PUBLICATIONS MATHÉMATIQUES DE L'I.H.É.S.

MICHAEL R. HERMAN

Inégalités a priori pour des tores lagrangiens invariants par des difféomorphismes symplectiques

Publications mathématiques de l'I.H.É.S., tome 70 (1989), p. 47-101

<http://www.numdam.org/item?id=PMIHES_1989__70__47_0>

© Publications mathématiques de l'I.H.É.S., 1989, tous droits réservés.

L'accès aux archives de la revue « Publications mathématiques de l'I.H.É.S. » (http://www.ihes.fr/IHES/Publications/Publications.html) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

INÉGALITÉS « A PRIORI » POUR DES TORES LAGRANGIENS INVARIANTS PAR DES DIFFÉOMORPHISMES SYMPLECTIQUES

par MICHAEL R. HERMAN

Dédié à René Thom pour son 65e anniversaire.

Introduction

Nous nous proposons, dans ce travail, de généraliser, en partie, la théorie de Birkhoff sur les courbes invariantes par les difféomorphismes monotones préservant les aires de l'anneau $\mathbf{A} = \mathbf{T}^1 \times \mathbf{R}$ (cf. [B_1], [H_3], I.4) au cas des tores de dimension n invariants par des difféomorphismes symplectiques du fibré cotangent $\mathbf{A}^n \cong T^*(\mathbf{T}^n)$ du tore de dimension n.

Ce travail comporte deux chapitres.

Le premier consiste à généraliser le premier théorème de Birkhoff sur les inégalités lipschitziennes, le second donne une généralisation perturbative du deuxième théorème de Birkhoff sur le fait que les courbes invariantes sont des graphes.

Le théorème de tores invariants de Kolmogorov-Arnold-Moser permet d'obtenir de « nombreux » tores de dimension n invariants par des difféomorphismes C^∞ exacts symplectiques qui sont des perturbations en topologie C^∞ de difféomorphismes complètement intégrables monotones.

La théorie de Birkhoff et ses généralisations doivent être considérées comme un premier pas vers une étude globale (*i.e.* non perturbative) de l'existence des tores invariants (donnés perturbativement par le théorème KAM). Pour $n = 1$, cette théorie a permis aussi, grâce aux travaux de Birkhoff [B_1] et J. Mather [M_5] (voir aussi Le Calvez [L_1]) de donner une assez bonne description de la dynamique dans les zones d'instabilité de Birkhoff. Il semble beaucoup moins évident (cf. [D]) de généraliser ces résultats, qui s'appuient sur des propriétés de séparation par compacts connexes de $\mathbf{A}$.

La théorie de Birkhoff peut être considérée comme donnant des inégalités *a priori* pour les tores invariants obtenus par le théorème KAM, de nature analytique pour le chapitre I et géométrique pour le chapitre II.

Nous avons inclus dans chaque chapitre une introduction qui en décrit plus en détail le contenu.

Les résultats de ce travail ont été annoncés, en partie, dans [H_1]. Nous avons extrait certains des résultats de [H_1] en incluant des démonstrations. Dans le présent article, nous avons inclus des exemples pour aider le lecteur. A deux occasions, nous avons seulement donné les idées de la démonstration pour ne pas trop allonger le texte.

Les paragraphes I.9 à 12 sont très fortement inspirés par une lettre de J. Mather [M_3].

Je remercie John Mather de m'avoir communiqué les démonstrations de ses résultats remarquables sur les mesures d'énergies minimales [M_2] [M_4] [M_6], Jean Lannes de m'avoir aidé sur plusieurs questions de topologie nécessaire au chapitre II, Adrien Douady pour l'idée de la démonstration de II.3.21, Jean-Christophe Yoccoz de m'avoir aidé à préciser et à simplifier la démonstration de la généralisation perturbative du deuxième théorème de Birkhoff présentée au chapitre II, et Mme P. Truc d'avoir tapé avec grand soin le manuscrit. Je voudrais remercier très chaleureusement l'Institute for Advanced Study, de Princeton, de m'avoir invité, et aidé, pour un séjour durant lequel ce travail a été rédigé et grâce auquel de nombreuses améliorations ont été obtenues.

Chapitre I

INÉGALITÉS LIPSCHITZIENNES : GÉNÉRALISATION DU PREMIER THÉORÈME DE BIRKHOFF

1. Introduction

Nous nous proposons de généraliser dans ce chapitre le premier théorème de Birkhoff.

Au paragraphe 2, nous définissons la propriété de monotonicité (ou de torsion) pour les difféomorphismes de $\mathbf{A}^n$. Au paragraphe 3, nous montrons que les tores invariants donnés par le théorème KAM sont lagrangiens, et au paragraphe 4, nous rappelons la théorie de Birkhoff quand $n = 1$. Le paragraphe 5 montre que si l'on veut généraliser la théorie de Birkhoff pour $n \geqslant 2$, il faut, même pour les difféomorphismes complètement intégrables (qu'on suppose toujours symplectiques), se restreindre aux tores invariants lagrangiens, et au paragraphe 6, qu'il faut faire des hypothèses de positivité de la torsion. Au paragraphe 7, nous définissons la notion de difféomorphisme symplectique monotone positif et nous démontrons que les tores invariants lagrangiens, qui sont des graphes de fonctions $\psi \in C^1(\mathbf{T}^n, \mathbf{R}^n)$, satisfont à la même inégalité *a priori* que celle qu'on obtient pour $n = 1$! Au paragraphe 8, nous définissons la notion de difféomorphisme globalement monotone positif. Cela suppose que le difféomorphisme soit exact symplectique et qu'il ait une fonction génératrice globale; notion définie en 8.1.

On définit en 8.13 un tore C^0-lagrangien (qu'on suppose toujours être un graphe). La généralisation du premier théorème de Birkhoff 8.14 s'énonce : si F est un difféomorphisme globalement positif qui laisse invariant un tore T, C^0-lagrangien, alors T est le graphe d'une fonction lipschitzienne $\psi \in \mathrm{Lip}(\mathbf{T}^n, \mathbf{R}^n)$ et $\| D\psi \|_{L^\infty}$ satisfait à la même inégalité que si $n = 1$. Avec le choix canonique des normes euclidiennes que nous faisons, il est important de noter que les inégalités ne dépendent pas de la dimension n. La démonstration est très semblable à celle de 7.2. L'introduction, quand $n = 1$, de la fonction $K(\theta, \Theta)$ que nous définirons en 9.1 est très naturelle (ainsi que nous l'expliquons au début de 9) et avait été faite antérieurement, dans le contexte de 12.1, par J. Mather dans [M_3] (dans [M_3], J. Mather utilise le premier théorème de Birkhoff). La démonstration de 8.14 au paragraphe 9 repose de façon essentielle sur 9.4 et 9.5 et consiste ensuite à prendre des dérivées au sens des distributions. Au paragraphe 9, nous

annonçons une unicité (utilisant des hypothèses très faibles) pour les tores invariants C^0-lagrangiens, unicité qui est démontrée au paragraphe 11. Au paragraphe 12 nous établissons le lien entre les tores C^0-lagrangiens et les orbites d'énergie minimale au sens d'Aubry [A_4] (voir aussi [K]) et les mesures d'énergie (ou d'action) minimale de Mather [M_2] [M_4] et [M_6]. Le paragraphe 12.1 généralise au cas $n \geqslant 1$ le résultat de Mather [M_3] obtenu dans [M_3] pour le cas $n = 1$, avec l'hypothèse plus faible que F soit seulement un difféomorphisme monotone.

Récemment, J. Mather [M_4] [M_6] a obtenu une généralisation des mesures d'énergie minimale [M_2] et a montré qu'elles vérifient une généralisation remarquable des théorèmes de Birkhoff [M_4], tout au moins pour chaque fermé invariant support d'une mesure de probabilité ergodique d'énergie minimale. L'approche que nous présentons avec les tores C^0-lagrangiens invariants implique, en partie, l'approche de J. Mather, ainsi que nous l'indiquons en 12.6. Il semble que les approches, tores invariants C^0-lagrangiens et mesures d'énergie minimales, soient distinctes : tout point d'un tore invariant C^0-lagrangien n'appartient pas forcément au support d'une mesure de probabilité invariante (d'énergie minimale) et inversement.

D'autres tentatives de généralisations des ensembles d'Aubry-Mather ([A_4] [M_1] [M_2], voir aussi l'exposé de Chenciner [C_1]) ont été avancées par D. Bernstein et A. Katok [BK] et V. Bangert [B]. L'exemple 3.3 montre que les orbites périodiques construites dans [BK] ne sont pas en général d'énergies localement minimales au sens d'Aubry. Il nous semble que la bonne généralisation des ensembles d'Aubry-Mather est celle de V. Bangert [B] et J. Mather [B_4]; cf. aussi [K].

2. Notations et définitions

On note $\mathbf{T}^n = \mathbf{R}^n/\mathbf{Z}^n$ le tore de dimension n, R_α la translation de $\mathbf{T}^n : \theta \mapsto \theta + \alpha$ par un vecteur $\alpha \in \mathbf{T}^n$.

On écrira $d\theta = d\theta_1 \otimes \ldots \otimes d\theta_n$ pour la mesure de Lebesgue de $\mathbf{R}^n$ ou la mesure de Haar de $\mathbf{T}^n$, et L^∞ pour l'espace $L^\infty(\mathbf{R}^n, d\theta, \mathbf{R})$.

On désigne par $T^*(\mathbf{T}^n)$ le fibré cotangent de $\mathbf{T}^n$; on pose $\mathbf{A} = T^*(\mathbf{T}^1)$ et l'on a canoniquement $\mathbf{A}^n \cong T^*(\mathbf{T}^n) \cong \mathbf{T}^n \times \mathbf{R}^n$.

Les coordonnées angles actions sur $\mathbf{A}^n$ sont notées (θ, r), avec $\theta = (\theta_1, \ldots, \theta_n)$, $r = (r_1, \ldots, r_n)$. On note $v = \sum_{j=1}^{n} r_j \, d\theta_j$ la 1-forme de Liouville et $w = -dv$ la forme symplectique canonique sur $\mathbf{A}^n$.

Un difféomorphisme F de classe C^1 de $\mathbf{A}^n$ est dit *symplectique* s'il vérifie $F^* w = w$ (ou encore la forme $F^* v - v$ est fermée; si F est C^1 on prend la différentielle extérieure d au sens des distributions, cf. [M_7]). On dit que F est *exact symplectique*, ou encore *globalement canonique* si la forme $F^* v - v$ est exacte.

On note F^k l'itéré k-ième ($k \in \mathbf{Z}$) d'un difféomorphisme F.

Définition. — *Un difféomorphisme* L *de* $\mathbf{A}^n$ *est dit* complètement intégrable *s'il est symplectique et de la forme*

$$L(\theta, r) = (\theta + \ell(r), r).$$

Le fait que L, de classe C^k, soit symplectique équivaut à ce qu'on ait $\ell(r) = \frac{\partial}{\partial r}\hat{\ell}(r)$, avec $\frac{\partial}{\partial r} = \left(\frac{\partial}{\partial r_1}, \ldots, \frac{\partial}{\partial r_n}\right)$, pour une fonction $\hat{\ell} \in C^{k+1}(\mathbf{R}^n, \mathbf{R})$.

Exemple. — Si ℓ est de la forme $\ell(r) = Br$ avec $B \in M_n(\mathbf{R})$, L est symplectique si et seulement si B est symétrique (on notera ${}^tB$ la transposée d'une matrice B).

On note $Sp(2n)$ le groupe symplectique (réel) de $\mathbf{R}^{2n}$, J l'élément $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ de $Sp(2n)$ (où 1 est la matrice unité de $M_n(\mathbf{R})$).

Soit F un difféomorphisme de classe C^1 de $\mathbf{A}^n$; on désigne par $DF(x) = \begin{pmatrix} a(x) & b(x) \\ c(x) & d(x) \end{pmatrix}$ sa matrice jacobienne en x dans les coordonnées (θ, r), et par $(DF(x))^{-1}$ la matrice inverse (alors que $DF^{-1}(x)$ désigne la matrice jacobienne de F^{-1} en x). Dans l'écriture ci-dessus, a, b, c, d sont des applications continues de $\mathbf{A}^n$ dans $M_n(\mathbf{R})$.

On a, pour tout $x \in \mathbf{A}^n$, les équivalences suivantes

(2.0) $\quad DF(x) \in Sp(2n);$

$\Updownarrow$

(2.1) $\quad {}^tDF(x)\, JDF(x) = J;$

$\Updownarrow$

(2.3) $\quad a(x)\, {}^tb(x) = b(x)\, {}^ta(x), \quad c(x)\, {}^td(x) = d(x)\, {}^tc(x),$

$\quad a(x)\, {}^td(x) - b(x)\, {}^tc(x) = 1;$

$\Updownarrow$

(2.4) $\quad {}^ta(x)\, c(x) = {}^tc(x)\, a(x), \quad {}^tb(x)\, d(x) = {}^td(x)\, b(x),$

$\quad {}^ta(x)\, d(x) - {}^tc(x)\, b(x) = 1;$

$\Updownarrow$

(2.5) $\quad [DF(x)]^{-1} = \begin{pmatrix} {}^td(x) & -{}^tb(x) \\ -{}^tc(x) & {}^ta(x) \end{pmatrix}.$

Définition. — *Un difféomorphisme* F *de classe* C^1 *de* $\mathbf{A}^n$ *est monotone si on a* $\det b(x) \neq 0$ *pour tout* $x \in \mathbf{A}^n$.

Un *tore* de classe C^k de $\mathbf{A}^n$ est l'image d'un plongement de classe C^k de $\mathbf{T}^n$ dans $\mathbf{A}^n$ *(on ne considérera dans la suite que des tores de dimension n)*. On désigne par j_T l'inclusion d'un tore T dans $\mathbf{A}^n$. Un tore T est *homotope à* $\{ r = 0 \}$ si on peut choisir le plongement ci-dessus homotope à l'inclusion canonique : $\mathbf{T}^n \to \mathbf{T}^n \times \{ 0 \} \subset \mathbf{T}^n \times \mathbf{R}^n = \mathbf{A}^n$. Un tore T est *lagrangien* s'il est de classe C^1 et $j_T^* w = 0$. Plus généralement une *sous-variété*

$N \subset \mathbf{A}^n$ est dite *lagrangienne* si en tout point $x \in N$ la dimension de N est égale à n et $j_N^* w = 0$.

Si $\psi \in C^0(\mathbf{T}^n, \mathbf{R}^n)$, le tore graphe de ψ est désigné par $\Gamma_\psi = \{(\theta, \psi(\theta)), \theta \in \mathbf{T}^n\}$.

On munira toujours $\mathbf{R}^n$ de la norme euclidienne $||(x_j)|| = (\sum_{j=1}^{n} x_j^2)^{1/2}$; les matrices $a \in M_n(\mathbf{R})$ sont équipées de la norme d'opérateurs

$$|| a || = \sup_{||v||=1} || av ||, \quad v \in \mathbf{R}^n.$$

Dans la suite, les matrices symétriques de $M_n(\mathbf{R})$ seront équipées des relations standard suivantes :

$$a > b \quad (\text{resp. } a \geqslant b)$$

équivaut à

$${}^t vav > {}^t vbv \quad (\text{resp. } {}^t vav \geqslant {}^t vbv) \text{ pour tout } v \in \mathbf{R}^n - \{0\}.$$

Une application $\psi : \mathbf{R}^n \to \mathbf{R}^n$ est lipschitzienne si on a $\sup_{x \neq y} \dfrac{|| \psi(x) - \psi(y) ||}{|| x - y ||} < +\infty$.

3. Tores invariants lagrangiens

3.1. Un des théorèmes des tores invariants de Kolmogorov-Arnold-Moser (théorème KAM) montre que si F est un difféomorphisme C^∞, exact symplectique, qui est une perturbation en topologie C^∞ compacte ouverte d'un difféomorphisme L de classe C^∞, complètement intégrable, monotone, alors F laisse invariants de « nombreux » tores de dimension n, de classe C^∞ homotopes à $\{r = 0\}$ et tels que F restreint à chaque tore est C^∞ conjugué à une translation ergodique de $\mathbf{T}^n$ qui satisfait à une condition diophantienne. Nous renvoyons le lecteur à $[B_2]$ pour des énoncés précis et de nombreuses références. Nous ajoutons les références récentes $[S_1]$ et $[S_2]$. La dernière donne une approche un peu différente.

Il est important, pour la suite, de noter que le théorème KAM n'impose pas d'hypothèse de positivité (cf. 6 à 8) sur le difféomorphisme C^∞ L, et n'impose rien d'autre à L que d'être complètement intégrable et *monotone*.

3.2. Les tores invariants obtenus par le théorème KAM sont lagrangiens, comme le montre la proposition suivante.

Proposition. — Soient F *un difféomorphisme symplectique de classe* C^1 *de* $\mathbf{A}^n$ *et* T *un tore de classe* C^1, *invariant par* F; *on suppose qu'il existe une translation ergodique* R_α *de* $\mathbf{T}^n$ *et un plongement* φ *de classe* C^1 *de* $\mathbf{T}^n$ *d'image* T *tels que* $\varphi^{-1} \circ F \circ \varphi$ *soit* C^1-*conjugué à* R_α. *Alors* T *est lagrangien.*

Démonstration. — Ecrivons $h^{-1} \circ \varphi^{-1} \circ F \circ \varphi \circ h = R_\alpha$, pour un difféomorphisme h de classe C^1 de $\mathbf{T}^n$. La 2-forme $j_T^* w$ est invariante par F; la 2-forme $w_1 = h^* \varphi^* j_T^* w$

est donc invariante par R_α. Comme la translation R_α est minimale (*i.e.* toutes ses orbites sont denses dans $\mathbf{T}^n$), on en déduit que w_1 est constante, *i.e.* s'écrit $\sum_{i<j} c_{ij}\, d\theta_i \wedge d\theta_j$, $c_{ij} \in \mathbf{R}$.

Par ailleurs, on a $w = dv$, donc w_1 est exacte d'après la théorie des formes cobordables [M_7]; finalement on a $w_1 = 0$ et $j_T^* w = 0$. ■

3.3. Dans la proposition ci-dessus il faut nécessairement supposer que la translation R_α soit ergodique (si $n \geqslant 2$).

Exemple. Soit $\psi(\theta)$ la fonction $\cos(2\pi\theta)$ sur $\mathbf{T}^1$, soit $\varepsilon > 0$ et considérons l'hamiltonien sur $\mathbf{A}^2$ $H_\varepsilon(r, \theta) = 1/2(r_1 - \varepsilon\psi(\theta_2))^2 + (1/2)\, r_2^2$. Soit F_s le temps s, $s > 0$, du flot $(F_s)_{s \in \mathbf{R}}$ du champ de vecteurs X_{H_ε} défini par $X_{H_\varepsilon} = \left(\frac{\partial H_\varepsilon}{\partial r}, -\frac{\partial H_\varepsilon}{\partial \theta}\right)$. Alors F_s est un difféomorphisme C^∞, exact symplectique, ayant un tore de points fixes

$$T = \Gamma_{\hat{\psi}}, \quad \hat{\psi}(\theta_1, \theta_2) = (\varepsilon\psi(\theta_2), 0)$$

et T *n'est pas un tore lagrangien.*

Ce phénomène, découvert par M. C. Arnaud [A_3], est intéressant à plus d'un titre : On montre (ce que nous ne ferons pas ici) que pour tout $c > 0$, si $\varepsilon > 0$ est assez petit et si $s = 1$, F_1 est monotone (et même monotone positif au sens du paragraphe 7) sur $H^{-1}([0, c])$.

Les points fixes de F_1, de nombre de rotation 0 dans le revêtement universel $\tilde{\mathbf{A}}^2$ de $\mathbf{A}^2$ (on choisit le relèvement $\tilde{F}_s$ de F_s à $\tilde{\mathbf{A}}^2$ en relevant l'isotopie (F_s), $s \in [0, 1]$, $\tilde{F}_{s=0} = \mathrm{Id}$) sont précisément les points de T.

Aucun point fixe de F_1, de vecteur de rotation 0 dans le revêtement universel de $\mathbf{A}^2$, ne définit un état d'énergie minimale au sens d'Aubry ([A_4] et 12.5). On peut même considérer le flot G_s associé au hamiltonien $H_\varepsilon(\theta, r) + \delta \sin(2\pi\, \theta_2)$.

Si $\varepsilon > 0$ est assez petit et $0 < |\delta| < \varepsilon$, $\delta \to 0$, alors $G_{s=1}$ n'a pas d'ensemble compact K invariant qui soit un graphe sur $p_1(K) \subset \mathbf{T}^2 \times \{0\}$ $(p_1(\theta, r) = \theta)$ et qui vérifie :

- K définit des états d'énergie localement minimale au sens d'Aubry.
- Le vecteur de rotation de $\tilde{G}_{s=1|\tilde{K}}$ existe en tout point et est égal à 0.

Les exemples ci-dessus sont des perturbations d'un difféomorphisme complètement intégrable. Si $\varepsilon \to 0$, les difféomorphismes F_1 et G_1 tendent en topologie C^∞ vers $(\theta, r) \to (\theta + r, r)$.

[Les résultats ci-dessus suivent en utilisant les paragraphes 10 à 12. On remarque que les flots hamiltoniens ci-dessus ont une intégrale première égale à r_1 et on vérifie que les points fixes de $G_{s=1}$ *ont une partie elliptique non triviale.*]

A propos de la non-existence d'orbites périodiques d'action minimale, le lecteur consultera les exemples de Hedlund de flots géodésiques sur $\mathbf{T}^3$; il en trouvera une présentation détaillée dans le travail de V. Bangert [B].

On peut généraliser ces exemples à $\mathbf{A}^n$ en considérant des hamiltoniens sur $\mathbf{A}^n$ de la forme

$$H_\varepsilon(\theta, r) = \hat{\ell}(r - \varepsilon\psi(\theta)), \quad \varepsilon > 0$$

où $\hat{\ell} \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ vérifie $\dfrac{\partial \hat{\ell}}{\partial r}(0) = 0$ et $\psi = (\psi_1, \ldots, \psi_n) \in C^\infty(\mathbf{T}^n, \mathbf{R}^n)$. Le flot du champ de vecteurs X_{H_ε} laisse fixe le tore $T = \Gamma_\psi$ qui est non lagrangien si la 1-forme $\sum_1^n \psi_j \, d\theta_j$ est non fermée.

3.4. *Question.* — *Soient* F *un difféomorphisme symplectique de* $\mathbf{A}^n$, *homotope à l'identité, et* T *un tore de classe* C^1, *homotope à* $\{ r = 0 \}$, *invariant par* F, *tel que la restriction de* F *à* T *soit minimale (i.e. toutes ses orbites sont denses dans* T*). Le tore* T *est-il nécessairement lagrangien ?*

La réponse est positive si $n = 1$ ou 2.

3.5. Soient F un difféomorphisme symplectique de $\mathbf{A}^n$, et x_0 un point fixe hyperbolique de F (*i.e.* on a $F(x_0) = x_0$ et $DF(x_0)$ n'a pas de valeur propre de module 1); alors les variétés stable $W^s_{x_0}$ et instable $W^u_{x_0}$ de F en x_0 sont des variétés lagrangiennes injectivement immergées dans $\mathbf{A}^n$.

3.6. *Remarque.* — Si un difféomorphisme symplectique F de classe C^1 de $\mathbf{A}^n$, homotope à l'identité, laisse invariant un tore lagrangien T de classe C^1, homotope à $\{ r = 0 \}$, alors F est exact symplectique; un exemple d'A. Douady montre que ce n'est pas nécessairement le cas pour $n \geqslant 2$, lorsque T n'est pas lagrangien. (En effet, on a

$$j_T^*(F^* v - v) = f^* \circ j_T^* v - j_T^* v \quad \text{où } f = F_{|T}.$$

Comme T est homotope à $\{ r = 0 \}$, le difféomorphisme f de T est homotope à l'identité. Le tore T étant lagrangien, la forme $j_T^* v$ est fermée; si T est seulement C^1, on utilise [M_7]. Il suit que la classe de $F^* v - v$ dans $H^1(\mathbf{A}^n, \mathbf{R})$ est nulle.)

4. La théorie de Birkhoff ($n = 1$)

Soit F un difféomorphisme de $\mathbf{A}$ monotone de classe C^1 et homotope à l'identité. En 1920, Birkhoff a esquissé la démonstration des deux théorèmes suivants ([B_1], voir [F], [H_3], [L_2] pour des démonstrations complètes).

Théorème **1.** — *Supposons qu'il existe une fonction continue* $\psi \in C^0(\mathbf{T}^1, \mathbf{R})$ *dont le graphe* $\Gamma_\psi \subset \mathbf{A}$ *soit un tore invariant par* F. *Alors la fonction* ψ *est lipschitzienne et on a*

$$\| D\psi \|_{L^\infty} \leqslant \sup_{x \in \Gamma_\psi} (| b^{-1}(x) \, a(x) |, | d(x) \, b^{-1}(x) |),$$

où $DF(x) = \begin{pmatrix} a(x) & b(x) \\ c(x) & d(x) \end{pmatrix}$ *et* $D\psi \in L^\infty$ *désigne la dérivée de* ψ *(qui existe presque partout puisque la fonction* ψ *est lipschitzienne).*

Dans le second théorème, on suppose de plus que F est symplectique (hypothèse qui n'est pas nécessaire pour le théorème 1), et vérifie

$$\sup_{x \in \mathbf{A}} (|\, b^{-1}(x)\, a(x)\,|, |\, d(x)\, b^{-1}(x)\,|) < +\infty.$$

Théorème **2.** — *Tout tore* C *de classe* C^0 *homotope à* $\{ r = 0 \}$ *et invariant par* F *est le graphe d'une fonction continue* $\psi \in C^0(\mathbf{T}^1, \mathbf{R})$.

Birkhoff montre en fait beaucoup plus [B_1]; nous n'en parlerons pas ici car ses résultats, qui utilisent des propriétés de séparation des compacts connexes dans $\mathbf{A}$, ne peuvent se généraliser pour $n \geqslant 2$ (cf. [D]).

On cherche à généraliser les théorèmes ci-dessus au cas où $n \geqslant 2$. On a les deux difficultés suivantes 5 et 6. La première difficulté montre, si on veut généraliser les théorèmes de Birkhoff, qu'il faut seulement considérer les tores lagrangiens, la seconde montre qu'il faut faire des hypothèses de positivité, hypothèses que nous énoncerons en 7 et 8.

5. Première difficulté

On prend $n = 2$ et on considère le difféomorphisme $L(\theta, r) = (\theta, +r, r)$ de $\mathbf{A}^2$, qui est monotone et complètement intégrable.

Soient un nombre rationnel $p/q \in \mathbf{Q}$, $\hat{\psi}$ une fonction $1/q$ périodique non constante dans $C^\infty(\mathbf{T}^1, \mathbf{R})$, et ψ la fonction : $(\theta_1, \theta_2) \mapsto (p/q, \hat{\psi}(\theta_1))$ de $\mathbf{T}^2$ dans $\mathbf{R}^2$. Le graphe $T = \Gamma_\psi$ de ψ est alors un tore invariant par L, mais n'est pas lagrangien. Considérons la partie, invariante par L, $E = \{(\theta, r) \in \mathbf{A}^2, r_1 = p/q\}$ de $\mathbf{A}^2$, la projection

$$p : (\theta_1, \theta_2, p/q, r_2) \mapsto (\theta_1, r_2)$$

de E sur $\mathbf{T}^1 \times \mathbf{R} = E_1$, et le diagramme commutatif

$$\begin{array}{ccc} E & \xrightarrow{L/E} & E \\ \downarrow p & & \downarrow p \\ E_1 & \longrightarrow & E_1 \\ (\theta_1, r_2) & \longmapsto & (\theta_1 + p/q, r_2) \end{array}$$

Si A est une partie de E_1 invariante par $R_{p/q} \times \mathrm{Id}$, alors $p^{-1}(A)$ est invariante par L. On obtient ainsi des tores de classe C^∞, homotopes à $r = 0$, invariants par L, qui ne sont pas des graphes. Le même phénomène a lieu si on remplace L par n'importe quel difféomorphisme monotone complètement intégrable de $\mathbf{A}^n$, dès que $n \geqslant 2$.

En effet, puisque L est monotone, $\ell : r \in \mathbf{R}^n \to \ell(r) \in \mathbf{R}^n$ est une immersion. On peut donc trouver $p/q \in \mathbf{Q}$ tel que $\{r_1 = p/q\} \cap \ell(\mathbf{R}^n) \neq \emptyset$. L'ensemble $V_{p/q} = \ell^{-1}(\{r_1 = p/q\})$ est une sous-variété de $\mathbf{R}^n$ de codimension 1, non vide et non compacte. Il suffit de

prendre $\theta_1 \in \mathbf{T}^1 \to \psi(\theta_1) \in V_{p/q}$, $1/q$ périodique, non constante et de considérer $T = \Gamma_\psi$ pour obtenir des tores non lagrangiens invariants par L.

Pour obtenir des tores qui ne soient pas des graphes, on considère $E = \mathbf{T}^n \times V_{p/q}$, $E_1 = \mathbf{T}^1 \times V_{p/q}$ et $p(\theta, r) = (\theta_1, r)$, l'autre partie de l'argument ci-dessus reste valable.

Conclusion. — La théorie de Birkhoff ne peut se généraliser pour $n \geqslant 2$ que si l'on impose aux tores invariants d'être lagrangiens.

Question. — Un tore de dimension n, homotope à $\{ r = 0 \}$, invariant par un difféomorphisme symplectique de classe C^∞, générique (pour la C^∞-topologie fine ou forte de Whitney), homotope à l'identité, est-il toujours lagrangien?

L'exemple suivant montre que l'hypothèse d'être homotope à l'identité est nécessaire pour que la réponse à la question ci-dessus ne soit pas négative.

Soit $A \in GL(n, Z)$ une matrice hyperbolique (par exemple $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ pour $n = 2$); le difféomorphisme $F : (\theta, r) \mapsto (A\theta, {}^tA^{-1} r)$ de $\mathbf{A}^n$ est symplectique; le tore $\mathbf{T}_0^n = \mathbf{T}^n \times \{ 0 \}$ est un ensemble *invariant hyperbolique* pour F. Par conséquent, une perturbation de classe C^1 de F laisse invariant un tore de dimension n voisin de $\mathbf{T}_0^n$ (cf. [S_5]) mais celui-ci n'est en général ni de classe C^1, ni C^0-lagrangien.

6. Deuxième difficulté : difféomorphismes complètement intégrables indéfinis

6.1. Soit $L(\theta, r) = (\theta + \ell(r), r)$ un difféomorphisme monotone complètement intégrable. Pour tout $r \in \mathbf{R}^n$, la forme quadratique $Q_L(r) : v \in \mathbf{R}^n \mapsto {}^tv\, D\ell(r)\, v$ est donc non dégénérée.

Lorsque $n = 1$, la forme quadratique $Q_L(r)$ est soit définie positive pour tout $r \in \mathbf{R}^n$, soit définie négative pour tout $r \in \mathbf{R}^n$; on dit alors que L est *monotone positif* ou *négatif*.

En écrivant $\ell(r) = \dfrac{\partial}{\partial r} \hat{\ell}(r)$, dire que L est monotone positif signifie que $\hat{\ell}$ est strictement convexe.

Lorsque $n \geqslant 2$, on dit que L est *monotone indéfini* si la forme quadratique $Q_L(r)$ est indéfinie pour tout $r \in \mathbf{R}^n$. La théorie de Birkhoff ne se généralise pas pour les tores invariants, même lagrangiens, par un difféomorphisme monotone indéfini. Nous allons seulement illustrer ce fait par un exemple et nous montrerons, ailleurs, que c'est vrai pour tout difféomorphisme L, monotone indéfini C^∞, même pour les perturbations C^∞ exactes symplectiques de L.

6.2. *Un exemple sur* $\mathbf{A}^2$. Posons $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, et définissons un difféomorphisme complètement intégrable monotone indéfini de $\mathbf{A}^2$ par la formule $L_B(\theta, r) = (\theta + Br, r)$.

Pour toute fonction $\varphi \in C^\infty(\mathbf{T}^1, \mathbf{R})$, considérons le difféomorphisme

$$G_\varphi : (\theta_1, \theta_2, r_1, r_2) \to (\theta_1, \theta_2, r_1 + \varphi'(\theta_1), r_2) \quad \text{de } \mathbf{A}^2,$$

et le composé $F = G_\varphi \circ L_B$, qui est exact symplectique.

Cherchons les tores invariants par F qui sont des graphes $\Gamma_{\hat{\psi}_a}$ d'une fonction $\hat{\psi}_a$ de la forme

$$\hat{\psi}_a(\theta_1, \theta_2) = (\psi_a(\theta_1), a), \tag{6.3}$$

où $a \in \mathbf{R}$ et $\psi_a \in C^\infty(\mathbf{T}^1, \mathbf{R})$.

L'invariance de Γ_{ψ_a} s'écrit :

$$\psi_a \circ R_a - \psi_a = \varphi' \circ R_a. \tag{6.4}$$

Lorsque a satisfait à une condition diophantienne, l'équation (6.4) admet une unique solution $\psi_a \in C^\infty(\mathbf{T}^1, \mathbf{R})$ de moyenne nulle. On note par CD l'ensemble des nombres réels qui satisfont à une condition diophantienne :

Il existe $\gamma > 0$ et $\beta \geqslant 0$ tel que pour tout $p/q \in \mathbf{Q}$ on ait $|\alpha - (p/q)| \geqslant \gamma q^{-2-\beta}$.

Si φ n'est pas constante, les tores $\Gamma_{\hat{\psi}_a}$ ne vérifient pas d'inégalités a priori *lorsqu'on laisse varier a dans* CD.

En effet, prenons $\varphi'(\theta_1) = \varepsilon \cos(2\pi\theta_1)$, avec $\varepsilon > 0$. Pour $a \notin \mathbf{Z}$, on a

$$\psi_a(\theta_1) = \frac{\varepsilon}{2} (\sin \pi a)^{-1} \sin(2\pi(\theta_1 + (a/2))).$$

Donc,

$$\psi_a(-a/2) = 0, \lim_{a \to 0} \| \psi_a \|_{C^0} = +\infty, \tag{6.5}$$

$$\lim_{a \to 0} | \psi_a'(-a/2) | = +\infty.$$

On peut choisir des difféomorphismes symplectiques de classe C^∞ H_a de $\mathbf{A}^n$, tendant vers l'identité dans la C^∞-topologie lorsque a tend vers 0, de manière que le tore $H_a(\Gamma_{\hat{\psi}_a})$, invariant par $H_a \circ F \circ H_a^{-1}$, ne soit pas un graphe.

6.6. *Comme φ est de classe C^∞, presque tout $x \in \mathbf{A}^2$ est sur un tore invariant de* F *lagrangien de classe* C^∞.

En effet pour tout $r_2 = a \in$ CD, on peut résoudre 6.4, la solution générale étant $\psi_a + c$, $c \in \mathbf{R}$ et donc chaque point de $\{ r_2 = a \}$ est sur un tore invariant de classe C^∞ de la forme 6.3.

6.7. *Proposition. — On suppose qu'une infinité de coefficients de Fourier de φ sont non nuls. Alors les points dont l'orbite par* F *est non bornée dans $\mathbf{A}^2$ forment une partie G_δ-dense de $\mathbf{A}^2$.*

Démonstration. — Il suit de [H_2, Chap. XIII] qu'il existe $G_1 \subset \mathbf{R}$, une partie G_δ-dense telle que si $a \in G_1$, l'équation 6.4 n'a pas de solution continue ψ_a (ni même mesurable si on suppose que φ est $\mathbf{R}$-analytique). L'ensemble

$$G_2 = \{ x \in \mathbf{A}^2, \sup_{j \in \mathbf{Z}} \| p_2(F^j(x)) \| = +\infty \}$$

est une partie G_δ de $\mathbf{A}^2$ où $p_2(\theta, r) = r$.

Si $r_2 = a \in G_1$, le résultat de Gottschalk et Hedlund [G] (voir aussi [H_2, IV]) implique que toute orbite $F_{|\{r_2 = a\}}$ est non bornée. On peut même montrer en utilisant [G] que si $a \in G_1$, alors $F_{|\{r_2 = a\}}$ a une orbite dense dans $\{ r_2 = a \}$. ■

6.8. Les exemples F considérés ci-dessus sont « *presque* » complètement intégrables, l'étude des tores invariants se ramenant à 6.7.

On peut donner de nombreux autres exemples : par exemple, un produit $F \times L_1$ avec F comme ci-dessus, et L_1 complètement intégrable. De façon similaire, des Hamiltoniens « presque » complètement intégrables sur $\mathbf{A}^{2n}$ sont donnés par des formules de la forme

$$H(\theta, r) = \sum_{j=1}^{n} r_j r_{j+n} + \varepsilon\varphi(\theta_1, \ldots, \theta_n),$$

avec $\varphi \in C^\infty(\mathbf{T}^n, \mathbf{R})$.

Il est important de rappeler au lecteur que le théorème des tores invariants de Kolmogorov-Arnold-Moser ne fait pas d'hypothèse de positivité sur la torsion. Dans le reste de ce travail, nous allons montrer que si l'on fait des hypothèses de positivité de la torsion, alors on peut généraliser, en partie, la théorie de Birkhoff.

7. Inégalité « a priori » pour les difféomorphismes monotones positifs

7.1. Soit F un difféomorphisme symplectique monotone de $\mathbf{A}^n$.

Ecrivant $DF(x) = \begin{pmatrix} a(x) & b(x) \\ c(x) & d(x) \end{pmatrix}$, la matrice $b(x)$ est alors inversible, et les matrices $b^{-1}(x)\, a(x)$, $d(x)\, b^{-1}(x)$ sont symétriques (cf. (2.3) et (2.4)).

On dit que F est *monotone positif* si l'une des deux conditions suivantes est vérifiée :

($*_1$) pour tout $x \in \mathbf{A}^n$, $b^{-1}(x)\, a(x) > 0$;

($*_2$) pour tout $x \in \mathbf{A}^n$, $d(x)\, b^{-1}(x) > 0$.

Des exemples simples montrent que ($*_1$) et ($*_2$) ne sont pas équivalentes. On dit que F est *monotone négatif* si F^{-1} est monotone positif.

7.2. *Proposition.* — *Si* F *est monotone positif, une fonction* $\psi \in C^1(\mathbf{T}^n, \mathbf{R}^n)$ *dont le graphe* $T = \Gamma_\psi$ *est un tore lagrangien invariant par* F *doit vérifier* :

$$\| D\psi \|_{L^\infty} \leqslant \sup_{x \in T} (\| b^{-1} a) (x) \|, \| db^{-1}(x) \|).$$

Démonstration. — Le difféomorphisme $H(\theta, r) = (\theta, r + \psi(\theta))$ est symplectique de classe C^1, car T est lagrangien. Notons f le difféomorphisme de classe C^1 de $\mathbf{T}^n$ vérifiant $H^{-1} \circ F \circ H(\theta, 0) = (f(\theta, 0)$, et $G(\theta)$ la matrice jacobienne de $H^{-1} \circ F \circ H$ en $(\theta, 0)$. En posant $\bar{a}(\theta) = a(\theta, \psi(\theta))$, $\bar{b}(\theta) = b(\theta, \psi(\theta))$, ..., on a

$$G = \begin{pmatrix} 1 & 0 \\ -D\psi \circ f & 1 \end{pmatrix} \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ D\psi & 1 \end{pmatrix}$$
$$= \begin{pmatrix} Df & \bar{b} \\ 0 & {}^t(Df)^{-1} \end{pmatrix};$$

d'après 2.1-2.5, $G(\theta)$ appartenant à $Sp(2n)$.

On en déduit :

$$\bar{b}^{-1}\bar{a} + D\psi = \bar{b}^{-1} Df,$$
$$\bar{d}\bar{b}^{-1} - D\psi \circ f = {}^t(Df)^{-1}\bar{b}^{-1}.$$

Posons $e = \bar{b}^{-1} Df$, $e_1 = {}^t(Df)^{-1}\bar{b}^{-1}$; les matrices $e(\theta)$, $e_1(\theta)$ sont symétriques et on a

(7.3) $$e_1 = {}^t\bar{b}^{-1} e^{-1} \bar{b}^{-1}.$$

Supposons que F vérifie la condition $(*_1)$.

Lemme. — *Pour tout* $\theta \in \mathbf{T}^n$, *on a* $e(\theta) > 0$.

Démonstration. — Soit $S \in C^2(\mathbf{T}^n, \mathbf{R})$ une fonction telle que $\psi - dS$ soit constante (une telle fonction existe car T est lagrangien). En un point θ_0 où S atteint son minimum sur $\mathbf{T}^n$, on a $D\psi(\theta_0) \geqslant 0$ et $\bar{b}^{-1}\bar{a}(\theta_0) > 0$ d'après $(*_1)$, d'où $e(\theta_0) > 0$. Comme $e(\theta)$ est toujours non dégénérée, on a $e_1(\theta) > 0$ pour tout $\theta \in \mathbf{T}^n$. ■

D'après le lemme et 7.3, on montre aussi $e(\theta)$ pour tout $\theta \in \mathbf{T}^n$.

Si F vérifie la condition $(*_2)$, on montre d'abord que e_1 est positive et on en déduit que e l'est aussi.

Les matrices $\bar{b}^{-1}\bar{a} + D\psi$ et $\bar{d}\bar{b}^{-1} - D\psi \circ f$ sont donc partout positives et on obtient :

$$-\| \bar{b}^{-1}\bar{a} \|_{L^\infty} 1 \leqslant D\psi \leqslant \| \bar{d}\bar{b}^{-1} \|_{L^\infty} 1,$$

où 1 désigne la matrice unité de $M_n(\mathbf{R})$, d'où la proposition. ■

8. Fonctions génératrices et généralisation du premier théorème de Birkhoff

8.1. Soit F un difféomorphisme exact symplectique et monotone de $\mathbf{A}^n$. On note $\tilde{\mathbf{A}}^n = \mathbf{R}^n \times \mathbf{R}^n$ le revêtement universel de $\mathbf{A}^n$, et on choisit un relèvement $\tilde{F}$ de F à $\tilde{\mathbf{A}}^n$; c'est un difféomorphisme symplectique de $\tilde{\mathbf{A}}^n$.

En notant $\tilde{\theta}$ la coordonnée dans $\tilde{\mathbf{T}}^n = \mathbf{R}^n$, on peut écrire $\tilde{F}$ sous la forme

$$\tilde{F}(\tilde{\theta}, r) = (A\tilde{\theta} + \varphi(\theta, r), \tilde{F}_2(\theta, r)),$$

avec $A \in GL(n, \mathbf{Z})$ et $\varphi \in C^1(\mathbf{A}^n, \mathbf{R}^n)$; on suppose que F est homotope à $(\theta, r) \to (A\theta, r)$.

L'hypothèse de monotonicité sur F signifie que, pour tout $\theta \in \mathbf{T}^n$, l'application $r \mapsto \varphi(\theta, r)$ est une immersion de $\mathbf{R}^n$ dans $\mathbf{R}^n$. On fera l'hypothèse suivante :

8.2. Pour tout $\theta \in \mathbf{T}^n$, l'application $r \mapsto \varphi(\theta, r)$ est un difféomorphisme de $\mathbf{R}^n$.

Dire que F est monotone équivaut à dire que pour tout θ, $r \mapsto \varphi(\theta, r)$ est une immersion, et si $n \geqslant 2$ cela n'implique pas, même si F est complètement intégrable, que $r \mapsto \varphi(\theta, r)$ est un difféomorphisme sur son image comme le montre

$$(r_1, r_2) \in \mathbf{R}^2 = \mathbf{C} \to e^{r_1 - ir_2} \in \mathbf{C}, \quad i = \sqrt{-1}.$$

La propriété 8.2, si F est monotone, est impliquée par la condition

$$\sup_{x \in \mathbf{A}^n} \| b^{-1}(x) \| < +\infty. \tag{8.3}$$

(Cela résulte d'un théorème classique dont le lecteur trouvera une démonstration dans [S_3, 1.22].)

Comme F est exact symplectique, on a $F^* v - v = dS$, pour une fonction $S \in C^1(\mathbf{A}^n, \mathbf{R})$. On écrit $\tilde{F}(\tilde{\theta}, r) = (\tilde{\Theta}, R)$; l'hypothèse 8.2 permet de prendre $\tilde{\theta}$ et $u = \tilde{\Theta} - A\tilde{\theta}$ comme coordonnées sur $\tilde{\mathbf{A}}^n$ et dans ces coordonnées, S est de classe C^2. Posons $H(\tilde{\theta}, \tilde{\Theta}) = S(\tilde{\theta}, \tilde{\Theta} - A\tilde{\theta})$.

La fonction H est de classe C^2 et on a les relations

$$r = -\frac{\partial}{\partial \tilde{\theta}} H(\tilde{\theta}, \tilde{\Theta}) \equiv -H_1(\tilde{\theta}, \tilde{\Theta}),$$

$$R = \frac{\partial}{\partial \tilde{\Theta}} H(\tilde{\theta}, \tilde{\Theta}) \equiv H_2(\tilde{\theta}, \tilde{\Theta}),$$

$$dH = \langle R, d\tilde{\Theta} \rangle - \langle r, d\tilde{\theta} \rangle.$$

La fonction H s'appelle une *fonction génératrice (globale)* de $\tilde{F}$. Elle existe et détermine $\tilde{F}$ si et seulement si F *est exact symplectique et vérifie la condition* 8.2.

8.4. La condition 8.2 équivaut à : Pour tout $\tilde{\theta} \in \mathbf{R}^n$, $\tilde{\Theta} \to H_1(\tilde{\theta}, \tilde{\Theta})$ est un difféomorphisme de $\mathbf{R}^n$.

8.5. La condition 8.2 implique aussi : Pour tout $\tilde{\Theta} \in \mathbf{R}^n$, $\tilde{\theta} \to H_2(\tilde{\theta}, \tilde{\Theta})$ est un difféomorphisme de $\mathbf{R}^n$.

Démonstration. — L'application de $\mathbf{R}^{2n}$ dans $\mathbf{R}^{2n}$: $G(\tilde{\theta}, \tilde{\Theta}) = (\tilde{\theta}, -H_1(\tilde{\theta}, \tilde{\Theta}))$ est par 8.2 un difféomorphisme $\tilde{F} \circ G(\tilde{\theta}, \tilde{\Theta}) = (\tilde{\Theta}, H_2(\tilde{\theta}, \tilde{\Theta}))$ puisque

$$\tilde{\Theta} = A\tilde{\theta} + \varphi(\tilde{\theta}, -H_1(\tilde{\theta}, \tilde{\Theta})) \tag{8.6}$$

et

$$R = H_2(\tilde{\theta}, \tilde{\Theta}) = F_2(\tilde{\theta}, -H_1(\tilde{\theta}, \tilde{\Theta})), \tag{8.7}$$

où $\tilde{F}(\tilde{\theta}, r) = (\tilde{\Theta}, F_2(\tilde{\theta}, r))$. ■

En dérivant 8.6 par rapport à $\widetilde{\Theta}$ on obtient

(8.8) $$1 = -\, b(\widetilde{\theta}, -\, H_1(\widetilde{\theta}, \widetilde{\Theta})) . H_{12}(\widetilde{\theta}, \widetilde{\Theta})$$

et par rapport à $\widetilde{\theta}$

(8.9) $$(b^{-1}.a)\,(\widetilde{\theta}, -\, H_1(\widetilde{\theta}, \widetilde{\Theta})) = H_{11}(\widetilde{\theta}, \widetilde{\Theta}).$$

En dérivant 8.7 par rapport à $\widetilde{\Theta}$, il suit

(8.10) $$H_{22}(\widetilde{\theta}, \widetilde{\Theta}) = d.b^{-1}(\widetilde{\theta}, -\, H_1(\widetilde{\theta}, \widetilde{\Theta})).$$

On adopte les conventions $\dfrac{\partial}{\partial\widetilde{\Theta}}\dfrac{\partial H}{\partial\widetilde{\theta}} = H_{12} = {}^tH_{21}$, $\dfrac{\partial^2}{\partial\widetilde{\theta}^2} H = H_{11}, \ldots$

8.11. *Définition.* — *On dit que* F *est* monotone globalement positif *si* F *a une fonction génératrice (globale, et par suite, d'après la définition, vérifie* 8.2*)* H *telle que*

$$\operatorname{Lim} \frac{H(\widetilde{\theta}, \widetilde{\Theta})}{1 + \|\widetilde{\theta} - \widetilde{\Theta}\|} = +\infty$$

pour tout $\widetilde{\theta} \in \mathbf{R}^n$ *(resp. pour tout* $\widetilde{\Theta}$*) lorsque* $\|\widetilde{\theta} - \widetilde{\Theta}\| \to +\infty$.

8.12. Supposons que F soit exact symplectique, vérifie 8.2 (ou 8.3) et qu'il existe une matrice (constante) symétrique $\Delta > 0$ vérifiant les deux conditions suivantes :

$(+_1)$ $$b^{-1}\,a \geqslant \Delta \quad \left(\Leftrightarrow \frac{\partial^2 H}{\partial\widetilde{\theta}^2}(\widetilde{\theta}, \widetilde{\Theta}) \geqslant \Delta\right);$$

$(+_2)$ $$db^{-1} \geqslant \Delta \quad \left(\Leftrightarrow \frac{\partial^2 H}{\partial\widetilde{\Theta}^2}(\widetilde{\theta}, \widetilde{\Theta}) \geqslant \Delta\right).$$

Alors F est monotone globalement positif et monotone positif.

Si F est complètement intégrable, $F(\theta, r) = (\theta + \ell(r), r)$, $\ell(r) = \dfrac{\partial\hat{\ell}(r)}{\partial r}$, alors F monotone positif équivaut à ce que la fonction $\hat{\ell}$ soit strictement convexe sur $\mathbf{R}^n$. Dans ce cas, F monotone globalement positif implique que F est monotone positif et que $r \mapsto \ell(r)$ est un difféomorphisme de $\mathbf{R}^n$; comme le montre $r \in \mathbf{R} \mapsto \hat{\ell}(r) = e^r \in \mathbf{R}$, monotone positif n'implique pas 8.2.

Si on considère sur $\mathbf{A}$ des difféomorphismes F de la forme $F = G_{\varphi_1} \circ L \circ G_{\varphi_2}$, où $\varphi_i \in C^\infty(\mathbf{T}^1, \mathbf{R})$, $i = 1, 2$, $G_{\varphi_i}(\theta, r) = (\theta, r + \varphi_i'(\theta))$ et $L(\theta, r) = (\theta + r, r)$, on obtient des exemples montrant que globalement monotone positif n'implique pas monotone positif.

8.13. *Définition.* — *Un tore* T *de classe* C^0 *de* $\mathbf{A}^n$, *qui est le graphe d'une fonction* $\psi = (\psi_1, \ldots, \psi_n) \in C^0(\mathbf{T}^n, \mathbf{R}^n)$, *est dit* C^0*-lagrangien si la 1-forme* $\sum_{j=1}^{n} \psi_j\, d\theta_j$ *est fermée au sens des distributions.*

De façon équivalente, on doit avoir $\psi = c + \frac{\partial \eta}{\partial \theta}$ avec $c = \int_{\mathbf{T}^n} \psi(\theta)\, d\theta$ et $\eta \in C^1(\mathbf{T}^n, \mathbf{R})$.

Si $(T_j)_{j \in \mathbf{N}}$ est une suite de tores graphes de ψ_j, C^0-lagrangiens, si ψ_j tend en topologie C^0 vers $\psi \in C^0(\mathbf{T}^n, \mathbf{R}^n)$, alors le graphe de ψ, $T = \Gamma_\psi$, est un tore C^0-lagrangien.

Le théorème suivant généralise le théorème 1 (paragraphe 4.1) de Birkhoff.

8.14. *Théorème. — Soit* F *un difféomorphisme symplectique de classe* C^1 *monotone globalement positif de* $\mathbf{A}^n$. *Toute fonction continue* $\psi \in C^0(\mathbf{T}^n, \mathbf{R}^n)$ *dont le graphe* $T = \Gamma_\psi$ *est* C^0*-lagrangien et invariant par* F, *est lipschitzienne, et on a*

$$\|D\psi\|_{L_\infty} \leqslant \sup_{x \in T} (\| b^{-1}(x)\, a(x) \|, \| d(x)\, b^{-1}(x) \|); \tag{8.15}$$

cette inégalité est conséquence des inégalités suivantes :

$$\bar{b}^{-1}\, \bar{a} + D\psi \geqslant 0, \quad \textit{presque partout}; \tag{8.16}$$

$$\bar{d}\, \bar{b}^{-1} - (D\psi) \circ f \geqslant 0, \quad \textit{presque partout.}$$

Les notations $\bar{a}$, $\bar{b}$, $\bar{d}$ sont celles de 7.2 et « presque partout » est relatif à la mesure de Haar $d\theta$ sur $\mathbf{T}^n$ et $D\psi$ désigne l'application linéaire tangente, au sens de Fréchet, de ψ, qui existe presque partout puisque ψ est lipschitzienne.

9. Démonstration du théorème 8.14

Elle s'inspire de la démonstration de 7.2; néanmoins, il faut procéder autrement, *même si on suppose que* ψ *est lipschitzienne*, car il n'est pas évident, dans ce dernier cas (avec les notations de 7.2), que $e > 0$ presque partout (on peut montrer que $e > 0$ sur un ensemble de mesure de Haar positive). Toujours avec les notations de 7.2, l'homéomorphisme $G^{-1} \circ F \circ G$ a pour « fonction génératrice formelle », si on suppose que F a une fonction génératrice H globale,

$$(\tilde{\theta}, \tilde{\Theta}) \mapsto K(\tilde{\theta}, \tilde{\Theta}),$$

où la fonction K est définie ci-dessous. La démonstration va consister à prendre les dérivées secondes de la fonction K au sens des distributions, et de montrer qu'on a suffisamment de « positivité » pour obtenir le théorème.

9.1. On pose

$$\hat{\psi}(\tilde{\theta}) = \langle c, \tilde{\theta} \rangle + \eta(\tilde{\theta})$$

si

$$\tilde{\theta} = (\tilde{\theta}_j) \in \mathbf{R}^n, \quad c = (c_j) = \int_{\mathbf{T}^n} \psi(\theta)\, d\theta, \quad \psi = c + \frac{\partial \eta}{\partial \theta},$$

où $\langle c, \tilde{\theta} \rangle = \Sigma c_j \tilde{\theta}_j$. On définit la fonction

$$K(\tilde{\theta}, \tilde{\Theta}) = H(\tilde{\theta}, \tilde{\Theta}) + \hat{\psi}(\tilde{\theta}) - \hat{\psi}(\tilde{\Theta}).$$

9.2. Pour $\widetilde{\Theta}$ fixé, la fonction $\widetilde{\theta} \mapsto K(\widetilde{\theta}, \widetilde{\Theta})$ tend vers $+\infty$ si $\|\widetilde{\theta}\| \to +\infty$, et pour $\widetilde{\theta}$ fixé, la fonction $\widetilde{\Theta} \mapsto K(\widetilde{\theta}, \widetilde{\Theta})$ tend vers $+\infty$ si $\|\widetilde{\Theta}\| \to +\infty$ (cela résulte du fait que F est globalement monotone, cf. 8.11).

On pose

$$\widetilde{f}(\widetilde{\theta}) = A\widetilde{\theta} + \varphi(\widetilde{\theta}, \psi(\widetilde{\theta})).$$

Le fait que le tore Γ_ψ soit invariant par F est équivalent à ce que $\widetilde{f}$ soit un homéomorphisme de $\mathbf{R}^n$ et

$$\psi(\widetilde{\theta}) = -H_1(\widetilde{\theta}, \widetilde{f}(\widetilde{\theta})) = H_2(\widetilde{f}^{-1}(\widetilde{\theta}), \widetilde{\theta}) \tag{9.3}$$

$$\left(\frac{\partial}{\partial\widetilde{\theta}} H \equiv H_1, \frac{\partial}{\partial\widetilde{\Theta}} H \equiv H_2\right).$$

9.4. *Lemme. — Pour tout* $\widetilde{\Theta}$, *la fonction* $\widetilde{\theta} \to K(\widetilde{\theta}, \widetilde{\Theta})$ *a un unique point critique, qui est le minimum de cette fonction, en* $\theta_0 = \widetilde{f}^{-1}(\widetilde{\Theta})$.

Démonstration. — On a $K_1(\widetilde{\theta}, \widetilde{\Theta}) = H_1(\widetilde{\theta}, \widetilde{\Theta}) + \psi(\widetilde{\theta})$ et par 9.3

$$K_1(\widetilde{f}^{-1}(\widetilde{\Theta}), \widetilde{\Theta}) = 0.$$

Soit $\theta_0 \in \mathbf{R}^n$ un point critique de la fonction $\widetilde{\theta} \to K(\widetilde{\theta}, \widetilde{\Theta})$. On a

$$H_1(\theta_0, \widetilde{\Theta}) + \psi(\theta_0) = 0$$

et d'autre part 9.3 donne

$$H_1(\theta_0, \widetilde{f}(\theta_0)) + \psi(\theta_0) = 0.$$

Comme $\widetilde{\Theta} \mapsto H_1(\theta, \widetilde{\Theta})$ est injective (cf. 8.4), on obtient $\widetilde{f}(\theta_0) = \widetilde{\Theta}$. L'unicité du point critique suit du fait que $\widetilde{f}$ est un homéomorphisme de $\mathbf{R}^n$. Il en résulte que la fonction $\widetilde{\theta} \to K(\widetilde{\theta}, \widetilde{\Theta})$ a bien un minimum en un point noté θ_m. Comme la fonction $\widetilde{\theta} \mapsto K(\widetilde{\theta}, \widetilde{\Theta})$ est de classe C^1 (puisque H est de classe C^2 et $\widehat{\psi}$ est de classe C^1 car Γ_ψ est C^0-lagrangien), θ_m est un point critique de cette fonction et donc $\theta_m = \widetilde{f}^{-1}(\widetilde{\Theta})$.

Il suit aussi que la fonction $\widetilde{\theta} \mapsto K(\widetilde{\theta}, \widetilde{\Theta})$ a un unique minimum en θ_0. ■

Par la même démonstration on a aussi le

9.5. *Lemme. — Pour tout* $\widetilde{\theta} \in \mathbf{R}^n$, *la fonction* $\widetilde{\Theta} \mapsto K(\widetilde{\theta}, \widetilde{\Theta})$ *a un unique point critique, qui est le minimum de cette fonction, au point* $\widetilde{\Theta}_0 = \widetilde{f}(\widetilde{\theta})$.

9.6. *Fin de la démonstration de* 8.14.

On se donne $v \in \mathbf{R}^n$, $s \in \mathbf{R}^*$ et $\widetilde{\theta} \in \mathbf{R}^n$. On pose

$$\Delta_{sv}(\widetilde{\theta}) = \frac{1}{s^2}[K(\widetilde{\theta} + sv, \widetilde{f}(\widetilde{\theta})) + K(\widetilde{\theta} - sv, \widetilde{f}(\theta)) - 2K(\widetilde{\theta}, \widetilde{f}(\theta))].$$

Par le lemme 9.4 on obtient

$$\Delta_{sv}(\widetilde{\theta}) \geqslant 0 \quad \text{pour tout } \widetilde{\theta},\ v \text{ et } s. \tag{9.7}$$

On peut écrire

$$\Delta_{sv}(\tilde{\theta}) = \frac{1}{s^2}[\mathrm{H}(\tilde{\theta} + sv, \tilde{f}(\tilde{\theta})) + \mathrm{H}(\tilde{\theta} - sv, \tilde{f}(\tilde{\theta})) - 2\mathrm{H}(\tilde{\theta}, \tilde{f}(\tilde{\theta}))]$$
$$+ \frac{1}{s^2}[\hat{\psi}(\tilde{\theta} + sv) + \hat{\psi}(\tilde{\theta} - sv) - 2\hat{\psi}(\tilde{\theta})].$$

On fait tendre s vers 0 et on prend la limite au sens des distributions (*i.e.* dans le dual topologique de l'espace vectoriel topologique $\mathrm{C}^\infty_{\mathrm{K}}(\mathbf{R}^n)$ des fonctions de classe C^∞ sur $\mathbf{R}^n$ à supports compacts avec sa topologie limite inductive), la limite étant prise (ce qui nous suffirait) pour la topologie faible induite par la dualité $\sigma((\mathrm{C}^\infty_{\mathrm{K}}(\mathbf{R}^n))', \mathrm{C}^\infty_{\mathrm{K}}(\mathbf{R}^n))$ [S_4].

On en déduit

$$\textbf{(9.8)} \qquad {}^t v\, \mathrm{D}^2\, \hat{\psi}\, v + {}^t v \mathrm{H}_{11}(\ ,\tilde{f})\, v = \nu^1_v,$$

où ν^1_v est une distribution $\geqslant 0$ et donc une mesure de Radon $\geqslant 0$ (cela résulte de 9.7), ${}^t v(\mathrm{H}_{11}(\ ,\tilde{f})\, v \in \mathrm{C}^0(\mathbf{R}^n, \mathbf{R})$ (H est de classe C^2) et $\mathrm{D}^2\, \hat{\psi} = \mathrm{D}\psi$ désigne la matrice dérivée seconde de $\hat{\psi}$, les dérivées partielles étant prises au sens des distributions.

Si u et $v \in \mathbf{R}^n$

$${}^t v\, \mathrm{D}^2\, \hat{\psi}\, u = \frac{1}{4}({}^t(u + v)\, \mathrm{D}^2\, \hat{\psi}\, (u + v) - {}^t(u - v)\, \mathrm{D}^2\, \hat{\psi}\, (u - v)),$$

l'égalité étant prise au sens des distributions (c'est vrai pour les fonctions de $\mathbf{R}^n$ de classe C^∞ à supports compacts et donc par dualité (ou par densité) pour les distributions).

On a montré la

9.9. *Proposition. — Toutes les dérivées partielles d'ordre 1 de* ψ *(i.e.* $\mathrm{D}\psi$*) sont des mesures de Radon.*

Du lemme 9.5, on déduit

$$\mathrm{K}(\tilde{f}^{-1}(\Theta), \Theta + vs) \geqslant \mathrm{K}(\tilde{f}^{-1}(\Theta), \Theta), \quad \Theta \in \mathbf{R}^n, \quad v \in \mathbf{R}^n, \quad s \in \mathbf{R}^*;$$

donc la fonction

$$\check{\Delta}_{sv}(\Theta) = \frac{1}{s^2}[\mathrm{K}(\tilde{f}^{-1}(\Theta), \Theta + vs) + \mathrm{K}(\tilde{f}^{-1}(\Theta), \Theta - sv) - 2\mathrm{K}(\tilde{f}^{-1}(\Theta), \Theta)]$$
$$= \frac{1}{s^2}[\mathrm{H}(\tilde{f}^{-1}(\Theta), \Theta + vs) + \mathrm{H}(\tilde{f}^{-1}(\Theta), \Theta - sv) - 2\mathrm{H}(\tilde{f}^{-1}(\Theta), \Theta)]$$
$$+ \frac{1}{s^2}[\hat{\psi}(\Theta + sv) + \hat{\psi}(\Theta - sv) - 2\hat{\psi}(\Theta)]$$

vérifie $\check{\Delta}_{sv} \geqslant 0$.

On fait tendre s vers 0 et on prend la limite au sens des distributions :

$$\textbf{(9.10)} \qquad {}^t v \mathrm{H}_{22}(\tilde{f}^{-1},\ )\, v - {}^t v\, \mathrm{D}^2\, \hat{\psi}\, v = \nu^2_v,$$

où ν^2_v est une mesure de Radon positive.

On pose

$$ {}^t v H_{11}(\ ,\tilde{f})\, v = a_v\, d\theta, \qquad a_v \in L^\infty \equiv L^\infty(\mathbf{R}^n, d\theta, \mathbf{R}) $$

$$ {}^t v H_{22}(\tilde{f}^{-1},\ )\, v = b_v\, d\theta, \qquad b_v \in L^\infty $$

et

$$ \mu_v = {}^t v\, D^2 \hat{\psi}\, v \quad (\mu_v \text{ est une mesure de Radon pour tout } v \in \mathbf{R}^n). $$

On écrit la *décomposition de Riesz* de μ_v

$$ \mu_v = \mu_v^+ - \mu_v^- $$

la décomposition étant unique. Les égalités 9.8 et 9.10 donnent

$$ a_v\, d\theta - \nu_v^1 = \mu_v^- - \mu_v^+; \tag{9.11} $$

$$ b_v\, d\theta - \nu_v^2 = \mu_v^+ - \mu_v^-. \tag{9.12} $$

9.11 donne

$$ \mu_v^- = (a_v\, d\theta - \nu_v^1)^+. $$

Ceci implique, par le théorème de Radon-Nykodim,

$$ \mu_v^- = \ell_1(x)\, a_v^+ . d\theta \tag{9.13} $$

où $\ell_1(x)$ est $a_v^+\, d\theta$ mesurable et vérifie $0 \leqslant \ell_1(x) \leqslant 1$ pour presque tout x par rapport à la mesure $a_v^+\, d\theta$.

Si μ est une mesure de Radon sur $\mathbf{R}^n$, et $h \geqslant 0$ est une fonction continue sur $\mathbf{R}^n$ à support compact, *i.e.* $h \in C_K^0(\mathbf{R}^n)$, alors [$B_3$]

$$ \mu^+(h) = \sup_{0 \leqslant g \leqslant h} \mu(g), \quad g \in C_K^0(\mathbf{R}^n), $$

d'où

$$ \mu_v^-(h) \leqslant \int_{\mathbf{R}^n} h a_v^+\, d\theta $$

puisque $\nu_v^1 \geqslant 0$, où on a posé $a_v^+ = \sup(a_v, 0) \in L^\infty$. Ceci implique 9.13.

Le même argument appliqué à 9.12 nous permet de conclure que

$$ \mu_v^+ = \ell_2\, b_v^+\, d\theta \tag{9.14} $$

où la fonction ℓ_2 est $b_v^+\, d\theta$ mesurable et vérifie $0 \leqslant \ell_2(x) \leqslant 1$ presque partout par rapport à la mesure $b_v^+\, d\theta$.

9.13 et 9.14 donnent, pour tout $v \in \mathbf{R}^n$,

$$ {}^t v\, D^2 \hat{\psi}\, v = {}^t v\, D\psi\, v \in L^\infty. $$

On a montré que toutes les dérivées partielles de ψ au sens des distributions sont dans L^∞.

Il est standard [S_6] (en utilisant des opérateurs de convolutions [S_5]) que cela implique que l'application $\psi : \mathbf{R}^n \to \mathbf{R}^n$ soit lipschitzienne.

Inégalités.

Comme ψ est lipschitzienne, ψ est presque partout, pour la mesure de Lebesgue, dérivable au sens de Fréchet [S_6].

On note $D\psi(x)$ l'application linéaire tangente définie pour $d\theta$ presque tout $x \in \mathbf{R}^n$.

La fonction matrice $x \to D\psi(x)$ a ses coefficients dans L^∞ et est égale, considérée comme matrice à coefficients distributions, à la matrice dérivée $D^2 \hat{\psi} = D\psi$ prise au sens des distributions.

L'identité 9.8 donne pour les matrices symétriques (l'inégalité étant prise pour l'ordre standard sur les matrices symétriques)

$$H_{11}(\theta, \tilde{f}(\theta)) + D\psi(\theta) \geqslant 0 \quad d\theta \text{ presque partout;}$$

et 9.10

$$H_{22}(\tilde{f}^{-1}(\theta), \theta) - D\psi(\theta) \geqslant 0 \quad d\theta \text{ presque partout.}$$

De 8.9 et 8.10 on tire

$$H_{11}(\theta, \tilde{f}(\theta)) = \bar{b}^{-1}(\theta)\, \bar{a}(\theta)$$

et

$$H_{22}(\theta, \tilde{f}(\theta)) = \bar{d}(\theta)\, \bar{b}^{-1}(\theta).$$

On obtient que l'application

$$\tilde{\theta} \mapsto \tilde{f}(\tilde{\theta}) = A\tilde{\theta} + \varphi(\tilde{\theta}, \psi(\tilde{\theta})), \quad \tilde{\theta} \in \mathbf{R}^n$$

est lipschitzienne.

Comme F^{-1} laisse invariant Γ_ψ, l'application

$$\tilde{\theta} \in \mathbf{R}^n \mapsto \tilde{f}^{-1}(\tilde{\theta}) = A^{-1}\tilde{\theta} + \varphi_-(\tilde{\theta}, \psi(\tilde{\theta})) \in \mathbf{R}^n$$

est aussi lipschitzienne où $\tilde{F}^{-1}(\tilde{\theta}, r) = (A^{-1}\tilde{\theta} + \varphi_-(\tilde{\theta}, r), R)$.

On en conclut que

(9.15) $\tilde{f}$ est un homéomorphisme bilipschitzien de $\mathbf{R}^n$.

Il suit que :

- B est un ensemble de mesure de Lebesgue nulle, $B \subset \mathbf{R}^n$ est équivalent à :
- $\tilde{f}(B)$ est de mesure de Lebesgue nulle.

On voit finalement que

$$(\bar{d}\bar{b}^{-1}) \circ \tilde{f}^{-1} - D\psi \geqslant 0 \quad d\theta \text{ presque partout}$$

équivaut à

$$\bar{d}\bar{b}^{-1} - D\psi \circ \tilde{f} \geqslant 0 \quad d\theta \text{ presque partout.} \quad \blacksquare$$

10. Unicité

10.1. Soit F un difféomorphisme symplectique monotone globalement positif de $\mathbf{A}^n$, homotope à l'identité. Pour $j = 1, 2$, on suppose que F laisse invariant un tore T_j, C^0-lagrangien, graphe d'une fonction $\psi_j \in C^0(\mathbf{T}^n, \mathbf{R}^n)$.

Choisissons un relèvement $\widetilde{F}$ de F à $\widetilde{\mathbf{A}}^n$. Notons $\pi : \mathbf{R}^n \to \mathbf{T}^n$ la projection canonique, et $p_1 : \mathbf{A}^n \to \mathbf{T}^n$, $\widetilde{p}_1 : \widetilde{\mathbf{A}}^n \to \mathbf{R}^n$ les projections sur la première coordonnée (θ ou $\widetilde{\theta}$). Le difféomorphisme $\widetilde{F}$ préserve les graphes $\widetilde{T}_j = \Gamma_{\widetilde{\psi}_j}$ des applications $\widetilde{\psi}_j = \psi_j \circ \pi$, pour $j = 1, 2$; l'application $\theta \mapsto p_1 \circ F(\theta, \psi_j(\theta))$ est un homéomorphisme f_j de $\mathbf{T}^n$, homotope à l'identité, et l'application $\widetilde{\theta} \mapsto \widetilde{p}_1 \circ \widetilde{F}(\widetilde{\theta}, \widetilde{\psi}_j(\widetilde{\theta}))$ est un relèvement $\widetilde{f}_j$ de f_j appartenant au groupe

$$D^0(\mathbf{T}^n) = \{\widetilde{f} \in \text{Homéo}(\mathbf{R}^n), \widetilde{f} - \text{Id est } \mathbf{Z}^n\text{-périodique}\}.$$

10.2. Pour un homéomorphisme $\widetilde{f} \in D^0(\mathbf{T}^n)$, induisant l'homéomorphisme f de $\mathbf{T}^n$, on désigne par $M(f)$ l'ensemble des mesures de probabilité sur $\mathbf{T}^n$ invariantes par f, et on définit *l'ensemble (des vecteurs) de rotation* de $\widetilde{f}$ par

$$R(\widetilde{f}) = \left\{ \int_{\mathbf{T}^n} (\widetilde{f}(\theta) - \theta)\, d\mu(\theta),\ \mu \in M(f) \right\} \subset \mathbf{R}^n.$$

L'ensemble $R(\widetilde{f})$ est compact convexe non vide puisque $M(f) \subset M = \{$ mesures de probabilité sur $\mathbf{T}^n \}$ a les mêmes propriétés (sur M on met la topologie de la convergence vague).

L'ensemble $R(\widetilde{f})$ est un invariant de conjugaison de $\widetilde{f}$ dans le groupe $D^0(\mathbf{T}^n)$. Lorsque $R(\widetilde{f})$ est un singleton $\{\alpha\}$, alors la suite $\dfrac{\widetilde{f}^k - \text{Id}}{k} = (1/k) \sum_{j=0}^{k-1} \eta \circ f^j$ (avec $k \in \mathbf{N}^*$ et $\eta = \widetilde{f} - \text{Id}$) converge uniformément vers α lorsque $k \to +\infty$. Cela résulte immédiatement du lemme de [H_4], p. 487. On dit alors que $\widetilde{f}$ a un vecteur de *rotation* (égal à α). C'est le cas notamment si :

- f est *uniquement ergodique*, c'est-à-dire que $M(f)$ a un seul élément;
- $n = 1$ (théorie du nombre de rotation).

Lorsque n est au moins égal à 2, $R(\widetilde{f})$ est un compact convexe non vide de $\mathbf{R}^n$, « en général » non réduit à un singleton (par exemple pour un difféomorphisme de la forme $(\theta_1, \theta_2) \mapsto (\theta_1, \theta_2 + \varphi(\theta_1))$ avec une fonction $\varphi \in C^\infty(\mathbf{T}^1, \mathbf{R})$ non constante).

10.3. On met sur $D^0(\mathbf{T}^n)$ la topologie de la convergence uniforme; le groupe $D^0(\mathbf{T}^n)$ est alors un groupe topologique polonais. On vérifie que l'application

$$(\widetilde{f}, \mu) \in D^0(\mathbf{T}^n) \times M \mapsto f_*(\mu) \in M$$

est continue, où $f_*(\mu)$ désigne la mesure de l'image directe de μ par f.

L'application $\widetilde{f} \in D^0(\mathbf{T}^n) \mapsto R(\widetilde{f}) \subset \mathbf{R}^n$ est une application multivoque semi-continue supérieurement. Cela vient de ce que l'application $\widetilde{f} \mapsto M(f) \subset M$ a la même propriété puisque le graphe de $M(f)$

$$\{(f, M(f)) \subset D^0(\mathbf{T}^n) \times M, f \in D^0(\mathbf{T}^n)\}$$

est fermé dans $D^0(\mathbf{T}^n) \times M$.

Il suit que si $\alpha \in \mathbf{R}^n$, l'ensemble $\{\widetilde{f} \in D^0(\mathbf{T}^n), \alpha \in R(\widetilde{f})\}$ est fermé dans $D^0(\mathbf{T}^n)$.

La propriété de semi-continuité supérieure de l'application $\tilde{f} \mapsto \mathrm{R}(\tilde{f})$ permet d'obtenir de nombreux résultats et a de nombreux avantages sur les définitions dépendant de la convergence ponctuelle de $(\tilde{f}^n - \mathrm{Id})/n$; (voir à ce propos la définition que nous avons donnée en 1975 [H_2, XIII] et la référence récente [M_8]).

10.4. *Théorème.* — *Si dans la situation de* 10.1, *on a* $\mathrm{R}(\tilde{f}_1) \cap \mathrm{R}(\tilde{f}_2) \neq \emptyset$, *alors* $\mathrm{T}_1 \cap \mathrm{T}_2 \neq \emptyset$.

10.5. *Corollaire.* — *Si* $\mathrm{R}(\tilde{f}_1) \cap \mathrm{R}(\tilde{f}_2) \neq \emptyset$, *et si toute orbite de* f_1 *est dense dans* $\mathbf{T}^n$, *alors* $\mathrm{T}_1 = \mathrm{T}_2$.

La démonstration de ce théorème s'inspire fortement du travail de J. Mather [M_3]. Au § 12 nous généraliserons le résultat de [M_3] : nous montrerons que si F est un difféomorphisme monotone globalement positif de $\mathbf{A}^n$ qui laisse invariant un tore T, C^0-lagrangien, qui est un graphe sur $\mathbf{T}^n \times \{0\}$, alors chaque $x \in \mathrm{T}$ définit pour F un état d'énergie minimale au sens d'Aubry, cf. [A_4], [C_1], [K] et 12.5. La démonstration que nous donnerons en 12 est fortement inspirée par le cas $n = 1$ traité par J. Mather [M_3].

11. Démonstration du théorème 10.4

Pour $j = 1, 2$ il existe $c_j \in \mathbf{R}$ vérifiant

$$c_j = \mathrm{H}(\tilde{\theta}, \tilde{f}_j, (\tilde{\theta})) + \hat{\psi}_j(\tilde{\theta}) - \hat{\psi}_j(\tilde{f}_j(\tilde{\theta})) \tag{11.1}$$

pour tout $\tilde{\theta} \in \mathbf{R}^n$.

En effet, $\tilde{f}_j$ est un homéomorphisme bilipschitzien (cf. 9.15). La fonction sur $\mathbf{R}^n$ $\tilde{\theta} \mapsto \mathrm{H}(\tilde{\theta}, \tilde{f}_j(\tilde{\theta})) + \hat{\psi}_j(\tilde{\theta}) - \hat{\psi}_j(\tilde{f}_j(\tilde{\theta}))$ est lipschitzienne. Il suit qu'elle est Lebesgue presque partout dérivable au sens de Fréchet et par 9.3 sa dérivée est presque partout nulle; on conclut que cette fonction est constante.

Du lemme 9.5 on déduit, pour tout $\tilde{\theta} \in \mathbf{R}^n$,

$$\mathrm{H}(\tilde{\theta}, \tilde{f}_j(\tilde{\theta})) + \hat{\psi}_j(\tilde{\theta}) - \hat{\psi}_j(\tilde{f}_j(\tilde{\theta})) < \mathrm{H}(\tilde{\theta}, \tilde{\Theta}) + \hat{\psi}_j(\tilde{\theta}) - \hat{\psi}_j(\tilde{\Theta}), \tag{11.2}$$

si $\tilde{\Theta} \neq \tilde{f}_j(\tilde{\theta})$.

11.3. Comme F est homotope à l'identité, on a $\mathrm{H}(\tilde{\theta}, \tilde{f}(\theta)) = \mathrm{S}(\tilde{\theta}, \tilde{f}(\tilde{\theta}) - \tilde{\theta})$ (cf. le paragraphe après 8.3) ; il en résulte que la fonction $\tilde{\theta} \mapsto \mathrm{H}(\tilde{\theta}, \tilde{f}(\tilde{\theta}))$ est $\mathbf{Z}^n$-périodique.

On peut écrire (cf. 9.1) pour $j = 1, 2$

$$\hat{\psi}_j = \ell_{[\psi_j]} + \eta_j, \quad \eta_j \in \mathrm{C}^1(\mathbf{T}^n, \mathbf{R})$$

où $[\psi_j] = \int_{\mathbf{T}^n} \psi_j(\theta)\, d\theta$ et $\ell_{[\psi_j]}(u) = \langle [\psi_j], u \rangle$, $u \in \mathbf{R}^n$.

Soit $\mu_j \in \mathrm{M}(f_j)$. On intègre 11.1 par rapport à la mesure μ_j, d'où, en utilisant 11.3,

$$c_j = \int_{\mathbf{T}^n} \mathrm{H}(\theta, \tilde{f}_j(\theta))\, d\mu_j(\theta) - \ell_{[\psi_j]}\Big(\int_{\mathbf{T}^n} (\tilde{f}_j(\theta) - \theta)\, d\mu_j(\theta)\Big). \tag{11.4}$$

Comme par hypothèse $R(\tilde{f}_1) \cap R(\tilde{f}_2) \neq \emptyset$, on peut trouver $\mu_j \in M(f_j)$, $j = 1, 2$, et $\alpha \in \mathbf{R}^n$ tels que $\alpha = \int_{\mathbf{T}^n} (\tilde{f}_j - \mathrm{Id})\, d\mu_j$. On pose $a_j = \int_{\mathbf{T}^n} H(\theta, \tilde{f}_j(\theta))\, d\mu_j(\theta)$.

11.5. *Lemme.* — *On a* $a_1 = a_2$.

Démonstration. — Nous allons montrer que $a_1 \leqslant a_2$. La relation 11.4 donne

$$c_1 = a_1 - \ell_{[\psi_1]}(\alpha)$$

et 11.2 donne

$$c_1 \leqslant H(\tilde{\theta}, \tilde{f}_2(\tilde{\theta})) + \hat{\psi}_1(\tilde{\theta}) - \hat{\psi}_1(\tilde{f}_2(\tilde{\theta})) \tag{11.6}$$

avec l'égalité si et seulement si $\tilde{f}_2(\tilde{\theta}) = \tilde{f}_1(\tilde{\theta})$.

En intégrant 11.6 par rapport à μ_2, on obtient

$$c_1 \leqslant a_2 - \ell_{[\psi_1]}(\alpha).$$

Le lemme suit en changeant les rôles des indices 1 et 2. ■

11.7. *Fin de la démonstration du théorème* 10.4.

Il résulte de 8.2 que

$$\bullet\ \tilde{\theta}_1 \in \mathbf{R}^n \quad \text{vérifie } \psi_1(\tilde{\theta}_1) \neq \psi_2(\tilde{\theta}_1)$$

équivaut à

$$\bullet\ \tilde{f}_1(\tilde{\theta}_1) \neq \tilde{f}_2(\tilde{\theta}_1)$$

et donc, si pour tout $\tilde{\theta}$, $\psi_1(\tilde{\theta}) \neq \psi_2(\tilde{\theta})$, alors $\tilde{f}_2(\tilde{\theta}) \neq \tilde{f}_1(\tilde{\theta})$ pour tout $\tilde{\theta} \in \mathbf{R}^n$.

Intégrant 11.6 par rapport à la mesure μ_2, en supposant que $\tilde{f}_1(\tilde{\theta}) \neq \tilde{f}_2(\tilde{\theta})$ pour tout $\tilde{\theta}$, on conclut

$$c_1 = a_1 - \ell_{[\psi_1]}(\alpha) < a_2 - \ell_{[\psi_1]}(\alpha)$$

ce qui contredit $a_1 = a_2$.

Il s'ensuit que $\Gamma_{\psi_1} \cap \Gamma_{\psi_2} = \emptyset$ est absurde. ■

11.8. *Remarque.* — Si μ_1 ou μ_2 a son support égal à $\mathbf{T}^n$, alors la démonstration ci-dessus montre que $\Gamma_{\psi_1} = \Gamma_{\psi_2}$. Cette hypothèse est vérifiée si f_1 ou f_2 est un homéomorphisme minimal de $\mathbf{T}^n$ (*i.e.* a toutes ses orbites denses). Ceci montre 10.5.

12. Compléments

Les compléments qui vont suivre sont importants si on veut mettre ce que nous avons fait en perspective avec les théories des états d'énergies minimales de S. Aubry [A_4] (voir aussi [K]) et les mesures d'énergies minimales de J. Mather [M_2] [M_4] et [M_6]. Néanmoins, pour ne pas trop allonger ce travail, nous renvoyons le lecteur à [M_2] [M_4] pour la définition des mesures d'énergie minimale.

12.1. On suppose que F est un difféomorphisme symplectique de $\mathbf{A}^n$ globalement monotone positif homotope à l'identité, et que F laisse invariant $T = \Gamma_\psi$, un tore C^0-lagrangien.

12.2. Soit $\tilde{\theta}_0 \in \mathbf{R}^n$, $\tilde{\theta}_i = \tilde{f}(\tilde{\theta}_{i-1})$, $i \geqslant 1$, i entier.

Il suit de 11.1 qu'il existe $c \in \mathbf{R}$, tel que pour tout $p \in \mathbf{N}^*$

$$c = \frac{1}{p}\sum_{i=0}^{p-1} H(\theta_i, \theta_{i+1}) + \frac{1}{p}(\eta(\theta_0) - \eta(\theta_p)) + \frac{1}{p}\ell_{[\psi]}(\theta_0 - \theta_p) \tag{12.3}$$

Soit $(\theta'_i)_{0 \leqslant i \leqslant p}$ une suite dans $\mathbf{R}^n$ avec $\theta'_0 = \theta_0$ et $\theta'_p = \theta_p$; alors

$$\sum_{i=0}^{p-1} H(\theta_i, \theta_{i+1}) \leqslant \sum_{i=0}^{p-1} H(\theta'_i, \theta'_{i+1}) \tag{12.4}$$

et on a *l'égalité si et seulement si*

$$\theta'_j = \theta_j, \quad j = 0, \ldots, p.$$

C'est une conséquence immédiate de 11.2.

On peut reformuler 12.4 en disant que chaque orbite de $x \in T$ définit un *état d'énergie minimale au sens d'Aubry* [A_4].

12.5. Un *état* (ou configuration) est une suite $(\theta_j)_{j \in \mathbf{Z}}$ dans $\mathbf{R}^n$. Un état $(\theta_j)_j$ est un *état d'équilibre* s'il vérifie, pour tout $i \in \mathbf{Z}$,

$$H_2(\theta_i, \theta_{i+1}) + H_1(\theta_{i+1}, \theta_{i+2}) = 0.$$

Il revient au même de considérer l'état d'équilibre $(\theta_j)_j$ et l'orbite de $\tilde{F}$ associé, $(\theta_i, r_i)_{i \in \mathbf{Z}}$, avec $r_i = -H_1(\theta_i, \theta_{i+1}) = H_2(\theta_{i-1}, \theta_i)$.

Un *état* $(\theta_j)_j$ est *d'énergie minimale* si pour tout entier relatif $k \leqslant \ell$ la fonction $(u_j) \in B_{k,\ell} \to \sum_{j=k}^{\ell} H(u_j, u_{j+1})$ définie sur

$$B_{k,\ell} = \{(u_j)_{k \leqslant j \leqslant \ell+1} \mid u_j \in \mathbf{R}^n, u_k = \theta_k, u_{\ell+1} = \theta_{\ell+1}\}$$

réalise son minimum pour $u_j = \theta_j$, $j = k, \ldots, \ell + 1$.

Remarque. — Soit $L : \mathbf{A}^2 \to \mathbf{A}^2$, $L(\theta, r) = (\theta + r, r)$, et T un tore invariant par L non lagrangien (on ne suppose même pas que T soit un graphe), cf. 5; alors tout $x \in T$ définit un état d'énergie minimale pour L!

12.6. On suppose que F vérifie 12.1 et on définit p_1, $\tilde{p}_1$ et π comme 10.1. On suppose que F laisse invariant un ensemble compact $K \subset \mathbf{A}^n$ qui est le graphe d'une fonction continue ψ_1 sur $p_1(K) = \mathbf{T}^n$. Soit g l'homéomorphisme induit par $F_{|K}$ sur $p_1(K)$:

$$g(x) \subset p_1(F(x, \psi_1(x))) \quad \text{pour } x \in p_1(K).$$

$\tilde{F}$ laisse invariant $\tilde{K} = \pi^{-1}(K) \subset \tilde{\mathbf{A}}^n$, qui est le graphe de $\psi_1 \circ \pi$ sur $\tilde{p}_1(\tilde{K})$. On note $\tilde{g}$ l'homéomorphisme de $p_1(\tilde{K})$ induit par $\tilde{F}_{|\tilde{K}}$.

La fonction $\tilde{g} - \mathrm{Id} : \tilde{p}_1(\tilde{\mathrm{K}}) \to \mathbf{R}^n$ est invariante par l'action de $\mathbf{Z}^n$ sur $\tilde{p}_1(\tilde{\mathrm{K}})$ puisque $\tilde{p}_1(\tilde{\mathrm{K}})$ l'est et que $\tilde{\mathrm{F}}(\tilde{\theta} + q, r) = \hat{q} + \tilde{\mathrm{F}}(\tilde{\theta}, r)$, $q \in \mathbf{Z}^n$, $\hat{q} = (q, 0) \in \mathbf{R}^n \times \mathbf{R}^n$. Il en est de même pour la fonction $x \in \tilde{p}_1(\tilde{\mathrm{K}}) \mapsto \mathrm{H}(x, \tilde{g}(x))$.

On suppose que $\mathrm{F}_{|\mathrm{K}}$ laisse invariante une mesure de probabilité μ ayant pour support K. On suppose de plus que

$$\int_{\mathrm{K}} (\tilde{g} - \mathrm{Id})\, d\nu = \alpha \in \mathrm{R}(\tilde{f}) = \text{ensemble de rotation de } \tilde{f},$$

où $\nu = (p_1)_* \mu$ est l'image par p_1 de la mesure μ et $\tilde{f}$ désigne l'homéomorphisme de $\tilde{\mathrm{F}}_{|\tilde{\mathrm{T}}}$ sur $\mathbf{R}^n$ induit par $\tilde{p}_1$; cf. 10.1 et 12.1.

Soit $\mu_1 \in \mathrm{M}(f)$ tel que $\alpha = \int_{\mathrm{T}^n} (\tilde{f} - \mathrm{Id})\, d\mu_1$.

Alors la démonstration que nous avons donnée en 11 montre que

$$c + \langle [\psi], \alpha \rangle = a = \int_{\mathrm{T}^n} \mathrm{H}(\theta, \tilde{f}(\theta))\, d\mu_1(\theta) \leqslant \int_{\mathrm{K}} \mathrm{H}(x, \tilde{g}(x))\, d\nu(x).$$

On obtient plus : comme $\psi_1 = -\mathrm{H}_1(\ , \tilde{g})$, $\mathrm{K} = \Gamma_{\psi_1}$ et que le support de ν est égal à $p_1(\mathrm{K})$, si, pour un $x \in p_1(\mathrm{K})$, $\psi(x) \neq \psi_1(x)$, alors

$$a < \int_{p_1(\mathrm{K})} \mathrm{H}(x, \tilde{g}(x))\, d\nu(x).$$

En effet, d'après 11.3,

$$c \leqslant \mathrm{H}(y, \tilde{g}(y)) + \hat{\psi}(y) - \hat{\psi}(\tilde{g}(y)) \quad \text{pour } y \in p_1(\mathrm{K}),$$

l'inégalité étant stricte en x et donc sur un ouvert de K.

Nous avons ainsi montré que si μ est une mesure d'énergie minimale au sens de J. Mather [M_2], [M_4], *vérifiant les hypothèses ci-dessus, alors*

$$\mathrm{K} \subset \mathrm{T} = \Gamma_\psi.$$

Si la mesure de probabilité μ est d'énergie minimale, comme ν et μ_1 ont même nombre de rotation, on a, par définition,

$$\int_{p_1(\mathrm{K})} \mathrm{H}(x, \tilde{g}(x))\, d\nu(x) \leqslant a.$$

Chapitre II

INÉGALITÉS « A PRIORI » GÉOMÉTRIQUES : GÉNÉRALISATION PERTURBATIVE DU DEUXIÈME THÉORÈME DE BIRKHOFF

1. Introduction

Nous nous proposons de donner, dans ce chapitre, une généralisation partielle et perturbative du deuxième théorème de Birkhoff. Nous montrons (théorème 2.2) que si F est un difféomorphisme symplectique de $\mathbf{A}^n$, qui est une perturbation en topologie C^1 d'un difféomorphisme complètement intégrable L monotone positif laissant invariante une sous-variété $N \subset \mathbf{A}^n$ de classe C^1, lagrangienne, compacte connexe ayant une classe de Maslov nulle et si $F_{|N} \equiv F_N : N \to N$ est récurrent par chaîne et N ne s'approche pas de l'infini, alors N est le graphe d'une fonction $\psi \in C^1(\mathbf{T}^n, \mathbf{R}^n)$.

Si on considère un tore T de dimension n donné par le théorème KAM, alors T est une variété lagrangienne de classe C^1 et F_T est récurrent par chaîne. Le tore T peut disparaître (cf. [H_3], [H_1]). L'inégalité *a priori* du chapitre I force T de rester, à sa première disparition, un tore C^0-lagrangien, graphe d'une fonction lipschitzienne. La disparition des tores invariants lagrangiens reste à l'heure actuelle mathématiquement inexpliquée.

Ce que la généralisation perturbative du deuxième théorème de Birkhoff affirme, c'est qu'un tore donné par le théorème KAM sera un graphe si sa classe de Maslov est nulle et s'il ne s'approche pas de l'infini.

Pour une généralisation plus forte du deuxième théorème de Birkhoff dans le cas des flots géodésiques sur $\mathbf{T}^2$, voir Byalyi et Polterovich [B_9] ainsi que [B_5] et le travail de Byalyi [B_6].

Nous allons décrire le contenu de ce chapitre.

L'idée fondamentale de Birkhoff consiste à faire agir l'application tangente TF de F sur $\mathbf{A} \times \mathbf{P}_1(\mathbf{R})$ et sur « les chemins négatifs » (cf. [H_3]). Si $n \geqslant 2$, nous avons montré au chapitre I qu'il faut seulement considérer les tores invariants lagrangiens, et faire des hypothèses de positivité de la torsion qui ont seulement un sens pour les difféomorphismes symplectiques. Toute la démonstration consiste à étudier l'action de $\widetilde{TF}_{|N} \equiv \widetilde{TF}_N$ sur $N \times \widetilde{\Lambda(n)}$, où $\Lambda(n)$ désigne la grassmannienne des sous-espaces vectoriels lagrangiens de dimension n de $(\mathbf{R}^{2n}, w)$ et $\widetilde{\Lambda(n)}$ le revêtement universel de $\Lambda(n)$. Nous travaillons sur

la sous-variété lagrangienne N car la notion de chemin négatif, nous semble-t-il, nécessite des propriétés de séparations particulières au cas $n = 1$.

Le paragraphe 3 consiste en des préliminaires de géométrie symplectique, la proposition qui fait tout marcher étant 3.24. C'est ici que l'hypothèse de positivité de L est transparente.

Au paragraphe 4 nous déterminons les conditions de proximité, la complication venant en partie de ce qu'on doit introduire deux voisinages W et V. La raison venant de l'exemple dans la remarque qui suit 4.9.

Au paragraphe 5 on introduit la situation abstraite : N est un ensemble compact connexe invariant par F et $\xi \to N$ est un fibré vectoriel lagrangien $\xi \hookrightarrow T\mathbf{A}^n_{|N}$ invariant par $TF_{|N} \equiv TF_N$. Pour le théorème, on considère le cas où $\xi \equiv TN$ est le fibré tangent de N.

Le fait que la classe de Maslov soit nulle est utilisé pour relever l'application $y \in N \xrightarrow{\bar{\sigma}} [\xi_y] \in \Lambda(n)$ en une application $\sigma : N \to \widetilde{\Lambda(n)}$, où $[\xi_y]$ dénote le plan lagrangien $\xi_y \subset T_y \mathbf{A}^n$ considéré comme élément de $\Lambda(n)$.

Au paragraphe 6 nous rappelons la définition et quelques propriétés des homéomorphismes récurrents par chaîne ainsi que celles de l'ensemble récurrent par chaîne d'un homéomorphisme d'un espace métrique non vide X.

Les propositions qui permettent de démontrer le théorème 2.2 se trouvent au paragraphe 7.

Au paragraphe 8 nous donnons un théorème de structure quand on ne suppose pas que F_N est récurrent par chaîne.

La démonstration du théorème 2.2 se trouve au paragraphe 9 et elle utilise une proposition 9.4 de V. I. Arnold (tirée de [B_4]). Nous avons inclus une démonstration complète de 9.4 pour assurer au lecteur que la proposition reste vraie en classe C^1.

Au paragraphe 10 nous donnons des compléments et des remarques : en 10.1 on indique des hypothèses pour éviter de demander que la sous-variété N ne s'approche pas de l'infini, et en 10.2 nous montrons comment retrouver, en partie, les inégalités *a priori* de 1.7. En 10.3, nous montrons comment la situation abstraite introduite au paragraphe 5 s'applique aux tores moustachus d'Arnold. La fin du paragraphe 10 est une discussion sur les constantes, ce qui est important pour les applications.

Les notations sont presque toujours les mêmes que celles du chapitre I, mais nous utilisons occasionnellement certaines lettres pour désigner des choses différentes.

2. Généralisation perturbative du deuxième théorème de Birkhoff

2.1. Pour $R > 0$ définissons $\mathbf{A}^n_R = \{(\theta, r) \in \mathbf{A}^n, \|r\| \leqslant R\}$. Soit N une sous-variété de $\mathbf{A}^n$ de classe C^1, lagrangienne, compact connexe.

Le fibré tangent de $\mathbf{A}^n$, $T\mathbf{A}^n$ est *canoniquement isomorphe* au fibré symplectique $\mathbf{A}^n \times (\mathbf{R}^{2n}, w)$. Soit τ_N l'application qui associe à $y \in N$ l'espace tangent de N en y,

$T_y N \subset (\mathbf{R}^{2n}, w)$, considéré comme élément de $\Lambda(n)$ ($T_y N$ est un sous-espace lagrangien de $(\mathbf{R}^{2n}, w)$) :

$$\tau_N(y) = [T_y N] \in \Lambda(n).$$

La classe de Maslov de la sous-variété lagrangienne N, notée Maslov(N), est l'image par τ_N^* dans $H^1(N, \mathbf{Z})$ du générateur 1 de $H^1(\Lambda(n), \mathbf{Z}) \cong \mathbf{Z}$ (générateur qui dépend du choix d'une orientation de $\mathbf{S}^1 = \{ z \in \mathbf{C}, | z | = 1 \}$).

La définition de Maslov(T) suppose que nous ayons trivialisé canoniquement le fibré tangent de $\mathbf{A}^n$.

Si $n \geqslant 1$, C. Viterbo a montré [V_1], [V_2] que Maslov(N) $= 0$ lorsque N est un tore lagrangien homotope à $\{ r = 0 \}$.

Pour des exemples de sous-variétés lagrangiennes de $\mathbf{A}^n$, le lecteur peut se reporter à l'article de M. Audin [A_5]. D'après D. Bennequin, c'est une question ouverte de savoir si Maslov(N) $= 0$ implique que N est homéomorphe à $\mathbf{T}^n$.

2.2. *Théorème.* — *On se donne* $0 < R < R_1$ *et un difféomorphisme complètement intégrable* L *de classe* C^1, *vérifiant*

$$D\ell(r) > 0 \quad \textit{pour} \ \| r \| \leqslant R_1,$$

où $L(\theta, r) = (\theta + \ell(r), r)$. *Il existe* $\varepsilon(L, R, R_1) > 0$ *tel que si* F *est un difféomorphisme symplectique de* $\mathbf{A}^n$, *de classe* C^1, N *une variété lagrangienne de classe* C^1 *compacte connexe, invariante par* F *(i.e.* $F(N) = N$*), vérifiant*

(2.3) $$\| F - L \|_{C^1(\mathbf{A}^n_{R_1})} < \varepsilon(L, R, R_1),$$

(2.4) $$N \subset \mathbf{A}^n_R,$$

(2.5) $$\text{Maslov}(N) = 0,$$

et

(2.6) *la restriction* $F_N \equiv F_{|N}$ *de* F *à* N *est récurrente par chaîne.*

Alors N *est le graphe d'une fonction* $\psi \in C^1(\mathbf{T}^n, \mathbf{R}^n)$.

Commentaires. — On ne suppose pas *a priori* que N est homéomorphe à $\mathbf{T}^n$.

• Nous définirons en 4.11 et 4.12 la condition de proximité en topologie C^1.

• Nous définirons en 6 la notion de récurrent par chaîne. Elle est impliquée par le fait que F_N est récurrent (*i.e.* l'ensemble non errant de F_N, $\Omega(F_N) = N$).

Cette condition est vérifiée par les tores invariants obtenus par le théorème des tores invariants de Kolmogorov-Arnold-Moser.

• La condition Maslov(N) $= 0$ est en général nécessaire : si $n = 1$, F peut laisser invariant un tore N de dimension $n = 1$ homotope à 0 dans $\mathbf{A}^n$ (par exemple un tore invariant associé à un point fixe elliptique générique quand F est de classe C^k, $k \geqslant 3$). Dans ce cas, N n'est pas un graphe et on a Maslov(N) $= 2$.

• *A priori* la démonstration donne $\varepsilon(L, R, R_1) \to 0$ si la dimension n tend vers $+\infty$, cf. 4.28.2 et 3 ainsi que 10.5.

• On peut supprimer l'hypothèse 2.4 si on fait des hypothèses à l'infini, cf. 10.1.

• Ainsi que nous l'indiquerons en 4.28.4 on peut déterminer le nombre $\varepsilon(L, R, R_1) > 0$ ne dépendant pas de R. On peut remplacer 2.3 par les conditions géométriques 4.14 à 4.25. Nous choisissons $0 < R < R_1$ pour avoir

$$\textbf{(4.14)} \qquad F^j(\mathbf{A}^n_R) \subset \mathbf{A}^n_{R_1}, \quad -1 \leqslant j \leqslant 2n,$$

et obtenir 4.25. En utilisant 4.14 on peut remplacer dans 2.2 l'hypothèse que F est un difféomorphisme de $\mathbf{A}^n$ par : F est un plongement symplectique de $\mathbf{A}^n_{R_1}$ dans $\mathbf{A}^n$ de classe C^1, vérifiant $F(\mathbf{A}^n_{R_1}) \supset \mathbf{A}^n_R$.

3. Préliminaires géométriques

Tout ce que nous allons faire jusqu'en 3.23 est probablement connu; faute de références, nous incluons des démonstrations. Dans la suite, on désigne par 1 l'élément neutre d'un groupe de Lie ou de $M_n(\mathbf{R})$.

3.1. On identifie $\mathbf{R}^n \times \mathbf{R}^n \cong \mathbf{R}^{2n} = \mathbf{R}^n \oplus i\mathbf{R}^n$ à $\mathbf{C}^n$, ou encore à $\mathbf{R}^n \oplus J\mathbf{R}^n$, où $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Sur $\mathbf{R}^{2n}$ on met la forme symplectique

$$w(v_1, v_2) = -\,{}^t v_1 J v_2, \quad v_j \in \mathbf{R}^n \times \mathbf{R}^n, \quad j = 1, 2.$$

On désigne par $\Lambda(n)$ la grassmanienne des sous-espaces vectoriels de dimension n, lagrangiens de $(\mathbf{R}^{2n}, w)$.

3.2. Le groupe unitaire $U(n) = Sp(2n) \cap GL(n, \mathbf{C})$ est un sous-groupe compact maximal de $Sp(2n)$:

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in Sp(2n) \cap GL(n, \mathbf{C}) \subset GL(2n, \mathbf{R})$$

si et seulement si $a + ib \in U(n) \subset GL(n, \mathbf{C})$; cf. I.2.0 à 2.5.

Le groupe $U(n)$ agit transitivement, à gauche sur $\Lambda(n)$ (cf. [A_1]), le stabilisateur du plan horizontal $h_0 = \mathbf{R}^n \oplus i\{0\} \in \Lambda(n)$ étant isomorphe à $O(n)$ par l'isomorphisme $a \in O(n) \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in Sp(2n)$. Par suite $\Lambda(n) = U(n)/O(n)$.

On a les suites exactes de fibrations :

$$\begin{array}{ccccc}
SO(n) & \longrightarrow & O(n) & \xrightarrow{\text{dét}} & \{-1, 1\} \\
\downarrow & & \downarrow & & \downarrow \\
SU(n) & \longrightarrow & U(n) & \xrightarrow{\text{dét}} & \mathbf{S}^1 \ni z \\
\downarrow & & \downarrow & & \downarrow \\
SU(n)/SO(n) & \longrightarrow & U(n)/O(n) & \xrightarrow{\text{dét}^2} & \mathbf{S}^1 \ni z^2
\end{array}$$

Pour alléger on écrit dét $= d_M$, le déterminant étant pris dans $GL(n, \mathbf{C})$.

En écrivant les suites exactes d'homotopie pour les fibrations ci-dessus on conclut que d_M induit un isomorphisme sur le groupe fondamental :

$$(d_M^2)_* : \pi_1(U(n)/O(n)) \xrightarrow{\cong} \pi_1(\mathbf{S}^1),$$

ce qui implique l'isomorphisme suivant sur le premier groupe de cohomologie :

$$(d_M^2)^*; H^1(\mathbf{S}^1, \mathbf{Z}) \xrightarrow{\cong} H^1(U(n)/O(n), \mathbf{Z}).$$

On choisit l'orientation sur $\mathbf{S}^1$ obtenue en orientant canoniquement $\mathbf{R}$ et en imposant que $t \mapsto e^{2\pi i t}$ préserve l'orientation, ce qui détermine l'isomorphisme $H^1(\mathbf{S}^1, \mathbf{Z}) \cong \mathbf{Z}$.

3.3. Le revêtement universel $\widetilde{\Lambda(n)}$ de $\Lambda(n)$ est canoniquement isomorphe à $\widetilde{U(n)}/SO(n)$. Plus précisément : soit $\widetilde{U(n)}$ le groupe revêtement universel de $U(n)$, $\widetilde{U(n)}$ est un groupe de Lie et $p : \widetilde{U(n)} \to U(n)$ est un homomorphisme continu.

On a les suites exactes de groupes

$$\{1\} \longrightarrow C \longrightarrow \widetilde{U(n)} \xrightarrow{p} U(n) \longrightarrow \{1\};$$

$$\begin{array}{ccc}
C & \xrightarrow[\cong]{\tilde{d}_M} & \mathbf{Z} \\
\downarrow & & \downarrow \\
\widetilde{U(n)} & \xrightarrow{\tilde{d}_M} & \mathbf{R} \\
{\scriptstyle p}\downarrow & & \downarrow{\scriptstyle e^{2\pi i}} \\
U(n) & \xrightarrow{d_M} & \mathbf{S}^1 \cong \mathbf{R}/\mathbf{Z}
\end{array}$$

avec $C = p^{-1}(\mathbf{Z})$.

On pose $\hat{O}(n) = p^{-1}(O(n))$; on a la suite exacte de groupes

$$\{1\} \longrightarrow SO(n) \longrightarrow \hat{O}(n) \xrightarrow{2\tilde{d}_M} \mathbf{Z} \longrightarrow \{1\},$$

où $SO(n)$ désigne la composante connexe de l'élément neutre du groupe $p^{-1}(SO(n))$; ce groupe est canoniquement isomorphe à $SO(n)$ puisque le revêtement

$$p : p^{-1}(SU(n)) \to SU(n)$$

est trivial.

On a $\Lambda(n) = \widetilde{U(n)}/\hat{O}(n)$ et $C_1 = \hat{O}(n)/SO(n) \xrightarrow[2\tilde{d}_M]{\cong} \mathbf{Z}$ agit à droite sur $\widetilde{U(n)}/SO(n)$ par $uSO(n)\, c = ucSO(n)$, $u \in \widetilde{U(n)}$, $c \in C_1$ (on utilise le fait que le groupe $SO(n)$ est distingué dans $\hat{O}(n)$).

De plus, on a le diagramme commutatif

$$\begin{array}{ccc} C_1 & \xrightarrow{2\tilde{d}_M} & \mathbf{Z} \\ \downarrow & & \downarrow \\ \widetilde{\Lambda(n)} \cong \widetilde{U(n)}/SO(n) & \xrightarrow{2\tilde{d}_M} & \mathbf{R} \\ \downarrow & & \downarrow{\scriptstyle e^{2\pi i}} \\ \Lambda(n) = \widetilde{\Lambda(n)}/C_1 & \xrightarrow{\tilde{d}_M^2} & \mathbf{S}^1 \end{array}$$

et les isomorphismes $\Lambda(n) = \widetilde{\Lambda(n)}/C_1 = \widetilde{U(n)}/\hat{O}(n)$.

3.4. On désigne par Q_q, $q \in \mathbf{Z} \cong C_1$, l'action à droite, sans point fixe, de C_1 sur $\widetilde{\Lambda(n)}$. Les difféomorphismes $Q_q : (x, q) \in \widetilde{\Lambda(n)} \times C_1 \to Q_q(x) \in \widetilde{\Lambda(n)}$ sont les automorphismes du revêtement $p : \widetilde{\Lambda(n)} \to \Lambda(n)$.

Cette action commute avec l'action à gauche de $\widetilde{U(n)}$ sur $\widetilde{U(n)}/SO(n)$:

$$(u, hSO(n)) \in \widetilde{U(n)} \times \widetilde{\Lambda(n)} \mapsto uhSO(n) \in \widetilde{\Lambda(n)}.$$

On pose $2\tilde{d}_M = \delta_M$ (l'indice M rappelant au lecteur le souvenir de Maslov).

Si $u \in \widetilde{U(n)}$ et $h \in \widetilde{\Lambda(n)}$ on a

$$\delta_M(u(h)) = \delta_M(u) + \delta_M(h); \tag{3.5}$$

$$\delta_M(Q_q(h)) = q + \delta_M(h), \quad q \in \mathbf{Z}. \tag{3.6}$$

L'application

$$\widetilde{U(n)} \xrightarrow{\tilde{d}_M} \mathbf{R}$$

est une fibration (triviale) de fibre $SU(n) \cong \ker(\tilde{d}_M)$.

Il suit que

$$\delta_M = 2\tilde{d}_M : \widetilde{\Lambda(n)} \cong \widetilde{U(n)}/SO(n) \to \mathbf{R} \tag{3.7}$$

est une fibration (triviale) de fibre $SU(n)/SO(n)$.

3.8. Le sous-groupe de $Sp(2n)$ laissant invariant le plan lagrangien $h_0 = \mathbf{R}^n \times \{0\}$ est le groupe suivant :

$$G_1 = \left\{ \begin{pmatrix} a & b \\ 0 & {}^t a^{-1} \end{pmatrix} \in Sp(2n),\ a \in GL(n, \mathbf{R}) \right\}.$$

On considère le sous-groupe de R_1

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & {}^t a^{-1} \end{pmatrix} \in R,\ a = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \right.$$

$$\left. \text{est une matrice triangulaire supérieure vérifiant } \lambda_j > 0,\ j = 1, \ldots, n \right\}.$$

3.9. *Lemme 1. — L'application*

$$\Phi : (u, r) \in U(n) \times G \mapsto ur \in Sp(2n)$$

est un difféomorphisme C^ω *(i.e.* **R***-analytique).*

Démonstration. — L'application Φ est **R**-analytique, injective car $U(n) \cap G = \{ 1 \}$. Elle est surjective car, si $A \in Sp(2n)$, puisque $U(n)$ agit transitivement sur $\Lambda(n)$, il existe $u \in U(n)$ tel que $u(A(h_0)) = h_0$, *i.e.* $uA \in G_1 : uA = \begin{pmatrix} a & b \\ 0 & {}^t a^{-1} \end{pmatrix}$. Par la décomposition d'Iwasawa de $GL(n, \mathbf{R})$, il suit que Φ est surjective.

Sur un voisinage V de A dans $Sp(2n)$, puisque $U(n) \to U(n)/O(n)$ est une fibration C^ω, on peut supposer que $u(b)$, pour $B \in V$, dépend **R**-analytiquement de B et comme la décomposition d'Iwasawa de $GL(n, \mathbf{R})$,

$$(u, v) ; O(n) \times \left\{ \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \lambda_j > 0,\ j = 1, \ldots, n \right\} \mapsto uv \in GL(n, \mathbf{R})$$

est un difféomorphisme C^ω, il suit que Φ est une immersion C^ω, surjective. Comme Φ est injective, c'est un difféomorphisme. ■

Remarque. — Φ définit une *décomposition d'Iwasawa* de $Sp(2n)$.

3.10. *Lemme 2. — L'application*

$$(r, u) \in G \times U(n) \mapsto ru \in Sp(2n)$$

est un difféomorphisme **R***-analytique.*

Démonstration. — La démonstration du lemme 1 montre que l'application $(u, r) \mapsto ur^{-1}$ est un difféomorphisme, donc l'application composée

$$(u, r) \mapsto (u^{-1}, r) \mapsto u^{-1} r^{-1} \mapsto (u^{-1} r^{-1})^{-1}$$

est aussi un difféomorphisme. ■

3.11. Soit $\widetilde{Sp}(2n)$ le groupe revêtement universel de $Sp(2n)$; en note

$$\pi : \widetilde{Sp}(2n) \to Sp(2n)$$

la projection canonique. On désigne par $\check{G}$ (resp. $\check{G}_1$) la composante connexe de l'élément neutre du groupe

$$\pi^{-1}(G) \subset \widetilde{Sp}(2n) \quad (\text{resp. } \pi^{-1}(G_1) \subset \widetilde{Sp}(2n)).$$

Il suit du lemme 3.9, par relèvement aux revêtements universels, que l'application

$$(\tilde{u}, r) \in \tilde{U}(n) \times \check{G} \mapsto \tilde{u}r \in \widetilde{Sp}(2n)$$

est un difféomorphisme de classe C^ω ainsi que l'application (lemme 3.10)

$$(\tilde{u}, r) \in \tilde{U}(n) \times \check{G} \mapsto r\tilde{u} \in \widetilde{Sp}(2n).$$

On a les difféomorphismes

$$\widetilde{Sp}(2n)/\check{G} \cong \tilde{U}(n),$$
$$\widetilde{Sp}(2n)/\check{G}_1 \cong \tilde{U}(n)/SO(n),$$
$$Sp(2n)/G_1 \cong \Lambda(n),$$

et $\widetilde{Sp}(2n)$ (resp. $Sp(2n)$) agit canoniquement à gauche sur $\widetilde{\Lambda(n)} \cong \widetilde{Sp}(2n)/\check{G}_1$ (resp. $\Lambda(n) \cong Sp(2n)/G_1$).

Si $A \in \widetilde{Sp}(2n)$ et $q \in \mathbf{Z}$ on a

$$\tilde{A} Q_q \tilde{A}^{-1} = Q_q,$$

où les Q_q sont les transformations du revêtement $C_1 \to \widetilde{\Lambda(n)} \to \Lambda(n)$, $C_1 \cong \mathbf{Z}$.

(En effet, sur $\Lambda(n)$, Q_q induit l'identité. On peut trouver un chemin continu $s \in [0, 1] \mapsto \tilde{A}_s \in \widetilde{Sp}(2n)$ vérifiant $\tilde{A}_0 = 1$, élément neutre de $\widetilde{Sp}(2n)$, et $\tilde{A}_1 = \tilde{A}$; c'est possible puisque $Sp(2n)$ est connexe (cf. 3.9). On a $\tilde{A}_s Q_q \tilde{A}_s^{-1} = Q_{q(s)}$, où $s \mapsto q(s) \in \mathbf{Z}$ est continu et $q(0) = q$.)

3.12. Soit $h_1 = \{0\} \times \mathbf{R}^n = Jh_0$. On définit pour $j = 0, 1$ les ouverts

$$U_j = \{ h \in \Lambda(n),\ h \text{ est un graphe sur } h_j \}$$

qui sont *denses* dans $\Lambda(n)$ ($\Lambda(n) - U_j$ est un ensemble $\mathbf{R}$-analytique non vide de $\Lambda(n)$).

Pour $j = 0, 1$, U_j est *contractile* puisque

$$h \in U_0 \Leftrightarrow h = \begin{pmatrix} 1 \\ m \end{pmatrix} = \{ (x, m(x)) \in \mathbf{R}^{2n},\ x \in \mathbf{R}^n \},$$

$$h \in U_1 \Leftrightarrow h = \begin{pmatrix} m \\ 1 \end{pmatrix} = \{ (m(x), x) \in \mathbf{R}^{2n},\ x \in \mathbf{R}^n \},$$

où $m \in M_n(\mathbf{R})$ est une matrice symétrique, *i.e.* ${}^t m = m$.

Dans la suite on utilisera les « *coordonnées projectives* » $\begin{pmatrix} 1 \\ m \end{pmatrix}$ et $\begin{pmatrix} m \\ 1 \end{pmatrix}$.

3.13. On désigne par $\widetilde{h}_0$ le relèvement de $h_0 \in \Lambda(n)$ à $\widetilde{\Lambda}(n)$ qui vérifie $\delta_M(\widetilde{h}_0) = 0$. Soit $\check{U}_0$ la composante connexe de $p^{-1}(U_0) \subset \widetilde{\Lambda}(n)$ contenant h_0 (le revêtement p : $\widetilde{\Lambda}(n) \to \Lambda(n)$ est trivial sur U_0).

3.14. *Proposition.* — *Soit* $h = \begin{pmatrix} 1 \\ m \end{pmatrix} \in U_0$ *et* $\widetilde{h} \in \check{U}_0$ *tel que* $p(\widetilde{h}) = h$; *alors*

$$\delta_M(\widetilde{h}) = \frac{1}{\pi} \sum_{j=1}^{n} \operatorname{Arctg} \lambda_j,$$

où λ_j *désigne les valeurs propres de la matrice symétrique* m.

Démonstration. — Soit $(v_1, \ldots, v_n)$ une base orthonormée de $\mathbf{R}^n$ de vecteurs propres de la matrice m (sur $\mathbf{R}^n$ on met la métrique $\| x \|^2 = \sum_1^n x_j^2$, $x = (x_1, \ldots, x_n)$). Pour $0 \leqslant s \leqslant 1$, on pose $h_s = \begin{pmatrix} 1 \\ sm \end{pmatrix}$. On détermine $u_s \in U(n)$ par les conditions

$$u_s(v_j) = e^{i \operatorname{Arctg} s\lambda_j} v_j, \quad j = 1, \ldots, n,$$
$$u_s J = J u_s,$$

et donc $u_s(h_0) = h_s$.

Il suit de la définition de u_s que

$$d_M^2(u_s) = e^{2\pi i \left(\frac{1}{\pi} \sum_1^n \operatorname{Arctg} s\lambda_j\right)}.$$

On peut relever le chemin $s \mapsto h_s \in U_0$ en un unique chemin $s \mapsto \widetilde{h}_s \in \check{U}_0$ vérifiant $\widetilde{h}_{s|s=0} = \widetilde{h}_0$, $\widetilde{h}_1 = h$. On peut aussi relever de façon unique le chemin

$$s \mapsto d_M^2(u_s) = d_M^2(h_s) \in \mathbf{S}^1$$

en

$$s \mapsto \varphi(s) = \frac{1}{\pi} \sum_1^n \operatorname{Arctg}(s\lambda_j) \in \mathbf{R},$$

l'unicité résultant de $\varphi(0) = 0$.

Comme $\delta_M(\widetilde{h}_0) = 0$, il résulte de l'unicité que $\delta_M(\widetilde{h}_s) = \varphi(s)$, et la proposition s'ensuit en faisant $s = 1$. ■

3.15. Soient, pour $0 \leqslant s \leqslant 1$, les matrices diagonales

$$w_s = \operatorname{diag}(e^{-\frac{i\pi}{2}s}, \ldots, e^{-\frac{i\pi}{2}s}) \in U(n).$$

On relève le chemin $s \mapsto w_s \in U(n)$ en $s \mapsto \widetilde{w}_s \in \widetilde{U}(n)$, de façon unique si on impose $\widetilde{w}_0 = 1 \in \widetilde{U}(n)$, 1 désignant l'élément neutre de $\widetilde{U}(n)$. On pose $\widetilde{w}_1 = \widetilde{J}^{-1}$; $\widetilde{J}^{-1}$ est un relèvement de J^{-1} à $\widetilde{U}(n)$.

3.16. On désigne par $\check{U}_1$ la composante connexe de $p^{-1}(U_1) \subset \widetilde{\Lambda(n)}$ vérifiant $\widetilde{J}^{-1}(\check{U}_1) = \check{U}_0$.

3.17. *Proposition.* — *Soit* $\tilde{h} \in \check{U}_1$ *tel que* $p(\tilde{h}) = \begin{pmatrix} m \\ 1 \end{pmatrix}$; *alors*

$$\delta_M(\tilde{h}) = \frac{n}{2} - \frac{1}{\pi} \sum_1^n \operatorname{Arctg} \lambda_j,$$

où les λ_j *désignent les valeurs propres de la matrice symétrique* $m \in M_n(\mathbf{R})$.

Démonstration. — On a

$$J^{-1} \begin{pmatrix} m \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -m \end{pmatrix} \in U_0, \quad J^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

D'après 3.5 et 3.6,

$$\delta_M(\tilde{J}^{-1}(\tilde{h})) = \delta_M(\tilde{J}^{-1}) + \delta_M(\tilde{h}),$$

et la proposition suit de 3.14 et du fait que

$$\delta_M(\tilde{J}^{-1}) = -\frac{n}{2}. \quad \blacksquare$$

3.18. Soient $r \in G \subset Sp(2n)$ et $\tilde{r}$ son unique relèvement à $\check{G} \subset \widetilde{Sp}(2n)$. On a

$$r(U_1) = U_1,$$
$$r(h_0) = h_0,$$

et

$$\tilde{r}(\tilde{h}_0) = \tilde{h}_0,$$

puisque l'application

$$r \in \check{G} \mapsto r(h_0) \in \{ Q_q(\tilde{h}_0), q \in \mathbf{Z} \}$$

est continue et $\check{G}$ est connexe et même contractile.

3.19. *Lemme 3.* — *L'action à gauche de* $\tilde{r}$ *sur* $\widetilde{\Lambda(n)}$ *laisse invariante chaque composante connexe de* $p^{-1}(U_1)$.

Démonstration. — Soit W une composante connexe de $p^{-1}(U_1)$. On a $\tilde{r}(W) = Q_q(W)$ pour un $q(\tilde{r}) \in \mathbf{Z} \cong C_1$. Comme $\check{G}$ est connexe et que $1 \in \check{G}$ induit l'identité sur $\tilde{\Lambda}(n)$, le lemme suit. $\blacksquare$

3.20. *Corollaire.* — *Pour tout* $\tilde{r} \in \check{G}$ *et* $\tilde{h} \in \tilde{\Lambda}(n)$ *on a*

$$| \delta_M(\tilde{r}(\tilde{h})) - \delta_M(\tilde{h}) | \leqslant n.$$

Démonstration. — Par le lemme 3.19 et 3.6 on peut supposer $\tilde{h} \in \check{U}_1$ et le corollaire suit de la proposition 3.17. $\blacksquare$

Pour une autre démonstration (plus générale) de la proposition suivante, nous renvoyons le lecteur à [J_1].

3.21. *Proposition.* — *Pour tout* $A \in \widetilde{Sp}(2n)$ *on a*

$$\operatorname*{Osc}_{h \in \widetilde{\Lambda}(n)} (\delta_M(A(h)) - \delta_M(h)) \leqslant n$$

où Osc *désigne l'oscillation, ou encore, si* $\varphi : \widetilde{\Lambda(n)} \to \mathbf{R}$ *est une fonction,*

$$\operatorname*{Osc}_{h \in \widetilde{\Lambda(n)}} (\varphi(h)) = \operatorname*{Sup}_{h \in \widetilde{\Lambda(n)}} \varphi(h) - \operatorname*{Inf}_{h \in \widetilde{\Lambda(n)}} \varphi(h).$$

Démonstration. — On écrit de façon unique (cf. 3.9)

$$A = \tilde{u}_A r_A, \quad \tilde{u}_A \in \widetilde{U}(n) \text{ et } r_A \in \check{G}.$$

Si $\tilde{u}_1 \in \widetilde{U}(n)$, on peut aussi décomposer de façon unique (3.10)

$$r_A \tilde{u}_1 = \tilde{u}_2 r_2, \quad (\tilde{u}_2, r_2) \in \widetilde{U}(n) \times \check{G}.$$

Comme $r_2(\tilde{h}_0) = \tilde{h}_0$ on a, d'après 3.5,

$$\delta_M(\tilde{u}_2) = \delta_M(r_A(h))$$

où $$h = \tilde{u}_1(\tilde{h}_0).$$

On a aussi $\delta_M(h) = \delta_M(\tilde{u}_1)$, d'où

$$\delta_M(\tilde{u}_2 \tilde{u}_1^{-1}) = \delta_M(r_A(h)) - \delta_M(h)$$

et donc, d'après le corollaire 3.20, pour tout $\tilde{u}_1$,

$$|\delta_M(\tilde{u}_2 \tilde{u}_1^{-1})| \leqslant n. \tag{3.22}$$

On choisit $\tilde{u}_1 \in \widetilde{U}(n)$ tel que $\tilde{u}_1(\tilde{h}_0) = h$. On a

$$\begin{aligned} \delta_M(A\tilde{u}_1(\tilde{h}_0)) &= \delta_M(\tilde{u}_A) + \delta_M(r_A \tilde{u}_1(\tilde{h}_0)) \\ &= \delta_M(\tilde{u}_A) + \delta_M(\tilde{u}_2), \end{aligned}$$

d'où $$\delta_M(A(h)) - \delta_M(h) = \delta_M(\tilde{u}_A) + \delta_M(\tilde{u}_2 \tilde{u}_1^{-1}),$$

et la proposition découle de 3.22. ■

Remarque. — La proposition précédente, si $n = 1$, est un cas particulier du fait suivant : si $\tilde{f} : \mathbf{R} \to \mathbf{R}$ est un homéomorphisme vérifiant $\tilde{f}(x + 1) = 1 + \tilde{f}(x)$, $x \in \mathbf{R}$, alors $\operatorname*{Osc}_{\mathbf{R}} (\tilde{f} - \mathrm{Id}) \leqslant 1$.

3.23. On se donne

$$B = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix} \in R$$

vérifiant $\ell > 0$ ($\ell \in M_n(\mathbf{R})$, ${}^t\ell = \ell$), et l'on note $\tilde{B}$ le relèvement de $B \in G$ à $\check{G}$. On a $\tilde{B}(\tilde{h}_0) = \tilde{h}_0$.

Pratiquement toute la démonstration repose sur la proposition suivante.

3.24. *Proposition.* — *Si* $h \in \widetilde{\Lambda}(n)$, *on a* $\delta_M(\widetilde{B}(h)) \leqslant \delta_M(h)$ *et l'égalité* $\delta_M(\widetilde{B}(h)) = \delta_M(h)$ *implique* $h = Q_k(\widetilde{h}_0)$ *pour un* $k \in \mathbf{Z}$.

Démonstration. — Soit $\check{U}_1$ la composante de $p^{-1}(U_1)$ définie en 3.16. Soit $h \in \check{U}_1$, $p(h) = \begin{pmatrix} m \\ 1 \end{pmatrix}$. On a

$$B\begin{pmatrix} m \\ 1 \end{pmatrix} = \begin{pmatrix} m+\ell \\ 1 \end{pmatrix},$$

d'où $B(U_1) = U_1$, ce qui implique, par 3.19, $\widetilde{B}(\check{U}_1) = \check{U}_1$.

On note

$$\lambda_1(m+\ell) \leqslant \lambda_2(m+\ell) \leqslant \ldots \leqslant \lambda_n(m+\ell)$$

les valeurs propres de la matrice $m + \ell$.

D'après la définition variationnelle des vecteurs propres d'une matrice symétrique [H_5, p. 181],

$$\lambda_j(m+\ell) \geqslant \lambda_j(m) + \lambda_1(\ell)$$

où $$\lambda_1(\ell) = \operatorname*{Inf}_{\|v\|=1} {}^t v \ell v \quad \text{et} \quad \lambda_1(\ell) > 0$$

car $\ell > 0$.

Par la proposition 3.17,

$$\begin{aligned} \delta_M(\widetilde{B}(h)) &= \frac{n}{2} - \frac{1}{\pi}\sum_1^n \operatorname{Arctg}(\lambda_j(m+\ell)) \\ &\leqslant \frac{n}{2} + \frac{1}{\pi}\sum_1^n \operatorname{Arctg}(-\lambda_j(m) - \lambda_1(\ell)) \\ &< \delta_M(h). \end{aligned} \tag{3.25}$$

Comme d'une part $\delta_M(\widetilde{B}(Q_q(h))) = q + \delta_M(\widetilde{B}(h))$ et d'autre part $p^{-1}(U_1)$ est dense dans $\Lambda(n)$ et que la fonction $h \in \widetilde{\Lambda}(n) \to \delta_M(\widetilde{B}(h)) \in \mathbf{R}$ est continue, on en déduit

$$\delta_M(\widetilde{B}(h)) \leqslant \delta_M(h) \quad \text{pour tout } h \in \widetilde{\Lambda}(n).$$

On suppose maintenant que $h \notin \{ Q_q(\widetilde{h}_0), q \in \mathbf{Z} \}$. On peut approcher $\widetilde{h}$ par une suite $(y_i)_{i \in \mathbf{N}}$ incluse dans une composante connexe de $p^{-1}(U_1)$, $p(y_i) = \begin{pmatrix} m_i \\ 1 \end{pmatrix}$.

Comme $h \neq Q_q(\widetilde{h}_0)$ pour tout $q \in \mathbf{Z}$, quand $i \to +\infty$, les matrices ont une valeur propre $\lambda_{j_i}(m_i)$ qui reste finie (*i.e.* ne tend pas vers $+\infty$ ou $-\infty$). En effet, dans le cas contraire on aurait, pour une sous-suite (y_{i_k}) de (y_i), $p(y_{i_k}) = \begin{pmatrix} 1 \\ m_{i_k}^{-1} \end{pmatrix}$, $m_{i_k}^{-1} \to 0$ si $i_k \to +\infty$, et par conséquent $p(h) = h_0$, ce qui contredirait $p(h) \neq h_0$.

D'après 3.25, on peut trouver $\varepsilon > 0$ et $i_0 > 0$ tels que, si $i \geqslant i_0$,

$$\delta_M(\widetilde{B}(y_i)) \leqslant \delta_M(y_i) - \varepsilon;$$

en faisant tendre i vers $+\infty$, on en conclut que

$$\delta_M(\widetilde{B}(h)) \leqslant \delta_M(h) - \varepsilon < \delta_M(h).$$

Comme $\delta_M(\widetilde{B}(Q_q(\widetilde{h}_0))) = \delta_M(\widetilde{h}_0) + q$, nous avons démontré la proposition. ■

3.26. *Corollaire. — On a*

$$\delta_M(h) \leqslant \delta_M(\widetilde{B}^{-1}(h)),$$

et l'égalité

$$\delta_M(h) = \delta_M(\widetilde{B}^{-1}(h))$$

implique $h = Q_q(\widetilde{h}_0)$ *pour un* $q \in \mathbf{Z}$.

Remarques :

(i) Sur $X = \widetilde{\Lambda(n)} - \{ Q_q(\widetilde{h}_0), q \in \mathbf{Z} \}$, la fonction δ_M est une fonction de Liapounoff pour $\widetilde{B}$ agissant sur X. C'est le point essentiel pour toute la suite, et tout particulièrement pour démontrer les propositions 7.2 et 7.5.

(ii) Si $B = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}$, $\ell \in M_n(\mathbf{R})$, ${}^t\ell = \ell$, $\det \ell \neq 0$ et la forme quadratique $v \in \mathbf{R}^n \mapsto {}^t v \ell v$ est indéfinie, alors B laisse invariants des plans lagrangiens $\begin{pmatrix} 1 \\ m \end{pmatrix}$, graphes sur h_0, où $m = {}^t m$, $m \neq 0$ (il suffit de considérer $m = \varepsilon P$ où $\varepsilon > 0$ et P est projecteur orthogonal sur un vecteur $v \neq 0$ isotrope pour ℓ).

4. Détermination de la proximité de F à L

4.1. On se donne L, un difféomorphisme de classe C^1 de $\mathbf{A}^n$ complètement intégrable monotone positif

$$L(\theta, r) = (\theta + \ell(r), r).$$

On a $D\ell(r) > 0$, $L(\mathbf{A}^n_{R_1}) = \mathbf{A}^n_{R_1}$, $R_1 > 0$. L'application tangente de L, $TL = L \times DL$

$$DL(\theta, r) = \begin{pmatrix} 1 & D\ell(r) \\ 0 & 1 \end{pmatrix}$$

agit sur le figré tangent $T(\mathbf{A}^n) = \mathbf{A}^n \times (\mathbf{R}^{2n}, w)$ et donc sur $\mathbf{A}^n \times \Lambda(n)$, le fibré $\mathbf{A}^n_{R_1} \times \Lambda(n)$ étant invariant par TL.

On peut relever de façon unique l'application continue

$$x \in \mathbf{A}^n \mapsto DL(x) \in G \subset Sp(2n)$$

en une application continue

$$x \mapsto \widetilde{DL}(x) \in \check{G} \subset Sp(2n).$$

On pose $\widetilde{\mathrm{TL}} = \mathrm{L} \times \widetilde{\mathrm{DL}}$. L'application $\widetilde{\mathrm{TL}}$ agit sur $\mathbf{A}^n_{\mathrm{R}_1} \times \widetilde{\Lambda(n)}$ et on a (cf. 3.13 et 3.18)

(4.2) $$\widetilde{\mathrm{DL}}(x)(\widetilde{h}_0) = \widetilde{h}_0, \quad \text{pour tout } x \in \mathbf{A}^n.$$

On fixe $\mathrm{R}_1 > \mathrm{R} > 0$.

4.3. *Proposition.* — *Il existe* $\eta > 0$ *et un ouvert connexe* V *de* $\check{\mathrm{U}}_0$ *relativement compact dans* $\check{\mathrm{U}}_0$, *avec* $\widetilde{h}_0 \in \mathrm{V} \subset \check{\mathrm{U}}_0 \subset \widetilde{\Lambda}(n)$, *tel que si* $h \notin \mathrm{Q}_k(\mathrm{V})$ *pour tout* $k \in \mathbf{Z}$, *alors pour tout* $x \in \mathbf{A}^n_{\mathrm{R}_1}$

(4.4) $$\delta_{\mathrm{M}}(\widetilde{\mathrm{DL}}(x)h) < \delta_{\mathrm{M}}(h) - \eta.$$

De plus, le diamètre de V *tend vers* 0 *lorsque* η *tend vers* 0.

Démonstration. — Cela suit immédiatement de 4.2 et 3.24 en utilisant la continuité de $\widetilde{\mathrm{TL}}$, la compacité de $\mathbf{A}^n_{\mathrm{R}_1}$ et le fait qu'il suffit de vérifier l'assertion sur un compact qui est l'adhérence d'un domaine fondamental de l'action de $(\mathrm{Q}_q)_{q \in \mathbf{Z}}$ (cf. 3.6 et 3.7). ■

4.5. *Proposition.* — *Il existe* $\eta_1 > 0$, $s > 0$ *et un ouvert* W, *connexe, relativement compact dans* V, $\mathrm{W} \subset \mathrm{V} \subset \widetilde{\Lambda}(n)$, $\widetilde{h}_0 \in \mathrm{W}$, *tels que l'on ait*

(4.6) $$\delta_{\mathrm{M}}(\mathrm{W}) =]-s, s[, \quad s < \eta,$$

(4.7) $$\widetilde{\mathrm{DL}}(x)(\overline{\mathrm{W}}) \subset \mathrm{V} \quad \textit{pour tout } x \in \mathbf{A}^n_{\mathrm{R}_1},$$

et si $h \notin \{\mathrm{Q}_q(\mathrm{W}), q \in \mathbf{Z}\}$,

(4.8) $$\delta_{\mathrm{M}}(\widetilde{\mathrm{DL}}(x)(h)) < \delta_{\mathrm{M}}(h) - \eta_1.$$

Démonstration. — Par la continuité de $\widetilde{\mathrm{TL}}$ et la compacité de $\mathbf{A}^n_{\mathrm{R}_1}$, 4.7 suit de 4.2 et 4.8 résulte de la proposition 3.24 par le même argument que dans la démonstration de la proposition précédente. Quitte à restreindre W, les conditions 4.6 sont satisfaites (en utilisant le fait que la fonction δ_{M} est ouverte). ■

4.9. On peut toujours supposer, quitte à restreindre W, que s vérifie

$$2s < 1$$

et

$$2s < \eta.$$

Remarque. — Lorsque $n \geqslant 2$, pour tout $\varepsilon > 0$, $\delta_{\mathrm{M}}^{-1}(]-\varepsilon, \varepsilon[)$ n'est pas inclus dans $\check{\mathrm{U}}_0$.

Pour démontrer le théorème, et tout particulièrement pour démontrer les propositions 7.2 et 7.5, on est amené à introduire les ouverts connexes V et W que nous avons déterminés par la proposition 4.5. Le fait d'avoir deux ouverts V et W alourdit la suite.

Malheureusement la proposition 7.2 ne nous suffirait pas. En effet, pour tout $\varepsilon > 0$, on peut trouver un tore $T_\varepsilon \subset \mathbf{A}^2$ de classe C^∞, lagrangien, vérifiant :

$$\operatorname*{Sup}_{x \in T_\varepsilon} | \det{}^2(\tau_{T_\varepsilon}(x)) - 1 | < \varepsilon \quad \text{où } \tau_{T_\varepsilon} \text{ est définie en } 2.1,$$

T_ε n'est pas un graphe sur $\mathbf{T}^2 \times \{ 0 \} \subset \mathbf{A}^2$,
T_ε tend en topologie de Hausdorff vers $\mathbf{T}^2 \times \{ 0 \}$, lorsque $\varepsilon \to 0$.

Nous indiquons seulement l'idée de la construction des tores T_ε : on construit $\hat{\varphi}_\varepsilon \in C^\infty(\mathbf{T}^2, \mathbf{R})$ vérifiant, pour ε tendant vers 0,

(i) si $\varphi_\varepsilon = \dfrac{\partial \hat{\varphi}_\varepsilon}{\partial \theta}$, $\| \varphi_\varepsilon \|_{L^\infty} \to 0$,

(ii) $\| \Delta \hat{\varphi}_\varepsilon \|_{L^\infty} \to 0$, où $\Delta = \left(\dfrac{\partial}{\partial \theta_1}\right)^2 + \left(\dfrac{\partial}{\partial \theta_2}\right)^2$ est le laplacien,

(iii) $\| D\varphi_\varepsilon \|_{L^\infty} \to +\infty$.

On obtient la famille $\hat{\varphi}_\varepsilon$ en modifiant et lissant, au voisinage de 0, la fonction $\varepsilon' \theta_1 \theta_2 \operatorname{Log}(\theta_1^2 + \theta_2^2)$, où $(\theta_1, \theta_2) \in \mathbf{R}^2 - \{ 0 \}$ et $\varepsilon' > 0$.

Les tores lagrangiens $\hat{T}_\varepsilon = \Gamma_{\varphi_\varepsilon}$ vérifient, si $\varepsilon \to 0$,

$$\operatorname*{Sup}_{x \in \hat{T}_\varepsilon} | \det{}^2(\tau_{\hat{T}_\varepsilon}(x)) - 1 | \to 0.$$

Cela résulte de (ii), 3.14 et de l'inégalité $| \operatorname{Arctg} \lambda_1 + \operatorname{Arctg} \lambda_2 | \leqslant | \lambda_1 + \lambda_2 |$.

(iii) implique que, si $\varepsilon \to 0$, le plan tangent de $\hat{T}_\varepsilon$ en un point x_ε contient presque un vecteur du plan vertical $\{ 0 \} \times \mathbf{R}^2 \subset T_{x_\varepsilon}(\mathbf{A}^2)$ et en considérant des difféomorphismes bien choisis G_ε tendant vers Id pour $\varepsilon \to 0$, on peut prendre $T_\varepsilon = G_\varepsilon(\hat{T}_\varepsilon)$.

4.10. On se donne F, un difféomorphisme symplectique de $\mathbf{A}^n$. On cherche $\varepsilon > 0$ et $\varepsilon' > 0$ tels que si

(4.11) $$\operatorname*{Sup}_{x \in \mathbf{A}^n_{R_1}} \| DF(x) - DL(x) \| < \varepsilon \quad (\| \ \| \text{ est la norme définie en I.2})$$

et

(4.12) $$\operatorname*{Sup}_{x \in \mathbf{A}^n_{R_1}} (d_0(F^{-1}(x), L^{-1}(x)) + d_0(F(x), L(x))) < \varepsilon',$$

où d_0 désigne la métrique sur $\mathbf{R}^n/\mathbf{Z}^n \times \mathbf{R}^n$ induite par passage au quotient de la métrique $\| x - y \|_\infty$ sur $\mathbf{R}^{2n}$, où $\|(z_j)\|_\infty = \operatorname*{Sup}_{1 \leqslant j \leqslant 2n} | z_j |$, alors on ait les conclusions du théorème 2.2.

4.13. Comme le groupe des homéomorphismes de $\mathbf{A}^n$ avec la topologie compacte ouverte est un groupe topologique, on en déduit :

4.14. Si $0 < \varepsilon' \leqslant \varepsilon'_1$, on peut supposer que

$$F^j(\mathbf{A}^n_R) \subset \mathbf{A}^n_{R_1} \quad \text{pour} \quad -1 \leqslant j \leqslant 2n.$$

4.15. Si $\varepsilon < \varepsilon_1$ alors on peut supposer que l'application

$$x \in \mathbf{A}^n_{R_1} \mapsto DF(x) \in Sp(2n) \tag{4.16}$$

est homotope à $x \mapsto DL(x)$ et on peut relever l'application $x \mapsto DF(x)$ en une application continue $x \in \mathbf{A}^n_{R_1} \mapsto \widetilde{DF}(x) \in \widetilde{Sp}(2n)$ vérifiant

$$\operatorname*{Sup}_{x \in \mathbf{A}^1_{R_1}} d_{Sp}(\widetilde{DF}(x), \widetilde{DL}(x)) < \varepsilon_2, \tag{4.17}$$

où d_{Sp} désigne une métrique sur $\widetilde{Sp}(2n)$ définissant sa topologie.

4.18. On pose $\widetilde{TF} = F \times \widetilde{DF}$ et

$$\widetilde{TF}^{-1} = (F^{-1}, \widetilde{DF}^{-1}) \equiv (F^{-1}, (\widetilde{DF})^{-1} \circ F^{-1}).$$

4.19. On peut trouver $\varepsilon_2 > 0$, tel que si 4.11 est vérifiée, alors, pour tout $x \in \mathbf{A}^n_{R_1}$, on a les conclusions suivantes :

4.20. Si $h \in \widetilde{\Lambda(n)}$, $h \notin \{ Q_k(V), k \in \mathbf{Z} \}$, alors

$$\delta_M(\widetilde{DF}(x)h) < \delta_M(h) - \eta.$$

$$\widetilde{DF}(x)(\overline{W}) \subset V. \tag{4.21}$$

4.22. Si $h \in \widetilde{\Lambda(n)}$ vérifie $h \notin \{ Q_k(W), k \in \mathbf{Z} \}$, alors

$$\delta_M(\widetilde{DF}(x)h) < \delta_M(h) - \eta_1.$$

Toutes ces conditions sont ouvertes pour la topologie de la convergence uniforme sur $C^0(\mathbf{A}^n_{R_1}, Sp(2n))$ et résultent de 4.6 et 4.8.

4.23. Comme $Q_q(\check{U}_0) \cap \check{U}_0 = \emptyset$ si $q \neq 0$ et $\overline{V} \subset \check{U}_0$ est compact, si F vérifie 4.17 avec $\varepsilon_2 > 0$ petit, alors

$$\widetilde{DF}(x)(\overline{V}) \cap \{ Q_q(V), q \in \mathbf{Z}^* \} = \emptyset. \tag{4.24}$$

Finalement si 4.11 est vérifiée avec $\varepsilon < \varepsilon_1$, avec ε_1 assez petit, alors toutes les conditions 4.20 à 4.24 sont satisfaites.

On a besoin de la relation 4.25 ci-dessous.

Comme $\widetilde{DL}^k(x)(\tilde{h}_0) = \tilde{h}_0$ pour tout $x \in \mathbf{A}^n$ et $k \in \mathbf{Z}$, $\delta_M(\widetilde{DL}^k(x)(\tilde{h}_0)) = 0$, quitte à supposer ε_1 et ε_1' assez petits, on a, pour tout $x \in \mathbf{A}^n_R$,

$$|\delta_M(\widetilde{DF}^{2n}(x)(\tilde{h}_0))| = |\delta_M(\widetilde{DF}^{2n}(x)(\tilde{h}_0)) - \delta_M(\widetilde{DL}^{2n}(x)(\tilde{h}_0))| < (1/2) \tag{4.25}$$

où $\widetilde{DF}^j$ est définie par $\widetilde{DF}^j = (F^j, \widetilde{DF}^j)$, $j \in \mathbf{Z}$ et donc $\widetilde{DF}^j = (\widetilde{DF} \circ F^{j-1}) \ldots (\widetilde{DF})$ si $j \geqslant 1$.

4.26. C'est seulement ici qu'on suppose 4.12 vérifiée (pour avoir 4.14); la condition 4.25 est une condition ouverte pour la topologie de la convergence uniforme sur $x \in \mathbf{A}^n_{R_1} \to \widetilde{DF}(x)$.

4.27. Dans la suite, on supposera que $\varepsilon > 0$, $\varepsilon' > 0$ sont choisis, $0 < \varepsilon \leqslant \varepsilon(L, R, R_1)$, $0 < \varepsilon' \leqslant \varepsilon'(L, R, R_1)$, où $\varepsilon(L, R, R_1)$ et $\varepsilon'(L, R, R_1)$ ne dépendent que de L, R et R_1, pour avoir 4.20 à 4.25.

4.28. *Remarques.*

1. Si $L(\theta, r) = (\theta + \ell(r), r)$ vérifie

$$\| D\ell \|_{C^0(\mathbf{R}^n)} < +\infty$$

et

$$\| (D\ell)^{-1} \|_{C^0(\mathbf{R}^n)} < +\infty,$$

on n'a pas besoin de fixer $R > 0$, quitte à supposer que les applications

$$x \in \mathbf{A}^n \mapsto DL(x) \in M_n(\mathbf{R}) \quad \text{et} \quad x \in \mathbf{A}^n \mapsto DF(x)$$

sont uniformément continues sur $\mathbf{A}^n$ (on munit $\mathbf{A}^n$ de la métrique définie en 4.12 et $M_n(\mathbf{R})$ de la norme $\| \ \|$) et que le difféomorphisme F de $\mathbf{A}^n$ est de classe C^1, symplectique, et satisfait à

$$\| DL - DF \|_{C^0(\mathbf{A}^n)} < \varepsilon$$

ainsi que

$$\| DL^{-1} - DF^{-1} \|_{C^0(\mathbf{A}^n)} < \varepsilon;$$

si $\varepsilon > 0$ est assez petit, on a pour tout $x \in \mathbf{A}^n$ les conclusions 4.20 à 4.25, où le nombre ε ne dépend que de L et de la dimension n.

2. Toutes les conditions 4.20 à 4.24 sont de nature géométrique et l'on n'utilise le nombre $\varepsilon_1 > 0$ que pour donner une condition suffisante pour qu'elles soient vérifiées. *A priori* le nombre $\varepsilon_1 > 0$, qui permet d'avoir 4.20 à 4.24, ne dépend que de L et de $R_1 > 0$, mais il dépend aussi de la dimension n puisque la fonction δ_M en dépend.

3. La condition 4.25 dépend explicitement de la dimension de n et elle semble être la condition la plus restrictive. *A priori* elle impose que $\varepsilon_1 = 0(n^{-2})$ quand $n \to +\infty$. Cette condition permet d'avoir 5.8 et n'est utilisée seulement qu'à ce moment. Voir à ce propos 5.9 et 10.5.

4. Ainsi que nous l'avons signalé en 4.26, les conditions

$$F^j(\mathbf{A}^n_R) \subset \mathbf{A}^n_{R_1}, \quad 0 \leqslant j \leqslant 2n,$$

sont seulement utilisées pour obtenir 4.25.

On peut supprimer la dépendance de $\varepsilon(L, R, R_1)$ sur R de la façon suivante : il suffira d'imposer que l'inégalité 4.25 soit vérifiée pour tout $x \in \bigcap_{j=0}^{2n} F^j(\mathbf{A}^n_{R_1})$. En effet,

4.25 sera seulement utilisé en 5.8 en des points $x \in N \subset \mathbf{A}^n_{R_1}$ où N est invariant par $F : F(N) = N$.

La condition $F^{-1}(\mathbf{A}^n_R) \subset \mathbf{A}^n_{R_1}$ sera seulement utilisée en 4.29.

4.29. Quitte à diminuer les nombres $\varepsilon(L, R, R_1)$ et $\varepsilon'(L, R, R_1)$ que nous avons déterminés en 4.27, on peut trouver $\eta_2 > 0$ tel que, pour tout $x \in \mathbf{A}^n_R$ et tout $h \in \widetilde{\Lambda(n)}$, $h \notin \{ Q_q(W), q \in \mathbf{Z} \}$, on ait

(4.30) $$\delta_M(DF^{-1}(x)h) > \delta_M(h) + \eta_2.$$

Cela suit du fait que

$$\delta_M((D\ell)^{-1}(x)h) > \delta_M(h) + \eta_2, \quad x \in \mathbf{A}^n_{R_1},$$

pour un $\eta_2 > 0$ ($\eta_2 > 0$ résulte de 3.26).

5. Ensembles N invariants par F

5.1. On suppose que $N \subset \mathbf{A}^n_R$ est *compact connexe* (non vide) *invariant* par F. On note F_N la restriction de F à N.

5.2. On se donne $\xi \to N$ un fibré vectoriel continu qui est un sous-fibré lagrangien de $T(\mathbf{A}^n)_{|N}$: pour tout $x \in N$, la fibre ξ_x de ξ en x est un sous-espace lagrangien de $(\mathbf{R}^{2n}, w)$. On note $\bar{\sigma}$ l'application

$$x \in N \mapsto \bar{\sigma}(x) = [\xi_x] \in \Lambda(n).$$

On fait *l'hypothèse* que l'on peut *relever* l'application $\bar{\sigma}$

$$\begin{tikzcd} & \widetilde{\Lambda(n)} \arrow[d, "p"] \\ N \arrow[ur, "\sigma"] \arrow[r, "\bar{\sigma}"'] & \Lambda(n) \end{tikzcd}$$

en une application continue $\sigma : N \to \widetilde{\Lambda(n)}$ vérifiant $p \circ \sigma = \bar{\sigma}$.

C'est le cas si et seulement si l'application

(5.3) $$d^2_M \circ \bar{\sigma} : N \to \mathbf{S}^1 \text{ est homotope à une constante.}$$

On dit, si 5.3 est vérifiée, que $\mathrm{Maslov}(\xi) = 0$.

Si N est homéomorphe à un cw complexe (ce qui est le cas si N est une variété de classe C^1), alors 5.3 équivaut à

(5.3′) $$0 = \mathrm{Maslov}(\xi) = \bar{\sigma}^*(1) \in H^1(N, \mathbf{Z}),$$

où $\bar{\sigma}^* : H^1(\Lambda(n), \mathbf{Z}) \cong \mathbf{Z} \to H^1(N, \mathbf{Z})$ désigne l'application induite en cohomologie par $\bar{\sigma}$ (on remarque que $\mathbf{S}^1$ est un espace d'Eilenberg-MacLane).

On note $\Gamma_\sigma = \{ (x, \sigma(x)) \in N \times \widetilde{\Lambda(x)}, x \in N \}$ le graphe de σ.

5.4. On suppose de plus que la restriction $TF_N = F_N \times DF_N$ de TF à N *laisse invariant* le fibré ξ.

5.5. *Exemple.* N est une sous-variété de classe C^1, connexe lagrangienne de $\mathbf{A}^n_R$ invariante par F vérifiant Maslov(N) = 0. On prend $\xi = T(N)$ = le fibré tangent de N, $\tau_N = \bar{\sigma}$. La condition

$$\mathrm{Maslov}(N) = 0 = \tau_N^*(1) \in H^1(N, \mathbf{Z})$$

est équivalente à 5.3.

5.6. Comme le fibré ξ est invariant par TF_N et que N est connexe, il existe $q \in \mathbf{Z}$ tel que

$$(\mathrm{id}_N \times Q_q) \circ \widetilde{TF}_N(\Gamma_\sigma) = \Gamma_\sigma. \tag{5.7}$$

On rappelle que l'on supposera dans la suite la condition 4.27 satisfaite.

5.8. *Proposition. — On a* $\widetilde{TF}_N(\Gamma_\sigma) = \Gamma_\sigma$ *i.e. dans l'entier q de* 5.7 *est nul.*

Démonstration. — Soit $x_0 \in N$ tel que

$$\delta_M(\sigma(x_0)) = \operatorname*{Sup}_{x \in N} \delta_M(\sigma(x)).$$

On a

$$\delta_M(\sigma(x_0)) \geqslant \delta_M(\sigma(F_N^{2n}(x_0)))$$

et en utilisant 3.6

$$\delta_M(\sigma(x_0)) \geqslant \delta_M(Q_{2nq} \circ \widetilde{DF}_N^{2n}(x_0)(\sigma(x_0))) = 2nq + \delta_M(\widetilde{DF}^{2n}(x_0)(\sigma(x_0))).$$

D'après 3.21,

$$\delta_M(\widetilde{DF}_N^{2n}(x_0)(\sigma(x_0))) \geqslant \delta_M(\sigma(x_0)) + \delta_M(\widetilde{DF}_N^{2n}(x_0)(\tilde{h}_0)) - \delta_M(\tilde{h}_0) - n.$$

En utilisant 4.25, il suit

$$\delta_M(\sigma(x_0)) \geqslant 2nq + \delta_M(\sigma(x_0)) - n - (1/2)$$

et donc, puisque $q \in \mathbf{Z}$ et $n \geqslant 1$, on obtient $q \leqslant 0$.

Soit $x_1 \in N$ tel que

$$\delta_M(\sigma(x_1)) = \operatorname*{Inf}_{x \in N} \delta_M(\sigma(x)).$$

On a

$$\delta_M(\sigma(x_1)) \leqslant \delta_M(\sigma(F_N^{2n}(x_1))) = 2nq + \delta_M(\widetilde{DF}_N^{2n}(x_1)\sigma(x_1))$$

et en utilisant 3.21 et 4.25,

$$\delta_M(\sigma(x_1)) \leqslant 2nq + \delta_M(\sigma(x_1)) + n + (1/2),$$

ce qui implique que $q \geqslant 0$.

On a bien montré que dans 5.7, $q = 0$. ■

5.9. *Remarque.* — C'est seulement ici qu'on a besoin de 4.25. Il est probable que le fait que $q = 0$ dans le cas où N est une variété lagrangienne et $\bar{\sigma} = \tau_N$ ait une interprétation géométrique qui permette de remplacer 4.25 par une condition géométrique sur N.

6. Ensemble récurrent par chaînes

Nous allons rappeler les définitions standard et quelques propriétés immédiates dont les démonstrations sont laissées au lecteur qui peut aussi se reporter à [C_2].

6.1. On se donne un espace métrique (X, d_0) non vide et $f: X \to X$ un homéomorphisme.

Une suite $(x_j)_{0 \leqslant j \leqslant s}$, $s \geqslant 1$ est une ε *f-chaîne entre x et y* : si $x_0 = x$, $x_s = y$ et $d_0(x_{i+1}, f(x_i)) < \varepsilon$, $i = 0, \ldots, s-1$.

On définit la *relation* de *préordre* sur X : $x \mathscr{R}_f y$ si tout $\varepsilon > 0$ il existe une ε f-chaîne entre x et y.

On dit que x est *récurrent par* f-chaînes si $x \mathscr{R}_f x$.

On note $\mathscr{R}(f) = \{ x \in X, x \mathscr{R}_f x \}$ *l'ensemble récurrent par chaînes* de f.

On dit que f est *récurrent par chaînes* si $\mathscr{R}(f) = X$.

On suppose que f et f^{-1} sont uniformément continus pour la métrique d_0 sur X. Alors $\mathscr{R}(f)$ est fermé invariant par f. L'ensemble $\mathscr{R}(f)$, si X est non compact, dépend, en général, du choix de la métrique d_0;

Exemple. — On prend $X = \mathbf{Z}, f(x) = x + 1$; avec la métrique $d_0(x, y) = | x - y |$, on a $\mathscr{R}_{d_0}(f) = \emptyset$, et avec la métrique induite d'une métrique sur le compactifié d'Alexandroff $\hat{\mathbf{Z}} = \mathbf{Z} \cup \{ \infty \}$, $\mathscr{R}(f) = \mathbf{Z}$.

6.2. On suppose que X est compact métrique. Alors $\mathscr{R}(f)$ est compact non vide (il contient l'ensemble non errant $\Omega(f)$). L'ensemble $\mathscr{R}(f)$ ne dépend pas du choix de la métrique d_0 sur X.

On montre [C_2] que :

- $f_{|\mathscr{R}(f)}$ est récurrent par chaîne.
- Si X est connexe et $\mathscr{R}(f) = X$, alors, pour tous x, y dans X, $x \mathscr{R}_f y$.

6.3. Le principal avantage de l'ensemble récurrent par chaînes, quand X est compact métrique, est que *l'application multivoque* :

$$f \in \text{Homéo}(X) \mapsto \mathscr{R}(f) \in \mathscr{K}(X)$$

est *semi-continue supérieurement*, où Homéo(X) est le groupe des homéomorphismes avec la topologie de la convergence uniforme et $\mathscr{K}(X)$ désigne l'ensemble des compacts de X avec la métrique de Hausdorff. Cette propriété est immédiate. C'est son principal avan-

tage sur d'autres définitions d'ensembles fermés invariants, par exemple l'ensemble non errant $\Omega(f)$.

On a, si h est un homéomorphisme de X,

(6.4) $$\mathscr{R}(h \circ f \circ h^{-1}) = h(\mathscr{R}(f)).$$

6.5. On utilise les notations de 6.1.

Lemme. — *Soient* $G \subset X$ *un sous-ensemble vérifiant* $V_\varepsilon(f(G)) \subset G$ *pour un* $\varepsilon > 0$, *où* $V_\varepsilon(G) = \{ x \in X$, *il existe* $y \in G$ *vérifiant* $d_0(x, y) < \varepsilon \}$, *et* $z \notin G$ *vérifiant* $d_0(z, V_\varepsilon(f(G))) > 0$ *et tel qu'il existe un entier* $q \geqslant 1$ *avec* $f^q(z) \in \mathrm{Int}(G)$, *où* $\mathrm{Int}(G)$ *désigne l'intérieur de* G; *alors* z *n'est pas récurrent par* f*-chaînes dans l'espace métrique* (X, d_0).

Démonstration. — On choisit $0 < \varepsilon' < \varepsilon$ assez petit et on suppose par l'absurde qu'il existe une ε' f-chaîne $(x_j)_{0 \leqslant j \leqslant s}$ entre z et z. Par la continuité de f on a nécessairement $s > q$ car $x_q \in \mathrm{Int}(G)$, si ε' est assez petit (x_q est voisin de $f^q(z)$). Comme $\varepsilon' < \varepsilon$ et $V_\varepsilon(f(G)) \subset G$, on a, pour tout $j > q$,

$$x_j \in V_\varepsilon(f(G)) \quad \text{et} \quad d_0(x_j, z) > (1/2)\, d_0(z, V_\varepsilon(f(G))) > 0.$$

Par l'absurde, z n'est pas récurrent par f-chaînes. ■

6.6. On se place dans les hypothèses de 5.8. On pose

$$\widetilde{TF}_\sigma = \widetilde{TF}_{N|\Gamma_\sigma}$$

et $\mathscr{R}(\widetilde{TF}_\sigma) \subset \Gamma_\sigma$ désigne son ensemble récurrent par chaîne. Comme $\widetilde{TF}_\sigma$ est conjugué à $F_N : N \to N$ par la première projection $p_{1|\Gamma_\sigma}$, on a, d'après 6.4,

$$p_1(\mathscr{R}(\widetilde{TF}_\sigma)) = \mathscr{R}(F_N).$$

On a trivialement : Si $x \in \mathscr{R}(\widetilde{TF}_\sigma)$, alors x est récurrent par $\widetilde{TF}_N$-chaînes

$$(\widetilde{TF}_N : N \times \widetilde{\Lambda(n)} \to N \times \widetilde{\Lambda(n)})$$

et ceci pour toute métrique sur $N \times \widetilde{\Lambda(n)}$ définissant sa topologie.

Pour éviter l'exemple de 6.1, nous introduisons la notion de bonne métrique sur $N \times \widetilde{\Lambda(n)}$: une métrique d_0 sur $N \times \widetilde{\Lambda(n)}$ est dite *une bonne métrique* si elle vérifie, pour tout $z_j = (x_j, y_j) \in N \times \widetilde{\Lambda(n)}$, $j = 1, 2$, la relation $d_0(z_1, z_2) \geqslant |\, \delta_M(y_1) - \delta_M(y_2) \,|$.

7. Propositions pour démontrer le théorème 2.2

7.1. On pose, pour $k \in \mathbf{Z}$,

$$Z_k = \{ y \in \widetilde{\Lambda(n)}, \,|\, \delta_M(y) - k \,| < s \}$$

où $\delta_M(W) =]-s, s[$, $0 < s < (1/2)$ (cf. 4.9).

Si $a_1, a_2 \in \mathbf{R}$, on note

$$Y_{(a_1, a_2)} = \{y \in \widetilde{\Lambda(n)}, \delta_M(y) \in (a_1, a_2)\}$$

et $Y_{a_1} = Y_{]-\infty, a_1]}$. D'après 3.7, $Y_{]-\infty, a_1]}$ est une variété connexe ayant pour bord la variété compacte connexe $\delta_M^{-1}(a_1)$.

On suppose dans la suite que 4.27 est vérifiée, ce qui implique que 5.8 est aussi vérifiée.

Les deux propositions suivantes (avec 9.4) vont impliquer le théorème 2.2.

7.2. *Proposition.* — *Soit* $z = (x, y) \in N \times \widetilde{\Lambda(N)}$ *avec* $y \notin Z_k$ *pour tout* $k \in \mathbf{Z}$, *alors* z *n'est pas récurrent par* $\widetilde{TF}_N$*-chaînes* ($\widetilde{TF}_N$ *agit sur* $N \times \widetilde{\Lambda(n)}$) *pour tout choix de bonne métrique sur* $N \times \widetilde{\Lambda(n)}$, *et* a fortiori $z \notin \mathscr{R}(\widetilde{TF}_\sigma)$; *cf.* 6.5.

7.3. *Démonstration.*

On pose $\delta_M(y) = a$. Si $\delta_M(y_1) = a$, alors $y_1 \notin \{Q_q(W), q \in \mathbf{Z}\}$ où l'ouvert W est défini en 4.5.

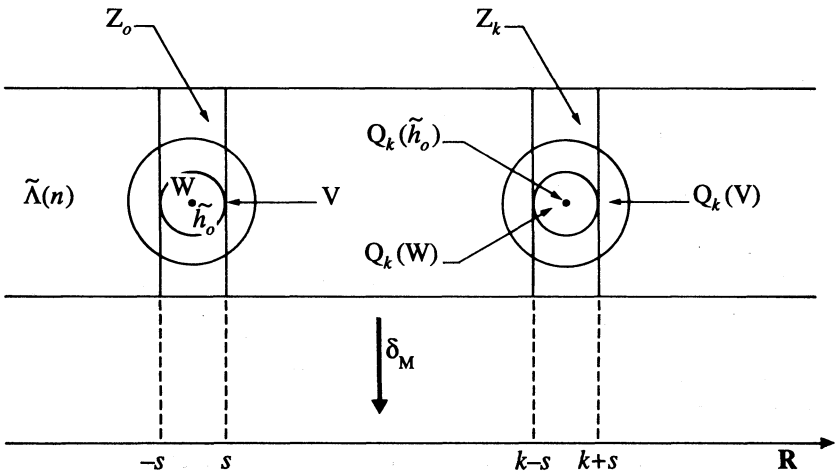
On a

$$\widetilde{TF}_N(N \times Y_a) \subset N \times Y_{a-\eta_1}. \tag{7.4}$$

En effet, d'après 4.22, pour tout $x \in N$,

$$\delta_M(\widetilde{DF}_N(x)\, y_1) < \delta_M(y_1) - \eta_1 = a - \eta_1.$$

La variété à bord $\widetilde{DF}_N(x)\,(Y_a)$ contient Y_{a_1} pour $a_1 \ll a$, et son bord est inclus dans $Y_{a-\eta_1}$ (l'ensemble $\widetilde{DF}_N(x)\,(Y_a) \cap Y_{]a-\eta_1, +\infty[}$ est ouvert et fermé, et borné dans $Y_{]a-\eta_1, +\infty[}$, donc vide).



On choisit d_0, une bonne métrique sur $X = N \times \widetilde{\Lambda(n)}$. On peut trouver $\varepsilon > 0$ et appliquer 6.5 avec $G = N \times Y_{]+\infty, a[}$ et donc z n'est pas récurrent par $\widetilde{TF}_N$-chaînes. (Nous avons supposé ici que d_0 est une bonne métrique pour éviter l'exemple de 6.1). ■

7.5. *Proposition.* — *Si* $z = (x, y) \in N \times (Z_k - Q_k(V))$, $k \in \mathbf{Z}$, *alors* z *n'est pas récurrent par* $\widetilde{TF}_N$*-chaînes pour tout choix de la bonne métrique sur* $N \times \widetilde{\Lambda(n)}$.

Démonstration. — Par 4.20 on a

$$\delta_M(\widetilde{DF}_N(x)y) < k + s - \eta < k - s$$

(on utilise que $2s < \eta$; cf. 4.9).

Comme $\widetilde{TF}_N(N \times Y_{k-s}) \subset N \times Y_{k-s-\eta_1}$, on a

$$T^2 F_N(z) \in \mathrm{Int}(G),$$

où $G = N \times Y_{k-s}$ et $z \notin G$ et la proposition suit de 6.5. ■

7.6. *Corollaire.* — *Soit* $z = (x, y) \in \Gamma_\sigma$ *récurrent par* $\widetilde{TF}_N$*-chaînes; alors il existe* $k \in \mathbf{Z}$ *tel que les*

$$\widetilde{TF}_N^q(z) = z_q = (x_q, y_q), \quad q = 0, 1, \ldots$$

vérifient

$$y_q \in Q_k(V) \cap Z_k, \quad q = 0, 1, \ldots$$

Démonstration. — D'après les propositions 7.2 et 7.5, $y \in Q_k(V) \cap Z_k$ pour un $k \in \mathbf{Z}$. Comme $\mathscr{R}(\widetilde{TF}_\sigma)$ est invariant par $\widetilde{TF}_N$, on en conclut que $y_q \in Q_{k_q}(V) \cap Z_{k_q}$ pour un $k_q \in \mathbf{Z}$, et la proposition suit en utilisant 4.24. ■

8. Conséquences

8.1. On peut écrire $\mathscr{R}(\widetilde{TF}_\sigma) = \widetilde{\mathscr{R}}_1 \amalg \widetilde{\mathscr{R}}_2 \amalg \ldots \amalg \widetilde{\mathscr{R}}_j$, où chaque $\widetilde{\mathscr{R}}_i$, $1 \leqslant i \leqslant j$, est un fermé, invariant par $\widetilde{TF}_N$, $\widetilde{\mathscr{R}}_{i_1} \cap \widetilde{\mathscr{R}}_{i_2} = \emptyset$, si $i_1 \neq i_2$, et de plus

$$\widetilde{\mathscr{R}}_i \subset Q_{k_i}(V) \cap Z_{k_i}$$

où $k_i \in \mathbf{Z}$. On suppose que $k_1 < k_2 < \ldots < k_j$.

Soit $\mathscr{R}_1 \amalg \mathscr{R}_2 \amalg \ldots \amalg \mathscr{R}_j$ la décomposition de $\mathscr{R}(F_N)$ obtenu par projection de la décomposition $(\widetilde{\mathscr{R}}_i)_{1 \leqslant i \leqslant j}$ de $\mathscr{R}(\widetilde{TF}_\sigma)$ par la première projection, cf. 6.5.

Comme $\widetilde{\mathscr{R}}_i \subset Q_{k_i}(V) \subset Q_{k_i}(\check{U}_0)$, on conclut que pour tout $1 \leqslant i \leqslant j$, le fibré $\xi_{|\mathscr{R}_i} \to \mathscr{R}_i$ est de la forme

$$\{(x, u, m_i(x)u) \in \mathscr{R}_i \times \mathbf{R}^n \times (\mathbf{R}^n)^*;\ x \in \mathscr{R}_i,\ u \in \mathbf{R}^n\},$$

où $x \in \mathscr{R}_i \mapsto m_i(x) \in M_n(\mathbf{R})$ est continue et vérifie ${}^t m_i = m_i$ (la continuité de m_i résultant de l'unicité de m_i et du fait que le fibré $\xi \to N$ est un fibré continu). On dit que le fibré vectoriel $\xi_{|\mathscr{R}_i}$ est le *graphe* de m_i, ou, ce qui revient au même, de la fonction continue $(x, u) \in \mathscr{R}_i \times \mathbf{R}^n \mapsto m_i(x)u \in (\mathbf{R}^n)^*$.

8.2. Si on fait l'hypothèse que

$$\mathscr{R}(F_N) = N \quad \textit{i.e.}\ f_N \text{ est récurrent par chaînes,}$$

comme N est connexe, et d'après 4.23, $Q_q(V) \cap Q_{q'}(V) = \emptyset$, si $q \neq q'$, on conclut que

$$\sigma(N) \subset Q_k(V) \quad \text{pour un } k \in \mathbf{Z}$$

et donc

$$\bar{\sigma}(N) \subset p(V) \subset U_0 \subset \Lambda(n).$$

Il suit que le fibré vectoriel $\xi \to N$ est le graphe d'une fonction continue

$$x \in N \mapsto m(x) \in M_n(\mathbf{R}), \qquad {}^t m = m.$$

8.3. *Proposition.* — *Quitte à diminuer les nombres* $\varepsilon(L, R, R_1)$ et $\varepsilon'(L, R, R_1)$ *que nous avons déterminés en* 4.27, *si on suppose que la décomposition* $(\widetilde{\mathscr{R}}_i)_{1 \leqslant i \leqslant j}$ *de* $\mathscr{R}(\widetilde{TF}_\sigma)$ *a un seul élément (i.e.* $j = 1$*), alors*

$$\Gamma_\sigma \subset N \times (Q_{k_1}(V) \cap Z_{k_1}).$$

8.4. *Remarque.* — La proposition ci-dessus s'applique si $\mathscr{R}(F_N)$ est connexe et implique que le fibré $\xi \to N$ est un graphe d'une fonction continue $m : N \to M_n(\mathbf{R})$.

Démonstration de 8.3. — On a

$$\mathscr{R}(\widetilde{TF}_\sigma) \subset N \times (Q_{k_1}(V) \cap Z_{k_1}).$$

Si on suppose que

$$\Gamma_\sigma \cap [(N \times Y_{k_1 - s}) \cup N \times (Z_{k_1} - Q_{k_1}(V))] \neq \emptyset,$$

alors on peut trouver $x \in \Gamma_\sigma$ tel que l'ensemble ω limite par $\widetilde{TF}_\sigma$ soit inclus dans $N \times Y_{k_1 - s - \eta_1}$, et donc $j = 1$ est absurde.

Si $\Gamma_\sigma \cap (N \times Y_{[k_1 + s, +\infty[}) \neq \emptyset$, alors, vu 4.29, il existe $x \in \Gamma_\sigma$ tel que l'ensemble α limite de x par $\widetilde{TF}_\sigma$ soit inclus dans $N \times Y_{[k_1 + s, +\infty[}$, ce qui contredit aussi le fait que $j = 1$. ■

9. Fin de la démonstration du théorème 2.2

9.1. On note à nouveau la première projection $p_1 : (\theta, r) \in \mathbf{T}^n \times \mathbf{R}^n \mapsto \theta \in \mathbf{T}^n$. On rappelle qu'on a les isomorphismes

$$\mathbf{A}^n \cong T^*(\mathbf{T}^n) \cong \mathbf{T}^n \times \mathbf{R}^n.$$

On se place sous les hypothèses du théorème.

D'après 5.5, on conclut, si 4.27 est vérifié et $\mathscr{R}(F_N) = N$ (*i.e.* f_N est récurrent par chaînes), qu'on peut appliquer 8.2, ce qui implique la proposition suivante.

9.2. *Proposition.* — *L'application* $p_N \equiv p_{1|N} : N \to \mathbf{T}^n$ *est une immersion et donc un revêtement.*

9.3. Le théorème 2.2 découle maintenant de la proposition suivante, due à V. I. Arnold, dont la démonstration est esquissée dans [B_4]. Pour la commodité du lecteur, nous donnons une démonstration complète et valable avec l'hypothèse que N est de classe C^1.

9.4. *Proposition.* — *Soit* $N \subset T^*(\mathbf{T}^n)$ *une sous-variété lagrangienne compacte connexe de classe* C^1 *telle que* p_N *soit un revêtement*; *alors* p_N *est un difféomorphisme.*

Démonstration. — On écrit la suite exacte d'homotopie

$$0 \to \pi_1(N) \xrightarrow{(p_N)_*} \pi_1(\mathbf{T}^n) \to G \to 0.$$

Dans la suite, comme tous les groupes fondamentaux seront abéliens, nous éviterons les points bases. Puisque N est compacte, G est un groupe abélien fini.

Soit A un endomorphisme continu surjectif du groupe $\mathbf{T}^n$ tel que

$$0 \to \pi_1(\mathbf{T}^n) \xrightarrow{A_*} \pi_1(\mathbf{T}^n)$$

vérifie

$$\mathfrak{I}A_* = \mathfrak{I}(p_N)_* \subset \pi_1(\mathbf{T}^n), \tag{9.5}$$

où $\mathfrak{I}$ désigne l'image. Soit $T^* A(\theta, r) = (A(\theta), {}^tA^{-1}(r))$.

L'application

$$T^* A : T^*(\mathbf{T}^n) \to T^*(\mathbf{T}^n)$$

est un revêtement fini symplectique qui est un endomorphisme de groupe de Lie, les automorphismes du revêtement étant de la forme $\theta \to (\theta + g, r)$ où $g \in \hat{G} \subset (\mathbf{Q}/\mathbf{Z})^n$, $\hat{G}$ désignant le groupe dual de G. On pose $Z = (T^* A)^{-1}(N)$.

Il suit de 9.5 qu'on peut relever l'application continue A

$$\begin{tikzcd} & N \arrow[d, "p_N"] \\ \mathbf{T}^n \arrow[r, "A"] \arrow[ur, "\varphi"] & \mathbf{T}^n \end{tikzcd}$$

en une application continue φ vérifiant $p_N \circ \varphi = A$. De plus $\varphi_* : \pi_1(\mathbf{T}^n) \xrightarrow{\cong} \pi_1(N)$ est un isomorphisme (A_* et $(p_N)_*$ sont injectifs).

On considère l'application

$$\varphi : \mathbf{T}^n \to N \subset T^*(\mathbf{T}^n).$$

Comme $\mathfrak{I}\varphi_* = \mathfrak{I}(T^* A)_* \subset \pi_1(T^*(\mathbf{T}^n))$, on peut relever φ

$$\begin{tikzcd} & T^*(\mathbf{T}^n) \arrow[d, "T^* A"] \\ \mathbf{T}^n \arrow[ur, "\varphi_A"] \arrow[r, "\varphi"] & T^*(\mathbf{T}^n) \end{tikzcd}$$

en une application continue φ_A vérifiant $(T^* A) \circ \varphi_A = \varphi$. Il suit que le revêtement $p_{1|Z} : Z \to \mathbf{T}^n$ est trivial. Comme $T^* A_{|Z} : Z \mapsto N$ est aussi un revêtement trivial et que N est compacte et connexe, on conclut :

(i) Z est une sous-variété lagrangienne de $T^*(\mathbf{T}^n)$.

(ii) Si le revêtement $p_N : N \to \mathbf{T}^n$ est non trivial, il existe $g \neq 0 \in \mathbf{T}^n$ et une composante connexe Z_0 de Z telle que

$$H(Z_0) \cap Z_0 = \emptyset, \quad \text{avec } H(\theta, r) = (\theta + g, r).$$

On écrit $Z_0 = \Gamma_\psi$ où $\psi = (\psi_1, \ldots, \psi_n) \in C^1(\mathbf{T}^n, \mathbf{R}^n)$ et comme Z_0 est une sous-variété lagrangienne de $T^*(\mathbf{T}^n)$ la 1-forme (de classe C^1) $\sum_1^n \psi_j \, d\theta_j$ est fermée.

On a $H(Z_0) = \Gamma_{\psi \circ R_{-g}}$ et $\psi \circ R_{-g} - \psi = dS$, où $S \in C^2(\mathbf{T}^n, \mathbf{R}^n)$ a un point critique, et on arrive à une contradiction avec (ii) si on suppose que le revêtement $p_N : N \to \mathbf{T}^n$ est non trivial. ■

10. Compléments et remarques

10.1. Dans le théorème 2.2 on peut supposer que $R = R_1 = +\infty$ (*i.e.* on n'a pas besoin d'imposer 2.4). On impose que toutes les conditions de la remarque 1 de 4.28 soient vérifiées (puisqu'on a 4.20 à 4.25, et c'est tout ce dont on a besoin par la suite).

Exemples. — On peut trouver $\varepsilon > 0$ ne dépendant que de la dimension n tel que si $\hat{\varphi} \in C^2(\mathbf{T}^n, \mathbf{R})$, $\varphi = \dfrac{\partial \hat{\varphi}}{\partial \theta}$, et si

$$\| D\varphi \|_{C^1(\mathbf{T}^n)} < \varepsilon_n,$$

alors on a les conclusions du théorème 2.2 pour F_φ où $F_\varphi(\theta, r) = (\theta + r, r + \varphi(\theta + r))$ pour toute variété lagrangienne $N \subset \mathbf{A}^n$ compacte connexe ayant une classe de Maslov nulle et telle que $F_{\varphi|N}$ soit récurrent par chaîne.

10.2. On suppose que $N = \Gamma_\psi$, $\psi \in C^1(\mathbf{T}^n, \mathbf{R}^n)$ est un tore lagrangien invariant par le difféomorphisme symplectique F, où $\| F - L \|_{C^1(\mathbf{A}^n_{R_1})} < \varepsilon$ et $N \subset \mathbf{A}^n_R$.

La démonstration que nous avons donnée montre géométriquement que $\| D\psi \|_{C^0} \to 0$, si $\varepsilon \to 0$. Cela suit aussi de I.7, qui donne un résultat plus précis, mais nous allons le montrer ici de la façon suivante : Si $\varepsilon \to 0$, on peut supposer que $\eta \to 0$

et que le voisinage V de $\tilde{h}_0$ déterminé en 4.3 tend vers $\tilde{h}_0 \in \check{U}_0$. (V tend vers $\tilde{h}_0$, ce qu'on écrit $V \to \tilde{h}_0$, et veut dire que le diamètre de V tend vers 0.)

On remarque que, puisque N est un graphe, Maslov(N) = 0, et on peut supposer que $\sigma = \tilde{\tau}_N$ a son image incluse dans $\check{U}_0$.

Ensuite, on peut remarquer que 8.3 s'applique avec $k_1 = 0$ puisque $\sigma(N) \subset \check{U}_0$ et $Q_q(\check{U}_0) \cap \check{U}_0 = \emptyset$, si $q \neq 0$. On conclut finalement que

$$\tau_N(N) \subset p_1(V) \to h_0, \quad \text{si } \varepsilon \to 0.$$

10.3. Ce que nous avons fait en 8.2 s'applique aux tores moustachus d'Arnold [A_2].

Exemple. — On suppose que N est de classe C^1 et est difféomorphe à $\mathbf{T}^{n_1}$, $0 \leqslant n_1 < n$, $F(N) = N$, $N \subset \mathbf{A}^n_{\mathbf{R}}$, et que F_N est C^1-conjugué à une translation ergodique de $\mathbf{T}^{n_1}$. La démonstration de I.3.2 montre que $j_N^* w = 0$. On suppose qu'il existe sur N un fibré (vectoriel) W^s dont la dimension de la fibre en chaque point est égale à $n - n_1$. On suppose de plus que le fibré W^s est invariant par TF_N et qu'il existe $C > 0$ et $0 < \lambda < 1$ vérifiant

$$\| TF_N^k(x)\, v \| \leqslant C\lambda^k \| v \| \quad \text{pour tout } x \in N,\ v \in W^s_x \text{ et } k \in \mathbf{N}.$$

Le fibré $\xi = T(N) \oplus W^s$ est donc un fibré vectoriel lagrangien sur N, invariant par TF_N.

C'est un fibré tangent le long de N, du feuilletage stable pour F asymptotique à N :

$$W^s_N = \{ x \in \mathbf{A}^n,\ \limsup_{k \to +\infty} \operatorname{Log} d_0(F^k(x), N) < 0 \}$$

où $d_0(x, N)$ désigne la distance de x à N et d_0 est la métrique sur $\mathbf{A}^n$ qu'on a définie en 4.12.

L'union des feuilles W^s_N a été appelée par V. I. Arnold *moustaches stables* du tore N (voir [A_2]).

Si on suppose que les fibrés W^s, W^u et T(N) ainsi que le difféomorphisme F et la conjugaison de F_N à une translation sont de classe C^r avec r grand, on peut montrer que W^s_N est une sous-variété de classe C^1 lagrangienne (cf. [Z]).

On suppose que 5.3′ est vérifiée :

$$0 = \text{Maslov}(\xi) \in H^1(N, \mathbf{Z}).$$

Alors 8.2 s'applique, si 4.27 est vérifiée, et il suit que le fibré ξ est le graphe de m et donc $p_{1|N} : N \to \mathbf{T}^n$ est une immersion.

De plus, comme en 10.2, on conclut que, si $\varepsilon \to 0$, le fibré ξ tend vers le fibré $N \times (\mathbf{R}^n \times \{0\}) \subset T(\mathbf{A}^n)$ ou de façon équivalente $\| m \|_{C^0(N)} \to 0$, $m : N \to M_n(\mathbf{R})$, ${}^t m = m$.

Les conclusions ci-dessus sont *en défaut* si on ne suppose pas Maslov(ξ) = 0 ainsi qu'on peut facilement le voir en considérant $F = L_1 \times F_1$, où $L_1 : \mathbf{A}^{n-1} \to \mathbf{A}^{n-1}$ est complètement intégrable, F_1 possède une orbite périodique elliptique rationnelle ($F_1 : \mathbf{A} \to \mathbf{A}$), et en supposant dans l'exemple que L est un difféomorphisme produit : $L = L_1 \times L_2$.

10.4. *Cas particulier de* 10.3.

On suppose que F est le temps $s = 1$ d'un flot $(F_s)_{s \in \mathbf{R}}$ d'un hamiltonien $H \in C^2(\mathbf{A}^n, \mathbf{R})$ (auquel on impose, pour simplifier, de définir un champ de vecteurs X_H complet).

On suppose que le flot F_s a une orbite périodique N comprise dans une surface d'énergie $H^{-1}(c)$, où $c \in \mathbf{R}$ est valeur régulière de H. De plus, on impose $N \subset \mathbf{A}^n_{\mathbf{R}}$ et que l'orbite $N \subset H^{-1}(c)$ soit normalement hyperbolique, ainsi que $\mathrm{Maslov}(TN \oplus W^s) = 0$, où W^s désigne le fibré stable de F_s le long de N.

Si 4.27 est vérifié, alors 10.3 permet de conclure que le fibré vectoriel $\xi = TN \oplus W^s$ est un graphe d'une fonction m et, si $\varepsilon \to 0$, le fibré ξ tend vers le fibré

$$N \times (\mathbf{R}^n \times \{0\}) \subset T(\mathbf{A}^n).$$

10.5. On suppose, comme en 10.4, que F est le temps $s = 1$ d'un flot hamiltonien F_s complet.

On relève en 4.15 l'isotopie $(TF_s)_{0 \leqslant s \leqslant 1}$ en $\widetilde{TF}_s$ avec $\widetilde{TF}_{s=0} = \mathrm{Id}$, et pour relever L on choisit l'isotopie $L_s(\theta, r) = (\theta + s\ell(r), r)$.

Avec ces choix de relèvements, on impose 4.17 pour obtenir 4.20 à 4.24, mais on *n'impose pas* 4.25. (On doit demander 4.30 si on veut avoir 8.3).

On avait besoin de 4.25 pour en déduire 5.8.

Dans le cas que nous considérons maintenant on a bien 5.8 (sans avoir besoin de 4.25). En effet, pour tout $s \in \mathbf{R}$, 5.7 s'applique.

Soit

$$G = \{ s \in \mathbf{R} \mid \widetilde{TF}_s(\Gamma_\sigma) = \Gamma_\sigma \}.$$

G est un sous-groupe de $\mathbf{R}$ puisque $\widetilde{TF}_s$ définit une action continue de $\mathbf{R}$ sur $\mathbf{A}^n \times \widetilde{\Lambda(n)}$. Le sous-groupe G est évidemment fermé et contient $\{ |s| < \varepsilon \}$ pour un $\varepsilon > 0$. (On utilise Γ_σ compact connexe et $(\mathrm{Id} \times Q_q)(\Gamma_\sigma) \cap \Gamma_\sigma = \emptyset$ si $q \neq 0$.)

Il suit que $G = \mathbf{R}$. En particularisant pour $s = 1$

$$\widetilde{TF} = \widetilde{TF}_{s=1},$$

$$\widetilde{TF}(\Gamma_\sigma) = \Gamma_\sigma,$$

ou encore, puisque $F(N) = N$,

$$\widetilde{TF}_N(\Gamma_\sigma) = \Gamma_\sigma.$$

Ceci démontre 5.8 dans le cas où nous nous sommes placés.

Ainsi que nous l'avons remarqué en 4.28.2, les conditions 4.20 à 4.24 sont de nature géométrique. La condition 4.25 est *a priori* nettement plus restrictive, ainsi que nous l'avons remarqué en 4.28.3. Ainsi que nous l'avons remarqué en 5.9, on doit probablement pouvoir éviter 4.25, ou tout au moins remplacer ces conditions par des conditions d'essence plus géométrique.

10.6. On suppose que F est un plongement symplectique de classe C^1, défini sur $\mathbf{A}^n_{R_0}$, $R_0 > 0$, laissant invariant $\mathbf{T}^n \times \{0\}$ et de la forme :

$$F(\theta, r) = (\theta + \alpha + Br + o(r), r + o(r))$$

où $\alpha \in \mathbf{T}^n$, $B \in M_n(\mathbf{R})$, ${}^tB = B$, $B > 0$ et le petit o voulant dire que les termes divisés par $\| r \|$ tendent vers 0 en topologie C^1, lorsque $\| r \| \to 0$.

On suppose que $R_1 = 2R$ et $L(\theta, r) = (\theta + \alpha + Br, r)$. Si dans le théorème 2.2, $R \to 0$, alors le nombre $\varepsilon(L, R, 2R)$ ne dépend que de la matrice $B > 0$ et des termes $o(r)$, et, si $R \to 0$, 2.3 est automatiquement vérifié. Si on suppose que $R \to 0$ et que les hypothèses 2.4 à 2.6 sont remplies, alors on a la conclusion du théorème 2.2 :

$$N = \Gamma_\psi, \quad \psi \in C^1(\mathbf{T}^n, \mathbf{R}^n).$$

De plus, si $R \to 0$, il suit de 10.2 que

$$\| D\psi \|_{C_0} \to 0.$$

En effet, si $R \to 0$, $F_{|\mathbf{A}^n_{2R}} \to L_{|\mathbf{A}^n_{2R}}$ en topologie C^1 (par définition des termes $o(r)$) et le fait que nous ayons supposé $R < R_1$ (on choisit ici $R_1 = 2R$) n'est seulement utilisé, au cours de la démonstration, que pour avoir 4.14, ce qui est automatique si $R \to 0$, puisque $F^j(\theta, r) = (\theta_j, r + no(R))$ si $|j| \leqslant 2n$ et $\| r \| \leqslant R$.

BIBLIOGRAPHIE

[A_1] V. I. Arnold, Characteristic class entering in quantization conditions, *Funct. Anal. Appl.*, **1** (1967), 1-13.

[A_2] V. I. Arnold et A. Avez, *Problèmes ergodiques de la mécanique classique*, Paris, Gauthier Villars, 1967.

[A_3] M. C. Arnaud, travail en préparation annoncé dans : Sur les points fixes des difféomorphismes exacts symplectiques de $\mathbf{T}^n \times \mathbf{R}^n$, *C.R. Acad. Sci. Paris*, t. **309** (1989), 191-194.

[A_4] S. Aubry and P. Y. Daeron, The discrete Frenkel-Kantova model and its extensions I. Exact results for ground states, *Physica*, **8D** (1983), 381-422.

[A_5] M. Audin, Fibrés normaux d'immersions en dimension double, points doubles d'immersions lagrangiennes et plongements totalement réels, *Comment. Math. Helv.*, **63** (1988), 593-623.

[B] V. Bangert, *Minimal geodesics*, Prepint, Univ. de Berne, 1987.

[BK] D. Bernstein and A. B. Katok, Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians, *Invent. Math.*, **88** (1987), 225-241.

[B_1] G. D. Birkhoff, Surface transformations and their dynamical applications, *Acta Math.*, **43** (1920), 1-119; *Collected Math. Papers*, vol. 2, p. 111-229.

[B_2] J.-B. Bost, Tores invariants des systèmes dynamiques hamiltoniens, *Séminaire Bourbaki*, Exposé n° 639, *Astérisque*, **133-134** (1986), 113-157.

[B_3] N. Bourbaki, *Intégration*, chap. 3, § 1, n° 6, 2e éd., Paris, Hermann, 1965.

[B_4] M.-L. Byalyi and L. V. Polterovich, Geodesic flows on the two-dimensional torus and phase transitions "commensurability-non commensurability", *Funct. Anal. Appl.*, **20** (1986), 260-266.

[B_5] M.-L. Byalyi and L. V. Polterovich, Lagrangian singularities of invariant tori of Hamiltonian systems with two degrees of freedom, *Invent. Math.*, **97** (1989), 291-303.

[B_6] M.-L. Byalyi, *Aubry Mather sets and Birkoff's theorem for geodesic flows on the two-dimensional torus*, Preprint 1988, Weizmann Inst.

[C_1] A. Chencinger, La dynamique au voisinage d'un point fixe elleptique conservatif : de Poincaré et Birkhoff à Aubry et Mather, *Séminaire Bourbaki*, exposé n° 622, *Astérisque*, **121-122** (1985), 147-170.

[C_2] C. C. Conley, The gradient structure of a flow : I, *Erg. Th. Dyn. Syst.*, **8*** (1988), 11-31.

[D] R. DOUADY, Stabilité ou instabilité des points fixes elliptiques, *Ann. Sci. Ec. Norm. Sup.*, 4e série, **21** (1988), 1-46.

[F] A. FATHI, « Une interprétation plus topologique de la démonstration du théorème de Birkhoff », Appendice du Chap. I de [H_3].

[G] W. H. GOTTSCHALK and G. A. HEDLUND, Topological Dynamics, *Am. Math. Soc.* (1955), § 14.

[H_1] M. R. HERMAN, Existence et non-existence de tores invariants par des difféomorphismes symplectiques, *Séminaire Equations aux dérivées partielles*, Exposé XIV, 1987-1988, Ecole Polytechnique.

[H_2] M. R. HERMAN, Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations, *Publ. Math. I.H.E.S.*, **49** (1979), 5-233.

[H_3] M. R. HERMAN, Sur les courbes invariantes par les difféomorphismes de l'anneau, vol. 1, *Astérisque*, **103-104** (1983).

[H_4] M. R. HERMAN, Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser, *Comment. Math. Helv.*, **58** (1983), 453-502.

[H_5] R. A. HORN et C. A. JOHNSON, *Matrix analysis*, New York, Cambridge Univ. Press (1985).

[J_1] R. A. JOHNSON, *m*-functions and Floquet exponents for linear differential systems, *Annali di Mat. Pura ed Appl.*, **147** (1987), 211-248.

[K] A. B. KATOK, *Minimal orbits for small perturbations of completely integrable Hamiltonian systems*, Preprint 1989, Cal. Tech.

[L_1] P. LE CALVEZ, Propriétés dynamiques des zones d'instabilités, *Ann. Sci. Ec. Norm. Sup.*, 4e série, **20** (1987), 443-464.

[L_2] P. LE CALVEZ, Propriétés générales des applications déviant la verticale, *Bull. Soc. Math. France*, **117** (1989), 69-102.

[M_1] J. MATHER, The existence of quasi-periodic orbits for twist homeomorphisms of the annulus, *Topology*, **21** (1982), 457-467.

[M_2] J. MATHER, Minimal measures, *Comm. Math. Helv.*, **64** (1989), 375-394.

[M_3] J. MATHER, *Letter to R. Mac Kay*, Feb. 21, 1984.

[M_4] J. MATHER, Minimal action measures for positive definite lagrangian systems, to appear, in *Proc. IXth Int. Conf. Math. Phys.*; version détaillée, Preprint, ETH, Zürich, 1989.

[M_5] J. MATHER, Exposés au Séminaire de Systèmes dynamiques au Centre de Mathématiques de l'Ecole Polytechnique (1985) ; *Variational construction of orbits of twist diffeomorphisms*, Preprint, ETH, Zürich, 1990.

[M_6] J. MATHER, Communications personnelles, 1988-1989.

[M_7] M. MOHSIN, Formes cobordables, *Ann. Sci. Ec. Norm. Sup.*, 3e série, **83** (1966), 201-213.

[M_8] M. MISIUREWIZ et K. ZIEMAN, *Rotation sets for maps of tori*, Preprint, Univ. Warwick (1988).

[S_1] D. SALAMON, *The Kolmogorov-Arnold-Moser theorem*, Preprint, ETH Zürich (1986).

[S_2] D. SALAMON and E. ZEHNDER, KAM theory in configuration space, *Comment. Math. Helv.*, **64** (1989), 84-142.

[S_3] J. J. SCHWARTZ, *Nonlinear functional analysis*, Courant Inst. of Math. Sci., 1965.

[S_4] L. SCHWARTZ, *Théories des distributions*, t. I et II, Paris, Hermann, 1950.

[S_5] M. SHUB, Stabilité globale des systèmes dynamiques, *Astérisque*, **56** (1978).

[S_6] E. M. STEIN, *Singular integrals and differentiability properties of functions*, Princeton, Princeton Univ. Press, 1970.

[V_1] C. VITERBO, Lettre 10 mai 1989.

[V_2] C. VITERBO, *A new obstruction to embedding Lagrangian tori*, Preprint, Berkeley, 1989, à paraître dans *Inventiones Math.*

[Z] E. ZEHNDER, Generalized implicit function theorems with applications to small divisor problems, II, *Comm. Pure Appl. Math.*, **29** (1976), 49-113.

Centre de Mathématiques
Ecole Polytechnique
F-91128 Palaiseau Cedex

Manuscrit reçu le 15 décembre 1989.