M P E J

Mathematical Physics Electronic Journal

ISSN 1086-6655
Volume 10, 2004

Paper 5
Received: Mar 2, 2004, Accepted: Mar 23, 2004
Editor: G. Gallavotti

WHISKERED AND LOW DIMENSIONAL TORI IN NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

R. DE LA LLAVE AND C. E. WAYNE

Abstract. We show that a nearly integrable hamiltonian system has invariant tori of all dimensions smaller than the number of degrees of freedom provided that certain nondegeneracy conditions are met. The tori we construct are generated by the resonances of the system and are topologically different from the orbits that are present in the integrable system. We also show that the tori we construct have stable and unstable manifolds and point out how to construct other types of interesting orbits.

The method of proof is a combination of different perturbation methods.

Preface

This paper proves the existence of whiskered tori in nearly integrable Hamiltonian systems. This paper was circulated widely in 1989, but was never formally published due to a combination of circumstances.

The paper, however has been widely quoted and some of the perturbation techniques described here have been used for a variety of purposes by several people. In spite of very dramatic progress in Hamiltonian mechanics in the intervening years we feel that this work may still be of interest for several reasons.

We note that the tori constructed in this paper have different topologies from the tori that are present in the original system. In some way they are higher dimensional analogues of *islands* in nearly integrable maps.

Within the last five years there have been significant advances in the understanding of "Arnold diffusion" – the process by which solutions in nearly integrable Hamiltonian systems my drift far from their initial values. One problem that must be confronted in analytical approaches to the Arnold diffusion problem is the so called "large gaps" problem. This is the appearance of gaps in the "transition chain" of whiskered tori near resonances of the integrable Hamiltonian.

It is precisely in such gaps that the whiskered tori of the present paper are constructed and we think that there is mounting evidence (numerical and heuristic in [HdlL00] and rigorous studies of diffusion [DdlLMS03]) secondary invariant tori play an important role in recurrence in Fermi-Pasta-Ulam as well as in phenomena of instability and diffusion. The later is in agreement with the widely believed intuition that diffusion in Hamiltonian systems is generated by resonances.

The paper that follows is largely unchanged from the version of 1989. The only significant changes are:

- We have eliminated Appendix A, since a expanded version was published in [dlLW95].
- We have shortened Section 3, so that we now just quote results in the literature rather than indicating a proof.
- We have added a sentence to the abstract.

Besides the changes above, we have corrected a few obvious typos and misspellings, but we decided not to introduce any other modifications.

In particular, we did not update the bibliography and, hence, we do not cover any of the developments that have occurred after 1989. Also, we have resisted the temptation to improve some of the results on regularity which in light of subsequent theoretical developments are not optimal.

1. Introduction

It is well known, (and probably true!) that certain trajectories of nearly integrable hamiltonian systems with more than two degrees of freedom will wander very far from their initial values. The best known mechanism proposed to explain that effect is the so–called "Arnol'd diffusion". This mechanism, first described in [Arn64], requires the existence of an abundance of "whiskered tori" in the system. These are invariant tori that are the intersection of an "arriving whisker" and a "departing whisker." All trajectories in the arriving whisker are asymptotic to the invariant torus as $t \to +\infty$, while those in the departing whisker are asymptotic to the torus as $t \to -\infty$. The intersection of arriving and departing whiskers corresponding to different invariant tori provides the means to transport trajectories from one part of the phase space to another.

In this paper we study the formation of whiskered tori in nearly integrable systems, whose hamiltonian can be expressed in terms of action-angle variables as the smooth function

$$H(I, \phi) = h(I) + \epsilon f(I, \phi), \tag{1}$$

with $I \in V$, an open subset of $\mathbb{R}^N$ and $\phi \in \mathbb{T}^N$. We assume that H is a sufficiently smooth function that also satisfies:

$$\left|\left|(h_{II})^{-1}\right|\right| < c, \qquad \forall I \in V. \tag{2}$$

Note that while the $\epsilon = 0$ problem has an abundance of invariant tori, it has no whiskered tori. In fact, there is no hyperbolicity in the unperturbed ($\epsilon = 0$) problem. This is one of the differences between the present paper, and [AA68], [Gra74],[Zeh76] where it is assumed that the $\epsilon = 0$ problem already contains invariant whiskered tori and hence is not integrable. In [HM82] it is assumed that the $\epsilon = 0$ hamiltonian has homoclinic connections. Even if this situation could happen in a system whose hamiltonian depends only on action variables, it cannot happen in systems with the complementary variables of the actions being angles as we consider. Notice also that the orbits in a homoclinic trajectory are hyperbolic. The present method, on the other hand, not only proves the existence of invariant tori for (1) but also elucidates the mechanism which gives rise to the hyperbolicity.

Unfortunately even though we can prove that (1) typically possesses an abundance of whiskered tori our methods, which are perturbative in nature, do not give sufficient control over the whiskers to prove that the whiskers corresponding to different tori intersect one another. Thus we cannot conclude that (1) actually exhibits Arnol'd diffusion. Indeed, the verification of the existence of Arnol'd diffusion by perturbative methods seems exceedingly difficult because the results of Nekhoroshev [Neh77] show that the effect of this diffusion for a C^∞ system of the form (1) should be zero to all orders in ϵ^n (see e.g [BG86, BGG85] for a more modern exposition). Therefore, establishing the existence of Arnol'd diffusion for concrete systems of the form (1) would require a perturbation theory that can establish results beyond all orders in ϵ.

Even if direct perturbation methods have not succeeded in establishing Arnol'd diffusion for concrete systems, it is still possible to use non–constructive arguments based on genericity. Douady [Dou88] has shown that generically, in the neighborhood of an elliptic fixed point one gets Arnol'd diffusion. It seems that his methods could be made to work with our analysis to conclude that a generic hamiltonian (1) exhibits Arnol'd diffusion on a certain scale.

Our results also shed some light on other aspects of the behavior of (1). There has recently been considerable interest in whether or not (1) possesses invariant tori of dimension greater than 1 (periodic orbits) and less than N-1. The invariant tori most commonly constructed by the K. A. M. theory have dimension N, but in [Mos67], Moser proved that one could also construct tori of dimension N-1 using essentially the same methods and he asked the question of whether it would be possible to construct tori of lower dimension. Graff [Gra74] showed that if the unperturbed tori were hyperbolic, one could develop a K.A.M. perturbation theory regardless of their dimension – see also [Zeh76] for another way of developing this perturbation theory –, but it was not until quite recently that Eliasson [Eli88], followed by Rüssman [R̈88] and Pöschel[Poschel89], showed that one could use a variant of K. A. M. methods to construct low dimensional tori for (1).

Our methods also give low dimensional tori for (1) – whiskered tori must have dimension less than $N-1$, and can in general have any dimension between 1 and $N-2$. Both the methods we use and the tori we construct, are quite different from those of [Eli89, R̈88, Pös89]. The low dimensional tori they construct could possibly have no hyperbolic direction even when $\epsilon \neq 0$, while those that we produce certainly do. The tori we construct, can be retracted to a periodic orbit, so all the closed paths contained in them have the same homotopy type in the original phase space. The tori of [Eli89, R̈88, Pös89] have as many homotopically different closed paths as their dimension. A perturbation that left the system (1) integrable would not produce tori with hyperbolic directions, therefore, our result must include non–degeneracy conditions on the perturbation as well as on the original system. The results of [Eli89, R̈88, Pös89]. only require non–degeneracy conditions on the unperturbed system.

As a byproduct of our method, we can also produce some other invariant sets using the rich theory of monotone area–preserving maps or the theory of behavior near elliptic points. We point out that, from the point of view of producing diffusion, the intersection of invariant manifolds of those sets can be as effective as that of whiskers of tori.

We now state our results somewhat more precisely and sketch the method of proof.

We will work in a neighborhood of a periodic orbit of the $\epsilon = 0$ case of (1), $I(t) = I^*$, $\phi(t) = h_I(I^*)t + \phi(0)$. Since the orbit is periodic, $Th_I(I^*) \in \mathbb{Z}^N$, where T is the period. If $\nu \in \mathbb{Z}^N$, we will have that $T\nu \cdot h_I(I^*) \in \mathbb{Z}$. Therefore,

$$|\nu \cdot h_I(I^*)| \geq c > 0 \tag{3}$$

with $c = 1/T$ whenever $\nu \cdot h_I(I^*) \neq 0$.

We will be concerned with "whiskered invariant tori" which we will take to mean d–dimensional invariant tori on which the motion of the hamiltonian system is differentiably conjugate to a rotation and on which the restriction of the symplectic form vanishes. Each point in the torus, will have $N-d$ dimensional "stable" and "unstable" manifolds consisting of points whose orbits are asymptotic (for positive and negative times respectively) to the orbit of the corresponding point in the torus with exponential bounds on the speed of approach. The union of all the stable manifolds of all the points of the torus will form a smooth N dimensional manifold on which the symplectic form vanishes. Notice that it is impossible to have any manifold of dimension greater than N on which the symplectic form vanishes. These manifolds of maximal dimension on which the symplectic form vanishes are usually called Lagrangian submanifolds and are known to enjoy remarkable topological properties.

Notice that since the motion on the torus is conjugate to a rotation the infinitesimal displacements tangent to the torus will not grow under positive or negative iteration. Since the symplectic form is to be preserved under time evolution, the conjugate directions to these infinitesimal displacements cannot grow under iteration either. It is easy to check that the symplectic conjugates to d independent vectors should also be independent. Moreover, since the symplectic form vanishes on the torus, the span of the vectors symplectically conjugate to tangents should not contain any tangent vector. A torus with the motion conjugate to a rotation and on which the symplectic form vanishes should, therefore, have $2d$ neutral directions. So "whiskered tori" have as many hyperbolic directions as possible under the other conditions. It is for such tori that there are perturbation theories; e.g. [Gra74],[Zeh76].

Only a part of the interaction term, f, in (1) is relevant for the construction of whiskered tori. Define the resonant part of the interaction by

$$\begin{aligned} f^R(I,\phi) &= \frac{1}{T}\int_0^T f(I,\phi+\omega(I^*)t)dt \\ &= \sum_{\substack{\nu\in\mathbb{N}^N \\ \nu\cdot\omega(I^*)=0}} \hat{f}_\nu(I)e^{2\pi i\nu\cdot\phi} \end{aligned} \tag{4}$$

where $\omega(I) = h_I(I)$. Since the components of $\omega(I^*)$ are rationally related there will be infinitely many $\nu \in \mathbb{Z}^N$ such that $\nu \cdot \omega(I^*) = 0$. Let ϕ^* be a point at which $f^R_\phi(I^*,\phi^*) = 0$. (Since $f^R(I,\phi)$ is a smooth function on a compact manifold such points must exist.) We have to consider the $N \times N$

matrix $h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$. In particular, suppose that this matrix has only a single zero eigenvalue. We then have,

Theorem 1.1. *Assume that h,I^*, $\omega(I^*)$ are as in* (1), (2),(3), *H sufficiently differentiable and that certain non-degeneracy conditions –detailed in section 3 – are satisfied. If the matrix $\beta = -h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$ has k positive eigenvalues, $N - (k + \ell + 1)$ negative eigenvalues, ℓ eigenvalues of non–zero imaginary part and one zero eigenvalue, and if all the negative eigenvalues are different, then, for all positive ϵ sufficiently small,* (1) *has whiskered tori of dimension $N - k - \ell$, with $k + \ell$ dimensional stable and unstable whiskers. If the positive eigenvalues of β are different, for ϵ sufficiently small and negative, the hamiltonian* (1) *has whiskered tori of dimension $k + 1$ with $N - k - 1$ dimensional stable and unstable manifolds.*

Remark 1.2. The matrix β is not necessarily symmetric, but nevertheless, in many important cases, all its eigenvalues are real. We recall, for example, the following result from [WR71] p. 303 (See also [GVL83] sec. 8.6)

Proposition 1.3. *Let A, B be symmetric matrices and let A be positive definite. Then, the eigenvalues of AB are real.*

Proof. By the Cholesky decomposition theorem (see [GVL83] sec. 5.2), we can find a lower triangular invertible matrix L such that $A = LL^T$. Introducing $w = L^{-1}v$, the eigenvalue problem $ABv = \lambda v$ is equivalent to the eigenvalue problem $L^T BLw = \lambda w$. Since the matrix $L^T BL$ is symmetric, all the eigenvalues are real. □

Remark 1.4. We note that $-h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$ will have exactly one zero eigenvalue for h and f in an open and dense set of functions. By an elementary application of Morse theory, in a C^2 open and dense set of functions f there will be points $\phi_1, \phi_2, \ldots, \phi_N$ at which $f^R_{\phi\phi}(I^*, \phi_j)$ has k positive eigenvalues $N - k - 1$ negative eigenvalues and one zero eigenvalue.

By the proof of Proposition 1.3, we see that, when $h_{II}(I^*)$ is positive definite, the eigenvalues of $h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$ are the eigenvalues of $L^T f^R_{\phi,\phi}(I^*, \phi)L$ for some conveniently chosen L. It follows from Sylvester's law of inertia (See [GVL83] p. 274 ff.) that, in this case, the number of positive and negative eigenvalues of $h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$ are the same as those of $f^R_{\phi\phi}(I^*, \phi^*)$. So that when $h_{II}(I^*)$ is positive definite, and $\phi_1, \ldots, \phi_N$ are as above, $h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*_j)$ has k positive eigenvalues $N - k - 1$ negative eigenvalues and one zero eigenvalue.

Remark 1.5. It will also follow from the details of the proof of Theorem 1.1 that the other nondegeneracy conditions we require will hold in a C^5 open and dense set. Notice that the open and dense sets we consider have quite a lot of structure. In certain regions, their complement is a manifold of strictly positive codimension. Thus we have

Corollary 1.6. *Define $\mathcal{P}^k$ to be the set of $h \in C^k$ such that h_{II} is a positive definite matrix for all values of $I \in V$. For h and f in a C^5 open and dense set of $\mathcal{P}^k \times C^k$, and ϵ sufficiently small* (1) *has whiskered tori of all dimensions between 1 and N-1.*

Zehnder outlined a method in [Zeh88] to show that a generic perturbation of an integrable system leads to the existence of invariant whiskered tori. His method consisted–roughly–in showing that one could add a generic perturbation to the original hamiltonian (1) in such a way that the original invariant torus becomes hyperbolic and then, that these invariant tori persist with the full perturbation.

Our method, on the other hand consists in a sequence of perturbation theories. First, an averaging method gives us the existence of perturbed periodic orbits as well as control on the eigenvalues of the time T map for a carefully chosen intermediate system. An elementary application of the implicit function theorem then gives the same result for our original system. In a second stage, we restrict ourselves to the center manifold and use K. A. M. theory to construct invariant tori near this periodic orbit. Finally, we use the perturbation theory for partially hyperbolic systems to show that these tori are in fact whiskered tori.

This procedure is fully constructive, so that given a specific system it is possible to decide by a finite computation whether our theorem applies or not. We also have explicit expressions for the expansions of the invariant tori and their stable and unstable manifolds. It looks plausible that by pushing these calculations further than we did, one could establish more properties of the tori and their whiskers. A numerical computation of these invariant objects also seems quite feasible when they are of low dimension.

We observe that to analyze the motion on the center stable manifold, one could use other methods than K. A. M. theory. When the center manifold is two dimensional, one could use the theory of monotone twist maps (see [Che85, Mat86, Ban88] for reviews) to produce invariant Cantor sets with minimal orbits. It seems that significant parts of this theory carry over to higher dimensions [Kat92]. Other interesting invariant sets such as homoclinic orbits, etc., have been known for a long time, and could also be shown to exist in these systems. From the point of view of explaining diffusion, the intersection of invariant manifolds corresponding to any of those sets could play the same role as the intersection of whiskers of tori–this is the reason we formulate the stable and unstable manifold theorems in enough generality to cover all these possibilities. It seems quite possible that one could adapt the method of [Dou88] in such a way that the role of the periodic points is played by these other invariant sets.

The use of the center manifold theorem has a technical shortcoming. We can only conclude that our invariant tori are finitely differentiable even if the original hamiltonian and its perturbation are C^∞ or C^ω. This restriction is unnatural and somewhat more elaborate methods can eliminate it. In future papers, we plan to come back to these issues.

Since the most delicate step of our construction–the K. A. M. computation–is done in a low dimensional manifold, our approach could work in some infinite dimensional systems such as partial differential equations for which there are good substitutes of the averaging method to produce periodic orbits. We plan to apply this approach to the Boussinesq equation in future work.

One particularly interesting result in the study of Arnol'd diffusion is the unpublished thesis of A. Delshams, which was brought to our attention by C. Simó after the main results of this paper were completed. Even if the methods and the model considered are different from those of this paper, the point of view is similar in that only constructive methods are used. That paper studies a hamiltonian system in the neighborhood of an elliptic fixed point at which the eigenvalues of the derivative of the hamiltonian are resonant. One then makes canonical transformations that reduce the hamiltonian to a canonical normal form up to an error which vanishes to high order in the distance to the fixed point. By inspection, the main term of the canonical form, is shown to have invariant tori of dimension $N-1$, where N is the number of degrees of freedom, and that they have hyperbolic directions. By invoking an appropriate perturbation theory for those objects it is possible to show that these invariant objects predicted by the main term of the normal form are present for the full system. (One needs to modify somewhat the exposition in [Del83] so that the perturbation theory that is invoked is that of [Gra74] or [Zeh76] rather than the perturbation theory for normally hyperbolic manifolds of [HPS77] since as we detail below the whiskered tori are not normally hyperbolic.) Since the invariant manifolds do not leave the neighborhood in which the normal form is defined, it is possible to perform the calculations required by the first order Melnikov's method and show that, in this approximation, the manifolds cross transversally. Unfortunately, as Delshams points out, this does not suffice to establish that the whiskers cross transversally in the real model since the errors incurred by first order Melnikov's method are much bigger than the first order results. Nevertheless, since the method is quite constructive, for concrete sytems, Delshams [Del88] has shown to us how it is possible to perform reliable computer calculations that, if taken at face value would establish the existence of Arnol'd diffusion in a certain scale for the concrete system being considered. (See [BGGS84] for related calculations.) The calculations needed fall within the framework of those that can be turned into rigorous proofs by a careful automatic analysis of the errors involved.

2. Whiskered Periodic Orbits

In the present section we will show how certain periodic orbits in the integrable problem persist in the nearly integrable case. We will also construct the stable, unstable, and center manifolds of this periodic orbit. Note that the integrable problem will typically have uncountably many periodic orbits of fixed period which fill an invariant torus. In the perturbed problem only finitely many periodic orbits will generally survive. We use the method of averaging to determine which of the orbits of the unperturbed problem continue to an orbit of an intermediate perturbed problem. We get sufficient control of the periodic orbits in the "averaged" problem to be able to conclude that we can use the implicit function theorem to control the remaining part of the perturbation, and to be able to apply stable, unstable and center manifold theorems.

The problem we will analyze is a nearly integrable hamiltonian expressed in action-angle variables,

$$H(I,\phi) = h(I) + \epsilon f(I,\phi), \tag{5}$$

with $I \in V \subset \mathbb{R}^N$ and $\phi \in \mathbb{T}^N \equiv \mathbb{R}^N/\mathbb{N}^N$. Using (2), the set of frequency vectors of the unperturbed problem $\Omega = \{\omega(I) \equiv \frac{\partial h}{\partial I}(I) \big| I \in V\}$ has non empty interior. In particular, it contains frequency vectors of the form $\frac{1}{T}(n_1,\ldots,n_N)$ with $T \in \mathbb{R}, n_j \in \mathbb{N}$. These frequencies correspond to periodic orbits for the unperturbed problem and satisfy:

$$|\nu \cdot \omega(I^*)| \geq c > 0 \tag{6}$$

with $c = 1/T$ whenever $\nu \cdot \omega(I^*) \neq 0$.

If I^* is an interior point of V one can compute the time-T map for this periodic orbit. Attempting to apply the implicit function theorem to construct an orbit for the perturbed problem fails though, essentially because there is no hyperbolicity in the unperturbed problem. This is where our method differs from that of [AA68] or [HM82], for instance. One of the strengths of the present approach is that it shows how the hyperbolicity present in the whiskered torus is generated.

Let f^R be the resonant part of the interaction defined in (4) and ϕ^* a point for which $f^R_\phi(I^*,\phi^*) = 0$. The main result of the present section is

Theorem 2.1. *If the matrix $\beta \equiv -h_{II}(I^*) f^R_{\phi\phi}(I^*,\phi^*)$ has only a single zero eigenvalue, and $|\epsilon|$ is sufficiently small, then the hamiltonian* (1) *has a periodic orbit, whose period differs from the periodic orbit of the unperturbed hamiltonian passing through I^* by at most $\mathcal{O}(\epsilon)$. Suppose further that k of the eigenvalues of the matrix above are positive, $N-(k+\ell+1)$ are negative, and ℓ of the eigenvalues have non-zero imaginary part. When ϵ is positive and the negative eigenvalues of β are all different the periodic orbit of* (1) *has $k+\ell$ dimensional stable and unstable manifolds, and a $2(N-(k+\ell))$ dimensional center manifold. When ϵ is negative and the positive eigenvalues of β are all different, the periodic orbit of* (1) *has $(N-(k+1))$ dimensional stable and unstable manifolds and a $2(k+1)$ dimensional center manifold.*

Remark 2.2. For fixed I^* there will be many points ϕ^* for which $f^R_\phi(I^*,\phi^*) = 0$. In fact these points will occur in one parameter families corresponding to motion along the unperturbed periodic orbit. It is easy to check that the number of positive and negative eigenvalues of $h_{II}(I^*) f^R_{\phi\phi}(I^*, \phi^*)$ is independent of the point ϕ^* we pick in such a family.

Our first task in proving Theorem 2.1 will be to clarify the existence of critical points of functions such as f^R and their properties. We prove:

Lemma 2.3. *Suppose $f(I, \phi)$ is a C^r function $r = 2, 3 \ldots, \infty, \omega$ on $V \times \mathbb{T}^N$, with $0 \in V \subset \mathbb{R}^N$, which satisfies:*

(i) *$f(I, \phi + t\omega^*) = f(I, \phi)$, for all t, where ω^* is an N-dimensional vector whose components are multiples of an integer vector.*
(ii) *There is a point ϕ_0 such that $f_\phi(0, \phi_0) = 0$ and $f_{\phi\phi}(0, \phi_0)$ has exactly one zero eigenvalue. (Note that condition (i) implies that $f_{\phi\phi}(0, \phi_0)$ always has at least one zero eigenvalue, with eigenvector ω^*.)*

Then there is a C^r function, $\phi(I)$, defined for $|I|$ sufficiently small such that

$$f_\phi(I, \phi(I)) = 0.$$

Proof. Setting $\tilde{\phi}$ to represent the orbit of the rational flow going though ϕ, we can identify the functions satisfying (i) with functions $\tilde{f}(I, \tilde{\phi})$ on $V \times \mathbb{T}^{N-1}$. (The space of orbits of the rational flow on $\mathbb{T}^N$ is $\mathbb{T}^{N-1}$.) We can see that $\tilde{f}_{\tilde{\phi}}(I, \tilde{\phi}) = 0$ is equivalent to $f_\phi(I, \phi) = 0$. By choosing an appropriate set of coordinates, in which the direction of ω^* is a coordinate vector, we can see that the matrix $f_{\phi\phi}(I, \phi)$ is just the matrix of $\tilde{f}_{\tilde{\phi},\tilde{\phi}}(I, \tilde{\phi})$ with an extra row and column of zeros. Therefore, the eigenvalues of $f_{\phi\phi}(I, \phi)$ are those of $\tilde{f}_{\tilde{\phi},\tilde{\phi}}(I, \tilde{\phi})$ and another zero.

Condition (ii) becomes simply that $\tilde{f}_{\tilde{\phi},\tilde{\phi}}(0, \tilde{\phi}_0)$ is non singular. Then, we can apply the implicit function theorem to the equation $\tilde{f}_{\tilde{\phi}}(I, \tilde{\phi}) = 0$ to find a C^r function $\bar{\phi}(I)$ satisfying $\tilde{f}_{\tilde{\phi}}(I, \bar{\phi}(I)) = 0$. Once we have this function, we can find the remaining coordinate as a function of I in any way we please. This ambiguity is real, because we are parameterizing orbits by a single point on the orbit. We will assume that we make the choice in a differentiable way. □

The same construction allows us to prove a lemma that we will use later

Lemma 2.4. *Consider the set of C^r functions $r = 2, 3, \ldots, \infty, \omega$ satisfying (i) as above. Then, for all the functions in an open dense set there are at least $\binom{N-1}{k}$ points $\tilde{\phi} \in \mathbb{T}^{N-1}$ satisfying $\tilde{f}_{\tilde{\phi}}(0, \tilde{\phi}) = 0$ and at which $\tilde{f}_{\tilde{\phi}\tilde{\phi}}(0, \tilde{\phi})$ has exactly k negative eigenvalues . Moreover, all those negative eigenvalues are different.*

Proof. This is an standard application of the Morse inequalities [Mil63] (the i^{th} Betti number for the n torus is $\binom{n}{i}$). The fact that the negative eigenvalues can be made different in an open dense set is a consequence of transversality theory [AR67]. □

We begin the proof of Theorem 2.1 by making some preliminary changes of variables. We first translate the origin of the coordinate system to I^* by setting $J = I - I^*$. (In a slight abuse of notation, we will denote the transformed hamiltonian as $H(J, \phi)$.) Given a function of $g(J, \phi)$ we define its power series with respect to J by $g(J, \phi) = \sum_\alpha g_\alpha(\phi) J^\alpha$. Here, $\alpha \in (\mathbb{N})^N$, $J^\alpha = \prod_i J_i^{\alpha_i}$, and $|\alpha| = \sum_i |\alpha_i|$.

We now make a canonical transformation to make more explicit the long time behavior. For the sake of convenience, we will use the Lie transform method (see e.g. [Car81]). The generating function of the transformation will be $\chi(J, \phi) = \sum_{\ell=0}^{4} \sum_{\substack{\alpha \\ |\alpha|=\ell}} \chi_\alpha(\phi) J^\alpha$. We will choose χ so that:

$$H + \{\chi, H\} = h(J) + \epsilon f^R(J, \phi) + \text{ higher order terms .}$$

Exactly what we mean by "higher order terms" is explained below.

Consider the coefficient of J^α in $\{\chi, h\}$. We have

$$\{\chi, h\}_\alpha = \{\chi_\alpha, \omega(I^*) \cdot J\} + \sum_{\substack{\beta,\gamma: \\ |\beta|+|\gamma|=|\alpha|+1, |\gamma|>1}} \{J^\beta \chi_\beta, J^\gamma h_\gamma\}_\alpha$$

Thus, we define χ_α to be the solution of

$$\omega(I^*) \frac{\partial \chi_\alpha}{\partial \phi} = \epsilon (f^R - f)_\alpha - \sum_{\substack{\beta,\gamma: \\ |\beta|+|\gamma|=|\alpha|+1, |\gamma|>1}} \{J^\beta \chi_\beta, J^\gamma h_\gamma\}_\alpha \tag{7}$$

(Note that in the sum on the right hand side of (7), $|\beta| \le |\alpha| - 1$, so we solve for χ_α by induction on $|\alpha|$.) In the sequel we will use this proposition for $r = 4$ only but the general case is not any harder.

Proposition 2.5. *If H is C^r,$r \ge 1$, then*

$$\chi_\alpha(\phi) = \frac{1}{T} \int_0^T \{\epsilon (f - f^R)_\alpha(\phi + \omega(I^*)t) + \sum_{\substack{\beta,\gamma: \\ |\beta|+|\gamma|=|\alpha|+1, |\gamma|>1}} \{J^\beta \chi_\beta(\phi + \omega(I^*)t), J^\gamma h_\gamma\}_\alpha\} t dt$$

is well defined and solves (7) for $0 \le |\alpha| \le r$. Moreover, f^R and χ are C^r functions. The linear operators that to each f associate f^R and χ are bounded from C^r to C^r.

Proof. The proof is by induction on $|\alpha|$. Assume that the proposition is true for $|\alpha| < m$. We prove it for any α with $|\alpha| = m$. (The proof also holds if $m = 0$, which allows us to start the induction.)

The boundedness in C^r follows from the fact that all the translates by different $\omega(I^*)t$ have the same C^r norm and we are superimposing them with a weight given by a function in L^1.

Let $g_\alpha(\phi) = \sum_{\substack{\beta,\gamma: \\ |\beta|+|\gamma|=|\alpha|+1, |\gamma|>1}} \{J^\beta \chi_\beta(\phi), J^\gamma h_\gamma\}_\alpha$.

Observe that

$$\sum_{i=1}^{N} \omega(I^*)_i \frac{\partial}{\partial \phi_i} \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} =$$

$$= \frac{d}{dt} \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} \ .$$

Proceeding formally we insert the definition of χ and obtain

$$\begin{aligned}
&\sum_{i=1}^{N} \omega(I^*)_i \frac{\partial}{\partial \phi_i} \frac{1}{T} \int_0^T \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} t dt \\
&= \frac{1}{T} \int_0^T \sum_{i=1}^{N} \omega(I^*)_i \frac{\partial}{\partial \phi_i} \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} t dt \\
&= \frac{1}{T} \int_0^T \left(\frac{d}{dt} \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} \right) t dt
\end{aligned}$$

Integrating by parts and using the definition of f^R and χ_β, with $|\beta| < m$ we have:

$$\frac{1}{T} \int_0^T \left\{ \epsilon (f - f^R)_\alpha (\phi + \omega(I^*)t) + g_\alpha(\phi + \omega(I^*)t) \right\} dt = 0$$

so we get the desired result.

The interchange of order of the integral and the derivative we used in the above argument is easily justified by observing that the resulting integral converges uniformly. $\square$

We now proceed to estimate the remainder of this formal calculation. We will show bounds which are uniform in a certain domain which, even if it tends to zero when ϵ tends to zero, is big enough for the following. Let $\Phi^1(J, \phi)$ be the time one map of the system whose hamiltonian is the function $\chi(J, \phi)$ constructed in the previous proposition. We are interested in the transformed hamiltonian,

$$H^R(J, \phi) = H \circ \Phi^1(J, \phi)$$

This may be rewritten as

$$\begin{aligned}
H^R(J, \phi) &= H(J, \phi) + \int_0^1 \{\chi, H\} \circ \Phi^t(J, \phi) dt \\
&= H(J, \phi) + \{\chi, H\} - \int_0^1 (t-1) \{\chi, \{\chi, H\}\} \circ \Phi^t(J, \phi) dt
\end{aligned} \tag{8}$$

Note that (7) implies that χ and all of its derivatives are $\mathcal{O}(\epsilon)$ so the term containing the integral in (8) is $\mathcal{O}(\epsilon^2)$ (along with all of its derivatives).

We now examine the term $\{\chi, H\}$. Given a function $g(J, \phi)$, define $g_{\geq}(J, \phi) = g(J, \phi) - \sum_{\alpha:|\alpha| \leq 4} g_\alpha(\phi) J^\alpha$. Also, set $f^{NR} = f - f^R$. Using the definition of χ we see that

$$\epsilon f(J, \phi) + \{\chi, H\}(J, \phi) = \epsilon f^R(J, \phi) + \epsilon f^{NR}_{\geq}(J, \phi) + \{\chi, H\}_{\geq}(J, \phi) \ .$$

Note that if $H \in C^5$, then $\epsilon f^{NR}_{\geq}$, $\{\chi, H\}_{\geq}$ and their derivatives of order two or less will be $\mathcal{O}(\epsilon^2)$ in any neighborhood of size $\mathcal{O}(\epsilon^\delta)$ about $J = 0$, if $\delta \geq 2/3$. If we define a new hamiltonian

$$H^N(J, \phi) = h(J) + \epsilon f^R(J, \phi) \ , \tag{9}$$

then we have:

Proposition 2.6. *On any neighborhood of size $\mathcal{O}(\epsilon^\delta)$ about $J = 0$, with $\delta \geq 2/3$*

$$||H^R - H^N||_{C^2} = \mathcal{O}(\epsilon^2) \ .$$

We will establish the existence of periodic orbits for the hamiltonian H^N and then, perform an elementary perturbation theory to establish the existence of periodic orbits for the hamiltonian H^R, in which we are really interested.

The equations of motion for the hamiltonian H^N are

$$\dot{J} = -H^N_\phi = -\epsilon f^R_\phi(J, \phi)$$

$$\dot{\phi} = H^N_J = h_J(J) + \epsilon f^R_J(J, \phi) = \omega(I^*) + h_{II}(I^*)J + R_{3J}(J) + \epsilon f^R_J(J, \phi) \ . \tag{10}$$

Here, $R_3(J) = \sum_{\alpha:|\alpha|\geq 3} h_\alpha J^\alpha$. Lemma 2.3 implies that there is a smooth function $\tilde{\phi}(J)$, defined for $|J|$ sufficiently small such that $f^R_\phi(J, \tilde{\phi}(J)) = 0$. Note that the domain on which $\tilde{\phi}(J)$ is defined is independent of ϵ. Now consider the function

$$\mathcal{H}(\epsilon, J) = H^N_J(J, \tilde{\phi}(J))$$

Then $\mathcal{H}(0, 0) = \omega(I^*)$, and $\mathcal{H}_J(0, 0) = h_{II}(I^*)$. Since this matrix is invertible, the implicit function theorem implies that there is some $\epsilon_0 > 0$ such that if $|\epsilon| < \epsilon_0$, there exists a smooth function $\tilde{J}(\epsilon)$ such that

$$-\epsilon f^R_\phi(\tilde{J}(\epsilon), \tilde{\phi}(\tilde{J}(\epsilon))) = 0$$

$$\omega(I^*) + h_{II}(I^*)\tilde{J}(\epsilon) + R_{3J}(\tilde{J}(\epsilon)) + \epsilon f^R_J(\tilde{J}(\epsilon), \tilde{\phi}(\tilde{J}(\epsilon))) = \omega(I^*).$$

Recall that $f^R(J, \phi + \omega(I^*)t) = f^R(J, \phi)$. Thus, we have proved

Proposition 2.7. *There exists a constant $\epsilon_0 > 0$ such that if $|\epsilon| < \epsilon_0$, the hamiltonian system $H^N(J, \phi)$ has a periodic orbit with frequency $\omega(I^*)$.*

We now proceed to compute the derivative of the time T map of the hamiltonian H^N around the periodic orbit constructed above and compute its eigenvalues as well as the norm of the inverse of this derivative.

We note that from (10), $J(t) = J(0) + \mathcal{O}(\epsilon)$. This in turn implies that

$$\phi(t) = \phi(0) + [\omega(I^*) + h_{II}(I^*)J(0) + R_{3J}(J(0))]t + \mathcal{O}(\epsilon).$$

We then feed this back into the equation for $\dot{J}$ in (10) and find

(11)
$$\begin{aligned} J(t) &= J(0) - \epsilon \int_0^t f_\phi^R(J(0), \phi(s))ds + \mathcal{O}(\epsilon^2) \\ &= J(0) - \epsilon \int_0^t f_\phi^R(J(0), \phi(0) + [h_{II}(I^*)J(0) + R_{3J}(J(0))]s)ds + \mathcal{O}(\epsilon^2). \end{aligned}$$

The last equation used the fact that $f_\phi^R(J, \phi + \omega(I^*)t) = f_\phi^R(J, \phi)$.

Finally, we substitute this expression for $J(t)$ into the equation for $\dot{\phi}$ and find

(12)
$$\begin{aligned} \phi(t) &= \phi(0) + \omega(I^*)t + \int_0^t h_{II}(I^*)J(s) + R_{3J}(J(s))ds + \\ &\quad + \epsilon \int_0^t f_J^R\left(J(0), \phi(0) + [h_{II}(I^*)J(0) + R_{3J}(J(0))]\, s)\right) ds + \mathcal{O}(\epsilon^2) \end{aligned}$$

Remark 2.8. It is important to stress that in the above expressions for $J(t)$ and $\phi(t)$ both the omitted terms and all their derivatives are $\mathcal{O}(\epsilon^2)$.

Remark 2.9. We also have to point out that we are assuming that the time t remains in the bounded interval $[0, T]$. For any given T, the $\mathcal{O}(\epsilon^2)$ terms can be chosen uniformly, but these choices will have to change if we change T. A more precise notation for the errors here would have been $\mathcal{O}(T\epsilon^2)$. Similar remarks will be valid for subsequent estimates that use this one and involve integrations with respect to time.

Now consider the map that sends

$$\mathcal{T} : \begin{pmatrix} J \\ \phi \end{pmatrix} \to \begin{pmatrix} J(T) \\ \phi(T) \end{pmatrix}$$

where T is the period of the periodic orbit constructed in Proposition 2.7. We know that any point on this orbit is a fixed point for $\mathcal{T}$. We now compute the derivative of $\mathcal{T}$ at this fixed point, (J^*, ϕ^*). We remark that $|J^*| \approx \mathcal{O}(\epsilon)$.

Then

(13)
$$\begin{aligned} \frac{\partial J(T)}{\partial J} &= 1 - \epsilon \int_0^T f_{\phi J}^R(J, \{\phi^* + [h_{II}(I^*)J + R_{3J}(J^*)]s\})ds \\ &\quad - \epsilon \int_0^T f_{\phi\phi}^R(J, \{\phi^* + [h_{II}(I^*)J + R_{3J}(J^*)]s\})[h_{II}(I^*) + R_{3JJ}(J)]sds \\ &\quad + \mathcal{O}(\epsilon^2) \\ &= 1 - \epsilon \int_0^T \left\{ f_{\phi J}^R(J, \phi) + f_{\phi\phi}^R(J, \phi) h_{II}(I^*)s \right\} ds + \mathcal{O}(\epsilon^2). \end{aligned}$$

so that

$$\begin{aligned} \left.\frac{\partial J(T)}{\partial J}\right|_{J^*,\phi^*} &= \mathbf{1} - \epsilon T f^R_{\phi J}(J^*,\phi^*) + \epsilon\frac{T^2}{2} f^R_{\phi\phi}(J^*,\phi^*) h_{II}(I^*) + \mathcal{O}(\epsilon^2) \\ &= 1 + \epsilon A + \mathcal{O}(\epsilon^2). \end{aligned} \tag{14}$$

By very similar computations, whose details we omit, we find that:

$$D_{(J^*,\phi^*)}\mathcal{T} = \begin{pmatrix} 1+\epsilon A & -\epsilon T f^R_{\phi\phi}(J^*,\phi^*) \\ T h_{II}(I^*) + \epsilon B & 1+\epsilon C \end{pmatrix} + \mathcal{O}(\epsilon^2). \tag{15}$$

where

$$C = T f^R_{J\phi}(J^*,\phi^*) + \frac{T^2}{2} h_{II}(I^*) f^R_{\phi\phi}(J^*,\phi^*)$$

and

$$\begin{aligned} B =& \frac{T^2}{2} h_{II}(I^*) f^R_{\phi J}(J^*,\phi^*) + \frac{T^3}{6} h_{II}(I^*) f^R_{\phi\phi}(J^*,\phi^*) h_{II}(I^*) \\ &+ T f^R_{JJ} + \frac{T^2}{2} f^R_{J\phi} h_{II}(I^*). \end{aligned}$$

If $\lambda_1,\dots,\lambda_{2N}$ are the eigenvalues of $D_{(J^*,\phi^*)}\mathcal{T}$, it is clear that $\lambda_j = 1+\mu_j$, where $\mu_1,\dots,\mu_{2N}$ are the eigenvalues of

$$U = \begin{pmatrix} \epsilon A & -\epsilon T f^R_{\phi\phi}(J^*,\phi^*) \\ T h_{II}(I^*) + \epsilon B & \epsilon C \end{pmatrix} + \mathcal{O}(\epsilon^2).$$

Furthermore, $\mu_1^2,\dots,\mu_{2N}^2$ are of the form $\mu_j^2 = \epsilon\nu_j^2$ where $\nu_1^2,\dots,\nu_{2N}^2$ are the eigenvalues of

$$V^2 = (1/\epsilon)U^2 = \begin{pmatrix} -T^2 f^R_{\phi\phi}(J^*,\phi^*) h_{II}(I^*) & 0 \\ T h_{II}(I^*)A + TC h_{II}(I^*) & -T^2 h_{II}(I^*) f^R_{\phi\phi}(J^*,\phi^*) \end{pmatrix} + \mathcal{O}(\epsilon).$$

Ignoring the $\mathcal{O}(\epsilon)$ correction terms, and noting that the spectrum of $f^R_{\phi\phi}(J^*,\phi^*)h_{II}(I^*)$ is the same as that of $h_{II}(I^*)f^R_{\phi\phi}(J^*,\phi^*)$ (Notice that $f^R_{\phi\phi}(J^*,\phi^*)h_{II}(I^*) = \left(h_{II}(I^*)f^R_{\phi\phi}(J^*,\phi^*)\right)^T$), we see that, if $\epsilon = 0$, $\nu_1^2 = \nu_{N+1}^2,\dots,\nu_N^2 = \nu_{2N}^2$, with $\nu_1^2,\dots,\nu_N^2$ the eigenvalues of $-T^2 f^R_{\phi\phi}(J^*,\phi^*)h_{II}(I^*)$, which we denote as $\kappa_1,\dots,\kappa_N$. By assumption exactly one of these eigenvalues is zero. For the sake of notation, choose it to be κ_N. The eigenvalues of V^2 will be algebraic functions of ϵ, so for ϵ small we have

$$\nu_1^2 = \kappa_1 + \mathcal{O}(\epsilon^\alpha),\dots,\nu_{2N}^2 = \kappa_N + \mathcal{O}(\epsilon^\alpha),$$

with $\alpha > 0$.

Working our way back to the eigenvalues of $D_{(J^*,\phi^*)}\mathcal{T}$ we see that

$$\begin{aligned} \lambda_j &= 1 + \sqrt{\epsilon\kappa_j + \mathcal{O}(\epsilon^{1+\alpha})} \\ \lambda_{j+N} &= 1 - \sqrt{\epsilon\kappa_j + \mathcal{O}(\epsilon^{1+\alpha})} \quad j = 1,\dots,N. \end{aligned}$$

Note that $\lambda_N = \lambda_{2N} = 1$, but but all other eigenvalues must be different from 1 by an amount of $\mathcal{O}(\epsilon^{1/2})$.

We would now like to apply the Implicit Function Theorem to the map $\mathcal{T}$ in order to prove that the transformed hamiltonian H^R, (of (9)) has a periodic orbit.

Define $H_\mu = H^N + \mu(H^R - H^N)$. Let $\mathcal{T}_\mu$ be the time T return map of H_μ. Then $\mathcal{T}_0$ has a fixed point, but the presence of the eigenvalues $\lambda_N = \lambda_{N+1} = 1$ in the spectrum of $D\mathcal{T}_0$ prevents us from applying the ordinary implicit function theorem. Moreover, since we are making perturbation theory from a degenerate situation where the Implicit Function Theorem indeed does not apply, we will have to be quite careful with the set up and the dependence on ϵ of all the relevant parameters.

It is easy to modify the Implicit Function Theorem along the lines of Theorem 5.6.6 of [AM78] (see also [Dui84] and [Hal80]) to eliminate this somewhat artificial problem. (Artificial, since one of the eigenvalues with modulus 1 results from motion along the periodic orbit and can be removed by taking a section transverse to the flow, while the other results from energy conservation.) Because one must keep careful control of the various constants in the problem the statement and proof of our theorem is somewhat more involved than that in [AM78] but we follow the notation of that reference.

Let (J^*, ϕ^*) be points on the periodic orbit of H^N. Locally in a neighborhood of (J^*, ϕ^*) we can regard $V \times T^N$ as $\mathbb{R}^{2N} \equiv \mathbb{R}^N \times \mathbb{R}^N$. Without loss of generality, we may assume that (J^*, ϕ^*) is at the origin of our coordinate system.

We now choose an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ for $\mathbb{R}^N$ with $\mathbf{e}_1 = \omega(I^*)$ and introduce new coordinates $\mathbf{x}$ in a neighborhood of (J^*, ϕ^*). Let $\mathbf{x}$ be coordinates with respect to the basis for $\mathbb{R}^{2N} \equiv \mathbb{R}^N \times \mathbb{R}^N$ in which we take $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ as a basis for each factor of $\mathbb{R}^N$.

Because energy is conserved, we will work in energy surfaces $\Sigma(E, \mu) = H_\mu^{-1}(E)$. For ϵ sufficiently small, (depending only on $\omega(I^*)$), $\Sigma(E, \mu)$ will be a smooth submanifold for all $\mu \in [0, 1]$, and E sufficiently close to $E_0 \equiv h(I^*)$. Furthermore, since the gradient of H^N is parallel to $(\omega(I^*), 0)$ (here, 0 is an N-vector) when $\epsilon = 0$, $\Sigma(E, \mu)$ will be a graph of a function of $(\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{2N})$ for ϵ sufficiently small, (where "sufficiently small" again depends only on $\omega(I^*)$). Denote this function by $\Phi(\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{2N}, E, \mu)$.

Let X_μ be the vector field and $\mathcal{T}_\mu^t$ the flow associated with the hamiltonian H_μ. Then the periodic orbits of $\mathcal{T}_\mu^t$ are zeros of

$$\Theta(\mathbf{x}, t, \mu) = \mathcal{T}_\mu^t(\mathbf{x}) - \mathbf{x} \tag{16}$$

If we define Π^E to be the projection of $\mathbf{x}$, onto its last $2N-1$ components, then since $\mathcal{T}_\mu^t$ leaves $\Sigma(E, \mu)$ invariant, zeros of (16) are equivalent to zeros of

$$\Psi(\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{2N}, t, E, \mu) = \Pi^E(\Theta(\Phi(\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{2N}, E, \mu), t, \mu)).$$

We now construct a Poincaré map for our system by noting that for $\epsilon = 0$, the vector field for our system is $X_0 = (0, \omega(I^*))$, so for ϵ sufficiently small, the hyperplane $\mathbf{x}_{N+1} = 0$ will remain transverse to the flow. With this in mind, we define, $\mathbf{y} = (\mathbf{x}_2, \dots, \mathbf{x}_N, \mathbf{x}_{N+2}, \dots, \mathbf{x}_{2N})$. Let Π^Y be the projection of $(\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{2N})$ onto $\mathbf{y}$. For ϵ sufficiently small, the flow $\mathcal{T}^t_\mu$ will map $B(r, 0) \to B(2r, 0)$, for all $\mu \in [0, 1]$. Let $\tilde{T}(y, \mu)$ be the return time. (In particular, $\tilde{T}(0, 0) = T$.) We then define the Poincaré map by

$$P(\mathbf{y}, E, \mu) = \Pi^Y \Psi(\mathbf{y}, \tilde{T}(y, \mu), \mu).$$

By the usual smooth dependence on initial conditions for differential equations, provided that $H^R, H^N \in C^3$ – which follows from $h, f \in C^5$ we have that $P \in C^2$ and the second derivative is uniformly bounded in a neighborhood of the line $\mu \in [0, 1]$. If $\mathbf{y}_0$ is the zero of $P(\mathbf{y}_0, E, \mu = 0)$ corresponding to the periodic orbit of H^N, we will also have that $||P(\mathbf{y}_0, E, 1)|| \le \mathcal{O}(\epsilon^2)$ because the difference between the vector fields corresponding to H^R and H^N is $\mathcal{O}(\epsilon^2)$ and we are only considering a finite interval of time.

Now, we would like to apply some version of the implicit function theorem to show that we can obtain solutions of $P(\mathbf{y}, E, \mu) = 0$ or equivalently of $-P(\mathbf{y}, E, \mu) + \mathbf{y} = \mathbf{y}$. As we will see, the standard version does not apply so that we will need to prove a version specific to our case that will use not only the sizes of the derivatives involved but also the specific form of the equations we computed before.

Next note that the computations leading up to (15) imply that

$$D_y P(\mathbf{y}_0, E_0, 0) = \begin{pmatrix} 0 & 0 \\ A & 0 \end{pmatrix} + \epsilon \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} + \epsilon^2 R_\epsilon \tag{17}$$

where A, α, β, γ and δ are constant matrices, R_ϵ is a continuous matrix valued function, and A and β are invertible. The important observation here is that β is invertible. This follows from the fact that $f^R_{\phi,\phi}(J^*, \phi^*)$ has a single eigenvalue equal to zero whose eigenvector is exactly $\omega(I^*)$, and the coordinates used for the Poincaré map were chosen to exclude this direction.

Our first task is to obtain estimates on $||D_y P(\mathbf{y}_0, E_0, 0)^{-1}||$. It is very tempting to try to derive estimates for this matrix just from the eigenvalues of $D_y P(\mathbf{y}_0, E_0, 0)$. which are easily seen to be $\mathcal{O}(\sqrt{\epsilon})$. Unfortunately, in our situation, the matrix is not selfadjoint and we are starting from a degenerate situation so that the separation between eigenspaces could be very small, making the estimates of the inverse much worse that those of the eigenvalues. A model for the situation is the matrix

$$M_\epsilon = \begin{pmatrix} 0 & \epsilon \\ 1 & 0 \end{pmatrix}$$

whose eigenvalues are $\mathcal{O}(\epsilon^{1/2})$ but whose inverse has norm $\mathcal{O}(\epsilon^{-1})$.

Unfortunately, such estimates on the size of the derivatives are not enough to conclude that the assumptions of the implicit function theorem are verified for ϵ small enough. The one dimensional example: $P_\epsilon(x) = (x - \epsilon)^2 + \epsilon^4$

verifies smallness assumptions on $P_\epsilon(0)$ and $DP_\epsilon(0)^{-1}$ similar to those of P in our case, but clearly, it has no zero. So it is clear that we will have to use the structure of the derivative we just computed.

Proposition 2.10. *Let $M(\epsilon)$ be a matrix valued function from $[-1,1]$ to $\mathcal{M}_{n\times n}$*

$$M(\epsilon) \equiv \begin{pmatrix} 0 & 0 \\ A & 0 \end{pmatrix} + \epsilon \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} + \epsilon^2 R_\epsilon \tag{18}$$

where A, α, β, γ and δ are constant matrices and R_ϵ is a continuous matrix valued function. Assume that A and β are invertible. Then, for $|\epsilon|$ small enough, $M^{qi}(\epsilon)$ defined by

$$M^{qi}(\epsilon) = \epsilon^{-1} \begin{pmatrix} 0 & 0 \\ \beta^{-1} & 0 \end{pmatrix} + \begin{pmatrix} -A^{-1}\delta\beta^{-1} & A^{-1} \\ 0 & -\beta^{-1}\alpha A^{-1} \end{pmatrix} \tag{19}$$

is invertible and satisfies

$$\begin{aligned} ||M(\epsilon)M^{qi}(\epsilon) - Id|| &\le K|\epsilon| \\ ||M^{qi}(\epsilon)M(\epsilon) - Id|| &\le K|\epsilon| \ . \end{aligned} \tag{20}$$

(Here, qi stands for quasi-inverse.) Moreover, suppose $||A - A_0|| \le \rho$, $||\alpha - \alpha_0|| \le \rho$, $||\beta - \beta_0|| \le \rho$, $||\gamma - \gamma_0|| \le \rho$ and $||\delta - \delta_0|| \le \rho$ with ρ sufficiently small (the smallness conditions depend only on A and β and are independent of ϵ). Then, if $M_0^{qi}(\epsilon)$ is defined by substituting for each of the components of M^{qi} the corresponding matrix with subscript zero, we have

$$||M_0^{qi}(\epsilon)M(\epsilon) - Id|| < K_1\rho + K_2\epsilon \tag{21}$$

Proof. The proof of the bounds is just a simple computation. Once we have the bounds, we observe that the imply that MM^{qi} and $M^{qi}M$ are invertible for $|\epsilon|$ small, since then, they are perturbations of the identity. □

The previous inequalities show that $M^{qi}(\epsilon)$ is an approximate inverse for $M(\epsilon)$ in the sense that it produces residuals up to order ϵ. It is possible to show that a true inverse exists, but it could differ from $M^{qi}(\epsilon)$ by $\mathcal{O}(\epsilon^0)$ because the corrections can be introduced as multiplicative factors of the form $Id + \mathcal{O}(\epsilon)$ which, however, has to be multiplied by $\mathcal{O}(\epsilon^{-1})$.

The crucial observation is (21). We will need that the derivatives we have to control have the form of M. Clearly (21) is not true for general perturbations of size ρ.

We now want to consider the existence of solutions of $P(\mathbf{y}, E, \mu) = 0$. Proceeding as in Newton's method, we define

$$N(y, E, \mu) \equiv -\left(D_{\mathbf{y}}P(\mathbf{y}_0, E, \mu)\right)^{qi} P(\mathbf{y}, E, \mu) + \mathbf{y} \ . \tag{22}$$

Since, by assumption $D_{\mathbf{y}}P(\mathbf{y}_0, E, \mu)^{qi}$ is invertible, the fixed points of N will give zeros of P. Clearly one has:

$$D_{\mathbf{y}}N(y, E, \mu) = -\left(D_{\mathbf{y}}P(\mathbf{y}_0, E, \mu)\right)^{qi} D_{\mathbf{y}}P(\mathbf{y}, E, \mu) + Id$$

Furthermore we have already observed that $D_{\mathbf{y}}P(\mathbf{y}, E, \mu)$ can be written in the form $M(\epsilon)$ considered before, with A, α, β, γ, δ being a function of $\mathbf{y}$,E,μ, and A and β invertible. (We computed the first order in ϵ quite explicitly).

Since A and β depend differentiably on the point $\mathbf{y}$, it is clear that they will be invertible for all $\mathbf{y}$,E,μ in a neighborhood of the fixed point. More precisely, in a neighborhood of radius $C_1|\epsilon|^{0.9}$ (C_1 will be chosen later) around $\mathbf{y}_0$ – the point which is a periodic orbit of H^N we have:

$$||D_{\mathbf{y}}P(\mathbf{y}, E, \mu) - D_{\mathbf{y}}P(\mathbf{y}_0, E, \mu)|| \leq K_1 C_1 |\epsilon|^{0.9}$$

We take the convention that K's will denote constants that depend on the system we are considering but which can be chosen uniformly in ϵ small enough. This bound is obvious if we realize that the second derivatives of P can be bounded uniformly for ϵ small enough.

Applying (21) and (20) we get:

$$||D_{\mathbf{y}}N(\mathbf{y}, E, \mu)|| \leq K_2 K_1 C_1 |\epsilon|^{0.9} + K_3|\epsilon| \tag{23}$$

Since $||P(\mathbf{y}_0, E_0, \mu)|| \leq K_4\epsilon^2$ we also have

$$||N(\mathbf{y}_0, E, \mu)|| < K_5|\epsilon|$$

We can see that, by choosing C_1 and then choosing $|\epsilon|$ small enough, we can ensure that the neighborhood of radius $C_1|\epsilon|^{0.9}$ around y_0 gets mapped into itself by N and that N is a contraction there.

Remark 2.11. Notice that to obtain the estimate in (23) we used (21) essentially, and that equation depends heavily on the specific form of the derivative.

Since N has a fixed point, the Poincaré mapping has a zero by (22), so we have proved:

Theorem 2.12. *For ϵ sufficiently small, $\mu \in [0, 1]$, and E such that $|E - E_0| \leq \mathcal{O}(\epsilon^2)$, then there exists a smooth function $y^* = y^*(\mu, E)$, such that $P(y^*(\mu, E), E, \mu) = 0$. In particular, for $\mu = 1$, this gives us a periodic orbit of the hamiltonian H^R in* (9).

Proof. Now, we finish the proof of Theorem 2.1. One has many periodic solutions for every value of μ corresponding to different values of the energy, E. To be definite we will choose the solution with $E = E_0$. Note that with this choice, it is easy to see from the implicit function theorem, and the smooth dependence of solutions of differential equations on parameters in the vector field that the period, $T(\mu)$, of the periodic orbit of X_μ satisfies $|T(\mu) - T| \leq \mathcal{O}(\epsilon^{3/2})$. If we again denote the eigenvalues of $T^2 f^R_{\phi\phi}(0, \phi^*)h_{II}$ by $\kappa_1, \ldots, \kappa_N$, we find (not surprisingly) that $\tilde{\lambda}_j$, the eigenvalues of the Poincaré mapping evaluated on the periodic orbit constructed in Theorem2.12, are given by

$$\tilde{\lambda}_j = 1 + \sqrt{\epsilon\kappa_j + \mathcal{O}(\epsilon^{1+\alpha})}$$
$$\tilde{\lambda}_{j+N} = 1 - \sqrt{\epsilon\kappa_j + \mathcal{O}(\epsilon^{1+\alpha})} \quad j = 1, \ldots, N. \tag{24}$$

(This calculation proceeds just as before, with the additional fact that $|T(\mu) - T| \leq \mathcal{O}(\epsilon^{3/2})$. We also recall that α is some positive constant.) We see that, when $\kappa_j\epsilon$ is positive, then, $\tilde{\lambda}_j, \tilde{\lambda}_{N+j}$ are stable and unstable eigenvalues respectively.

We claim that, under the assumption that all the negative κ_j's are different, the corresponding $\tilde{\lambda}_j, \tilde{\lambda}_{N+j}$ have modulus one for ϵ positive. In effect, for a symplectic transformation, if $\tilde{\lambda}$ is an eigenvalue, so are $1/\tilde{\lambda}$, $\tilde{\lambda}^*$. If $\tilde{\lambda}^* \neq 1/\tilde{\lambda}$, we would have two eigenvalues of the linearization which agree up to order ϵ^2, which contradicts the assumption that all the positive eigenvalues of $h_{II}(I^*)\Psi^R_{\phi\phi}(J(\epsilon), \phi(J(\epsilon))))$ are different for ϵ small enough. The same argument can be used to show that, if all positive κ_j are different, the corresponding $\tilde{\lambda}_j, \tilde{\lambda}_{N+j}$ have modulus one for ϵ negative and small.

If some of the κ_j have a non-zero imaginary part, independently of the sign of ϵ, neither of the two determinations of $\sqrt{\epsilon\kappa_j}$ is real or purely imaginary. Then, one of $\tilde{\lambda}_j, \tilde{\lambda}_{N+j}$ will be a stable eigenvalue and the other will be unstable.

This ends the proof of Theorem 2.1. □

Remark 2.13. The problem of computing an expansion in ϵ of a periodic orbit of (1) and its Lyapounov exponents is considered in [Poi99] chapters 74, 79. There, Poincaré shows it is possible to match powers of $\epsilon^{1/2}$ to get a consistent formal expansion. The results of this section show that indeed the orbits can be continued – something that a formal expansion does not establish.

Moreover, our methods can be used to show that the expansions of Poincaré are asymptotic. The idea is that if we substitute the orbit predicted by the asymptotic expansion into the equation (22), we can compute that the remainder is of the order predicted by the next term in the asymptotic expansion while the contraction is uniform. Thus, there is a true orbit whose distance from the one predicted by the asymptotic expansion is of the order of the remainder in the expansion.

3. Analysis on the Center Manifold

3.1. **Reduction to the center manifold.**

We recall the center manifold theorem.

Theorem 3.1. *Let f be a C^r, $r = 2, 3, \ldots$ mapping of $\mathbb{R}^N$ leaving the origin fixed. Assume that $Df(0)$ has eigenvalues of modulus 1 and let Π^c be the invariant subspace associated to these eigenvalues. There exists a C^r*

manifold W^c, tangent to Π^c at the origin, and invariant under f. Moreover if f_μ is a C^l family of C^r mappings, $Df_\mu(0) = Df_0(0)$, then the center manifold for each f_μ can be chosen to be a C^l family of C^r manifolds.

Remark 3.2. We recall that, in distinction to the stable and unstable manifolds, the center manifolds are not unique because there are mappings with infinitely many invariant center manifolds.

The C^∞ analogue of Theorem 3.1 is also false in a very strong sense. There are C^∞ mappings which have infinitely many finitely differentiable invariant manifolds, but none of them is C^∞. In certain cases, it is possible to find the Taylor expansion of the center manifold by matching powers. This shows that, for a class of mappings analytic center manifolds are unique. The same argument performed in a quantitative way can be used to show that there exist analytic mappings without analytic center manifolds. One such example, due to Lanford, can be found in [MM76] p. 44. In some cases [Pös86] contains a proof of existence of analytic center manifolds. The results of [Pös86] however, do not apply to real valued, let alone symplectic, mappings since they include non-resonance assumptions that are never verified by real valued mappings.

The existence of smooth families is important for us since it will allow smooth perturbations. Of course, given the non-uniqueness of the center manifolds for each value of the parameters, one could well choose a discontinuous family.

Remark 3.3. In this section, we will produce invariant objects in any of the invariant manifold. Since these invariant objects are locally unique, it will follow that these whiskered tori lie in the intersection of all the possible center manifolds

Proof. There are many standard proofs of the center manifold theorem.

One that produces dependence on parameters and that is well suited for our purposes is to deduce the the center manifold theorem from the "pseudostable manifold theorem"

A proof of the pseudostable manifold including dependence on parameters and optimal regularity is obtained in [dlLW95]. Indeed, a preliminary version of [dlLW95] was an appendix in the original version of this paper.

Theorem 3.4. *Let f be a C^r, $r = 2, 3, \dots$ mapping of $\mathbb{R}^N$ leaving the origin fixed and with $||f - Df(0)||_{C^r}$ sufficiently small. Let α be a real number bigger than 1. Assume that $Df(0)$ has no eigenvalues of modulus in the interval $[\alpha, \alpha^k]$, $\quad k < r$ and let Π^α be the invariant subspace associated to the eigenvalues of modulus less than α. Then there exists a C^k manifold W^α, tangent to Π^α at the origin, and invariant under f. This manifold is unique. Moreover if f_μ is a C^l family of C^r mappings, $Df_\mu(0) = Df_0(0)$, then the center manifold for each f_μ can be chosen to be a C^l family of C^k manifolds.*

Using Theorem 3.4, we can prove Theorem 3.1 as follows: First, given f, consider $\tilde{f}(x) = \Phi(x)Df(0)x + (1 - \Phi(x))f(x)$, where Φ is C^∞ function taking the value 1 in a neighborhood of the origin and the value zero outside of another one. By choosing Φ appropriately, we can arrange that the hypothesis about smallness of $f - Df(0)$ in Theorem 3.4 is satisfied. If f_μ is a differentiable family of differentiable functions, one can choose a Φ that works for all of them.

We now take α very close to 1 in such a way that α^r is smaller than the modulus of all the eigenvalues of $Df(0)$ of modulus bigger than 1 and applying Theorem 3.4 we can conclude that there is a W^α manifold for $\tilde{f}$. Now, we can repeat the analysis for $\tilde{f}^{-1}$ restricted to W^α. Again, we observe that if we have a family, then $\tilde{f}_\mu^{-1}$ restricted to W_μ^α depends smoothly on parameters after we take coordinates. The resulting manifold is the center manifold claimed in Theorem 3.1.

The possibility of having several center manifolds comes from the fact that different choices of cutoff functions could lead to different α–manifolds. □

It is clear that the hypotheses of Theorem 3.1 are verified for the return mapping of the time T map of the orbit constructed in Theorem 2.1. So, we will construct orbits in the center manifold using the theory of motion near an elliptic fixed point.

Lemma 3.5. *If the mapping f is symplectic, and satisfies the hypotheses of Theorem 3.1 we can find a symplectic structure in W^c which is preserved by the restriction of f to W^c.*

Proof. Denote by i the immersion of W^c into $\mathbb{R}^N$ and by ω the symplectic form in R^N. Since $i^*f^* = f^*i^*$ by the invariance of the manifold, $i^*\omega$ is invariant under f and $d(i^*\omega) = i^*(d\omega) = 0$. We also observe that $i^*\omega(0)$ is non–degenerate because it agrees there with $\omega|_{\Pi^c}$. By the implicit function theorem, it is non-degenerate in a neighborhood of this point.

□

Lemma3.5 applies to the mapping $\mathcal{T}$ that we used in the previous section because it is the restriction of a symplectic mapping to a symplectic manifold. Denote the restriction of $\mathcal{T}$ to the center manifold by $\mathcal{T}^c$.

Remark 3.6. The importance of this lemma is that there are many methods to analyze the behavior of a symplectic mapping in a neighborhood of an elliptic periodic point. For example, one can produce periodic orbits in any neighborhood of the origin [BL33, Wei73, Mos76, Mos77]. There are also variational methods to produce quasiperiodic orbits when the manifold is of dimension two. (These were originally worked out in [Mat82] for mappings of the cylinder, but there is an adaptation to neighborhoods of elliptic fixed points in [Che85], which also contains a very nice exposition of many related results.) Other related methods to produce quasiperiodic minimal orbits in a neighborhood of elliptic points in higher dimensions can be found in [BK87, Kat92]. (The adaptation requires non–resonance assumptions on the eigenvalues along the neutral direction.) In two dimensions, it is known that, generically, there are many homoclinic points in any neighborhood of an elliptic point [Zeh73]. Some other properties that are true for generic systems are discussed in [New77, Tak70], The task of adapting these theorems to the situation described in Theorem 2.1 is rather straightforward and we leave it to the reader.

Remark 3.7. Since there are many results about generic properties of mappings near elliptic fixed points, it is worthwhile pointing out that there is a machinery to transform statements of genericity for germs of diffeomorphisms in a neighborhood of a fixed point into genericity statements for Hamiltonian vector fields. As we mentioned, the mapping that to the pair (h, f) associates the time T map is differentiable. (Recall that h and f are respectively the integrable piece and perturbation in our hamiltonian function.) Furthermore the map that associates to a mapping an invariant center manifold can be chosen to be differentiable. Therefore, the map that sends a hamiltonian to the germ of the mapping in the center manifold can be chosen to be differentiable. It can also be shown that this mapping is onto ([Mos86, Dou82]). A simple application of point set topology will show that residual sets in the space of germs, correspond to residual sets in an open set of hamiltonians and perturbations. This construction will be useful in following sections to conclude that certain nondegeneracy conditions which we will need hold in open and dense sets of flows. It will suffice to show that they hold in open and dense sets of maps.

As we showed in section 2, the dimension of the center manifold of the fixed point of the Poincaré mapping depends on the number of positive, negative, and complex eigenvalues, as well as whether or not ϵ is positive or negative. In order not to have to keep track of these various possibility, we will throughout this section use $2M$ to denote the dimension of the center-manifold, independently of which case we are in.

When restricted to the center manifold of the periodic orbit, we can take a return map to a symplectic manifold. The periodic orbit corresponds to an elliptic fixed point.

Applying the standard KAM theorem on existence of quasi-periodic orbits around a fixed point, we obtain:

Theorem 3.8. *Denote by $2M$ the dimension of the center manifold.*
Assume that

(i) *The negative eigenvalues $\kappa_1 \dots \kappa_{N-(k+1)}$ of the matrix considered in Theorem 2.1 satisfy*

$$\sum_{i=1}^{N-(k+1)} \kappa_i \nu_i \neq 0 \quad \text{whenever} \quad 0 < \sum_{i=1}^{N-(k+1)} |\nu_i| \leq \ell. \tag{25}$$

We assume that $\ell \geq 4$.

(ii) *The Birkhoff Normal form up to order 4 has full rank derivatives.*

(iii) h, f *are differentiable enough, depending just on* M.

Then, for ϵ *small enough and positive, there are invariant tori arbitrarily close to the orbit constructed in Theorem 2.1. In each of the tori, the motion is smoothly conjugate to an irrational rotation. The invariant tori in a neighborhood of size* ρ *in the center manifold cover the whole set except a measure* $O(\rho^{\frac{1}{2}(\ell-3)}\rho^{2M})$

Condition (ii) above just amounts to some algebraic expression of the derivatives up to order 4 being different from zero. Hence, it holds for very generic Hamiltonians.

Remark 3.9. An analogous result holds for ϵ small and negative, provided that the positive eigenvalues of the matrix considered in Theorem 2.1 satisfy relations analogous to (25).

Remark 3.10. To make more precise the conditions on the derivatives of the Hamiltonian required in hypothesis (ii) of Theorem 3.8, note that because of hypothesis (i) we can construct the Birkhoff normal form for the mapping in a neighborhood of the origin and it will have the form, in symplectic polar coordinates:

$$\begin{aligned} r'_j &= r_j + \mathcal{O}(r^2) \\ \phi'_j &= \phi_j + (\sum_{\ell=1}^{N-(k+1)} \mathcal{B}_{j\ell} r_\ell) + \mathcal{O}(r^2)\ . \end{aligned}$$

The "twist matrix", $\mathcal{B}$, is given (via the normal form construction) by a complicated but explicit algebraic expressions in terms of the derivatives of f and h of order 4 or less. The requirement in hypothesis (ii) of the theorem is simply that this matrix be non-singular. Note that this will clearly be the case for an open and dense set of Hamiltonians. Furthermore, given a particular Hamiltonian it is possible, at least in principle, to compute $\mathcal{B}$ and determine whether or not (ii) is satisfied.

The result of Theorem 3.8 follows from standard arguments although we are unaware of anyplace in the literature where it is written out in detail. The paper [Pös82], however, does contain a complete proof of the analogue of this result for perturbations of flows. The proof of [Pös82] can be adapted without difficulty to the case where one has a periodic perturbation of the flow (and hence to the case of mappings). The fact that the theorem for flows can be adapted for periodic perturbations in a rather automatic fashion is discussed in [Zeh76].

4. Hyperbolic Invariant Manifolds

In this section we prove the existence of contracting and expanding manifolds – the whiskers – for the invariant tori constructed in the previous sections. Actually, we will show the existence of stable and unstable manifolds for all invariant sets in a neighborhood of the periodic orbits, not just the invariant tori. In the case of invariant tori, we will show that the whiskers depend differentiably on the point in the torus.

The invariant manifolds we construct are going to be perturbations of those of the periodic orbit. The techniques we use will work for general dynamical systems and we will not make any use of the hamiltonian structure in this section–except for the final theorem where we show that some manifolds are Lagrangian.

The best known invariant manifold theorems are those for the so–called "normally hyperbolic" systems of [HPS77]. This theory does not apply to our situation because the tori we constructed in Theorem 3.8, have infinitesimal displacement vectors which do not grow under forward or backward iteration and are not tangent to the manifold. Nevertheless, we will be able to use the theory of "ρ-hyperbolic" sets which is developed in [HPS77] chapter 5. Similar results are contained in [Fen72, Fen74]. The main difference is that the later references are concerned with invariant sets which are manifolds, but make less restrictive assumptions in the uniformity of the hyperbolic behavior.

The proof is divided into three parts. First, we make a perturbation theory of the infinitesimal situation: using the hyperbolicity of the periodic orbit, we conclude hyperbolicity of the invariant torus. Second–following the references above–we transform the infinitesimal results into results that hold under smallness assumptions: we can construct an invariant manifold for each orbit. Then, those assumptions will have to be checked in the cases that we are interested in. The results we obtain so far in this section are quite general and apply to any invariant sets e.g. Aubry–Mather sets, invariant circles with Liouville rotation number and any other invariant set in the center manifold as discussed in the remarks after Theorem 3.8. Thirdly, we will show that, for invariant tori with the motion on them conjugate to a rotation we can construct invariant manifolds for the whole torus: in other words, the stable manifolds we constructed for each of the orbits fit together in a smooth manner. A very similar theory has been developed in [Zeh76] Section 7. The main difference is that our theory is based on real variable methods and optimized for finitely differentiable systems while the theory of [Zeh76] is based on analyticity considerations and uses complex extensions. We will perform a more detailed comparison in the remarks at the end of the section.

Notation

In order to manipulate d–dimensional subbundles of the tangent bundle of an n–dimensional manifold it is quite convenient to associate to each of them a $(n-d)$–form in such a way that the kernel of this form is precisely the subbundle. The advantage of working with forms is that the space of all forms has much more structure – in particular it is a linear space and we can use techniques based on partitions of unity.

If S is a d–dimensional subbundle of TM and M is a Riemannian manifold of dimension n, we can associate to S an $(n-d)$–form Ψ_S as follows:

Choose a coordinate patch in which S is trivial and in this coordinate patch choose $v_1, v_2, \ldots, v_{n-d}$ orthonormal vectors which are orthogonal to S. Then, set $\Psi_S = v_1 \wedge v_2 \wedge \ldots \wedge v_{n-d}$. This form is uniquely determined by S–up to a sign– in the coordinate patch since it is the only $(n-d)$ form of norm one that vanishes when contracted with vectors in S. Conversely, S is characterized as the span of the vectors which, when contracted with Ψ_S, give 0. Therefore, Ψ_S is a geometric object even if we used coordinates to define it. By elementary perturbation theory for matrices, if two C^k forms are sufficiently C^k–close, we can apply the range theorem and conclude that the associated subbundles are also C^k–close.

By the uniqueness in each coordinate patch, this form has to agree in the overlap of two different patches–up to a sign. For a fixed set of coordinate patches, we can– if necessary by passing to a double cover, obtain a globally defined form representing S. Passing to a double cover will not affect any of the following considerations, so we will deal with subbundles by studying their associated forms.

We will say that a subbundle S is C^k if its associated form is, and we will refer to the C^k norm of the Ψ_S as the C^k norm of S. Notice that a C^k norm can only be defined when we fix a set of coordinate patches. It is clear that all such norms will be equivalent. Even if the norms of the operators we consider are affected and we need some of them to be contractions, the arguments do not require modification, because they would remain contractions if we impose slightly stronger smallness conditions.

Since two forms differing in a scalar multiple have the same null space, it is convenient for us to take the convention of considering only forms normalized to have norm 1. It is possible to define a push forward in the space of normalized forms by just using the usual push forward for the form and, then, dividing by the norm. This operator is not linear, since the space of normalized forms is not linear.

Lemma 4.1. *Let Ω be a compact set and let S be a continuous subbundle of $T\Omega$. Let $\mathcal{V}$ be a sufficiently small neighborhood of Ω. (How small depends only on Ω). Then, there is a continuous subbundle S^* of $T\mathcal{V}$ which extends S. Moreover, if S is C^k, $k = 1, \ldots, \infty$ in the sense of Whitney, this extension can be taken to be C^k, with a C^k norm as close as we wish to that of S.*

Proof. Let Ψ_S be a form on Ω defining S. Let $V_1, \ldots, V_N$ be coordinate patches covering Ω and on which S is trivial.

For a fixed i we denote by $v^i_1, \ldots, v^i_n$ vector fields that, at each point x in V_i span T_x – the tangent space at x.

The functions $\phi_{j_1,\ldots,j_{n-d}}(x) = \Psi_S(x)(v^i_{j_1}(x), \ldots, v^i_{j_{n-d}}(x))$ are defined in $\Omega \cap V_i$. By the Borel–Whitney extension theorem (see e.g the appendix in

[AR67]), they can be extended to functions $\tilde{\phi}_i$ defined on V_i and, hence, we can define a form Ψ^i_S on V_i by setting $\Psi^i_S(x)(v^i_{j_1}(x), \dots, v^i_{j_{n-d}}(x)) = \phi^i_{j_1,\dots,j_{n-d}}(x)$ and using the linearity to define the result on other vector fields.

If $\Theta_1, \dots, \Theta_N$ are C^∞ functions so that $(\Theta_1)^{n-d+1}, \dots, (\Theta_N)^{n-d+1}$ constitute a partition of unity subordinate to $V_1, \dots, V_N$, then the $(n-d)$–form $\widetilde{\Psi}_S$ defined by :

$$\widetilde{\Psi}_S(v_1, \dots, v_{n-d}) = \sum_i \Theta_i \Psi^i_S(\Theta_i v_1, \dots, \Theta_i v_{n-d})$$

will agree with Ψ_S on Ω and be as smooth on V_i as Ψ_S is on Ω

By reading carefully the proof of the Borel–Whitney theorem, it is also possible to check that the C^k norm of the functions $\tilde{\phi}_i$ can be made as close to the C^k norm of the restriction, ϕ_i as desired. This translates immediately into the possibility of making the C^k norm of the extended subbundle S^* as close as desired to the C^k norm of the subbundle S. □

Lemma 4.2. *Let Ω be a compact invariant set for the flow Φ_t generated by the C^2 vector field X. Assume moreover that there is a splitting invariant under Φ_{t*}:*

$$T\Omega = S \oplus U \oplus N$$

such that:

$$\begin{aligned} ||D\Phi_t(x)|_S|| &\le Ce^{-\lambda t} \ , t \ge 0 \ , x \in \Omega \\ ||D\Phi_t(x)|_U|| &\le Ce^{\lambda t} \ , t \le 0 \ , x \in \Omega \\ ||D\Phi_t(x)|_N|| &\le Ce^{\mu|t|} \ , t \in \mathbb{R} \ , x \in \Omega \end{aligned}$$

for some $C > 0$, $\lambda > \mu > 0$. Let Ω' be another invariant set for X with the property that, for every point $x' \in \Omega'$ we can find another point $x \in \Omega$ with $d(x, x') \le \epsilon$ for some sufficiently small $\epsilon > 0$. Then, we can find another continuous splitting :

$$T\Omega' = S' \oplus U' \oplus N' \tag{26}$$

such that:

$$\begin{aligned} ||D\Phi_t(x)|_{S'}|| &\le C'e^{-\lambda' t} \ , t \ge 0 \ , x \in \Omega' \\ ||D\Phi_t(x)|_{U'}|| &\le C'e^{\lambda' t} \ , t \le 0 \ , x \in \Omega' \\ ||D\Phi_t(x)|_{N'}|| &\le C'e^{\mu'|t|} \ , t \in \mathbb{R} \ , x \in \Omega' \end{aligned} \tag{27}$$

for some $C' > 0$, $\lambda' > \mu' \ge 0$.

Proof.

Let $\mathcal{V}$ be an open neighborhood of Ω.

Fix T in such a way that $Ce^{-\lambda T} Ce^{\mu T}$ is sufficiently small.

By Lemma 4.1, we can extend the subbundles S, U, N to $\mathcal{V}$, having, thus a new splitting of $T\mathcal{V}$. The new splitting will be invariant under Φ_t only on Ω, but we will, nevertheless have :

$$|| \left(\Psi_S - \Phi_{T*}\Psi_S \right) |_{\Omega'} ||_{C^0} \leq \epsilon'$$

where Ψ_S is the normalized form defining S on Ω and ϵ' can be made as small as we wish by making ϵ small enough.

The reason is that on Ω we have that

$$\Psi_S - \Phi_{T*}\Psi_S = 0$$

and, under the assumptions of the lemma, the left hand side is differentiable.

If Φ_T is C^2, then the push–forward on normalized $(n-d)$–forms equipped with the supremum norm is C^1. Given the inequalities that define T and the invariance of S, we see that, on Ω, Φ_T is a contraction. By requiring that ϵ is small enough – depending only on Ω and Φ_t, we can ensure that Φ_{T*} is a contraction acting on the space of $(n-d)$ forms on $T\Omega'$. Notice that the same argument establishes that Φ_t is a contraction for all t big enough. We can assume that T is such that Φ_t is a contraction for all $t \geq T$.

Since Ω' is invariant, we can iterate Φ_T and $\tilde{\Psi}^S_T = \lim_{n\to\infty} \Phi_{nT*}\Psi_S$ will be the unique fixed point. Moreover, we will have $||(\Psi_S - \tilde{\Psi}^S_T)|_{\Omega'}||_{C^0} \leq 2\epsilon'$.

We want to show that $\tilde{\Psi}^S_T$ – which in principle is invariant only under Φ_T is actually invariant under all Φ_t, $t \in \mathbb{R}$. This will complete the proof of the theorem with respect to S, since the inequalities claimed in Lemma 4.2 are obvious from the ones we have studied so far.

To show that Ψ^S_T is invariant under Φ_t for all $t \in \mathbb{R}$ it suffices to show that this $\tilde{\Psi}^S_T$ is invariant under Φ_t, for all $t \in [T, T']$, $T' > T$ because, then, we can use the group property of the evolution $(\Phi_{t+t'} = \Phi_t \circ \Phi_{t'})$ to conclude that $\tilde{\Psi}^S_T$ is invariant under all Φ_t. If t_0 is slightly bigger than T, the same argument we have used to produce the fixed point applies. Φ_{t_0*} is a contraction and has a fixed point $\tilde{\Psi}^S_{t_0}$. If $T = (n/m)t_0$, $n, m \in \mathbb{Z}$, we have that that $\tilde{\Psi}^S_T$ and $\tilde{\Psi}^S_{t_0}$ are both fixed points of $\Phi_{mT*} = \Phi_{nt_0*}$, which is a contraction, so that both agree. Moreover, Φ_t is continuous in t. The fixed points Ψ_t depend continuously on t, but we have shown that they are constant on the rationals. This finishes the proof of the claim for the subbundle S.

The same proof applies for the subbundle U just changing the direction of time.

The proof for N is obtained along the same lines. We just have to observe that the contraction in the space of normalized forms only depends on the " gap " between the rates of growth for vectors in the different splittings. We could then consider repeating the same proof for the splitting $TV = S \oplus \tilde{U}$; $\tilde{U} = U \oplus N$ and $TV = \tilde{S} \oplus U$; $\tilde{S} = S \oplus N$. Then, we can set $N = \tilde{U} \cap \tilde{S}$. $\square$

Lemma 4.3. *If S,U,N are continuous bundles satisfying* (26), (27), *and* $\beta > 0$ *is such that* $\sup_{x\in M} ||D\Phi_t(x)|| < Ce^{\beta t} \quad \forall t \in R$ *then* S, U, $S \oplus N$, $U \oplus N$, N *are* C^{α_1}, $\alpha_1 = (\lambda - \mu)/\beta$.

Proof.
This theorem can be reduced to Theorem 6.1 of [HP70]. Another very similar proof can be found in [Fen77] Theorem 2, which is stated only for invariant sets that are smooth manifolds but the proof carries over to compact invariant sets.

We recall briefly the argument omitting some details on the precise definition of Hölder norms.

In a C^0 neighborhood of Ψ_S, Φ_t will be a contraction by a factor $Ce^{-(\lambda-\mu)t}$.

Working in a set of coordinates, we can define a Hölder norm and obtain that for Ψ in this neighborhood,

$$||\Phi_{t*}\Psi||_{C^\alpha} \equiv \sup_{x,y} ||(\Phi_{t*}\Psi)(x) - (\Phi_{t*}\Psi)(y)||/d(x,y)^\alpha <$$
$$Ce^{(\lambda-\mu)t} \sup_{x,y} ||\Psi(x) - \Psi(y)||/d(x,y)^\alpha \sup_{x,y} \left(\frac{d(x,y)}{d(\Phi_t(x), \Phi_t(y))} \right)^\alpha .$$

□

Lemma 4.4. *Let* Ω *be a closed invariant set under the* C^r *flow* Φ_t. *Assume that, we can decompose* $T\Omega$ *into bundles S,U,N satisfying the hypothesis of* (27). *Then, for each point in* Ω, *we can find* C^r *manifolds* W_x^S, W_x^U *characterized by:*

$$\begin{aligned} &y \in W_x^S \quad (\text{resp.} W_x^U) \iff \\ &d(\Phi_t(y), \Phi_t(x)) < Ce^{-\lambda t} d(x,y), \quad t \geq 0 \text{ (resp. } t \leq 0). \end{aligned} \tag{28}$$

Moreover, the r–jets of these manifolds depend in a Hölder fashion on the point. The tangent space of W_x^S *at* x *is* S_x *(analogously for* U*). The set of these manifolds extends to a foliation in a neighborhood of* Ω.

Remark 4.5. Notice that if Ω is a compact invariant set contained in a neighborhood V, then, $W_\Omega^U \cap V$ will also be an invariant set. If we have a set Ω of small enough diameter all of whose points have a splitting then using the characterization in (28):

$$y \in W_x^S \iff x \in W_y^S$$

so that the set of W_x^S extends to a foliation.

Proof. Except for the claim of Hölder continuity of the jets, this is a particular case of theorem (5.5) of [HPS77]. It is also very similar to Theorem 1 of [Fen77] except that [Fe3] does not assume that the hyperbolic behavior is uniform but does assume that the invariant sets are smooth manifolds.

With small modifications, the methods of [HPS77, Fen77] would suffice to prove Lemma 4.4. We do not present the details here, but rather, we will prove a stronger result for the case of most interest to us, namely when the invariant set is a torus and the flow on it is a linear flow with an irrational frequency vector. For this particular case, we will prove that the jets of W_x^s, W_x^u depend smoothly on the point x on the torus. □

Notation. We will sometimes refer to W_x^S as the the stable manifold of x. This is an slight abuse of notation since there could be points not in W_x^S whose orbit is asymptotic to the orbit of x. Nevertheless, they have to approach the orbit of x much slower that those in W_x^S as the characterization of W_x^S shows. We will prefer to use the name whiskers.

Remark 4.6. It is possible to use the invariant bundles $S \oplus N$, $U \oplus N$ and N to produce other invariant structures which are perturbations of them. See Theorem 5.5 of [HPS77]. Nevertheless, the regularity theory of these objects is more complicated than that of the objects obtained by perturbing S or U. They can also fail to be foliations.

Putting together the results we have proved so far, we can obtain easily the main result of this section:

Theorem 4.7. *Let H_ϵ be a hamiltonian as in* (1). *Let O be a periodic orbit as in Theorem 2.1 existing for $\varepsilon \neq 0$. Let $\mathcal{N}$ a sufficiently small neighborhood of O. Let Ω be an invariant set in $W_O^c \cap \mathcal{N}$. Then, for each point $x \in \Omega$ there are two smooth k–dimensional manifolds W_x^S and W_x^U which are characterized by* (28).

Remark 4.8. We could take Ω in the above theorem to be one of the invariant tori produced in Theorem 3.8 but we could also take Ω to be any invariant set contained in the neighborhood of the origin in the center manifold e.g. Aubry–Mather sets, invariant tori with Liouville rotations, periodic orbits, horseshoes, homoclinic tangles, or the union of several of those.

Remark 4.9. The argument so far does not allow us to conclude that the invariant whisker for the whole torus Ω , defined by $W_\Omega^S \equiv \cup_{x\in\Omega} W_x^S$, is even a C^1 manifold. It could, in principle, happen that even if the stable manifolds for each point in the torus are C^r they fit together in a fashion which is only Hölder. This situation, could happen for a general system and there are examples of this situation in [Fen74] section I.H. Nevertheless, for the particular case of invariant tori, whose motion is differentiably conjugate to a rotation such as those in Theorem 3.8, we will show that the union of all the whiskers of points in Ω forms a smooth manifold. Notice that we do not require that the rotation on the torus satisfies Diophantine properties.

What we will show is that the invariant bundle S is smooth when restricted to Ω if the motion in Ω is smoothly equivalent to a rotation and then prove a theorem that says that the foliation is almost as smooth as the bundle. Although the later theorem is only proved in the context we need it – this avoids to have to set up a complicated notation – it is a general theorem and applies to dynamical systems leaving invariant a smooth manifold.

An analytic version of these results can be found in [Zeh76] Section 7. We will compare the two approaches in some remarks at the end of the section.

Lemma 4.10. *Let Ω be a C^r invariant torus as in Theorem 3.8 sufficiently close to the periodic orbit and such that the motion on it is conjugate by a C^r change of variables to a linear rotation of angular frequency ω (not necessarily Diophantine) and let $S \oplus N \oplus U$ be a decomposition of the tangent bundle to the ambient space as in* (26)*. Then, S is C^r.*

Proof. We can perform a C^r change of coordinates (we do not require that it is symplectic) in a neighborhood of the torus in such a way that the equations of motion in the torus are motion at constant speed with frequency ω.

It is well known that if $x(t)$ is a solution of $\dot{x} = X(x)$, the flow Φ_s will act on tangent vectors based at $x(t)$ by multiplying them by a matrix $M_t(s)$ obtained by solving the so called "*equations of variation*"

$$\begin{gathered} \dot{M}_t(s) = DX(x(s))M_t(s) \\ M_t(t) = Id \end{gathered}$$

(See e.g. [Hal80] p.95.)

In our case, the equation of variation becomes

$$\dot{M} \equiv \sum_{i=1}^{k} \omega_i \frac{\partial M(x)}{\partial x_i} = \Gamma(x) M(x) \tag{29}$$

and the bundle S can be identified with the vectors that decrease exponentially fast.

Notice that by assuming that the torus is sufficiently close to the periodic orbit, we can assume that $\Gamma(x)$ is close to a constant matrix $\bar{\Gamma}$. Furthermore, by choosing an appropriate system of coordinates, we can assume that the matrix $\bar{\Gamma}$ is already in normal form. We will adopt the convention that the first k coordinates in our space are close to the contracting eigenspace.

We claim that there exists a C^r matrix valued function Υ on the torus such that it is C^r close to the identity and

$$\dot{\Upsilon}\Upsilon^{-1} + \Upsilon\Gamma\Upsilon^{-1} = \begin{pmatrix} \tilde{\Gamma}_S & 0 \\ 0 & \tilde{\Gamma}_{U+N} \end{pmatrix} \tag{30}$$

with $\tilde{\Gamma}_S$ close to the contractive part and $\tilde{\Gamma}_{U+N}$ is close to the expansive part of Γ.

If we set $\tilde{M} = \Upsilon M$, (29) becomes :

$$\dot{\tilde{M}} = \left(\dot{\Upsilon}\Upsilon^{-1} + \Upsilon\Gamma\Upsilon^{-1}\right)\tilde{M}$$

In this new equation, given that Υ satisfies (30), it is quite obvious that the vectors which decrease exponentially fast are those that have the coordinates corresponding to the last block equal to zero. The bundle S will be the range under Υ^{-1} of a constant bundle and hence smooth.

Remark 4.11. We emphasize that the matrix valued functions Υ, $\tilde{\Gamma}_S$, $\tilde{\Gamma}_{S+N}$ solving (30) are not unique. For example, if we multiply any solution Υ by block diagonal matrices like those on the right hand side of (30), the result will still be a solution. Therefore, the smallness and proximity statements in the conclusions of the claim refer only to the specific solution that we will construct.

The claim will be established by constructing an iterative procedure and showing it is a contraction. We will use, roughly, a Newton method that, assuming that the blocks off diagonal are small produces an Υ that kills them off up to quadratic terms.

We start by discussing the *"linear approximation"* to equation (30) and show that it has solutions satisfying suitable bounds. After that, we will describe the iterative procedure that uses it.

The linear approximation will be derived by setting $\Upsilon = \mathrm{Id} + \hat{\Upsilon}$, $\Gamma = \bar{\Gamma} + \hat{\Gamma}$, substituting in (30) and keeping only linear terms.

The linear approximation to the L . H . S . of (30) will be:

$$\dot{\hat{\Upsilon}} - \bar{\Gamma}\Upsilon + \Upsilon\bar{\Gamma} \ . \tag{31}$$

We will determine $\hat{\Upsilon}$ in such a way that the off-diagonal blocks of this expression are zero in the linear approximation. As we will see, this does not determine $\hat{\Upsilon}$ completely. The only ambiguity, though, is related to the fact that the diagonal terms are not determined. This causes no problem because the result we want to establish does not make any claims about the diagonal blocks except that they are C^r and close to constant.

For simplicity we will only discuss the case when $\bar{\Gamma}$ is a perturbation of a diagonal matrix, which is true, for instance, if the real eigenvalues of the Poincaré map we considered in Section 2 are all different. For the purposes of this discussion, we can allow complex valued matrices. The case when $\bar{\Gamma}$ contains Jordan blocks is notationally more complicated but not essentially different. We will remark at the end on which modifications are necessary.

In the case that $\bar{\Gamma}$ is diagonal, we have $\Gamma = \mathrm{Diagonal}(\lambda_1, \lambda_2, \ldots, \lambda_n) + \hat{\Gamma}$ and equating to zero the off-diagonal elements in (31) we obtain:

$$\dot{\hat{\Upsilon}}_{ij} + (\lambda_j - \lambda_i)\hat{\Upsilon}_{ij} = -\hat{\Gamma}_{ij} \tag{32}$$

with $i = 1, \ldots, k$ and $j = k+1, \ldots, n$ or $j = 1, \ldots, k$ and $i = k+1, \ldots, n$.

We will show that (32) admits solutions by showing that the related equation

$$\hat{\Upsilon}_{ij}(x) = \left(\hat{\Upsilon}_{ij}(x - t\omega) - \int_0^t \hat{\Gamma}_{ij}(x + (s-t)\omega)e^{-(\lambda_i - \lambda_j)s}ds\right) e^{(\lambda_i - \lambda_j)t} \tag{33}$$

can be solved for $\hat{\Upsilon}$ given $\hat{\Gamma}$ and that the solutions of (33) are also solutions of (32).

The heuristic motivation for considering (33) is that it is the equation that one obtains by applying the formula for the solution of an inhomogeneous linear equation to (32). Therefore, if $\hat{\Upsilon}$ solves (32), then it solves (33) for any value of t. Conversely, if $\hat{\Upsilon}$ solves (33) for all t, taking derivatives we obtain that it also solves (32).

We will first show that (33) has a solution for some particular t and then will show that the solution thus obtained, solves (32). For the moment, we will take $t = +1$ when $\Re(\lambda_i - \lambda_j) > 0$ or $t = -1$ when $\Re(\lambda_i - \lambda_j) < 0$. Notice that there are no cases with $\Re(\lambda_j - \lambda_i) = 0$ by the assumptions we made on the spectrum of M and the definition of the splitting.

Since the only way that $\hat{\Upsilon}$ enters in the R.H.S of (33) is as a translated version multiplied by a number smaller than one, it is clear that the R.H.S. of (33) considered as a function of $\hat{\Upsilon}$ is a contraction when $\hat{\Upsilon}$ is given the C^r norm. Applying the contraction mapping theorem, we obtain that (33) has a C^r solution. Furthermore, the C^r norm of $\hat{\Upsilon}_{ij}$ is bounded by a constant times the C^r norm of $\hat{\Gamma}$.

We now have to show that this solution of (33) actually solves (32). Since the R.H.S. of (33) depends on t, it could, in principle happen that the solution we have produced for $t \pm 1$, fails to be a solution for all t. Once we know that (32) is satisfied, we conclude that $\hat{\Upsilon}$ solves (33) for all t.

Since $\hat{\Upsilon}$ is C^r, we can apply to it the operator $D_\omega \equiv \sum_i \omega_i \frac{\partial}{\partial x_i}$. Using (33) for a fixed t, we obtain

(34)
$$D_\omega \hat{\Upsilon}_{ij}(x) = \left(D_\omega \hat{\Upsilon}_{ij}(x - t\omega) + \int_0^t D_\omega \hat{\Gamma}_{ij}(x + (s-t)\omega) e^{-(\lambda_i - \lambda_j)s} ds \right) e^{(\lambda_i - \lambda_j)t}$$

We now observe that $D_\omega \hat{\Gamma}_{ij}(x + (s-t)\omega) = \frac{\partial}{\partial s} \hat{\Gamma}_{ij}(x + (s-t)\omega)$. Substituting this into (34) and integrating by parts, we obtain:

$$\begin{aligned} D_\omega \hat{\Upsilon}_{ij}(x) &= e^{(\lambda_i - \lambda_j)t} D_\omega \hat{\Upsilon}_{ij}(x - t\omega) - \hat{\Gamma}_{ij}(x) + \hat{\Gamma}_{ij}(x - t\omega) e^{(\lambda_i - \lambda_j)t} \\ &\quad - (\lambda_i - \lambda_j) \int_0^t \hat{\Gamma}_{ij}(x + (s-t)\omega) e^{-(\lambda_i - \lambda_j)s} ds e^{(\lambda_i - \lambda_j)t} \end{aligned} \tag{35}$$

Using now the expression for the integral given by (33), we obtain that the function $\mathbb{R}(x) \equiv D_\omega \hat{\Upsilon}_{ij}(x) + (\lambda_i - \lambda_j)\hat{\Upsilon}_{ij}(x) - \hat{\Gamma}_{ij}(x)$ satisfies:

$$\mathbb{R}(x) = e^{(\lambda_i - \lambda_j)t} \mathbb{R}(x - t\omega)$$

From that, we can see, e.g. taking Fourier coefficients and observing that $|e^{(\lambda_i - \lambda_j)t}| \neq 1$ that $\mathbb{R}(x) = 0$. That is, $\hat{\Upsilon}$ satisfies (32).

Notice also that, as we pointed out, the C^r norm of $\hat{\Upsilon}_{ij}$ is bounded by a constant times the norm of $\hat{\Gamma}_{ij}$. Using (32), we obtain that so is the C^r norm of $\dot{\hat{\Upsilon}}$.

This finishes the discussion of the first order equation.

To study the full equation we observe that if we write $\Upsilon = \tilde{\Upsilon}_1 \Upsilon_1$, then (30) becomes:

$$\dot{\Upsilon}_2 \tilde{\Upsilon}_1^{-1} + \tilde{\Upsilon}_1 \left(\Upsilon_1 \Gamma \Upsilon_1^{-1} \dot{\Upsilon}_1 \Upsilon_1 \right) \tilde{\Upsilon}_1^{-1} = \begin{pmatrix} \tilde{\Gamma}_S & 0 \\ 0 & \tilde{\Gamma}_{U+N} \end{pmatrix} \tag{36}$$

Notice that if we choose Υ_1 to be any matrix valued function we wish, we can determine $\tilde{\Upsilon}_1$ by solving an equation of the same form of (30) but with Γ replaced by $\Gamma_1 \equiv \Upsilon_1 \Gamma \Upsilon_1^{-1} \dot{\Upsilon}_1 \Upsilon_1$. If we choose Υ_1 to be the solution of (32), then $\tilde{\Upsilon}_1$ will satisfy an equation of the form (30) in which the off–diagonal terms of Γ_1 have a C^r norm which is of the order of the square of the C^r norm of those of Γ. We can then write $\tilde{\Upsilon}_1 = \tilde{\Upsilon}_2 \Upsilon_2$ where Υ_2 is determined by solving the linearized equation. As before, $\tilde{\Upsilon}_2$ will satisfy an equation of the form (30) but with Γ_1 replaced by Γ_2 whose off diagonal terms have C^r norm of the order of the square of the C^r norm of the off–diagonal terms of Γ_1.

The procedure can be iterated indefinitely and, provided that the original non–diagonal terms were small enough, the off diagonal terms of Γ_n will converge to zero quadratically. Using the bounds for $||\hat{\Upsilon}||_{C^r}$ in terms of the C^r norm of the off–diagonal terms of Γ obtained in the discussion of (32), it is easy to show that $\Upsilon = \lim_{n\to\infty} \Upsilon_n \cdots \Upsilon_2 \Upsilon_1$ converges and solves (30). □

Now we prove that the union of all the S manifolds for all the points in the invariant torus form a smooth manifold. In order to avoid having to introduce a complicated notation, we will prove only a version that applies to our situation.

Theorem 4.12. *Let Ω be a C^r d–dimensional torus contained in $\mathbb{R}^\nu$, and Φ_t be a C^r flow. Assume that the flow Φ_t leaves Ω invariant and that the restriction of Φ_t to Ω is C^r conjugate to a rotation of angular frequency ω. Assume that, in Ω, there is a decomposition $T\mathbb{R}^\nu = S \oplus N \oplus U$ which is invariant under Φ_t and which satisfies* (26), (27) *and that, moreover, S is a C^r bundle. Then, the mapping that to each $x \in \Omega$ associates W_x^S – constructed in Lemma 4.4 – is C^{r-2} when the space of curves is given the the C^1 topology. The set $W_\Omega^S = \cup_{x\in\Omega} W_x^S$ is a C^{r-3} manifold.*

Remark 4.13. The regularity we have claimed in the theorem is not optimal. It will follow from the proof that we can give the space of curves stronger topologies and still conclude that the mapping is differentiable. It is also possible to get better regularity results by extending the techniques we presented here. Since the main concern of this paper is to obtain the geometric results on existence of invariant sets, we have relegated these questions to remarks sketching the proofs or pointing to references in the literature where the extra techniques are used.

Proof. The proof will be done in a specific system of coordinates to simplify the calculations. It will be useful to take a system of coordinates with special properties. We will also adopt some notational conventions that will simplify the exposition of the proof.

We can perform a C^r change of variables in a neighborhood of Ω that allows us to identify this neighborhood with $\mathbb{T}^d \times \mathbb{I}^s \times \mathbb{I}^{n+u}$, where $\mathbb{I} = [-1, 1]$ and s,n,u are the dimensions of S,N,U). We can construct the change of variables in such a way that:

(i) Ω is identified with the set of points of the form $(\phi, 0, 0)$. When there is no danger of confusion we will talk of the point $\phi \in \Omega$.

(ii) Φ_t acting on Ω is given by $\phi \mapsto \phi + \omega t$. We will adopt the convention that ϕ refers to variables in $\mathbb{T}^d$, σ will refer to variables in $\mathbb{I}^s$, τ will refer to variables in $\mathbb{I}^{n+u}$ and γ will refer to variables in $\mathbb{T}^d \times \mathbb{I}^{n+u}$.

(iii) Φ_t can then be written as

$$\Phi_t(\sigma, \gamma) = (A_s\sigma + N_s(\sigma, \gamma), A_\gamma\gamma + (\phi + \omega t, 0, 0) + N_\gamma(\sigma, \gamma)). \tag{37}$$

Where A_s,A_γ are linear operators acting respectively on S and $N+U$ respectively. A_γ may depend on σ but A_s is a constant matrix. We furthermore have that $N_\gamma\big|_\Omega = 0$, $N_s\big|_\Omega = 0$, $DN_\gamma\big|_\Omega = 0$, $DN_s\big|_\Omega = 0$. (iv) By adjusting the choice of scales, we can also assume that $||N_s||_{C^1}$, $||N_\gamma||_{C^1}$ are arbitrarily small.

Once this system of coordinates is chosen, we will construct the stable manifold going though ϕ as the graph of a function w_ϕ. That is, as the set of points of the form:

$$(\sigma, w_\phi(\sigma)) \tag{38}$$

for a suitably chosen function $w_\phi : nterval^s \mapsto \mathbb{T}^d \times \mathbb{T}^{n+u}$.

Proceeding formally, we will derive a functional equation for w_ϕ. Then, we will show that this equation has a solutions and that the graph of this function indeed is invariant.

Applying (37) to a point in the graph of w_ϕ we obtain:

$$\Phi_t(\sigma, w_\phi(\sigma)) = \big(A_s\sigma + N_s(\sigma, w_\phi(\sigma)), \quad A_\gamma w_\phi(\sigma) + (\phi + \omega t, 0, 0) + N_\gamma(\sigma, w_\phi(\sigma))\big).$$

Since we want that $\Phi_t(W^S_\phi) \subset W^S_{\phi+\omega t}$ we should have:

$$w_{\phi+\omega t}\left(A_s\sigma + N_s\left(\sigma, w_\phi(\sigma)\right)\right) \;=\; A_\gamma w_\phi(\sigma) + N_\gamma(\sigma, w_\phi(\sigma))).$$

This can be conveniently written as:

$$w_\phi\left(\sigma, \gamma\right) = A_\gamma^{-1}\left[w_{\phi+\omega t}\left(A_s\sigma + N_s\left(\sigma, w_\phi\left(\sigma\right)\right)\right) - N_\gamma\left(\sigma, w_\phi\left(\sigma\right)\right)\right] \tag{39}$$

Conversely, once we check that the domains make sense, by reading the derivation above backwards, it is possible to show that if (39) is satisfied, the set (38) satisfies the condition of invariance. Moreover, all sets which are graphs which get mapped into each other are graphs of functions satisfying (39). The condition that W^S_ϕ passes though ϕ is equivalent to $w_\phi(0) = 0$.

We will show that there are solutions to (39) by considering it as a fixed point problem. We will consider the right hand side of (39) as the definition of an operator $\mathcal{T}$ acting on w_ϕ and will show that $\mathcal{T}$ is a contraction in a carefully chosen space of functions. A more careful analysis of the operator and the spaces will establish the regularity claimed in the lemma. Another geometric argument – which is also an ingredient in the proof of the characterization of the stable manifolds by rates of convergence of orbits that we gave in Lemma 4.4 will allow us to conclude that the manifolds we constructed are the invariant manifolds of Lemma 4.4 we cannot assume that $||N||_{C^r}$ is small – the possibility of choosing such a system of coordinates requires , in general, non–resonance conditions on the eigenvalues – we will have to pay a lot of attention to the norms we choose in such a way as to be able to use the advantages of the smooth norms.

Notice that $\mathcal{T}$ is – in some sense that will be made precise later – a perturbation of $\tilde{\mathcal{T}}[w_\phi](\sigma) \equiv A_\gamma^{-1} w_{\phi+\omega t}(A_s \sigma)$. This operator is not a contraction in C^0, but it is a contraction in a norm which is the supremum of the derivative, because the derivative picks up an extra contractive factor A_s. Unfortunately, if we took the norm given by the supremum of the derivative, one could only make sense of the smallness of the perturbation by requiring C^2 estimates since w appears among the arguments of w. A possible compromise, between the desire of getting $\tilde{\mathcal{T}}$ to be a contraction and not using many derivatives is the following definition. This definition appears in the context of analytic spaces in [Zeh76] (7.28). The justification we use and some generalizations are discussed in [dlL97].

Definition 4.14.

$$||w_\phi||_\beta \equiv \sup_\phi \sup_\sigma ||w_\phi(\sigma)||/||\sigma|| \tag{40}$$

The following propositions make the previously outlined strategy precise.

Proposition 4.15. *Provided that N's are C^r and small enough in C^1 it is possible to find $\varepsilon_1, \dots, \varepsilon_r > 0$ in such a way that*

$$\begin{aligned} \chi_{\varepsilon_1,\dots,\varepsilon_r} = \{ & w_\phi : \mathbb{T}^d \times \mathbb{I}^s \mapsto \mathbb{T}^d \times \mathbb{I}^{n+s} \text{ s. t.} \\ & \text{a) } C^0 \text{ in } \phi, C^r \text{ in } s \\ & \text{b) } w_\phi(0) = 0 \\ & \text{c) } \sup_{\phi \in \mathbb{T}^d} \sup_{s \in \mathbb{I}^s} ||D^i w_\phi(s)|| < \varepsilon_i \;\; i = 1, \dots r \} \end{aligned}$$

satisfies $\mathcal{T}(\chi_{\varepsilon_1,\dots,\varepsilon_r}) \subset \chi_{\varepsilon_1,\dots,\varepsilon_r}$.

Proof. It is easy to check that if w_ϕ satisfies conditions a) and b), so will $\mathcal{T}(w_\phi)$. If we take ℓ ($1 \le \ell \le r$) derivatives with respect to σ in the expression of $\mathcal{T}$ and apply the chain rule and the rule for derivatives of the product as often as possible, we obtain a sum of terms one of which is

$$A_\gamma^{-1} D^\ell w_{\phi+\omega t} \left(A_s \sigma + N_s \left(\sigma, w_\phi\left(\sigma\right)\right)\right) A_s^{\otimes \ell}.$$

All the other terms in the sum are the product of derivatives of N of order between 1 and ℓ and derivatives of w of order between 1 and ℓ. The important fact to notice is, that every term containing an ϵ_ℓ, must also contain at least one factor of the C^1 norm of N, and hence can be made small. If we now apply the triangle inequality as often as possible, we obtain bounds of the form

$$\sup_\phi \sup_\sigma ||D^\ell w_\phi(\sigma)|| \le ||A_\gamma^{-1}|| \cdot ||A_s||^\ell \varepsilon_\ell + R_\ell(\varepsilon_1, \dots, \varepsilon_\ell), \tag{41}$$

where R_ℓ is a polynomial expression involving the ε's and the supremum of the derivatives of N. For ε in a bounded set, R_ℓ can be made as small as desired by assuming bounds on the norm of the derivatives of N. Since $||A_\gamma^{-1}|| \cdot ||A_s||^\ell < 1$, we can recursively choose $\varepsilon_{\ell+1} = 2R_\ell(\varepsilon_1, \dots, \varepsilon_\ell)/(1 - ||A_\gamma^{-1}|| \cdot ||A_s||^{\ell+1})$ by imposing smallness conditions on N and its derivatives, we can ensure that the R.H.S. of (41) is smaller than $\varepsilon_{\ell+1}$. Moreover, it is clear that, by imposing further smallness conditions in DN, we can obtain that the ε's are small. □

Proposition 4.16. *Assume that the N's that enter in the definition of $\mathcal{T}$ have sufficiently small C^1 norm. Then, $\mathcal{T}$ will be a contraction in $\chi_{\varepsilon_1,\dots,\varepsilon_r}$ when the w_ϕ are topologized with the norm $|| \ ||_\beta$ we introduced above.*

Proof. By the mean value theorem we have:

$$||N_\gamma(\sigma, w_\phi(\sigma)) - N_\gamma\left(\sigma, w'_\phi(\sigma)\right)|| \le ||N||_{C^1} ||w_\phi(\sigma) - w'_\phi(\sigma)||.$$

Dividing by $||\sigma||$ and taking suprema, we obtain that $w_\phi \mapsto N_\gamma(., w_\phi)$ is a contraction in $|| \ ||_\beta$ by a factor that can be made as small as we wish. In effect:

$$\begin{aligned}
&||w_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w_\phi(\sigma))) - w'_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w'_\phi(\sigma)))|| \le \\
&||w_{\phi+\omega t}(A_s\sigma, +N_s(\sigma, w_\phi(\sigma))) - w'_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w_\phi(\sigma)))|| \\
&\quad + ||w'_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w'_\phi(\sigma))) - w'_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w_\phi(\sigma)))||
\end{aligned}$$

The first term of the inequalities of the right hand side can be bounded by

$$||w - w'||_\beta ||A_s\sigma + N_s(\sigma, w_\phi(\sigma))||$$

which, in turn, can be bounded by

$$||w - w'||_\beta(||A_s|| + ||w'||_{C^1}||N_s||_{C^1}||\sigma||).$$

Putting all these estimates together, dividing by $||\sigma||$ and taking sups, we obtain that $\mathcal{T}$ has a Lipschitz constant which is $||A_\gamma^{-1}|| \cdots ||A_s||$ plus terms that can be made arbitrarily small by assuming that $||N||$ is sufficiently small in C^1. □

Proposition 4.17. *If, under the same hypotheses as all the theorems before, the N's are C^r, the mappings w_ϕ solving (39) are $C^{r-1+Lipschitz}$.*

Proof. It suffices to show that the closure of $\chi_{\varepsilon_1,\dots,\varepsilon_r}$ in the $|| \ ||_\beta$ norm is contained in $C^{r-1+Lipschitz}$. This is an easy consequence of Ascoli–Arzelá theorem. A much stronger theorem, valid even in infinite dimensional spaces, is proved as Lemma 2.5 in [LI73]. □

Remark 4.18. By carrying along in the proof estimates on the modulus of continuity of the derivatives of higher order it is possible to obtain that the function is C^r. This would allow us to improve the results claimed in Theorem 4.12 Notice that in the statement of Theorem 4.12 we only claimed C^{r-1} for w_ϕ even if we have proved $C^{r-1+Lipschitz}$.

Our next step is to show that the mapping $\phi \mapsto w_\phi$ is differentiable when the space of mappings w_ϕ is given the topology induced by the $|| \ ||_\beta$ norm.

What we would like to do is to apply the implicit function theorem to $\mathcal{T}$ entering in (39). Unfortunately, this is not possible since the operator that to two functions associates their composite – which is an important ingredient in the construction of $\mathcal{T}$ – is not differentiable in the spaces of norm C^r. The key observation is that the operator $\mathcal{T}$ is differentiable at the fixed points since Proposition 4.17 shows that the fixed points are more differentiable than a straightforward application of a contraction mapping principle would imply. This observation can be exploited by first finding a candidate for a derivative and then, showing that these formal candidates are indeed derivatives. Here we give all the details for the case we are considering. In Appendix A we show how this circle of ideas can be transformed into an abstract implicit function theorem with smooth dependence on parameters.

Proceeding formally, if we take ℓ derivatives with respect to ϕ in (39) we obtain:

(42)
$$\begin{aligned} D^\ell_\phi w_\phi(\sigma,\gamma) = A_\gamma^{-1}[&D^\ell_\phi w_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w_\phi(\sigma))) \\ &+ D_\sigma w_{\phi+\omega t}(A_s\sigma + N_s(\sigma, w_\phi(\sigma)))D_\gamma N_s(\sigma, w_\phi(\sigma))D^\ell_\phi w_\phi(\sigma) \\ &- D_\gamma N_\gamma(\sigma, w_\phi(\sigma))D^\ell_\phi w_\phi(\sigma)] + R_\ell \end{aligned}$$

where R_ℓ is an expression that involves only derivatives of w with respect to ϕ of order strictly less than ℓ.

Proposition 4.19. *It is possible to find recursively functions $D^\ell_\phi w_\phi$, $\ell = 1,\dots,r-1$ that satisfy the equations* (42).

Proof. Proceeding by induction we can assume that all the $D^\ell_\phi w_\phi$ of order less than ℓ have been found, hence R_ℓ is a fixed function. If we call $\mathcal{T}_\ell\left[D^\ell_\phi w_\phi\right]$ all the terms that involve $D^\ell_\phi w_\phi$ in the right hand side of (42) we see that we are considering a fixed point problem of the form

$$D^\ell_\phi w_\phi = \mathcal{T}_\ell\left[D^\ell_\phi w_\phi\right] + R_\ell.$$

Proceeding as in in the proof of Proposition 4.16, we can obtain that $\mathcal{T}_\ell$ is a contraction, hence there is a fixed point. $\square$

Proposition 4.20. *The function w solving* (39) *constructed in Proposition 4.16 is C^{r-1} with respect to ϕ and the derivatives are given by the functions $D^\ell_\phi w_\phi$ constructed in Proposition 4.19 solving* (42)*.*

Proof. Fix Δ a vector in the torus – which we will assume to be small – We consider the function $\tilde{w}$ defined by:

$$\tilde{w}_\phi(\sigma) = \sum_{\ell=0}^{r-1} \frac{1}{\ell!} D^\ell_\phi w_{\phi-\Delta}(\sigma)\Delta^{\otimes \ell} \tag{43}$$

We claim that

$$||\mathcal{T}[\tilde{w}] - \tilde{w}||_\beta \Delta^{-(r-1)} = \mathcal{O}(||\Delta||) \tag{44}$$

Then, by uniqueness of the fixed point solving (39) we get that $||w - \tilde{w}||_\beta \Delta^{-(r-1)}$ tends to zero with Δ. This is well known to imply that w_ϕ is differentiable with respect to ϕ (See e.g. [AR67]).

In order to establish (44) we observe that if we take ℓ derivatives with respect to Δ in $\mathcal{T}[\tilde{w}] - \tilde{w}$ and evaluate at $\Delta = 0$ they vanish for $\ell = 1, \dots, r-1$. We also observe that the expression $\mathcal{T}[\tilde{w}] - \tilde{w}$ is uniformly C^r in Δ for $||\Delta||$ small enough. We, therefore, have bounds $||\mathcal{T}[\tilde{w}] - \tilde{w}|| \le \mathcal{O}(||\Delta||^r)$ where C depends on the C^r norm of the expression we are differentiating and can be chosen uniformly for all Δ small enough. It is also easy to see that $\mathcal{T}[\tilde{w}] - \tilde{w}$ vanishes up to first order in σ since all of the terms do. We therefore have bounds : $||\mathcal{T}[\tilde{w}] - \tilde{w}|| \le C||\Delta||^r||\sigma||$ $\square$

Remark 4.21. We have also shown that the mapping $\phi \mapsto w_\phi$ is differentiable when the w's are given the topology induced by $||\ \ ||_\beta$. The fact that the derivatives are given by the formulas (42) shows that that they are also differentiable at zero.

In order to conclude that W_Ω is a differentiable manifold we need that $w_\phi(s)$ is a differentiable function of (ϕ, s). The previous argument shows that $w_\phi(s)$ is a differentiable function of ϕ when s is kept fixed and a differentiable function of s when ϕ is kept fixed. The fact that having continuous partial derivatives is the same as being C^1 is classic. Nevertheless, for higher derivatives the situation is unpleasantly different. There are functions which are C^2 when restricted to any coordinate line, but fail to be C^2 because they have no mixed partial derivatives. (See [Kra83] for one such example due to Yudovich).

The theory of when functions which are smooth when restricted to coordinate lines are smooth has natural answers in the so–called Λ_α spaces. See [Ste70, Kra83] for some equivalent definitions. We just point out that, for α not an integer, these spaces are the same as the usual C^α spaces.

When α is an integer, the spaces are different from the usual C^α and from $C^{\alpha-1+Lipschitz}$.

The following theorem can be found in [Ste70, Kra83].

Theorem 4.22. *Let f be a continuous function in $[-1,1]^N$ with the property that,for every $j = 1,\dots,N$ the one variable function that we obtain by fixing $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N$ satisfies:*

$$||f(x_1,\dots,x_{j-1},\cdot,x_{j+1},\dots,x_N)||_{\Lambda_\alpha([-1,1])} \leq C \tag{45}$$

where C is a constant independent of $x_1,\dots,x_{j-1},x_{j+1},\dots,x_N$ and of j. Then, $f \in \Lambda_\alpha([-1,1]^N)$. Moreover C_f, the best C that can be used in (45) *defines a norm, which is equivalent to the Λ_α norm.*

Proof. [Kra83] has two very nice proofs. One uses a characterization of the Λ_α spaces by the properties of approximation by complex analytic functions and another one based on real variable methods. The proof of [Ste70] is based on estimates of the Poisson kernel acting on functions in Λ_α. □

This finishes the proof of Theorem 4.12 □

We now proceed to point out different possible extensions of the techniques and to refer to related papers in the literature.

Remark 4.23. It is possible to improve the regularity conclusions in Proposition 4.17 to C^r by using an argument very similar to the one we used in the proof of Proposition 4.20.

Remark 4.24. Notice that in order to get the contraction argument working, we have only used that $||A_\gamma^{-1}|| \cdot ||A_s|| < 1$.

This suggests that we can generalize the argument for all the bundles that include some kind of separation between rates of growth. In particular, the bundles N, $S \oplus N$ and $U \oplus N$.

The argument to conclude that the bundle was smooth on the invariant torus only used the existence of a gap in the spectrum of the action on vectors and that the motion on the torus was a rotation.

We have used that $||A_s|| < 1$ to get the mapping $\mathcal{T}$ defined on a space of functions on an interval. We can eliminate this requirement if we cut–off the nonlinearity as was done in the proof of the pseudo–stable manifold and then use w's that are defined everywhere. As in the case of the pseudo–stable manifold, this can lead to non–uniqueness, since the fixed point could depend on the cut–off used.

Remark 4.25. Theorems of the type Theorem 4.22 have have been generalized to situations in which we assume regularity of restrictions to less well behaved sets than coordinate planes. ([dlLMM86, HK90, Jou86, Jou88]). The method of proof in [Jou86] is very general and it is remarked in [dlL92] that it applies to Cantor sets for which every point in the set is the limit of a sequence of points also in the set converging not faster than exponentially. This happens very frequently for invariant hyperbolic sets. In that case, applying a theory like the one we have developed it is possible to show that if the bundle S is smooth when restricted to a set Ω with this property then W^S_Ω is a subset of a smooth manifold.

Remark 4.26. An alternative theory of invariant manifolds for analytic tori on which the flow is analytically conjugate to a rotation is presented in [Zeh76] Section 7.

The reduction of the existence of the invariant manifold theorem to a fixed point problem is the same as the one presented here. The analytic theory is considerably simpler because $\mathcal{T}$ is differentiable when acting on a space of analytic functions satisfying smallness assumptions. The key observation is that we can choose analyticity domains that get mapped strictly into themselves by $z \to A_s z + N_s(z, w_\phi(z))$ provided that $||w||$ is sufficiently small. In that case, $\mathcal{T}[w](z) = w(A_s z + N_s(z, w_\phi(z)))$ is differentiable and even compact as can be checked using Cauchy estimates.

Since Λ_α functions can be characterized by speed of the approximation by analytic functions, it is possible very often to develop a theory of finitely differentiable regularity by systematically approximating the differentiable problem by analytic ones: The regularity of the data in the original problem gets translated as the possibility of approximating them by analytic functions which are not very big. By applying a quantitative version of the contraction mapping theorem, it is possible to conclude that there is a sequence of analytic solutions which are not very big either. One can deduce, using the characterization of differentiable functions by the ease of approximation that the solution is differentiable.

Such an approach is discussed systematically in [Zeh75] and in some remarks in [Zeh76].

Some results among the ones we presented here can indeed be obtained using the approach based on analytic regularity outlined above. Nevertheless, we believe that the real variable method outlined here has some advantages that justify our writing it in detail: (*i*) Even for the case of stable and unstable manifolds near a fixed point, the method based in analytic smoothing seems to lose more derivatives. (*ii*) Since cutting off is impossible for analytic functions, the theory based on analytic regularity can only produce stable manifolds and not the center manifolds whose construction we have sketched. (*iii*) The real variable method can deal with invariant Cantor sets provided that the motion on them satisfies certain properties of expansiveness.

The last result we consider is due to Zehnder [Zeh76], whose proof we reproduce for the sake of completeness.

Theorem 4.27. *If Ω is a torus such as those in Theorem 3.8 then W^S_Ω, W^U_Ω are Lagrangian submanifolds.*

Proof. Since the dimension is N, it suffices to show that the symplectic form – which we denote by ω – vanishes.

To show that the symplectic form vanishes it suffices to show that for every pair of vectors v_1, v_2 tangent to W^S_Ω we have that $\omega(v_1, v_2) = 0$.

By the conservation of the symplectic form under hamiltonian evolution we have that $\omega(v_1, v_2) = \omega(\Phi_t^* v_1, \Phi_t^* v_2)$.

Using the fact that $||\Phi_t^* v|| \leq Ce^{-\mu t}||v||$ we see that the component along S is smaller than $Ce^{-\lambda t}||\Pi^S v||$. Thus $\Phi_t^* v_1$ and $\Phi_t^* v_2$ are vectors of length not bigger than $Ce^{-\mu t}$ Since the manifolds are differentiable, they differ from vectors tangent to the torus by an amount not bigger than $Ce^{-\lambda t}$.

Since the symplectic form vanishes on the torus, we see that $|\omega(\Phi_t^* v_1, \Phi_t^* v_2)| \leq Ce^{(\mu-\lambda)t}$. Since t is arbitrary, we have proved the theorem. $\square$

5. Acknowledgements

It is a pleasure to thank E. Formanek for useful conversations, and C. Simó for calling the work of Delshams to our attention and providing us with a copy of it. We also want to thank H. Elliason, J. Pöschel and A. Delshams for their comments about a version of this paper.

C.E.W. wishes to thank T. Spencer for an invitation to the Institute for Advanced Study, Princeton, during which this work was completed. R.L. wants to thank the Dept. of Mathematics at Penn State for repeated invitations that made this work possible.

The work of R.L. was partially supported by NSF Grants #DMS 85-04984 and #DMS89–15118 and that of C.E.W. by NSF Grant #DMS 88-02118 and #DMS 01-03915. A substantial part of this work was done while R.L. was affiliated with Princeton University and C.E.W. with Penn State.

References

[AA68] V. I. Arnol'd and A. Avez. *Ergodic Problems of Classical Mechanics.* W. A. Benjamin, New York-Amsterdam, 1968.

[AM78] R. Abraham and J. E. Marsden. *Foundations of Mechanics.* Benjamin/Cummings, Reading, Mass., 1978.

[AR67] Ralph Abraham and Joel Robbin. *Transversal mappings and flows.* W. A. Benjamin, Inc., New York-Amsterdam, 1967.

[Arn64] V.I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[Ban88] V. Bangert. Mather sets for twist maps and geodesics on tori. In *Dynamics reported, Vol. 1*, pages 1–56. Teubner, Stuttgart, 1988.

[BG86] Giancarlo Benettin and Giovanni Gallavotti. Stability of motions near resonances in quasi-integrable Hamiltonian systems. *J. Statist. Phys.*, 44(3-4):293–338, 1986.

[BGG85] G. Benettin, L. Galgani, and A. Giorgilli. A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. *Celestial Mech.*, 37(1):1–25, 1985.

[BGGS84] G. Benettin, L. Galgani, A. Giorgili, and J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B (11)*, 79(2):201–223, 1984.

[BK87] David Bernstein and Anatole Katok. Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. *Invent. Math.*, 88(2):225–241, 1987.

[BL33] G.D. Birkhoff and D. C. Lewis. On the periodic motions near a given periodic motion of a dynamical system. *Annali di Mat.*, 12:117–133, 1933.

[Car81] John R. Cary. Lie transform perturbation theory for Hamiltonian systems. *Phys. Rep.*, 79(2):129–159, 1981.

[Che85] Alain Chenciner. La dynamique au voisinage d'un point fixe elliptique conservatif: de Poincaré et Birkhoff à Aubry et Mather. *Astérisque*, (121-122):147–170, 1985. Seminar Bourbaki, Vol. 1983/84.

[DdlLMS03] A. Delshams, R. de la Llave, and T. M.-Seara. A geometric mechanism for diffusion in hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. MP_ARC # 03-182, 2003.

[Del83] A. Delshams. *Por qué la difusión de Arnold es genérica en sistemas de mas de dos grados de libertad.* PhD thesis, Univ. de Barcelona, 1983.

[Del88] A. Delshams. Private communication, 1988.

[dlL92] R. de la Llave. A renormalization group explanation of numerical observations of analyticity domains. *J. Statist. Phys.*, 66(5-6):1631–1634, 1992.

[dlL97] Rafael de la Llave. Invariant manifolds associated to nonresonant spectral subspaces. *J. Statist. Phys.*, 87(1-2):211–249, 1997.

[dlLMM86] R. de la Llave, J. M. Marco, and R. Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math. (2)*, 123(3):537–611, 1986.

[dlLW95] Rafael de la Llave and C. Eugene Wayne. On Irwin's proof of the pseudostable manifold theorem. *Math. Z.*, 219(2):301–321, 1995.

[Dou82] R. Douady. Une démonstration directe de l'équivalence des théorèmes de tores invariants pour difféomorphismes et champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 295(2):201–204, 1982.

[Dou88] R. Douady. Stabilité ou instabilité des points fixes elliptiques. *Ann. Sci. École Norm. Sup. (4)*, 21(1):1–46, 1988.

[Dui84] J. J. Duistermaat. Bifurcation of periodic solutions near equilibrium points of Hamiltonian systems. In *Bifurcation theory and applications (Montecatini, 1983)*, pages 57–105. Springer, Berlin, 1984.

[Eli88] L. H. Eliasson. Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(1):115–147 (1989), 1988.

[Eli89] L. H. Eliasson. Hamiltonian systems with linear normal form near an invariant torus. In *Nonlinear Dynamics (Bologna, 1988)*, pages 11–29. World Sci. Publishing, Teaneck, NJ, 1989.

[Fen72] Neil Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1971/1972.

[Fen77] N. Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen74] N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[Gra74] Samuel M. Graff. On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, 15:1–69, 1974.

[GVL83] Gene H. Golub and Charles F. Van Loan. *Matrix computations.* Johns Hopkins University Press, Baltimore, MD, 1983.

[Hal80] Jack K. Hale. *Ordinary differential equations.* Robert E. Krieger Publishing Co. Inc., Huntington, N.Y., second edition, 1980.

[HdlL00] A. Haro and R. de la Llave. New mechanisms for lack of equipartion of energy. *Phys. Rev. Lett.*, 89(7):1859–1862, 2000.

[HK90] S. Hurder and A. Katok. Differentiability, rigidity and Godbillon-Vey classes for Anosov flows. *Inst. Hautes Études Sci. Publ. Math.*, 72:5–61 (1991), 1990.

[HM82] P.J. Holmes and J.E. Marsden. Melnikov's method and Arnol'd diffusion for perturbations of integrable Hamiltonian systems. *J. Math. Phys.*, 23(4):669–675, 1982.

[HP70] M.W. Hirsch and C.C. Pugh. Stable manifolds and hyperbolic sets. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 133–163. Amer. Math. Soc., Providence, R.I., 1970.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds.* Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 583.

[Jou86] Jean-Lin Journé. On a regularity problem occurring in connection with Anosov diffeomorphisms. *Comm. Math. Phys.*, 106(2):345–351, 1986.

[Jou88] J.-L. Journé. A regularity lemma for functions of several variables. *Rev. Mat. Iberoamericana*, 4(2):187–193, 1988.

[Kat92] A. Katok. Minimal orbits for small perturbations of completely integrable Hamiltonian systems. In *Twist mappings and their applications*, pages 183–199. Springer, New York, 1992.

[Kra83] S. G. Krantz. Lipschitz spaces, smoothness of functions, and approximation theory. *Exposition. Math.*, 1(3):193–260, 1983.

[LI73] Oscar E. Lanford III. Bifurcation of periodic solutions into invariant tori: the work of Ruelle and Takens. In Ivar Stakgold, Daniel D. Joseph, and David H. Sattinger, editors, *Nonlinear problems in the Physical Sciences and biology: Proceedi ngs of a Battelle summer institute*, pages 159–192, Berlin, 1973. Springer-Verlag. Lecture Notes in Mathematics, Vol. 322.

[Mat82] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21(4):457–467, 1982.

[Mat86] John Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, (63):153–204, 1986.

[Mil63] J. Milnor. *Morse theory.* Princeton University Press, Princeton, N.J., 1963.

[MM76] J. E. Marsden and M. McCracken. *The Hopf bifurcation and its applications.* Springer-Verlag, New York, 1976. With contributions by P. Chernoff, G. Childs, S. Chow, J. R. Dorroh, J. Guckenheimer, L. Howard, N. Kopell, O. Lanford, J. Mallet-Paret, G. Oster, O. Ruiz, S. Schecter, D. Schmidt and S. Smale, Applied Mathematical Sciences, Vol. 19.

[Mos67] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[Mos76] J. Moser. Periodic orbits near an equilibrium and a theorem by Alan Weinstein. *Comm. Pure Appl. Math.*, 29(6):724–747, 1976.

[Mos77] J. Moser. Proof of a generalized form of a fixed point theorem due to G. D. Birkhoff. In *Geometry and topology (Proc. III Latin Amer. School of Math., Inst. Mat. Pura Aplicada CNPq, Rio de Janeiro, 1976)*, pages 464–494. Lecture Notes in Math., Vol. 597. Springer, Berlin, 1977.

[Mos86] Jürgen Moser. Monotone twist mappings and the calculus of variations. *Ergodic Theory Dynamical Systems*, 6(3):401–413, 1986.

[Neh77] N. N. Nehorošev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Uspehi Mat. Nauk*, 32(6(198)):5–66, 287, 1977. English translation: *Russian Math. Surveys*, 32(6):1–65, 1977.

[New77] Sheldon E. Newhouse. Quasi-elliptic periodic points in conservative dynamical systems. *Amer. J. Math.*, 99(5):1061–1087, 1977.

[Poi99] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*, volume 1, 2, 3. Gauthier-Villars, Paris, 1892–1899.

[Pös82] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653–696, 1982.

[Pös86] J. Pöschel. On invariant manifolds of complex analytic mappings near fixed points. *Exposition. Math.*, 4(2):97–109, 1986.

[Pös89] Jürgen Pöschel. On elliptic lower-dimensional tori in Hamiltonian systems. *Math. Z.*, 202(4):559–608, 1989.

[R8̈8] H. Rüssmann. Private communication, 1988.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press, Princeton, N.J., 1970. Princeton Mathematical Series, No. 30.

[Tak70] Floris Takens. Hamiltonian systems: Generic propeties of closed orbits and local perturbations. *Math. Ann.*, 188:304–312, 1970.

[Wei73] Alan Weinstein. Normal modes for nonlinear Hamiltonian systems. *Invent. Math.*, 20:47–57, 1973.

[WR71] J.H. Wilkinson and C. Reinsch, editors. *Handbook for automatic computation. Vol. II.* Springer-Verlag, New York, 1971. Linear algebra, Die Grundlehren der Mathematischen Wissenschaften, Band 186.

[Zeh73] E. Zehnder. Homoclinic points near elliptic fixed points. *Comm. Pure Appl. Math.*, 26:131–182, 1973.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

[Zeh88] H. Zehnder. Private communication, 1988.

Dept. of Mathematics, Univ. of Texas, Austin TX 78712-1802
E-mail address, R. de la Llave: `llave@math.utexas.edu`

Dept. of Mathematics & Statistics and Center for BioDynamics, Boston Univ.
E-mail address, C.E.Wayne: `cew@math.bu.edu`