Large normally hyperbolic cylinders in *a priori* stable Hamiltonian systems

Patrick Bernard*

December 2009

Patrick Bernard,
CEREMADE, UMR CNRS 7534
Pl. du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16, France
patrick.bernard@ceremade.dauphine.fr

A major problem in dynamical systems consists in studying the Hamiltonian systems on $\mathbb{T}^n \times \mathbb{R}^n$ of the form

$$H(q,p) = h(p) - \epsilon^2 G(t,q,p), \quad (t,q,p) \in \mathbb{T} \times \mathbb{T}^n \times \mathbb{R}^n. \tag{H}$$

Here ϵ should be considered as a small perturbation parameter, we put a square because the sign of the perturbation will play a role in our discussion. In the unperturbed system ($\epsilon = 0$) the momentum variable p is constant.

We want to study the dynamics of the perturbed system in the neighborhood of a torus $\{p = p_0\}$, corresponding to a resonant frequency. There is no loss of generality in assuming that the frequency is of the form

$$\partial h(p_0) = (\omega, 0) \in \mathbb{R}^m \times \mathbb{R}^r.$$

If the restricted frequency ω is non-resonant in $\mathbb{R}^m$, then it is expected that the averaged system

$$H_a(q,p) = H_a(q_1,q_2,p_1,p_2) = h(p) - \epsilon^2 V(q_2) \tag{H_a}$$

should locally approximate the dynamics of (H) near $p = p_0 = (p_1^0, p_2^0)$, where $q = (q_1, q_2) \in \mathbb{T}^r \times \mathbb{T}^m$ and $p = (p_1, p_2) \in \mathbb{R}^r \times \mathbb{R}^m$, and where

$$V(q_2) = \int G(t, q_1, q_2, p_0)\, dt\, dq_1.$$

We make the following hypothesis on the averaged system:

Hypotheses 1. *The function h is convex with positive definite Hessian and the averaged potential V has a non-degenerate local maximum at $q_2 = 0$.*

*membre de l'IUF

Under Hypothesis 1, the averaged system has an invariant manifold of equations

$$(\partial_{p_2} h = 0, q_2 = 0) \in \mathbb{T}^n \times \mathbb{R}^n.$$

Because h has positive definite Hessian, the equation $\partial_{p_2} h(p_1, p_2) = 0$ is non-degenerate and it defines a smooth m-dimensional manifold in $\mathbb{R}^n$ wich can also be described parametrically by the relation $p_2 = P_2(p_1)$ for some function $P_2 : \mathbb{R}^m \longrightarrow \mathbb{R}^r$. Therefore, the corresponding invariant manifold can be written in a parametric form as

$$(q_1, 0, p_1, P_2(p_1)); (q_1, p_1) \in \mathbb{T}^m \times \mathbb{R}^m,$$

it is a cylinder. Moreover, this manifold is Normally hyperbolic in the sense of [9]. It is then pretty well understood that some small pieces of this manifold persist in the initial system (which is in some sense a perturbation of the averaged system), meaning that some normally hyperbolic invariant manifold close to that cylinder exist. It can be seen as the center manifold of a "whiskered" (or partially hyperbolic) torus. This whiskered torus is the continuation in the full systems of the invariant torus

$$\{(t, q_1, 0, p_0), \quad (t, q_1) \in \mathbb{T} \times \mathbb{T}^m\}$$

which exists in the averaged system. The name whiskered comes from the fact that this torus has hyperbolic normal directions, this name (as well as the corresponding object) was introduced by Arnold in [1]. The existence of a whiskered torus in the original system was proved in [18], and it is well understood, see for example [4] that such a torus must be contained in an invariant cylinder which is normally hyperbolic. Proving the existence of whiskered tori involves KAM theory, which is quite demanding in terms of regularity, while the existence of the invariant cylinder relies on the softer theory of normal hyperbolicity. The idea of embedding whiskered tori into a normally hyperbolic cylinder and to use the theory of normal hyperbolicity in the context of Arnold diffusion is more recent than the paper of Arnold. To the best of our knowledge, it appears first in Moeckel [16]. It was understood even more recently that normally hyperbolic invariant cylinders can be used to produce diffusion even in the absence of whiskered tori. The approaches used so far to study whiskered tori and their center manifolds rely on a rescaling of the momentum p, and produce an invariant cylinder of size $O(\epsilon)$. In the present paper, we ignore whiskered tori and focus on the normally hyperbolic invariant cylinder:

Theorem 1. *Assume that H is smooth (or at least C^r for a sufficiently large r) and satisfies Hypothesis 1. Assume that ω is Diophantine, and fix $\kappa > 0$. Then there exists an open ball $B \subset \mathbb{R}^m$ containing p_1^0, a neighborhood U of 0 in $\mathbb{T}^r$, a positive number ϵ_0 and, for $\epsilon < \epsilon_0$ two C^1 functions*

$$Q_2^\epsilon : \mathbb{T} \times \mathbb{T}^m \times B \longrightarrow U \subset \mathbb{T}^r \quad \text{and} \quad P_2^\epsilon : \mathbb{T} \times \mathbb{T}^m \times B \longrightarrow \mathbb{R}^r$$

such that the annulus

$$A^\epsilon = \{(t, q_1, Q_2^\epsilon(t, q_1, p_1), p_1, P_2^\epsilon(t, q_1, p_1)), \quad (t, q_1, p_1) \in \mathbb{T} \times \mathbb{T}^m \times B\}$$

is weakly invariant for (H) (in the sense that the Hamiltonian vectorfield is tangent to it). We have $P_2^\epsilon \longrightarrow P_2^0$ uniformly as $\epsilon \longrightarrow 0$, where P_2^0 is the function $(t, q_1, p_1) \longmapsto P_2(p_1)$. Moreover, we have $\|P_2^\epsilon - P_2^0\|_{C^1} \leqslant \kappa$, and $\|Q_2^\epsilon\|_{C^1} \leqslant \kappa/\epsilon$. Each strongly invariant set of (H) (in the sense that it contains the full orbit of each of its points, for example, a whiskered torus) contained in the domain

$$\mathcal{D}^\epsilon := \mathbb{T} \times \mathbb{T}^m \times U \times B \times \{p_2 \in \mathbb{R}^r : \|p_2\| \leqslant \epsilon\}$$

is contained in A^ϵ is for $\epsilon < \epsilon_0$. The cylinder A^ϵ is symplectic and normally hyperbolic in the following sense: Each strongly invariant set K (meaning that K contains the full orbit of each of its points) contained in A^ϵ is partially hyperbolic, and its center space at each point is the tangent space to A^ϵ.

The novelty here is that the ball B does not depend on ϵ. Easy examples show that we can't expect a fine control of the asymptotic behavior of Q_2^ϵ in terms of the averaged system only except if we restrict to smaller domains depending on ϵ. This asymptotic behavior will also depend on the averaged systems at other frequencies. However, the very weak estimates we have are sufficient to describe the restricted dynamics. Let $\phi : \mathbb{T}^n \times \mathbb{R}^n \longrightarrow \mathbb{T}^n \times \mathbb{R}^n$ be the time-one flow of H, and let $A_0^\epsilon \subset \mathbb{T}^n \times \mathbb{R}^n$ be the restriction of the invariant annulus to the section $\{t = 0\}$,

$$A_0^\epsilon = \big\{(q_1, Q_2^\epsilon(0, q_1, p_1), p_1, P_2^\epsilon(0, q_1, p_1)), \quad (q_1, p_1) \in \mathbb{T}^m \times B\big\}.$$

Then A_0^ϵ is somewhat invariant for ϕ (although there are some difficulties near the boundary) in a sense that will be given more precisely below. We define the map $\Phi : \mathbb{T}^m \times B \longrightarrow \mathbb{T}^m \times B$ as the restriction of ϕ to A_0^ϵ seen in coordinates (q_1, p_1), more precisely

$$\Phi(q_1, p_1) = (q_1, p_1) \circ \phi\big(q_1, Q_2^\epsilon(0, q_1, p_1), p_1, P_2^\epsilon(0, q_1, p_1)\big).$$

Note that this map is well-defined on $\mathbb{T}^m \times B$. Let us finally consider an open ball $B_0 \subset \mathbb{R}^m$ which contains p_0 and whose closure is contained in B, and set

$$A_{00}^\epsilon = \big\{(q_1, Q_2^\epsilon(0, q_1, p_1), p_1, P_2^\epsilon(0, q_1, p_1)), \quad (q_1, p_1) \in \mathbb{T}^m \times B_0\big\}.$$

Proposition 1. *The map Φ is converging uniformly (when $\epsilon \longrightarrow 0$) on $\mathbb{T}^m \times B_0$ to the map*

$$\Phi_0 : \begin{pmatrix} q_1 \\ p_1 \end{pmatrix} \longmapsto \begin{pmatrix} q_1 + \partial_{p_1} h\big(p_1, P_2(p_1)\big) \\ p_1 \end{pmatrix},$$

which gives the unperturbed dynamics on the invariant cylinder of the averaged system. Moreover, we have $\phi(A_{00}^\epsilon) \subset A_0^\epsilon$ when ϵ is small enough. Finally, given $\eta > 0$, we can choose the ball B_0 small enough so that the inequality

$$\|d\Phi - d\Phi_0\|_{C^0} \leqslant \eta$$

holds on $\mathbb{T}^m \times B_0$ when ϵ is small enough.

The frequency map

$$p_1 \longmapsto \Omega_0(p_1) := \partial_{p_1} h\big(p_1, P_2(p_1)\big)$$

has positive torsion in the sense that

$$\partial_{p_1} \Omega_0 = \partial_{p_1}^2 h(p_1, P_2(p_1))$$

is a positive definite symmetric matrix for all $p_1 \in \mathbb{R}^m$. As a consequence, when ϵ is small enough, the restricted map Φ has positive torsion in a neighborhood (independent of ϵ) of $\mathbb{T}^m \times \{p_1^0\}$, in the sense that

$$\partial_{p_1}(q_1 \circ \Phi)_{(q_1,p_1)} \rho_1 \cdot \rho_1 > 0 \quad \forall \rho_1 \in \mathbb{R}^m$$

for all $q_1 \in \mathbb{T}^m$ and $p_1 \in B_0$, provided that B_0 has been chosen small enough. The map Φ is symplectic with respect to the symplectic form obtained by restriction of the ambient symplectic form to A_0^ϵ. It is part of the statement of Theorem 1 that this form is non-degenerate on A_0^ϵ. Note that this symplectic form is not $dq_1 \wedge dp_1$ in general.

In the case $m = 1$, (but for any dimension n) one can combine these results with existing techniques on the *a priori* unstable situation, like the variational methods coming from Mather Theory (see [14, 2]), developped for the *a priori* unstable situation in [3, 6, 7] or more geometric methods like [13] (The papers [10, 18] also treat the *a priori* unstable situation, but it seems to me at first sight that they require too strong informations on the restricted dynamics to be applicable here). One can then hope to obtain, under additional non-degeneracy assumptions, the existence of restricted Arnold diffusion in the following sense: There exists $\delta > 0$ and ϵ_0 such that, for each $\epsilon \in]0, \epsilon_0[$ there exists an orbit $(q_\epsilon(t), p_\epsilon(t))$ with the following property: The image $p_\epsilon(\mathbb{R})$ is not contained in any ball of radius δ in $\mathbb{R}^n$. Once again, the key point here is that δ can be chosen independent of ϵ. Specifying the needed "non-degeneracy assumptions" will require some further work, but I believe it will not require any method beyond those which are already available.

Of course, finding "global" Arnold diffusion, as announced in [15], that is orbits wondering in the whole phase space along different resonant lines (or far away along a given resonant line) requires a specific study of relative resonances (when the restricted frequency ω is resonant), where the existence of normally hyperbolic invariant cylinders can't be obtained by the method used in the present paper.

Let us close this introduction with a remark on uniqueness. In general, there is no uniqueness statement for the normally invariant cylinder we obtain. However, in the case $m = 1$, we can obtain a stronger result: Let $[p_1^-, p_1^+] \subset B \subset \mathbb{R}$ be an interval such that both $\Omega_0(p_1^-)$ and $\Omega_0(p_1^+)$ are Diophantine. Then, there exists whiskered tori $\mathbb{T}_-^\epsilon$ and $\mathbb{T}_+^\epsilon$ of dimension 2 in $\mathbb{T} \times \mathbb{T}^n \times \mathbb{R}^n$ which are close to the unperturbed tori

$$T_-^0 = \big\{ \big(t, q_1, 0, p_1^-, P_2(p_1^-)\big) : \quad (t, q_1) \in \mathbb{T} \times \mathbb{T} \big\}$$

and

$$T_+^0 = \big\{ \big(t, q_1, 0, p_1^+, P_2(p_1^+)\big) : \quad (t, q_1) \in \mathbb{T} \times \mathbb{T} \big\}.$$

The whiskered tori $\mathbb{T}_\pm^\epsilon$ are contained in the annulus A^ϵ. They bound a compact part $A_=^\epsilon$ of A^ϵ which is then stronly invariant in the sense that it contains the full orbit of each of its points. The annulus $A_=^\epsilon$ is then unique in the sense that if $\tilde{A}^\epsilon$ is another normally hyperbolic cylinder given by Theorem 1 (with the same domain B), then it must contain $A_=^\epsilon$. The cylinder $A_=^\epsilon$ is a normally hyperbolic invariant cylinder in the genuine sense. If the interval $[p^-, p^+]$ has been chosen small enough, then the restricted map $\Phi : A_=^\epsilon \longrightarrow A_=^\epsilon$ is a C^1 area preserving twist map (for the appropriate area form). When $m > 1$ one should not expect the same kind of properties, since Arnold diffusion may occur inside the invariant cylinder.

1 Averaging

In order to apply averaging methods, it is easier to consider the extended phase space

$$(t, e, q, p) \in \mathbb{T} \times \mathbb{R} \times \mathbb{T}^n \times \mathbb{R}^n$$

where the Hamiltonian flow can be seen as the Hamiltonian flow of the autonomous Hamiltonian function

$$\tilde{H}(t, e, q, p) = h(p) + e - \epsilon^2 G(t, q, p)$$

on one of its energy surfaces, for example $\tilde{H} = 0$. Then, we consider a smooth solution $f(t, q)$ of the Homological equation

$$\partial_t f + \partial_q f \cdot (\omega, 0) = G(t, q, p_0) - V(q_2).$$

Such a solution exists because ω is Diophantine, as can be checked easily by power series expansion. It is unique up to an additive constant. We consider the smooth symplectic diffeomorphism

$$\psi^\epsilon : (t, e, q, p) \longmapsto (t, e + \epsilon^2 \partial_t f(t, q), q, p + \epsilon^2 \partial_q f(t, q))$$

and use the same notation for the diffeomorphism $(t, q, p) \longmapsto (t, q, p + \epsilon^2 \partial_q f(t, q))$. We have

$$\tilde{H} \circ \psi^\epsilon = h(p) + e - \epsilon^2 V(q_2) - \epsilon^2 R(t, q, p) + O(\epsilon^4),$$

where $R(t, q, p) = G(t, q, p) - G(t, q, p_0)$. In other words, by the time-dependent symplectic change of coordinates ψ^ϵ, we have reduced the study of H to the study of the time-dependent Hamiltonian

$$H_1(t, q, p) = h(p) - \epsilon^2 V(q_2) - \epsilon^2 R(t, q, p) + O(\epsilon^4)$$

where $R = O(p - p_0)$. As a consequence, Theorem 1 holds for H if it holds for H_1. More precisely, assume that there exists an invariant cylinder

$$\tilde{A}^\epsilon = (t, q_1, \tilde{Q}_2^\epsilon(t, q_1, p_1), p_1, \tilde{P}_2^\epsilon(t, q_1, p_1))$$

for H_1, with $\|\tilde{Q}_2^\epsilon\|_{C^1} \leqslant \kappa/2\epsilon$ and $\|P_2^\epsilon - P_2^0\|_{C^1} \leqslant \kappa/2$. Then the annulus $A^\epsilon := \psi^\epsilon(\tilde{A}^\epsilon)$ is invariant for H. Since ψ^ϵ is ϵ^2-close to the identity, while $\|Q_2^\epsilon\|_{C^1} \leqslant \kappa/2\epsilon$, the annulus A^ϵ has the form

$$A^\epsilon = (t, q_1, Q_2^\epsilon(t, q_1, p_1), p_1, P_2^\epsilon(t, q_1, p_1))$$

for C^1 functions Q_2^ϵ, P_2^ϵ which satisfy $\|Q_2^\epsilon\|_{C^1} \leqslant \kappa/\epsilon$ and $\|P_2^\epsilon - P_2^0\|_{C^1} \leqslant \kappa$. We will prove that Theorem 1 holds for H_1 in section 4. We first expose some useful tools.

2 Normally hyperbolic manifolds

We shall now present a version of the classical theory of Normally hyperbolic manifolds adapted for our purpose. On $\mathbb{R}^{n_z} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$, let us consider the time dependent vectorfield

$$\begin{aligned} \dot{z} &= Z(t, z, x, y) \\ \dot{x} &= A(z)x \\ \dot{y} &= -B(z)y. \end{aligned}$$

We assume that the function

$$Z : \mathbb{R} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \longrightarrow \mathbb{R}^{n_z}$$

is C^1-bounded in the domain

$$\mathbb{R} \times \mathbb{R}^{n_z} \times \{x \in \mathbb{R}^{n_x} : \|x\| < 1\} \times \{y \in \mathbb{R}^{n_y} : \|y\| < 1\}, \tag{D}$$

and that the matrices A and B are C^1-bounded functions of z. Moreover, we assume that there exists constants $a > b > 0$ such that

$$A(z)x \cdot x \geqslant a\|x\|^2 \quad , \quad B(z)y \cdot y \geqslant a\|y\|^2$$

for all x, y, z, and such that

$$\|\partial_{(t,z)} Z(t, z, x, y)\| \leqslant b$$

for all (t, z, x, y) belonging to (D). We consider the perturbed vectorfield

$$\begin{aligned} \dot{z} &= Z(t, z, x, y) + R_z(t, z, x, y) \\ \dot{x} &= A(z)x \qquad + R_x(t, z, x, y) \\ \dot{y} &= -B(z)y \quad\;\; + R_x(t, z, x, y). \end{aligned}$$

where $R = (R_z, R_x, R_y)$ is seen as a small perturbation.

Theorem 2. *There exists $\epsilon > 0$ such that, when $\|R\|_{C^1} < \epsilon$, the maximal invariant set of the perturbed vectorfield contained in the domain (D) is a graph of the form*

$$\{(t, z, X(t, z), Y(t, z)), \quad (t, z) \in \mathbb{R} \times \mathbb{R}^{n_z}\}$$

where X and Y are C^1 maps. This graph is normally hyperbolic, and it is contained in the domain

$$\mathbb{R}^{n_z} \times \{x \in \mathbb{R}^{n_x} : \|x\| \leqslant (2/a)\|R\|_{C^0}\} \times \{y \in \mathbb{R}^{n_y} : \|y\| \leqslant (2/a)\|R\|_{C^0}\}.$$

In other words, we have

$$\|(X, Y)\|_{C^0} \leqslant (2/a)\|R\|_{C^0}.$$

The C^1 norm of (X, Y) is converging to zero when the C^1 norm of the perturbation converges to zero.

Proof. The invariant space $\mathbb{R}^{n_z}$ is normally hyperbolic in the sense of [8, 9]. As a consequence, the standard theory applies and implies the existence of functions X and Y such that the graph $(t, z, X(t, z), Y(t, z))$ is invariant, normally hyperbolic, and contained in (D). Note that we are slightly outside of the hypotheses of the statements in [9] because our unperturbed manifold is not compact. However, the results actually depend on uniform estimates rather than on compactness (see [11], Appendix B, for example, see also [5]), and we assumed such uniform estimates.

Let us now prove the estimate on (X, Y). We have the inequality

$$\dot{x} \cdot x \geqslant a\|x\|^2 + x \cdot R_x \geqslant a\|x\|(\|x\| - \|R_x\|_{C^0}/a)$$

which implies that

$$\dot{x} \cdot x \geqslant \|x\|\|R_x\|_{C^0}$$

if

$$2\|R_x\|_{C^0}/a \leqslant \|x\| \leqslant 1,$$

hence this domain can't intersect the invariant graph. Similar considerations show that the domain $2\|R_y\|_{C^0}/a \leqslant \|y\| \leqslant 1$ can't intersect the graph. □

3 Hyperbolic Linear System

Let us consider the linear Hamiltonian system on $\mathbb{R}^n \times \mathbb{R}^n$ generated by the Hamiltonian

$$H(q, p) = \frac{1}{2}\langle Bp, p\rangle - \frac{1}{2}\langle Aq, q\rangle,$$

where both A and B are positive definite symmetric matrices. We recall that this system can be reduced to

$$G(x, y) = \langle Dx, y\rangle,$$

where D is a positive definite symmetric matrix, by a linear symplectic change of variables $(q, p) \longrightarrow (x, y)$. In order to do so, we consider the symmetric positive definite matrix

$$L := \left(A^{-1/2}(A^{1/2}BA^{1/2})^{1/2}A^{-1/2}\right)^{1/2},$$

which is the only symmetric and positive definite solution of the equation $L^2AL^2 = B$. Considering the change of variables

$$x = \frac{1}{\sqrt{2}}(Lp + L^{-1}q) \quad ; \quad y = \frac{1}{\sqrt{2}}(Lp - L^{-1}q)$$

or equivalently

$$q = \frac{1}{\sqrt{2}}L(x-y) \quad ; \quad p = \frac{1}{\sqrt{2}}L^{-1}(x+y),$$

an elementary calculation shows that we obtain the desired form for the Hamiltonian in coordinates (x,y), with

$$D = LAL = L^{-1}BL^{-1}.$$

As a consequence, the equations of motions in the new variables take the block-diagonal form

$$\dot{x} = Dx \quad ; \quad \dot{y} = -Dy.$$

In the original coordinates (q,p) the stable space (which is the space $x=0$) is the space $\{(q,-L^2q), q \in \mathbb{R}^n\}$ while the unstable space is $\{(q,L^2q), q \in \mathbb{R}^n\}$.

4 Proof of Theorem 1

We now prove Theorem 1 for the Hamiltonian

$$H_1(t,q,p) = h(p) - \epsilon^2 V(q_2) - \epsilon^2 R(t,q,p) + O(\epsilon^{2+\gamma}),$$

where $R = O(p-p_0)$ and $\gamma > 0$ ($\gamma = 2$ in our situation). We assume that Hypothesis 1 holds. We lift all the angular variables to the universal covering, and see H_1 as a Hamiltonian of the variables

$$(t,q,p) = (t,q_1,q_2,p_1,p_2) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^m \times \mathbb{R}^r$$

which is one-periodic in t,q. We assume that $p_0 = 0$.

We will need some notations. We set $A := \partial^2 V(0)$, it is a symmetric positive definite matrix. We will denote by $B(p_1)$ a matrix which depends smoothly on p_1, is uniformly positive definite, is constant outside of a neighborhood of $p_1 = 0$ in $\mathbb{R}^m$, and coincides with $\partial^2_{p_2} h(p_1, P_2(p_1))$ in a neighborhood of $p_1 = 0$. We will denote by $\tilde{P}_2(p_1)$ a compactly supported smooth function $\tilde{P}_2 : \mathbb{R}^m \longrightarrow \mathbb{R}^r$ which coincides with P_2 around $p_1 = 0$. Finally, we will denote by $h_0(p_1)$ a smooth compactly supported function which is equal to $h(p_1, P_2(p_1))$ around $p_1 = 0$.

It is useful to introduce two new positive parameters α and δ. We always assume that

$$0 < \epsilon < \delta < \alpha < 1.$$

In the sequel, we shall chose α small, then δ small with respect to α, and work with ϵ small enough with respect to α and δ. The parameter δ represents the size of the normally hyperbolic cylinder we intend to find. We will denote by $\underline{\chi}$ a smooth function of its arguments which may depend (in an unexplicited way) on the parameters ϵ, δ, but which is C^2-bounded, uniformly in ϵ, δ. The notation χ will be used in a similar way when only C^1 bounds are assumed.

Lemma 2. *There exists a smooth Hamiltonian function $H_2(t,q,p)$ (which depends on the parameters ϵ, δ) of the form*

$$H_2 = h_0(p_1) + \frac{1}{2}B(p_1)\cdot(p_2 - \tilde{P}_2(p_1))^2 - \frac{\epsilon^2}{2}A \cdot q_2^2$$
$$+ \epsilon^3 \underline{\chi}\big(p_1, (p_2 - \tilde{P}_2(p_1))/\epsilon\big) + \epsilon^2\delta^{3/2}\underline{\chi}(q_2/\sqrt{\delta}) + \epsilon^2\delta\underline{\chi}(t,q,p/\delta) + \epsilon^{2+\gamma}\underline{\chi}(t,q,p)$$

which coincides with H_1 on the domain

$$\big\{\|q_2\| \leqslant \sqrt{\delta}, \|p_1\| \leqslant \delta, \|p_2 - P_2(p_1)\| \leqslant \epsilon\big\}.$$

PROOF. Let us expand the function h with respect to p_2 at the point $P_2(p_1)$:

$$h(p_1,p_2)=h(p_1,P_2(p_1))+\frac{1}{2}\partial^2_{p_2}h(p_1,P_2(p_1))\cdot(p_2-P_2(p_1))^2+\underline{S}(p)\cdot(p_1-P_2(p_1))^3$$

where $\underline{S}(p)$ is a 3-linear form on $\mathbb{R}^r$ depending smoothly on p. We consider a 3-form $S(p)$ which depends smoothly on p, is compactly supported, and is equal to $\underline{S}(p)$ near $p=0$. Let $i:\mathbb{R}^k\longrightarrow\mathbb{R}^k$ (for any k) be a compactly supported smooth map which is equal to the identity on the unit ball. Then the function

$$\begin{aligned}&h_0(p_1)+\frac{1}{2}B(p_1)\cdot(p_1-P_2(p_1))^2+\epsilon^3S(p)\cdot\big(i\big[(p_2-\tilde{P}_2(p_1))/\epsilon\big]\big)^3\\=&h_0(p_1)+\frac{1}{2}B(p_1)\cdot(p_1-P_2(p_1))^2+\epsilon^3\underline{\chi}\big(p_1,(p_2-\tilde{P}_2(p_1))/\epsilon\big)\end{aligned}$$

is equal to h if p belongs to a given neighborhood of 0 (independant of ϵ,δ) and satisfies $\|p_2-\tilde{P}_2(p_1)\|\leqslant\epsilon$. Similarly, we write

$$V(q_2)=\frac{1}{2}A\cdot q_2^2+W(q_2)\cdot q_2^3$$

for some 3-linear form $W(q_2)$. It is equal to

$$\frac{1}{2}A\cdot q_2^2+\delta^{3/2}W(q_2)\cdot(i(q_2/\sqrt{\delta}))^3=\frac{1}{2}A\cdot q_2^2+\delta^{3/2}\underline{\chi}\big(q_2/\sqrt{\delta}\big)$$

on $\{\|q_2\|\leqslant\sqrt{\delta}\}$. Finally, we observe that the function $R(t,q,p)$ can be written in the form

$$R(t,q,p)=L(t,q,p)\cdot p$$

and is equal to the function

$$\delta L(t,q,p)\cdot i(p/\delta)=\delta\underline{\chi}(t,q,p/\delta)$$

on $\{\|p\|\leqslant\delta\}$. Collecting all terms proves the Lemma. □

We will now prove the existence of a normally hyperbolic invariant graph for H_2 contained in the region

$$\{\|q_2\|\leqslant\sqrt{\delta},\|p_2-\tilde{P}_2(p_1)\|\leqslant\epsilon\}$$

Its intersection with $\{\|p_1\|<\delta\}$ will give a weakly invariant manifold for H_1 (meaning that the Hamiltonian vectorfield of H_1 is tangent to it). In order to simplify the following equations, we set

$$h_2(p):=h_0(p_1)+\frac{1}{2}B(p_1)\cdot(p_2-\tilde{P}_2(p_1))^2.$$

The Hamiltonian vectorfield of H_2 can be written

$$\begin{aligned}\dot{q}_1&=\partial_{p_1}h_2(p) &&+\epsilon^2\chi\big(p_1,(p_2-\tilde{P}_2(p_1))/\epsilon\big)+\epsilon^2\chi(t,q,p/\delta)\\\dot{p}_1&=0 &&+\epsilon^2\delta\chi(t,q,p/\delta)\\\dot{q}_2&=B(p_1)(p_2-P_2(p_1))&&+\epsilon^2\chi(p_1,(p_2-\tilde{P}_2(p_1))/\epsilon)+\epsilon^2\chi(t,q,p)\\\dot{p}_2&=\epsilon^2Aq_2 &&+\epsilon^2\delta\chi(q_2/\sqrt{\delta})+\epsilon^2\delta\chi(t,q_1,q_2,p/\delta)\end{aligned}$$

recalling the convention that $\chi(.)$ always denotes a C^1 function of its arguments, depending on ϵ and δ, but bounded in C^1 independently of δ and ϵ. Motivated by section 2, we set

$$L(p_1)=\big(A^{-1/2}(A^{1/2}B(p_1)A^{1/2})^{1/2}A^{-1/2}\big)^{1/2},$$

and perform the change of variables $(t, q_1, p_1, q_2, p_2) \longrightarrow (\tau, \theta, r, x, y)$ given by:

$$\tau = \epsilon t, \quad \theta = \epsilon\alpha q_1, \quad r = p_1,$$
$$x = L(p_1)(p_2 - \tilde{P}_2(p_1)) + \epsilon L^{-1}(p_1)q_2, \quad y = L(p_1)(p_2 - \tilde{P}_2(p_1)) - \epsilon L^{-1}(p_1)q_2,$$

recalling that α is a fixed positive parameter. Equivalently, this can be written

$$t = \tau/\epsilon, \quad q_1 = \theta/\epsilon\alpha, \quad p_1 = r, \quad q_2 = L(r)(x-y)/2\epsilon, \quad p_2 = \tilde{P}_2(r) + L^{-1}(r)(x+y)/2.$$

In the new coordinates, the principal part of the vectorfield takes the form (denoting $\acute{f}$ for $df/d\tau$)

$$\acute{\theta} = \alpha\Omega(r, x, y), \quad \acute{r} = 0, \quad \acute{x} = D(r)x, \quad \acute{y} = -D(r)y,$$

with

$$\Omega(r, x, y) := \partial_{p_1} h_2\big(r, \tilde{P}_2(r) + L^{-1}(r)(x+y)/2\big)$$

and

$$D(r) := L(r)AL(r) = L^{-1}(r)B(r)L^{-1}(r).$$

The equality above holds because $L(r)$ solves the equation $L^2(r)AL^2(r) = B(r)$. Let us detail the calculations leading to the expressions of $\acute{x} := dx/d\tau$ (the calculation for $\acute{y}$ is similar):

$$\begin{aligned}
\epsilon\acute{x} = \dot{x} =& L(p_1)\big(\dot{p}_2 - \partial_{p_1}P_2 \cdot \dot{p}_1\big) + \epsilon L^{-1}(p_1)\dot{q}_2 + \big(\partial_{p_1}L \cdot \dot{p}_1\big)\big(p_2 - P_2(p_1)\big) + \epsilon\big(\partial_{p_1}(L^{-1}) \cdot \dot{p}_1\big)q_2 \\
=& \epsilon^2 L(p_1)Aq_2 + \epsilon L^{-1}(p_1)B(p_1)(p_2 - P_2(p_1)) \\
&+ \epsilon^2\delta\chi(t, q, p/\delta, x, y) + \epsilon^3\chi(p_1, (x+y)/\epsilon) + \epsilon^2\delta\chi(q_2/\sqrt{\delta}) + \epsilon^{2+\gamma}\chi(t, q, p, x, y) \\
=& \epsilon L(r)AL(r)(x-y)/2 + \epsilon L^{-1}(r)B(r)L^{-1}(r)(x+y)/2 \\
&+ \epsilon^2\delta\chi(\tau/\epsilon, \theta/\epsilon, r/\delta, x/\delta, y/\delta, x/\epsilon, y/\epsilon) + \epsilon^2\delta\chi(r/\sqrt{\delta}, x/\sqrt{\delta}\epsilon, y/\sqrt{\delta}\epsilon) \\
=& \epsilon D(r)x + \epsilon^2\delta\chi(\tau/\epsilon, \theta/\epsilon, r/\delta, x/\delta, y/\delta, x/\epsilon, y/\epsilon) + \epsilon^2\delta\chi(r/\sqrt{\delta}, x/\sqrt{\delta}\epsilon, y/\sqrt{\delta}\epsilon).
\end{aligned}$$

The function $\Omega(r, x, y)$ is C^1-bounded on

$$\big\{(r, x, y), \quad \|x\| \leqslant 1, \|y\| \leqslant 1\big\}.$$

We can choose $\alpha < 1$ once and for all in order that the principal part of the vectorfield satisfies the hypotheses of Theorem 2. The full vectorfield can be written in the new coordinates, (with the notation $\acute{f} := df/d\tau$):

$$\begin{aligned}
\acute{\theta} &= \alpha\Omega(r, x, y) + \epsilon^2\chi(\tau/\epsilon, \theta/\alpha\epsilon, r, x/\epsilon, y/\epsilon) \\
\acute{r} &= 0 \qquad\qquad + \epsilon\delta\chi(\tau/\epsilon, \theta/\alpha\epsilon, r/\delta, x/\delta, y/\delta, x/\epsilon, y/\epsilon) \\
\acute{x} &= D(r)x \quad\; + \epsilon\delta\chi(\tau/\epsilon, \theta/\alpha\epsilon, r/\delta, x/\delta, y/\delta, x/\epsilon, y/\epsilon) + \epsilon\delta\chi(r/\sqrt{\delta}, x/\sqrt{\delta}\epsilon, y/\sqrt{\delta}\epsilon) \\
\acute{y} &= -D(r)y \;\; + \epsilon\delta\chi(\tau/\epsilon, \theta/\alpha\epsilon, r/\delta, x/\delta, y/\delta, x/\epsilon, y/\epsilon) + \epsilon\delta\chi(r/\sqrt{\delta}, x/\sqrt{\delta}\epsilon, y/\sqrt{\delta}\epsilon).
\end{aligned}$$

In this expression, we observe that the uniform norm of the perturbation is $O(\epsilon\delta)$ while the C^1 norm is $O(\sqrt{\delta})$ (recall that $0 < \epsilon < \delta < 1$). We can apply Theorem 2 and find a unique bounded normally hyperbolic invariant graph

$$(\tau, \theta, X(\tau, \theta, r), r, Y(\tau, \theta, r)).$$

Moreover Theorem 2 also implies that

$$\|(X, Y)\|_{C^0} \leqslant C\epsilon\delta.$$

Since the invariant graph we have obtained is the maximal invariant set contained in the domain $\{\|x\| \leqslant 1, \|y\| \leqslant 1\}$, and since the vectorfield is ϵ-periodic in t and $\alpha\epsilon$-periodic in q_1, we conclude that the functions X and Y are ϵ-periodic in t and $\alpha\epsilon$-periodic in q_1. In the initial coordinates, we have an invariant graph

$$(t, q_1, Q_2^\epsilon(t, q_1, p_1), p_1, P_2^\epsilon(t, q_1, p_1))$$

with

$$Q_2^\epsilon(t, q_1, p_1) = L(p_1)\big(X(\epsilon t, \epsilon q_1, p_1) - Y(\epsilon t, \epsilon q_1, p_1)\big)/2\epsilon$$

and

$$P_2^\epsilon(t, q_1, p_1) = P_2(p_1) + L^{-1}(p_1)\big(X(\epsilon t, \epsilon q_1, p_1) + Y(\epsilon t, \epsilon q_1, p_1)\big)/2.$$

The functions Q_2^ϵ and P_2^ϵ are 1-periodic in (t, q_1). The invariant graph we have obtained is normally hyperbolic for the flow of H_2, and its strong stable and strong unstable directions have the same dimension r. It follows from general results on partial hyperbolicity in a symplectic context (see *e. g.* [12], Proposition 1.8.3 [1]) that it is a symplectic manifold. This means that the restriction to the invariant graph of the ambient symplectic form is a symplectic form. Observing that

$$\|Q_2^\epsilon\|_{C^0} \leqslant C\delta, \quad \|P_2^\epsilon\|_{C^0} \leqslant C\epsilon\delta,$$

we infer that the annulus

$$\big\{(t, q_1, Q_2^\epsilon(t, q_1, p_1), p_1, P_2^\epsilon(t, q_1, p_1)) : t \in \mathbb{T}, q_1 \in \mathbb{T}^m, p_1 \in \mathbb{R}^m, \|p_1\| < \delta\big\} \subset \mathbb{T} \times \mathbb{T}^n \times \mathbb{R}^n$$

is contained in the domain

$$\{\|q_2\| \leqslant \sqrt{\delta}, \|p_1\| \leqslant \delta, \|p_2\| \leqslant \epsilon\}$$

where $H_2 = H_1$, provided δ has been chosen small enough. It is thus a weakly invariant cylinder for H_1 *i.e.* the extended Hamiltonian vectorfield of H_1 on $\mathbb{T} \times \mathbb{T}^n \times \mathbb{R}^n$ is tangent to this annulus at each point. Orbits may still from the cylinder through its boundary. We finish with the estimates on the C^1 norms. Since the C^1 size of the perturbation is $O(\sqrt{\delta})$, we can make it as small as we want by chosing δ small. We can thus assume that $\|(X, Y)\|_{C^1}$ is small, and this implies the desired C^1 estimates on P_2^ϵ and Q_2^ϵ. We have proved Theorem 1 for H_1, we conclude from Section 1 that Theorem 1 holds for H.

5 Proof of Proposition 1

Let (q_1, p_1) be given in $\mathbb{T}^m \times B$, and let $(q_1(t), q_2(t), p_1(t), p_2(t))$ be the orbit (under H) of the point

$$\big(q_1, Q_2^\epsilon(0, q_1, p_1), p_1, P_2^\epsilon(0, q_2, p_2)\big).$$

We have the Hamilton equations

$$\dot{q}_1(t) = \partial_{p_1} H(t, q_1(t), q_2(t), p_1(t), p_2(t))$$
$$\dot{p}_1(t) = -\partial_{q_1} H(t, q_1(t), q_2(t), p_1(t), p_2(t)).$$

They imply that $\dot{p}_1 = O(\epsilon^2)$, and we conclude that $p_1(t) \in B$ for all $t \in [0, 1]$ if $p_1 \in B_0$, provided ϵ is small enough. This implies the inclusion

$$\phi(A_{00}^\epsilon) \subset A_0^\epsilon,$$

[1] In this text, the equality of dimensions (that obviously holds here) is stated as a conclusion, although it should be taken as an assumption

and it also implies that

$$\big(q_1(t), q_2(t), p_1(t), p_2(t)\big) = \big(q_1(t), Q_2^\epsilon(t, q_1(t), p_1(t)), p_1(t), P_2^\epsilon(t, q_1(t), p_1(t))\big)$$

for each $t \in [0, 1]$. The Hamilton equations then take the form

$$\begin{aligned}\dot q_1(t) &= \partial_{p_1} h\big(p_1(t), P_2^\epsilon(t, q_1(t), p_1(t))\big) - \epsilon^2 \partial_{p_1} G\big(t, q_1(t), Q_2^\epsilon(t, q_1(t), p_1(t)), P_2^\epsilon(t, q_1(t), p_1(t))\big)\\ \dot p_1(t) &= \qquad\qquad + \epsilon^2 \partial_{q_1} G\big(t, q_1(t), Q_2^\epsilon(t, q_1(t), p_1(t)), P_2^\epsilon(t, q_1(t), p_1(t))\big).\end{aligned}$$

The map Φ is thus the time-one flow of the vectorfield

$$\begin{pmatrix} q_1 \\ p_1 \end{pmatrix} \longmapsto \begin{pmatrix} \partial_{p_1} h\big(p_1, P_2^\epsilon(t, q_1, p_1)\big) - \epsilon^2 \partial_{p_1} G\big(t, q_1, Q_2^\epsilon(t, q_1, p_1), P_2^\epsilon(t, q_1, p_1)\big) \\ \epsilon^2 \partial_{q_1} G\big(t, q_1, Q_2^\epsilon(t, q_1, p_1), P_2^\epsilon(t, q_1, p_1)\big) \end{pmatrix}$$

which converges uniformly to the vectorfield

$$\begin{pmatrix} q_1 \\ p_1 \end{pmatrix} \longmapsto \begin{pmatrix} \partial_{p_1} h\big(p_1, P_2(p_1)\big) \\ 0 \end{pmatrix}$$

when $\epsilon \longrightarrow 0$ on $\mathbb{T}^m \times B$. We conclude that Φ is converging uniformly to Φ_0 (as defined in Proposition 1). Moreover, we see that the C^1 distance between these two vectorfields is $O(\kappa)$, so it can be made arbitrarily small by taking B_0 small enough. The same statement then holds for the time-one flows Φ and Φ_0.

References

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

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

[3] P. Bernard: The dynamics of pseudographs in convex Hamiltonian systems, Journ. AMS, **21** No. 3 (2008) 625–669.

[4] S.V. Bolotin, D.V. Treschev: Remarks on the definition of hyperbolic tori of Hamiltonian systems *Regular and Chaotic dynamics*, **5** (2000), no. 4, 401–412.

[5] M. Chaperon, Stable manifolds and the Perron-Irwin method, *Erg. Th. Dyn. Sys.*, **24** (2004), 1359-1394.

[6] C.-Q. Cheng, J. Yan, Existence of diffusion orbits in a priori unstable Hamiltonian systems, *J. Differential Geom.* **67** (2004), no. 3, 457–517.

[7] C.-Q. Cheng, J. Yan, Arnold Diffusion in Hamiltonian systems: the a priori unstable case, *J. Differential Geom.* **82**, (2009), no. 2, 229-277.

[8] N. Fenichel:Persistence and smoothness of invariant manifolds for flows, *Indiana Univ. Math. J.*, **21**, (1971), 193–226.

[9] M.W. Hirsch, C.C. Pugh, M. Shub *Invariant manifolds*, Lecture notes in Math. Springer Berlin, New York, (1977).

[10] A. Delshams, R. de la Llave, T. M. Seara : *A Geometric Mechanism for diffusion in Hamiltonian Systems Overcoming the Large Gap Problem: Heuristics and Rigorous Verification on a Model,* Mem. A.M.S. **179** (2006), no 844.

[11] A. Delshams, R. de la Llave, T. M. Seara : Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows, *Adv. in Math.* **202** (2006) 64-188.

[12] P. Lochack, J. P. Marco, D. Sauzin : *On the Splitting of Invariant Manifolds in Multidimensional Near-Integrable Hamiltonian Systems*, Mem. A.M.S. **163** (2003) no. 775.

[13] M. Gidea, C. Robinson: Obstruction argument for transition chains of Tori interspersed with gaps, preprint.

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

[15] J. N. Mather: Arnold diffusion: announcement of results, J. Math. Sci. (N. Y.) **124** (2004), no. 5, 5275–5289.

[16] R. Moeckel: Transition Tori in the Five-Body Problem, JDE **129** (1996), 290-314.

[17] D. Treschev: Hyperbolic tori and asymptotic surfaces in Hamiltonian systems Russ. J. Math. Phys, **2** (1994) no. 1, 93-110.

[18] D. Treschev, Evolution of slow variables in a priori unstable Hamiltonian systems. *Nonlinearity* **17** (2004), no. 5, 1803–1841.