KAM theory and a partial justification of Greene's criterion for non-twist maps

Amadeu Delshams
Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya
Diagonal 647, 08028 Barcelona, Spain
amadeu@ma1.upc.es

Rafael de la Llave
Department of Mathematics
The University of Texas at Austin
Austin, TX 78712-1082, U.S.A.
llave@math.utexas.edu

Abstract. We consider perturbations of integrable, area preserving non-twist maps of the annulus (those are maps that violate the twist condition in a very strong sense: $\partial q'/\partial p$ changes sign). These maps appear in a variety of applications, notably transport in atmospheric Rossby waves.

We show in suitable 2-parameter families the persistence of critical circles (invariant circles whose rotation number is the maximum of all the rotation numbers of points in the map) with Diophantine rotation number. The parameter values with critical circles of frequency ω_0 lie on a one dimensional analytic curve. In contrast with recent progress in KAM theorem with degeneration in the frequency map, the curves we consider have rotation numbers on the boundary of the range of the frequency.

Furthermore, we show a partial justification of Greene's criterion: If critical analytic critical curves with Diophantine rotation number ω_0 exist, the residue of periodic orbits with rotation number converging to ω_0 converges to zero exponentially fast. We also show that if analytic curves exist, there should be periodic orbits approximating them and indicate how to compute them.

These results justify conjectures put forward on the basis of numerical evidence in D. del Castillo et al., *Phys. D.* **91**, 1–23 (1996). The proof of both results relies on the successive application of an iterative lemma which is valid also for $2d$-dimensional exact symplectic diffeomorphisms. The proof of this iterative lemma is based on the deformation method of singularity theory.

1. Introduction

The goal of this paper is to provide rigorous proofs of several phenomena discovered empirically in [CGM1]. In that paper, the authors consider the two parameter family of area preserving maps, called there the "quadratic standard map".

$$T_{\omega,\varepsilon}(p,q) = \big(p + \varepsilon \sin(2\pi q), q - (p + \varepsilon \sin 2\pi q)^2 + \omega \pmod 1\big)\,. \tag{1.1}$$

One motivation for such study is that qualitatively similar maps appear naturally in the study of geostrophic flows and indeed in many problems in hydrodynamics and in other applications, mentioned briefly later.

The map $T_{\omega,0}$ describes a situation where particles in a fluid are moving in a laminar flow whose velocity is faster in the middle but slower as we move away from the center of the stream. This is a very common situation in fluid motion, where often the motion slows down as we move closer to edges. The map $T_{\omega,0}$ is an integrable map, since the curves $p = \mathrm{cte}$ are preserved and the motion in them is a rigid rotation. In many applications, it is natural to consider q as an angle. For example, in the description of the jet stream, q corresponds to the longitude and p is a range of latitudes.

The extra term modified by ε is representative of the maps that arise when one considers the physical effect of a periodic oscillation transverse to the channel flow. Such phenomena occur frequently in hydrodynamics when channel flows are destabilized through a Hopf bifurcation. This happens in jet flows in the atmosphere due to Rossby waves. We refer to [C] and [CM] for a detailed description of the fluid mechanics motivation of such models. In particular for the justification of the use of a two dimensional approximation. In this interpretation, the invariant circles are complete barriers for the mixing of the material in the pole — one of the edges of p — with the material near the equator — the other edge of p. This barriers are important in the creation of "ozone holes" since they isolate the ozone created in the tropics from the regions near the poles.

The general theory we will develop will not depend on the exact form for the map, but on qualitative features that can be verified in the realistic models. Of course, the map (1.1) is a concrete model introduced for the purpose of discovering qualitative features through numerical calculations.

We also point out that models having non-twist maps have appeared with other motivations. For example, they appear in celestial mechanics in problems such as the "critical inclination" [K] and in the study of billiards with a boundary moving periodically in time [KMOP1,KMOP2] or in the study of the motion of particles in magnetic fields [ZZUSC]. As a matter of fact, since the iterates of a twist map are not, in general, twist maps, we expect that they appear also as descriptions of regions of iterates of twist maps. (See. e.g. [BST,Si].)

These non twist maps exhibit a very rich phenomenology that is only now started to be explored. The papers cited above as well as [VG,HH1,HH2,Si,Ha1,Ha2] contain descriptions and studies of a wealth of phenomena such as "scaling relations", "reconnection", "meandering curves" etc., that deserve to be investigated further. Notably in [Si,Ha1], there are studies of new phenomena that happen in higher dimensional non-twist maps.

In this paper, we will develop rigorous techniques that can produce results on two problems of the ones mentioned above: The existence of critical invariant circles and the validity of Greene's criterion. Needless to say, we hope that the techniques that we develop for this purpose (e.g. finding appropriate normal forms and quantitative error estimates of them in neighborhoods) can eventually be used in the study of some of these other phenomena.

The paper [CGM1] finds numerically — among other results — numerical evidence for the following:

Claim 1.1. *Let $\omega_0 = (\sqrt{5}-1)/2$. Then, for $|\varepsilon| \ll 1$ there is a smooth curve $\omega(\varepsilon)$ with $\omega(0) = \omega_0$ such that*

a) If $\omega > \omega(\varepsilon)$, then $T_{\omega,\varepsilon}$ admits two invariant circles with rotation number ω_0.

b) If $\omega < \omega(\varepsilon)$, then $T_{\omega,\varepsilon}$ admits no invariant circles with rotation number ω_0.

c) If $\omega = \omega(\varepsilon)$, then $T_{\omega(\varepsilon),\varepsilon}$ admits an invariant circle with rotation number ω_0.

The circle in c), moreover is "critical", that is, there exists a change of variables in its neighborhood in such a way that

$$h^{-1} \circ T_{\omega(\varepsilon),\varepsilon} \circ h(A,\varphi) = (A, \varphi + \omega_0 + \kappa A^2) + O(A^3) \qquad \kappa \neq 0$$

(in fact, $\kappa < 0$ for the example in (1.1)).

Moreover, they also find numerical evidence for:

Claim 1.2. *Greene's criterion applies.*

We recall that the Greene's criterion, introduced in [Gr], asserts that there exists a smooth invariant circle with rotation number ω_0 if and only if

$$\operatorname{Res}(O_{m,n}) := \tfrac{1}{4}\left[\operatorname{tr}\left(DT^n_{\omega,\varepsilon}\left(O_{m,n}\right)\right) - 2\right] \underset{m/n\to\omega_0}{\longrightarrow} 0$$

for any sequence of periodic orbits O_{mn} of type m/n converging to ω_0.

More interestingly, a followup paper [CGM2] goes on to find scaling relations for the invariant circles with rotation number ω_0 of $T_{\omega(\varepsilon),\varepsilon}$ as ε goes to a critical value where they cease to exist. Unfortunately, we do not have any rigorous result for this non-perturbative regime and, hence, we just call attention to its existence.

In this paper, we will prove rigorous results that justify the experimental results we stated in detail above. We will state and prove a result that justifies Claim 1.1 and another one that justifies one of the implications in Claim 1.2. Namely that, if there exists an invariant circle, the residue goes to zero.

The converse (that is, if the residue goes to zero, one can find a periodic orbit), to our knowledge, remains an open problem even for twist maps. However we call attention to the work of [KO], which proves that if there are periodic orbits of twist maps which are, in a precise sense, well distributed, one can find a KAM circle with rotation number related to that of the periodic orbit. We also note that if the renormalization group picture can be justified, at least to a certain extent, the Greene's criterion will be also justified and indeed several improvements on that give precise asymptotic of the residue (see [McK].)

It is worth remarking that an easy argument, which we will detail later in Proposition 4.4, shows that if there is a critical invariant circle as above, indeed it is approximated by periodic orbits of type m/n with m/n converging to ω_0. Hence, this criterion is rather effective.

We also point out that, for the maps $T_{\omega,\varepsilon}$ as in (1.1), the criterion can be implemented numerically very efficiently. These maps are reversible and, for reversible maps, the search for periodic orbits in some critical lines — not all of the periodic orbits of the map — can be reduced to finding zeros of one dimensional functions, a tractable numerical task. Moreover, it is possible to show that if there are critical circles, it should be possible to find periodic orbits in these critical lines converging to them. We will return to these questions later, in Proposition 4.4.

About the method of proof we note that there are two basic methods in KAM theory. One is based on applying successive transformations and another one is based on solving functional equations that express invariance.

Both methods have complementary advantages. The functional equation method leads to very crisp proofs and they are more natural for numerical implementations. On the other hand, the methods based on transformation theory yield more information about the behavior of the map on a neighborhood of the invariant torus.

Since in this paper we wanted to discuss the partial justification of the Greene's criterion, we certainly needed a method based on the transformation theory and it was natural to use the same method for the proof of the persistence of the invariant tori.

In the future, we plan to come back to the functional method, specially in connection with a numerical implementation.

The proof we present here will be based on the deformation method. This method was introduced in the study of singularities of mappings [TL,Mat] and it is very well suited for study of equivalence of maps in situations where geometric structures are present [BLW], like families of exact symplectic diffeomorphisms. One can use it also for the regular KAM theorem [Ll]. In our case, the use of the deformation method is very natural since the unknown involves a family of maps.

Note that in this situation we are trying to study the persistence of invariant circles whose frequency is on the boundary of the frequencies that are present on the integrable map. This is in contrast with the KAM theory, where the non-degeneracy conditions — the so called twist condition or the more sophisticated Rüssmann conditions (see [BHS], chapter 4) — imply that the frequency under study is in the interior of the frequencies of the invariant circles in the integrable case.

Since the frequency we want to study is on the boundary of the frequencies, it is not difficult to consider a perturbation of the integrable case in which there is no invariant circle with the frequency we want. (It suffices to consider an integrable perturbation in which we just add — or subtract — an extra rotation so that all the invariant circles persist, but their rotation number is changed.)

Speaking heuristically, what we will do is to consider the regular perturbation theory supplemented with a choice of $\omega(\varepsilon)$. The regular perturbation theory may force the ω_0 out of the range of frequencies, but we will find the extra rotation $\omega(\varepsilon)$ that puts it on the boundary. Since in this method of proof one needs to consider families all the time, the use of the deformation method seems particularly well justified.

On a more technical level, we note that the proof will be based on an iterative lemma (Lemma 3.6), that describes how it is possible to obtain transformations that reduce the system to integrable. Moreover, we will present bounds on the error of this reduction depending on the domain. This iterative lemma can be applied repeatedly in different ways depending on how one plays the trade off between domain loss and accuracy. One can try to make the error decrease very fast at the price that the domain decreases very fast or one can make the error decrease slowly on a larger domain. In this way, one can obtain a unified approach towards KAM theory and towards exponentially small estimates, which we will show justify Greene's criterion. This approach has precedents in [DG1]. Since the iterative lemma as well as the deformation method are widely applicable, we have developed it in an arbitrary dimension. The geometric considerations that lead to the KAM theorem for critical circles and to the Greene's criterion seem to be different in higher dimensions, so we have postponed the discussion of this part.

Now we turn to making all these ideas more precise.

Definition 1.3. *We say that a circle S, invariant under an area preserving map T of $\mathbb{R} \times \mathbb{T}^1 \equiv M$, is a critical invariant circle if there exists a canonical transformation $h : [-\delta, \delta] \times \mathbb{T}^1 \to M$ in such a way that*

$$h^{-1} \circ T \circ h(A, \varphi) = (A, \varphi + \omega_0 + \kappa A^2) + O(A^3)$$

with $\kappa \neq 0$ and $h(\{0\} \times \mathbb{T}^1) = S$.

Remark. The definition of a critical circle includes in its hypothesis that the motion on the circle is conjugate to a rotation of ω_0. We will not include the ω_0 in the notation since it will be understood from the context.

We also recall — and we will develop it in more detail later in Lemma 4.2 — that there is an analogue of Birkhoff normal form in a neighborhood of an invariant circle with a Diophantine rotation. (In the twist map case, this was also considered in [OS,FL].) Given $N \in \mathbb{N}$, it is possible to find coefficients $\kappa_1, \ldots, \kappa_N$ and a canonical transformation h such that

$$h^{-1} \circ T \circ h(A, \varphi) = (A, \varphi + \omega_0 + \kappa_1 A + \kappa_2 A^2 + \cdots + \kappa_M A^M) + O(A^{M+1}) \tag{1.2}$$

The coefficients $\kappa_1, \ldots, \kappa_N$ are uniquely defined and are properties of the invariant circle. In this language, critical circles are those for which $\kappa_1 = 0$, $\kappa_2 \neq 0$.

Definition 1.4. *We will call a circle non-degenerate when the normal form (1.2) does not vanish identically. That is, we can find $M \in \mathbb{N}$ such that $\kappa_1 = \cdots = \kappa_{M-1} = 0$, $\kappa_M \neq 0$.*

Our result to justify Claim 1.1 is:

Theorem 1.5. *Let ω_0 be a Diophantine number, (that is, $|k \cdot \omega_0 - m|^{-1} \leq C|k|^{\theta-1}, \forall k \in \mathbb{Z}, m \in \mathbb{Z} \setminus \{0\}$), $f_{\omega,\varepsilon}$ be a family of mappings from $\mathbb{R}^1 \times \mathbb{T}^1$ to itself satisfying*

i) $f_{\omega,\varepsilon}(p, q)$ is analytic in

$$|\omega - \omega_0| < \rho_0 , \quad |\varepsilon| < \rho_0 , \quad |\mathrm{Im}\ q| < \beta_0 , \quad |p| < \rho_0 .$$

and takes real values for $\omega, \varepsilon, p, q$ real

ii) $f_{\omega,\varepsilon}$ is exact symplectic for all ω, ε

iii) $f_{\omega,0}(p,q) = (p, q + \Gamma(\omega,p))$ *with*

$$\Gamma(\omega_0,0) = \omega_0, \qquad \frac{\partial}{\partial p}\Gamma(\omega_0,0) = 0$$
$$\frac{\partial^2}{\partial p^2}\Gamma(\omega_0,0) = t < 0, \qquad \frac{\partial}{\partial \omega}\Gamma(\omega_0,0) = s > 0$$

Then, we can find a $\delta > 0$ and an analytic function ω defined for $|\varepsilon|$ sufficiently small and taking real values for ε real in such a way that

a) $f_{\omega(\varepsilon),\varepsilon}$ *has exactly one critical invariant circle in* $[-\delta,\delta] \times \mathbb{T}^1$

b) if $\omega < \omega(\varepsilon)$, $f_{\omega,\varepsilon}$ *has no points in* $[-\delta,\delta] \times \mathbb{T}^1$ *with rotation number* ω_0, *and if* $\omega > \omega(\varepsilon)$ *there are two invariant circles of* $f_{\omega,\varepsilon}$ *in* $[-\delta,\delta] \times \mathbb{T}^1$ *which are not critical.*

Remark. It is possible to change hypothesis *iii*) of Theorem 1.5, to be that t is positive. It suffices to change the inequalities between ω, $\omega(\varepsilon)$ in part *b*) of the conclusions and the proof goes through without change. Similarly, if s is negative in *iii*).

The precise meaning in which Greene's criterion can be justified is the following. We will show

Theorem 1.6. *Let $f_{\omega,\varepsilon}$ be an analytic area preserving diffeomorphism of the annulus. Assume that $f_{\omega,\varepsilon}$ admits an analytic invariant circle Γ on which the motion is analytically conjugate to a rotation with Diophantine number ω_0 and which is non-degenerate in the sense of Definition 1.4.*

Then, we can find $C_1, C_2, \mu > 0$ (depending on ω_0, the map and the torus) such that for any sequence of periodic orbits O_n of type p_n/q_n which are converging to Γ and such that $|\omega_0 - p_n/q_n| \le 1/q_n$, we have

$$\mathrm{Res}(O_n) \le C_1 \exp\left(-C_2|\omega_0 - p_n/q_n|^{-\mu}\right) \tag{1.3}$$

We will also show that there is one such sequence of periodic orbits converging to the non-degenerate circle. Of course, when the circle is critical, depending on the sign of $\omega - p_n/q_n$ we will find either one or two periodic orbits. For more general non-degenerate circles, when M is even we will find one or two periodic orbits of type p_n/q_n depending on the sign of $\omega - p_n/q_n$ and when M is odd we will find one irrespective of the sign or $\omega - p_n/q_n$.

Remark. For a Diophantine number, $|\omega_0 - p_n/q_n|^{-1} \ge C q_n^\tau$ for some $\tau \ge 2$, if we take p_n/q_n to be the convergents of the continued fraction expansion of ω_0. Hence, the conclusion of Theorem 1.6 can be written as

$$\mathrm{Res}(O_n) \le C_1 \exp(-C_2 q_n^{\mu'})$$

Remark. The proof that the residue goes to zero faster than any power is significantly easier than the proof with an explicit rate.

Remark. We note that the paper [CGM2] finds scaling relations that suggest that there is a renormalization group description of these invariant circles in the case of a golden mean rotation number $\omega_0 = (\sqrt{5}-1)/2$, with the KAM circles corresponding to a trivial fixed point. If this was the case (to our knowledge nobody has yet worked out a precise formulation and computed the trivial fixed points), the residue of a periodic orbit of type F_n/F_{n+1} would go to zero super-exponentially fast in n, since for the Fibonacci numbers $F_0 = F_1 = 1$, $F_{n+1} = F_n + F_{n-1}$, one has $|\omega_0 - F_n/F_{n+1}|^{-1} \approx C\omega_0^{-2n}$.

2. The deformation method

In this section we recall the basis of the deformation method for symplectic maps. This method was introduced in singularity theory [TL,Mat], but it was remarked later that it can be used very effectively to obtain structure theorems for volume preserving maps of a manifold [Mo1], or for symplectic maps [W] giving a very direct proof of Darboux theorem. More details and other applications can be found in [LMM,Ll,BLW] and in several other places.

In this section, the dimension of the space will not play a role, so we will consider M a $2d$-dimensional manifold.

We recall that a 2-form ϖ on M is a symplectic form if it is closed and has full rank. (Of course, the fact that ϖ has full rank implies that the dimension of M is even, this is why we chose the notation $2d$ for it.) We will be specially interested in the case when ϖ is exact. That is, there exist a 1-form ϑ such that $\varpi = d\vartheta$.

A diffeomorphism f is symplectic when $f_*\varpi = \varpi$. For ϖ exact, this is equivalent to $d(f_*\vartheta - \vartheta) = 0$. We say that a symplectic map f is exact when $f_*\vartheta - \vartheta = dS$ for some function S, called the primitive function of f.

Given a family of diffeomorphisms f_ε, we denote by $\mathcal{F}_\varepsilon$ the vector field defined by

$$\frac{d}{d\varepsilon} f_\varepsilon = \mathcal{F}_\varepsilon \circ f_\varepsilon \tag{2.1}$$

and refer to $\mathcal{F}_\varepsilon$ as the generator of f_ε. Note that a family determines the generator and, conversely, by the uniqueness theorem for O.D.E.'s, a family is determined by its initial point f_0 and its generator, when the generator is $\mathcal{C}^1$. (We will always assume that this is the case.)

The main idea of the deformation method is to work always with the generators, which, when the families are differentiable enough so that the uniqueness theorem for O.D.E.'s applies, is equivalent to working with the families. When the diffeomorphisms are symplectic, further simplifications are possible. Using Cartan's formula for Lie derivatives and that ϖ is closed we obtain

$$\begin{aligned} \frac{d}{d\varepsilon} f_{\varepsilon *}\varpi &= f_{\varepsilon *}(d(i(\mathcal{F}_\varepsilon)\varpi) + i(\mathcal{F}_\varepsilon)\, d\varpi) = f_{\varepsilon *}(d(i(\mathcal{F}_\varepsilon)\varpi)) \\ \frac{d}{d\varepsilon} f_{\varepsilon *}\vartheta &= d(f_{\varepsilon *}(i(\mathcal{F}_\varepsilon))\vartheta) + f_{\varepsilon *}(i(\mathcal{F}_\varepsilon)\varpi) \end{aligned} \tag{2.2}$$

If f_ε is symplectic, $\frac{d}{d\varepsilon} f_{\varepsilon *}\varpi = 0$, and then we see that

$$d(i(\mathcal{F}_\varepsilon)\varpi) = 0 \tag{2.3}$$

If f_ε is exact symplectic, $d\left(\frac{d}{d\varepsilon} S_\varepsilon\right) - f_{\varepsilon *} d(i(\mathcal{F}_\varepsilon)\varpi) = f_{\varepsilon *}(i(\mathcal{F}_\varepsilon)\vartheta)$ and, therefore,

$$i(\mathcal{F}_\varepsilon)\varpi = dF_\varepsilon \tag{2.4}$$

with $F_\varepsilon = \left(\frac{d}{d\varepsilon} S_\varepsilon\right) \circ f_\varepsilon - i(\mathcal{F}_\varepsilon)\vartheta$.

Conversely, if $\mathcal{F}_\varepsilon$ satisfies (2.3) or (2.4) and f_0 is symplectic or exact symplectic, the family f_ε is symplectic or exact symplectic as can be seen integrating (2.2).

Along this paper, we will refer to F_ε as the Hamiltonian for the family f_ε. Note that given f_ε, (2.4) determines F_ε up to a function of zero differential hence, constant on each connected component of its domain of definition. This justifies calling F_ε "the Hamiltonian" if we think of Hamiltonians as equivalent when they differ in a function with zero differential. This identification is natural since two Hamiltonian differing by a function with zero differential generate the same dynamics.

Conversely, for a $\mathcal{C}^2$ Hamiltonian F_ε, given that ϖ is full rank, (2.4) determines $\mathcal{F}_\varepsilon$, and it is $\mathcal{C}^1$. This $\mathcal{F}_\varepsilon$ and f_0 determine f_ε by the uniqueness result for O.D.E.'s.

Hence, for sufficiently smooth families it is equivalent to work with the Hamiltonians and the initial points of the families.

The main idea of the deformation method for exact symplectic maps is to reformulate all the problems in terms of Hamiltonians. As it turns out, the equations involving generators are linear. This is to be expected since we can heuristically think of generators as infinitesimal transformations and all the equations among infinitesimal quantities are linear. Moreover, using Hamiltonians, the otherwise complicated constraint of the transformations being exact symplectic is implemented automatically, and the resulting equations only involve functions. Hence, rather than dealing with non-linear equations among diffeomorphisms satisfying non-linear constraints, we just have to deal with a linear equation among functions.

We will follow the convention of denoting families in lower case f_ε, their generators in calligraphic font $\mathcal{F}_\varepsilon$ and the Hamiltonians in upper case F_ε.

Proposition 2.1. *Let $f_\varepsilon, g_\varepsilon$ be exact symplectic families and k an exact symplectic diffeomorphism. Then, the Hamiltonian of the families formed out of them are given in the following table.*

family	*Hamiltonian*
$f_\varepsilon \circ g_\varepsilon$	$F_\varepsilon + f_{\varepsilon *} G_\varepsilon = F_\varepsilon + G_\varepsilon \circ f_\varepsilon^{-1}$
f_ε^{-1}	$-F_\varepsilon \circ f_\varepsilon$
$g_\varepsilon^{-1} \circ f_\varepsilon \circ g_\varepsilon$	$F_\varepsilon \circ g_\varepsilon - G_\varepsilon \circ g_\varepsilon + G_\varepsilon \circ f_\varepsilon^{-1} \circ g_\varepsilon$
$k^{-1} \circ f_\varepsilon \circ k$	$F_\varepsilon \circ k$
$f_\varepsilon \circ k$	F_ε

The computations needed to work out this table can be found in [LMM,BLW]. In the latter paper one can find similar tables for volume preserving or contact families.

Since in perturbation theory one does not always have a family of diffeomorphisms but just two diffeomorphisms that are close, it is worth remarking that given two symplectic diffeomorphisms that are close, one can always interpolate them by a family with small Hamiltonian. If the two maps are exact, the family can be chosen to be exact. This is an immediate consequence of the general fact that symplectic (or exact symplectic) maps form a Banach manifold (see [W]). We just sketch a direct construction whose details appear in [BLW]. An alternative, old fashioned proof can be obtained using generating functions. (Interpolate the generating functions.) Unfortunately, since it is impossible to obtain generating functions that are globally defined, one has to also use partitions of unity and fragmentation lemmas and the proof becomes cumbersome.

Given f_0, f_1 symplectic and close enough, we can find a family of diffeomorphisms f_ε interpolating between them (e.g., $f_\varepsilon(x) = \exp_{f_0(x)} \varepsilon \exp^{-1}_{f_0(x)} f_1(x)$ where exp is the Riemannian exponential map). The family f_ε will not be symplectic. In general, $f_{\varepsilon *}\varpi = \varpi_\varepsilon$ where ϖ_ε is a family of symplectic forms. Note that, by our assumptions $\varpi_0 = \varpi_1 = \varpi$. Using Moser's construction [Mo1] — we refer to [LMM,BLW] for the elementary justification of the smooth dependence on parameters in Moser's construction — we can find h_ε close to the identity in such a way that $h_{\varepsilon *}\varpi_\varepsilon = \varpi$. Moreover, $h_0 = h_1 = \mathrm{Id}$. Then $\tilde f_\varepsilon = h_\varepsilon \circ f_\varepsilon$ satisfies $\tilde f_0 = f_0$, $\tilde f_1 = f_1$, $\tilde f_{\varepsilon *}\varpi = \varpi$. If $\varpi = d\vartheta$ then $\varpi_\varepsilon = d\vartheta_\varepsilon$ with $\vartheta_\varepsilon = f_{\varepsilon *}\vartheta$. Also $(h_\varepsilon \circ f_\varepsilon)_* \vartheta - \vartheta$ is closed. It is then possible to choose g_ε close to the identity in such a way that $(g_\varepsilon \circ h_\varepsilon \circ f_\varepsilon)_* \vartheta - \vartheta$ is exact (e.g., on the annulus choose translations in the radial direction and in another manifolds choose a displacement in a neighborhood of paths that generate the homology.)

We have, therefore, established

Lemma 2.2. *Let f_0 be a $\mathcal{C}^\infty$ (resp. $\mathcal{C}^\omega$) symplectic (resp. exact symplectic) diffeomorphism of a manifold.*

If f_1 is a symplectic (resp. exact symplectic) diffeomorphism close to f_0 we can find a $\mathcal{C}^\infty$ (resp. $\mathcal{C}^\omega$) family f_ε of symplectic (resp. exact symplectic) diffeomorphisms interpolating between f_0 and f_1.

Moreover, we can arrange that the generators and therefore the Hamiltonians of the isotopy are arbitrarily small in the $\mathcal{C}^\infty$ (resp. $\mathcal{C}^\omega$) topology by assuming that f_1 is arbitrarily close to f_0.

3. Proof of Theorem 1.5 using the deformation method

3.1. Heuristic discussion

The proof we present here starts with the observation that the result would be obvious if we had a family of the form

$$i_{\omega,\varepsilon}(p,q) = (p, q + \Omega(\omega,\varepsilon,p)) \tag{3.1}$$

in which the p is conserved and the q is translated by $\Omega(\omega,\varepsilon,p)$, which depends on p and on external parameters and is close to the frequency $\Gamma(\omega,p)$ satisfying hypothesis *iii*) of Theorem 1.5. We will refer to such families as integrable.

If we require that the set $p = p_0$ is an invariant circle with rotation ω_0, we obtain the implicit equation

$$\Omega(\omega, \varepsilon, p_0) = \omega_0 \tag{3.2}$$

The possibility of finding solutions of (3.2) is described by singularity theory and the phenomenon of a critical invariant circle corresponds to the situation when $\Omega(\omega, \varepsilon, p_0) - \omega_0$ has a fold:

$$\Omega(\omega, \varepsilon, p_0) - \omega_0 = 0, \qquad \partial_p \Omega(\omega, \varepsilon, p_0) = 0$$

The equation for $\omega(\varepsilon)$ is precisely the equation for the edge of a fold. We will parameterize the folding surface (3.2) as the set of points $(\Upsilon(\varepsilon, p), \varepsilon, p)$ for an appropriate function Υ:

$$\Omega(\omega, \varepsilon, p) = \omega_0 \iff \omega = \Upsilon(\varepsilon, p) \tag{3.3}$$

Then, a critical invariant circle takes place at $p = p_0 = p_0(\varepsilon)$ if $\partial_p \Upsilon(\varepsilon, p_0) = 0$, and $\omega(\varepsilon) = \Upsilon(\varepsilon, p_0(\varepsilon))$.

A standard technique in KAM theory is to make changes of variables so that in the new system of coordinates, the properties of the map are apparent from its expression. In the present case, we try to find g_ε in such a way that

$$\tilde{f}_{\omega,\varepsilon} = g_\varepsilon^{-1} \circ f_\varepsilon \circ g_\varepsilon \tag{3.4}$$

has the desired form (3.1).

Unfortunately, in general it is not possible to obtain a change of variables reducing to (3.1) in the whole phase space. We only know how to do it approximately in a subset of the domain in (ω, ε, p) for which $\Omega(\omega, \varepsilon, p) = \omega_0$.

Hence we will use an iterative scheme in which at step n, the system will be (described in the notation of the deformation method by the initial point of the isotopy and the generating Hamiltonian)

$$f^n_{\omega,0}(p, q) = (p, q + \Gamma(\omega, p)); \qquad F^n_{\omega,\varepsilon}(p, q) = I^n_{\omega,\varepsilon}(p) + E^n_{\omega,\varepsilon}(p, q) \tag{3.5}$$

where $E^n_{\omega,\varepsilon}$ is "small" in a neighborhood of $\{p = 0\}$.

The Hamiltonian $I^n_{\omega,\varepsilon}(p)$ corresponds to a deformation of the form

$$i^n_{\omega,\varepsilon}(p, q) = (p, q + \Omega^n(\omega, \varepsilon, p)) \tag{3.6}$$

where

$$\Omega^n(\omega, \varepsilon, p) = \Gamma(\omega, p) + \int_0^\varepsilon ds \frac{\partial}{\partial p} I^n_{\omega,s}(p) \tag{3.7}$$

when we assume that $i_{\omega,0} = f_{\omega,0}$. Hence, the $I^n_{\omega,\varepsilon}$ should be thought of as the integrable part of the Hamiltonian $F^n_{\omega,\varepsilon}$. We will think of $E^n_{\omega,\varepsilon}$ as an error term that is to be made smaller and smaller in the iterative process.

Remark. We note that the decomposition of a Hamiltonian into an integrable part and an small part is not uniquely defined. A particularly natural one would be to take the integrable part to be the average over the q. Nevertheless, we will not be assuming that this natural decomposition is taken, just that such a decomposition exists.

Remark. Note that when we consider perturbations of an integrable system, we can write the integrable part in Ω and, hence, assume that $I^0_{\omega,\varepsilon}(p) = 0$.

The main ingredient of the proof of Theorem 1.5 will be an algorithm that, given a family as in (3.5), finds a transformation $g^n_{\omega,\varepsilon}$ defined in a neighborhood of the surface $\Omega^n(\varepsilon, \omega, p) = \omega_0$ such that setting $f^{n+1}_{\omega,\varepsilon} = (g^n_{\omega,\varepsilon})^{-1} \circ f^n_{\omega,\varepsilon} \circ g^n_{\omega,\varepsilon}$ we have

$$F^{n+1}_{\omega,\varepsilon}(p, q) = I^{n+1}_{\omega,\varepsilon}(p) + E^{n+1}_{\omega,\varepsilon}(p, q)$$

where $E^{n+1}_{\omega,\varepsilon}$ is much smaller than $E^n_{\omega,\varepsilon}$ and $I^{n+1}_{\omega,\varepsilon}$ differs little from $I^n_{\omega,\varepsilon}$ in a domain which will be chosen appropriately (a smaller neighborhood of the surface $\Omega^{n+1} = \omega_0$.)

Since Ω^{n+1} is close to Ω^n, the folding surfaces defined by $\Omega^{n+1} = \omega_0$ and by $\Omega^n = \omega_0$ are very close. Quantitative estimates will show that the $E^n_{\omega,\varepsilon}$'s decrease super-exponentially and that the $g^n_{\omega,\varepsilon}$'s differ from the identity by a super-exponentially small quantity in neighborhoods of the surfaces $\Omega^{n+1} = \omega_0$. As it turns out, we will have to choose these neighborhoods to become super-exponentially thinner. The transformations will be defined in these thin slivers in the ω, ε, p coordinates and in domains in q which include complex extensions of $\mathbb{T}^1$ so that the size of the size of the imaginary extension of the domain remains bounded from below.

Similarly, the functions Ω^n converge to a function Ω^∞. Therefore, the surfaces $\hat{\Omega}^n \equiv (g^1\circ\cdots\circ g^n)^{-1}\{\Omega^n = \omega_0\}$ converge to a surface $\hat{\Omega}^\infty$. Since each of the surfaces $\{\Omega^n = \omega_0\}$ is foliated by smooth circles invariant by $F \circ g^1 \circ \cdots \circ g^n$ up to super-exponentially small errors, it follows that $\hat{\Omega}^\infty$ is foliated by smooth circles invariant by F.

For the benefit of experts, we point out that an alternative method to prove Theorem 1.5 could have been to use the non-degeneracy in ω to prove a KAM theorem for all small enough ε and p. (That is, we fix ε and p, but allow ourselves to choose the ω). Even if not all methods to prove KAM theorems would have worked, it seems that methods based on the "translated curve method" works since one can use the ω to adjust the frequency. Then, one needs to prove the analytic dependence of the circle on the parameter ε and to prove that there is indeed a fold.

The method we develop in this paper seems more appealing since one has an understanding of the folding surface at all the stages of the iteration and it is certainly not longer to write in all detail.

Moreover, we can use much of the technology developed along these lines, to prove the partial converse of Greene's theorem. In particular, Lemma 3.6 is the crux of the iterative step in the proofs of both problems. The difference between the KAM theorem and the proof of the exponentially small estimates that imply Greene's criterion, lies only in different choices on how we iterate the method. In the KAM theorem, we lose domain very fast and drive the errors to zero very fast. In the exponentially small estimates, we reduce the domains more slowly and do not obtain convergence but the estimates are valid in a larger domain.

We also call attention to the fact that Lemma 3.6 is valid in any dimension. It is only the geometric considerations about domains that one uses to conclude Theorem 1.5 and Theorem 1.6 that require the fact that we are working in an annulus. We think that this restriction can be lifted with some small amount of extra effort.

3.2. Notation and elementary estimates

Since the iterative step will rely on making transformations on functions in such a way that the errors become smaller, we will need to define appropriate norms. We will also need to be able to manipulate sets where our transformations will be defined. (As usual in KAM theory, one has to consider functions defined in decreasing sets.) In this section, we collect the definitions of the norms, parameterizations of sets that we will use later as well as some elementary lemmas and propositions dealing with them.

Since Lemma 3.6 is valid in any number of dimensions, we will be considering maps in $\mathbb{R}^d \times \mathbb{T}^d$ till the end of Section 3.5.

We recall the standard definition that $\omega_0 \in \mathbb{R}^d$ is said to be Diophantine of exponent θ if we can find a $C > 0$ such that for all $k \in \mathbb{Z}^d, m \in \mathbb{Z}$ we have

$$|k \cdot \omega_0 - m|^{-1} \le C|k|^{\theta-1} \tag{3.8}$$

This is the definition of Diophantine vectors that appears naturally KAM theory for maps. (The definition that appears naturally in the KAM theory for flows is slightly different.)

Besides the above standard definition, in this paper we will use the following notations.

We will denote by $I_{a,b}$ the real interval $[a, b]$, by $B_{x,c}$ the closed ball in $\mathbb{R}^d$ with center $x \in \mathbb{R}^d$ and radius $c > 0$, and by $\mathbb{T}^d$ the d-dimensional torus $\mathbb{R}^d/\mathbb{Z}^d$.

We will also denote by $I_{a,b,\delta} = \{z \in \mathbb{C} \mid d(z, I_{a,b}) \le \delta\}$, $B_{x,c,\delta} = \{z \in \mathbb{C}^d \mid d(z, I_{a,b}) \le \delta\}$. Similarly we will denote by $\mathbb{T}^d_\beta$ the complex extensions on the torus $\mathbb{T}^d$ of a distance β.

Given a set $U = B_{x_1,c_1,\delta} \times I_{a_2,b_2,\delta} \times B_{x_3,b_3,\delta}$ and a function $\Omega : U \to \mathbb{C}^d$, we will denote for $\alpha, \beta > 0$

$$
\begin{aligned}
\Sigma_{\beta,U} &= \{(\omega,\varepsilon,p,q) \mid (\omega,\varepsilon,p) \in U, |\mathrm{Im}\ q| \le \beta\} = U \times \mathbb{T}^d_\beta \\
\Sigma_{\Omega,\alpha,\beta,U} &= \{(\omega,\varepsilon,p,q) \in \Sigma_{\beta,U} \mid |\Omega(\omega,\varepsilon,p) - \omega_0| \le \alpha\}
\end{aligned}
\tag{3.9}
$$

The way to think about $\Sigma_{\Omega,\alpha,\beta,U}$ is as the Cartesian product of a thin film — of width α, which will be extremely small in the proof — around a portion of surface given by the equation $\Omega(\omega,\varepsilon,p) = \omega_0$ and a complex extension of width β of the torus. The parameter U just limits which portion of the surface we are considering and it plays a somewhat minor role.

Note that, for the sake of notation, we are suppressing some of the parameters on which $\Sigma_{\Omega,\alpha,\beta,U}$ depends. Notably ω_0. We hope that this does not lead to confusion in the proof since the values of these parameters will be kept fixed. The ω_0 will be that appearing in Theorem 1.5 and, hence, will not change throughout the proof.

We will introduce the notation U_σ to denote a domain formed by restricting the domain only in the variable p by an amount $\sigma > 0$, that is, $U = B_{x_1,c_1,\delta} \times I_{a_2,b_2,\delta} \times B_{x_3,b_3,\delta-\sigma}$.

This will be used later since we need to reduce the domains in phase space (to guarantee that compositions make sense) but the domains in parameters are not affected.

Given a complex domain Σ, we will denote by $\|F\|_\Sigma \equiv \sup_{x\in\Sigma} |F(x)|$ and by χ^Σ the Banach space of functions analytic in Σ (analytic in the interior and continuous up to the boundary) equipped with the norm $\|\cdot\|_\Sigma$. In particular, for $\Sigma = \Sigma_{\beta,U}$, $\Sigma = \Sigma_{\Omega,\alpha,\beta,U}$ of the form (3.9), for typographical reasons, we will write $\|\cdot\|_{\Sigma_{\beta,U}}$ as $\|\cdot\|_{\beta,U}$ and $\|\cdot\|_{\Sigma_{\Omega,\alpha,\beta,U}}$ as $\|\cdot\|_{\Omega,\alpha,\beta,U}$.

For a function $F : U \times \mathbb{T}^d_\beta \to \mathbb{C}$, where $U = B_{x_1,c_1,\delta} \times I_{a_2,b_2,\delta} \times B_{x_3,b_3,\delta}$. we define the partial Fourier expansion

$$
F_{\omega,\varepsilon}(p,q) = \sum_{k\in\mathbb{Z}^d} \hat{F}_{\omega,\varepsilon;k}(p) e^{2\pi i\, k\cdot q}
$$

The coefficients are unique in the regularity classes we will be considering.

For this kind of functions depending on parameters, we will use the notation ∇ to denote the derivatives with respect to the variables, not with respect to the parameters. Hence

$$
\nabla F_{\omega,\varepsilon}(p,q) = \left(\frac{\partial}{\partial p} F_{\omega,\varepsilon}(p,q), \frac{\partial}{\partial q} F_{\omega,\varepsilon}(p,q) \right)
$$

In the cases that we will need to consider derivatives with respect to the parameters, we will write them explicitly.

We recall that the well known Cauchy inequalities allow us to bound derivatives (in a domain) and Fourier coefficients of a function in terms of its size in a (slightly larger) domain.

Lemma 3.1. *Let $U = B_{x_1,c_1,\delta} \times I_{a_2,b_2,\delta} \times B_{x_3,b_3,\delta}$, $\tilde{U} \subset U$ be a domain that is at a distance $\sigma > 0$ from the complement of U, and $F : U \times \mathbb{T}^d_\beta \to \mathbb{C}$ analytic. Then,*

$$
\begin{aligned}
\|\nabla^m F\|_{\beta-\sigma,\tilde{U}} &\le K\sigma^{-m} \|F\|_{\beta,\tilde{U}} \\
\|\partial^m_\omega F\|_{\beta,\tilde{U}}, \|\partial^m_\varepsilon F\|_{\beta,\tilde{U}} &\le K\sigma^{-m} \|F\|_{\beta,\tilde{U}} \\
|\hat{F}_{\omega,\varepsilon;k}(p)| &\le K e^{-2\pi\beta|k|} \|F\|_{\beta,\{(\omega,\varepsilon,p)\}}
\end{aligned}
$$

The well known proof is based on expressing the Fourier coefficients or derivatives as integrals over paths and deforming them in the complex domain. It can be found in many reference books and we will not reproduce it here.

3.3. The iterative step

In this subsection, we will specify the iterative step of the algorithm and we develop quantitative estimates that will later lead to the possibility of iterating it and showing it converges. Most of these estimates will be used also in Theorem 1.6 on the partial justification of Greene's criterion.

We recall that for the purposes of the iterative lemma Lemma 3.6, the dimension of the space will be irrelevant, so we will state the results in the $2d$-annulus $\mathbb{R}^d \times \mathbb{T}^d$.

At the beginning of the iterative step, we will be given a family of exact symplectic maps $f_{\omega,\varepsilon}$ defined on a subset of $\mathbb{R}^d \times \mathbb{T}^d$ endowed with the standard symplectic structure.

$$f_{\omega,0}(p,q) = (p, q + \Gamma(\omega,p)) \qquad F_{\omega,\varepsilon}(p,q) = I_{\omega,\varepsilon}(p) + E_{\omega,\varepsilon}(p,q) \tag{3.10}$$

where $F_{\omega,\varepsilon}$, the Hamiltonian of the deformation $f_{\omega,\varepsilon}$, is defined in a set $\Sigma_{\Omega,\alpha,\beta,U}$ of the type described in (3.9), with

$$U = B_{\omega_0,\gamma,\delta} \times I_{[-1,1],\delta} \times B_{0,\gamma,\delta}$$

for some $\gamma > 0$, $0 < \delta < 1$, where ω_0 is a Diophantine vector (e.g. it satisfies (3.8)) and

$$\Omega(\omega,\varepsilon,p) = \Gamma(\omega,p) + \int_0^\varepsilon ds \frac{\partial}{\partial p} I_{\omega,s}(p) \tag{3.11}$$

Since $\Omega(\omega,0,p) = \Gamma(\omega,p)$, from the hypotheses of Theorem 1.5 we will also assume that Ω is non-degenerate, that is, that we have

$$\left\|(\partial_\omega \Omega)^{-1}\right\|_U \leq A, \qquad \left\|(\partial_p^2 \Omega)^{-1}\right\|_U \leq B \tag{3.12}$$

The goal of the iterative step is to determine $g_{\omega,\varepsilon}$, $g_{\omega,0} = \mathrm{Id}$, in such a way that $\tilde{f}_\varepsilon = g_{\omega,\varepsilon}^{-1} \circ f_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}$ has Hamiltonian

$$\tilde{F}_{\omega,\varepsilon}(p,q) = \tilde{I}_{\omega,\varepsilon}(p) + \tilde{E}_{\omega,\varepsilon}(p,q) \tag{3.13}$$

where $\tilde{I}_{\omega,\varepsilon}$, $\tilde{E}_{\omega,\varepsilon}$ will be defined in an slightly smaller domain than $I_{\omega,\varepsilon}$, $E_{\omega,\varepsilon}$ and where $\tilde{E}_{\omega,\varepsilon}$ is much smaller than $E_{\omega,\varepsilon}$ and $\tilde{I}_{\omega,\varepsilon} - I_{\omega,\varepsilon}$ is of the same order of magnitude than $E_{\omega,\varepsilon}$ with all these functions defined in an slightly smaller domain than the original ones.

According to Proposition 2.1, the Hamiltonian of $g_{\omega,\varepsilon}^{-1} \circ f_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}$ is

$$F_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - G_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} + G_{\omega,\varepsilon} \circ f_{\omega,\varepsilon}^{-1} \circ g_{\omega,\varepsilon} \tag{3.14}$$

Heuristically, assuming that $G_{\omega,\varepsilon}$ and $E_{\omega,\varepsilon}$ are small and of the same order — and therefore that $g_{\omega,\varepsilon} - \mathrm{Id}$ and $f_{\omega,\varepsilon} - i_{\omega,\varepsilon}$ are small, where $i_{\omega,\varepsilon}$ is the integrable part of $f_{\omega,\varepsilon}$ as in (3.6) — the main terms in (3.14) are

$$F_{\omega,\varepsilon} - G_{\omega,\varepsilon} + G_{\omega,\varepsilon} \circ i_{\omega,\varepsilon}^{-1}$$

Hence, to make the new error $\tilde{E}_{\omega,\varepsilon}$ zero in this linear approximation, we need to determine $G_{\omega,\varepsilon}$ in such a way that these main terms give just an integrable system (which we will call $\tilde{I}_{\omega,\varepsilon}$). This is formulated as the equation for $G_{\omega,\varepsilon}$, $\tilde{I}_{\omega,\varepsilon}$, given $F_{\omega,\varepsilon}$:

$$\tilde{I}_{\omega,\varepsilon}(p) = F_{\omega,\varepsilon}(p,q) - G_{\omega,\varepsilon}(p,q) + G_{\omega,\varepsilon} \circ i_\omega^{-1}(p,q)$$

Equivalently, we look for an approximate solution of

$$\Delta_{\omega,\varepsilon}(p) = E_{\omega,\varepsilon}(p,q) - G_{\omega,\varepsilon}(p,q) + G_{\omega,\varepsilon} \circ i_\omega^{-1}(p,q) \tag{3.15}$$

where $\Delta_{\omega,\varepsilon}(p) := \tilde{I}_{\omega,\varepsilon}(p) - I_{\omega,\varepsilon}(p)$.

This approximate solution will be used to construct a $g_{\omega,\varepsilon}$, which will lead to a Hamiltonian which is much closer to integrable.

Indeed, the approximate solution of (3.15) will be chosen as an exact solution of

$$\Delta_{\omega,\varepsilon}(p) = E_{\omega,\varepsilon}(p,q) - G_{\omega,\varepsilon}(p,q) + G_{\omega,\varepsilon}(p,q-\omega_0) \tag{3.16}$$

which can be solved by taking Fourier coefficients. We will show that, if we restrict ourselves to a domain $\Sigma_{\Omega,\alpha,\tilde{\beta},\tilde{U}}$, with α very small, the solutions of (3.16) solve (3.15) up to errors that can be controlled by α. Then, the system will be reduced very approximately to a new integrable one. If the frequency function Ω is non-degenerate, we can apply the implicit function theorem and express the domain in terms of the new frequency function $\tilde{\Omega}$. We call attention that it is only in this last step that the non-degeneracy of the frequency function is used.

To justify the above heuristic argument, we will just find the $g_{\omega,\varepsilon}$ obtained by the procedure detailed above and estimate rigorously the remainder after we conjugate the original problem with it. This task will take most of the present section. We will collect all the estimates systematically and, at the end of the section we will formulate the final result precisely. Once we have these results, we will also need to estimate how the integrable part has changed and, in particular, how much the folding surface Σ and its parameterization Υ introduced in (3.3) have changed. This is the task we will undertake in the next section. Then, in a subsequent section, we will show that the procedure can be iterated indefinitely (when some of the arbitrary choices are made appropriately), and that the transformations converge to a limiting transformation that reduces the system to integrable.

3.4. The iterative step. Estimates

In this subsection, we present detailed quantitative estimates for the iterative step that we described informally in the previous section.

Following standard practice, we denote by K sufficiently large positive constants that depend only on the dimension, the number ω_0 and other elements that remain constant during the proof and denote by K^{-1} all sufficiently small positive constants. We will also need to assume that some quantities related to the integrable part of the system remain bounded under the iteration. We will use K_1, K_2 for these constants that depend on the integrable part. The constants K may depend on these K_1, K_2 but not viceversa. When we discuss the iteration, we will see that these K_1, K_2 are chosen in the first step and then, they remain unaltered. In particular, we will need to assume that the constants A and B that quantify the non-degeneracy assumptions (3.12) satisfy

$$A \le K_1, \quad B \le K_2 \tag{3.17}$$

Recall that the goal was, given a Hamiltonian with an error term E, defined in a set $\Sigma_{\Omega,\alpha,\beta,U}$ of the form defined in (3.9), perform a transformation that has an error term $\tilde{E}$ which is much smaller even if defined in a smaller set $\Sigma_{\tilde{\Omega},\tilde{\alpha},\tilde{\beta},\tilde{U}}$.

As it turns out, we will take a number σ and take $\tilde{\beta} = \beta - \alpha - 4\sigma$, $\tilde{U} = U_{4\sigma}$. At the n step σ_n will be $\sigma_0 2^{-n}$ but α will have to decrease super-exponentially.

Our goal will be to show that, under appropriate hypothesis, which we will assume inductively, we can perform the transformation and obtain estimates of the form

$$\|\tilde{E}\|_{\tilde{\Omega},\tilde{\alpha},\tilde{\beta},\tilde{U}} \le K\sigma^{-\tau}\|E\|_{\Omega,\alpha,\beta,U}(\|E\|_{\Omega,\alpha,\beta,U} + \tilde{\alpha}) \tag{3.18}$$

for some fixed positive number τ (we will show later that it suffices to take $\tau = 2\nu + 3$ where $\nu = \theta + d - 1$, and θ is the Diophantine exponent of ω_0).

We will also establish that Υ and $\tilde{\Upsilon}$ — the parameterizations (3.3) of the surfaces $\Omega = \omega_0$ and $\tilde{\Omega} = \omega_0$ respectively —, are defined in very similar domains and differ by an small amount

$$\|\Upsilon - \tilde{\Upsilon}\|_{\tilde{U}} \le K\sigma^{-1}\|E\|_{\Omega,\alpha,\beta,U} \tag{3.19}$$

The proof will be conveniently divided into two parts. In the first one, we obtain estimates in terms of the old domains parameterized by Ω and α. In this first part — culminated in Lemma 3.6 — we will not need to use any non-degeneracy hypothesis in Ω and indeed ω and ε will just go along for the ride. In a second part of the inductive step, we adjust the domains to the new frequency map. This part will require that we assume that Ω is non-degenerate and we will have to lose some domain in ω. This division is natural since the first part is exactly the same as that used in the proof of Theorem 1.6.

Remark. For the experts in KAM theory, we call attention to the fact that the right hand side of (3.18) is not quadratic in $\|E\|_{\Omega,\alpha,\beta,U}$ — the size of the error. Nevertheless, the linear term is multiplied by the number $\tilde{\alpha}$. As we will see in the following subsection, as $\tilde{\alpha}$ goes to zero super-exponentially with the number of steps taken, it is possible to recover the super-exponential convergence of KAM theory that beats the small divisors.

As is customary in KAM theory, in order to be able to carry out the iterative step, we will need to assume that certain quantities are sufficiently small with respect to others — so that for example, compositions have domains that match, implicit function theorems can be applied, etc. As it will turn out all the conditions necessary to perform the iterative step will be implied by smallness conditions of $\|E\|_{\Omega,\alpha,\beta,U}$ with respect to other quantities. Since the iterative step implies that this goes to zero extremely fast, the conditions will be recovered from one step to the next.

Hence, for the proof of Theorem 1.5, the main result of this subsection will be Lemma 3.7 below, which states that, under some explicit conditions, the iterative step can be performed and that the result satisfies (3.18) and (3.19).

Since the proof of Lemma 3.7 will consist in walking through the steps outlined before and just record the conditions needed for them to go through, it is natural to start with the proof of the lemma and postpone its precise statement.

Using Proposition 2.1, the Hamiltonian of $g_{\omega,\varepsilon}^{-1} \circ f_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}$ — if it is possible to define all the compositions — is $I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} + E_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - G_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} + G_{\omega,\varepsilon} \circ f_{\omega,\varepsilon}^{-1} \circ g_{\omega,\varepsilon}$, which adding and subtracting appropriate terms becomes

$$
\begin{aligned}
&\overline{I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}} + (I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - \overline{I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}}) \\
&+ \overline{E_{\omega,\varepsilon}} \\
&+ (E_{\omega,\varepsilon} - \overline{E_{\omega,\varepsilon}}) + (E_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - E_{\omega,\varepsilon}) \\
&- G_{\omega,\varepsilon} + (-G_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} + G_{\omega,\varepsilon}) \\
&+ G_{\omega,\varepsilon} \circ T^0 + (G_{\omega,\varepsilon} \circ i_{\omega,\varepsilon}^{-1} - G_{\omega,\varepsilon} \circ T^0) + (G_{\omega,\varepsilon} \circ f_{\omega,\varepsilon}^{-1} \circ g_{\omega,\varepsilon} - G_{\omega,\varepsilon} \circ i_{\omega,\varepsilon}^{-1})
\end{aligned}
\tag{3.20}
$$

where we have used the notation $\overline{}$ to indicate average over the q variables and $T^0(p,q) = (p, q - \omega_0)$.

The main idea will be to show that it is possible to choose $G_{\omega,\varepsilon}$ in such a way that the first terms in the last three lines of (3.20) add to zero. That is,

$$
E_{\omega,\varepsilon} - \overline{E_{\omega,\varepsilon}} - G_{\omega,\varepsilon} + G_{\omega,\varepsilon} \circ T^0 = 0 \tag{3.21}
$$

and that this $G_{\omega,\varepsilon}$ satisfies estimates which will guarantee that the compositions we used are indeed defined. (We call attention to the fact that (3.21) is the linearized equation that always appears in KAM theory.) Then, the transformed system will have an integrable part $\tilde{I}_{\omega,\varepsilon} = \overline{I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}} + \overline{E_{\omega,\varepsilon}}$ and the other terms appearing in (3.20) will be the error part of the new Hamiltonian. We will estimate them and show that, in a precise sense, they will be smaller than the other ones.

Remark. For the experts in KAM theory, we note that this procedure has two error terms that are linear in G — and hence first order in E —, namely $(G_{\omega,\varepsilon} \circ i_{\omega,\varepsilon}^{-1} - G_{\omega,\varepsilon} \circ T^0)$ and $(I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - \overline{I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}})$ —recall that I will not be converging to zero.

Even if full details will be given later, we advance that for the first term, in the domains that we are considering, $i_{\omega,\varepsilon}^{-1}$ and T^0 are indeed close and the distance is measured by $\tilde{\alpha}$. The mean value theorem will give an estimate that contains the factor $\|E\|\tilde{\alpha}$ multiplied by the small divisors. This is the estimate that appears in one of the terms in (3.18). The second term will turn out to be quadratic because of the fact that

$g_{\omega,\varepsilon}$ is exact symplectic. This is the only place in all the estimates where we use that the maps are exact symplectic.

As usual in KAM theory, we start by obtaining bounds on $G_{\omega,\varepsilon}$ and we will use them to obtain bounds on all the other terms.

Lemma 3.2. *For any $E_{\omega,\varepsilon}(p,q)$ defined in $\Sigma_{\Omega,\alpha,\beta,U}$, we can find unique $\Delta_{\omega,\varepsilon}(p)$, $G_{\omega,\varepsilon}(p,q)$ satisfying*

$$\Delta_{\omega,\varepsilon}(p) = E_{\omega,\varepsilon}(p,q) - G_{\omega,\varepsilon}(p,q) + G_{\omega,\varepsilon}(p,q-\omega_0)$$
$$\int_{\mathbb{T}^d} G_{\omega,\varepsilon}(p,q)\,dq = 0$$

Moreover, these Δ, G satisfy

$$\|G\|_{\beta-\sigma,U} \le K\sigma^{-\nu}\|E\|_{\beta,U}, \qquad \|\Delta\|_{\beta,U} \le \|E\|_{\beta,U} \tag{3.22}$$

where $\nu = \theta + d - 1$.

Proof. The proof is quite standard. We note that integrating in q we have

$$\Delta_{\omega,\varepsilon}(p) = \overline{E_{\omega,\varepsilon}}(p) := \int_{\mathbb{T}^d} dq\, E_{\omega,\varepsilon}(p,q) \tag{3.23}$$

hence, the first estimate in (3.22) follows.

If we take Fourier transforms in the variable q we obtain:

$$\hat{G}_{\omega,\varepsilon;k}(p) = \frac{1}{(e^{-2\pi i k\cdot\omega_0} - 1)}\hat{E}_{\omega,\varepsilon;k}(p) \tag{3.24}$$

By the Cauchy estimates of Lemma 3.1, we have $|\hat{E}_{\omega,\varepsilon;k}(p)| \le Ke^{-2\pi\beta|k|}\|E\|_{\beta,U}$ and, by the Diophantine assumptions, $|e^{-2\pi ik\omega_0} - 1|^{-1} \le C|k|^{\theta-1}$. Hence,

$$|\hat{G}_{\omega,\varepsilon;k}(p)| \le K|k|^{\theta-1}e^{-2\pi\beta|k|}\|E\|_{\beta,U}$$

and, therefore

$$\begin{aligned}\|G\|_{\beta-\sigma,U} &\le \sum_{k\in\mathbb{Z}^d}|\hat{G}_{\omega,\varepsilon;k}(p)|e^{2\pi(\beta-\sigma)|k|} \le K\left(\sum_{k\in\mathbb{Z}^d}|k|^{\theta-1}e^{-2\pi\sigma|k|}\right)\|E\|_{\beta,U}\\ &\le K\left(\sum_{l\in\mathbb{N}}|l|^{\theta-1+d-1}e^{-2\pi\sigma l}\right)\|E\|_{\beta,U} \le K\sigma^{-\nu}\|E\|_{\beta,U}\end{aligned}$$

where $\nu = \theta + d - 1$.

We refer to [SM] for more details but point out that it is possible to obtain better exponents in σ (see e.g., [Ru]). Of course, since the rest of the proof goes through for any exponent, this does not affect the subsequent reasoning.

■

A small generalization of these estimates is:

Proposition 3.3. *With the notation of Lemma 3.2*

$$\|\nabla^m G\|_{\beta-\sigma,U_\sigma} \le K\sigma^{-\nu-m}\|E\|_{\beta,U} \tag{3.25}$$

Proof. Using Lemma 3.1 and (3.24) we obtain that, for $(\omega,\varepsilon,p)\in U_\sigma$, we have

$$|\partial_p^i\hat{G}_{\omega,\varepsilon;k}(p)| \le K\sigma^{-(i+\theta-1)}e^{-2\pi|k|\beta}\|E\|_{\beta,U} \tag{3.26}$$

Similarly, we have

$$|\partial_q^j(\hat{G}_{\omega,\varepsilon;k}(p)e^{2\pi i k\cdot q})| \le K|k|^j|\hat{G}_{\omega,\varepsilon;k}(p)| \le K|k|^{j+\theta-1}e^{-2\pi|k|\beta}\,\|E\|_{\beta,U} \tag{3.27}$$

On $\Sigma_{\beta-\sigma,U_\sigma}$ we have $|\mathrm{Im}q| \le \beta-\sigma$ and hence $|e^{2\pi i k\cdot q}| \le e^{2\pi|k|(\beta-\sigma)}$. Therefore, using the above estimates (3.26) and (3.27) in the same way as in Lemma 3.2, we obtain the desired result.

■

Now, we can prove estimates for the flow of $G_{\omega,\varepsilon}$

Proposition 3.4. *Assume that the conditions of Proposition 3.3 are met and that, furthermore*

$$K\sigma^{-\nu-1}\,\|E\|_{\beta,U} \le \sigma/2 \tag{3.28}$$

Then:

i) for $(\omega,\varepsilon,p,q) \in \Sigma_{\beta-2\sigma,U_{2\sigma}}$, the flow $g_{\omega,\varepsilon}(p,q)$ generated by the Hamiltonian $G_{\omega,\varepsilon}$ is well defined, and $(\omega,\varepsilon,g_{\omega,\varepsilon}(p,q)) \in \Sigma_{\beta-\sigma,U_\sigma}$

ii) $\|g-\mathrm{Id}\|_{\beta-2\sigma,U_{2\sigma}} \le \|\nabla G\|_{\beta-\sigma,U_\sigma} \le K\sigma^{-\nu-1}\,\|E\|_{\beta,U}$

Proof. It follows from hypothesis (3.28), Proposition 3.3 and the local existence theorem for solutions of O.D.E.'s.

■

From now on, we will assume that (3.28) holds, and we will proceed to estimate the terms in (3.20).

By Proposition 3.4, the compositions $G_{\omega,\varepsilon}\circ g_{\omega,\varepsilon}$, $E_{\omega,\varepsilon}\circ g_{\omega,\varepsilon}$ are well defined on $\Sigma_{\beta-2\sigma,U_{2\sigma}}$. Using the mean value theorem and Cauchy inequalities from Lemma 3.1, we can bound

$$\|G-G\circ g\|_{\beta-2\sigma,U_{2\sigma}} \le \|\nabla G\|_{\beta-\sigma,U_\sigma}\ \|g-\mathrm{Id}\|_{\beta-2\sigma,U_{2\sigma}} \le K\sigma^{-2\nu-2}\,\|E\|_{\beta,U}^2 \tag{3.29}$$

$$\|E-E\circ g\|_{\beta-2\sigma,U_{2\sigma}} \le \|\nabla E\|_{\beta-\sigma,U_\sigma}\ \|g-\mathrm{Id}\|_{\beta-2\sigma,U_{2\sigma}} \le K\sigma^{-\nu-2}\,\|E\|_{\beta,U}^2 \tag{3.30}$$

These estimates show that two of the terms in (3.20) are quadratically small in the original error.

Now, we turn to estimate the last term in (3.20), which, as we will show, will also be quadratic in $\|E\|$. The reason is that $f_{\omega,\varepsilon}$ and $i_{\omega,\varepsilon}$ satisfy differential equations whose difference can be controlled by $\|E\|$ and the same initial conditions. Hence, $\|f^{-1}-i^{-1}\| \le K\|E\|$ under some mild extra assumptions that guarantee that domains match etc., and we can now apply the mean value theorem. The precise details are a walk through the standard proof of the existence and uniqueness for O.D.E.'s, as we detail below.

First, we recall that $i_{\omega,\varepsilon}$ has the form (3.6): $i_{\omega,\varepsilon}(p,q)=(p,q+\Omega(\omega,\varepsilon,p))$, with $\Omega(\omega,\varepsilon,p)$ given in (3.11), and we note that $i_{\omega,\varepsilon}^{-1}(p,q)=(p,q-\Omega(\omega,\varepsilon,p))$. Hence, for

$$\|i-T_0\|_U = \left\|i^{-1}-T_0\right\|_U = \|\Omega-\omega_0\|_U \le \alpha \tag{3.31}$$

we have

$$(\omega,\varepsilon,p,q)\in\Sigma_{\beta-\alpha,U} \Longrightarrow \left(\omega,\varepsilon,i_{\omega,\varepsilon}(p,q)\right), \left(\omega,\varepsilon,i_{\omega,\varepsilon}^{-1}(p,q)\right) \in \Sigma_{\beta,U} \tag{3.32}$$

Assuming

$$\left\|\frac{\partial\Omega}{\partial p}\right\|_U \le K_3 \tag{3.33}$$

(where without loss of generality, we assume, to simplify some formulas that $K_3>1$), we can bound

$$\|\nabla i\|_U = \left\|\nabla i^{-1}\right\|_U \le K \tag{3.34}$$

We recall now that $f_{\omega,\varepsilon}$ is the solution of

$$
\begin{aligned}
f_{\omega,\varepsilon}(x) &= f_{\omega,0}(x) + \int_0^\varepsilon ds\, \mathcal{F}_{\omega,s} \circ f_{\omega,s}(x) \\
&= f_{\omega,0}(x) + \int_0^\varepsilon ds\, \left[\mathcal{I}_{\omega,\varepsilon} \circ f_{\omega,s}(x) + \mathcal{E}_{\omega,\varepsilon} \circ f_{\omega,s}(x)\right]
\end{aligned} \tag{3.35}
$$

while $i_{\omega,\varepsilon}$ satisfies $i_{\omega,\varepsilon}(x) = i_{\omega,0}(x) + \int_0^\varepsilon ds\, \mathcal{I}_{\omega,s} \circ i_{\omega,s}(x)$, with $f_{\omega,0}(x) = i_{\omega,0}(x)$. By hypothesis (3.28), using standard arguments of O.D.E.'s based on the Gronwall inequality, we get that for $(\omega, \varepsilon, p, q) \in \Sigma_{\beta-\alpha-2\sigma, U_{2\sigma}}$, the flow $f_{\omega,\varepsilon}(p,q)$ is well defined, and satisfies

$$
(\omega, \varepsilon, p, q) \in \Sigma_{\beta-\alpha-2\sigma, U_{2\sigma}} \Longrightarrow \left(\omega, \varepsilon, f_{\omega,\varepsilon}(p,q)\right) \in \Sigma_{\beta-\sigma, U_\sigma} \tag{3.36}
$$

$$
\|f - i\|_{\beta-\alpha-2\sigma, U_{2\sigma}} \le e^{K_3} \|\nabla E\|_{\beta-\sigma, U_\sigma} \le K\sigma^{-1} \|E\|_{\beta,U} \tag{3.37}
$$

From (3.34), and Lemma 3.1 applied to (3.37), we can bound $\nabla f_{\omega,\varepsilon}$:

$$
\|\nabla f\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le \|\nabla i\|_U + \|\nabla(f - i)\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le K \tag{3.38}
$$

Applying the Implicit Function Theorem to the estimates above, it turns out that for $(\omega, \varepsilon, p, q) \in \Sigma_{\beta-\alpha-2\sigma, U_{2\sigma}}$, $f^{-1}_{\omega,\varepsilon}(p,q)$ is well defined, satisfies $\left(\omega, \varepsilon, f^{-1}_{\omega,\varepsilon}(p,q)\right) \in \Sigma_{\beta-\sigma, U_\sigma}$ and

$$
\left\|f^{-1} - i^{-1}\right\|_{\beta-\alpha-2\sigma, U_{2\sigma}} \le K\sigma^{-1}\| \, \|E\|_{\beta,U} \tag{3.39}
$$

As before, from (3.34), and Lemma 3.1 applied to (3.39), we can bound $\nabla f^{-1}_{\omega,\varepsilon}$:

$$
\left\|\nabla f^{-1}\right\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le K \tag{3.40}
$$

Using the mean value theorem, (3.40) and the bounds on $g_{\omega,\varepsilon} - \mathrm{Id}$ established in Proposition 3.4, we obtain:

$$
\left\|f^{-1} - f^{-1} \circ g\right\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le K\sigma^{-\nu-1} \|E\|_{\beta,U} \tag{3.41}
$$

Putting together (3.39) and (3.41), by the triangle inequality, we obtain

$$
\left\|f^{-1} \circ g - i^{-1}\right\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le K\sigma^{-\nu-1} \|E\|_{\beta,U} \tag{3.42}
$$

Using the mean value theorem, the estimates in Proposition 3.3 and (3.42), we can bound the last term in (3.20) as

$$
\left\|G \circ f^{-1} \circ g - G \circ i^{-1}\right\|_{\beta-\alpha-3\sigma, U_{3\sigma}} \le K\sigma^{-2\nu-2} \|E\|^2_{\beta,U} \tag{3.43}
$$

Now, we turn our attention to the first term in (3.20). It will depend on the approximate expression $g^0_{\omega,\varepsilon} = \mathrm{Id} + \int_0^\varepsilon ds\, \mathcal{G}_{\omega,s}$ for $g_{\omega,\varepsilon}$:

$$
g^0_{\omega,\varepsilon}(p,q) = \left(p - \int_0^\varepsilon ds\, \frac{\partial}{\partial q} G_{\omega,s}(p,q) \ , q + \int_0^\varepsilon ds\, \frac{\partial}{\partial p} G_{\omega,s}(p,q)\right) \tag{3.44}
$$

Proposition 3.5. *Under our standing hypotheses, we have*

$$
\left\|g - g^0\right\|_{\beta-2\sigma, U_{2\sigma}} \le K\sigma^{-2\nu-3} \|E\|^2_{\beta,U}
$$

Proof. Note that our standing assumptions imply

$$\left\|g^0 - \mathrm{Id}\right\|_{\beta-\sigma,U_\sigma} \le \|\nabla G\|_{\beta-\sigma,U_\sigma} \le K\sigma^{-\nu-1}\|E\|_{\beta,U}$$

and consequently $\left(\omega,\varepsilon,g^0_{\omega,\varepsilon}(p,q)\right) \in \Sigma_{\beta-\sigma,U_\sigma}$ for $(\omega,\varepsilon,p,q) \in \Sigma_{\beta-\alpha-2\sigma,U_{2\sigma}}$.

We can write $g_{\omega,\varepsilon}$ as the solution of a fixed point problem. Namely,

$$g_{\omega,\varepsilon} = \mathrm{Id} + \int_0^\varepsilon ds\, \mathcal{G}_{\omega,s} \circ g_{\omega,s} \equiv \mathcal{T}(g)_{\omega,\varepsilon}$$

and we have the identity

$$\mathcal{T}(g^0)_{\omega,\varepsilon} - g^0_{\omega,\varepsilon} = \int_0^\varepsilon ds\, [\mathcal{G}_{\omega,s} \circ g^0_{\omega,s} - \mathcal{G}_{\omega,s}]$$

If we estimate the integrand of the R.H.S. by the mean value theorem, we have

$$\left\|\mathcal{T}(g^0) - g^0\right\|_{\beta-2\sigma,U_{2\sigma}} \le \left\|\nabla^2 G\right\|_{\beta-\sigma,U_\sigma} \left\|g^0 - \mathrm{Id}\right\|_{\beta-\sigma,U_\sigma} \le K\sigma^{-2\nu-3}\|E\|^2_{\beta,U} \tag{3.45}$$

We also obtain, under (3.28), that $\mathcal{T}$ is a contraction of factor $1/2$. Hence, there is a fixed point of $\mathcal{T}$ whose distance from $g^0_{\omega,\varepsilon}$ is not bigger than $1/(1-1/2) = 2$ times the R.H.S. of (3.45).

■

We note that, because $I_{\omega,\varepsilon}(x)$ does not depend on q, denoting by Π_p, Π_q the projections on the p and q components respectively, we have for $x = (p,q)$

$$\begin{aligned} I_{\omega,\varepsilon}(g_{\omega,\varepsilon}(x)) &= I_{\omega,\varepsilon}(\Pi_p g_{\omega,\varepsilon}(x)) \\ &= I_{\omega,\varepsilon}(p) + \partial_p I_{\omega,\varepsilon}(p)\Pi_p\left[g_{\omega,\varepsilon}(x) - x\right] + R_2\left(\omega,\varepsilon,x,g_{\omega,\varepsilon}(x)\right) \\ &= I_{\omega,\varepsilon}(p) + \partial_p I_{\omega,\varepsilon}(p)\Pi_p\left[g^0_{\omega,\varepsilon}(x) - x\right] \\ &\quad + \partial_p I_{\omega,\varepsilon}(p)\Pi_p\left[g_{\omega,\varepsilon}(x) - g^0_{\omega,\varepsilon}(x)\right] + R_2\left(\omega,\varepsilon,x,g_{\omega,\varepsilon}(x)\right) \end{aligned} \tag{3.46}$$

where we have denoted by R_2 the remainder of the second order Taylor expansion in p.

Note that $\Pi_p\left[g^0_{\omega,\varepsilon}(x) - x\right] = \partial_q G_{\omega,\varepsilon}(x)$ (see (3.44)) and that $\overline{\partial_q G_{\omega,\varepsilon}} = 0$ since q is a periodic variable. Hence, observing that $\partial_p I$ is independent of q, we obtain

$$\overline{\partial_p I \int_0^\varepsilon ds\, \partial_q G_{\omega,s}} = 0 \tag{3.47}$$

That is, the second term in the R.H.S. of the formula of (3.46) has zero average. We call attention to the fact that this is the only part in the whole proof of the estimates where we use the exact symplectic character of the deformation, which is equivalent to the fact that G is a function on the annulus and not just on the universal cover.

Since $I_{\omega,\varepsilon}$ depends only on p we have that $\overline{I_{\omega,\varepsilon}} = I_{\omega,\varepsilon}$.

Under the assumption

$$\left\|\nabla^2 I\right\|_{\beta-\sigma,U_\sigma} \le K_4 \tag{3.48}$$

we can bound the last two terms in (3.46) by terms that are quadratic in $\|E\|$.

Since the last two terms in (3.46) are the only ones that contribute to $I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon} - \overline{I_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}}$, we obtain from Proposition 3.5

$$\left\|I \circ g - \overline{I \circ g}\right\|_{\beta-2\sigma,U_{2\sigma}} \le K\sigma^{-2\nu-3}\|E\|^2_{\beta,U} \tag{3.49}$$

The only term in (3.20) that remains to be estimated is $G_{\omega,\varepsilon} \circ i_{\omega,\varepsilon}^{-1} - G_{\omega,\varepsilon} \circ T^0$. We note that, by (3.31), we have

$$\left\| i^{-1} - T^0 \right\|_U \leq \alpha$$

Therefore, using the estimates in Proposition 3.3

$$\left\| G \circ i^{-1} - G \circ T^0 \right\|_{\sigma, U_{\beta-\alpha-\sigma}} \leq K\sigma^{-\nu-1} \alpha \left\| E \right\|_{\beta,U} \tag{3.50}$$

If we add the estimates in (3.29), (3.30), (3.43), (3.49) and (3.50), for the terms that has to be bounded in (3.20), and claim them only in the domain $\Sigma_{\beta-\alpha-4\sigma, U_{4\sigma}}$, which is smaller than any of the domains in which we have bounds, we obtain

$$\left\| \tilde{E} \right\|_{\beta-\alpha-4\sigma, U_{4\sigma}} \leq K\sigma^{-\tau} \left\| E \right\|_{\beta,U} \left(\left\| E \right\|_{\beta,U} + \alpha \right) \tag{3.51}$$

where $\tau := 2\nu + 3$ and $\| \Omega - \omega_0 \|_U \leq \alpha$.

We also notice that from Proposition 3.4 and (3.36), it follows that if $(\omega, \varepsilon, p, q) \in \Sigma_{\beta-\alpha-4\sigma, U_{4\sigma}}$, then $\left(\omega, \varepsilon, g_{\omega,\varepsilon}^{-1} \circ f_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}(p, q)\right) \in \Sigma_{\beta-\sigma, U_\sigma}$.

On the set $\Sigma_{\Omega,\alpha,\beta\alpha-4\sigma,U_{4\sigma}}$ introduced in (3.9), equation (3.51) reads as

$$\|\tilde{E}\|_{\Omega,\alpha,\beta-\alpha-4\sigma,U_{4\sigma}} \leq K\sigma^{-\tau} \|E\|_{\Omega,\alpha,\beta,U} (\|E\|_{\Omega,\alpha,\beta,U} + \alpha) \tag{3.52}$$

This is very similar to the estimates desired in (3.18) and it only differs from them in the fact that the norm in the L.H.S. of (3.52) is referred to the domain specified by Ω and not by $\tilde{\Omega}$.

To remedy that, we will estimate the change in Ω and the attendant change in the parameterizations Υ of the surface and the domain Σ. Using that the frequency function Ω is non-degenerate, this will allow us to transform the expression of the domain in which we have improved estimates into an expression involving the new frequency function.

We will find it convenient to state formally what we have already accomplished without using non-degeneracy conditions in Ω. We call attention that this lemma will also play an important role in the proof of Theorem 1.6. Later, we will prove Lemma 3.7 that takes into account the change in the frequency function and which indeed uses the non-degeneracy assumptions in Ω.

Lemma 3.6. *Given the Hamiltonian $F = I + E$ of $f_{\omega,\varepsilon}$ introduced in (3.10), choose G, Δ as given by Lemma 3.2, and consider the new Hamiltonian $\tilde{F} = \tilde{I} + \tilde{E}$ of $g_{\omega,\varepsilon}^{-1} \circ f_{\omega,\varepsilon} \circ g_{\omega,\varepsilon}$ as given in (3.13). Assume that σ is such that (3.28), (3.33), and (3.48) are met, and let $\tau = 2\nu + 3$. Then*

$$\|\tilde{E}\|_{\Omega,\alpha,\beta-\alpha-4\sigma,U_{4\sigma}} \leq K\sigma^{-\tau} \|E\|_{\Omega,\alpha,\beta,U} (\|E\|_{\Omega,\alpha,\beta,U} + \alpha) \tag{3.53}$$

$$\|\Delta\|_{\Omega,\alpha,\beta,U} \leq \|E\|_{\Omega,\alpha,\beta,U}, \qquad \|\nabla\Delta\|_{\Omega,\alpha,\beta-\sigma,U_\sigma} \leq K\sigma^{-1} \|E\|_{\Omega,\alpha,\beta,U} \tag{3.54}$$

The way of interpreting these estimates is that (3.53) indicates that, after the transformation, the resulting Hamiltonian is essentially an integrable one (albeit in a smaller domain): the right hand side of (3.53) consists on two terms, one of which is quadratic in $\|E\|$ and the other one contains $\|E\|\alpha$. If we choose α sufficiently small, we will be able to make the right hand side of (3.53) much smaller than the original one. This will overcome the small divisors $\sigma^{-\tau}$.

We call attention to the fact that Lemma 3.6 does not need the non-degeneracy assumption on Ω and that does not lose any domain in the parameters. This lemma will a basic tool for the estimates of the inductive steps both in the proof of the KAM theorem and in the justification of Greene's criterion. The difference between the two results will be that that the inductive steps will have different domain loses and that we will have to apply them repeatedly in different ways, losing domain at different rates.

3.5. The KAM inductive step. Geometry of domains

To complete the work for the bounds of the inductive step in the KAM theorem, we need to study the change in Ω, the surface Σ defined by $\Omega = \omega_0$ and its natural parameterization Υ defined in (3.3). In particular, we will need to provide estimates for the changes of the bounds in (3.12) that quantify the non-degeneracy assumptions. Since we are also taking into account the derivative of Ω with respect to ω, instead of (3.33), we are going to assume:

$$\left\| \frac{\partial \Omega}{\partial p} \right\|_U \leq K_3, \qquad \left\| \frac{\partial \Omega}{\partial \omega} \right\|_U \leq K_3 \tag{3.55}$$

Again, we emphasize that most of the results in this section are true for arbitrary d. The only exception is *iv)* in Lemma 4.7 below.

Given the estimates that we have on Δ, it will be very easy to estimate the change in Ω and all the other estimates will follow by an application of the implicit function theorem. We note that since Δ is small, and Ω depends linearly on the integrable part, the change in Ω will be of the same order of magnitude and hence also small. All the changes in the surface and in the parameterization will be small and hence can be estimated by $\|E\|$ possibly multiplied by some factors that come from the fact that we have to involve derivatives and control them by Cauchy estimates.

More precisely, we have:

Lemma 3.7. *Let Ω be the frequency function (3.11) for the family $f_{\omega,\varepsilon}$ (3.10) defined on $\Sigma_{\Omega,\alpha,\beta,U}$ as in (3.9). Let Δ be given by (3.23) and let σ be a positive number. Assume that (3.17), (3.28), (3.55), and (3.48), hold. Consider $\tilde{\Omega}$, the new frequency function defined by*

$$\tilde{\Omega}(\omega, \varepsilon, p) = \Omega(\omega, \varepsilon, p) + \int_0^\varepsilon ds \, \frac{\partial}{\partial p} \Delta(\omega, s, p) \tag{3.56.}$$

Denote by Υ and $\tilde{\Upsilon}$ the parameterizations (3.3) corresponding to Ω and $\tilde{\Omega}$.

Then, for any $\tilde{\alpha} \leq \alpha$ satisfying

$$K \sigma^{-1} \|E\|_{\Omega,\alpha,\beta,U} \leq \tilde{\alpha} \tag{3.57}$$

we have:

i) $\left\| \Omega - \tilde{\Omega} \right\|_{U_\sigma} \leq K\sigma^{-1} \|E\|_{\Omega,\alpha,\beta,U} \leq \tilde{\alpha}$

ii) *For $\tilde{\alpha}$ as before, $\tilde{\beta} = \beta - \alpha - 4\sigma$, $\tilde{U} \equiv U_{4\sigma}$, we have:*

$$\Sigma_{\tilde{\Omega},\tilde{\alpha},\tilde{\beta},\tilde{U}} \subset \Sigma_{\Omega,2\alpha,\beta-4\sigma,U_{4\sigma}}$$

iii)

$$\left\| \left(\frac{\partial}{\partial \omega} \tilde{\Omega} \right)^{-1} \right\|_{\tilde{U}} \leq \left\| \left(\frac{\partial}{\partial \omega} \tilde{\Omega} \right)^{-1} \right\|_{U_{4\sigma}} \leq \left\| \left(\frac{\partial}{\partial \omega} \Omega \right)^{-1} \right\|_U + K\sigma^{-2} \|E\|_{\Omega,\alpha,\beta,U}$$

iv) *When $d = 1$,*

$$\left\| \left(\frac{\partial^2}{\partial p^2} \tilde{\Omega} \right)^{-1} \right\|_{\tilde{U}} \leq \left\| \left(\frac{\partial^2}{\partial p^2} \tilde{\Omega} \right)^{-1} \right\|_{U_{4\sigma}} \leq \left\| \left(\frac{\partial^2}{\partial p^2} \Omega \right)^{-1} \right\|_U + K\sigma^{-3} \|E\|_{\Omega,\alpha,\beta,U}$$

v) $\|\Upsilon - \tilde{\Upsilon}\|_{\tilde{U}} \leq K\sigma^{-1} \|E\|_{\Omega,\alpha,\beta,U}$

vi) *The inequalities (3.18) hold. That is, for $\tau = 2\nu + 3$*

$$\|\tilde{E}\|_{\tilde{\Omega},\tilde{\alpha},\tilde{\beta},\tilde{U}} \leq K\sigma^{-\tau} \|E\|_{\Omega,\alpha,\beta,U} (\|E\|_{\Omega,\alpha,\beta,U} + \tilde{\alpha})$$

Proof. Part i) follows immediately from the formula (3.56) for $\tilde{\Omega}$ and the estimates that we have for Δ in Lemma 3.2. The last inequality in i) is just a restatement of (3.28), which is one of the hypothesis of the lemma.

Part ii) follows because of (3.57).

Parts iii) and iv) follow because we can use Cauchy estimates to estimate the derivatives of Δ. Then, we can use Cauchy estimates to bound the derivatives of Ω.

The existence of $\tilde{\Upsilon}$ and its estimates are a very simple consequence of the implicit function theorem. Recall the well known result that if an analytic function Φ satisfies $|\Phi(0)| \le \varepsilon$ and $|\Phi'|^{-1} \le a$ on a ball around zero of radius $a\varepsilon$ there is one and only one zero in this ball. Moreover, if Φ depends analytically on parameters, the zero depends analytically on parameters. We can apply this result to $\Phi(s) = \Omega(s + \Upsilon(\varepsilon, p), \varepsilon, p) - \omega_0$ and then, the result follows.

Part vi) is a consequence of the estimates in Lemma 3.6 and part ii) of this Lemma.

∎

Notice that the only places where we had to consider derivatives with respect to ω are iii) and v). Hence, this will be easy to adapt to the situation in the justification of the Greene's criterion where there is some degeneracy in the frequency function.

Remark. Notice also that it is only in these non-degeneracy assumptions that we have to consider the one-dimensional properties of the map. It seems that with some appropriate notion of critical circle in higher dimensions (one has to consider invariant tori with 'degenerate torsion'), one could develop an analogous converging KAM process, and a subsequent geometrical interpretation could provide the structure of invariant objects nearby the critical torus.

3.6. Iteration of the KAM inductive step. Convergence

In this subsection, we verify that if we start with a sufficiently small perturbation E, the iterative step can be repeated infinitely many times and, moreover, converges to a solution. The estimates are very similar to those in the paper [Ru2] on the translated curve method. Again, we will assume that $d = 1$.

The main idea is that the loss of domain has to be fast — say exponentially fast — in the variables q so that we have some domain left. On the other hand, we have to decrease super-exponentially fast the variable α which controls the thickness of the approximations to the surface Σ. This will achieve that the $\|E\|$ decreases super-exponentially and that, as a consequence, the process can be iterated indefinitely.

We will choose α_n, σ_n, and show that if $\|E^0\|_{\Omega^0,\alpha_0,\beta_0,U^0}$ is small enough, the iterative step described in the previous section can be repeated indefinitely and the transformations converge to a solution that indeed solves the problem.

We point out that these smallness conditions can always be adjusted by switching to another variable $\varepsilon' = \varepsilon\lambda$. If we choose λ small enough, the remainder is made arbitrarily small while all the other parameters in the problem are left unaltered. (That is, when we have families, we can obtain the smallness conditions by considering ε restricted to an small domain.) Of course, when our families are obtained by interpolating between two diffeomorphisms, as in Lemma 2.2, the smallness assumptions in the family can be accomplished by assuming that the diffeomorphisms we are interpolating are close.

We will start by picking $\sigma_n = \sigma^* 2^{-n}$, where we pick $\sigma^* < \beta_0/8$ so that β_n defined in Lemma 3.7 by $\beta_{n+1} = \beta_n - 4\sigma_n$ is bounded away from zero, and $\sigma^* < \delta/8$ so that all the domains $U^{n+1} = U^n_{\sigma_n}$ contain the open domain $U^0_{2\sigma^*}$. Now, we will show that it is possible to choose α_n in such a way that if $\|E^0\|_{\Omega^0,\alpha_0,\beta_0,U^0}$ is small enough, the process can be iterated indefinitely and it converges.

Introducing the notation $e_n = \|E_n\|_{\Omega^n,\alpha_n,\beta_n,U^n}$, $a_n = \alpha_{n+1}$, $A = 2^\tau$, $C = K/\sigma^*$, the recursion equation in vi) of Lemma 3.7 becomes

$$e_{n+1} \le CA^n\, e_n(e_n + a_n) \tag{3.58}$$

We claim that

Lemma 3.8. *If e_0 is small enough, it is possible to choose $0 < \rho < 1$ in such a way that setting $a_n = \rho^{2^n}(AB)^{-n}$, for $B > 1$, the conditions for Lemma 3.7 are satisfied for all n and*

$$e_n \le \frac{a_n}{C\,2^n} = \frac{\rho^{2^n}}{C(2AB)^n}. \tag{3.59}$$

Proof. Assume that (3.59) holds for a certain n and that we have chosen a_n as indicated and that the iterative step can be applied at this step.

Then, by (3.58) we have

$$\begin{aligned} e_{n+1} &\le \frac{\rho^{2^n}}{C(2AB)^n}\left(\frac{\rho^{2^n}}{C(2AB)^n} + \frac{\rho^{2^n}}{(AB)^n}\right) CA^n \\ &= \frac{\rho^{2^{n+1}}}{C(2AB)^{n+1}}\frac{2ABC}{B^n}\left(\frac{1}{C\,2^n}+1\right) \le \frac{\rho^{2^{n+1}}}{C(2AB)^{n+1}}\frac{4ABC}{B^n}. \end{aligned} \tag{3.60}$$

If $n > N_0(A, B, C)$ we have that

$$\frac{4ABC}{B^n} \le 1 \tag{3.61}$$

so that indeed the formula (3.59) holds for $n + 1$.

We also observe that, if a_n and e_n are of the form that we claimed, there is an $N_1(A, B, C) \ge N_0$ so that all the hypotheses (3.17), (3.28), (3.55), (3.48), (3.57) are satisfied for $n > N_1$.

Therefore, it suffices to ensure that e_0 is so small that the iterative step can be performed N_1 times and that the inequalities (3.59) hold for $n \le N_1$. Then, the argument in (3.60) will show that (3.59) continue to hold, and that the hypotheses needed to perform the iterative step and (3.61) hold.

■

Clearly, from (3.59), we obtain that the error of the solution goes to zero on the surfaces. Similarly, using the estimates in Lemma 3.7 we can show that the parameterizations Υ of the surface converge. (It suffices to check that the increments are summable.)

Moreover, defining $h^n_{\omega,\varepsilon} = g^0_{\omega,\varepsilon} \circ \cdots \circ g^n_{\omega,\varepsilon}$ we have that

$$\begin{aligned} \|h^n - h^{n-1}\|_{\Omega^n,\alpha_n,\beta_n,U^n} &= \|h^{n-1}\circ g^n - h^{n-1}\|_{\Omega^n,\alpha_n,\beta_n,U^n} \\ &\le \sigma_{n-1}^{-1} K \|h^{n-1}\|_{\Omega^{n-1},\alpha_{n-1},\beta_{n-1},U^{n-1}} \|g^n - \mathrm{Id}\,\|_{\Omega^n,\alpha_n,\beta_n,U^n} \end{aligned} \tag{3.62}$$

From (3.62) and the estimates in *ii)* of Proposition 3.4, it is immediate to show by induction that $\|h^n\|_{\Omega^n,\alpha_n,\beta_n,U^n}$ remains bounded independently of n. Then, using *ii)* of Proposition 3.4, the RHS of (3.62) is summable in n. Hence $h^n_{\omega,\varepsilon}$ converges in the limiting domain $\Sigma_{\Omega^\infty,\alpha_\infty,\beta_\infty,U^\infty}$, with $\alpha_\infty = 0$, consisting on the points $(\omega, \varepsilon, p, q)$ with $(\omega, \varepsilon, p) \in U^\infty = U^0_{2\sigma^*}$ such that $\Omega^\infty(\omega, \varepsilon, p) = \omega_0$ and $|\mathrm{Im} q| \le \beta_\infty = \beta_0 - 2\sigma^* \ge \beta_0/2$.

This finishes the proof of Theorem 1.5.

4. Partial justification of Greene's criterion

To assess numerically the existence of invariant circles, the most frequently used method is the so-called Greene's criterion, formulated in [Gr] for two dimensional maps.

This criterion asserts that a smooth invariant circle with motion smoothly conjugate to a rotation ω exists if and only if it is possible to find a sequence of periodic orbits of type m/n whose "residue" (that is, the trace of the derivative of the return map minus 2) converges to zero as the n/m converges to ω_0.

As it turns out, this criterion has not been proved to hold, nevertheless, parts of it can be established rigorously.

For standard KAM tori, Mather (see [McK] Section 1.3.2.4) suggested a method to prove that if KAM tori existed, the residue should go to zero faster than any power of $|\omega - p_n/q_n|$. This method was implemented in [FL,McK2] for two dimensional maps to show that the residue is smaller than $\exp(-c|\omega - p_n/q_n|^{-\alpha})$ for some $\alpha > 0$.

The main goal of this section is to prove one of the implications of Greene's criterion for critical circles. We will prove that if a critical circle exists, then any sequence of periodic orbits converging to it has residual converging to zero. We will also show that, if a critical circle exists, indeed there is at least one such sequence. Actually, for any p/q such that $p/q < \omega, |p/q - \omega| \ll 1$, we can find at least 2 periodic orbits of type p/q and, under mild non-degeneracy conditions, at least 4.

Again, we will assume in this section that $d = 1$. We note that for higher dimensional maps, in [T1] and [T2] there are versions of Greene's criterion for higher dimensional twist maps (a rigorous justification of one of the implications and numerical evidence respectively). There are some differences between the proofs in higher dimensional cases and the case considered here of $d = 1$ and we will comment on them after the proof of our results.

The main part of the proof will consist in showing that, in a neighborhood of the invariant circle, it is possible to find changes of variables that reduce the system almost to integrable. Once we have that, the result will follow word for word the result in [FL].

Of course, the estimates near the invariant torus are a more general result than that of the Greene's criterion and they allow to control not only the behavior of the periodic orbits, but also other dynamical objects. Other papers in which similar estimates are obtained for non-degenerate circles are [OS,PW,JV,DG2].

Most of the work has been done already in Section 3. The estimates that we will use are the same as those of the iterative step and the only difference is that we will be in the iterative step that makes different choices. This unified approach between KAM theorem and exponentially small estimates appears also in [DG1].

4.1. Preliminary estimates and notation.

We will be considering area preserving maps f which are defined in a neighborhood of $[-\delta, \delta] \times \mathbb{T}^1$ to itself.

These maps will have the form

$$f(p, q) = (p, q + \omega_0 + \kappa p^M) + O(p^{M+1})$$

for some $\kappa \neq 0$.

By Lemma 2.2, we can find an f_ε in such a way that the $f_0(p, q) = (p, q + \omega_0 + \kappa p^M)$. The Hamiltonian of this deformation will be $F_\varepsilon = O(p^{M+1})$.

We will write for these type of families $F_\varepsilon(p, q) = I_\varepsilon(p) + E_\varepsilon(p, q)$, where again I_ε will be thought of as the integrable part. We will denote by i_ε the deformation with initial point f_0 and with Hamiltonian I_ε: $i_\varepsilon(p, q) = \big(p, q + \omega_0 + \kappa p^M + \int_0^\varepsilon ds\, \partial_p I_s(p)\big)$.

We note that these families are a particular case of the families we have considered in Section 3. (In particular, $\kappa < 0$ and $M = 2$ for the example (1.1).) In that section, we allowed a dependence in another parameter ω. The families we consider here can be considered as embedded in families depending on ω but such that the dependence on ω is trivial. Clearly, all the results of Section 3 that do not rely on the dependence on ω being non-trivial will go through as stated using the elementary device of writing the extra variable ω and noticing that the functions we consider do not depend on ω. We will use this completely elementary device without too much of an explicit mention.

For the purposes of this section, it will be sufficient to use particular cases of the neighborhoods $\Sigma_{\Omega,\alpha,\beta,U}$. Since all the objects we will consider will not depend on ω, we will not need to consider objects that depend on this, in particular we can suppress U from the notation.

We will also introduce the simplified domains

$$\Sigma_\delta = \{(p, q, \varepsilon) \mid |p| \leq \delta, \quad |\mathrm{Im}\ q| \leq \delta, \quad d(\varepsilon, [0, 1)) \leq \delta\}$$

and, given a family of functions $H_\varepsilon(p, q)$, we will denote by

$$\|H\|_\delta = \sup_{(p,q,\varepsilon) \in \Sigma_\delta} |H_\varepsilon(p, q)|$$

Since we will be working with functions that vanish at the origin to a high order, it is worth remarking that Cauchy bounds can be improved for them.

Proposition 4.1. *Let $H_\varepsilon(p,q)$ be such that $H_\varepsilon(p,q) = p^n J_\varepsilon(p,q)$. Then, provided that the norms are defined,*

i) $\|J\|_\delta = \delta^{-n}\|H\|_\delta$

and, for $\delta' < \delta$, we have

ii) $\|H\|_{\delta'} \le (\delta'/\delta)^n \|H\|_\delta$

iii) $\|\nabla H\|_{\delta'} \le (n/\delta' + (\delta-\delta')^{-1})(\delta'/\delta)^n\|H\|_\delta$

Proof. By the maximum modulus principle

$$\|H\|_\delta = \sup_{\Sigma_\delta} |H_\varepsilon(p,q)| = \sup_{\substack{|p|=\delta \\ |\mathrm{Im}\ q|\le\delta \\ d(\varepsilon,[0,1])\le\delta}} |H_\varepsilon(p,q)| = \delta^n \sup_{\Sigma_\delta} |J_\varepsilon(p,q)| = \delta^n \|J\|_\delta$$

This proves i). Then,

$$\|H\|_{\delta'} = \delta'^n \|J\|_{\delta'} \le \delta'^n \|J\|_\delta = (\delta'/\delta)^n \|H\|_\delta$$

Furthermore,

$$\begin{aligned}\|\nabla(p^n J_\varepsilon)\|_{\delta'} &= \|(np^{n-1}J_\varepsilon + p^n\partial_p J_\varepsilon, p^n\partial_q J_\varepsilon)\|_{\delta'} \le n\delta'^{(n-1)}\delta^{-n}\|H\|_\delta + \delta'^n\|\nabla J_\varepsilon\|_{\delta'} \\ &\le n\delta'^{-1}(\delta'/\delta)^n\|H\|_\delta + \delta'^n(\delta-\delta')^{-1}\|J_\varepsilon\|_\delta \\ &\le (n\delta'^{-1} + (\delta-\delta')^{-1})(\delta'/\delta)^n\|H\|_\delta\end{aligned}$$

■

4.2. Reduction of maps to integrable in a neighborhood of a Diophantine circle

The key step in the proof of Theorem 1.6 is the following. Once we prove this result, the proof will be the same as in [FL].

Lemma 4.2. *Let ω_0 be a Diophantine number, M an integer. Let f be an analytic area preserving map of the form*

$$f(p,q) = \left(p, q + \omega_0 + \kappa p^M\right) + O\left(p^{M+1}\right)$$

for some $\kappa \neq 0$. Then,

i) For every $N \in \mathbb{N}$ we can find a symplectic and analytic canonical transformation such that

$$g_N^{-1} \circ T \circ g_N(p,q) = (p, q + \Omega_N(p)) + R_N(p,q) \tag{4.1}$$

with Ω_N analytic, $\Omega_N(p) = \omega_0 + \kappa p^M + O(p^{M+1})$, and $|R_N(p,q)| \le C_N|p|^N$.

ii) Moreover, we can find $\mu_1, \mu_2 > 0$ depending only on M and the Diophantine properties of ω_0, such that for sufficiently small δ, choosing $N = K\delta^{\mu_1}$, we have

$$\|R_N\|_\delta \le K\exp(-K^{-1}\delta^{-\mu_2}) \tag{4.2}$$

Remark. We note that Lemma 4.2, besides giving some control on the periodic orbits that we will use to prove Theorem 1.6, also provides control over other orbits. Notably, it shows that critical circles are approximated by KAM circles. Indeed, the density of KAM circles in a neighborhood of size δ of a critical circle will be bigger than $1 - C_1\exp(-C_2\delta^{-\alpha})$ for some positive C_1, C_2, α.

Remark. We observe that the first part of the claim, the reduction to an integrable form could go through with less differentiability. If we only want that $g_N \in \mathcal{C}^4$ (which we will show is enough to show that the residue goes to zero faster than $|\omega_0 - m/n|^{N/M}$) it would suffice to assume that f is $\mathcal{C}^r$ with r depending on N and the Diophantine properties of ω_0. Of course, the quantitative estimates (4.2) depend on the analyticity properties. The first part of the claim is much easier to prove, since, as we will see, only entails matching powers of p in an equation that expresses the desired result. We note that this is enough to show using the methods that we will develop later that if there is a finitely differentiable circle, then the residue of a periodic orbit of type m/n is smaller than a power of $|\omega_0 - m/n|$. This power can be made as large as we want by assuming that the differentiability is high enough.

Proof of i). If we denote by $f_0(p,q) = (p, q + \omega_0 + \kappa p^M)$, by Lemma 2.2 we can find an analytic family f_ε that interpolates between f_0 and f. The Hamiltonian of this family F^0_ε will be an analytic function of (p,q,ε) in a complex neighborhood of Σ_δ.

To prove that we can find g_N so that (4.1) holds, we proceed by induction in N and assume that for some $N \geq 2$ we can write our Hamiltonian as

$$F^N_\varepsilon(p,q) = I^N_\varepsilon(p) + E^N_\varepsilon(p,q) \tag{4.3}$$

with

$$E^N_\varepsilon(p,q) = p^N R^N_\varepsilon(p,q)$$

We seek Hamiltonians $G^N_\varepsilon(p,q) = p^N S^N_\varepsilon(q)$ determined in such a way that the family g^N_ε with this Hamiltonian and starting in the identity is such that

$$j_\varepsilon = (g^N_\varepsilon)^{-1} \circ f^N_\varepsilon \circ g^N_\varepsilon \tag{4.4}$$

has a Hamiltonian which is integrable up to a higher order error in p.

We note that

$$g^N_\varepsilon(p,q) = \big(p + p^N \Delta_p(p,q), q + p^{N-1}\Delta_q(p,q)\big) \tag{4.5}$$

where Δ_p, Δ_q are analytic functions. Therefore, the compositions needed to define j_ε in (4.4) make sense in a sufficiently small neighborhood of the circle.

From Proposition 2.1 and (4.5), we can compute the Hamiltonian of j_ε

$$J_\varepsilon = I^N_\varepsilon \circ g^N_\varepsilon + (p + p^N\Delta_p)^N \cdot R^N_\varepsilon \circ g_\varepsilon - G_\varepsilon \circ g^N_\varepsilon + G^N_\varepsilon \circ (f^N_\varepsilon)^{-1} \circ g^N_\varepsilon \tag{4.6}$$

Expanding the above formula and denoting $R^N_\varepsilon(p,q) = \sum_{i\geq 0} p^i R^{N,i}_\varepsilon(q)$ — and analogously for other functions — we obtain

$$J_\varepsilon(p,q) = I^N_\varepsilon(p) + p^N \overline{R^{N,0}_\varepsilon} + p^N\left\{\left(R^{N,0}_\varepsilon(q) - \overline{R^{N,0}_\varepsilon}\right) - S^N_\varepsilon(q) + S^N_\varepsilon(q-\omega_0)\right\} + O(p^{N+1})$$

Using Lemma 3.2 we now that we can find an analytic S^N so that the term in braces is zero in the domain where the function is defined, which includes a strip around the torus. By the form of the functions, all the compositions needed to define j_ε will be defined in a sufficiently small strip around of the torus.

This establishes the first part of the claim, the fact that we can reduce to any order.

Remark. Rather than using an inductive argument, as we have done, it is possible to show that (4.1) holds to all orders by matching terms in (4.6). We note that the terms of order p^{N+m} have the form:

$$R^{N,m}_\varepsilon(q) - S^{N,m}_\varepsilon(q) + S^{N,m}_\varepsilon(q-\omega_0) + \tilde{R}^{N,m-1}_\varepsilon(p,q)$$

where $\tilde{R}^{N,m-1}_\varepsilon$ is a polynomial expression in $R^{N,i}_\varepsilon$, $S^{N,i}_\varepsilon$, $i \leq m-1$ and their derivatives and the derivatives of I. Again, we can use Lemma 3.2 to prove that a solution exists to all orders in p^n.

This method clearly shows that the coefficients of the expansion in the reduction are uniquely determined by the map and the torus, and are independent of the procedure. For example, in [OS], a different procedure using generating functions is used for twist maps and one can find the remark that the coefficients of this normal form are unique. (For the situation we are considering here, generating functions are not so convenient since the mixed variables are not a good system of coordinates in a neighborhood of the invariant torus. Nevertheless, the formalism that we developed above allows us to reach the same uniqueness conclusions.)

To obtain the estimates on the remainders of the reduction, we use an slightly different procedure. We use (3.20) and determine G_ε in exactly the same way as in section 3. We can apply Lemma 3.6 — which does not depend on Ω being non degenerate — to obtain, with the notation introduced there, (3.53) and (3.54) provided that the inductive hypothesis hold.

Lemma 4.3. *Let Ω be the frequency function (3.11) defined in $\Sigma_{\Omega,\alpha,\beta,U}$ as in (3.9). Let Δ be defined as in (3.23) and let σ be a positive number. Assume that (3.17), (3.28), (3.33), (3.48), hold. Consider $\tilde\Omega$, the new frequency function defined by*

$$\tilde\Omega(\omega,\varepsilon,p) = \Omega(\omega,\varepsilon,p) + \int_0^\varepsilon ds\, \frac{\partial}{\partial p}\Delta(\omega,s,p) \tag{4.7.}$$

Then, for any $\tilde\alpha \le \alpha$ satisfying

$$K\sigma^{-1}\|E\|_{\Omega,\alpha,\beta,U} \le \tilde\alpha. \tag{4.8}$$

we have:

i) $\left\|\Omega - \tilde\Omega\right\|_{U_\sigma} \le K\sigma^{-1}\|E\|_{\Omega,\alpha,\beta,U} \le \tilde\alpha$

ii) For $\tilde\alpha$ as before, $\tilde\beta = \beta - \alpha - 4\sigma$, $\tilde U \equiv U_{4\sigma}$, we have:

$$\Sigma_{\tilde\Omega,\tilde\alpha,\tilde\beta,\tilde U} \subset \Sigma_{\Omega,2\alpha,\beta-4\sigma,U_{4\sigma}}$$

iii)

$$\left\|\left(\frac{\partial^M}{\partial p^M}\tilde\Omega\right)^{-1}\right\|_{\tilde U} \le \left\|\left(\frac{\partial^M}{\partial p^M}\tilde\Omega\right)^{-1}\right\|_{U_{4\sigma}} \le \left\|\left(\frac{\partial^M}{\partial p^M}\Omega\right)^{-1}\right\|_{U} + K\sigma^{-M-1}\|E\|_{\Omega,\alpha,\beta,U}$$

iv) The inequalities (3.18) hold. That is, for $\tau = 2\nu + 3$

$$\|\tilde E\|_{\tilde\Omega,\tilde\alpha,\tilde\beta,\tilde U} \le K\sigma^{-\tau}\|E\|_{\Omega,\alpha,\beta,U}(\|E\|_{\Omega,\alpha,\beta,U} + \tilde\alpha)$$

The only difference between the proofs of Lemma 3.7 and Lemma 4.3 is that in Lemma 4.3 we do not need to worry about the non-degeneracy in Ω with respect to ω. Item *iii)* in Lemma 4.3 is just an slight generalization of the standard implicit function theorem.

■

We also note that if E is $O(p^L)$, then G is also $O(p^L)$ and, as a consequence, all the terms in the decomposition of $\tilde E$ according to (3.20) are $O(p^{2L-1})$ except $(G_{\omega,\varepsilon}\circ i^{-1}_{\omega,\varepsilon} - G_{\omega,\varepsilon}\circ T^0)$ which is only $O(p^{L+M})$. We note that, for high enough L, $2L-1 > L+M$ so that, for large enough L the order of tangency grows by M in each step.

We can therefore assume that if we have performed n steps, the resulting non-integrable part is $O(p^{Mn-A})$ where A is a number that may depend only on M and not on n. The number A takes into account that in the first steps of the iteration it could happen that $2L-1$ is smaller than $L+M$.

Remark. One could have obtained slightly more sophisticated estimates taking advantage of the fact that the functions we are considering vanish with powers of A and we can use the sharper Proposition 4.1 instead of Lemma 3.1. As it turns out, this does not make an appreciable difference in the final answer and it would require that the estimates leading to Lemma 3.6 are redone.

Proof of part *ii)* of Lemma 4.3: Iteration of the inductive step.

Now we discuss the possibility and the effect of iterating the inductive step. Since the goals are quite different than in the iteration leading to the KAM theorem, the choices that we will make in domain losses etc. will be also quite different. In our case, we are not interested in having some analyticity domain left (the existence of an analytic torus is part of the assumptions); rather, we are interested in obtaining control of the remainders in a wide domain.

We will take take

$$\alpha_{n+1} = cn^{-\eta}, \qquad \beta_n = cn^{-\gamma} \tag{4.9}$$

with η, $\gamma > 0$ chosen in such a way that

$$(\gamma+1)\tau - \eta < 0 \tag{4.10}$$

Note that then, $\eta > \tau\gamma \geq \gamma+1$ so that the domains in the p variable are smaller than those in the q variable. Moreover, $\sigma_n = (\beta_n - \beta_{n+1})/4 = c\gamma n^{-\gamma-1} + O(n^{-\gamma-2})$ and we can bound $\sigma_n^{-\tau} \leq Kn^{\tau(\gamma+1)}$. Note also that it also follows from (4.10) that $\eta > \tau$ and that given any $\eta > \tau$ we can chose $\gamma > 0$ in such a way that (4.10) is satisfied.

We claim that if the iterative step can be iterated N times, and c as in (4.9) is sufficiently small, we have:

$$\|E^N\|_{\Omega^N,\alpha_N,\beta_N,U^N} \leq (N!)^{(\gamma+1)\tau-\eta} \tag{4.11}$$

We can proceed by induction. Note that if (4.11) were true, we could, for $N > N_0(c)$, obtain the bound $\|E^N\|_{\Omega^N,\alpha_N,\beta_N,U^N} + \alpha_{N+1} \leq K\alpha_{N+1}$. Then,

$$\|E^{N+1}\|_{\Omega^{N+1},\alpha_{N+1},\beta_{N+1},U^{N+1}} \leq (N!)^{(\gamma+1)\tau-\eta} Kc(N+1)^{(\gamma+1)\tau-\eta}$$

which implies the result for $N+1$ when c is small enough.

We note, as in the proof of the KAM theorem, that all but one of the hypotheses of the iterative step are satisfied provided that $\|E^n\|$ is much smaller than σ_n to a fixed power. The only condition that involves the α is (4.8). Namely,

$$K\sigma^{-1}\|E\|_{\Omega,\alpha,\beta,U} \leq \tilde{\alpha}$$

We note that, if we fix c, we have the hypothesis satisfied for $N > N_1 > N_0$. If we assume that the error is sufficiently small to start with — which can be assumed if we start in a neighborhood sufficiently small —, then, we can perform the N_1 steps and then, the iteration can continue. Therefore, if the initial error $\|E^0\|$ is sufficiently small, we can iterate indefinitely. Notice that since E^0 vanishes up to order M in p it suffices to choose c sufficiently small.

Moreover, the estimates *iii)* of Lemma 4.3 tell us that we can bound from below $\left|\left(\frac{\partial^M}{\partial p^M}\Omega\right)^{-1}\right|$ independently of the number of iterates. Then, the domain Σ_δ is contained in all the domains of the form $\Sigma_{\Omega^N,K\delta^{1/M},K\delta^{1/M},U^N}$ provided that U^N contains a neighborhood of the map.

With the choices of α_n β_n that we have made above in (4.9), we see that we can repeat the iterative step described in Lemma 4.3 and obtain control in a 2δ neighborhood of the circle while $cN^{-\eta} \geq K\delta^{1/M}$. That is, $N \leq K^{-1}\delta^{-1/(M\eta)}$.

As we have seen in (4.11), for N large enough — which is implied by δ small enough — we have:

$$\|E^N\|_{\Omega^N,\alpha_N,\beta_N,U^N} \leq (N!)^{(\gamma+1)\tau-\eta} \leq \exp\left(-K^{-1}\delta^{-1/(M\eta)}|\log(\delta^{-1/(M\eta)})| + K\right)$$

By worsening slightly the power of δ in the first term, we can suppress the logarithm to simplify the expression

$$\|E^N\|_{\Omega^N,\alpha_N,\beta_N,U^N} \leq (N!)^{(\gamma+1)\tau-\eta} \leq \exp\left(-K^{-1}\delta^{-1/(M\eta)-\zeta}\right) \tag{4.12}$$

for some small $\zeta > 0$. Now, we note that the system $f_{\omega,\varepsilon}$ is obtained by solving up to time 1 the system

$$\frac{d}{d\varepsilon}x = \mathcal{F}_\varepsilon^N(x) = \mathcal{I}_\varepsilon^N(x) + \mathcal{E}_\varepsilon^N(x)$$

Applying Cauchy bounds to (4.12), we can obtain bounds for $\mathcal{E}^{\mathcal{N}}$ in a δ neighborhood of the origin which are of the same form as (4.12) with an slightly bigger ζ and some bigger K.

Note that, by definition, $\mathcal{I}^N$ generates an integrable flow. Hence, applying the usual estimates for the dependence of the solutions on the vector field, we obtain the result claimed in Lemma 4.2.

Note that the argument we have given shows that we can take $\mu_1 = 1/M(\eta)$ and μ_2 any number strictly smaller. Since we only needed $(\gamma+1)\tau - \eta < 0$, we can choose η any number bigger than τ and then choose γ. Of course, the constants will be worse.

■

4.3. Proof of Theorem 1.6 using Lemma 4.2

A possible proof can be made following the argument in [FL].

We note that Theorem 1.6 makes statements about the trace of derivatives of F^n at fixed points of F^n. Since the trace of the derivative of a map at a fixed point is invariant under changes of coordinates, we can study the derivatives of this map in the coordinates provided by Lemma 4.2.

First, we need to obtain some idea of where the periodic orbits could be. We will need to show that if $|\omega_0 - m/n|$ is small, then the orbit is very close to the invariant circle so that, in the coordinates provided by Lemma 4.2, the orbit is close to being the orbit of an integrable system. Note that for the orbit of an integrable system, the derivative is upper triangular with a diagonal which is the identity (hence, for an integrable system the trace of the derivative is 2 and the residue is 0). A second part of the argument is a perturbation argument that shows that if the system is close to integrable, the trace of the derivative is close to 2 and, hence, the residue is small.

The first part of the argument is accomplished by the following

Proposition 4.4. *For m/n sufficiently close to ω_0, any orbit of type m/n should be contained in annuli of radii $r \pm O(r^{1+\varepsilon})$ where r satisfies $\omega_0 + \kappa_M r^M = m/n$.*

We see that, when M is odd, we find one such r, namely $r = ((\omega_0 - m/n)/\kappa_M)^{1/M}$. When M is even, if $(\omega_0 - m/n)/\kappa_M$ is positive we can find two such r, namely $r = \pm\,(m/n - \omega_0)/\kappa_M)^{1/M}$ and when $(m/n - \omega_0)/\kappa_M$ is negative, we can find none. (In general, for each of the values of r that guarantee the existence of periodic orbits, they will appear in pairs: elliptic and hyperbolic.)

The argument will also show that, when we cannot find any r solving the equation, there are no periodic orbits of type m/n in a sufficiently small neighborhood of the non-degenerate circle.

Proof. If we apply the first claim of Lemma 4.2 to order $2M+2$, we obtain that, in an appropriate system of coordinates, our map can be written as

$$(p, q) \mapsto (p, q + \Omega(p)) + O\left(p^{2M+2}\right) \tag{4.13}$$

with $\Omega(p) = \kappa p^M + O(p^{M+1})$.

In the set $I = [(9/10)r, (11/10)r] \times \mathbb{T}^1$, the mapping (4.13) can be considered as a perturbation of an integrable system.

We note that the frequencies present in the integrable system in the domain considered are

$$\omega_0 + \kappa r^M[(9/10)^M, (11/10)^M] + O(r^{M+1}).$$

Note also that $\frac{d\Omega}{dp} > \kappa M r^{M-1} + O(r^M)$. This lower bound on the derivative is called the twist constant.

We recall that, by standard arguments in Diophantine approximation, we can find ω^* such that $\forall\ i \in \mathbb{Z}, \forall\ j \in \mathbb{N}$, $|\omega^* - i/j|^{-1} \leq Cj^{5/4}$ in any interval of length bigger than KC^{-1}. (It suffices to fix i, j and consider the length of the interval of ω for which the desired inequality fails. See e.g. [AA] p. 252.)

Hence, we can find two frequencies $\omega_\pm$ such that

a) They are Diophantine with exponent $5/4$ and with constant $C = r^{-M}$
b) $\omega_- < m/n < \omega_+$
c) $\omega_+ - \omega_- \leq Kr^M$

We now recall the quantitative version of the twist mapping theorem [He] that states that if we perturb an integrable system with twist constant σ defined in a range of A of diameter D, by a perturbation of $\mathcal{C}^4$ size ρ, the invariant circles corresponding to a Diophantine frequency of constant C persist provided that $C^2\rho \; \rho\sigma^{-1}/D$ are sufficiently small. Moreover, $\mathcal{C}^1$ distance of these invariant tori to the unperturbed ones can be bound by $\rho\sigma^{-1}$.

If we apply this to the circles of frequencies $\omega_\pm$ in the domain indicated, we see that $\rho = O(r^{2M+2})$, $C = O(R^{-M})$, $\sigma^{-1} = O(r^{-M+1})$, and $D \geq 2/10r$.

Hence, we conclude that these circles with frequency $\omega_\pm$ persist. Since in a sufficiently small neighborhood of the invariant circle, the map is a twist map, all the orbits with rotation number in $[\omega_-, \omega_+]$ have to be contained in the annulus bounded by these two invariant circles. In particular those of rotation number m/n.

This finishes the proof of Proposition 4.4.

■

For the cases where we can find an r such that the rotation number of the integrable part is m/n, we can apply Lemma 4.2 with $\delta = 2r$ with r as above to obtain that $\|R_N\|_\delta$ vanishes to order $K^{-1}|\omega - m/n|^{-\mu_1/M}$ and has size smaller than $K\exp(-K^{-1}|\omega - m/n|^{-\mu_2/M})$.

The improved Cauchy estimates, Proposition 4.1, give us that the entries on the matrix DR are smaller than

$$\begin{aligned} 2^{-K^{-1}|\omega_0 - m/n|^{-\mu_1/M}} K \exp\left(-K^{-1}|\omega_0 - m/n|^{-\mu_2/M}\right) \\ \leq K \exp\left(-K^{-1}|\omega_0 - m/n|^{-\mu_3}\right) \end{aligned} \tag{4.14}$$

for some $\mu_3 > 0$.

We also note that the derivatives of the integrable part are of the form $DI = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ with a bounded independently of the number of iterates that we need to take in Lemma 4.2.

If we have a periodic orbit of type m/n, by the chain rule we have $DF^n(x) = DF(x_{n-1}) \cdots DF(x)$, where $x_i = F^i(x)$. Note that $DF(x_i) = DI(x_i) + DR(x_i)$.

Therefore, we can apply the following lemma, which appears as Lemma 3.4 of [FL].

Lemma 4.5. *Let $\{A_i\}_{i=1}^N$ be a set of 2×2 matrices of the form $A_i = \begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}$ with $\sup_{1 \leq i \leq N} |a_i| \leq A$.*

Let $\{B_i\}_{i=1}^N$ satisfy

$$\sup_{\substack{1 \leq i \leq N \\ j,k=1,2}} |(B_i)_{jk} - (A_i)_{jk}| \leq \varepsilon \quad \text{with } \varepsilon \leq A\ .$$

Then $B = B_1 \cdots B_N$ satisfies

$$|\operatorname{Tr} B - 2| \leq 2\left[\left(1 + 3\sqrt{A}\sqrt{\varepsilon}\right)^N - 1\right]$$

Applying Lemma 4.5 with $A_i = DI(x_i)$, $B_i = DF(x_i)$, we obtain that for sufficiently large n, recalling that Theorem 1.6 includes in the assumptions that $|\omega_0 - m/n| \leq 1/n$ and that therefore $K\exp(-K^{-1}|\omega_0 - m/n|^{\mu_3})$ tends to zero

$$\begin{aligned} |\operatorname{Tr} DF^n(x) - 2| &\leq 2\left[\left(1 + K\exp(-K^{-1}|\omega_0 - m/n|^{\mu_3})\right)^n - 1\right] \\ &\leq nK\exp(-K^{-1}|\omega_0 - m/n|^{\mu_3}) \leq K\exp(-K^{-1}|\omega_0 - m/n|^{\mu_4}) \end{aligned}$$

This concludes the proof of Theorem 1.6.

■

We also remark that the argument that we gave to locate the periodic orbits also shows that if we have a non-degenerate critical circle, then it is approximated by periodic orbits.

In the cases that we can find an approximate r (i.e. in the case of odd M or, when M is even, that the sign of $\omega_0 - m/n$ is chosen correctly) we see that we can apply Poincaré last geometric theorem to $F^n - (0, m)$

and obtain one periodic orbit. In the case that M is even and the signs are right, since we can find two rings we can obtain two periodic orbits.

A different line of argument that produces sharper results under stronger hypothesis is the following:

In a annulus $p \in [r - r^{1+\varepsilon}, r + r^{1+\varepsilon}]$ the map is an small perturbation of an integrable map that is non-degenerate. If this perturbation satisfies some non-degeneracy assumptions, one can find two periodic orbits of type m/n. One of them is hyperbolic an another one is elliptic. The first order calculations of these periodic orbits is sometimes called subharmonic Melnikov theory. Formal expansions, including non-degeneracy assumptions that imply that the expansions predict one pair of elliptic and hyperbolic periodic orbits can be found in [Po] §74, §79. A justification of these expansions for finitely differentiable functions that shows that, under the formal conditions derived in [Po] one can find indeed the periodic orbits with the character predicted by the expansions can be found in [LW] chapter 2, or in [Po] \$ 39

Remark. Note that the above argument only requires estimates about the trace of the derivative. The fact that the trace of the derivative can be studied requires that $g_N \in \mathcal{C}^1$. The argument that we used to show that, in the coordinates given by g_N, the periodic orbit of period m/n is at a distance not more that $|\omega_0 - m/n|^{1/M}$ requires the twist mapping theorem with Lipschitz estimates and hence that $g_N \in \mathcal{C}^4$. The rest of the argument applying Lemma 4.5 only requires that the map $g_N \in \mathcal{C}^1$. Hence we see that if $g_N \in \mathcal{C}^4$ we have that $|\mathrm{Res}(O_{m,n})| \leq K|\omega_0 - m/n|^{N/M}$. Therefore, as we remarked before, to show that the residue goes to zero faster than a power, one only needs finite differentiability and for $\mathcal{C}^\infty$ mappings one can show that the residue goes to zero faster than any power.

Remark. In higher dimensions, under the non-degeneracy hypothesis of KAM theorem — which are weaker than twist hypothesis — an argument similar to the one given above has been developed in [T1]. The reduction to integrable normal form up to a very small error can be carried out. Similarly, there is an analogue of Lemma 4.5 that shows that products of sufficiently small perturbations of Jordan Blocks with identity in the diagonal, still have characteristic polynomials close to $(t-1)^{2d}$. Therefore, if there is a periodic orbit in a neighborhood of the torus, not only the trace but all the other coefficients of the characteristic polynomial have to converge to those of the Jordan normal form. One important element from our present argument that does not generalize to higher dimensions is the application of the twist mapping theorem to conclude that the distance of the periodic orbits to the invariant circle is bounded by the difference of the rotation numbers. Nevertheless, it is possible to show that if there is an invariant torus, there are periodic orbits that approximate it well and that the characteristic polynomial of the derivative converges to $(t-1)^{2d}$. It has been argued — and implemented numerically in [T2] — that this convergence of the coefficients of the characteristic polynomial of the derivative can considered as a test of the presence of a KAM torus.

We think that it should be possible to extend the methods presented here to establish one of the implications of Greene's criterion for some invariant torus that satisfy some hypothesis of non-degeneracy weaker than the twist hypothesis.

5. Acknowledgements

We thank D. Del Castillo and P. Morrison for many discussions about their work and for encouragement. We also A. Haro and C. Simó for helpful discussions and for making their unpublished results available to us. This work has been partially supported by the NATO grant CRG950273. Research by A.D. is also supported by EC grant ERBCHRXCT-940460, the Spanish grant DGICYT PB94-0215 and the Catalan grant CIRIT 1996SGR–00105. Research by R.L. is also supported by NSF grants. We also thank TICAM, UPC and IMA for invitations that made possible this collaboration.

6. References

[AA] V.I. Arnol'd, A. Avez: "*Ergodic problems of classical mechanics*", Benjamin (1968).

[BHS] H.W. Broer, G.B. Huitema, M.B. Sevryuk: Quasi-periodic motions in families of dynamical systems (Order amidst chaos). *Lecture Notes in Math.* **1645**, (1996).

[BLW] A. Banyaga, R. de la Llave, C.E. Wayne: Cohomology equations near hyperbolic points and geometric versions of Sternberg linearization theorems. *J. Geom. Anal.* **4**, 613–649 (1996).

[BST] H. Broer, C. Simó, J.C. Tatjer: Towards global models near homoclinic tangencies of dissipative diffeomorphisms. *Nonlinearity* **11**, 667–770 (1998).

[C] D. del-Castillo-Negrete: Univ. of Texas Thesis. (1995).

[CGM1] D. del-Castillo-Negrete, J.M. Greene, P.J. Morrison: Area preserving nontwist maps: periodic orbits and transition to chaos. *Phys. D.* **91**, 1–23 (1996).

[CGM2] D. del-Castillo-Negrete, J.M. Greene, P.J. Morrison: Renormalization and transition to chaos in area preserving nontwist maps. *Phys. D.* **100**, 311–329 (1997).

[CM] D. del-Castillo-Negrete, P.J. Morrison: Chaotic transport by Rossby waves in shear flow. *Phys. Fluids* **A 5**, 948–965 (1993).

[DG1] A. Delshams, P. Gutiérrez: Effective stability and KAM theory. *J. Differential Equations,* **128**, 415–490 (1996).

[DG2] A. Delshams, P. Gutiérrez: Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system. *J. Differential Equations,* **131**, 277–303 (1996).

[FL] C. Falcolini, R. de la Llave: A rigorous partial justification of Greene's criterion. *J. Statist. Phys.* **67**, 609–643 (1992).

[Gr] J.M. Greene: A method for determining a stochastic transition. *J. Math. Phys.* **20**, 1183–1201 (1979).

[Ha1] A. Haro: Univ. of Barcelona Thesis. (1998).

[Ha2] A. Haro: Converse KAM theory for monotone positive symplectomorphisms. *Preprint* (1998).

[He] M.R. Herman: Sur les courbes invariantes par les difféomorphismes de l'anneau. *Astérique* **103–104**, (1983).

[HH1] J.E. Howard, S.M. Hohs: Stochasticity and reconnection in Hamiltonian systems. *Phys. Rev.* **A 29**, 418–421 (1984).

[HH2] J.E. Howard, J. Humphreys: Nonmonotonic twist maps. *Phys. D.* **80**, 62–72 (1995).

[JV] A. Jorba, J. Villanueva: On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems. *Nonlinearity* **10**, 783–822 (1997).

[K] W.T. Kyner: Rigorous and formal stability of orbits about an oblate planet. *Mem. Amer. Math. Soc.* **81**, 1–27 (1968).

[KMOP1] J. Koiller, R. Markarian, S. Oliffson-Kamphorst, S. Pinto de Carvalho: Time-dependent billiards. *Nonlinearity* **8**, 983–1003 (1995).

[KMOP2] J. Koiller, R. Markarian, S. Oliffson-Kamphorst, S. Pinto de Carvalho: Static and time-dependent perturbations of the classical elliptical billiard. *J. Statist. Phys.* **83**, 127–143 (1996).

[KO] Y. Katznelson, D.S. Ornstein: A new method for twist theorems. *J. Anal. Math.* **60**, 157–208 (1983).

[LMM] R. de la Llave, J.M. Marco, R. Moriyón: Canonical perturbation theories of Anosov diffeomorphisms and regularity results for the Livsic cohomology equation. *Ann. of Math.* **123**, 537–611 (1986).

[Ll] R. de la Llave: Introduction to KAM theory. *Preprint #93-8 in* `mp_arc` (1993).

[LW] R. de la Llave, C.E. Wayne: Whiskered and low dimensional invariant tori for near integrable Hamiltonian systems. *Preprint* (1989).

[Mat] J. Mather: Stability of C^∞ mappings II: Infinitesimal stability implies stability. *Ann. of Math.* **89**, 254–291 (1969).

[McK] R.S. McKay: Princeton Univ. Thesis. (1982).

[McK2] R.S. McKay: On Greene's residue criterion. *Nonlinearity* **5**, 161–187 (1992).

[Mo1] J. Moser: On the volume elements of a manifold. *Trans. Amer. Math. Soc.* **120**, 286–294 (1965).

[OS] A. Olvera, C. Simó: Normal forms close to invariant circles of twist maps. In "*European Conference in Iteration Theory (ECIT 87)*", (C. Alsina, J. Llibre, Ch. Mira, C. Simó, G. Targonski, R. Thibault, editors). World Scientific 438–443 (1989).

[Po] H. Poincaré: "*Les méthodes nouvelles de la mécanique céleste*", Gauthier Villars (1899).

[PW] A. Perry, S. Wiggins: KAM tori are very sticky: rigorous lower bounds on the time to move away from an invariant Lagrangian torus with linear flow. *Phys. D* **71**, 102–121 (1994).

[Ru] H. Rüssmann: On optimal estimates for the solutions of linear difference equations on the circle. *Celestial Mech. Dynam. Astronom.* **14**, 33–37 (1976).

[Ru2] H. Rüssmann: On a new proof of Moser's twist mapping theorem. *Celestial Mech. Dynam. Astronom.* **14**, 19–31 (1976).

[Si] C. Simó: Invariant curves of analytic perturbed nontwist area preserving maps. *Preprint* (1998).

[SM] C.L. Siegel, J. Moser: "*Lectures on Celestial Mechanics*", Springer-Verlag, New York (1971).

[TL] H.I. Levine: Singularities of differentiable mappings, notes of a course by R. Thom. In "*Proceedings of Liverpool Singularities— Symposium. I (1969/1970)*", (C.T.C. Wall, editor). Springer Verlag, Berlin, *Lecture Notes in Math.* **192**, (1971).

[T1] S. Tompaidis: Approximation of invariant surfaces by periodic orbits in high-dimensional maps: some rigorous results. *Experiment. Math.* **5**, 197–209 (1996).

[T2] S. Tompaidis: Numerical study of invariant sets of a quasiperiodic perturbation of a symplectic map. *Experiment. Math.* **5**, 211–230 (1996).

[VG] T.P. Valkering, S.A. van Gils: Bifurcation of periodic orbits near a frequency maximum in near-integrable driven oscillators with friction. *Z. Angew. Math. Phys.* **44**, 103–130 (1993).

[W] A. Weinstein: Lagrangian submanifolds and Hamiltonian systems. *Ann. of Math.* **98**, 377–410 (1973).

[ZZSUC] G. Zaslavsky: Stochastic web and diffusion of particles in a magnetic field. *Soviet Phys. JETP* **64**, 294–303 (1987). Translated from *Zh. Èksper. Teoret. Fiz.*, **91**, 500–516 (1986).