La dynamique des systèmes hamiltoniens presque intégrables

Patrick Bernard

Université Joseph Fourier, Grenoble I

La dynamique des systèmes hamiltoniens presque intégrables

Mémoire Présenté pour obtenir

l'habilitation à diriger des recherches

spécialité

Mathématiques

par

Patrick Bernard

Soutenu le 20 decembre 2004 devant le jury composé de

Alain CHENCINER
Yves COLIN de VERDIÈRE
Albert FATHI
Lucien GUILLOU
Raoul ROBERT
Eric SÉRÉ

Au vu des rapports de Alain CHENCINER, Craig EVANS et Albert FATHI.

LA DYNAMIQUE DES SYSTÈMES HAMILTONIENS PRESQUE INTÉGRABLES

PATRICK BERNARD

Si l'oiseau ne chante pas
c'est mauvais signe
signe que le tableau est mauvais
mais s'il chante c'est bon signe
signe que vous pouvez signer
Alors vous arrachez tout doucement
une des plumes de l'oiseau
et vous écrivez votre nom dans un coin du tableau.

J. Prévert, Pour Faire le Portrait d'un Oiseau.

Quel sera l'état du système solaire dans cinq cent millions d'années ?

On peut donner de l'importance à cette question en réalisant que de faibles variations des paramètres de l'orbite terrestre ont de désagréables conséquences climatiques. En un mot : l'excentricité nuit à votre santé ! Sur de telles échelles de temps, les connaissances actuelles ne permettent d'ailleurs pas totalement d'exclure des évènements bien plus spectaculaires comme une collision entre deux planètes. Il est vrai qu'à observer les civilisations humaines, en particulier celles qui se considèrent comme les plus évoluées, basculer dans la surenchère technologique et militaire et dans l'appauvrissement culturel et philosophique, à constater l'asservissement croissant des sciences et des techniques aux logiques de pouvoir et d'affrontement, la marginalisation des humanismes, la pérénité de la violence et de la pauvreté, le pillage des ressources naturelles au seul profit des vanités les plus futiles de minorités aveugles et arrogantes, on est en droit de douter que la dynamique céleste soit un facteur déterminant pour l'avenir de l'humanité. La Terre survivra à ses habitants et continuera à tourner.

Décrire et comprendre le mouvement des planètes, c'est avant tout répondre à une énigme qui s'est posée à la curiosité des hommes dès qu'ils ont observé le ciel nocturne. C'est tenter de prolonger une longue histoire qui a été une des motivations majeures au développement des mathématiques, mathématiques qui ensuite ont souvent pris leur envol et trouvé d'autres applications et d'autres liens à la réalité. C'est rajouter une poussière de connaissance à l'édifice scientifique et culturel, et, en dépit de la candeur de la question posée, prendre le risque (infime) de contribuer indirectement à la résolution violente et définitive de quelques conflits humains à venir, mais aussi une chance (non moins infime) de participer à l'élaboration d'un monde mieux civilisé. Enfin, plus pragmatiquement, c'est en ce qui me concerne remplir les conditions nécessaires à l'avancement de ma carrière académique.

Les succès de la mécanique céleste sont spectaculaires. On sait prédire l'avenir (et retrouver le passé) du système solaire sur plusieurs milliers d'années, on sait prévoir une éclipse à la minute près plusieurs années à l'avance (comparer aux prévisions météorologiques !) on sait jouer avec l'attraction gravitationnelle des planètes pour accélérer une sonde spatiale. Laskar explique l'écartement entre les planètes comme étant précisément l'écart nécessaire pour qu'elles ne se heurtent pas sur les longues périodes de temps. Le lecteur choqué par ce principe finaliste consultera [31] pour une explication plus évolutionniste. Confirmer une telle théorie serait un beau succès de la mécanique classique sur une question qui demeure largement énigmatique. Mais les mathématiques ne permettent pour l'instant pas de justifier rigoureusement ces jolies spéculations. En fait, les quelques succès des mathématiques de la dynamique céleste cachent une multitude de questions restées sans réponse. L'illusion de dominer cette science tient avant tout à ce que nos vies sont misérablement courtes à l'échelle des mouvements célestes, si bien que la complexité de ces derniers ne nous est accessible que par des mesures sophistiquées, indirectes, et peu spectaculaires. Pourtant, prévoir exactement l'état du système solaire dans cinq cent millions d'années reste aussi inaccessible à la science que de prévoir avec certitude le temps qu'il fera dans un mois. À la question " la Terre et Mars vont elles se heurter ?" il est possible que les mathématiques répondent un jour "non", mais il est peu probables qu'elles répondent jamais "oui". Par contre, nous pouvons tenter de répondre positivement à la question "est-il envisageable que la Terre et Mars se heurtent ?", ou plus précisément "peut on imaginer un système planétaire qui soit presque indistinguable du nôtre, mais dans lequel la Terre et Mars se heurteront ?". Alors, la décision serait prise par les multiples facteurs incontrôlables, par les petits chiffres significatifs des conditions initiales, par les imperfections du modèle. Les travaux que je présenterai dans ce petit texte vont dans le sens d'une réponse positive à cette question (au moins si l'on accorde au soleil une vie éternelle). Il est toutefois temps de préciser que l'application à la mécanique céleste des méthodes que je décrirai présente des difficultés particulières, que je n'ai pas du tout la prétention de savoir résoudre. Si la mécanique céleste est sans doute le champ d'application le plus poétique et le plus ancestral de le théorie qualitative des systèmes dynamiques, ce n'est ni le plus facile, ni le plus direct, ni, en général, le plus concluant.

Les systèmes physiques qui conservent l'énergie sont modélisés par des systèmes dynamiques hamiltoniens. Il est en général très difficile d'en décrire les solutions exactement, que ce soit par des expressions explicites ou par un calcul numérique. On cherche donc, depuis Poincaré, des informations qualitatives. Il est intéressant, par exemple de distinguer des comportements possibles et des comportements impossibles. On se demande ainsi, étant donnés deux états du système, si il est possible que le système évolue d'un voisinage du premier état vers un voisinage du second état. La question concernant la possibilité de collision entre la Terre et Mars est de cette nature.

Un autre exemple de ce type de question est de grande importance historique. La mécanique statistique consiste à décrire les assemblées composées de nombreuses particules, c'est à dire certains systèmes hamiltoniens à grand nombre de degrés de liberté, par quelques quantités macroscopiques. Le physicien Boltzmann avait justifié mathématiquement cette approche en supposant que les orbites du système remplissent équitablement tout l'espace des phases (ou plus précisément l'ensemble d'énergie constante, puisque le système préserve l'énergie), c'est à dire que pour la plupart des conditions initiales, le système visitera toute la surface d'énergie. C'est l'hypothèse ergodique de Boltzmann. Pendant longtemps, une question centrale de la mécanique hamiltonienne a consisté à justifier cette hypothèse, c'est à dire à montrer que *la plupart* des systèmes hamiltoniens à grand nombre de degrés de liberté la satisfont. Dans les années 50, des calculs numériques menés par Fermi, Pasta et Ulam d'un côté, et un important théorème de mathématiques découvert par Kolmogorov de l'autre côté (de l'océan pacifique) ont remis en question cette hypothèse. On sait maintenant qu'il existe des systèmes qui ne sont pas ergodiques, et qui ne peuvent pas être rendus ergodiques par une petite perturbation.

On peut cependant formuler une hypothèse plus faible, mais voisine. Étant donnés deux états d'un système hamiltonien, peut on trouver une orbite entre un voisinage de l'un et un voisinage de l'autre ? On dit qu'un système pour lequel la réponse est positive est transitif, ou bien qu'il satisfait l'hypothèse quasi-ergodique de Boltzmann. Il s'avère que les considérations qui permettent de rejeter l'hypothèse ergodique de Boltzmann ne rejettent pas l'hypothèse quasi-ergodique. La conjecture "la plupart des systèmes hamiltoniens à grand nombre de degrés de libertés sont transitifs" apparaît donc comme un bon paradigme scientifique. Cette conjecture est due à V.I. Arnold, qui a aussi donné des pistes pour la prouver. Pour finir sur deux réserves, précisons que le rapport entre la mécanique statistique et l'hypothèse quasi-ergodique de Boltzmann est loin d'être clair, et que la preuve de la conjecture de quasi-ergodicité dans toute sa généralité semble encore hors de portée. Il est donc intéressant de considérer des cas particuliers.

Il existe une classe particulière de systèmes hamiltoniens qui ne vérifient ni l'hypothèse ergodique ni l'hypothèse quasi-ergodique : les systèmes intégrables. On obtient un tel système en négligeant les interactions entre planètes dans notre système planétaire. L'évolution est alors extrêmement simple (il a cependant fallu des siècles de recherche pour arriver à cette conclusion). Chaque planète tourne de manière périodique sur une orbite elliptique immuable dont le soleil occupe un foyer. Dans un tel système, il est clair que l'hypothèse quasi-ergodique de Boltzmann n'est pas satisfaite. C'est en étudiant les perturbations de tels systèmes que l'hypothèse ergodique a été invalidée. En effet les perturbations de systèmes intégrables (que nous appellerons systèmes presque intégrables) ne sont pas ergodiques. Par contre l'argument qui permet d'exclure l'ergodicité (sur lequel nous reviendrons dans la suite) ne permet pas d'exclure la transitivité. Ils est donc naturel de se demander si les systèmes presque intégrables sont transitifs. Je vais présenter la dynamique des systèmes presque intégrables dans la perspective de cette question. Comme l'attraction qu'exerce les autres planètes sur une planète donnée est bien plus faible que l'attraction solaire, le système solaire peut être vu comme la perturbation du système intégrable obtenu en négligeant les interactions entre planètes. C'est donc un exemple de système presque intégrable. Je précise encore que c'est à plusieurs titres un mauvais exemple, qui ne vérifie pas les hypothèses que nous ferons dans la suite.

Je l'ai dit, le présent texte a pour objectif principal de me permettre d'obtenir le diplôme d'Habilitation à Diriger des Recherches. J'y présente donc certaines de mes activités de recherche, et je n'ai aucun scrupule à mettre largement en avant mes propres travaux par rapport à ceux, pourtant parfois d'une bien plus grande profondeur, de mes collègues, amis, concurrents et maîtres. J'adopterai d'ailleurs une nomenclature particulière pour les désigner. Dans la suite, [B*], où * est un symbole quelconque, désigne un article dont je suis signataire, alors que [n], où n est un entier naturel, désigne un article dont je ne suis pas signataire. Par exemple, [1] désigne un article essentiel de V. I. Arnold, qu'il considère lui même comme sa contribution majeure à la théorie des systèmes dynamiques, et dont les 4 pages sont sans doute à lire avant la centaine de pages de $\cup_*[B*]$. Je précise aussi que mes travaux s'appuient très largement sur ceux de John Mather et de Albert Fathi.

.

Cher Lucien, je le sais bien, que je te dois mon poste à Grenoble. Ton bureau est sans aucun doute la pièce de l'Institut Fourier où j'ai passé le plus de temps (après le mien, tout de même). J'y ai appris pas mal de choses, sur toutes les questions mathématiques qui me sont passées par la tête, sur la conjugaison du verbe avoir, sur le diable qui se cache dans les détails, et sur les dessous des commissions de recrutement. Une pièce toujours ouverte pour discuter est ce qui fait la différence entre un endroit où on s'ennuie et un endroit où on se plait. Tu es le symbole d'un laboratoire dans lequel je me suis senti parfaitement accueilli, largement protégé des affres d'une tâche d'enseignement chronophage, souvent ingrate, et parfois absurde, et chaleureusement encouragé dans mon travail de recherche. Je n'ai donc pas été surpris que tu acceptes de te plier en quatre pour venir à ma soutenance, qui ne pouvait pas décemment avoir lieu sans toi. Merci.

Cher Albert, je crois que tu es la seule personne ayant lu tous mes articles. Tu es donc un rempart à mes angoisses métaphysiques. Ton soutien m'a aussi protégé des problèmes beaucoup plus concrets dans lesquels se trouvent nombre de jeunes mathématiciens privés de poste. J'ai la reconnaissance du ventre. Merci d'être encore là pour cette soutenance.

Cher Raoul, je n'aurais pas osé usurper de ton temps pour un exposé dans lequel je ne résoudrai aucun des problèmes à un million (de dollards, ça ne vaut plus rien...) s'il avait neigé. Heureusement, la météo m'affranchit de tout scrupule. On se retrouvera en Janvier pour les choses sérieuses.

Cher Yves, je me rejouis d'avoir profité de tes dernières années d'exercice. Quand je pense qu'à ton age il me restera sans doute au moins dix ans... Enfin ce n'est pas encore fini, et je ne m'attends pas à ce que tu quittes la scène sans quelques come back flamboyants.

Cher Alain, tu es, dans ce jury, celui que je connais le moins. Je ne t'en suis que plus reconnaissant d'avoir accepté de me consacrer du temps. J'espere que ce bref passage à Grenoble ne te laissera pas une aussi mauvaise impression qu'un fameux hotel Radisson au pays de GW.

Cher Éric, maintenant que des centaines de kilomètres nous séparent, nous ne nous voyons plus aussi souvent. Je me rejouis que cette habilitation nous en donne une occasion. Cela me fait bien plaisir de constater que tu continues de t'intéresser à ma carrière mathématique, et j'espère que d'autres occasions nous permettrons de retravailler ensemble.

Cher Pierre, il y a presque dix ans maintenant, je suis rentré dans ton bureau, à Paris, alors que je cherchais mon chemin mathématique. Maintenant que je suis (presque) habilité, je n'ai plus besoin d'essayer de faire croire que j'ai compris ce que tu m'as dit ce jour-là sur les systèmes presque intégrables et la diffusion d'Arnold. Pourtant, les question que tu m'as alors posées motivent encore la grande majorité de mon travail, et je redoute le jour, visiblement assez proches, où elles seront résolues et où il me faudra trouver d'autres buts.

Dear Craig, thank you very much for writing this report about my work. Thank you also for having given my first opportunity to visit California, and to begin with its best place, the so called Democratic Republic of Berkeley.

Tout ceci manque un peu de féminité. Merci, aussi, à toutes et à tous les autres.

.

1. Moyennisation

Les penseurs selon lesquels les astres se meuvent de façon cyclique ne sont pas les plus profonds : qui regarde au-dedans de soi-même comme à l'intérieur d'un immense univers et porte en soi des voies lactées sait aussi combien irrégulières sont toutes les voies lactées : elles conduisent jusqu'au fond du chaos et du labyrinthe de l'existence.

F. Nietzsche *in* Le Gai Savoir

1.1. **Présentation.** Considérons un hamiltonien $H : \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T} \longrightarrow \mathbb{R}$ sous la forme

$$H(q,p,t) = h_0(p) + \epsilon H_1(q,p,t).$$

Il est préférable de considérer que H_1 dépend aussi de ϵ, mais nous n'écrivons pas explicitement cette dépendance. Si $\epsilon = 0$, on dit que le hamiltonien est totalement intégrable. Dans ce cas, les équations du mouvement sont

$$\dot{p} = 0, \dot{q} = \omega_0(p) := dh_0(p),$$

si bien que les variables p sont constantes au cours du mouvement, et que sur chaque tore invariant $\mathbb{T}^d \times \{p_0\} \times \mathbb{T}$, la dynamique est déterminée par le vecteur constant $\omega_0(p_0)$. Un tel système est loin d'être transitif. On cherche à décrire l'évolution des variables d'action $p \in \mathbb{R}^d$ lorsque ϵ est petit. Dans le cadre de la mécanique céleste, on peut penser que le terme H_1 correspond aux interactions entre planètes. La dépendance explicite en temps est introduite pour des raisons techniques, elle ne change pas fondamentalement la nature du problème. On peut aussi l'expliquer par exemple de la façon suivante. Dans le systèmes solaire, les grosses planètes Saturne et Jupiter ont des mouvement beaucoup plus réguliers que les petites planètes intérieures dont la Terre. Une bonne façon de modéliser le système est alors de déterminer dans un premier temps les orbites de ces grosses planètes, puis ensuite de considérer le système formé par les petites planètes sous l'influence de ces grosses planètes. Ce sous système non isolé sera alors modélisé par un hamiltonien dépendant explicitement du temps – il est vrai qu'alors c'est la périodicité qui n'est plus justifiée.

Une fréquence $\omega \in \mathbb{R}^d$ est dite résonante si il existe un vecteur $k \in \mathbb{Z}^d_*$ tel que $\langle k, \omega \rangle \in \mathbb{Z}$. On note ici $\mathbb{Z}^d_*$ pour $\mathbb{Z}^d - \{0\}$. Le module de résonance de ω,

$$Z(\omega) = \{k \in \mathbb{Z}^d / \langle k, \omega \rangle \in \mathbb{Z}\}$$

est un sous groupe de $\mathbb{Z}^d$, on note $R(\omega)$ l'espace vectoriel engendré par $Z(\omega)$ dans $\mathbb{R}^d$.

En raison de phénomènes de moyennisation, on peut décomposer l'évolution des variables d'action p en plusieurs parties :

- Une oscillation d'amplitude ϵ
- Une évolution *assez lente*, qui est assujettie à la contrainte $\dot{p} \in R(\omega_0(p))$, dont la vitesse est de l'ordre de ϵ, et qui est donnée par des formes normales "assez simples".
- Une évolution *très lente*, sans contrainte, correspondant aux restes des formes normales, de vitesse $\mu << \epsilon$ (souvent $\mu \leqslant e^{-C\epsilon^{-a}}$).

Notons en particulier que le mouvement *assez lent* est réduit à une oscillation d'amplitude ϵ si la fréquence ω n'est pas résonante. Or la plupart des fréquences sont non-résonantes (les fréquences non-résonantes forment un G_δ dense de mesure totale). Lorsque le hamiltonien

h_0 est non-linéaire, c'est à dire lorsque la fréquence $\omega_0(p)$ n'est pas constante, on s'attend au phénomène suivant : lorsque p évolue, alors $\omega_0(p)$ évolue aussi et devient non résonante, de sorte que $R(\omega_0(p))$ s'annule, et que l'évolution *assez lente* s'arrête. Cette description est extrêmement simpliste puisqu'elle met de côté des estimations quantitatives essentielles, mais le phénomène se produit effectivement pour la plupart des hamiltoniens non perturbés, c'est le contenu du théorème de Nekhoroshhev.

C'est en particulier le cas lorsque h_0 est convexe de Hessienne définie positive en chaque point. Nous donnerons un joli exemple de Herman qui illustre les pièges pouvant survenir lorsque h_0 n'est pas convexe.

Lorsque le théorème de Nekhoroshev s'applique, le mouvement *assez lent* est d'amplitude $o(1)$ (c'est à dire tendant vers 0 lorsque ϵ tend vers 0). L'éventuelle transitivité ne peut donc résulter que du mouvement très lent. L'évolution des variables d'action p est donc la superposition d'oscillations d'amplitude $o(1)$ et d'une éventuelle évolution *très lente* de vitesse $O(\mu)$.

On a même plus. Pour beaucoup de conditions initiales (correspondant à des fréquences "très non résonantes") l'évolution *très lente* est elle-même bornée, si bien que les variables p correspondant à ces trajectoires restent à distance $O(\epsilon)$ de leur valeur initiale. C'est une conséquence du théorème KAM.

1.2. **Fréquences.** Le rang r de $Z(\omega)$ est l'ordre de résonance de ω. L'exemple principal de résonance de rang r est constitué par les fréquences de la forme $\omega = (\omega_1, 0)$ où $\omega_1 \in \mathbb{R}^{d-r}$ est non résonante. Cet exemple est presque universel. En effet, si ω est une fréquence résonante, alors il existe une matrice $A \in Gl_d(\mathbb{Z})$ telle que $A\omega = (\omega_1, \omega_2)$, où $\omega_1 \in \mathbb{R}^{d-r}$ est non-résonant et où $\omega_2 \in \mathbb{R}^r$ est une fréquence rationnelle (c'est à dire dont toutes les composantes sont rationnelles). La matrice A donne lieu à un difféomorphisme de $\mathbb{T}^d$ qui transporte le champ de vecteur constant ω sur le champ de vecteur constant $A\omega$. Parmi les fréquences non résonantes, on distingue les fréquences diophantiennes. La fréquence ω est dite diophantienne si il existe des réels positifs C et τ tels que

$$d(\langle \omega, k \rangle, \mathbb{Z}) \geqslant C\|k\|^{-d(1+\tau)}$$

pour tout $k \in \mathbb{Z}^d_*$. L'ensemble des fréquences diophantiennes est à la fois négligeable au sens topologique, et prépondérant au sens de la mesure, c'est à dire que son complémentaire est de mesure de Lebesgue nulle dans $\mathbb{R}^n$. Pour étudier les résonances, il est utile de considérer les fréquences résonantes-diophantiennes. Ce sont les fréquences de la forme

$$\omega = A(\omega_1, \omega_2)$$

avec $A \in Gl_d(\mathbb{Z})$, $\omega_1 \in \mathbb{R}^{d-r}$ diophantienne, et $\omega_2 \in \mathbb{Q}^r$ rationnelle.

Les définitions que nous avons données sont basées sur une estimation des *petits diviseurs* $d(\langle \omega, k \rangle, \mathbb{Z})$. Pierre Lochak a montré dans [33] que d'autres approches peuvent être utilisées avec succès pour décrire les propriétés des fréquences utiles à la moyennisation. Par exemple, il a défini les *périodes* de ω comme étant la suite $T_i(\omega)$ des entiers définis par récurrence par $T_0(\omega) = 1$ et

$$T_i(\omega) = \min\{T \in \mathbb{N}, d(T\omega, \mathbb{Z}^d) < d(T_{i-1}(\omega)\omega, \mathbb{Z}^d)\}.$$

J'ai montré dans [B3] que, si il existe un réel positif C et un réel $\tau \in]0, 1/(d-1)[$ tels que

$$T_{i+1}(\omega) \leqslant C T_i(\omega)^{1+\tau},$$

alors la fréquence ω est résonante-diophantienne, elle est donc diophantienne si elle n'est pas résonante.

1.3. **Changements de variables.** Nous travaillons dans le contexte des hamiltoniens dépendant périodiquement du temps. Il faut pour moyenner commencer par définir une bonne notion de changement de variables, ce qui est légèrement délicat dans notre contexte. En effet, il est déplaisant de ne pas avoir, dans la définition qui suit, Φ^*H sous forme plus explicite. On note λ la forme de Liouville $\lambda = pdq$.

DÉFINITION 1.1. *Un changement de variables est un difféomorphisme $\Phi : T^*M \times \mathbb{T} \longrightarrow T^*M \times \mathbb{T}$ de le forme*

$$\Phi(z,t) = (\phi_t(z), t)$$

où chaque $\phi_t : T^*M \longrightarrow T^*M$ est un difféomorphisme exact. Si $H(z,t)$ est un hamiltonien, on définit le hamiltonien

$$\Phi^*H(z,t) = H \circ \Phi(z,t) + \partial_t S(z,t),$$

où $S(z,t)$ est une fonction sur $T^*M \times \mathbb{T}$ telle que $\phi_t^*\lambda - \lambda = \partial_z S$. Le changement de variables Φ conjugue alors le flot du flot du hamiltonien $H(z,t)$ et celui du hamiltonien $\Phi^*H(z,t)$.

En effet, on a alors

$$\Phi^*(\lambda - Hdt) = \phi_t^*\lambda - H \circ \Phi dt = \lambda - (H \circ \Phi + \partial_t S)dt + dS,$$

donc Φ^{-1} envoie les caractéristiques de la forme $\lambda - Hdt$, qui sont les orbites de H, sur les caractéristiques de la forme $\lambda - \Phi^*Hdt$, qui sont les orbites de Φ^*H □

1.4. **Cas linéaire.** Considérons le hamiltonien

$$H(q,p,t) = \langle \omega, p\rangle + \epsilon H_1(q,p,t).$$

Un hamiltonien $N(q,p)$ est dit sous forme normale si

$$\partial_q N \in R(\omega)$$

en chaque point.

THÉORÈME 1. *Si H est analytique, et si ω est résonante diophantienne, alors il existe un changement de variables analytique Φ, ϵ-proche de l'identité, tel que*

$$\Phi^*H(q,p,t) = \langle \omega, p\rangle + \epsilon N(q,p) + \mu(\epsilon)R(q,p,t)$$

où N est sous forme normale, de taille $O(1)$, et où R est un reste de taille $O(1)$. De plus, on a $\mu(\epsilon) \leqslant e^{-C\epsilon^{-a}}$.

La dynamique de Φ^*H est alors donnée par l'équation

$$\dot p = \partial_q(\Phi^*H) = \partial_q N + \partial_q R.$$

Le terme N engendre l'évolution *assez lente*, qui s'effectue à vitesse $O(\epsilon)$ dans le direction $\dot p \in R(\omega)$. Le terme R engendre une évolution *très lente*, à vitesse μ exponentiellement petite. Si ω est diophantienne, alors le terme N est constant, donc les variables p restent dans le voisinage de leurs valeurs initiales pendant un temps très long de l'ordre de $1/\mu$. Si ω est résonante-diophantienne, alors les variables p s'écartent en général linéairement de leur valeur initiale, à vitesse ϵ.

Pour fixer les idées, précisons un peu les choses dans le cas d'une fréquence ω de la forme $\omega = (\omega_1, 0) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, avec ω_1 diophantienne. Notons alors $q = (q_1, q_2) \in \mathbb{T}^{d_1} \times \mathbb{T}^{d_2}$, et

$p = (p_1, p_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Le hamiltonien N est sous forme normale si et seulement si il ne dépend pas de q_1. Les variables p_1 sont alors des intégrales premières de l'évolution du système *assez lent*

$$H = \langle \omega_1, p_1 \rangle + \epsilon N(q_2, p).$$

C'est une façon d'exprimer la contrainte $\dot{p} \in R(\omega)$, puisqu'ici $R(\omega) = \mathbb{R}^{d_2}$.

1.5. **Cas non-linéaire.** Considérons maintenant un hamiltonien non perturbé plus général $h_0(p)$. Le paragraphe précédent montre l'importance de la fréquence

$$\omega_0(p) := dh_0(p) : \mathbb{R}^d \longrightarrow \mathbb{R}^d.$$

Donnons un résultat modèle, pris dans [35], 2.1.4.

THÉORÈME 2. *Fixons p_0 de sorte que $\omega_0(p_0)$ est résonante-diophantienne. Soit B_ϵ une boule de $\mathbb{R}^d$, centrée en p_0, et de rayon $O(\sqrt{\epsilon})$. Alors il existe un changement de variables local*

$$\Phi_\epsilon : \mathbb{T}^d \times B_\epsilon \times \mathbb{T} \longrightarrow \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T},$$

qui est ϵ-proche de l'identité, tel que

$$\Phi^* H(q, p, t) = \langle \omega, p \rangle + \epsilon N(q, p) + \mu(\epsilon) R(q, p, t)$$

où N est sous forme normale, de taille $O(1)$, et où R est un reste de taille $O(1)$. De plus, on a $\mu(\epsilon) \leqslant e^{-C\epsilon^{-a}}$.

Ce résultat décrit la décomposition locale de la dynamique des variables p entre oscillations d'amplitude ϵ dues au changement de variables, évolution *assez lente* due au terme résonant et évolution *très lente* due au reste. L'évolution *assez lente* a lieu dans la direction $R(\omega)$, c'est à dire, comme ω n'est plus constante, qu'elle est soumise à la contrainte

$$\dot{p} \in R(\omega(p)).$$

On s'attend à ce que, dans de nombreux cas, cette contrainte interdise de grandes évolutions. Il ne faut pas oublier que l'on a $R(\omega) = 0$ pour un ensemble de mesure totale de fréquences ω. L'idée intuitive est la suivante. Lorsque $\omega(p)$ est résonant, alors l'évolution *assez lente* existe, c'est à dire qu'un transfert d'énergie a lieu entre les modes synchrones. Cependant, en raison de la non-linéarité, c'est à dire de la non constance de l'application $\omega(p)$, il est probable que les modes se désynchronisent lorsque ce transfert d'énergie a lieu. Cela se traduit par le fait que $\omega(p)$ quitte la résonance $\{\omega \in \mathbb{R}^d / Z(\omega) = Z(\omega_0(p_0))\}$ de sorte que $R(\omega)$ s'annule. Ces considérations sont trop simplistes pour rendre justice à la difficulté des preuves, pour lesquelles des analyses quantitatives très fines sont nécessaires. C'est Nekhoroshev qui a montré que le mouvement *assez lent* se réduit à des petites oscillations d'amplitude $o(1)$ pour la plupart des hamiltoniens h_0, [45]. La condition suffisante introduite pas Nekhoroshev sur le hamiltonien non perturbé est appelée raideur. Nous ne l'énoncerons pas ici, elle est assez compliquée, même si des travaux de Ilyashenko [27] et Niederman [47] l'ont largement clarifiée.

THÉORÈME 3. [45] *Si H est analytique et si la partie non perturbée satisfait la condition de raideur, alors le mouvement assez lent se réduit à des oscillations d'amplitudes $O(\epsilon^b)$, avec $b \in]0, 1[$. Autrement dit, il existe une fonction $\mu(\epsilon) \leqslant e^{-C\epsilon^{-a}}$ telle que l'inégalité $\|p(t) - p(t_0)\| \leqslant O(\epsilon^b)$ reste valable tant que $|t| \leqslant 1/\mu(\epsilon)$.*

Les hamiltoniens convexes satisfont la condition de raideur. Ils sont ceux pour lesquelles les constantes a et b dans l'énoncé sont les plus grandes.

Dans [33], Pierre Lochak a obtenu les estimations considérées comme (presque) optimales des exposants a et b dans le cas convexe. On peut alors prendre $a = b = 1/2d$. Il faut donc noter que les estimations sont d'autant meilleures que le nombre de degrés de liberté est petit, ce qui est très naturel. Pöschel [49] a retrouvé ces estimations par la méthode originale de Nekhoroshev, qui se simplifie fortement dans le cas convexe. Récemment, Niedermann [46] a considérablement amélioré l'estimation des exposants dans le cas d'un hamiltonien raide général en utilisant la méthode introduite par Lochak.

Donnons maintenant deux exemples simples. Le premier illustre les raisons pour lesquelles le mouvement lent est confiné dans le cas convexe, le second, dû à Michel Hermann, montre les pièges qui peuvent survenir lorsque h_0 ne satisfait pas la condition de raideur.

1.6. **Exemple d'Arnold tronqué.** Considérons le hamiltonien

$$H(q_1, q_2, p_1, p_2) = (p_1^2 + p_2^2)/2 + \epsilon \cos(2\pi q_2).$$

On a donc $\omega(p) = p$, et la perturbation correspond à une forme normale au voisinage des fréquences du type $(\omega_1, 0)$. Le système est le produit non couplé d'un oscillateur libre $H_1(q_1, p_1) = p_1^2/2$ et d'un pendule pesant $H_2(q_2, p_2) = p_2^2/2 + \epsilon \cos(2\pi q_2)$. Un argument élémentaire de conservation de l'énergie montre alors que la variation Δp_2 de la variables p_2 est majorée par $2\sqrt{\epsilon}$.

1.7. **Exemple de Herman.** [54] Considérons le hamiltonien

$$H(q_1, q_2, p_1, p_2) = p_1 p_2 + \epsilon V(q_1).$$

avec $V : \mathbb{T} \longrightarrow \mathbb{R}$. Les équations associées sont :

$$\dot{p}_2 = 0, \quad \dot{p}_1 = -dV, \quad \dot{q}_1 = p_2, \quad \dot{q}_2 = p_1.$$

Considérons une condition initiale vérifiant $p_2(0) = 0$ et $V'(q_1(0)) \neq 0$. La trajectoire correspondante est

$$p_1(t) = p_1(0) - tV'(q_1(0)), \quad p_2(t) = 0, \quad q_1(t) = q_1(0).$$

On a donc ici un mouvement *assez lent* (à vitesse ϵ) non borné, tel que les actions $p(t)$ s'écartent linéairement de leur position initiale, et ce malgré la non-linéarité. Pour comprendre ce phénomène, on remarque que $\omega(p_1, p_2) = (p_2, p_1)$. En conséquence, lorsque $p_2 = 0$, on a $\omega_1 = 0$ et donc

$$\mathbb{R} \times \{0\} \subset R(\omega(p_1, 0))$$

On peut donc créer un mouvement *assez lent* dans la direction $\dot{p} \in \mathbb{R} \times \{0\}$. Lors d'une telle évolution, le système ne quitte pas la résonance $p_2 = 0$, si bien que le mouvement assez lent peut se poursuivre indéfiniment.

Exercice : Comprendre pourquoi cet exemple ne contredit pas le théorème KAM. [1]

[1] La condition de non-dégénérescence isoénergétique n'est pas satisfaite.

1.8. **KAM.** Nous avons vu que le mouvement *assez lent* se réduit en général à des petites oscillations d'amplitude $o(1)$. Le fameux théorème suivant, dont les premières versions dues à Kolmogorov, [30], sont antérieures au théorème de Nekhoroshev, montre que, pour beaucoup de conditions initiales, l'évolution des variables d'action se réduit éternellement à des oscillations d'amplitude $O(\epsilon)$. Considérons le hamiltonien C^∞

$$H(q,p,t) = h(p) + \epsilon H_1(q,p,t).$$

THÉORÈME 4. [48],[9] *Il existe*

- *Un changement de variables Φ_ϵ, qui est ϵ-proche de l'identité,*
- *Une fonction $g_\epsilon : \mathbb{R}^d \longrightarrow \mathbb{R}$, qui sera aussi vue comme la fonction sur $\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T} = T^*\mathbb{T}^d \times \mathbb{T}$ donnée par $(q,p,t) \longmapsto g_\epsilon(p)$,*
- *Un fermé F_ϵ de $\mathbb{R}^d$, qui est non-vide lorsque ϵ est assez petit,*

tels que le hamiltonien $\Phi_\epsilon^* H$ a un contact d'ordre infini avec g_ϵ en chaque point de $\mathbb{T}^d \times F_\epsilon \times \mathbb{T}$. En conséquence, chaque tore $\Phi_\epsilon(\mathbb{T}^d \times \{f\} \times \mathbb{T})$, avec $f \in F_\epsilon$, est invariant par le flot de H. Le changement de variables Φ_ϵ conjugue le flot de g sur $\mathbb{T}^d \times F_\epsilon \times \mathbb{T}$ avec le flot de H sur $\Phi_\epsilon(\mathbb{T}^d \times F_\epsilon \times \mathbb{T})$.

Pour ϵ assez petit, l'ensemble F_ϵ est de mesure positive mais d'intérieur vide, si bien que l'ensemble fermé invariant

$$\mathcal{F}_\epsilon = \Phi_\epsilon(\mathbb{T}^d \times F_\epsilon \times \mathbb{T})$$

est de mesure positive mais de complémentaire dense.

C'est ce théorème qui exclut l'hypothèse ergodique de Boltzmann, puisqu'il donne l'existence d'un ensemble invariant de mesure positive. En revanche, comme le complémentaire de cet ensemble est un ouvert dense, la transitivité n'est pas remise en question. En fait l'ensemble invariant $\mathcal{F}_\epsilon$ est important au sens de la mesure, mais négligeable au sens topologique, ce qui est un pendant direct de la propriété similaire satisfaite par les fréquences diophantiennes.

Lorsque $d = 1$, le complémentaire de $\mathcal{F}_\epsilon$ est découpé par $\mathcal{F}_\epsilon$ en petites composantes connexes. On a donc stabilité perpétuelle des variables d'action dans ce cas, et l'hypothèse quasi-ergodique est invalidée.

Toutefois, dès que $d \geqslant 2$, le complémentaire de $\mathcal{F}_\epsilon$ est un ouvert dense connexe. La dynamique est alors transitive sur tout l'espace des phases si elle est transitive sur cet ouvert connexe (ce qui, bien sûr, reste à établir).

1.9. **Conclusion.** Il y a trois raisons pour considérer les hamiltoniens h_0 convexes.

- Parce que ce cas comprend la situation "naturelle" où $h_0 = |p|^2/2$ est une énergie cinétique, et où $H_1 = V(q,t)$ est une énergie potentielle.
- Parce que les hamiltoniens convexes sont ceux pour lesquels le théorème de Nekhoroshev donne les meilleures estimées, ce sont donc d'une certaine façon les cas où la transitivité est la plus lente.
- Parce que nous allons utiliser des méthodes variationnelles qui pour l'instant ne peuvent être mises en place que dans ce cadre.

L'évolution locale au voisinage de $\{p = p_0\}$ est régie par des hamiltoniens de la forme

$$\Phi^* H = h_0(p) + \epsilon N(q,p) + \mu(\epsilon) R(q,p,t),$$

où N est sous forme normale, c'est à dire satisfait $\partial_q N \in R(\omega(p_0))$.

Le théorème de Nekhoroshev nous dit alors qu'il n'est pas possible d'obtenir une évolution de taille $O(1)$ des actions en considérant seulement le mouvement *assez lent* engendré par le hamiltonien tronqué $h_0(p) + \epsilon N(q,p)$. Ceci a deux conséquences :

– Toute évolution d'ordre $O(1)$ des variables d'action p est *très lente*, c'est à dire se fait à vitesse $O(\mu)$.
– Pour mettre en évidence une évolution de taille $O(1)$ des variables d'action, il faudra tenir compte des restes, sur lesquels on ne sait pas grand chose.

Résumons quelques questions naturelles ouvertes à ce stade.

La question qui nous intéressera principalement est de comprendre si la transitivité topologique est une propriété possible dans ce contexte, et si c'est une propriété "fréquente", c'est à dire si elle se produit pour une perturbation typique. Plus généralement, on cherchera à décrire la dynamique topologique des variables d'action. À l'heure actuelle, le résultat le plus intéressant est sans conteste celui qu'a annoncé John Mather dans [44] : Pour $d = 2$, et pour une perturbation typique, les variables d'action ont des évolutions de taille $O(1)$. Je pense que les résultats que j'ai obtenus dans [B8] auront aussi des applications déterminantes dans ce contexte, et qu'ils permettront d'étendre et de clarifier les travaux remarquables de John Mather.

Dans le contexte analytique, des questions plus qualitatives sont aussi posées par le théorème de Nekhoroshev. Il s'agit de comprendre quelle est la vitesse maximale d'évolution des variables d'action, c'est à dire quel est l'exposant optimal a dans l'énoncé du théorème. Pour majorer l'exposant a, il faut considérer des perturbations pour lesquelles on arrive à minorer la vitesse d'évolution des variables d'action. On peut considérer cette question comme presque résolue, du côté de la minoration par les travaux de Pierre Lochak et les optimisations qui ont suivies, et du côté de la majoration par la construction récente, exploitant des germes semés par Michel Herman avant son départ trop précoce, par Marco et Sauzin dans [39], puis Lochak et Marco [34], de superbes nouveaux exemples. Pour résumer, l'exposant optimal satisfait

$$\frac{1}{2d} \leqslant a \leqslant \frac{1}{2d-4}.$$

Il faut remarquer que la minoration des vitesses d'évolution des variables d'action sur des exemples explicites est une très belle question qui a motivé de nombreux travaux. On peut commencer par le travail fondateur de Arnold, puis penser au traitement variationnel par U. Bessi [5], [6] de l'exemple d'Arnold, qui est sans doute le premier à obtenir des estimations raisonnables. J'ai utilisé ces méthodes dans [B1] pour donner la première estimation polynomiale dans le cas initialement hyperbolique, nous y reviendrons. Il est toutefois utile de préciser que tous ces travaux font appel à des structures particulières des exemples choisis. De manière générale, à ce stade de leur développement, les méthode de construction qui permettre d'obtenir des estimations précises des vitesses sont des méthodes qui ne se généralisent pas au cas de perturbations générales. Inversement, les méthodes récemment proposées pour traiter des perturbations générales, sur lesquelles nous reviendrons, ne fournissent pas d'estimation des vitesses. Ceci ouvre donc une question intermédiaire entre les précédentes :

Quelle est la vitesse d'évolution des variables d'action pour une perturbation typique, est-elle du même ordre que pour une perturbation ad hoc, ou est-elle bien plus faible ?

2. Diffusion d'Arnold

Une société qui obéirait à une législation émanant d'une académie scientifique, non parce qu'elle en aurait compris elle-même le caractère rationnel auquel cas l'existence de l'académie deviendrait inutile, mais parce que cette législation émanant de la science s'imposerait au nom de la science qu'on vénérerait sans la comprendre, une telle société serait une société non d'hommes, mais de brutes.

M. Bakounine

De la section précédente, on retiendra que la dynamique au voisinage d'une valeur $\{p = p_0\}$ des variables d'action est régie par un hamiltonien de la forme

$$H = h_0(p) + \epsilon N(q,p) + \mu R(q,p,t),$$

où $\partial_q N \in R(\omega_0(p_0))$, et où μ est très petit devant ϵ. Ceci donne tout son poids à l'exemple introduit par Arnold dans [1] :

$$H = (p_1^2 + p_2^2)/2 + \epsilon \cos 2\pi q_2 + \mu(\cos 2\pi q_2)(\cos 2\pi q_2 + \cos 2\pi t).$$

Théorème 5. [1] *Fixons des réels (indépendants de ϵ) $0 < A < B$. Il existe une fonction strictement positive $\mu_0(\epsilon)$ telle que, pour $0 < \mu < \mu_0(\epsilon)$, il existe une trajectoire*

$$(q_1(t), q_2(t), p_1(t), p_2(t))$$

et un temps $T > 0$ tel que

$$p_1(0) \leqslant A \ ,\ p_1(T) \geqslant B.$$

On peut prendre $\mu_0(\epsilon) \approx e^{-\sqrt{\epsilon}}$.

2.1. **Traitement de l'exemple d'Arnold.**

2.1.1. *Le système tronqué.* Commençons par quelques remarques sur le hamiltonien tronqué obtenu pour $\mu = 0$:

$$H_0(q,p) = H_1(q_1,p_1) + H_2(q_2,p_2) = p_1^2/2 + p_2^2/2 + \epsilon \cos 2\pi q_2.$$

Ce système est le produit non couplé de l'oscillateur libre de hamiltonien H_1 et du pendule pesant de hamiltonien H_2. La variable p_1 est donc une intégrale première, et la conclusion du Théorème est donc exclue pour $\mu = 0$.

Rappelons que le point $q_2 = 0, p_2 = 0$ est un point fixe hyperbolique du pendule pesant $H_2(q_2,p_2) = p_2^2/2 + \epsilon \cos 2\pi q_2$. Les variétés stable et instable de ce système intégrable coïncident, elles forment le niveau d'énergie $H_2 = \epsilon$. En conséquence, dans le système produit de hamiltonien $H_0 = H_1 + H_2$, il y a une famille à un paramètre T_ω de cercles invariants,

$$T_\omega = \{p_1 = \omega, q_2 = 0, p_2 = 0\} \subset \mathbb{T}^2 \times \mathbb{R}^2.$$

Chacun de ces cercles invariants est hyperbolique au sens qu'il admet une variété stable de dimension 2 et une variété instable de dimension 2, qui sont simplement les relevées des variétés stable et instable du point fixe hyperbolique de H_2. Il faut remarquer cependant que ces variétés ne s'intersectent pas transversalement le long de T_ω, si bien qu'il ne s'agit que d'une hyperbolicité partielle. Comme on va introduire une perturbation dépendant explicitement du temps, il est utile de relever les cercles invariants à l'espace des phases étendu et de considérer les tores $T_\omega \times \mathbb{T} \subset \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T}$, que nous noterons encore T_ω. Les variétés stables et instables de ces tores invariants de l'espace de phase étendu (qui sont de dimension 2) sont de dimension 3.

2.1.2. *Transversalité des variétés invariantes.* Lorsque $\mu \neq 0$, la perturbation est choisie de sorte que les tores T_ω sont laissés invariants par le flot du hamiltonien total H. Cependant les variétés stables et instables de ces tores invariants sont perturbées, et elle n'ont plus aucune raison de coïncider. On a effectivement :

PROPOSITION 1. *Il existe une fonction $\mu_0(\epsilon) > 0$ telle que, si $0 < \mu < \mu_0$, alors les tores invariants T_ω du système total H admettent des variétés stables et instables de dimension 3, qui se coupent transversalement dans $\mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T}$ le long d'une trajectoire homocline à T_ω. On peut prendre $\mu_0 \approx e^{-\sqrt{\epsilon}}$.*

Quelques commentaires sur la preuve : Il faut d'abord vérifier que les tores demeurent hyperboliques, et que leurs variétés stables et instables ne sont que perturbées par l'introduction du reste. Ceci résulte directement de l'observation que la variété M formée du l'union des tores invariants T_ω est normalement hyperbolique au sens de [26]. Elle demeure donc normalement hyperbolique sous l'effet d'une petite perturbation, et ses variétés stable et instable ne sont que légèrement déformées. Pour faire fonctionner cette étape, il suffit de prendre $\mu = o(\epsilon)$.

La question de l'existence d'une orbite homocline est alors bien comprise. Elle est très simple dans ce cadre, et est généralisée dans [18], [7] et [B2]. Il s'agit d'une propriété globale, qui ne demande aucune condition de petitesse de μ.

Le point clef, qui nécessite l'exponentielle petitesse de μ, est la transversalité. Comme la transversalité est un phénomène générique, on pourrait penser que cette étape n'est pas si cruciale. En pratique, la transversalité ne peut être montrée sur l'exemple explicite que pour $\mu \leqslant e^{-\sqrt{\epsilon}}$, et c'est donc ici qu'intervient cette condition dans la preuve. En voici la raison. Suivant Poincaré et Melnikov, Arnold a démontré la transversalité en écrivant un développement en puissances de μ des variétés stable et instable. Il s'avère que le premier terme de ce développement peut être calculé par une intégrale explicite. Très grossièrement, on obtient alors un développement de l'angle entre les variétés (dont on cherche à montrer la non-nullité) de la forme

$$\alpha = \alpha_1 \mu + O(\mu^2).$$

Le coefficient α_1 peut être calculé explicitement, et il est de l'ordre de $e^{-\sqrt{\epsilon}}$. Le développement permet donc de conclure la non-nullité de α lorsque μ est petit devant cette quantité, de sorte que le terme $O(\mu^2)$ est petit devant le terme principal $\alpha_1\mu$. Même si cette preuve ne permet pas de l'affirmer, il est fort probable que la transversalité reste valable pour la plupart des valeurs de $\mu \leqslant O(\epsilon)$ même au-delà de $\mu_0(\epsilon)$.

Comme nous allons le voir, l'ordre de grandeur de l'angle entre les variétés stable et instable est fortement relié à l'existence et à la vitesse de l'évolution des variables d'action. Pour cette raison, il a donné lieu à de nombreux travaux, citons [35] pour une plus ample discussion.

2.1.3. *Chaîne de transition.* On vient d'établir l'existence, pour le système total avec μ assez petit, d'une famille T_ω de tores invariants hyperboliques, tels que les variétés stable W^+_ω et instable W^-_ω s'intersectent transversalement le long d'une orbite homocline (mais pas le long du tore T_ω !) pour chaque ω.

Un argument de persistance des intersections transverses montre alors que la variété instable $W^-_{\omega_0}$ du tore T_{ω_0} intersecte transversalement la variété stable W^+_ω du tore T_ω lorsque ω est assez proche de ω_0 (l'aspect quantitatif dépend ici directement de la valeur de l'angle de transversalité évoqué précédemment). On obtient donc des orbites hétéroclites entre tores voisins. Comme le tore T_ω est contenu dans $\{p_1 = \omega\}$, on a ainsi fait bouger la variable p_1. Il reste un joli argument de globalisation pour obtenir une évolution plus ample de cette variable.

Si on se donne deux valeurs ω et ω' de la variable p_1, alors on peut trouver, au vu de ce qui précède, une suite $\omega_i, 1 \leqslant i \leqslant N$ telle que $\omega_0 = \omega$, $\omega_N = \omega'$, et telle que W_i^- intersecte transversalement W_{i+1}^+ pour tout i. La famille de tores T_{ω_i} correspondante est appelée une chaîne de transition.

Un argument topologique simple permet alors à Arnold de prouver l'existence d'orbites joignant un voisinage du tore T_ω à un voisinage du tore $T_{\omega'}$. Ceci termine la démonstration du théorème pour peu que l'on ait choisi $\omega < A < B < \omega'$.

La dynamique associée à une chaîne de transition n'est pas complètement élémentaire. Pour donner une estimation de la vitesse d'évolution de la variable p_1, il est nécessaire de raffiner l'argument initial de Arnold. Ces raffinements ont été donnés bien ultérieurement, par exemple dans [38] et [15]. Ces travaux étudient dans quelle mesure la dynamique associée à une intersection homocline transverse des variétés stable et instable d'un tore invariant hyperbolique ressemble à la dynamique bien connue associée à l'intersection transverse des variétés stable et instable d'un point fixe hyperbolique, et en particulier ils généralisent le λ-lemme à ce contexte. Comme on l'imagine, on peut alors montrer que, dans une chaîne de transition, la variété instable du premier tore, W_0^-, intersecte transversalement la variété stable du dernier tore, W_N^+, et estimer la vitesses d'évolution de la variable p_1. Par exemple, les résultats d'estimation polynomiale des vitesses dans le cas initialement hyperbolique (c'est à dire lorsque $\epsilon = 1$) que j'avais obtenues variationnellement dans [B1] ont été retrouvées par ce type de méthode dans [16].

2.1.4. *Remarque sur la vitesse.* Une idée qui a rapidement fait son chemin est que la diffusion d'Arnold, c'est à dire l'évolution des variable d'actions, est un phénomène exponentiellement lent. Cela signifie que, dans l'exemple d'Arnold, et pour $\mu \leqslant O(\epsilon)$, l'évolution, si elle a lieu, de la variable p_1 se fait à vitesse au plus $e^{-\epsilon^{-a}}$.

Cependant, on peut fixer ϵ, et étudier cette vitesse en fonction du second paramètre μ. C'est ce que l'on appelle le cas a priori instable, ou initialement hyperbolique. Dans ce cas, la vitesse de diffusion est polynomiale. Plus précisément, elle est de l'ordre de $-\mu/\ln(\mu)$, comme ceci est par exemple montré dans [3] (j'avais préalablement obtenu la minoration μ^2 dans [B1]). L'idée suivant laquelle la vitesse est polynomiale dans le cas a priori instable a eu quelques difficultés à s'imposer. Je crois avoir contribué à cette prise de conscience grâce à mon article [B1], qui est le premier où une vitesse polynomiale est obtenue. Le mérite en revient surtout à Pierre Lochak, qui avait prédit le phénomène, et m'a téléguidé vers ce beau résultat.

2.2. **Universalité et particularités de l'exemple d'Arnold, généralisations.** Considérons une perturbation du hamiltonien $h_0(p) = \|p\|^2/2$. On a alors $\omega_0(p) = p$, et on est ramené par un changement de variable moyennisant au voisinage de $p_0 = (\omega_1, 0)$ à

$$H = \|p\|^2/2 + \epsilon N(q_2, p) + \mu R(q, p, t)$$

Pour simplifier, on translate les variables d'action, ce qui revient à étudier

$$H = \langle \omega_1, p_1 \rangle + \|p\|^2/2 + \epsilon N(q_2, p) + \mu R(q, p, t)$$

au voisinage de $p_0 = 0$. En notant $V(q_2)$ la fonction $N(q_2, 0)$, on obtient

$$H(q,p,t) = \langle \omega_1, p_1 \rangle + \|p_1\|^2/2 + \|p_2\|^2/2 + \epsilon V(q_2) + \epsilon O(\|p\|) + O(\|p\|^3) + \mu R,$$

et les termes principaux de notre hamiltonien sous forme normale constituent le produit d'un oscillateur $H_1(q_1, p_1) = \langle \omega_1, p_1 \rangle + \|p_1\|^2/2$ et d'un système naturel $H_2(q_2, p_2) = \|p_2\|^2/2 + V(q_2)$. On voit donc apparaître la structure de l'exemple d'Arnold d'une manière très générale.

Cet exemple a toutefois des spécificités :

La première est liée au choix des dimensions. On a $d_2 = 1$. Ceci permet au système H_2 d'être intégrables, et rend possible une étude relativement explicite des problèmes de transversalité.

La seconde est que la fonction V choisie, en l'occurrence $V(q_2) = \cos 2\pi q_2$, admet un unique maximum non dégénéré. C'est ce maximum qui donne le point fixe hyperbolique de H_2 sur lequel toute la construction est basée. Cette particularité peut sembler innocente dans la mesure où c'est une propriété générique du potentiel V. Toutefois, pour étudier globalement la diffusion, il faudra considérer des familles de formes normales, et donc des familles de fonctions V. Or la propriété mentionnée est de codimension 1, et n'est donc pas générique dans une famille.

La troisième particularité de l'exemple d'Arnold est de loin la plus importante : la perturbation R est choisie de manière à ne pas perturber les tores invariants hyperboliques. Ceci est bien sur très spécifique, et il est difficile de faire marcher la méthode avec une perturbation plus générale. Pour une perturbation générale, la variété M formée par l'union des tores T_ω, et qui est normalement hyperbolique, serait déformée mais pas détruite. Cela signifie qu'il y aurait dans le système perturbé une variété invariante normalement hyperbolique M_μ proche de M. Par contre, et c'est la le point clef, la variété M_μ ne serait pas feuilletée en tores invariants du hamiltonien perturbé. Une généralisation naturelle et bien plus universelle de l'exemple d'Arnold est donc l'exemple suivant :

$$H(q,p,t) = \langle \omega_1, p_1 \rangle + \|p_1\|^2/2 + \|p_2\|^2/2 + \epsilon V(q_2) + \mu R(q, p_1, t). \tag{1}$$

Le reste R est ici choisi indépendant de q_2 de façon à ce que la variété $M = \{p_2 = q_2 = 0\}$ soit globalement préservée, ce qui est une hypothèse simplificatrice sans grande portée. Par contre, la dynamique interne à cette variété est perturbée, de sorte que les tores T_ω ne sont pas invariants par le hamiltonien H. La dynamique interne à la variété invariante $M = \mathbb{T} \times \mathbb{R} \times \mathbb{T}$ est donnée par la restriction du hamiltonien,

$$H(q_1, 0, p_1, 0, t) = \langle \omega_1, p_1 \rangle + \|p_1\|^2/2 + \mu R(q_1, 0, p_1, t).$$

C'est donc une perturbation quelconque de la dynamique complètement intégrable de $\|p_1\|^2/2$. Le théorème KAM s'applique, voir [25] pour la première application à ce type de systèmes, il donne l'existence d'une famille de tores invariants, mais pas d'un feuilletage : il y a des trous entre les tores invariants. Sans rentrer dans les détails, on comprend bien que l'existence ou non de chaînes de transition dépend alors d'une compétition sur les ordres de grandeurs, entre d'un côté l'angle de transversalité entre les variétés stable et instable d'un tore donné, et de l'autre l'écart entre deux tores voisins. Or ce calcul d'ordre de grandeur est défavorable : on ne peut pas en général construire de chaînes de transitions au sens d'Arnold dans l'exemple généralisé, sauf dans certains cas spécifiques, comme par exemple dans [8].

Obtenir les conclusions du théorème d'Arnold pour le hamiltonien généralisé (1), par exemple en supposant que $d_1 = 1 = d_2$ et que le potentiel $V(q_2)$ admet un unique maximum non dégénéré, a longtemps été considéré comme un problème majeur. C'est le *Large*

Gap Problem. Précisons toutefois que beaucoup de travail a été nécessaire pour identifier et isoler clairement cette difficulté.

Ce problème peut être considéré comme résolu grâce à de nombreux travaux récents, [17], [51], [10], [11] [52], [53], mais aussi [B8]. On résumera ces travaux en disant qu'ils prouvent l'énoncé modèle suivant.

THÉORÈME 6. *Supposons que $d_1 = d_2 = 1$, que le potentiel V a un unique maximum non dégénéré, et que certaines conditions de non-dégénérescence supplémentaire sont satisfaites par la perturbation. Alors, les conclusions du théorème d'Arnold sont valables lorsque $\mu > 0$ est assez petit par rapport à ϵ.*

Ils est sans doute utile de donner quelques précisions de nature historique. L'approche *géométrique* d'Arnold a été poursuivie. Notre présentation de l'exemple d'Arnold bénéficie d'ailleurs de ces travaux ultérieurs, qui ont mis en lumière le caractère normalement hyperbolique de la variété M (cette notion est postérieure à l'article d'Arnold). Dans le remarquable article [17] (et [51] ?), des nouveaux types de tores invariants (dit secondaires) sont considérés. Une étude fine basée sur la moyennisation et sur le théorème KAM a alors permis aux auteurs de construire des chaînes de transitions généralisées impliquant ces tores secondaires. C'est une des solutions du *Large Gap Problem*. Son mérite est qu'elle met clairement en évidence un mécanisme géométrique proche de celui d'Arnold. De plus, elle ne nécessite pas la convexité du hamiltonien. Par contre, elle est très lourde à mettre en place techniquement, et semble impossible à étendre à des contextes moins perturbatifs. Il faut remarquer que le passage des chaînes de transitions aux chaînes de transitions généralisées fait perdre toute estimation des vitesses. Il semble que, bien qu'elle soit plus géométriquement explicite, cette méthode ne soit pas la plus prometteuse pour estimer les vitesses. La méthode résumée dans [28] permet aussi peut être de traiter le hamiltonien (1).

Parallèlement à ces travaux, des méthodes variationnelles se sont développées sous l'impulsion principale de John Mather, [42], mais aussi de Ugo Bessi qui, dans [5] et [6], a obtenu les premières applications des méthodes variationnelles à des cas concrets de diffusion d'Arnold. Il étudie l'exemple d'Arnold, fournit une preuve très élémentaire (et qui est probablement la première preuve complète) du théorème principal de [1], et obtient les premières estimations raisonnables des vitesses. C'est d'ailleurs sa méthode que j'utilise dans [B1] pour obtenir une estimation polynomiale dans le cas instable. Cependant, malgré des développements intéressants, comme dans [4], les méthodes issues des travaux de Bessi ne permettent toujours pas de résoudre le *Large Gap Problem*, sans doute en raison du principe qui veut qu'une méthode trop explicite est vouée à l'échec dans l'étude du hamiltonien (1).

L'approche de Mather est sans doute plus universelle. Il l'a peu a peu raffinée dans plusieurs travaux essentiellement non publiés, comme [43], jusqu'à des résultats remarquables annoncés dans [44], qui constituent sans aucun doute l'avancée la plus profonde dans ce sujet. J'espère que mon article [B8] permettra de clarifier les preuves de ces résultats, et de les étendre. Pour revenir au *Large Gap Problem*, Xia a annoncé une solution dans le succinct [52] dès 1998. Les détails, donnés plus tard dans [53], s'appuient sur [42] mais aussi sur des travaux antérieurs de Mather, [40], spécifiques au cas $d = 1$. Le superbe article [42] de John Mather est très elliptique sur certains points. Je me suis efforcé de la clarifier dans [B4]. Ma motivation initiale était de comprendre pourquoi les résultats de [42] ne sont pas optimaux quand $d = 1$ (fait remarqué par l'auteur lui même). J'ai obtenu un résultat proche de celui de Mather, mais optimal en dimension $d = 1$, et j'ai précisé certains points laissés obscurs dans [42]. Ce travail de clarification a été utilisé dans [10] puis [11], qui proposent une autre solution au *Large Gap Problem.*

Les méthodes que je développe dans [B8], sur lesquelles je reviendrai, s'appliquent elles aussi au *Large Gap Problem*, et permettent de montrer le théorème plus général :

THÉORÈME 7. [B8] *Considérons le hamiltonien (1), avec $d_1 = 1$, où V admet un nombre fini de maxima. Supposons de plus que certaines conditions de non-dégénerescence sont satisfaitespar la perturbation, et que μ est assez petit devant ϵ. Alors, pour toute suite biinfinie $P_i \in \mathbb{R}, i \in \mathbb{Z}$, il existe une suite croissante de temps $t_i, i \in \mathbb{Z}$ telle que*

$$p_1(t_i) = P_i.$$

Les améliorations offertes par ce théorème sont les suivantes :

On y autorise toute valeur de d_2, et donc de $d = d_1 + d_2$ (par contre, la contrainte $d_1 = 1$ est essentielle). Dans les autres traitements du hamiltonien (1) qui sont disponibles, on a $d_1 = d_2 = 1$.

On autorise des potentiels V a plusieurs maxima, ce qui est essentiel en vue de l'application aux systèmes presque intégrables généraux (car il faudra considérer des familles de hamiltoniens sous forme normale).

On obtient une conclusion qui est plus fine que celles qui sont accessibles par les autres méthodes.

Je rappelle pour finir que les méthodes de [B8], comme celles de [42] et de [B4], ne sont pas intrinsèquement perturbatives, d'autres types d'applications sont donc possibles. Par contre, elles reposent crutialement sur la convexité du hamiltonien non perturbé h_0.

3. La fonction α de Mather.

La première fois qu'Aurélien vit Bérénice, il la trouva franchement laide.

Aragon, *in* Aurélien

3.1. **Le contexte.** Nous allons maintenant aborder plus en détails l'approche variationnelle. La présente section sera plus précise que les précédentes, et consistera principalement en la description sous toutes ses facettes du premier objet de cette théorie, la fonction α de Mather. Commençons par rappeler les hypothèses standard, introduites par Mather dans [41]. Les hamiltoniens $H : T^*\mathbb{T}^d \times \mathbb{T} \longrightarrow \mathbb{R}$ seront supposés C^2. On note λ la forme de Liouville $\lambda = pdq$, et $X(q,p,t)$ le champ de vecteur hamiltonien associé à H. On supposera toujours que les hypothèses suivantes sont vérifiées :

(1) Complétude. Le champ de vecteurs hamiltonien X engendre un flot complet sur $\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T}$.

(2) Convexité. Pour tout (q,t) dans $\mathbb{T}^d \times \mathbb{T}$, la fonction $p \longrightarrow H(q,p,t)$ est convexe sur $\mathbb{R}^d$, de Hessienne définie positive, c'est à dire que $\partial_p^2 H > 0$.

(3) Surlinéarité. Pour tout (q,t) dans $\mathbb{T}^d \times \mathbb{T}$, on a $\lim_{\|p\| \longrightarrow \infty} H(q,p,t)/\|p\| = \infty$.

On associe au hamiltonien H le Lagrangien $L \in C^2(T\mathbb{T}^d \times \mathbb{T}, \mathbb{R})$ donné par

$$L(q,v,t) = \sup_{p \in \mathbb{R}^d} \langle p, v \rangle - H(q,p,t).$$

Les applications

$$\partial_p H : T^*\mathbb{T}^d \times \mathbb{T} \longrightarrow T\mathbb{T}^d \times \mathbb{T}$$

et

$$\partial_v L : T\mathbb{T}^d \times \mathbb{T} \longrightarrow T^*\mathbb{T}^d \times \mathbb{T}$$

sont des difféomorphismes inverses, on les appelle transformées de Legendre. On a $L(q,v,t) = \partial_v L(q,v,t)(v) - H(q, \partial_v L(t,x,v), t)$ et $H(q,p,t) = \partial_p H(q,p,t)(p) - L(q, \partial_p H(t,x,v), t)$. Le Lagrangien L vérifie

(1) Convexité. Pour tout (q,t) dans $\mathbb{T}^d \times \mathbb{T}$, la fonction $v \longrightarrow L(q,v,t)$ est convexe sur $\mathbb{R}^d$, de Hessienne définie positive, c'est à dire que $\partial_v^2 L > 0$.

(2) Surlinéarité. Pour tout (q,t) dans $\mathbb{T}^d \times \mathbb{T}$, on a $\lim_{\|v\| \longrightarrow \infty} L(q,v,t)/\|v\| = \infty$.

3.2. **Mesures.** Commençons par donner la définition originale de Mather. Soit μ une mesure de probabilité sur $T\mathbb{T}^d \times \mathbb{T}$. La mesure (Borélienne) μ est dite invariante si la mesure μ^* sur $T^*\mathbb{T}^d \times \mathbb{T}$ obtenue en transportant μ par la transformée de Legendre est invariante par le flot hamiltonien. La mesure μ est dite fermée si elle vérifie

$$\int_{T\mathbb{T}^d \times \mathbb{T}} \partial_q f(v) + \partial_t f d\mu = 0$$

pour toute fonction $f \in C^\infty(\mathbb{T}^d \times \mathbb{T}, \mathbb{R})$. John Mather a montré dans [41] que les mesures invariantes sont fermées. On définit la classe d'homologie $[\mu] \in H^1(\mathbb{T}^d, \mathbb{R}) = \mathbb{R}^d$ d'une mesure fermée μ par

$$\langle [\mu], [\omega] \rangle = \int_{T\mathbb{T}^d \times \mathbb{T}} \omega d\mu,$$

pour toute forme fermée ω sur $\mathbb{T}^d$ (vue comme une fonction sur $T\mathbb{T}^d \times \mathbb{T}$), où $[\omega] \in H^1(\mathbb{T}, \mathbb{R}) = \mathbb{R}^d$ est la classe de cohomologie de ω, et où $\langle ., . \rangle : H_1(\mathbb{T}^d, \mathbb{R}) \times H^1(\mathbb{T}^d, \mathbb{R}) \longrightarrow \mathbb{R}$ est le produit de dualité, qui s'identifie au produit scalaire standard sur $\mathbb{R}^d$.

DÉFINITION 3.1. *On définit la fonction $\alpha : \mathbb{R}^d \longrightarrow \mathbb{R}$ de Mather par*

$$\alpha(c) = \max_{\mu} \left(\langle [\mu], c \rangle - \int_{T\mathbb{T}^d \times \mathbb{T}} L d\mu \right)$$

où le maximum est pris sur l'ensemble des mesures de probabilité invariantes à support compact.

L'existence du maximum dans cette définition est un théorème de Mather, [41]. Les mesures réalisant ce maximum sont appelées *mesures (c-)minimales*, parce qu'elles minimisent l'action $\int L d\mu$ à homologie fixée. Les mesures minimales ont la propriété remarquable d'être concentrées sur des graphes Lipschitz. Plus précisément, on a le théorème suivant de John Mather.

THÉORÈME 8. [41] *Pour chaque cohomologie c, il existe un champ de vecteurs Lipschitzien $v(q,t) : \mathbb{T}^d \times \mathbb{T} \longrightarrow \mathbb{R}^d$ tel que toutes les mesures c-minimales sont supportées sur le graphe de v.*

Il est utile de donner une variante de la définition de Mather.

PROPOSITION 2. *On a*

$$\alpha(c) = \max_{\mu} \langle [\mu], c \rangle - \int_{T\mathbb{T}^d \times \mathbb{T}} L d\mu$$

où le maximum est pris sur l'ensemble des mesures de probabilité fermées à support compact.

L'intérêt de cette variante de la définition est que l'ensemble de mesures sur lequel est maximisée la fonctionnelle est indépendant du hamiltonien. La possibilité de maximiser par rapport à une ensemble universel (indépendant du hamiltonien) a d'abord été remarquée par Mañe, [36]. Je déduis dans [B5] la simple mais utile conséquence suivante :

COROLLAIRE 1. *La correspondance*

$$H \longmapsto \alpha$$

qui, au hamiltonien H, associe la fonction α, est croissante.

On a de plus le remarquable résultat suivant, qui généralise légèrement une observation de Mañe, puis une élaboration de Bangert [2]. Les détails concernant ce résultat dans le présent contexte sont donnés dans l'article [BB].

THÉORÈME 9. *Les mesures réalisant le maximum dans la proposition 2 sont les mesures c-minimales de Mather.*

L'invariance est donc obtenue comme conséquence du caractère minimisant, et non supposée à priori comme dans les travaux de Mather.

3.3. **Orbites périodiques.** Donnons maintenant une définition équivalente que j'ai introduite dans [B5]. Toute orbite périodique $z(t) : \mathbb{R} \longrightarrow T^*\mathbb{T}^d$ a une homologie $[z] \in H^1(\mathbb{T}^d, \mathbb{Z}) = \mathbb{Z}^d$, une période $T \in \mathbb{Z}$, et une action

$$\mathcal{A}(z) := \int_z \lambda - H dt.$$

Le spectre d'action marqué est l'ensemble des triplets

$$([z], T, \mathcal{A}(z)) \in \mathbb{Z}^d \times \mathbb{N} \times \mathbb{R}$$

où z décrit l'ensemble des orbites périodiques de H de période entière T. Le spectre d'action marqué est donc un sous-ensemble A de $\mathbb{Z}^d \times \mathbb{N} \times \mathbb{R}$.

PROPOSITION 3. *On a*

$$\alpha(c) = \sup_{(w,T,a)\in A} \frac{\langle w, c\rangle - a}{T}.$$

Notons que cette fois le supremum n'est pas atteint en général. L'avantage de cette définition est qu'elle met en lumière le caractère symplectique de la fonction α. Soit $\Phi(z,t) = (\phi_t(z), t)$ un changement de variable. Le morphisme $\phi_t^* : H^1(T^*\mathbb{T}^d, \mathbb{R}) \longrightarrow H^1(T^*\mathbb{T}^d, \mathbb{R})$ est indépendant de t, et il engendre naturellement via l'isomorphisme $\pi^* : H^1(\mathbb{T}^d, \mathbb{R}) \longrightarrow H^1(T^*\mathbb{T}^d, \mathbb{R})$ un morphisme

$$\Phi^* : H^1(\mathbb{T}^d, \mathbb{R}) \longrightarrow H^1(\mathbb{T}^d, \mathbb{R})$$

COROLLAIRE 2. *On a*

$$\alpha_{\Phi^* H} \circ \Phi^* = \alpha_H$$

En particulier, si Φ est proche de l'identité, alors $\alpha_{\Phi^* H} = \alpha_H$.

3.4. **Équation de Hamilton-Jacobi.**

PROPOSITION 4. *Pour chaque cohomologie $c \in \mathbb{R}^d$, on a*

$$\alpha(c) = \min_{u\in C^1(\mathbb{T}^d\times\mathbb{T},\mathbb{R})} \ \max_{(q,t)\in\mathbb{T}^d\times\mathbb{T}} \partial_t u(q,t) + H(q, c + \partial_x u(q,t), t).$$

Cette proposition contient deux résultat récents. Le premier est la valeur de l'infimum, montrée dans [12], mais aussi due à Fathi. Le second est le fait que le minimum existe. Il est du à Fathi et Siconolfi, [24], dans le cas autonome, et nous traitons le cas non autonome dans [BB].

Considérons l'équation de Hamilton-Jacobi

$$\partial_t u + H(q, c + \partial_q u, t \bmod 1) = 0,$$

d'inconnue $u : \mathbb{T}^d \times [0,\infty) \longrightarrow \mathbb{R}$. Le problème de Cauchy avec une donnée initiale u_0 continue est bien posé pour les solutions de viscosités. Autrement dit, pour toute condition initiale u_0 continue, il existe une et une seule solution de viscosité $u(q,t) : \mathbb{T}^d \times [0,\infty) \longrightarrow \mathbb{R}$ de l'équation de Hamilton-Jacobi. Elle est donnée par (et on peut prendre cette formule comme définition)

$$u(q,t) = \min u_0(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s), s) + \langle c, \dot\gamma(s)\rangle ds$$

où le minimum est pris sur l'ensemble des courbes $\gamma \in C^1([0,t], \mathbb{T}^d)$ vérifiant $\gamma(t) = q$.

J'ai montré le théorème suivant dans [B7].

PROPOSITION 5. *Si $u(q,t) : \mathbb{T}^d \times [0,\infty) \longrightarrow \mathbb{R}$ est une solution de viscosité de l'équation de Hamilton-Jacobi, alors la fonction $u(q,t) + t\alpha(c)$ est bornée. En particulier, on a*

$$\alpha(c) = -\lim_{t\longrightarrow\infty} u(q,t)/t.$$

Le facile corollaire suivant, qui est bien antérieur, aide sans doute à comprendre le sens de la fonction α. Son extension dans le cas presque intégrable sera discuté à la fin de cette section.

COROLLAIRE 3. *Si il existe un réel a et une solution $u \in C^1(\mathbb{T}^d \times \mathbb{T}, \mathbb{R})$ de l'équation*

$$\partial_t u + H(q, c + \partial_q u, t) = a,$$

alors $a = \alpha(c)$. En particulier, si $H(q,p,t) = h_0(p)$ est totalement intégrable, alors $\alpha(c) = h_0(c)$.

Il est tentant de considérer l'équation de Hamilton-Jacobi corrigée

$$\partial_t u + H(q, c + \partial_q u, t \bmod 1) = \alpha(c), \tag{2}$$

dont les solutions de viscosité sont bornées. L'opérateur $\mathcal{T}_c$ qui, à toute condition initiale $u_0 \in C(\mathbb{T}^d, \mathbb{R})$, associe la valeur $u_1 \in C(\mathbb{T}^d, \mathbb{R})$ de la solution de viscosité au temps 1 est donné explicitement par

$$\mathcal{T}_c u(q) = \min_{\gamma(1)=q} u_0(\gamma(0)) + \int_0^1 L(\gamma(s), \dot\gamma(s), s) + \langle c, \dot\gamma(s)\rangle ds - \alpha(c).$$

Nous reviendrons plus tard sur le lien concret entre cet opérateur et la dynamique. Le comportement asymptotique du semi-groupe $\mathcal{T}_c^n$ a été étudié entre autre dans cette perspective. Albert Fathi a montré dans [20], [21], [19] qu'il existe au moins un point fixe, qu'il a appelé solution KAM faible. Pour comprendre cette terminologie, il faut rappeler qu'un point fixe C^1 donne un tore Lagrangien invariant. Lorsque le hamiltonien est autonome, il a montré de plus dans [22] que l'ensemble des points fixes est un attracteur global de la dynamique de $\mathcal{T}_c$, c'est à dire que toute suite $n \longmapsto \mathcal{T}_c^n(u)$ converge uniformément vers un point fixe. Ceci peut être assimilé à une forme de λ-lemme.

Dans le cas non-autonome, la situation est beaucoup plus complexe. En règle générale, le comportement asymptotique du semi-groupe ne se réduit plus à des points fixes. Par exemple, le semi-group admet souvent des points périodiques de période (minimale) strictement supérieure à 1. J'ai étudié ce comportamant asymptotique dans [B4], [BR] et [B7]. Lorsque $d = 1$, je montre qu'il se réduit à des orbites périodiques. Par contre, en dimension supérieure, il n'existe aucun résultat. Je suppose que l'ensemble ω-limite du semi-groupe peut devenir très complexe lorsque la dimension augmente, mais aucun exemple ne vient pour l'instant confirmer cette hypothèse.

Certaines considérations de [B4] m'avaient alors convaincu que le comportement asympotique du semi-groupe $\mathcal{T}_c^n$ jouait un role essentiel dans la compréhension de la structure des objets mis en place pour étudier variationnellement la dynamique des variables d'action. Ceci était lié à certaines difficultés passées sous silence dans [42]. Suite à la nouvelle approche de [B8], cela ne me paraît plus aussi clair. Je mentionne cependant un résultat de nature dynamique obtenu dans [B7] comme corollaire de l'étude du comportement asymptotique de $\mathcal{T}_c^n$ lorsque $d = 1$. Des résultats similaires ont été obtenus préalablement par Fathi dans le cas autonome, en toute dimension. Il s'agit en un certain sens d'une réciproque au célèbre théorème de Birkhoff selon lequel les courbes invariantes homotopiquement non triviales sont des graphes.

PROPOSITION 6. [B7] *Pour $d = 1$, si il existe une courbe C de $T^*\mathbb{T}$ qui est un graphe au dessus de la base, et une suite non bornée t_n d'entiers tels que les images par le flot hamiltonien au temps t_n de la courbe C sont des graphes, alors il existe un entier T tel que la courbe C est invariante par le flot au temps T. De plus, ou bien $T = 1$, ou bien la courbe C contient une orbite périodique.*

3.5. **Familles de tores invariants.** Les considérations suivantes sont, essentiellement, issues de [B5]. Soit $H : T^*\mathbb{T}^d \times \mathbb{T} \longrightarrow \mathbb{R}$ un hamiltonien. Supposons que la conclusion du Théorème KAM est satisfaite, c'est à dire qu'il existe

- Un changement de variables Φ, homotope à l'identité,
- Une fonction $g : \mathbb{R}^d \longrightarrow \mathbb{R}$, qui sera aussi vue comme une fonction sur $\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T} = T^*\mathbb{T}^d \times \mathbb{T}$,
- Un fermé F de $\mathbb{R}^d$,

tels que le hamiltonien Φ^*H a un contact d'ordre infini avec g en chaque point de $\mathbb{T}^d \times F \times \mathbb{T}$, ce qui implique en particulier que Φ conjugue le flot de g sur $\mathbb{T}^d \times F \times \mathbb{T}$ avec le flot de H sur $\Phi(\mathbb{T}^d \times F \times \mathbb{T})$. Alors, la fonction α_H a un contact d'ordre infini avec g en chaque point de F. La preuve, suivant [B5] est une conséquence triviale des observations préalables sur la fonction α.

En effet, si Φ^*H a un contact d'ordre infini avec g sur $\mathbb{T}^d \times F \times \mathbb{T}$, c'est qu'il existe deux fonctions g^+ et g^- sur $\mathbb{R}^d$, ayant chacune un contact d'ordre infini avec g sur F, et telles que

$$g^- \leqslant \Phi^*H \leqslant g^+.$$

En vertu de la croissance de $H \longmapsto \alpha$, on a alors

$$g^- = \alpha_{g^-} \leqslant \alpha_{\Phi^*H} = \alpha_H \leqslant \alpha_{g^+} = g^+.$$

4. Une relation d'équivalence

Imagine si ceci
un jour ceci
un beau jour
imagine
si un jour
un beau jour ceci
cessait
imagine

S. Beckett

J'introduis dans [B8] une relation d'équivalence sur $\mathbb{R}^d$ (profitant de l'identification entre l'espace des variables d'action et le groupe de cohomologie), que je note $\lhd\!\rhd$. Avant de donner la définition de cette relation, donnons une conséquence primordiale :

Soit $P_i \in \mathbb{R}^d, i \in \mathbb{Z}$, une suite biinfinie. Supposons que les valeurs P_i appartiennent à une même classe d'équivalence pour la relation $\lhd\!\rhd$. Alors il existe une suite croissante t_i de temps, et une trajectoire hamiltonienne $(q(t), p(t))$ telle que

$$p(t_i) = P_i.$$

4.1. **Pseudopgraphes recouvrants.**

Soit $\mathcal{C}$ l'ensemble des fonctions semi-concaves sur $\mathbb{T}^d$. Ce sont les fonctions $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ telles que la fonction $x \longmapsto \tilde{f}(x) - K\|x\|^2$ est concave sur $\mathbb{R}^d$ pour K assez grand, où $\tilde{f}$ est le relèvement de f à $\mathbb{R}^d$. Nous noterons $\mathcal{C}/\mathbb{R}$ le quotient de $\mathcal{C}$ par les fonctions constantes agissant par addition. On définit alors l'ensemble

$$\mathbb{P} = \mathbb{R}^d \times \mathcal{C}/\mathbb{R}.$$

Les éléments de $\mathbb{P}$ sont appelés pseudographes recouvrants. À tout pseudographe recouvrant $(c, u) \in \mathbb{P}$, on associe le sous-ensemble

$$\Gamma = \Gamma_{c,u} = \{(q, c + du(q)), du(q) \text{ existe}\} \subset T^*\mathbb{T}^d.$$

Comme la correspondance $(c, u) \longmapsto \Gamma_{c,u}$ est injective, on identifie les pseudographes aux sous-ensembles Γ associés. L'ensemble des pseudographes recouvrant contient donc l'ensemble des graphes lagrangiens continus, qui s'identifie à $\mathbb{R}^d \times C^1(\mathbb{T}^d, \mathbb{R})/\mathbb{R} \subset \mathbb{P}$. En munissant $\mathcal{C}$ de la norme uniforme, on munit naturellement $\mathbb{P}$ d'une norme et d'une topologie. La restriction de la projection canonique à un pseudographe recouvrant Γ est injective, son image est de mesure totale. Notons $\phi : T^*\mathbb{T}^d \longrightarrow T^*\mathbb{T}^d$ le flot hamiltonien au temps 1. Le terme *recouvrant* est justifiée par la remarque suivante : la restriction de la projection canonique à l'image $\phi(\Gamma)$ d'un pseudographe recouvrant est surjective. En dimension $d = 1$, les pseudographes recouvrants sont les graphes de fonctions $u : \mathbb{T} \longrightarrow \mathbb{R}$ n'ayant que des discontinuités décroissantes. Ils ont été utilisé dans [29] pour une très élégante présentation de certains résultats connus de la dynamique en dimension 1. L'article [B8] peut être vu comme une tentative partielle de généraliser [29] à la dimension supérieure. Toutefois, une telle généralisation impose des changements radicaux.

DÉFINITION 4.1. [B8] *Soient c et c' dans $\mathbb{R}^d$ et $N \in \mathbb{N}$. On dit que $c \lhd_N \rhd c'$ si, pour tout pseudographe recouvrant $\Gamma \in \mathbb{P}$ de cohomologie c (respectivement c'), il existe un pseudographe recouvrant $\Gamma' \in \mathbb{P}$ de cohomologie c' (respectivement c) tel que*

$$\bar{\Gamma}' \subset \bigcup_{i=1}^{N} \phi^i(\Gamma),$$

où les pseudographes Γ et Γ' sont vus comme sous ensembles de $T^*\mathbb{T}^d$ et où $\bar{\Gamma}'$ désigne l'adhérence de Γ' dans $T^*\mathbb{T}^d$. On dit que $c \lhd\rhd c'$ si il existe $N \in \mathbb{N}$ tel que $c \lhd_N \rhd c'$.

Cette définition oublie la valeur de N, qui est le paramètre crucial pour une éventuelle estimation des temps. Dans un second temps, On pourra étudier les relations $c \lhd_N \rhd c'$ et ainsi chercher à obtenir des estimations des vitesses.

4.2. **Opérateur d'évolution.** Soit $\mathcal{T}_c$ l'opérateur qui, à toute condition initiale $u_0 \in C(\mathbb{T}^d, \mathbb{R})$, associe la valeur $u_1 \in C(\mathbb{T}^d, \mathbb{R})$ au temps 1 de la solution de viscosité de l'équation de Hamilton-Jacobi (2). On peut montrer que $u_1 \in \mathcal{C}$. Il existe un unique opérateur $\Psi : \mathbb{P} \longrightarrow \mathbb{P}$ tel que

$$\Psi(\Gamma_{c,u}) = \Gamma_{c,\mathcal{T}_c(u)}.$$

L'opérateur Ψ est lié à la dynamique par l'inclusion fondamentale

$$\overline{\Psi(\Gamma)} \subset \phi(\Gamma).$$

Aux solutions KAM faible de Fathi, correspondent des points fixes de Ψ. On note $\mathbb{V}_c$ l'ensemble des points fixes de Ψ de cohomologie c. C'est un compact non-vide dans $\mathbb{P}$. Si Γ est un point fixe de Ψ, alors on a

$$\bar{\Gamma} \subset \phi(\Gamma).$$

On définit alors un compact invariant de ϕ par

$$\tilde{\mathcal{I}}(\Gamma) = \bigcap_{i \in \mathbb{N}} \phi^{-i}(\bar{\Gamma}).$$

C'est une nouvelle approche, due à Fathi, pour définir les ensembles d'Aubry $\tilde{\mathcal{A}}(c)$ et de Mañé $\tilde{\mathcal{N}}(c)$ qui avaient été préalablement introduits par Mather dans [41]. On a

$$\tilde{\mathcal{A}}(c) = \bigcap_{\Gamma \in \mathbb{V}_c} \tilde{\mathcal{I}}(\Gamma) \text{ et } \tilde{\mathcal{N}}(c) = \bigcup_{\Gamma \in \mathbb{V}_c} \tilde{\mathcal{I}}(\Gamma).$$

Ce sont des ensembles compacts non vides invariants pas ϕ. Mañé a montré que les mesures c-minimisantes de Mather sont précisément les mesures invariantes qui sont concentrées sur l'ensemble d'Aubry. Comme il existe des mesures c-minimisantes, on en déduit que l'ensemble d'Aubry est non-vide, ce qui n'est pas évident de part sa définition. L'ensemble d'Aubry est un graphe Lipschitz. Je rappelle que les orbites de $\tilde{\mathcal{N}}(c) - \tilde{\mathcal{A}}(c)$ sont biasymptotiques à $\tilde{\mathcal{A}}(c)$. Il peut être utile pour fixer les idées de penser, dans le cadre de la construction d'Arnold, que l'ensemble $\tilde{\mathcal{A}}(c)$ est un des tores T_ω, que le point fixe (il n'y en a qu'un dans ce cas) $\Gamma \in \mathbb{V}_c$ est un morceau de la variété instable de ce tore, et que les orbites de $\tilde{\mathcal{N}}(c) - \tilde{\mathcal{A}}(c)$ sont les homoclines à ce tore sur lesquelles se base la construction d'Arnold. Je précise que ceci n'est qu'un guide intuitif (et faux). Il faudrait être bien plus précis, par exemple sur la correspondance entre c et ω, pour le rendre rigoureux, et il faudrait aussi travailler dans un revêtement, sinon $\tilde{\mathcal{N}}(c) - \tilde{\mathcal{A}}(c)$ est vide.

Pour terminer, mentionnons qu'il existe une partition naturelle, introduite par Mather et Mañé, de l'ensemble d'Aubry $\tilde{\mathcal{A}}(c)$ en sous-ensembles compacts invariants appelés *classes statiques* par Mañé. Voir par exemple [13] pour une définition et des discussions relatives à cette partition, mais aussi [B8].

Chaque fois que $\tilde{\mathcal{X}}$ est un sous-ensemble de $T^*\mathbb{T}^d$, nous désignerons par $\mathcal{X}$ la projection de $\tilde{\mathcal{X}}$ sur la base $\mathbb{T}^d$.

4.3. **Résultats.** Nous pouvons maintenant énoncer les résultats principaux de [B8]. Pour les preuves, et pour plus de détails, consulter cet article, qui est essentiellement auto-contenu.

4.3.1. *Évolution assez lente.* Pour tout sous ensemble K de $\mathbb{T}^d$, on définit le sous-espace vectoriel $R(K) \subset \mathbb{R}^d$ constitué des classes de cohomologie des formes fermées sur $\mathbb{T}^d$ dont le support est disjoint de K. On définit alors, pour tout $c \in \mathbb{R}^d$, les sous-espace

$$R(c) = \bigcap_{\Gamma \in \mathbb{V}_c} R(\mathcal{I}(\Gamma)) \subset \mathbb{R}^d.$$

Le choix de la même lettre pour désigner le sous-espace $R(c)$ ci-dessus et pour désigner le sous-espace $R(\omega)$ associé aux résonances d'une fréquence ω est intentionnel. Préciser la relation entre ces deux objets est un travail à faire. En un sens vague, le théorème suivant montre que le mouvement *assez lent* est pris en compte par la relation $\lhd\!\rhd$. Ce résultat est à rapprocher des résultats principaux de [42] et [B4].

THÉORÈME 10. [B8] *Pour chaque $c_0 \in \mathbb{R}^d$, il existe $\epsilon > 0$ tel que toute classe $c \in \mathbb{R}^d$ vérifiant $c - c_0 \in R(c_0)$ et $\|c - c_0\| \leqslant \epsilon$ satisfait*

$$c_0 \lhd\!\rhd c.$$

4.3.2. *Évolution très lente.* Le théorème suivant prend en compte l'évolution *très lente* des variables d'action (éventuellement après passage à un revêtement, voir [B8]).

THÉORÈME 11. [B8] *Si il existe un nombre fini de classes statiques dans $\tilde{\mathcal{A}}(c)$, et si il existe un nombre fini (non nul) d'orbites dans $\tilde{\mathcal{N}}(c) - \tilde{\mathcal{A}}(c)$, alors la classe de cohomologie c est dans l'intérieur de sa classe de $\lhd\!\rhd$-équivalence.*

5. Conclusion

Les pensées sont des ombres de nos sentiments–toujours obscures, plus vides, plus simples que ceux-ci.

F. Nietzsche *in* Le Gai Savoir

Le lecteur a sans doute constaté un certain flou quand à la manière dont les divers aspects présentés, se combinent pour revenir à la question principale, qui était de décrire la dynamique globale des variables d'action dans un système presque intégrable. Ceci se manifeste particulièrement lorsqu'il s'agit de faire le lien entre les objets définis variationnellement dans les dernières sections, et les objets définis dans les premières sections. Du travail reste à faire pour préciser ces liens.

On voudrait par exemple montrer que, pour une perturbation générale, le complémentaire de l'ensemble $F \subset \mathbb{R}^d$ apparaissant dans le théorème KAM forme une seule classe de $\lhd\rhd$-équivalence. Le mérite de mes travaux de [B8], est de ramener cette question globale à la vérification locale des hypothèses abstraites. Cette vérification devrait pouvoir être faite en utilisant les formes normales locales, dans l'esprit de mon article [B2], et du superbe travail antérieur de Bolotin [7], où l'existence locale d'orbites homoclines (à des tores invariants dans l'article de Bolotin, et à des ensemble d'Aubry, qui y sont appelés ensembles de Peierls dans mon article) est montrée en utilisant une combinaison de méthodes de formes normales et de méthodes variationnelles. Il y a toutefois quelques difficultés. Par exemple, la théorie des formes normales est symplectique, alors que les théories de [B8] utilisent fortement la structure additionnelle fibrée de l'espace des phases. Pour pouvoir idéalement appliquer les résultats de formes normales, il faudra donc comprendre dans quelle mesure les objets apparaissant dans les hypothèses abstraites de [B8] sont transformées par les changements de variables. L'article [B5] va dans cette direction, voir aussi le livre [50]. Il faudra ensuite comprendre dans quelle mesure ces hypothèses abstraites sont vérifiées pour des perturbations typiques.

Pour terminer sur une note positive, je voudrais préciser que les systèmes presque intégrables ne sont pas les seules applications envisageables des méthodes de [B8]. Je voudrais aussi m'avancer à suggérer que les difficultés mentionnées ci-dessus ne sont pas très profondes, je m'attend à ce qu'elles soient levées rapidement. L'énoncé donné par Mather dans [44] ouvre d'ailleurs la voie.

Je propose donc la conjecture suivante : Pour une perturbation typique d'un système complètement intégrable convexe, il existe une classe de $\lhd\rhd$-équivalence qui est un ouvert dense de $\mathbb{R}^d$.

Références

Mes articles cités dans le texte (disponibles sur
http ://www-fourier.ujf-grenoble.fr/~pbernard/publi.html)

[B1] P. Bernard, Perturbation d'un hamiltonien partiellement hyperbolique, Comptes Rendus de l'Academie des sciences, serie I, 323 (1996).

[B2] P. Bernard, Homoclinic orbits to invariant sets of quasi-integrable exact maps, Ergodic Theory and Dynamical systems **20** (2000), 1583-1601.

[B3] P. Bernard, Une propriete de transfert en approximation diophantienne, Annales scientifiques de l'universite de Toulouse, (6) 12 (2003), no. 4, p.453-463.

[B4] P. Bernard, Connecting orbits of time dependent Lagrangian systems. *Ann. Inst. Fourier* **52** (2002), 1533–1568.

[B5] P. Bernard, The action spectrum near positive definite invariant tori. *Bull. Soc. Math. France* **131** (2003), no. 4, 603–616.

[BR] P. Bernard, J.M. Rocquejoffre, Convergence to time-periodic solutions in Hamilton-Jacobi equations on the circle, Communications in PDE ,**29** (2004),n. 3-4, 457-469.

[B7] P. Bernard, The asymptotic behaviour of solutions of forced Burgers equation on the circle, Nonlinearity,**18** (2005), 101-124.

[B8] P. Bernard, The dynamics of peudographs in convex hamiltonien systems, préprint.

[BB] P. Bernard, B. Buffoni, Optimal mass transportation and Mather theory, préprint.

Mes autres articles (aussi disponibles sur
http ://www-fourier.ujf-grenoble.fr/~pbernard/publi.html)

[B10] P. Bernard, Homoclinic orbits in families of hypersurfaces with hyperbolic periodic orbits, Journal of Differential Equations **180**, 2 (2002), 427-452.

[B11] P. Bernard, Homoclinic orbit to a center manifold, Calculus of Variations **17**, 2 (2003), 121-157.

[BRS] P. Bernard, C. Grotta Ragazzo, P. A. Salamao, Homoclinic orbits near saddle center fixed points of hamiltonian systems with two degrees of freedom, Astérisque 286 , Geometric Methods in Dynamics (I) - Volume in honor of Jacob Palis, (2003), 151-165.

[BL] P. Bernard, E. Lombardi, Homoclinic orbits in some bifurcations, en phase terminale.

Autres références

[1] V. I. Arnold, Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, **5** (1964), 581–585.

[2] V. Bangert, Minimal measures and minimizing closed normal one-currents. Geom. Funct. Anal. **9**(1999), 413–427.

[3] M. Berti, P. Bolle, A functional analysis approach to Arnold diffusion. Ann. Inst. H. Poincaré Anal. Non Linéaire **19** (2002), no. 4, 395–450.

[4] M. Berti, L. Biasco, P. Bolle, Drift in phase space : a new variational mechanism with optimal diffusion time. *J. Math. Pures Appl.* (9) **82** (2003), no. 6, 613–664.

[5] U. Bessi, An approach to Arnold's diffusion through the calculus of variations. *Nonlinear Anal.* **26** (1996), no. 6, 1115–1135.

[6] U. Bessi, Arnold's example with three rotators, Nonlinearity **10** (1997), 763–781.

[7] S. V. Bolotin, Homoclinic orbits to invariant tori of Hamiltonian systems, AMS Transl. **168** (1995), 21–90.

[8] S. Bolotin, D. Treschev, Unbounded growth of energy in nonautonomous Hamiltonian systems. *Nonlinearity* **12** (1999), no. 2, 365–388.

[9] J.B.Bost, Tores invariants des systèmes dynamiques Hamiltoniens, Asérisque **133–134** (196), 113–157.

[10] C.-Q. Cheng, J. Yan, Existence of diffusion orbits in a priori unstable Hamiltonian systems, à paraitre dans J. Diff. Geometry.

[11] C.-Q. Cheng, J. Yan, Arnold Diffusion in Hamiltonian systems : the a priori unstable case, préprint.

[12] G. Contreras, R. Iturriaga, G. Paternain, P. Paternain, Lagrangian graphs, minimizing measures and Mañé's critical values. Geom. Funct. Anal. **8**, 788–809

[13] G. Contreras, G. P. Paternain, Connecting orbits between static classes for generic Lagrangian systems, *Topology* **41** (2002), no. 4, 645–666.

[14] G. Contreras, J. Delgado, R. Iturriaga, Lagrangian flows : the dynamics of globally minimizing orbits. II. *Bol. Soc. Brasil. Mat. (N.S.)* **28** (1997), no. 2, 155–196.

[15] J. Cresson, Symbolic dynamics and Arnold diffusion. J. Differential Equations **187** (2003), no. 2, 269–292

[16] J. Cresson, Temps d'instabilité des systèmes Hamiltoniens initialement hyperboliques, C. R. Acad. Sci. Paris Sér. I Math. **332** (2001), no. 9, 831–834.

[17] A. Delshams, R. de la Llave, T. Seara, A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem : announcement of results. *Electron. Res. Announc. Amer. Math. Soc.* **9** (2003), 125–134 (electronic).

[18] L.H.Eliasson, Biasymptotic solutions of perturbed integrable Hamiltonian systems, Bull. Soc. Bras. Mat. **25** (1994), 57–76.

[19] A. Fathi, Book in preparation.

[20] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, (French) [A weak KAM theorem and Mather's theory of Lagrangian systems] *C. R. Acad. Sci. Paris Ser. I Math.* **324** (1997), no. 9, 1043–1046.

[21] A. Fathi, Solutions KAM faibles conjuguées et barrières de Peierls. (French) [Weakly conjugate KAM solutions and Peierls's barriers] *C. R. Acad. Sci. Paris Ser. I Math.* **325** (1997), no. 6, 649–652.

[22] A. Fathi, Orbites hétéroclines et ensemble de Peierls. (French) [Heteroclinic orbits and the Peierls set] *C. R. Acad. Sci. Paris Ser. I Math.* **326** (1998), no. 10, 1213–1216.

[23] A. Fathi, Sur la convergence du semi-groupe de Lax-Oleinik. C. R. Acad. Sci. Paris Sér. I Math, **327** (1998), no. 3, 267–270.

[24] A. Fathi, Siconolfi, Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation, Invent. Math.**155** (2004), no. 2, 363–388.

[25] S. M. Graff, On the conservation of hyperbolic invariant tori for Hamiltonian systems, J. of Diff. Eq. **15** (1974), 1–69.

[26] M. W. Hirsch, C. C. Pugh, M. Shub , Invariant manifolds, L. N. in Math. **583**, Springer Verlag, 1977.

[27] I. S. Ilyashenko, A steepness test for analytic functions. Russian Math. Surveys **41** (1986), 229–230.

[28] V. Kaloshin, Geometric proofs of Mather's connecting and accelerating theorems, preprint.

[29] Y. Katznelson, D. Ornstein, Twist maps and Aubry-Mather sets. Lipa's legacy, *Contemp. Math.*, Amer. Math. Soc., Providence, RI, **211** (1997), 343–357.

[30] A. N. Kolmogorov, On the preservation of conditionally periodic motions, Doklady AN **98** (1954), 527–530. Reprinted in Selected works of A.N.Kolmogorov.

[31] J. Laskar, On the Spacing of Planetary systems, Phys. Review Letters **84** (2000), no 15, 3240-3243.

[32] P. Lochak, Canonical perturbation theory via simultaneous approximation, Russian Math. Surveys 47 (1992), 57–133.

[33] P. Lochak, Arnold diffusion ; a compendium of remarks and questions, in *Hamiltonian systems with three or more degrees of freedom*, Kluwer Acad. Publ., Dordrecht, 1999, 168–183.

[34] P. Lochak, J. P. Marco, Diffusion times and stability exponents for nearly integrable analytic systems, preprint.

[35] P. Lochak, J. P. Marco, D. Sauzin On the splitting of invariant manifolds in multidimensional near-integrable Hamiltonian systems. Mem. Amer. Math. Soc. **163** (2003), no. 775,

[36] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems. Nonlinearity **9** (1996), 273–310.

[37] R. Mañé Lagrangian flows : The dynamics of globally minimizing orbits, Bol. Soc. Bras. Mat, **28** (1997) 141-153

[38] J. P. Marco,Transition le long des chaînes de tores invariants pour les systèmes Hamiltoniens analytiques. Ann. Inst. H. Poincaré Phys. Théor. **64** (1996), no. 2, 205–252.

[39] J. P. Marco, D. Sauzin, Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems. *Publ. Math. I. H. E. S.* **96** (2002), 199–275 (2003).

[40] J. N. Mather, Variational construction of orbits of twist diffeomorphisms, Journal A.M.S. **4** (1991), no 2, 207-263.

[41] J. N. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, Math. Z. **207** 169–207 (1991)

[42] J. N. Mather Variational construction of connecting orbits, Ann. Inst. Fourier, **43** (1993), 1349-1368.

[43] J. N. Mather Variational construction of trajectories for time-periodic Lagrangian systems on the two torus, unpublished manuscript.

[44] J. N. Mather Arnold diffusion I : Announcement of results. Preprint.

[45] N.N.Nekhoroshev, An exponential estimate for the time of stability of nearly integrable Hamiltonian systems, Russian Math. Surveys 32 (1977), 1–65.

[46] L. Niederman Exponential stability for small perturbations of steep integrable Hamiltonian systems. Ergodic Theory Dynam. Systems **24** (2004), no. 2, 593–608.

[47] L. Niederman, preprint.

[48] J. Pöschel, Integrability of Hamiltonien systems on Cantor sets, Comm. Pure Appl. Math. **35** (1982), 653–695.

[49] J. Pöschel, Nekhoroshev estimates for quasi-convex Hamiltonian systems, Math. Z. **213** (1993), 187–216.

[50] K. F. Siburg, *The principle of least action in geometry and dynamics.* Lecture Notes in Mathematics, **1844**, Springer-Verlag, Berlin, 2004.

[51] D. Treschev, Evolution of slow variables in near-integrable Hamiltonian systems. *Progress in nonlinear science, vol. 1, Lerman and Shilnikov editors, Russian Academy of Sciences, Institute of Applied Physics, NizhniĭNovgorod,* (2002).

[52] Z. Xia, Arnold diffusion : a variational construction, in *Proceedings of the International Congress of Mathematicians* , Vol. II (Berlin, 1998), Extra Vol. II, 867-877 (electronic), (1998).

[53] Z. Xia, Arnold Diffusion and instabilities in Hamiltonian dynamics, preprint.

[54] J. C.Yoccoz, Travaux de Herman sur les tores invariants, séminaire Bourbaki 754, Astérisque **201–203** (1992), 143–165.