynamique hors équilibre de systèmes avec interactions à longue portée

Julien Barré

► **To cite this version:**

Julien Barré. Mécanique statistique et dynamique hors équilibre de systèmes avec interactions à longue portée. Physique [physics]. Ecole normale supérieure de lyon - ENS LYON, 2003. Français. NNT : . tel-00430889

HAL Id: tel-00430889

<https://theses.hal.science/tel-00430889>

Submitted on 10 Nov 2009

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.

École Normale Supérieure de Lyon

Soutenue le 8 juillet 2003

JULIEN BARRÉ

**MÉCANIQUE STATISTIQUE ET
DYNAMIQUE HORS ÉQUILIBRE DE
SYSTÈMES AVEC INTERACTIONS À
LONGUE PORTÉE**

Jury : Thierry Dauxois (Directeur de thèse)

Peter Holdsworth

Giovanni Jona-Lasinio

Marc Mézard (Rapporteur)

Luca Peliti (Rapporteur)

Stefano Ruffo (Directeur de thèse)

Remerciements

Je suis très reconnaissant à Marc Mézard et Luca Peliti d'avoir pris le temps de lire cette thèse en détail, et d'avoir accepté d'en être rapporteur, et je remercie Giovanni Jona-Lasinio de participer au jury : je suis très honoré de l'intérêt qu'ils ont porté à mon travail. Ce sont entre autres des discussions avec Peter Holdsworth, et un de ses articles, qui ont motivé le début de cette thèse : je suis donc très heureux qu'il puisse également participer au jury.

Avec mes deux directeurs de thèse, j'ai bénéficié d'un encadrement "en or". Toujours disponible, toujours enthousiaste, Thierry a su me guider dans mes essais et hésitations : j'ai eu durant ces trois années l'impression très agréable d'être à la fois libre et soutenu. Les nombreuses discussions avec Stefano ont toujours été très enrichissantes : j'espère avoir retenu un peu de ses connaissances étendues en dynamique non linéaire, théorie du chaos et mécanique statistique. Grâce à lui, le groupe de physiciens florentins parmi les ingénieurs est très vivant, et constitue un excellent "milieu de culture" pour un jeune étudiant ! Bref, j'ai eu beaucoup de plaisir à travailler avec Thierry et Stefano, et aussi parfois à jouer au foot, manger une bistecca, ou discuter des malheurs de la Fiorentina et de la recherche italienne... je les remercie donc chaleureusement, et j'espère pouvoir continuer avec eux toutes ces activités.

Toujours grâce à mes directeurs, j'ai pu rencontrer beaucoup de chercheurs ayant des intérêts voisins, et nouer des collaborations : ce fut un des bonheurs de cette thèse. Ma première pensée va bien sûr à Freddy Bouchet, dont les idées profondes ont eu une grande influence sur ce travail. Discuter avec lui fut toujours un plaisir, qui j'espère se renouvellera souvent; je le remercie aussi pour m'avoir souvent fait partager sa vivante vie de famille ! J'ai pu aussi au cours de mes séjours à Florence rencontrer David Mukamel et travailler avec lui : la clarté et la précision de ses idées et de ses exposés resteront pour moi un exemple, et je le remercie de m'en avoir fait profiter. Sa gentillesse, sa disponibilité et son impressionnante dextérité numérique font de Yoshi Yamaguchi un collaborateur très agréable : je le remercie donc avec plaisir. Ma gratitude va aussi à Giovanni de Ninno, qui ne s'est pas lassé d'essayer de combler mes nombreuses lacunes au sujet du laser à électrons libres.

Beaucoup d'autres discussions m'ont aidé et éclairé au cours de ces trois ans, notamment avec Luca Celardo, Pierre-Henri Chavanis, Pascal Degiovanni, Yves Elskens, Duccio Fanelli, Jean Farago, Marie-Christine Firpo, Olivier Fliegans, Dieter Gross, Julien Javaloyès, François Leyvraz, Julien Michel, Mathias Perrin, Antonio Politi, Atsushi Taruya... Je salue aussi tous ceux, à Lyon et à Florence, grâce à qui ces trois années n'ont pas été agréables que scientifiquement : le groupe DOCS, l'équipe de foot du labo, mes camarades thésards, et spécialement mes cobureaux...

Une thèse en cotutelle comme celle-ci aurait pu poser des problèmes administratifs : grâce à la gentillesse et la compétence que j'ai toujours trouvées au laboratoire

de Physique et aux services de l'ENS Lyon, ainsi qu'au Dipartimento di Energetica de Florence, il n'y eut aucune difficulté de cette sorte. Il n'y eut aucune difficulté financière non plus, grâce cette fois à l'aide de la région Rhône-Alpes. Je remercie aussi l'équipe enseignante de l'ENS Lyon, notamment Hervé Gayvallet et François Gieres, grâce auxquels l'enseignement fut un plaisir, et put se combiner avec les contraintes de la cotutelle.

Enfin, si la joie de travailler m'a rarement quitté tout au long de cette thèse, c'est à ma famille et mes amis que je le dois. Je réserve bien sûr une place particulière à Magali, à qui, en plus de tout le reste, je dois une reconnaissance spéciale pour sa relecture d'un œil acéré du manuscrit...

Contents

Introduction	5
I Mécanique statistique à l'équilibre	9
Introduction	10
1 Le modèle Blume-Emery-Griffiths champ moyen	14
2 Champ moyen et théorie des grandes déviations	19
2.1 Introduction à la théorie des grandes déviations	19
2.2 Une méthode générale	22
2.3 Exemples	27
2.3.1 Retour sur le modèle Blume-Emery-Griffiths	27
2.3.2 Le modèle HMF (Heisenberg Mean Field)	28
2.3.3 Un modèle de laser CARL	30
2.3.4 Le modèle d'Ising, avec interactions en $1/r^\alpha$	34
2.4 Remarques sur la méthode des grandes déviations	39
2.4.1 Le problème du scaling	39
2.4.2 Limites de la méthode des grandes déviations	41
2.4.3 Lien avec les potentiels de Kac	41
2.5 Revue de résultats	42
2.5.1 Systèmes auto-gravitants	42
2.5.2 Turbulence bidimensionnelle	43
2.5.3 Physique des plasmas	44
3 Étude générale des problèmes variationnels et classification des transitions de phases	46
3.1 Étude générale	46
3.1.1 Inéquivalence d'ensembles	46
3.1.2 États métastables	48
3.2 Classification	51
3.2.1 Notions générales sur les singularités	52
3.2.2 Réduction du cas microcanonique à un problème sans contrainte	54
3.2.3 Cas d'un problème sans symétrie	54
3.2.4 Cas d'un problème avec symétrie	68

3.2.5	Conclusion	82
	Conclusion	84
	II Dynamique hors équilibre	88
	Introduction	89
4	Formation de structures hors équilibre : étude du modèle HMF anti-ferromagnétique	95
4.1	Introduction	95
4.2	La dynamique effective	97
4.2.1	Description variationnelle d'une dynamique adiabatique : étude d'un exemple simple	98
4.2.2	Les variables rapides	100
4.2.3	Les variables lentes	102
4.3	Formation de la structure aux temps courts	105
4.3.1	Divergence de la densité	107
4.3.2	Description des caustiques	107
4.3.3	Lien avec la classification des singularités	109
4.3.4	Estimation de la distribution angulaire	111
4.4	Comportement du double amas aux temps longs	112
4.4.1	Étude du Hamiltonien effectif	113
4.4.2	Comparaison avec le Hamiltonien complet	116
4.5	Conclusion	118
5	Relaxation lente à l'équilibre : étude du modèle HMF ferromagnétique	119
5.1	Introduction	119
5.2	Étude de l'équation de Vlasov associée	122
5.2.1	États stationnaires de l'équation de Vlasov	123
5.2.2	Stabilité des états stationnaires	124
5.2.3	Stabilité linéaire	129
5.3	Convergence vers un état stationnaire	131
5.3.1	Deux conditions nécessaires de stationnarité	132
5.3.2	Tests numériques	134
5.3.3	Étude numérique de l'évolution lente	135
5.3.4	Une approche analytique de la relaxation lente	138
5.4	Discussion	139
	Conclusion	141

III	Application à un modèle de laser à électrons libres	143
	Introduction	144
6	Étude statistique à l'équilibre et hors équilibre d'un modèle de Laser à Électrons Libres.	145
6.1	La physique du LEL et l'obtention du Hamiltonien	145
6.1.1	Le rayonnement synchrotron	146
6.1.2	Interaction entre un électron et le champ	146
6.1.3	Obtention d'un Hamiltonien unidimensionnel	147
6.1.4	Étude linéaire	149
6.2	Mécanique statistique d'équilibre	150
6.2.1	La solution	151
6.2.2	Comparaison avec les simulations numériques	153
6.3	Approche hors équilibre	155
6.3.1	Un état métastable	155
6.3.2	Mécanique statistique de l'équation de Vlasov	155
6.3.3	Solution pour le modèle du LEL	158
6.3.4	Comparaison avec les simulations numériques	160
	Conclusion	163
	Conclusion et perspectives	165
	A. Potentiels en loi de puissance et longue portée	169
	B. Approximation uniforme d'un Hamiltonien discret par un Hamiltonien continu	170
	C. Valeurs et vecteurs propres de la matrice T	176
	D. Stabilité formelle d'un système dynamique	177
	Liste des publications	187

Introduction

Cette thèse est consacrée à l'étude à l'équilibre et hors équilibre des systèmes avec interactions à longue portée : nous allons chercher à mettre en évidence et exploiter l'universalité des comportements physiques induits par la présence d'interactions à longue portée. Cette notion de longue portée est assez vague, et son emploi dans la littérature recouvre des situations très différentes. Nous commençons donc ici par définir plus précisément ce que nous entendons par là.

Qu'appellera-t-on "longue portée" ?

Considérons un système macroscopique, et imaginons qu'on le sépare en deux. L'énergie du système total est alors égale à la somme des énergies de chaque partie, plus l'énergie d'interaction entre les deux parties. Cette énergie d'interface est proportionnelle, lorsque les interactions au sein du système sont à courte portée, à la surface de contact entre les deux parties : pour un système macroscopique, elle sera donc négligeable devant les énergies volumiques. On voit bien cependant que ce raisonnement n'est plus valable si les interactions sont à suffisamment longue portée. Considérons par exemple un modèle d'Ising comportant N spins, avec interactions de portée infinie (indépendante de la distance); le Hamiltonien est :

$$H = -J \sum_{i \neq j} S_i S_j . \quad (1)$$

Lorsque tous les spins sont orientés positivement, l'énergie totale est $E = -JN(N-1)$; séparons maintenant le système en deux sous-systèmes contenant chacun $N' = N/2$ spins. Chaque sous-système, lorsqu'il est pris isolément de l'autre, a une énergie $E' = -JN(N-2)/4$: on constate que $E \neq 2E'$. Remarquons aussi que l'emploi d'une constante de couplage J/N renormalisée par le nombre de spins, comme il est d'usage pour ces modèles de type champ moyen, a pour effet de fournir des énergies d'ordre N , mais ne soigne pas le défaut d'additivité. On prendra dorénavant le critère de non additivité suivant comme définition de la "longue portée" :

Un système macroscopique sera dit "avec interactions à longue portée" si on ne peut pas exprimer son énergie comme somme des énergies de sous systèmes macroscopiques indépendants.

Avec cette définition, une interaction entre particules décroissant en loi de puissance avec la distance, comme $1/r^\alpha$, est "à longue portée" dès que l'exposant α est plus petit que D , la dimension spatiale du système (voir appendice A). Ceci inclut par exemple la gravitation et exclut les interactions de Van der Waals. Ce critère *ne correspond pas* à la terminologie des phénomènes critiques, selon laquelle sont "à longue portée" les potentiels tels que $\alpha < D + 2 - \eta$, η étant un exposant critique dépendant du système, mais en général faible, si bien que le domaine de "longue portée" de la communauté des phénomènes critiques est plus étendu que celui considéré dans cette thèse. Nos interactions à longue portée sont ainsi parfois appelées "interactions non intégrables" dans la littérature [110].

Des exemples physiques

Les interactions à longue portée, telles qu'on vient de les définir, ne sont pas des curiosités mathématiques, et se rencontrent en fait dans un assez grand nombre de systèmes physiques. On peut donner bien sûr comme premier exemple les systèmes en interaction gravitationnelle pour lesquels le potentiel est attractif et décroît avec la distance comme $1/r$. Ils ont évidemment été étudiés depuis longtemps dans le contexte astrophysique : Padmanabhan fait par exemple dans [94] une revue des concepts et résultats de la mécanique statistique des systèmes auto-gravitants. Nous reviendrons plus en détail par la suite sur ce sujet.

Un autre exemple, moins immédiat, est donné par la mécanique des fluides bidimensionnels. En effet, l'énergie du système apparaît dans ce cas comme une énergie d'interaction entre vortex, et cette interaction est à longue portée : par exemple, $V(r)$ est proportionnel à $\ln r$ pour un domaine infini. La turbulence bidimensionnelle est donc une application importante de la théorie des systèmes avec interactions à longue portée, d'autant plus importante qu'il faut inclure dans cette catégorie les flots géophysiques, rendus quasi bidimensionnels par la rotation des planètes. Notons une différence cependant avec les systèmes gravitationnels : les équations de la mécanique des fluides sont continues (bien qu'une approche consiste à les discrétiser et à considérer un système de points vortex), alors que les systèmes auto-gravitants sont intrinsèquement particuliers (bien que dans certaines conditions, ils soient très bien décrits par les équations continues de Vlasov-Poisson, voir chapitre 5).

On pourrait penser qu'il est possible d'ajouter à ces deux exemples les systèmes en interaction coulombienne. En fait, l'interaction est écrantée pour les systèmes neutres, si bien qu'ils restent additifs; les méthodes et résultats que nous allons décrire ne peuvent donc s'y appliquer; nous reviendrons sur ce problème par la suite. La physique des plasmas fournit néanmoins beaucoup d'exemples de systèmes non additifs, à partir d'une description *effective* du système. Plus précisément, même si l'équilibre statistique global d'un système est additif, il est possible qu'il soit pertinent (pour certaines conditions expérimentales, ou échelles de temps) de ne considérer l'équilibre statistique que d'une partie des degrés de liberté du système. Cette description effective est alors parfois non additive : il en est ainsi de la modélisation de l'interaction de filaments de courant par des modèles de points vortex [79], proches donc de la turbulence bidimensionnelle, ou encore de la description d'un plasma par l'interaction entre ondes de Langmuir et particules résonantes avec ces ondes [5] : dans ce cas, l'onde, degré de liberté global pour le système, couple entre elles des parties éloignées du système, et crée ainsi la non additivité, bien que les interactions de particule à particule puissent être négligeables.

Ce dernier exemple permet d'ajouter toute une classe de modèles susceptibles de bénéficier de l'étude générale des systèmes non additifs : ceux qui mettent en jeu le couplage entre ondes et particules; l'interaction à longue portée effective est alors transmise par l'onde (ou les ondes), variable globale. Nous développerons ce point de vue à la partie III, en étudiant un modèle de laser à électrons libres.

On s'aperçoit donc que la définition de la longue portée par la non additivité conduit

à écarter certains systèmes (coulombiens par exemple), mais élargit les possibilités d'application de l'étude à toute une classe de nouveaux problèmes, qui présentent tous des structures à grande échelle (onde, gros vortex par exemple). Cerner plus précisément le type de systèmes que nous pouvons appeler, selon la définition ci-dessus, "avec interactions à longue portée" n'est pas une tâche facile; c'est aussi l'un des objectifs de ce travail.

Universalité des comportements et des méthodes ?

La non additivité est à l'origine pour ces systèmes de comportements inhabituels, tant en ce qui concerne leur thermodynamique à l'équilibre que leurs propriétés de relaxation et leur dynamique hors équilibre. Ainsi, la séparation de phases au sens habituel est impossible, et ceci remet en question par exemple l'équivalence entre les ensembles statistiques canonique et microcanonique. De même, la dynamique est nécessairement cohérente à l'échelle du système entier, ce qui modifie la perception habituelle de la relaxation vers l'équilibre. Ces différents aspects ont bien sûr déjà été étudiés en détail dans chaque domaine particulier : les systèmes auto-gravitants (voir [94] pour une revue), la turbulence bidimensionnelle (voir par exemple [36]), la physique des plasmas (voir par exemple [50])... Cependant, ce n'est que plus récemment, semble-t-il, que des travaux cherchent à mettre en évidence une possible universalité des comportements pour ces systèmes avec interactions à longue portée. En ce qui concerne la mécanique statistique d'équilibre et ses anomalies, on peut citer ici le travail précurseur de Hertel et Thirring, sur un modèle jouet [60]; la similarité des méthodes de résolution des différents modèles a été exploitée par exemple dans les travaux de Spohn *et al.* [52, 89], et Kiessling *et al.* [75, 76, 77]. En ce qui concerne la dynamique, Chavanis, Sommeria et Robert [33, 34, 36] ont développé les analogies entre turbulence bidimensionnelle et systèmes auto-gravitants, en se fondant sur la proximité formelle entre les équations d'Euler et de Vlasov. C'est un petit peu dans l'esprit de ces travaux que nous avons essayé de travailler au cours de cette thèse, avec l'idée de poursuivre jusqu'à ses limites ce processus de généralisation.

Une telle démarche présente un intérêt du point de vue fondamental bien sûr; mais surtout, une meilleure compréhension de l'universalité des comportements, son origine et ses limites, peut stimuler le transfert de concepts et méthodes entre les différents domaines concernés par les interactions à longue portée. Une telle approche peut aussi permettre de mettre en évidence de nouveaux domaines d'application.

Méthodologie et plan de la thèse

Nous proposons donc dans cette thèse d'essayer de comprendre l'étendue, et les limites, de l'universalité des comportements des systèmes non additifs. Notre but est d'identifier des méthodes et d'obtenir des résultats généraux qui pourront être transférés aux différents domaines d'application. Pour cela, nous partons de modèles simples, sur lesquels les calculs analytiques peuvent être poussés très loin, et les simulations numériques sont faciles. Ils permettent donc de tester commodément idées et méthodes. Notre étude se divise en deux grandes parties : la mécanique statistique à

l'équilibre d'une part, et la dynamique, et mécanique statistique, hors équilibre d'autre part. Nous chercherons et mettrons en évidence dans chaque cas, à partir de l'étude des modèles jouets, un concept unificateur : ce sera la théorie de champ moyen (en un sens élargi) pour l'étude à l'équilibre, et l'équation de Vlasov, avec ses succès et ses limites, dans le cas de la dynamique hors équilibre.

La partie I est donc consacrée à la mécanique statistique à l'équilibre des systèmes avec interactions à longue portée. Nous donnons d'abord un exemple de modèle jouet simple, exactement soluble, qui met en évidence certains comportements caractéristiques : chaleur spécifique négative, inéquivalence d'ensembles par exemple. Puis nous passerons du particulier au général, en introduisant une méthode puissante de résolution pour la mécanique statistique des interactions à longue portée, et en proposant une classification des types de transitions de phases possibles, et des situations inhabituelles associées. En ce qui concerne l'équilibre, l'approche par modèles jouets a pu donc être menée jusqu'à son terme. À la partie II nous nous intéressons à la dynamique et mécanique statistique hors équilibre de ces systèmes : nous pourrons dans ce cas aussi proposer des conclusions générales à partir de l'étude de modèles jouets, bien que ce soit plus difficile. Nous appliquons ensuite ces résultats, à l'équilibre et hors équilibre, à un modèle de laser à électrons libres plus proche de la réalité expérimentale (partie III) : c'est là un exemple de retour vers les applications; notre étude générale préalable permettra alors l'utilisation de méthodes statistiques dans un contexte nouveau.

Part I

Mécanique statistique à l'équilibre

Introduction

Lorsque les interactions sont à longue portée, avec la définition donnée en introduction, les systèmes concernés sont très spéciaux du point de vue thermodynamique. En effet, l'énergie n'est pas additive et, de ce fait, beaucoup de résultats habituels de mécanique statistique et de thermodynamique ne sont plus valables.

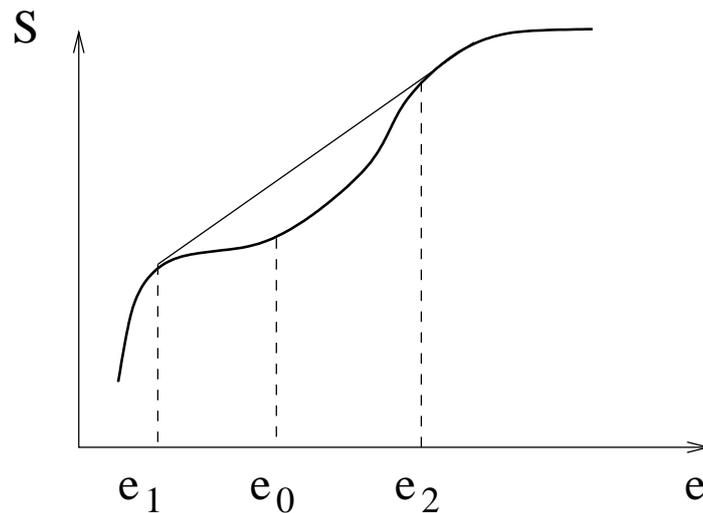


Figure 1: Schéma d'une entropie non concave dans le cas d'un système additif : pour une énergie e_0 , il y aura séparation de phases.

Donnons un exemple. Le modèle le plus simple avec interactions à longue portée est sans doute le modèle d'Ising champ moyen, avec pour Hamiltonien (les S_i sont des spins qui valent $+1$ ou -1) :

$$H = -\frac{J}{N} \sum_{i,j=1}^N S_i S_j . \quad (2)$$

La constante de couplage J est renormalisée par un facteur bien choisi dépendant de N (ici simplement N), de façon à ce que le système se comporte de manière extensive : son énergie totale est proportionnelle à N . Sans cet artifice, la limite thermodynamique n'existe pas au sens habituel, et il faut recourir à d'autres limites pour étudier les grands systèmes (voir chapitre 2).

Néanmoins, même lorsque l'interaction est renormalisée de telle sorte que le système

soit rendu (pseudo-)extensif, il reste non additif; une conséquence importante est qu'il ne peut pas se séparer en deux phases. En effet, le raisonnement justifiant la séparation de phases est en général le suivant. Imaginons un système dont l'entropie en fonction de l'énergie $S(e)$ n'est pas concave (voir figure 1), et considérons une énergie e_0 située sous la double tangente. Si e_1 et e_2 sont les points d'appui de la double tangente, on a $e_0 = xe_1 + (1-x)e_2$, avec $x \in]0, 1[$. Divisons le système en deux phases, d'énergie e_1 et e_2 , avec des proportions x et $1-x$; l'entropie du système ainsi divisé est $xS(e_1) + (1-x)S(e_2) > S(e_0)$: cette division est donc favorable. Évidemment, cet argument repose entièrement sur l'additivité. De même, la démonstration courante de l'équivalence de l'ensemble microcanonique avec l'ensemble canonique (ou ses généralisations) n'est plus valable. Dans ce contexte, et en pensant aux situations où le sens physique d'un éventuel thermostat est loin d'être évident (astrophysique, turbulence 2D...), il est tentant de revenir à la mécanique statistique microcanonique. Or celle-ci réserve bien des surprises : la chaleur spécifique peut être négative, la température en fonction de l'énergie peut être discontinue... L'existence de chaleurs spécifiques négatives, et donc d'inéquivalences possibles entre les ensembles statistiques, est connue depuis longtemps des astrophysiciens, mais a mis plus longtemps pour s'imposer parmi les physiciens, habitués de l'ensemble canonique : après un article de M. Lax sur l'inéquivalence des ensembles dans le modèle sphérique de Berlin et Kac [83], qui aborde déjà le problème de façon assez générale, Hertel et Thirring introduisent un modèle simple inspiré de la gravitation, exactement soluble dans les ensembles canonique et microcanonique, qui met en évidence des zones de chaleur spécifique négative [60]. Par la suite, ce genre d'anomalies a été étudié dans les divers domaines où la longue portée intervient. L'importance de la physique statistique microcanonique, ainsi que ses différences avec la physique statistique canonique, plus habituelle, a été également mise en valeur ces dix dernières années par D. Gross, même en l'absence d'interactions à longue portée, dans le contexte des systèmes à petit nombre de degrés de liberté [58] (physique nucléaire, agrégats atomiques par exemple).

Nous adopterons pour étudier ces problèmes une démarche simple, qui consiste à partir des idées de base de la mécanique statistique (volume de l'espace des phases, entropie de Boltzmann), pour essayer d'en tirer le maximum d'information. Cette démarche naturelle est évidemment loin d'être nouvelle, et la mécanique statistique des systèmes avec interactions à longue portée a une longue histoire : nous faisons à la section 2.5 une rapide revue des principaux résultats connus, pour les différents domaines concernés (astrophysique, turbulence bidimensionnelle, plasmas).

Motivés par les particularités des systèmes avec interactions à longue portée, certains auteurs ont proposé d'appliquer à ces systèmes une nouvelle définition de l'entropie, intrinsèquement non additive [107] : il s'agit de remplacer l'expression habituelle de l'entropie informationnelle de Gibbs, pour un ensemble de probabilités p_i ,

$$S_G = - \sum_i p_i \ln p_i, \quad (3)$$

par l'entropie de Tsallis dépendant du paramètre q ,

$$S_q = \frac{1 - \sum_i p_i^q}{q - 1}, \quad (4)$$

et de déduire à partir de là toute une nouvelle thermodynamique, dépendant de ce paramètre q . S_q est dite non additive, car la q -entropie (pour $q \neq 1$) de la réunion de deux systèmes indépendants (au sens probabiliste) n'est pas égale à la somme des entropies des deux systèmes pris indépendamment. S_q se réduit à S_G pour q tendant vers 1. L'idée d'utiliser une telle entropie pour décrire des situations d'équilibre semble abandonnée, mais nous reviendrons au chapitre 5 sur l'entropie de Tsallis, puisqu'il a été plus récemment proposé qu'elle puisse décrire des situations hors équilibre.

Le principe de l'étude de la mécanique statistique d'équilibre est donc très simple : il tient tout entier dans l'hypothèse de Boltzmann d'une équiprobabilité des différentes configurations microscopiques accessibles au système. Les problèmes que nous allons aborder au cours de ce premier chapitre sont les suivants :

- **Comment mettre en pratique ce fameux postulat d'équiprobabilité ? En d'autres termes, quelles sont les méthodes de calcul à notre disposition ?**
- **Comme souvent en mécanique statistique, tous les calculs ou presque sont effectués dans une approximation de champ moyen. Dans certains cas (par exemple les modèles explicitement champ moyen avec interactions indépendantes de la distance), cette approche est exacte. Quelle confiance lui accorder dans les autres cas ?**
- **Les anomalies possibles de la thermodynamique microcanonique (chaleur spécifique négative, inéquivalence d'ensembles...) sont connues. A quel point sont elles générales pour les systèmes avec interactions à longue portée ? Doit-on s'attendre à d'autres anomalies ?**

Les deux premiers points ont suscité beaucoup d'études, approchées ou rigoureuses, dont nous essayons de rendre compte à la section 2.5. Pour aborder ces questions, nous avons utilisé une méthode de "modèles jouets" : nous commençons par étudier des modèles simples, pour lesquels les calculs analytiques peuvent être poussés très loin, et on cherche ensuite à en tirer des idées et des méthodes générales, qui peuvent trouver de nouveaux champs d'applications.

Nous commençons donc par l'étude à l'aide de méthodes élémentaires d'un modèle très simple (chapitre 1), qui nous permettra d'introduire les concepts que nous utiliserons par la suite. Puis nous présenterons une méthode générale pour calculer les équilibres statistiques canonique et microcanonique (section 2); les autres modèles jouets deviennent exemples (section 2.3), et mettent en évidence la phénoménologie particulière de ces systèmes avec interactions à longue portée. Après avoir passé en

revue les résultats généraux déjà connus au sujet de la thermodynamique microcanonique (section 3.1), nous établissons finalement une classification des transitions de phase et des situations d'inéquivalence d'ensembles pour ces systèmes (section 3.2).

Chapter 1

Le modèle Blume-Emery-Griffiths champ moyen

Avec l'idée de disposer de modèles simples, exactement solubles, et montrant les particularités de la thermodynamique des systèmes non additifs, nous avons étudié un système champ moyen de spins 1. Ce système, dit de Blume-Emery-Griffiths [19] (ou plus exactement de Blume-Capel, sous la forme étudiée dans ce chapitre), fut introduit pour modéliser les mélanges binaires $\text{He}^3\text{-He}^4$; il est intéressant en tant que modèle jouet, car son diagramme de phase canonique est déjà connu [19], et présente une phénoménologie assez riche: une ligne de transition de phase du deuxième ordre et une ligne du premier ordre, séparées par un point tricritique. Que deviennent ces transitions de phase dans l'ensemble microcanonique ? Ce travail a été publié dans [15], et repris et détaillé dans un article de synthèse [16].

Le modèle est défini sur un réseau, et chaque site est occupé par une variable de spin 1, $S_i = 0, \pm 1$. Le Hamiltonien s'écrit:

$$H = \Delta \sum_{i=1}^N S_i^2 - \frac{J}{2N} \left(\sum_{i=1}^N S_i \right)^2 \quad (1.1)$$

où $J > 0$ est une constante de couplage ferromagnétique, et Δ contrôle la différence d'énergie entre les états magnétiques ($S_i = \pm 1$) et non magnétique ($S_i = 0$). Nous avons choisi de renormaliser l'interaction à longue portée par $1/N$, ce qui ne fait pas disparaître le caractère non additif du système, mais élimine les problèmes de scaling sur lesquels nous reviendrons plus tard.

La solution canonique

Pour Δ/J petit, le système se rapproche d'un modèle d'Ising champ moyen, et subit une transition de phase du deuxième ordre lorsque β varie. Inversement, à $T = 0$ ($\beta = +\infty$) et $2\Delta/J = 1$, les phases paramagnétique $S_i = 0, \forall i$ et ferromagnétique $S_i = 1, \forall i$ sont dégénérées : il y a transition du premier ordre entre ces deux états fondamentaux.

La solution canonique complète est connue depuis longtemps [19], et s'obtient par les méthodes habituelles; la fonction de partition s'écrit :

$$Z(\beta, N) = \sum_{S_i} \exp \left(-\beta \Delta \sum_i S_i^2 + \frac{\beta J}{2N} \left(\sum_i S_i \right)^2 \right). \quad (1.2)$$

On utilise alors la transformation Gaussienne

$$\exp \left(\frac{\beta N J m^2}{2} \right) = \sqrt{\frac{N}{\pi \beta J}} \int_{-\infty}^{+\infty} dv \exp \left(-\frac{N v^2}{2\beta J} + N m v \right), \quad (1.3)$$

pour pouvoir effectuer la somme sur toutes les configurations :

$$Z(\beta, N) = \sqrt{\frac{N}{\pi \beta J}} \int_{-\infty}^{+\infty} dv \exp \left(-\frac{N v^2}{2\beta J} \right) [1 + 2e^{-\beta \Delta} \cosh v]^N. \quad (1.4)$$

Cette dernière intégrale peut alors être évaluée par la méthode du col dans la limite $N \rightarrow \infty$. On obtient alors l'énergie libre par particule :

$$F(\beta) = -\frac{1}{\beta} \min_v \left(\frac{v^2}{2\beta J} - \ln [1 + 2e^{-\beta \Delta} \cosh v] \right). \quad (1.5)$$

La ligne de transition du deuxième ordre est donnée par l'expression

$$\beta J = \frac{1}{2} e^{\beta \Delta} + 1. \quad (1.6)$$

Le point tricritique séparant cette ligne de la ligne de transitions du premier ordre est situé en $\Delta/J = (\ln 4)/3$, $\beta J = 3$. La ligne de transition du premier ordre doit être obtenue numériquement. Le diagramme de phase canonique est résumé schématiquement par la figure 1.1.

La solution microcanonique

Nous passons maintenant à la solution microcanonique du modèle : nous cherchons donc directement à déterminer l'entropie du système pour une énergie fixée. Notons respectivement N_+ , N_- , N_0 le nombre de spins $+1$, -1 et 0 d'une configuration microscopique donnée. On note q le moment quadrupolaire et m la magnétisation par spin :

$$q = \frac{1}{N} \sum_i S_i^2 = \frac{N_+ + N_-}{N}, \quad (1.7)$$

$$m = \frac{1}{N} \sum_i S_i = \frac{N_+ - N_-}{N}. \quad (1.8)$$

L'énergie par particule, renormalisée par Δ par commodité, s'écrit facilement à l'aide de q et m :

$$e = \frac{H}{\Delta N} \left(q - \frac{J}{2\Delta} m^2 \right). \quad (1.9)$$

Comme $N_0 + N_+ + N_- = N$, la donnée de q et m suffit à connaître N_0, N_+, N_- , et des arguments combinatoires permettent de calculer le nombre de configurations microscopiques avec q et m fixés :

$$\Omega(q, m) = \frac{N!}{N_+! N_-! N_0!}. \quad (1.10)$$

On définit alors s par $\ln \Omega(q, m) = Ns(q, m) + o(N)$; la formule de Stirling donne

$$s(q, m) = -\frac{q+m}{2} \ln \frac{q+m}{2} - \frac{q-m}{2} \ln \frac{q-m}{2} - (1-q) \ln(1-q) - \ln 3. \quad (1.11)$$

L'entropie microcanonique s'obtient alors en maximisant s (1.11) à e (1.9) constant. En exprimant la contrainte sous la forme $q = e + km^2$, avec $k = J/2\Delta$, on se ramène à un problème variationnel à une seule variable :

$$S(e) = \sup_m (s(e + km^2, m)). \quad (1.12)$$

La température microcanonique est alors donnée par l'expression $\Delta\beta = \partial S/\partial e$. Comme dans le cas canonique, l'équation de la ligne de transition du deuxième ordre est calculable : elle correspond à l'annulation du terme en m^2 dans le développement autour de $m = 0$; on trouve exactement l'expression (1.6). Cette ligne critique s'arrête en un point tricritique, qui correspond à l'annulation des termes en m^2 et m^4 dans le développement autour de $m = 0$. On trouve numériquement $k \simeq 1.0813$ et $\beta\Delta \simeq 1.3998$. Ces valeurs sont très proches, mais *différentes* des valeurs canoniques $k = 3/\ln(16) \simeq 1.0820$ et $\beta\Delta = \ln 4 \simeq 1.3995$. La ligne du deuxième ordre s'étend un petit peu plus loin dans le cas microcanonique. Dans la région située entre les deux points tricritiques, la transition est du premier ordre dans l'ensemble canonique, mais reste continue dans l'ensemble microcanonique : une représentation schématique du diagramme de phase dans cette zone est donnée figure 1.2.

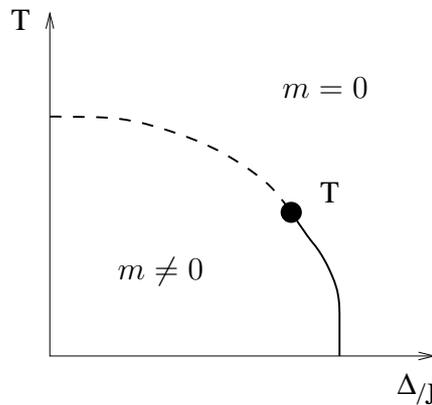


Figure 1.1: Diagramme de phase canonique du modèle BEG champ moyen. Pour Δ/J petit, la transition est du deuxième ordre (ligne pointillée); quand Δ/J augmente, la transition devient du premier ordre (ligne continue); ces deux régimes sont séparés par un point tricritique (T). Pour $\Delta/J > 1/2$, il n'y a plus de transition.

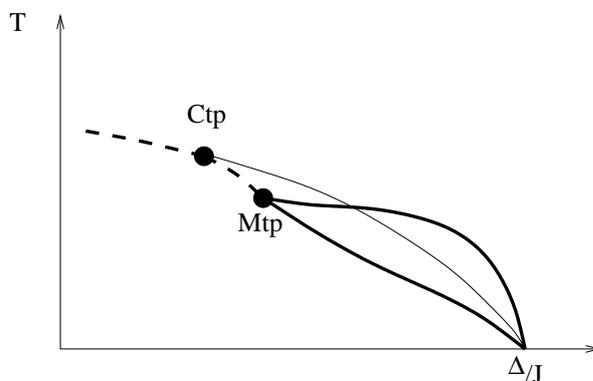


Figure 1.2: Représentation schématique du diagramme de phase dans la zone des deux points tricritiques, canonique (Ctp) et microcanonique (Mtp). En gras tireté, la ligne de transition de phase du second ordre, commune aux deux ensembles à gauche de Ctp; en trait continu fin, la transition de phase canonique du premier ordre; en trait gras continu, les deux températures de la transition de phase microcanonique du premier ordre.

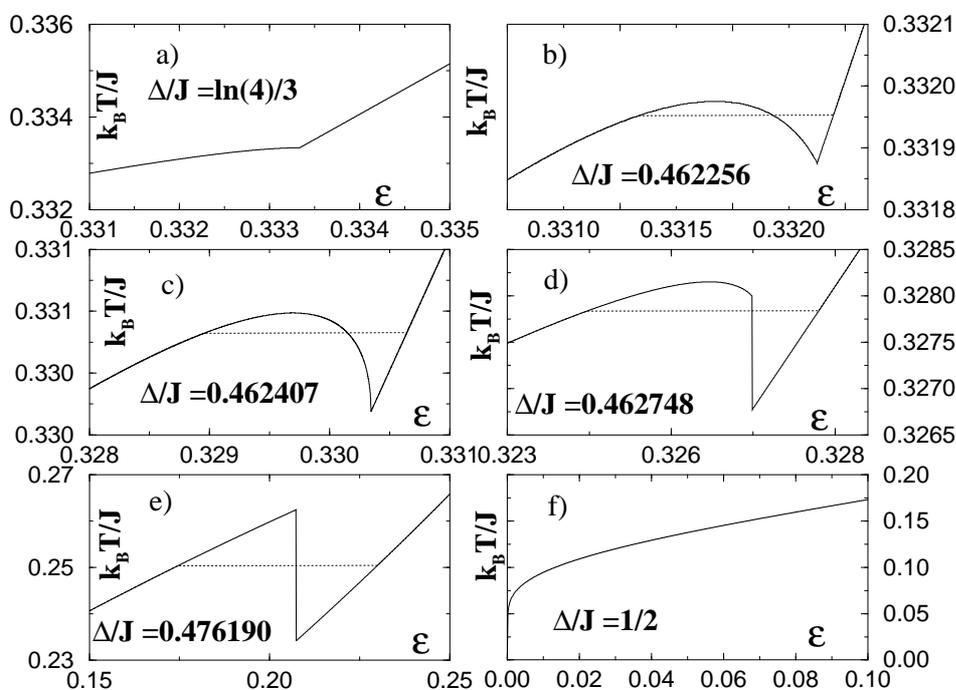


Figure 1.3: Courbes caloriques $T(\epsilon)$, pour différentes valeurs du paramètre $\Delta/J = 2/k$. Les lignes continues représentent l'ensemble microcanonique, et les lignes pointillées l'ensemble canonique, lorsqu'il est différent de l'ensemble microcanonique.

Au delà du point tricritique microcanonique, la température subit une discontinuité au passage de l'énergie critique microcanonique; les deux lignes de la figure 1.2 représentent la température de chaque côté du saut. Toutes les transitions disparaissent à $T = 0$, $\Delta/J = 1/2$.

Les courbes $T(e)$, obtenues numériquement à partir de l'équation (1.12), permettent de mieux comprendre la situation : voir figure 1.3. Au point tricritique canonique la courbe $T(e)$ a une tangente horizontale, ce qui correspond à une chaleur spécifique divergente, figure 1.3a (la chaleur spécifique par particule est définie par $\partial e/\partial T$). Au delà, une zone de chaleur spécifique négative apparaît, invisible dans l'ensemble canonique, car remplacée par une transition du premier ordre, figure 1.3b. Au point tricritique microcanonique, la courbe $T(e)$ est verticale, et la chaleur spécifique approche 0 par valeurs inférieures figure 1.3c. Au delà apparaît un saut de température, figure 1.3d; c'est une discontinuité de la dérivée première de l'entropie, on appellera donc ce point transition du premier ordre microcanonique, par analogie avec la terminologie habituelle ¹. Si on continue à augmenter Δ/J , la zone de chaleur spécifique négative disparaît complètement figure 1.3e.

Le modèle Blume-Emery-Griffiths champ moyen est soluble analytiquement, par des méthodes élémentaires, dans les ensembles canonique et microcanonique, et permet donc une comparaison détaillée des deux ensembles. La phénoménologie autour du point tricritique est particulièrement intéressante, puisqu'elle met en évidence l'apparition de l'inéquivalence entre les deux ensembles, avec des zones de chaleur spécifique négative, et des discontinuités de température.

Nous discuterons à nouveau ces résultats à la lumière de la classification des transitions de phase pour les systèmes avec interactions à longue portée, à la section 3.2. Nous introduisons à la section suivante une méthode générale pour l'étude des propriétés d'équilibre des systèmes avec interactions à longue portée, qui permet de résoudre des modèles bien plus compliqués que celui que nous venons de présenter.

¹Il faut faire attention en comparant avec la littérature, car cette terminologie n'est pas générale, certains auteurs appelant transition du premier ordre toute zone de chaleur spécifique négative.

Chapter 2

Champ moyen et théorie des grandes déviations

L'approximation du champ moyen consiste à évaluer le champ subi par une particule en supposant que toutes les autres sont dans leur état "moyen". Pour un système avec interactions à longue portée, un grand nombre de particules vont participer à ce champ, et les fluctuations de sa valeur réelle autour de sa valeur donnée par le champ moyen devraient être rendues très faibles par la loi des grands nombres. On conçoit donc qu'on puisse obtenir ainsi une très bonne approximation du comportement réel du système. En fait, on peut même montrer pour un certain nombre de systèmes que la solution donnée par le champ moyen est exacte dans la limite d'un très grand nombre de particules. L'objet de cette partie est d'introduire l'outil mathématique qui sera nécessaire par la suite : la théorie des grandes déviations. D'une part, elle permettra de démontrer dans certains cas l'exactitude du champ moyen, d'autre part, et c'est le plus intéressant pour nous, elle fournira une méthode puissante pour calculer effectivement les états d'équilibre microcanoniques et canoniques. La théorie des grandes déviations a été utilisée par Michel et Robert [90] pour justifier rigoureusement la mécanique statistique de l'équation d'Euler à deux dimensions; la méthode a ensuite été généralisée et systématisée par Ellis *et al.* [45] pour étudier d'autres modèles de fluides bidimensionnels ou géophysiques. Sous cette forme, elle peut s'appliquer à un grand nombre de systèmes avec interactions à longue portée.

Après une introduction générale aux idées de la théorie des grandes déviations à la section 2.1, nous expliquons la démarche générale pour résoudre un problème donné, telle qu'elle est décrite par Ellis dans [45]. La section 2.3 sera consacrée à l'étude d'un grand nombre d'exemples.

2.1 Introduction à la théorie des grandes déviations

Nous donnons dans ce paragraphe une très rapide introduction à la théorie des grandes déviations, à ses idées et à ses méthodes. On ne cherchera pas en général à donner aux énoncés des formulations mathématiquement précises. On peut trouver un traitement

rigoureux et complet dans [42] par exemple; la référence [44] est dédiée aux applications de la théorie des grandes déviations à la physique statistique, d'un point de vue de mathématicien.

Qu'est ce qu'un principe de grande déviation ?

Considérons une somme de N variables aléatoires indépendantes X_k , de même distribution, de moyenne nulle, ainsi que la moyenne empirique S_N :

$$S_N = \frac{1}{N} \sum_{k=1}^N X_k. \quad (2.1)$$

La loi des grands nombres nous dit que S_N tend vers la moyenne des X_k , soit 0, lorsque N tend vers l'infini. On peut être plus précis, en utilisant le théorème central limite (si les hypothèses sont vérifiées) : on sait que la fonction $P(\sqrt{N}S_N = x)$ tend vers une Gaussienne en x (on a pris des variables aléatoires de moyenne nulle). Les fluctuations de S_N sont donc d'ordre $1/\sqrt{N}$. On peut vouloir être plus précis, et chercher à savoir ce qui se passe dans les queues de la distribution : quelle est la probabilité d'une fluctuation d'ordre un ? C'est à dire que vaut $P(S_N = x)$? C'est à ce type de questions que cherche à répondre la théorie des grandes déviations.

Donnons un exemple. On lance une pièce pour chaque variable X_k ; $X_k = 1$ si c'est pile, $X_k = -1$ si c'est face. Dans ce cas, $P(S_N = x)$ se calcule facilement. Par calcul combinatoire, on a

$$P(S_N = x) = \frac{N!}{\left(\frac{1+x}{2}N\right)! \left(\frac{1-x}{2}N\right)! 2^N}, \quad (2.2)$$

et, en utilisant la formule de Stirling

$$\ln P(x) \sim -N \left(\frac{1+x}{2} \ln \frac{1+x}{2} + \frac{1-x}{2} \ln \frac{1-x}{2} + \ln 2 \right) \quad (2.3)$$

$$\sim -NI(x). \quad (2.4)$$

Plus précisément, on peut montrer, pour tout segment $]x_1, x_2[$,

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \ln P(x \in]x_1, x_2[) = \max_{x \in]x_1, x_2[} I(x). \quad (2.5)$$

On dit que S_N vérifie un principe de grande déviation, avec fonctionnelle de grande déviation I . Dans le vocabulaire de la mécanique statistique, $I(x)$ est l'opposé de l'entropie associée à une configuration de moyenne x . On voit que les valeurs de x telles que $I(x) > 0$ sont exponentiellement supprimées avec N ; de plus pour que la probabilité soit normalisable, il faut $I(x) \geq 0$, et $\inf I(x) = 0$ (en général, I atteint 0 pour $x = \langle X_k \rangle$).

C'est un exemple de principe de grande déviation, et toute la théorie consiste à chercher ce genre d'estimations, dans des cas plus compliqués. Un outil simple et efficace est fourni par le théorème de Cramér.

Comment obtenir un principe de grande déviation : le théorème de Cramér

Le théorème de Cramér sera constamment utilisé par la suite; il permet de résoudre le problème précédent pour des variables X_k à valeurs dans \mathbb{R}^d , de distribution rapidement décroissante quelconque (mais toujours la même pour chaque variable). On pose encore

$$S_N = \frac{1}{N} \sum_{k=1}^N X_k. \quad (2.6)$$

Alors $P(S_N = \mathbf{x})$ vérifie un principe de grande déviation similaire à celui de (2.4) :

$$\ln P(S_N = \mathbf{x}) \sim -NI(\mathbf{x}). \quad (2.7)$$

On sait donc que les grandes déviations sont supprimées exponentiellement en N , et le théorème nous permet de calculer la fonctionnelle de grande déviation $I(\mathbf{x})$. On définit pour cela la fonction suivante (dont le logarithme est parfois appelé fonction génératrice des cumulants) :

$$\Psi(\lambda) = \langle e^{\lambda \cdot X_1} \rangle, \quad (2.8)$$

où λ est un élément de \mathbb{R}^d , \cdot est le produit scalaire habituel, et $\langle \cdot \rangle$ désigne la moyenne sur la distribution de X_1 (ou de tout X_k , puisqu'elles sont identiquement distribuées). Alors la fonction cherchée $I(\mathbf{x})$ est la transformée de Legendre de $\ln \Psi$:

$$I(\mathbf{x}) = \sup_{\lambda \in \mathbb{R}^d} (\lambda \cdot \mathbf{x} - \ln \Psi(\lambda)) . \quad (2.9)$$

Ce théorème est valable si la distribution des X_k décroît suffisamment vite à l'infini, de telle sorte que Ψ soit définie sur \mathbb{R}^d tout entier. Si ce n'est pas le cas, des versions plus restrictives existent (voir [42] pour une formulation précise et une démonstration). On dispose donc d'une méthode générale pour évaluer ces fonctionnelles de grande déviation I , alors que les méthodes combinatoires échouent par exemple dès que la distribution des X_k est continue.

Sans entrer dans les détails, donnons une justification heuristique à ce résultat, dans le cas où $X_k \in \mathbb{R}$. La probabilité d'obtenir $S_N = x$ est donnée par le volume dans l'espace des phases occupé par les points compatible avec $S_N = x$. Appelons $d\mu$ la distribution de probabilité de chacun des X_k . Alors

$$P(S_N = x) = \int d\mu(X_1) \dots d\mu(X_N) \delta(S_N - x) \quad (2.10)$$

$$\begin{aligned} &= \frac{1}{2\pi i} \int_{\Gamma} d\lambda e^{-N\lambda x} \int d\mu(X_1) \dots d\mu(X_N) e^{\lambda(X_1 + \dots + X_N)} \\ &= \frac{1}{2\pi i} \int_{\Gamma} d\lambda e^{-N\lambda x} [\langle e^{\lambda X} \rangle]^N, \end{aligned} \quad (2.11)$$

où la fonction δ de Dirac a été représentée par une intégrale complexe sur un chemin Γ perpendiculaire à l'axe réel dans le plan des λ complexes. Pour évaluer la dernière intégrale par la méthode du col, il faut déterminer les points critiques de $\lambda x - \ln \Psi(\lambda)$, avec $\Psi(\lambda) = \langle e^{\lambda X} \rangle$, ce qui "justifie" l'expression (2.9).

Une utilisation d'un principe de grande déviation

Donnons un exemple d'utilisation d'un principe de grande déviation. Supposons que le joueur à pile ou face des paragraphes précédents lance N fois la pièce, calcule x , nombre de piles moins nombre de faces divisé par N , et gagne $\exp Nh(x)$ centimes d'euros, avec h une certaine fonction de x . Quelle est son espérance de gain ? On peut écrire

$$\begin{aligned} \ln \langle e^{Nh(x)} \rangle &= \ln \int P(x) e^{Nh(x)} dx \\ &\simeq \ln \int e^{NI(x)} e^{Nh(x)} dx \\ &\simeq N \max_x (I(x) + h(x)). \end{aligned} \quad (2.12)$$

L'espérance de gain du joueur est alors donnée par l'exponentielle de cette dernière expression. La théorie des grandes déviations permet de justifier rigoureusement ce genre de calculs. Si on pense maintenant à la fonction h comme à une énergie, on voit que les applications en physique statistique ne sont pas loin.

2.2 Une méthode générale

Michel et Robert ont utilisé des techniques de grandes déviations [90] pour justifier la mécanique statistique des flots bidimensionnels proposée quelques années auparavant par Miller [91] et Robert et Sommeria [99]. Ellis et ses collaborateurs ont ensuite développé cette approche et proposé une méthode générale pour étudier et résoudre un grand nombre de systèmes avec interactions à longue portée. Nous exposons cette méthode dans ce paragraphe, telle qu'elle est présentée dans [45], en mettant l'accent sur l'utilisation plutôt que sur la rigueur mathématique. Ellis cite en général des applications liées à la turbulence bidimensionnelle; il y en a beaucoup d'autres et nous appliquerons dans ce paragraphe chaque étape sur un exemple plus simple : le modèle de Potts à 3 états, avec interactions de portée infinie. Nous avons choisi cet exemple parce qu'il a déjà été utilisé en tant que modèle jouet pour illustrer la thermodynamique des interactions à longue portée [66] ou celle des systèmes à petit nombre de degrés de liberté [57], et parce qu'il se généralise facilement en augmentant le nombre d'états ou en autorisant des sites vides. La section suivante sera consacrée à l'étude de plusieurs autres exemples.

Le Hamiltonien du modèle de Potts à trois états s'écrit :

$$H_N^{Potts} = -\frac{J}{2} \sum_{i,j=1}^N \delta_{S_i, S_j} - 1. \quad (2.13)$$

Chaque site i du réseau est occupé par un "spin" S_i , qui peut être dans l'un des trois états a , b , ou c . Deux spins contribuent pour $-J$ à l'énergie s'ils sont dans le même état, pour 0 sinon. La somme porte sur toutes les paires de sites (i, j) , puisque l'interaction est supposée à portée infinie. Contrairement à l'usage lorsqu'il s'agit d'approximer le comportement d'interactions à courte portée par un modèle champ moyen, nous ne renormalisons pas la force de l'interaction par $1/N$, ce qui rendrait le système "pseudo

extensif"; pour un système auto-gravitant par exemple, cette renormalisation n'a pas lieu d'être a priori.

Étape 1 : Décrire le système à l'aide d'une variable globale.

Plus précisément, appelons Ω_N l'espace des phases avec N particules; le système est défini par son Hamiltonien

$$\begin{aligned} H_N &: \Omega_N \rightarrow \mathbb{R} \\ \omega_N &\mapsto H_N(\omega_N) \end{aligned}$$

La première étape consiste à associer à chaque configuration microscopique ω_N une variable globale, ou champ moyen, $\mu(\omega_N)$, et un nouvel Hamiltonien \tilde{H}_N de telle sorte que

$$H_N(\omega_N) = \tilde{H}_N(\mu(\omega_N)) + R_N(\omega_N).$$

où le reste $R_N(\omega_N)$ est négligeable par rapport à \tilde{H}_N quand N tend vers l'infini. Prenons l'exemple du modèle de Potts. La bonne variable globale est $\mu = (n_a, n_b)$, avec $(n_a, n_b, n_c = 1 - n_a - n_b)$ représentant respectivement la densité de spins dans les états a, b, c . En effet, on peut écrire

$$H_N^{\text{Potts}} = -\frac{JN^2}{2}(n_a^2 + n_b^2 + n_c^2). \quad (2.14)$$

Dans ce cas, le reste R_N est nul pour toute configuration microscopique ω , μ est de dimension finie, et \tilde{H}_N est le polynôme du membre de droite de l'équation précédente. Ce n'est pas toujours aussi simple : μ peut être une densité locale de masse en astrophysique, une densité moyenne locale de la vorticit  en turbulence 2D (ou un objet encore plus compliqu , lorsqu'il s'agit de m canique statistique d'une  quation continue comme l' quation d'Euler). La variable globale est alors de dimension infinie. On verra dans la partie 2.3 des exemples de ce type, o  R_N n'est pas nul. Du fait des interactions   longue port e, H_N n'a pas ici le scaling habituel en N ; on reviendra sur cette question par la suite.

Bien s r, il n'est pas toujours possible d' crire le Hamiltonien d'un syst me   l'aide d'une variable champ moyen, m me de dimension infinie et en s'autorisant une erreur R_N . Ce sera possible en g n ral pour les syst mes   longue port e sur un r seau, le pas du r seau r gularisant l' ventuelle singularit    courte distance.

 tape 2 : Obtenir une fonctionnelle d'entropie pour la variable globale.

On ne consid re pas ici l' nergie. M me ainsi, les valeurs de la variable champ moyen μ ne sont pas  quiprobables. En effet, un grand nombre de configurations microscopiques peuvent donner lieu   un m me μ , et ce nombre d pend de la valeur de μ . On d finit donc une fonctionnelle d'entropie $s(\mu)$, comme le logarithme du nombre de configurations microscopiques $\Omega_N(\mu)$ correspondant   μ , divis  par N :

$$\Omega_N(\mu) \propto e^{Ns(\mu)}. \quad (2.15)$$

Ici intervient la théorie des grandes déviations, d'abord pour nous assurer que la fonctionnelle s existe bien, et ensuite pour nous permettre de la calculer.

Donnons l'exemple du modèle de Potts. En posant N_a, N_b, N_c le nombre de spins a, b, c d'une configuration microscopique, on peut, cette fois encore, calculer s en utilisant des arguments combinatoires et la formule de Stirling. Utilisons plutôt le théorème de Cramér. μ s'écrit sous la forme

$$\mu = \left(\frac{1}{N} \sum_i \delta_{S_i, a}, \frac{1}{N} \sum_i \delta_{S_i, b} \right). \quad (2.16)$$

μ est ainsi exprimée comme une moyenne de variables aléatoires indépendantes à valeurs dans \mathbb{R}^2 . On peut appliquer le théorème de Cramér :

$$\begin{aligned} \Psi(\lambda_a, \lambda_b) &= \frac{1}{3} \sum_{S=a,b,c} (e^{\lambda_a \delta_{S,a} + \lambda_b \delta_{S,b}}) \\ &= \frac{1}{3} (e^{\lambda_a} + e^{\lambda_b} + 1). \end{aligned}$$

La fonctionnelle de grande déviation est donc donnée par le problème d'optimisation :

$$I(\mu) = \sup_{\lambda_a, \lambda_b} (\lambda_a n_a + \lambda_b n_b - \ln \Psi(\lambda_a, \lambda_b)). \quad (2.17)$$

On peut résoudre explicitement, et on obtient :

$$I(\mu) = n_a \ln n_a + n_b \ln n_b + (1 - n_a - n_b) \ln(1 - n_a - n_b) + \ln 3. \quad (2.18)$$

La fonctionnelle d'entropie $s(\mu)$ est égale à l'opposé de I ; on obtiendrait évidemment le même résultat par la méthode combinatoire. Le terme $\ln 3$ assure que le maximum de l'entropie est exactement 0.

Étape 3 : Écrire les problèmes variationnels.

- La solution microcanonique

Lorsque les étapes 1 et 2 sont achevées, c'est à dire lorsque :

-on a une approximation du Hamiltonien par une fonctionnelle dépendant d'une variable globale μ , du type (2.14); on suppose que $\tilde{H}_N \propto N^p$ à grand N ;

-cette variable globale vérifie un principe de grande déviation, i.e. $\Omega_N(\mu) \propto e^{N s(\mu)}$ pour une certaine fonctionnelle s ;

le théorème 3.2 de la référence [46] assure que la solution microcanonique, dans la limite $N \rightarrow \infty$, $H_N/N^p = cste$, est donnée exactement par la solution du problème variationnel

$$S(e) = \sup_{\mu} \left(s(\mu) \mid \tilde{H}_N(\mu)/N^p = e \right). \quad (2.19)$$

Ce résultat est exact (dans la limite $N \rightarrow \infty$, $H_N/N^p = cste$), et est précisément celui donné par une solution champ moyen "intuitive"; c'est en ce sens qu'on peut dire que le champ moyen devient exact pour ces systèmes. Remarquons que l'expression ci-dessus permet de calculer l'entropie par particule $S(e)$; comme le montre l'expression (2.15),

l'entropie totale est extensive, c'est à dire proportionnelle à N dans cette limite : il n'y a pas ici de différence avec les interactions à courte portée.

Dans le cas du modèle de Potts avec interactions à portée infinie, $p = 2$ et la solution microcanonique s'écrit

$$S(e) = \sup_{n_a, n_b} \left(-n_a \ln n_a - n_b \ln n_b - (1 - n_a - n_b) \ln(1 - n_a - n_b) \right) - \frac{J}{2} (n_a^2 + n_b^2 + (1 - n_a - n_b)^2) = e \quad (2.20)$$

Résoudre ce problème d'optimisation permet de retrouver la courbe de la figure 6 de la référence [66] : voir figure 2.1. Notons que l'énergie maximale du système pour $J = 1$ est $e_{max} = -1/6$.

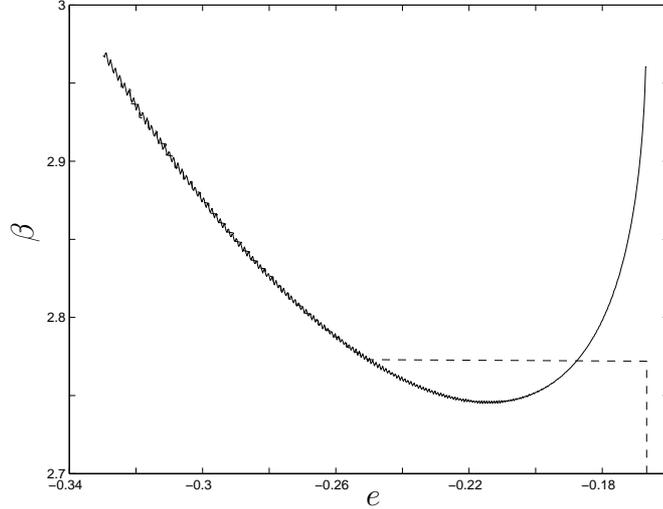


Figure 2.1: Courbe calorique $\beta(e)$ du modèle de Potts à 3 états champ moyen. En trait plein la solution microcanonique, obtenue à partir de l'équation (2.20); les petites oscillations sont dues aux imprécisions numériques lors du calcul de la dérivée de $S(e)$. Pour $e > -0.215$, $d\beta/de > 0$, la chaleur spécifique est donc négative. En tireté la solution canonique, obtenue à partir de l'équation (2.24). Elle coïncide avec la solution microcanonique pour $e < -0.255$ environ, et subit ensuite une transition de phase du premier ordre (discontinuité de e).

- La solution canonique

A une constante de normalisation près, la fonction de partition $Z(\beta, N)$ d'un système est la moyenne de $e^{-\beta H_N}$, prise sur toutes les configurations microscopiques équiprobables :

$$Z(\beta, N) \propto \langle e^{-\beta H_N} \rangle . \quad (2.21)$$

Si on remplace H_N par son expression approchée en fonction de μ , \tilde{H}_N :

$$Z(\beta, N) \simeq \left\langle \exp \left(-N \frac{\beta \tilde{H}_N(\mu)}{N} \right) \right\rangle . \quad (2.22)$$

Cette expression est très similaire à l'équation (2.12), et peut être évaluée de la même façon, en se ramenant à un problème d'optimisation.

Revenons à l'exemple du modèle de Potts. On a $\tilde{H}_N/N \propto N$, donc si β est indépendant de N , le terme énergétique l'emporte toujours à grand N : l'état fondamental est toujours infiniment plus probable que les autres, quel que soit β . En réalité, si on veut explorer avec l'ensemble canonique la même gamme d'énergies qu'avec l'ensemble microcanonique, c'est à dire $H_N/N^p = e$, il faut choisir pour β une loi d'échelle $\beta \propto N^{1-p}$. Si on prend $\beta = \beta_0 N^{1-p}$, la solution est donnée rigoureusement dans la limite $N \rightarrow \infty$ par la minimisation de la fonctionnelle $f(\mu) = \beta_0 \tilde{H}_N(\mu)/N^p - s(\mu)$ (c'est une application du théorème 2.4 de la référence [46]). En posant $h(\mu) = \tilde{H}_N(\mu)/N^p$, on a pour l'énergie libre par particule (à un facteur β près) : ¹

$$F(\beta_0) = \inf_{\mu} (\beta_0 h(\mu) - s(\mu)) . \quad (2.23)$$

Cette fois encore, c'est le résultat qu'on aurait attendu par une approximation champ moyen. Pour le modèle de Potts, cette solution s'écrit

$$F(\beta) = \sup_{n_a, n_b} \left(n_a \ln n_a + n_b \ln n_b + (1 - n_a - n_b) \ln(1 - n_a - n_b) - \frac{\beta J}{2} (n_a^2 + n_b^2 + (1 - n_a - n_b)^2) \right) . \quad (2.24)$$

Le résultat de cette maximisation, ainsi que sa comparaison avec la solution microcanonique est montrée figure 2.1, par l'intermédiaire de la courbe calorique $\beta(e)$. La solution canonique coïncide avec la solution microcanonique pour $e < -0.255$, ce qui correspond à $\beta > \beta_c = 2.75$ environ; si l'on augmente encore la température, le système subit une transition de phase du premier ordre et c'est la phase homogène qui devient stable. Pour tous les $\beta < \beta_c$, on a donc $e = e_{max} = -1/6$, ce qui explique la branche verticale de la solution canonique. L'étude systématique des points communs et différences entre solutions canonique et microcanonique est l'objet des parties 3.1 et 3.2.

Ceci termine l'exposé de la méthode générale pour résoudre la mécanique statistique d'équilibre des systèmes à longue portée. Retrouver de manière rigoureuse les prédictions du champ moyen pour le modèle de Potts avec interactions à portée infinie n'est bien sûr pas surprenant; mais on va voir à la section 2.3.4 que cette méthode s'applique aussi dans des situations plus délicates, par exemple lorsque les interactions décroissent en loi de puissance avec la distance. Dans ces cas là, il est remarquable que l'on puisse obtenir des solutions exactes.

D'autre part, ce paragraphe a mis en lumière l'importance de la manière avec laquelle on fait tendre N vers l'infini. Nous y reviendrons à la section 2.4.

¹on appellera toujours dans cette thèse "énergie libre" une quantité qui est en réalité F/β , par commodité pour comparer les deux ensembles, et aussi parce que cette quantité F/β doit toujours être minimisée, alors que la véritable énergie libre doit être maximisée si $\beta < 0$. On omettra aussi les facteurs k_B de l'entropie.

2.3 Exemples

Cette partie est consacrée à l'étude de nombreux modèles simples à l'aide de la méthode exposée ci-dessus. Nous avons déjà résolu le modèle Blume-Emery-Griffiths (BEG) champ moyen par des méthodes élémentaires, au chapitre 1. Nous en donnons ici une solution par la méthode des grandes déviations. Le modèle BEG est un modèle de spins, sans dynamique. Nous étudions ensuite les modèles HMF en dimensions un et deux (Heisenberg Mean Field; ce sont des modèles de type XY, champ moyen, auxquels on a ajouté un terme d'énergie cinétique), qui ont une véritable dynamique. La méthode des grandes déviations donne la solution microcanonique de ces modèles, qui n'est pas accessible autrement (paragraphe 2.3.2). Nous donnons également un exemple d'application de la méthode des grandes déviations à un modèle de laser, un petit peu plus compliqué que les précédents (paragraphe 2.3.3). Enfin, nous terminons par l'étude du modèle d'Ising avec interactions décroissant en $1/r^\alpha$ avec la distance (paragraphe 2.3.4). Ceci nous permettra de répondre définitivement à certaines questions soulevées récemment par l'étude de ces modèles, qui restaient en partie ouvertes faute d'avoir accès à la solution microcanonique [3, 31, 101]: la densité d'états croît-elle de façon exponentielle avec N ? Pourquoi la solution de certains de ces modèles avec interactions en $1/r^\alpha$ est-elle indépendante de la valeur de α (tant que $\alpha < D$, dimension de l'espace)?

2.3.1 Retour sur le modèle Blume-Emery-Griffiths

Nous appliquons pas à pas les trois étapes de la méthode, au modèle BEG défini par le Hamiltonien (1.1).

Étape 1 :

Le modèle est explicitement champ moyen, donc l'énergie se réécrit facilement à l'aide d'une variable globale μ . Ici, $\mu = (q, m) \in \mathbb{R}^2$, avec $q = 1/N \sum S_i^2$ et $m = 1/N \sum S_i$. On a alors $H = N(\Delta q - J/2m^2) = Nh(q, m)$. Le comportement du système ne dépend que du rapport Δ/J .

Étape 2 :

Appliquons le théorème de Cramér pour calculer l'entropie associée à μ :

$$\begin{aligned} \Psi(\lambda, \rho) &= \langle e^{\lambda S_i^2 + \rho S_i} \rangle \\ &= \frac{1 + 2e^\lambda \cosh \rho}{3}. \end{aligned} \quad (2.25)$$

Pour calculer $s(q, m)$, il faut maintenant résoudre $\partial \ln \Psi / \partial \lambda = q$ et $\partial \ln \Psi / \partial \rho = m$, pour λ et ρ . Le calcul peut se faire explicitement :

$$\lambda = \ln \frac{q\sqrt{1-r^2}}{2(1-q)}, \quad (2.26)$$

$$\rho = \ln \sqrt{\frac{1+r}{1-r}}, \quad (2.27)$$

où r est le rapport m/q . En substituant dans (2.9), on obtient s :

$$s(q, m) = -\frac{q+m}{2} \ln \frac{q+m}{2} - \frac{q-m}{2} \ln \frac{q-m}{2} - (1-q) \ln(1-q) - \ln 3 . \quad (2.28)$$

Étape 3 :

La solution microcanonique pour une énergie par spin e est donnée par

$$S(e) = \sup_{q,m} (s(q, m) | h(q, m) = e) , \quad (2.29)$$

et l'énergie libre s'écrit (à un facteur β près)

$$F(\beta) = \inf_{q,m} \left(\frac{q+m}{2} \ln \frac{q+m}{2} + \frac{q-m}{2} \ln \frac{q-m}{2} + (1-q) \ln(1-q) + \beta \Delta q - \beta \frac{J}{2} m^2 \right) . \quad (2.30)$$

En écrivant $q = e + km^2$ (comme plus haut, $k = J/2\Delta$; on peut également supposer sans restriction que $\Delta = 1$), on se ramène bien au problème d'optimisation trouvé précédemment (1.12). Il est moins évident de voir que les deux solutions canoniques sont bien identiques; ce sont pourtant strictement les mêmes.

2.3.2 Le modèle HMF (Heisenberg Mean Field)

Le modèle HMF est une simplification d'un système de plans massifs (ou de plans chargés dans le cas répulsif) se déplaçant en une dimension, avec conditions aux limites périodiques : on ne garde que la première harmonique du potentiel $V(x) = \pm|x|$ (le signe \pm dépendant du cas, Coulombien ou gravitationnel). La taille du système est fixée à 2π , et la masse des plans à 1. Le Hamiltonien s'écrit

$$H = \sum_{i=1}^N \frac{p_i^2}{2} + \frac{C}{2N} \sum_{i,j} \cos(\theta_i - \theta_j) . \quad (2.31)$$

La somme pour l'énergie potentielle s'étend sur tous les couples i, j ; les interactions sont attractives si $C > 0$, répulsives si $C < 0$. La renormalisation par N de l'énergie potentielle est équivalente à une renormalisation du temps par \sqrt{N} ; nous l'adoptons car elle rend plus facile la comparaison entre des systèmes de différentes tailles, en rendant l'énergie (artificiellement) extensive.

Introduit à l'origine pour étudier quelles étaient les propriétés des systèmes auto-gravitants retenues par ce modèle très simple [65], il est devenu un modèle jouet très utilisé pour tester les idées générales concernant les interactions à longue portée, à l'équilibre et hors équilibre (voir [4, 82], et pour une revue récente [41]). Nous l'étudierons en détail au chapitre suivant, du point de vue de la dynamique hors équilibre.

L'équilibre canonique s'obtient facilement par les méthodes traditionnelles, dans les

cas attractifs et répulsifs, mais la solution microcanonique restait inaccessible. Or il est important de détecter d'éventuelles inéquivalences d'ensembles, pour comparer les prévisions statistiques avec les simulations de dynamique moléculaire, effectuées en général à énergie constante.

Étape 1 :

Le Hamiltonien se réécrit sous la forme

$$H = \sum_{i=1}^N \frac{p_i^2}{2} + \frac{NC}{2}(M_x^2 + M_y^2) \quad (2.32)$$

où $M_1 = M_x + iM_y = \sum_k e^{i\theta_k}/N$. Il y a deux quantités conservées, H et la quantité de mouvement totale $P = \sum_i p_i$. On va donc prendre comme variable globale $\mu = (u, v, M_x, M_y)$, où $u = \sum_i p_i^2/N$ et $v = \sum_i p_i/N$. Les deux quantités conservées s'expriment exactement à l'aide de μ .

Étape 2 :

Le théorème de Cramér permet d'associer facilement une entropie à la partie configurationnelle :

$$\Psi(\lambda_x, \lambda_y) = \frac{1}{2\pi} \int_0^{2\pi} e^{\lambda_x \cos \theta + \lambda_y \sin \theta} d\theta = I_0 \left(\sqrt{\lambda_x^2 + \lambda_y^2} \right), \quad (2.33)$$

où I_0 est la fonction de Bessel modifiée d'ordre 0. On peut simplifier en utilisant la symétrie de rotation sur θ : l'entropie $s_{conf}(M_x, M_y)$ ne dépend que de $|M_1| = \sqrt{M_x^2 + M_y^2}$. On obtient

$$s_{conf}(|M_1|) = - \max_{\lambda} (\lambda |M_1| - \ln I_0(\lambda)). \quad (2.34)$$

On ne peut pas appliquer immédiatement le théorème de Cramér à la partie cinétique; en effet, p_i , qui joue le rôle de la variable aléatoire X_k , est à priori distribué uniformément sur \mathbb{R} , ce qui n'a pas de sens. En réalité, lorsqu'on calcule l'entropie associée à la partie cinétique $s_{cin}(u, v)$, p_i est borné puisque u est fixé. Une application "formelle" du théorème de Cramér, oubliant les normalisations, donne le bon résultat, à une constante (infinie !) près, qui n'est pas importante dans le processus d'optimisation :

$$\Psi(\lambda_u, \lambda_v) = \int_{\mathbb{R}} e^{\lambda_u x^2 + \lambda_v x} dx = e^{-\lambda_v^2/4\lambda_u} \sqrt{\frac{\pi}{-\lambda_u}}. \quad (2.35)$$

La maximisation sur λ_u, λ_v donne finalement

$$s_{cin}(u, v) = \frac{1}{2} \ln(u - v^2) + cste. \quad (2.36)$$

Étape 3 :

On peut maintenant écrire les problèmes variationnels canonique et microcanonique (on suppose par exemple que la quantité de mouvement totale P est nulle) :

$$S(e) = \sup_{u, |M_1|} \left(\frac{1}{2} \ln u + s_{conf}(|M_1|) \left| \frac{1}{2} u + \frac{C}{2} |M_1|^2 = e \right. \right), \quad (2.37)$$

et

$$F(\beta) = \inf_{u, |M_1|} \left(\frac{\beta}{2} (u + C|M_1|^2) - \frac{1}{2} \ln u - s_{conf}(|M_1|) \right). \quad (2.38)$$

La résolution numérique permet bien sûr de retrouver les résultats canoniques déjà connus, et montre d'autre part que dans ce cas, les deux ensembles sont équivalents, que l'interaction soit attractive ou répulsive. Nous reviendrons sur ce modèle à la partie II, dans les cas attractif et répulsif.

Le modèle HMF en dimension deux

Toujours dans l'esprit de disposer de modèles simples sur lesquels on peut effectuer de longues simulations et tester ses idées, une généralisation à deux dimensions du modèle HMF a été introduite. Antoni et Torcini [7] résolvent ce modèle dans l'ensemble canonique, comparent les résultats avec des simulations de dynamique moléculaire à énergie constante et mettent en évidence une phénoménologie assez riche, avec inéquivalence d'ensembles. Néanmoins, ils ne peuvent résoudre le modèle dans l'ensemble micro-canonique. La méthode des grandes déviations s'applique sans difficulté à ce problème, et permettrait certainement de retrouver les résultats numériques dans les zones où la solution canonique échoue.

2.3.3 Un modèle de laser CARL

Un laser de type CARL (Collective Atomic Recoil Laser), comme les lasers habituels, émet une lumière cohérente par émission stimulée d'une transition atomique. Cependant, les comportements collectifs des degrés de liberté *externes* (position, quantité de mouvement) des atomes jouent un rôle fondamental dans le processus d'amplification, alors qu'on ne considère habituellement que les degrés de liberté *internes* (polarisation, inversion de population). Ce type de laser a été proposé récemment [22], et est actuellement l'objet d'études analytiques, numériques et expérimentales [95]. Nous proposons ici une étude statistique d'un modèle simplifié du processus d'amplification : cette approche est tout à fait nouvelle, et peut être complémentaire de l'approche dynamique usuelle. Il s'agit cependant à ce stade d'une illustration de la puissance de la méthode des grandes déviations dans des situations très diverses, plutôt qu'une réelle application à la physique de ces lasers, qui demanderait une étude plus approfondie. Le chapitre 6 est consacré à une analyse statistique détaillée à l'équilibre et hors équilibre d'un autre type de laser, le laser à électrons libres.

Présentation du modèle

On étudie un modèle de laser CARL dans une cavité parfaite (sans perte), et en négligeant les termes d'émission spontanée. Les atomes sont vus comme des particules qui se déplacent et ont une structure interne à deux niveaux; la mécanique quantique de ces deux niveaux est traitée dans l'approximation semi-classique, comme on le fait pour obtenir les équations de Maxwell-Bloch d'un laser conventionnel. Les atomes sont donc caractérisés par leur "position" $\theta_j \in [0, 2\pi]$ (il s'agit plutôt de leur phase

par rapport à l'onde, qui dépend de leur position), leur quantité de mouvement p_j , leur polarisation $\mathbf{S}_j = (S_j^x, S_j^y)$, et leur "inversion de population" D_j ($D_j = 1$ -resp. 0- si l'atome est dans son état excité -resp. fondamental-). Le champ est décrit par deux quantités complexes $\mathbf{f} = (f^x, f^y)$ et $\mathbf{b} = (b^x, b^y)$, représentant les ondes se propageant dans la cavité, respectivement dans le sens positif et négatif. Les équations du mouvement sont alors (nous ne détaillons pas leur obtention, ni le sens physique des différents paramètres) :

$$\dot{\theta}_j = \Omega p_j \quad (2.39a)$$

$$\dot{p}_j = -\frac{\gamma}{4} [\mathbf{S}_j^* (\mathbf{f} e^{i\theta_j} - \mathbf{b} e^{-i\theta_j}) + c.c.] \quad (2.39b)$$

$$\dot{\mathbf{S}}_j = D_j (\mathbf{f} e^{i\theta_j} + \mathbf{b} e^{-i\theta_j}) \quad (2.39c)$$

$$\dot{D}_j = -\frac{\gamma}{2} [\mathbf{S}_j^* (\mathbf{f} e^{i\theta_j} + \mathbf{b} e^{-i\theta_j}) + c.c.] \quad (2.39d)$$

$$\dot{\mathbf{f}} = i\Delta \mathbf{f} + \frac{C}{N} \sum_j \mathbf{S}_j e^{-i\theta_j} \quad (2.39e)$$

$$\dot{\mathbf{b}} = i\Delta \mathbf{b} + \frac{C}{N} \sum_j \mathbf{S}_j e^{i\theta_j} . \quad (2.39f)$$

Ces équations du mouvement conservent pour chaque atome la norme du vecteur de Bloch (c'est la traduction de la normalisation de la fonction d'onde) :

$$|\mathbf{S}_j|^2 + D_j^2 = 1 . \quad (2.40)$$

Il y a également trois quantités conservées globales; l'énergie

$$H = Ne = \sum_j \Omega \frac{p_j^2}{2} + \frac{\gamma}{4} \sum_j [-i \mathbf{S}_j^* (\mathbf{f} e^{i\theta_j} + \mathbf{b} e^{-i\theta_j}) + c.c.] + \frac{\Delta \sqrt{\gamma}}{2} \sum_j D_j , \quad (2.41)$$

l'impulsion

$$P = N\sigma = \frac{C}{\gamma} \sum_j p_j + \frac{N}{4} (|\mathbf{f}|^2 - |\mathbf{b}|^2) , \quad (2.42)$$

et le nombre d'excitations

$$\mathcal{N} = Nn = N (|\mathbf{f}|^2 + |\mathbf{b}|^2) + \frac{2C}{\sqrt{\gamma}} \sum_j D_j . \quad (2.43)$$

De plus, les équations du mouvement conserve l'élément de volume de l'espace des phases dV

$$dV = \prod_j d\theta_j dp_j dS_j^x dS_j^y df^x df^y db^x db^y . \quad (2.44)$$

Il est donc raisonnable d'étudier la mécanique statistique de ce système, en calculant le volume de son espace des phases à l'aide de la mesure ci-dessus, et en tenant compte des contraintes imposées par les quantités conservées.

Solution statistique**Étape 1 :**

On commence par utiliser la conservation de la norme du vecteur de Bloch pour poser

$$S_j^x = \sin \phi_j \cos \alpha_j \quad (2.45a)$$

$$S_j^y = \sin \phi_j \sin \alpha_j \quad (2.45b)$$

$$D_j = \cos \phi_j , \quad (2.45c)$$

où $\alpha_j \in [0, 2\pi[$ et $\phi_j \in [0, \pi[$. On peut alors réécrire l'énergie sous la forme :

$$\begin{aligned} Ne = & \sum_j \Omega \frac{p_j^2}{2} + \frac{\gamma}{2} \left[f^x \sum_j \sin \phi_j \sin(\theta_j - \alpha_j) + f^y \sum_j \sin \phi_j \cos(\theta_j - \alpha_j) \right. \\ & \left. - b^x \sum_j \sin \phi_j \sin(\theta_j + \alpha_j) + b^y \sum_j \sin \phi_j \cos(\theta_j + \alpha_j) \right] \\ & + \frac{\Delta\sqrt{\gamma}}{2} \sum_j \cos \phi_j . \end{aligned} \quad (2.46)$$

On pose maintenant $\mathbf{f} = f e^{i\varphi_f}$, $\mathbf{b} = b e^{i\varphi_b}$, et $q_j^\pm = \theta_j \pm \alpha_j$. L'expression de l'énergie se simplifie :

$$\begin{aligned} Ne = & \sum_j \Omega \frac{p_j^2}{2} + \frac{\gamma}{2} \left[f \sum_j \sin \phi_j \sin(q_j^- + \varphi_f) - b \sum_j \sin \phi_j \sin(q_j^+ - \varphi_b) \right] \\ & + \frac{\Delta\sqrt{\gamma}}{2} \sum_j \cos \phi_j . \end{aligned} \quad (2.47)$$

On peut donc choisir comme variable globale $\mu = (u, v, s_1, s_2, d, f, b)$, avec

$$u = \frac{1}{N} \sum_j p_j^2 \quad (2.48a)$$

$$v = \frac{1}{N} \sum_j p_j \quad (2.48b)$$

$$s_1 = \frac{1}{N} \sum_j \sin \phi_j \sin(q_j^- + \varphi_f) \quad (2.48c)$$

$$s_2 = \frac{1}{N} \sum_j \sin \phi_j \sin(q_j^+ - \varphi_b) \quad (2.48d)$$

$$d = \frac{1}{N} \sum_j \cos \phi_j . \quad (2.48e)$$

d représente l'inversion de population moyenne, et s_1 et s_2 quantifient la cohérence qui se développe entre la position des atomes et leur polarisation. Toutes les quantités

conservées s'expriment à l'aide de μ :

$$e = \Omega \frac{u}{2} + \frac{\gamma}{2} (f s_1 - b s_2) + \frac{\Delta \sqrt{\gamma}}{2} d \quad (2.49a)$$

$$\sigma = \frac{C}{\gamma} v + \frac{1}{4} (f^2 - b^2) \quad (2.49b)$$

$$n = \frac{2C}{\sqrt{\gamma}} d + f^2 + b^2 . \quad (2.49c)$$

La conservation des normes des vecteurs de Bloch a automatiquement été prise en compte par la paramétrisation utilisée.

Étape 2 :

Il faut maintenant associer une entropie à μ . La partie cinétique se traite comme pour le modèle HMF :

$$s_{cin}(u, v) = \frac{1}{2} \ln (u - v^2) . \quad (2.50)$$

La partie configurationnelle $s_{conf}(s_1, s_2, d)$ est un petit peu plus délicate, mais ne nécessite que l'application du théorème de Cramér :

$$\Psi(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{8\pi^2} \int_0^\pi \sin \phi d\phi \int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta e^{\lambda_3 \cos \phi + \sin \phi (\lambda_1 \sin(\theta - \alpha + \varphi_f) + \lambda_2 \sin(\theta + \alpha - \varphi_b))} . \quad (2.51)$$

Après changement de variables et intégration sur q_j^+ et q_j^- , on obtient :

$$\Psi(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{2} \int_0^\pi \sin \phi d\phi e^{\lambda_3 \cos \phi} I_0(\lambda_1 \sin \phi) I_0(\lambda_2 \sin \phi). \quad (2.52)$$

L'entropie configurationnelle est alors :

$$s_{conf}(s_1, s_2, d) = - \max_{\lambda_1, \lambda_2, \lambda_3} [\lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 d - \ln \Psi(\lambda_1, \lambda_2, \lambda_3)] \quad (2.53)$$

On voit que les phases des champs n'interviennent plus.

Étape 3 :

On peut exprimer u , v et d à l'aide de e , σ , n , s_1 , s_2 , f et b en utilisant les expressions des quantités conservées pour obtenir un problème variationnel sans contraintes :

$$S(e, \sigma, n) = \max_{s_1, s_2, f, b} \left[\frac{1}{2} \ln(u - v^2) + s_{conf}(s_1, s_2, d) \right] , \quad (2.54)$$

avec

$$u = \frac{2}{\Omega} e - \frac{2\gamma}{\Omega} (f s_1 - b s_2) - \frac{\Delta \gamma}{2\Omega C} n + \frac{\Delta \gamma}{2\Omega C} (f^2 + b^2) \quad (2.55a)$$

$$v = \frac{\gamma}{C} \sigma - \frac{1}{4} (f^2 - b^2) \quad (2.55b)$$

$$d = \frac{\sqrt{\gamma}}{2C} (n - f^2 - b^2) . \quad (2.55c)$$

Ce problème d'optimisation doit a priori être résolu numériquement. On peut néanmoins essayer de pousser plus loin les calculs analytiques, en développant autour de $(s_1, s_2) = (0, 0)$: si $(s_1, s_2) = (0, 0)$ maximise l'entropie, aucune cohérence spatiale ne se développe, et les ondes ne sont pas amplifiées [95]. On suppose aussi que $\sigma = 0$ (quantité de mouvement totale nulle). On peut alors montrer que $v = 0$, $f = b$ et $s_1 = -s_2$. Pour une certaine gamme de paramètres, le système présente une transition de phase du deuxième ordre, lorsqu'on fait varier l'énergie e ou le nombre d'excitations n : l'état avec une onde d'amplitude négligeable est favorable à haute énergie ou à bas n . La figure 2.2 résume ce comportement. Le but de cet exemple est plutôt de montrer la

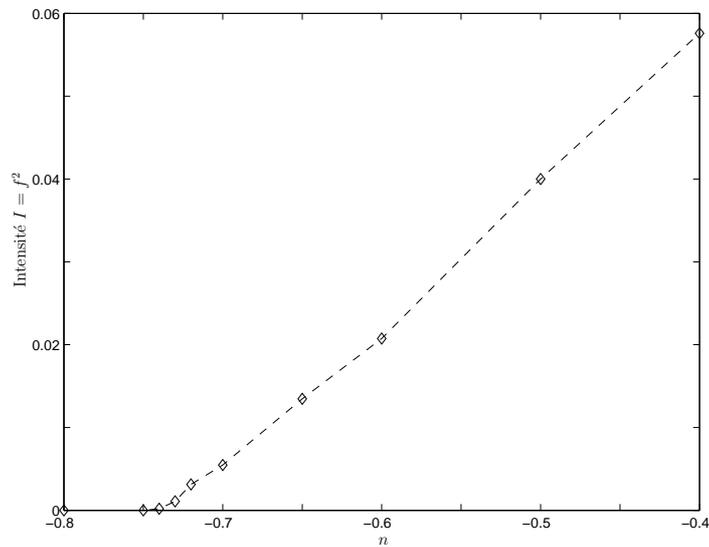


Figure 2.2: Intensité d'équilibre de l'onde $I = f^2$ en fonction du nombre d'excitations n . Le système subit une transition de phase du deuxième ordre en $x_c \sim -0.75$. Paramètres : $\Delta = 0$, $\gamma = C = \Omega = 1$. Quantités conservées : $\sigma = 0$, $e = 0.125$.

puissance de la méthode des grandes déviations, même pour des modèles un peu plus compliqués que les précédents. Une analyse détaillée de la dynamique et de la physique du système serait bien sûr nécessaire pour une véritable application à la physique des lasers de type CARL. Nous ferons cette démarche pour un laser à électrons libres, à la partie III.

2.3.4 Le modèle d'Ising, avec interactions en $1/r^\alpha$

Le problème

On n'a vu jusqu'ici que des modèles pour lesquels les interactions sont indépendantes de la distance. L'énergie s'exprime alors exactement à l'aide de quelques variables globales. Ce n'est pas le cas pour beaucoup d'interactions physiquement intéressantes (en $1/r$ pour la gravité, en $\ln r$ pour la turbulence en dimension 2). Les modèles jouets avec interactions décroissant lentement avec la distance en $1/r^\alpha$ ont donc suscité un certain intérêt récemment [2, 3, 23, 31, 32]. Numériquement, Anteneodo *et al.* [3] ont montré

que le comportement thermodynamique de ces modèles est indépendant de l'exposant α , pour des conditions aux limites périodiques et une fois qu'un réechelonnement correct de l'énergie est choisi. Il ne s'agit pas seulement ici d'un comportement critique identique, mais bien d'une correspondance parfaite sur toute la gamme d'énergie, ou de température selon l'ensemble choisi (les courbes magnétisation/température sont identiques quel que soit α pour un modèle d'Ising, par exemple).

Ils ont aussi été étudiés en tant que candidats possibles à l'application de la statistique de Tsallis : certains auteurs ont soutenu que pour ces systèmes non extensifs, le nombre d'états ne se comportait peut être pas de façon exponentielle en N , le nombre de degrés de liberté, et qu'il était donc nécessaire de changer la définition habituelle de l'entropie $S = k_B \ln \Omega$ [101].

Campa *et al.* [31], et indépendamment (dans un cadre plus général) Vollmayr-Lee *et al.* [110] ont résolu ce type de modèles dans l'ensemble canonique, expliquant en partie l'invariance observée en fonction de α , et démontrant que la statistique de Tsallis ne pouvait en aucun cas s'appliquer à la thermodynamique d'équilibre (canonique au moins) de ces systèmes. Néanmoins, leur méthode ne permet pas d'obtenir des solutions microcanoniques; or on sait que c'est important, puisque des inéquivalences d'ensembles peuvent survenir. Ainsi, Salazar *et al.* ont étudié très récemment [101], par des simulations Monte-Carlo microcanoniques, un système avec interactions en $1/r^\alpha$. Ils retrouvent l'invariance lorsque α varie de 0 à la dimension du système, et le fait que le nombre d'états est bien exponentiel en N . Pour clore le sujet, nous donnons ici une solution microcanonique exacte du modèle d'Ising à une dimension, en appliquant la méthode d'Ellis. Il n'est pas difficile d'étendre les calculs à un modèle sur réseau quelconque, en dimension quelconque (dans [12] est donnée la solution microcanonique de ce modèle, mais l'exactitude du champ moyen n'y est pas démontrée rigoureusement). Nous revenons ensuite sur l'invariance de la solution en fonction de α , lorsque les conditions aux limites sont périodiques. L'étude de ce modèle sera également un moyen de mettre en évidence l'importance (intuitivement claire) des conditions aux limites, lorsque les interactions sont à longue portée.

Le Hamiltonien du modèle d'Ising avec interactions en $1/r^\alpha$ est :

$$H_N = -J \sum_{i>j} \frac{S_i S_j}{|i-j|^\alpha} . \quad (2.56)$$

Chaque spin peut prendre les valeurs ± 1 . Dans l'expression ci-dessus, on a choisi des conditions aux limites libres; on peut aussi choisir des conditions périodiques (spins sur un cercle), en prenant pour $|i-j|$ la plus petite distance le long du cercle entre i et j .

La solution

Étape 1 :

Cette fois-ci, il n'est pas possible d'exprimer exactement H_N à l'aide d'un nombre fini de variables réelles. Il faut choisir une fonction comme variable globale, que nous allons construire par coarse-graining.

On divise le réseau en K boîtes, chacune contenant $n = N/K$ sites, et on décrit le

système par la magnétisation moyenne dans chaque boîte. Dans la limite $N \rightarrow \infty$, $K \rightarrow \infty$, $K/N \rightarrow 0$, le système est donc décrit par une fonction continue $m(x)$; on choisit par commodité $x \in [0, 1]$.

Il faut maintenant vérifier que l'on peut exprimer H_N à l'aide de $m(x)$. Le calcul est facile mais long, et est détaillé à l'appendice A. On obtient le résultat suivant, pour $\alpha < 1$:

$$\begin{aligned} H_N &= -N^{2-\alpha} \frac{J}{2} \int_0^1 dx \int_0^1 dy \frac{m(x)m(y)}{|x-y|^\alpha} + o(N^{2-\alpha}) \\ &= \tilde{H}_N[m(x)] + o(N^{2-\alpha}), \end{aligned} \quad (2.57)$$

l'estimation étant uniforme sur toutes les configurations, comme demandé par les hypothèses des théorèmes de Ellis *et al.* Ce résultat utilise bien sûr de manière essentielle l'hypothèse $\alpha < 1$.

Pour des conditions aux limites périodiques, il faut le modifier légèrement, en remplaçant $|x-y|$ par la plus petite distance, modulo 1, entre x et y .

Étape 2 :

On veut établir un principe de grande déviation pour la variable globale $m(x)$. De façon équivalente, il faut estimer la probabilité d'obtenir un certain profil $m(x)$ à partir d'une configuration microscopique $(S_i)_{i=1\dots N}$, en supposant à priori toutes les configurations microscopiques équiprobables.

Considérons d'abord une seule boîte du coarse-graining; puisque la boîte contient n sites, la magnétisation dans cette boîte m_k vérifie un principe de grande déviation, $P(m_k) \propto e^{ns(m_k)}$, avec

$$s(m) = -\frac{1+m}{2} \log \frac{1+m}{2} - \frac{1-m}{2} \log \frac{1-m}{2} - \log 2. \quad (2.58)$$

Intuitivement, on peut donc écrire la probabilité d'obtenir (m_1, \dots, m_n) sous la forme

$$\begin{aligned} P(m_1, m_2, \dots, m_n) &= P(m_1)P(m_2) \dots P(m_n) \\ &\simeq e^{ns(m_1)} \dots e^{ns(m_n)} \\ &\simeq e^{N \int_0^1 s(m(x)) dx} \\ &\simeq e^{S[m(x)]}. \end{aligned} \quad (2.59)$$

$S[m(x)]$ est alors la fonctionnelle d'entropie associée à la variable globale $m(x)$ que l'on cherchait. Une nouvelle fois, les techniques de la théorie des grandes déviations permettent de justifier rigoureusement ces calculs [46]. Notons que $s(m(x))$ est une quantité d'ordre 1, qu'on intègre entre 0 et 1; la fonctionnelle $S[m]$ est donc proportionnelle à N .

Étape 3 :

On peut maintenant écrire facilement le problème variationnel associé à l'ensemble microcanonique :

$$S(e) = \sup_{m(x)} \left(S[m(x)] \left| -\frac{J}{2} \int_0^1 dx \int_0^1 dy \frac{m(x)m(y)}{|x-y|^\alpha} = e \right. \right) \quad (2.60)$$

Il suffit d'introduire le multiplicateur de Lagrange β associé à la contrainte pour obtenir le problème variationnel canonique. Il s'agit ici d'une optimisation dans un espace fonctionnel : cette dernière étape doit en général être effectuée numériquement. Néanmoins, soulignons une nouvelle fois que la réduction du problème statistique au problème variationnel est exacte (dans la limite $N \rightarrow \infty$, $H \propto N^{2-\alpha}$). Des exemples de profils de magnétisation pour des conditions aux limites libres sont données figure 2.3.

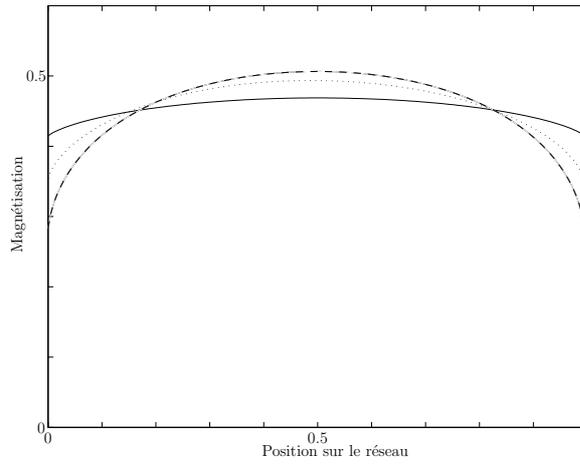


Figure 2.3: Profils de magnétisations pour le modèle d'Ising avec interactions décroissant en $1/r^\alpha$. La densité d'énergie $e = H/N^{2-\alpha}$ est fixée à $e = 0.1$. Les valeurs de α sont $\alpha = 0.2$ (ligne continue), $\alpha = 0.5$ (ligne pointillée), $\alpha = 0.8$ (ligne tiretée). Les conditions aux limites sont libres.

Puisque $S[m] \propto N$, on a prouvé par le calcul que l'entropie habituelle ($S(e)$ est l'entropie de Boltzmann) est bien proportionnelle à N . Il est évident en inspectant la méthode que ceci n'est pas lié au modèle particulier, et s'applique aussi au modèle sur lequel sont faites les simulations de la référence [101].

Les calculs effectués sont rigoureux, mais il est toujours satisfaisant de vérifier sur une simulation numérique qu'on ne s'est pas trompé. Il est aussi intéressant de savoir à quelle vitesse la solution pour N fini converge vers la solution analytique $N \rightarrow \infty$. La figure 2.4 montre l'entropie en fonction de l'énergie pour le modèle d'Ising avec $\alpha = 0.8$, et des conditions aux limites périodiques. La convergence est déjà excellente pour $N \sim 500 - 1000$.

L'invariance en fonction de α

Revenons maintenant sur l'autre point discuté dans la littérature : l'invariance du comportement d'équilibre en fonction de α , pour des conditions aux limites périodiques et pour un rééchelonnement bien choisi de l'énergie. Cette invariance a été observée numériquement [3], puis prouvée analytiquement dans le cadre canonique [31, 110]. On va ici la montrer dans le cadre microcanonique, ce qui apporte une information lorsque les ensembles ne sont pas équivalents.

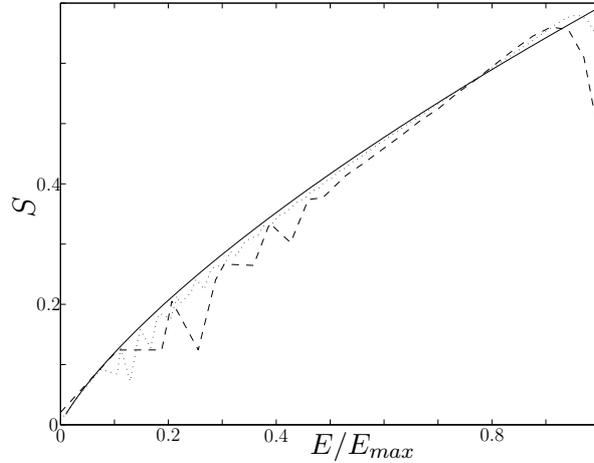


Figure 2.4: Courbe de l'entropie en fonction de l'énergie pour le modèle d'Ising, $\alpha = 0.8$, conditions aux bords périodiques. Les courbes tiretée ($N = 34$) et pointillée ($N = 100$) sont obtenues par une méthode Monte-Carlo microcanonique. La courbe continue est la solution théorique. Pour $N \simeq 1000$ et au delà, les courbes numériques et théorique se superposent parfaitement. Nous remercions chaleureusement R. Salazar pour avoir bien voulu nous fournir les données issues de ses calculs numériques.

On doit résoudre (2.60), avec des conditions aux limites périodiques : on remplace donc $|x - y|^\alpha$ par $d(x, y)^\alpha$, où $d(x, y)$ désigne la plus petite distance entre $x \bmod 1$ et $y \bmod 1$. Par exemple $d(0.2, 0.9) = 0.3$. On pourrait choisir d'autres types de conditions aux limites périodiques, en considérant par exemple les interactions d'un spin avec toutes ses images; nous choisissons celle-ci qui a été couramment utilisée récemment dans la littérature. On écrit l'énergie sous la forme $h = H/\tilde{N} = (m, \mathcal{L}m)$ (produit scalaire dans $L^2([0, 1])$), avec $\tilde{N} = CN^{2-\alpha}$ (C est une constante que nous allons choisir plus tard), et

$$[\mathcal{L}m](x) = \int_0^1 \frac{m(y)}{d(x, y)^\alpha} dy. \quad (2.61)$$

Les extrema de l'équation (2.60) donnant la solution microcanonique du problème vérifient donc, en introduisant un paramètre de Lagrange β , qui doit être choisi pour assurer la conservation de l'énergie :

$$\frac{\delta S}{\delta m(x)} = \beta \frac{\delta h}{\delta m(x)}. \quad (2.62)$$

Cette équation se réécrit

$$\tanh(m(x)) = \frac{\beta J}{C} \int_0^1 \frac{m(y)}{d(x, y)^\alpha} dy. \quad (2.63)$$

Le profil d'équilibre dépend a priori fortement du choix de α (voir les exemples de profils pour des conditions au bord libres représentés figure 2.3). Cependant, pour

des conditions aux limites périodiques, il est facile de voir que le membre de droite de l'équation (2.63) est en réalité indépendant de x , égal à une constante D ; donc les profils $m(x) = m$ homogènes sont solutions de (2.63), pourvu qu'ils satisfassent l'équation

$$\tanh(m) = \beta J m, \quad (2.64)$$

où on a choisi $C = D$. Cette équation est la même quel que soit α , et en particulier c'est celle du cas purement champ moyen $\alpha = 0$. Ceci explique ² l'invariance de la solution en fonction de α . Toutefois, ce résultat dépend essentiellement des conditions aux limites périodiques utilisées, qui ne sont pas très réalistes : la portée physique du résultat paraît donc très limitée. D'un point de vue formel cependant, puisque ce résultat est valable pour tous les systèmes sur réseau (la justification est la même que pour le modèle d'Ising traité ici), il permet d'affirmer que les solutions exactes obtenues à $\alpha = 0$ se généralisent à $\alpha \neq 0$, moyennant une renormalisation de l'énergie : c'est le cas du modèle BEG par exemple. On dispose ainsi d'un "réservoir" de modèles résolus pour toute valeur de α plus petite que la dimension.

2.4 Remarques sur la méthode des grandes déviations

En raison du caractère non additif des systèmes qu'elle étudie, la physique statistique des interactions à longue portée est parfois considérée comme problématique du point de vue conceptuel. Nous essayons de clarifier ce problème de non additivité, qui se manifeste par exemple dans les scalings inhabituels que nous avons dû employer dans les sections précédentes ($H/N^2 = cste$ pour le modèle de Potts, $H/N^{2-\alpha} = cste$ pour le modèle d'Ising...). Par la suite, nous essayons de préciser les limites de la méthode des grandes déviations telle qu'elle est présentée plus haut : elle ne peut s'appliquer à tous les systèmes avec interactions à longue portée. Nous terminons cette section par une petite digression vers l'approche de type "potentiel de Kac".

2.4.1 Le problème du scaling

La théorie des grandes déviations permet d'obtenir des solutions exactes lorsque $N \rightarrow \infty$, dans une certaine limite de scaling : par exemple lorsqu'on garde $H/N^{2-\alpha}$ constant dans le cas du modèle d'Ising avec exposant α . En réalité, on s'intéresse en physique à des systèmes souvent très grands, mais finis. Pour utiliser ces résultats asymptotiques, il faut donc calculer la valeur du paramètre de scaling e pour le système à étudier (par exemple $e = H/N^{2-\alpha}$), obtenir la solution à $N \rightarrow \infty$ pour cette valeur e du paramètre de scaling, et espérer que les effets de taille finie ne changent pas trop le résultat. Se pose alors une question importante : si on choisit un autre paramètre

²Cette démonstration n'est pas vraiment rigoureuse : il faudrait vérifier qu'il n'existe pas de solution non homogène de l'équation (2.63), qui ait une entropie plus élevée que celle de la solution homogène.

de scaling à garder constant lorsque $N \rightarrow \infty$, par exemple $e_0 = H/N$, on pourrait obtenir des résultats différents, et on fera donc une prévision différente pour notre système à N fini. Par exemple, si on fixe pour le modèle d'Ising avec interactions en $1/r^\alpha$, $e_0 = H/N$, on en conclura que pour $N \rightarrow \infty$, le profil de magnétisation $m(x)$ est nul (pour tout e_0). Pourtant, le système à N fini n'a bien sûr qu'une seule physique statistique. Laquelle est la bonne ?

En fait, plus qu'un résultat pour $N \rightarrow \infty$, ce qui est important physiquement c'est d'obtenir un résultat qui dépend peu de N : on pourra alors l'appliquer sans crainte à des systèmes finis. Prenons l'exemple du modèle d'Ising. Une fois le paramètre de scaling $e = H/N^{2-\alpha}$ fixé, N a disparu du problème; il est donc vraisemblable que le résultat final dépendra peu de N , comme espéré. Tout autre choix du scaling ne permettrait pas d'éliminer N du problème, et le résultat obtenu serait beaucoup plus sensible à la valeur de N . Il y a donc un scaling naturel pour le système, et une seule prédiction de la physique statistique; ce scaling est seulement différent de la limite thermodynamique habituelle pour des systèmes avec interactions à courte portée. Pour des systèmes sur réseau de dimension D , ce scaling est donné par $H \sim N^{2-\alpha/D}$. Pour des systèmes hors réseau, il faut être un petit peu plus précis : on prend un volume fixé, avec N particules identiques; on considère une énergie d'interaction à deux corps, proportionnelle au produit des masses des deux corps (ce peut être bien sûr aussi des charges électriques, des charges de vorticité...). On suppose *la masse totale M fixée*; l'énergie potentielle totale est de la forme :

$$E_p = \frac{M^2}{N^2} \sum_{i,j} V(|\vec{r}_i - \vec{r}_j|) . \quad (2.65)$$

On prend alors la limite $N \rightarrow \infty$. Pour un système continu comme un écoulement bidimensionnel, cette limite est tout à fait naturelle. En effet, la mécanique statistique de l'équation d'Euler en dimension deux s'obtient en discrétisant le domaine considéré, et en faisant tendre le pas de discrétisation vers 0 : ainsi, dans un domaine fini fixé, on considère un nombre tendant vers l'infini de cellules dont la taille (donc la vorticité, ou la masse dans la terminologie ci-dessus) tend vers 0. Lorsqu'on approxime l'équation d'Euler par un système discret de points vortex, c'est donc bien aussi cette limite qui est naturelle. Si on s'intéresse à un problème intrinsèquement discret, comme un plasma, ou une galaxie, ce scaling correspond à la limite d'un système comprenant une infinité de constituants, tout en maintenant constant des paramètres comme l'échelle de temps microscopique (la pulsation plasma par exemple : voir [5], et aussi le modèle HMF au paragraphe 2.3.2). Ce scaling spécifique se manifeste parfois par la renormalisation de la constante d'interaction par un facteur dépendant de N (voir le paragraphe 2.3.2 par exemple, ou, pour un système auto-gravitant, par l'introduction d'un paramètre de scaling $e = HR/GM^2$, où H et M sont l'énergie et la masse totales, R le rayon du système et G la constante gravitationnelle) : il s'agit en fait à chaque fois de la même démarche. Un certain nombre de simulations numériques se sont attachées à vérifier la validité de ce scaling, sur des modèles plus compliqués que le modèle d'Ising [23], ainsi que sur les systèmes auto-gravitants [69, 96].

En conclusion, dès que l'on utilise un scaling rendant l'énergie pseudo extensive (par renormalisation de l'interaction, ou en fixant $e = H/N$ comme dans le modèle α -Ising), le champ moyen est justifié (parfois par une démonstration rigoureuse, parfois par un argument plus heuristique). Cependant, le scaling à utiliser pour un problème à longue portée donné, et donc la validité du champ moyen, peut dépendre du régime physique que l'on veut explorer : un exemple est donné au paragraphe suivant.

2.4.2 Limites de la méthode des grandes déviations

Nous avons passé jusqu'ici sous silence un certain nombre de systèmes avec interactions à longue portée pour lesquels la méthode des grandes déviations ne marche pas, et dont l'étude statistique ne peut se réduire à un problème variationnel de type champ moyen. Telle que nous l'avons présentée à la section 2.2, la méthode des grandes déviations repose sur l'introduction de variables globales pour le système, ou, au moins, de variables moyennées localement à gros grains : magnétisation moyenne, énergie cinétique totale... Cette description permet de décrire les structures à grande échelle, dont la taille est du même ordre de grandeur que la taille totale du système. Elle est donc inopérante lorsque les phénomènes intéressants ont lieu à l'échelle microscopique. Cela peut notamment être le cas en présence de forces répulsives à longue portée; l'approche champ moyen, ou grandes déviations, prévoit alors souvent comme équilibre statistique l'absence de structure à grande échelle et les éventuels effets intéressants à petite échelle doivent être étudiés autrement. Un exemple est donné par les modèles d'Ising attractifs à courte portée, et répulsifs à longue portée étudiés par Tarjus *et al.* [105]. La création de structures à grande échelle coûterait une énergie infinie du fait de la répulsion à longue portée, qui joue ici presque le rôle d'une contrainte. La compétition entre répulsion et attraction fait naître des structures intéressantes (alternance de bandes magnétisées + et -, d'une largeur de quelques sites), compatibles avec une magnétisation nulle à grande échelle, et qui ne sont bien sûr pas capturées par la méthode exposée plus haut. Bien que possédant une interaction à longue portée, ce système est en fait additif. C'est un exemple où le scaling "champ moyen" n'est pas pertinent physiquement (voir paragraphe précédent 2.4.1).

Des modèles avec interactions à longue portée (infinie en fait) ont également été introduits dans le cadre de la physique des systèmes désordonnés. Un exemple célèbre est le modèle de Sherrington-Kirkpatrick pour les verres de spins. Pour ces modèles non plus, la méthode des grandes déviations, telle qu'elle est décrite plus haut en tout cas, n'est pas directement applicable. En effet, la recherche d'une variable globale pour décrire ces systèmes est, malgré la portée infinie des interactions, très difficile.

2.4.3 Lien avec les potentiels de Kac

Le fait que des interactions à portée tendant vers l'infini (par rapport à l'échelle microscopique), mais à intensité tendant vers 0, conduisent à une description champ moyen exacte d'un système est connu depuis longtemps. C'est l'idée de base des potentiels de Kac, de la forme [71] :

$$V_\gamma(r) = -\gamma \exp(-\gamma r) \tag{2.66}$$

où γ est un paramètre mesurant la portée et l'intensité de l'interaction, de telle sorte que $\int V_\gamma(r) dr = cste$. Ces potentiels n'ont pas été introduits pour étudier les interactions à longue portée, mais au contraire pour disposer de modèles exactement solubles de systèmes à courte portée. En effet, en prenant d'abord la limite d'une taille infinie du système $L \rightarrow \infty$, puis d'une portée infinie de l'interaction $\gamma \rightarrow 0$, on décrit un système dont les interactions sont à portée infiniment grandes à l'échelle microscopique, mais infiniment courtes par rapport à la taille du système; elles ne sont donc pas à longue portée selon la définition que nous avons adoptée (en particulier, ces systèmes avec potentiels de Kac sont additifs, et ne présentent pas d'inéquivalence d'ensembles). Ces techniques de potentiels de Kac ont fourni par exemple le premier modèle pour lequel une transition de phase liquide-gaz de type Van der Waals a pu être prouvée exactement (Lebowitz et Penrose [84]). On peut penser que des méthodes fondées sur la théorie des grandes déviations permettent de retrouver certains résultats obtenus avec les potentiels de Kac, en effectuant les limites dans l'autre sens : d'abord $L \rightarrow \infty$ avec $\gamma/L = cste$ (pour permettre l'utilisation d'une moyenne à gros grains et les outils des grandes déviations), puis $\gamma/L \rightarrow 0$, pour explorer le même régime (portée des interactions infiniment longue par rapport à l'échelle microscopique, et infiniment courte à l'échelle macroscopique); ces questions sont discutées à l'aide d'autres outils dans la référence [110].

2.5 Revue de résultats

On a vu dans les sections précédentes que la théorie des grandes déviations fournit une méthode puissante, à la fois de démonstration et de calcul, pour un grand nombre de systèmes avec interactions à longue portée. Pour des systèmes plus compliqués, et plus réalistes, que ceux traités jusqu'ici, par exemple des particules hors réseau, appliquer la théorie des grandes déviations est beaucoup moins immédiat. Beaucoup de ces systèmes ne sont traités qu'en introduisant explicitement une approximation de champ moyen, conduisant à des problèmes variationnels du type (2.19) ou (2.23); les résultats exacts connus démontrent l'exactitude de cette approche champ moyen (pour $N \rightarrow \infty$ et un scaling adapté de l'énergie), et on peut conjecturer que dans les autres cas, le champ moyen est soit exact, soit une excellente approximation. Nous essayons dans les paragraphes qui suivent de passer en revue les principaux résultats connus concernant les différents domaines d'application : les systèmes auto-gravitants, la turbulence bidimensionnelle et la physique des plasmas. En particulier, nous soulignons que dans tous les cas, le problème statistique se ramène à un problème variationnel.

2.5.1 Systèmes auto-gravitants

Il est presque immédiat de montrer qu'un système de points matériels auto-gravitants dans un volume donné a un espace des phases accessible infini : en raison de la divergence à courte distance, le système a tendance à se concentrer infiniment tout en augmentant son énergie cinétique (voir par exemple [94]). Il n'y a donc en toute

rigueur aucun équilibre possible, et le problème est hors équilibre. On peut néanmoins écrire des problèmes variationnels champ moyen du type (2.19) ou (2.23) : utilisons comme paramètre d'ordre la fonction de distribution dans l'espace des phases à une particule (6 dimensions) $f(\mathbf{x}, \mathbf{p})$. On peut alors définir une fonctionnelle d'entropie

$$s[f] = - \int d\mathbf{x} \int d\mathbf{p} f \ln f \quad (2.67)$$

et écrire une approximation pour l'énergie (après renormalisations appropriées) :

$$h[f] = \int d\mathbf{x} \int d\mathbf{p} \frac{p^2}{2} f(\mathbf{x}, \mathbf{p}) - g \int d\mathbf{x} d\mathbf{y} \frac{\rho(x)\rho(y)}{|\mathbf{x} - \mathbf{y}|}. \quad (2.68)$$

On est ramené à un problème d'optimisation sur f , mais la réduction n'est cette fois-ci pas rigoureuse. La divergence de la fonction de partition canonique ou microcanonique se manifeste alors par le fait qu'il n'existe pas d'optimum global pour ces problèmes d'optimisation sur f , même s'il peut dans certains cas exister un optimum local, métastable (voir [94] pour une discussion détaillée).

On peut résoudre cette difficulté de plusieurs façons. Une première méthode consiste à considérer la mécanique statistique, non pas du système particulaire, mais de l'équation de Vlasov-Poisson associée. C'est une bonne approximation de la dynamique réelle sur des temps "courts" (voir partie II); de plus, comme cette notion de temps courts dépend du nombre de particules impliquées, l'équation de Vlasov-Poisson peut être pertinente pour décrire la dynamique des galaxies [35, 87]. Dans ce cas, Michel et Robert ont justifié rigoureusement la réduction de la mécanique statistique du système à un problème variationnel, en utilisant les grandes déviations [90, 100].

Une autre méthode consiste à introduire explicitement une régularisation de l'interaction à courte distance. Je ne connais pas alors de démonstration rigoureuse de la validité du champ moyen utilisant la théorie des grandes déviations, ou une autre technique, à moins que la régularisation ne soit introduite explicitement par l'intermédiaire d'un réseau : on peut alors se ramener au cas du modèle d'Ising du paragraphe 2.3.4. Des justifications sans réseau convaincantes fondées sur les intégrales de chemin sont données par Horwitz et Katz [62] et plus récemment par De Vega et Sanchez [43]. Kiessling [77] donne une fonctionnelle d'entropie un petit peu différente à maximiser sous contraintes. Dans tous les cas, le champ moyen réduit à nouveau la mécanique statistique à des problèmes variationnels du type (2.19) et (2.23). Ces problèmes ont été étudiés en détail (voir par exemple [36, 94] pour des revues).

2.5.2 Turbulence bidimensionnelle

Une première approche statistique de la turbulence bidimensionnelle, due à Onsager [93], consiste à discrétiser le flot continu par un nombre fini de points vortex, avec une interaction logarithmique entre eux. Messer et Spohn [89] obtiennent de façon générale la solution canonique de modèles de type "points vortex", sous la forme d'un problème variationnel, mais pour un potentiel non singulier (leurs résultats sont donc pertinents

pour des interactions gravitationnelles ou dans un plasma régularisées). Kiessling [75] étend leur résultats au cas singulier, mais intéressant pour les vortex, de l'interaction logarithmique; Eyink et Spohn [52] traitent le cas de l'ensemble microcanonique, et Kiessling et Lebowitz [78] s'intéressent au problème de l'inéquivalence des ensembles. Pour tous ces travaux, les résultats sont obtenus rigoureusement. Notons qu'il n'y a pas dans ce cas de divergence de la fonction de partition microcanonique : en effet, bien que l'interaction soit également singulière à courte distance, la conservation de l'énergie et l'absence d'énergie cinétique empêchent deux vortex d'être infiniment proches l'un de l'autre.

Au lieu de considérer des points-vortex, il est aussi possible, et cela donne en pratique de meilleurs résultats, de conserver le caractère continu du flot. La dynamique possède alors un nombre infini de quantités conservées (Casimirs), de même que l'équation de Vlasov des systèmes auto-gravitants³. Michel et Robert [90, 100] ont justifié l'approche champ moyen dans ce cas; Ellis et ses collaborateurs ont généralisé la méthode [45, 46] (c'est une version heuristique de cette méthode qui a été expliquée section 2.2).

Les solutions des problèmes variationnels obtenus (du type (2.19) et (2.23)) ont également été étudiées en détail dans un certain nombre de situations : écoulement bidimensionnel incompressible [91, 99], flot quasi-géostrophique [24, 25], modèle Shallow-Water en rotation [37].

2.5.3 Physique des plasmas

L'interaction Coulombienne est évidemment à longue portée, mais la situation pour les plasmas est en général très différente de celle de systèmes gravitationnels, du fait de l'écrantage possible des charges (voir le paragraphe 2.4.2). En fait, comme expliqué en introduction, l'interaction à longue portée pour les plasmas n'est pas toujours directement l'interaction Coulombienne : elle peut apparaître dans une description effective du système qui prend en compte les différentes échelles de temps. Antoni *et al.* calculent explicitement cette description effective, à partir d'un plasma unidimensionnel avec interactions Coulombiennes répulsives [5]. Ils considèrent d'un côté les particules du bulk, qu'ils décrivent collectivement en termes d'ondes de Langmuir, et de l'autre les particules résonantes avec ces ondes de Langmuir, traités de façon individuelle. Ils obtiennent un Hamiltonien effectif couplant M ondes à N particules résonantes. Les ondes elles mêmes fournissent un couplage global, si bien que la description effective est à longue portée et la mécanique statistique d'équilibre conduit aux problèmes variationnels habituels (voir [55] pour le cas à une onde).

De façon similaire, Joyce et Montgomery [70], ainsi que Smith et O'Neil [103], étudient une description effective d'un plasma par des modèles de points vortex, semblables à ceux de la turbulence bidimensionnelle. L'analyse statistique de ces modèles conduit donc à des problèmes variationnels de type (2.19) et (2.23).

Le modèle des oscillateurs de Coulomb, qui contient une répulsion électrostatique et un potentiel quadratique confinant, en deux dimensions, décrit la dynamique de faisceaux

³Les travaux de Chavanis notamment ont mis en évidence et développé cette analogie entre turbulence bidimensionnelle et systèmes stellaires, sous tendue par l'analogie mathématique entre l'équation d'Euler pour la vortacité et l'équation de Vlasov-Poisson [33].

dans les accélérateurs de particules [18]. Les structures d'équilibre microcanoniques et canoniques peuvent être décrites, dans un cadre champ moyen, par des problèmes variationnels de type (2.19) et (2.23), dont il doit être possible de donner une justification rigoureuse.

Citons également l'étude de Kiessling et Neukirch [79] d'un plasma toroïdal auto-confiné, dans lequel l'interaction entre les filaments de courant est à longue portée. Ces auteurs définissent une fonctionnelle d'entropie et la maximisent, concluant à la présence de zones de chaleur spécifique négative.

Ceci termine cette très rapide revue des applications de la mécanique statistique d'équilibre des interactions à longue portée, ainsi que des résultats exacts et approchés obtenus dans les différents domaines. Dans tous les cas, la mécanique statistique d'équilibre peut se réduire à l'étude de problèmes variationnels de type (2.19) et (2.23). Ceci justifie l'étude générale de ces problèmes variationnels dans les parties suivantes.

Chapter 3

Étude générale des problèmes variationnels et classification des transitions de phases

3.1 Étude générale des problèmes variationnels

On a vu que l'étude des états d'équilibre des systèmes avec interactions à longue portée pouvait se réduire, dans un grand nombre de cas, à deux problèmes variationnels, pour l'ensemble microcanonique et pour l'ensemble canonique (l'étude d'autres ensembles statistiques conduirait bien sûr à d'autres problèmes variationnels similaires). Ceci trace le chemin pour une étude générale de la mécanique statistique d'équilibre de ces systèmes, c'est à dire indépendamment du contexte physique, de manière à obtenir des conclusions générales sur les deux ensembles et les liens entre eux. Nous rassemblons d'abord dans cette partie les résultats déjà obtenus dans cette voie. La partie suivante 3.2 sera consacrée à notre propre contribution.

Rappelons pour commodité ici les deux problèmes variationnels, microcanonique :

$$S(e) = \sup_{\mu} (s(\mu) \mid h(\mu) = e) , \quad (3.1)$$

et canonique :

$$F(\beta) = \inf_{\mu} (-s(\mu) + \beta h(\mu)) . \quad (3.2)$$

3.1.1 Inéquivalence d'ensembles

Il semble connu depuis longtemps que l'inéquivalence d'ensembles est directement reliée à la non concavité de la courbe de l'entropie en fonction de l'énergie $S(e)$; les auteurs de [78] utilisent par exemple cela comme définition de l'inéquivalence d'ensembles. Nous adopterons une définition plus intuitive : les deux ensembles sont équivalents si à chaque état d'équilibre microcanonique solution de (3.1), $\mu_m(e)$, correspond un état d'équilibre canonique, $\mu_c(\beta)$, solution de (3.2), et réciproquement. À partir de cette définition, et des deux problèmes variationnels, nous allons montrer :

1. Un état d'équilibre canonique $\mu_c(\beta)$ est toujours un état d'équilibre microcanonique pour une certaine énergie e .
2. Un état d'équilibre microcanonique $\mu_m(e)$ est un état d'équilibre canonique pour un certain β si, et seulement si, la fonction $S(e)$ coïncide au point e avec son enveloppe concave.

La première preuve précise de ce résultat a été donnée par Ellis *et al.* [46]. Nous donnons dans ce qui suit une version simplifiée de la preuve. Au paragraphe suivant, nous compléterons l'étude par l'analyse de la métastabilité.

Démonstration :

On peut tout d'abord remarquer que les problèmes (3.1) et (3.2) sont étroitement liés : le problème canonique est simplement obtenu à partir du problème microcanonique en libérant la contrainte par l'intermédiaire d'un nouveau paramètre extérieur β . Les conditions d'extrémalité au premier ordre sont les mêmes dans les deux cas : $-ds + \beta dh = 0$ (ds et dh représentent respectivement les différentielles des fonctionnelles entropie et énergie). Dans le cas microcanonique, β est un multiplicateur de Lagrange à déterminer en utilisant la contrainte, dans le cas canonique, c'est un paramètre libre. Les deux problèmes ont donc les mêmes extrema, mais on ne sait rien a priori sur le type de ces extrema : maximum, minimum, global, local, point selle...C'est là que réside tout le problème.

Définissons la transformée de Legendre d'une fonction $S(e)$ par

$$S^*(\beta) = \inf_e \{\beta e - S(e)\}. \quad (3.3)$$

Nous utilisons dans la suite un résultat classique sur les transformées de Legendre : la double transformée de Legendre d'une fonction, S^{**} , est égale à l'enveloppe concave de la fonction de départ S [44]. Donc $S^{**} = S$ si et seulement si S est concave.

Prouvons d'abord que F est la transformée de Legendre de S . Pour cela, effectuons la minimisation de (3.2) en deux temps, en se restreignant d'abord à chaque "surface" d'énergie constante : $F(\beta) = \inf_e \inf_\mu \{-s(\mu) + \beta h(\mu) \mid h(\mu) = e\}$. En utilisant la définition de S , on obtient $F(\beta) = \inf_e \{\beta e - S(e)\}$, c'est à dire $F = S^*$.

Montrons maintenant le point 1 : un état d'équilibre canonique μ_c pour un certain β est toujours un état d'équilibre microcanonique pour l'énergie correspondante $e = h(\mu_c)$. Il suffit pour cela de montrer que $S(e) = s(\mu_c)$. Par définition de S , on a $S(e) \geq s(\mu_c)$. D'autre part, comme μ_c est un équilibre canonique associé à l'inverse de la température β , $F(\beta) = \beta e - s(\mu_c)$. En utilisant le fait que F est la transformée de Legendre de S , on a aussi $F(\beta) \leq \beta e - S(e)$. On en déduit donc que $S(e) \leq s(\mu_c)$, et finalement $S(e) = s(\mu_c)$, ce qui termine la démonstration du point 1.

Un équilibre canonique est donc toujours un équilibre microcanonique. La réciproque est fautive, comme l'indique le point 2 : un équilibre microcanonique μ_m pour une certaine énergie e est aussi un équilibre canonique si et seulement si S est égale à son enveloppe concave au point e , c'est à dire $S(e) = S^{**}(e)$.

Montrons d'abord que c'est une condition suffisante, et supposons donc que $S(e) = S^{**}(e)$. Alors $S(e) = \inf_\lambda \{\lambda e - F(\lambda)\}$. On prend β comme étant une valeur de λ

pour laquelle cet infimum est atteint. Alors $S(e) = \beta e - F(\beta)$. Or μ_m est un équilibre microcanonique, donc $S(e) = s(\mu_m)$. Finalement, on obtient $F(\beta) = -s(\mu_m) + \beta h(\mu_m)$, ce qui prouve que μ_m est bien un équilibre canonique.

On suppose maintenant que $S(e) \neq S^{**}(e)$, c'est à dire $S(e) < S^{**}(e)$, puisque S^{**} est toujours l'enveloppe concave de S . Alors pour tout λ , $F(\lambda) < \lambda h(\mu_m) - s(\mu_m)$: μ_m ne peut donc être une solution canonique pour aucune température.

Nous terminons ainsi la démonstration des points 1 et 2; le problème de l'inéquivalence d'ensembles est ainsi ramené à l'étude de la concavité de la fonction S . Néanmoins, il est intéressant de pousser l'analyse plus loin, et de s'intéresser aux états métastables canoniques et microcanoniques et à leurs liens.

3.1.2 États métastables

On appellera état métastable microcanonique (resp. canonique) un maximum local de la fonctionnelle d'entropie sous contrainte (resp. un minimum local de la fonctionnelle d'énergie libre). Ces états métastables jouent un rôle un petit peu marginal dans les systèmes avec interactions à courte portée : par exemple, on ne peut pas descendre très en dessous de 0 degré sans que l'eau gèle, et une petite perturbation suffit à déclencher la solidification. La situation change totalement en présence d'interactions à longue portée. En effet, le passage de l'état métastable à l'état stable nécessite un réarrangement global du système, et non plus seulement un réarrangement local qui pourrait se propager (un germe de glace, ou une bulle de vapeur qui auraient atteint la taille critique). Les états métastables sont donc en réalité extraordinairement stables (leur temps de vie augmente exponentiellement avec N [81]), et ne perdent leur intérêt thermodynamique que lorsqu'ils se déstabilisent (point spinodal), c'est à dire deviennent un point selle au lieu d'un extremum local. Il est donc important du point de vue pratique d'étudier ces états métastables.

Les deux premiers résultats ci-dessous permettent d'étudier la stabilité locale d'une solution (ils forment le théorème de Katz [74]); le troisième, une généralisation de la construction de Maxwell est utile pour déterminer la stabilité globale, et donc le point de la transition thermodynamique réelle :

1. On considère la courbe $\beta(e)$ obtenue par solution de l'équation $-ds + \beta dh = 0$; cette courbe représente les états d'équilibre canoniques et microcanoniques, mais aussi les états métastables et tous les points selles des problèmes variationnels (3.1) et (3.2). À chaque fois que cette courbe passe dans le sens trigonométrique un point avec tangente verticale, l'état microcanonique correspondant perd un mode de stabilité . Il en regagne un lorsque ce passage est dans le sens antitrigonométrique.
2. De même, à chaque fois que la courbe $\beta(e)$ passe dans le sens trigonométrique (resp. antitrigonométrique) un point avec tangente horizontale, l'état canonique correspondant perd (resp. gagne) un mode de stabilité.
3. La construction de Maxwell bien connue dans le cas canonique pour déterminer

le point de transition de phase est valable également dans l'ensemble microcanonique (la transition est alors bien sûr à une énergie donnée au lieu d'une température donnée).

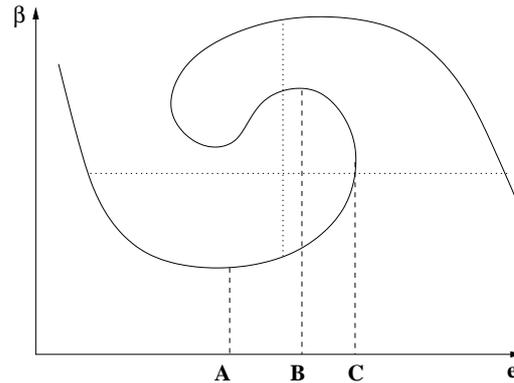


Figure 3.1: Illustration schématique de la méthode de Katz pour déterminer la stabilité locale des solutions canonique et microcanonique. On suppose que la branche à gauche de la figure est stable. Alors l'état représenté par la courbe $\beta(e)$ se déstabilise dans l'ensemble canonique au point A , et dans l'ensemble microcanonique au point C ; il perd un deuxième mode de stabilité canonique au point B . Par la suite, il regagne ce mode de stabilité, puis redevient stable microcanoniquement au passage de la seconde tangente verticale, et enfin redevient stable canoniquement au passage de la dernière tangente horizontale. Les droites pointillées représentent les transitions canonique et microcanonique (voir figure 3.2).

Perdre un mode de stabilité signifie par exemple passer d'une situation d'extremum local à un point selle avec une direction instable. Les deux premiers résultats (c'est à dire le théorème de Katz) permettent donc, au vu de la courbe *complète* $\beta(e)$ (c'est à dire donnée par toutes les solutions de $-ds + \beta ds = 0$) de connaître les points de déstabilisation. Un exemple schématique est donné à la figure 3.1. En application, on peut, sur la figure 1.3 pour le modèle BEG, déterminer les points auxquels les solutions canoniques ou microcanoniques se déstabilisent. On voit aussi là une limite du théorème : en cas de bifurcation (transition du deuxième ordre), on ne peut plus déterminer si facilement la stabilité. Les points 1 et 2 sont montrés dans [74]; nous donnons une démonstration fondée sur la géométrie différentielle, valable en dimension finie, dans [16]. Nous ne répéterons pas ces preuves ici.

Le troisième point en revanche concerne la stabilité globale : il permet de déterminer la température, ou l'énergie, à laquelle deux extrema globaux coexistent et échangent leur stabilité : c'est le véritable point de transition thermodynamique.

Démontrons le point 3, c'est à dire la construction de Maxwell dans le cas d'une transition microcanonique discontinue (soit microcanonique du premier ordre selon notre terminologie). Commençons pour cela par la construction de Maxwell ordinaire, dans l'ensemble canonique. On considère la courbe $\beta(e)$ de la figure 3.2a; la condition

de Maxwell $A_1 = A_2$ d'égalité des aires de part et d'autre de la courbe s'écrit

$$\int_{e_1}^{e_3} [\beta(e) - \beta_c] de = 0 , \quad (3.4)$$

où β_c est l'inverse de la température au point de transition du premier ordre. En utilisant $\beta = ds/de$, on obtient

$$s(e_3) - s(e_1) - \beta_c(e_3 - e_1) = 0 , \quad (3.5)$$

ce qui se réécrit $f(e_1) = f(e_3)$: la construction de Maxwell a bien permis d'identifier le point de transition du premier ordre, où les énergies libres des deux phases sont égales.

Étudions maintenant une transition microcanonique du premier ordre, en supposant que la courbe $\beta(e)$ a l'allure de la figure 3.2b (ce serait le cas pour le modèle BEG, si on avait tracé également les lignes métastables et instables sur la figure 1.3). On inverse la relation $\beta(e)$ au voisinage de e_c pour obtenir $e(\beta)$. La condition d'égalité des aires s'écrit

$$\int_{\beta_1}^{\beta_3} [e(\beta) - e_c] d\beta = 0 . \quad (3.6)$$

en utilisant $e = d(\beta f)/d\beta$, on obtient

$$\beta_3 f(\beta_3) - \beta_1 f(\beta_1) - e_c(\beta_3 - \beta_1) = 0 , \quad (3.7)$$

ce qui implique $s(\beta_3) = s(\beta_1)$: c'est bien la condition d'égalité des entropies des deux phases au point de transition microcanonique du premier ordre. L'équation (3.6) n'est valable que pour la configuration simple de la figure 3.2b; pour des courbes plus complexes, qui s'enroulent plusieurs fois par exemple, la construction est encore valable à condition d'évaluer précisément les différentes aires.

Les résultats connus concernant l'étude générale de ces problèmes variationnels permettent donc d'importantes conclusions, valables indépendamment du modèle particulier considéré :

1. On peut déterminer à priori les points de déstabilisation des solutions canoniques et microcanoniques, en examinant la courbe $\beta(e)$ (théorème de Katz).
2. On peut déterminer les températures ou énergies de transition grâce à la construction de Maxwell.
3. Enfin, on sait que l'inéquivalence entre les ensembles est directement liée à la concavité de la courbe $S(e)$ [46].

On peut aller beaucoup plus loin : dans la partie suivante, toujours en partant des deux problèmes variationnels, nous expliquons comment *classifier* toutes les transitions de phase qui peuvent survenir dans les ensembles canonique et microcanonique, les liens entre elles, et les modes d'apparition de l'inéquivalence d'ensembles.

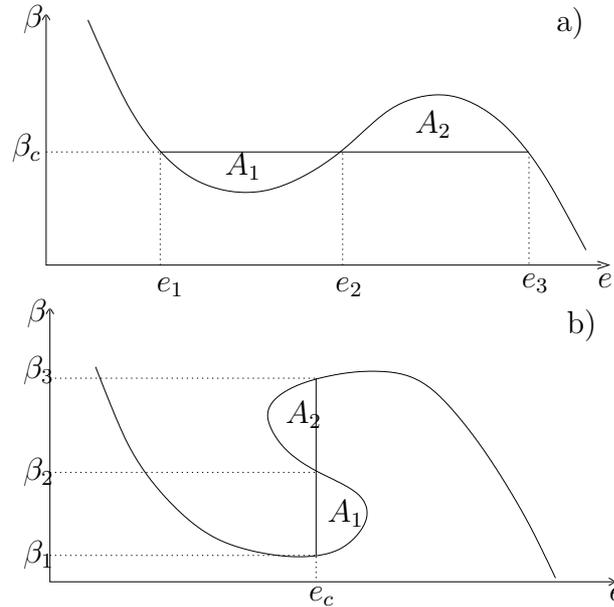


Figure 3.2: Constructions de Maxwell canonique (a) et microcanonique (b). Dans chaque cas, les aires A_1 et A_2 sont égales.

3.2 La classification des transitions de phase et des situations d'inéquivalence d'ensembles

On appellera transition de phase dans l'ensemble microcanonique (resp. canonique) toute brisure d'analyticité de la fonction $S(e)$ (resp. de l'enveloppe concave de $S(e)$). Pour l'ensemble canonique, c'est équivalent à une brisure d'analyticité de la fonction énergie libre : on retrouve bien la définition habituelle. Pour classifier les situations d'inéquivalence d'ensembles, on a vu à la partie précédente qu'il fallait s'intéresser aux propriétés de concavité de $S(e)$; la concavité de $S(e)$ (c'est-à-dire croissance ou décroissance de $\beta(e)$) est par ailleurs directement reliée à la stabilité thermodynamique de la solution (cf. partie précédente également). Notre tâche consiste donc à classifier les brisures d'analyticité de $S(e)$, et ses propriétés de concavité, ainsi que leur évolution lorsqu'on fait varier un ou plusieurs paramètres extérieurs.

Pour cela, nous exploitons les idées générales de la théorie des singularités [9] (ou théorie des catastrophes). Nous expliquons d'abord rapidement au paragraphe suivant 3.2.1 les notions de stabilité d'une situation (une petite modification du modèle induit-elle un changement qualitatif ?), et de codimension (comment la stabilité d'une situation change-t-elle lorsqu'on augmente le nombre de paramètres extérieurs ?). Nous appliquons ensuite ces idées au cas qui nous intéresse, les transitions de phase et les situations d'inéquivalence d'ensembles dans les systèmes à longue portée. Ces résultats font l'objet de l'article [26].

Établir la classification de toutes les transitions de phases possibles nécessitera une étude précise et parfois fastidieuse. Nous utiliserons des résultats mathématiques lorsque ce sera possible, et nous expliquerons les résultats par des dessins. Enfin, nous montrerons que toutes les transitions de phase discutées jusqu'ici dans la littérature

sur les interactions à longue portée sont recensées dans cette classification, et nous prédirons plusieurs situations nouvelles, qui devraient être découvertes au fur et à mesure que de nouveaux modèles seront étudiés.

3.2.1 Notions générales sur les singularités

Introduisons quelques idées générales de la théorie des singularités sur un exemple simple. Prenons une fonction réelle et considérons ses zéros. Dans le cas où aucun zéro n'est double (figure 3.3a), une petite perturbation de la fonction ne change pas qualitativement la situation : le nombre de zéros reste le même, ils sont juste un petit peu déplacés. On dira que cette situation est *générique*. Sur la figure 3.3b en revanche, un zéro est double. Dans ce cas, une petite perturbation de la fonction détruit cette situation, et à la place du zéro double, on trouve soit deux zéros simples, soit aucun zéro. Cette situation de zéro double, instable par petite perturbation des courbes, est dite non générique.

Ajoutons maintenant un paramètre extérieur au problème : on étudie donc les zéros d'une famille de fonctions. Considérons une famille de fonctions dont un représentant a un zéro double (figure 3.3c) : cette fois-ci, cette situation est stable vis à vis d'une petite perturbation de la famille de fonctions. On voit bien que de plus en plus de phénomènes deviennent génériques lorsqu'on augmente le nombre de paramètres extérieurs. Les phénomènes génériques sans paramètres extérieurs (par exemple, un zéro simple pour une fonction) seront appelés *de codimension 0*, avec un paramètre extérieur *de codimension 1*, et ainsi de suite. Ce terme de codimension se réfère au diagramme représentant les différentes situations du système dans l'espace des paramètres de dimension n . En effet, les singularités de codimension 0 sont représentées sur ce diagramme par des ensembles de dimension n , de codimension 1 par des ensembles de dimension $n - 1$, etc. Un exemple est donné à la figure 3.3d. Le but d'une classification est de déterminer tous les phénomènes génériques pour une certaine codimension.

On a vu que le comportement thermodynamique d'un système avec interactions à longue portée est dicté par les propriétés d'analyticité et de convexité de la fonction $S(e)$. On appellera donc dans la suite singularité toute rupture d'analyticité, changement de concavité ou changement de l'enveloppe concave de la fonction $S(e)$. Le but de cette section est la classification de ces singularités. Nous décrivons d'abord au paragraphe 3.2.3 une courbe générique $S(e)$; certaines valeurs spéciales de l'énergie apparaissent : changement de concavité, point de non dérivabilité... L'énergie étant considérée comme un paramètre interne, ces singularités sont génériques en l'absence de paramètre extérieur : on les appelle donc "de codimension 0"; elles sont classifiées au paragraphe 3.2.3b. Les singularités apparaissant en présence d'un paramètre extérieur (de codimension 1) sont classifiées au paragraphe 3.2.3c.

La notion de situation générique dépend, comme on l'a vu, du nombre de paramètres extérieurs du problème. En fait, elle dépend aussi de la nature des fonctions considérées. Par exemple, la situation "avoir un zéro en $x = 0$ " n'est pas générique pour l'ensemble des fonctions continues. En revanche, elle l'est bien sûr pour l'ensemble

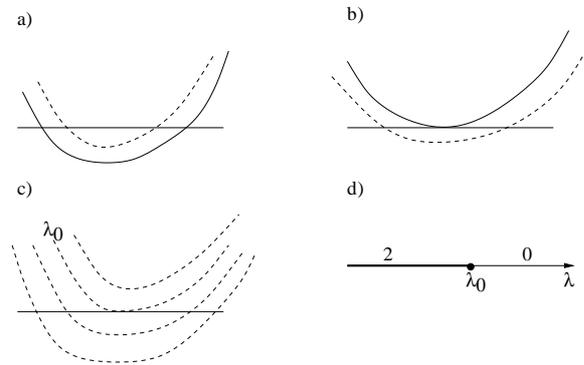


Figure 3.3: En a), une situation générique pour la fonction en trait plein : une petite perturbation (tireté) ne change pas le nombre de zéros. En b), une situation non générique. En c), une situation générique pour une famille de fonctions f_λ ; une famille de fonctions proche de f_λ aura toujours un élément avec un zéro double. En d), l'illustration de la notion de codimension : les situations génériques sans paramètre extérieur (deux zéros ou aucun) sont représentées par des lignes (codimension 0), avec un paramètre extérieur (un zéro double) par un point (codimension 1).

des fonctions impaires. De manière analogue, un modèle physique peut imposer des symétries particulières pour les fonctionnelles entropie $s(\mu)$ et énergie $h(\mu)$; dans ce cas, de nouvelles transitions de phase génériques en codimension 0 et 1 apparaissent. Nous les classifions à la section 3.2.4. Pour établir la classification, nous utilisons dans la suite des résultats mathématiques de la théorie des singularités concernant d'une part le maximum d'une fonction dépendant d'un paramètre, et d'autre part la convexification d'une fonction [1, 20, 29, 114]. Pour pouvoir utiliser ces résultats, nous discutons au paragraphe suivant la possibilité de remplacer le problème microcanonique de maximisation sous contrainte par un problème de maximisation dépendant d'un paramètre. Dans le cas d'une fonctionnelle symétrique, nous ne connaissons pas de résultats mathématiques sur lesquels nous appuyer; nous nous contentons donc d'une représentation géométrique du problème pour obtenir de façon heuristique, mais systématique, la classification. Notons enfin qu'une approximation champ moyen de la thermodynamique habituelle d'un système avec interactions à courte portée conduit bien sûr aussi à un problème d'optimisation, par exemple d'une énergie libre de Landau dépendant d'un paramètre d'ordre. Certaines transitions de phase que nous allons trouver ont donc déjà été étudiées dans ce contexte, et fait l'objet de classification (par exemple pour les mélanges binaires [1]). Cependant, d'une part il ne s'agit pas seulement, pour les systèmes à longue portée, d'une approximation de champ moyen; et d'autre part, lorsque les interactions sont à courte portée, les différents ensembles statistiques sont équivalents, et l'étude est faite en général dans l'ensemble canonique. Nous nous attachons ici à décrire les phénomènes parallèlement dans les deux ensembles, ce qui permet d'étudier les liens entre eux et leurs différences; nous nous intéresserons ainsi particulièrement aux différentes façons dont l'équivalence d'ensembles peut se briser.

3.2.2 Réduction du cas microcanonique à un problème sans contrainte

Nous voulons utiliser des résultats mathématiques valables pour l'étude de problèmes d'optimisation sans contrainte, qui ne peuvent donc pas s'appliquer directement au cas microcanonique. Commençons donc par une remarque technique. Pour s'affranchir de la contrainte $h(\mu) = e$, il faut "inverser" cette relation, c'est à dire écrire une coordonnée de μ en fonction de e et des autres coordonnées de μ : $\mu_1 = \phi(e, \tilde{\mu})$. Dans beaucoup de situations physiques, on peut exprimer l'énergie sous la forme $e = e_c + e_p(\tilde{\mu})$, où e_c est l'énergie cinétique, e_p l'énergie potentielle, et μ est le couple $(e_c, \tilde{\mu})$ (voir par exemple 2.3.2). Alors, en remplaçant e_c par $e - e_p(\tilde{\mu})$, on obtient la fonctionnelle d'entropie à maximiser $\tilde{s}(e, \tilde{\mu}) = s(e - e_p(\tilde{\mu}), \tilde{\mu})$. Il s'agit bien d'un problème de maximisation sans contrainte, paramétré par e , comme attendu.

Dans les autres cas, la simplification est moins simple. Tant que la différentielle $dh(\mu)$ est non nulle, on peut inverser localement la relation $h(\mu) = e$, en utilisant le théorème des fonctions implicites. Ce résultat n'est que local, et ne suffit pas pour conclure : une étude spécifique à chaque cas est nécessaire. Néanmoins, nous avons fait apparaître l'obstacle principal à la procédure de simplification : un point où la différentielle $dh(\mu)$ s'annule. Nous ne connaissons pas d'exemple physique pour lequel ce phénomène (susceptible de créer une discontinuité de l'entropie $S(e)$!) est important. Dans la suite, on supposera toujours que le problème microcanonique peut être réduit à un problème sans contrainte.

3.2.3 Cas d'un problème sans symétrie

a. Les points génériques des courbes $S(e)$

Dans notre problème, les singularités peuvent venir de trois sources différentes : du processus de maximisation, des propriétés de concavité (puisque la concavité de $S(e)$ est importante pour notre problème, on appelle singularités les changements de concavité de S), et du processus de concavification.

On appellera point générique d'une courbe $S(e)$ un point qui n'est singulier pour aucune de ces trois sources de singularité. Déterminons les différents types de points génériques :

- **Singularités de maximisation**

Le maximum de plusieurs fonctions régulières est une fonction continue, mais pas forcément dérivable. Un point générique vis à vis de la maximisation est un point où S est régulière.

- **Propriétés de convexité**

La concavité de S change lorsque sa dérivée seconde s'annule. Un point générique vis à vis des propriétés de concavité est donc un point où d^2S/de^2 n'est pas nul. Ceci définit deux types de points génériques : les concaves et les convexes.

- **La concavification**

Le processus de concavification découpe l'axe de l'énergie en intervalles sur

lesquels soit S coïncide avec son enveloppe concave, soit S est remplacée par un segment de droite. Il y a donc deux types de points génériques vis à vis de la concavification, selon qu'ils appartiennent ou non à l'enveloppe concave de S .

En combinant ces propriétés, on obtient les différents types de points génériques :

- le point concave qui appartient à l'enveloppe concave (intervalles A sur la figure 3.4).
- le point concave qui n'appartient pas à l'enveloppe concave (intervalles B sur la figure 3.4).
- le point convexe, qui n'appartient pas à l'enveloppe concave (intervalle C sur la figure 3.4).

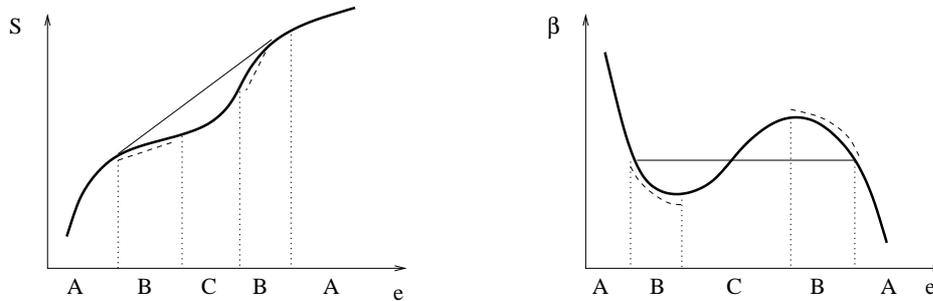


Figure 3.4: Illustration des trois types de points génériques sur une courbe $S(e)$ et la courbe $\beta(e)$ associée. En gras, la solution microcanonique. En trait fin, la solution canonique (elle n'est pas indiquée lorsqu'elle coïncide avec la précédente). En trait tireté fin à côté du trait gras, les zones où la solution microcanonique est métastable canoniquement.

Les points convexes ne peuvent appartenir à l'enveloppe concave, donc nous avons classifié tous les points génériques. Les deux ensembles sont équivalents aux points de type A, et inéquivalents aux points de type B et C. La chaleur spécifique microcanonique est négative aux points de type C. L'axe des énergies se divise en intervalles de type A, B ou C, séparés par des points spéciaux. Ces points spéciaux sont justement les singularités de codimension 0. Au paragraphe suivant, nous classifions systématiquement ces singularités de codimension 0.

b. Les singularités de codimension 0

Comme expliqué au paragraphe précédent, les singularités de codimension 0 sont les points spéciaux le long d'une courbe $S(e)$, rencontrés en faisant varier l'énergie et aucun paramètre extérieur. Elles sont de codimension 0 par rapport à l'une des trois sources de singularités (maximisation, propriétés de convexité et concavification) et génériques par rapport aux deux autres. Nous les classifions donc en employant la même méthode qu'au paragraphe précédent : nous commençons par énumérer les singularités de codimension 0 par rapport à chacune des trois sources, et nous les

combinons ensuite avec les différents types de points génériques A, B, C, pour obtenir toutes les situations possibles.

Les trois sources de singularités

- **Singularités de maximisation et propriétés analytiques**

Les résultats de la théorie des singularités pour un problème d'optimisation dépendant de paramètres [29] nous assurent qu'il n'y a qu'un seul type de singularité de codimension 0. Il correspond à un échange de stabilité/métastabilité entre deux branches différentes de solutions du problème variationnel. En un tel point, S est continue, mais sa dérivée $\beta = dS/de$ présente un saut *positif*. Nous appellerons ces points *transitions microcanoniques du premier ordre*.

- **Propriétés de convexité**

Il n'y a qu'un seul type de singularité de codimension 0 due aux propriétés de convexité : le point d'inflexion, où $d^2S/de^2 = 0$ (donc $d\beta/de = 0$), et la troisième dérivée est non nulle. D'après le critère de Katz (section 3.1), c'est un point où un minimum local de la fonctionnelle énergie libre $f(\mu, \beta) = \beta h(\mu, \beta) - s(\mu)$ devient un point selle; nous l'appellerons donc *perte de stabilité canonique*.

- **Propriétés de concavification**

L'enveloppe concave de S est formée de portions où elle coïncide avec S , séparées par des segments de droite. Le seul type de point spécial, ou singularité de codimension 0, dû à la concavification est donc un point de jonction entre une portion concave et un segment de droite de l'enveloppe concave. Ces points correspondent à une discontinuité de la première dérivée de l'énergie libre $f(\beta)$, et nous les appellerons donc *transitions canoniques du premier ordre*.

Construction des singularités de codimension 0

Il ne nous reste plus qu'à construire toutes les singularités de codimension 0, par combinaison d'une situation de codimension 0 vis à vis d'une des trois sources de singularités avec une situation générique vis à vis des deux autres.

- Considérons d'abord la *transition microcanonique du premier ordre*. Au point de rencontre des deux branches de S , la discontinuité de la dérivée dS/de est positive, et l'angle est rentrant. Ce point n'appartient donc jamais à l'enveloppe concave de S : il est toujours dans la zone d'inéquivalence. Les deux branches qui se croisent peuvent chacune être soit concave, soit convexe : ceci définit quatre types de transition microcanonique du premier ordre, mais les types concave-convexe et convexe-concave sont les mêmes, à une symétrie $e \rightarrow -e$ près. Ils sont représentés à la figure 3.5.
- Le point d'inflexion, ou *perte de stabilité canonique*, joint une zone concave et une zone convexe de S , et ne peut donc appartenir à son enveloppe concave : il est toujours dans la zone d'inéquivalence, et il ne définit qu'un seul type de points (voir la figure 3.5).

- Enfin, la *transition canonique du premier ordre* apparaît toujours dans une zone localement concave de S ($d^2S/de^2 < 0$); il ne définit donc qu'un seul type de point (voir la figure 3.5).

Nous avons donc terminé l'énumération des singularités de codimension 0. En résumé, elles comprennent les transitions microcanoniques du premier ordre (quatre types différents), canoniques du premier ordre (un seul type), et les pertes de stabilité canoniques (un seul type). Une courbe générique $S(e)$ contient ces singularités comme points isolés; on pourrait y ajouter les pertes de stabilité microcanoniques, mais elles ne sont pas visibles sur la courbe $S(e)$, puisqu'elles concernent des états métastables microcanoniques. Parmi ces transitions, seule la transition canonique du premier ordre est visible dans l'ensemble canonique, les autres appartiennent à la zone d'inéquivalence.

Au paragraphe suivant, nous énumérons de la même manière toutes les singularités de codimension 1, susceptibles d'apparaître lorsqu'on peut faire varier un paramètre extérieur. Nous donnerons ensuite des exemples tirés de la littérature, chaque fois que cela sera possible.

c. Les singularités de codimension 1

Nous employons la même méthode pour classifier systématiquement les singularités de codimension 1. Grâce à la théorie des singularités, nous énumérons d'abord les situations de codimension 1 pour chacune des trois sources de singularité, indépendamment des deux autres; puis nous construisons toutes les singularités de codimension 1 par combinaison : une situation de codimension 1 vis à vis d'une source avec une situation générique vis à vis des deux autres, ou une situation de codimension 0 vis à vis de deux sources avec une situation générique vis à vis de la troisième. Pour plus de clarté en ce qui concerne les singularités de maximisation, nous donnons des exemples les plus simples possibles pour l'entropie $S(e, \lambda)$ dans les différents cas (λ est le paramètre extérieur qui peut varier) : ces exemples sont en quelque sorte des formes réduites, génériques au voisinage de la singularité.

Les trois sources de singularités

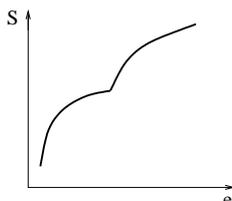
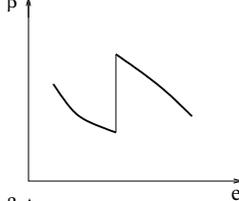
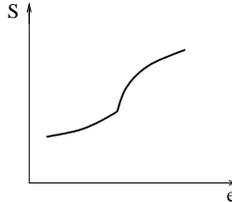
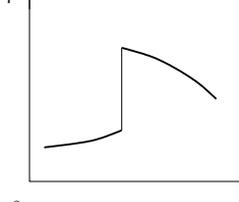
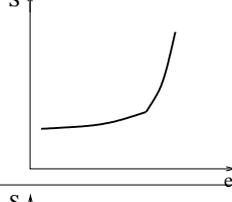
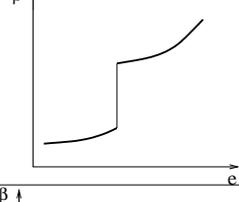
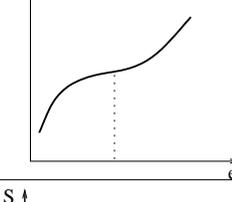
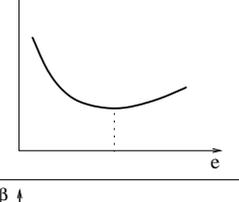
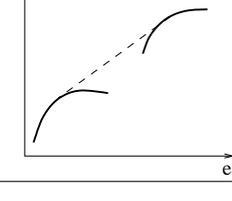
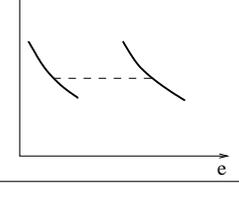
- **Singularités de maximisation et propriétés analytiques**

Toujours en utilisant la référence [29], nous savons que les singularités de maximisation de codimension 1 sont de trois types :

- Le croisement de deux transitions microcanoniques du premier ordre; cette situation survient lorsque trois branches de solutions du problème variationnel ont la même entropie. Nous appellerons cette singularité le *point triple microcanonique*. On peut écrire une forme réduite pour cette situation

$$S(e, \lambda) = \max_m \left(-(m - m_1)^2 + a; -(m - m_2)^2 + b; -(m - m_3)^2 \right), \quad (3.8)$$

où m_1, m_2, m_3 sont fixés, et a et b sont des combinaisons linéaires de e et λ ; la fonctionnelle $s(m, e, \lambda)$ a trois maxima distincts en m_1, m_2, m_3 . Ici

Singularités	Concavité	Concavification	$S(e)$	$\beta(e)$
Singularités de maximisation Transition microcanonique du premier ordre	CC	instable		
	VC	instable		
	VV	instable		
Concavité Déstabilisation canonique		instable		
Concavification Premier ordre canonique		Premier ordre		

C: conCave V: conVexe — : Microcanonique
 ----- : Canonique

Figure 3.5: Résumé de toutes les singularités de codimension 0, avec les différentes configurations de concavité. Toutes les zones convexes (β croissant) sont instables canoniquement.

comme dans la suite, on n'écrit explicitement la forme de s qu'au voisinage des maxima. Le point triple correspond à $a = b = 0$: voir la figure 3.6. On peut remarquer que quelles que soient les orientations de e et λ au voisinage du point critique, la singularité apparaît toujours de la même façon : on passe toujours d'une seule transition pour $\lambda < \lambda_c$ à deux pour $\lambda > \lambda_c$, ou bien de deux à une, mais jamais par exemple trois transitions ne peuvent apparaître ensemble (c'est une propriété bien connue des points triples en général).

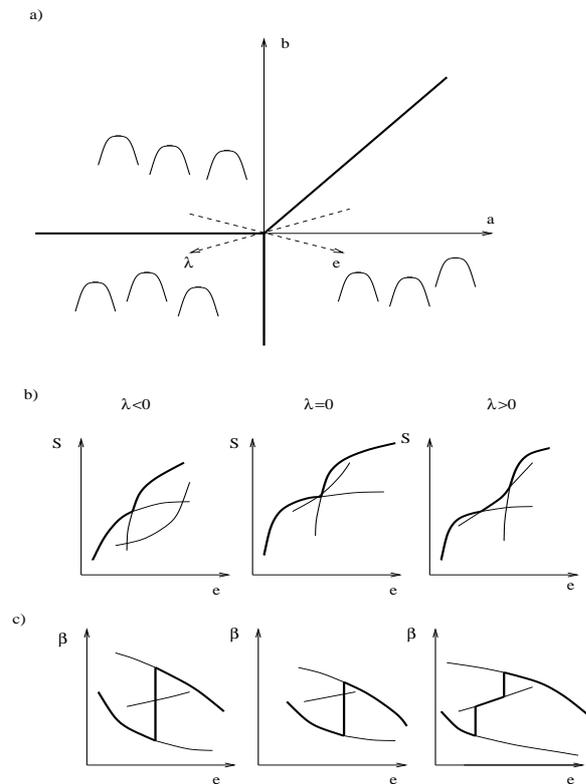


Figure 3.6: Illustration du **point triple microcanonique**, à partir de la forme réduite (3.8). En gras sur la figure du haut, les trois lignes de transitions du premier ordre; l'allure de la fonctionnelle $s(m)$ est schématisée dans les différents domaines. Les flèches en tireté indiquent des directions possibles pour e et λ autour du point critique. Sur les figures b et c sont représentées l'entropie $S(e)$ et la température $\beta(e)$ pour trois valeurs du paramètre extérieur : nous avons pris l'exemple du cas CVC (concave-convexe-concave).

- L'apparition d'une transition microcanonique du premier ordre; en ce point, $d\beta/de = d^2S/de^2 = +\infty$. Nous appellerons ce point *point critique microcanonique*. Une forme réduite est

$$S(e, \lambda) = \max_m (-m^4 + 2am^2 + bm), \quad (3.9)$$

où a et b sont toujours des combinaisons linéaires de e et λ . $b = 0, a > 0$ est une ligne de transition du premier ordre microcanonique, qui se termine au

point critique $a = b = 0$: le schéma de la figure 3.7 permet de comprendre la situation. On peut remarquer que, de même que pour le point triple, la singularité apparaît toujours de la même façon sur les courbes $S(e)$ et $\beta(e)$, quelles que soient les orientations de e et λ au voisinage du point critique.

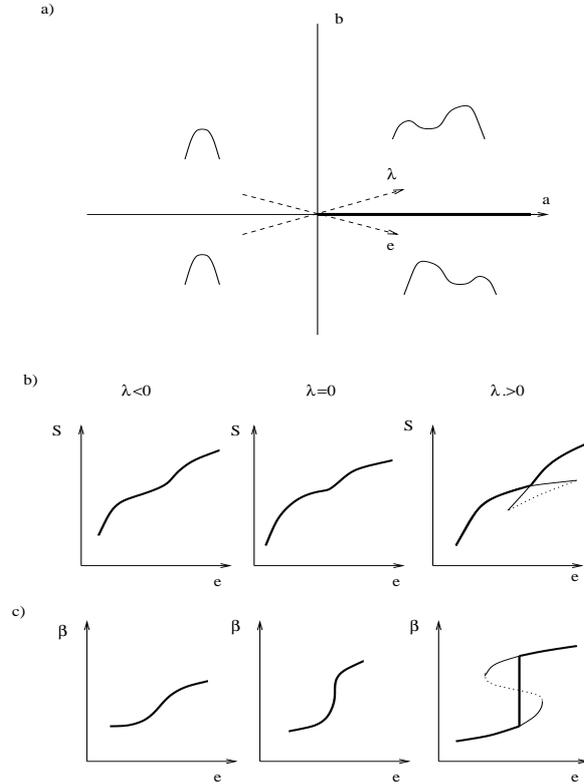


Figure 3.7: Illustration du **point critique microcanonique**, à partir de la forme réduite (3.9). Sur la figure a), la ligne de transition du premier ordre est représentée en gras ($b = 0, a > 0$); elle se termine au point critique $a = b = 0$. Les flèches en trait tireté indiquent des directions possibles pour e et λ autour du point critique. Sur les figure b) et c) sont représentées l'entropie $S(e)$ et la température $\beta(e)$ pour trois valeurs du paramètre extérieur. En trait gras la solution microcanonique, en trait fin les branches métastables microcanoniquement, et en pointillé les branches instables microcanoniquement, qui sont des points selles de la fonctionnelle s .

- L'apparition simultanée de deux transitions microcanoniques du premier ordre; cela correspond à un point où deux branches du problème variationnel sont tangentes l'une avec l'autre. Suivant la terminologie en usage pour les mélanges binaires [1], nous appellerons ce point, *point azéotropique*. Une forme réduite est

$$S(e, \lambda) = \max_m \left(-(m - m_1)^2; -(m - m_2)^2 + (\lambda - e^2) \right). \quad (3.10)$$

Dans le plan, cette singularité se présente en fait comme une ligne de transition du premier ordre habituelle, mais à laquelle une droite $\lambda = cste$ est tangente : voir la figure 3.8.

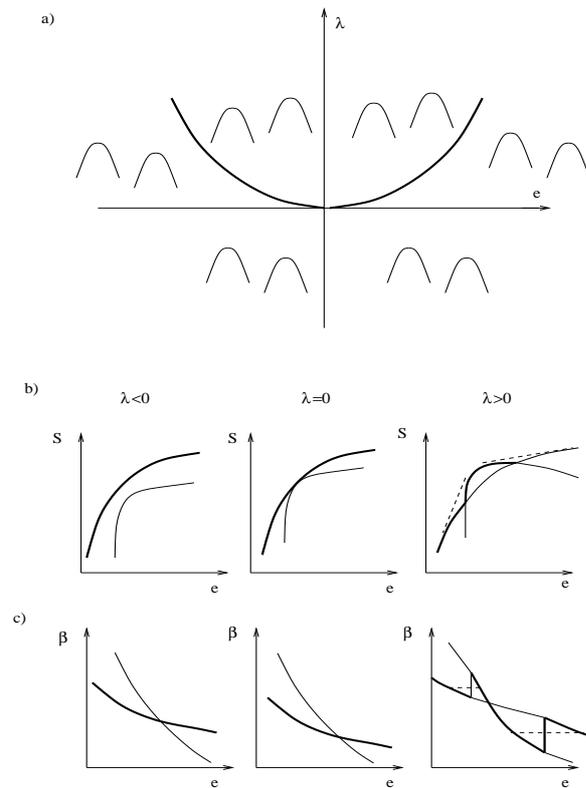


Figure 3.8: Illustration de l'**azéotropie**, à partir de la forme réduite (3.10). Sur la figure a), la courbe $\lambda = 0$ est tangente à la ligne de transitions du premier ordre; l'allure des deux branches de la fonctionnelle d'entropie est schématisée dans les différents domaines. La figure b) représente l'entropie $S(e)$ pour trois valeurs du paramètre λ ; la figure c) représente la courbe $\beta(e)$ pour ces trois mêmes valeurs. En trait tireté, l'enveloppe concave de $S(e)$, c'est-à-dire la solution canonique.

- **Propriétés de convexité**

En codimension 1, nous pouvons nous déplacer le long de la courbe $S(e)$ et faire varier un paramètre extérieur. Cela nous permet de trouver des points où les deuxième et troisième dérivées de $S(e)$ s'annulent. Ceci correspond à un point d'inflexion avec tangente horizontale sur la courbe $\beta(e)$. Nous l'appellerons *changement de concavité*.

- **Propriétés de concavification**

L'enveloppe concave de S est formée d'une succession de régions concaves et de segments de droite; en faisant varier un paramètre extérieur, la structure de cette enveloppe peut changer, par apparition ou disparition de nouveaux segments. Les singularités de codimension 1 associées à la concavification sont précisément ces points où la structure change. Ils sont de deux types :

- L'apparition d'un segment de droite dans une région précédemment concave de l'enveloppe. Ceci impose que d^2S/de^2 et d^3S/de^3 s'annulent ensemble. Il s'agit donc d'un changement de concavité comme décrit plus haut, avec la propriété supplémentaire d'être sur l'enveloppe concave. Nous appellerons ce type de changement de concavité *point critique canonique*.
- La brisure d'un segment de droite de l'enveloppe concave en deux segments de droite séparés par une zone concave. Cette situation apparaît lorsque le graphe de $S(e)$ admet une tangente triple; nous appellerons cela un *point triple canonique*.

Construction des singularités de codimension 1

Nous construisons maintenant par combinaison toutes les singularités de codimension 1.

- Le *point triple microcanonique* met en jeu trois branches de solutions différentes, chacune pouvant être soit concave, soit convexe. Comme la transition du premier ordre microcanonique, le point triple microcanonique est toujours invisible sur l'enveloppe concave de S , dans la zone d'inéquivalence d'ensembles. Il y a donc a priori huit types de tels points, selon la concavité des trois branches de solution, notés CCC, CCV, CVC, CVV, VCC, VCV, VVC, VVV (les 3 lettres symbolisent -C pour conCave, V pour conVexe- la concavité des 3 branches impliquées, par ordre croissant d'énergie; voir la figure 3.6 pour un exemple). En fait, les types CCV et VCC d'une part, CVV et VVC d'autre part, sont symétriques par changement de $e \rightarrow -e$. Il n'y a donc que 6 types différents.
- Le *point critique microcanonique* apparaît nécessairement dans une zone concave de S , invisible sur l'enveloppe concave; il ne définit donc qu'un seul type de point. Voir la figure 3.7.
- Le *point azéotropique* met en jeu deux branches; cependant, comme elles sont tangentes, il n'est pas possible que la branche inférieure soit convexe et la branche supérieure concave. Ceci laisse trois cas possibles, que l'on désigne par CC, VC, VV (C pour conCave, V pour conVexe, la première lettre pour la branche

supérieure, la seconde pour la branche inférieure). Seul le cas CC (concave-concave) est visible dans l'ensemble canonique. Voir la figure 3.8.

- Il y a deux types de *changement de concavité* : l'apparition d'une zone convexe dans une zone concave (notée C), et l'apparition d'une zone concave dans une zone convexe (notée V). Le premier type peut être visible sur l'enveloppe concave, c'est à dire dans l'ensemble canonique. Dans ce cas, il s'agit du *point critique canonique*.
- Le *point triple canonique* correspond à une triple tangente au graphe de $S(e)$; cette triple tangente relie nécessairement trois zones concaves de S , il n'y a donc qu'un seul type de tel point.
- Nous nous intéressons maintenant à la combinaison de deux singularités de codimension 0. Une *perte de stabilité canonique* (point d'inflexion) peut croiser une *transition microcanonique du premier ordre*, avec deux cas différents selon la concavité de la branche qui n'a pas de point d'inflexion. La transition du premier ordre canonique apparaît toujours seule; il n'y a donc pas d'autre combinaison possible de singularités de codimension 0.

La figure 3.9 illustre les cas du croisement d'une *perte de stabilité canonique* avec une *transition microcanonique du premier ordre*, du *point critique canonique*, ainsi que de *l'azéotropie de type VC*.

Dans le cadre des singularités de codimension 1, on peut étudier l'apparition de l'inéquivalence d'ensembles, c'est à dire comment se brise la concavité de la courbe $S(e)$ lorsqu'on fait varier un paramètre extérieur. On s'aperçoit en étudiant la classification que cette concavité ne peut se briser que de deux façons différentes : à un point critique canonique, ou à un point azéotropique de type concave-concave (CC). Nous avons terminé la classification des singularités de codimension 1. Au paragraphe suivant nous cherchons dans la littérature des exemples physiques pour illustrer ces différentes situations, et nous discutons l'intérêt de celles qui n'ont jamais été observées.

d. Illustration sur des exemples physiques

Le système de particules auto-gravitantes est le plus connu et le plus étudié des systèmes avec interactions à longue portée. Nous discutons donc d'abord en détail ce cas, en lien avec la classification.

Discussion du système auto-gravitant

Nous reproduisons les courbes caloriques (température-énergie) d'un système auto-gravitant, à la figure 3.10. La première correspond à des particules ponctuelles [36] (c'est un résultat classique pour les systèmes auto-gravitants). La deuxième illustre le cas de fermions auto-gravitants [36]. Nous donnons les courbes température-énergie car il est plus facile de voir les transitions sur celles-ci que sur les courbes d'entropie $S(e)$.

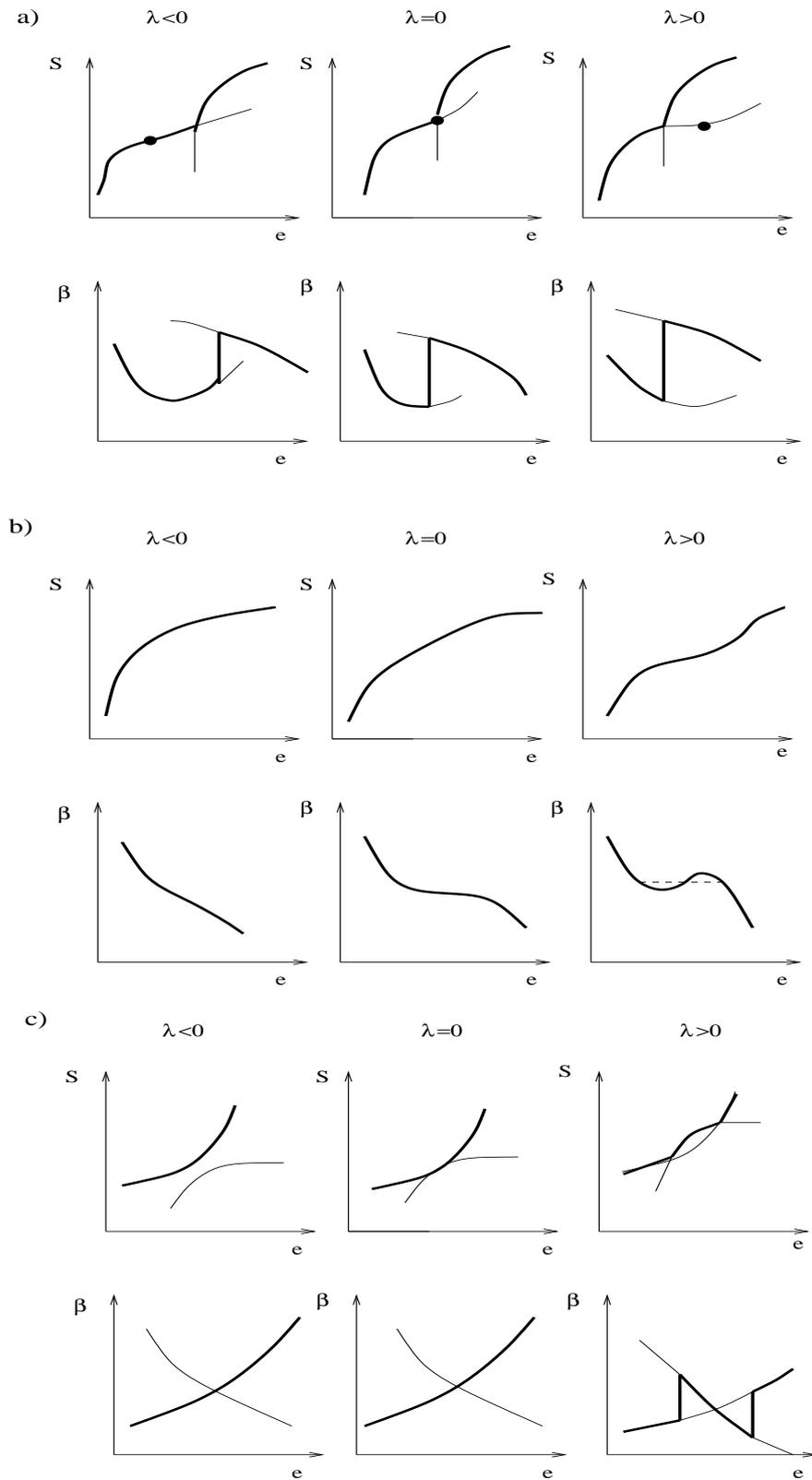


Figure 3.9: a) : croisement d'une **perte de stabilité canonique** (repérée par le point) avec une **transition microcanonique du premier ordre**. b) : un **point critique canonique**; on voit l'apparition d'une zone d'inéquivalence d'ensembles. c) : **azéotropie** de type VC; la figure 3.8 montre une azéotropie de type CC.

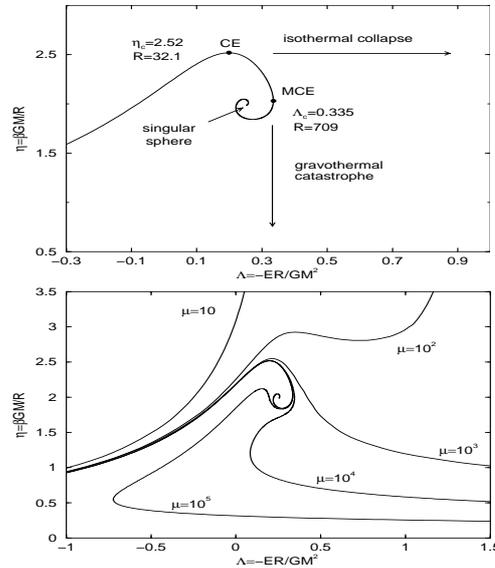


Figure 3.10: Paramètre β en fonction de l'opposé de l'énergie $-E$ pour un système auto-gravitant (les variables sont renormalisées à l'aide de la masse totale M , du rayon du système R , et de la constante de gravitation G). La courbe supérieure est obtenue sans régularisation à courte distance. Les courbes inférieures correspondent au cas de fermions auto-gravitants [36]. Le paramètre $1/\mu$ joue le rôle d'une coupure aux petites distances. Nous remercions P.H. Chavanis pour nous avoir fourni ces figures.

Le cas de particules ponctuelles est particulier : en effet, à cause de la singularité du potentiel en $1/r$ à petite distance, aucun équilibre strict n'est possible, et on ne peut trouver au mieux que des états métastables. Sur la figure 3.10 en haut, on voit que, en diminuant l'énergie, on rencontre d'abord une *perte de stabilité canonique* (point CE; c'est une application du critère de Katz : la courbe calorique passe une tangente horizontale), puis une *perte de stabilité microcanonique* (point MCE; c'est encore une application du critère de Katz). Au delà de l'énergie du point MCE, aucun équilibre métastable n'existe, et le système s'effondre sur lui même (c'est la catastrophe gravotherme).

Considérer des fermions, ou une interaction modifiée à courte distance, permet de régulariser la situation : un état stable existe alors quelle que soit l'énergie. De plus, en modifiant cette distance de coupure, on a accès aux singularités de codimension 1. Sur la figure 3.10 en bas, le paramètre $1/\mu$ joue le rôle d'une distance de coupure; la courbe calorique est représentée pour différentes valeurs de $1/\mu$. Les résultats sont résumés par le diagramme de phase schématisé dans le plan $(-E, 1/\mu)$ de la figure 3.11. On voit que le système présente tous les types de situations de codimension 0 : transitions de phases *microcanonique du premier ordre* (obtenues par construction de Maxwell généralisée), *canonique du premier ordre*, *déstabilisation canonique* (et microcanonique), ainsi que certaines situations de codimension 1 : *point critique microcanonique*, *point critique canonique* (qui marque l'entrée dans la zone d'inéquivalence d'ensembles), *croisement entre une transition microcanonique du premier ordre et une*

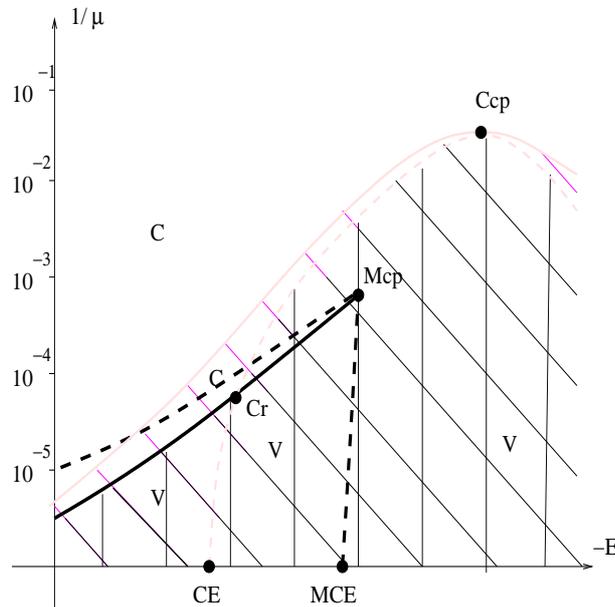


Figure 3.11: Diagramme de phase schématisé pour un système de fermions auto-gravitants. Pour $1/\mu = 0$, on retrouve l'effondrement isotherme (CE) et la catastrophe gravotherme (MCE) du cas sans distance de coupure. La ligne grasse continue est une ligne de transitions microcanoniques du premier ordre (entre une phase gazeuse et une phase concentrée), qui se termine en un **point critique microcanonique** (Mcp). Les lignes grasses grises correspondent à une transition canonique du premier ordre (singularité de concavification); elles se terminent à un **point critique canonique** (Ccp). La ligne grise tiretée est une ligne de déstabilisation canonique; elle croise la transition microcanonique du premier ordre en Cr. La partie hachurée correspond à la zone d'inéquivalence d'ensembles, la partie doublement hachurée à la région de chaleur spécifique négative. Les lettres C et V indiquent la concavité de la branche d'entropie maximale. La ligne grasse tiretée est un ligne de perte de stabilité microcanonique (point du type de MCE); cette transition n'est pas décrite dans notre classification, car elle intervient sur une branche microcanoniquement métastable.

déstabilisation canonique. Lorsque la coupure à petite distance n'est pas imposée par le principe de Pauli, mais par exemple par une modification explicite du potentiel (avec un coeur dur par exemple), la phénoménologie est identique [36].

Autres exemples

Hertel et Thirring [60] ont introduit un modèle exactement soluble inspiré de l'astrophysique pour illustrer les concepts de chaleur spécifique négative et d'inéquivalence d'ensembles. La figure 3.12 montre clairement les 3 types de points génériques, et toutes les singularités de codimension 0 : transition *canonique du premier ordre*, *microcanonique du premier ordre* (entre une branche convexe et une branche concave), et une *déstabilisation canonique*.

D'autres exemples ont été trouvés en dynamique des fluides bidimensionnels. Kiessling

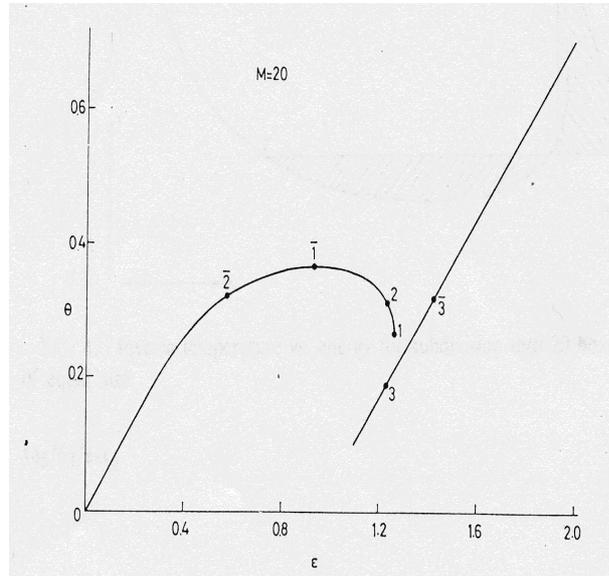


Figure 3.12: Température (θ) en fonction de l'énergie pour le modèle de Hertel et Thirring [60]. On reconnaît une transition canonique du premier ordre (entre les points $\bar{2}$ et $\bar{3}$), une transition microcanonique du premier ordre de type Convexe-Concave (entre les points 2 et 3), et une déstabilisation canonique (au point $\bar{1}$). La branche entre les points $\bar{2}$ et $\bar{1}$ est métastable canoniquement; celle entre les points $\bar{1}$ et 2 est instable canoniquement, à chaleur spécifique négative.

et Lebowitz [78] ont par exemple prouvé, pour le modèle des points vortex, qu'une inéquivalence d'ensembles apparaît pour certaines géométries. Nous reviendrons au modèle des points vortex lorsque nous étudierons les systèmes présentant une symétrie (paragraphe 3.2.4). Ellis *et al.* ont également donné un exemple d'inéquivalence entre les ensembles canonique et microcanonique dans le cadre du modèle Quasi-Géostrophique (géophysique des fluides); on voit sur la figure 2 de la référence [47] une *déstabilisation canonique* et une zone de chaleur spécifique négative. La figure 3 de la même référence, reproduite ici figure 3.13, montre une phénoménologie analogue en faisant varier la circulation (une autre quantité conservée du modèle), au lieu de l'énergie : de fait, nous avons parlé d'ensembles canonique et microcanonique par commodité, mais les résultats généraux sur les problèmes variationnels, ainsi que la classification des transitions, sont valables si l'on considère une autre quantité conservée que l'énergie.

En conclusion, toutes les situations de codimension 0 ont été observées dans la littérature sur différents modèles, ainsi que plusieurs situations de codimension 1. Cependant, un certain nombre d'autres singularités de codimension 1 n'ont jamais, semble-t-il, été trouvées.

Situations non -encore- observées

Parmi les singularités de codimension 1 que nous avons classifiées, beaucoup n'ont

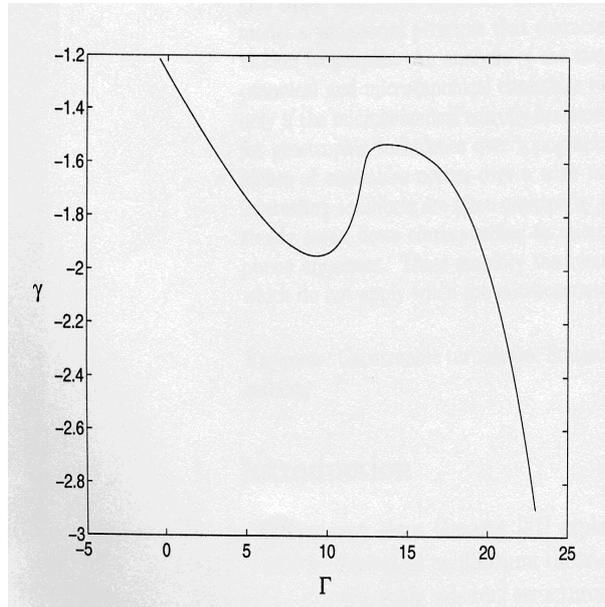


Figure 3.13: Relation entre la circulation totale Γ et le multiplicateur de Lagrange associé γ , pour un modèle de turbulence quasi-géostrophique bidimensionnelle. La circulation totale est une constante du mouvement. La région $d\gamma/d\Gamma > 0$ n'est pas accessible dans l'ensemble γ fixé. Les points tels que $d\gamma/d\Gamma = 0$ sont des déstabilisations canoniques.

jamais été décrites dans la littérature consacrée aux systèmes avec interactions à longue portée : le **point triple microcanonique**, le **point triple canonique**, l'**azéotropie**, le **changement de concavité convexe V**.

Bien sûr, le point triple canonique par exemple est courant dans les études de systèmes à courte portée, et ne présente pas de particularité associée à la longue portée. En revanche, le point triple microcanonique, puisqu'il autorise des concavités quelconques pour les branches impliquées, peut avoir une phénoménologie très particulière.

L'azéotropie également est bien connue, par exemple pour les mélanges binaires, mais elle est associée dans le contexte des interactions à longue portée (dans le cas concave-concave CC) à l'apparition d'une zone d'inéquivalence d'ensembles : il s'agit d'un nouveau mode d'apparition de l'inéquivalence.

Il est probable que certaines de ces nouvelles situations seront observées, au fur et à mesure que des systèmes plus complexes avec interactions à longue portée seront étudiés .

3.2.4 Cas d'un problème avec symétrie

La classification précédente a permis de reproduire avec succès la phénoménologie d'un certain nombre de systèmes, mais pas de tous. En particulier, les différentes situations associées aux transitions du deuxième ordre n'apparaissent pas, alors qu'elles se rencontrent souvent dans l'étude de modèles de physique des plasmas [103, 7] ou de systèmes de spins [15]. La classification se veut pourtant exhaustive. D'où vient le

problème ?

Jusqu'ici, nous avons supposé uniquement une certaine régularité (deux fois dérivables par exemple) pour les fonctionnelles entropie et énergie, et rien d'autre; il n'y avait en particulier aucune référence au système physique considéré. En réalité, il arrive très souvent que les symétries de ce problème physique imposent de restreindre la classe des fonctionnelles admissibles. Dans ce cas, certaines singularités nouvelles peuvent apparaître et créer une phénoménologie beaucoup plus riche, même en se restreignant aux codimensions 0 et 1. Nous donnons un exemple au paragraphe suivant de cette influence de la symétrie, puis nous complétons notre classification, en codimension 0 et 1.

a. Influence de la symétrie sur un exemple simple

Illustrons ce rôle de la symétrie sur un exemple simple. On considère une fonctionnelle $s(m, e)$, avec $m \in \mathbb{R}$, paire en m : $s(m, e) = s(-m, e)$. En $m = 0$, toutes les dérivées par rapport à m d'ordre impair s'annulent : $\partial_m^{(2n+1)} s|_{(0,e)} = 0$.

Un point critique microcanonique est défini comme un point (m_c, e_c) où $\partial_m^{(n)} s|_{(m_c, e_c)} = 0$, pour $n = 1, 2, 3$. Sans hypothèse particulière sur s , cela suppose de résoudre trois équations, avec deux variables m et e . Il faut donc en général ajouter un paramètre extérieur pour observer cette situation, qui est par conséquent de codimension 1. En revanche, si s est paire, trouver un point critique en $m = 0$ ne suppose la résolution que d'une seule équation, avec une variable à disposition, e . Le point critique est donc cette fois de codimension 0 : c'est la transition (microcanonique) du second ordre.

Pour étudier la classification sans symétrie, nous pouvons nous appuyer sur des résultats mathématiques rigoureux [29]; nous ne connaissons pas de tels résultats en présence de symétrie, et nous ferons souvent appel à la représentation des singularités par leur forme réduite pour justifier et illustrer notre propos. Donnons un exemple de forme réduite dans ce cas très simple du point critique (c'est la même chose que le développement habituel de l'énergie libre de Landau, dans le cas microcanonique). On appelle λ le paramètre extérieur. Lorsqu'il n'y a pas de symétrie, on peut prendre comme forme réduite

$$s(m, e, \lambda) = -m^4 + b(e, \lambda)m^2 + a(e, \lambda)m + s_0(e, \lambda), \quad (3.11)$$

où a et b sont simplement des combinaisons linéaires de e et λ ; s_0 est le terme d'ordre m^0 , qui aura son importance plus tard. Ajouter un terme en m^3 n'apporterait rien. L'entropie $S(e, \lambda)$ s'obtient par maximisation :

$$S(e, \lambda) = \max_m s(m, e, \lambda). \quad (3.12)$$

Le point critique correspond à $a = b = 0$; il y a une ligne de transitions microcanoniques du premier ordre en $a = 0, b > 0$.

Lorsqu'il y a une symétrie de parité par rapport à m , une forme réduite est donnée par

$$s(m, e) = -m^4 + b(e)m^2 + s_0(e), \quad (3.13)$$

puisque les dérivées d'ordre impair s'annulent toujours en $m = 0$. Comme on ne s'intéresse qu'au voisinage de la singularité, b peut être pris comme linéaire en e ; la transition du deuxième ordre correspond à l'annulation de b .

Dans la suite, pour identifier et classifier toutes les transitions microcanoniques possibles, pour une codimension donnée, ainsi que leur lien avec l'ensemble canonique, nous allons employer ce critère du "nombre d'équations à satisfaire", et raisonner à l'aide de formes réduites.

La symétrie du système physique n'est pas nécessairement la parité; nous supposons qu'on peut trouver au moins une variable m_1 telle que toutes les dérivées impaires par rapport à m_1 de la fonctionnelle d'entropie $s(m_1, m_2, e, \lambda)$ s'annulent en $m_1 = 0$: $\partial_{m_1}^{(2n+1)} s|_{m_1=0} = 0$. Par exemple, une symétrie de rotation entre dans ce cadre : toutes les dérivées impaires par rapport au rayon r s'annulent au centre. Il faut en général étudier la symétrie cas par cas.

b. Les singularités de codimension 0 avec symétrie

Nous classifions dans ce paragraphe les nouvelles singularités de codimension 0 apparaissant du fait de la symétrie. De manière générale, les singularités concernant les propriétés de concavité (point d'inflexion, point critique canonique...), ainsi que celles concernant la concavification (premier ordre canonique, point triple canonique...) ne sont pas modifiées; en revanche, on a vu que de nouvelles singularités liées au processus de maximisation apparaissent.

En codimension 0, il y a une seule nouvelle singularité : la *transition de phase du deuxième ordre*. Analysons la forme réduite

$$S(e) = \max_m (-m^4 - 2em^2 + s_0(e)). \quad (3.14)$$

Pour $e > 0$, $m = 0$ et $S(e) = s_0(e)$; pour $e < 0$, la branche symétrique est instable, et $S(e) = e^2 + s_0(e)$. De manière générale, nous supposons toujours que la branche symétrique est stable aux grandes énergies (ce n'est ici qu'une convention, mais qui correspond au cas le plus fréquent dans la pratique). On voit alors qu'au point de transition, $S(e)$ et dS/de sont continues, alors que d^2S/de^2 subit *un saut négatif* (la branche non symétrique est toujours plus convexe que la branche symétrique). Selon la taille du saut et la concavité de s_0 au voisinage du point de transition, on voit qu'il y a trois types de transitions du deuxième ordre possibles, représentés à la figure 3.14 :

1. branche bifurquée concave, branche symétrique concave (notée CC : la première lettre correspond à la branche petite énergie, non symétrique),
2. Convexe-Concave (VC),
3. Convexe-Convexe (VV).

Puisque la branche non symétrique est plus convexe que la branche symétrique, l'association Concave-Convexe est impossible. Seule la transition Concave-Concave (CC) est visible dans l'ensemble canonique : c'est la transition canonique du deuxième ordre habituelle.

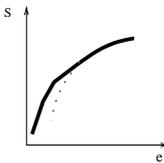
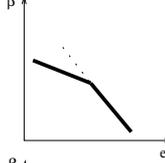
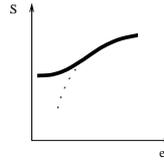
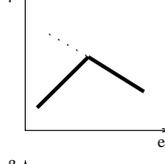
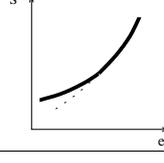
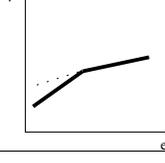
Singularités	Concavité	Concavification	S(e)	$\beta(e)$
Singularités de maximisation Transition microcanonique du second ordre	CC	second ordre		
	VC	instable		
	VV	instable		

Figure 3.14: Nouvelles singularités de codimension 0 pour les systèmes avec symétrie : les trois types de transition du deuxième ordre.

c. Les singularités de codimension 1 avec symétrie

Nous passons aux nouvelles singularités de codimension 1. Elles vont apparaître d’une part du fait des nouvelles singularités de maximisation (à classifier), et d’autre part par la combinaison d’une transition du deuxième ordre avec une autre singularité de codimension 0. Nous commençons par identifier les nouvelles singularités de maximisation (il n’y a pas de nouvelle singularité liée aux propriétés de concavité ou à la concavification).

Nouvelles singularités de maximisation en codimension 1

Comme expliqué plus haut, nous ne pouvons plus nous appuyer sur des résultats mathématiques précis dans ce cas. Nous utiliserons donc pour classifier ces singularités le critère heuristique (mais systématique) du “nombre d’équations à satisfaire”, dont un exemple a été donné plus haut. Cinq nouvelles singularités doivent être considérées :

- On développe la fonctionnelle s jusqu’à l’ordre 6 en m , ce qui fournit la forme réduite

$$S(e, \lambda) = \max_m \left(-m^6 - \frac{3bm^4}{2} - 3am^2 + s_0(e, \lambda) \right), \quad (3.15)$$

où a et b sont des combinaisons linéaires de e et λ , qui tendent vers 0 au voisinage de la singularité. ¹ Dans le plan (e, λ) , cette singularité connecte une ligne de transitions du premier ordre avec une ligne du deuxième ordre : nous l’appellerons *point tricritique microcanonique*, conformément à la terminologie

¹Nous avons mis en évidence le terme d’ordre zéro s_0 ; il n’intervient pas dans le processus de maximisation, et n’aura d’importance qu’à l’étape suivante, pour identifier la concavité des différentes branches (voir le paragraphe 3.2.4). Dans toute la suite de ce paragraphe, on l’oubliera.

habituelle. La figure 3.15 représente la singularité, à partir de la forme réduite (3.15). L'analyse de cette même forme réduite permet de montrer que, pour $\lambda = \lambda_c$, on a, au voisinage du point tricritique et pour la branche de basse énergie $e < e_c$, $\beta(e) \sim -C(e_c - e)^{1/2}$ (toujours avec notre convention : la branche de basse énergie est la branche asymétrique $m \neq 0$). Ceci prouve que la branche de basse énergie est nécessairement convexe autour du point tricritique.

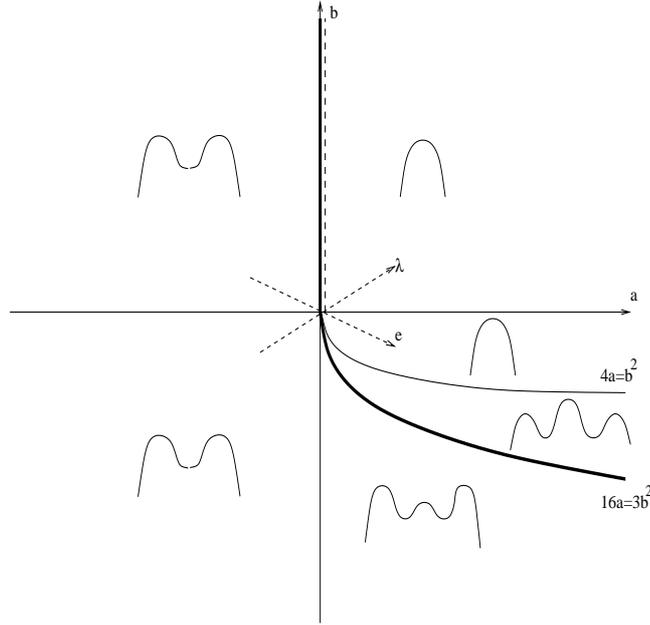


Figure 3.15: Exemple d'un **point tricritique microcanonique** à partir de la forme réduite (3.15). Les courbes schématisées dans les différents domaines montrent l'allure de la courbe $s_{a,b}(m)$. La courbe $4a = b^2$ correspond à la disparition de deux maxima locaux; la courbe (en gras) $16a = 3b^2$, $b < 0$ est la ligne du premier ordre; la courbe en gras et tireté est la ligne du deuxième ordre, avec un saut de concavité négatif en passant du côté tireté. a et b sont des combinaisons linéaires de e et λ ; un exemple de directions possibles pour e et λ est indiqué par les flèches tiretées.

- Nous avons pour le point tricritique utilisé nos deux paramètres (e et λ) pour annuler les dérivées seconde et quatrième de s . On peut aussi annuler uniquement la dérivée seconde, et ajuster la hauteur d'une autre branche de manière à avoir un maximum dégénéré. Une forme réduite est

$$S(e, \lambda) = \max_m (-m^4 - 2am^2; -(m - m_0)^2 + b). \quad (3.16)$$

Dans le plan (e, λ) cette singularité est à l'intersection de deux lignes du premier ordre et d'une ligne du second ordre (les deux lignes du premier ordre sont tangentes) : voir la figure 3.16. Trois phases sont concernées : la phase symétrique $m = 0$, la phase non symétrique $m \simeq 0$, $m \neq 0$, et la phase non symétrique $m = m_0$. On appellera cette singularité un *point triple 112 microcanonique*, pour point triple impliquant deux transitions du premier ordre et une transition

du deuxième ordre². Comme il y a trois branches, la singularité ne sera pas vue

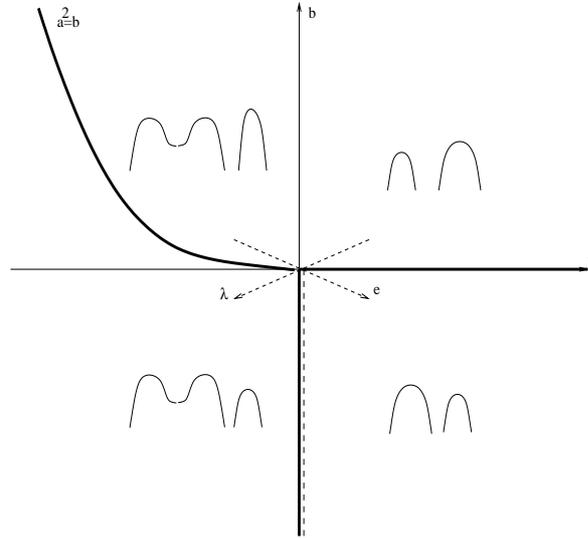


Figure 3.16: Exemple d'un **point triple 112 microcanonique**, à partir de la forme réduite (3.16). Les courbes schématisées montrent l'allure des deux branches de la courbe $s_{a,b}(m)$ dans les différents domaines (à gauche, la branche subissant la transition du deuxième ordre). Les courbes (en gras) $a = b^2$, $a < 0$ et $b = 0$, $a > 0$ sont des lignes du premier ordre. La courbe en gras et tireté est une ligne du deuxième ordre, avec la même convention que plus haut : le saut de concavité est négatif en passant du côté tireté. a et b sont des combinaisons linéaires de e et λ ; un exemple de directions possibles pour e et λ est indiqué par les flèches tiretées.

de la même façon selon la direction de e et λ , c'est à dire selon la façon dont est traversée la singularité. Il y a deux cas possibles. Quand λ traverse sa valeur critique, on peut observer sur la courbe $S(e)$: soit le passage d'une seule transition du premier ordre à deux transitions, une du premier ordre suivie, quand e augmente d'une autre du deuxième ordre (c'est le cas de figure avec les flèches tiretées de la figure 3.16, noté $1 \rightarrow 1 + 2$), soit le passage d'une seule transition du premier ordre à deux transitions, d'abord du deuxième ordre puis du premier quand e augmente (noté $1 \rightarrow 2 + 1$). Une configuration du type $2 \rightarrow 1 + 1$ nécessiterait que l'axe de l'énergie soit tangent aux deux lignes du premier ordre au point triple 112; ce serait donc une singularité de codimension 2.

- Si la fonctionnelle s est paire selon deux directions, on peut, en ajustant e et λ , annuler les dérivées seconde de s selon deux directions. Une forme réduite est donnée par

$$S(e, \lambda) = \max_m (-m_1^4 - m_2^4 - 2am_1^2 - 2bm_2^2 - cm_1^2m_2^2), \quad (3.17)$$

où c est une constante finie, $c > -1$ pour que le maximum existe. Les termes impairs en m_1 ou m_2 ont été enlevés, mais le terme $m_1^2m_2^2$ doit être conservé;

²David Mukamel nous a signalé que la dénomination standard de ce type de singularité dans l'ensemble canonique est "critical end point".

deux cas différents sont à considérer, selon la valeur de c .

Pour $c < 1$, les deux transitions du deuxième ordre se croisent quand λ traverse sa valeur critique λ_c . Pour $\lambda < \lambda_c$ et $\lambda > \lambda_c$, on observe toujours deux transitions du deuxième ordre. Quatre phases sont impliquées : $(m_1 = 0, m_2 = 0)$, $(m_1 \neq 0, m_2 = 0)$, $(m_1 = 0, m_2 \neq 0)$, $(m_1 \neq 0, m_2 \neq 0)$. Nous appellerons donc cette singularité *point quadruple du second ordre microcanonique* (nous ne connaissons pas d'appellation consacrée pour cette singularité, ni d'ailleurs pour celles qui vont suivre).

Pour $c > 1$, la situation change, et la phase doublement asymétrique $(m_1 \neq 0, m_2 \neq 0)$ n'est jamais stable. Seules trois phases sont donc impliquées; la transition entre les deux phases asymétriques, $(m_1 \neq 0, m_2 = 0)$ et $(m_1 = 0, m_2 \neq 0)$, est du premier ordre, les deux autres sont du deuxième ordre : voir la figure 3.17.

Nous appellerons donc cette singularité un *point triple 122 microcanonique*.

À nouveau, ces deux singularités peuvent se manifester de différentes façons selon l'orientation de e et λ par rapport aux lignes de transition : voir la figure 3.17. Nous détaillerons les cas possibles au paragraphe suivant.

- Enfin, la dernière singularité est liée à un comportement particulier des courbes $\lambda = cste$: en codimension 1, il peut exister une telle courbe tangente à une ligne de transition du deuxième ordre : c'est l'analogue de l'azéotropie pour le deuxième ordre. Une forme réduite est donnée par

$$S(e, \lambda) = \max_m (-m^4 + 2(e^2 - \lambda)m^2). \quad (3.18)$$

Cette singularité est représentée dans le plan (λ, e) , à la figure 3.18.

Construction des nouvelles singularités de codimension 1

Nous avons énuméré les cinq nouvelles singularités de maximisation. L'étude des formes réduites a permis de mettre en évidence qu'une même singularité de maximisation peut parfois se manifester de plusieurs façons différentes selon les orientations relatives des lignes de transition et des axes de l'énergie et du paramètre λ . Nous énumérons donc pour chaque singularité de maximisation les différentes situations possibles.

Puis nous classifions toutes les nouvelles singularités de codimension 1 soit en combinant une de ces nouvelles singularités de maximisation avec un point générique vis à vis des propriétés de convexité et de concavification (cela revient en fait à étudier la concavité des différentes branches impliquées), soit en combinant une transition de phase du deuxième ordre avec une autre situation de codimension 0. Cela nous permet de discuter la relation de chacune de ces singularités avec l'ensemble canonique, et de trouver les nouveaux modes d'apparition de l'inéquivalence d'ensembles, ce qui est un des principaux buts de l'étude.

Cette partie de la classification est un petit peu fastidieuse, car les propriétés de concavité imposent de distinguer beaucoup de sous-cas. L'étude de ces propriétés est cependant nécessaire pour déterminer quelles sont les transitions visibles dans l'ensemble canonique, et donc comment apparaît l'inéquivalence.

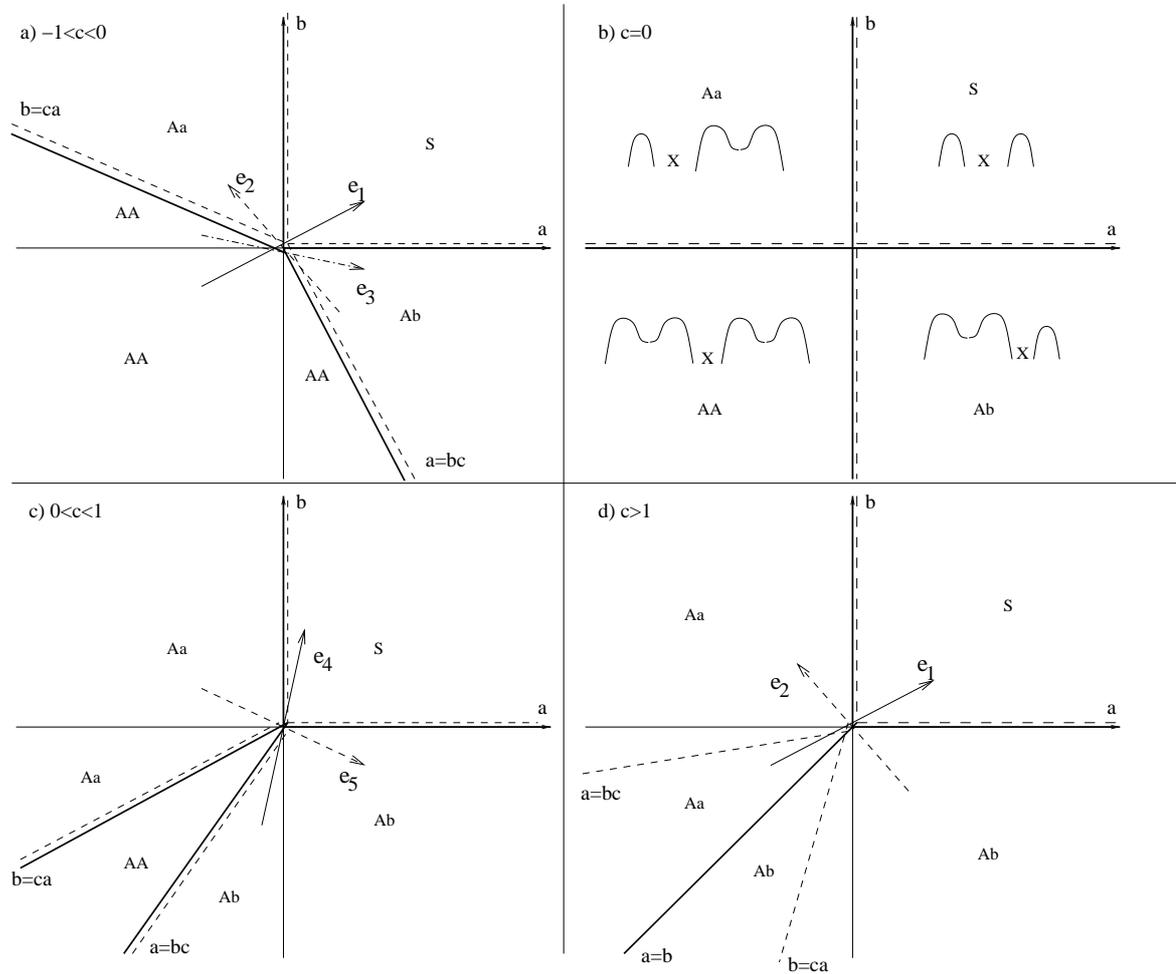


Figure 3.17: Le **point quadruple du second ordre** et le **point triple 122 microcanonique** à partir de la forme réduite $S(e, \lambda) = \max_{m_1, m_2} \{-m_1^4 - 2am_1^2 - m_2^4 - 2bm_2^2 - 2cm_1^2m_2^2\}$, où a et b sont des combinaisons linéaires de e et λ . Quatre phases sont impliquées, notées S (symétrique $m_1 = m_2 = 0$), Aa (antisymétrique $m_1 \neq 0, m_2 = 0$), Ab (antisymétrique $m_1 = 0, m_2 \neq 0$), et AA (antisymétrique $m_1 \neq 0, m_2 \neq 0$). Leur entropie est respectivement $S_S = 0$, $S_{Aa} = a^2$, $S_{Ab} = b^2$ et $S_{AA} = (b^2 + a^2 - 2cab) / (1 - c^2)$. Les figures a), b), c) correspondent au **point quadruple du second ordre microcanonique**, pour respectivement $-1 < c < 0$, $c = 0$, $0 < c < 1$. La figure d), avec $c > 1$, correspond au **point triple 122 microcanonique**. La figure b) montre l'allure de $s_{a,b}(m_1, m_2)$ dans les différentes zones, la courbe de gauche correspondant à la dépendance en m_1 , celle de droite à la dépendance en m_2 . Deux exemples de directions possibles pour e sont indiqués par les flèches continues et tiretées sur les figures a) et d).

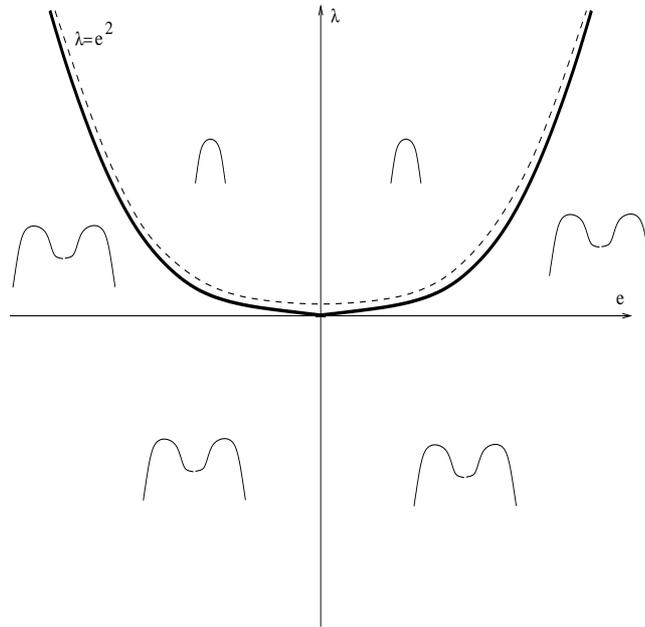


Figure 3.18: **Azéotropie du deuxième ordre**, à partir de la forme réduite (3.18). Les schémas montrent l'allure de $s_{e,\lambda}(m)$ dans les différentes zones. La ligne de transition du deuxième ordre est représentée en gras et tireté; le saut de concavité est négatif en passant du côté tireté.

- La branche basse énergie (asymétrique) d'un *point tricritique microcanonique* est toujours convexe (on l'a vu plus haut par l'étude de la forme réduite). Il y a donc deux types de points tricritiques microcanoniques, selon la concavité de la branche de haute énergie, que l'on appelle VV et VC. Aucun n'est visible dans l'ensemble canonique, à cause de la branche de basse énergie convexe.
- Pour chacun des deux types de *point triple 112 microcanonique* que nous avons identifiés sur la figure 3.16, il y a 6 possibilités, obtenues par combinaison des trois types de transition du deuxième ordre de la figure 3.14, avec les deux concavités possibles de la branche additionnelle. À cause de la transition microcanonique du premier ordre, aucun n'est visible canoniquement. Nous donnons une liste des cas pour le type $1 \rightarrow 2+1$; la première lettre correspond à la branche responsable de la transition du premier ordre, les deux dernières à la transition du deuxième ordre : C-CC, V-CC, C-VC, V-VC, C-VV et V-VV. Le cas du type $1 \rightarrow 1+2$ est analogue.
- Le *point quadruple du second ordre microcanonique* met en jeu quatre phases, notées S, Aa, Ab et AA (voir la figure 3.17). La phase doublement asymétrique AA est toujours la plus convexe, et la phase symétrique S la plus concave; entre les deux, le saut de concavité entre les branches Aa et Ab peut a priori être positif ou négatif. Ceci ne laisse que 6 combinaisons possibles de concavité pour les branches AA, Aa, Ab, S (dans cet ordre) : VVVV, VVVC, VVCC, VCVC, VCCC, CCCC.

Lorsque l'on traverse la singularité, la façon dont se manifeste la transition (c'est à dire la succession des phases), dépend de la direction de e : voir les trois directions e_1, e_2, e_3, e_4 sur les figures 3.17a) et c). Suivant la direction e_1 , la succession des phases est AA-Ab-S d'un côté de la transition, et AA-Aa-S de l'autre; suivant e_2 , la succession est Ab-S-Aa→Ab-AA-Aa; suivant e_3 , la succession est AA-Ab→AA-Aa-S-Ab (ce n'est possible que si le paramètre c de la forme réduite vérifie $-1 < c < 0$) : dans ce cas, trois transitions apparaissent à partir d'une seule; enfin, suivant e_4 , la succession des phases est Ab-S→Ab-AA-Aa-S (ce n'est possible que si le paramètre c vérifie $0 < c < 1$) : dans ce cas également, trois transitions apparaissent à partir d'une seule. Toutes les autres directions possibles pour e et λ sont équivalentes à l'une des précédentes, éventuellement en utilisant les changements de variables $e \rightarrow -e, \lambda \rightarrow -\lambda$, et le fait que les branches Aa et Ab jouent un rôle symétrique. Par exemple, suivant la direction e_5 de la figure 3.17c), la succession des phases est Aa-AA-Ab→Aa-S-Ab, ce qui est en fait équivalent à e_2 , après les changements de variables $e \rightarrow -e, \lambda \rightarrow -\lambda$. Pour chacun de ces quatre cas, nous étudions les concavités possibles des différentes branches.

En utilisant les six configurations de concavité mises en évidence plus haut, nous identifions cinq types de transitions AA-Ab-S→AA-Aa-S : VVV→VVV, VVC→VVC, VCC→VVC, VCC→VCC, CCC→CCC (le cas VVC→VCC est aussi possible, mais équivalent, par symétrie, au cas VCC→VVC). Seul le cas CCC→CCC est visible canoniquement, et il n'est pas associé à l'apparition de l'inéquivalence d'ensembles (les ensembles sont équivalents avant et après). Le point quadruple du second ordre microcanonique AA-Ab-S→AA-Aa-S, CCC→CCC est donc associé à un *point quadruple du second ordre canonique* AA-Ab-S→AA-Aa-S.

De même, il y a cinq types de transitions Ab-AA-Aa→Ab-S-Aa : VVV→VVV, VVV→VCV, CVV→CCV, CVC→CCC, CCC→CCC (le cas VVC→VCC est aussi possible, mais équivalent, par symétrie, au cas CVV→CCV). Seuls les types CCC→CCC et CVC→CCC sont visibles canoniquement, et ce dernier est associé à l'apparition d'une zone d'inéquivalence d'ensembles (ou à sa disparition, lorsque traversé dans ce sens là) : la branche convexe AA est remplacée par une transition du premier ordre dans l'ensemble canonique. Nous concluons que le point quadruple du second ordre microcanonique Ab-AA-Aa→Ab-S-Aa, CCC→CCC est associé à un *point quadruple du second ordre canonique* Ab-AA-Aa→Ab-S-Aa, et que le point quadruple du second ordre microcanonique Ab-AA-Aa→Ab-S-Aa, CVC→CCC est associé à un *point triple 122 canonique* Ab-Aa→Ab-S-Aa.

Considérons le type AA-Ab→AA-Aa-S-Ab; la direction de e (par exemple e_3 sur la figure 3.17a) implique que la branche Aa est plus convexe que la branche Ab (il faut utiliser l'expression de l'entropie pour ces deux branches, respectivement a^2 et b^2). Restent donc possibles cinq configurations de concavité : VV→VVVV, VV→VVCV, VC→VVCC, VC→VCCC, CC→CCCC. Seule la dernière est visible canoniquement; elle n'est pas associée à l'apparition de l'inéquivalence, et correspond à un *point quadruple du second ordre canonique* AA-Ab→AA-Aa-S-

Ab.

Enfin, analysons le type $\text{Ab-S} \rightarrow \text{Ab-AA-Aa-S}$. Cette fois la direction de e (e_4 sur la figure 3.17c) implique que la branche Ab est plus convexe que la branche Aa. Restent cinq configurations : $\text{VV} \rightarrow \text{VVVV}$, $\text{VC} \rightarrow \text{VVVC}$, $\text{VC} \rightarrow \text{VVCC}$, $\text{CC} \rightarrow \text{CVCC}$, $\text{CC} \rightarrow \text{CCCC}$. Le cas $\text{CC} \rightarrow \text{CCCC}$ est visible canoniquement, n'est pas associé à l'apparition de l'inéquivalence, et correspond à un *point quadruple du second ordre canonique* $\text{Ab-S} \rightarrow \text{Ab-AA-Aa-S}$. Le cas $\text{CC} \rightarrow \text{CVCC}$ est également visible canoniquement, et est associé à l'apparition de l'inéquivalence d'ensembles : la branche convexe est remplacée par une transition du premier ordre dans l'ensemble canonique. C'est un *point triple 122 canonique* $\text{Ab-S} \rightarrow \text{Ab-Aa-S}$ (c'est l'équivalent canonique, après symétrie, du point triple 122 microcanonique $\text{Aa-Ab-S} \rightarrow \text{Aa-S}$, discuté ci-dessous).

- Le *point triple 122 microcanonique* est schématisé sur la figure 3.17d). Il met en jeu trois phases Aa, Ab et S; la transition entre Aa et Ab est du premier ordre, les deux autres sont du deuxième ordre. On peut traverser la singularité de deux façons, notées e_1 et e_2 sur la figure.

Suivant e_1 , la succession des phases est Aa-Ab-S puis Aa-S (la transition du premier ordre disparaît). On peut montrer que la branche Aa est plus convexe que la branche Ab, ce qui laisse quatre configurations possibles pour les concavités : $\text{VVV} \rightarrow \text{VV}$, $\text{VVC} \rightarrow \text{VC}$, $\text{VCC} \rightarrow \text{VC}$, $\text{CCC} \rightarrow \text{CC}$. Seule la dernière est visible canoniquement, et est associée à l'apparition d'une zone d'inéquivalence, autour de la transition du premier ordre microcanonique. À partir de la forme réduite, on montre que ce cas correspond à un *point triple 122 canonique* $\text{Aa-Ab-S} \rightarrow \text{Aa-S}$.

Suivant e_2 , la transition du premier ordre est remplacée par deux transitions du deuxième ordre; la succession des phases est : Aa-Ab \rightarrow Aa-S-Ab. Quatre configurations de concavité sont possibles : $\text{VV} \rightarrow \text{VVV}$, $\text{VV} \rightarrow \text{VCV}$, $\text{VC} \rightarrow \text{VCC}$, $\text{CC} \rightarrow \text{CCC}$ ($\text{CV} \rightarrow \text{CCV}$ est possible aussi, mais symétrique de $\text{VC} \rightarrow \text{VCC}$). Seul le cas $\text{CC} \rightarrow \text{CCC}$ est visible canoniquement, et est associé à l'apparition d'une zone d'inéquivalence. C'est un *point triple canonique 122* $\text{Aa-Ab} \rightarrow \text{Aa-S-Ab}$.

- L'*azéotropie du deuxième ordre* se manifeste par l'apparition simultanée de deux transitions du deuxième ordre. Après la singularité, les sauts de la concavité d^2S/de^2 sont négatif puis positif, et exactement opposés, puisque les branches basse et haute énergies sont les mêmes. De plus, ces sauts s'annulent exactement au point de singularité, ce qui montrent qu'au voisinage de ce point les deux branches mises en jeu ont la même concavité. Il n'y a donc que deux configurations de concavité : CC et VV; la première est visible canoniquement, c'est une *azéotropie du deuxième ordre canonique*.
- On passe maintenant à la superposition d'une transition du deuxième ordre avec une autre situation de codimension 0. Elle peut d'abord se superposer avec un point d'inflexion sur une des deux branches.
Si le point d'inflexion est sur la branche basse énergie de la transition du deuxième

ordre (branche asymétrique, plus convexe, selon notre convention), cette transition passe du type CC au type VC. Cette situation est visible canoniquement et correspond à l'apparition d'une transition du premier ordre canonique à partir d'une transition du deuxième ordre. C'est un *point tricritique canonique*, et il est associé à l'apparition d'une zone d'inéquivalence.

Si le point d'inflexion est sur la branche haute énergie, la situation passe de VC à VV. Cette configuration n'est pas visible canoniquement.

- Enfin, une transition du deuxième ordre concave-concave peut croiser un point frontière d'un segment de droite de l'enveloppe concave de $S(e)$. Il n'y a qu'une seule configuration de concavité possible, visible dans l'ensemble canonique : on voit canoniquement le croisement d'une transition du premier ordre avec une transition du deuxième ordre, c'est donc un *point triple 112 canonique*. Il apparaît à la limite d'une zone d'inéquivalence, mais n'est pas associé à l'apparition d'une transition du premier ordre canonique, donc pas à l'apparition d'une zone d'inéquivalence.

Table 3.1: Ce tableau résume les nouvelles singularités de codimension qui apparaissent en raison de la symétrie. Dans la deuxième colonne est indiqué le type de singularité microcanonique; les deux nombres de la troisième colonne précisent respectivement le nombre de façons différentes de traverser la singularité dans le plan (e, λ) , et le nombre total de configurations possibles lorsque la concavité des différentes branches est prise en compte. La quatrième colonne indique la nature et le nombre des singularités précédentes visibles dans l'ensemble canonique. La dernière colonne enfin énumère les configurations qui sont associées à l'apparition de l'inéquivalence d'ensembles.

Type de singularité	Singularité Microcanonique	Nbr	Canonique (Nbr)	Inéquivalence
Maximisation	Tricritique	1/2	Non	Non
	Triple 112	1/12	Non	Non
	Quadruple 2 ^e ordre	4/20	Quadruple 2 ^e ordre (4)	Non
	Triple 122	2/8	Triple 122 (2)	2
	Azéotropie 2 ^e ordre	1/2	Triple 122 (2)	2
Convexité	2 ^e ordre / Inflexion	1/2	Azéotropie 2 ^e ordre (1)	Non
Convexité	2 ^e ordre / Inflexion	1/2	Tricritique (1)	1
Concavification	2 ^e ordre / Concavification	1/1	Tricritique (1)	1
Concavification	2 ^e ordre / Concavification	1/1	Triple 112 (1)	Non

Nous avons fini l'énumération de toutes les situations de codimension 1 possibles, dans le cas d'un système présentant une symétrie (de type parité). Nous avons trouvé cinq nouvelles singularités de maximisation (point tricritique, point triple 112, point triple 122, point quadruple du deuxième ordre, azéotropie du deuxième ordre), et nous obtenons, en prenant en compte les différentes possibilités de traverser ces singularités, neuf nouvelles transitions de phases microcanoniques. Ces transitions de phase peuvent être combinées à de nombreuses configurations de convexité différentes. Il y a

également deux nouvelles singularité formées par superposition de deux situations de codimension 0 : une transition du deuxième ordre avec un point d'inflexion, et une transition du deuxième ordre avec la limite d'une zone de concavification. Tout cela est résumé dans le tableau 3.1. Notons que la plupart des singularités, comme déjà dans le cas sans symétrie, se trouvent dans la zone invisible canoniquement; néanmoins, chacune des transitions microcanoniques a un équivalent canonique (ce qui est logique : en effet, si nous avons construit dans cette classification la solution canonique comme concavification de la solution microcanonique, elle est également donnée par un problème d'optimisation, formellement identique au problème microcanonique; la classification est donc cohérente).

Nous avons aussi identifié deux nouvelles voies vers l'inéquivalence d'ensembles : le point tricritique canonique, et le point triple 122 canonique (qui peut être associé dans l'ensemble microcanonique à un point triple 122 ou à un point quadruple du deuxième ordre).

Comme nous l'avons fait plus haut pour le cas non symétrique (paragraphe 3.2.3), nous passons en revue au paragraphe suivant les différentes configurations déjà observées sur des modèles particuliers dans la littérature, et nous mettons en évidence celles qui ne l'ont pas encore été.

d. Illustrations de la classification sur des exemples

Nous discutons d'abord en détail le cas du modèle Blume-Emery-Griffiths champ moyen, introduit au chapitre 1.

Le modèle Blume-Emery-Griffiths

L'énergie et l'entropie de ce modèle sont bien sûr paires par rapport à m , la magnétisation. On doit donc s'attendre à observer certaines des nouvelles singularités dues à la symétrie. En faisant varier le rapport des constantes de couplage Δ/J , on peut avoir accès aux singularités de codimension 1.

Reprenons la représentation des courbes caloriques pour ce modèle (voir la figure 1.3). Sur la figure a), la branche basse énergie de la transition du deuxième ordre a une tangente horizontale; c'est un exemple de *point tricritique canonique*. Il correspond à l'apparition d'une zone d'inéquivalence d'ensembles, comme déjà noté. Sur la figure b), on voit une transition du deuxième ordre avec une branche basse énergie convexe. Puis, sur la figure c), la tangente de la branche basse énergie devient verticale : c'est un *point tricritique microcanonique*, associé à l'apparition d'une transition microcanonique du premier ordre (figure d) : cette transition est de type VC, figure e) : elle est de type CC). Ces résultats sont résumés sur le diagramme de phase schématique de la figure 3.19. Le modèle illustre donc certaines des nouvelles situations de codimension 0 et 1 dues à la symétrie du système.

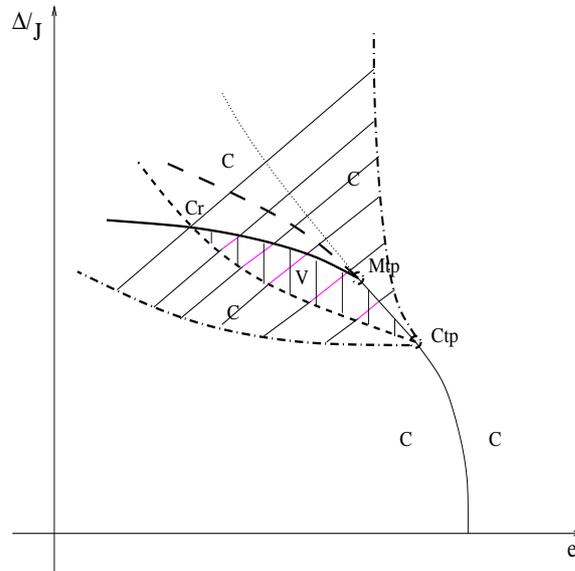


Figure 3.19: Diagramme de phase schématique pour le modèle de Blume-Emery-Griffiths champ moyen. Pour $\Delta/J \rightarrow 0$, le modèle se réduit à un modèle d'Ising, et présente une transition de phase du deuxième ordre (ligne continue fine). Au **point tricritique canonique** (Ctp) naît une zone d'inéquivalence, et la transition du deuxième ordre devient invisible dans l'ensemble canonique. À partir de ce point apparaissent également une ligne de transitions du premier ordre canoniques (ligne grasse et mi-tirée mi-pointillée) et une ligne de déstabilisations canoniques (ligne grasse et tirée court). La ligne du deuxième ordre se termine au **point tricritique microcanonique** (Mtp), où elle rencontre une ligne de transitions du premier ordre microcanoniques (ligne grasse et continue) et une ligne de déstabilisations microcanoniques (ligne grasse et tirée long). La ligne pointillée est la continuation de la ligne du deuxième ordre, elle est invisible dans les deux ensembles. La zone hachurée est la région d'inéquivalence d'ensembles, celle doublement hachurée est la région de chaleur spécifique négative. Cr est un **croisement entre une transition microcanonique du premier ordre et une déstabilisation canonique**.

Autres exemples

Les seules nouvelles singularités de codimension 0 sont les transitions du deuxième ordre. Celles du type CC sont très courantes (voir par exemple le modèle BEG, au paragraphe précédent, l'interaction d'ondes de Langmuir et de particules résonantes dans les plasmas [55], le modèle HMF vu plus haut au paragraphe 2.3.2...). La configuration VC a été observée au paragraphe précédent pour le modèle BEG, dans un modèle de point vortex en physique des plasmas [103] (figure 3.20), ainsi que pour un plasma toroïdal autoconfiné [79] (figure 3.21). À notre connaissance, la configuration VV n'a jamais été observée.

Passons maintenant aux singularités de codimension 1, et à un de nos objectifs principaux : l'identification de tous les modes d'apparition de l'inéquivalence d'ensembles. Il y a deux nouveaux modes pour les systèmes avec symétrie. Le premier est le *point*

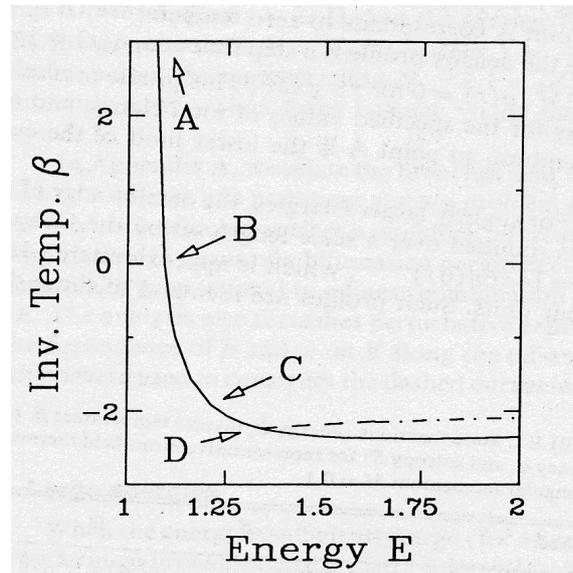


Figure 3.20: β en fonction de l'énergie pour un gaz bidimensionnel de points vortex confinés dans un cercle [103]. Au point D survient une transition de phase du deuxième ordre; la branche tiretée (qui brise la symétrie de rotation) est stable, et convexe. Il s'agit donc d'une transition de type VC (notons que, contrairement à nos conventions, la branche bifurquée non symétrique est stable à haute énergie dans ce cas).

tricritique canonique, au passage duquel une transition du deuxième ordre CC devient VC; il a été observé au moins deux fois, dans un modèle jouet pour les systèmes auto-gravitants (le HMF en dimension 2 [7]), et le modèle BEG (voir paragraphe précédent). Ce mode d'apparition de l'inéquivalence d'ensembles par point tricritique canonique a d'ailleurs été étudié en détail par Leyvraz *et al.* [86]. En revanche, l'apparition de l'inéquivalence autour d'un *point triple 122 canonique* n'a jamais été observée.

En ce qui concerne les autres singularités de codimension 1, très peu ont à notre connaissance déjà été trouvées. On peut citer la référence [7], où un *point triple canonique 112* est décrit. Nous nous attendons bien sûr à ce que l'étude de nouveaux modèles fasse apparaître certaines des situations manquantes.

3.2.5 Conclusion

En s'appuyant sur le fait que l'approche champ moyen est, pour les systèmes non additifs, soit exacte, soit probablement une excellente approximation, nous avons dressé dans cette partie une classification de toutes les transitions de phase pouvant survenir pour ces systèmes, avec ou sans symétrie, en ne faisant varier que l'énergie (codimension 0), ou également un paramètre extérieur (codimension 1). Nous connaissons donc tous les composants élémentaires qui constituent les diagrammes de phase des systèmes avec une contrainte (l'énergie en général) et un paramètre extérieur.

Cette classification donne un cadre unique pour comprendre tous les phénomènes thermodynamiques inhabituels des systèmes non additifs, en astrophysique, turbulence bidimensionnel et géophysique, physique des plasmas... Tous les diagrammes de

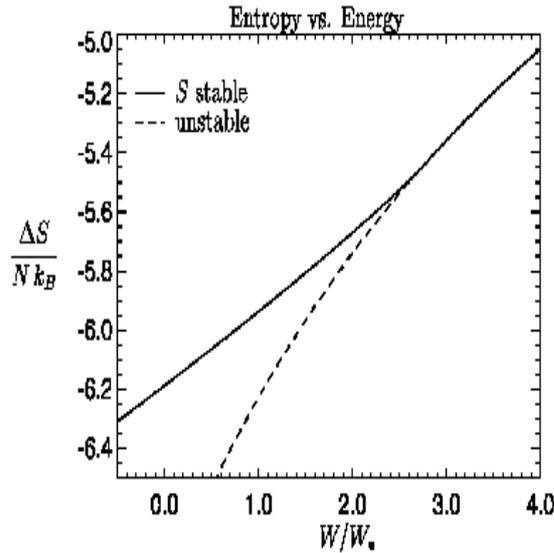


Figure 3.21: Courbe de l'entropie e fonction de l'énergie pour un plasma toroïdal magnétiquement auto-confiné [79]. La branche bifurquée à basse énergie est convexe, il s'agit donc encore d'une transition microcanonique du deuxième ordre de type VC.

phase obtenus jusqu'ici dans la littérature (numériquement ou analytiquement) sont représentés dans la classification. En revanche, certaines singularités que nous avons classifiées n'ont jamais été observées sur un modèle précis. C'est le cas notamment de l'azéotropie et du point triple 122 canonique, qui correspondent à deux nouveaux modes d'apparition de l'inéquivalence d'ensembles. D'autres singularités microcanoniques, qui se caractérisent par des propriétés de concavité parfois exotiques, sont aussi dans ce cas.

Ce travail peut se poursuivre logiquement dans plusieurs voies : en recherchant les singularités de codimension 2, pour décrire les diagrammes de phase des systèmes avec deux paramètres; ou, ce qui est peut être plus intéressant, en étudiant le cas de deux quantités conservées (qui peuvent être l'énergie et la circulation en dynamique des fluides, l'énergie et le moment cinétique total en astrophysique...).

Conclusion

Nous terminons ici cette première partie consacrée à la mécanique statistique d'équilibre des systèmes à longue portée. En nous appuyant notamment sur les travaux d'Ellis et ses collaborateurs, nous avons introduit au début de cette partie la méthode des grandes déviations, très adaptée aux interactions à longue portée, et assez peu connue, semble-t-il, hors du cercle des physiciens mathématiciens. Revenons d'abord sur ce que nous a appris cette méthode.

La méthode des grandes déviations, résumé

- Elle fournit une méthode de calcul puissante pour déterminer explicitement les états d'équilibre d'un certain nombre de systèmes, dans l'ensemble canonique ou microcanonique. Nous avons ainsi pu voir à la section 2.3 que beaucoup de modèles jouets étudiés auparavant numériquement, ou uniquement de façon canonique, étaient exactement solubles; nous avons pu également donner des exemples mettant en évidence la phénoménologie particulière de ces systèmes à longue portée (le modèle BEG notamment, même si sa solution ne nécessite pas l'emploi de techniques de grandes déviations).
- D'autre part, la méthode des grandes déviations fournit également des conclusions intéressantes du point de vue formel : elle réduit l'étude statistique des systèmes traités à des problèmes d'optimisation. C'est la base pour une étude générale de la phénoménologie de l'équilibre des systèmes avec interactions à longue portée, développée aux sections 3.1 et 3.2.
- Enfin, si les systèmes auto-gravitants et la turbulence bidimensionnelle sont étudiés depuis longtemps par des méthodes statistiques, il n'en est pas de même pour les interactions ondes-particules par exemple; la méthode des grandes déviations, peu connue en physique des plasmas, ou en physique des lasers, pourrait trouver là un terrain d'application totalement nouveau. Nous avons étudié un exemple au chapitre 6.

À première vue, on pouvait s'attendre à ce que la présence d'interactions à longue portée complique techniquement les calculs statistiques. On s'aperçoit au contraire que la longue portée, en rendant le champ moyen exact dans une certaine limite de scaling, simplifie beaucoup la théorie et permet une approche générale poussée. Répétons ici que ces remarques ne sont pas valables pour tous les systèmes à longue portée, et notamment pour certains systèmes en interaction Coulombienne, ou systèmes désordonnés : la méthode des grandes déviations, puisqu'elle introduit une variable globale,

ou moyennée à gros grains, ne peut décrire que les structures à grande échelle (de même ordre de grandeur que le système entier); dans sa forme simple présentée plus haut, elle échoue à décrire les systèmes pour lesquels la physique intéressante ne se fait pas à cette échelle, comme les plasmas dont la densité de charge moyennée à grande échelle est nulle.

Problèmes conceptuels de la physique statistique des interactions à longue portée

Un des objectifs de cette partie consacrée à la mécanique statistique d'équilibre des systèmes à longue portée était également d'essayer de clarifier certains problèmes de fond. En effet, cette physique statistique d'équilibre est parfois considérée comme problématique du point de vue conceptuel, en raison du caractère non additif des systèmes considérés. Elle a ainsi motivé l'introduction de théories "exotiques", comme celle de l'entropie de Tsallis présentée rapidement en introduction. En réalité, nous espérons avoir convaincu qu'il suffit de revenir aux premiers principes de la physique statistique pour obtenir une description cohérente de l'équilibre de ces systèmes : le postulat d'équiprobabilité en phase de Boltzmann. Cette idée de base est bien sûr très loin d'être nouvelle, et est utilisée depuis longtemps en astrophysique, ou en turbulence bidimensionnelle ³, mais n'avait apparemment pas, jusqu'à récemment, conquis tous les physiciens s'intéressant aux interactions à longue portée. On peut se reporter pour ce débat à la critique par Vollmayr-Lee et Luijten de l'utilisation de l'entropie de Tsallis dans des situations d'équilibre [110].

Par rapport aux systèmes habituels à courte portée, la principale différence concerne alors le scaling : il faut choisir un scaling différent de la limite thermodynamique standard, dépendant de la portée des interactions. D'autre part, on ne peut pas considérer une petite partie d'un grand système : chaque système doit être étudié comme un tout; les questions d'additivité de l'entropie ou de l'énergie pour deux systèmes que l'on rapproche n'ont alors pas beaucoup de sens. De ce fait, la phénoménologie est, comme on l'a vu, très différente de celle des systèmes à courte portée (inéquivalence d'ensembles, chaleur spécifique négative...), mais cela ne pose pas de problème conceptuel particulier.

L'étude générale des problèmes variationnels

La méthode des grandes déviations permet dans un certain nombre de cas de justifier rigoureusement l'approche champ moyen. De plus, elle permet de supposer que, lorsqu'aucune justification rigoureuse n'est disponible (systèmes auto-gravitants par exemple), cette approximation du champ moyen, constamment utilisée dans la littérature, est excellente.

Cette approche champ moyen réduit, exactement ou approximativement, la mécanique statistique d'équilibre à un problème variationnel, ce qui ouvre la voie à une étude générale de la phénoménologie de ces systèmes. Nous avons détaillé à la section 3.1

³Ce point de vue a été défendu également avec succès par D. Gross et ses collaborateurs dans le cadre des systèmes à petit nombre de particules, parfois vus également comme problématiques du point de vue de la physique statistique [59].

les résultats déjà connus, tirés des travaux de Katz et Ellis notamment. À la section suivante 3.2, nous avons poussé beaucoup plus loin cette étude générale et proposons une classification des transitions de phases et des différentes situations d'inéquivalence d'ensembles pouvant survenir pour ces systèmes. Cette classification s'appuie lorsque cela est possible sur les résultats mathématiques existants de la théorie des singularités. Les conclusions principales sont les suivantes :

- En codimension 0, c'est à dire sans paramètre extérieur et en faisant varier uniquement l'énergie ou la température, toutes les situations possibles ont déjà été observées sur les différents modèles étudiés en détail dans la littérature.
- La situation est tout à fait différente en codimension 1, c'est à dire en présence d'un paramètre extérieur que l'on peut faire varier : la classification a mis en évidence une variété insoupçonnée de phénoménologies possibles dans ce cas, dont beaucoup n'ont pas encore été observées sur un modèle précis. Deux nouveaux modes d'apparition d'inéquivalence d'ensembles ont notamment été découverts.

Cette classification est pour le moment limitée à un seul paramètre extérieur, et une seule quantité conservée. Elle peut naturellement être généralisée dans ces deux directions : en ajoutant un second paramètre extérieur, on aurait accès aux singularités de codimension 2; en considérant la possibilité de deux quantités conservées, on pourrait étudier le phénomène d'inéquivalence d'ensembles simultanément sur trois ensembles, par exemple microcanonique, canonique et grand canonique. Cette situation à deux quantités conservées est importante par exemple en astrophysique, où l'on peut souvent considérer qu'en plus de l'énergie, le moment cinétique total d'un amas d'étoiles ou d'une galaxie est conservé.

Remarquons enfin que l'inéquivalence d'ensembles *n'est pas* un phénomène spécifique aux interactions à longue portée : voir par exemple la référence [98], même si les auteurs ne présentent pas leurs résultats sous cet angle. Une étude de ce phénomène hors du cadre des interactions à longue portée reste à faire.

Que nous apprend la mécanique statistique d'équilibre ?

On voit donc que l'étude de la mécanique statistique d'équilibre des systèmes avec interactions à longue portée peut être poussée très loin. Mais que nous apprend-elle sur le comportement physique réel des systèmes physiques considérés ? Elle nous dit ceci : l'immense majorité des états microscopiques accessibles au système, en tenant compte des contraintes (quantités conservées), correspond à un état macroscopique bien défini μ^* , que l'on peut déterminer par le calcul. μ^* est, selon le cas, la densité de masse, la distribution de vorticit , la magn tisation d' quilibre du syst me... En d'autres termes, si l'on choisit au hasard un  tat microscopique pour le syst me sur la surface d' nergie constante (cas de l'ensemble microcanonique standard), il correspond presque certainement   l' tat macroscopique μ^* .

Malheureusement, la situation physique est souvent plut t celle-ci : le syst me part d'une condition initiale qui n'est pas en g n ral   l' quilibre, il  volue pendant un certain temps, et on cherche   conna tre son  tat. Sauf cas particulier (un syst me

intégrable par exemple), le système finira sans doute par atteindre son état d'équilibre statistique; mais la question du temps de relaxation nécessaire pour atteindre cet équilibre est alors cruciale, et nous allons voir à la partie suivante que la compréhension de ces systèmes avec interactions à longue portée peut requérir effectivement une étude hors équilibre, et la réintroduction du temps dans le problème.

Terminons ce chapitre par une remarque sur la méthodologie utilisée : dans le cadre de la mécanique statistique d'équilibre, la méthode des "modèles jouets" a pu être menée jusqu'au bout. Des concepts et outils généraux ont été mis en évidence sur des modèles simples, puis généralisés; finalement, la meilleure compréhension des phénomènes a permis de suggérer de nouveaux domaines d'application potentielle, comme l'interaction onde-particules dans certains lasers (voir le paragraphe 2.3.3 et le chapitre 6).

Part II

Dynamique hors équilibre

Introduction

On l'a vu au premier chapitre : on dispose d'outils puissants pour traiter la mécanique statistique d'équilibre des systèmes à longue portée (au moins d'une large classe d'entre eux). Ces résultats d'équilibre nous renseignent sur la structure de l'immense majorité des états microscopiques accessibles au système, et sont donc importants. Néanmoins, avant de pouvoir tirer de ces résultats des conclusions quant au comportement réel des systèmes considérés, il est nécessaire d'en connaître les propriétés de relaxation. De manière générale, il s'avère que les temps de relaxation pour les systèmes à longue portée peuvent être très longs, et augmenter avec le nombre de particules, ce qui rend indispensable une compréhension de la dynamique hors équilibre. La formation de structures cohérentes hors équilibre [4, 109], et, récemment, des phénomènes de vieillissement [115] ont aussi été observés.

Lors de l'étude des propriétés d'équilibre, nous avons commencé par nous appuyer sur des modèles jouets (modèles BEG champ moyen, HMF, α -Ising...), facilement solubles analytiquement, pour dégager des concepts généraux. Une approche très générale a finalement été rendue possible par l'étendue de la validité du champ moyen. Nous allons essayer de procéder de la même façon pour étudier les propriétés dynamiques, et on aimerait bien entendu trouver dans ce cas un concept unificateur aussi simple que le champ moyen. À première vue, on peut penser le trouver dans ce qui est en quelque sorte l'équivalent dynamique du champ moyen, l'équation de Vlasov.

L'équation de Vlasov

a. Convergence vers la dynamique de Vlasov

La mécanique statistique d'équilibre de particules en interaction rencontre très vite des difficultés de calcul importantes, même pour des modèles apparemment simples (modèle d'Ising par exemple). La situation de la théorie cinétique, qui se propose de déduire, à partir des équations microscopiques, l'évolution d'observables macroscopiques, n'est pas plus facile : il est en général impossible de tenir compte des corrélations que la dynamique fait naître entre les particules. A priori, la théorie cinétique décrit un système par une distribution de probabilité sur l'espace des phases à N particules, $f_N(\mathbf{r}_1, \mathbf{p}_1, \dots, \mathbf{r}_N, \mathbf{p}_N, t)$; cette distribution contient toute l'information sur les corrélations. L'approximation la plus simple consiste alors à négliger ces corrélations, et à décrire le système par une seule distribution de probabilité, à une particule, $f(\mathbf{r}, \mathbf{p}, t)$; la distribution de probabilité à N particules f_N est alors donnée par

$$f_N(\mathbf{r}_1, \mathbf{p}_1, \dots, \mathbf{r}_N, \mathbf{p}_N, t) = f(\mathbf{r}_1, \mathbf{p}_1, t) \dots f(\mathbf{r}_N, \mathbf{p}_N, t). \quad (3.19)$$

L'évolution de cette fonction à une particule se fait alors sous l'effet du potentiel moyen dans le système d'une part, des collisions entre particules d'autre part :

$$\frac{\partial f}{\partial t} + \mathbf{p} \cdot \nabla_{\mathbf{r}} f - \nabla_{\mathbf{r}} U \cdot \nabla_{\mathbf{p}} f = C(f), \quad (3.20)$$

où U est le potentiel, et $C(f)$ représente l'évolution due aux collisions. L'équation de Vlasov s'obtient à partir de celle-ci en négligeant le terme de collisions. Il s'agit en général d'une mauvaise approximation, mais lorsque l'interaction entre particules est à longue portée, le potentiel dépend lui même de façon non locale de f , et ce terme produit une évolution non linéaire complexe, qui peut être dominante. De fait, il existe des résultats exacts très généraux permettant de démontrer la convergence de la dynamique particulaire véritable vers la dynamique de l'équation de Vlasov, pour un nombre de particules N tendant vers l'infini.

Plus précisément, le théorème de Braun et Hepp [28] dit : soit un système classique de N particules, interagissant via un potentiel

$$E_p = \frac{1}{N} \sum_{1 \leq i < j \leq N} \Phi(\mathbf{x}_i - \mathbf{x}_j), \quad (3.21)$$

avec Φ paire ($\Phi(\mathbf{x}) = \Phi(-\mathbf{x})$), et suffisamment régulière (par exemple deux fois dérivable et à dérivées bornées). Alors pour tout temps t et pour toute erreur admissible fixée ε , il y a un nombre de particules N pour lequel les dynamiques particulaire et de Vlasov coïncident jusqu'au temps t , avec une erreur maximale ε . Il *ne dit pas* : pour toute erreur ε , il existe N tel que pour tout t les dynamiques coïncident à ε près⁴...

Comme pour les études à l'équilibre, la renormalisation par $1/N$ de l'interaction dans (3.21) peut a priori paraître étrange; mais elle correspond, comme discuté au paragraphe 2.4.1, à une limite d'un nombre de particules tendant vers l'infini, prise en gardant constante l'échelle de temps microscopique du système (la pulsation plasma par exemple); ce peut être la limite physiquement intéressante. Enfin, remarquons que les hypothèses de régularité sur le potentiel Φ ne sont pas satisfaites pour les systèmes auto-gravitants; on retrouve là, comme pour l'équilibre, le fait que les difficultés de ce potentiel sont plutôt dues au comportement à courte distance.

L'équation de Vlasov hérite de la dynamique particulaire la conservation des mêmes quantités conservées : énergie totale, quantité de mouvement... D'autre part, le passage à une description continue crée une infinité de nouvelles quantités conservées, appelées Casimirs. En effet, l'équation de Vlasov n'est qu'une équation d'advection (non linéaire) de la densité f , si bien que toutes les quantités du type

$$I[\varphi] = \int \varphi(f(\mathbf{r}, \mathbf{p}, t)) \, d\mathbf{r} d\mathbf{p}, \quad (3.22)$$

où φ est une fonction quelconque, sont conservées. On peut exprimer cela de façon équivalente en disant que l'aire associée à chaque niveau de f est conservée. Ces

⁴Un énoncé mathématiquement précis du théorème nécessiterait d'entrer dans les détails de ce que l'on entend par erreur, convergence... Nous avons choisi de rester à ce niveau heuristique.

nouvelles quantités conservées vont jouer un rôle important dans la dynamique particulaire. D'autre part, il est connu que l'équation possède énormément de solutions stationnaires ou se déplaçant à vitesse constante (modes BGK, pour Bernstein, Greene et Kruskal). Ces deux faits, la présence d'une infinité de nouvelles quantités conservées et l'existence de très nombreux états stationnaires, peuvent donner une idée intuitive de la raison pour laquelle la relaxation à l'équilibre des systèmes avec interactions à longue portée est de façon générale si lente : nous reviendrons sur ce point au chapitre 5 et détaillerons cette idée sur l'exemple du modèle HMF.

b. Relaxation collisionnelle, relaxation violente

Puisqu'on a négligé les collisions entre particules, et créé de ce fait une infinité de nouvelles quantités conservées, on pourrait penser que la dynamique de Vlasov ne relaxe pas vers un équilibre statistique. En réalité, la dynamique d'advection non linéaire de Vlasov crée rapidement une filamentation de plus en plus fine de la densité f , si bien qu'une distribution d'équilibre peut apparaître pour la densité \bar{f} , moyenne à gros grains de la densité réelle f . C'est le principe de la relaxation violente, introduit par Lynden-Bell en 1968 dans le contexte astrophysique [87]; elle agit sur des temps beaucoup plus courts que la relaxation due aux collisions entre particules (appelée relaxation collisionnelle), et conduit à un équilibre statistique *différent* de l'équilibre collisionnel complet, car il contient les traces des nouvelles quantités conservées⁵. Nous n'entrons pas ici dans les détails de la mécanique statistique de l'équation de Vlasov, dont nous donnerons un exemple au chapitre 6.

On peut remarquer que l'équation de Vlasov est très proche de l'équation d'Euler de la turbulence 2D qui est aussi une équation d'advection non linéaire, de la vorticit   cette fois, et conserve   galement une infinit   de Casimirs. Cette remarque a permis    Chavanis, Sommeria et al. [33] d'  tendre aux syst  mes auto gravitants les r  sultats de m  canique statistique d'  quilibre obtenus pour la turbulence 2D [91, 99, 90], et de d  velopper la profonde analogie entre ces domaines de la physique a priori tr  s   loign  s.

L'  quation de Vlasov peut elle jouer un r  le analogue, pour la compr  hension de la dynamique,    celui du champ moyen pour l'  quilibre ? Le r  sultat de Braun et Hepp incite    penser que oui. En r  alit  , le th  or  me montre la convergence vers la dynamique de Vlasov lorsque $N \rightarrow \infty$, pour tout *intervalle de temps born  *. Autrement dit, la dynamique des particules et la dynamique de Vlasov co  ncident sur un temps qui diverge avec N ; mais cette divergence peut   tre lente, et une   tude attentive des   chelles de temps est n  cessaire pour chaque cas : le temps pour lequel l'  quation de Vlasov est une bonne approximation de la dynamique, le temps de relaxation de l'  quation de Vlasov elle m  me vers son   quilibre, et le temps de relaxation vers l'  quilibre statistique v  ritable (appel   temps de relaxation collisionnel).

- Une inspection des d  monstrations du th  or  me de Braun et Hepp, ou d'autres th  or  mes analogues pour d'autres syst  mes [54], donne une estimation du temps

⁵Dans le cas des syst  mes auto gravitants, la prise en compte des Casimirs permet en fait de d  finir rigoureusement un   quilibre statistique, qui n'existe pas    proprement parler pour le syst  me particulaire; nous avons d  j   fait allusion    ce fait au cours du premier chapitre.

t_v sur lequel l'équation de Vlasov coïncide avec la dynamique particulaire. On trouve a priori $t_v \sim \ln N$. On verra cependant au chapitre 5 que, lorsque la condition initiale correspond à un état stationnaire de l'équation de Vlasov, ce temps peut être bien plus grand, d'ordre N .

- Le temps de relaxation de l'équation de Vlasov vers son équilibre statistique ne dépend bien sûr pas du nombre de particules N ; il est néanmoins connu que cette relaxation peut être très ralentie, voire stoppée, et la dynamique être piégée dans l'un des nombreux états stationnaires de l'équation de Vlasov. Ainsi, même pour un système "réellement" continu (un fluide par exemple), où le temps de relaxation de l'équation de Vlasov n'est pas en compétition avec d'autres temps⁶, la détermination de l'état stationnaire pertinent de la dynamique de Vlasov peut être très difficile. Nous reviendrons aussi sur ce point au chapitre 5, par l'étude d'un exemple.
- Enfin, le temps de relaxation collisionnel dépend fortement de N . Il a été estimé pour les systèmes auto gravitants par Chandrasekhar $t_c \sim N/\ln N$. On verra, toujours au chapitre 5, qu'il peut être différent, sur certains modèles simples; de manière générale, $t_c \gg t_v$.

Il est connu depuis longtemps en astrophysique que, puisque ces temps dépendent du nombre de particules N , l'échelle de temps, et donc la description (Vlasovienne ou particulaire) pertinente dépend de la taille du système. Ainsi, le temps de relaxation collisionnel d'une galaxie est gigantesque, beaucoup plus grand que l'âge de l'univers. On pense donc que c'est le processus de relaxation violente (introduit pour cela d'ailleurs par Lynden-Bell) qui est dominant. En revanche, le temps de relaxation des amas globulaires, qui comptent beaucoup moins d'étoiles, est plus petit, si bien que la relaxation collisionnelle et l'équilibre statistique "véritable" pourraient être pertinents dans ce cas. Il est aussi tout à fait possible que l'échelle de temps intéressante soit précisément entre les deux : dans ce cas, le système n'a pas encore relaxé vers l'équilibre statistique complet, mais les effets de granularité, ou de taille finie, jouent déjà un rôle important. Nous essaierons d'étudier précisément le comportement d'un modèle jouet dans cette zone intermédiaire, au chapitre 5.

L'approche de Tsallis

La description, par définition délicate, de ce régime intermédiaire a fait l'objet très récemment de nombreuses études en lien avec l'entropie de Tsallis : certains auteurs pensent en effet que l'état du système dans cette région, qui ne maximise pas l'entropie habituelle (ce serait alors l'équilibre), maximise l'entropie de Tsallis. Le modèle jouet HMF ferromagnétique, déjà présenté au chapitre 2, a notamment fait l'objet de simulations numériques détaillées; les distributions de vitesses obtenues ont

⁶Nous voulons dire qu'il n'existe pas de temps lié au nombre de particules; il peut exister en général d'autres temps, liés par exemple à la dissipation.

alors été fittées par une distribution “de Tsallis”

$$P(v) = \left[1 - \frac{\beta(1-q)v^2}{2} \right]^{1/(1-q)} ; \quad (3.23)$$

Latora *et al.* utilisent la température obtenue numériquement, et ajustent le paramètre q ainsi qu’un paramètre de coupure aux grandes vitesses [82].

D’autre part, des simulations numériques plus réalistes d’un système de points auto gravitants en dimension 3 ont montré une phénoménologie un petit peu similaire : après une évolution rapide, le système est assez bien décrit par une distribution “de Tsallis”, puis son évolution, plus lente, peut être caractérisée par une dépendance temporelle du paramètre de Tsallis $q(t)$ [106]. Il n’y a plus ici un unique paramètre q , mais une famille qui décrit l’évolution lente du système avant l’équilibre statistique final.

À l’heure actuelle, nous ne connaissons pas d’argument théorique pour expliquer une sorte d’universalité des distributions de Tsallis comme résultat de la relaxation violente, ou dans cette région d’évolution intermédiaire où les effets granulaires se font sentir, mais la relaxation collisionnelle n’est pas achevée; les preuves numériques, au moins dans le cas du modèle HMF ferromagnétique, nous semblent peu convaincantes et limitées à des conditions initiales très particulières. Nous suivons au chapitre 5 une approche qui nous semble plus logique, et qui consiste, fidèles à notre stratégie de modèles jouets, à étudier en détail le cas du HMF ferromagnétique avec les outils à notre disposition : les distributions de Tsallis, si elles sont pertinentes, doivent apparaître à la fin de l’étude, plutôt qu’être postulées a priori. Pour terminer sur ce sujet, précisons que l’entropie de Tsallis a été appliquée à un très grand nombre de problèmes différents, dont la relaxation lente des systèmes avec interactions à longue portée n’est qu’une petite partie (le site [108] fournit une bibliographie complète); notre étude ne peut apporter d’éléments que concernant cette petite partie.

Plan du chapitre

Nous nous intéressons de manière générale dans ce chapitre à la relaxation vers l’équilibre des systèmes avec interactions à longue portée, aux raisons qui la ralentissent, et aux phénomènes hors équilibre qu’elle engendre.

Nous concentrons notre étude sur deux modèles jouets, dont l’équilibre statistique a déjà été étudié au chapitre 1 : les modèles HMF antiferromagnétique et ferromagnétique. Le but est bien sûr de dégager ensuite des concepts généraux.

Le HMF antiferromagnétique présente la formation d’intéressantes structures hors équilibre (paragraphe 2.3.2), qui restaient inexplicées; nous les étudions en détail au chapitre 4, et montrons qu’elles sont dues à la présence de deux échelles de temps bien séparées dans la dynamique. Le HMF antiferromagnétique donne ainsi un exemple entièrement soluble analytiquement de dynamique à N particules en interaction, où la séparation des échelles de temps conduit à la formation de structures hors équilibre, stables sur des temps très longs. Il s’agit là d’un phénomène assez fréquent en physique, et il est donc intéressant de disposer d’un modèle dynamique à N particules sur lequel les calculs peuvent aller très loin. Cependant, nous ne nous attendons pas à ce que

ce phénomène de séparation des échelles temporelles soit une explication générale à la relaxation lente pour les systèmes avec interactions à longue portée.

En revanche, l'étude du HMF ferromagnétique au chapitre 5 nous permettra de faire des conjectures plus générales. Nous commencerons par présenter les études (numériques) et les résultats obtenus précédemment pour ce système, qui a déjà une longue carrière de modèle jouet. Puis, nous montrerons comment le comportement aux temps courts peut être précisément étudié à l'aide de l'équation de Vlasov, et présenterons un scénario général pour la dynamique de ce genre de systèmes; nous le discuterons en liaison avec les différentes théories proposées, et les résultats numériques obtenus pour le HMF ferromagnétique et d'autres modèles. Nous appliquerons les idées et résultats issus de cette étude au chapitre 6, sur un modèle de laser à électrons libres.

Chapter 4

Formation de structures hors équilibre : étude du modèle HMF antiferromagnétique

4.1 Introduction

Nous avons déjà introduit le modèle HMF antiferromagnétique au paragraphe 2.3.2. Rappelons ici le Hamiltonien :

$$H = \sum_{i=1}^N \frac{p_i^2}{2} + \frac{1}{2N} \sum_{i,j} \cos(\theta_i - \theta_j). \quad (4.1)$$

Les positions θ_i des particules sont comprises entre 0 et 2π , et obéissent aux équations du mouvement :

$$\ddot{\theta}_i = M_1 \sin(\theta_i - \varphi_1), \quad (4.2)$$

où nous notons

$$\mathbf{M}_1 = M_1 e^{i\varphi_1} = \frac{1}{N} \sum_j \exp(i\theta_j). \quad (4.3)$$

L'énergie potentielle peut se réécrire

$$E_p = \frac{1}{2N} \left| \sum_j e^{i\theta_j} \right|^2 = \frac{N}{2} |\mathbf{M}_1|^2. \quad (4.4)$$

Les configurations à M_1 faible sont donc énergétiquement favorisées. Depuis son introduction par Antoni et Ruffo en 1995 [4], il est connu que le modèle HMF antiferromagnétique donne naissance à des structures cohérentes de durée de vie très longue. La solution statistique canonique du modèle, qui prévoit une distribution homogène des particules sur le cercle quelle que soit la température, ne les explique pas. À basse énergie, le système s'organise spontanément en un double amas de particules, suivant une dynamique particulière : voir la figure 4.1. Ce double amas, dont la distribution après un temps de relaxation assez grand est montrée à la figure 4.2, est stable sur des

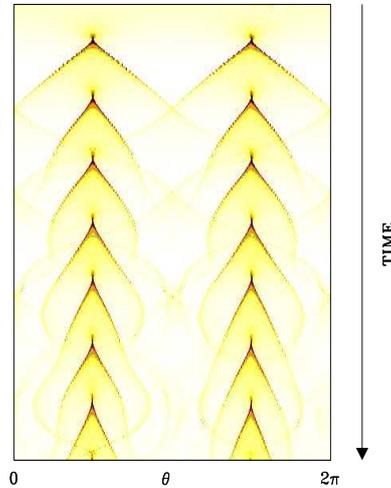


Figure 4.1: Formation du double amas aux temps courts : évolution de la densité de particules en niveaux de gris (plus le gris est foncé, plus la densité est élevée). À partir d’une condition initiale où les particules sont distribuées de façon homogène, on observe une rapide concentration, puis l’apparition quasi périodique des chevrons, qui se resserrent lorsque le temps augmente.

temps très longs (aucune décroissance n’est observée numériquement pour N grand); la présence de cette structure modifie la chaleur spécifique du système. Ces phénomènes ont été étudiés numériquement en détail par Dauxois, Ruffo et Holdsworth [39].

On pouvait a priori supposer qu’il s’agissait d’une manifestation de l’inéquivalence d’ensembles, et qu’une analyse statistique microcanonique expliquerait la présence du double amas. En réalité, les deux ensembles sont équivalents dans ce cas là (voir le paragraphe 2.3.2), et certaines expériences numériques tendent à montrer qu’il s’agit d’un effet dynamique hors équilibre, stable sur des temps très longs. D’une part, le double amas se forme à petite énergie pour une grande variété de conditions initiales, mais pas pour toutes : la distribution spatiale initiale peut être quelconque, mais il est nécessaire que la dispersion initiale des vitesses soit assez faible. D’autre part, des simulations numériques plus récentes avec un petit nombre de particules ont montré une dégradation aux temps longs du double amas [85]. Une explication de la formation de cette structure hors équilibre, de sa stabilité et de ses propriétés manquait encore.

La réponse repose sur la présence de deux échelles de temps dans la dynamique des particules : une rapide oscillation de petite amplitude (de type plasma, due à leur répulsion), et un mouvement lent de grande amplitude. En exploitant ce fait, nous obtenons par une méthode de moyenne, à la section 4.2, un Hamiltonien effectif qui décrit le mouvement lent des particules. Ce Hamiltonien effectif est lui même de type champ moyen, il fait donc partie de nos “systèmes à longue portée”.

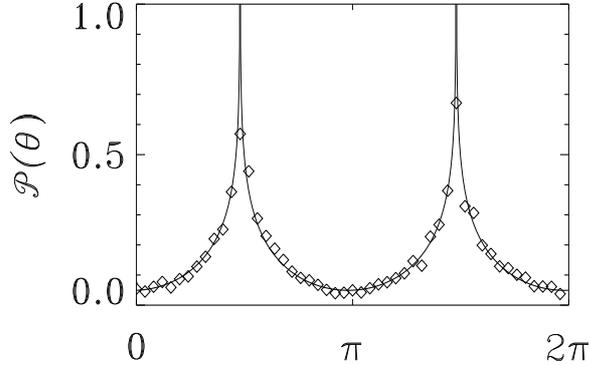


Figure 4.2: Distribution des particules en fonction de la position $\mathcal{P}(\theta)$, pour $N = 10^4$, une énergie par particule $e \simeq 10^{-5}$, et après un temps de relaxation assez long : environ 10^5 fois la période plasma.

Nous divisons ensuite le problème en deux parties. Dans un premier temps, nous nous intéresserons au comportement du système aux temps courts (section 4.3), avec l’objectif de décrire précisément les structures en chevrons de la figure 4.1. Dans un second temps, nous étudierons le comportement aux temps longs (section 4.4). Cette fois-ci, il s’agit d’expliquer la stabilité du double amas, sa forme (figure 4.2), la chaleur spécifique modifiée du système...

Selon Antoni *et al.* [6], ce modèle jouet permet de reproduire qualitativement, et même quantitativement, les propriétés dynamiques d’un système Coulombien à une dimension, pour des énergies supérieures à celles que nous allons explorer dans la suite. Néanmoins, l’intérêt de son étude ne repose pas sur ses applications physiques directes, mais plutôt sur les phénomènes qu’il met en valeur : d’une part, la présence de plusieurs échelles de temps, la possibilité d’étudier une dynamique effective sont des phénomènes courants en physique des plasmas [5], et d’autre part l’étude du Hamiltonien effectif fera apparaître beaucoup de caractéristiques dynamiques générales des systèmes avec interactions à longue portée.

Nous avons publié l’étude de la formation de la structure dans [11]. Nous employons dans cet article une méthode un petit peu différente de celle que nous allons présenter ici, fondée sur l’utilisation de l’équation de Vlasov. Les calculs nécessaires sont plus simples, mais les résultats sont moins précis et ne peuvent pas vraiment être exploités pour l’étude des temps longs. Cette étude aux temps longs est rendue possible grâce à la description Hamiltonienne effective; nous l’avons publiée dans [13, 14].

4.2 La dynamique effective

La procédure d’obtention du Hamiltonien effectif repose sur l’existence de deux échelles de temps bien séparées dans la dynamique des particules, lorsque l’énergie du système

est faible : le temps rapide est intrinsèque, lié à la fréquence plasma du système, le temps lent dépend de l'énergie. La figure 4.3 met bien en évidence que la trajectoire de chaque particule est la superposition d'oscillations rapides de petite amplitude avec un mouvement lent de grande amplitude. Nous allons employer la procédure suivante :

- introduire explicitement des variables rapides et lentes en utilisant pour θ_i la forme

$$\theta_i = \theta_i^0(\tau) + \varepsilon f_i(t, \tau), \quad (4.5)$$

avec $\tau = \varepsilon t$. Les f_i représentent les petites et rapides oscillations, et les θ_i sont les variables lentes que l'on veut étudier; ε est donc proportionnel à l'amplitude des petites oscillations;

- obtenir les équations du mouvement pour les variables rapides f_i et les résoudre explicitement (paragraphe 4.2.2);
- réintroduire cette solution dans l'expression de l'action du système (nous allons employer une formulation Lagrangienne), et moyenner cette action sur le temps rapide, pour obtenir une dynamique restreinte aux variables lentes (paragraphe 4.2.3).

L'idée de base est en fait celle de toutes les méthodes variationnelles : on cherche des solutions aux problèmes en se restreignant à un espace de fonctions particulier, ici celles de la forme (4.5).

On conçoit que la méthode soit d'autant meilleure que la séparation des échelles de temps est bien vérifiée, c'est à dire que ε est petit. Or ε est lié à l'amplitude des petites oscillations, si bien que l'énergie de ces petites oscillations est d'ordre ε^2 . On peut donc estimer $\varepsilon \sim \sqrt{e}$, où e est l'énergie par particule dans le système. Les simulations de [39] qui mettent en évidence le phénomène correspondent à des énergies par particule de l'ordre de 10^{-4} ou moins, donc à $\varepsilon \simeq 10^{-2}$. On s'attend donc à ce que la procédure décrite ci-dessus soit très bonne.

Raisonnement en restant au niveau du Lagrangien (ou du Hamiltonien), plutôt que sur les équations du mouvement, a plusieurs avantages. D'une part cette approche, inspirée par le livre de Witham [111], fournit une formulation Hamiltonienne pour les variables lentes, et d'autre part elle donne accès naturellement à une quantité "quasi-conservée", l'invariant adiabatique. Nous donnons au paragraphe suivant 4.2.1 un exemple d'application simple de la méthode, pour insister sur les concepts plutôt que sur les calculs un petit peu techniques qui vont suivre.

4.2.1 Description variationnelle d'une dynamique adiabatique : étude d'un exemple simple

Deux oscillateurs couplés Considérons deux oscillateurs couplés, l'un rapide (y) et l'autre lent (x), décrit par le Lagrangien

$$L(x, \dot{x}, y, \dot{y}) = \frac{\dot{x}^2}{2} + \frac{\dot{y}^2}{2} - \frac{\varepsilon^2 x^2}{2} - \frac{y^2}{2} - \frac{x^2 y^2}{2}. \quad (4.6)$$

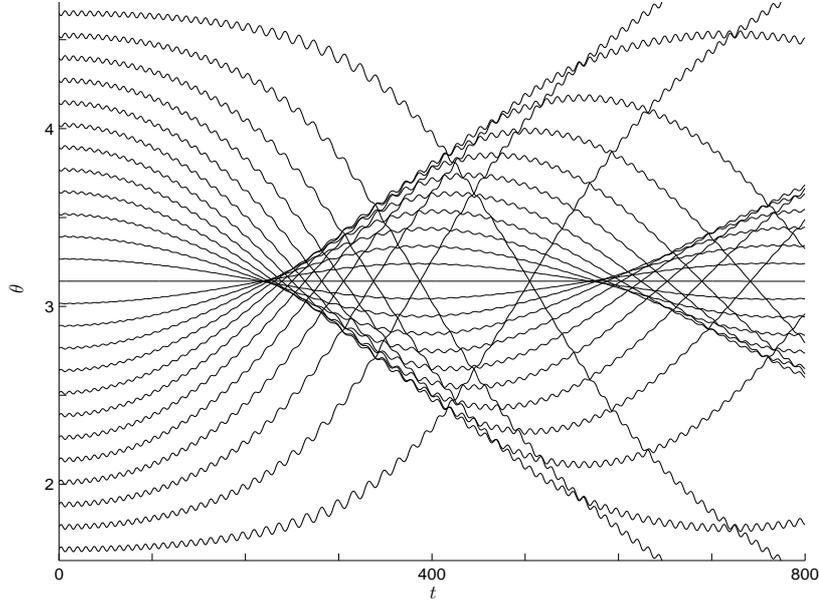


Figure 4.3: Trajectoire des particules pour le Hamiltonien d'origine 4.1. Énergie par particule $e = 1.3 \cdot 10^{-5}$.

On veut décrire le mouvement lent de x , c'est-à-dire celui qui ne dépend que de τ . Les oscillations de y sont rapides, on les suppose de petite amplitude; on utilise donc l'ansatz :

$$y(t) = \varepsilon a(\tau) \sin \varphi(t), \quad (4.7)$$

où $\tau = \varepsilon t$ est l'échelle de temps lente, la phase $\varphi(t)$ évolue rapidement, mais $d\varphi/dt$ est une variable lente. Au lieu d'effectuer un développement asymptotique sur les équations du mouvement, comme il est plus habituel de faire, nous conservons la formulation variationnelle (suivant Witham [111]). On introduit donc l'ansatz (4.7) dans le Lagrangien (4.6), puis on moyenne sur la variable rapide φ . Après simplification d'un facteur ε^2 , on obtient

$$\mathcal{L}(x, \dot{x}, \varphi, \dot{\varphi}, a) = \langle L \rangle_t = \frac{\dot{x}^2}{2} + \frac{a^2 \dot{\varphi}^2}{4} - \frac{x^2}{2} - \frac{a^2 x^2}{4} - \frac{a^2}{4}. \quad (4.8)$$

Cette procédure de moyenne sur le temps rapide produit nécessairement des Lagrangiens cycliques vis à vis des phases rapides, même pour des exemples plus compliqués. Ici, (4.8) est indépendant de φ , ce qui fournit naturellement un invariant (l'invariant adiabatique) :

$$\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{1}{2} a^2 \dot{\varphi} = P. \quad (4.9)$$

Passons maintenant en formalisme Hamiltonien pour obtenir une formulation de la dynamique effective de x . La dérivée temporelle de a n'intervient pas; il n'y a donc pas de transformation de Legendre à effectuer sur cette variable. On obtient

$$\mathcal{H}(x, p_x, P, a) = P \dot{\varphi} + p_x \dot{x} - \mathcal{L} = \frac{p_x^2}{2} + \frac{P^2}{a^2} + \frac{x^2}{2} + \frac{a^2 x^2}{4} + \frac{a^2}{4}, \quad (4.10)$$

où P est une constante, et p_x le moment conjugué de x . L'équation "du mouvement" pour a s'écrit simplement

$$\frac{\partial \mathcal{H}}{\partial a} = 0, \quad (4.11)$$

ce qui fournit une expression de a en fonction de P et x , $a^4 = 4P^2/(1+x^2)$, ainsi que de la fréquence $\dot{\varphi} = \sqrt{1+x^2}$. En introduisant l'expression de a dans l'équation (4.10), on obtient le Hamiltonien effectif décrivant la dynamique de x :

$$\mathcal{H}_{eff}(x, p_x) = \frac{p_x^2}{2} + \frac{x^2}{2} + P\sqrt{1+x^2}. \quad (4.12)$$

L'invariant adiabatique P apparaît comme un paramètre dans la dynamique effective décrite par le Hamiltonien (4.12). L'obtention d'une formulation Hamiltonienne pour la dynamique effective sera très utile dans l'étude du modèle HMF : elle nous permettra d'utiliser les méthodes de la mécanique statistique.

Formalisme Hamiltonien On peut se demander pourquoi passer au formalisme Hamiltonien en cours de calcul, et pourquoi ne pas, par exemple, obtenir un Lagrangien effectif en réintroduisant dans (4.8) les expressions de $\dot{\varphi}$ et a obtenues. En fait, cette procédure ne donne pas le bon résultat, comme on peut s'en convaincre sur un exemple très simple. Prenons un Lagrangien qui ne dépend que d'une variable a , et de la dérivée d'une variable x , $L(\dot{x}, a)$. Les équations du mouvement sont simplement $\partial L/\partial \dot{x} = p = cste$ et $\partial L/\partial a = 0$. Invertissons la première pour obtenir \dot{x} en fonction de p et a : $\dot{x} = f(p, a)$. Si on réintroduit cela dans le Lagrangien de départ, on obtient un Lagrangien effectif pour a :

$$L_{eff}(a) = L(f(p, a), a). \quad (4.13)$$

On voit bien que l'équation du mouvement pour a déduite de L_{eff} n'est pas la même que celle déduite de L ...

En revanche, si on écrit le Hamiltonien

$$H_{eff}(a) = pf(p, a) - L(f(p, a), a), \quad (4.14)$$

l'équation du mouvement $\partial H_{eff}/\partial a = 0$ est correcte. Le formalisme Hamiltonien, en rendant formellement indépendants chaque coordonnée et son moment conjugué, permet donc ces substitutions. Nous allons utiliser cette propriété dans ce qui suit.

4.2.2 Les variables rapides

Revenons maintenant au modèle HMF. La méthode utilisée est très proche de celle illustrée ci-dessus, mais les calculs sont plus ardues. Le Lagrangien du système est

$$L(\theta_i, \dot{\theta}_i) = \sum_{i=1}^N \frac{\dot{\theta}_i^2}{2} - \frac{1}{2N} \sum_{i,j} \cos(\theta_i - \theta_j). \quad (4.15)$$

Les équations du mouvement sont données par la minimisation de l'action $S = \int L dt$. On veut introduire dans (4.15) l'ansatz (4.5); commençons par quelques remarques

et définitions.

Les petites oscillations, qui ont une amplitude d'ordre ε , et le mouvement lent, qui évolue sur des temps d'ordre ε , ont tous deux une énergie d'ordre ε^2 . La magnétisation totale est donc d'ordre ε , de même que la magnétisation lente $\mathbf{M}_1^0 = \left(\sum_j e^{i\theta_j^0} \right) / N$. On peut donc poser

$$\varepsilon M_{1x}^0 = \frac{1}{N} \sum_i \cos(\theta_i^0 + \psi) \quad (4.16)$$

$$\varepsilon M_{1y}^0 = \frac{1}{N} \sum_i \sin(\theta_i^0 + \psi), \quad (4.17)$$

où ψ est une phase que nous allons définir dans ce qui suit.

On s'intéresse à la formation du double amas de particules; il est alors naturel d'introduire un indicateur de son importance : $\mathbf{M}_2 = \left(\sum_j e^{2i\theta_j^0} \right) / N$. Si $|\mathbf{M}_2| = 1$, les particules sont regroupées en deux points séparés de π sur le cercle; si $|\mathbf{M}_2| = 0$, le double amas est inexistant. On choisit la phase ψ telle que

$$|\mathbf{M}_2| = \left| \frac{1}{N} \sum_i \exp(2i\theta_i^0) \right| = \frac{1}{N} \sum_i \cos(2\theta_i^0 + 2\psi). \quad (4.18)$$

On introduit maintenant l'ansatz (4.5) dans le Lagrangien (4.15); on développe toutes les expressions jusqu'à l'ordre ε^2 , pour obtenir un nouveau Lagrangien L_2 qui dépend des θ_i^0 , des f_i , et de leur dérivée par rapport au temps :

$$\begin{aligned} L_2 &= L_2 \left(\theta_i^0, \frac{d\theta_i^0}{d\tau}, f_i, \frac{df_i}{dt} \right) \\ &= \frac{\varepsilon^2}{2} \sum_i \left[\left(\frac{d\theta_i^0}{d\tau} \right)^2 + 2 \frac{d\theta_i^0}{d\tau} \frac{df_i}{dt} + \left(\frac{df_i}{dt} \right)^2 \right] \\ &\quad - \frac{\varepsilon^2 N}{2} \left[(M_{1x}^0)^2 + (M_{1y}^0)^2 \right] + \frac{\varepsilon}{N} \sum_{i,j} f_i \sin(\theta_i^0 - \theta_j^0) \\ &\quad - \frac{\varepsilon^2}{2N} \sum_{i,j} f_i f_j \cos(\theta_i^0 - \theta_j^0). \end{aligned} \quad (4.19)$$

À partir de cette expression, on obtient les équations du mouvement pour les variables rapides, en considérant τ comme une constante :

$$\ddot{f}_i = M_{1x}^0 \sin(\theta_i^0 + \psi) - M_{1y}^0 \cos(\theta_i^0 + \psi) - \frac{1}{N} \sum_j f_j \cos(\theta_i^0 - \theta_j^0). \quad (4.20)$$

C'est une équation linéaire du second ordre, à coefficients constants par rapport au temps rapide. Il faut donc pour la résoudre diagonaliser la matrice :

$$\mathbf{T} = [T_{ij}] = \frac{1}{N} [\cos(\theta_i^0 - \theta_j^0)]; \quad (4.21)$$

on montre que \mathbf{T} n'a que deux valeurs propres non nulles (Appendice B), définissant les deux fréquences ω_{\pm} :

$$\omega_{\pm} = \sqrt{\frac{1 \pm |M_2|}{2}}. \quad (4.22)$$

Les deux vecteurs propres associés sont

$$\mathbf{X}_+ = [\cos(\theta_i^0 + \psi)]_{i=1\dots N} \quad (4.23)$$

$$\mathbf{X}_- = [\sin(\theta_i^0 + \psi)]_{i=1\dots N}. \quad (4.24)$$

On en déduit (par la méthode de variation de la constante par exemple) la solution générale pour les oscillations rapides f_i

$$\begin{aligned} f_i(t, \tau) &= \left[\sqrt{2}A_+ \sin(\omega_+ t + \varphi_+) - \frac{M_{1y}^0}{\omega_+^2} \right] \cos(\theta_i^0 + \psi) \\ &+ \left[\sqrt{2}A_- \sin(\omega_- t + \varphi_-) + \frac{M_{1x}^0}{\omega_-^2} \right] \sin(\theta_i^0 + \psi). \end{aligned} \quad (4.25)$$

Les termes correspondant aux valeurs propres nulles n'apparaissent pas, car ils divergeraient avec le temps et contrediraient l'uniformité du développement en ε .

4.2.3 Les variables lentes

La solution précédente pour les f_i a été obtenue en prenant τ constant. En réalité, les θ_i^0 , les amplitudes A_{\pm} , les décalages $a_+ = -M_{1y}^0/\omega_+^2$, $a_- = M_{1x}^0/\omega_-^2$ et les fréquences propres dépendent de τ . On va donc utiliser pour les f_i le nouvel ansatz

$$\begin{aligned} f_i &= \left[\sqrt{2}A_+(\tau) \sin(\phi_+(t)) + a_+(\tau) \right] \cos(\theta_i^0 + \psi) \\ &+ \left[\sqrt{2}A_-(\tau) \sin(\phi_-(t)) + a_-(\tau) \right] \sin(\theta_i^0 + \psi), \end{aligned} \quad (4.26)$$

où les $\phi_{\pm}(t)$ sont des variables rapides, mais $d\phi_{\pm}/dt$, $A_{\pm}(\tau)$, et $a_{\pm}(\tau)$ sont des variables lentes. On introduit (4.26) dans le Lagrangien (4.19), et on moyenne par rapport aux variables rapides $\phi_{\pm}(t)$. On laisse de côté un facteur ε^2 , et on obtient le Lagrangien moyenné

$$\begin{aligned} L_{eff} &= \frac{1}{2} \sum_{i=1}^N \left(\frac{d\theta_i^0}{d\tau} \right)^2 + \frac{N}{2} \left[A_+^2 \omega_+^2 \dot{\phi}_+^2 + A_-^2 \omega_-^2 \dot{\phi}_-^2 \right] - N \left[\frac{M_{1x}^0{}^2 + M_{1y}^0{}^2}{2} \right. \\ &\left. + M_{1y}^0 a_+ \omega_+^2 - M_{1x}^0 a_- \omega_-^2 + (A_+^2 + a_+^2) \frac{\omega_+^4}{2} + (A_-^2 + a_-^2) \frac{\omega_-^4}{2} \right], \end{aligned} \quad (4.27)$$

où les ω_{\pm} sont les fréquences introduites au paragraphe précédent. Puisqu'on a moyenné sur les $\phi_{\pm}(t)$, ces phases n'apparaissent plus dans l'expression du Lagrangien (4.27), ce qui implique que leur moments conjugués P_{\pm} sont conservés :

$$P_{\pm} = \frac{\partial L_{eff}}{\partial \dot{\phi}_{\pm}} = N A_{\pm}^2 \omega_{\pm}^2 \dot{\phi}_{\pm}. \quad (4.28)$$

Les expressions ci-dessus sont l'analogie de (4.9) dans l'exemple des deux oscillateurs. On passe maintenant en formalisme Hamiltonien (suivant la remarque du paragraphe 4.2.1); comme (4.27) ne dépend pas des dérivées temporelles de A_+ , A_- , a_+ , et a_- , on ne fait pas de transformée de Legendre sur ces variables. On obtient, en appelant p_i^0 les moments conjugués des θ_i^0 :

$$\begin{aligned}
H_{eff} &= P_+ \dot{\phi}_+ + P_- \dot{\phi}_- + \sum_{i=1}^N p_i^0 \dot{\theta}_i^0 - L_{eff} & (4.29) \\
&= \frac{P_+^2}{2NA_+^2\omega_+^2} + \frac{P_-^2}{2NA_-^2\omega_-^2} + \sum_{i=1}^N \frac{p_i^{02}}{2} \\
&\quad + N \left[\frac{M_{1x}^{02} + M_{1y}^{02}}{2} + M_{1y}^0 a_+ \omega_+^2 - M_{1x}^0 a_- \omega_-^2 \right. \\
&\quad \left. + (A_+^2 + a_+^2) \frac{\omega_+^4}{2} + (A_-^2 + a_-^2) \frac{\omega_-^4}{2} \right] . & (4.30)
\end{aligned}$$

Comme dans l'exemple, on écrit maintenant les équations du mouvement associées à A_\pm et a_\pm . On en déduit l'expression des fréquences rapides (qui est bien celle attendue après l'étude de la matrice \mathbf{T}) :

$$\frac{d\phi_\pm}{dt} = \omega_\pm , \quad (4.31)$$

des amplitudes :

$$A_\pm^2 = \frac{P_\pm}{N\omega_\pm^3} . \quad (4.32)$$

et des variables a_\pm :

$$a_+ = -\frac{M_{1y}^0}{\omega_+^2} \quad \text{et} \quad a_- = \frac{M_{1x}^0}{\omega_-^2} . \quad (4.33)$$

On peut maintenant réintroduire toutes ces expressions dans le Hamiltonien (4.30), et on obtient finalement le Hamiltonien effectif cherché, qui ne dépend que des variables lentes, et contient les invariants adiabatiques comme paramètres :

$$H_{eff} = \sum_i \frac{p_i^{02}}{2} + P_+ \omega_+ + P_- \omega_- . \quad (4.34)$$

Une fois les invariants P_\pm déterminés par les conditions initiales, la dynamique lente est en principe donnée par (4.34). Il faut néanmoins souligner une difficulté lorsque $|\mathbf{M}_2|$ s'annule : dans ce cas, les deux valeurs propres ω_\pm se croisent. ψ , qui est définie comme la moitié de la phase du nombre complexe \mathbf{M}_2 , subit une discontinuité de $\pi/2$ en ce point, et les deux vecteurs propres \mathbf{X}_\pm (définis en fonction de ψ , voir l'étude de la matrice \mathbf{T}) s'intervertissent. La figure 4.4 présente une illustration numérique de ce phénomène.

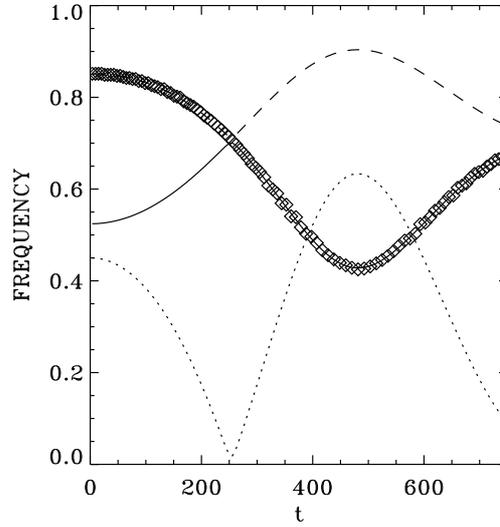


Figure 4.4: Les losanges représentent l'évolution de la fréquence, obtenu numériquement, en fonction du temps, pour $N = 200$ particules à partir du Hamiltonien d'origine (4.1). Nous avons ajouté ω_+ (ligne tiretée) et ω_- (ligne continue), ainsi que $|M_2|$ (ligne pointillée). On voit clairement l'échange des deux fréquences lorsque $|M_2|$ passe par zéro.

En fait, dans la plupart des expériences numériques de l'article [39] que nous voulons expliquer, une seule des fréquences rapides était excitée. En effet, les auteurs utilisent des conditions initiales avec $|M_2| = 0$; les deux valeurs propres de \mathbf{T} , sont alors dégénérées, et le système choisit le vecteur propre \mathbf{X}_- . Dorénavant, nous allons donc nous concentrer sur le cas où seule ω_- est excitée, et négliger ce phénomène de croisement de valeurs propres. En laissant de côté l'indice $-$ pour P et le 0 pour les p_i , on obtient

$$H_{eff} = \sum_i \frac{p_i^2}{2} + P \sqrt{\frac{1 - |M_2|}{2}} . \quad (4.35)$$

Avant de passer à l'étude détaillée du Hamiltonien effectif (4.35), on peut vérifier que les expressions obtenues pour la fréquence et l'amplitude des oscillations rapides (équations (4.31) et (4.32)) sont valables en comparant avec les simulations numériques; la figure 4.5 montre que l'accord est parfait. Nous sommes maintenant prêts pour étudier la formation aux temps courts du double amas et les structures en chevron de la figure 4.1 (partie 4.3); puis une étude statistique du Hamiltonien (4.35) nous permettra de comprendre la stabilité et les propriétés de la structure dans sa phase quasi stationnaire (partie 4.4).

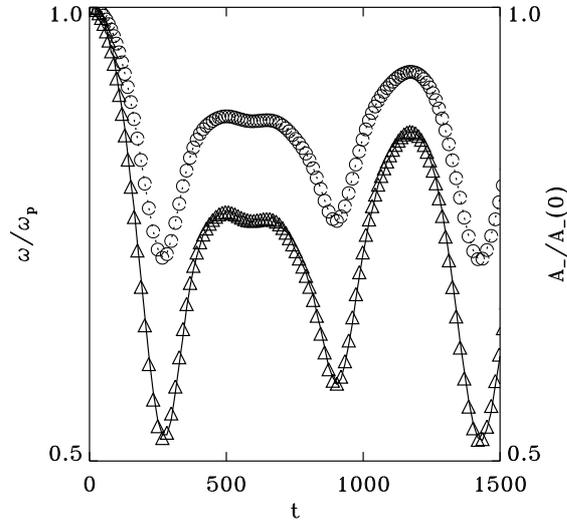


Figure 4.5: Comparaison entre les fréquences d’oscillation numérique (triangles) et analytique (ligne continue pour ω_- donné par l’équation (4.22)). Même comparaison pour les amplitudes d’oscillation : résultats numériques (cercles) et analytiques (ligne pointillée donnée par l’équation (4.32)). La simulation est faite pour une énergie par particule $e = 3.6 \cdot 10^{-5}$, jusqu’au temps $t_{\max} = 1500$. La fréquence est renormalisée par $\omega_p = \sqrt{2}/2$ et l’amplitude par $A_-(0)$, sa valeur à $t = 0$.

4.3 Formation de la structure aux temps courts

Nous utilisons dans cette partie la description effective de la dynamique pour donner une explication complète de la dynamique du système aux temps courts, c’est-à-dire notamment l’apparition récurrente des “chevrons”, visibles sur la figure 4.1.

Pour comparer la dynamique de (4.34) avec la dynamique du Hamiltonien complet, il faut analyser les conditions initiales utilisées pour attribuer à P la bonne valeur. On considère les conditions initiales les plus utilisées dans [39] : une distribution aléatoire des positions sur le cercle (avec $\mathbf{M}_1 \simeq 0$, donc une énergie très faible), et des vitesses initiales nulles. Dans ce cas, toute l’énergie est contenue initialement dans les petites oscillations d’amplitude εA , ce qui correspond à une énergie par particule à $\tau = 0$: $e = \varepsilon^2 A^2 \omega^2 / 2$ (on calcule l’énergie associée aux petites oscillations à partir de l’expression (4.25)). ε est défini jusqu’ici à une constante près, que l’on choisit de telle sorte que $A^2 \omega^2 = 1$, pour ces conditions initiales : ceci revient à poser $\varepsilon = \sqrt{2e}$. De l’équation (4.32) à $\tau = 0$, on tire alors $P/N = 1/\sqrt{2}$.

Les équations du mouvement correspondant au Hamiltonien (4.35) sont

$$\ddot{\theta}_i = -\frac{P}{N\sqrt{2}\sqrt{1-|\mathbf{M}_2|}} \sin 2(\theta_i + \psi). \quad (4.36)$$

Elles couplent toujours de façon non linéaire les N particules, et ne peuvent pas être résolues exactement. La figure 4.6 compare leur solution numérique avec la résolution directe des équations du mouvement complètes (4.2) : en dehors de la suppression des

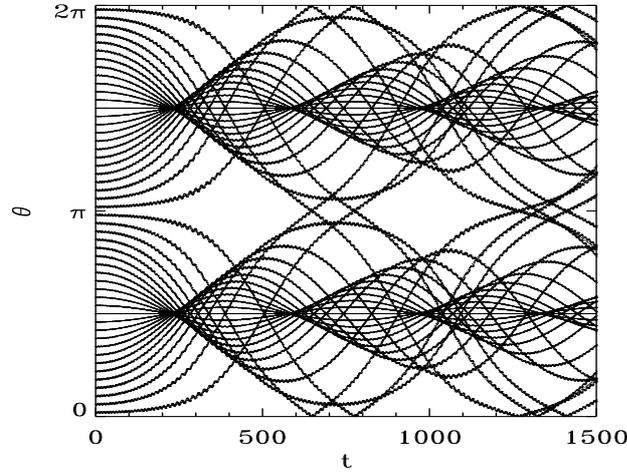


Figure 4.6: Évolution spatio-temporelle de $N = 50$ particules. Les résultats du Hamiltonien initial (4.1) et du Hamiltonien effectif (4.35) sont superposés : ils sont presque indiscernables, sauf en ce qui concerne les petites oscillations. On a pris comme condition initiale des particules parfaitement équidistribuées sur le cercle, avec une vitesse initiale $p_i \propto \sin \theta_i$. L'énergie par particule est $e = 2.5 \cdot 10^{-5}$.

oscillations rapides, les deux jeux de trajectoires sont presque indiscernables. Pour les temps courts, le système est donc parfaitement décrit par le Hamiltonien (4.34). On voit aussi sur la figure 4.6 apparaître les régions de forte densité correspondant aux chevrons : ce sont des zones où les trajectoires des particules se concentrent. Plus précisément ce sont les *enveloppes* des trajectoires des particules : c'est à dire les courbes auxquelles toutes les trajectoires des particules sont tangentes (ce sont les analogues des caustiques en optique).

Pour décrire quantitativement ces enveloppes, remarquons que si une solution exacte des équations du mouvement (4.36) est hors de portée, on peut donner une approximation pour les temps courts. En effet, tant que \mathbf{M}_2 est presque constant, ces équations (4.36) sont celles de N trajectoires de pendule simple, avec différentes conditions initiales. Puisque les simulations de [39], comme indiqué plus haut, sont effectuées avec des particules initialement distribuées au hasard sur le cercle, nous allons donc faire dans ce qui suit l'approximation $\mathbf{M}_2 = 0$. On est donc ramené à l'étude d'une famille de trajectoires définies par l'équation

$$\frac{d^2\theta}{d\tau^2} + \frac{1}{2} \sin(2\theta) = 0, \quad (4.37)$$

et par N conditions initiales avec vitesse nulle et répartition quasi homogène des angles sur $[0, 2\pi]$. La figure 4.3 compare cette famille de trajectoires pendulaires (4.37) avec la dynamique du Hamiltonien d'origine (4.30). L'accord qualitatif reste correct.

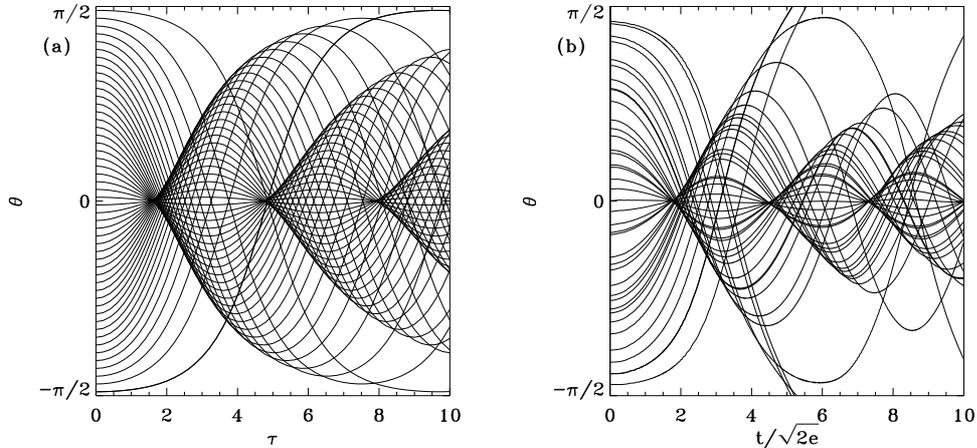


Figure 4.7: (a) : Trajectoires des particules dans le potentiel supposé constant $V = -(\cos 2\theta)/4$. (b) : Trajectoires des particules évoluant selon le vrai Hamiltonien (4.30). On peut voir les trois premières apparitions des chevrons. Deux phénomènes ne sont pas reproduits par les trajectoires simplifiées de la figure (a) : les petites oscillations, à peine visibles sur la figure (b), qui ont été moyennées, et la présence de particules non piégées, au voisinage des maxima du potentiel $V = -(\cos 2\theta)/4$.

4.3.1 Divergence de la densité

Les oscillations dans le potentiel effectif $V(\theta) = -(\cos 2\theta)/4$ que ressentent les particules sont quasi harmoniques au voisinage des deux minima. Comme toutes les particules ont une vitesse initiale nulle, ceci implique que toutes les trajectoires voisines des minima vont passer exactement à ce minimum au même moment, créant une divergence de la densité en ces points et à ce temps τ_d . τ_d se calcule facilement : c'est le quart de la période des oscillations quasi harmoniques au fond du puits; en unités de temps originales

$$t_d = \frac{\tau_d}{\varepsilon} = \frac{\pi}{\sqrt{8\varepsilon}}. \quad (4.38)$$

De plus, on peut prévoir que cette divergence apparaît périodiquement, chaque demi période : $t_n = (2n - 1)t_d$. L'accord quantitatif avec les simulations du modèle complet est bon pour la première divergence : voir la figure 4.8. On peut néanmoins remarquer que les divergences ont tendance par la suite à survenir un peu plus tôt dans le modèle complet : ceci est dû au fait que nous avons négligé la variation de \mathbf{M}_2 , nous reviendrons sur ce point au paragraphe suivant.

4.3.2 Description des caustiques

Après chaque divergence de la densité décrite au paragraphe précédent naissent les chevrons, ou enveloppes des trajectoires. Nous décrivons maintenant leur forme. Nous paramétrons chaque trajectoire $\theta(\tau)$ solution de l'équation (4.37) par l'angle initial θ_0 . La famille des trajectoires est donc décrite par une équation $F(\tau, \theta, \theta_0) = 0$;

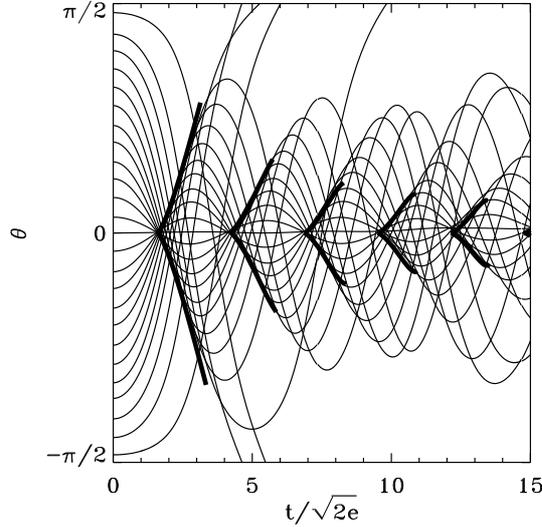


Figure 4.8: Superposition des enveloppes calculées avec les trajectoires de 50 particules réparties de façon homogène entre 0 et 2π à $t = 0$, pour une énergie par particule $e = 3 \cdot 10^{-6}$. L'intervalle entre deux apparitions de chevrons a été réduit d'un facteur 0.84 par rapport à sa valeur théorique (voir texte); le premier chevron n'a pas été déplacé.

F peut être exprimée à l'aide de fonctions elliptiques. L'enveloppe de cette famille de trajectoires dans le plan (τ, θ) est donc donnée par les deux équations (la méthode pour calculer l'équation de l'enveloppe est standard [116]) :

$$F(\tau, \theta, \theta_0) = 0 \quad (4.39)$$

$$\frac{\partial F}{\partial \theta_0}(\tau, \theta, \theta_0) = 0. \quad (4.40)$$

Ces équations ne sont bien sûr pas solubles telles quelles, nous allons donc utiliser une approximation de la fonction F au voisinage des naissances des chevrons, pour $\tau = \tau_d$. Nous remplaçons la véritable trajectoire des particules par une droite :

$$\theta(\tau, \theta_0) = v_0(\tau - \tau_0), \quad (4.41)$$

où τ_0 est le temps auquel la trajectoire commençant en θ_0 passe au fond du puits, $\theta = 0$; bien sûr, τ_0 tend vers τ_d quand θ_0 tend vers 0. v_0 est la vitesse de la particule lorsqu'elle passe en $\theta = 0$. τ_0 et v_0 sont des fonctions de θ_0 . τ_0 est le quart de la période des oscillations d'une particule d'angle initial θ_0 dans le potentiel $V = (1 - \cos 2\theta)/4$; les tables nous donnent $\tau_0 = K_0 = K(\sin \theta_0)$, où K est l'intégrale elliptique complète du premier type. La conservation de l'énergie pour chaque particule (nous sommes toujours dans l'approximation $\mathbf{M}_2 = cste$) permet d'obtenir $v_0 = \sin \theta_0$.

Les équations (4.39) et (4.40) qui donnent l'enveloppe se simplifient alors en

$$\theta(\tau, \theta_0) = v_0(\tau - \tau_0) \quad (4.42)$$

$$0 = v_0'(\tau - \tau_0) - v_0\tau_0'. \quad (4.43)$$

où les $'$ symbolisent la dérivation par rapport à θ_0 . On tire facilement de là une équation paramétrique pour l'enveloppe, c'est à dire τ et θ en fonction de θ_0 , valable pour θ_0 petit (en effet les θ_0 plus grands contribuent à l'enveloppe plus loin du fond du puits, et l'approximation (4.41) n'est plus valable). On voit aussi qu'il est facile d'obtenir un développement limité en θ_0 , en considérant des améliorations de (4.41) prenant en compte les dérivées suivantes, et en développant K_0 à l'ordre correspondant en θ_0 . Cette procédure donne

$$\tau = \tau_s + 3a\theta_0^2 + \left(5b + \frac{2a}{3}\right)\theta_0^4 + \left(7c + \frac{4b}{3} + \frac{4a}{15} - \frac{8a^3}{3}\right)\theta_0^6 + O(\theta_0^8) \quad (4.44)$$

$$\theta = 2a\theta_0^3 + \left(4b + \frac{a}{3}\right)\theta_0^5 + \left(6c + \frac{2b}{3} + \frac{31a}{180} - 4a^3\right)\theta_0^7 + O(\theta_0^9), \quad (4.45)$$

avec $(a, b, c) = \frac{\pi}{8} \left(1, \frac{11}{48}, \frac{173}{2880}\right)$. Il est facile d'obtenir aussi l'équation pour les chevrons successifs : ils ont leur origine pour le $n^{\text{ème}}$ chevron en $\tau_n = (2n - 1)\tau_d$; il suffit donc de remplacer dans l'équation (4.41) $\tau_0 = K_0$ par $(2n - 1)K_0$. À l'ordre le plus bas, on trouve ainsi

$$\theta = 2a(2n - 1) \left(\frac{\tau - \tau_n}{3a(2n - 1)} \right)^{3/2} \quad (4.46)$$

donc

$$\theta \propto \frac{(t - (2n - 1)t_s)^{3/2}}{\sqrt{2n - 1}}. \quad (4.47)$$

Le facteur $1/\sqrt{2n - 1}$ explique le rétrécissement des chevrons au fur et à mesure que le temps augmente. La figure 4.8 montre que les équations (4.44), (4.45) et leurs généralisations aux n successifs décrivent très bien le processus de formation des chevrons.

En réalité, les intervalles entre deux occurrences des chevrons ont été réduits sur la figure 4.8 d'un facteur 0.84 par rapport à leur valeur théorique de $2\tau_d$: c'est la signature du fait que \mathbf{M}_2 est en réalité non nul, et que par conséquent, les particules n'évoluent pas dans un potentiel $V = (1 - \cos 2\theta)/4$, mais $V = (1 - \cos 2\theta)/(4\sqrt{1 - |\mathbf{M}_2|})$. Lorsque \mathbf{M}_2 est non nul, le potentiel devient plus raide, et la fréquence des oscillations augmente; une réduction du temps τ_d d'un facteur 0.84 correspond à $|\mathbf{M}_2| \simeq 0.5$, ce qui est précisément la valeur observée dans les simulations.

4.3.3 Lien avec la classification des singularités

On peut se demander si cette formation de structures par l'intermédiaire de caustiques, ainsi que sa forme caractéristique en chevrons, est générale, ou bien étroitement liée au modèle particulier que nous étudions ici. Une fois encore, la théorie des singularités nous aide à répondre à la question. Considérons un ensemble de particules en une dimension, qui initialement ont une dispersion en vitesses très faible, si bien qu'on peut caractériser le système par une densité de particules $\rho(x)$ et par un champ de vitesses $v(x)$. Supposons que les particules évoluent librement, sans collision, en conservant leur vitesse initiale; alors la position au temps t d'une particule initialement

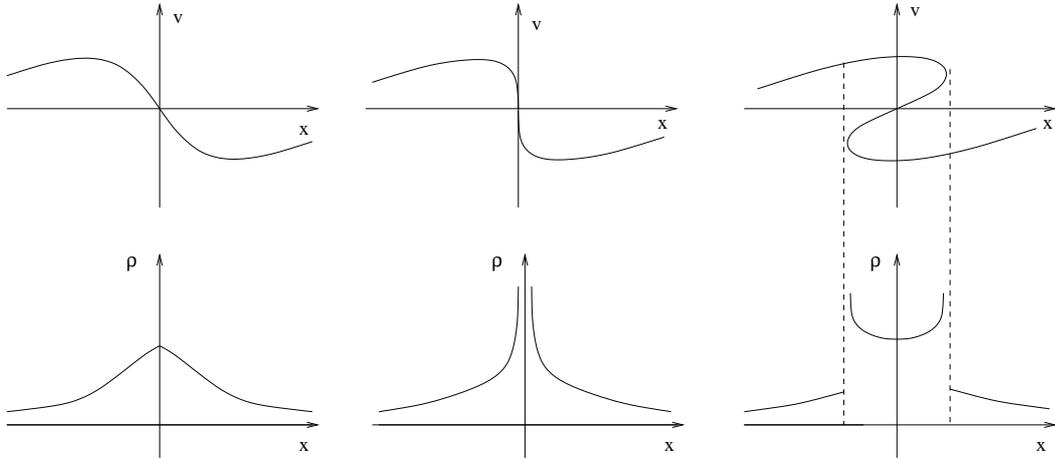


Figure 4.9: Évolution du profil de vitesse et de la densité. La première divergence apparaît au point d'inflexion, avec $v' < 0$; de gauche à droite : avant la première divergence, au moment de la divergence (singularité A_3 selon la classification d'Arnold [10]), et après celle-ci (deux singularités A_2).

en x_0 est

$$x = x_0 + tv(x_0) . \quad (4.48)$$

Les particules plus rapides vont rattraper les plus lentes, si bien que le profil de vitesse se raidit, et la densité augmente aux points où le profil est raide : figure 4.9. Plus précisément, en utilisant la conservation de la masse, on en déduit la densité de particules au point x , au temps t :

$$\rho(x, t) = \rho(x_0) \frac{dx_0}{dx} , \quad (4.49)$$

où x_0 se calcule en fonction de x et t en inversant l'équation (4.48). On peut expliciter le terme dx_0/dx :

$$\rho(x, t) = \frac{\rho(x_0)}{1 + tv'(x_0)} . \quad (4.50)$$

On voit alors que la densité diverge d'abord pour x_0 tel que $v'(x_0)$ est minimum et négatif; on peut calculer facilement que $\rho(x) \sim |x|^{-2/3}$ autour de cette divergence. Puis le profil de vitesse devient trivalué, et la densité diverge autour des deux points de rebroussement du profil; la divergence est plus faible, $\rho(x) \sim |x|^{-1/2}$. Enfin, la position de ces deux nouvelles divergences est caractérisée par le même exposant $3/2$ que nous avons trouvé pour les chevrons, à l'équation (4.46). La situation est résumée par la figure 4.9. Ces résultats sont obtenus dans une situation très simple, de particules indépendantes, avec une vitesse constante. Cependant, et c'est là l'intérêt et la puissance de la théorie des singularités, il est possible de montrer qu'une situation plus compliquée, par exemple la présence d'un potentiel, ou d'un potentiel créé par les particules elles-mêmes, ne change pas qualitativement les conclusions (référence [10], p. 31); ainsi, la formation des deux types de singularités, les exposants sont universels. Ce sont bien ceux que l'on retrouve dans notre problème.

Notons pour finir que la formation de ces singularités est liée à notre condition initiale sans dispersion locale en vitesse : nous avons défini un champ de vitesse $v(x)$. La présence d'une "température locale" estompe les divergences, jusqu'à les faire disparaître lorsqu'elle devient trop importante [102]. La formation de ces structures spatiales par l'intermédiaire de caustiques a été étudiée en détail par Zeldovich et Shan-darin dans le contexte astrophysique [102] : ce mécanisme pourrait être à l'origine des inhomogénéités de densité à grande échelle dans l'univers (en trois dimensions, d'autres structures que A_2 et A_3 sont possibles [10]).

4.3.4 Estimation de la distribution angulaire

Simplifier, comme on l'a fait, la dynamique effective, en supposant $|\mathbf{M}_2| = 0$, peut sembler grossier, mais cela est tout à fait suffisant pour décrire le chevrons, et donne également des résultats intéressants en ce qui concerne la distribution de probabilité des particules $\rho(\theta)$. Les auteurs de [39] ont trouvé que la distribution, après une évolution initiale rapide, se stabilisait, en très bonne approximation, autour de

$$\rho_1(\theta) = \frac{1}{2\pi} (1 - \log(2|\sin \theta|)) , \quad (4.51)$$

sans avoir aucune théorie conduisant à ce résultat.

L'approximation du potentiel $V = (1 - \cos 2\theta)/4$ permet d'obtenir un résultat similaire, notamment autour du fond du puits. En effet, puisque les oscillations sont anharmoniques, au bout d'un certain temps, tout effet de cohérence entre les particules, même de conditions initiales proches, aura disparu (sauf très près du fond du puits où les oscillations sont quasi isochrones); on peut donc estimer la densité $\rho(\theta)$ en additionnant les contributions de chaque trajectoire. Le long de la trajectoire d'une particule, on a

$$\dot{\theta} = \sqrt{\frac{\cos 2\theta - \cos 2\theta_0}{2}} . \quad (4.52)$$

Or le temps $dt(\theta)$ passé par une particule autour d'une position θ est inversement proportionnel à sa vitesse $\dot{\theta}(\theta, \theta_0)$ en ce point. Par conséquent, la trajectoire paramétrée par sa condition initiale θ_0 donne une contribution à la densité

$$\rho_{\theta_0}(\theta)d\theta = \frac{dt(\theta)d\theta}{\int_{-\theta_0}^{\theta_0} dt(\theta)} = \frac{2\sqrt{2} d\theta}{T(\theta_0)\sqrt{\cos 2\theta - \cos 2\theta_0}} , \quad (4.53)$$

où $T(\theta_0)$ est la période de la trajectoire. On calcule alors la densité de probabilité en additionnant toutes ces contributions, en tenant compte du fait que la distribution initiale est homogène en θ_0 . On obtient

$$\rho(\theta) \propto \int_{2\theta}^{\pi} \frac{2\sqrt{2} du}{T(u)\sqrt{\cos 2\theta - \cos 2u}} , \quad (4.54)$$

pour $\theta \in [0, \pi/2]$; le reste se déduit par symétrie et π -périodicité. Cette distribution reproduit la divergence logarithmique de la densité au fond du puits donnée par (4.51),

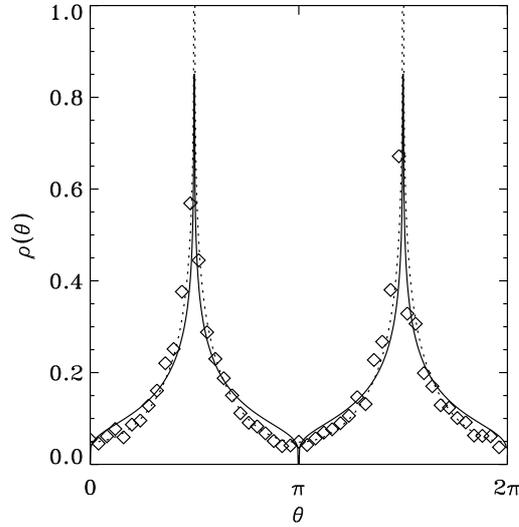


Figure 4.10: Distribution $\rho(\theta)$ pour les temps “intermédiaires” (voir texte). Comparaison entre les résultats numériques (losanges) et la formule analytique (4.54) (ligne continue), pour $e \simeq 10^{-4}$. La ligne pointillée correspond à la formule (4.51).

et est en accord correct avec les résultats numériques pour des temps “moyennement longs” (c’est à dire de l’ordre de quelques centaines de périodes des oscillations au fond du puits) : voir figure 4.10. Le désaccord est important surtout au niveau des maxima du potentiel $V(\theta)$; c’était à prévoir si on tient compte du fait que V n’est pas constant mais oscille en réalité un petit peu.

La simplification de la dynamique effective en une collection de particules indépendantes se déplaçant dans un potentiel en double puits constant a donc permis de décrire quantitativement la dynamique aux temps courts (formation et récurrence des chevrons), ainsi qu’en partie aux temps intermédiaires (densité $\rho(\theta)$). Néanmoins, il est clair que l’on peut faire mieux aux temps intermédiaires et longs en considérant la dynamique effective complète. Au paragraphe suivant, nous verrons que cela nous permet d’étudier le comportement du double amas aux temps longs, d’expliquer la valeur $|\mathbf{M}_2| \simeq 0.5$ obtenue dans [39] pour diverses conditions initiales, ainsi que la relation température-énergie modifiée $T \simeq 1.3e$.

4.4 Comportement du double amas aux temps longs

La dynamique effective est auto-cohérente : le potentiel dans lequel évoluent les particules dépend de l’état global du système, par l’intermédiaire de \mathbf{M}_2 . Aux temps longs, il est clair que l’approximation d’une collection de trajectoires indépendantes dans un potentiel fixe ne sera pas suffisante : les particules peuvent échanger de l’énergie par l’intermédiaire du potentiel fluctuant. Pour étudier le comportement du double amas aux temps longs, nous allons donc commencer par une étude statistique détaillée du Hamiltonien effectif (paragraphe 4.4.1), avant de comparer les résultats avec les simu-

lations numériques du modèle initial. C'est l'un des avantages de la procédure utilisée pour obtenir la dynamique effective : puisqu'elle est formulée de manière hamiltonienne, elle se prête facilement à une étude statistique.

4.4.1 Étude du Hamiltonien effectif

Mécanique statistique du Hamiltonien effectif L'énergie potentielle du Hamiltonien effectif (4.35) ne dépend que de la variable globale \mathbf{M}_2 ; la méthode des grandes déviations pour calculer les propriétés d'équilibre, telle qu'elle est décrite dans la première partie de cette thèse, s'applique donc sans problème. On verra que pour le Hamiltonien effectif (4.35), les ensembles canonique et microcanonique sont équivalents; on calculera néanmoins dans l'ensemble microcanonique, car cela facilite les comparaisons avec les simulations numériques, effectuées à énergie constante. Pour plus de clarté, détaillons les différentes étapes de la résolution par la méthode des grandes déviations.

Étape1 :

Il y a deux quantités conservées : l'énergie et la quantité de mouvement totale. En fait, une quantité de mouvement totale non nulle crée une rotation globale du système, découplée du reste de la dynamique; on peut donc sans restriction considérer la quantité de mouvement totale nulle, et ne s'intéresser qu'à l'énergie. Le système étant invariant par rotation, toutes les orientations de \mathbf{M}_2 sont équivalentes; on supposera donc que $|\mathbf{M}_2| = (\sum \cos 2\theta_i) / N$, et on posera $M_2 = |\mathbf{M}_2|$.

La variable globale est alors facile à choisir : on prendra $\mu = (u, M_2)$, où $u = (\sum p_i^2) / N$. L'énergie s'écrit alors

$$H_{eff} = N \left(\frac{u}{2} + \tilde{P} \sqrt{\frac{1 - M_2}{2}} \right), \quad (4.55)$$

avec $\tilde{P} = P/N$.

Étape2 :

L'entropie associée à une valeur de u est comme d'habitude donnée par (2.36) :

$$s_{cin}(u) = \frac{1}{2} \ln u. \quad (4.56)$$

Pour calculer l'entropie configurationnelle associée à M_2 , il faut utiliser le théorème de Cramér :

$$\begin{aligned} \Psi(\lambda) &= \ln \langle e^{\lambda \cos 2\theta_i} \rangle \\ &= \ln I_0(\lambda), \end{aligned} \quad (4.57)$$

où I_0 est la fonction de Bessel modifiée d'ordre 0. Ce qui donne pour l'entropie

$$s_{conf}(M_2) = - \max_{\lambda} (\lambda M_2 - \ln I_0(\lambda)). \quad (4.58)$$

Étape 3 :

Pour obtenir la configuration d'équilibre à une énergie E , pour une valeur P de l'invariant adiabatique, il suffit de résoudre le problème variationnel

$$s(E, P) = \max_{M_2} \left(\frac{1}{2} \ln \left(\frac{P}{N} \right) + \frac{1}{2} \ln \left(\frac{2E}{P} - \sqrt{\frac{1 - M_2}{2}} \right) + s_{conf}(M_2) \right). \quad (4.59)$$

La valeur d'équilibre M_2^* ne dépend donc que du rapport E/P . On voit bien sur l'équation (4.59) la compétition entre les deux termes : M_2 grand est défavorable pour l'entropie configurationnelle, mais permet une plus grande énergie cinétique, favorable pour l'entropie s_{cin} .

La solution de ce problème variationnel est donnée par la ligne continue de la figure 4.12. Il n'y a pas de transition de phase : M_2^* est non nul quelle que soit l'énergie; néanmoins, M_2^* tend vers 0 pour les très grandes valeurs du rapport E/P : les particules sont presque libres dans cette limite.

Une fois que la valeur d'équilibre M_2^* est connue, il est facile de compléter la description de l'équilibre statistique. La température est donnée par

$$T = \langle u \rangle = \frac{2}{N} \left\langle E - P \sqrt{\frac{1 - |M_2|^*}{2}} \right\rangle. \quad (4.60)$$

La distribution des vitesses est alors une Maxwellienne avec température T , et la distribution angulaire une Boltzmannienne $\rho(\theta) \propto e^{-V(\theta)/T}$, avec le potentiel

$$V(\theta) = \frac{P}{N2\sqrt{2}\sqrt{1 - |M_2|^*}} (1 - \cos(2\theta + 2\psi)). \quad (4.61)$$

Relaxation à l'équilibre du Hamiltonien effectif Nous avons maintenant à disposition une description complète de l'équilibre statistique du Hamiltonien effectif. On voudrait bien sûr en tirer des conclusions concernant le modèle de départ, HMF antiferromagnétique. Malheureusement, le Hamiltonien effectif est un Hamiltonien de type champ moyen, et, comme indiqué en introduction, la dynamique de ces systèmes, dans certaines conditions, ne relaxe que très lentement vers l'équilibre. Nous allons donc commencer par étudier les échelles de temps de la relaxation à l'équilibre du Hamiltonien effectif, avant de les comparer avec celles du modèle complet. Ceci nous permettra de déterminer s'il est raisonnable d'espérer observer l'équilibre statistique du modèle effectif dans la dynamique réelle.

Les propriétés de relaxation à l'équilibre des systèmes avec interactions à longue portée (dont le Hamiltonien effectif ou le système complet sont des exemples) posent un problème important en lui-même. Nous nous concentrons ici sur le cas du Hamiltonien effectif (4.34), et approcherons le problème de manière plus générale au chapitre suivant 5.

La figure 4.11 illustre l'approche à l'équilibre de la dynamique effective, pour trois tailles différentes du système, avec comme conditions initiales des vitesses nulles et

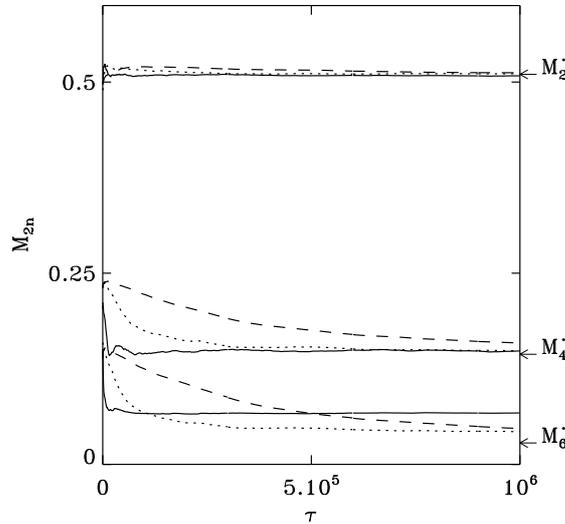


Figure 4.11: Relaxation à l'équilibre des trois premiers moments pairs (moyennés sur le temps) $\langle M_{2n} \rangle$, à partir de simulations numériques de la dynamique effective (4.35). La ligne continue (respectivement pointillée et tiretée) correspond aux résultats pour $N = 200$ (resp. $N = 800$ et $N = 3200$). $\langle M_2 \rangle$, $\langle M_4 \rangle$ et $\langle M_6 \rangle$ sont représentés de haut en bas. Ils convergent aux temps longs vers les valeurs d'équilibre $M_2^* = 0.510$, $M_4^* = 0.144$ et $M_6^* = 0.028$, corrigées par les effets de taille finie.

des angles distribués de façon homogène sur le cercle (ceci correspond aux conditions initiales typiques utilisées dans [4, 39]). On définit les moments successifs de la distribution angulaire

$$M_k = \left| \frac{1}{N} \sum_j e^{ik\theta_j} \right|. \quad (4.62)$$

Nous avons représenté $\langle M_{2n} \rangle(\tau)$, les moyennes temporelles de M_{2n} du temps initial 0 au temps τ , pour n égal à 1, 2 et 3. Les fluctuations temporelles sont donc éliminées. On observe d'abord des effets de taille finie : par exemple, pour le petit système $N = 200$, M_6 converge vers une valeur plus élevée que M_6^* , la valeur d'équilibre à N infini.

On observe aussi que M_2 atteint rapidement sa valeur d'équilibre, quelle que soit la taille N du système, alors que le temps de relaxation de M_4 et M_6 augmente rapidement avec N (on n'étudiera pas plus précisément ici cette dépendance, qui rappelle beaucoup les phénomènes observés dans d'autres modèles comme le HMF ferromagnétique, et qui vont être étudiés plus précisément au chapitre suivant 5).

En conclusion, on s'attend à ce que pour un grand système, *seules certaines caractéristiques* de l'équilibre statistique effectif soient observables. Plus précisément, on devrait observer la valeur d'équilibre M_2^* , ainsi que les grandeurs qui lui sont liées : la température (voir l'équation (4.60), ou la capacité calorifique. En revanche, la relaxation des moments d'ordre plus élevé est très lente; la distribution des particules $\rho(\theta)$ (de même que la distribution des vitesses) doit donc significativement différer de

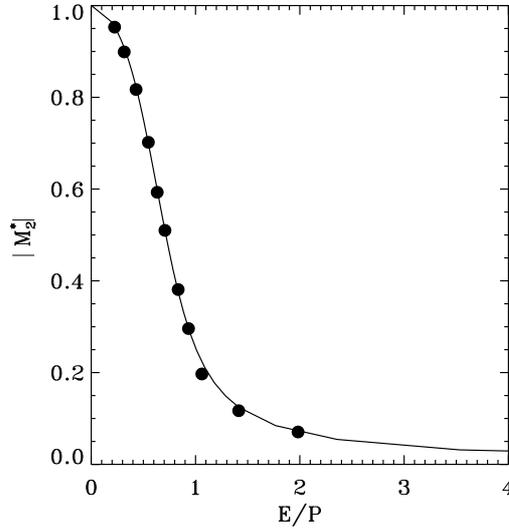


Figure 4.12: Prédiction statistique de la valeur de M_2 en fonction du rapport E/P entre l'énergie du Hamiltonien effectif et l'invariant adiabatique. La ligne continue montre la prédiction analytique, et les cercles correspondent aux résultats numériques pour le modèle complet (4.1).

la Maxwellienne prévue par la mécanique statistique d'équilibre : on a déjà vu que c'était le cas, sur la figure 4.10.

4.4.2 Comparaison avec le Hamiltonien complet

Nous allons maintenant utiliser l'étude précédente du Hamiltonien effectif pour comprendre la dynamique du modèle complet de départ, et expliquer les expériences numériques de [4, 39].

Comme on vient de le voir, la valeur du moment M_2 dans le modèle complet doit être donnée par sa valeur d'équilibre dans le Hamiltonien effectif. Les conditions initiales utilisées dans [39] pour la simulation du modèle original (vitesses nulles, distribution aléatoire, quasi homogène des particules) correspondent à un rapport $E/P = \sqrt{2}/2$. D'après la figure 4.12, cela donne une valeur d'équilibre $M_2^* \simeq 0.51$; il est à noter que cette valeur ne doit pas dépendre de l'énergie par particule dans le système complet (à condition que celle-ci reste petite, pour assurer la séparation des échelles de temps) : tout ceci est en parfait accord avec les résultats numériques de [39].

Il est aussi possible d'expliquer maintenant la relation modifiée $T \simeq 1.3e$, également rapportée dans [39]. La mécanique statistique standard conduit en effet à la relation $T = 2e$. Cependant, en présence des petites oscillations, il faut revoir l'application du principe d'équipartition de l'énergie. L'énergie totale du système complet se divise en trois parties : l'énergie potentielle des petites oscillations, l'énergie cinétique des petites oscillations, et l'énergie cinétique du mouvement lent. Les deux premières

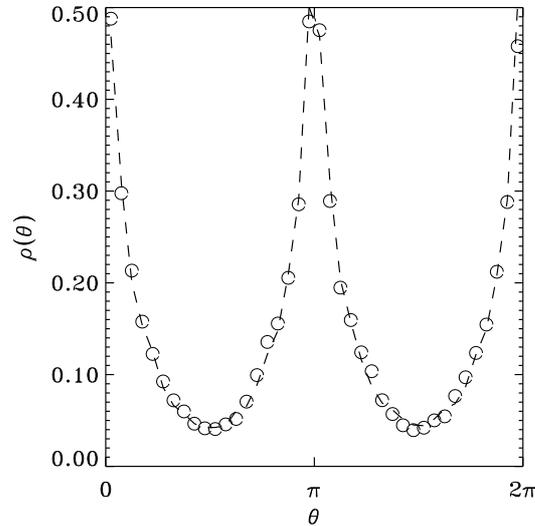


Figure 4.13: Comparaison de la distribution de probabilité angulaire des particules $\rho(\theta)$ obtenue d'une part avec le Hamiltonien d'origine (4.1) (cercles), d'autre part avec le Hamiltonien effectif (4.35) (ligne tiretée). Les courbes ont été calculées pour $N = 10^3$ particules, et sont moyennées sur des temps intermédiaires correspondant à $\tau = 10^3, \dots, 10^4$. L'énergie dans le Hamiltonien d'origine est $e = 2.5 \cdot 10^{-5}$.

parties sont égales en moyenne et forment l'énergie potentielle du Hamiltonien effectif; la troisième partie est l'énergie cinétique du Hamiltonien effectif. En utilisant ces remarques, le fait que la température du système complet est donnée par deux fois l'énergie cinétique du système complet, et la valeur d'équilibre $M_2^* \simeq 0.51$, on obtient précisément $T = 1.3e$.

Toutes les simulations de [4, 39] sont faites avec des conditions initiales correspondant pour le Hamiltonien effectif à un rapport $E/P = \sqrt{2}/2$. On peut choisir des conditions initiales de façon à faire varier ce rapport (par exemple en choisissant M_2 non nul à $t = 0$), et ainsi tester la validité de toute la courbe $M_2^*(E/P)$. La figure 4.12 montre que l'accord est excellent pour toutes les conditions initiales testées.

En ce qui concerne la distribution angulaire aux temps longs $\rho(\theta)$, on a vu au paragraphe précédent qu'il n'était pas possible d'espérer la décrire par la distribution d'équilibre, la Maxwellienne. On peut néanmoins essayer d'améliorer le résultat de l'équation (4.54), en comparant la densité $\rho(\theta)$ obtenue au temps t par la dynamique complète, avec la densité $\rho_{eff}(\theta)$ obtenue par la dynamique effective au temps τ/ε . Le résultat est montré sur la figure 4.13 : les deux densités coïncident parfaitement, ce qui montre que les structures hors équilibre de la dynamique complète sont parfaitement décrites par la dynamique effective.

En conclusion, la description du système à l'aide du Hamiltonien effectif a été très fructueuse, et a permis d'expliquer presque toutes les observations numériques

surprenantes des articles [4, 39] : la formation et la récurrence des chevrons aux temps courts, la stabilisation à une valeur asymptotique de M_2 différente de la valeur d'équilibre statistique, la distribution de probabilité angulaire $\rho(\theta)$... Certains points ne sont pas totalement éclaircis. Ainsi, toutes les conditions initiales de faible énergie ne conduisent pas à la formation du double amas : par exemple, si la distribution initiale des particules est aléatoire et les vitesses nulles, le double amas se forme; en revanche, si la distribution initiale est parfaitement homogène et les vitesses aléatoires, le double amas ne se forme pas. À partir de notre analyse, nous pouvons dire que le double amas se forme dès que les oscillations rapides sont présentes et prennent une part extensive de l'énergie; néanmoins, nous n'avons pas de description précise de cette classe de conditions initiales. Nous ne savons pas non plus comment se déstabilise le double amas lorsque ε se rapproche de 1. Enfin, nous avons laissé de côté le problème du croisement des deux valeurs propres non nulles, parce qu'il n'entraîne pas en jeu dans les cas qui nous occupaient; une étude plus détaillée de ce sujet (analogue au croisement de deux niveaux en mécanique quantique) pourrait être intéressante.

4.5 Conclusion

Le modèle HMF antiferromagnétique donne un exemple d'un phénomène général, que l'on trouve dans beaucoup de systèmes physiques : des oscillations très rapides interagissant avec un mouvement lent et modifiant ses propriétés; on peut citer le pendule de Kapitza, l'interaction des ondes de gravité rapides avec le mouvement vortical pour le modèle Shallow-Water en rotation [49], ou certains problèmes de physique des plasmas [5]. En outre la dynamique à N particules du modèle HMF antiferromagnétique peut être traitée complètement, jusqu'à obtenir un Hamiltonien effectif pour les degrés de liberté lents, et cette dynamique effective peut elle-même être étudiée en détail : c'est la tâche qu'on aimerait accomplir dans des cas plus compliqués et physiquement plus intéressants. En ce sens, le HMF antiferromagnétique est un bon modèle jouet qui nous permet de mettre en évidence l'intérêt de l'approche variationnelle dans la procédure de moyennisation.

Néanmoins, il ne semble pas que l'apparition d'oscillations rapides et de deux échelles de temps très bien séparées soit réellement générique parmi les systèmes avec interactions à longue portée : on trouve ce phénomène en présence d'interactions répulsives, comme en physique des plasmas par exemple, mais pas pour les systèmes auto-gravitants. En revanche, le phénomène de relaxation lente vers l'équilibre a été observé dans un très grand nombre de systèmes (y compris notre Hamiltonien effectif (4.35)), et semble être en un certain sens générique : voir l'introduction de cette partie II. Toujours dans l'idée de disposer de modèles jouets pour lesquels une analyse poussée est possible, nous allons étudier au chapitre suivant ce phénomène sur le modèle HMF ferromagnétique, qui a déjà été très souvent employé à cet effet dans la littérature.

Chapter 5

Relaxation lente à l'équilibre : étude du modèle HMF ferromagnétique

5.1 Introduction

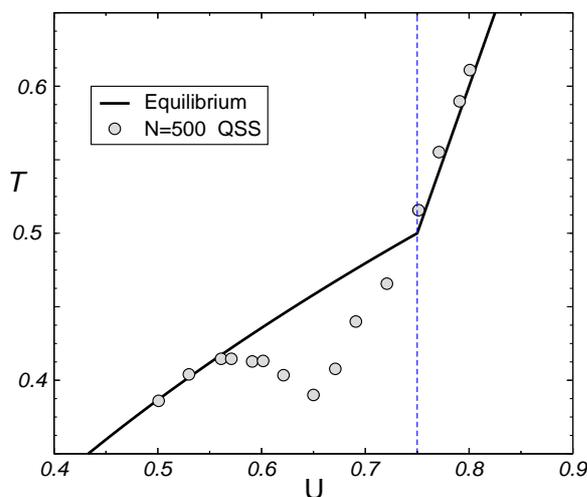


Figure 5.1: Modèle HMF ferromagnétique. La ligne continue est l'état d'équilibre canonique; la ligne tiretée verticale indique la position de la transition du deuxième ordre. Les symboles circulaires sont le résultat de simulations de dynamique moléculaire, à partir de conditions initiales hors équilibre, pour des temps d'intégration assez longs.

Le modèle HMF ferromagnétique ne diffère du modèle antiferromagnétique que par le signe de la constante de couplage :

$$H = \sum_j \frac{p_j^2}{2} - \frac{1}{2N} \sum_{i,j} \cos(\theta_i - \theta_j); \quad (5.1)$$

elle est maintenant négative, si bien qu'il est énergétiquement favorable pour le système de se "magnétiser", c'est à dire d'avoir $\mathbf{M}_1 = (\sum e^{i\theta_j})/N$ non nul (au contraire du cas

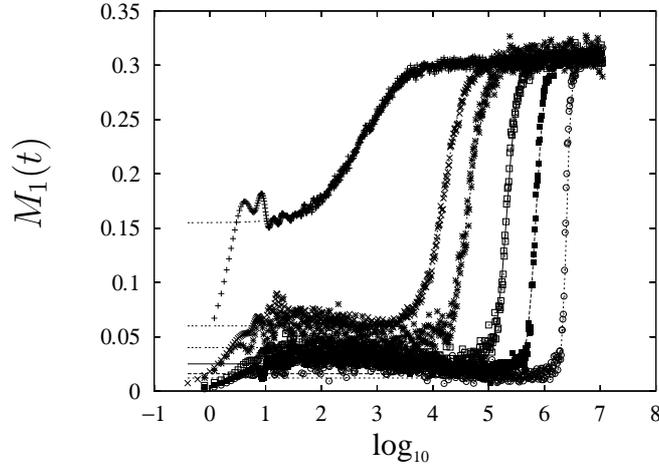


Figure 5.2: L'évolution temporelle de la magnétisation $M_1(t)$ pour $N = 100(1000)$, $1000(100)$, $2000(8)$, $5000(8)$, $10000(8)$ et $20000(4)$ de gauche à droite; entre parenthèses le nombre de réalisations sur lequel les courbes sont moyennées.

antiferromagnétique). La solution canonique est facile à calculer [4], et nous avons vu au paragraphe 2.3.2 que l'ensemble microcanonique lui était équivalent. Néanmoins, pour certaines conditions initiales, le système semble ne pas relaxer vers l'équilibre, voir la figure 5.1. Une étude plus précise montre que le système relaxe en fait vers l'équilibre, mais de plus en plus lentement à mesure que le nombre de particules augmente; pour N assez grand, on ne pourra donc pas voir l'approche de l'équilibre sur les simulations. La figure 5.2 illustre ce comportement : la magnétisation reste initialement proche de 0, et ne relaxe vers sa valeur d'équilibre $M_1 = |\mathbf{M}_1| \neq 0$ que pour des temps qui augmentent avec N .

Nous avons introduit au début de cette partie II l'équation de Vlasov, et suggéré qu'elle pouvait fournir un élément d'analyse utile et général pour comprendre la relaxation lente et les états quasi stationnaires. Plus précisément, suivant les idées de relaxations violente et collisionnelle, on peut s'attendre au scénario suivant pour l'évolution du système :

1. Le système évolue d'abord rapidement en suivant à peu près la dynamique de Vlasov, donc avec une échelle de temps indépendante du nombre de particules.
2. Le système reste piégé au voisinage de l'un des innombrables états stationnaires stables de l'équation de Vlasov (voir section 5.2). Si la théorie de la relaxation violente est parfaitement valable, cet état est l'équilibre statistique de l'équation de Vlasov; en réalité, il ne s'agit pas en général exactement de cet état d'équilibre.
3. Le système évolue ensuite lentement, sous l'effet des collisions ou, de façon équivalente, des fluctuations autour de cet état stationnaire stable : cette fois-ci, l'échelle de temps dépend clairement du nombre de particules. Au cours de cette évolution lente, on peut supposer que le système se déplace parmi les états stationnaires stables de l'équation de Vlasov.

4. Enfin, le système atteint un état stationnaire stable de Vlasov particulier : l'équilibre statistique complet (différent de l'équilibre statistique de l'équation de Vlasov, car les Casimirs ne sont pas conservés lors de la relaxation collisionnelle).

Bien sûr, le système pourrait rester piégé pour d'autres raisons que les états stationnaires stables invoqués ici et décrits dans l'introduction de cette partie II; l'étude du HMF antiferromagnétique en a donné un exemple : le piégeage résulte dans ce cas de la présence de deux échelles de temps bien séparées. Néanmoins, en l'absence d'autres informations, ces états stationnaires stables sont l'hypothèse la plus naturelle. Nous allons dans la suite essayer de tester, au moins en partie, ce scénario. Cela ne résout bien sûr pas tous les problèmes : ainsi, les états stationnaires stables sont très nombreux (voir le paragraphe 5.2); comprendre les déviations à l'équilibre statistique de Vlasov, c'est à dire quel état précisément choisit le système après la phase de relaxation violente, est une tâche difficile. D'autre part, il faudrait aussi savoir comment évolue par la suite le système parmi ces états stationnaires de Vlasov. Nous discuterons rapidement ces questions.

Ce problème de relaxation lente pour le système HMF a déjà suscité beaucoup d'études : ce modèle est en effet considéré comme un laboratoire pour tester les idées que l'on espère valables de façon plus générale. Ainsi, Latora *et al.* l'ont étudié en détail à partir de conditions initiales particulières : toutes les particules sont placées au même endroit sur le cercle (soit $M_1 = 1$), et les vitesses sont distribuées selon une distribution rectangulaire, dite "water-bag". L'énergie est contrôlée en faisant varier la largeur de cette distribution. Ils appellent la zone $M_1 \simeq 0$, visible sur la figure 5.2¹ état quasi stationnaire, et remarquent que la distribution des vitesses de cet état quasi stationnaire peut être fittée par une distribution de Tsallis, à condition de bien choisir le paramètre de Tsallis q ainsi qu'une vitesse de coupure pour éviter une divergence aux grandes énergies [82] : leurs données sont présentées figure 5.3.

Pour ces auteurs, il s'agit d'un exemple paradigmatique de l'application de la statistique de Tsallis. Le débat est actuellement très vif sur le sujet. Montemurro et Zanette contestent même la notion d'état quasi stationnaire [115] : selon eux, il s'agit d'une relaxation lente et continue vers l'état d'équilibre, pour laquelle on ne peut définir aucun état quasi stationnaire. Notre but dans cette partie est plutôt d'illustrer l'intérêt et les limites de l'équation de Vlasov pour étudier ce phénomène de relaxation lente, mais nous reviendrons sur cette controverse.

Le plan de ce chapitre est le suivant : nous commençons par une étude de l'équation de Vlasov associée au modèle HMF; en particulier, nous décrivons tous les états stationnaires (section 5.2.1), et étudions la stabilité de toute une classe d'entre eux. Les méthodes utilisées sont générales; la simplicité du modèle HMF permet de pousser assez loin les calculs analytiques. Nous confrontons ensuite ces résultats de stabilité valables pour l'équation de Vlasov avec des résultats numériques obtenus pour la dy-

¹Précisons que cette figure n'a pas été obtenue avec les conditions initiales de Latora *et al.*; à l'exception des temps très courts cependant, l'allure de l'évolution de $M_1(t)$ pour les conditions initiales de Latora *et al.* est analogue.

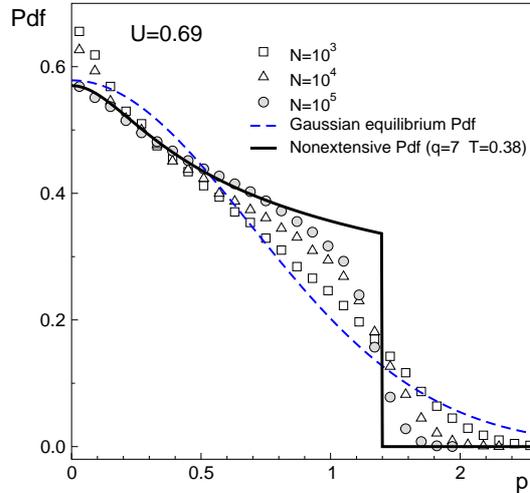


Figure 5.3: Distribution des vitesses (Pdf) pour le modèle HMF au temps $t = 1000$, c'est à dire après la relaxation violente, et avant la relaxation collisionnelle. L'énergie par particule est 0.69, et les conditions initiales celles de Latora *et al.* (voir texte). La distribution d'équilibre et le meilleur fit par une distribution de Tsallis sont indiqués. Cette figure est tirée de [82].

namique particulière. Enfin, nous illustrons, de façon numérique, les étapes 2, 3 et 4 du scénario : approche d'un état stationnaire stable de Vlasov, puis lente évolution vers l'équilibre, parmi les états stationnaires stables de Vlasov (section 5.3). Cette étude du modèle HMF a donné lieu à une prépublication [113]. Finalement, nous introduisons brièvement l'approche récente de F. Bouchet [27], qui vise à décrire analytiquement cette lente évolution parmi les états stationnaires stables de Vlasov, tâche jusqu'ici hors de portée. Nous discutons ensuite ces résultats à la section 5.4 (leur généralité, le lien avec les études sur d'autres modèles plus réalistes), et nous revenons sur la controverse autour de l'utilisation de la statistique de Tsallis dans le modèle HMF.

5.2 Étude de l'équation de Vlasov associée

La première étape du scénario consiste en une évolution rapide, sur une échelle de temps indépendante du nombre de particules, sous l'effet de la dynamique de Vlasov. Dans cette section, nous allons essayer de tester cette première phase, ce qui implique de comprendre un petit peu la dynamique Vlasovienne; nous allons donc commencer par décrire les états stationnaires de cette équation.

Écrivons l'équation de Vlasov dans le cas du modèle HMF. Les équations du mouvement peuvent se mettre sous la forme

$$\frac{d\theta_j}{dt} = p_j, \quad \frac{dp_j}{dt} = -M_x \sin \theta_j + M_y \cos \theta_j, \quad (5.2)$$

avec M_x et M_y les parties réelles et imaginaires de la magnétisation complexe \mathbf{M}_1 . Pour passer à la description Vlasovienne, il faut approximer la distribution discrète

$$\frac{1}{N} \sum_{j=1}^N \delta(\theta - \theta_j(t), p - p_j(t)) \quad (5.3)$$

par une densité continue sur l'espace des phases à une particule $f(\theta, p, t)$. Alors M_x et M_y peuvent s'écrire approximativement en fonction de f :

$$\overline{M_x}[f] \equiv \int \cos(\theta) f(\theta, p, t) d\theta dp \quad \text{et} \quad \overline{M_y}[f] \equiv \int \sin(\theta) f(\theta, p, t) d\theta dp. \quad (5.4)$$

Dans cette approximation (qui devient exacte dans la limite $N \rightarrow \infty$, au sens indiqué plus haut), les particules ressentent toutes le même potentiel

$$V(\theta)[f] = 1 - \overline{M_x}[f] \cos \theta - \overline{M_y}[f] \sin \theta. \quad (5.5)$$

Dans la suite, nous omettrons de surligner M_x et M_y pour indiquer l'approximation continue, lorsqu'aucune confusion ne sera possible. L'équation de Vlasov gouvernant l'évolution de la densité f s'écrit alors

$$\frac{\partial f}{\partial t} + p \frac{\partial f}{\partial \theta} - \frac{dV}{d\theta}[f] \frac{\partial f}{\partial p} = 0. \quad (5.6)$$

Comme expliqué en introduction de cette partie II, l'équation de Vlasov hérite de la dynamique particulière la conservation de l'énergie et de la quantité de mouvement totales

$$H[f] = \int \frac{p^2}{2} f(\theta, p, t) d\theta dp - \frac{M_x^2 + M_y^2}{2} \quad \text{et} \quad P = \int p f(\theta, p, t) d\theta dp, \quad (5.7)$$

et conserve de plus une infinité de nouvelles quantités, appelées Casimirs; ainsi, toutes les quantités de la forme

$$C_\varphi[f] = \int \varphi(f(\theta, p, t)) d\theta dp, \quad (5.8)$$

pour φ une fonction continue quelconque, sont conservées.

5.2.1 États stationnaires de l'équation de Vlasov

Nous étudions dans ce paragraphe les états stationnaires de l'équation de Vlasov. Pour une densité f stationnaire, la magnétisation M_1 est constante. Or l'équation de Vlasov à potentiel V constant est linéaire (voir les équations (5.5) et (5.6)); les solutions de cette équation aux dérivées partielles linéaire du premier ordre sont alors fournis par la méthode des caractéristiques; les solutions stationnaires sont les densités f constantes le long des caractéristiques de l'équation, qui sont les lignes de niveau de l'énergie ϵ

$$\epsilon(\theta, p) = \frac{p^2}{2} + V(\theta) = \frac{p^2}{2} + 1 - M_x \cos \theta - M_y \sin \theta. \quad (5.9)$$

Remarquons que l'énergie totale donnée par (5.7) est différente de la somme de ces énergies individuelles; c'est dû au fait que le potentiel V est en fait auto-cohérent et dépend de f , mais cela ne remet pas en cause la validité de la méthode des caractéristiques. Les solutions stationnaires de l'équation de Vlasov sont donc données par $f(\theta, p) = \Phi(\epsilon(\theta, p))$, où Φ est une fonction quelconque. Le cas particulier $\Phi = \exp(-\beta\epsilon)$ correspond à l'équilibre statistique.

On peut s'attendre à ce que ces très nombreux états stationnaires empêchent d'une part l'équation de Vlasov d'atteindre son état d'équilibre statistique, et d'autre part soient à l'origine de la relaxation lente, ou des états quasi stationnaires, observés sur la dynamique particulière. Nous étudions au paragraphe suivant la stabilité de ces états stationnaires.

5.2.2 Stabilité des états stationnaires

La méthode que nous allons employer dans ce paragraphe est tout à fait générale. Pour un système dynamique quelconque, tout extremum f_0 d'une quantité conservée $F[f]$ est un point stationnaire de la dynamique. De plus, si f_0 est un maximum ou un minimum strict de cette quantité conservée, f_0 est dite *formellement stable*. Holm et al. discutent dans la référence [61] les différents types de stabilité : stabilité spectrale (les valeurs propres de la dynamique linéarisée n'ont pas de partie réelle positive), stabilité linéaire (la dynamique linéarisée reste confinée au voisinage du point fixe), stabilité formelle (voir ci-dessus), stabilité non linéaire (la dynamique réelle reste confinée autour du point fixe). La stabilité formelle implique la stabilité spectrale et la stabilité linéaire². Les deux réciproques sont fausses : la stabilité spectrale n'implique ni la stabilité formelle ni la stabilité linéaire, et celle-ci n'implique pas la stabilité formelle. Enfin, la stabilité formelle n'implique la stabilité non linéaire que pour une dynamique de dimension finie. L'appendice C est consacrée à la démonstration de quelques résultats simples concernant la stabilité formelle en dimension finie. Nous allons dans le cas de l'équation de Vlasov étudier la stabilité formelle de différents états stationnaires, en utilisant les quantités conservées pour construire la fonction $F[f]$ (énergie et Casimirs). C'est une méthode standard en dynamique des fluides bidimensionnels et en physique des plasmas, qui permet ensuite d'obtenir des résultats de stabilité non linéaire [61]. Elle a été utilisée par exemple par Kandrup dans le cadre des systèmes auto-gravitants [72].

Considérons le maximum de la fonctionnelle

$$F[f] = C_s[f] - \beta H[f] - \mu \int f(\theta, p, t) d\theta dp. \quad (5.10)$$

β et μ sont des paramètres; H (l'énergie (5.7)), C_s (fonctionnelle de Casimir (5.8)) et $\int f = 1$ sont des quantités conservées par la dynamique, donc aussi F . On pourrait aussi ajouter à la fonctionnelle un terme dépendant de la quantité de mouvement

²en dimension infinie, il faut préciser la norme utilisée; dans ce cas, on utilise la norme donnée par les variations secondes de F autour de f_0 , qui sont définies positives par hypothèse (ou définies négatives, on prend alors l'opposé).

totale, qui est aussi une quantité conservée : on obtiendrait ainsi des distributions stationnaires dans des référentiels en mouvement; on ne s'intéresse pas à ce cas dans la suite. On suppose aussi que la fonction s qui définit le Casimir est strictement concave. Les points critiques de ces fonctionnelles données par l'équation (5.10) fournissent des états stationnaires de Vlasov. Remarquons que pour $s(x) = -x \ln x$, le premier terme de F est l'entropie de Boltzmann d'une distribution f , et F est l'opposé de la fonctionnelle d'énergie libre. L'équilibre statistique apparaît comme un cas particulier parmi tous les états stationnaires de l'équation de Vlasov; cette présentation permet donc de prouver la stabilité dynamique (au moins formelle) de l'état d'équilibre statistique, lorsque celui-ci est un minimum de la fonctionnelle d'énergie libre. Lorsque ce n'est pas le cas (c'est à dire dans une situation d'inéquivalence entre les ensembles canonique et microcanonique), l'état d'équilibre statistique, maximum d'entropie, est encore dynamiquement stable, mais il faut employer des arguments plus sophistiqués pour le prouver [47].

Écrivons les premières variations pour f_0 , point critique de F :

$$s'(f_0) - \beta \left(p^2/2 - \int \cos(\theta - \alpha) f_0(\alpha, p, t) d\alpha dp \right) - \mu = 0. \quad (5.11)$$

Ce qui donne une expression pour f_0 :

$$f_0(\theta, p) = \Phi(\beta\epsilon(\theta, p) + \mu), \quad (5.12)$$

où ϵ est donné par (5.9) et Φ est l'inverse de la fonction s' dérivée de s . On peut donc caractériser avec cette méthode toutes les solutions stationnaires de l'équation de Vlasov avec Φ strictement décroissante. Calculons maintenant les variations secondes de F pour savoir à quelle condition f_0 est un maximum local :

$$\Delta_2 F[\delta f, \delta f] = \int s''(f_0(\theta, p)) (\delta f(\theta, p))^2 d\theta dp + \beta ((M_x[\delta f])^2 + (M_y[\delta f])^2). \quad (5.13)$$

Comme s est concave, le premier terme est négatif, alors que le second est clairement positif.

On s'intéresse dans la suite à la stabilité des états homogènes, pour lesquels $M_x = M_y = 0$, et $f_0(\theta, p) = f_0(p) = \Phi(\beta p^2/2 + \mu)$. On décompose alors la perturbation en série de Fourier

$$\delta f(\theta, p) = \sum_n c_n(p) \cos(n\theta) + s_n(p) \sin(n\theta). \quad (5.14)$$

À partir de (5.13) et après intégration sur θ , on obtient l'expression pour les variations secondes

$$\Delta_2 F[\delta f, \delta f] = \int s''(f_0(p)) \sum_{n>1} (c_n^2(p) + s_n^2(p)) dp + 2G(c_1(p)) + 2G(s_1(p)), \quad (5.15)$$

avec

$$G(c(p)) \equiv \int s''(f_0(p)) c^2(p) dp + \frac{\beta}{2} \left(\int c(p) dp \right)^2.$$

Les termes mettant en jeu c_n et s_n pour $n > 1$ dans (5.15) sont définis négatifs (rappelons que s est concave). Par conséquent, les variations secondes de F sont définies négatives si et seulement si G est définie négative. Pour étudier le signe de G , utilisons l'inégalité de Cauchy-Schwartz

$$\begin{aligned} \left(\int c(p) dp \right)^2 &= \left(\int \frac{c(p) \sqrt{-s''(f_0(p))}}{\sqrt{-s''(f_0(p))}} dp \right)^2 \\ &\leq \left(\int s''(f_0(p)) c^2(p) dp \right) \left(\int \frac{1}{s''(f_0(p))} dp \right). \end{aligned} \quad (5.16)$$

Cette inégalité permet de prouver que

$$G(c(p)) \leq \int s''(f_0(p)) c^2(p) dp \left(1 + \frac{\beta}{2} \left(\int \frac{1}{s''(f_0(p))} dp \right) \right).$$

On pose alors

$$I[f_0] = 1 + \beta/2 \left(\int 1/s''(f_0(p)) dp \right). \quad (5.17)$$

Puisque s'' est strictement négatif, on conclut que si $I[f_0]$ est positif, alors G est négatif et donc f_0 est formellement stable. Inversement, si $I[f_0]$ est négatif, il suffit de prendre comme cas particulier de perturbation $c(p) = -1/s''(f_0(p))$ pour voir que G n'est pas négatif : nous avons donc un critère pour tester la stabilité de f_0 .

On utilise maintenant le fait que $s'(f_0(p)) = \beta p^2/2 + \mu$; en dérivant par rapport à p , on trouve $\beta/s''(f_0(p)) = f'_0(p)/p$. Le critère de stabilité s'écrit alors sans faire intervenir s :

$$I[f_0] = 1 + \frac{1}{2} \int_{-\infty}^{\infty} \frac{f'_0(p)}{p} dp. \quad (5.18)$$

Résumons; on a prouvé l'équivalence

$$f_0 \text{ formellement stable} \iff I[f_0] > 0, \quad (5.19)$$

ce qui fournit un critère simple pour tester la stabilité de toutes les solutions stationnaires de l'équation de Vlasov, avec distribution angulaire homogène. Nous l'appliquons maintenant sur plusieurs distributions.

Ex. 1 : Commençons par une distribution gaussienne (c'est à dire l'équilibre statistique dès que la densité d'énergie $e = E/N$ est supérieure à $e_c = 3/4$).

$$f_{\text{gauss}}(p) = \sqrt{\frac{\beta}{2\pi}} e^{-\beta p^2/2}.$$

On obtient alors

$$I[f_{\text{gauss}}] = 1 - \frac{1}{2}\beta,$$

ce qui fournit une température critique $\beta_c^* = 2$ et donc une densité d'énergie critique $e_c^* = 3/4$, qui coïncide avec le point de transition du second ordre (Inagaki [64] avait déjà remarqué ce fait).

Ex. 2 : Considérons maintenant la distribution dite “water bag”

$$f_{\text{waterbag}}(p) = \begin{cases} 0 & (p < -\bar{p}) \\ 1/(2\bar{p}) & (-\bar{p} < p < \bar{p}) \\ 0 & (\bar{p} < p) \end{cases}, \quad (5.20)$$

avec $\bar{p} = \sqrt{6e - 3}$. Cette distribution est utilisée par exemple par [112, 97]; associée à une distribution angulaire initiale non homogène (tous les θ_i à $t = 0$ sont égaux), elle a été étudiée en détail [82].

On calcule $I[f_{\text{waterbag}}]$ en utilisant

$$f'_{\text{waterbag}}(p) = \frac{1}{2\bar{p}} (\delta(p + \bar{p}) - \delta(p - \bar{p})). \quad (5.21)$$

Alors

$$I[f_{\text{waterbag}}] = 1 - \frac{1}{2} \frac{1}{\bar{p}^2},$$

ce qui fournit la largeur critique de la distribution $\bar{p}_c^2 = 1/2$. Or la densité d'énergie $e = E/N$ vérifie $e = \langle p^2/2 \rangle + (1 - M_1^2)/2$ avec $\langle p^2 \rangle = \bar{p}^2/3$. L'énergie critique est donc

$$e_c^* = \frac{1}{2} + \frac{1}{12}.$$

Cette énergie critique est plus faible que le point de transition thermodynamique; en d'autres termes, pour $7/12 < e < 3/4$, la distribution “water bag”, avec magnétisation nulle, est stable pour l'équation de Vlasov, bien que l'équilibre statistique prédise $M_1 \neq 0$.

Ex. 3 : Enfin, introduisons à partir des deux exemples précédents une famille de distributions interpolants entre la gaussienne et la distribution rectangulaire “water bag” :

$$f_a(p) = (1 - a)f_{\text{waterbag}}(p) + af_{\text{gaussian}}(p). \quad (5.22)$$

Pour $0 \leq a \leq 1$. L'énergie critique pour f_a s'obtient par combinaison linéaire des énergies critiques des exemples 1 et 2 (car I est linéaire en f)

$$U_c^*(a) = (1 - a) \left(\frac{1}{2} + \frac{1}{12} \right) + a \frac{3}{4} = \frac{7}{12} + a \frac{1}{6}.$$

On a donc trouvé des solutions stationnaires stables de l'équation de Vlasov, qui ne vont relaxer vers l'équilibre que sous l'influence des fluctuations de N fini, qui différencient la densité continue f de la distribution discrète : il est alors logique que cette relaxation soit d'autant plus lente que N est grand. Remarquons que ces résultats permettent de comprendre qualitativement la zone d'énergie de la figure 5.1 pour laquelle la relaxation vers l'équilibre est lente : elle correspond approximativement à la zone où l'état $M_1 = 0$ avec distribution de vitesses “water bag” est stable, bien que $e < e_c$.

Pour valider ces calculs, nous les comparons maintenant avec les simulations numériques. Les équations du mouvement des particules sont intégrées à l'aide du meilleur

algorithme symplectique d'ordre 4 [88], et le pas de temps est choisi de telle sorte que l'erreur relative en énergie reste inférieure à 10^{-7} . Les conditions initiales sont : particules réparties de façon aléatoire entre 0 et 2π (donc magnétisation quasi nulle $M_1 \sim 1/\sqrt{N}$), et vitesses distribuées selon la distribution à tester; le moment total est fixé à 0. Pour une distribution instable, on s'attend à une croissance de la magnétisation jusqu'à une valeur finie; pour une distribution stable, on s'attend à ce que $M_1(t)$ fluctue avec une amplitude d'autant plus petite que N est grand. L'évolution temporelle de $M_1(t)$ aux temps courts est montrée figure 5.4, pour une distribution initiale “water bag” et différentes énergies. Pour tester le critère de stabilité, considérons

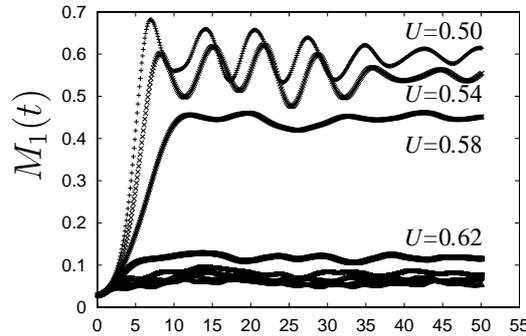


Figure 5.4: Évolution temporelle de $M_1(t)$ pour $N = 1000$. Les résultats sont moyennés sur 100 réalisations.

le premier pic de $M_1(t)$, visible sur la figure 5.4; la figure 5.5 représente la hauteur de ce premier pic en fonction de l'énergie et de N , pour une distribution initiale “water bag” : la présence d'une énergie critique $e_c = 7/12$ est très claire sur cette figure.

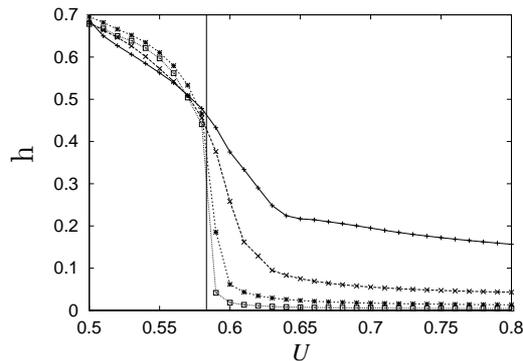


Figure 5.5: La hauteur du premier pic de la magnétisation $M_1(t)$ en fonction de la densité d'énergie e , pour différentes valeurs de N (10^2 , 10^3 , 10^4 et 10^5). Les résultats sont moyennés sur 100 réalisations (20 pour $N = 10^5$).

La figure 5.6 teste la validité du critère de stabilité d'une distribution initiale de vitesses (5.19), pour des distributions mixtes gaussiennes et “water bag”; $a = 0$ (resp. $a = 1$) correspond à une distribution initiale “water bag” (resp. gaussienne) pure. Bien que la limite entre stabilité et instabilité, en regardant la hauteur du premier

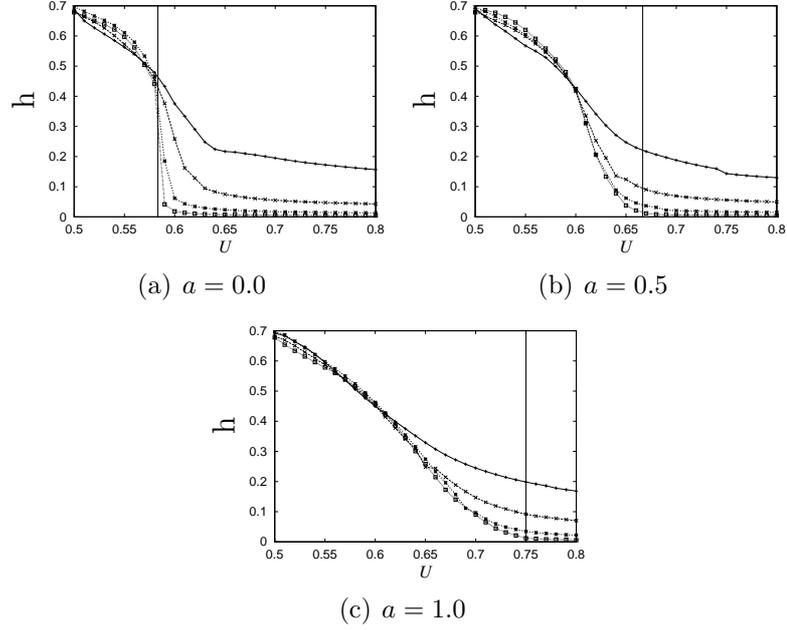


Figure 5.6: Même figure que 5.5, pour différentes distributions initiales.

pic, soit moins nette pour $a = 0.5$ et $a = 1$ que dans le cas purement “water bag”, les résultats sont cohérents avec la prédiction théorique : la hauteur du premier pic de $M_1(t)$ tend vers 0 lorsque N tend vers l’infini, pour $e > e_c$. Pour $e < e_c$, avec e proche de e_c , l’instabilité semble très faible : l’analyse linéaire montre en effet (voir paragraphe suivant et [4, 64]) que le taux de croissance de l’instabilité croît comme $(e_c - e)^{1/2}$ dans le cas “water bag”, et seulement de façon linéaire en $(e_c - e)$ dans le cas gaussien.

5.2.3 Stabilité linéaire

La propriété de stabilité formelle que nous avons montrée est plus forte que la stabilité linéaire, et de plus la méthode est très générale. Néanmoins, l’analyse linéaire apporte une information supplémentaire : la valeur du taux de croissance de l’instabilité dans le cas instable. Nous calculons ci-dessous ce taux de croissance, dans le cas “water bag” [4].

Nous linéarisons donc l’équation de Vlasov (5.6) autour d’une solution stationnaire homogène, qui ne dépend donc que de p , $f_0(p)$. La perturbation δf dépend de (p, θ, t) ; l’équation linéarisée pour δf est

$$\frac{\partial \delta f}{\partial t} + p \frac{\partial \delta f}{\partial \theta} - (\sin \theta M_x[\delta f] - \cos \theta M_y[\delta f]) \frac{\partial f}{\partial p} = 0. \quad (5.23)$$

Nous décomposons à nouveau δf en série de Fourier

$$\delta f = \sum_n (c_n(p, t) \cos n\theta + s_n(p, t) \sin n\theta), \quad (5.24)$$

et nous écrivons l'équation d'évolution de chaque composante :

$$\frac{\partial c_n}{\partial t} = -nps_n \quad \forall n > 1 \quad (5.25)$$

$$\frac{\partial s_n}{\partial t} = npc_n \quad \forall n > 1 \quad (5.26)$$

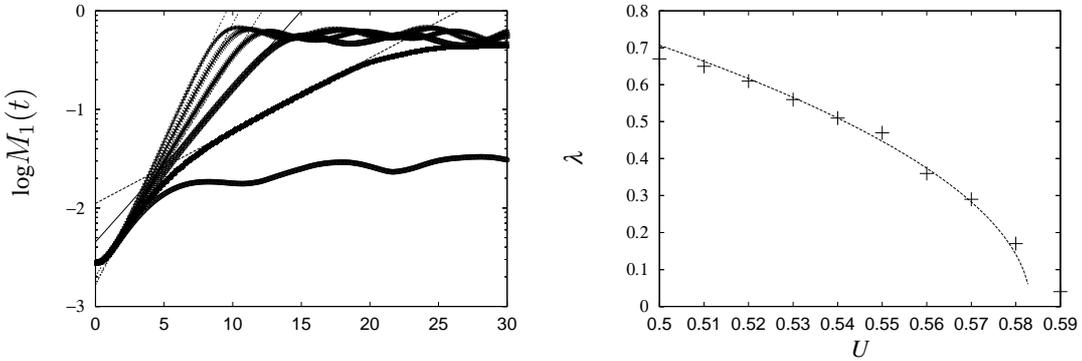
$$\frac{\partial c_1}{\partial t} = -ps_1 - \pi \frac{\partial f}{\partial p} \int du s_1(u) \quad (5.27)$$

$$\frac{\partial s_1}{\partial t} = pc_1 + \pi \frac{\partial f}{\partial p} \int du c_1(u). \quad (5.28)$$

Il est facile de voir que les composantes avec $n > 1$ ne peuvent pas être instables. En effet, en écrivant $v_n = c_n + is_n$, on a $\dot{v}_n = ipv_n$, où \dot{v} indique la dérivée temporelle de v . Les vecteurs propres et valeurs propres (généralisés) sont donc respectivement toutes les fonctions δ de Dirac et tous les ip , pour $p \in \mathbb{R}$. Nous passons donc maintenant aux composantes $n = 1$. Dans le cas “water bag”, la dérivée de f fait intervenir des fonctions δ , si bien que les équations pour c_1 et s_1 deviennent

$$\dot{c}_1 = -ps_1 - \frac{1}{4\bar{p}} \left[\int du s_1(u) \right] (\delta(p + \bar{p}) - \delta(p - \bar{p})) \quad (5.29)$$

$$\dot{s}_1 = pc_1 + \frac{1}{4\bar{p}} \left[\int du c_1(u) \right] (\delta(p + \bar{p}) - \delta(p - \bar{p})) . \quad (5.30)$$



(a) Courbe $M_1(t)$, échelle semi logarithmique (b) Comparaison théorie/numérique pour le taux de croissance

Figure 5.7: (a) Courbe $M_1(t)$, en échelle semi logarithmique; $N = 10^5$; moyenne sur 20 réalisations. (b) Comparaison entre le taux de croissance calculé par l'équation (5.31) et le taux de croissance estimé numériquement.

Il s'agit toujours d'un problème aux valeurs propres de dimension infinie. Nous nous restreignons maintenant aux fonctions c_1 et s_1 combinaisons linéaires de $\delta(p + \bar{p})$ et $\delta(p - \bar{p})$ (cet espace vectoriel est stable sous l'action de l'opérateur à diagonaliser). Le problème n'a plus alors que 4 dimensions, et les valeurs propres peuvent être calculées explicitement en fonction de \bar{p} , c'est à dire en fonction de l'énergie par particule e . Toutes les valeurs propres sont imaginaires pures si $e > 7/12$ (on retrouve le seuil

obtenu plus haut); si $e < 7/12$, la plus grande valeur propre, qui contrôle le taux de croissance, est

$$\lambda = \sqrt{3.5 - 6U}. \quad (5.31)$$

Or la croissance exponentielle de la perturbation entraîne la croissance exponentielle de la magnétisation $M_1(t)$. La figure 5.7a) montre en effet une croissance exponentielle de $M_1(t)$, avec un taux bien décrit par l'équation (5.31) : voir la figure 5.7b).

Résumons les enseignements de cette étude des états stationnaires de l'équation de Vlasov; ils ont une portée plus large que le modèle jouet HMF, sur lequel nous les avons illustrés (nous reviendrons sur ce point à la section 5.4) :

- On connaît les états stationnaires de l'équation de Vlasov : ce sont les distributions sur l'espace (θ, p) qui ne dépendent que de l'énergie ϵ (5.9). On peut les caractériser comme points critiques de fonctionnelles construites à partir des quantités conservées du système, en incluant les Casimirs. Le cas particulier de l'équilibre statistique collisionnel correspond à un point critique de la fonctionnelle énergie libre, qui est aussi construite à partir des Casimirs. L'équilibre statistique de l'équation de Vlasov est un autre cas particulier.
- La stabilité de ces états stationnaires peut être étudiée en considérant les variations secondes de ces fonctionnelles. Dans le cas du modèle HMF, nous avons pu déterminer de cette façon un critère simple de stabilité pour toutes les solutions stationnaires homogènes de Vlasov.
- La présence de ces très nombreux états stationnaires stables de l'équation de Vlasov donne une explication à la relaxation très lente vers l'équilibre de la dynamique particulière, pour certaines conditions initiales.

5.3 Convergence vers un état stationnaire de l'équation de Vlasov

Lorsque la condition initiale ne correspond pas à une solution stationnaire stable de Vlasov (par exemple une distribution “water bag” d'énergie inférieure à $7/12$), il est naturel de penser que le système se rapprochera d'une solution stationnaire stable, après une période d'évolution rapide sous l'effet de la dynamique de Vlasov (étape 2 du scénario de la section 5.1). Nous testons maintenant numériquement cette hypothèse.

On a vu qu'une distribution sur l'espace des phases $f(\theta, p)$ est stationnaire pour l'équation de Vlasov si elle ne dépend que de l'énergie ϵ (5.9) de chaque particule. On peut donc imaginer une procédure pour tester la stationnarité : diviser l'espace des phases (θ, p) en petites boîtes, chacune correspondant à une énergie ϵ à $\Delta\epsilon$ près; compter le nombre de particules n_b dans chaque boîte; mettre un point pour chaque boîte sur un graphe (ϵ, n_b) ; si les points obtenus se distribuent sur une courbe unidimensionnelle, la densité $f(\theta, p)$ s'écrit sous la forme $f = \Phi(\epsilon(\theta, p))$, et f est stationnaire; si en revanche les points se distribuent en deux dimensions, f n'est pas stationnaire.

Malheureusement, cette procédure nécessite une statistique très précise dans un espace des phases à deux dimensions, si bien qu'elle n'est pas utilisable en pratique. Nous allons donc utiliser des tests de stationnarité plus simples, bien qu'a priori moins précis (ce ne sont que des conditions nécessaires de stabilité, pas suffisantes).

5.3.1 Deux conditions nécessaires de stationnarité

Moments de la distribution d'énergie - Test I Considérons par exemple une distribution f et la distribution associée des énergies individuelles des particules (5.9) $\Phi(\epsilon, t)$. Si f est stationnaire, Φ l'est aussi, ainsi que tous ses moments

$$\sigma_k(t) = \langle \epsilon^k \Phi(\epsilon) \rangle, \quad k = 1, 2, \dots \quad (5.32)$$

Lorsque donc l'un des $d\sigma_k/dt$ est non nul, la distribution correspondante est non stationnaire. Ces moments σ_k sont faciles à calculer à partir des positions et vitesses des N particules; on appellera dans la suite ce test "test I". Notons aussi qu'il serait valable pour toute autre équation de Vlasov.

Symétrie de la distribution f - Test II Si l'on considère une distribution stationnaire f , on peut supposer sans perte de généralité, que la partie imaginaire de la magnétisation M_y est nulle. Alors, d'après (5.9), $\epsilon(\theta, p)$ est paire par rapport à θ , ce qui implique que f l'est aussi. On peut fonder sur cette remarque un autre test de stationnarité, qui ne nécessite pas d'échantillonnage de l'espace des phases à deux dimensions. En effet, cette parité implique la parité de la fonction de distribution intégrée

$$\tilde{f}(\theta, t) \equiv \int f(\theta, p, t) dp = \tilde{f}(-\theta, t). \quad (5.33)$$

Malheureusement, la distribution intégrée sur p peut être paire sans que la distribution f le soit; un exemple typique est donné par la figure 5.8. Pour mieux conserver l'asymétrie de la distribution et donc améliorer le test, on restreint l'intégration aux $p < 0$, ou aux $p > 0$:

$$\tilde{f}_+(\theta, t) = \int_{p>0} f(\theta, p, t) dp, \quad \tilde{f}_-(\theta, t) = \int_{p<0} f(\theta, p, t) dp. \quad (5.34)$$

Alors, sur l'exemple de la figure 5.8, f_{\pm} restent asymétriques. Pour quantifier l'asymétrie, on introduit alors les deux quantités

$$A_+(t) = \int [\tilde{f}_+(\theta, t) - \tilde{f}_+(-\theta, t)]^2 d\theta = \int d\theta \int_{p>0} dp [f(\theta, p, t) - f(-\theta, p, t)]^2, \quad (5.35)$$

$$A_-(t) = \int [\tilde{f}_-(\theta, t) - \tilde{f}_-(-\theta, t)]^2 d\theta = \int d\theta \int_{p<0} dp [f(\theta, p, t) - f(-\theta, p, t)]^2, \quad (5.36)$$

et on étudie l'évolution de $A_{\pm}(t)/A(t)$, où $A(t)$ est

$$A(t) = \int d\theta \left[\int dp f(\theta, p, t) - f(-\theta, -p, t) \right]^2. \quad (5.37)$$

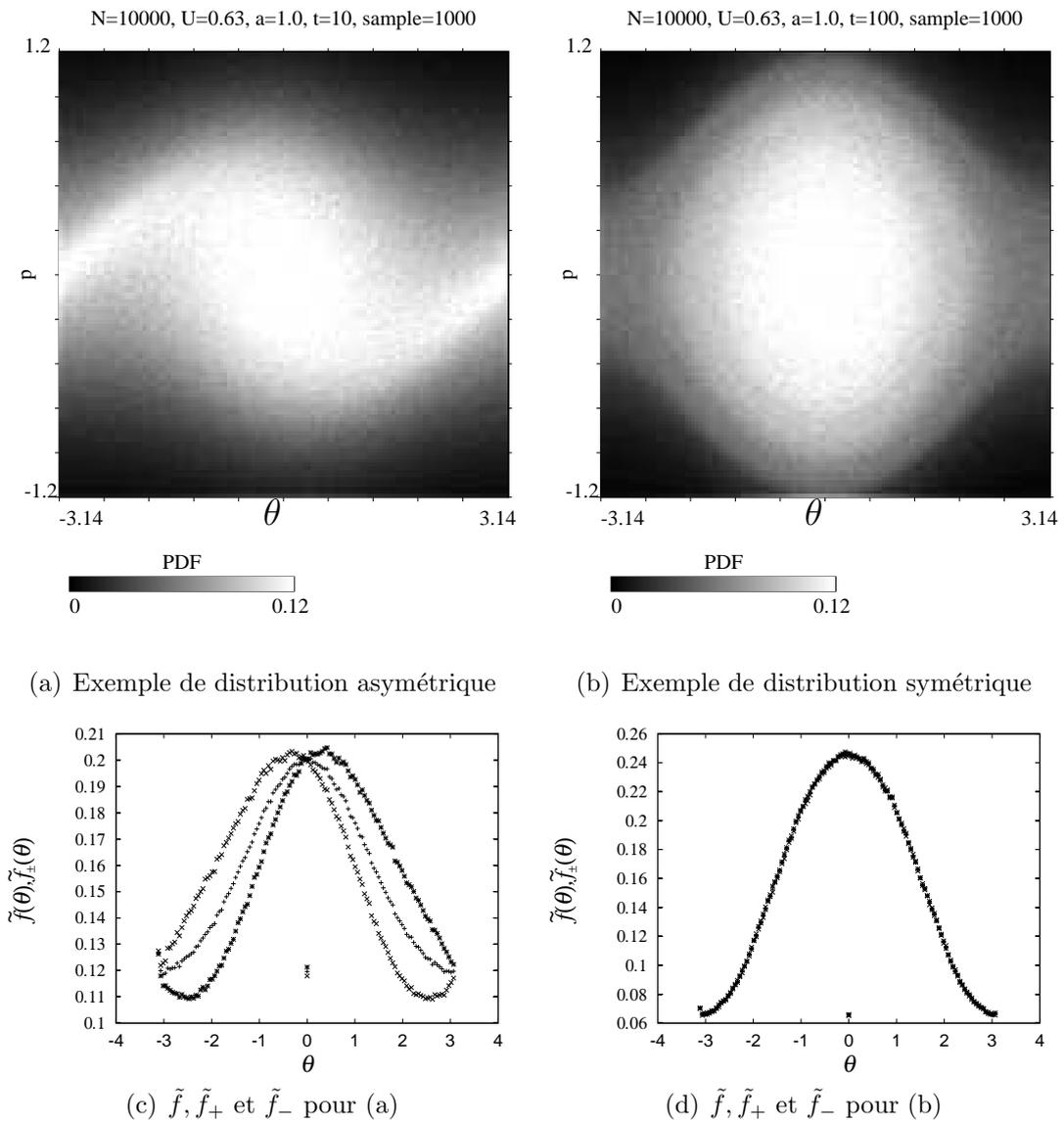


Figure 5.8: Exemples de distributions $f(\theta, p, t)$ asymétrique (a) et symétrique (b). (c) et (d) présentent les distributions intégrées $\tilde{f}(\theta, t)$ (courbe du milieu), $\tilde{f}_+(\theta, t)$ (à gauche) et $\tilde{f}_-(\theta, t)$ (à droite), pour les distributions des figures (a) et (b). Les trois courbes sont presque identiques sur la figure (d). $N = 10^4$, $U = 0.63$, $a = 1.0$; les données sont moyennées sur 10^3 réalisations. $t = 10$ pour les figures (a) et (c), et $t = 10^2$ pour les figures (b) et (d).

On étudie dans la suite des conditions initiales qui vérifient

$$f(\theta, p, t = 0) = f(-\theta, -p, t = 0) ; \quad (5.38)$$

cette symétrie est conservée par la dynamique et donc valable pour tout temps t . La quantité $A(t)$ est donc toujours très petite, non nulle du fait des effets de statistique finie et des erreurs numériques. Les quantités A_{\pm} en revanche sont très petites pour une distribution stationnaire, et pas nécessairement pour une distribution quelconque. Pour une distribution non stationnaire, on peut donc s'attendre à ce que A_{\pm}/A prenne des valeurs très grandes. Nous appellerons dans la suite ce test "test II".

5.3.2 Tests numériques

Nous utilisons maintenant les deux tests présentés ci-dessus pour vérifier le scénario : la dynamique particulière doit vite se rapprocher d'une distribution Vlasov-stable.

La figure 5.9 présente les résultats des tests I et II pour des conditions initiales de type "water bag" (avec distribution initialement homogène des particules). Lorsque la condition initiale est Vlasov-instable, comme pour la figure a), les deux tests prennent des valeurs nettement différentes de 0 jusqu'à $t \simeq 80$, ce qui correspond à la zone où M_1 croît rapidement ou oscille beaucoup, puis des valeurs quasi nulles. Bien qu'il ne s'agisse que de conditions suffisantes de stationnarité, il est tentant de conclure que toute l'évolution ultérieure s'effectue parmi les états stationnaires stables de Vlasov. La figure b) permet la comparaison avec une condition initiale Vlasov-stable (toujours "water bag" et distribution spatiale initiale homogène, mais de plus haute énergie). Si l'on excepte les petites oscillations du test I aux temps courts, qui disparaissent probablement lorsque N tend vers l'infini, les résultats sont à nouveau compatibles avec une évolution parmi les états stationnaires stables de Vlasov, conformément au point 3 du scénario de l'introduction. La figure 5.10 présente les mêmes calculs numériques, cette fois-ci pour une condition initiale gaussienne (et toujours une distribution initialement homogène des particules). Les résultats sont à nouveau compatibles avec les points 2 et 3 du scénario : convergence rapide vers un état Vlasov-stable, puis lente évolution parmi ces états. Néanmoins, pour $e = 0.69$, c'est à dire assez proche du seuil de stabilité $e = 0.75$, la zone non stationnaire est peu visible.

Nous avons jusqu'ici utilisé des conditions initiales correspondant à un état stationnaire de Vlasov, stable ou instable. La figure 5.11 illustre le cas de conditions initiales qui ne sont pas un état stationnaire de Vlasov : toutes les particules sont initialement au même point (donc $M_1(0) = 1$), avec une distribution de vitesse "water bag" (ce sont les conditions initiales couramment utilisés par Latora *et al.* [82]). Là encore, après l'évolution violente des temps courts, les deux tests prennent des valeurs quasi nulles, et il est tentant de conclure que la distribution est Vlasov stable dans toute la suite de l'évolution. En fait, la situation est probablement un petit peu plus compliquée. En effet, la condition initiale utilisée est unidimensionnelle dans l'espace des phases (p, θ) (toutes les particules ont le même angle initial). Cette ligne initiale se déforme et se distend jusqu'à devenir "bidimensionnelle", comme doivent l'être les solutions stationnaires (régulières) de Vlasov, mais certaines petites portions restent cohérentes et quasi unidimensionnelles pour des temps assez longs [48, 97]. Pour ces conditions

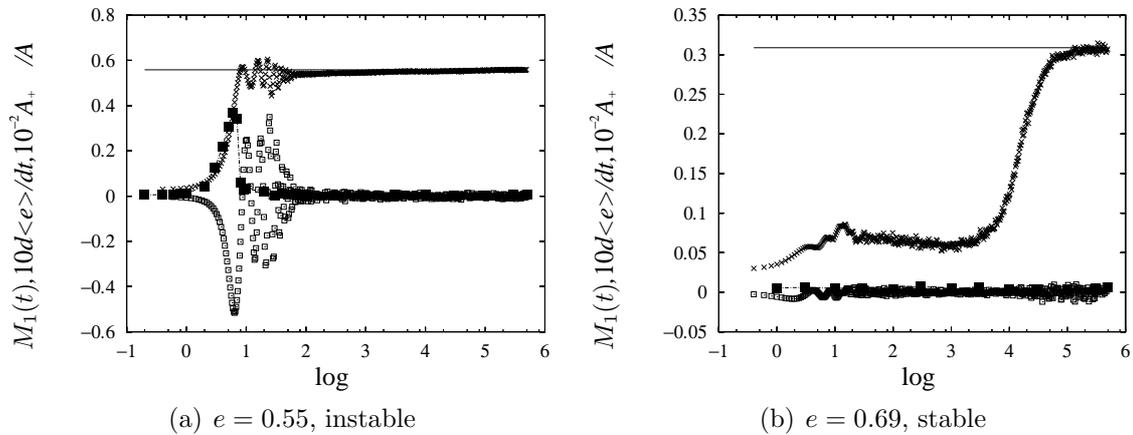


Figure 5.9: Tests de non stationnarité pour une distribution initiale des vitesses “water bag”; $N = 10^3$ particules, 100 réalisations. Noter l’échelle logarithmique pour les temps. La figure (a) (resp. (b)) montre l’évolution temporelle de la magnétisation $M_1(t)$ (croix), le test I, $d\langle e \rangle / dt$ (carrés blancs), et le test II, $A_+(t)/A(t)$ (carrés noirs), pour une énergie par particule $e = 0.55$ (resp. $e = 0.69$). Les quantités $d\langle e^j \rangle / dt$ (resp. $A_\pm(t)/A(t)$) sont multipliées par un facteur 10 (resp. 10^{-2}) pour des raisons graphiques. La ligne horizontale représente la valeur d’équilibre de la magnétisation.

initiales singulières, le scénario suggéré n’est donc pas tout à fait correct. Nous reviendrons dans la discussion 5.4 sur ce point, qui est important pour faire le lien avec les nombreuses études de Latora *et al.*

Il ressort de ces tests que, pour la plupart des conditions initiales utilisées, et probablement pour toutes les conditions initiales “bidimensionnelles”, les points 2 et 3 du scénario sont confirmés. Il est intéressant de remarquer que A. Taruya [106] obtient des résultats concordants, obtenus sur des simulations numériques précises de systèmes auto-gravitants en trois dimensions : après une relaxation rapide, l’évolution peut être paramétrée par une évolution parmi les états stationnaires stables de l’équation de Vlasov correspondant; le système s’approche finalement de l’équilibre statistique. Bien sûr, deux points importants restent dans l’ombre : d’une part quel est l’état stationnaire de Vlasov “choisi” par la relaxation violente, et d’autre part la dynamique précise qui régit la lente évolution parmi les états stationnaires de Vlasov est à priori difficile à connaître. Au paragraphe suivant, nous l’étudions numériquement, puis nous introduisons rapidement une approche analytique très récente de ce problème, due à F. Bouchet.

5.3.3 Étude numérique de l’évolution lente

La figure 5.12 illustre la lente convergence des distributions angulaires et des vitesses vers l’état d’équilibre statistique. Pour $e = 0.55$, la condition initiale est instable, et l’évolution est rapide entre $t = 1$ et $t = 10$; elle évolue ensuite très lentement jusqu’à l’équilibre, atteint pour $t \sim 10^5$. Pour $e = 0.69$, la condition initiale est stable; la dynamique est donc lente dès le départ, et l’équilibre est atteint aussi pour $t \sim 10^5$.

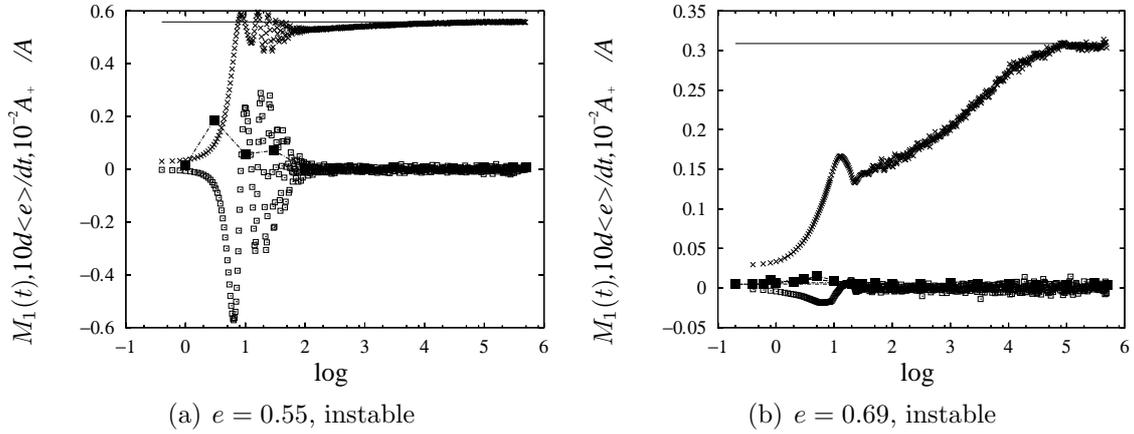


Figure 5.10: Mêmes courbes que sur les figures 5.9, mais pour une condition initiale gaussienne.

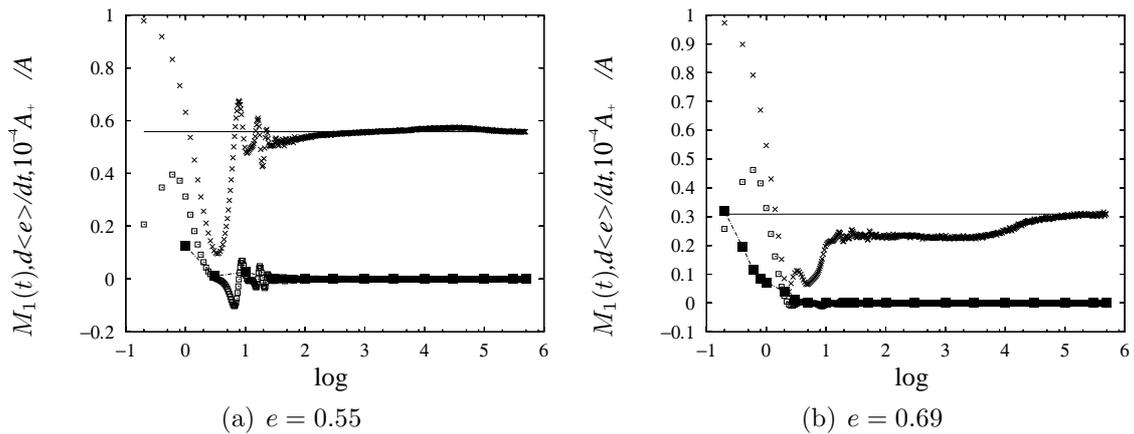


Figure 5.11: Mêmes courbes que sur les figures 5.9, mais avec une condition initiale "water bag" pour les moments p_j , et toutes les particules au même point : $\theta_j = 0$, pour tout j (donc $M_1(0) = 1$). Les quantités $A_{\pm}(t)/A(t)$ sont multipliées par 10^{-4} pour des raisons graphiques.

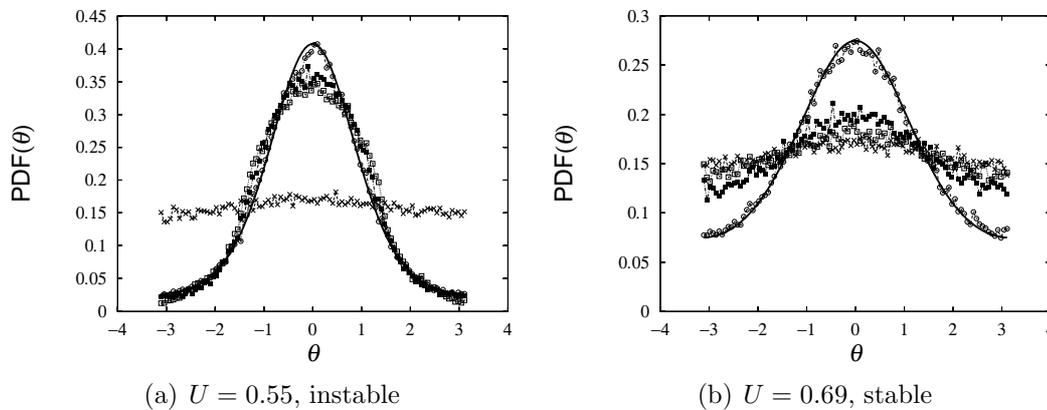


Figure 5.12: Évolution temporelle de la distribution $f(\theta, t)$ pour une distribution initiale des vitesses de type “water bag”; $N = 1000$, et les courbes sont moyennées sur 100 réalisations. Sont indiquées pour $U = 0.55$ (a) et $U = 0.69$ (b) les distributions aux temps $t = 1$ (croix), $t = 10$ (carré blanc), $t = 10^4$ (carré noir), $t = 10^5$ (cercle blanc). La ligne continue représente la distribution d’équilibre $\propto \exp(M_1^{eq} \cos \theta / T)$.

Bien que la dynamique soit lente, et la magnétisation M_1 parfois “gelée” hors équilibre, il semble que l’évolution du système, c’est à dire la déformation de la densité $\Phi(\epsilon)$ soit continue : il est alors difficile de définir un état quasi stationnaire. C’est aussi la conclusion à laquelle arrivent Montemurro et Zanette [115] (ainsi d’ailleurs que Taruya [106]).

Nous nous intéressons pour terminer aux échelles de temps mises en jeu. Nous avons dit à plusieurs reprises que le temps de relaxation, donc la durée de vie des états hors équilibre, augmente avec le nombre de particules du système. En pratique, il est bien sûr important de savoir plus précisément quelle est cette dépendance, pour savoir dans quelle régime se trouve un système donné. L’astrophysique fournit un bon exemple : on pense en effet que la structure des galaxies vient plutôt de la relaxation violente, tandis que la relaxation collisionnelle et donc le véritable équilibre statistique pourrait jouer un rôle pour les amas globulaires, plus petits.

Pour le modèle HMF, le théorème de Braun et Hepp dit que la dynamique de Vlasov et la dynamique particulaire restent proches sur un temps d’ordre $\ln N$; on pourrait donc a priori estimer un premier temps de relaxation $t_{rel} \sim \ln N$. En fait, une fois que le système est piégé par un état stationnaire de Vlasov, il y reste pour des temps beaucoup plus longs : il n’est déstabilisé que par les fluctuations de N fini, ce qui pourrait cette fois faire penser à un temps de relaxation en $t_{rel} \sim N^{1/2}$. Un calcul analytique récent de F. Bouchet portant sur les fluctuations dynamiques autour de l’équilibre [27] suggère en fait que ces fluctuations n’agissent que sur un temps d’ordre N (le calcul de Bouchet n’est valable a priori qu’autour de l’équilibre, mais il pourrait se généraliser à des situations hors équilibre). Ce temps $t_{rel} \sim N$ est effectivement celui obtenu numériquement par Latora et collaborateurs pour la durée de vie de “l’état quasi stationnaire” [82] (ils ne précisent pas comment exactement ce temps est obtenu). Cependant, Zanette et Montemurro [115], avec pourtant les mêmes conditions initiales, montrent que $M_1(t)$ atteint un minimum avant de remonter vers