Ann. Scient. Éc. Norm. Sup.,
4e série, t. 34, 2001, p. 159 à 221.

A NEW METHOD FOR MEASURING THE SPLITTING OF INVARIANT MANIFOLDS

BY DAVID SAUZIN

ABSTRACT. – We study the so-called Generalized Arnol'd Model (a weakly hyperbolic near-integrable Hamiltonian system), with $d+1$ degrees of freedom ($d \geqslant 2$), in the case where the perturbative term does not affect a fixed invariant d-dimensional torus. This torus is thus independent of the two perturbation parameters which are denoted ε ($\varepsilon > 0$) and μ.

We describe its stable and unstable manifolds by solutions of the Hamilton–Jacobi equation for which we obtain a large enough domain of analyticity. The splitting of the manifolds is measured by the partial derivatives of the difference ΔS of the solutions, for which we obtain upper bounds which are exponentially small with respect to ε.

A crucial tool of the method is a *characteristic vector field*, which is defined on a part of the configuration space, which acts by zero on the function ΔS and which has constant coefficients in well-chosen coordinates.

It is in the case where $|\mu|$ is bounded by some positive power of ε that the most precise results are obtained. In a particular case with three degrees of freedom, the method leads also to lower bounds for the splitting.

RÉSUMÉ. – Nous étudions le système hamiltonien à $d+1$ degrés de liberté ($d \geqslant 2$), proche de l'intégrable et faiblement hyperbolique, appelé *modèle d'Arnol'd généralisé*, dans le cas où le terme perturbatif n'affecte pas un tore invariant de dimension d. Ce tore est donc indépendant des deux paramètres de perturbation qui sont notés ε ($\varepsilon > 0$) et μ.

Ses variétés stable et instable sont décrites par deux solutions de l'équation de Hamilton–Jacobi, pour lesquelles nous obtenons un domaine d'analyticité assez étendu. L'écart des variétés est mesuré par les dérivées partielles de la différence ΔS des solutions, et nous obtenons des bornes supérieures exponentiellement petites.

L'outil essentiel de la méthode est un *champ de vecteurs caractéristique* défini sur une partie de l'espace de configuration, qui annule la fonction ΔS et dont les coefficients sont constants dans un système de coordonnées bien choisi.

C'est dans le cas où $|\mu|$ est majoré par une puissance positive de ε que les résultats les plus précis sont obtenus. Dans un cas particulier avec trois degrés de liberté, la méthode fournit même des bornes inférieures pour l'écart des variétés.

1. Presentation of the problem

1.1. Introduction

The present paper is devoted to the exposition of a new method for studying the phenomenon of "exponentially small splitting". It is concerned with the stable and unstable manifolds of a partially hyperbolic invariant torus (a "whiskered torus") of a near-integrable Hamiltonian

system. We restrict our attention to systems for which the invariant torus is given right from the beginning: it is not affected by changes of the perturbation parameters, thus we can refrain from resorting to KAM theory to find an invariant torus; in our opinion this helps to isolate the mechanism which produces exponential smallness.

We shall illustrate the method on the case of the following Hamiltonian function:

$$H_{\varepsilon,\mu}(q,\varphi,p,I) = \omega\cdot I + \frac{1}{2}\alpha I^2 + \frac{1}{2}p^2 + \varepsilon(\cos q - 1) + \mu\varepsilon F(q,\varphi) \tag{1.1}$$

with $d+1$ degrees of freedom ($d \geqslant 2$), the conjugate variables being:

$$(p,I) \in \mathbb{R}\times\mathbb{R}^d \quad \text{and} \quad (q,\varphi) \in \mathbb{T}\times\mathbb{T}^d, \quad \text{with } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z},$$

and the various parameters being: a vector $\omega \in \mathbb{R}^d$, a real diagonal matrix $\alpha = \operatorname{diag}(\alpha_1,\dots,\alpha_d)$ (the notation αI^2 means $\sum \alpha_j I_j^2$), two small real parameters $\varepsilon > 0$ and μ, and a function F real-analytic on $\mathbb{T}\times\mathbb{T}^d$.

Let $\mathrm{d}(p\mathrm{d}q + I\mathrm{d}\varphi)$ be the usual symplectic two-form. The corresponding Hamiltonian system is integrable for $\mu = 0$, since it decouples then as the product of a simple pendulum and d independent rotators. We shall refer to that situation as to the "unperturbed" one.

Let us use the notation

$$\mathbb{T}_h = \big\{\phi \in \mathbb{C}/2\pi\mathbb{Z};\ |\Im\mathrm{m}\,\phi| < h\big\} \quad \text{for } h > 0.$$

We shall require two assumptions on F:

(A1) for all $\varphi \in \mathbb{T}^d$, $F(0,\varphi) = 0$ and $\partial_q F(0,\varphi) = 0$;

(A2) there exists $h_0 > 0$ such that F extends analytically to $(\mathbb{C}/2\pi\mathbb{Z})\times\mathbb{T}^d_{h_0}$.

The first hypothesis amounts to saying that F vanishes at order 2 on $\{q = 0\}$. As a result, the d-dimensional torus

$$\mathcal{T} = \big\{(0,\varphi,0,0),\ \ \varphi \in \mathbb{T}^d\big\}$$

is invariant by the Hamiltonian vector field

$$\begin{cases} \dot q = p, \\ \dot p = \varepsilon\sin q - \mu\varepsilon\partial_q F, \end{cases} \qquad \begin{cases} \dot\varphi_j = \omega_j + \alpha_j I_j, \\ \dot I_j = -\mu\varepsilon\partial_{\varphi_j} F, \end{cases} \tag{1.2}$$

independently of the parameters ε and μ. The restriction of its flow to $\mathcal{T}$ is quasiperiodic with ω as frequency-vector. We shall see that $\mathcal{T}$ is partially hyperbolic:[1] it admits $(d+1)$-dimensional stable and unstable manifolds, denoted by $\mathcal{W}^+_{\varepsilon,\mu}$ and $\mathcal{W}^-_{\varepsilon,\mu}$, the first one being the union of all the orbits which are positively asymptotic to $\mathcal{T}$, and the second one the union of all the orbits which are negatively asymptotic to $\mathcal{T}$. These manifolds depend analytically on μ and coincide for $\mu = 0$. In general there is no reason why they should coincide for nonzero μ, but it turns out that they are exponentially close one to the other with respect to ε as ε tends to zero. This is the exponentially small phenomenon that we want to study. In the sequel we shall omit the indices ε,μ when referring to the manifolds $\mathcal{W}^+$ and $\mathcal{W}^-$. Their intersection consists of orbits, which are called homoclinic (or biasymptotic) orbits; we shall see that this intersection is not empty.

[1] The reader is referred to [3], [14] and [28,29] for results on partially hyperbolic tori.

In the particular case of an even perturbation, i.e. when

$$\forall (q,\varphi)\in \mathbb{T}\times\mathbb{T}^d,\quad F(-q,-\varphi)=F(q,\varphi),$$

one checks easily that the symmetry $(q,\varphi,p,I)\mapsto(2\pi-q,-\varphi,p,I)$ sends any orbit onto an orbit and reverses the time-parametrization on it, thus this symmetry exchanges $\mathcal{W}^+$ and $\mathcal{W}^-$. In that situation, one obtains a homoclinic orbit by considering the intersection of $\mathcal{W}^+$ with the $(d+1)$-plane $\{q=\pi,\varphi=0\}$.

Note also that the assumptions (A1) and (A2) are met if for instance F is of the form

$$F(q,\varphi)=(1-\cos q)m(\varphi),$$

where the function m is analytic on $\mathbb{T}^d_{h_0}$.

1.2. Historical remarks [2]

The Hamiltonian function (1.1) is a natural generalization of the example considered by V.I. Arnol'd in his famous note [1]. This system was introduced in [17] (*see* also [18]) with the purpose of studying the speed of Arnol'd diffusion. It is sometimes referred to as the "generalized Arnol'd model". This model was designed to embody the main features of a near-integrable Hamiltonian in the vicinity of a simple resonance; indeed, if ω is non-resonant, the integrable Hamiltonian which is obtained when ε vanishes displays a simple resonance at $(p,I)=(0,0)$. In fact, only the case where the perturbation has the special form $F(q,\varphi)=(1-\cos q)m(\varphi)$ was considered in [17], but the emphasis was already put on the importance of including arbitrarily high harmonics in the perturbation and the Poincaré–Melnikov approximation of the splitting was discussed (see below).

In [11], among other things a "rotator-pendulum model" is studied, with an even trigonometric polynomial of (q,φ) for the perturbation F, but without the assumption (A1). The existence of an invariant hyperbolic torus is proved and "quasiflat" upper bounds are obtained for the splitting of its whiskers by direct perturbative methods. The proof is rather involved and we must say that unfortunately we were not able to follow it in all details. In [12], still for a polynomial perturbation but with a number of harmonics tending to infinity as ε decreases, it is claimed that the same methods lead to results of the kind we are interested in.

The case where $\alpha=0$ is considered in [27] and then in [4]; it may be called the *isochronous* or *linear* case. In that case we can forget about the variables I_j and consider the Hamiltonian vector field associated to (1.1) as a quasi-periodic perturbation of a simple pendulum. More precisely, if $\alpha=0$, one can associate to the Hamiltonian vector field (1.2) a *reduced vector field*

$$\begin{cases}\dot q=p,\\ \dot p=\varepsilon\sin q-\mu\varepsilon\partial_q F(q,\varphi),\\ \dot\varphi_j=\omega_j\end{cases}\tag{1.3}$$

(the original vector field was invariant under the translations $(q,\varphi,p,I)\mapsto(q,\varphi,p,I+\text{constant})$, and (1.3) is indeed its reduction under that group of transformations), or a non-autonomous quasi-periodic second-order differential equation

$$\ddot q=\varepsilon\sin q-\mu\varepsilon\partial_q F(q,\omega_1 t,\dots,\omega_d t).$$

[2] We confine ourselves to the case $d\geqslant 2$, but of course a number of references should be quoted for the two-degree-of-freedom case.

Every solution of (1.2) projects onto a solution of (1.3). Note that the invariant tori which we are interested in project onto *normally* hyperbolic invariant tori for (1.3). However, even if assumptions (A1) and (A2) are satisfied, not all the homoclinic orbits of the reduced vector field are the projection of some homoclinic orbit of the original system. But it is legitimate in that case to concentrate on the reduced system and to consider (1.2) as an auxiliary system, a mere way of putting (1.3) into Hamiltonian form.

In [27] a general averaging theorem is proved and applied, among other things, to a specific example which can be written in the form (1.1) with $d=2$, $\alpha=0$, $\omega=\left(1,\frac{1+\sqrt{5}}{2}\right)$, $\mu=\varepsilon^{-1}$, and $\partial_q F$ even trigonometric polynomial in φ which does not depend on q. The splitting of the invariant manifolds is studied (for the reduced system) and shown to be exponentially small. After a change of variables (one step of averaging), a Poincaré–Melnikov approximation is derived and bounded from above *and from below*; numerical evidence is then produced which indicates that the size of the whole splitting is correctly predicted by that first-order approximation.

The model which is studied in [4] corresponds to $d=2$, $\alpha=0$, $\omega=\left(1,\frac{1+\sqrt{5}}{2}\right)$, $\mu=\varepsilon^p$ ($p>3/2$), and $F(q,\varphi)=m(\varphi)\cos q$ where the function m is analytic in a "strip" $\mathbb{T}_{r_1}\times\mathbb{T}_{r_2}$ but cannot be analytically continued to a larger strip because of some hypothesis on its high harmonics. The torus $\{q=0,p=0,\varphi\in\mathbb{T}^2\}$ is invariant for the corresponding system (1.2), and the splitting of its invariant manifolds is shown to be correctly predicted by the Poincaré–Melnikov approximation whose asymptotics is precisely computed (*see* Section 4 below).

In [25] strong results are stated for the anisochronous case, with an even perturbation, but unfortunately an error has been discovered in that article (a correction is expected).

The present article is strongly related to a joint work with P. Lochak and J.-P. Marco [21], to which the reader is referred for further bibliographical notes.

1.3. The method

It is not so easy to compare the existing methods and results, in particular because each author has his own way of parametrising the invariant manifolds and then of measuring the distance between them. As for us, we shall use particular solutions of the Hamilton–Jacobi equation in order to describe these manifolds, as in the article [26] which was itself inspired by [22].

1.3.1. Let us consider the Hamiltonian system associated to (1.1), under the assumptions (A1) and (A2) of p. 160. For $\mu=0$, the stable and unstable manifolds of the invariant torus $\mathcal{T}=\{(0,\varphi,0,0),\ \varphi\in\mathbb{T}^d\}$ coincide and are given by the separatrix of the pendulum; we find it convenient to write them as

$$\mathcal{W}^-_{|\mu=0}=\left\{(q,\varphi,p,I)\mid q\in]-2\pi,2\pi[,\ \varphi\in\mathbb{T}^d,\ p=2\varepsilon^{1/2}\sin\frac{q}{2},\ I=0\right\},$$
$$\mathcal{W}^+_{|\mu=0}=\left\{(q,\varphi,p,I)\mid q\in]0,4\pi[,\ \varphi\in\mathbb{T}^d,\ p=2\varepsilon^{1/2}\sin\frac{q}{2},\ I=0\right\},$$

distinguishing them quite arbitrarily only by their domain of definition. We should give different names to the tori $\{0\}\times\mathbb{T}^d\times\{0\}\times\{0\}$ and $\{2\pi\}\times\mathbb{T}^d\times\{0\}\times\{0\}$ as well, but we shall not do it; from now on we shall consider that the phase space is $\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}\times\mathbb{R}^d$, which we identify with the cotangent bundle of the *configuration space* $\mathbb{R}\times\mathbb{T}^d$ where the variables (q,φ) live (the cotangent bundle being endowed with its canonical exact symplectic structure). Thus, above a point of the configuration space, covectors are identified with vectors of $\mathbb{R}\times\mathbb{R}^d$. Each of the unperturbed invariant manifolds is an exact Lagrangian graph, i.e. the graph of the differential of

a function defined on a part of $\mathbb{R} \times \mathbb{T}^d$:

$$\mathcal{W}^{\pm}_{|\mu=0} = \mathcal{G}\mathrm{r}(\mathrm{d}S_0) = \big\{(q, \varphi, \partial_q S_0(q,\varphi), \partial_\varphi S_0(q,\varphi)\big\}$$

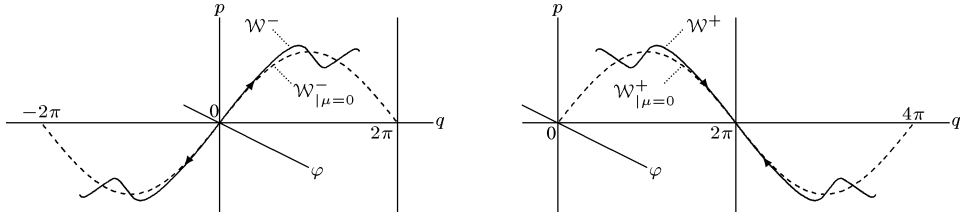
with $S_0(q,\varphi) = S_0(q) = 4\varepsilon^{1/2}\big(\cos\frac{q}{2} - 1\big)$ considered as a function either on $]-2\pi, 2\pi[\times \mathbb{T}^d$ or on $]0, 4\pi[\times \mathbb{T}^d$. We shall represent the perturbed invariant manifolds too as graphs over this space (at least parts of them which do not lie too far from the torus $\mathcal{T}$, i.e. local or "semi-local" stable and unstable manifolds):

$$\mathcal{W}^- = \mathcal{G}\mathrm{r}(\mathrm{d}S^-), \qquad \mathcal{W}^+ = \mathcal{G}\mathrm{r}(\mathrm{d}S^+),$$

where S^- and S^+ are functions on some parts of the configuration space, which depend on (ε, μ) and which will be uniquely determined (up to an additive constant) as the solutions of the Hamilton–Jacobi equation

$$H_{\varepsilon,\mu}\big(q, \varphi, \partial_q S(q,\varphi), \partial_\varphi S(q,\varphi)\big) = 0 \tag{1.4}$$

such that $\mathcal{G}\mathrm{r}(\mathrm{d}S^{\pm})$ contains the torus $\mathcal{T}$. (The right-hand side of (1.4) must vanish since $\mathcal{T}$ itself has zero energy.)



PROPOSITION 1.1. – *For any $q_0 \in]0, 2\pi[$, there exists a positive constant μ_0 such that the Hamilton–Jacobi equation* (1.4) *admits a unique solution $S^- = S^-(q, \varphi; \varepsilon, \mu)$ real-analytic with respect to all its arguments for*

$$q \in]-q_0, q_0[, \quad \varphi \in \mathbb{T}^d, \quad \varepsilon > 0, \quad |\mu| < \mu_0,$$

and such that $S^-_{|\mu=0}$ coincides with S_0 and $S^- - S_0$ vanishes at order 2 on $\{q = 0\}$; and a unique solution $S^+(q, \varphi; \varepsilon, \mu)$ real-analytic with respect to all its arguments for

$$q \in]2\pi - q_0, 2\pi + q_0[, \quad \varphi \in \mathbb{T}^d, \quad \varepsilon > 0, \quad |\mu| < \mu_0,$$

and such that $S^-_{|\mu=0}$ coincides with S_0 and $S^+ - S_0$ vanishes at order 2 on $\{q = 2\pi\}$.

COROLLARY 1.1. – *The invariant torus $\mathcal{T}$ admits stable and unstable manifolds which are locally the graphs of the differentials of the previous functions S^+ and S^- (differentials with respect to the variables q and φ):*

$$\mathcal{W}^{\pm} = \mathcal{G}\mathrm{r}(\mathrm{d}S^{\pm}).$$

Let us define the function

$$\Delta S = S^+ - S^-, \quad \text{for } q \in]2\pi - q_0, q_0[\ (q_0 > \pi),\ \varphi \in \mathbb{T}^d,\ \varepsilon > 0,\ |\mu| < \mu_0.$$

The differential of ΔS is related to the distance between $\mathcal{W}^-$ and $\mathcal{W}^+$: above a point $Q=(q,\varphi)$, the vector between $(Q,\mathrm{d}S^-(Q))$ and $(Q,\mathrm{d}S^+(Q))$ can be identified with $d(\Delta S)(Q)$. The critical points of ΔS are thus projections of homoclinic intersections. [3]

In fact, the Hessian matrix of ΔS at a critical point can be taken as a measure of the splitting of the manifolds at the corresponding homoclinic point. For that reason it is interesting to estimate the first and second-order partial derivatives of the function ΔS (still with respect to the variables q and φ). The function ΔS itself is a difference of Lagrangian actions, which contains an arbitrary additive constant (constant with respect to q and φ but function of ε and μ).

We stated the corollary apart just in order to emphasize the fact that, since the invariant manifolds are represented by functions on the configuration space, our problem is reduced to the study of the function ΔS.

The proposition and its corollary will be a consequence of Proposition 3.1 *and Theorem* 3.1 *below* (we shall even obtain information on the complex extension of the domain of analyticity of S^- and S^+: for instance S^- will be proved to be analytic for complex $q,\varphi,\varepsilon,\mu$, with ε belonging to some sector and $|\mu|$ small enough). Perhaps we should say that we expected the invariant manifolds to be graphs because of their being close (for small $|\mu|$) to $\mathcal{W}^{\pm}_{|\mu=0}$ which is a graph, and to be Lagrangian because of their being asymptotic to $\mathcal{T}$ which is an isotropic submanifold of the phase space (*see* Remark 5.1).

Note that in the isochronous case ($\alpha=0$), it is not true that all the partial derivatives of ΔS necessarily vanish at a point corresponding to a homoclinic orbit of the *reduced* system (1.3): the only condition is that $\partial_q \Delta S$ should vanish at such a point, since only $\partial_q S^{\pm}$ is needed for the description of the stable or unstable manifold of the reduced system (but, due to the conservation of the energy, there is a priori a relation between the other partial derivatives of ΔS at such a point – see the next paragraph).

1.3.2. The geometrical tool of our method is a vector field of the configuration space which acts by zero on the function ΔS. We call it the *characteristic vector field of the pair* (S^-,S^+).

PROPOSITION 1.2. – *Fix any $q_0\in]\pi,2\pi[$ and consider the functions S^- and S^+ of Proposition* 1.1. *The vector field*

$$D=\frac{1}{2}\partial_q(S^++S^-)\frac{\partial}{\partial q}+\left(\omega+\frac{1}{2}\alpha\partial_\varphi(S^++S^-)\right)\cdot\frac{\partial}{\partial\varphi}$$

is defined and analytic for $q\in]2\pi-q_0,q_0[$, $\varphi\in\mathbb{T}^d$, $\varepsilon>0$, $|\mu|<\mu_0$, and the function $\Delta S=S^+-S^-$ satisfies

$$D\cdot\Delta S=0.$$

[3] We may call $\mathcal{W}^-=\mathcal{G}\mathrm{r}(\mathrm{d}S^-)$ *semi-local* unstable manifold in opposition to the local unstable manifold $\mathcal{W}^-_{\mathrm{loc}}$ and to the global unstable manifold. The local manifold is defined by the use of some small enough neighbourhood $\mathcal{V}$ of $\mathcal{T}$ (we can assume in particular that $\mathcal{V}$ is contained in $\mathrm{T}^*(]-q_0,q_0[\times\mathbb{T}^d)$); it consists of all the points in $\mathcal{V}$ whose trajectories are negatively asymptotic to $\mathcal{T}$. The global manifold is the union of the trajectories of the points of $\mathcal{W}^-_{\mathrm{loc}}$, i.e. $\{\phi_H^t(M);\ M\in\mathcal{W}^-_{\mathrm{loc}},\ t\in\mathbb{R}\}$ if we denote by ϕ_H^t the time-t map of the Hamiltonian flow, whereas

$$\mathcal{W}^-=\{\phi_H^t(M);\ M\in\mathcal{W}^-_{\mathrm{loc}},t\in\mathbb{R}\text{ such that }\phi_H^t(M)\in\mathrm{T}^*(]-q_0,q_0[\times\mathbb{T}^d)\}.$$

Analogously we may call $\mathcal{W}^+$ semi-local stable manifold. For that reason the homoclinic points obtained as intersections of $\mathcal{W}^-$ and $\mathcal{W}^+$ may be called "primary" homoclinic points.

Proof. – We present this property in Section 2.1 in a slightly more general context, but it follows from the Hamilton–Jacobi equation by a straightforward computation – take the difference between equation (1.4) for S^+ and equation (1.4) for S^-:

$$0 = \omega \cdot \partial_\varphi S^+ + \frac{1}{2}\alpha(\partial_\varphi S^+)^2 + \frac{1}{2}(\partial_q S^+)^2 + \varepsilon(\cos q - 1) + \mu\varepsilon F,$$
$$0 = \omega \cdot \partial_\varphi S^- + \frac{1}{2}\alpha(\partial_\varphi S^-)^2 + \frac{1}{2}(\partial_q S^-)^2 + \varepsilon(\cos q - 1) + \mu\varepsilon F. \qquad \square$$

Note that for $\mu = 0$ this vector field reduces to the characteristic vector field of the unperturbed invariant manifold:

$$D_{|\mu=0} = \frac{\mathrm{d}S_0}{\mathrm{d}q}\frac{\partial}{\partial q} + \omega \cdot \frac{\partial}{\partial \varphi}.$$

The invariance of the function ΔS under the flow of the vector field D is a simple manifestation of the conservation of energy and of the exact symplectic features of the problem that we have tried to take into account as much as possible. This fact has important consequences for us, since our goal is to study ΔS and we discover now that this function is determined by its restriction to any global section of the configuration space which is transversal to D, and such a section is a torus of dimension d. (All this seems very related to the approach of [5], where a "splitting potential" is introduced which is also a function on $\mathbb{T}^d$, but we have not yet completely elucidated the connection between that recent method and ours.) However, we do not want to fix once for all a particular section, since there is no privileged choice – except maybe in the case of an even perturbation, where D is conjugate to its opposite by the symmetry $(q, \varphi) \mapsto (2\pi - q, \varphi)$ and the section $\{q = \pi\}$ may look more natural since $(\pi, 0)$ is a critical point of ΔS.

Proposition 1.2 (together with a detail from Theorem 3.1 about the dependence of $S^\pm$ on ε and μ) allows us to obtain very easily the following geometrical result, which is a particular case of a theorem by L.H. Eliasson ([9], [5]):

COROLLARY 1.2. – *There exists a positive constant μ_0' such that, for $\varepsilon > 0$ and $|\mu| \leqslant \mu_0'$, the Hamiltonian system associated to $H_{\varepsilon,\mu}$ admits at least $d + 1$ distinct homoclinic orbits.*

Proof. – Let us choose $q_0 \in]\pi, 2\pi[$ and μ_0 so that Proposition 1.1 applies and let us pick any $q_* \in]2\pi - q_0, q_0[$. We shall use the notation $\mathfrak{S} = \{(q_*, \varphi),\ \varphi \in \mathbb{T}^d\}$: this set is a global section of the configuration space and is isomorphic to the torus $\mathbb{T}^d$.

The characteristic vector field D depends analytically on μ and is transversal to the section $\mathfrak{S}$ for $\mu = 0$; thus it is still transversal to that section for μ small enough, as long as the function $\frac{1}{2}\partial_q(S^+ + S^-)$ does not vanish on $\mathfrak{S}$. One can ensure that to be the case for $|\mu| \leqslant \mu_0'$ with a positive number μ_0' which does not depend on ε, since Theorem 3.1 provides bounds for the partial derivatives of $\varepsilon^{-1/2}(S^\pm - S_0) = \mathrm{O}(\mu)$ which do not depend on ε.

Now, for $\varepsilon > 0$ and $|\mu| \leqslant \mu_0'$, if we denote by χ the restriction to $\mathfrak{S}$ of the function ΔS, we observe that any critical point of χ is necessarily a critical point of ΔS (because of the equation $D \cdot \Delta S = 0$, viewed as a relationship between the partial derivatives of ΔS). And χ, being a real-analytic function on a torus of dimension d, admits at least $d + 1$ critical points according to a theorem by Ljusternik and Schnirelman [2]. According to Corollary 1.1, those critical points yield homoclinic orbits. $\square$

The Ljusternik–Schnirelman theorem was already used in [9] to prove the existence of homoclinic orbits, but in a more general context and in a slightly different manner (*see* also [5]).

1.3.3. The analytical tool of our method is already present in [4] and consists in a lemma (Lemma 2.1) from which we deduce that, in order to obtain an exponentially small upper

bound for ΔS and its derivatives, it is sufficient to study the analytic continuation of the flow of D. More precisely, it is sufficient to *straighten* the vector field D, i.e. to conjugate it to its unperturbed form $D_{|\mu=0}$ which has constant coefficients in a proper set of coordinates, namely the coordinates (u,φ) which are defined according to formula (1.7) below.

The straightening of D will be achieved in Proposition 3.3. As a result there exist coordinates (v,θ) of the configuration space, which depend on (ε,μ) and differ from the coordinates (u,φ) only by $\mathrm{O}(\mu)$, such that ΔS can be written in these coordinates as a function $\chi(v,\theta;\varepsilon,\mu)$ which is periodic in the angles θ_j and satisfies $(\partial_v+\varepsilon^{-1/2}\omega\cdot\partial_\theta)\chi=0$. Lemma 2.1 then implies that its Fourier coefficients satisfy, for $0<\rho<\frac{\pi}{2}$ and $0<h<h_0$, the inequalities

$$|\chi_k(v;\varepsilon,\mu)|\leqslant \mathrm{const}\varepsilon^{1/2}|\mu|\exp\bigl(-\rho\varepsilon^{-1/2}|k\cdot\omega|-h|k|\bigr)$$

for $k\in\mathbb{Z}^d\setminus\{0\}$, $v\in\mathbb{R}$, $\varepsilon>0$, $\mu\in[-\mu_0,\mu_0]$. The constant $a(\varepsilon,\mu)$ is nothing but the mean value of χ, which does not depend on v.

In these inequalities, the traditional "small divisors" $|k\cdot\omega|$ do not appear as divisors but as coefficients of $-\varepsilon^{-1/2}$ in the argument of an exponential, hence a difficulty which we call the problem of "small exponents" and which we explain in Section 2.2. In order to overcome it, we shall impose a Diophantine condition on ω.

1.3.4. An advantage of this method is the fact that it deals as much as possible with functions on the configuration space which has dimension $d+1$. For instance the straightening of the vector field D consists in finding a kind of flow-box coordinates in that space; we need not study the Hamiltonian flow in the $(2d+2)$-dimensional phase space outside the invariant manifolds.

We have restricted ourselves to the case where $d\geqslant 2$ (at least two fast frequencies) although the present method would apply as well in the case where $d=1$. In fact, if there is only one fast frequency, the problem is simpler because there are no "small exponents", and for some technical reasons it is easier to solve the Hamilton–Jacobi equation, but the results would require a slightly different presentation in that case (and it would be worthwhile to compare them with the results obtained in [6], [7], or [13] – this will be the subject of some other article).

1.4. General results

The proofs of the statements below are spread over Sections 2, 3, 5 and 6.

1.4.1. The first result claims that the invariant manifolds $\mathcal{W}^-$ and $\mathcal{W}^+$ are exponentially close one to the other: it provides an upper bound for the partial derivatives of ΔS of order 1 or 2. We shall use the notation $|k|=|k_1|+\cdots+|k_d|$ if $k\in\mathbb{Z}^d$.

THEOREM 1.1. – *Consider the Hamiltonian system* (1.1) *with F satisfying the assumptions* (A1) *and* (A2) *of p.* 160 *and ω satisfying the Diophantine condition*

$$\forall k\in\mathbb{Z}^d\setminus\{0\},\quad |k\cdot\omega|\geqslant\gamma|k|^{1-\tau} \tag{1.5}$$

for some fixed positive numbers γ and τ ($\tau\geqslant d$). Denote by ΔS the difference S^+-S^- of the two functions determined in Proposition 1.1 *and by w_* the number*

$$w_*=\bigl(1+(\tau-1)^{-1}\bigr)\biggl((\tau-1)\gamma h_0^{\tau-1}\frac{\pi}{2}\biggr)^{1/\tau}.$$

For any $w \in]0, w_*[$ and for any closed subinterval $[q_1, q_2]$ of $]0, 2\pi[$, there exist positive numbers μ_0 and C such that the inequalities

$$|d(\Delta S)|, \left|d^2(\Delta S)\right| \leqslant C\varepsilon^{1/2}|\mu| \exp\left(-w\varepsilon^{-1/2\tau}\right) \tag{1.6}$$

hold for

$$q \in [q_1, q_2], \quad \varphi \in \mathbb{T}^d, \quad \varepsilon > 0, \quad \mu \in [-\mu_0, \mu_0].$$

Of course the inequalities (1.6) mean that each of the first and second-order derivatives with respect to the variables q or φ_j of ΔS is bounded by the right-hand side. Note that all the variables are required to be real for that exponentially small bound to hold. To quote the words of [17], we could say that "the most important feature in the formula (1.6) is the exponent $-1/2\tau$ of ε". Indeed, in the case where $\tau = d$, this exponent coincides with the one which is involved in the lower bound for the exponentially long time of stability for trajectories starting in the vicinity of a simply resonant surface (*see* [19] for this version of Nekhoroshev theorem with *local exponents* of stability).

In order of importance, the coefficient w inside the exponential of the formula (1.6) comes after the exponent $-1/2\tau$ of ε. We shall see how to let it reach the value w_* in order to obtain a smaller upper bound.

Remark 1.1. – In view of Corollary 1.1, what we are interested in is really $d(\Delta S)$ and $d^2(\Delta S)$. But the function ΔS itself is exponentially close to a constant: under the hypotheses of Theorem 1.1, there exists a real-analytic function $a(\varepsilon, \mu)$ such that, for all $w \in]0, w_*[$ and for all closed subinterval $[q_1, q_2]$ of $]0, 2\pi[$,

$$\begin{aligned} \exists \mu_0,\ C > 0 \text{ such that} \quad & \forall (q, \varphi, \varepsilon, \mu) \in [q_1, q_2] \times \mathbb{T}^d \times \mathbb{R}^{*+} \times [-\mu_0, \mu_0], \\ & |\Delta S(q, \varphi; \varepsilon, \mu) - \mu a(\varepsilon, \mu)| \leqslant C\varepsilon^{1/2}|\mu| \exp\left(-w\varepsilon^{-1/2\tau}\right). \end{aligned}$$

1.4.2. In order to go farther and to obtain a better information on the asymptotics of ΔS with respect to ε, it is natural to try to isolate the first-order approximation ΔS_1 with respect to μ of that function, which is usually called the Poincaré–Melnikov approximation.

DEFINITION 1.1. – The *Poincaré–Melnikov approximation of* ΔS is the function

$$\Delta S_1(q, \varphi; \varepsilon) = \partial_\mu(\Delta S)(q, \varphi; \varepsilon, 0).$$

Thus we can write $\Delta S(q, \varphi; \varepsilon, \mu) = \mu \Delta S_1(q, \varphi; \varepsilon) + \mathrm{O}(\mu^2)$, and our goal is to study that remainder "$\mathrm{O}(\mu^2)$": is it smaller than the Poincaré–Melnikov approximation itself? Of course, in order to provide a true answer to that question, we would need to know how large ΔS_1 is exactly, and this turns out to be a difficult problem. Proposition 1.3 below shows that ΔS_1 can be expressed directly as an integral involving the perturbation function F and bounded from *above* by an exponentially small quantity depending on F, but in the general situation we do not know how to obtain a *lower* bound for ΔS_1 (the problem is more tractable when $d = 2$; *see* Section 4). We shall thus content ourselves with proving that the remainder $\Delta S - \mu \Delta S_1$ is smaller than a quantity which can be compared to the known upper bound of ΔS_1, although this is not completely satisfactory.

Let us define the change of variable

$$q = q_0(u) = 4 \arctan \mathrm{e}^u. \tag{1.7}$$

The variable u is nothing but the time along the separatrix of the pendulum and it will prove very useful in the sequel. In particular, it will be essential to see it as a complex variable, and to try to obtain the largest possible domains of analyticity with respect to it for all the functions defined on the configuration space. We shall put a tilde over the symbol denoting such a function in order to indicate that we have performed the change of variable $q = q_0(u)$.

The function $q_0(u)$ extends analytically to the universal covering of $\mathbb{C} \smallsetminus (\frac{\mathrm{i}\pi}{2} + \mathrm{i}\pi\mathbb{Z})$ with logarithmic singularities only, and it defines a (uniform) analytic $2\pi\mathrm{i}$-periodic function [4] in

$$\mathcal{C} = \mathbb{C} \smallsetminus \left(\left[\frac{\mathrm{i}\pi}{2}, \frac{3\mathrm{i}\pi}{2} \right] + 2\mathrm{i}\pi\mathbb{Z} \right).$$

For $0 < \delta < \pi/2$, we shall denote by $\mathcal{C}_\delta$ a subset of $\mathcal{C}$ which contains $\mathbb{R}$:

$$\mathcal{C}_\delta = \left\{ u \in \mathbb{C} \,\middle|\, \operatorname{dist}\left(u, \left[\frac{\mathrm{i}\pi}{2}, \frac{3\mathrm{i}\pi}{2} \right] + 2\mathrm{i}\pi\mathbb{Z} \right) \geqslant \delta \right\}. \tag{1.8}$$

Because of the assumptions (A1) and (A2), the function

$$\widetilde{F}(u, \varphi) = F\big(q_0(u), \varphi\big)$$

is analytic in $\mathcal{C} \times \mathbb{T}^d_{h_0}$ and for all $\delta, \sigma > 0$ (with $\delta < \pi/2$ and $\sigma < h_0$), there exists a number

$$A = A(\delta, \sigma) \geqslant 1$$

such that

$$\forall (u, \varphi) \in \mathcal{C}_\delta \times \overline{\mathbb{T}}^d_{h_0 - \sigma}, \quad |\widetilde{F}(u, \varphi)| \leqslant A \mathrm{e}^{-2|\,\Re\mathrm{e}\, u|} \tag{1.9}$$

(we have used the notation $\overline{\mathbb{T}}_h = \{\phi \in \mathbb{C}/2\pi\mathbb{Z};\ |\,\Im\mathrm{m}\,\phi| \leqslant h\}$ if $h > 0$). We shall consider this function $A(\cdot, \cdot)$ as a datum of the problem in the same way as the function F itself; it is in fact a manner of measuring the size of F, or the strength of the singularities of $\widetilde{F}$ on the imaginary axis for the variable u and on the boundary of $\overline{\mathbb{T}}^d_{h_0}$ for the variable φ. One may keep in mind the typical example of a function like $A(\delta, \sigma) = \mathrm{const}\delta^{-n}\sigma^{-m}$, with $n, m \in \mathbb{N}^*$, which would correspond to polar singularities (*cf.* the notion of *order of the perturbation along the separatrix* in [7]).

Here is an example taken from [4]: if

$$F(q, \varphi) = (1 - \cos q) \frac{\cos\varphi_1}{\cosh h_0 - \cos\varphi_1} \cdots \frac{\cos\varphi_d}{\cosh h_0 - \cos\varphi_d},$$

one chooses $A(\delta, \sigma) = \mathrm{const}\delta^{-2}\sigma^{-d}$ (observe that $1 - \cos q_0(u) = 2\cosh^{-2} u$ has a second-order pole).

[4] The image of $\mathcal{C}$ by q_0 is the vertical strip $\{q \in \mathbb{C}|\ -\pi < \Re\mathrm{e}\, q < 3\pi\}$ except for the points 0 and 2π which are obtained only at the limit when $\Re\mathrm{e}\, u$ tends to $-\infty$ or $+\infty$ respectively (the singular points $\mathrm{i}\pi/2$ and $3\mathrm{i}\pi/2$ correspond to $\Im\mathrm{m}\, q = +\infty$ and $\Im\mathrm{m}\, q = -\infty$ respectively, and the left and right sides of the cut $]\frac{\mathrm{i}\pi}{2}, \frac{3i\pi}{2}[$ correspond to the vertical boundaries $\Re\mathrm{e}\, q = -\pi$ and $\Re\mathrm{e}\, q = 3\pi$ respectively).

PROPOSITION 1.3. – *The Poincaré–Melnikov approximation can be expressed in the variable u as*

$$\Delta\widetilde{S}_1(u,\varphi;\varepsilon)=\varepsilon^{1/2}\int_{-\infty}^{+\infty}F\big(q_0(u+\zeta),\varphi+\varepsilon^{-1/2}\zeta\omega\big)\,\mathrm{d}\zeta.$$

Assume that ω satisfies the Diophantine condition (1.5) *and use the notation*

$$w_*=\big(1+(\tau-1)^{-1}\big)\left((\tau-1)\gamma h_0^{\tau-1}\frac{\pi}{2}\right)^{1/\tau}.$$

For any closed subinterval $[q_1,q_2]$ of $]0,2\pi[$, there exist positive constants ε_0 and b such that, for $(r_0,r)\in\mathbb{N}\times\mathbb{N}^d$ with $1\leqslant r_0+|r|\leqslant 3$,

$$\big|(\partial_q)^{r_0}(\partial_\varphi)^r\Delta S_1\big|\leqslant bA\big(\varepsilon^{1/2\tau},\varepsilon^{1/2\tau}\big)\varepsilon^{-\frac{r_0+|r|+d-\tau}{2\tau}}\exp\big(-w_*\varepsilon^{-1/2\tau}\big)\tag{1.10}$$

whenever

$$(q,\varphi)\in[q_1,q_2]\times\mathbb{T}^d\quad\textit{and}\quad\varepsilon\in\,]0,\varepsilon_0].$$

Of course, $(\partial_\varphi)^r$ means $(\partial_{\varphi_1})^{r_1}\cdots(\partial_{\varphi_d})^{r_d}$ and $|r|$ means $r_1+\cdots+r_d$. This proposition may be viewed as a refined version of the result which appears already in [17]. Observe that the analyticity width h_0 enters into the upper bound (1.10) through w_*; in fact, if F is a trigonometric polynomial (and thus any h_0 is allowed), one can obtain a much smaller bound, with $-1/2$ in place of $-1/2\tau$ as an exponent for ε inside the exponential.

THEOREM 1.2. – *Under the hypotheses and notations of Theorem* 1.1, *for any closed subinterval $[q_1,q_2]$ of $]0,2\pi[$ there exist positive constants ε_0 and b such that, for $(r_0,r)\in\mathbb{N}\times\mathbb{N}^d$ with $r_0+|r|=1$ or 2,*

$$\begin{aligned}&\big|(\partial_q)^{r_0}(\partial_\varphi)^r(\Delta S-\mu\Delta S_1)(q,\varphi;\varepsilon,\mu)\big|\\&\quad\leqslant b|\mu|^2A\big(\varepsilon^{1/2\tau},\varepsilon^{1/2\tau}\big)^2\varepsilon^{-\frac{r_0+|r|+3d+3-\tau}{2\tau}}\exp\big(-w_*\varepsilon^{-1/2\tau}\big)\end{aligned}\tag{1.11}$$

whenever

$$(q,\varphi)\in[q_1,q_2]\times\mathbb{T}^d,\quad\varepsilon\in\,]0,\varepsilon_0],\quad\mu\in[-\mu_0(\varepsilon),\mu_0(\varepsilon)],$$

with

$$\mu_0(\varepsilon)=b^{-1}A\big(\varepsilon^{1/2\tau},\varepsilon^{1/2\tau}\big)^{-1}\varepsilon^{\frac{d+2}{2\tau}}.$$

Remark 1.2. – As for the case $(r_0,r)=(0,0)$, there exists a real-analytic function $a(\varepsilon,\mu)$ such that $\Delta S-\mu\Delta S_1-\mu^2a$ satisfies the same kind of an inequality:

$$\big|(\Delta S-\mu\Delta S_1)(q,\varphi;\varepsilon,\mu)-\mu^2a(\varepsilon,\mu)\big|\leqslant b|\mu|^2A\big(\varepsilon^{1/2\tau},\varepsilon^{1/2\tau}\big)^2\varepsilon^{-\frac{3d+4-\tau}{2\tau}}\exp\big(-w_*\varepsilon^{-1/2\tau}\big).$$

As we said before, what is not satisfactory is the fact that in general ΔS_1 does not admit a *lower bound* which would be of the same kind as the upper bound (1.10) and thus comparable to the upper bound (1.11) for the remainder $\Delta S-\mu\Delta S_1$. We must insist on this point: *Theorem* 1.2 *does not solve the question of the preponderance of the Poincaré–Melnikov approximation over the remainder.* But in some sense the problem is reduced to the estimation of ΔS_1, because the method that we use is quite adaptable: if an argument is given for obtaining a better upper bound

of ΔS_1, one can also try to incorporate it to the method in order to bound the remainder. This is done in Section 4 in a particular case with $d = 2$, where the coefficient w_* is replaced by an oscillating function of ε.

In inequalities (1.10) or (1.11) we focused on the exponent of ε and the coefficient inside the exponential, but we did not pay much attention to the prefactor (the quantity in front of the exponential) which could be slightly decreased easily.

1.4.3. We believe that our method can be applied to Hamiltonian systems more general than (1.1). The first generalization that we envisage would consist in taking a perturbation F which depends on (q, φ, p, I) (and also ε and μ) and not only on (q, φ), but which still satisfies assumptions analogous to (A1) and (A2). The characteristic vector field would still be defined according to Section 2.1, but the technical details (especially the proof of the analyticity of $S^{\pm}$) would need to be rewritten. One could also consider the case where the invariant torus $\mathcal{T}$ depends on ε and μ, i.e. the case where (A1) is no longer satisfied and some KAM-type result [5] is needed at the beginning to find an invariant hyperbolic torus before to study its whiskers. In fact, one would demand from such a KAM theorem the largest possible domain of analyticity for the parametrization of the torus, and in order to go on with the method, one would put the torus at the origin by a symplectic change of coordinates (or one would exploit its isotropy and look for its invariant manifolds as (non-exact) Lagrangian graphs – the characteristic vector field is still defined because the default of exactness is the same for both manifolds [21]). Finally, it would be interesting to investigate more general models, e.g. including a coupling term between the variables p and I like for a general simple resonance [5].

2. Characteristic vector fields

2.1. Geometrical aspect

In all the Section 2.1 we suppose that M is a differentiable manifold of dimension n (configuration space), T^*M its cotangent bundle (phase space) and $H : \mathrm{T}^*M \to \mathbb{R}$ a Hamiltonian function.

Let $\pi : \mathrm{T}^*M \to M$ be the natural projection. The canonical exact symplectic structure of T^*M is induced by the Liouville form λ, which is defined as follows: for $\alpha \in \mathrm{T}^*M$, $\lambda(\alpha) = \alpha \circ \mathrm{T}_\alpha \pi$, where α is considered as a linear map from $\mathrm{T}_{\pi(\alpha)}M$ to $\mathbb{R}$ and $\mathrm{T}_\alpha \pi$ denotes the linear tangent map to π (from $\mathrm{T}_\alpha(\mathrm{T}^*M)$ to $\mathrm{T}_{\pi(\alpha)}M$). A local system of coordinates $(Q_1, \ldots, Q_n)$ of M induces a local system of coordinates $(Q_1, \ldots, Q_n, P_1, \ldots, P_n)$ of T^*M in which λ takes the usual form $P_1 \mathrm{d}Q_1 + \cdots + P_n \mathrm{d}Q_n$.

The Hamiltonian vector field X_H associated to H is characterized by the property $\mathrm{d}H = -\iota_{X_H} \mathrm{d}\lambda$. For a Lagrangian submanifold of T^*M to be invariant by X_H, it is necessary and sufficient that some constant-energy hypersurface contains it. [6]

[5] This kind of result usually requires non-degenerate torsion, i.e. $\alpha_j \neq 0$ for $j = 1, \ldots, d$. But [20] shows how to deal with the case where $\alpha_1 = \cdots = \alpha_m = 0$ ($m \leqslant d$) and F depends only on $(q, \varphi, p, I_{m+1}, \ldots, I_d)$. The case where $\alpha = 0$ and F does not depend on I is the easiest one, since the normal hyperbolicity which is then present in the reduced system (1.3) provides an invariant torus for (1.3) (without any KAM technique) which can be lifted in an invariant torus for (1.2).

[6] A submanifold $\mathcal{W}$ is said to be Lagrangian if the restriction to $\mathcal{W}$ of the symplectic 2-form $\mathrm{d}\lambda$ vanishes identically and if the dimension of $\mathcal{W}$ is n (maximal dimension for the previous property): at each point of $\mathcal{W}$, the tangent space is equal to its symplectic orthogonal. And if H is constant on $\mathcal{W}$, at each point of $\mathcal{W}$ the Hamiltonian vector field belongs to the symplectic orthogonal of the tangent space to $\mathcal{W}$, thus to the tangent space itself, and conversely.

If α is a one-form of M, we denote by $\operatorname{im}\alpha$ its image in T^*M (wiewing α as a section of π):

$$\operatorname{im}\alpha = \{\alpha(x),\ x \in M\} \subset \mathrm{T}^*M.$$

It is a submanifold of T^*M and π induces a diffeomorphism between $\operatorname{im}\alpha$ and M, by which $\lambda_{|\operatorname{im}\alpha}$ can be identified to α itself (this property characterizes the Liouville form). Thus $\operatorname{im}\alpha$ is Lagrangian if and only if α is closed, and exact Lagrangian if and only α is exact.[7] We are particularly concerned with the latter case; what we have denoted $\mathcal{G}\mathrm{r}(\mathrm{d}S)$ previously should be written $\operatorname{im}\mathrm{d}S$ in this intrinsic formulation.

In the case where α is closed, $\operatorname{im}\alpha$ is invariant by the Hamiltonian vector field X_H if and only if H is constant on it, i.e. if and only if the function $H \circ \alpha$ is constant on M: this is the Hamilton–Jacobi equation. In that situation the *characteristic vector field* of $\operatorname{im}\alpha$ is usually defined to be the vector field of M which corresponds via π to the restriction of the Hamiltonian vector field to $\operatorname{im}\alpha$. It may be written $T\pi \circ X_H \circ \alpha$.

We propose the following generalization of that construction:

DEFINITION 2.1. – Given any pair (α_0, α_1) of 1-forms of M, we call *characteristic vector field of* (α_0, α_1) the vector field of M obtained as

$$D = \int_0^1 D_t\,\mathrm{d}t, \quad \text{where for all } t \in [0,1],\ \ D_t = T\pi \circ X_H \circ \big(\alpha_0 + t(\alpha_1 - \alpha_0)\big).$$

In the exact case, i.e. $\alpha_0 = \mathrm{d}S_0$, $\alpha_1 = \mathrm{d}S_1$, with S_0, S_1 functions on the configuration space, D will be called *characteristic vector field of* (S_0, S_1) as well.

Thus, if we consider a point x of the configuration space, the fiber T^*_xM intersects the manifolds $\operatorname{im}\alpha_0$ and $\operatorname{im}\alpha_1$ at the points $\alpha_0(x)$ and $\alpha_1(x)$, this determines a "vertical" segment $\Sigma(x)$ between both manifolds above x: $\Sigma(x) = \{\alpha_0(x) + t(\alpha_1 - \alpha_0)(x),\ t \in [0,1]\}$, and the characteristic vector field at x is nothing but the projection onto the configuration space of the Hamiltonian vector field averaged along $\Sigma(x)$.

In a local system of coordinates $(Q_1, \ldots, Q_n)$ of M we can write

$$D = \sum_{1 \leqslant j \leqslant n} D_j \frac{\partial}{\partial Q_j}, \quad D_j(Q) = \int_0^1 \frac{\partial H}{\partial P_j}\big(\alpha_t(Q)\big)\mathrm{d}t,$$

setting $\alpha_t = (1-t)\alpha_0 + t\alpha_1$ for $t \in [0,1]$ and using the induced canonical system of coordinates (Q, P) of T^*M. If the Hamiltonian function is quadratic in the momenta $P_1, \ldots, P_n$, the vector field is merely the arithmetic mean of D_0 and D_1, the ordinary characteristic vector fields associated to α_0 and α_1, hence the definition of D in Proposition 1.2 in the case of $H_{\varepsilon,\mu}$, $\alpha_0 = \mathrm{d}S^-$, $\alpha_1 = \mathrm{d}S^+$.

PROPOSITION 2.1. – *Let D be the characteristic vector field of a pair (α_0, α_1) of 1-forms of M. Its action on the difference $\alpha_1 - \alpha_0$ may be described as*

$$\langle \alpha_1 - \alpha_0, D \rangle = H \circ \alpha_1 - H \circ \alpha_0.$$

[7] A submanifold $\mathcal{W}$ is said to be exact Lagrangian if its dimension is n and if the restriction to $\mathcal{W}$ of the Liouville form λ is exact.

Proof. – Fix $x \in M$, denote by H_x the restriction of H to $\mathrm{T}^*_x M$ and consider the path $\alpha_t = \alpha_0(x) + t(\alpha_1 - \alpha_0)(x)$ in $\mathrm{T}^*_x M$:

$$(H \circ \alpha_1 - H \circ \alpha_0)(x) = \int_0^1 \frac{\mathrm{d}(H_x(\alpha_t))}{\mathrm{d}t}\,\mathrm{d}t = \int_0^1 \left\langle \mathrm{d}H_x(\alpha_t), \frac{\mathrm{d}\alpha_t}{\mathrm{d}t} \right\rangle \mathrm{d}t.$$

But if $t \in [0,1]$, using the canonical isomorphism between the vector space $\mathrm{T}^*_x M$ and its tangent space $\mathrm{T}_{\alpha_t}(\mathrm{T}^*_x M)$ at α_t (as well as the dual isomorphism), we can identify $\frac{\mathrm{d}\alpha_t}{\mathrm{d}t} \in \mathrm{T}_{\alpha_t}(\mathrm{T}^*_x M)$ with $(\alpha_1 - \alpha_0)(x) \in \mathrm{T}^*_x M$, and $\mathrm{d}H_x(\alpha_t) \in \mathrm{T}^*_{\alpha_t}(\mathrm{T}^*_x M)$ with $D_t(x) = \mathrm{T}_{\alpha_t}\pi \cdot X_H(\alpha_t) \in \mathrm{T}_x M$ (the last identification may be checked in local coordinates). □

COROLLARY 2.1. – *If S_0 and S_1 are solutions of the Hamilton–Jacobi equation associated to the same energy* (*i.e. if $\mathcal{G}\mathrm{r}(\mathrm{d}S_0)$ and $\mathcal{G}\mathrm{r}(\mathrm{d}S_1)$ are both contained in the same constant-energy hypersurface*), *their characteristic vector field acts by zero on their difference, i.e.*

$$\big\langle \mathrm{d}(S_1 - S_0), D \big\rangle = 0$$

with the notations of the previous definition.

2.2. Analytical aspect

The linear homogeneous partial differential equation $D \cdot \Delta S = \langle \mathrm{d}(\Delta S), D\rangle = 0$ obtained in Proposition 1.2 is thus a particular case of the previous corollary. According to the following lemma, the knowledge of a large domain of analyticity for ΔS and for the flow of D will be of importance to us: in the case where all but one of the coordinates are angular variables and if new coordinates can be found in which D has constant coefficients, complex extension is sufficient to ensure exponential smallness with respect to large frequency-vectors.

We shall use the notations $]\cdot,\cdot[$ and $[\cdot,\cdot]$ for open and closed segments of the complex plane.

LEMMA 2.1. – *Let $\chi(v, \theta_1, \ldots, \theta_d)$ be a function 2π-periodic with respect to the variables θ_j, analytic in $]-\mathrm{i}\rho_0, \mathrm{i}\rho_0[\times \mathbb{T}^d_{h_0}$ for some $\rho_0, h_0 > 0$. Suppose that, for some $\Omega \in \mathbb{R}^d$, it satisfies the partial differential equation*

$$\left(\frac{\partial}{\partial v} + \Omega \cdot \frac{\partial}{\partial \theta}\right)\chi = 0.$$

Then the function χ extends analytically to $\{|\,\Im\mathrm{m}\, v| < \rho_0\} \times \mathbb{T}^d_{h_0}$ and its Fourier coefficients with respect to the angles θ_j satisfy the following inequalities, for all positive $\rho < \rho_0$ and $h < h_0$,

$$\forall k \in \mathbb{Z}^d,\ \forall v \in \mathbb{R}, \quad |\chi_k(v)| \leqslant \mathrm{e}^{-h|k| - \rho|k\cdot\Omega|} \sup_{[-\mathrm{i}\rho, \mathrm{i}\rho] \times \overline{\mathbb{T}}^d_h} |\chi|$$

with the notations $\chi_k(v) = (2\pi)^{-d} \int_{\mathbb{T}^d} \chi(v,\theta)\mathrm{e}^{-\mathrm{i}k\cdot\theta}\mathrm{d}\theta$, $|k| = |k_1| + \cdots + |k_d|$ if $k = (k_1, \ldots, k_d)$, and $\overline{\mathbb{T}}_h = \{\phi \in \mathbb{C}/2\pi\mathbb{Z};\ |\,\Im\mathrm{m}\,\phi| \leqslant h\}$.

A version of this lemma was already given in [4]; it is the quasiperiodic generalization of a lemma by Lazutkin [16] on the Fourier coefficients of a periodic function.

Proof. – The function

$$\Psi(\theta) = \chi(0, \theta)$$

is analytic and 2π-periodic with respect to the variables θ_j. Because of the partial differential equation, we have

$$\chi(v,\theta) = \Psi(\theta - v\Omega),$$

hence a relation between the Fourier coefficients:

$$\chi_k(v) = \Psi_k \mathrm{e}^{-\mathrm{i}vk\cdot\Omega}. \tag{2.1}$$

Let us denote by B the supremum of $|\chi|$ over $[-\mathrm{i}\rho,\mathrm{i}\rho]\times\overline{\mathbb{T}}^d_h$. The Cauchy inequalities

$$\forall v\in[-\mathrm{i}\rho,\mathrm{i}\rho],\quad |\chi_k(v)|\leqslant B\mathrm{e}^{-h|k|},$$

when specialized to $v=\pm i\rho$ (according to the sign of $k\cdot\Omega$), show that

$$\forall k\in\mathbb{Z}^d,\quad |\Psi_k|\leqslant B\mathrm{e}^{-h|k|-\rho|k\cdot\Omega|}.$$

Thus Eq. (2.1) implies that

$$\forall k\in\mathbb{Z}^d,\ \forall v\in\mathbb{C},\quad |\chi_k(v)|\leqslant B\mathrm{e}^{-h|k|-(\rho-|\,\Im\mathrm{m}\,v|)|k\cdot\Omega|}$$

and the Fourier series $\sum\chi_k(v)\mathrm{e}^{\mathrm{i}k\cdot\theta}$ converges for $|\,\Im\mathrm{m}\,v|\leqslant\rho$ and $|\,\Im\mathrm{m}\,\theta_j|<h$. □

In the situation described by the lemma, if moreover $\Omega=\varepsilon^{-1/2}\omega$, each Fourier coefficient of χ of nonzero index is thus exponentially small with respect to ε (as soon as ω is non-resonant). It is natural to try to deduce from that fact a result of exponential smallness for the whole oscillatory part of the function χ. But we are now faced with a difficulty which is typical of the case $d\geqslant 2$ and which we could name the *problem of small exponents*: even if ω is non-resonant, the coefficient $\varepsilon^{-1/2}|k\cdot\omega|$ which enters into the argument of the exponential in the bound of $|\chi_k|$ may reach arbitrary small values as k varies. Yet if we impose a Diophantine condition on ω, there will be a balance between the terms $\varepsilon^{-1/2}|k\cdot\omega|$ and $h|k|$: for the former to be small, the latter must be large. We shall thus recover some exponential smallness for the oscillatory part of χ. This phenomenon was clearly identified in [17] or [18], and then in [27], [4] and [25].

COROLLARY 2.2. – *Let $\chi(v,\theta,\varepsilon)$ be analytic for $(v,\theta)\in]-\mathrm{i}\rho_0,\mathrm{i}\rho_0[\times\mathbb{T}^d_{h_0}$ and $\varepsilon>0$. Suppose that there exists $\omega\in\mathbb{R}^d$ satisfying the Diophantine condition*

$$\forall k\in\mathbb{Z}^d\smallsetminus\{0\},\quad |k\cdot\omega|\geqslant\gamma|k|^{1-\tau}$$

for some fixed positive numbers γ and τ ($\tau\geqslant d$), and such that χ is a solution of the partial differential equation

$$\left(\frac{\partial}{\partial v}+\varepsilon^{-1/2}\omega\cdot\frac{\partial}{\partial\theta}\right)\chi=0.$$

For all positive $\rho<\rho_0$ and $h<h_0$, we shall use the notations

$$B(\rho,h,\varepsilon)=\sup_{(v,\theta)\in[-\mathrm{i}\rho,\mathrm{i}\rho]\times\overline{\mathbb{T}}^d_h}|\chi(v,\theta,\varepsilon)|\quad\textit{and}\quad \nu=\big(1+(\tau-1)^{-1}\big)\big((\tau-1)\gamma\big)^{1/\tau}.$$

The function χ extends analytically to $\{|\,\Im\mathrm{m}\,v|<\rho_0\}\times\mathbb{T}^d_{h_0}\times\{\varepsilon>0\}$, its mean value on the torus $a(\varepsilon)=(2\pi)^{-d}\int_{\mathbb{T}^d}\chi(v,\theta,\varepsilon)\,\mathrm{d}\theta$ does not depend on the variable v and

(i) *for all* $\delta \in]0,\rho_0[$, $\sigma \in]0,\frac{h_0}{2}[$, $\varepsilon > 0$, $(v,\theta) \in \mathbb{R}\times\mathbb{T}^d$,

$$|\chi(v,\theta,\varepsilon) - a(\varepsilon)| \leqslant \frac{c}{\sigma^d} B(\rho_0-\delta, h_0-\sigma, \varepsilon) \exp\big(-w(\delta,\sigma)\varepsilon^{-1/2\tau}\big),$$

where $w(\delta,\sigma) = \nu(\rho_0-\delta)^{1/\tau}(h_0-2\sigma)^{(\tau-1)/\tau}$ *and* $c>0$ *depends only on the dimension* d*;*

(ii) *there exists a positive number* ε_0 *which depends only on* ρ_0 *and* h_0 *such that, for* $\delta_0,\sigma_0 \in]0,10]$ *and* $\varepsilon \in]0,\varepsilon_0]$, $(v,\theta)\in\mathbb{R}\times\mathbb{T}^d$,

$$\big|\chi(v,\theta,\varepsilon) - a(\varepsilon)\big| \leqslant c'\varepsilon^{-d/2\tau} B\big(\rho_0-\delta_0\varepsilon^{1/2\tau}, h_0-\sigma_0\varepsilon^{1/2\tau},\varepsilon\big)\exp\big(-w_*\varepsilon^{-1/2\tau}\big),$$

where $w_* = \nu\rho_0^{1/\tau}h_0^{(\tau-1)/\tau}$ *and* $c'>0$ *depends only on* $d,\ \tau,\ \gamma,\ \rho_0,\ h_0$.

Remark 2.1. – Unfortunately this result is not optimal. We focused on the exponent $-1/2\tau$ of ε and the coefficient w_* inside the exponential, but we do not know how to improve them under general hypotheses – in Section 4 we shall see how to replace w_* by something larger in a particular situation with $d=2$. On the other hand the prefactor $c'\varepsilon^{-d/2\tau}B(\cdots)$ could be slightly decreased by an appropriate modification of the proof below.

Proof. – The fact that the function $a(\varepsilon)$ does not depend on v is an obvious consequence of the partial differential equation.

(i) Let $\rho=\rho_0-\delta$, $h=h_0-\sigma$ and $h_1=h_0-2\sigma$. For each $\varepsilon>0$, we obtain from Lemma 2.1 the following bounds for the Fourier coefficients of χ:

$$\begin{aligned}\forall k\in\mathbb{Z}^d,\ \forall v\in\mathbb{R},\quad |\chi_k(v,\varepsilon)| &\leqslant B(\rho,h,\varepsilon)\mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}\\ &\leqslant B(\rho,h,\varepsilon)\mathrm{e}^{-\sigma|k|}\mathrm{e}^{-h_1|k|-\gamma\rho\varepsilon^{-1/2}|k|^{1-\tau}}\end{aligned}$$

(we have used the Diophantine condition in the second inequality). On one hand, one checks easily that

$$\forall x,y>0,\quad y+\gamma x y^{1-\tau} \geqslant \nu x^{1/\tau},$$

so, if $k\in\mathbb{Z}^d\setminus\{0\}$,

$$h_1|k| + \rho\varepsilon^{-1/2}|k\cdot\omega| \geqslant w(\delta,\sigma)\varepsilon^{-1/2\tau}.$$

On the other hand,

$$\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \mathrm{e}^{-\sigma|k|} = -1 + \coth^d\left(\frac{\sigma}{2}\right) < \frac{c}{\sigma^d}$$

with a positive constant c depending only on the dimension d.

Thus, if $v\in\mathbb{R}$, $\theta\in\mathbb{T}^d$ and $\varepsilon>0$,

$$\big|\chi(v,\theta,\varepsilon)-a(\varepsilon)\big| = \left|\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \chi_k(v,\varepsilon)\mathrm{e}^{\mathrm{i}k\cdot\theta}\right| \leqslant \frac{c}{\sigma^d}B(\rho,h,\varepsilon)\exp\big(-w(\delta,\sigma)\varepsilon^{-1/2\tau}\big).$$

(ii) With the choice $(\delta,\sigma)=(\delta_0\varepsilon^{1/2\tau},\sigma_0\varepsilon^{1/2\tau})$, we get

$$w(\delta,\sigma)\varepsilon^{-1/2\tau} \geqslant w_*\varepsilon^{-1/2\tau} - c'',\quad 0<\varepsilon<\varepsilon_0,$$

where ε_0 depends only on ρ_0,h_0 and c'' depends only on ν,ρ_0,h_0 (because $w(\delta,\sigma) = w_* + \mathrm{O}(\delta,\sigma)$). It is thus sufficient to take $c'=c\mathrm{e}^{c''}$. □

Remark 2.2. – There is no necessity of confining the variables v and θ to the real domain: we could have worked in a set of the form $\{|\,\Im\mathrm{m}\, v| \leqslant \rho'\} \times \overline{\mathbb{T}}^d_{h'}$ with $\rho' < \rho$ and $h' < h$, but then the coefficients $w(\delta, \sigma)$ and w_* would have been smaller, whereas we were interested in the largest possible coefficients (even if they were not optimal). On the other hand we could have replaced $\varepsilon^{-d/2\tau}$ in front of the upper bound in (ii) by a smaller term, but we prefered a simpler result.

3. Exponential closeness of the invariant manifolds

3.1. Hamilton–Jacobi algorithm

Let us return to our Hamiltonian (1.1). The notations will be slightly simplified by a rescaling of time and variables (the time t being multiplied by $\varepsilon^{1/2}$, and the action-like variables p and I divided by the same factor): it is equivalent to study the Hamiltonian system generated by

$$H_{z,\mu}(q, \varphi, p, I) = z\omega \cdot I + \frac{1}{2}\alpha I^2 + \frac{1}{2}p^2 + \cos q - 1 + \mu F(q, \varphi), \tag{3.1}$$

with a large frequency-vector $z\omega$, where

$$z = \varepsilon^{-1/2}.$$

The unperturbed solution of the Hamilton–Jacobi equation is now

$$S_0(q, \varphi) = S_0(q) = 4\left(\cos\frac{q}{2} - 1\right).$$

After having determined and studied particular solutions of the Hamilton–Jacobi equation for the system (3.1), we shall have to multiply them by $\varepsilon^{1/2}$ in order to come back to the original system (1.1). Still, we shall use the same notation $S^{\pm}$ before and after this rescaling.

Our first task is to prove Proposition 1.1. To begin with we shall see how the Taylor expansions of S^- and S^+ with respect to μ are determined: we shall work with formal series in μ whose coefficients are functions of (q, φ) but also on ε, but the dependence on ε will be understood.

PROPOSITION 3.1. – *Fix $\omega \in \mathbb{R}^d$ and suppose that the function F satisfies the assumptions* (A1) *and* (A2)*. For all $z > 0$, there exists a unique sequence $S^-_1, S^-_2, \ldots$ of real-analytic functions of (q, φ) defined for q close to 0 and $\varphi \in \mathbb{T}^d$, vanishing at order 2 on $\{q = 0\}$, and such that*

$$S^- = S_0(q) + \sum_{n \geqslant 1} \mu^n S^-_n(q, \varphi)$$

satisfies formally the Hamilton–Jacobi equation $H_{z,\mu}(q, \varphi, \partial_q S, \partial_\varphi S) = 0$. In fact, these functions extend to $]-2\pi, 2\pi[\times \mathbb{T}^d$ and depend analytically on $z = \varepsilon^{-1/2}$.

For all $z > 0$, there exists a unique sequence $S^+_1, S^+_2, \ldots$ of real-analytic functions of (q, φ) defined for q close to 2π and $\varphi \in \mathbb{T}^d$, vanishing at order 2 on $\{q = 2\pi\}$, and such that

$$S^+ = S_0(q) + \sum_{n \geqslant 1} \mu^n S^+_n(q, \varphi)$$

satisfies formally the Hamilton–Jacobi equation. In fact, these functions extend to $]0, 4\pi[\times \mathbb{T}^d$ and depend analytically on $z = \varepsilon^{-1/2}$.

Proof. – Of course the two parts of the proposition are analogous, and we shall content ourselves with the proof of the first one. Let us denote by $\mathcal{B}^-$ the space of all analytic functions vanishing at order 2 at $q=0$:

$$\mathcal{B}^- = \big\{U(q,\varphi) \text{ analytic for } q \text{ close to } 0 \text{ and } \varphi\in\mathbb{T}^d, \text{ with } U(0,\varphi)=\partial_q U(0,\varphi)=0\big\},$$

and use the notation $D_0 = \frac{\mathrm{d}S_0}{\mathrm{d}q}\frac{\partial}{\partial q} + z\omega\cdot\frac{\partial}{\partial\varphi}$. Let us call

$$T=\sum_{n\geqslant 0}\mu^n S^-_{n+1}(q,\varphi)$$

the formal expansion that we are looking for: $S_0+\mu T$ is solution of the Hamilton–Jacobi equation if and only if

$$D_0 T = -F(q,\varphi) - \frac{1}{2}\mu\big[\alpha(\partial_\varphi T)^2 + (\partial_q T)^2\big].$$

We observe that F belongs to $\mathcal{B}^-$. Thus, expanding the equation in powers of μ, we can determine inductively the functions S^-_{n+1} and conclude the proof by applying a simple lemma (whose proof is left to the reader):

LEMMA 3.1. – *The operator D_0 induces an automorphism of $\mathcal{B}^-$. The change of variable $w=\tan\frac{q}{4}$ gives to it the form $D_0 = w\frac{\partial}{\partial w} + z\omega\cdot\frac{\partial}{\partial\varphi}$ and allows to express its inverse $E^- = (D_{0|\mathcal{B}^-})^{-1}$ as*

$$(E^-U)(w,\varphi) = -\int_0^{-\infty} U\big(w\mathrm{e}^\zeta, \varphi+z\zeta\omega\big)\mathrm{d}\zeta.$$

Note that, if U depends analytically on z for $z>0$, this is also the case of E^-U: indeed, if z is allowed to move in a sector which contains $\mathbb{R}^{*+}$, we still can change the half-line of integration $[0,-\infty[$ into $[0,-z^{-1}\infty[$ in order to keep $z\zeta$ real and the new formula will provide the analytic continuation of E^-U with respect to z. □

Here are the induction formulas that we obtain:

$$S^-_1 = -E^-F,$$
$$S^-_{n+1} = -\frac{1}{2}E^-\Bigg(\sum_{n_1+n_2=n-1}\big[\alpha\partial_\varphi S_{n_1+1}\cdot\partial_\varphi S_{n_2+1} + \partial_q S_{n_1+1}\partial_q S_{n_2+1}\big]\Bigg), \quad n\geqslant 1.$$

They define what we call the Hamilton–Jacobi algorithm.

DEFINITION 3.1. – The vector field (or differential operator) of the configuration space

$$D_0 = \frac{\mathrm{d}S_0}{\mathrm{d}q}\frac{\partial}{\partial q} + z\omega\cdot\frac{\partial}{\partial\varphi}$$

will be called the *unperturbed characteristic vector field.* Note that after the change of variable (1.7) it reduces to

$$\widetilde{D}_0 = \frac{\partial}{\partial u} + z\omega\cdot\frac{\partial}{\partial\varphi}.$$

Proof of Proposition 1.3 *of p.* 169. – It follows from the proof of Proposition 3.1 that $S_1^- = -E^- F$ and similarly $S_1^+ = -E^+ F$, which means

$$\widetilde{S}_1^{\pm}(u,\varphi) = S_1^{\pm}\big(q_0(u),\varphi\big) = \int_0^{\pm\infty} \widetilde{F}(u+\zeta,\varphi+z\zeta\omega)\,\mathrm{d}\zeta$$

when one uses the variable u. The difference is the familiar Poincaré–Melnikov integral, as announced in Proposition 1.3. Let us prove the inequalities (1.10) for the function $\Delta\widetilde{S}_1(u,\varphi;\varepsilon) = \widetilde{S}_1^+ - \widetilde{S}_1^-$.

This function is obviously analytic in $\big\{|\,\Im\mathrm{m}\,u| < \frac{\pi}{2}\big\} \times \mathbb{T}^d_{h_0} \times \mathbb{R}^{*+}$. Using the function A to measure the size of F according to the formula (1.9) of p. 168, we obtain easily the following bounds: if $(r_0, r) \in \mathbb{N} \times \mathbb{N}^d$ with $r_0 + |r| \leqslant 3$,

$$\big|(\partial_u)^{r_0}(\partial_\varphi)^r \Delta\widetilde{S}_1\big| \leqslant 24A\left(\frac{\delta}{2},\frac{\sigma}{2}\right)\delta^{-r_0}\sigma^{-|r|} \quad \text{in } \left\{|\,\Im\mathrm{m}\,u| \leqslant \frac{\pi}{2}-\delta\right\} \times \overline{\mathbb{T}}^d_{h_0-\sigma} \times \mathbb{R}^{*+}, \tag{3.2}$$

where $(\partial_\varphi)^r = (\partial_{\varphi_1})^{r_1}\cdots(\partial_{\varphi_d})^{r_d}$, $|r| = r_1 + \cdots + r_d$. Since $\widetilde{D}_0 \cdot (\Delta\widetilde{S}_1) = 0$, we can apply Corollary 2.2 to the function $\Delta\widetilde{S}_1$ or to its partial derivatives (because $\widetilde{D}_0$ has constant coefficients); if we assume $1 \leqslant r_0 + |r|$, there is no mean value to substract in the results of exponential smallness we get: the first part of the corollary yields

$$\big|(\partial_u)^{r_0}(\partial_\varphi)^r \Delta\widetilde{S}_1\big| \leqslant 24cA\left(\frac{\delta}{2},\frac{\sigma}{2}\right)\delta^{-r_0}\sigma^{-d-|r|}\exp\big(-w\varepsilon^{-1/2\tau}\big) \quad \text{in } \mathbb{R}\times\mathbb{T}^d\times\mathbb{R}^{*+}, \tag{3.3}$$

with $w = (1+(\tau-1)^{-1})((\tau-1)\gamma)^{1/\tau}(\frac{\pi}{2}-\delta)^{1/\tau}(h_0-2\sigma)^{(\tau-1)/\tau}$, and the second part (with $\delta_0 = \sigma_0 = 2$) yields

$$\big|(\partial_u)^{r_0}(\partial_\varphi)^r \Delta\widetilde{S}_1\big| \leqslant 24c'A\big(\varepsilon^{1/2\tau},\varepsilon^{1/2\tau}\big)\varepsilon^{-\frac{r_0+|r|+d}{2\tau}}\exp\big(-w_*\varepsilon^{-1/2\tau}\big),$$

for $(u,\varphi) \in \mathbb{R}\times\mathbb{T}^d$ and $z = \varepsilon^{-1/2} > 0$ large enough, with

$$w_* = \big(1+(\tau-1)^{-1}\big)\big((\tau-1)\gamma h_0^{\tau-1}\big)^{1/\tau}\left(\frac{\pi}{2}\right)^{1/\tau}.$$

This allows us to end the proof of Proposition 1.3, since the inverse change of variable $q \mapsto u$ has its derivatives bounded in any closed subinterval $[q_1,q_2]$ of $]0,2\pi[$. We just have to multiply by $\varepsilon^{1/2}$ because of the rescaling from (3.1) to (1.1). □

Remark 3.1. – *Link with parametric resurgence and Gevrey properties.* In [26] the operator $\widetilde{D}_0$ was studied in the case where $d = 1$ which is much simpler, and it was shown that its inversion led to divergent series in the parameter ε. More precisely, the obtained series were expanded in powers of $\varepsilon^{1/2}$ and were *resurgent* in $z = \varepsilon^{-1/2}$. Here we can at least show that Gevrey properties take place in the variable $x = z^{1/\tau} = \varepsilon^{-1/2\tau}$, in the sense that formel Borel transforms with respect to x will be convergent.

Since this remark is a little far from our topic, we shall only give a statement without proving it. We recall that the formal Borel transform with respect to x is defined by the formula:

$$\chi(x)=\sum_N \chi^{[N]}x^{-N} \;\mapsto\; \widehat{\chi}(\xi)=\sum_N \chi^{[N]}\frac{\xi^{N-1}}{\Gamma(N)};$$

in our situation the index N will run through $\tau\mathbb{N}^*$ and $\xi\widehat{\chi}(\xi)$ will converge to a regular holomorphic germ in the variable ξ^τ, which means that $\widehat{\chi}(\xi)$ will be analytic with ramification at the origin. Thus we may consider that $\xi\in\mathbb{C}_\bullet$, the Riemann surface of the logarithm. For simplicity we shall consider the case where the variable u moves only inside disks.

PROPOSITION 3.2. – *Let $\rho_0,h_0>0$ and denote by $\mathfrak{D}_{\rho_0}$ the open disk of centre 0 and radius ρ_0. Suppose that a function $\psi(u,\varphi)$ is analytic in $\mathfrak{D}_{\rho_0}\times\mathbb{T}^d_{h_0}$ and extends continuously to the closure of that domain, and suppose that $\omega\in\mathbb{R}^d$ satisfies the Diophantine condition*

$$\forall k\in\mathbb{Z}^d\smallsetminus\{0\},\quad |k\cdot\omega|\geqslant\gamma|k|^{1-\tau}.$$

The partial differential equation $\widetilde{D}_0\chi=\psi$ admits a unique formal solution of the type

$$\chi=\sum_{n\geqslant 1}\chi^{[n]}(u,\varphi)z^{-n},$$

and its formal Borel transform with respect to $x=z^{1/\tau}$ converges and defines a function $\widehat{\chi}(u,\varphi,\xi)$ analytic in

$$\mathfrak{E}=\big\{(u,\varphi,\xi)\in\mathfrak{D}_{\rho_0}\times\mathbb{T}^d_{h_0}\times\mathbb{C}_\bullet;\;\; |\xi|\leqslant W(u,\varphi)\big\}$$

with $W(u,\varphi)=\gamma^{1/\tau}(\rho_0-|u|)^{1/\tau}(h_0-\max_j|\Im\mathrm{m}\,\varphi_j|)^{(\tau-1)/\tau}$.

In fact the formal Fourier coefficients $\chi_k(u;z)$ of χ have a Borel transform with respect to z which is convergent and can be expressed as $-\frac{1}{\mathrm{i}k\cdot\omega}\psi_k\big(u-\frac{\zeta}{\mathrm{i}k\cdot\omega}\big)$, and this leads to parametric resurgence (with respect to z), e.g. if the data ψ_k are meromorphic functions, but the associated Fourier series may be divergent (because the singularities in the ζ-plane may accumulate the origin). One could present things using the apparatus of *accelero-summation* [8]: the deceleration operator from z to $x=z^{1/\tau}$ yields entire functions $\widehat{\chi}_k(u,\xi)$ as individual Fourier coefficients and restores the convergence of the whole Fourier series. Moreover, this operator should allow one to express these Borel transforms with respect to x, $\widehat{\chi}_k(u,\xi)$, as integrals involving the Borel transforms of the χ_k with respect to z and some kernel. This could lead to a better knowledge of the analytic continuation of $\widehat{\chi}$ with respect to ξ.

Thinking of the Hamilton–Jacobi equation and its particular solutions S^- and S^+, we suspect the existence of a formal solution (formal power series in $\varepsilon^{1/2}$) to which both S^- and S^+ are Gevrey asymptotic (Gevrey-1 in the variable $x=\varepsilon^{-1/2\tau}$) and which must be considered as resurgent in the variable $z=\varepsilon^{-1/2}$ (even if the Borel transform with respect to z does not converge as a Fourier series). This could shed another light on their exponential proximity. (Notice that a relationship between Gevrey properties and exponential smallness was recently studied in [23].)

3.2. Analyticity of the solutions

Let us now define complex domains in the u-plane, $\mathcal{D}^-_{u_1,\delta}$ and $\mathcal{D}^+_{u_1,\delta}$, in which the Taylor expansions with respect to μ of the solutions $\widetilde{S}^-$ or $\widetilde{S}^+$ will be proved to converge. For $u_1\in\mathbb{R}$

and $0<\delta<\pi/2$, we consider all the open sectors

$$\big\{u=-u_1+\xi \mathrm{e}^{\mathrm{i}\beta};\ \xi>0,\ \beta\in\,]-\beta_1,\beta_1[\big\}$$

for $\beta_1\in\,]0,\pi/2]$, and select among them the largest one which is contained in $\mathcal{C}_\delta$: we call $\mathcal{D}^+_{u_1,\delta}$ this sector; if $u_1>-\delta$, its vertex is $-u_1$ and its aperture $2\beta_1$ is determined by the equation

$$\frac{\pi}{2}\cos\beta_1=\delta+u_1\sin\beta_1,$$

but if $u_1\leqslant-\delta$, it is the open half-plane $\{\Re\mathrm{e}\,u>-u_1\}$ and $\beta_1=\pi/2$.

Similarly, $\mathcal{D}^-_{u_1,\delta}$ is the largest of the sectors

$$\big\{u=u_1-\xi \mathrm{e}^{\mathrm{i}\beta};\ \xi>0,\ \beta\in\,]-\beta_1,\beta_1[\big\},\quad \beta_1\in\,]0,\pi/2],$$

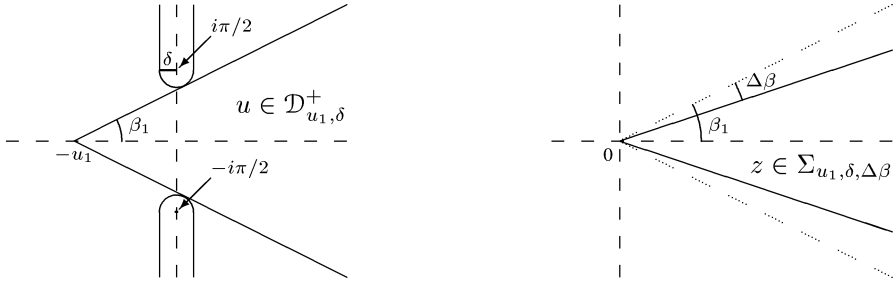
which is contained in $\mathcal{C}_\delta$; it is the opposite sector:

$$\mathcal{D}^-_{u_1,\delta}=\big\{-u;\ u\in\mathcal{D}^+_{u_1,\delta}\big\}.$$

We define also a complex domain for the variable z: for $\Delta\beta\in\,]0,\pi/2[$, with moreover $\Delta\beta<\arctan\frac{\pi}{2u_1}$ if $u_1>0$,

$$\Sigma_{u_1,\delta,\Delta\beta}:=\big\{z=\xi \mathrm{e}^{\mathrm{i}\beta};\ \xi>0,\ \beta\in\,]-\beta_1+\Delta\beta,\ \beta_1-\Delta\beta[\big\},$$

where the half-aperture $\beta_1\in\,]0,\pi/2]$ of the sectors $\mathcal{D}^-_{u_1,\delta}$ and $\mathcal{D}^+_{u_1,\delta}$ is supposed to be strictly larger than $\Delta\beta$; this will be the case if δ is small enough with respect to u_1 and $\Delta\beta$.



THEOREM 3.1. – *For all $u_1\in\mathbb{R}$ and $\Delta\beta\in\,]0,\pi/2[$ with $\Delta\beta<\arctan\frac{\pi}{2u_1}$ if $u_1>0$, and for all small enough $\delta,\sigma>0$, there exist positive numbers μ_1 and $\{B_\partial\}$ such that the series of Proposition* 3.1 *expressed in the variable u*

$$\widetilde{S}^\pm=\widetilde{S}_0(u)+\sum_{n\geqslant1}\mu^n\widetilde{S}^\pm_n(u,\varphi;z)$$

converges to a function $\widetilde{S}^\pm(u,\varphi;z,\mu)$ analytic with respect to all its arguments for

$$u\in\mathcal{D}^\pm_{u_1,\delta},\quad \varphi\in\mathbb{T}^d_{h_0-\sigma},\quad z=\varepsilon^{-1/2}\in\Sigma_{u_1,\delta,\Delta\beta},\quad |\mu|<\mu_1,$$

satisfying in the closure of that domain the inequalities

$$\big|\partial\big(\widetilde{S}^\pm-\widetilde{S}_0-\mu\widetilde{S}^\pm_1\big)\big|\leqslant B_\partial|\mu|^2\mathrm{e}^{-(\pm2\,\Re\mathrm{e}\,u)},$$

where ∂ stands for one of the operators $(\partial_u)^{r_0}(\partial_\varphi)^r$ with $(r_0,r)\in\mathbb{N}\times\mathbb{N}^d$, $r_0+|r|\leqslant 2$.
More precisely, if $\delta\leqslant\sigma$, one can take $\mu_1=b_1^{-1}A(\delta/2,\sigma/2)^{-1}\delta\sigma^{d+1}$ and

$$B_\partial=\begin{cases} b_1A(\delta/2,\sigma/2)^2\delta^{-2-\sup\{1,r_0\}}\sigma^{-2d-1} & \text{if } |r|=0,\\ b_1A(\delta/2,\sigma/2)^2\delta^{-2}\sigma^{-2d-|r|} & \text{if } |r|\geqslant 1,\end{cases}$$

where the positive number b_1 depends only on u_1 and $\Delta\beta$, and the function A measures the size of F according to the formula (1.9).

Of course, $(\partial_\varphi)^r$ means $(\partial_{\varphi_1})^{r_1}\cdots(\partial_{\varphi_d})^{r_d}$ and $|r|=r_1+\cdots+r_d$. *The proof of this theorem is delayed to Section* 5, because it involves some technicalities although it is not difficult: appropriate Banach spaces are defined and the ordinary fixed-point theorem is applied. The key-point lies in Lemma 5.2 which asserts the boundedness of the inverse to the operator D_0 (this inverse is E^+ or E^-, depending on whether one studies the stable or the unstable manifold, with the notation of Lemma 3.1; in other words the "homological" equation – or rather the linearized equation – which is $\widetilde{D}_0\chi=\psi$ in the variables (u,φ), can be solved by $\chi=E^\pm\psi$, where $E^\pm$ is a bounded operator of the suitable Banach space: no small divisor appears because of the presence of the variable u beside the angular variables φ_j).

Proof of Proposition 1.1 *and of its corollary.* – Proposition 1.1 is indeed a consequence of what has been done up to now: if we consider S^- for instance, by letting δ and u_1 vary we see that the variable u can reach (provided that $|\mu|$ is kept small enough) any compact subset of

$$\mathcal{D}^-=\bigcup_{\delta>0,u_1\in\mathbb{R}}\mathcal{D}^-_{u_1,\delta}=\{u\in\mathbb{C}\mid \Re\mathrm{e}\,u<0 \text{ or } |\Im\mathrm{m}\,u|<\pi/2\}$$

whose image by q_0 is $\{-\pi<\Re\mathrm{e}\,q<2\pi\}\setminus\{0\}$; the 2πi-periodicity in u of $\widetilde{S}^-$ and its exponential decrease for $\Re\mathrm{e}\,u$ tending to $-\infty$ ensure that it defines a function S^- analytic for q belonging to any compact subset of $\{-\pi<\Re\mathrm{e}\,q<2\pi\}$ provided that $|\mu|$ is small enough. We would obtain any compact subset of $\{-2\pi<\Re\mathrm{e}\,q<\pi\}$ simply by repeating the previous work with the change of variable $q=-2\pi+4\arctan\mathrm{e}^{-u}$ instead of $q=q_0(u)$.

Now, the manifold $\mathcal{W}^-=\mathcal{G}\mathrm{r}(\mathrm{d}S^-)$ is Lagrangian and contained in the zero-energy level of the Hamiltonian $H_{z,\mu}$ (since S^- satisfies the Hamilton–Jacobi equation), it is thus invariant by the Hamiltonian flow. This manifold contains the torus $\mathcal{T}$ because $\mathrm{d}S^-$ vanishes for $q=0$ – it is in fact its unstable manifold; more information on the dynamics on it is given in Section 5. □

3.3. Exponentially small upper bounds

Theorems 1.1 and 1.2 will derive from the following result. We use here the variable u and we recall that the solutions of the Hamilton–Jacobi equation for system (3.1) must be multiplied by $\varepsilon^{1/2}$ in order to yield solutions corresponding to (1.1).

THEOREM 3.2. – *Consider the Hamiltonian system* (3.1) *with F satisfying the assumptions* (A1) *and* (A2)*, and ω satisfying the Diophantine condition*

$$\forall k\in\mathbb{Z}^d\setminus\{0\},\quad |k\cdot\omega|\geqslant\gamma|k|^{1-\tau}$$

for some fixed positive numbers γ and τ ($\tau\geqslant d$). Denote by $\Delta\widetilde{S}$ the difference $\widetilde{S}^+-\widetilde{S}^-$ of the two functions determined above.

For all $u_0 > 0$ and for all small enough $\delta, \sigma > 0$, there exist a real-analytic function $a(\varepsilon, \mu)$ and positive numbers $\mu_0, \{C_\partial\}$ such that, if we use the notation

$$w = \big(1 + (\tau-1)^{-1}\big)((\tau-1)\gamma)^{1/\tau}\left(\frac{\pi}{2} - \delta\right)^{1/\tau}(h_0 - 2\sigma)^{(\tau-1)/\tau}, \tag{3.4}$$

the inequalities

$$\big|\partial\big((\Delta\widetilde{S} - \mu\Delta\widetilde{S}_1)(u, \varphi; \varepsilon^{-1/2}, \mu) - \mu^2 a(\varepsilon, \mu)\big)\big| \leqslant C_\partial |\mu|^2 \exp\big(-w\varepsilon^{-1/2\tau}\big) \tag{3.5}$$

hold for

$$u \in [-u_0, u_0], \quad \varphi \in \mathbb{T}^d, \quad \varepsilon > 0, \quad \mu \in [-\mu_0, \mu_0]$$

and $\partial = (\partial_u)^{r_0}(\partial_\varphi)^r$ with $r_0 + |r| \leqslant 2$.

If $2\delta \leqslant \sigma$, one can take $\mu_0 = b_0^{-1} A(\delta/4, \sigma/4)^{-1}\delta^2\sigma^d$ and

$$C_\partial = \begin{cases} b_0 A(\delta/4, \sigma/4)^2 \delta^{-3}\sigma^{-3d-1} & \text{if } r_0 + |r| = 0, \\ b_0 A(\delta/4, \sigma/4)^2 \delta^{-2-r_0}\sigma^{-3d-1-|r|} & \text{if } r_0 + |r| = 1 \text{ or } 2, \end{cases}$$

where the positive number b_0 depends only on u_0 and the function A measures the size of F according to formula (1.9)*.*

Theorem 3.2 *implies Theorem* 1.1. – The number w defined by (3.4) can be made arbitrarily close to w_* by an appropriate choice of δ and σ. Then inequalities (3.5) and (3.3) together with the assumption $|\mu| \leqslant \mu_0$ produce bounds for the partial derivatives of the function $\Delta\widetilde{S}$, which translate into the bounds (1.6) for the partial derivatives of ΔS since the change of variable $q \mapsto u$ has its derivatives bounded in $[q_0(-u_0), q_0(u_0)]$. □

Theorem 3.2 *implies Theorem* 1.2. – Fix any $u_0 > 0$ and choose $b, \sigma_{\max}$ such that, for all $\varepsilon, \delta, \sigma > 0$ satisfying $2\delta \leqslant \sigma \leqslant \sigma_{\max}$, inequalities (3.5) hold for $u \in [-u_0, u_0]$, $\varphi \in \mathbb{T}^d$ and $\mu \in [-\mu_0, \mu_0]$, with the numbers μ_0 and C_∂ which are indicated at the end of Theorem 3.2. Let us specialize the result to the case where $\delta = 4\varepsilon^{1/2\tau}$ and $\sigma = 8\varepsilon^{1/2\tau}$, for ε small enough (in order to ensure $2\delta \leqslant \sigma \leqslant \sigma_{\max}$). As noticed earlier (at the end of the proof of the last statement of Corollary 2.2), we have in that case

$$w\varepsilon^{-1/2\tau} \geqslant w_*\varepsilon^{-1/2\tau} - \text{const},$$

hence the result in the variables (u, φ). We transfer it to the variables (q, φ) by the same argument as above. □

Proof of Theorem 3.2. –

(a) Let us call D the characteristic vector field of (S^-, S^+). Since the dependence on the action-like variables in our Hamiltonian (3.1) is so simple (it is quadratic), we find

$$D \cdot (\Delta S) = 0 \quad \text{with } D = \frac{1}{2}\partial_q(S^+ + S^-)\frac{\partial}{\partial q} + \left(z\omega + \frac{1}{2}\alpha\partial_\varphi(S^+ + S^-)\right)\cdot\frac{\partial}{\partial\varphi}$$

(so, in our case, D is the characteristic vector field associated to the Lagrangian manifold

$$\mathcal{G}\mathrm{r}\left(\frac{1}{2}\mathrm{d}(S^+ + S^-)\right),$$

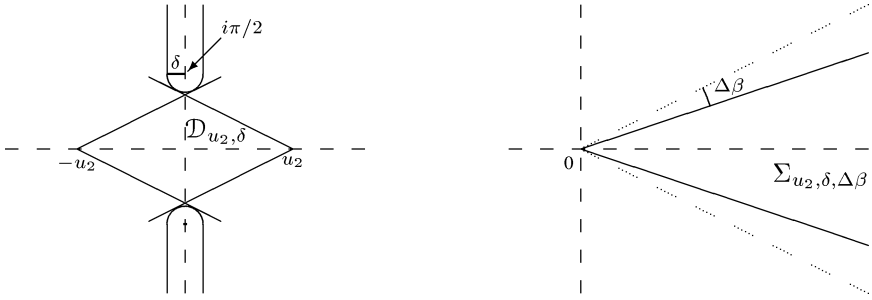
i.e. the projection onto the configuration space of the Hamiltonian vector field restricted to this "averaged" manifold which is generally not invariant!).

The components of D are real-analytic in the intersection of the domains of analyticity of S^- and S^+, and for $\mu = 0$ we find the unperturbed characteristic field D_0 again. In the same manner that we had at our disposal the coordinates (u, φ) in the configuration space to straighten D_0, i.e. to conjugate it to a vector field $\widetilde{D}_0$ with constant coefficients, we shall "straighten" D. Let us start by writing it in the coordinates (u, φ):

$$\widetilde{D} \cdot (\Delta \widetilde{S}) = 0, \tag{3.6}$$

$$\widetilde{D} = \frac{1}{2}\left(\frac{\mathrm{d}q_0}{\mathrm{d}u}\right)^{-2} \partial_u(\widetilde{S}^+ + \widetilde{S}^-)\frac{\partial}{\partial u} + \left(z\omega + \frac{1}{2}\alpha\partial_\varphi(\widetilde{S}^+ + \widetilde{S}^-)\right) \cdot \frac{\partial}{\partial \varphi}.$$

We shall henceforth use the notation $\mathcal{D}_{u_2,\delta} = \mathcal{D}^+_{u_2,\delta} \cap \mathcal{D}^-_{u_2,\delta}$.



PROPOSITION 3.3. – *For all $u_2 > 0$, for all $\Delta\beta \in]0, \arctan\frac{\pi}{2u_2}[$ and for all small enough $\delta, \sigma > 0$, there exist positive numbers μ_2 and $\{M_\partial\}$, and there exists a real-analytic change of coordinates*

$$(u, \varphi) = (v, \theta) + \mu\mathcal{U}(v, \theta; z, \mu) \quad \Longleftrightarrow \quad (v, \theta) = (u, \varphi) + \mu\mathcal{V}(u, \varphi; z, \mu)$$

satisfying the following properties:
– it conjugates $\widetilde{D}$ and $\frac{\partial}{\partial v} + z\omega \cdot \frac{\partial}{\partial \theta}$;
– the mapping $\mathrm{Id} + \mu\mathcal{U}$ induces a bijection between the domain

$$v \in \overline{\mathcal{D}}_{u_2,\delta}, \quad \theta \in \overline{\mathbb{T}}^d_{h_0-\sigma}$$

and its image for $|\mu| \leqslant \mu_2$, $z \in \overline{\Sigma}_{u_2,\delta,\Delta\beta}$; for those values of (v, θ, z, μ), the components $\mathcal{U}_j$ of $\mathcal{U}$ are analytic with respect to all their arguments and satisfy

$$|\partial \mathcal{U}_j| \leqslant M_\partial,$$

where ∂ stands for one of the operators $(\partial_v)^{r_0}(\partial_\theta)^r$ with $(r_0, r) \in \mathbb{N} \times \mathbb{N}^d$, $r_0 + |r| \leqslant 2$;
– the mapping $\mathrm{Id} + \mu\mathcal{V}$ induces a bijection between the domain

$$u \in \overline{\mathcal{D}}_{u_2,\delta}, \qquad \varphi \in \overline{\mathbb{T}}^d_{h_0-\sigma}$$

and its image for $|\mu| \leqslant \mu_2$, $z \in \overline{\Sigma}_{u_2,\delta,\Delta\beta}$; for those values of (u, φ, z, μ), the components $\mathcal{V}_j$ of $\mathcal{V}$ are analytic with respect to all their arguments and satisfy

$$|\partial \mathcal{V}_j| \leqslant M_\partial,$$

where ∂ stands for one of the operators $(\partial_u)^{r_0}(\partial_\varphi)^r$ with $(r_0, r) \in \mathbb{N} \times \mathbb{N}^d$, $r_0 + |r| \leqslant 2$.

More precisely, if $2\delta \leqslant \sigma$, one can take

$$\mu_2 = b_2^{-1} A(\delta/4, \sigma/4)^{-1} \delta^2 \sigma^d \quad \textit{and} \quad M_\partial = b_2 A(\delta/4, \sigma/4) \delta^{-1-r_0} \sigma^{-d-|r|},$$

where the positive number b_2 depends only on u_2 and $\Delta\beta$, and the function A measures the size of F according to the formula (1.9).

The proof of Proposition 3.3 *is postponed to Section 6.* – However, we already notice that it claims only the existence of global flow-box coordinates for D in some complex neighbourhood of $[2\pi - q_2, q_2] \times \mathbb{T}^d$, for $\pi < q_2 < 2\pi$ and $|\mu|$ small enough. This is not surprising, at least in the real domain, since there exist global transversal sections to D as we have already noticed earlier (Section 1.3.2). At a technical level, only the Banach fixed-point theorem is needed, the key-point lying in Lemma 6.2 which provides a bounded inverse to the right for $\widetilde{D}_0$.

(b) Let us see now how we can deduce Theorem 3.2 from what has been done up to now. The idea is simply to apply Corollary 2.2 to the function

$$\chi(v, \theta; \varepsilon, \mu) = \Delta\widetilde{S} \circ (\mathrm{Id} + \mu\mathcal{U}) - \mu\Delta\widetilde{S}_1$$

and to its partial derivatives, and then to transfer the exponentially small bounds which we shall obtain from the real part of the (v, θ)-domain to the real part of the (u, φ)-domain. We shall end up with exponentially small bounds for

$$(\Delta\widetilde{S} - \mu\Delta\widetilde{S}_1)\big(u, \varphi; \varepsilon^{-1/2}, \mu\big) = \chi \circ (\mathrm{Id} + \mu\mathcal{V}) + \mu\big[\Delta\widetilde{S}_1 \circ (\mathrm{Id} + \mu\mathcal{V}) - \Delta\widetilde{S}_1\big]$$

and its derivatives (but only for real values of ε and μ since we want realness to be preserved by the transfer). We only need to write carefully each step of the process in order to keep track of the dependence of the bounds on δ and σ.

We thus suppose that the hypotheses of Theorem 3.2 are fulfilled and that we are given $u_0 > 0$. We define

$$u_1 = \frac{\pi}{\sqrt{3}}, \quad u_2 = \frac{\pi}{2\sqrt{3}}, \quad \Delta\beta = \inf\left\{\arctan\frac{\pi}{2u_0},\ \beta\left(u_1, \frac{\pi}{8}\right),\ \beta\left(u_2, \frac{\pi}{4}\right)\right\},$$

where we have used the notation $\beta(u_i, \delta)$ for the half-aperture of the sectors $\mathcal{D}^-_{u_i,\delta}$ and $\mathcal{D}^+_{u_i,\delta}$, and

$$\sigma_0 = \frac{1}{2}h_0, \quad \delta_0 = \frac{1}{2}\inf\left\{\sigma_0, \frac{\pi}{2}\cos\Delta\beta - u_0 \sin\Delta\beta\right\}.$$

These definitions are quite arbitrary. In particular, they are meant to ensure that, if $0 < \delta \leqslant \frac{\pi}{4}$, $\Sigma_{u_1,\delta/2,\Delta\beta}$, $\Sigma_{u_2,\delta,\Delta\beta}$ and $\Sigma_{u_0,\delta_0,\Delta\beta}$ are well defined sectors which contain $\mathbb{R}^{*+}$ and

$$[-\mathrm{i}\rho, \mathrm{i}\rho] \subset \mathcal{D}_{u_2,\delta}, \quad \rho = \frac{\pi}{2} - 2\delta.$$

(c) Let δ and σ be positive numbers such that $\delta \leqslant \frac{\pi}{4}$, and small enough for Theorem 3.1 to apply with $(u_1, \Delta\beta, \delta/2, \sigma/2)$ and for Proposition 3.3 to apply with $(u_2, \Delta\beta, \delta, \sigma)$. Let $\rho = \frac{\pi}{2} - 2\delta$ and define w according to formula (3.4).

– Proposition 3.3 provides an analytic vectorial function $\mathcal{U}$ whose components are bounded by some $M = M(u_2, \Delta\beta, \delta, \sigma)$ in $\overline{\mathcal{D}}_{u_2,\delta} \times \overline{\mathbb{T}}^d_{h_0-\sigma} \times \overline{\Sigma}_{u_2,\delta,\Delta\beta} \times \overline{D}_{\mu_2}$, where $\mu_2 = \mu_2(u_2, \Delta\beta, \delta, \sigma)$

is some positive number and $\overline{D}_{\mu_2}$ denotes the closed disk of centre 0 and radius μ_2. If $(z,\mu)\in\mathbb{R}^+\times[-\mu_2,\mu_2]$, the image of $\overline{\mathcal{D}}_{u_2,\delta}\times\overline{\mathbb{T}}^d_{h_0-\sigma}$ by $\mathrm{Id}+\mu\mathcal{U}$ is contained in $\overline{\mathcal{D}}_{u_2+\frac{|\mu|M}{\sin\beta_2},\delta-|\mu|M}\times\overline{\mathbb{T}}^d_{h_0-(\sigma-|\mu|M)}$, where $\beta_2=\beta(u_2,\delta)$. In particular, if

$$|\mu|\leqslant\mu_3:=\inf\left\{\mu_2,\frac{\delta}{2M},\frac{\sigma}{2M}\right\},$$

this image is contained in $\overline{\mathcal{D}}_{u_1,\delta/2}\times\overline{\mathbb{T}}^d_{h_0-\sigma/2}$ (since $u_2=\frac{\pi}{2\sqrt{3}}$ and $\delta\leqslant\frac{\pi}{4}$, we have $\sin\beta_2\geqslant\frac{\sqrt{3}}{4}$ and $u_2+\frac{\delta}{2\sin\beta_2}\leqslant u_2+\frac{\pi}{2\sqrt{3}}=u_1$). We can summarize by

$$\boxed{\mathrm{Id}+\mu\mathcal{U}:\mathcal{E}_2=\mathcal{D}_{u_2,\delta}\times\mathbb{T}^d_{h_0-\sigma}\ \to\ \mathcal{E}_1=\mathcal{D}_{u_1,\delta/2}\times\mathbb{T}^d_{h_0-\sigma/2}}$$

if we forget about the parameters (ε,μ) and focus on the variables.

– The definition of an analytic function

$$\chi=\Delta\widetilde{S}\circ(\mathrm{Id}+\mu\mathcal{U})-\mu\Delta\widetilde{S}_1 \tag{3.7}$$

is allowed by Theorem 3.1 which shows that $\Delta\widetilde{S}$ is analytic in $\mathcal{E}_1\times\Sigma_{u_1,\delta/2,\Delta\beta}\times D_{\mu_1}$ and bounded in the closure of that domain for some $\mu_1=\mu_1(u_1,\Delta\beta,\delta/2,\sigma/2)$:

$$\boxed{\Delta\widetilde{S}:\mathcal{E}_1\ \to\ \mathbb{R},\quad\big|\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1\big|\leqslant 2\mathrm{e}^{2u_1}B|\mu|^2\quad\text{in }\overline{\mathcal{E}}_1.}$$

Thus the function χ is bounded in $\overline{\mathcal{E}}_2\mathbb{R}^+\times[-\mu_4,\mu_4]$ where $\mu_4=\inf\{\mu_3,\mu_1\}$. The identity

$$\chi=(\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1)\circ(\mathrm{Id}+\mu\mathcal{U})+\mu\big[\Delta\widetilde{S}_1\circ(\mathrm{Id}+\mu\mathcal{U})-\Delta\widetilde{S}_1\big] \tag{3.8}$$

shows that $|\chi|$ is bounded by $C'|\mu|^2$ in that domain, where

$$C'=2\mathrm{e}^{2u_1}B+M\sup_{\overline{\mathcal{D}}_{u_1,\delta/2}\times\overline{\mathbb{T}}^d_{h_0-\sigma/2}}\big|\mathrm{d}(\Delta\widetilde{S}_1)\big|.$$

We retain

$$\boxed{\chi:\mathcal{E}_2\ \to\ \mathbb{R},\quad|\chi|\leqslant C'|\mu|^2\quad\text{in }\overline{\mathcal{E}}_2.}$$

– Since $D\cdot(\Delta S)=0$ and $\widetilde{D}_0\cdot(\Delta\widetilde{S}_1)=0$, the function χ defined by formula (3.7) satisfies the partial differential equation

$$\left(\frac{\partial}{\partial v}+z\omega\cdot\frac{\partial}{\partial\theta}\right)\chi=0$$

and so do all its partial derivatives. Since $[-\mathrm{i}\rho,\mathrm{i}\rho]\subset\mathcal{D}_{u_2,\delta}$, Corollary 2.2 applies and we find

$$\boxed{\big|\chi-\mu^2a\big|\leqslant c\sigma^{-d}C'|\mu|^2\exp\big(-w\varepsilon^{-1/2\tau}\big)\quad\text{in }\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}^{*+}\times[-\mu_4,\mu_4],}$$

where a is an analytic function of $z=\varepsilon^{-1/2}$ and μ only which is defined by the formula

$$\mu^2 a(z,\mu)=(2\pi)^{-d}\int_{\mathbb{T}^d}\chi(v,\theta;z,\mu)\,\mathrm{d}\theta.$$

– In order to come back to the variables (u,φ) we apply Proposition 3.3 again, but with $(u_0,\Delta\beta,\delta_0,\sigma_0)$: we get an analytic vectorial function $\mathcal{V}$ whose components are bounded by $M'=M(u_0,\Delta\beta,\delta_0,\sigma_0)$ in $\overline{\mathcal{D}}_{u_0,\delta_0}\times\overline{\mathbb{T}}^d_{h_0-\sigma_0}\times\overline{\Sigma}_{u_0,\delta_0,\Delta\beta}\times\overline{D}_{\mu_2'}$, for some $\mu_2'=\mu_2(u_0,\Delta\beta,\delta_0,\sigma_0)$. We choose

$$\mu_0=\inf\{\mu_4,\mu_2'\}.$$

Since realness is preserved by our change of variables (for real parameters), we have

$$\boxed{\mathrm{Id}+\mu\mathcal{V}\colon[-u_0,u_0]\times\mathbb{T}^d\ \to\ \mathbb{R}\times\mathbb{T}^d}$$

and the identity

$$\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1=\chi\circ(\mathrm{Id}+\mu\mathcal{V})+\mu\big[\Delta\widetilde{S}_1\circ(\mathrm{Id}+\mu\mathcal{V})-\Delta\widetilde{S}_1\big] \tag{3.9}$$

yields the inequality

$$\big|\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1-\mu^2 a\big|\leqslant\big(c\sigma^{-d}C'\exp\big(-w\varepsilon^{-1/2\tau}\big)+M'\sup_{\mathbb{R}\times\mathbb{T}^d}\big|\mathrm{d}(\Delta\widetilde{S}_1)\big|\big)|\mu|^2$$

in $[-u_0,u_0]\times\mathbb{T}^d\times\mathbb{R}^{*+}\times[-\mu_0,\mu_0]$. And it is clear that bounds of the type (3.5) may be obtained too, since at each step of the process we could bound the appropriate partial derivatives too.

(d) Let us now suppose that $2\delta\leqslant\sigma\leqslant 1$ and compute explicit values that the numbers μ_0 and C_∂ may assume. We shall denote by A the number $A(\delta/4,\sigma/4)$ (which is not smaller than 1).

According to the above chain of reasoning, we can take

$$\mu_0=\inf\left\{\mu_2,\frac{\delta}{2M},\frac{\sigma}{2M},\mu_1,\mu_2'\right\}.$$

According to Proposition 3.3, we can take

$$\mu_2=\mu_2(u_2,\Delta\beta,\delta,\sigma)=b_2^{-1}A^{-1}\delta^2\sigma^d,\qquad M=M(u_2,\Delta\beta,\delta,\sigma)=b_2A\delta^{-1}\sigma^{-d},$$

where the number b_2 depends only on u_0, and $\mu_2'=\mu_2(u_0,\Delta\beta,\delta_0,\sigma_0)$ depending only on u_0. According to Theorem 3.1, we can take

$$\mu_1=\mu_1(u_1,\Delta\beta,\delta/2,\sigma/2)=b_1^{-1}A^{-1}\delta\sigma^{d+1},\quad b_1\text{ depending only on }u_0.$$

Hence the choice $\mu_0=b_0^{-1}A^{-1}\delta^2\sigma^d$ with a number b_0 which depends only u_0.

Let us number from 0 to d the coordinates in $\mathbb{C}\times(\mathbb{C}/2\pi\mathbb{Z})^d$ and denote by ∂_i the operator of partial differentiation with respect to the ith variable; ∂ will denote a differential operator $\partial_0^{r_0}\partial_1^{r_1}\cdots\partial_d^{r_d}$ with $r_0+r_1+\cdots+r_d\leqslant 2$. According to Proposition 3.3, the components of $\mathcal{U}$ satisfy

$$|\partial\mathcal{U}_j|\leqslant M_\partial=b_2A\delta^{-1-r_0}\sigma^{-d-|r|}\quad\text{in }\overline{\mathcal{E}}_2\times\mathbb{R}^+\times[-\mu_0,\mu_0],\ j=0,1,\ldots,d$$

(the above number M coincides with M_{Id}). And according to Theorem 3.1, the function $\Psi = \Delta\widetilde{S} - \mu\Delta\widetilde{S}_1$ satisfies

$$|\partial\Psi| \leqslant 2\mathrm{e}^{2u_1} B_\partial |\mu|^2 \quad \text{in } \overline{\mathcal{E}}_1 \times \mathbb{R}^+ \times [-\mu_0, \mu_0],$$

with

$$B_\partial = \begin{cases} b_1 A^2 \delta^{-2-\sup\{1,r_0\}} \sigma^{-2d-1} & \text{if } |r| = 0, \\ b_1 A^2 \delta^{-2} \sigma^{-2d-|r|} & \text{if } |r| \geqslant 1. \end{cases}$$

The formula (3.8) can be rewritten as

$$\chi = \chi_1 + \chi_2, \quad \chi_1 = \Psi \circ (\mathrm{Id} + \mu\mathcal{U}), \quad \chi_2 = \mu^2 \Phi \cdot \mathcal{U}$$

with

$$\Phi = \int_0^1 \mathrm{d}(\Delta\widetilde{S}_1) \circ (\mathrm{Id} + \mu x \mathcal{U})\, \mathrm{d}x.$$

The following lemma (whose proof is left to the reader) will help us to compute numbers C'_∂ such that

$$|\partial\chi| \leqslant C'_\partial |\mu|^2 \quad \text{in } \overline{\mathcal{E}}_2 \times \mathbb{R}^+ \times \overline{D}_{\mu_0}.$$

LEMMA 3.2. – *Let $\mathfrak{E}_1$ and $\mathfrak{E}_2$ be subsets of $\mathbb{C} \times (\mathbb{C}/2\pi\mathbb{Z})^d$, and number from 0 to d the coordinates in that space. Let Ψ be a function analytic in $\mathfrak{E}_1$ and $\mathcal{U}$ an analytic vectorial function such that the image of $\mathfrak{E}_2$ by $\mathrm{Id} + \mu\mathcal{U}$ is contained in $\mathfrak{E}_1$. Suppose that Ψ and the components of $\mathcal{U}$ satisfy the inequalities*

$$\begin{cases} |\partial\Psi| \leqslant B_\partial & \text{in } \mathfrak{E}_1, \\ |\partial\mathcal{U}_j| \leqslant M_\partial & \text{in } \mathfrak{E}_2, \quad j = 0, 1, \ldots, d, \end{cases}$$

for $\partial = \partial_0^{r_0}\partial_1^{r_1}\cdots\partial_d^{r_d}$ with $r_0 + r_1 + \cdots + r_d \leqslant 2$. Then the function $\Psi \circ (\mathrm{Id} + \mu\mathcal{U})$ satisfies the inequalities

$$\big|\partial\big(\Psi \circ (\mathrm{Id} + \mu\mathcal{U})\big)\big| \leqslant C_\partial \quad \text{in } \mathfrak{E}_2$$

with

$$\begin{aligned} C_{\mathrm{Id}} &= B_{\mathrm{Id}}, \\ C_{\partial_j} &= B_{\partial_j} + (d+1)|\mu| M_{\partial_j} \sup_\ell \big\{B_{\partial_\ell}\big\}, \\ C_{\partial_j\partial_k} &= B_{\partial_j\partial_k} + (d+1)|\mu| \big(M_{\partial_j} \sup_\ell \big\{B_{\partial_k\partial_\ell}\big\} + M_{\partial_k} \sup_\ell \big\{B_{\partial_j\partial_\ell}\big\} + M_{\partial_j\partial_k} \sup_\ell \big\{B_{\partial_\ell}\big\}\big) \\ &\quad + (d+1)^2 |\mu|^2 M_{\partial_j} M_{\partial_k} \sup_{\ell,\ell'} \big\{B_{\partial_\ell\partial_{\ell'}}\big\}. \end{aligned}$$

We apply the lemma to $\chi_1 = \Psi \circ (\mathrm{Id} + \mu\mathcal{U})$ and to $\Phi_x = \mathrm{d}(\Delta\widetilde{S}_1) \circ (\mathrm{Id} + \mu x \mathcal{U})$ (bounds for the functions $\partial\partial_j(\Delta\widetilde{S}_1)$ in $\mathfrak{E}_1 = \overline{\mathcal{E}}_1$ are given by formula (3.2)). Using the assumption $|\mu| \leqslant b_0^{-1} A^{-1} \delta^2 \sigma^d$, we obtain

$$C'_\partial = \begin{cases} b'_0 A^2 \delta^{-3} \sigma^{-2d-1} & \text{if } r_0 + |r| = 0, \\ b'_0 A^2 \delta^{-2-r_0} \sigma^{-2d-1-|r|} & \text{if } r_0 + |r| = 1 \text{ or } 2. \end{cases}$$

By Corollary 2.2, we deduce that

$$\big|\partial\big(\chi-\mu^2 a\big)\big| \leqslant c\sigma^{-d}C'_\partial|\mu|^2\exp\big(-w\varepsilon^{-1/2\tau}\big) \quad \text{in } \mathbb{R}\times\mathbb{T}^d\times\mathbb{R}^+\times[-\mu_0,\mu_0],$$

and the formula (3.9) can be rewritten as

$$\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1-\mu^2 a=\chi_1^*+\chi_2^*, \quad \chi_1^*=(\chi-\mu^2 a)\circ(\mathrm{Id}+\mu\mathcal{V}), \quad \chi_2^*=\mu^2\Phi^*\cdot\mathcal{V}$$

with

$$\Phi^*=\int_0^1 \mathrm{d}(\Delta\widetilde{S}_1)\circ(\mathrm{Id}+\mu x\mathcal{V})\,\mathrm{d}x.$$

We apply again Lemma 3.2, but with $\mathfrak{E}_1=\mathfrak{E}_2=\mathbb{R}\times\mathbb{T}^d$ and $\mathcal{V}$ (or $x\mathcal{V}$) in place of $\mathcal{U}$ (its partial derivatives are bounded by a number which depends only on u_0, and we use the bounds for the functions $\partial\partial_j(\Delta\widetilde{S}_1)$ in $\mathfrak{E}_1$ which are given by the formula (3.3)). We end up with inequalities $\big|\partial(\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1-\mu^2 a)\big|\leqslant C_\partial|\mu|^2\exp\big(-w\varepsilon^{-1/2\tau}\big)$, where we can take for the numbers C_∂ the values which were announced in Theorem 3.2. □

Remark 3.2. – Proposition 3.3 allows one to present things slightly differently and to obtain an interesting intermediary result without any Diophantine condition on ω.

Let $\mathfrak{C}_2=]q_0(-u_2),q_0(u_2)[\times\mathbb{T}^d$ be the part of the configuration space which corresponds to the real part of the domain $(v,\theta)\in\mathcal{D}_{u_2,\delta}\times\mathbb{T}^d_{h_0-\sigma}$. The change of coordinates $\mathrm{Id}+\mu\mathcal{U}$ induces a mapping f according to the formula

$$(u,\varphi)=(\mathrm{Id}+\mu\mathcal{U})(v,\theta) \iff (q_0(u),\varphi)=f(q_0(v),\theta),$$

which is a diffeomorphism between $\mathfrak{C}_2$ and some domain $\mathfrak{C}_1$. We may consider f as a change of coordinates as well, its reciprocal will be denoted g:

$$(q,\varphi)=f(Q,\theta) \iff (Q,\theta)=g(q,\varphi).$$

According to the classical construction f admits an exact symplectic lift defined as

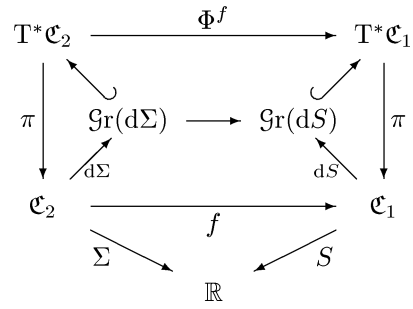
$$\Phi^f:\begin{cases}\mathrm{T}^*\mathfrak{C}_2\longrightarrow\mathrm{T}^*\mathfrak{C}_1,\\ \beta\longmapsto\Phi^f(\beta)={}^{\mathrm{t}}(\mathrm{T}_{f\circ\pi(\beta)}g)\cdot\beta,\end{cases}$$

where π denotes the projection $\mathrm{T}^*\mathfrak{C}_2\to\mathfrak{C}_2$.

$$\begin{array}{ccc}\mathrm{T}^*\mathfrak{C}_2 & \xrightarrow{\Phi^f} & \mathrm{T}^*\mathfrak{C}_1\\ \pi\downarrow & & \downarrow\pi\\ \mathfrak{C}_2 & \underset{g}{\overset{f}{\rightleftarrows}} & \mathfrak{C}_1\end{array}$$

Not only Φ^f is a lift of f (i.e. $\pi\circ\Phi^f=f\circ\pi$) which preserves the Liouville form λ, but its action on exact Lagrangian graphs is easily described: if S is a function on $\mathfrak{C}_1$,

$$\mathcal{G}\mathrm{r}(\mathrm{d}S)=\Phi^f\big(\mathcal{G}\mathrm{r}(\mathrm{d}\Sigma)\big) \quad \text{with } \Sigma=S\circ f.$$



Thus we have two reciprocal transformations Φ^f and Φ^g, that we can consider as an exact symplectic change of coordinates in $\mathrm{T}^*\mathfrak{C}_1$, under which $\mathcal{W}^- = \mathcal{G}\mathrm{r}(\mathrm{d}S^-)$ and $\mathcal{W}^+ = \mathcal{G}\mathrm{r}(\mathrm{d}S^+)$ become the exact Lagrangian graphs associated to new generating functions defined on $\mathfrak{C}_2$:

$$\mathcal{W}^\pm = \Phi^f\big(\mathcal{G}\mathrm{r}(\mathrm{d}\Sigma^\pm)\big) \quad \text{with } \Sigma^\pm = S^\pm \circ f.$$

The interest of all this is the possibility of applying Lemma 2.1 to the function $\Delta\Sigma = \Sigma^+ - \Sigma^-$ (or rather to the functions $\partial(\Delta\widetilde{\Sigma} - \mu\Delta\widetilde{S}_1)$), since f was chosen precisely in order to yield $D_0 \cdot \Delta\Sigma = 0$.

Of course all these transformations depend on ε and μ, but we notice that, since $f = \mathrm{Id} + \mathrm{O}(\mu)$, the function $\chi = \Delta\Sigma - \mu\Delta S_1$ (which is the same as in the formula (3.7) up to the change of variable $Q = q_0(v)$) is $\mathrm{O}(\mu^2)$. Using the same inequalities as above we end up with the following result:

THEOREM 3.3. – *Let $Q_2 \in]\pi, 2\pi[$ and let δ, σ be small enough positive numbers. Let $\rho = \frac{\pi}{2} - 2\delta$, $h = h_0 - \sigma$ and $\mathfrak{C}_2 =]2\pi - Q_2, Q_2[\times \mathbb{T}^d$. There exist a subdomain $\mathfrak{C}_1$ of $]0, 2\pi[\times \mathbb{T}^d$, an exact symplectic diffeomorphism Φ between $\mathrm{T}^*\mathfrak{C}_2$ and $\mathrm{T}^*\mathfrak{C}_1$ and functions Σ^- and Σ^+ on $\mathfrak{C}_2$ (which depend analytically on (ε, μ)) such that*

$$\mathcal{W}^\pm = \Phi\big(\mathcal{G}\mathrm{r}(\mathrm{d}\Sigma^\pm)\big), \quad \Sigma^+ - \Sigma^- = \mu\Delta S_1 + \chi,$$

where the Fourier coefficients χ_k^∂ of the partial derivatives $\partial\chi$ of the function χ satisfy

$$\big|\chi_k^\partial(Q;\varepsilon,\mu)\big| \leqslant C'_\partial |\mu|^2 \mathrm{e}^{-\rho\varepsilon^{-1/2}|k\cdot\omega| - h|k|}$$

for

$$k \in \mathbb{Z}^d, \quad Q \in]2\pi - Q_2, Q_2[, \quad \varepsilon > 0, \quad \mu \in [-\mu_2, \mu_2]$$

and $\partial = (\partial_Q)^{r_0}(\partial_\theta)^r$ with $r_0 + |r| \leqslant 2$.

If $2\delta \leqslant \sigma$, one can take $\mu_2 = b_2^{-1} A(\delta/4, \sigma/4)^{-1}\delta^2\sigma^d$ and

$$C'_\partial = \begin{cases} b_2 A(\delta/4,\sigma/4)^2 \delta^{-3}\sigma^{-2d-1} & \text{if } r_0 + |r| = 0, \\ b_2 A(\delta/4,\sigma/4)^2 \delta^{-2-r_0}\sigma^{-2d-1-|r|} & \text{if } r_0 + |r| = 1 \text{ or } 2, \end{cases}$$

where the positive number b_2 depends only on Q_2.

4. Results for the three-degree-of-freedom case ($d = 2$)

In this section we restrict ourselves to a very specific case in order to benefit from some arithmetic results contained in [4] and to adapt them to our situation. Our goal here is to obtain

in the upper bound of the splitting an exponential term of the same kind as in [4], which will be optimal in some cases. We shall be concerned not only by the function ΔS and its partial derivatives, but also by its Hessian matrix: we saw in Proposition 1.2 and its corollary that ΔS is determined by its restriction to any section $\{q = q_*\}$ of the configuration space and that it admits critical points. In fact the Hessian (i.e. the determinant of the Hessian matrix) of such a restriction $\Delta S_{|q=q_*}$ at a critical point provides a symplectic measure of the splitting along the corresponding homoclinic orbit [21], whereas the gradient of ΔS only measures the distance of the invariant manifolds above a given point of the configuration space. We shall see cases where lower bounds are available for the gradient of ΔS and the Hessian of its restrictions.

4.1. Statement of the results

Notations. – In this section, $\tau = d = 2$ and $\omega = (1, \Gamma)$ with $\Gamma = \frac{1+\sqrt{5}}{2}$. We denote by $\{x\}$ the distance of a real number x to $\mathbb{Z}$:

$$\forall x \in \mathbb{R}, \quad \{x\} = \operatorname{dist}(x, \mathbb{Z}) \in [0, 1/2],$$

and for any $\rho, h > 0$ we define a continuous $(4 \log \Gamma)$-periodic function $w_{\rho,h}$ on $\mathbb{R}$ by the formula

$$\forall X \in \mathbb{R}, \quad w_{\rho,h}(X) = C_{\rho,h} \cosh\left(\left\{ \frac{X - X_{\rho,h}}{4 \log \Gamma} \right\} \log \Gamma \right),$$

where

$$C_{\rho,h} = 5^{-1/4} \Gamma \sqrt{4\rho h}, \qquad X_{\rho,h} = 2 \log\left(\frac{\rho\sqrt{5}}{h} \right).$$

THEOREM 4.1. – *Consider the Hamiltonian system* (1.1) *with F satisfying the assumptions* (A1) *and* (A2)*, the difference ΔS of the solutions of the Hamilton–Jacobi equation associated to the stable and unstable manifolds and its Poincaré–Melnikov approximation ΔS_1. We shall use the notation*

$$w_* = w_{\pi/2, h_0}.$$

Under the above hypotheses and notations, for any closed subinterval $[q_1, q_2]$ of $]0, 2\pi[$ there exist positive constants ε_0 and b such that the inequalities

$$\left| (\partial_q)^{r_0} (\partial_\varphi)^r \Delta S_1 \right| \leqslant b A\big(\varepsilon^{1/4}, \varepsilon^{1/4}\big) \varepsilon^{-\frac{r_0 + |r| - 2}{4}} \exp\big(-w_*(\log \varepsilon) \varepsilon^{-1/4}\big) \tag{4.1}$$

and

$$\begin{aligned} &\left| (\partial_q)^{r_0} (\partial_\varphi)^r (\Delta S - \mu \Delta S_1)(q, \varphi; \varepsilon, \mu) \right| \\ &\quad \leqslant b |\mu|^2 A\big(\varepsilon^{1/4}, \varepsilon^{1/4}\big)^2 \varepsilon^{-\frac{r_0 + |r| + 5}{4}} \exp\big(-w_*(\log \varepsilon) \varepsilon^{-1/4}\big) \end{aligned} \tag{4.2}$$

hold for

$$(q, \varphi) \in [q_1, q_2] \times \mathbb{T}^2, \quad \varepsilon \in]0, \varepsilon_0], \quad \mu \in [-\mu_0(\varepsilon), \mu_0(\varepsilon)],$$

with $r_0 + |r| = 1$ or 2 and $\mu_0(\varepsilon) = b^{-1} A(\varepsilon^{1/4}, \varepsilon^{1/4})^{-1} \varepsilon$.

The improvement with respect to Proposition 1.3 and Theorem 1.2 consists in the replacement of the fixed coefficient previously denoted by w_* by a new coefficient $w_*(\log \varepsilon)$ which oscillates between two positive values (the new coefficient being never smaller than the fixed value it

assumed in Section 1.4). Moreover, the upper bound (4.2) may be compared to a lower bound of ΔS_1 which is available when F satisfies further assumptions:

THEOREM 4.2 [4]. – *Suppose that $K, a > 0$ and*

$$F(q,\varphi) = m(\varphi_1,\varphi_2)(1-\cos q)$$

with an analytic function m whose Fourier coefficients satisfy:
- *for all $k \in \mathbb{Z}^2$, $|m_k| \leqslant K\mathrm{e}^{-h_0|k|}$;*
- *for all $n \in \mathbb{N}^*$, $\left|m_{\pm k^{(n)}}\right| \geqslant a\mathrm{e}^{-h_0|k^{(n)}|}$, if we denote by $k^{(n)} = (-\mathcal{F}_n, \mathcal{F}_{n-1})$ the Fourier numbers which correspond to the Fibonacci sequence $\mathcal{F}_0 = 1$, $\mathcal{F}_1 = 1$, $\mathcal{F}_n = \mathcal{F}_{n-1} + \mathcal{F}_{n-2}$ for $n \geqslant 2$.*

Then there exists a positive constant b such that

$$\max_{\varphi\in\mathbb{T}^2}\left\{\left|(\partial_q)^{r_0}(\partial_\varphi)^r \Delta S_1(q,\varphi;\varepsilon)\right|\right\} \geqslant b^{-1}\varepsilon^{-\frac{r_0+|r|-1}{4}}\exp\left(-w_*(\log\varepsilon)\varepsilon^{-1/4}\right) \tag{4.3}$$

for $\varepsilon > 0$ small enough and $r_0 + |r| = 1$ or 2.

Note that the left-hand side in inequality (4.3) does not depend on the variable q because of the invariance of ΔS_1 under D_0. The first hypothesis on the Fourier coefficients of m could be replaced by the condition $A(\delta,\sigma) = \mathrm{const}\delta^{-2}\sigma^{-2}$ which is slightly weaker.

Thus the size of the Poincaré–Melnikov approximation is always bounded as indicated by the inequality (4.1) of Theorem 4.1, and under the assumptions of Theorem 4.2 it is not "abnormally small" (i.e. it is not exponentially small with an exponent which would be smaller than $-1/4$ or with a coefficient which would be larger than $w_*(\log\varepsilon)$): in that case one can take $A(\varepsilon^{1/4},\varepsilon^{1/4}) = \mathrm{const}\,\varepsilon^{-1}$ and inequalities (4.2) and (4.3) indicate as a range of values of μ for which $\mu\Delta S_1$ is the dominant part of ΔS an interval $\{\mu\in\mathbb{R};\ |\mu| \leqslant \mathrm{const}\,\varepsilon^{7/2}\}$ at least.

Theorem 4.1 constitutes a generalization of one of the results of [4], inasmuch as that paper was confined to the case where $\alpha = 0$.

Were the first-order approximation abnormally small (e.g. because the right harmonics are lacking in F), we could still try our luck with the second-order approximation, or look for the first finite-order approximation of ΔS of the "right" exponentially small size. The Hamilton–Jacobi algorithm above gives a way of determining the functions S^- and S^+ at each finite order with respect to μ (there still remains the problem of bounding from below the partial derivatives of the functions $\Delta S_n = S_n^+ - S_n^-$), and the method for proving Theorem 4.1 may be adapted to bound the partial derivatives of the remainder

$$\Delta S - \mu\Delta S_1 - \cdots - \mu^N \Delta S_N$$

at any order N. But perhaps a better solution would be to perform a change of variables in order to modify the form of the perturbation as in [27].

Let us now introduce some other notations in order to deal with Hessians.

Notations. – For any $\rho, h > 0$ we define a continuous $(4\log\Gamma)$-periodic function $s_{\rho,h}$ on $\mathbb{R}$ by the formula

$$\forall X\in\mathbb{R},\quad s_{\rho,h}(X) = C_{\rho,h}\Gamma^{3/2}\cosh\left(\left(\frac{1}{2} - \left\{\frac{X - X_{\rho,h}}{4\log\Gamma}\right\}\right)\log\Gamma\right),$$

with the same numbers $C_{\rho,h}$ and $X_{\rho,h}$ as above. We set $s_* = s_{\pi/2,h_0}$ and $X_* = X_{\pi/2,h_0}$.

For $\kappa \in]0,1/4[$, we define

$$L_\kappa = \left\{ X \in \mathbb{R} \,\middle|\, \kappa \leqslant \left\{ \frac{X - X_*}{4 \log \Gamma} \right\} \leqslant \frac{1}{2} - \kappa \right\}.$$

Notice that $\mathbb{R} \setminus L_\kappa$ is the union of the intervals of length $4\kappa \log \Gamma$ which are centered on the solutions X of the equations $w_*(X) = C_*$ or $w_*(X) = C_* \cosh \frac{\log \Gamma}{2}$ (i.e. the points of minimum or maximum of w_*).

Notice also that

$$s_{\rho,h}(X) = C_{\rho,h} \left[\cosh\left(\left\{ \frac{X - X_{\rho,h}}{4 \log \Gamma} \right\} \log \Gamma \right) + \cosh\left(\left(1 - \left\{ \frac{X - X_{\rho,h}}{4 \log \Gamma} \right\} \right) \log \Gamma \right) \right],$$

hence $2w_{\rho,h}(X) \leqslant s_{\rho,h}(X)$. According to Theorem 4.1, we know that the Hessian of the restriction of ΔS to any section $\{q = q_*\}$ is exponentially small with at least a coefficient $2w_*(\log \varepsilon)$ in front of $\varepsilon^{-1/4}$, but we shall see that under the hypotheses of Theorem 4.2 it can reach $s_*(\log \varepsilon)$. But we exclude some intervals for ε by requiring $\log \varepsilon \in L_\kappa$.

THEOREM 4.3. – *Let $q_* \in]0, 2\pi[$ and $\kappa \in]0, 1/4[$. Under the hypotheses of Theorem 4.2, for $\varepsilon > 0$ small enough and such that $\log \varepsilon \in L_\kappa$, and for $\mu \in \mathbb{R}$ such that $|\mu| \leqslant \text{const}\, \varepsilon^{13/4}$, the function*

$$\varphi \in \mathbb{T}^2 \mapsto \Delta S(q_*, \varphi; \mu, \varepsilon)$$

is a Morse function with exactly four distinct critical points, at which the absolute value of its Hessian is bounded from above and from below by expressions of the form

$$\text{const} |\mu|^2 \varepsilon^{1/2} \exp\big(- s_*(\log \varepsilon) \varepsilon^{-1/4} \big).$$

The results for the three-degree-of-freedom case can be generalized to some extent: in all this section the ratio of the two components of the frequency-vector ω is kept fixed to the value Γ in order to use the arithmetic arguments of [4], but [25] provides other arithmetic tools to deal with the case of any constant-type ratio.

4.2. Proof of Theorems 4.1 and 4.2

The only difference with the previous sections lies in what we have called the analytic part of the method (Section 2.2), which can be improved: the accurate knowledge of the arithmetic properties of $\omega = (1, \Gamma)$ makes it possible to replace Corollary 2.2 by a refined result, which we essentially take from [4] (Lemma 4, p. 50).

LEMMA 4.1. – *Let $\rho_0, h_0 > 0$. Suppose that a function $\chi(v, \theta, \varepsilon)$ is real-analytic for $(v, \theta) \in]-\mathrm{i}\rho_0, \mathrm{i}\rho_0[\times \mathbb{T}^d_{h_0}$ and $\varepsilon > 0$, and satisfies the partial differential equation*

$$\left(\frac{\partial}{\partial v} + \varepsilon^{-1/2} \omega \cdot \frac{\partial}{\partial \theta} \right) \chi = 0.$$

For all positive $\rho < \rho_0$ and $h < h_0$ we define

$$B(\rho, h, \varepsilon) = \sup_{(v,\theta) \in [-\mathrm{i}\rho, \mathrm{i}\rho] \times \overline{\mathbb{T}}^d_h} |\chi(v, \theta, \varepsilon)|.$$

The function χ extends analytically to $\{|\Im\mathrm{m}\, v|<\rho_0\}\times\mathbb{T}^d_{h_0}\times\{\varepsilon>0\}$, its mean value $a(\varepsilon)=(2\pi)^{-2}\int_{\mathbb{T}^2}\chi(v,\theta,\varepsilon)\mathrm{d}\theta$ does not depend on v and

(i) *there exists $c>0$ such that, for $\varepsilon>0$ small enough, $(v,\theta)\in\mathbb{R}\times\mathbb{T}^2$, $\frac{1}{2}\rho_0<\rho<\rho_0$ and $\frac{1}{2}h_0<h<h_0$,*

$$|\chi(v,\theta,\varepsilon)-a(\varepsilon)|\leqslant cB(\rho,h,\varepsilon)\exp\big(-w_{\rho,h}(\log\varepsilon)\varepsilon^{-1/4}\big);$$

(ii) *for all $\delta_0,\sigma_0>0$, there exists $c'>0$ such that*

$$\forall\varepsilon>0 \text{ small enough},\ \forall(v,\theta)\in\mathbb{R}\times\mathbb{T}^2,$$
$$|\chi(v,\theta,\varepsilon)-a(\varepsilon)|\leqslant c'B\big(\rho_0-\delta_0\varepsilon^{1/4},h_0-\sigma_0\varepsilon^{1/4},\varepsilon\big)\exp\big(-w_{\rho_0,h_0}(\log\varepsilon)\varepsilon^{-1/4}\big);$$

Proof. – (i) The proof is given in [4], p. 50–51. We reproduce it in order to show where the function $w_{\rho,h}$ comes from.

We begin like for the proof of Corollary 2.2 (we are in the same position, but with additional hypotheses) by applying Lemma 2.1 for each $\varepsilon>0$; it yields the following bounds for the Fourier coefficients of χ:

$$\forall k\in\mathbb{Z}^2,\ \forall v\in\mathbb{R},\quad |\chi_k(v,\varepsilon)|\leqslant B(\rho,h,\varepsilon)\mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}.$$

Thus, if $v\in\mathbb{R}$, $\theta\in\mathbb{T}^2$ and $\varepsilon>0$,

$$|\chi(v,\theta,\varepsilon)-a(\varepsilon)|=\Big|\sum_{k\in\mathbb{Z}^2\smallsetminus\{0\}}\chi_k(v,\varepsilon)\mathrm{e}^{\mathrm{i}k\cdot\theta}\Big|\leqslant B(\rho,h,\varepsilon)\Sigma,$$

with

$$\Sigma=\sum_{k\in\mathbb{Z}^2\smallsetminus\{0\}}\mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}.$$

We decompose this sum according to a partition of $\mathbb{Z}^2\smallsetminus\{0\}$:

$$\Sigma=\sum_{|k\cdot\omega|\geqslant 1/2}+\sum_{|k\cdot\omega|<1/2\text{ and }|k_2|\geqslant\varepsilon^{-1/2}}+\sum_{|k\cdot\omega|<1/2\text{ and }|k_2|<\varepsilon^{-1/2}}.$$

Observe that in the last two terms the condition $|k\cdot\omega|=|k_1+\Gamma k_2|<1/2$ implies that the index k is determined by its second component k_2. Now, one checks easily that each of the first two terms can be bounded by an expression of the form

$$\mathrm{const}_2\,\mathrm{e}^{-\mathrm{const}_1\varepsilon^{-1/2}}$$

with $\mathrm{const}_1=\rho/2$ for the first term, $\mathrm{const}_1=h(\Gamma+1)$ for the second one, and const_2 depending only on h_0 in both cases. In order to conclude, it is thus sufficient to show that

$$\sum_{|k\cdot\omega|<1/2\text{ and }0<k_2<\varepsilon^{-1/2}}\mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}\leqslant c\exp\big(-w_{\rho,h}(\log\varepsilon)\varepsilon^{-1/4}\big),\tag{4.4}$$

with some $c>0$ independent of ρ and h (the restriction to positive values of k_2 is allowed by parity).

The left-hand side in (4.4) is a finite sum which contains no more than $\varepsilon^{-1/2}$ terms; let us look for the largest one. The Fibonacci sequence $(\mathcal{F}_n)_{n\geqslant 0}$ appears here in relation with the best approximations of Γ. It is recalled in [4] that, if $|k\cdot\omega|<1/2$ and $k_2\geqslant 1$, there are only two possibilities:

– either there exists $n\in\mathbb{N}^*$ such that $k_2=\mathcal{F}_{n-1}$, then necessarily $k=k^{(n)}$ (i.e. $k_1=-\mathcal{F}_n$) and
$$|k\cdot\omega|=\Gamma^{-n},\quad |k|=C\big(\Gamma^{n+2}-(-1)^n\Gamma^{-n-2}\big),$$
with $C=5^{-1/2}$;
– or k_2 does not belong to the sequence $(\mathcal{F}_n)$ and
$$|k\cdot\omega|\geqslant\frac{\Gamma C}{|k|}.$$

The second possibility leads immediately to a "very small" contribution:

$$k_2\notin\{\mathcal{F}_n\}\implies \mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}\leqslant\exp\big(-\varepsilon^{-1/4}\sqrt{4\Gamma C\rho h}\big), \tag{4.5}$$

with $\sqrt{4\Gamma C\rho h}>\sqrt{\Gamma^3 C\rho h}=\max_{X\in\mathbb{R}}\big\{w_{\rho,h}(X)\big\}$.

Whereas in the first case we can compute

$$\mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}=\exp\big(-\big(\rho\varepsilon^{-1/2}+(-1)^{n+1}Ch\Gamma^{-2}\big)\Gamma^{-n}-Ch\Gamma^2\cdot\Gamma^n\big)=\exp\big(-A_x(n)\big)$$

with

$$\begin{cases} A_x(n)=2x^{1/4}\Gamma\sqrt{C\rho h}\cosh\left(\dfrac{1}{4}\log x-\dfrac{1}{2}\log\left(\dfrac{\Gamma^2Ch}{\rho}\right)-n\log\Gamma\right),\\[2ex] x=\varepsilon^{-1}\left(1+\dfrac{(-1)^{n+1}\Gamma^{-2}Ch}{\rho}\varepsilon^{1/2}\right)^2, \end{cases} \tag{4.6}$$

(by use of the identity $at+bt^{-1}=2\sqrt{ab}\cosh\big(\frac{1}{2}\log a-\frac{1}{2}\log b+\log t\big)$ for $a,b,t>0$).

Observe that x depends slightly on n but takes only two values if ε is fixed, and

$$x^{1/4}=\varepsilon^{-1/4}\big(1+\mathrm{O}\big(\varepsilon^{1/2}\big)\big),\qquad \log x=-\log\varepsilon+\mathrm{O}\big(\varepsilon^{1/2}\big). \tag{4.7}$$

It is thus interesting to study, for a fixed value of $x\gg 1$, the sequence

$$A_x(n)=x^{1/4}C_{\rho,h}\cosh\left(\frac{\log x+X_{\rho,h}}{4}-(n+1)\log\Gamma\right).$$

Let n_x be the integral part of $\frac{\log x+X_{\rho,h}}{4\log\Gamma}$: the sequence $(A_x(n))_{n<n_x}$ is decreasing and the sequence $(A_x(n))_{n>n_x}$ is increasing, thus

$$\min_{n\geqslant 0}\{A_x(n)\}=\big(A_x(n_x-1)\text{ or }A_x(n_x)\big)=x^{1/4}w_{\rho,h}(-\log x), \tag{4.8}$$

but apart from these two exceptional terms we get again "very small" contributions:

$$n\notin\{n_x-1,n_x\}\implies A_x(n)\geqslant x^{1/4}C_{\rho,h}\cosh(\log\Gamma),$$

thus

$$k_2=\mathcal{F}_{n-1}\text{ with }n\notin\{n_x-1,n_x\}\implies \mathrm{e}^{-h|k|-\rho\varepsilon^{-1/2}|k\cdot\omega|}\leqslant c\exp\bigl(-\varepsilon^{-1/4}C_{\rho,h}\cosh(\log\Gamma)\bigr), \tag{4.9}$$

with $C_{\rho,h}\cosh(\log\Gamma)>C_{\rho,h}\cosh\left(\frac{1}{2}\log\Gamma\right)=\max_{X\in\mathbb{R}}\bigl\{w_{\rho,h}(X)\bigr\}$, and with a constant c stemming from the use of (4.7) to replace $x^{1/4}$ by $\varepsilon^{-1/4}$.

We finally obtain the inequality (4.4) by observing that its left-hand side is bounded by

$$c'\varepsilon^{-1/2}\exp\bigl(-W\varepsilon^{-1/4}\bigr)+c''\exp\bigl(-w_{\rho,h}(\log\varepsilon)\varepsilon^{-1/4}\bigr),$$

where $W>\max w_{\rho,h}$ in the first term thanks to (4.5) and (4.9) (with $W-\max w_{\rho,h}$ bounded from below by some positive number which does not depend on ρ and h), and the second term takes into account the two exceptional indices (in order to replace $-\log x$ by $\log\varepsilon$, one can use the estimate (4.7) and the fact that the function $w_{\rho,h}$ is Lipschitzian with a Lipschitz constant uniform in ρ, h).

(ii) We apply the first part of the lemma with $\rho=\rho_0-\delta_0\varepsilon^{1/4}$ and $h=h_0-\sigma_0\varepsilon^{1/4}$, and we observe that

$$w_{\rho,h}(X)=\frac{C_{\rho,h}}{C_{\rho_0,h_0}}w_{\rho_0,h_0}(X+X_{\rho_0,h_0}-X_{\rho,h}),$$

$\frac{C_{\rho,h}}{C_{\rho_0,h_0}}=1+\mathrm{O}(\varepsilon^{1/4})$, $X_{\rho_0,h_0}-X_{\rho,h}=\mathrm{O}(\varepsilon^{1/4})$ and the function w_{ρ_0,h_0} is Lipschitzian, thus $w_{\rho,h}(X)=w_{\rho_0,h_0}(X)+\mathrm{O}(\varepsilon^{1/4})$ where the involved constant depends only on $\rho_0,h_0,\delta_0,\sigma_0$. □

Proof of Theorem 4.1. – In order to prove the inequalities (4.1) it is sufficient to adapt the proof of Proposition 1.3 which was given on p. 177: the inequalities (3.2) are still valid but one can now use Lemma 4.1 (instead of Corollary 2.2 which was used there) – we let the reader check the details.

In order to prove the inequalities (4.2) we substitute to Theorem 3.2 the following assertion:

For all $u_0>0$ there exist positive numbers ε_0 and b_0 such that, for all small enough $\delta,\sigma>0$ satisfying $2\delta\leqslant\sigma$, the inequalities

$$\begin{aligned}&\bigl|(\partial_u)^{r_0}(\partial_\varphi)^r(\Delta\widetilde{S}-\mu\Delta\widetilde{S}_1)(u,\varphi;\varepsilon,\mu)\bigr|\\&\quad\leqslant b_0|\mu|^2A(\delta/4,\sigma/4)^2\delta^{-2-r_0}\sigma^{-5-|r|}\exp\bigl(-w_{\rho,h}(\log\varepsilon)\varepsilon^{-1/4}\bigr)\end{aligned} \tag{4.10}$$

hold for

$$(u,\varphi)\in[-u_0,u_0]\times\mathbb{T}^2,\quad \varepsilon\in]0,\varepsilon_0],\quad \mu\in[-\mu_0,\mu_0],$$

with $r_0+|r|=1$ or 2, and $\rho=\frac{\pi}{2}-\delta$, $h=h_0-\sigma$, $\mu_0=b_0^{-1}A(\delta/4,\sigma/4)^{-1}\delta^2\sigma^2$.

Such a result is sufficient to conclude, since the choice $(\delta,\sigma)=(4\varepsilon^{1/4},8\varepsilon^{1/4})$ yields

$$w_{\frac{\pi}{2}-\delta,h_0-\sigma}(X)=w_*(X)+\mathrm{O}\bigl(\varepsilon^{1/4}\bigr)$$

as noticed previously. To prove the assertion we proceed as on p. 181–187, but where Corollary 2.2 was used we apply Lemma 4.1 instead. □

Proof of Theorem 4.2. – See [4], p. 41–43, where it is shown that the chain of reasoning which leads to Lemma 4.1 is optimal in the considered case. Indeed, if we use the variable u, we can

write

$$\Delta\widetilde{S}_1(u,\varphi;\varepsilon)=\varepsilon^{1/2}\int_{-\infty}^{+\infty}\frac{2}{\cosh^2(u+\zeta)}m\big(\varphi+\varepsilon^{-1/2}\zeta\omega\big)\mathrm{d}\zeta$$

and the theorem of residues allows us to compute each Fourier coefficient of the function

$$\chi^{\partial}=\varepsilon^{-1/2}\partial(\Delta\widetilde{S}_1),\quad \partial=(\partial_u)^{r_0}(\partial_\varphi)^r.$$

One finds

$$\chi_k^{\partial}(u;\varepsilon)=(-1)^{r_0}\mathrm{i}^{r_0+|r|}\frac{2\pi(\varepsilon^{-1/2}k\cdot\omega)^{r_0+1}k^r}{\sinh(\frac{\pi}{2}\varepsilon^{-1/2}k\cdot\omega)}m_k\mathrm{e}^{-\mathrm{i}\varepsilon^{-1/2}uk\cdot\omega}. \tag{4.11}$$

Since

$$\forall k\in\mathbb{Z}^2,\ \forall(u,\varepsilon)\in\mathbb{R}\times\mathbb{R}^{*+},\quad \big|\chi_k^{\partial}(u;\varepsilon)\big|\leqslant\max_{\varphi\in\mathbb{T}^2}\big\{\big|\chi^{\partial}(u,\varphi;\varepsilon)\big|\big\},$$

it is sufficient to estimate the size of a well chosen Fourier coefficient. The largest one was identified in the course of the proof of Lemma 4.1: it corresponds to the minimum of the sequence which we had denoted A_x, i.e. $k=k^{(n)}$ with $n=n_x$ or n_x-1 according to the formulas (4.6) and (4.8) (we have now $\rho=\pi/2$ and $h=h_0$). With that index k,

$$\big|\chi_k^{\partial}(u;\varepsilon)\big|\geqslant\mathrm{const}\big(\varepsilon^{-1/2}|k\cdot\omega|\big)^{r_0+1}|k_1|^{r_1}|k_2|^{r_2}\exp\left(-\frac{\pi}{2}\varepsilon^{-1/2}|k\cdot\omega|-h_0|k|\right)$$

thanks to the second hypothesis on the Fourier coefficients of m. We recognize in the exponential term $\exp(-A_x(n))=x^{1/4}w_*(-\log x)$; we conclude by observing that $\big|n-\frac{\log x}{4\log\Gamma}\big|$ is bounded, thus

$$|k\cdot\omega|=\Gamma^{-n}\geqslant\mathrm{const}\,\mathrm{e}^{-\frac{\log x}{4}}\geqslant\mathrm{const}\,\varepsilon^{1/4}$$

and, for $j=1$ or 2,

$$|k_j|\geqslant\mathrm{const}\,\Gamma^n\geqslant\mathrm{const}\,\mathrm{e}^{\frac{\log x}{4}}\geqslant\mathrm{const}\,\varepsilon^{-1/4}.\qquad\square$$

4.3. Proof of Theorem 4.3

Our goal is to study the critical points of the function

$$\boxed{f_{\varepsilon,\mu}=\Delta S_{|q=q_*}:\varphi\in\mathbb{T}^2\ \mapsto\ f_{\varepsilon,\mu}(\varphi)=\Delta\widetilde{S}(u_*,\varphi;\varepsilon,\mu)}$$

with u_* defined by $q_*=q_0(u_*)$ (we shall often omit ε,μ in subscript). We shall use the same scheme (and the same notations) as in the proof of Theorem 3.2. In particular, we fix some $u_0>|u_*|$, and we work with $\sigma=2\delta=8\varepsilon^{1/4}$ and

$$\rho=\frac{\pi}{2}-8\varepsilon^{1/4},\qquad h=h_0-8\varepsilon^{1/4}.$$

The subsets $\mathcal{E}_1$ and $\mathcal{E}_2$ of $\mathbb{C}\times(\mathbb{C}/2\pi\mathbb{Z})^d$ now depend on ε and we take $A(\delta,\sigma)=\mathrm{const}\,\varepsilon^{-1}$.

– The mutually reciprocal transformations $\mathrm{Id}+\mu\mathcal{U}:\mathcal{E}_2\to\mathcal{E}_1$ and $\mathrm{Id}+\mu\mathcal{V}:[-u_0,u_0]\times\mathbb{T}^d\to\mathbb{R}\times\mathbb{T}^d$ provided by Proposition 3.3 allow to define

$$\Delta\widetilde{\Sigma}=\Delta\widetilde{S}\circ(\mathrm{Id}+\mu\mathcal{U})\text{ solution of }\big(\partial_v+\varepsilon^{-1/2}\omega\cdot\partial_\theta\big)\Delta\widetilde{\Sigma}=0$$

for $\varepsilon>0$ and $\mu\in[-\mu_0(\varepsilon),\mu_0(\varepsilon)]$, with $\mu_0(\varepsilon)=\mathrm{const}\,\varepsilon^2$. The function $\Delta\widetilde{\Sigma}$ is in fact determined by its restriction

$$\boxed{g_{\varepsilon,\mu}=\Delta\widetilde{\Sigma}_{|v=0}:\theta\in\mathbb{T}^2\ \mapsto\ g_{\varepsilon,\mu}(\theta)=\widetilde{\Delta}\Sigma(0,\theta;\varepsilon,\mu)}$$

according to the rule $\Delta\widetilde{\Sigma}(v,\theta)=g(\theta-\varepsilon^{-1/2}v\omega)$. Thus we can define a diffeomorphism $\Theta_{\varepsilon,\mu}$ of $\mathbb{T}^2$ by

$$\Theta_{\varepsilon,\mu}(\varphi)=\varphi-\varepsilon^{-1/2}u_*\omega+\mu\big(\mathcal{V}_\theta-\varepsilon^{-1/2}\mathcal{V}_v\omega\big)(u_*,\varphi;\varepsilon,\mu)$$

which satisfies

$$\boxed{f_{\varepsilon,\mu}=g_{\varepsilon,\mu}\circ\Theta_{\varepsilon,\mu}.}$$

As soon as $|\mu|$ is small enough, independently of ε, the Jacobian $\mathrm{Jac}\Theta$ of $\Theta_{\varepsilon,\mu}$ at any point has a determinant larger than $1/2$ (because the definition of $\mathcal{V}$ is required here only in the real domain, we need not bound the partial derivatives of its components near the complex singularities). Thus, for $\mu\in[-\mu_0(\varepsilon),\mu_0(\varepsilon)]$, there is a one-to-one correspondence between the critical points of $f_{\varepsilon,\mu}$ and $g_{\varepsilon,\mu}$, any critical point φ_0 of f with Hessian $\mathcal{H}_f(\varphi_0)$ giving rise to a critical point θ_0 of g with Hessian $\mathcal{H}_g(\theta_0)=\mathcal{H}_f(\varphi_0)(\mathrm{Jac}\Theta)^{-2}(\varphi_0)$, and the problem is reduced to the study of the critical points of $g_{\varepsilon,\mu}$.

We shall apply to $g_{\varepsilon,\mu}$ the following lemma, in which $\|\mathrm{d}R\|$ is a notation for $\max_{\mathbb{T}^2}\{|\partial_{\theta_1}R|,|\partial_{\theta_2}R|\}$ and $\|\mathrm{d}^2R\|$ for $\max_{\mathbb{T}^2}\{|\partial^2_{\theta_1}R|,|\partial_{\theta_1}\partial_{\theta_2}R|,|\partial^2_{\theta_2}R|\}$, and whose proof is left to the reader:

LEMMA 4.2. – *Let $A,\ A'\in\mathbb{R}^*$, $\varpi,\varpi'\in\mathbb{R}$, $n\in\mathbb{N}^*$ and $k=k^{(n-1)}$, $k'=k^{(n)}$. Consider the function*

$$\theta\in\mathbb{T}^2\ \mapsto\ P(\theta)=A\cos(k\cdot\theta+\varpi)+A'\cos(k'\cdot\theta+\varpi')$$

and a perturbation $R:\mathbb{T}^2\to\mathbb{R}$ which is a C^2 function such that

$$\|\mathrm{d}R\|\leqslant\frac{\min\{|A|,|A'|\}}{10|k'_1|},\quad\big\|\mathrm{d}^2R\big\|\leqslant\frac{\min\{|A|,|A'|\}}{20|k'_1|^2}.$$

Then the function $P+R$ is a Morse function on $\mathbb{T}^2$, which possesses exactly four distinct critical points. Furthermore, at each critical point θ_0 of $P+R$, the Hessian $\mathcal{H}_{P+R}(\theta_0)$ satisfies

$$\frac{1}{3}|AA'|\leqslant|\mathcal{H}_{P+R}(\theta_0)|\leqslant3|AA'|.$$

For that purpose, we shall identify inside the Fourier expansion of $g_{\varepsilon,\mu}$ the two pairs of largest coefficients (they go in pairs of terms with opposite indices due to realness) and show that the corresponding indices are of the form

$$k_\varepsilon=k^{(n_\varepsilon-1)},\quad k'_\varepsilon=k^{(n_\varepsilon)},\quad n_\varepsilon=\left[\frac{X_*-\log\varepsilon}{4\log\Gamma}\right]$$

($[\cdot]$ denotes the integral part of a real number). Note that isolating the dominant pair (with indices say k_ε and $-k_\varepsilon$) would amount to providing a "monomial" approximation

$$g_{\varepsilon,\mu}(\theta) \approx A_\varepsilon \sin(k_\varepsilon \cdot \theta + \varpi_\varepsilon)$$

which cannot be a Morse function on $\mathbb{T}^2$ for dimensional reasons: we need at least a "binomial" approximation with independent pairs of indices, as in Lemma 4.2.

– Let us now state a refined version of Lemma 4.1 which will produce such approximations. We shall use the notations

$$\forall X \in \mathbb{R}, \quad w'_*(X) = C_* \cosh\left(\left(1 - \left\{\frac{X - X_*}{4\log\Gamma}\right\}\log\Gamma\right)\right)$$

and $W = C_* \cosh(\log\Gamma)$. Observe that $s_* = w_* + w'_*$.

LEMMA 4.3. – *Under the same hypotheses and notations as in Lemma* 4.1*, there exists* $c > 0$ *such that, for* $\varepsilon > 0$ *small enough such that* $\log\varepsilon \in L_\kappa$*,* $\frac{1}{2}\rho_0 < \rho < \rho_0$ *and* $\frac{1}{2}h_0 < h < h_0$*, one can write*

$$\chi(0,\theta,\varepsilon) = A_{k_\varepsilon}\cos(k_\varepsilon \cdot \theta + \varpi_{k_\varepsilon}) + A_{k'_\varepsilon}\cos(k'_\varepsilon \cdot \theta + \varpi_{k'_\varepsilon}) + R_\varepsilon(\theta)$$

with a perturbation R_ε *which satisfies*

$$\|\mathrm{d}R_\varepsilon\|,\ \left\|d^2 R_\varepsilon\right\| \leqslant cB(\rho,h,\varepsilon)\exp\bigl(-W\varepsilon^{-1/4}\bigr),$$

and real numbers A_k, ϖ_k *depending on* ε *and defined by the Fourier coefficients of* $\chi_{|v=0}$ *as*

$$\chi_k(0,\varepsilon) = 2A_k \mathrm{e}^{\mathrm{i}\varpi_k}.$$

Moreover the absolute value of one of the numbers $A_{k_\varepsilon}, A_{k'_\varepsilon}$ *is bounded by*

$$cB(\rho,h,\varepsilon)\exp\bigl(-w_*(\log\varepsilon)\varepsilon^{-1/4}\bigr),$$

and the absolute value of the other is bounded by $cB(\rho,h,\varepsilon)\exp(-w'_*(\log\varepsilon)\varepsilon^{-1/4})$*, depending on whether* n_ε *or* $n_\varepsilon + 1$ *is closest to* $\frac{X_* - \log\varepsilon}{4\log\Gamma}$.

For the sake of simplicity, we shall assume in the sequel that n_ε is closer to $\frac{X_* - \log\varepsilon}{4\log\Gamma}$ than $n_\varepsilon + 1$. The proof of Lemma 4.3 follows the same lines as the proof of Lemma 4.1 and we omit it. We only indicate that the restriction $\log\varepsilon \in L_\kappa$ is meant to ensure gaps in the hierarchy of Fourier indices, in the sense that

$$\begin{aligned}
h|k_\varepsilon| + \rho\varepsilon^{-1/2}|k_\varepsilon \cdot \omega| &\approx w_*(\log\varepsilon)\varepsilon^{-1/4},\\
h|k'_\varepsilon| + \rho\varepsilon^{-1/2}|k'_\varepsilon \cdot \omega| &\approx w'_*(\log\varepsilon)\varepsilon^{-1/4},\\
h|k^{(n)}| + \rho\varepsilon^{-1/2}|k^{(n)} \cdot \omega| &\geqslant W\varepsilon^{-1/4} \quad \text{if } n \notin \{n_\varepsilon - 1, n_\varepsilon\},\\
\max w_* < \min w'_* &< \max w'_* < W.
\end{aligned}$$

Moreover, when applied to the function $\Delta\widetilde{S}_1$, the chain of reasoning which leads to Lemma 4.3 yields also bounds from below for the main coefficients:

$$\Delta\widetilde{S}_1(0,\theta,\varepsilon) = A^{(1)}_{k_\varepsilon}\cos\bigl(k_\varepsilon \cdot \theta + \varpi^{(1)}_{k_\varepsilon}\bigr) + A^{(1)}_{k'_\varepsilon}\cos\bigl(k'_\varepsilon \cdot \theta + \varpi^{(1)}_{k'_\varepsilon}\bigr) + R^{(1)}_\varepsilon(\theta)$$

with $\|\mathrm{d}R_\varepsilon^{(1)}\|,\ \|d^2R_\varepsilon^{(1)}\| \leqslant \text{const}\,\varepsilon^{1/2}\mathrm{e}^{-W\varepsilon^{-1/4}}$ and

$$\big|A_{k_\varepsilon}^{(1)}\big| \asymp \text{const}\,\varepsilon^{1/4}\mathrm{e}^{-w_*(\log\varepsilon)\varepsilon^{-1/4}}, \quad \big|A_{k'_\varepsilon}^{(1)}\big| \asymp \text{const}\,\varepsilon^{1/4}\mathrm{e}^{-w'_*(\log\varepsilon)\varepsilon^{-1/4}},$$

where he symbol $\asymp$ means that the left-hand side can be bounded from above and from below by expressions like those in the right-hand side (use formula (4.11) and the inequalities (3.2)).

When applied to the function $\chi = \Delta\widetilde{\Sigma} - \mu\Delta\widetilde{S}_1$, which satisfies

$$|\chi| \leqslant \text{const}\,\varepsilon^{-3}|\mu|^2 \quad \text{in } \overline{\mathcal{E}}_2,$$

Lemma 4.3 yields

$$g_{\varepsilon,\mu}(\theta) - \mu\Delta\widetilde{S}_1(0,\theta,\varepsilon) = \mu^2\big(A_{k_\varepsilon}^{(2)}\cos\big(k_\varepsilon\cdot\theta+\varpi_{k_\varepsilon}^{(2)}\big) + A_{k'_\varepsilon}^{(2)}\cos\big(k'_\varepsilon\cdot\theta+\varpi_{k'_\varepsilon}^{(2)}\big) + R_{\varepsilon,\mu}^{(2)}\big)$$

with

$$\big|A_{k_\varepsilon}^{(2)}\big| \leqslant \text{const}\,\varepsilon^{-3}\mathrm{e}^{-w_*(\log\varepsilon)\varepsilon^{-1/4}}, \quad \big|A_{k'_\varepsilon}^{(2)}\big| \leqslant \text{const}\,\varepsilon^{-3}\mathrm{e}^{-w'_*(\log\varepsilon)\varepsilon^{-1/4}},$$

and

$$\big\|\mathrm{d}R_\varepsilon^{(2)}\big\|,\ \big\|d^2R_\varepsilon^{(1)}\big\| \leqslant \text{const}\,\varepsilon^{-3}\mathrm{e}^{-W\varepsilon^{-1/4}}.$$

– We are now in a position to apply Lemma 4.2 to

$$\boxed{g_{\varepsilon,\mu}(\theta) = A_{\varepsilon,\mu}\cos(k_\varepsilon\cdot\theta+\varpi_{\varepsilon,\mu}) + A'_{\varepsilon,\mu}\cos(k'_\varepsilon\cdot\theta+\varpi'_{\varepsilon,\mu}) + R_{\varepsilon,\mu}(\theta)}$$

with

$$A_{\varepsilon,\mu} = \mu A_{k_\varepsilon}^{(1)} + \mu^2 A_{k_\varepsilon}^{(2)}, \quad A'_{\varepsilon,\mu} = \mu A_{k'_\varepsilon}^{(1)} + \mu^2 A_{k'_\varepsilon}^{(2)}, \quad R_{\varepsilon,\mu} = \mu R_\varepsilon^{(1)} + \mu^2 R_{\varepsilon,\mu}^{(2)}.$$

This yields the desired information on the number of critical points and the Hessian at them for $g_{\varepsilon,\mu}$, and thus for $f_{\varepsilon,\mu}$. □

5. Proof of Theorem 3.1: domain of analyticity for the solution $S^\pm$

5.1. Method

We shall content ourselves with the case of $\widetilde{S}^+$. Let us fix once for all $u_1, \Delta\beta, \delta, \sigma$; we suppose of course that the half-aperture of the sector $\mathcal{D}^+_{u_1,\delta}$ is strictly larger than $\Delta\beta$.

In the particular cases where $d = 1$ or $\alpha = 0$, we could treat directly the Hamilton–Jacobi equation, but for the general situation we can only propose an indirect method. The idea is that the Hamilton–Jacobi equation does not tell much – at least explicitly – about the dynamics on the invariant manifold under concern, and for that reason we could not rephrase it as a fixed-point integral equation involving only bounded operators (or perhaps we should say that all the information is contained in the Hamilton–Jacobi equation but we do not know how to extract it in an efficient manner). More specifically, this unboundedness phenomenon takes place with respect to the angular variables φ_j and not u: when we solve the linearized equation $\widetilde{D}_0\chi = \psi$ in the space of the functions decreasing exponentially fast when u tends to $+\infty$, we find $\chi = E^+\psi$ where E^+ is a bounded operator of the suitable Banach space, but beside the operator $\widetilde{D}_0$

the Hamilton–Jacobi involves separately the partial derivatives of the unknown function, which means that we should study the operators $\partial_u \circ E^+$ and $\partial_{\varphi_j} \circ E^+$ too; the first of them is a bounded operator, but not the others when $d \geqslant 2$, hence the difficulty (this difficulty does not arise if $\alpha = 0$). But we shall see that the operators $\partial_{\varphi_j} \circ E^+$ are no longer involved when the Hamilton–Jacobi equation is bypassed through dynamical considerations.

The alternative strategy which we propose consists in finding first a parametrization of the invariant manifold

$$\mathcal{W}^+ \colon \begin{cases} q = q_0(u), \\ p = \dot{q}_0(u) + \mu P(u, \theta; z, \mu), \\ \varphi = \theta + \mu \Phi(u, \theta; z, \mu), \\ I = \mu J(u, \theta; z, \mu), \end{cases} \tag{5.1}$$

which makes use of the plain variable u and of new angular variables θ_j with respect to which the pull-back of the restricted Hamiltonian vector field is straightened, i.e. can be written

$$\begin{cases} \dot{u} = 1 + \mathrm{O}(\mu), \\ \dot{\theta} = z\omega. \end{cases}$$

Requiring moreover that the functions P, Φ_j, J_j decrease exponentially fast when u goes to $+\infty$, we shall be able to apply the fixed-point theorem in some Banach space (the operators $\partial_{\theta_j} \circ E^+$ won't appear thanks to the straightening condition) and to obtain the analytic extension of the solution to the domain

$$u \in \mathcal{D}^+_{u_1,\delta}, \quad \theta \in \mathbb{T}^d_{h_0 - \frac{7\sigma}{10}}, \quad z \in \Sigma_{u_1,\delta,\Delta\beta}, \quad |\mu| < \mu_1',$$

for some positive μ_1'.

This means that we shall replace the representation of the invariant manifold as a Lagrangian graph by a parametrization which will carry a better control of the dynamics on it because it will correspond to the stable foliation:[8] by fixing $\theta \in \mathbb{T}^d$ in the formulas (5.1) and letting u vary, one would obtain the parametrization of a curve $\mathcal{W}^+_{\underline{\theta}}$ which is a part of the stable manifold of the point $\underline{\theta} = (2\pi, \theta, 0, 0) \in \mathcal{T}$, i.e.

$$\mathcal{W}^+_{\underline{\theta}} = \big\{ M \in \mathrm{T}^*\mathfrak{C} \mid \operatorname{dist}\big(\phi^t_H(M), \phi^t_H(\underline{\theta})\big) \text{ tends to } 0 \text{ exponentially fast as } t \text{ tends to } +\infty \big\},$$

denoting $\mathfrak{C} =]2\pi - q_0, 2\pi[\times \mathbb{T}^d$ a part of the configuration space, $\mathrm{T}^*\mathfrak{C}$ the corresponding part of the phase space, and φ^t_H the time-t map of the Hamiltonian flow associated to $H_{\varepsilon,\mu}$.

Then we shall solve the inversion problem

$$\varphi = \theta + \mu \Phi(u, \theta; z, \mu) \iff \theta = \varphi + \mu \Theta(u, \varphi; z, \mu)$$

and find functions Θ_j analytic for

$$u \in \mathcal{D}^+_{u_1,\delta}, \quad \varphi \in \mathbb{T}^d_{h_0 - \frac{9\sigma}{10}}, \quad z \in \Sigma_{u_1,\delta,\Delta\beta}, \quad |\mu| < \mu_2',$$

[8] In fact our method is close to the usual one for proving the stable manifold theorem in finite differentiability (*see* [30], [15]), and somewhat easier since the analyticity with respect to θ comes for free.

for some positive μ_2'. For $|\mu|$ small enough, we shall do the substitution:

$$\mathcal{W}^+\colon \begin{cases} q = q_0(u), \\ p = \dot{q}_0(u) + \mu P(u, \varphi + \mu\Theta(u,\varphi;z,\mu);z,\mu), \\ \varphi = \varphi, \\ I = \mu J(u, \varphi + \mu\Theta(u,\varphi;z,\mu);z,\mu), \end{cases}$$

and there will remain only to perform an integration in order to recover the function $\widetilde{S}^+$.

5.2. Fixed-point theorem and various Banach algebras

We shall use twice (to find the parametrization and to solve the inversion problem) the following classical fixed-point theorem:

THEOREM 5.1. – *Let $\mathbb{B}$ be a Banach space and μ^*, c', c'' positive numbers. Denote by $\|\cdot\|$ the norm of $\mathbb{B}$ and suppose that for every complex number μ such that $|\mu| \leqslant \mu^*$, a mapping $\mathcal{F}_\mu$ is defined from the ball $\mathcal{A} = \{X \in \mathbb{B};\ \|X\| \leqslant 2c'\}$ to $\mathbb{B}$, which satisfies*

$$\|\mathcal{F}_\mu(0)\| \leqslant c' \quad \textit{and} \quad \forall X, X' \in \mathcal{A},\ \|\mathcal{F}_\mu(X') - \mathcal{F}_\mu(X)\| \leqslant c''|\mu|\|X' - X\|.$$

Then, for $|\mu| \leqslant \inf\{\mu^*, \frac{1}{2c''}\}$, the mapping $\mathcal{F}_\mu$ admits a unique fixed point $\underline{X}_\mu$ in $\mathcal{A}$, the sequence of iterates $(\mathcal{F}_\mu^m(0))_{m\geqslant 0}$ is well defined and converges to $\underline{X}_\mu$ in $\mathbb{B}$ uniformly with respect to μ, and

$$\|\underline{X}_\mu\| \leqslant 2c', \quad \|\underline{X}_\mu - \mathcal{F}_\mu(0)\| \leqslant 2c'c''|\mu|.$$

Let us define the Banach algebras[9] which our unknown functions will belong to. We begin with spaces for their Fourier coefficients. Let $\mathcal{B}_0^+$ be the Banach algebra of all analytic functions in the domain $\mathcal{D}^+_{u_1,\delta} \times \Sigma_{u_1,\delta,\Delta\beta}$ which extend continuously to bounded functions on its closure $\overline{\mathcal{D}}^+_{u_1,\delta} \times \overline{\Sigma}_{u_1,\delta,\Delta\beta}$, equipped with the supremum norm. If $s \in \mathbb{N}$, we shall use the notations

$$\mathcal{B}_s^+ = \big\{\psi \in \mathcal{B}_0^+;\ \|\psi\|_s < \infty\big\}, \quad \|\psi\|_s := \sup_{(u,z)\in\overline{\mathcal{D}}^+_{u_1,\delta}\times\overline{\Sigma}_{u_1,\delta,\Delta\beta}} \big\{\mathrm{e}^{su_1 + s\,\Re\mathrm{e}\,u}|\psi(u,z)|\big\}.$$

Since $u_1 + \Re\mathrm{e}\,u \geqslant 0$ for all u in $\overline{\mathcal{D}}^+_{u_1,\delta}$ this defines a decreasing sequence of Banach algebras $\mathcal{B}_s^+$ (with weighted norms $\|\cdot\|_s$). We define also the Banach algebras

$$\underline{\mathcal{B}}_s^+ \big\{\psi \in \mathcal{B}_0^+;\ \|\psi\|_{\underline{s}} < \infty\big\}, \quad \|\psi\|_{\underline{s}} := \|\psi\|_s + \left\|\frac{\partial\psi}{\partial u}\right\|_s.$$

If $\psi \in \mathcal{B}_{s+1}^+$, $\|\psi\|_s \leqslant \|\psi\|_{s+1}$; if $\psi \in \mathcal{B}_s^+$ and $\chi \in \mathcal{B}_t^+$, $\psi\chi \in \mathcal{B}_{s+t}^+$ and $\|\psi\chi\|_{s+t} \leqslant \|\psi\|_s\|\chi\|_t$. And these properties hold also for the sequence $(\underline{\mathcal{B}}_s^+)_{s\geqslant 0}$.

The corresponding Banach algebras $\mathcal{B}_s^+(h)$ and $\underline{\mathcal{B}}_s^+(h)$ are defined according to the following construction: if $\mathcal{B}$ is a Banach algebra (with norm $\|\cdot\|$) and h a positive number, we consider the Fourier series in d angular variables with coefficients in $\mathcal{B}$ and introduce the notation

$$\|\psi\|_h = \sum_{k\in\mathbb{Z}^d} \mathrm{e}^{|k|h}\|\psi_k\| \quad \text{if } \psi = \sum_{k\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}k\cdot\theta}\psi_k. \tag{5.2}$$

[9] All the Banach algebras that we consider are commutative Banach algebras over $\mathbb{C}$ in which the product of two elements ψ, χ satisfies the inequality $\|\psi\chi\| \leqslant \|\psi\|\|\chi\|$.

Here of course, $k\cdot\theta = k_1\theta_1 + \cdots + k_d\theta_d$ and $|k| = |k_1| + \cdots + |k_d|$. It defines a Banach norm on the algebra

$$\mathcal{B}(h) = \left\{ \psi = \sum_{k\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}k\cdot\theta}\psi_k \text{ with } \psi_k \in \mathcal{B} \text{ such that } \|\psi\|_h < \infty \right\}. \tag{5.3}$$

Any Fourier series ψ in $\mathcal{B}(h)$ converges to a continuous mapping ψ from $\overline{\mathbb{T}}_h^d$ to $\mathcal{B}$ which satisfies the inequality $\sup_{\overline{\mathbb{T}}_h^d} \|\psi\| \leqslant \|\psi\|_h$. Moreover, if $h' < h$, it also satisfies (*see* [24])

$$\|\psi\|_{h'} \leqslant \coth^d\left(\frac{h-h'}{2}\right) \sup_{\overline{\mathbb{T}}_h^d} \|\psi\|. \tag{5.4}$$

One checks easily the inequality for the product:

$$\forall \psi, \chi \in \mathcal{B}(h), \quad \|\psi\chi\|_h \leqslant \|\psi\|_h \|\chi\|_h,$$

and if $0 < h < h^*$, the Cauchy inequalities:

$$\forall \psi \in \mathcal{B}(h^*),\ \forall r \in \mathbb{N}^d, \quad \left\|(\partial_\theta)^r \psi\right\|_h \leqslant r_1! \cdots r_d! (h^* - h)^{-|r|} \|\psi\|_{h^*}.$$

(*See* Appendix A.)

Such a construction was already used in [24,25]. The reason why we too use this kind of norm for Fourier series will appear in Lemma 5.2. We shall denote by $\|\cdot\|_{s,h}$ and $\|\cdot\|_{\underline{s},h}$ the norms of $\mathcal{B}_s^+(h)$ and $\underline{\mathcal{B}}_s^+(h)$. These are spaces of functions analytic in the domain $\mathcal{D}_{u_1,\delta}^+ \times \mathbb{T}_h^d \times \Sigma_{u_1,\delta,\Delta\beta}$ and bounded in its closure.

For the parametrization of $\mathcal{W}^+$ and the inverse change of angular variables which we described in Section 5.1, we shall obtain

$$X = (P, J, \Phi) \in \underline{\mathcal{B}}_1^+(h_1) \times \left[\underline{\mathcal{B}}_2^+(h_1)\right]^d \times \left[\underline{\mathcal{B}}_2^+(h_1)\right]^d, \quad \Theta \in \left[\underline{\mathcal{B}}_2^+(h_2)\right]^d,$$

with $h_1 = h_0 - \frac{7\sigma}{10}$, $h_2 = h_0 - \frac{9\sigma}{10}$ (and X and Θ will depend analytically on μ). When dealing with products of Banach spaces of this kind, we shall use as a norm the supremum of the norms of the components. Let us give already two lemmas which will prove useful:

LEMMA 5.1. – *Let $\mathcal{B}$ be a Banach algebra and $0 < h < h^*$. Suppose $G \in \mathcal{B}(h^*)$ with $\partial_\theta G \in [\mathcal{B}(h^*)]^d$, and let $\mu \in \mathbb{C}$ and $\mathcal{A} = \left\{\psi \in [\mathcal{B}(h)]^d;\ \|\mu\psi\|_h \leqslant \frac{h^*-h}{2}\right\}$. The formula*

$$\forall \psi \in \mathcal{A}, \quad \mathcal{G}(\psi) = G \circ (\mathrm{Id} + \mu\psi) := \sum_{r\in\mathbb{N}^d} \frac{\mu^{|r|}}{r_1! \cdots r_d!} \psi_1^{r_1} \cdots \psi_d^{r_d} (\partial_\theta)^r G$$

defines a mapping from $\mathcal{A}$ to $\mathcal{B}(h)$ which satisfies:

$$\forall \psi, \psi' \in \mathcal{A}, \quad \|\mathcal{G}(\psi') - \mathcal{G}(\psi)\|_h \leqslant 2^d d|\mu| \|\partial_\theta G\|_{h^*} \|\psi' - \psi\|_h.$$

(Proof in Appendix A.)

For the second one, we introduce the notation

$$p_0(u) = \dot{q}_0(u) = \frac{2}{\cosh u}$$

and we note that $p_0(u)^{-1} = \frac{1}{2}\mathrm{e}^u(1+\mathrm{e}^{-2u})$, thus

$$\forall u \in \overline{\mathcal{D}}^+_{u_1,\delta}, \quad |p_0(u)^{-1}| \leqslant \lambda \mathrm{e}^{u_1+\Re\mathrm{e}\, u} \quad \text{and} \quad \left|\frac{\mathrm{d}(p_0^{-1})}{\mathrm{d}u}(u)\right| \leqslant \lambda \mathrm{e}^{u_1+\Re\mathrm{e}\, u},$$

with $\lambda = \sup\left\{1, \frac{1}{2}\cosh u_1\right\}$ (λ depends only on u_1).

LEMMA 5.2. – *Let $h>0$ and $s \geqslant 1$. The formula*

$$\forall \psi \in \mathcal{B}^+_s(h), \quad (E^+\psi)(u,\theta,z) = -\int_0^{z^{-1}\infty} \psi(u+\zeta, \theta+z\zeta\omega, z)\,\mathrm{d}\zeta$$

defines an operator $E^+ : \mathcal{B}^+_s(h) \to \underline{\mathcal{B}}^+_s(h)$ which is an inverse to the right for $\widetilde{D}_0 = \frac{\partial}{\partial u} + z\omega\cdot\frac{\partial}{\partial\theta}$ (i.e. $\widetilde{D}_0 \circ E^+$ is defined and equal to Id*). Moreover, for all $\psi \in \mathcal{B}^+_s(h)$,*

$$\|E^+\psi\|_{\underline{s},h} \leqslant \kappa\|\psi\|_{s,h} \quad \textit{and} \quad \|p_0^{-1}E^+\psi\|_{\underline{s-1},h} \leqslant \lambda\kappa\|\psi\|_{s,h}$$

with $\kappa = 4/\sin\Delta\beta$ ($\kappa > 4$ and κ depends only on $\Delta\beta$).

Proof. – We suppose $h>0$, $s\geqslant 1$ and $\psi = \sum_{k\in\mathbb{Z}^d}\psi_k(u,z)\mathrm{e}^{\mathrm{i}k\cdot\theta} \in \mathcal{B}^+_s(h)$. Let β_1 the half-aperture of the sector $\mathcal{D}^+_{u_1,\delta}$. For $k\in\mathbb{Z}^d$, we define a function χ_k:

$$\forall (u,z) \in \overline{\mathcal{D}}^+_{u_1,\delta} \times \overline{\Sigma}_{u_1,\delta,\Delta\beta}, \quad \chi_k(u,z) = -\int_0^{z^{-1}\infty} \psi_k(u+\zeta,z)\mathrm{e}^{\mathrm{i}z\zeta k\cdot\omega}\,\mathrm{d}\zeta \tag{5.5}$$

(observe that $|\arg\zeta| = |\arg z^{-1}| \leqslant \beta_1$, thus $u+\zeta \in \overline{\mathcal{D}}^+_{u_1,\delta}$). This integral converges, since $s\geqslant 1$, $|\psi_k(u+\zeta,z)| \leqslant \mathrm{e}^{-su_1 - s\,\Re\mathrm{e}\,u}\|\psi_k\|_s \mathrm{e}^{-s\,\Re\mathrm{e}\,\zeta}$ and $\Re\mathrm{e}\,\zeta = |\zeta|\cos(\arg z)$ with $\cos(\arg z) \geqslant \sin\Delta\beta$ (because $|\arg z| \leqslant \frac{\pi}{2} - \Delta\beta$). Moreover, the inequalities

$$|\chi_k(u,z)| \leqslant \mathrm{e}^{-su_1 - s\,\Re\mathrm{e}\,u}\|\psi_k\|_s \sin^{-1}\Delta\beta$$

show that $E^+\psi = \chi = \sum_{k\in\mathbb{Z}^d}\chi_k(u,z)\mathrm{e}^{\mathrm{i}k\cdot\theta}$ belongs to $\mathcal{B}^+_s(h)$ and

$$\|\chi\|_{s,h} \leqslant \frac{\kappa}{4}\|\psi\|_{s,h}$$

with $\kappa = 4\sin^{-1}\Delta\beta$. Clearly, χ satisfies the partial differential equation

$$\widetilde{D}_0\chi = \psi$$

in $\mathcal{D}^+_{u_1,\delta}\times\mathbb{T}^d_h\times\Sigma_{u_1,\delta,\Delta\beta}$.

Let $(u,z) \in \overline{\mathcal{D}}^+_{u_1,\delta}\times\overline{\Sigma}_{u_1,\delta,\Delta\beta}$. The Cauchy theorem allows us to change the half-line of integration in (5.5). If $k\cdot\omega > 0$, we increase the slope of the half-line in order to take advantage of the decrease of the exponential $\mathrm{e}^{\mathrm{i}z\zeta k\cdot\omega}$:

$$\chi_k(u,z) = -\int_0^{+\infty} \psi_k\big(u+\xi\mathrm{e}^{\mathrm{i}\beta_1}\big)\mathrm{e}^{\mathrm{i}z\mathrm{e}^{\mathrm{i}\beta_1}\xi k\cdot\omega}\mathrm{e}^{\mathrm{i}\beta_1}\,\mathrm{d}\xi,$$

and we obtain a new bound:

$$\|\chi_k\|_s \leqslant \frac{1}{s\cos\beta_1 + |z|k\cdot\omega\sin(\beta_1+\arg z)}\|\psi_k\|_s.$$

Since $0<\beta_1\leqslant\frac{\pi}{2}$ and $\Delta\beta\leqslant\beta_1+\arg z\leqslant\pi-\Delta\beta$, we end up with

$$\|zk\cdot\omega\chi_k\|_s \leqslant \frac{\kappa}{4}\|\psi_k\|_s.$$

Similarly, if $k\cdot\omega<0$, we decrease the slope of the half-line of integration, and the previous inequality holds true in all cases (even if $k\cdot\omega=0$). But $\partial_u\chi_k=\psi_k-\mathrm{i}zk\cdot\omega\chi_k$ because of the partial differential equation, thus for all $k\in\mathbb{Z}^d$, $\|\partial_u\chi_k\|_s\leqslant\frac{\kappa}{2}\|\psi_k\|_s$ and

$$\|\partial_u\chi\|_{s,h}\leqslant\frac{\kappa}{2}\|\psi\|_{s,h}.$$

Lastly, the inequalities $(\forall u\in\overline{\mathcal{D}}^+_{u_1,\delta})\ |p_0(u)^{-1}|, \left|\frac{\mathrm{d}(p_0^{-1})}{\mathrm{d}u}(u)\right|\leqslant\lambda\mathrm{e}^{u_1+\Re\mathrm{e}\,u}$ imply that the function $p_0^{-1}\chi$ belongs to $\mathcal{B}^+_{s-1}(h)$ and

$$\left\|p_0^{-1}\chi\right\|_{s-1,h}\leqslant\lambda\|\chi\|_{s,h},$$

and its derivative $\partial_u(p_0^{-1}\chi)=\frac{\mathrm{d}(p_0^{-1})}{\mathrm{d}u}\chi+p_0^{-1}\partial_u\chi$ belongs to $\mathcal{B}^+_{s-1}(h)$ too with

$$\left\|\partial_u(p_0^{-1}\chi)\right\|_{s-1,h}\leqslant\lambda\|\chi\|_{s,h}+\lambda\|\partial_u\chi\|_{s,h},$$

since $\left|\frac{\mathrm{d}(p_0^{-1})}{\mathrm{d}u}(u)\right|\leqslant\lambda\mathrm{e}^{u_1+\Re\mathrm{e}\,u}$. Hence

$$\left\|p_0^{-1}\chi\right\|_{\underline{s-1},h}\leqslant\lambda\big[2\|\chi\|_{s,h}+\|\partial_u\chi\|_{s,h}\big]\leqslant\lambda\kappa\|\psi\|_{s,h}. \qquad \square$$

5.3. The (u,θ)-parametrization

Let $P, J_1,\dots,J_d,\Phi_1,\dots,\Phi_d$ be analytic functions of $(u,\theta_1,\dots,\theta_d,z)$ 2π-periodic in the θ_j and consider the manifold

$$\mathcal{W}=\big\{(q,\varphi,p,I)=\big(q_0(u),\theta+\mu\Phi(u,\theta,z),p_0(u)+\mu P(u,\theta,z),\mu J(u,\theta,z)\big)\big\}.$$

A necessary and sufficient condition for the Hamiltonian vector field

$$\begin{cases}\dot q=p,\\ \dot p=\sin q-\mu\partial_q F(q,\varphi),\end{cases}\qquad\begin{cases}\dot\varphi=z\omega+\alpha I,\\ \dot I=-\mu\partial_\varphi F(q,\varphi),\end{cases}$$

to leave $\mathcal{W}$ invariant and for its pull-back to be of the form

$$\begin{cases}\dot u=1+\mathrm{O}(\mu),\\ \dot\theta=z\omega,\end{cases}$$

is that $X=(P,J,\Phi)$ satisfy

$$\text{(5.6)}\qquad \begin{cases} \widetilde{D}_0(p_0P) = -\partial_u\widetilde{F}(u,\theta+\mu\Phi(u,\theta,z)) - \mu P\partial_u P, \\ \widetilde{D}_0 J = -\partial_\varphi\widetilde{F}(u,\theta+\mu\Phi(u,\theta,z)) - \mu p_0^{-1}P\partial_u J, \\ \widetilde{D}_0\Phi = \alpha J - \mu p_0^{-1}P\partial_u\Phi. \end{cases}$$

PROPOSITION 5.1. – *Let $h_1 = h_0 - \frac{7\sigma}{10}$. There exists a positive number μ_1' such that system* (5.6) *admits a solution X in $\mathbb{B} = \underline{\mathcal{B}}_1^+(h_1)\times[\underline{\mathcal{B}}_2^+(h_1)]^d\times[\underline{\mathcal{B}}_2^+(h_1)]^d$ for $|\mu|\leqslant\mu_1'$. Moreover, the solution $X=(P,J,\Phi)$ depends analytically on μ and there exist $B_1,C_1>0$ such that*

$$\|P\|_{\underline{1},h_1}\leqslant C_1,\quad \|J\|_{\underline{2},h_1}\leqslant C_1,\quad \|\Phi\|_{\underline{2},h_1}\leqslant C_1,$$

$$\|P+p_0^{-1}E^+\partial_u\widetilde{F}\|_{\underline{1},h_1}\leqslant B_1|\mu|,\quad \|J+E^+\partial_\varphi\widetilde{F}\|_{\underline{2},h_1}\leqslant B_1|\mu|.$$

If $\delta\leqslant\sigma\leqslant\lambda^{-2}$, one can take

$$\mu_1' = b_1^{-1}A^{-1}\delta\sigma,\quad B_1 = b_1A^2\delta^{-2}\sigma^{-1},\quad C_1 = b_1A\delta^{-1},$$

where $A=\sigma^{-d}A(\delta/2,\sigma/2)$ and the positive number b_1 depends only on u_1 and $\Delta\beta$.

Proof. – Let us define a mapping $\mathcal{F}_\mu$ of $\mathbb{B}$ in itself by the formulas

$$\mathcal{F}_\mu(P,J,\Phi) = (P^*,J^*,\Phi^*),$$

$$\begin{cases} P^* = -p_0^{-1}E^+(\partial_u\widetilde{F}(u,\theta+\mu\Phi(u,\theta,z)) + \mu P\partial_u P), \\ J^* = -E^+(\partial_\varphi\widetilde{F}(u,\theta+\mu\Phi(u,\theta,z)) + \mu p_0^{-1}P\partial_u J), \\ \Phi^* = E^+(\alpha J^* - \mu p_0^{-1}P\partial_u\Phi). \end{cases}$$

Our goal is to apply Theorem 5.1, since a fixed point of $\mathcal{F}_\mu$ is a solution of (5.6). We first note that $\mathcal{F}_\mu(0)=X^{(1)}=(P^{(1)},J^{(1)},\Phi^{(1)})$ with

$$P^{(1)} = -p_0^{-1}E^+\partial_u\widetilde{F},\quad J^{(1)} = -E^+\partial_\varphi\widetilde{F},\quad \Phi^{(1)} = E^+\big(\alpha J^{(1)}\big).$$

In order to indicate explicit constants in the course of the demonstration, we shall assume $\delta\leqslant\sigma\leqslant\lambda^{-2}$ right from the beginning. But it will be clear that, if this is not the case, there still exist constants which satisfy the desired inequalities: only the formulas for them may then differ.

LEMMA 5.3. – *Let $A=\sigma^{-d}A(\delta/2,\sigma/2)$ and $h_1=h_0-\frac{7\sigma}{10}$. There exists a positive constant c* (*which depends only on u_1*) *such that, for any $j,j'\in\{1,\dots,d\}$, the following inequalities hold*:

$$\|\partial_{\varphi_j}\widetilde{F}\|_{2,h_1+\frac{\sigma}{10}}\leqslant cA\sigma^{-1}\quad \textit{and}\quad \|\partial_u\widetilde{F}\|_{2,h_1+\frac{\sigma}{10}}\leqslant cA\delta^{-1},$$

$$\|\partial_{\varphi_j}\partial_{\varphi_j'}\widetilde{F}\|_{2,h_1+\frac{\sigma}{10}}\leqslant cA\sigma^{-2}\quad \textit{and}\quad \|\partial_u\partial_{\varphi_j}\widetilde{F}\|_{2,h_1+\frac{\sigma}{10}}\leqslant cA\delta^{-1}\sigma^{-1}.$$

(Proof in Appendix B.)

The lemma implies that

$$\|\partial_u\widetilde{F}\|_{2,h_1}\leqslant cA\delta^{-1},\quad \|\partial_\varphi\widetilde{F}\|_{2,h_1}\leqslant cA\sigma^{-1}\leqslant cA\delta^{-1},$$

and thus, by virtue of Lemma 5.2,

$$\big\|P^{(1)}\big\|_{\underline{1},h_1}\leqslant\lambda\kappa cA\delta^{-1},\quad \big\|J^{(1)}\big\|_{\underline{2},h_1}\leqslant\kappa cA\delta^{-1}.$$

$$\big\|\Phi^{(1)}\big\|_{\underline{2},h_1}\leqslant\sup\{|\alpha_1|,\dots,|\alpha_d|\}\kappa^2cA\delta^{-1}.$$

Let $b_1'=\lambda\kappa^2c\sup\{1,|\alpha_1|,\dots,|\alpha_d|\}$: this positive number depends only on u_1 and $\Delta\beta$ and allows to bound the first approximation $X^{(1)}$ by

$$\big\|X^{(1)}\big\|\leqslant c_1'=b_1'A\delta^{-1}.$$

By applying Lemma 5.1 to the Banach algebra $\mathcal{B}_2^+$, we see that whenever

$$|\mu|\leqslant\mu_1^*=\frac{\sigma}{40c_1'}\quad\text{and}\quad\|\Phi\|_{2,h_1}\leqslant 2c_1',$$

the composition of $\partial_u\widetilde{F}$ or $\partial_\varphi\widetilde{F}$ with $\mathrm{Id}+\mu\Phi$ defines an element of $\mathcal{B}_2^+(h_1)$, thus the mapping $\mathcal{F}_\mu$ is well defined on

$$\mathcal{A}=\{X\in\mathbb{B};\ \|X\|\leqslant 2c_1'\}$$

provided that $|\mu|\leqslant\mu_1^*$.

Let us now compute a constant $c''=c_1''$ which will satisfy the assumptions of Theorem 5.1. Suppose that $|\mu|\leqslant\mu_1^*$ and $X,X'\in\mathcal{A}$. Let $X^*=\mathcal{F}_\mu(X),X'^*=\mathcal{F}_\mu(X')$. We must study the components P'^*-P^*, J'^*-J^*, $\Phi'^*-\Phi^*$ of X'^*-X^*.

Lemmas 5.1 and 5.3 show that

$$\big\|\partial_u\widetilde{F}\circ(\mathrm{Id}+\mu\Phi')-\partial_u\widetilde{F}\circ(\mathrm{Id}+\mu\Phi)\big\|_{2,h_1}\leqslant 2^ddcA\delta^{-1}\sigma^{-1}|\mu|\|X'-X\|\tag{5.7}$$

and

$$\big\|\partial_\varphi\widetilde{F}\circ(\mathrm{Id}+\mu\Phi')-\partial_\varphi\widetilde{F}\circ(\mathrm{Id}+\mu\Phi)\big\|_{2,h_1}\leqslant 2^ddcA\sigma^{-2}|\mu|\|X'-X\|.\tag{5.8}$$

We observe also that the identity $P'\partial_uP'-P\partial_uP=(P'-P)\partial_uP'+P\partial_u(P'-P)$ implies

$$\|P'\partial_uP'-P\partial_uP\|_{2,h_1}\leqslant\|P'-P\|_{1,h_1}\|P'\|_{\underline{1},h_1}+\|P\|_{1,h_1}\|P'-P\|_{\underline{1},h_1}\leqslant 4c_1'\|P'-P\|_{\underline{1},h_1},$$

hence

$$\|P'\partial_uP'-P\partial_uP\|_{2,h_1}\leqslant 4b_1'A\delta^{-1}\|X'-X\|.\tag{5.9}$$

And the identity $p_0^{-1}P'\partial_uJ'-p_0^{-1}P\partial_uJ=p_0^{-1}(P'-P)\partial_uJ'+p_0^{-1}P\partial_u(J'-J)$ together with the inequalities $|p_0^{-1}(u)|\leqslant\lambda\mathrm{e}^{u_1+\Re\mathrm{e}\,u}$ $(\forall u\in\mathcal{D}^+_{u_1,\delta})$ imply

$$\begin{aligned}&\big\|p_0^{-1}P'\partial_uJ'-p_0^{-1}P\partial_uJ\big\|_{2,h_1}\\&\quad\leqslant\lambda\|P'-P\|_{1,h_1}\|J'\|_{\underline{2},h_1}+\lambda\|P\|_{1,h_1}\|J'-J\|_{\underline{2},h_1}\leqslant 4\lambda c_1'\|X'-X\|.\end{aligned}$$

Hence

$$\big\|p_0^{-1}P'\partial_uJ'-p_0^{-1}P\partial_uJ\big\|_{2,h_1}\leqslant 4\lambda b_1'A\delta^{-1}\|X'-X\|,\tag{5.10}$$

and similarly,

$$\big\|p_0^{-1}P'\partial_u\Phi' - p_0^{-1}P\partial_u\Phi\big\|_{\underline{2},h_1} \leqslant 4\lambda b_1' A\delta^{-1}\|X'-X\|. \tag{5.11}$$

Now, in view of the inequalities (5.7) and (5.9), Lemma 5.2 yields

$$\begin{aligned}\|P'^* - P^*\|_{\underline{1},h_1} &\leqslant \lambda\kappa\big(2^d dc + 4b_1'\sigma\big)A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\| \\ &\leqslant \lambda^{-1}\kappa\big(2^d d\lambda^2 c + 4b_1'\big)A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\|.\end{aligned}$$

And, in view of the inequalities (5.8) and (5.10),

$$\begin{aligned}\|J'^* - J^*\|_{\underline{2},h_1} &\leqslant \kappa\big(2^d dc\delta\sigma^{-1} + 4\lambda b_1'\sigma\big)A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\| \\ &\leqslant \lambda^{-1}\kappa\big(2^d d\lambda c + 4b_1'\big)A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\|.\end{aligned} \tag{5.12}$$

Finally, because of the inequalities (5.12) and (5.11),

$$\begin{aligned}\|\Phi'^* - \Phi^*\|_{\underline{2},h_1} &\leqslant \lambda\kappa\big[\sup\{|\alpha_j|\}\lambda^{-1}\kappa\big(2^d d\lambda c + 4b_1'\big) + 4\lambda b_1'\sigma\big]A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\| \\ &\leqslant \kappa^2\sup\{1,|\alpha_j|\}\big(2^d d\lambda c + 8b_1'\big)A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\|,\end{aligned}$$

and $\|X'^* - X^*\| \leqslant b_1'' A\delta^{-1}\sigma^{-1}|\mu|\|X'-X\|$ with a positive number

$$b_1'' = \kappa^2\sup\{1,|\alpha_1|,\dots,|\alpha_d|\}\big(2^d d\lambda^2 c + 8b_1'\big)$$

which depends only on u_1 and $\Delta\beta$.

So, we can take $c_1'' = b_1'' A\delta^{-1}\sigma^{-1}$ and Theorem 5.1 provides a solution $X = (P, J, \Phi)$ in $\mathcal{A}$, analytic in μ for

$$|\mu| \leqslant \inf\left\{\mu_1^*, \frac{1}{2c_1''}\right\} = \inf\left\{\frac{1}{40b_1'}, \frac{1}{2b_1''}\right\}A^{-1}\delta\sigma.$$

Moreover,

$$\big\|P - P^{(1)}\big\|_{\underline{1},h_1},\ \big\|J - J^{(1)}\big\|_{\underline{2},h_1} \leqslant \big\|X - X^{(1)}\big\| \leqslant 2c_1'c_1''|\mu| = 2b_1'b_1''A^2\delta^{-2}\sigma^{-1}|\mu|$$

and

$$\|P\|_{\underline{1},h_1},\quad \|J\|_{\underline{2},h_1},\quad \|\Phi\|_{\underline{2},h_1} \leqslant \|X\| \leqslant 2c_1' = 2b_1' A\delta^{-1};$$

one can thus easily choose b_1 large enough, but depending only on u_1 and $\Delta\beta$, so to ensure the three inequalities announced in Proposition 5.1. The uniform convergence of the sequence $(\mathcal{F}_\mu^m(0))_{m\geqslant 0}$ guarantees the analyticity with respect to μ of the solution.

5.4. Elimination of the variable θ

Let us now focus on the functions $\Phi_1,\dots,\Phi_d$ discovered in the previous section. As functions of (u,θ,z) they belong to $\underline{\mathcal{B}}_2^+(h_0 - \frac{7\sigma}{10})$, but they also depend analytically on μ. They are defined for $|\mu| \leqslant \mu_1'$ and satisfy $\|\Phi\|_{\underline{2},h_0-\frac{7\sigma}{10}} \leqslant C_1$.

PROPOSITION 5.2. – *Let $h_2 = h_0 - \frac{9\sigma}{10}$. There exist positive numbers μ_2', C_2 such that, for $|\mu| \leqslant \mu_2'$, the close-to-identity change of angular variables*

$$\varphi = \theta + \mu\Phi(u,\theta,z,\mu)$$

admits an inverse

$$\theta = \varphi + \mu\Theta(u, \varphi, z, \mu)$$

with Θ *belonging to* $\mathcal{B} = [\underline{\mathcal{B}}_2^+(h_2)]^d$*, depending analytically on* μ*, and satisfying the inequality*

$$\|\Theta\|_{\underline{2},h_2} \leqslant C_2.$$

If $\delta \leqslant \sigma \leqslant \lambda^{-2}$*, one can take*

$$\mu_2' = b_2^{-1} A^{-1} \delta\sigma, \quad C_2 = b_2 A \delta^{-1},$$

where $A = \sigma^{-d} A(\delta/2, \sigma/2)$ *and the positive number* b_2 *depends only on* u_1 *and* $\Delta\beta$*.*

Proof. – The equation to be satisfied by Θ can be written

$$\Theta = \mathcal{F}_\mu(\Theta) := -\Phi \circ (\mathrm{Id} + \mu\Theta)$$

(here of course composition must be understood with respect to the variables θ_j only, the rest of the variables (u, z, μ) being considered as a parameter).

We shall assume $\delta \leqslant \sigma \leqslant \lambda^{-2}$, so that we can take $\mu_1' = b_1 A^{-1}\delta\sigma$ and $C_1 = b_1 A\delta^{-1}$ with $A = \sigma^{-d} A(\delta/2, \sigma/2)$. In view of applying our fixed-point theorem, we first observe that $\mathcal{F}_\mu(0) = -\Phi$ and

$$\|\Phi\|_{\underline{2},h_2} \leqslant \|\Phi\|_{\underline{2},h_2+\frac{2\sigma}{10}} \leqslant c_2' = b_1 A\delta^{-1},$$

and that according to Lemma 5.1, $\mathcal{F}_\mu(\Theta)$ is defined and belongs to $\mathcal{B}$ as soon as $\|\Theta\|_{\underline{2},h_2} \leqslant 2c_2'$ and

$$|\mu| \leqslant \mu_2^* = \inf\left\{\frac{\sigma}{40c_2'}, \mu_1'\right\}.$$

Let us suppose that $|\mu| \leqslant \mu_2^*$ and $\|\Theta\|_{\underline{2},h_2}$, $\|\Theta'\|_{\underline{2},h_2} \leqslant 2c_2'$. According to Lemma 5.1,

$$\|\mathcal{F}_\mu(\Theta') - \mathcal{F}_\mu(\Theta)\|_{\underline{2},h_2} \leqslant 2^d d|\mu| \|\partial_\theta \Phi\|_{\underline{2},h_2+\frac{\sigma}{10}} \|\Theta' - \Theta\|_{\underline{2},h_2},$$

but the Cauchy inequalities show that $\|\partial_\theta\Phi\|_{\underline{2},h_2+\frac{\sigma}{10}} \leqslant 10 b_1 A\delta^{-1}\sigma^{-1}$. Thus we can define

$$c_2'' = 10 \cdot 2^d d b_1 A \delta^{-1}\sigma^{-1}$$

and apply Theorem 5.1. □

5.5. Substitution and integration

Let us suppose $\delta \leqslant \sigma \leqslant \lambda^{-2}$ and consider the functions

$$\begin{cases} \mathcal{P}(u, \varphi, z, \mu) = P(u, \varphi + \mu\Theta(u, \varphi, z, \mu), z, \mu), \\ \mathcal{J}(u, \varphi, z, \mu) = J(u, \varphi + \mu\Theta(u, \varphi, z, \mu), z, \mu). \end{cases}$$

Since $\|\Theta\|_{\underline{1},h_0-\frac{9\sigma}{10}} \leqslant \|\Theta\|_{\underline{2},h_0-\frac{9\sigma}{10}} \leqslant C_2 = b_2 A\delta^{-1}$, we can apply Lemma 5.1 with $h = h_0 - \frac{9\sigma}{10}$, $h^* = h_0 - \frac{8\sigma}{10}$, $\mathcal{B} = \underline{\mathcal{B}}_1^+$ or $\underline{\mathcal{B}}_2^+$:

$$\mathcal{P} \in \underline{\mathcal{B}}_1^+\left(h_0 - \frac{9\sigma}{10}\right), \quad \mathcal{J} \in \underline{\mathcal{B}}_2^+\left(h_0 - \frac{9\sigma}{10}\right) \quad \text{for } |\mu| \leqslant \inf\left\{\mu_1', \mu_2', \frac{\sigma}{20C_2}\right\}$$

and these functions depend analytically on μ. Moreover,

$$\begin{cases} \|\mathcal{P}-P\|_{\underline{1},h_0-\frac{9\sigma}{10}} \leqslant 2^d d|\mu| \|\partial_\theta P\|_{\underline{1},h_0-\frac{8\sigma}{10}} C_2, \\ \|\mathcal{J}-J\|_{\underline{2},h_0-\frac{9\sigma}{10}} \leqslant 2^d d|\mu| \|\partial_\theta J\|_{\underline{2},h_0-\frac{8\sigma}{10}} C_2. \end{cases}$$

Using the Cauchy inequalities and Proposition 5.1 with $C_1 = b_1 A\delta^{-1}$, we find

$$\|\partial_\theta P\|_{\underline{1},h_0-\frac{8\sigma}{10}}, \|\partial_\theta J\|_{\underline{2},h_0-\frac{8\sigma}{10}} \leqslant 10\sigma^{-1}C_1,$$

so

$$\|\mathcal{P}-P\|_{\underline{1},h_0-\frac{9\sigma}{10}},\ \|\mathcal{J}-J\|_{\underline{2},h_0-\frac{9\sigma}{10}} \leqslant b_3 A^2\delta^{-2}\sigma^{-1}|\mu| \tag{5.13}$$

with $b_3 = 10\cdot 2^d d b_1 b_2$.

Proposition 5.1 provides also the inequalities

$$\big\|P-P^{(1)}\big\|_{\underline{1},h_0-\frac{7\sigma}{10}},\ \big\|J-J^{(1)}\big\|_{\underline{2},h_0-\frac{7\sigma}{10}} \leqslant b_1 A^2\delta^{-2}\sigma^{-1}|\mu|.$$

We thus end up with the inequalities

$$\big\|\mathcal{P}-P^{(1)}\big\|_{\underline{1},h_0-\frac{9\sigma}{10}},\ \big\|\mathcal{J}-J^{(1)}\big\|_{\underline{2},h_0-\frac{9\sigma}{10}} \leqslant b_4 A^2\delta^{-2}\sigma^{-1}|\mu|, \tag{5.14}$$

where we choose $b_4 \geqslant b_3 + b_1$ large enough (depending only on u_1 and $\Delta\beta$) so that all this holds for

$$|\mu| \leqslant b_4^{-1}A^{-1}\delta\sigma.$$

PROPOSITION 5.3. – *The series of Proposition* 3.1

$$\widetilde{S}^+ = \widetilde{S}_0(u) + \sum_{n\geqslant 1} \mu^n \widetilde{S}_n^+(u,\varphi,z)$$

converges for $|\mu| \leqslant b_4^{-1}A^{-1}\delta\sigma$ *to*

$$\widetilde{S}^+(u,\varphi,z,\mu) = \widetilde{S}_0(u) - \mu \int_u^{+\infty} p_0(u')\mathcal{P}(u',\varphi,z,\mu)\,\mathrm{d}u'$$

and

$$\partial_u(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+) = \mu p_0\big(\mathcal{P}-P^{(1)}\big), \quad \partial_\varphi(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+) = \mu\big(\mathcal{J}-J^{(1)}\big).$$

Theorem 3.1 *follows from this proposition*: one checks that there exists a positive number b_5 which depends only on u_1 such that $\|p_0\|_1 \leqslant b_5\delta^{-1}$ and $\|p_0'\|_1 \leqslant b_5\delta^{-2}$ (*see* Appendix B), thus the first inequality in (5.14) implies that for $|\mu| \leqslant b_4^{-1}A^{-1}\delta\sigma$,

$$\begin{cases} \big\|\partial_u(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)\big\|_{2,h_0-\frac{9\sigma}{10}} \leqslant b_4 b_5 A^2\delta^{-3}\sigma^{-1}|\mu|^2, \\ \big\|\partial_u^2(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)\big\|_{2,h_0-\frac{9\sigma}{10}} \leqslant b_4 b_5 A^2(1+\delta^{-1})\delta^{-3}\sigma^{-1}|\mu|^2. \end{cases} \tag{5.15}$$

The second one can be rewritten

$$\begin{cases} \big\|\partial_\varphi(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)\big\|_{2,h_0-\frac{9\sigma}{10}} \leqslant b_4 A^2\delta^{-2}\sigma^{-1}|\mu|^2, \\ \big\|\partial_u\partial_\varphi(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)\big\|_{2,h_0-\frac{9\sigma}{10}} \leqslant b_4 A^2\delta^{-2}\sigma^{-1}|\mu|^2. \end{cases}$$

The Cauchy inequalities yield

$$\big\|\partial_\varphi^2(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)\big\|_{2,h_0-\sigma} \leqslant 10 b_4 A^2\delta^{-2}\sigma^{-2}|\mu|^2,$$

and, by integration, the first inequality in (5.15) implies also

$$\big|(\widetilde{S}^+ - \widetilde{S}_0 - \mu\widetilde{S}_1^+)(u,\varphi;z,\mu)\big| \leqslant \frac{1}{2} b_4 b_5 A^2\delta^{-3}\sigma^{-1}|\mu|^2 \mathrm{e}^{-2u_1 - 2\,\Re\mathrm{e}\,u}$$

for $u \in \overline{\mathcal{D}}^+_{u_1,\delta}$, $\varphi \in \overline{\mathbb{T}}^d_{h_0-\sigma}$ and $z \in \overline{\Sigma}_{u_1,\delta,\Delta\beta}$. □

Proof of Proposition 5.3. – The formal manifold $\mathcal{G}\mathrm{r}(\mathrm{d}S^+)$ may be written

$$\mathcal{W}^*\colon \begin{cases} q = q_0(u), \\ p = p_0(u) + \mu\sum_{n\geqslant 0}\mu^n\mathcal{P}_n^*(u,\varphi,z), \\ I = \mu\sum_{n\geqslant 0}\mu^n\mathcal{J}_n^*(u,\varphi,z), \end{cases}$$

with coefficients $\mathcal{P}_n^* = p_0^{-1}\partial_u\widetilde{S}_{n+1}^+$, $\mathcal{J}_n^* = \partial_\varphi\widetilde{S}_{n+1}^+$ exponentially decreasing at $+\infty$ with respect to u, and it is invariant by the Hamiltonian vector field

$$\begin{cases} \dot q = p, \\ \dot p = \sin q - \mu\partial_q F(q,\varphi), \end{cases} \qquad \begin{cases} \dot\varphi = z\omega + \alpha I, \\ \dot I = -\mu\partial_\varphi F(q,\varphi). \end{cases}$$

The same is true for the Taylor expansion with respect to μ of $\mathcal{W}$:

$$\mathcal{W}\colon \begin{cases} q = q_0(u), \\ p = p_0(u) + \mu\sum_{n\geqslant 0}\mu^n\mathcal{P}_n(u,\varphi,z), \\ I = \mu\sum_{n\geqslant 0}\mu^n\mathcal{J}_n(u,\varphi,z), \end{cases}$$

with $\mathcal{P} = \sum\mu^n\mathcal{P}_n$ and $\mathcal{J} = \sum\mu^n\mathcal{J}_n$.

This means that we have got two formal solutions, $(\mathcal{P}^*,\mathcal{J}^*)$ and $(\mathcal{P},\mathcal{J})$, of the system

$$\begin{cases} \widetilde{D}_0(p_0\mathcal{P}) = -\partial_u\widetilde{F} - \mu(\mathcal{P}\partial_u\mathcal{P} + p_0\alpha\mathcal{J}\cdot\partial_\varphi\mathcal{P}), \\ \widetilde{D}_0\mathcal{J} = -\partial_\varphi\widetilde{F} - \mu(p_0^{-1}\mathcal{P}\partial_u\mathcal{J} + \alpha\mathcal{J}\cdot\partial_\varphi\mathcal{J}), \end{cases}$$

which expresses the invariance by the Hamiltonian vector field. But this system has a unique formal solution whose coefficients decrease exponentially with respect to u at $+\infty$, as is easily checked by using the invertibility of $\widetilde{D}_0$.

Therefore, we get for all $n \geqslant 0$ the identities

$$p_0^{-1}\partial_u\widetilde{S}_{n+1}^+ = \mathcal{P}_n, \quad \partial_\varphi S_{n+1}^+ = \mathcal{J}_n,$$

the first of which can be integrated:

$$\widetilde{S}^+_{n+1}(u,\varphi,z) = -\int_u^{+\infty} p_0(u')\mathcal{P}_n(u',\varphi,z)\,\mathrm{d}u'. \qquad \square$$

Remark 5.1. – The fact that $\mathcal{W}$ is Lagrangian stems from the isotropy of the torus $\mathcal{T}$. Indeed, the symplectic 2-form is preserved by the flow, but on $\mathcal{W}$, all the trajectories lead to $\mathcal{T}$ where it vanishes identically, thus it vanishes on $\mathcal{W}$ too: $\mathcal{W}$ is isotropic – and Lagrangian since it has dimension $d+1$.

6. Proof of Proposition 3.3: straightening of the characteristic vector field D

In order to prove Proposition 3.3, we shall apply twice the fixed-point Theorem 5.1: once to find $\mathcal{U}$ and once to invert $\mathrm{Id}+\mu\mathcal{U}$; then we shall deduce from the Cauchy inequalities bounds for the derivatives of $\mathcal{U}$ and $\mathcal{V}$ (each time the domain which we work in will shrink a little). We first define the appropriate Banach algebras.

6.1. Definitions and initial bounds

Let $\mathcal{D}$ and Σ be open subsets of $\mathbb{C}$. We denote by $\mathcal{B}_{\mathcal{D}\times\Sigma}$ the Banach algebra of all analytic functions in $\mathcal{D}\times\Sigma$ which extend continuously to bounded functions on the closure $\overline{\mathcal{D}}\times\overline{\Sigma}$ of that domain, equipped with the supremum norm, and by $\underline{\mathcal{B}}_{\mathcal{D}\times\Sigma}$ the subalgebra of the functions whose partial derivative with respect to the first variable belongs to $\mathcal{B}_{\mathcal{D}\times\Sigma}$ too. For $h>0$, we define then the Banach algebras $\mathcal{B}_{\mathcal{D}\times\Sigma}(h)$ and $\underline{\mathcal{B}}_{\mathcal{D}\times\Sigma}(h)$ according to the construction of Section 5.2 (*see* formulas (5.2) and (5.3)).

Let us fix $u_2>0$, $\Delta\beta\in\left]0,\arctan\frac{\pi}{2u_2}\right[$ and $\delta,\sigma>0$. The last two parameters will be supposed small enough with respect to u_2 and $\Delta\beta$, even if we do not mention it explicitly (for instance the half-aperture of the sectors $\mathcal{D}^-_{u_2,\delta}$ and $\mathcal{D}^+_{u_2,\delta}$ must be larger than $\Delta\beta$). We define four domains

$$\mathcal{D}^{(3)} = \mathcal{D}_{u^{(3)},\frac{6\delta}{6}} \subset \mathcal{D}^{(2)} = \mathcal{D}_{u^{(2)},\frac{5\delta}{6}} \subset \mathcal{D}^{(1)} = \mathcal{D}_{u^{(1)},\frac{4\delta}{6}} \subset \mathcal{D}^{(0)} = \mathcal{D}_{u^{(0)},\frac{3\delta}{6}}$$

by the conditions

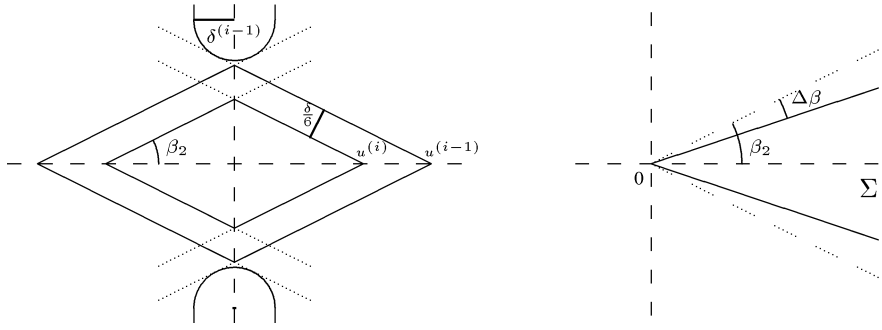
$$\mathcal{D}^{(i)} = \mathcal{D}_{u^{(i)},\delta^{(i)}}, \quad u^{(i)} = u_2 + \frac{(3-i)\delta}{6\sin\beta_2}, \quad \delta^{(i)} = \frac{3+\mathrm{i}}{6}\delta, \quad i=0,1,2,$$

where β_2 denotes the half-aperture of the sectors $\mathcal{D}^-_{u_2,\delta}$ and $\mathcal{D}^+_{u_2,\delta}$, so that

$$\mathcal{D}^{(3)} = \mathcal{D}_{u_2,\delta}, \quad \left\{u\in\mathbb{C} \mid \mathrm{dist}\big(u,\mathcal{D}^{(i)}\big) < \frac{\delta}{6}\right\} \subset \mathcal{D}^{(i-1)}, \quad i=3,2,1,$$

Since the sectors $\mathcal{D}^{\pm}_{u^{(i)},\delta}$ have the same aperture, there is only one corresponding sector for the variable z:

$$\Sigma_{u^{(i)},\,\delta^{(i)},\,\Delta\beta} = \Sigma, \quad i = 0,\,1,\,2,\,3.$$



The interest of these definitions is the property that for all u in $\mathcal{D}^{(i)}$, the disk of centre u and radius $\delta/6$ is contained in $\mathcal{D}^{(i-1)}$, which allows us to benefit from the Cauchy inequalities with respect to u. Observe that, if say $\delta \leqslant 1$, the numbers $u^{(0)}$, $u^{(1)}$ and $u^{(2)}$ are bounded by a number u_1 which depends only on u_2 (one can take $u_1 = \frac{\pi}{2}\cot\beta$, where β is the half-aperture of the sector $\mathcal{D}^{\pm}_{u_2,1}$).

We shall write $\mathcal{B}^{(i)}$ or $\underline{\mathcal{B}}^{(i)}$ instead of $\mathcal{B}_{\mathcal{D}^{(i)}\times\Sigma}$ or $\underline{\mathcal{B}}_{\mathcal{D}^{(i)}\times\Sigma}$ (the norm of $\mathcal{B}^{(i)}$, resp. $\underline{\mathcal{B}}^{(i)}$, will be denoted $\|\cdot\|^{(i)}$, resp. $\|\cdot\|^{(\underline{i})}$) and we shall use four different values of h:

$$h^{(3)} = h_0 - \sigma < h^{(2)} < h^{(1)} < h^{(0)} = h_0 - \frac{\sigma}{2}, \quad h^{(i)} = h_0 - \frac{3+i}{6}\sigma.$$

We shall obtain, for $|\mu|$ small enough,

$$\mathcal{U} \in \underline{\mathcal{B}}^{(1)}\big(h^{(1)}\big) \times \big[\underline{\mathcal{B}}^{(1)}\big(h^{(1)}\big)\big]^d, \quad \mathcal{V} \in \mathcal{B}^{(2)}\big(h^{(2)}\big) \times \big[\mathcal{B}^{(2)}\big(h^{(2)}\big)\big]^d.$$

When dealing with products of Banach spaces of this kind, we shall use as a norm the supremum of the norms of the components.

Section 5 provides initial bounds in $\mathcal{B}^{(0)}(h^{(0)})$ for the vector field D that we want to straighten:

LEMMA 6.1. – *In the coordinates (u,φ), the characteristic vector field D can be expressed as*

$$\widetilde{D} = \widetilde{D}_0 + \mu\left(\widetilde{D}_u\frac{\partial}{\partial u} + \widetilde{D}_\varphi\cdot\frac{\partial}{\partial\varphi}\right),$$

with

$$\begin{cases} \widetilde{D}_u = \mu^{-1}\left(\dfrac{\mathrm{d}q_0}{\mathrm{d}u}\right)^{-2}\partial_u\left(\dfrac{1}{2}\big(\widetilde{S}^+ + \widetilde{S}^-\big) - \widetilde{S}_0\right), \\ \widetilde{D}_\varphi = \mu^{-1}\alpha\partial_\varphi\left(\dfrac{1}{2}\big(\widetilde{S}^+ + \widetilde{S}^-\big) - \widetilde{S}_0\right). \end{cases}$$

If $\Delta\beta, \delta, \sigma$ are small enough, there exists a positive number μ_0 such that the functions $\widetilde{D}_u, \widetilde{D}_{\varphi_1}, \ldots, \widetilde{D}_{\varphi_d}$ are analytic for

$$u \in \mathcal{D}^{(0)}, \quad \varphi \in \mathbb{T}^d_{h^{(0)}}, \quad z \in \Sigma, \quad |\mu| < \mu_0.$$

Moreover, these functions and their partial derivatives belong to $\mathcal{B}^{(0)}(h^{(0)})$*, and there exist positive numbers* C_0, C_1 *such that*

$$\|\widetilde{D}_u\|^{(0)}_{h^{(0)}},\ \|\widetilde{D}_{\varphi_j}\|^{(0)}_{h^{(0)}} \leqslant C_0, \quad \|\mathrm{d}\widetilde{D}_u\|^{(0)}_{h^{(0)}},\ \|\mathrm{d}\widetilde{D}_{\varphi_j}\|^{(0)}_{h^{(0)}} \leqslant C_1, \quad 1 \leqslant j \leqslant d.$$

If $\delta \leqslant \sigma \leqslant 1$*, one can take*

$$\mu_0 = b_0^{-1} A^{-1} \delta\sigma, \quad C_0 = b_0 A \delta^{-1}, \quad C_1 = b_0 A \delta^{-1}\sigma^{-1},$$

where $A = \sigma^{-d} A(\delta/4, \sigma/4)$ *and* b_0 *depends only on* u_2 *and* $\Delta\beta$*.*

(Proof in Appendix B.)

6.2. Straightening of D by $\mathrm{Id} + \mu\mathcal{U}$

We now prove the first part of Proposition 3.3:

PROPOSITION 6.1. – *There exist positive numbers* μ_1 *and* M*, and there exist real-analytic functions* $\mathcal{U}_u, \mathcal{U}_{\varphi_1}, \ldots, \mathcal{U}_{\varphi_d}$ *satisfying the following properties*:
– the vector fields $\widetilde{D}$ *and* $\widetilde{D}_0 = \frac{\partial}{\partial v} + z\omega \cdot \frac{\partial}{\partial \theta}$ *are conjugated by the mapping*

$$(u, \varphi) = (v, \theta) + \mu\mathcal{U}(v, \theta; z, \mu),$$

with $\mathcal{U} = (\mathcal{U}_u, \mathcal{U}_{\varphi_1}, \ldots, \mathcal{U}_{\varphi_d})$*,*
– these functions are analytic with respect to all their arguments; for $|\mu| \leqslant \mu_1$*, they belong to* $\underline{\mathcal{B}}^{(1)}\big(h^{(1)}\big)$ *with* $\|\mathcal{U}_u\|^{(\underline{1})}_{h^{(1)}},\ \|\mathcal{U}_{\varphi_j}\|^{(\underline{1})}_{h^{(1)}} \leqslant M$*.*
More precisely, if $\delta \leqslant \sigma \leqslant 1$*, one can take*

$$\mu_1 = b_1^{-1} A^{-1} \delta^2 \quad \textit{and} \quad M = b_1 A \delta^{-1},$$

where $A = \sigma^{-d} A(\delta/4, \sigma/4)$ *and* b_1 *depends only on* u_2 *and* $\Delta\beta$*.*

Proof. – If we omit the variables z and μ, the equations to be solved can be written

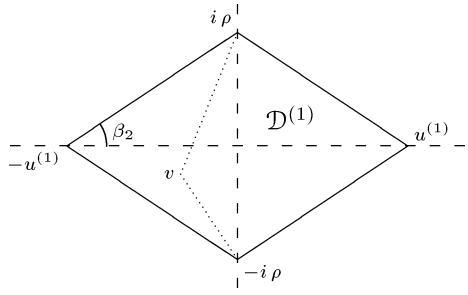
$$\begin{cases} \widetilde{D}_0 \mathcal{U}_u = \widetilde{D}_u \circ (\mathrm{Id} + \mu\mathcal{U}), \\ \widetilde{D}_0 \mathcal{U}_{\varphi_j} = \widetilde{D}_{\varphi_j} \circ (\mathrm{Id} + \mu\mathcal{U}). \end{cases} \tag{6.1}$$

In order to apply fixed-point Theorem 5.1, we shall define an inverse to the right for the operator $\widetilde{D}_0$ in $\mathcal{B}^{(1)}(h^{(1)})$.

The domain $\mathcal{D}^{(1)}$ is a lozenge whose corners are $u^{(1)}$, $\mathrm{i}\rho$, $-u^{(1)}$, $-\mathrm{i}\rho$, where

$$\rho = u^{(1)} \tan\beta_2 = \frac{\pi}{2} - \frac{2\delta}{3\cos\beta_2}$$

(β_2 denotes the half-aperture of the sectors $\mathcal{D}^-_{u^{(1)},\delta}$ and $\mathcal{D}^+_{u^{(1)},\delta}$).



LEMMA 6.2. – *Let* $\psi = \sum_{k\in\mathbb{Z}^d} \psi_k(v,z)\mathrm{e}^{\mathrm{i}k\cdot\theta} \in \mathcal{B}^{(1)}(h^{(1)})$. *The formulas*

$$\forall k\in\mathbb{Z}^d,\quad \chi_k(v,z) = \begin{cases} -\mathrm{e}^{-\mathrm{i}zvk\cdot\omega}\displaystyle\int_v^{\mathrm{i}\rho} \mathrm{e}^{\mathrm{i}z\zeta k\cdot\omega}\psi_k(\zeta,z)\,\mathrm{d}\zeta & \text{if } k\cdot\omega>0,\\ \mathrm{e}^{-\mathrm{i}zvk\cdot\omega}\displaystyle\int_{-\mathrm{i}\rho}^{v} \mathrm{e}^{\mathrm{i}z\zeta k\cdot\omega}\psi_k(\zeta,z)\,\mathrm{d}\zeta & \text{if } k\cdot\omega<0,\\ \displaystyle\int_0^v \psi_k(\zeta,z)\,\mathrm{d}\zeta & \text{if } k\cdot\omega=0, \end{cases}$$

define a function $\chi = \sum_{k\in\mathbb{Z}^d}\chi_k(v,z)\mathrm{e}^{\mathrm{i}k\cdot\theta}$ *of* $\underline{\mathcal{B}}^{(1)}(h^{(1)})$ *such that* $\widetilde{D}_0\chi$ *coincides with* ψ. *The correspondence* $\psi\mapsto\chi = E_\rho\psi$ *defines an operator* $E_\rho : \mathcal{B}^{(1)}(h^{(1)}) \to \underline{\mathcal{B}}^{(1)}(h^{(1)})$ *which satisfies a bound*

$$\|E_\rho\psi\|^{(\underline{1})}_{h^{(1)}} \leqslant \kappa\|\psi\|^{(1)}_{h^{(1)}},$$

where κ *depends only on* u_2 *and* $\Delta\beta$.

(This lemma is slightly reminiscent of [10], Lemma 3.3, which deals with a difference equation.)

Proof of Lemma 6.2. – We note that if $v\in\mathcal{D}^{(1)}$,

$$\begin{cases} [-\mathrm{i}\rho, v]\subset\mathcal{D}^{(1)}, & |v+\mathrm{i}\rho|^2 \leqslant \pi^2+u_1^2,\\ [v,+\mathrm{i}\rho]\subset\mathcal{D}^{(1)}, & |v-\mathrm{i}\rho|^2 \leqslant \pi^2+u_1^2, \end{cases}$$

where the number u_1 depends only on u_2 and is larger than $u^{(1)}$.

Suppose that $k\cdot\omega>0$, $(v,z)\in\overline{\mathcal{D}}^{(1)}\times\overline{\Sigma}$ and $\zeta\in[v,+\mathrm{i}\rho]$. We have

$$\left|\mathrm{e}^{-\mathrm{i}zvk\cdot\omega+\mathrm{i}z\zeta k\cdot\omega}\right| = \mathrm{e}^{-k\cdot\omega\,\Im\mathrm{m}(z\zeta-zv)} \leqslant 1,$$

since $\Im\mathrm{m}(zv)\leqslant\Im\mathrm{m}(z\zeta)\leqslant\Im\mathrm{m}(\mathrm{i}z\rho)$. Thus $|\chi_k(v,z)|\leqslant|v-\mathrm{i}\rho|\|\psi_k\|^{(1)}$, and

$$\|\chi_k\|^{(1)} \leqslant \left(\pi^2+u_1^2\right)^{1/2}\|\psi_k\|^{(1)}.$$

In fact, the previous inequality holds also if $k\cdot\omega<0$ or $k\cdot\omega=0$. Thus $\chi\in\mathcal{B}^{(1)}\big(h^{(1)}\big)$ and

$$\|\chi\|^{(1)}_{h^{(1)}}\leqslant\big(\pi^2+u_1^2\big)^{1/2}\|\psi\|^{(1)}_{h^{(1)}}.$$

Since for all $k\in\mathbb{Z}^d$, $\partial_u\chi_k+\mathrm{i}zk\cdot\omega\chi_k=\psi_k$, the function χ satisfies the partial differential equation

$$\widetilde{D}_0\chi=\psi$$

in $\mathcal{D}^{(1)}\times\mathbb{T}^d_{h^{(1)}}\times\Sigma$.

But we can derive another bound by using the inequality

$$\Im\mathrm{m}(zw)\geqslant|zw|\sin\Delta\beta,$$

where $w=\mathrm{i}\rho-v$, in the case where $k\cdot\omega>0$: if we parametrize the segment of integration by $\zeta=v+\xi w$, $\xi\in[0,1]$, we obtain

$$|\chi_k(v,z)|\leqslant\|\psi_k\|^{(1)}|w|\int_0^1\mathrm{e}^{-\xi k\cdot\omega\,\Im\mathrm{m}(zw)}\mathrm{d}\xi\leqslant\frac{|w|}{k\cdot\omega\,\Im\mathrm{m}(zw)}\|\psi_k\|^{(1)},$$

hence

$$\|zk\cdot\omega\chi_k\|^{(1)}\leqslant\frac{1}{\sin\Delta\beta}\|\psi_k\|^{(1)}.$$

In the case where $k\cdot\omega<0$, we would obtain the same inequality by using $w=v+\mathrm{i}\rho$. Finally, in all cases,

$$\|\partial_u\chi_k\|^{(1)}\leqslant\left(1+\frac{1}{\sin\Delta\beta}\right)\|\psi_k\|^{(1)}$$

because of the partial differential equation, thus χ belongs to $\underline{\mathcal{B}}^{(1)}\big(h^{(1)}\big)$ and we have the desired bound. □

Remark 6.1. – One checks easily that, if ψ is real-analytic, $\chi=E_\rho\psi$ is real-analytic too.

End of the proof of Proposition 6.1. – System (6.1) is thus equivalent to the equation $\mathcal{F}_\mu(\mathcal{U})=\mathcal{U}$ where the mapping $\mathcal{F}_\mu$ is defined by the formulas

$$\mathcal{F}_\mu(\mathcal{U})=\mathcal{U}^*\quad\begin{cases}\mathcal{U}^*_u=E_\rho\big[\widetilde{D}_u\circ(\mathrm{Id}+\mu\mathcal{U})\big],\\ \mathcal{U}^*_{\varphi_j}=E_\rho\big[\widetilde{D}_{\varphi_j}\circ(\mathrm{Id}+\mu\mathcal{U})\big],\quad 1\leqslant j\leqslant d.\end{cases}$$

In order to specify its domain of definition and to study that mapping, we apply the following lemma, which is analogous to Lemma 5.1 but with the variable v involved as well as the angular variables θ_j:

LEMMA 6.3. – *Let $\mathcal{D}$ and $\mathcal{D}^*$ be open subsets of $\mathbb{C}$ and h a positive number. Suppose that for each v in $\overline{\mathcal{D}}$, the closed disk of centre v and radius $\frac{\delta}{6}$ is contained in $\overline{\mathcal{D}}^*$, and let $h^*=h+\frac{\sigma}{6}$.*

Suppose $G\in\mathcal{B}_{\mathcal{D}^*\times\Sigma}(h^*)$ with $\mathrm{d}G\in\big[\mathcal{B}_{\mathcal{D}^*\times\Sigma}(h^*)\big]^{d+1}$ and let $\mu\in\mathbb{C}$ and

$$\mathcal{A}=\left\{\mathcal{U}=(\chi,\psi)\in\mathcal{B}_{\mathcal{D}\times\Sigma}(h)\times\big[\mathcal{B}_{\mathcal{D}\times\Sigma}(h)\big]^d;\ \|\mu\mathcal{U}\|_h\leqslant\inf\left\{\frac{\delta}{12},\frac{\sigma}{12}\right\}\right\}.$$

The formula

$$\forall \mathcal{U}=(\chi,\psi)\in\mathcal{A},\quad \mathcal{G}(\mathcal{U}) = G\circ(\mathrm{Id}+\mu\mathcal{U}) := \sum_{(r_0,r)\in\mathbb{N}\times\mathbb{N}^d} \frac{\mu^{r_0+|r|}}{r_0!r_1!\cdots r_d!}\chi^{r_0}\psi_1^{r_1}\cdots\psi_d^{r_d}(\partial_v)^{r_0}(\partial_\theta)^r G$$

defines a mapping from $\mathcal{A}$ to $\mathcal{B}_{\mathcal{D}\times\Sigma}(h)$ which satisfies:

$$\forall \mathcal{U},\mathcal{U}'\in\mathcal{A},\quad \|\mathcal{G}(\mathcal{U}')-\mathcal{G}(\mathcal{U})\|_h \leqslant 2^{d+1}(d+1)|\mu|\|\mathrm{d}G\|^*_{h^*}\|\mathcal{U}'-\mathcal{U}\|_h.$$

(Proof in Appendix A.)

From now on we shall assume that $\delta\leqslant\sigma\leqslant 1$ in order to indicate explicit constants (but it will be clear that, if this is not the case, there still exist constants which satisfy the desired inequalities). Specializing the previous lemma with $\mathcal{D}=\mathcal{D}^{(1)}\subset\mathcal{D}^*=\mathcal{D}^{(0)}$ and $h=h^{(1)}<h^*=h^{(0)}$, we find that our mapping $\mathcal{F}_\mu$ is defined at least for $\|\mu\mathcal{U}\|^{(1)}_{h^{(1)}}\leqslant\frac{\delta}{12}$ (and $|\mu|<\mu_0$).

By virtue of Lemmas 6.1 and 6.2, $\mathcal{F}_\mu(0)=(E_\rho\widetilde{D}_u, E_\rho\widetilde{D}_\varphi)$ satisfies

$$\|\mathcal{F}_\mu(0)\|^{(\underline{1})}_{h^{(1)}}\leqslant c' := \kappa b_0 A\delta^{-1}.$$

In view of applying Theorem 5.1 we define the Banach space $\underline{\mathbb{B}}^{(1)}=\underline{\mathcal{B}}^{(1)}(h^{(1)})\times[\underline{\mathcal{B}}^{(1)}(h^{(1)})]^d$, denote by $\|\cdot\|^{(\underline{1})}_{h^{(1)}}$ its norm, and we restrict the domain of definition of $\mathcal{F}_\mu$ to $\{\mathcal{U}\in\underline{\mathbb{B}}^{(1)};\ \|\mathcal{U}\|^{(\underline{1})}_{h^{(1)}}\leqslant 2c'\}$ assuming

$$|\mu|\leqslant\mu^* := \frac{\delta}{12}\cdot\frac{1}{2c'} = \frac{1}{24\kappa b_0}A^{-1}\delta^2.$$

We observe now that, for $\|\mathcal{U}\|^{(\underline{1})}_{h^{(1)}}, \|\mathcal{U}'\|^{(\underline{1})}_{h^{(1)}}\leqslant 2c'$, the images $\mathcal{U}^*=\mathcal{F}_\mu(\mathcal{U})$ and $\mathcal{U}'^*=\mathcal{F}_\mu(\mathcal{U}')$ satisfy

$$\|\mathcal{U}'^*-\mathcal{U}^*\|^{(\underline{1})}_{h^{(1)}}\leqslant c''|\mu|\|\mathcal{U}'-\mathcal{U}\|^{(1)}_{h^{(1)}}.$$

with $c''=2^{d+1}(d+1)\kappa C_1=2^{d+1}(d+1)\kappa b_0 A\delta^{-1}\sigma^{-1}$. We thus obtain a solution $\mathcal{U}\in\underline{\mathbb{B}}^{(1)}$ which satisfies

$$\|\mathcal{U}\|^{(\underline{1})}_{h^{(1)}}\leqslant b_1 A\delta^{-1}\quad\text{for } |\mu|\leqslant\mu_1=b_1^{-1}A^{-1}\delta^2$$

with a positive number b_1 which depends only on u_2 and $\Delta\beta$. And one checks easily that $\mathcal{U}$ is real-analytic thanks to Remark 6.1. □

6.3. Inversion of $\mathrm{Id}+\mu\mathcal{U}$ and end of the proof of Proposition 3.3

PROPOSITION 6.2. – *There exist positive numbers μ_2 and N, with $\mu_2\leqslant\mu_1$, and there exists a real-analytic vectorial function $\mathcal{V}=(\mathcal{V}_v,\mathcal{V}_{\theta_1},\ldots,\mathcal{V}_{\theta_d})$ which satisfies the following properties*:
– the components of $\mathcal{V}$ are analytic with respect to (u,φ,z,μ); for $|\mu|\leqslant\mu_2$, they belong to $\mathcal{B}^{(2)}(h^{(2)})$ and $\|\mathcal{V}_v\|^{(2)}_{h^{(2)}}, \|\mathcal{V}_{\theta_j}\|^{(2)}_{h^{(2)}}\leqslant N$;
– for $u\in\mathcal{D}^{(2)}$, $\varphi\in\mathbb{T}^d_{h^{(2)}}$, $z\in\Sigma$ and $|\mu|\leqslant\mu_2$, the point $(u,\varphi)+\mu\mathcal{V}(u,\varphi,z,\mu)$ belongs to $\overline{\mathcal{D}}^{(1)}\times\overline{\mathbb{T}}^d_{h^{(1)}}$ and its image by $\mathrm{Id}+\mu\mathcal{U}$ coincides with (u,φ).

More precisely, if $2\delta\leqslant\sigma\leqslant 1$, one can take

$$\mu_2=b_2^{-1}A^{-1}\delta^2\quad\text{and}\quad N=b_2A\delta^{-1},$$

where $A = \sigma^{-d} A(\delta/4, \sigma/4)$ and b_2 depends only on u_2 and $\Delta\beta$.

Proof. – The equation to be solved is $\mathcal{V} = \mathcal{F}_\mu(\mathcal{V})$, where

$$\mathcal{F}_\mu(\mathcal{V}) := -\mathcal{U} \circ (\mathrm{Id} + \mu\mathcal{V}).$$

We use Lemma 6.3 with a slight adaptation: we define $\mathcal{D} = \mathcal{D}^{(2)} \subset \mathcal{D}^* = \mathcal{D}^{(1)}$ and $h = h^{(2)} < h^* = h + \frac{\sigma}{12}$. Since $h^{(1)} = h^* + \frac{\sigma}{12}$, we can use the Cauchy inequalities for $\mathcal{U}$ in order to obtain bounds in $\mathbb{B}^* := \mathcal{B}^*(h^*) \times [\mathcal{B}^*(h^*)]^d$:

$$\|\partial_u \mathcal{U}\|^*_{h^*} \leqslant \|\mathcal{U}\|^{(1)}_{h^{(1)}} \leqslant M, \quad \|\partial_\theta \mathcal{U}\|^*_{h^*} \leqslant 12\sigma^{-1} M. \tag{6.2}$$

We shall suppose $\sigma \leqslant 1$ and we retain that $\|\mathrm{d}\mathcal{U}\|^*_{h^*} \leqslant 12\sigma^{-1}M$. According to Lemma 6.3, $\mathcal{F}_\mu(\mathcal{V})$ is defined as soon as $\mathcal{V} \in \mathbb{B}^{(2)} := \mathcal{B}^{(2)}(h^{(2)}) \times \big[\mathcal{B}^{(2)}\big(h^{(2)}\big)\big]^d$ and $\|\mu\mathcal{V}\|^{(2)}_{h^{(2)}} \leqslant \inf\{\delta/12, \sigma/24\}$. From now on we shall suppose $2\delta \leqslant \sigma$.

Let us check the hypotheses of Theorem 5.1: the first approximation is $\mathcal{F}_\mu(0) = -\mathcal{U}$ and

$$\|\mathcal{F}_\mu(0)\|^{(2)}_{h^{(2)}} \leqslant c' := M = b_1 A \delta^{-1}.$$

We can restrict $\mathcal{F}_\mu$ to the set $\big\{\mathcal{V} \in \mathbb{B}^{(2)};\ \|\mathcal{V}\|^{(2)}_{h^{(2)}} \leqslant 2c'\big\}$ if we suppose

$$|\mu| \leqslant \mu^* := \frac{\delta}{12} \cdot \frac{1}{2c'} = \frac{1}{24 b_1} A^{-1} \delta^2.$$

And if $\|\mathcal{V}\|^{(2)}_{h^{(2)}}$, $\|\mathcal{V}'\|^{(2)}_{h^{(2)}} \leqslant 2c'$, we find

$$\|\mathcal{F}_\mu(\mathcal{V}') - \mathcal{F}_\mu(\mathcal{V})\|^{(2)}_{h^{(2)}} \leqslant c''|\mu| \|\mathcal{V}' - \mathcal{V}\|^{(2)}_{h^{(2)}}$$

with $c'' = 2^{d+1}(d+1) \cdot 12\sigma^{-1}M = 12 \cdot 2^{d+1}(d+1) b_1 A \delta^{-1}\sigma^{-1}$. We thus obtain a solution $\mathcal{V} \in \mathbb{B}^{(2)}$ which satisfies $\|\mathcal{V}\|^{(2)}_{h^{(2)}} \leqslant b_2 A \delta^{-1}$ for $|\mu| \leqslant \mu_2 = b_2^{-1} A^{-1} \delta^2$, with some $b_2 = b_2(u_2, \Delta\beta)$. Of course $\mathcal{V}$ is real-analytic since $\mathcal{U}$ is. □

According to Proposition 6.2, for $z \in \overline{\Sigma}$ and $|\mu| \leqslant \mu_2$, the image of $\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}}$ by $\mathrm{Id} + \mu\mathcal{V}$ is contained in $\overline{\mathcal{D}}^{(1)} \times \overline{\mathbb{T}}^d_{h^{(1)}}$ and $(\mathrm{Id} + \mu\mathcal{U}) \circ (\mathrm{Id} + \mu\mathcal{V}) = \mathrm{Id}$. Moreover, we can ensure the injectivity of $\mathrm{Id} + \mu\mathcal{U}$ on $\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}}$ by taking b_2 large enough, thanks to the inequalities (6.2). This means that $\mathrm{Id} + \mu\mathcal{U}$ induces a bijection between $(\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}}) \cap (\mathrm{Id} + \mu\mathcal{U})^{-1}(\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}})$ and its image, the reciprocal being $\mathrm{Id} + \mu\mathcal{V}$.

Since $\|\mathcal{U}\|^{(1)}_{h^{(1)}} \leqslant M$ and $\|\mathcal{V}\|^{(2)}_{h^{(2)}} \leqslant N$, the sets $(\mathrm{Id} + \mu\mathcal{U})^{-1}(\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}})$ and $(\mathrm{Id} + \mu\mathcal{U})(\overline{\mathcal{D}}^{(2)} \times \overline{\mathbb{T}}^d_{h^{(2)}})$ contain $\overline{\mathcal{D}}^{(3)} \times \overline{\mathbb{T}}^d_{h^{(3)}}$ as soon as $|\mu| M \leqslant \inf\{\frac{\delta}{6}, \frac{\sigma}{6}\}$ and $|\mu| N \leqslant \inf\{\frac{\delta}{6}, \frac{\sigma}{6}\}$. Thus, for $|\mu| \leqslant \inf\{M^{-1}, N^{-1}\} \cdot \inf\{\frac{\delta}{6}, \frac{\sigma}{6}\}$, we have by restriction a bijection $\mathrm{Id} + \mu\mathcal{U}$ between $\overline{\mathcal{D}}^{(3)} \times \overline{\mathbb{T}}^d_{h^{(3)}}$ and its image, and a bijection $\mathrm{Id} + \mu\mathcal{V}$ between $\overline{\mathcal{D}}^{(3)} \times \overline{\mathbb{T}}^d_{h^{(3)}}$ and its image. The Cauchy inequalities provide the desired bounds for their partial derivatives in $\overline{\mathcal{D}}^{(3)} \times \overline{\mathbb{T}}^d_{h^{(3)}}$.

Appendix A. Fourier norms and composition lemmas

Let $\mathcal{B}$ be a Banach algebra, denote by $\|\cdot\|$ its norm, and let $h > 0$.

$(\mathcal{B}(h), \|\cdot\|_h)$ *is a Banach space.* Suppose indeed that $(\psi^{(n)})_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{B}(h)$. For each $k\in\mathbb{Z}^d$, the sequence $(\psi_k^{(n)})_{n\geqslant 0}$ of Fourier coefficients of index k is a Cauchy sequence in $\mathcal{B}$ and admits thus a limit ψ_k. Let us check that $\psi=\sum_{k\in\mathbb{Z}^d}\psi_k \mathrm{e}^{\mathrm{i}k\cdot\theta}$ belongs to $\mathcal{B}(h)$: if K is fixed in $\mathbb{N}^*$, we can choose n large enough so that

$$\sum_{|k|\leqslant K}\bigl\|\psi_k-\psi_k^{(n)}\bigr\|\mathrm{e}^{|k|h}\leqslant 1,$$

and the partial sum $\sum_{|k|\leqslant K}\|\psi_k\|\mathrm{e}^{|k|h}\leqslant 1+\|\psi^{(n)}\|_h$ is thus bounded independently of K ($A_n=\|\psi^{(n)}\|_h$ tends to some limit $A\geqslant 0$ since $|A_n-A_m|\leqslant\|\psi^{(n)}-\psi^{(m)}\|_h$, thus A_n is bounded independently of n), and $\psi\in\mathcal{B}(h)$. Let us check that $\|\psi-\psi^{(n)}\|_h$ tends to zero: let $\varepsilon>0$; we fix $N\in\mathbb{N}$ such that, for all $n,m\geqslant N$, $\|\psi^{(n)}-\psi^{(m)}\|_h\leqslant\varepsilon$, and we fix $K\in\mathbb{N}^*$ such that

$$\sum_{|k|\geqslant K}\bigl\|\psi_k^{(N)}\bigr\|\mathrm{e}^{|k|h}\leqslant\varepsilon,\quad \sum_{|k|\geqslant K}\|\psi_k\|\mathrm{e}^{|k|h}\leqslant\varepsilon.$$

We observe that for $n\geqslant N$,

$$\sum_{|k|\geqslant K}\bigl\|\psi_k^{(n)}\bigr\|\mathrm{e}^{|k|h}\leqslant\sum_{|k|\geqslant K}\bigl\|\psi_k^{(N)}\bigr\|\mathrm{e}^{|k|h}+\bigl\|\psi^{(n)}-\psi^{(N)}\bigr\|_h\leqslant 2\varepsilon,$$

and

$$\sum_{k\in\mathbb{Z}^d}\bigl\|\psi_k-\psi_k^{(n)}\bigr\|\mathrm{e}^{|k|h}\leqslant\sum_{|k|<K}\bigl\|\psi_k-\psi_k^{(n)}\bigr\|\mathrm{e}^{|k|h}+\sum_{|k|\geqslant K}\|\psi_k\|\mathrm{e}^{|k|h}+\sum_{|k|\geqslant K}\bigl\|\psi_k^{(n)}\bigr\|\mathrm{e}^{|k|h}.$$

But for n large enough the first sum in the right-hand side is less than ε, thus $\|\psi-\psi^{(n)}\|_h\leqslant 4\varepsilon$. □

The inequality for the product $\|\psi\chi\|_h\leqslant\|\psi\|_h\|\chi\|_h$ is obvious. Let us prove the *Cauchy inequalities*:

$$\bigl\|(\partial_\theta)^r\psi\bigr\|_h\leqslant r_1!\cdots r_d!(h^*-h)^{-|r|}\|\psi\|_{h^*}$$

if $0<h<h^*$, $\psi\in\mathcal{B}(h^*)$ and $r\in\mathbb{N}^d$. We have

$$\bigl\|(\partial_\theta)^r\psi\bigr\|_h=\sum_{k\in\mathbb{Z}^d}\bigl\|k_1^{r_1}\cdots k_d^{r_d}\psi_k\bigr\|\mathrm{e}^{|k|h}\leqslant M_1\cdots M_d\|\psi\|_{h^*},$$

with $M_j=\sup_{k_j\in\mathbb{N}}\bigl\{k_j^{r_j}\mathrm{e}^{-(h^*-h)k_j}\bigr\}$ for $j=1,\dots,d$, and we observe that

$$M_j\leqslant(h^*-h)^{-r_j}\sup_{t\geqslant 0}\bigl\{t^{r_j}\mathrm{e}^{-t}\bigr\}\leqslant(h^*-h)^{-r_j}r_j^{r_j}\mathrm{e}^{-r_j}\leqslant(h^*-h)^{-r_j}r_j!.\qquad\square$$

Proof of Lemma 5.1. – We suppose that we are given $G\in\mathcal{B}(h^*)$ and $\psi\in[\mathcal{B}(h)]^d$ which satisfy the assumptions of the lemma. We apply the Cauchy inequalities to G and we observe that

$$\left\|\frac{\mu^{|r|}}{r_1!\cdots r_d!}\psi_1^{r_1}\cdots\psi_d^{r_d}(\partial_\theta)^r G\right\|_h\leqslant\|\mu\psi\|_h^{|r|}(h^*-h)^{-|r|}\|G\|_{h^*}\leqslant 2^{-|r|}\|G\|_{h^*},$$

hence the convergence in $\mathcal{B}(h)$ of the series which defines $\mathcal{G}(\psi)$, and the inequality $\|\mathcal{G}(\psi)\|_h \leqslant 2^d\|G\|_{h^*}$.

We suppose now that we are given ψ and ψ' in $\mathcal{A}$. According to what has just been proved, for each $t\in[0,1]$, $\partial_\theta G\circ(\mathrm{Id}+\mu[t\psi+(1-t)\psi'])$ is bounded in $\mathcal{B}(h)$ by $2^d\|\partial_\theta G\|_{h^*}$. But

$$\mathcal{G}(\psi')-\mathcal{G}(\psi)=\mu(\psi'-\psi)\cdot\int_0^1\partial_\theta G\circ\big(\mathrm{Id}+\mu[t\psi+(1-t)\psi']\big)\mathrm{d}t,$$

hence the result. □

Proof of Lemma 6.3. – The proof is the same as for Lemma 5.1, except that there is one more variable involved. The Cauchy inequalities that we use are

$$\big\|(\partial_v)^{r_0}(\partial_\theta)^r G\big\|_h\leqslant r_0!r_1!\cdots r_d!\left(\frac{\delta}{6}\right)^{-r_0}\left(\frac{\sigma}{6}\right)^{-|r|}\|G\|^*_{h^*}. \qquad \square$$

Appendix B. Initial bounds

We first adopt the notations of Section 5. The numbers u_1, δ, σ and $\Delta\beta$ are given, with $\delta\leqslant\sigma\leqslant\lambda^{-2}$.

Proof of Lemma 5.3. – According to inequality (1.9),

$$\forall(u,\varphi)\in\mathcal{C}_{\frac{\delta}{2}}\times\overline{\mathbb{T}}^d_{h_0-\frac{\sigma}{2}},\quad |\widetilde{F}(u,\varphi)|\leqslant A(\delta/2,\sigma/2)\mathrm{e}^{-2|\,\Re\mathrm{e}\,u|}.$$

The Cauchy inequalities with respect to u show that

$$\forall(u,\varphi)\in\mathcal{C}_{\delta}\times\overline{\mathbb{T}}^d_{h_0-\frac{\sigma}{2}},\quad\begin{cases}\mathrm{e}^{2u_1+2\,\Re\mathrm{e}\,u}|\widetilde{F}(u,\varphi)|\leqslant\mathrm{e}^{2u_1}A(\delta/2,\sigma/2),\\ \mathrm{e}^{2u_1+2\,\Re\mathrm{e}\,u}|\partial_u\widetilde{F}(u,\varphi)|\leqslant\mathrm{e}^{2u_1+1}A(\delta/2,\sigma/2)\frac{2}{\delta}.\end{cases}$$

(We have used the fact that $\delta\leqslant1$.) Since $\mathcal{D}^+_{u_1,\delta}\subset\mathcal{C}_\delta$, the inequalities (5.4) provide a constant c_0 which depends only on the dimension d such that

$$\|\widetilde{F}\|_{2,h_0-\frac{11}{20}\sigma}\leqslant c_0\mathrm{e}^{2u_1}A\quad\text{and}\quad\|\partial_u\widetilde{F}\|_{2,h_0-\frac{11}{20}\sigma}\leqslant c_0\mathrm{e}^{2u_1}A\delta^{-1},$$

where $A=\sigma^{-d}A(\delta/2,\sigma/2)$. The Cauchy inequalities with respect to the angles now allow to bound $\|\partial_{\varphi_j}\widetilde{F}\|_{2,h_0-\frac{6}{10}\sigma}$, $\|\partial_u\widetilde{F}\|_{2,h_0-\frac{6}{10}\sigma}$, $\|\partial_{\varphi_j}\partial_{\varphi'_j}\widetilde{F}\|_{2,h_0-\frac{6}{10}\sigma}$ and $\|\partial_u\partial_{\varphi_j}\widetilde{F}\|_{2,h_0-\frac{6}{10}\sigma}$. □

Bounds for p_0 and p'_0. – We can write $p_0(u)=\mathrm{e}^{-u}f(u)$ where $f(u)=4(1+\mathrm{e}^{-2u})^{-1}$ is $2\pi\mathrm{i}$-periodic and meromorphic, with simple poles at $\mathrm{i}\pi/2$ and $-\mathrm{i}\pi/2$. A compacity argument shows that, if u belongs to the intersection of $\mathcal{C}_{\delta/2}$ and the strip $\{-u_1\leqslant\Re\mathrm{e}\,u\leqslant1\}$, $f(u)$ satisfies an inequality

$$|f(u)|\leqslant b\delta^{-1},$$

where the positive number b depends only on u_1. On the other hand, for $\Re\mathrm{e}\,u\geqslant1$, $|f(u)|$ is bounded by $|1-\mathrm{e}^{-2}|^{-1}$. Thus

$$\forall u\in\mathbb{C}_{\delta/2},\quad\Re\mathrm{e}\,u\geqslant-u_1\implies|p_0(u)|\leqslant b'\delta^{-1}\mathrm{e}^{-\Re\mathrm{e}\,u},$$

where b' depends only on u_1. The Cauchy inequalities allow us to conclude that

$$\|p_0\|_1 \leqslant b_5\delta^{-1} \quad \text{and} \quad \|p_0'\|_1 \leqslant b_5\delta^{-2}$$

for some b_5 which depends only on u_1. □

Proof of Lemma 6.1. – We now adopt the notations of Section 6. The first assertion of Lemma 6.1 is a simple restatement of formula (3.7). Let us denote by p_0 the function $\frac{\mathrm{d}q_0}{\mathrm{d}u}$. According to Proposition 5.3, we had

$$\partial_u(\widetilde{S}^+ - \widetilde{S}_0) = \mu p_0 \mathcal{P}^+ := \mu p_0 \mathcal{P}, \quad \partial_\varphi(\widetilde{S}^+ - \widetilde{S}_0) = \mu \mathcal{J}^+ := \mu \mathcal{J},$$

and if we use again the bounds of Section 5 but with $(u^{(0)}, \delta/2, \sigma/2)$ replacing (u_1, δ, σ), especially the bounds for $\mathcal{P} - P$ and $\mathcal{J} - J$ in inequality (5.13) and the bounds for P and J in Proposition 5.1, we find

$$\|\mathcal{P}^+\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\mathcal{J}^+\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\partial_u\mathcal{P}^+\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\partial_u\mathcal{J}^+\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}} \leqslant b' A\delta^{-1}$$

for $|\mu| \leqslant \mu_0 = {b'}^{-1}A^{-1}\delta\sigma$, where $A = \sigma^{-d}A(\delta/4, \sigma/4)$ and b' depends only on u_2 and $\Delta\beta$.

The functions $\partial_u(\widetilde{S}^+ - \widetilde{S}_0) = \mu p_0 \mathcal{P}^-$ and $\partial_\varphi(\widetilde{S}^+ - \widetilde{S}_0) = \mu \mathcal{J}^-$ satisfy the same kind of inequalities, and we have

$$\widetilde{D}_u = \frac{1}{2}p_0^{-1}\big(\mathcal{P}^+ + \mathcal{P}^-\big), \quad \widetilde{D}_{\varphi_j} = \frac{1}{2}\alpha_j\big(\mathcal{J}_j^+ + \mathcal{J}_j^-\big), \quad j = 1, \ldots, d,$$

where the function p_0^{-1} and its derivative are bounded in $\mathcal{D}^{(0)}$ by a number which depends only on u_2. We thus obtain

$$\|\widetilde{D}_u\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\widetilde{D}_\varphi\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\partial_u\widetilde{D}_u\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}}, \quad \|\partial_u\widetilde{D}_\varphi\|^{(0)}_{h^{(0)}+\frac{\sigma}{20}} \leqslant b'' A\delta^{-1}$$

with some new $b'' = b''(u_2, \Delta\beta)$ and we can conclude by the Cauchy inequalities. □

Acknowledgements

I am indebted to Pierre Lochak, who is at the origin of the collaboration [21], for having introduced me to this multidimensional problem and for having suggested the use of the Hamilton–Jacobi equation. I wish to thank Amadeu Delshams for the enlightening conversations we had during my stay in Barcelona in May 1998, Jean-Pierre Marco, Pierre Lochak and Jacky Cresson for precious discussions and help, and Alain Chenciner for his interest to the present work. Some of the results obtained here were announced at the international workshop *Exponentially small splitting, diffusion and celestial mechanics* (Aussois, June 1998), the organizers of which I would like to thank.

This work was supported by the CEE contract ERB-CHRX-CT94-0460.

REFERENCES

[1] ARNOL'D V.I., Instability of dynamical systems with many degrees of freedom, *Dokl. Akad. Nauk SSSR* **156** (1964) 9–12 (in Russian). [English translation: *Soviet Math. Dokl.* **5** (1964) 581–585].

[2] BOTT R., Lectures on Morse theory, old and new, *Bull. Amer. Math. Soc.* **7** (2) (1982) 331–358.
[3] BRUNO A.D., *Local Methods in Nonlinear Differential Equations*, Springer-Verlag, Berlin, 1989.
[4] DELSHAMS A., GELFREICH V.G., JORBA À., SEARA T.M., Exponentially small splitting of separatrices under fast quasi-periodic forcing, *Comm. Math. Phys.* **189** (1997) 35–71.
[5] DELSHAMS A., GUTIÉRREZ P., Homoclinic orbits to invariant tori in Hamiltonian systems, in: Jones C., Wiggins S., Khibnik A., Dumortier F., Terman D. (Eds.), *Multiple-Time-Scale Dynamical Systems*, IMA Vol. in Math. and its Appl., Springer-Verlag, Berlin, 1998.
[6] DELSHAMS A., SEARA T.M., An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum, *Comm. Math. Phys.* **150** (1992) 433–463.
[7] DELSHAMS A., SEARA T.M., Splitting of separatrices in Hamiltonian systems with one and half degrees of freedom, *Math. Phys. Elec. J.* **3** (1997), paper 4.
[8] ÉCALLE J., Six lectures on transseries, analysable functions and the constructive proof of Dulac's conjecture, in: Schlomiuk D. (Ed.), *Bifurcations and Periodic Orbits of Vector Fields*, Kluwer Academic, Dordrecht, 1993, pp. 75–184.
[9] ELIASSON L.H., Biasymptotic solutions of perturbed integrable Hamiltonian systems, *Bol. Soc. Bras. Mat.* **25** (1) (1994) 57–76.
[10] FRUCHARD A., SCHÄFKE R., Exponentially small splitting of separatrices for difference equations with small step size, *J. Dynam. Control Syst.* **2** (2) (1996) 193–238.
[11] GALLAVOTTI G., Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable systems. A review, *Rev. Math. Phys.* **6** (3) (1994) 343–411.
[12] GALLAVOTTI G., GENTILE G., MASTROPIETRO V., Melnikov's approximation dominance. Some examples, *Rev. Math. Phys.* **11** (4) (1999) 451–461.
[13] GELFREICH V.G., Melnikov method and exponentially small splitting of separatrices, *Physica D* **101** (1997) 227–248.
[14] GRAFF S.M., On the conservation of hyperbolic invariant tori for Hamiltonian systems, *J. Differential Equations* **15** (1974) 1–69.
[15] HIRSCH M.W., PUGH C.C., SHUB M., *Invariant Manifolds*, Lect. Notes in Math., Vol. **583**, Springer-Verlag, Berlin, 1977.
[16] LAZUTKIN V.F., *Splitting of separatrices for the Chirikov's standard map*, Preprint VINITI 6372-84, 1984 (in Russian).
[17] LOCHAK P., Effective speed of Arnol'd diffusion and small denominators, *Phys. Lett. A* **143** (1990) 39–42.
[18] LOCHAK P., Canonical perturbation theory via simultaneous approximation, *Russian Math. Surveys* **47** (1992) 57–133.
[19] LOCHAK P., Hamiltonian perturbation theory: periodic orbits, resonances and intermittency, *Nonlinearity* **6** (1993) 885–904.
[20] LOCHAK P., Tores invariants à torsion évanescente dans les systèmes hamiltoniens proches de l'intégrable, *C.R. Acad. Sci. Paris, Série I* **327** (1998) 833–836.
[21] LOCHAK P., MARCO J.-P., SAUZIN D., *On the splitting of the invariant manifolds in multidimensional near-integrable Hamiltonian systems*, Prépublication 220 de l'Institut de mathématiques de Jussieu, 1999.
[22] POINCARÉ H., *Les méthodes nouvelles de la mécanique céleste, Vol. 2*, Gauthier-Villars, Paris, 1893.
[23] POPOV G., *Invariant tori, effective stability and quasimodes with exponentially small error terms*, Preprint, 1999.
[24] PÖSCHEL J., Nekhoroshev estimates for quasi-convex Hamiltonian sytems, *Math. Z.* **213** (1993) 187–216.
[25] RUDNEV M., WIGGINS S., Existence of exponentially small separatrix splittings and homoclinic connections between whiskered tori in weakly hyperbolic near-integrable Hamiltonian systems, *Physica D* **114** (1998) 3–80.
[26] SAUZIN D., Résurgence paramétrique et exponentielle petitesse de l'écart des séparatrices du pendule rapidement forcé, *Ann. Inst. Fourier* **45** (1995) 453–511.

[27] SIMÓ C., Averaging under fast quasiperiodic forcing, in: Seimenis J. (Ed.), *Hamiltonian Mechanics: Integrability and Chaotic Behaviour*, NATO Adv. Sci. Inst. Ser. B Phys., Vol. **331**, Plenum Press, New York, 1994, pp. 13–34.

[28] TRESCHEV D.V., A mechanism for the destruction of resonance tori of Hamiltonian systems, *Math. USSR-Sbornik* **68** (1) (1991) 181–203.

[29] TRESCHEV D.V., Hyperbolic tori and asymptotic surfaces in Hamiltonian systems, *Russian J. Math. Phys.* **2** (1) (1994) 93–110.

[30] YOCCOZ J.-C., Introduction to hyperbolic dynamics, in: Branner B., Hjorth P. (Eds.), *Real and Complex Dynamical Systems*, NATO Adv. Sci. Inst. Ser. C Math. and Phys. Sciences, Vol. **464**, Kluwer Academic, Dordrecht, 1995, pp. 265–291.

(Manuscript received April 22, 1999;
accepted, after revision, January 28, 2000.)

David SAUZIN
Astronomie et systèmes dynamiques,
CNRS – Institut de Mécanique céleste,
77, avenue Denfert-Rochereau,
75014 Paris, France
E-mail: sauzin@bdl.fr